Adaptive Low-Rank Methods and Model Reduction for High-Dimensional Parabolic PDEs

Von der Fakultät für Mathematik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Manfred Faldum

aus

Mainz, Deutschland

Berichter: Prof. Dr. Markus Bachmayr
Prof. Dr. Christof Melcher
Prof. Dr. Karsten Urban

Tag der mündlichen Prüfung: 23.09.2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Eidesstattliche Erklärung

Manfred Faldum

erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:

   Markus Bachmayr und Manfred Faldum. „A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic Partial Differential Equations“. In: *Journal of Complexity* 82 (Juni 2024), S. 101839. ISSN: 0885064X. DOI: 10.1016/j.jco.2024.101839

   Manfred Faldum. „A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic PDEs“. In: *Oberwolfach Reports*. Hrsg. von Markus Bachmayr, Albert Cohen, Angela Kunoth und Olga Mula. Bd. 37. Computation and Learning in High Dimensions. 2025, S. 2021–2023. DOI: 10.4171/OWR/2025/37

   Manfred Faldum, Markus Bachmayr, Volkmar Schulz und Franziska Schrank. „Efficient Solvers for Coupled Brown-Néel Fokker-Planck Equations“. In: *International Journal on Magnetic Particle Imaging IJMPI* 11.1 Suppl 1 (März 2025). DOI: 10.18416/IJMPI.2025.2503011

Aachen, den 09.03.2026

(Manfred Faldum)

Abstract

High-dimensional parabolic partial differential equations (PDEs) pose significant computational challenges due to the curse of dimensionality, in which the complexity of determining numerical solutions to the PDE grows exponentially with the dimension. In this work, we develop and analyze efficient numerical methods to address two distinct sources of high dimensionality in parabolic PDEs: high-dimensional spatial domains and high-dimensional parameter spaces. The methods for both kinds of problems are based on space-time variational formulations and low-rank approximations.

For solving linear parabolic PDEs on high-dimensional spatial domains, we propose a space-time adaptive low-rank method. The main idea is to represent the numerical solution using a sparse wavelet expansion in the temporal variable, combined with adaptive low-rank approximations for the spatial coefficients. This specific structure allows for the independent adaptation of spatial ranks for each temporal basis function, offering greater flexibility than treating time as a tensor component. Based on a space-time variational formulation, we obtain a representation of the solution to the PDE as a bi-infinite matrix-vector equation. The numerical solution is computed using a perturbed Richardson iteration, in which the operator is applied approximately with controlled accuracy. Theoretical analysis using the concept of Riesz bases demonstrates that this method converges in the underlying space-time norm. To maintain the low-rank structure within the solver, exponential sum approximations are used to construct separable representations of the scaling matrices required for Riesz basis preconditioning. Furthermore, under natural regularity assumptions on the solution and compressibility assumptions on the operator, the method is proven to circumvent the curse of dimensionality, i.e., computational complexity that does not grow exponentially in the number of spatial dimensions. Numerical experiments on the high-dimensional heat equation validate the theoretical findings, demonstrating the effectiveness of the method up to spatial dimension $d = 128$.

The second part of this work addresses the high dimensionality arising from a parameter dependence. We consider a specific real-world problem, which is a tracer-based tomographic imaging method called magnetic particle imaging (MPI). The coupled Brown-Néel relaxation model, describing the stochastic dynamics of magnetic nanoparticles, leads to a parameter-dependent Fokker-Planck equation. This equation is defined on a manifold, which is the product of two unit spheres and time. The parameters in this case are the physical properties of the magnetic nanoparticles and of the magnetic field. In particular, in the context of MPI we are interested in computing the solution of the Fokker-Planck equation for numerous parameter combinations. However, determining the numerical solution directly is infeasible due to the large number of possible combinations. Hence, we use a model order reduction technique, namely the reduced basis (RB) method, to efficiently find a parameter-to-solution map. For typical MPI parameters, the Fokker-Planck equation is convection-dominated. To address this, we use a hybrid discontinuous Galerkin (HDG) finite element method within the RB method, which ensures stability. We present the construction of the RB using the HDG method for this Fokker-Planck equation posed on a manifold. Moreover, we state the efficient offline and online decomposition. Preliminary numerical results demonstrate that the RB method achieves significant computational speed-ups while maintaining accurate solutions, enabling the possibility for parameter studies and model-based reconstruction for MPI.

Zusammenfassung

Hochdimensionale parabolische Differentialgleichungen (PDEs) sind aufgrund des Fluchs der Dimensionalität sehr herausfordernd, da die Komplexität zur Bestimmung numerischer Lösungen exponentiell mit der Dimension des Problems ansteigt. In dieser Arbeit entwickeln und analysieren wir effiziente numerische Methoden zur Behandlung zweier unterschiedlicher Arten hochdimensionaler parabolischer PDEs: hochdimensionaler räumlicher Gebiete und hochdimensionaler Parameterräume. Die Methoden für beide Problemstellungen basieren auf Raum-Zeit-Variationsformulierungen und Niedrigrangapproximationen.

Zur Lösung linearer parabolischer PDEs auf hochdimensionalen räumlichen Gebieten entwickeln wir eine adaptive Raum-Zeit-Niedrigrangmethode. Die Grundidee besteht darin, die numerische Lösung als Wavelet-Entwicklung in der Zeitvariablen darzustellen und diese mit adaptiven Tensorniedrigrangapproximationen für die räumlichen Koeffizienten zu kombinieren. Diese Struktur ermöglicht eine separate Anpassung der räumlichen Ränge für jede zeitliche Basisfunktion und ist somit flexibler als die Behandlung der Zeit als Tensorkomponente. Ausgehend von einer Raum-Zeit-Variationsformulierung erhalten wir eine Darstellung der Lösung als unendlich großes Gleichungssystem. Die numerische Lösung wird mittels einer gestörten Richardson-Iteration berechnet, bei der der Operator approximativ, jedoch mit kontrollierter Genauigkeit, angewendet wird. Eine theoretische Analyse unter Verwendung von Riesz-Basen zeigt die Konvergenz des Verfahrens in der zugrundeliegenden Raum-Zeit-Norm. Um die Niedrigrangstruktur während des Lösens zu erhalten, werden Exponentialsummenapproximationen der Skalierungsmatrizen für die Riesz-Basis-Vorkonditionierung verwendet. Zudem wird unter natürlichen Regularitätsannahmen für die Lösung und Kompressibilitätsannahmen für den Operator nachgewiesen, dass das Verfahren den Fluch der Dimensionalität umgeht. Numerische Experimente für die hochdimensionale Wärmeleitungsgleichung bestätigen die theoretischen Ergebnisse und zeigen die Durchführbarkeit des Verfahrens bis zur räumlichen Dimension $d = 128$.

Der zweite Teil befasst sich mit der Hochdimensionalität durch Parameterabhängigkeit. Dabei betrachten wir eine konkrete Anwendung namens Magnetpartikelbildgebung (MPI), ein tracerbasiertes tomographisches Bildgebungsverfahren. Das gekoppelte Brown-Néel -Relaxationsmodell, das die stochastische Dynamik magnetischer Nanopartikel beschreibt, führt zu einer parameterabhängigen Fokker-Planck-Gleichung. Diese ist auf einem Produkt aus zwei Einheitskugeln und der Zeit definiert. Die Parameter sind die physikalischen Eigenschaften der magnetischen Nanopartikel und des Magnetfeldes. Die direkte Bestimmung der Lösung der Fokker-Planck-Gleichung für die hohe Anzahl verschiedener Parameterkombinationen in MPI ist numerisch nicht realisierbar. Daher verwenden wir ein Verfahren zur Modellreduktion, die Reduzierte-Basis-Methode (RB-Methode), um effizient eine Abbildung von Parameter auf Lösungsmenge zu finden. Für typische MPI-Parameter ist die Fokker-Planck-Gleichung konvektionsdominant. Um Stabilität zu gewährleisten, verwenden wir innerhalb der RB-Methode eine hybride diskontinuierliche Galerkin (HDG) Finite-Elemente-Methode. Wir geben die Konstruktion der reduzierten Basis unter Verwendung der HDG-Methode für diese auf einer Mannigfaltigkeit definierte Fokker-Planck-Gleichung an. Zudem beschreiben wir die effiziente Zerlegung in Offline- und Online-Phasen. Erste numerische Ergebnisse zeigen, dass die RB-Methode eine signifikante Beschleunigung der Rechenzeit bei gleichbleibender Genauigkeit der Lösungen ermöglicht, was Parameterstudien und eine modellbasierte Rekonstruktion für MPI erlaubt.

Previous Publications

Parts of this work have been published already in a journal and in conference proceedings:

The first part of this work on the space-time adaptive low-rank method for parabolic PDEs on high-dimensional spatial domains, in particular, the results of Chapter 4 and parts of Chapter 3, has already been published in

Markus Bachmayr and Manfred Faldum. "A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic Partial Differential Equations". In: *Journal of Complexity* 82 (June 2024), p. 101839. ISSN: 0885064X. DOI: 10.1016/j.jco.2024.101839

The results of this work are the product of joint discussions. The author provided most of the technical proofs, conducted the numerical implementation and experiments, and authored the majority of the text. Minor changes and corrections have been made compared to the article, in particular, the treatment of the initial value operator in Section 4.3.4. The modification leads to different constants in the complexity result in Theorem 4.69. Moreover, the different treatment of the initial-value operator leads to a slightly modified proof of this theorem.

A brief overview of the methodology of the space-time adaptive low-rank method for high-dimensional parabolic PDEs, in particular, the method described in Chapters 2 to 4, has been published in

Manfred Faldum. "A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic PDEs". In: *Oberwolfach Reports*. Ed. by Markus Bachmayr, Albert Cohen, Angela Kunoth, and Olga Mula. Vol. 37. Computation and Learning in High Dimensions. 2025, pp. 2021–2023. DOI: 10.4171/OWR/2025/37

A brief overview of the proposed reduced basis method for the parameter-dependent Fokker-Planck equation in the context of magnetic particle imaging, in particular, the method described in Chapter 6, has been published in

Manfred Faldum, Markus Bachmayr, Volkmar Schulz, and Franziska Schrank. "Efficient Solvers for Coupled Brown-Néel Fokker-Planck Equations". In: *International Journal on Magnetic Particle Imaging IJMPI* 11.1 Suppl 1 (Mar. 2025). DOI: 10.18416/IJMPI.2025.2503011

The results of this work are also the product of joint discussions. The author conducted the numerical implementation and experiments and authored the majority of the text. In this article, only a brief introduction is given; the full details are developed in this work in Chapter 6. Moreover, the article presents numerical results from a simplified model used for preliminary testing of the method. The results in Section 6.4 are based on the more accurate model described in Chapter 5.

Acknowledgements

I want to sincerely thank my supervisor, Prof. Markus Bachmayr, for his excellent support and guidance during my PhD, which I deeply value. I am very grateful for the constructive discussions and insightful ideas, which were crucial for the development of this work. I am also grateful for his quick responses to my questions and for the opportunity to attend many conferences and workshops, which were instrumental to my research and to establishing a network in the scientific community.

I want to thank Prof. Christof Melcher and Prof. Karsten Urban for serving on my defense committee and for their review of this thesis.

I would also like to thank Dr. Henrik Eisenmann and Prof. Mathias Oster for many helpful and fruitful discussions about my research, which were invaluable to my progress. Moreover, I thank Dr. Simon Boisserée, Dr. Henrik Eisenmann, Prof. Andreas Faldum, Prof. Mathias Oster, Dr. Michael Rom, and Stephanie Schwab for proofreading parts of the manuscript and for their helpful comments.

I gratefully acknowledge the financial support through Deutsche Forschungsgemeinschaft of TRR146 during my time in Mainz and of the GRK 2326 (Eddy) for part of my time in Aachen.

My time in Aachen was greatly enriched by my colleagues at Eddy and the IGPM. I am very grateful for the great working atmosphere and the many social events, which made my time in Aachen very enjoyable. In particular, I want to thank Simon for the great times we have shared, which already started in Mainz; Henrik for the enjoyable chess games; Michael for the visits to the Alemannia games and the bicycle tours; and Mathias for the many memorable times in the lounge. Along with Hannah, Huqing, Matti, Sophie, and Stephanie, you all made for many wonderful evenings.

Additionally, I want to thank my friends from BV Aachen, especially Alex, Judith, Julia, and Sven. Playing badminton with you was always a lot of fun and a great way to relax after work. I am also enjoyed the board game nights with you, which were always a great time.

A huge thanks to my friends from Heidesheim, especially to Basti, Katha, Lara, Laura, Matthias, Michi, Müller, Stefan, and Steven. I am very grateful for all the wonderful times we have shared over the years, especially the many enjoyable trips we have taken together. You always managed to cheer me up during stressful times and were always there for me. You are by far the best friends anyone could wish for.

Special thanks go to my brother Stefan, Stephanie, and Steven. You are the most important people in my life, and I am incredibly grateful to have you. I am also happy that I shared my education in mathematics and computer science with all three of you, which made this journey truly special.

Finally, I want to thank my family for their great support during my PhD and all the beautiful times we shared together. A special thanks to my dad, with whom I always loved discussing math, and to my mom, who supported me in every way possible.

Contents

1 Introduction **1**
1.1 Outline . . . 4
1.2 Notation . . . 5

2 High-Dimensional Approximation **7**
2.1 Wavelets in One Dimension . . . 7
2.1.1 Construction of Wavelets . . . 8
2.1.2 Approximation Properties . . . 9
2.1.3 Multiwavelets . . . 10
2.2 Nonlinear Approximation in Sequence Spaces . . . 11
2.3 Tensor-Product Spaces . . . 13
2.3.1 Hilbert Spaces . . . 13
2.3.2 Tensor-Product Wavelet Bases . . . 15
2.4 Low-Rank Tensor Formats . . . 16
2.4.1 Canonical Tensor Format . . . 17
2.4.2 Tucker Format . . . 17
2.4.3 Hierarchical Tensor Format . . . 19
2.4.4 Hierarchical Singular Value Decomposition . . . 22
2.4.5 Operations in Hierarchical Tensor Format . . . 23
2.5 Wavelet Representation of Operators . . . 26

3 Parabolic Differential Equations and Space-Time Variational Formulation **31**
3.1 Space-Time Variational Formulation . . . 32
3.1.1 Infinite Matrix Representation of the Operator . . . 33
3.1.2 Second-Order Problems on High-Dimensional Spatial Domains 36
3.2 Alternative Variational Formulations . . . 40
3.2.1 Alternative Handling of Initial Condition . . . 41
3.2.2 Saddle Point Formulation . . . 41

4 Adaptive Low-Rank Method **43**
4.1 Adaptive Methods for Parabolic PDEs . . . 44
4.1.1 Near-Optimal Low-Rank Recompression . . . 45
4.1.2 Contractions and Coarsening . . . 49
4.1.3 Combination of Tensor Recompression and Coarsening . . . 54
4.2 Low-Rank Preconditioning . . . 57
4.2.1 Scaling Matrix $\mathbf{D}_{\mathcal{X}}$ . . . 58
4.2.2 Scaling Matrix $\mathbf{D}_{\mathcal{Y}}$ . . . 70
4.2.3 Relation to Low-Dimensional Scaling Matrices . . . 72
4.3 Adaptive Approximation of Rescaled Low-Rank Operators . . . 73
4.3.1 Spatial Operator . . . 74
4.3.2 Temporal Operator . . . 78
4.3.3 Combination of Operators . . . 81
4.3.4 Initial Value Operator . . . 94

4.4 Space-Time Adaptive Richardson Iteration . . . 107
4.4.1 A Convergent Solver . . . 107
4.4.2 Computational Complexity . . . 111
4.5 Numerical Experiments . . . 136
4.5.1 Vanishing Initial Condition . . . 137
4.5.2 Vanishing Source Term . . . 142

5 Magnetic Particle Imaging 149
5.1 Mathematical Models . . . 152
5.1.1 Relaxation Mechanism . . . 153
5.1.2 Coupled Brown-Néel Model . . . 154
5.2 Fokker-Planck Equation . . . 156

6 Reduced Basis Method for Convection-Dominated Parametric Fokker-Planck Equation 159
6.1 Reduced Basis Method . . . 160
6.1.1 Offline Phase . . . 161
6.1.2 Online Phase . . . 163
6.2 Hybrid Discontinuous Galerkin Finite Element Method . . . 167
6.2.1 Basic Assumptions and Notation . . . 168
6.2.2 Variational Formulation . . . 169
6.2.3 Mesh-Dependent Norm and Error Estimate . . . 172
6.2.4 Magnetic Particle Imaging Fokker-Planck Equation . . . 173
6.3 Finite Elements on Surfaces . . . 182
6.3.1 Triangulation of the Sphere . . . 182
6.3.2 Raviart-Thomas Functions . . . 186
6.4 Numerical Experiments . . . 189
6.4.1 Solutions for Fixed Parameters . . . 190
6.4.2 Reduced Basis Method . . . 195

7 Conclusion and Outlook 201
7.1 High-Dimensional Spatial Domain . . . 201
7.2 Parametric Fokker-Planck Equation . . . 202

A Supplementary Proofs 205

List of Abbreviations 213

List of Figures 215

List of Tables 217

Bibliography 219

Chapter 1

Introduction

High-dimensional parabolic partial differential equations (PDEs) arise in various applications. An important example is the Kolmogorov forward equation, which provides a deterministic description of stochastic processes by describing the time evolution of the corresponding probability density function. When working with many-particle systems, we consider the joint probability density function of all particles, which leads to a dimensionality proportional to the number of particles. Another example of high-dimensional parabolic PDEs arises in option pricing in financial mathematics, where the Black–Scholes equation describes the price of an option as a function of multiple underlying assets, leading to a high-dimensional PDE. A third example is problems in optimal control, for instance, the Hamilton-Jacobi-Bellman equation. The latter two problems typically involve additional nonlinearities.

All of these examples exhibit high-dimensional structure arising from the spatial domain. Another source of high dimensionality is the presence of additional parameters, where the solution $u(x, t; \mu)$ of the PDE not only depends on the spatial and temporal variables but also on parameters $\mu = (\mu_1, \dots, \mu_p) \in \mathcal{P}$ in a high-dimensional parameter set $\mathcal{P}$. In many applications of parametric PDEs, we are interested in the solution of the PDE for many parameter combinations $\mu \in \mathcal{P}$. This means that we want to approximate the parameter-to-solution map $\mu \mapsto u(\cdot, \cdot; \mu)$.

In this work, we investigate examples for both sources of high dimensionality: PDEs on high-dimensional spatial domains and parametric PDEs. For the parametric case, we are particularly interested in one specific PDE arising from a tracer-based tomographic imaging technique called magnetic particle imaging (MPI) introduced in [GW05]. In MPI, we focus on accurately modeling the dynamics of magnetic nanoparticles in an applied magnetic field. These dynamics can be characterized by a stochastic differential equation (SDE) [Wei+09]. The corresponding Fokker-Planck equation describes the evolution of the probability density function of the particle dynamics. The parameters in this case are physical properties of the particles and the magnetic field, which are typically not known exactly and vary across different particles and physical setups of the MPI device.

Numerical methods designed for low-dimensional domains, such as standard finite element methods, are infeasible for PDEs on high-dimensional domains because they suffer from the so-called *curse of dimensionality*, which means that the computational complexity grows exponentially with the number of spatial dimensions. Hence, in this work, we aim for a method based on sparse and low-rank representations that offers scalability to high dimensions while ensuring reliable, deterministic a posteriori error control. A central objective of the method is to circumvent the curse of dimensionality, meaning we want to derive bounds on the computational complexity that do not grow exponentially with the number of spatial dimensions.

Generally, we consider different approaches to address the two types of high dimensionality. However, both approaches are closely connected. For the underlying

parabolic PDEs, in both the high-dimensional spatial and parametric context, we use space-time variational formulations. Moreover, for both types of problems, we apply a low-rank approach to address high dimensionality. For high-dimensional parabolic PDEs, we use a low-rank tensor method to approximate the solution. In contrast, for the parametric PDE, we apply a reduced basis (RB) method, which can be interpreted as a low-rank approximation that separates the domain and parameter spaces.

Parabolic PDEs on High-Dimensional Spatial Domains

There exists a vast literature on numerical methods for high-dimensional PDEs. One possible approach is to exploit the sparsity patterns of the coefficients in suitable tensor-product bases. Examples of such methods are sparse grid [BG04] and adaptive wavelet methods [CDD01]. These methods can be used for problems with moderate dimensionality up to $d \approx 20$. However, as shown in [DSS09], relying only on sparsity in the coefficient tensor is, in general, not sufficient to overcome the curse of dimensionality.

Another approach is to represent the coefficients of a product basis in a suitable low-rank tensor format. In this case, we exploit additional structure in the solution beyond sparsity. For elliptic PDEs, adaptive methods using low-rank structure have been developed in [BD15; BD16; AU20]. The methods have been shown to overcome the curse of dimensionality under natural assumptions on the solution. Additionally, they provide a posteriori error control and convergence guarantees.

There are several options to extend these methods to time-dependent problems. One option is to use a time-stepping approach, applying the low-rank method for elliptic PDEs at each time step. In combination with fixed spatial discretizations, such methods have been used in [DKO12; CVK16]. However, the evolution of the ranks in the low-rank representation during time stepping is difficult to control, which can lead to a significant increase in computational complexity. Recent progress in this regard is presented in [BDS25], which adaptively adjusts the ranks to keep them proportional to the optimal approximation ranks. An alternative is dynamical low-rank methods [KL07; KL10], where we consider the evolution on the manifold of fixed tensor ranks. Dynamical low-rank methods have also been used for parabolic PDEs in [Con20; Bac+21; BEU25]. However, based on the restriction to fixed ranks, these methods do not provide a posteriori error control and convergence guarantees.

Another approach to time-stepping schemes is space-time variational formulations. Again, there have been several approaches using sparsity patterns, both in time and space, e.g., adaptive wavelet methods [SS09; CS11; KSU15; RS19]. Moreover, approaches have been developed that combine wavelets in time with finite elements in space in [And16; SVVW22].

Space-time variational formulations have also been combined with low-rank approximations in [AT15; Boi+19] with fixed discretization in space and time. In these works, the time variable is treated as a tensor component, and a low-rank representation is used for all variables. However, efficient low-rank separability between spatial and temporal variables is not guaranteed. In scenarios with strong convection or moving sources, the solution often develops complex structures in which space and time are tightly coupled. Consequently, accurately capturing these features requires a substantial increase in rank.

To tackle this problem, we consider a representation that is a *sparse* wavelet expansion in time combined with a *low-rank* representation in the spatial variables.

The corresponding low-rank representation for the spatial coefficients can be chosen independently for each temporal wavelet basis function θ_{ν_t}. This means we use approximations of the form

$$u(t, x_1, \dots, x_d) \approx \sum_{\nu_t \in \Lambda_t} \theta_{\nu_t}(t) u_{\nu_t}(x_1, \dots, x_d)$$

with a finite index set Λ_t representing the sparsity in the temporal variable and

$$u_{\nu_t}(x_1, \dots, x_d) = \sum_{\nu_1 \in \Lambda_{\nu_t}^{(1)}} \cdots \sum_{\nu_d \in \Lambda_{\nu_t}^{(d)}} \mathbf{u}_{\nu_t, \nu_1, \dots, \nu_d} \Phi_{\nu_1, \dots, \nu_d}(x_1, \dots, x_d),$$

where $\Phi_{\nu_1,\dots,\nu_d}$ are product wavelet functions for the spatial domain. The coefficient vectors $\mathbf{u}_{\nu_t} = (\mathbf{u}_{\nu_t,\nu_1,\dots,\nu_d})_{(\nu_1,\dots,\nu_d)}$ are stored in a low-rank format for each $\nu_t \in \Lambda_t$. Based on this representation, we consider potentially different low-rank representations for each temporal wavelet basis function. For example, in the case of $d = 2$, we have

$$\mathbf{u}_{\nu_t,\nu_1,\nu_2} = \sum_{k=1}^{r_{\nu_t}} \mathbf{u}^{(1)}_{\nu_t,k,\nu_1} \mathbf{u}^{(2)}_{\nu_t,k,\nu_2}, \quad (\nu_1, \nu_2) \in \Lambda^{(1)}_{\nu_t} \times \Lambda^{(2)}_{\nu_t},$$

where r_{ν_t} is the rank of the approximation and $(\mathbf{u}^{(i)}_{\nu_t,k,\nu})_\nu \in \mathbb{R}^{\Lambda^{(i)}_{\nu_t}}$ are the corresponding coefficient vectors for $i = 1, 2$ and $\nu_t \in \vee_t$. If the rank is small, we only need to store $r_{\nu_t}(\#\Lambda^{(1)}_{\nu_t} + \#\Lambda^{(2)}_{\nu_t})$ entries, which can be substantially less than the $\#\Lambda^{(1)}_{\nu_t} \cdot \#\Lambda^{(2)}_{\nu_t}$ entries needed for a full representation of $\mathbf{u}_{\nu_t}$. In particular, for d spatial dimensions, the full representation has $\prod_{i=1}^{d} \#\Lambda^{(i)}_{\nu_t}$ entries, which grows exponentially in d, leading to a curse of dimensionality.

To compute a numerical solution of the parabolic PDE in the combined sparse and low-rank form, we construct a space-time adaptive low-rank solver analogously to the method for elliptic problems in [BD16] using a perturbed Richardson iteration. As a result, several new challenges arise compared to the elliptic case. One major challenge is the interactions between the different low-rank representations of the temporal wavelet basis functions, caused by the application of the linear operator representing the time derivative. This requires more restrictive compressibility assumptions. Another challenge is the treatment of the initial value in the parabolic PDE. In this work, we prove that the method yields approximations that converge in the space-time norm associated to the space-time variational formulation, using the concepts of Riesz bases. Moreover, we show that the method can overcome the curse of dimensionality under natural low-rank approximability assumptions on the solution. Finally, we present numerical results for the proposed method applied to a high-dimensional heat equation on the unit cube. The numerical results validate the theoretical findings and show the effectiveness of the method in spatial dimensions up to $d = 128$.

Parametric Parabolic PDEs

If the parameter set is large, calculating the solution of a parametric PDE for many parameter combinations $\mu \in \mathcal{P}$ directly can be very expensive. Hence, we use a model order reduction technique, namely the RB method, to tackle the parameter dependence. The goal of the RB method is to find a low-dimensional approximation of the solution manifold. Then we approximate the solution for a new parameter in

this low-dimensional space, i.e.,

$$u(\cdot,\cdot;\mu) \approx \sum_{i=1}^{n} c_i(\mu) u^{(i)}. \tag{1.1}$$

Hence, we assume that u can be approximated well by a low-rank representation separating the domain and the parameter space. The RB method consists of two phases. In the offline phase, we determine the basis functions $u^{(i)}$ by computing accurate approximations of the PDE for selected parameter combinations. These parameter combinations are selected so that the resulting basis functions approximate the solution manifold well. In the online phase, we compute the coefficients $c_i(\mu)$ in (1.1) for a new parameter combination μ by solving a low-dimensional problem. Since we only have to solve a low-dimensional problem in the online phase, we can achieve a significant reduction in computational complexity compared to solving the PDE directly. The idea behind the RB method is to invest computational effort in the offline phase to compress the solution manifold, enabling fast evaluation in the online phase. The determination of the components $u^{(i)}$ of the low-dimensional subspace depends on the underlying variational formulation of the PDE. In this work, we use a variational formulation based on a hybrid discontinuous Galerkin (HDG) method [ES10], which is well-suited for the Fokker-Planck equation related to the coupled Brown-Néel model for MPI [Wei18]. This Fokker-Planck equation is posed on a manifold, which is the product of two unit spheres and time. Therefore, we have to adapt the variational formulation to this geometry. Moreover, we need to make further modifications to meet the requirements of the RB method, for example, to ensure the affine decomposition of the variational formulation.

There are two different ways to address the time variable. One option is to use a time-stepping scheme; the other option is to use a space-time discretization. A comparison of the two approaches was made in [GMU17]. To avoid problems arising from the convection-dominated nature of the MPI Fokker-Planck equation combined with the RB method, we use a space-time discretization rather than a time-stepping scheme. Moreover, we can achieve better efficiency with the space-time formulation than with a time-stepping scheme. RB methods for parabolic PDEs using space-time formulations have also been investigated in different settings, for example, in [UP12; UP13; Yan14]

We present numerical results for the HDG method with fixed parameter combinations and preliminary results for the RB method with varying parameters. The results indicate that we can achieve good accuracy with a moderate number of snapshots, which is a promising result for the application of the RB method in MPI. This allows model-based image reconstruction in MPI as an alternative to system matrix calibration. Moreover, it enables parameter studies and optimal experimental design in MPI.

1.1 Outline

In Chapter 2, we introduce preliminaries for approximation in high dimensions to tackle the curse of dimensionality. We start by introducing Riesz bases as well as one-dimensional wavelet bases. Correspondingly, we consider nonlinear approximation in sequence spaces. Then we examine tensor-product spaces and the construction of high-dimensional product wavelet bases. We then turn to tensor formats to efficiently represent the coefficient vector of a function in a product wavelet basis, and

we consider wavelet representations of operators and their compressibility properties.

In Chapter 3, we investigate space-time formulations and their corresponding reformulations as bi-infinite matrix equations. One special focus is to show that the condition number of this bi-infinite matrix is independent of the number of spatial dimensions for a suitable choice of basis functions.

In Chapter 4 we propose and analyze an adaptive wavelet low-rank method for parabolic PDEs. One basic idea is to treat time not as a tensor component but as a sparse expansion in time, with low-rank approximations only in the spatial variables. Consequently, we define rank-truncation and coarsening operators for this representation. We investigate the adaptive approximate application of the operators of the space-time formulation. Additionally, we investigate diagonal preconditioning using exponential sums. As the main point, we show that a perturbed version of the Richardson iteration can circumvent the curse of dimensionality under natural assumptions. Finally, we present numerical results for the proposed method.

In Chapter 5, we describe the modeling of MPI and the main challenges concerning the parameter dependence. Moreover, we investigate one specific model for the dynamics of the magnetic nanoparticles based on a coupled system of SDEs and state the corresponding Fokker-Planck equation.

In Chapter 6, we introduce a method to tackle the parameter dependence of the Fokker-Planck equation using RB. We present a suitable variational formulation for the convection-dominated Fokker-Planck equation. Moreover, we discuss properties of finite elements on surfaces. Finally, we present numerical results for the parametric Fokker-Planck equation using the RB method.

In Chapter 7, we summarize the results for both sources of high dimensionality and give an outlook on possible future research directions.

1.2 Notation

In this section, we summarize the notation used throughout this thesis.

We denote the standard sets of positive natural numbers by $\mathbb{N}$ and use $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. We write $A \lesssim B$ if there exists a constant $C > 0$ independent of the relevant quantities such that $A \leq CB$. If it is not clear from the context, we specify the quantities on which the constant C may depend. Furthermore, we write $A \gtrsim B$ if $B \lesssim A$ and $A \eqsim B$ if both $A \lesssim B$ and $A \gtrsim B$. Moreover, for $\mathsf{k}, \mathsf{l} \in \mathbb{N}^n$, we write $\mathsf{k} \leq \mathsf{l}$ if $k_i \leq l_i$ for all $i = 1, \dots, n$.

Let $d \in \mathbb{N}$ and let $\Omega \subset \mathbb{R}^d$ be a domain. Then we denote by $L_p(\Omega)$ the usual Lebesgue spaces for $1 \leq p \leq \infty$ with norm $\|\cdot\|_{L_p(\Omega)}$. For $k \in \mathbb{N}_0$, we denote by $H^k(\Omega) = W_2^k(\Omega)$ the Sobolev spaces of order k with norm $\|\cdot\|_{H^k(\Omega)}$ and seminorm $|\cdot|_{H^k(\Omega)}$. Moreover, for $s \in \mathbb{R}$, we denote by $H^s(\Omega)$ the corresponding Sobolev spaces of fractional order s.

Let M be a manifold in $\mathbb{R}^d$. We denote by ∇_M the surface gradient of M and by div_M the surface divergence on M. Correspondingly, we denote by $L_2(M)$ the space of square-integrable functions on M and by $H^1(M)$ the Sobolev space of order one on M.

As index sets, we use $\vee_1$ for the one-dimensional wavelet bases and $\vee_{\mathsf{x}} = \bigtimes_{i=1}^d \vee_1$ for the full spatial index set in d dimensions. Furthermore, for the temporal variable, we use the index set $\vee_{\mathsf{t}}$, and for the high-dimensional space-time index set, we use $\vee = \vee_{\mathsf{t}} \times \vee_{\mathsf{x}}$.

For a Hilbert space H with inner product $\langle \cdot, \cdot \rangle_H$ and norm $\|\cdot\|_H$, we denote by H' its continuous dual space with dual norm $\|\cdot\|_{H'}$.

For a countable index set $\hat{\vee}$, we denote by $\ell_2(\hat{\vee})$ the space of square-summable sequences with norm $\|\cdot\|_{\ell_2(\hat{\vee})}$. If the index set is clear from the context, we write ℓ_2 and $\|\cdot\|$.

Chapter 2

High-Dimensional Approximation

Approximating high-dimensional functions is a challenging task due to the curse of dimensionality. Throughout this chapter, we introduce several key concepts to address this challenge. Tensor-product wavelet bases are used to represent functions in high-dimensional domains. Additionally, low-rank tensor formats are introduced to efficiently approximate the coefficient tensor of the function.

The outline of the chapter is as follows: In Section 2.1, we introduce wavelets in one dimension and their approximation properties. In Section 2.2, we present nonlinear approximation in sequence spaces, focusing on best N-term approximation and the corresponding approximation classes. Next, in Section 2.3, we provide examples of high-dimensional spaces as tensor products of low-dimensional spaces. Then we discuss tensor-product wavelet bases. In Section 2.4, we examine different low-rank tensor methods, focusing on the hierarchical tensor format. We also explain how to perform operations explicitly in this format. Finally, in Section 2.5, we cover the representation of operators in wavelet coordinates and their compressibility properties.

2.1 Wavelets in One Dimension

In this work, we are particularly interested in high-dimensional problems. To find a suitable basis for their numerical treatment, we begin by introducing wavelet bases in one dimension. Then we extend these bases to high-dimensional tensor-product wavelet bases in Section 2.3.2.

For the one-dimensional wavelet basis, we are generally interested in approximation properties in function spaces. Therefore, we first introduce the concept of a *Riesz basis* in separable Hilbert spaces. This concept provides a relationship between the norms of the coefficients in the basis representation and the norm of the function itself.

Definition 2.1. Let $\mathcal{I}$ be a countable set and H a separable Hilbert space. We call a family of functions $\{\psi_i\}_{i\in\mathcal{I}}$ a Riesz basis of H if the finite linear combinations are dense in H and if there exists constants $c, C > 0$ such that for any finite set $\Lambda \subset \mathcal{I}$, we have

$$c\Big(\sum_{\lambda\in\Lambda} u_\lambda^2\Big)^{\frac{1}{2}} \le \Big\|\sum_{\lambda\in\Lambda} u_\lambda\psi_\lambda\Big\|_H \le C\Big(\sum_{\lambda\in\Lambda} u_\lambda^2\Big)^{\frac{1}{2}} \quad \text{for all } (u_\lambda)_{\lambda\in\Lambda} \subset \mathbb{R}.$$

A special case of a Riesz basis is an orthonormal basis. In this case, the constants are $c = C = 1$.

Let $\{\psi_i\}_{i\in\mathcal{I}}$ be a Riesz basis and let $u, v \in H$ be given with basis expansions

$$u = \sum_{i\in\mathcal{I}} u_i\psi_i, \quad v = \sum_{i\in\mathcal{I}} v_i\psi_i,$$

where we write $\mathbf{u} = (u_i)_{i\in\mathcal{I}}$ and $\mathbf{v} = (v_i)_{i\in\mathcal{I}}$. Then as a direct consequence of the Riesz basis property we have

$$\|u - v\|_H \eqsim \|\mathbf{u} - \mathbf{v}\|_{\ell_2}, \tag{2.1}$$

meaning we have an isomorphism between H and $\ell_2(\mathcal{I})$. Based on this relation, we are able to quantify the approximation errors of functions in H by the approximation of the corresponding coefficient sequences in ℓ_2.

2.1.1 Construction of Wavelets

Now, we move to the construction of wavelets in $\mathbb{R}$. These are of particular interest in this work, as they can be used to construct Riesz bases for different function spaces. At the end of this section, we also comment on how to construct wavelets on intervals. The multiresolution analysis is our starting point. We follow the definition in [Coh03].

Definition 2.2. Let $(V_j)_{j\in\mathbb{Z}}$ be a sequence of closed subspaces of $L_2(\mathbb{R})$. This sequence is called a *multiresolution analysis* if the following properties hold:

(i) The subspaces are nested, i.e., $V_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$.

(ii) $f \in V_j$ if and only if $f(2\cdot) \in V_{j+1}$.

(iii) For all $f \in L_2(\mathbb{R})$ we have

$$\lim_{j\to\infty} \|f - P_j f\|_{L_2} = 0, \quad \lim_{j\to-\infty} \|P_j f\|_{L_2} = 0,$$

where P_j is the L_2-orthogonal projection onto V_j.

(iv) There exists a function $\varphi \in V_0$ such that $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis of V_0.

The function φ in property (iv) is called the *scaling function* of the multiresolution analysis. Based on properties (ii) and (iv), we obtain a basis $\{\varphi_{j,k}\}_{k\in\mathbb{Z}}$ of V_j by dilatation and translation of the scaling function, that is, $\varphi_{j,k} = 2^{j/2}\varphi(2^j \cdot -k)$.

In numerical analysis, we are often interested in orthonormal bases. Hence, in the following, we aim to construct an orthonormal basis of V_j instead of using $\{\varphi_{j,k}\}_{k\in\mathbb{Z}}$. We define W_j as the orthogonal complement of V_j in V_{j+1}, which means $V_{j+1} = V_j \oplus W_j$. Applying this relation recursively, we get the representation

$$V_j = V_0 \oplus \bigoplus_{l=0}^{j-1} W_l.$$

Hence, if we construct orthonormal bases of the spaces V_0 and W_l for $l = 0, \dots, j-1$, then we also obtain an orthonormal basis of V_j.

To achieve this, we define the *mother wavelet* function. First, we state that, by property (i), there exists a sequence of coefficients $\{h_k\}_{k\in\mathbb{Z}}$ such that

$$\varphi = \sum_{k\in\mathbb{Z}} h_k \varphi_{1,k}.$$

We call this relation the *refinement equation*. Based on the refinement equation, we define the mother wavelet by

$$\psi = \sum_{k\in\mathbb{Z}} (-1)^k h_{1-k} \varphi_{1,k}. \tag{2.2}$$

Let the scaling function satisfy the conditions (i)-(iii) in Theorem 2.2 and additionally the orthogonality property of its translates

$$\langle \varphi_{0,k}, \varphi_{0,k'} \rangle = \delta_{k,k'}.$$

Then as an immediate consequence of the refinement equation and the choice of the coefficients in the representation (2.2), the mother wavelet is orthogonal to the scaling function and its translates, since for each $l \in \mathbb{Z}$, we have

$$\langle \varphi_{0,l}, \psi \rangle = \Big\langle \sum_{k\in\mathbb{Z}} h_{k-2l}\varphi_{1,k}, \psi \Big\rangle = \sum_{k\in\mathbb{Z}} (-1)^k h_{1-k} h_{k-2l} = 0,$$

where we used in the last step that the odd and even terms cancel out. Moreover, we obtain

$$\langle \psi(\cdot - l), \psi \rangle = \sum_{k\in\mathbb{Z}} h_{1-k-2l} h_{1-k} = \sum_{k\in\mathbb{Z}} h_k h_{k+2l} = \langle \varphi, \varphi_{0,l} \rangle = \delta_{0,l}.$$

Hence, by Theorem 2.2, the space W_0 is generated by its orthonormal basis $\{\psi(\cdot - k)\}_{k\in\mathbb{Z}}$. Furthermore, the spaces W_j are generated by the orthonormal bases $\{\psi_{j,k}\}_{k\in\mathbb{Z}}$, where $\psi_{j,k} = 2^{j/2}\psi(2^j \cdot - k)$.

Based on the definition of the spaces and property (iii) in Theorem 2.2 we have

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L_2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j=j_0}^{\infty} W_j$$

for each $j_0 \in \mathbb{Z}$. For notational simplicity, we introduce the index set

$$\Lambda_{j_0} = \{(j_0, k, 0) : k \in \mathbb{Z}\} \cup \{(j, k, 1) : j \geq j_0, k \in \mathbb{Z}\}. \tag{2.3}$$

Let $\lambda = (j, k, e) \in \Lambda_{j_0}$, then correspondingly we define the shorthand notation $|\lambda| = j$ for the level of the wavelet function. For $\lambda \in \Lambda_{j_0}$, we set

$$\psi_\lambda = \begin{cases} \varphi_{j_0,k} & \text{if } \lambda = (j_0, k, 0), \\ \psi_{j,k} & \text{if } \lambda = (j, k, 1), \quad j \geq j_0. \end{cases} \tag{2.4}$$

By construction, the set of functions $\{\psi_\lambda\}_{\lambda\in\Lambda_{j_0}}$ form an orthonormal basis of $L_2(\mathbb{R})$. Throughout this work, we assume $j_0 = 0$ and define $\vee_1 = \Lambda_0$ for simplicity.

In this work, we are interested in solving PDEs on bounded domains Ω, in particular on the high-dimensional unit cube $(0,1)^d$. Hence, we need a Riesz basis for $L_2(0,1)$ instead of $L_2(\mathbb{R})$. There are several ways to construct wavelet bases on intervals. One is to restrict wavelet systems on $\mathbb{R}$ and carefully treat the boundary conditions. For details, see [Mey91; CDV93; DKU99] and [Mal09, Section 7.5] for an overview. Alternatively, we can construct a multiresolution directly on the interval using spline functions, as in [CQ92]. In Section 2.3.2, we describe how to construct high-dimensional wavelet bases on the high-dimensional unit cube as tensor products of one-dimensional wavelet bases on the interval.

2.1.2 Approximation Properties

There is an extensive body of literature on the approximation properties of wavelets. In this work, we focus on the approximation properties in Sobolev spaces. As seen

in (2.1), the Riesz basis property is an important tool for quantifying approximation in function spaces.

Theorem 2.3. *Let $\varphi \in H^\tau(\mathbb{R})$ for some $\tau > 0$. Furthermore, let φ be a compactly supported L_2-orthonormal scaling function of a multiresolution analysis. Then $\{2^{-s|\lambda|}\psi_\lambda\}$ forms a Riesz basis of $H^s(\mathbb{R})$ for $-\tau < s < \tau$.*

This statement appears for example in [Bac12, Theorem 3.5] and is a special case of [Coh03, Theorems 3.7.7 and 3.8.1].

Remark 2.4. In the case of bounded intervals, the same result holds true. We will use this statement in particular for the case $s = 1$. The statement is based on the fact that

$$\|\psi_\lambda\|_{H^1}^{-1}\|\psi_\nu\|_{H^1}^{-1}\langle\psi_\lambda,\psi_\nu\rangle_{H^1}$$

decays exponentially in the level distance $||\lambda| - |\nu||$. Moreover, by definition of the wavelet basis, we directly infer that $\|\psi_\lambda\|_{H^1} \eqsim 2^{|\lambda|}$. For a specific choice of wavelets the exponential decay in the level distance will be investigated in more detail in Theorem 2.30.

We use spline wavelets for numerical purposes, so we study the approximation properties of polynomials in Sobolev spaces.

Definition 2.5. A wavelet function ψ has *m vanishing moments* if

$$\int_{\mathbb{R}} x^k\psi(x)\,\mathrm{d}x = 0 \quad \text{for } k = 0,\dots,m-1.$$

The vanishing moment property is of special interest in the numerical treatment of PDEs if we use piecewise polynomial wavelet bases. In the variational formulation, we must integrate the product of two wavelet functions. If the level difference is large, then $\psi_{j,k}$ is polynomial on the support of $\psi_{j',k'}$. Due to the vanishing moment property, the integral may vanish. This results in a sparse matrix representation of the operator.

Theorem 2.6 (Deny-Lions)**.** *Let $1 \le p < \infty$ and let $\Omega \subset \mathbb{R}^d$ be a bounded domain. There is a constant C, depending on Ω, d, p and k, such that for all $f \in H^k(\Omega)$ we have*

$$\inf_{g\in\mathbb{P}_{k-1}} \|f-g\|_{H^k} \le C|f|_{H^k}.$$

This property determines how well a wavelet basis approximates functions with local smoothness. Hence, it shows how rapidly the entries in the basis expansion of a sufficiently regular function decay. We will discuss this further in Theorem 2.9.

For a more general discussion of approximation properties of wavelets, for example, in Besov spaces, we refer to [Coh03; Mal09].

2.1.3 Multiwavelets

As we will discuss in Section 3.1.2 and Section 4.4, the use of L_2-orthonormal bases is crucial for our numerical treatment of high-dimensional PDEs. Additionally, a high number of vanishing moments of the wavelet functions is necessary for the operators to be compressible. One example of such a wavelet that is compactly supported is the Daubechies wavelet [Dau88]. However, another important point is that integrals over the wavelet functions can be evaluated efficiently. Hence, spline basis

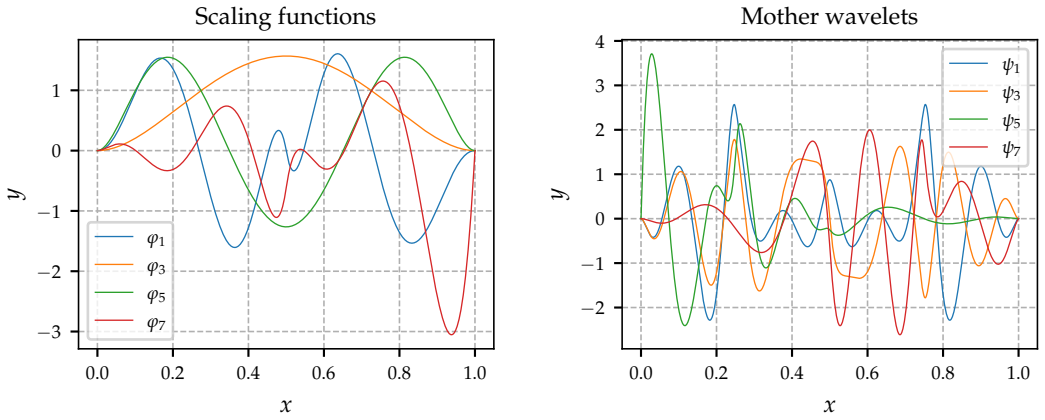


FIGURE 2.1: Selection of scaling functions (left) and mother wavelets (right) of the C^1 piecewise polynomial multiwavelets with Dirichlet boundary conditions from [DGH99]

functions are a good choice for the wavelet functions, since integrals over piecewise polynomials can be evaluated efficiently using quadrature rules. Unfortunately, no spline wavelet functions exist that are L_2-orthonormal, compactly supported and sufficiently regular. However, there exist *multiwavelets* [GHM94] that have these properties.

For multiwavelets, a multiresolution analysis is defined not by a single scaling function, but by a family of scaling functions. The space V_0 is given by the translates of the family

$$\varphi_l, \quad l = 1, \dots, L$$

for some $L \in \mathbb{N}$ with $L > 1$. The corresponding complements are then defined by the span of the L different wavelet functions ψ_l. This yields a basis of the form

$$\{\psi_{l,\nu} : (l, \lambda) \in \{1, \dots, L\} \times \Lambda\},$$

where the functions $\psi_{l,\lambda}$ are defined accordingly to (2.4) for the functions φ_l and ψ_l.

For the construction of compactly supported $L_2(0,1)$-orthonormal, piecewise polynomial multiwavelets we refer to [DGH96; DGH99]. The multiwavelets defined in [DGH96] are piecewise polynomials of degree 1, and the multiwavelets in [DGH99] are C^1-functions which are piecewise polynomials of degree 6. A selection of the scaling functions and mother wavelets of the multiwavelets from [DGH99] is visualized in Figure 2.1.

2.2 Nonlinear Approximation in Sequence Spaces

As seen in (2.1), the approximation of functions with Riesz bases is equivalent to the approximations of sequence spaces up to constants. This allows us to use approximation techniques in sequence spaces. In particular, we are interested in the best N-term approximation, i.e., the approximation of a sequence by another sequence with at most N nonvanishing entries. For a countable index set $\hat{\vee}$, we define the space

$$\Sigma_N = \{\mathbf{v} \in \ell_2(\hat{\vee}) : \#\operatorname{supp}(\mathbf{v}) \leq N\}.$$

Approximations in this set are *nonlinear* because Σ_N is a nonlinear space. In general, for $\mathbf{v}, \mathbf{w} \in \Sigma_N$, we have $\mathbf{v} + \mathbf{w} \in \Sigma_{2N}$. Best N-term approximations are realized by *restriction* operators. For $\Lambda \subset \hat{\vee}$, the restriction operator $\mathrm{R}_\Lambda \mathbf{v}$ equals $\mathbf{v}$ on Λ and zero elsewhere, i.e., on $\hat{\vee} \setminus \Lambda$.

$$(\mathrm{R}_\Lambda \mathbf{v})_\nu = \begin{cases} v_\nu & \text{if } \nu \in \Lambda, \\ 0 & \text{else.} \end{cases} \tag{2.5}$$

Definition 2.7. For any countable index set $\hat{\vee}$ and $s > 0$, we define

$$\|\mathbf{v}\|_{\mathcal{A}^s(\hat{\vee})} = \sup_{N \in \mathbb{N}_0} (N+1)^s \inf_{\substack{\Lambda \subset \hat{\vee} \\ \#\Lambda \le N}} \|\mathbf{v} - \mathrm{R}_\Lambda \mathbf{v}\|_{\ell_2(\hat{\vee})}$$

and the corresponding approximation class

$$\mathcal{A}^s(\hat{\vee}) = \{\mathbf{v} \in \ell_2(\hat{\vee}) : \|\mathbf{v}\|_{\mathcal{A}^s(\hat{\vee})} < \infty\}. \tag{2.6}$$

In what follows, we will write $\|\cdot\| = \|\cdot\|_{\ell_2(\hat{\vee})}$ if the index set $\hat{\vee}$ is clear from the context.

The space $\mathcal{A}^s(\hat{\vee})$, endowed with the quasinorm $\|\cdot\|_{\mathcal{A}^s}$, is a quasi-Banach space. This follows directly from the properties in Theorem 2.8. A quasinorm satisfies all the properties of a norm except that the triangle inequality is replaced by the weaker condition

$$\|\mathbf{v} + \mathbf{w}\|_{\mathcal{A}^s} \le C(\|\mathbf{v}\|_{\mathcal{A}^s} + \|\mathbf{w}\|_{\mathcal{A}^s})$$

for some $C \ge 1$. The specification of the constant C, along with additional properties of the approximation class, is provided in the following proposition.

Proposition 2.8. *For $\mathbf{v}, \mathbf{w} \in \mathcal{A}^s(\hat{\vee})$, we have:*

(i) $\|\mathbf{v}\| \le \|\mathbf{v}\|_{\mathcal{A}^s}$.

(ii) $\|\mathbf{v} + \mathbf{w}\|_{\mathcal{A}^s} \le 2^s(\|\mathbf{v}\|_{\mathcal{A}^s} + \|\mathbf{w}\|_{\mathcal{A}^s})$.

(iii) *If $\#\operatorname{supp}(\mathbf{v}) \le N$, we have $\|\mathbf{v}\|_{\mathcal{A}^s} \le N^s \|\mathbf{v}\|$.*

(iv) *If $\mathbf{v}_N$ is a best N-term approximation of $\mathbf{v}$, then $\|\mathbf{v}\|_{\mathcal{A}^s} \ge (N+1)^s \|\mathbf{v} - \mathbf{v}_N\|$.*

(v) $\|\mathrm{R}_\Lambda \mathbf{v}\|_{\mathcal{A}^s} \le \|\mathbf{v}\|_{\mathcal{A}^s}$ *for any $\Lambda \subset \hat{\vee}$.*

These statements follow directly from the definition of the $\mathcal{A}^s$-quasinorm. For this reason, we omit the proofs here.

To illustrate the ideas of best N-term approximations, we present an example of a function whose coefficient sequence belongs to $\mathcal{A}^s$.

Example 2.9. Assume we have a sufficiently regular one-dimensional function u on a bounded interval I. Furthermore, assume the wavelet function ψ has m vanishing moments as well as $|\operatorname{supp} \psi_\lambda| \approx 2^{-|\lambda|}$ for all functions in the L_2-orthonormal wavelet basis $\{\psi_\lambda\}_{\lambda \in \vee_1}$ and that

$$\#\{\lambda \in \vee_1 : |\lambda| = j\} \approx 2^j, \quad j \in \mathbb{N}.$$

Let $u = \sum_{\lambda \in \vee_1} u_\lambda \psi_\lambda$ for $\mathbf{u} = (u_\lambda)_{\lambda \in \vee_1}$. Using the L_2-orthonormality of the wavelet basis, the vanishing moments, as well as the Cauchy-Schwarz inequality, we obtain

$$|u_\lambda| = \left| \int_I \psi_\lambda(x) u(x) \, \mathrm{d}x \right| \le \inf_{p \in \mathbb{P}_{m-1}} \|u - p\|_{L_2(\operatorname{supp} \psi_\lambda)}.$$

Moreover, using the chain rule and Theorem 2.6, we have

$$|u_\lambda| \lesssim 2^{-|\lambda|m}|u|_{H^m(\operatorname{supp}\psi_\lambda)}.$$

Defining $\Lambda_j = \{\lambda \in \vee : |\lambda| \le j\}$, one can easily check that

$$\|\mathbf{u} - \mathrm{R}_{\Lambda_j}\mathbf{u}\| \lesssim 2^{-mj}.$$

Furthermore, we have $\#\Lambda_j \eqsim 2^{j+1}$. Hence, we obtain

$$\sup_{N\in\mathbb{N}_0}(N+1)^m \inf_{\substack{\Lambda\subset\hat{\vee}\\ \#\Lambda\le N}} \|\mathbf{u} - \mathrm{R}_\Lambda\mathbf{u}\| \lesssim \sup_{j\in\mathbb{N}_0}(\#\Lambda_{j+1}+1)^m\|\mathbf{u} - \mathrm{R}_{\Lambda_j}\mathbf{u}\| < \infty$$

and therefore $\mathbf{u} \in \mathcal{A}^s(\vee)$ for $s = m$.

This example demonstrates that sufficiently regular functions possess coefficient sequences in the approximation class $\mathcal{A}^s$ when the wavelet basis has s vanishing moments, provided certain additional assumptions are met as the L_2-orthonormality of the wavelet basis and the exponential decay of the support sizes in the level.

2.3 Tensor-Product Spaces

One way to address the curse of dimensionality is to use low-rank approximations. For this, we need a space that can be represented by lower-dimensional spaces. A common example is the Lebesgue space L_2 on the high-dimensional unit cube $(0,1)^d$, where we have

$$L_2((0,1)^d) \cong \bigotimes_{i=1}^{d} L_2(0,1) \tag{2.7}$$

which is a product of one-dimensional function spaces.

2.3.1 Hilbert Spaces

In this work, we are particularly interested in approximations in Hilbert spaces on high-dimensional domains. With this motivation, in the following section, we discuss the properties of tensor-product Hilbert spaces that are of interest to us. For this purpose, we follow the works of [Hac19; Bac23].

Tensor-product spaces can be defined for general Hilbert spaces. Here, we define the tensor-product space $H_1 \otimes H_2$ of the Hilbert spaces H_1 and H_2 as the closure of the sum of elementary tensors

$$\sum_{i=1}^{n} v_i \otimes w_i, \quad v_i \in H_1, w_i \in H_2$$

with respect to the canonical inner product

$$\langle v_1 \otimes w_1, v_2 \otimes w_2\rangle_{H_1\otimes H_2} = \langle v_1, v_2\rangle_{H_1}\langle w_1, w_2\rangle_{H_2}.$$

The corresponding canonical norm on this inner product is a cross norm, that is

$$\|v_1 \otimes w_1\|_{H_1\otimes H_2} = \|v_1\|_{H_1}\|w_1\|_{H_2}.$$

For details on the connection of algebraic and topological tensor products, we refer to [Hac19, Chapter 5].

Next, we consider tensor operators on tensor-product spaces. Let G_1 and G_2 be Hilbert spaces. Let $A_1 : H_1 \to G_1$ and $A_2 : H_2 \to G_2$ be bounded operators. Then the tensor operator $A_1 \otimes A_2 : H_1 \otimes H_2 \to G_1 \otimes G_2$ is defined on elementary tensors by

$$(A_1 \otimes A_2)\Big(\sum_{i=1}^{n} v_i \otimes w_i\Big) = \sum_{i=1}^{n} (A_1 v_i) \otimes (A_2 w_i)$$

and on general elements by the unique extension of this definition to the closure of the elementary tensors.

Theorem 2.10. *Let H_1, H_2, G_1 and G_2 be separable Hilbert spaces and $A_1 : H_1 \to G_1$ and $A_2 : H_2 \to G_2$ be bounded linear operators. Then the tensor operator $A_1 \otimes A_2 : H_1 \otimes H_2 \to G_1 \otimes G_2$ is a bounded linear operator with*

$$\|A_1 \otimes A_2\|_{H_1 \otimes H_2 \to G_1 \otimes G_2} = \|A_1\|_{H_1 \to G_1} \|A_2\|_{H_2 \to G_2}$$

If A_1 and A_2 are isomorphisms, then $A_1 \otimes A_2$ is an isomorphism as well with the inverse $A_1^{-1} \otimes A_2^{-1}$.

For a proof of the first statement, we refer to [LC85, Lemma 1.30] and for the second statement to [SU09, Lemma B.1].

For approximation purposes, we are interested in Riesz bases. The following theorem shows how to construct a Riesz basis of a tensor-product space from Riesz bases of single separable Hilbert spaces, i.e., Hilbert spaces with a countable basis.

Theorem 2.11. *Let H_1 and H_2 be separable Hilbert spaces, $\hat{\vee}_1$ and $\hat{\vee}_2$ be countable index sets and $\{\psi_\nu^{(1)}\}_{\nu \in \hat{\vee}_1}$ and $\{\psi_\nu^{(2)}\}_{\nu \in \hat{\vee}_2}$ be Riesz bases of H_1 and H_2, respectively, with constants c_1, c_2, C_1, C_2. Then*

$$\{\psi_{\nu_1}^{(1)} \otimes \psi_{\nu_2}^{(2)}\}_{\nu_1 \in \hat{\vee}_1, \nu_2 \in \hat{\vee}_2} \tag{2.8}$$

is a Riesz basis of the tensor-product space $H_1 \otimes H_2$ with constants $c = c_1 c_2$ and $C = C_1 C_2$.

Proof. This statement is an immediate consequence of Theorem 2.10 applied to the bounded linear operators $A_i : \ell_2(\hat{\vee}_i) \to H_i$ with

$$A_i((c_\nu)_{\nu \in \hat{\vee}_i}) = \sum_{\nu \in \hat{\vee}_i} c_\nu \psi_\nu^{(i)}, \quad i = 1, 2,$$

where $\|A_i\| \le C_i$ and $\|A_i\|^{-1} \le c_i^{-1}$. □

As a direct consequence of this theorem, we have that the tensor-product basis as in (2.8) constructed by two orthonormal bases is an orthonormal basis of the tensor-product space $H_1 \otimes H_2$.

This construction can be naturally extended to tensor-product spaces of d Hilbert spaces $H_1 \otimes \cdots \otimes H_d$ with the corresponding canonical inner product

$$\langle v_1 \otimes \cdots \otimes v_d, w_1 \otimes \cdots \otimes w_d \rangle_{H_1 \otimes \cdots \otimes H_d} = \prod_{i=1}^{d} \langle v_i, w_i \rangle_{H_i}$$

and the corresponding cross norm

$$\|v_1 \otimes \cdots \otimes v_d\|_{H_1 \otimes \cdots \otimes H_d} = \prod_{i=1}^{d} \|v_i\|_{H_i}.$$

Building on the general discussion above, we now specify the Hilbert spaces of particular interest in this work: the sequence space ℓ_2 and the function spaces L_2 and H^1.

Theorem 2.12. *Let $\Omega_1, \dots \Omega_d \subset \mathbb{R}$ and $\hat{\vee}_1, \dots, \hat{\vee}_d$ be countable index sets. Then*

$$L_2(\Omega_1) \otimes \cdots \otimes L_2(\Omega_d) \cong L_2(\Omega_1 \times \cdots \times \Omega_d)$$
$$\ell_2(\hat{\vee}_1) \otimes \cdots \otimes \ell_2(\hat{\vee}_d) \cong \ell_2(\hat{\vee}_1 \times \cdots \times \hat{\vee}_d).$$

For a proof of the first statement, we refer to [LC85, Lemma 1.39] and for the second statement to [Hac19, Remark 4.43].

In contrast to the two previous Hilbert spaces, the Sobolev space $H^1(\Omega)$ with $\Omega = \Omega_1 \times \cdots \times \Omega_d$ cannot be characterized as a tensor-product space. However, the space can be represented as intersections of d tensor-product spaces as

$$H^1(\Omega) \cong H^1(\Omega_1) \otimes L_2(\Omega_2) \otimes \cdots \otimes L_2(\Omega_d) \cap \cdots \cap L_2(\Omega_1) \otimes \cdots \otimes L_2(\Omega_{d-1}) \otimes H^1(\Omega_d)$$

with the canonical intersection norm

$$\|v_1 \otimes \cdots \otimes v_d\|_{H^1(\Omega)} = \Big(\sum_{i=1}^{d} \|v_i\|^2_{H^1(\Omega_i)} \prod_{j \neq i} \|v_j\|^2_{L_2(\Omega_j)} \Big)^{\frac{1}{2}}. \tag{2.9}$$

An analogous statement is true for the Sobolev space $H^1_0(\Omega)$ with Dirichlet boundary conditions.

In certain situations, a Riesz basis for $H^1(\Omega)$ can still be constructed by forming a product basis of $L_2(\Omega)$ and renormalizing it in $H^1(\Omega)$. One example are tensor-product wavelet bases, with sufficiently regular scaling functions. For details, we refer to [Bac12, Remark 3.16]. For this construction, one challenge is that the $H^1(\Omega)$-norm is not a cross norm, due to the intersection structure in (2.9). Often, this is handled by approximating the scaling factor with exponential sums. This is analyzed further in Section 4.2.

2.3.2 Tensor-Product Wavelet Bases

Because a separable structure of high-dimensional basis functions is crucial for our treatment of high-dimensional PDEs, we proceed by considering high-dimensional wavelet bases as tensor products of low-dimensional wavelet bases, in particular, as tensor products of one-dimensional wavelet bases.

One possible construction of tensor-product wavelet bases is directly via a multiresolution analysis, as in Theorem 2.2. For each high-dimensional wavelet function, we have a fixed level j, a vector of translation indices $k \in \mathbb{Z}^d$, and a binary vector $e \in \{0,1\}^d$ which encodes if the function is a scaling function or a wavelet function. The corresponding high-dimensional wavelet function is then given by

$$\Psi_{j,k,e} = \bigotimes_{i=1}^{d} \varphi_{j,k_i,e_i}.$$

In this case, the low-dimensional wavelet functions all have the same level j. Hence, the basis function $\Psi_{j,k,e}$ has isotropic support, meaning $\operatorname{diam}(\operatorname{supp}(\Psi_{j,k,e})) \eqsim 2^{-j}$ and $|\operatorname{supp}(\Psi_{j,k,e})| \eqsim 2^{-dj}$. In this approach, low-dimensional scaling functions are used at higher levels, unlike in the one-dimensional case.

Another approach is based on the recursive application of Theorem 2.11. For simplicity, we assume that our low-dimensional spaces are all the same. So, let $\hat{\vee} =$

$\times_{i=1}^{d} \hat{\vee}_1$ and represent $\nu \in \hat{\vee}$ by a vector of levels, a vector of dilations, and an encoding vector, we get

$$\Psi_{j,k,e} = \bigotimes_{i=1}^{d} \varphi_{j_i,k_i,e_i}, \quad (j_i, k_i, e_i) \in \hat{\vee}_1.$$

In this case, scaling functions are only used at the lowest level, i.e., $e_i = 0$ only for $j_i = 0$. In this case, the basis function support is anisotropic, meaning the support size may differ in each spatial direction.

2.4 Low-Rank Tensor Formats

Let u be a finitely supported function on a high-dimensional domain represented with a tensor-product wavelet basis by

$$u = \sum_{\nu_1 \in \Lambda_1} \cdots \sum_{\nu_d \in \Lambda_d} u_{\nu_1,\ldots,\nu_d} \psi_{\nu_1}^{(1)} \otimes \cdots \otimes \psi_{\nu_d}^{(d)}$$

with $\#\Lambda_i = n$ for $i = 1, \ldots, d$ and $n \geq 2$. Then the coefficient tensor $\mathbf{u} = (u_{\nu_1,\ldots,\nu_d})$ has n^d entries. Hence, the storage costs for the naive representation of the coefficient tensor grow exponentially in the number of dimensions d. In this section, we introduce low-rank tensor formats for efficiently representing the coefficient tensor $\mathbf{u}$.

As an introduction, we start with low-rank approximation in two dimensions. Let $\mathbf{u} \in \ell_2(\hat{\vee}_1 \times \hat{\vee}_2)$ be possibly infinitely supported. Then based on the Hilbert-Schmidt decomposition (or singular-value decomposition in the finite case), there exist orthonormal systems $\{\mathbf{U}_k^{(1)}\}_{k \in \hat{\vee}_1}$, $\{\mathbf{U}_k^{(2)}\}_{k \in \hat{\vee}_2}$ and a non-negative non-increasing sequence $(\sigma_i)_{i \in \mathbb{N}} \in \ell_2(\mathbb{N})$ such that

$$\mathbf{u} = \sum_{k=1}^{\infty} \sigma_i \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)}. \tag{2.10}$$

The values σ_i are called *singular values*. If there exists a $j \in \mathbb{N}$ with $\sigma_j = 0$, then $\sigma_i = 0$ for all $i \geq j$ and the sum in (2.10) is finite. We call r the *rank* of $\mathbf{u}$ if r is the maximal value with $\sigma_r \neq 0$. If $\sigma_i > 0$ for all $i \in \mathbb{N}$, then we say that $\operatorname{rank}(\mathbf{u}) = \infty$.

The decomposition (2.10) directly provides us the following best-approximation property, which is an extension of the Eckart-Young theorem for the finite case

$$\Big\| \mathbf{u} - \sum_{k=1}^{r} \sigma_i \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)} \Big\| = \min_{\operatorname{rank}(\mathbf{v}) \leq r} \|\mathbf{u} - \mathbf{v}\|. \tag{2.11}$$

Hence, truncation of the sum in (2.10) yields the best approximation over all sequences of rank r. Furthermore, based on the orthogonality, the error can be directly quantified by

$$\Big\| \mathbf{u} - \sum_{k=1}^{r} \sigma_i \mathbf{U}_k^{(1)} \otimes \mathbf{U}_k^{(2)} \Big\| = \Big(\sum_{k=r+1}^{\infty} \sigma_i^2 \Big)^{\frac{1}{2}}. \tag{2.12}$$

Next we discuss possible formats for $d > 2$ and their advantages and disadvantages.

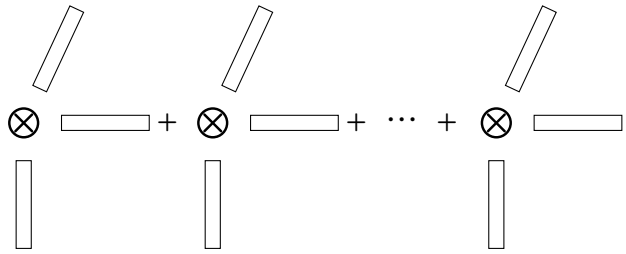


FIGURE 2.2: Graphical representation of the canonical polyadic format in three dimensions.

2.4.1 Canonical Tensor Format

The natural extension of the two-dimensional case to $d > 2$ is the canonical polyadic format, which is a representation of the form

$$\mathbf{u} = \sum_{k=1}^{r} a_k \mathbf{U}_k^{(1)} \otimes \cdots \otimes \mathbf{U}_k^{(d)}$$

with $a_k \in \mathbb{R}$ and $\|\mathbf{U}_k^{(i)}\| = 1$. The smallest r for which such a decomposition of $\mathbf{u}$ exists is called the *canonical rank* of $\mathbf{u}$. An example of the canonical polyadic format in three dimensions is illustrated in Figure 2.2.

A major drawback of this format is that the best approximation problem is ill-posed. For matrices, we generally have the following property. Let $\mathbf{M}$ be a matrix and $(\mathbf{M}_l)_{l\in\mathbb{N}}$ a sequence of matrices with $\mathbf{M}_l \to \mathbf{M}$. Then we have

$$\liminf_{l\to\infty} \operatorname{rank}(\mathbf{M}_l) \geq \operatorname{rank}(\mathbf{M}),$$

which is a direct consequence of (2.11) and (2.12). However, this property does not carry over to higher dimensions, as the following example for tensors of order three demonstrates.

Example 2.13. Let

$$\mathbf{u} = \mathbf{v} \otimes \mathbf{v} \otimes \mathbf{w} + \mathbf{v} \otimes \mathbf{w} \otimes \mathbf{v} + \mathbf{w} \otimes \mathbf{v} \otimes \mathbf{v}$$

with $\|\mathbf{v}\| = \|\mathbf{w}\| = 1$ and $|\langle \mathbf{v}, \mathbf{w} \rangle| \neq 1$, as well as

$$\mathbf{u}_n = n\big(\mathbf{v} + \tfrac{1}{n}\mathbf{w}\big) \otimes \big(\mathbf{v} + \tfrac{1}{n}\mathbf{w}\big) \otimes \big(\mathbf{v} + \tfrac{1}{n}\mathbf{w}\big) - n\mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v}.$$

Then one can easily check that $\mathbf{u}_n \to \mathbf{u}$, but $\operatorname{rank}(\mathbf{u}_n) = 2 < 3 = \operatorname{rank}(\mathbf{u})$ for all $n \in \mathbb{N}$. For details, we refer to [Hac19, Proposition 9.10].

The set of tensors with $\operatorname{rank}(\mathbf{u}) \leq k$ is generally not closed. Thus, as in the example above, specific tensors can be approximated arbitrarily well by tensors of smaller rank, but a minimum does not exist. This smaller rank is called the border rank of the tensor.

2.4.2 Tucker Format

To circumvent the ill-posedness of the approximation problem, we trace the problem back to the matrix case. Therefore, we apply the singular value decomposition (SVD) to *matricizations* of the tensor. Matricizations are defined by re-interpreting tensors

as matrices. For $\mathbf{u} \in \ell_2(\mathbb{N}^d)$, let

$$\mathrm{mat}_{\{i\}}(\mathbf{u}) = (u_\nu)_{\nu_i,(\nu_1,\dots,\nu_{i-1},\nu_{i+1},\dots,\nu_d)} \in \ell_2(\mathbb{N}) \otimes \ell_2(\mathbb{N}^{d-1})$$

with the natural interpretation for $i = 1$ and $i = d$. If we perform an SVD of these matricizations, we get

$$\mathrm{mat}_{\{i\}}(\mathbf{u}) = \sum_{k_i=1}^{r_i} \sigma_{k_i}^{(i)} \mathbf{U}_{k_i}^{(i)} \otimes \mathbf{V}_{k_i}^{(i)}.$$

Combining all matricizations yields a representation

$$\mathbf{u} = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} a_{\mathsf{k}} \mathbf{U}_{k_1}^{(1)} \otimes \cdots \otimes \mathbf{U}_{k_d}^{(d)}$$

with

$$a_{\mathsf{k}} = a_{k_1,\dots,k_d} = \left\langle \mathbf{u}, \mathbf{U}_{k_1}^{(1)} \otimes \cdots \otimes \mathbf{U}_{k_d}^{(d)} \right\rangle.$$

We call $\mathsf{r} = (r_1, \dots, r_d)$ the *multilinear rank*. In general, for infinitely supported vectors, the ranks can be infinite, so that $\mathsf{r} \in (\mathbb{N}_0 \cup \{\infty\})^d$.

Due to the definition via the matricizations, the best approximation

$$\min\left\{\|\mathbf{u} - \mathbf{v}\| : \mathrm{rank}(\mathrm{mat}_{\{i\}}(\mathbf{v})) \le s_i \text{ for all } i = 1, \dots, d\right\}$$

exists for all $\mathsf{s} = (s_1, \dots, s_d)$. Let $\mathsf{s} \le \mathsf{r}$, where the inequality is understood componentwise. Then by truncating to multilinear rank s by

$$\mathcal{T}_s(\mathbf{u}) = \sum_{k_1=1}^{s_1} \cdots \sum_{k_d=1}^{s_d} a_{\mathsf{k}} \mathbf{U}_{k_1}^{(1)} \otimes \cdots \otimes \mathbf{U}_{k_d}^{(d)} \tag{2.13}$$

we get the *quasi-optimality*

$$\begin{aligned}
\|\mathbf{u} - \mathcal{T}_{\mathsf{s}}(\mathbf{u})\|^2 &\le \|\mathbf{u} - \mathcal{T}_{s_1,r_2,\dots,r_d}(\mathbf{u})\|^2 + \cdots + \|\mathbf{u} - \mathcal{T}_{r_1,\dots,r_{d-1},s_d}(\mathbf{u})\|^2 \\
&= \sum_{i=1}^{d} \sum_{k_i=s_i+1}^{r_i} (\sigma_{k_i}^{(i)})^2 \\
&= \sum_{i=1}^{d} \min\left\{\|\mathbf{u} - \mathbf{v}\| : \mathrm{rank}(\mathrm{mat}_{\{i\}}(\mathbf{v})) \le s_i, \mathrm{rank}(\mathrm{mat}_{\{j\}}(\mathbf{v})) \le r_j \text{ for } i \ne j\right\} \\
&\le \sum_{i=1}^{d} \min\left\{\|\mathbf{u} - \mathbf{v}\| : \mathrm{rank}(\mathrm{mat}_{\{j\}}(\mathbf{v})) \le s_j \text{ for all } j = 1, \dots, d\right\} \\
&= d \min\left\{\|\mathbf{u} - \mathbf{v}\| : \mathrm{rank}(\mathrm{mat}_{\{j\}}(\mathbf{v})) \le s_j \text{ for all } j = 1, \dots, d\right\}.
\end{aligned}$$

With this truncation strategy, we no longer obtain the best approximation. Still, we do get an optimal approximation up to the constant $C = \sqrt{d}$, which is sufficient for our purposes.

However, the format still has a significant disadvantage. The *core-tensor* $\mathbf{a} = (a_{\mathsf{k}})_{\mathsf{k}}$ in (2.13) has size $\prod_{i=1}^{d} r_i$, which grows exponentially in the number of dimensions even for fixed ranks. Hence, we still suffer from the curse of dimensionality. This can only be circumvented if only a few ranks are larger than one, which is a very restrictive assumption in practice.

Figure 2.3 illustrates an example for the representation of the Tucker-format in three dimensions.

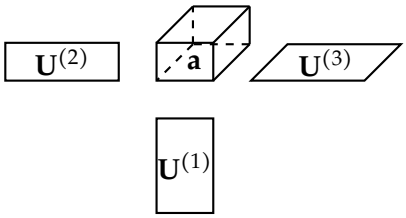


FIGURE 2.3: Graphical representation of the Tucker-format in three dimensions.

2.4.3 Hierarchical Tensor Format

To tackle the curse of dimensionality in the core tensor of the Tucker format, we introduce the hierarchical tensor format. For this introduction, we mainly follow [HK09; BS17; Hac19].

To define hierarchical tensors, we need general matricizations. For an index set $\alpha \subset \alpha^* = \{1, \dots, d\}$, we define the complement $\alpha^c = \alpha^* \setminus \alpha$. The corresponding matricization for $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ is given by

$$\mathrm{mat}_\alpha(\mathbf{u}) = (\mathbf{u}_\nu)_{\nu_\alpha \in \mathbb{N}^{\#\alpha}, \nu_{\alpha^c} \in \mathbb{N}^{d-\#\alpha}},$$

where $\nu_\beta = (\nu_i)_{i \in \beta}$ for $\beta \subset \alpha^*$. We define the corresponding rank by $\mathrm{rank}_\alpha(\mathbf{u}) = \mathrm{rank}(\mathrm{mat}_\alpha(\mathbf{u}))$. Furthermore, we define for $\alpha \subset \alpha^*$ and $\mathbf{W} \colon \ell_2(\mathbb{N}^{\#\alpha}) \to \ell_2(\mathbb{N}^{\#\alpha})$ the α-mode product

$$\mathbf{u} \times_\alpha \mathbf{W} = \mathrm{mat}_\alpha^{-1}(\mathbf{W}\,\mathrm{mat}_\alpha(\mathbf{u})). \tag{2.14}$$

The hierarchical tensor format is based on a binary dimension tree $\mathbb{T}_d$, which is a subset of the power set of α^* of the following form. We start with the root element $\alpha^* \in \mathbb{T}_d$ and divide recursively. For each $\alpha \in \mathbb{T}_d$ with $\#\alpha > 1$, there exists precisely two children $\alpha_1, \alpha_2 \in \mathbb{T}_d$ with $\alpha = \alpha_1 \cup \alpha_2$. If $\#\alpha_i > 1$, we continue the division until all nodes are singletons. Hence, we have $\{1\}, \dots, \{d\} \in \mathbb{T}_d$. Additionally, $\mathbb{T}_d$ has a tree structure, because for $\alpha_1, \alpha_2 \in \mathbb{T}_d$ with $\alpha_1 \neq \alpha_2$, we have $\alpha_1 \subset \alpha_2$ or $\alpha_2 \subset \alpha_1$ or $\alpha_1 \cap \alpha_2 = \emptyset$.

Definition 2.14. Let $\mathbb{T}_d$ be a fixed binary tree.

(i) The set of non-leaf nodes is defined by

$$\mathcal{N}(\mathbb{T}_d) = \{\alpha \in \mathbb{T}_d : \#\alpha > 1\}.$$

(ii) The *effective edges* of $\mathbb{T}_d$ are defined by

$$\mathbb{E}_d = \{\{\alpha, \alpha^c\} : \alpha \in \mathbb{T}_d \setminus \{\alpha^*\}\}.$$

Let $e = \{\alpha, \beta\} \in \mathbb{E}_d$, then we define the representative $[e] = \alpha$ if $\alpha \in \mathbb{T}_d$ and $[e] = \beta$ otherwise. If $\alpha, \beta \in \mathbb{T}_d$, we make an arbitrary choice.

(iii) For $\mathsf{r} = (r_e)_{e \in \mathbb{E}_d} \in (\mathbb{N}_0 \cup \{\infty\})^{\#\mathbb{E}_d}$ let

$$\mathcal{H}(\mathsf{r}) = \{\mathbf{u} \in \ell_2(\mathbb{N}^d) : \mathrm{rank}_{[e]}(\mathbf{u}) \leq r_e \text{ for all } e \in \mathbb{E}_d\} \tag{2.15}$$

the set of all tensors with rank at most r.

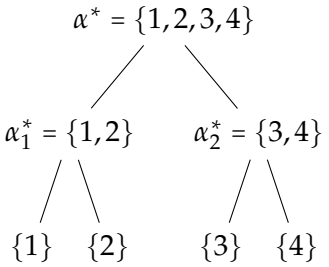


FIGURE 2.4: Example of a binary dimension tree $\mathbb{T}_4$.

(iv) Additionally, we define the vector of matricization ranks given by the binary tree by

$$\operatorname{rank}(\mathbf{u}) = (\operatorname{rank}_{[e]}(\mathbf{u}))_{e\in\mathbb{E}_d} \in \mathcal{R} \tag{2.16}$$

with $\mathcal{R} = (\mathbb{N}_0 \cup \{\infty\})^{\#\mathbb{E}_d}$.

As a direct consequence of the requirements of $\mathbb{T}_d$, the dimension tree has $2d-1$ nodes with d leaf nodes and $\#\mathcal{N}(\mathbb{T}_d) = d-1$ non-leaf nodes. Moreover, by the definition of the effective edges, we obtain $\#\mathbb{E}_d = 2d-3$, since the root node is not counted and the children of the root node are only counted once. An example for a dimension tree for $d = 4$ is illustrated in Figure 2.4.

Based on the Hilbert-Schmidt decomposition, for each $\alpha \in \mathbb{T}_d \setminus \{\alpha^*\}$ there exists an orthonormal basis $\{\mathbf{U}_k^\alpha\}_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{u})}$ of $\overline{\operatorname{range}(\operatorname{mat}_\alpha(\mathbf{u}))}$. These orthonormal bases are called *mode frames*. For notational simplicity, we write

$$\mathbf{U}_k^{(i)} = \mathbf{U}_k^{\{i\}} \quad \text{for } i = 1,\dots,d.$$

Let $\alpha \in \mathcal{N}(\mathbb{T}_d) \setminus \{\alpha^*\}$ with children $\alpha_1, \alpha_2 \in \mathbb{T}_d$. Then we have

$$\mathbf{U}_k^\alpha = \sum_{l_1=1}^{\operatorname{rank}_{\alpha_1}(\mathbf{u})} \sum_{l_2=1}^{\operatorname{rank}_{\alpha_2}(\mathbf{u})} \mathbf{b}_{k,l_1,l_2}^\alpha \mathbf{U}_{l_1}^{\alpha_1} \otimes \mathbf{U}_{l_2}^{\alpha_2}, \tag{2.17}$$

where the three-dimensional transfer tensor $\mathbf{b}^\alpha$ is given by

$$\mathbf{b}_{k,l_1,l_2}^\alpha = \langle \mathbf{U}_k^\alpha, \mathbf{U}_{l_1}^{\alpha_1} \otimes \mathbf{U}_{l_2}^{\alpha_2} \rangle.$$

This property is based on the nestedness property

$$\overline{\operatorname{range}\operatorname{mat}_\alpha(\mathbf{u})} \subseteq \overline{\operatorname{range}\operatorname{mat}_{\alpha_1}(\mathbf{u})} \otimes \overline{\operatorname{range}\operatorname{mat}_{\alpha_2}(\mathbf{u})}.$$

For details on this property, we refer to [Hac19, Section 11.2.2]. Additionally, by the nestedness property, we have the rank restriction

$$\operatorname{rank}_\alpha(\mathbf{u}) \le \operatorname{rank}_{\alpha_1}(\mathbf{u}) \operatorname{rank}_{\alpha_2}(\mathbf{u}).$$

Now we collected all prerequisites to describe the hierarchical tensor format. We start with a decomposition of the root node

$$\mathbf{u} = \sum_{l_1=1}^{\operatorname{rank}_{\alpha_1^*}(\mathbf{u})} \sum_{l_2=1}^{\operatorname{rank}_{\alpha_2^*}(\mathbf{u})} \mathbf{b}_{l_1,l_2}^{\alpha^*} \mathbf{U}_{l_1}^{\alpha_1^*} \otimes \mathbf{U}_{l_2}^{\alpha_2^*}, \tag{2.18}$$

where α_1^* and α_2^* are the children of the root node. Applying the decomposition

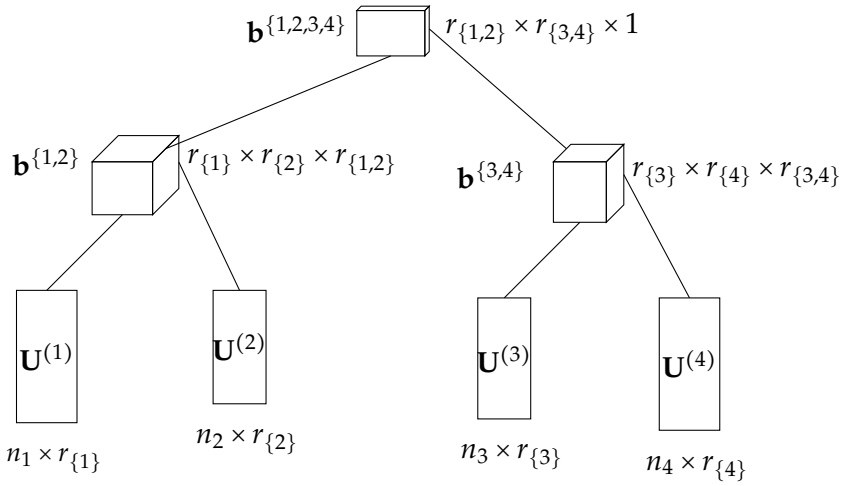


FIGURE 2.5: Representation of a hierarchical tensor in four dimensions with a balanced dimension tree with mode frame sizes n_i for $i = 1, \dots, 4$ and ranks r_e for $e \in \mathbb{E}_4$.

(2.17) recursively to α_1^*, α_2^*, and to their children, we get a representation of $\mathbf{u}$ by the set of all mode frames $\{\mathbf{U}_k^{\{i\}}\}_{k=1,\dots,\operatorname{rank}_{\{i\}}(\mathbf{u})}$ for $i = 1, \dots, d$ and the set of all transfer tensors $\mathbf{b}^\alpha$ for $\alpha \in \mathcal{N}(\mathbb{T}_d)$.

Example 2.15. For $d = 4$ and the balanced dimension tree in Figure 2.4, we have for $r_{1,2} = \operatorname{rank}_{\{1,2\}}(\mathbf{u})$, $r_{3,4} = \operatorname{rank}_{\{3,4\}}(\mathbf{u})$ and $r_i = \operatorname{rank}_{[i]}(\mathbf{u})$ for $i = 1, \dots, 4$ the decomposition

$$\mathbf{u} = \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_2} \sum_{k_3=1}^{r_3} \sum_{k_4=1}^{r_4} \sum_{l_1=1}^{r_{1,2}} \sum_{l_2=1}^{r_{3,4}} \mathbf{b}_{l_1,l_2}^{\{1,2,3,4\}} \mathbf{b}_{l_1,k_1,k_2}^{\{1,2\}} \mathbf{b}_{l_2,k_3,k_4}^{\{3,4\}} \mathbf{U}_{k_1}^{(1)} \otimes \mathbf{U}_{k_2}^{(2)} \otimes \mathbf{U}_{k_3}^{(3)} \otimes \mathbf{U}_{k_4}^{(4)}.$$

Figure 2.5 shows a corresponding graphical representation of this decomposition.

Remark 2.16. In general, one can also specify tensors in a hierarchical format, where the mode frames are not orthogonal, leading to a potentially redundant representation in comparison to the orthogonal representation. As we will see in the next section, these representations naturally arise as intermediate results when performing operations in hierarchical tensor format. In this case, we denote the mode sizes of the transfer tensors as *representation ranks*, which may in general exceed the hierarchical ranks.

Remark 2.17. The storage costs, i.e., the number of parameters to be stored to represent a hierarchical tensor $\mathbf{u}$ of size $n_1 \times \cdots \times n_d$ with ranks $r_\alpha = \operatorname{rank}_\alpha(\mathbf{u})$, is given by

$$\sum_{\alpha \in \mathcal{N}(\mathbb{T}_d)} r_\alpha r_{\alpha_1} r_{\alpha_2} + \sum_{i=1}^{d} r_{\{i\}} n_i,$$

where α_1 and α_2 are the children of α. If the ranks are relatively small, these storage costs are a tremendous reduction compared to the full tensor with $\prod_{i=1}^d n_i$ entries. In comparison to the Tucker format, we do not suffer from the curse of dimensionality in the core tensor, since the transfer tensors are only three-dimensional for all $\alpha \in \mathbb{T}_d \setminus \{\alpha^*\}$. However, we need additional conditions on the matricizations.

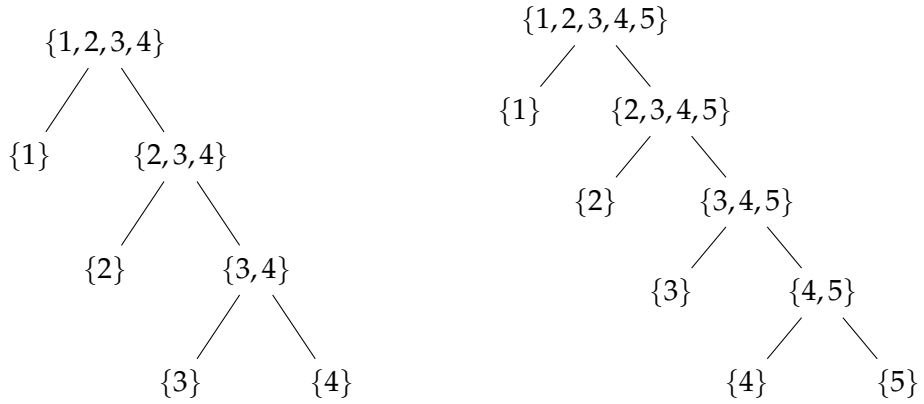


FIGURE 2.6: Example of linear dimension trees $\mathbb{T}_d$ for dimensions $d = 4$ and $d = 5$.

Remark 2.18. A special case is given by hierarchical tensors with linear dimension trees

$$\mathbb{T}_d = \{\alpha^*\} \cup \bigcup_{i=1}^{d-1} \big\{\{i\}, \{i+1, \dots, d\}\big\}. \tag{2.19}$$

Figure 2.6 illustrates examples of such dimension trees for dimensions $d = 4$ and $d = 5$. In this case, the hierarchical tensor format is closely related to the tensor train format, which is also known as matrix product states. For a comparison of the notions of these formats, we refer to [GH11]. We refer to [Ose11] for an introduction to tensor trains and to [Vid03] for an introduction to matrix product states.

2.4.4 Hierarchical Singular Value Decomposition

By Theorem 2.17, it is important to keep the ranks of hierarchical tensors as low as possible. Hence, we need a way to truncate hierarchical tensors to lower ranks. For the practical realization of this truncation, we need our hierarchical tensor to be in a specific form.

Definition 2.19. A hierarchical tensor $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ is a *hierarchical singular value decomposition (HSVD)* if the mode frames $\{\mathbf{U}_k^\alpha\}_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{u})}$ are the left singular vectors of the matricization $\operatorname{mat}_\alpha(\mathbf{u})$ such that the corresponding vector of singular values $(\sigma_k^\alpha(\mathbf{u}))_{k=1,\dots,\operatorname{rank}_\alpha(\mathbf{u})}$ is non-decreasing. Again, for notational convenience we write $\sigma_k^{(i)}(\mathbf{u}) = \sigma_k^{\{i\}}(\mathbf{u})$ for $i = 1, \dots, d$.

The existence of such a hierarchical representation for all tensors in $\ell_2(\mathbb{N}^d)$ is guaranteed by the Hilbert-Schmidt decomposition as in (2.10).

In [Gra10], a procedure for transforming a given hierarchical tensor with orthonormal mode frames into an HSVD representation was introduced. The computational complexity for performing this transformation to a finitely supported hierarchical tensor $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ with mode sizes $n_1 \times \cdots \times n_d$ is given by

$$\mathcal{O}\Big(d \max_{\alpha \in \mathbb{T}_d} \operatorname{rank}_\alpha(\mathbf{u})^4 + \sum_{i=1}^{d} \operatorname{rank}_{\{i\}}(\mathbf{u})^2 n_i\Big). \tag{2.20}$$

For more details on the procedure as well as the computational complexity, we refer to [Gra10; Hac19].

2.4.5 Operations in Hierarchical Tensor Format

For our method, we focus exclusively on hierarchical tensor representations. Hence, we need to be able to perform several operations directly in the hierarchical tensor format representation. For an overview of operations in hierarchical tensor format, we refer to [Tob12; Hac19]. In particular, we need to be able to perform scalar multiplications, add hierarchical tensors, apply operators in hierarchical format, perform rank truncation, and compute norms. For the latter, based on (2.18), we have, by the orthogonality of the mode frames, the property

$$\|\mathbf{u}\|_{\ell_2} = \left\|\mathbf{b}^{\alpha^*}\right\|_{\ell_2}.$$

If $\mathbf{u}$ has non-orthogonal mode frames according to Theorem 2.16, we first need to orthogonalize $\mathbf{u}$ and then calculate the norm as above. The orthogonalization procedure is described in the following section. For all the operations in this section, we are particularly interested in how these operations affect the ranks of hierarchical tensors.

2.4.5.1 Scalar Multiplication, Addition and Orthogonalization

Having in mind the representation (2.18) of a hierarchical tensor, we can perform scalar multiplications by simply multiplying the root transfer tensor $\mathbf{b}^*$ with the scalar. Hence, this operation does not change the ranks of the hierarchical tensor.

We will explain the principal idea of the addition of two hierarchical tensors $\mathbf{u}$ and $\mathbf{v}$ by the examples of low-rank matrices. Assume, we have two low-rank matrices $\mathbf{A}_1 = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^\top$ and $\mathbf{A}_2 = \mathbf{U}_2\mathbf{\Sigma}_2\mathbf{V}_2^\top$, where $\mathbf{U}_1 \in \mathbb{R}^{n\times r_1}$, $\mathbf{U}_2 \in \mathbb{R}^{n\times r_2}$, $\mathbf{V}_1 \in \mathbb{R}^{m\times r_1}$, $\mathbf{V}_2 \in \mathbb{R}^{m\times r_2}$ are orthogonal matrices and $\mathbf{\Sigma}_1 \in \mathbb{R}^{r_1\times r_1}$ and $\mathbf{\Sigma}_2 \in \mathbb{R}^{r_2\times r_2}$ are diagonal matrices with the singular values on the diagonal and where r_1 and r_2 are minimal. Then we have

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 = \begin{pmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{pmatrix}\begin{bmatrix}\mathbf{\Sigma}_1 & 0\\ 0 & \mathbf{\Sigma}_2\end{bmatrix}\begin{pmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{pmatrix}^\top = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top.$$

From this representation of the sum, we directly see that

$$\operatorname{rank}(\mathbf{A}) \le r_1 + r_2 = \operatorname{rank}(\mathbf{A}_1) + \operatorname{rank}(\mathbf{A}_2).$$

Since $\mathbf{U}$ and $\mathbf{V}$ are not necessarily orthogonal anymore, the above representation is not a singular value decomposition in general. However, one can efficiently compute a SVD of $\mathbf{A}$ by computing QR decompositions $\mathbf{U} = \mathbf{Q}_\mathbf{U}\mathbf{R}_\mathbf{U}$ and $\mathbf{V} = \mathbf{Q}_\mathbf{V}\mathbf{R}_\mathbf{V}$ and then performing an SVD on the smaller matrix $\mathbf{R}_\mathbf{U}\mathbf{\Sigma}\mathbf{R}_\mathbf{V}^\top = \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^\top$. From this, we obtain

$$\mathbf{A} = (\mathbf{Q}_\mathbf{U}\tilde{\mathbf{U}})\tilde{\mathbf{\Sigma}}(\mathbf{Q}_\mathbf{V}\tilde{\mathbf{V}})^\top,$$

which is a singular value decomposition of $\mathbf{A}$ since the product of orthogonal matrices is again orthogonal. Since $\mathbf{R}_\mathbf{U}\mathbf{\Sigma}\mathbf{R}_\mathbf{V}^\top \in \mathbb{R}^{r_1+r_2\times r_1+r_2}$, the computational complexity of the SVD is a fixed multiple of $(r_1+r_2)^3$. Moreover, the computational complexity of the two QR decompositions is a fixed multiple of $(r_1+r_2)^2(n+m)$. Hence, the total complexity scales linearly in the sum of the matrix dimensions n and m. For the summation of hierarchical tensors, we can use the same ideas. As for the low-rank matrices, we can simply concatenate the mode frames and additionally concatenate

the transfer tensors. As a simple consequence, for all $e \in \mathbb{E}_d$ we have

$$\operatorname{rank}_{[e]}(\mathbf{u}+\mathbf{v}) \le \operatorname{rank}_{[e]}(\mathbf{u}) + \operatorname{rank}_{[e]}(\mathbf{v})$$

and thus

$$\operatorname{rank}(\mathbf{u}+\mathbf{v}) \le \operatorname{rank}(\mathbf{u}) + \operatorname{rank}(\mathbf{v}).$$

By concatenating the mode frames and transfer tensors, similarly to low-rank matrices, we do not necessarily obtain a representation with orthogonal mode frames. However, as before, we can compute QR decompositions of the mode frames and adapt the transfer tensors from leaves to the root node in a suitable way to obtain such a representation. For details on this procedure, we refer to [Gra10; HK09]. The computational complexity of this orthogonalization procedure is

$$\mathcal{O}\Bigg(\sum_{\alpha\in\mathcal{N}(\mathbb{T}_d)} (r_\alpha^2 r_{\alpha_1} r_{\alpha_2} + r_\alpha r_{\alpha_1}^2 r_{\alpha_2} + r_\alpha r_{\alpha_1} r_{\alpha_2}^2) + \sum_{i=1}^{d} r_{\{i\}}^2 n_i\Bigg),$$

where the transfer tensors $\mathbf{b}^\alpha$ are of representation ranks $r_\alpha \times r_{\alpha_1} \times r_{\alpha_2}$ with the adaption $r^{\alpha^*} = 1$ and the mode frames $\mathbf{U}^{\{i\}}$ are of size $n_i \times r_{\{i\}}$ for $i = 1,\dots,d$.

2.4.5.2 Application of Operator in Hierarchical Tensor Format

In this section, we describe the application of operators in the hierarchical tensor format. To be able to do this, the operator $\mathbf{A}\colon \ell_2(\mathbb{N}^d) \to \ell_2(\mathbb{N}^d)$ needs to be in a hierarchical tensor format as well, that is

$$\mathbf{A} = \sum_{\mathsf{n}\in \mathrm{K}_d(R)} c_{\mathsf{n}} \bigotimes_{i=1}^{d} \mathbf{A}_{n_i}^{(i)}, \tag{2.21}$$

where $\mathrm{K}_d(R) = \times_{i=1}^d \{1,\dots,R\}$, $\mathbf{A}_{n_i}^{(i)}\colon \ell_2(\mathbb{N}) \to \ell_2(\mathbb{N})$ for $n_i = 1,\dots,R$ and $i = 1,\dots,d$ and the finitely supported tensor $\mathbf{c} = (c_{\mathsf{n}})_{\mathsf{n}} \in \ell_2(\mathbb{N}^d)$ is in hierarchical tensor format represented by transfer tensors $(\mathbf{b}_{\mathbf{A}}^\alpha)_{\alpha\in\mathcal{N}(\mathbb{T}_d)}$. In this case, the low-dimensional matrices $\mathbf{A}_{n_i}^{(i)}$ play the role of the mode frames in the sense of (2.17). The extension to a different number R_i of low-dimensional operators $\mathbf{A}_{n_i}^{(i)}$ per dimension is straightforward and omitted for simplicity. We denote by

$$r_{\mathbf{A}} = \max\left\{ \max\{r, r_1, r_2\} : \mathbf{b}_{\mathbf{A}}^\alpha \in \mathbb{R}^{r\times r_1\times r_2}, \alpha \in \mathcal{N}(\mathbb{T}_d)\right\}$$

the rank of the operator $\mathbf{A}$ which is the maximal mode size of all transfer tensor of $\mathbf{A}$.

Let $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ be a hierarchical tensor represented by mode frames $\{\mathbf{U}_k^{\{i\}}\}_{k=1,\dots,n_i}$ for $i = 1,\dots,d$ and transfer tensors $(\mathbf{b}_{\mathbf{u}}^\alpha)_{\alpha\in\mathcal{N}(\mathbb{T}_d)}$. Then $\mathbf{v} = \mathbf{A}\mathbf{u}$ can be represented by the mode frames $\{\mathbf{V}_k^{(i)}\}_{k=1,\dots,Rn_i}$, where

$$\mathbf{V}_{(l-1)n_i+m}^{(i)} = \mathbf{A}_l^{(i)} \mathbf{U}_m^{(i)}, \quad l = 1,\dots,R, \quad m = 1,\dots,n_i$$

and the transfer tensors

$$\mathbf{b}_{\mathbf{v}}^\alpha = \mathbf{b}_{\mathbf{A}}^\alpha \otimes \mathbf{b}_{\mathbf{u}}^\alpha \quad \text{for all } \alpha \in \mathcal{N}(\mathbb{T}_d),$$

where $\otimes$ denotes the Kronecker product of three-dimensional tensors in this case.

This representation of $\mathbf{v}$ is not necessarily in hierarchical tensor format and potentially redundant, since the mode frames are not orthogonal in general. However, as described in the previous section, we can efficiently orthogonalize the representation to obtain a hierarchical tensor representation of $\mathbf{v}$. From the obtained representation with non-orthogonal mode frames and the definition of the Kronecker product, we can directly infer

$$\operatorname{rank}(\mathbf{A}\mathbf{u}) \leq r_{\mathbf{A}} \operatorname{rank}(\mathbf{u}),$$

where $r_{\mathbf{A}}$ is the maximal rank of the operator $\mathbf{A}$ in hierarchical tensor format in (2.21). For more details on the application of operators in hierarchical tensor format, we refer to [Tob12, Section 3.8].

Example 2.20. Let $\hat{\vee}_i$ be one-dimensional index sets for $i = 1, \dots, d$ and $\hat{\vee} = \bigtimes_{i=1}^{d} \hat{\vee}_i$. Let $\mathbf{A} \in \mathbb{R}^{\hat{\vee} \times \hat{\vee}}$ be the representation of an operator on $\ell_2(\hat{\vee})$ of the form

$$\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{I}_2 \otimes \cdots \otimes \mathbf{I}_d + \cdots + \mathbf{I}_1 \otimes \mathbf{I}_2 \otimes \cdots \otimes \mathbf{A}_d,$$

where $\mathbf{A}_i \in \mathbb{R}^{\hat{\vee}_i \times \hat{\vee}_i}$ is the representation of an operator on $\ell_2(\hat{\vee}_i)$ for $i = 1, \dots, d$ and $\mathbf{I}_i$ is the representation of the identity on $\ell_2(\hat{\vee}_i)$. This operator can be represented in hierarchical tensor representation (2.21) with $\mathbf{A}_1^{(i)} = \mathbf{I}_i$ and $\mathbf{A}_2^{(i)} = \mathbf{A}_i$ for $i = 1, \dots, d$ and the transfer tensors

$$(\mathbf{b}^{\alpha}_{1,i,j})_{i,j=1,2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (\mathbf{b}^{\alpha}_{2,i,j})_{i,j=1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \alpha \in \mathcal{N}(\mathbb{T}_d) \setminus \{\alpha^*\},$$

as well as

$$(\mathbf{b}^{\alpha^*}_{1,i,j})_{i,j=1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Hence, the rank of this operator is $r_{\mathbf{A}} = 2$.

This operator is of particular interest in this work, as the Laplacian takes this form when expressed L_2-orthonormal product bases. For more details on this example, as well as for other examples of operators in hierarchical tensor format, we refer to [Tob12].

2.4.5.3 Truncation in Hierarchical Singular Value Decomposition

As seen in the previous sections, adding and applying operators in hierarchical tensor format increases the rank of the resulting hierarchical tensor. Moreover, as seen in Theorem 2.17, it is essential for the efficiency of hierarchical tensor methods to keep the ranks as low as possible.

For the hierarchical tensor format, the best approximation for given rank bounds always exists (see e.g. [Bac23, Theorem 2.31]), but it is not efficiently computable in practice. Therefore, we use again the concept of quasi-best approximations. To obtain quasi-best approximation, we truncate the HSVD. For the description of the truncation procedure, we follow the lines of [Bac23, Section 2.7.1]. We will state some results without proof and refer to [Bac23; Gra10; HK09] for details. Let $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ be in HSVD form with orthonormal mode frames $\{\mathbf{U}_k^{\alpha}\}_{k=1,\dots,\operatorname{rank}_{\alpha}(\mathbf{u})}$. Let $\mathbf{U}^{\alpha,r}$ be the matrix containing the first r columns of the mode frame $\mathbf{U}^{\alpha}$. Then we define the orthogonal projector $\mathrm{P}^{e}_{\mathbf{u},r}$ by

$$\mathrm{P}^{e}_{\mathbf{u},r}\mathbf{v} = \mathbf{v} \times_{[e]} (\mathbf{U}^{[e],r}(\mathbf{U}^{[e],r})^{\top}) \quad \text{for } e \in \mathbb{E}_d, \mathbf{v} \in \ell_2(\mathbb{N}^d)$$

with the mode product $\times_{[e]}$ defined in (2.14).

Furthermore, let $e_1, \dots, e_l$ with $l = \#\mathbb{E}_d$ be an ordering of all effective edges $\mathbb{E}_d$ such that the distance of $[e_i]$ to the root node is smaller or equal as the distance of $[e_j]$ to the root node for $i \le j$. Then with $\mathsf{r} = (r_e)_{e\in\mathbb{E}_d} \subset \mathbb{N}$, we define the projection

$$\mathrm{P}_{\mathbf{u},\mathsf{r}} = \mathrm{P}^{e_l}_{\mathbf{u},r_{e_l}} \cdots \mathrm{P}^{e_1}_{\mathbf{u},r_{e_1}}. \tag{2.22}$$

This operator projects onto truncated mode frames starting from the root node and realizes a rank-truncation of $\mathbf{u}$ to ranks of at most r.

Lemma 2.21. *Let $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ and $\mathsf{r} = (r_e)_{e\in\mathbb{E}_d} \subset \mathbb{N}$ with $\mathsf{r} \le \operatorname{rank}(\mathbf{u})$. Then we have $\operatorname{rank}(\mathrm{P}_{\mathbf{u},\mathsf{r}}\mathbf{u}) \le \mathsf{r}$.*

The ordering of the effective edges in the definition of $\mathrm{P}_{\mathbf{u},\mathsf{r}}$ is essential to ensure the rank bound. For more details, we refer to [Gra10].

As the final part, we state the quasi-best-approximation property of the HSVD truncation.

Theorem 2.22 ([Gra10])**.** *Let $\mathbf{u} \in \ell_2(\mathbb{N}^d)$ and $\mathsf{r} = (r_e)_{e\in\mathbb{E}_d} \subset \mathbb{N}$ with $\mathsf{r} \le \operatorname{rank}(\mathbf{u})$. Then we have*

$$\begin{aligned}\|\mathbf{u} - \mathrm{P}_{\mathbf{u},\mathsf{r}}\mathbf{u}\| &\le \Big(\sum_{e\in\mathbb{E}_d}\sum_{k>r_e} \big(\sigma_k^{[e]}(\mathbf{u})\big)^2\Big)^{\frac{1}{2}} \\ &\le \sqrt{2d-3}\,\min\big\{\|\mathbf{u}-\mathbf{v}\| : \operatorname{rank}(\mathbf{v}) \le \mathsf{r}\big\}.\end{aligned}$$

The proof of this statement is based on the definition of $\mathrm{P}_{\mathbf{u},\mathsf{r}}$ as a sequence of $l = \#\mathbb{E}_d = 2d-3$ projections. The main idea is similar to the proof of the quasi-best approximation property of the Tucker format in Section 2.4.2.

2.5 Wavelet Representation of Operators

Next, we discuss the representation of operators in wavelet bases arising in the context of variational formulations of PDEs. Furthermore, we discuss the compressibility of these operators.

Let H be a separable Hilbert space with a Riesz wavelet basis $\{\psi_\nu\}_{\nu\in\hat{\vee}}$ with constants $c, C > 0$ according to Theorem 2.1. Furthermore, let $A : H \to H'$ be a boundedly invertible operator. Then the bi-infinite matrix

$$\mathbf{A} = \big((A\psi_\nu)(\psi_\mu)\big)_{\mu,\nu\in\hat{\vee}}$$

is boundedly invertible on $\ell_2(\hat{\vee})$ with

$$\|\mathbf{A}\| \le C^2\|A\|, \quad \|\mathbf{A}^{-1}\| \le c^{-2}\|A^{-1}\|.$$

A proof of this statement is straightforward and can be found in [Dah01, Section 4]. In a slightly different setting, we show this result in Theorem 3.6.

An example is the operator corresponding to the weak formulation of the Laplacian. Let $\Omega = (0,1)^d$ and let $\{\psi_\nu\}_{\nu\in\vee_1}$ be a $L_2(0,1)$-orthonormal basis. Then with $\Psi_\nu = \bigotimes_{i=1}^d \psi_{\nu_i}$ as described in Section 2.3,

$$\big\{\|\Psi_\nu\|^{-1}_{H^1_0(\Omega)}\Psi_\nu\big\}_{\nu\in\vee_{\mathrm{x}}} \tag{2.23}$$

is a Riesz basis of $H^1_0(\Omega)$, where $\vee_{\mathrm{x}} = \times_{i=1}^d \vee_1$. The operator $A : H^1_0(\Omega) \to H^{-1}(\Omega)$ defined by

$$(Av)(w) = \int_\Omega \nabla v(x)\cdot\nabla w(x)\,\mathrm{d}x,$$

has the representation

$$\mathbf{A} = \bar{\mathbf{D}}\mathbf{T}_{\mathrm{x}}\bar{\mathbf{D}}$$

with respect to the basis (2.23), where

$$\mathbf{T}_{\mathrm{x}} = \mathbf{T}_2 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 + \cdots + \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{T}_2 \tag{2.24}$$

and where $\mathbf{I}_1$ is the identity operator on $\ell_2(\vee_1)$, $\bar{\mathbf{D}}$ is a diagonal scaling matrix with $\bar{\mathbf{D}}_{\nu,\nu} = \|\Psi_\nu\|^{-1}_{H^1_0(\Omega)}$, and

$$(\mathbf{T}_2)_{\nu,\mu} = \int_0^1 \psi'_\nu(x)\psi'_\mu(x)\,\mathrm{d}x. \tag{2.25}$$

This representation is a direct consequence of the product structure of the basis, the definition of the operator A, as well as the L_2-orthonormality of the one-dimensional basis.

To work efficiently with these kinds of operators, they need specific compressibility properties. For the compressibility of operators, we recall the following definition from [CDD01].

Definition 2.23. Let $s^* > 0$ and let $\hat{\vee}$ be a countable index set. An operator $\mathbf{M}\colon \ell_2(\hat{\vee}) \to \ell_2(\hat{\vee})$ is called *s^*-compressible* if for any $0 < s < s^*$, there exist summable positive sequences $(\alpha_j)_{j\in\mathbb{N}_0}, (\beta_j)_{j\in\mathbb{N}_0}$ such that for each $j \in \mathbb{N}_0$ there exists a $\mathbf{M}_j$ with at most $\alpha_j 2^j$ nonzero entries per row and column satisfying $\|\mathbf{M} - \mathbf{M}_j\| \le \beta_j 2^{-sj}$. The sequences arising from a given s^*-compressible operator are denoted by $\alpha(\mathbf{M})$ and $\beta(\mathbf{M})$.

Example 2.24. Clearly, each diagonal operator $\tilde{\mathbf{D}}$ is s^*-compressible for any $s^* > 0$ with $\tilde{\mathbf{D}}_j = \tilde{\mathbf{D}}$, $\alpha_j = 2^{-j}$ and $\beta_j = 0$. Hence, the identity operator is s^*-compressible for any $s^* > 0$.

We work with operators in low-rank format as in (2.24), which can be written in hierarchical format. For the application of operators in hierarchical format onto tensors in hierarchical format, as seen in Section 2.4.5.2, we only need to apply the low-dimensional components of the operator. Thus, we need to discuss only the compressibility of the factors in the low-rank representation.

As we will see in Section 4.3.2, the compressibility according to Theorem 2.23 will not be sufficient in certain situations, since an operator might be applied to a sum of many low-rank tensors. Hence, if we have an exponential number of nonzero entries per row/column in the approximations $\mathbf{M}_j$, the ranks might also increase exponentially by the addition properties discussed in Section 2.4.5.1. Due to this issue, we need a more restrictive compressibility class.

Definition 2.25. Let $\hat{\vee}, \tilde{\vee}$ be countable index sets. An operator $\mathbf{M}\colon \ell_2(\hat{\vee}) \to \ell_2(\tilde{\vee})$ is called *super-compressible* if there exist a summable positive sequence $(\beta_j)_{j\in\mathbb{N}_0}$ and a $c > 0$ such that for each $j \in \mathbb{N}_0$, there exist a $\mathbf{M}_j$ with $\|\mathbf{M} - \mathbf{M}_j\| \le \beta_j 2^{-j}$ and for each $j \in \mathbb{N}_0$ the operator $\mathbf{M}_{j+1} - \mathbf{M}_j$ has only c nonzero entries per row and column.

Remark 2.26. Assume we have an approximation sequence $(\tilde{\mathbf{M}}_j)_{j\in\mathbb{N}_0}$ for the operator $\mathbf{M}$ with $\|\mathbf{M} - \tilde{\mathbf{M}}_j\| \le \beta_j 2^{-sj}$ for some $0 < s < 1$ with summable $(\beta_j)_{j\in\mathbb{N}_0}$ and $\tilde{\mathbf{M}}_{j+1} - \tilde{\mathbf{M}}_j$ has only c nonzero entries per row and column. Then the operator $\mathbf{M}$ is super-compressible. The subsequence $(\mathbf{M}_j)_{j\in\mathbb{N}_0}$ with $\mathbf{M}_j = \tilde{\mathbf{M}}_{kj}$ for $k > s^{-1}$ fulfills the condition $\|\mathbf{M} - \mathbf{M}_j\| \le \beta_{kj} 2^{-j}$ and $\mathbf{M}_{j+1} - \mathbf{M}_j$ has only kc nonzero entries per row and column.

Remark 2.27. By taking suitable subsequences of the approximation operators in the Theorem 2.25 of super-compressibility, one can match any rate s in Theorem 2.23. To do this, we have to modify the sequence $(\beta_j)_j$ and the constant $c > 0$ correspondingly. Furthermore, for simplicity we can set $\mathbf{M}_0 = 0$ and $\beta_0 = \|\mathbf{M}\|$. Then as an immediate consequence of the definition, $\mathbf{M}_J$ has at most cJ nonzero entries in each row and column for each $J \in \mathbb{N}_0$.

In some situations, the compressibility of the operators alone will not be sufficient for our purposes. Additionally, we need to know which entries in the approximation vanish. Therefore, we make additional regularity assumptions in the following sense.

Definition 2.28. Let $\hat{\vee}$ be a countable index set as in (2.3) and $\mathbf{M}\colon \ell_2(\hat{\vee}) \to \ell_2(\hat{\vee})$ be s^*-compressible or super-compressible with approximations $\mathbf{M}_j$. We say that these approximations have *level decay* if there exists a $\gamma > 0$ such that $(\mathbf{M}_j)_{\nu\mu} = 0$ for $||\nu| - |\mu|| > \gamma j$.

This property was also used in the high-dimensional elliptic case in [BD16].

A standard tool to show the compressibility of operators is the following lemma.

Lemma 2.29 (Schur). *Let $\vee, \hat{\vee}$ be countable index sets and $\mathbf{M} = (m_{\nu,\mu})_{\nu\in\vee,\mu\in\hat{\vee}}$ be a bi-infinite matrix and $(\omega_\nu)_{\nu\in\vee}$ and $(\hat{\omega}_\mu)_{\mu\in\hat{\vee}}$ be positive sequences. If*

$$\sum_{\nu\in\vee} |m_{\nu,\mu}|\omega_\nu \le c\hat{\omega}_\mu, \quad \sum_{\mu\in\hat{\vee}} |m_{\nu,\mu}|\hat{\omega}_\mu \le \hat{c}\omega_\nu$$

for all $\nu \in \vee$ and $\mu \in \hat{\vee}$, then $\mathbf{M}$ is a bounded operator from $\ell_2(\hat{\vee})$ to $\ell_2(\vee)$ with $\|\mathbf{M}\| \le \sqrt{c\hat{c}}$.

For completeness, we give a proof in Chapter A.

In the final part of this section, we give two examples of compressible operators in wavelet coordinates. To show the compressibility of certain operators, we specify the assumptions on the wavelet basis. In this work, we restrict ourselves to L_2-orthonormal, compactly supported spline wavelet bases. Following [DSS09, Section 3] we consider a L_2-orthonormal wavelet basis $\{\psi_\nu\}_{\nu\in\vee_1}$ for $L_2(0,1)$ such that $\{2^{-|\nu|}\psi_\nu\}_{\nu\in\vee_1}$ is a Riesz basis for $H_0^1(0,1)$. Furthermore, we assume the following:

(i) $|\operatorname{supp}\psi_\nu| \eqsim 2^{-|\nu|}$.

(ii) Each wavelet ψ_ν is a piecewise polynomial of degree s with a fixed number of pieces.

(iii) Each point in $[0,1]$ is contained in the support of at most a bounded number of wavelets per level.

(iv) $\|\psi_\nu\|_{L_\infty} \lesssim 2^{\frac{|\nu|}{2}}$ and $\|\psi_\nu'\|_{L_\infty} \lesssim 2^{\frac{3}{2}|\nu|}$.

(v) Wavelets on level larger than zero whose support does not intersect the boundary have $s-1$ vanishing moments.

Example 2.30. Let the assumptions (i)-(v) above hold. Furthermore, let $\bar{\mathbf{D}}_1 \in \mathbb{R}^{\vee_1\times\vee_1}$ be the bi-infinite diagonal matrix with $\bar{\mathbf{D}}_{1,\nu,\nu} = \|\psi_\nu\|^{-1}_{H_0^1(0,1)}$. Then the operator $\mathbf{M} = \bar{\mathbf{D}}_1\mathbf{T}_2\bar{\mathbf{D}}_1$ is super-compressible with level decay, where $\mathbf{M}_j = \bar{\mathbf{D}}_1\mathbf{T}_{2,j}\bar{\mathbf{D}}_1$ and where $\mathbf{T}_{2,j}$ is defined by dropping all entries $(\mathbf{T}_2)_{\nu,\mu}$ with $||\nu| - |\mu|| > 3j$.

Proof. Based on the vanishing moment property, we have for ψ_μ with $\operatorname{supp}(\psi_\mu) \subset (0,1)$ that

$$\int_0^1 \psi_\nu'(x)\psi_\mu'(x)\,\mathrm{d}x = -\int_0^1 \psi_\nu(x)\psi_\mu''(x)\,\mathrm{d}x = 0 \tag{2.26}$$

if ψ_μ is polynomial on $\operatorname{supp}(\psi_\nu)$, i.e., $\operatorname{supp}(\psi_\nu)$ is a subset of one of the pieces of ψ_μ. Moreover, by property (i) and (iii) only a constant number of wavelets overlap with multiple pieces of ψ_μ per level. Hence, we only have a fixed maximal number c of nonzero entries per row/column per level in $\mathbf{T}_{2,j}$. Furthermore, we have the estimates

$$\begin{aligned}\left|\int_0^1 \psi_\nu'(x)\psi_\mu'(x)\,\mathrm{d}x\right| &\le |\operatorname{supp}(\psi_\nu)\cap\operatorname{supp}(\psi_\mu)|\|\psi_\nu'\|_{L_\infty}\|\psi_\mu'\|_{L_\infty}\\ &\lesssim 2^{-\max\{|\nu|,|\mu|\}}2^{\frac{3}{2}|\nu|}2^{\frac{3}{2}|\mu|}\\ &\eqsim 2^{-\frac{1}{2}||\nu|-|\mu||}\|\psi_\nu\|_{H_0^1}\|\psi_\mu\|_{H_0^1}.\end{aligned}$$

By the definition of $\mathbf{M}_j$, for fixed $\nu \in \vee_1$ we have

$$\begin{aligned}\sum_{\mu\in\vee_1}|(\mathbf{M}-\mathbf{M}_j)_{\nu,\mu}| &= \sum_{\substack{\mu\in\vee_1\\ \left||\mu|-|\nu|\right|>3j}}|(\mathbf{M})_{\nu,\mu}|\\ &= \sum_{k=3j+1}^{\infty}\sum_{\substack{\mu\in\vee_1\\ \left||\mu|-|\nu|\right|=k}}|(\mathbf{M})_{\nu,\mu}|\\ &\lesssim c\sum_{k=3j+1}^{\infty}2^{-\frac{1}{2}k}\\ &\lesssim c2^{-\frac{3}{2}j}.\end{aligned}$$

Combining this with Theorem 2.29 yields the desired approximation property of $\mathbf{M}_j$. For more details, we refer to [DSS09, Proposition 3.3]. □

As one can see in (2.26), $s-2$ vanishing moments would be sufficient for this example. However, in the following examples, we will need $s-1$ vanishing moments.

Example 2.31. The operator $\mathbf{M}$ defined by

$$(\mathbf{M})_{\nu,\mu} = \|\psi_\mu\|_{H_0^1}^{-1}\int_0^1 \psi_\nu'(x)\psi_\mu(x)\,\mathrm{d}x$$

is super-compressible with level decay, where $\mathbf{M}_j$ is defined by dropping all entries $(\mathbf{M})_{\nu,\mu}$ where $||\nu|-|\mu|| > j$. To show the compressibility, we can proceed as in the last example. The level decay property is a direct consequence of the definition of $\mathbf{M}_j$.

In the analysis of the adaptive application of operators in Section 4.3, we will only use that the one-dimensional operator $\bar{\mathbf{D}}_1\mathbf{T}_2\bar{\mathbf{D}}_1$ is s^*-compressible, but as seen in Theorem 2.30, the operator is even super-compressible with level decay. However, this is not necessary for our analysis, so we keep this as general as possible. The use of super-compressible operators will be essential in Section 4.3.2, where we consider an operator of similar form as in Theorem 2.31.

For the analysis of the compressibility of operators in wavelet coordinates for more general wavelet bases, we refer to [Ste04; Ste09].

Chapter 3

Parabolic Differential Equations and Space-Time Variational Formulation

We are interested in approximating solutions to parabolic PDEs, which depend on time as well as on spatial variables in a high dimensional domain. There are various numerical methods to solve such PDEs, however, many are not well-suited for high-dimensional problems due to the curse of dimensionality. To overcome this issue, [DKO12] combined classical time-stepping schemes with low-rank methods for elliptic problems. In this approach the computational complexity and the evolution of the tensor ranks during the time-stepping are difficult to control. In order to avoid growing ranks, [Bac+21; Con20] considered dynamical low-rank methods on the manifold of fixed-rank tensors together with time stepping schemes. However, one cannot provide error guarantees in this setting.

One way to provide error bounds and rank control is to use space-time formulations. In the space-time formulation, we use functions defined on the full space-time domain as approximations of the solution. This provides more flexibility than evolution at different time points, resulting in better rank control. Furthermore, the space-time formulation induces a natural space-time norm on the solution space, which allows us to derive rigorous error bounds.

First, we provide a brief introduction to parabolic PDEs. Let V, H be separable Hilbert spaces and let V be densely embedded in H. Furthermore, let $I = [0, T]$ be the time interval with final point $T \in (0, \infty)$. For almost every time point a bilinear form $a(t; \cdot, \cdot) : V \times V \to \mathbb{R}$ is given, such that for all $v, \hat{v}$ in V and almost every $t \in I$ the mapping $t \mapsto a(t; v, \hat{v})$ is measurable. Additionally, we assume that the bilinear form is bounded and coercive, i.e., for almost every $t \in I$ we have

$$\begin{aligned} |a(t; v, \hat{v})| &\leq a_{\max} \|v\|_V \|\hat{v}\|_V && \text{for all } v, \hat{v} \in V, \\ a(t; v, v) + \lambda_0 \|v\|_H^2 &\geq a_{\min} \|v\|_V^2 && \text{for all } v \in V \end{aligned} \tag{3.1}$$

for some constants $0 < a_{\min} \leq a_{\max} < \infty$ and $\lambda_0 \in \mathbb{R}$. Using this bilinear form, we can define for almost every $t \in I$ a linear operator $A(t) \in \mathcal{L}(V, V')$ by

$$\langle A(t)v, \hat{v} \rangle_{V',V} = a(t; v, \hat{v}), \tag{3.2}$$

where V' denotes the dual space of V. The linear parabolic PDE for given $g \in L_2(I, V')$ and $h \in H$ is then given by

$$\begin{aligned} \partial_t u(t) + A(t)u(t) &= g(t) && \text{in } V', \\ u(0) &= h && \text{in } H. \end{aligned} \tag{3.3}$$

The function g is called the *force term* and h is the *initial value*. In this work, we consider second-order elliptic operators A, in particular, the ones in the following two examples.

Example 3.1. The simplest example of a parabolic PDE is the heat equation. In this case, $A = -\Delta$ is the sum of the second derivatives, $H = L_2(\Omega)$ and $V = H_0^1(\Omega)$. Here, the operator is a sum of separable operators as seen in Equation (2.24).

Example 3.2. In the context of MPI, the linear operator A is given by the spatial differential operator of the corresponding Fokker-Planck equation, i.e.

$$A(t)u(t) = \operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2}(bu - \mathbf{M}\nabla_{\mathbb{S}_1\times\mathbb{S}_2}u)(t).$$

In this case, we consider a parabolic PDE on a manifold, which is the product of two unit spheres $\mathbb{S}\times\mathbb{S}$.

3.1 Space-Time Variational Formulation

A variety of space-time variational formulations for the linear parabolic problem (3.3) exists. For later use, we will choose a formulation compatible with adaptivity and low-rank methods.

Therefore, we use the classical space-time weak formulation analyzed in the context of adaptive methods in [SS09] and in previous works such as [DL00, Ch. XVIII].

For this formulation, we have different trial and test spaces. The trial space is given by

$$\mathcal{X} = L_2(I;V)\cap H^1(I;V') = \left\{v\in L_2(I;V) : \partial_t v\in L_2(I;V')\right\}, \tag{3.4}$$

and the test space by

$$\mathcal{Y} = L_2(I;V)\times H. \tag{3.5}$$

The corresponding norms are chosen naturally for intersection spaces and product spaces as

$$\begin{aligned}\|v\|_{\mathcal{X}} &= \left(\|v\|^2_{L^2(I;V)} + \|\partial_t v\|^2_{L_2(I;V')}\right)^{\frac{1}{2}},\\ \|(w_1,w_2)\|_{\mathcal{Y}} &= \left(\|w_1\|^2_{L^2(I;V)} + \|w_2\|^2_H\right)^{\frac{1}{2}}\end{aligned} \tag{3.6}$$

for $v\in\mathcal{X}$ and $(w_1,w_2)\in\mathcal{Y}$. The weak formulation of the parabolic PDE (3.3) is then given by the following variational problem: find $u\in\mathcal{X}$ such that

$$b(u,v) = f(v) \quad \text{for all } v\in\mathcal{Y}, \tag{3.7}$$

where the bilinear form $b\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{R}$ is defined by

$$b(v,w) = \int_I \langle\partial_t v(t), w_1(t)\rangle_{V'\times V} + a\big(t;v(t),w_1(t)\big)\,\mathrm{d}t + \langle v(0),w_2\rangle_H \tag{3.8}$$

and the functional $f\colon\mathcal{Y}\to\mathbb{R}$ is given by

$$f(w) = \int_I \langle g(t), w_1(t)\rangle_H\,\mathrm{d}t + \langle h, w_2\rangle_H.$$

Theorem 3.3. *Let the operator $B\in\mathcal{L}(\mathcal{X},\mathcal{Y}')$ be defined by $B(v) = b(v,\cdot)$ for the bilinear form b defined in (3.8) and $\mathcal{X}$, $\mathcal{Y}$ defined in (3.4) and (3.5), respectively. Then the operator B is boundedly invertible.*

We can prove the theorem using the Banach-Nečas-Babuška theorem, showing boundedness of the bilinear form, the inf-sup condition, and surjectivity. For details of the proof and in particular for specific bounds of $\|B\|$ and $\|B^{-1}\|$, we refer to [SS09].

Remark 3.4. If we consider a coercive spatial part, i.e., $\lambda_0 = 0$ in (3.1), then the bounds on $\|B\|$ and $\|B^{-1}\|$ are independent of the final time point T. This statement can be seen from [SS09, Thm 5.1]. An example for a parabolic PDE with coercive spatial part is the heat equation in Theorem 3.1. Alternatively, if we use a norm of the trial space $\mathcal{X}$ that includes the initial value at $t = 0$, one can also show T-independence for $\lambda_0 > 0$. For details, we refer to [TV16].

3.1.1 Infinite Matrix Representation of the Operator

The construction of the adaptive method in Chapter 4 is based on reformulating problem (3.7) as a bi-infinite matrix equation. As explained in Section 2.1, the choice of a suitable basis is crucial for establishing a relationship between the error in the function space norm and the sequence norm of the corresponding coefficient sequences. For this reason, we will use Riesz bases for the spaces $\mathcal{X}$ and $\mathcal{Y}$. By suitable combination of Riesz bases of $L_2(I), H^1(I), H, V$ and V', we can construct Riesz bases for the spaces $\mathcal{X}$ and $\mathcal{Y}$. To do this, we follow the construction of [SS09]. Let $\Sigma_{L_2} = \{\theta_\nu\}_{\nu\in\vee_t}$ be a Riesz basis of $L_2(I)$, such that $\Sigma_{H^1} = \{\|\theta_\nu\|^{-1}_{H^1(I)}\theta_\nu\}_{\nu\in\vee_t}$ is a Riesz basis of $H^1(I)$. As seen in Theorem 2.4, such bases can be constructed by using wavelets. Furthermore, let $\Sigma_H = \{\Psi_\nu\}_{\nu\in\vee_x}$ be a Riesz basis of H such that $\Sigma_V = \{\|\Psi_\nu\|^{-1}_V\Psi_\nu\}_{\nu\in\vee_x}$ is a Riesz basis of V and $\Sigma_{V'} = \{\|\Psi_\nu\|^{-1}_{V'}\Psi_\nu\}_{\nu\in\vee_x}$ is a Riesz basis of V'.

Additionally, let $\bar{s}^{\mathcal{X}}$ be a sequence of positive numbers on $\vee = \vee_t \times \vee_x$ with

$$\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \eqsim \|\theta_{\nu_t}\Psi_{\nu_x}\|^{-1}_{\mathcal{X}} \eqsim \left(\|\Psi_{\nu_x}\|^2_V + \|\theta_{\nu_t}\|^2_{H^1(I)}\|\Psi_{\nu_x}\|^2_{V'}\right)^{-\frac{1}{2}} \tag{3.9}$$

and $\bar{s}^{\mathcal{Y}}$ be a sequence of positive numbers on $\vee_x$ with

$$\bar{s}^{\mathcal{Y}}_{\nu_x} \eqsim \|\Psi_{\nu_x}\|^{-1}_V. \tag{3.10}$$

Due to the Riesz basis property, we additionally obtain

$$\bar{s}^{\mathcal{Y}}_{\nu_x} \eqsim \|\theta_{\nu_t}\Psi_{\nu_x}\|^{-1}_{L_2(I;V)}$$

for all $\nu_t \in \vee_t$.

Next, we construct Riesz bases of $\mathcal{X}$ and $\mathcal{Y}$ following [SS09, Section 6] by appropriate rescaling of the tensor products of the bases of the spatial and temporal parts.

Lemma 3.5. *Based on the previous construction, the set*

$$\Sigma_{\mathcal{X}} = \left\{(t,x) \mapsto \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\theta_{\nu_t}(t)\Psi_{\nu_x}(x) : \nu_t \in \vee_t, \nu_x \in \vee_x\right\}$$

is a Riesz basis of $\mathcal{X}$ and

$$\Sigma_{\mathcal{Y}} = \left\{(t,x) \mapsto (\bar{s}^{\mathcal{Y}}_{\nu_x}\theta_{\nu_t}(t)\Psi_{\nu_x}(x), 0) : \nu_t \in \vee_t, \nu_x \in \vee_x\right\} \cup \left\{x \mapsto (0, \Psi_{\nu_x}) : \nu_x \in \vee_x\right\}$$

is a Riesz basis of $\mathcal{Y}$. The corresponding Riesz constants fulfil

$$c_{\mathcal{X}} \gtrsim \min\{c_{\Sigma_{L_2}} c_{\Sigma_V}, c_{\Sigma_{H^1}} c_{\Sigma_{V'}}\}, \qquad C_{\mathcal{X}} \lesssim \max\{C_{\Sigma_{L_2}} C_{\Sigma_V}, C_{\Sigma_{H^1}} C_{\Sigma_{V'}}\},$$
$$c_{\mathcal{Y}} \gtrsim \min\{c_{\Sigma_{L_2}} c_{\Sigma_V}, c_{\Sigma_H}\}, \qquad C_{\mathcal{Y}} \lesssim \max\{C_{\Sigma_{L_2}} C_{\Sigma_V}, C_{\Sigma_H}\},$$

where the additional constants depend on the constants in (3.9) and (3.10).

Proof. Let $\Lambda \subset \vee$ be finite and $(u_\lambda)_{\lambda\in\vee} \subset \mathbb{R}$. Without loss of generality we can choose $\Lambda = \Lambda_t \times \Lambda_x$ as a product set with $\Lambda_t \subset \vee_t$ and $\Lambda_x \subset \vee_x$ by filling (u_λ) with additional zeros. Then using the Riesz basis property of Σ_{L_2} and Σ_V as well as (3.9), we have

$$\begin{aligned}
\Big\| \sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Psi_{\nu_x} \Big\|^2_{L_2(I;V)} &= \Big\| \sum_{\nu_x\in\Lambda_x} \|\Psi_{\nu_x}\|_V^{-1} \Psi_{\nu_x} \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_V \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Big\|^2_{L_2(I;V)} \\
&= \int_I \Big\| \sum_{\nu_x\in\Lambda_x} \|\Psi_{\nu_x}\|_V^{-1} \Psi_{\nu_x} \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_V \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t}(t) \Big\|^2_V \,\mathrm{d}t \\
&\geq c^2_{\Sigma_V} \sum_{\nu_x\in\Lambda_x} \int_I \Big(\sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_V \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t}(t) \Big)^2 \mathrm{d}t \\
&= c^2_{\Sigma_V} \sum_{\nu_x\in\Lambda_x} \Big\| \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_V \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Big\|^2_{L_2(I)} \\
&\geq c^2_{\Sigma_V} c^2_{\Sigma_{L_2}} \sum_{\nu_x\in\Lambda_x} \sum_{\nu_t\in\Lambda_t} \big(u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_V \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \big)^2
\end{aligned} \tag{3.11}$$

Furthermore, we have

$$\begin{aligned}
\Big\| \partial_t \sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Psi_{\nu_x} \Big\|^2_{L_2(I;V')} &= \Big\| \sum_{\nu_x\in\Lambda_x} \|\Psi_{\nu_x}\|_{V'}^{-1} \Psi_{\nu_x} \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \partial_t \theta_{\nu_t} \Big\|^2_{L_2(I;V')} \\
&\geq c^2_{\Sigma_{V'}} \sum_{\nu_x\in\Lambda_x} \Big\| \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \partial_t \theta_{\nu_t} \Big\|^2_{L_2(I)}
\end{aligned}$$

as well as

$$\begin{aligned}
\Big\| \sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Psi_{\nu_x} \Big\|^2_{L_2(I;V')} &= \Big\| \sum_{\nu_x\in\Lambda_x} \|\Psi_{\nu_x}\|_{V'}^{-1} \Psi_{\nu_x} \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Big\|^2_{L_2(I;V')} \\
&\geq c^2_{\Sigma_{V'}} \sum_{\nu_x\in\Lambda_x} \Big\| \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Big\|^2_{L_2(I)}.
\end{aligned}$$

Combining both estimates yields

$$\begin{aligned}
&\Big\| \partial_t \sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Psi_{\nu_x} \Big\|^2_{L_2(I;V')} + \Big\| \sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Psi_{\nu_x} \Big\|^2_{L_2(I;V')} \\
&\quad\geq c^2_{\Sigma_{V'}} \sum_{\nu_x\in\Lambda_x} \Big\| \sum_{\nu_t\in\Lambda_t} u_{\nu_t,\nu_x} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \theta_{\nu_t} \Big\|^2_{H^1(I)} \\
&\quad\geq c^2_{\Sigma_{V'}} c^2_{\Sigma_{H^1}} \sum_{\nu_x\in\Lambda_x} \sum_{\nu_t\in\Lambda_t} \big(u_{\nu_t,\nu_x} \|\theta_{\nu_t}\|_{H^1(I)} \|\Psi_{\nu_x}\|_{V'} \bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} \big)^2.
\end{aligned} \tag{3.12}$$

Since V is densely embedded in H and by the Gelfand triple identification

$$V \hookrightarrow H \cong H' \hookrightarrow V',$$

we know that there exists a constant $C_{\mathrm{emb}} > 0$ such that $\|v\|_{V'} \le C_{\mathrm{emb}}\|v\|_V$ for all $v \in V$. Hence, we have

$$C_{\mathrm{emb}}\Big\|\sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x}\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\theta_{\nu_t}\Psi_{\nu_x}\Big\|_{L_2(I;V)} \ge \Big\|\sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x}\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\theta_{\nu_t}\Psi_{\nu_x}\Big\|_{L_2(I;V')}.$$

Combining this with (3.11), (3.12) and with the definition of the $\mathcal{X}$-norm from (3.6), we obtain

$$\begin{aligned}
&(1+C_{\mathrm{emb}}^2)\Big\|\sum_{(\nu_t,\nu_x)\in\Lambda} u_{\nu_t,\nu_x}\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\theta_{\nu_t}\Psi_{\nu_x}\Big\|_{\mathcal{X}}^2\\
&\quad\ge \min\{c_{\Sigma_{L_2}}c_{\Sigma_V}, c_{\Sigma_{H^1}}c_{\Sigma_{V'}}\}\sum_{\nu_x\in\Lambda_x}\sum_{\nu_t\in\Lambda_t} u^2_{\nu_t,\nu_x}(\|\Psi_{\nu_x}\|_V^2 + \|\theta_{\nu_t}\|^2_{H^1(I)}\|\Psi_{\nu_x}\|^2_{V'})(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x})^2\\
&\quad\gtrsim \min\{c_{\Sigma_{L_2}}c_{\Sigma_V}, c_{\Sigma_{H^1}}c_{\Sigma_{V'}}\}\sum_{\nu_x\in\Lambda_x}\sum_{\nu_t\in\Lambda_t} u^2_{\nu_t,\nu_x},
\end{aligned}$$

where we used (3.9) in the last step. Since Λ and $(u_\lambda)_\lambda$ were chosen arbitrary, the first statement follows. The remaining inequalities can be proven analogously. □

For the Riesz bases defined in Theorem 3.5, with $\vee' = \vee \cup \vee_x$ we introduce the notation

$$\Sigma_{\mathcal{X}} = \big\{\Psi^{\mathcal{X}}_\nu : \nu \in \vee\big\}, \quad \Sigma_{\mathcal{Y}} = \big\{\Psi^{\mathcal{Y}}_\nu : \nu \in \vee'\big\}.$$

The bi-infinite matrix representation is now defined by inserting the two Riesz bases into the bilinear form, which yields

$$\bar{\mathbf{B}} = b(\Sigma_{\mathcal{X}}, \Sigma_{\mathcal{Y}}) = \Big(b\big(\Psi^{\mathcal{X}}_\nu, \Psi^{\mathcal{Y}}_\mu\big)\Big)_{\mu\in\vee',\nu\in\vee}. \tag{3.13}$$

Corollary 3.6. *The bi-infinite matrix* $\bar{\mathbf{B}}\colon \ell_2(\vee) \to \ell_2(\vee) \times \ell_2(\vee_x)$ *defined in* (3.13) *is an isomorphism with*

$$\|\bar{\mathbf{B}}\| \le \|B\| C_{\mathcal{X}} C_{\mathcal{Y}}, \quad \|\bar{\mathbf{B}}^{-1}\| \le \|B^{-1}\| c_{\mathcal{X}}^{-1} c_{\mathcal{Y}}^{-1},$$

where $c_{\mathcal{X}}, c_{\mathcal{Y}}, C_{\mathcal{X}}, C_{\mathcal{Y}}$ *are the constants from Theorem 3.5.*

Proof. Using

$$\begin{aligned}
|\langle \bar{\mathbf{B}}\mathbf{v}, \mathbf{w}\rangle| &= \Big|\Big\langle B\Big(\sum_{\nu\in\vee} v_\nu \psi^{\mathcal{X}}_\nu\Big), \sum_{\nu\in\vee'} w_\nu \psi^{\mathcal{Y}}_\nu\Big\rangle\Big|\\
&\le \|B\|\Big\|\sum_{\nu\in\vee} v_\nu \psi^{\mathcal{X}}_\nu\Big\|_{\mathcal{X}}\Big\|\sum_{\nu\in\vee'} w_\nu \psi^{\mathcal{Y}}_\nu\Big\|_{\mathcal{Y}}\\
&\le \|B\| C_{\mathcal{X}} C_{\mathcal{Y}} \|\mathbf{v}\|\|\mathbf{w}\|,
\end{aligned}$$

we get the first statement. To bound the inverse norm, we use that

$$\|\tilde{A}\|_{\mathcal{Y}'} = \sup_{0\neq w\in\mathcal{Y}} \frac{\tilde{A}(w)}{\|w\|_{\mathcal{Y}}} \le \sup_{0\neq w\in\mathcal{Y}} \frac{\sum_{\nu\in\vee'} \tilde{A}\big(\psi^{\mathcal{Y}}_\nu\big)w_\nu}{c_{\mathcal{Y}}\|\mathbf{w}\|_{\ell_2}} = c_{\mathcal{Y}}^{-1}\big\|\big(\tilde{A}\big(\psi^{\mathcal{Y}}_\nu\big)\big)_{\nu\in\vee'}\big\|_{\ell_2}$$

for all $\tilde{A} \in \mathcal{Y}'$. By the definition of the bi-infinite matrix, we have

$$\begin{aligned}\|\bar{\mathbf{B}}\mathbf{v}\| &= \Big\|\Big(\Big\langle B\Big(\sum_{\nu\in\vee} v_\nu \Psi_\nu^{\mathcal{X}}\Big), \Psi_\mu^{\mathcal{Y}}\Big\rangle\Big)_{\mu\in\vee'}\Big\|_{\ell_2} \\ &\geq c_{\mathcal{Y}}\Big\|B\Big(\sum_{\nu\in\vee} v_\nu \Psi_\nu^{\mathcal{X}}\Big)\Big\|_{\mathcal{Y}'} \\ &\geq c_{\mathcal{Y}}\|B^{-1}\|^{-1}\Big\|\sum_{\nu\in\vee} v_\nu \Psi_\nu^{\mathcal{X}}\Big\|_{\mathcal{Y}'} \\ &\geq c_{\mathcal{X}} c_{\mathcal{Y}}\|B^{-1}\|^{-1}\|\mathbf{v}\|_{\ell_2}.\end{aligned}$$

From this, we infer the second statement. □

This statement means that the condition number of $\bar{\mathbf{B}}$ depends on the bounds on the linear operator as well as on the constants of the Riesz bases.

Let $u = \sum_{\nu\in\vee} u_\nu \Psi_\nu^{\mathcal{X}}$ and $\mathbf{u} = (u_\nu)_{\nu\in\vee}$. Then u is a solution of the variational formulation (3.7) if and only if

$$\sum_{\nu\in\vee} u_\nu b\big(\Psi_\nu^{\mathcal{X}}, \Psi^{\mathcal{Y}}\big) = b\big(u, \Psi^{\mathcal{Y}}\big) = f\big(\Psi^{\mathcal{Y}}\big) \quad \text{for all } \Psi^{\mathcal{Y}} \in \Sigma_{\mathcal{Y}}. \tag{3.14}$$

This is equivalent to $\mathbf{u}$ satisfying the bi-infinite matrix equation

$$\bar{\mathbf{B}}\mathbf{u} = \bar{\mathbf{f}} \quad \text{with } \bar{\mathbf{f}} = \Big(f\big(\Psi_\nu^{\mathcal{Y}}\big)\Big)_{\nu\in\vee'}. \tag{3.15}$$

Thus, we can represent the solution of the variational formulation (3.7) as the solution of a bi-infinite matrix equation.

Next, we investigate the structure of the bi-infinite matrix. We have

$$\bar{\mathbf{B}} = \begin{pmatrix}\bar{\mathbf{B}}_1 \\ \bar{\mathbf{B}}_2\end{pmatrix} = \begin{pmatrix}\bar{\mathbf{D}}_{\mathcal{Y}} & 0 \\ 0 & \mathbf{I}_{\mathrm{x}}\end{pmatrix}\begin{pmatrix}\mathbf{T} \\ \mathbf{T}_0\end{pmatrix}\bar{\mathbf{D}}_{\mathcal{X}}, \tag{3.16}$$

where $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ are diagonal scaling matrices with

$$\begin{aligned}(\bar{\mathbf{D}}_{\mathcal{X}})_{(\nu_{\mathrm{t}},\nu_{\mathrm{x}}),(\nu_{\mathrm{t}},\nu_{\mathrm{x}})} &= \bar{s}^{\mathcal{X}}_{\nu_{\mathrm{t}},\nu_{\mathrm{x}}} \qquad \text{for } (\nu_{\mathrm{t}},\nu_{\mathrm{x}}) \in \vee, \\ (\bar{\mathbf{D}}_{\mathcal{Y}})_{(\nu_{\mathrm{t}},\nu_{\mathrm{x}}),(\nu_{\mathrm{t}},\nu_{\mathrm{x}})} &= \bar{s}^{\mathcal{Y}}_{\nu_{\mathrm{x}}} \qquad \text{for } (\nu_{\mathrm{t}},\nu_{\mathrm{x}}) \in \vee\end{aligned} \tag{3.17}$$

and $\mathbf{I}_{\mathrm{x}} = (\delta_{\nu_{\mathrm{x}},\nu'_{\mathrm{x}}})_{\nu_{\mathrm{x}},\nu'_{\mathrm{x}}\in\vee_{\mathrm{x}}}$ is the identity matrix on $\ell_2(\vee_{\mathrm{x}})$. Furthermore, $\mathbf{T}$ and $\mathbf{T}_0$ are given by

$$\begin{aligned}\mathbf{T} &= \big(b(\theta_{\nu_{\mathrm{t}}} \otimes \Psi_{\nu_{\mathrm{x}}}, (\theta_{\nu'_{\mathrm{t}}} \otimes \Psi_{\nu'_{\mathrm{x}}}, 0))\big)_{(\nu'_{\mathrm{t}},\nu'_{\mathrm{x}}),(\nu_{\mathrm{t}},\nu_{\mathrm{x}})\in\vee'}, \\ \mathbf{T}_0 &= \big(b(\theta_{\nu_{\mathrm{t}}} \otimes \Psi_{\nu_{\mathrm{x}}}, (0, \Psi_{\nu'_{\mathrm{x}}}))\big)_{\nu'_{\mathrm{x}}\in\vee_{\mathrm{x}},(\nu_{\mathrm{t}},\nu_{\mathrm{x}})\in\vee}.\end{aligned} \tag{3.18}$$

The right-hand side $\bar{\mathbf{f}}$ has the representation

$$\bar{\mathbf{f}} = \begin{pmatrix}\bar{\mathbf{D}}_{\mathcal{Y}}\mathbf{g}_1 \\ \mathbf{g}_2\end{pmatrix} = \begin{pmatrix}\bar{\mathbf{D}}_{\mathcal{Y}} & 0 \\ 0 & \mathbf{I}_{\mathrm{x}}\end{pmatrix}\begin{pmatrix}\mathbf{g}_1 \\ \mathbf{g}_2\end{pmatrix} \tag{3.19}$$

with

$$\mathbf{g}_1 = \big(f((\theta_{\nu_{\mathrm{t}}} \otimes \Psi_{\nu_{\mathrm{x}}}, 0))\big)_{(\nu_{\mathrm{t}},\nu_{\mathrm{x}})\in\vee'}, \quad \mathbf{g}_2 = (f((0, \Psi_{\nu_{\mathrm{x}}})))_{\nu_{\mathrm{x}}\in\vee_{\mathrm{x}}}. \tag{3.20}$$

3.1.2 Second-Order Problems on High-Dimensional Spatial Domains

Next, we consider the case of a second-order elliptic operator $A(t)$ on the high-dimensional unit cube $\Omega = (0,1)^d$ with spatial dimension d. For the following, we

assume $V = H_0^1(\Omega)$, $H = L_2(\Omega)$, $V' = H^{-1}(\Omega)$ and $A(t)\colon V \to V'$ to be given by

$$A(t)v = -\operatorname{div}_x(\mathbf{M}(t)\nabla_x v) + q(t)\cdot\nabla_x v + c(t)v, \tag{3.21}$$

with suitable diffusion matrix $\mathbf{M}(t) \in \mathbb{R}^{d\times d}$, convection term $q(t) \in \mathbb{R}^d$ and reaction term $c(t) \in \mathbb{R}$.

One primary requirement for our later method is the independence of the condition number of the bi-infinite matrix $\mathbf{B}$ from the dimension d. Based on Theorem 3.5, a possible dependence on the dimension is due to the constants of the Riesz basis of the three spaces H, V, and V'. To avoid such a dependence, we use the following construction of the Riesz basis. Let $\{\psi_\nu\}_{\nu\in\vee_1}$ be an orthonormal basis of $L_2(0,1)$, such that $\{\|\psi_\nu\|^{-1}_{H_0^1(0,1)}\psi_\nu\}_{\nu\in\vee_1}$ is a Riesz basis of $H_0^1(0,1)$ with constants $c_{H_0^1}$ and $C_{H_0^1}$ and $\{\|\psi_\nu\|^{-1}_{H^{-1}(0,1)}\psi_\nu\}_{\nu\in\vee_1}$ is a Riesz basis of $H^{-1}(0,1)$ with constants $c_{H^{-1}}$ and $C_{H^{-1}}$, where $\vee_1$ is a countable index set. Then we set

$$\vee_x = \bigtimes_{i=1}^d \vee_1, \qquad \Psi_\nu = \bigotimes_{i=1}^d \psi_{\nu_i}. \tag{3.22}$$

The space $H = L_2(\Omega)$ has the product structure (2.7) from which it directly follows that $\{\Psi_\nu\}_{\nu\in\vee_x}$ is an orthonormal basis of $L_2(\Omega)$. Thus, we have $c_{\Sigma_H} = C_{\Sigma_H} = 1$. The norm of $V = H_0^1(\Omega)$ is not a cross norm, but V is an intersection of d product spaces similar to $H^1(\Omega)$ in (2.9), i.e.,

$$H_0^1(\Omega) = H_0^1(0,1)\otimes\bigotimes_{i=2}^d L_2(0,1) \cap \cdots \cap \bigotimes_{i=1}^{d-1} L_2(0,1)\otimes H_0^1(0,1).$$

Using the fact that the one-dimensional bases are $L_2(0,1)$-orthonormal and the definition of the intersection norm as in (2.9), we arrive at

$$\|\Psi_\nu\|_{H_0^1(\Omega)} = \Big(\sum_{i=1}^d \|\psi_{\nu_i}\|^2_{H_0^1(0,1)}\Big)^{\frac{1}{2}}. \tag{3.23}$$

From this, we infer the following result from [BF24, Proposition 2.3], also stated in similar form in [DSS09, Section 2].

Proposition 3.7. *Let $H = L_2(\Omega)$ and $V = H_0^1(\Omega)$ so that $V' = H^{-1}(\Omega)$, and let $\{\Psi_\nu\}_{\nu\in\vee_x}$ be given by (3.22). Then we have*

(i) $\{\|\Psi_\nu\|_V^{-1}\Psi_\nu\}_{\nu\in\vee_x}$ *is a Riesz basis of V with constants independent of d.*

(ii) We have

$$\|\Psi_\nu\|_{V'} \eqsim \|\Psi_\nu\|_V^{-1}$$

for all $\nu \in \vee_x$ with constants independent of d and ν.

(iii) $\{\|\Psi_\nu\|_V\Psi_\nu\}_{\nu\in\vee_x}$ *is a Riesz basis of V' with constants independent of d.*

Proof. By the product structure and the L_2-orthonormality of the one-dimensional basis, we obtain

$$\begin{aligned}\Big\|\sum_{\nu\in\Lambda} c_\nu \|\Psi_\nu\|^{-1}_{H^1_0(\Omega)}\Psi_\nu\Big\|^2_{H^1_0(\Omega)} &= \sum_{i=1}^{d}\Big\|\sum_{\nu\in\Lambda} c_\nu \|\Psi_\nu\|^{-1}_{H^1_0(\Omega)}\psi_{\nu_i}\Big\|^2_{H^1_0(0,1)}\\ &\le \sum_{i=1}^{d} C^2_{H^1_0}\sum_{\nu\in\Lambda} c_\nu^2 \|\Psi_\nu\|^{-2}_{H^1_0(\Omega)}\|\psi_{\nu_i}\|^2_{H^1_0(0,1)}\\ &= C^2_{H^1_0}\sum_{\nu\in\Lambda} c_\nu^2,\end{aligned}$$

where we additionally used that $\{\|\psi_\nu\|^{-1}_{H^1_0(0,1)}\psi_\nu\}_{\nu\in\vee_1}$ is a Riesz basis of $H^1_0(0,1)$ in the second line and (3.23) in the last line. Analogously, we can show the lower bound

$$\Big\|\sum_{\nu\in\Lambda} c_\nu \|\Psi_\nu\|^{-1}_{H^1_0(\Omega)}\Psi_\nu\Big\|^2_{H^1_0(\Omega)} \ge c_{H^1_0}\sum_{\nu\in\Lambda} c_\nu^2,$$

which proves the first statement.

Let Ψ_ν be in V. Then we have

$$\begin{aligned}\|\Psi_\nu\|_{V'} &= \sup_{0\neq v\in V}\frac{\langle\Psi_\nu, v\rangle_H}{\|v\|_V}\\ &= \sup_{\substack{\Lambda\subset\vee_x \text{ finite}\\ 0\neq\mathbf{v}\in\mathbb{R}^\Lambda}}\frac{\langle\Psi_\nu, \sum_{\nu'\in\Lambda} v_{\nu'}\Psi_{\nu'}\rangle_H}{\|\sum_{\nu'\in\Lambda} v_{\nu'}\Psi_{\nu'}\|_V}\\ &= \sup_{\substack{\Lambda\subset\vee_x \text{ finite}\\ 0\neq\mathbf{v}\in\mathbb{R}^\Lambda}}\frac{v_\nu}{\|\sum_{\nu'\in\Lambda} v_{\nu'}\Psi_{\nu'}\|_V}\\ &\le \sup_{\substack{\Lambda\subset\vee_x \text{ finite}\\ 0\neq\mathbf{v}\in\mathbb{R}^\Lambda}}\frac{v_\nu}{c_{H^1_0}\Big(\sum_{\nu'\in\Lambda} v^2_{\nu'}\|\Psi_{\nu'}\|^2_V\Big)^{\frac{1}{2}}}\\ &= \frac{1}{c_{H^1_0}}\|\Psi_\nu\|^{-1}_V,\end{aligned}$$

where we used the Gelfand triple property, the density of finite linear combinations, the L_2-orthonormality of the basis, and the Riesz basis property from (i). Using the reverse inequality for the Riesz basis, we get in the same way

$$\|\Psi_\nu\|_{V'} \ge \frac{1}{C_{H^1_0}}\|\Psi_\nu\|^{-1}_V,$$

which shows (ii).

For the last statement, we can proceed similarly to (ii). Let $\Lambda \subset \vee_x$ be a finite subset and $\mathbf{v} \in \mathbb{R}^\Lambda$. Then we have

$$\begin{aligned}
\Big\| \sum_{\nu\in\Lambda} v_\nu \|\Psi_\nu\|_V \Psi_\nu \Big\|_{V'} &= \sup_{0\neq w\in V} \frac{\Big\langle \sum_{\nu\in\Lambda} v_\nu \|\Psi_\nu\|_V \Psi_\nu, w \Big\rangle_H}{\|w\|_V} \\
&= \sup_{\substack{\tilde\Lambda\subset\vee_x \text{ finite} \\ 0\neq \mathbf{w}\in\mathbb{R}^{\tilde\Lambda}}} \frac{\Big\langle \sum_{\nu\in\Lambda} v_\nu \|\Psi_\nu\|_V \Psi_\nu, \sum_{\nu'\in\tilde\Lambda} w_{\nu'} \|\Psi_{\nu'}\|_V^{-1} \Psi_{\nu'} \Big\rangle_H}{\Big\| \sum_{\nu'\in\tilde\Lambda} w_{\nu'} \|\Psi_{\nu'}\|_V^{-1} \Psi_{\nu'} \Big\|_V} \\
&= \sup_{\substack{\tilde\Lambda\subset\vee_x \text{ finite} \\ 0\neq \mathbf{w}\in\mathbb{R}^{\tilde\Lambda}}} \frac{\sum_{\nu\in\Lambda\cap\tilde\Lambda} v_\nu w_\nu}{\Big\| \sum_{\nu'\in\tilde\Lambda} w_{\nu'} \|\Psi_{\nu'}\|_V^{-1} \Psi_{\nu'} \Big\|_V} \\
&\leq \sup_{\substack{\tilde\Lambda\subset\vee_x \text{ finite} \\ 0\neq \mathbf{w}\in\mathbb{R}^{\tilde\Lambda}}} \frac{\sum_{\nu\in\Lambda\cap\tilde\Lambda} v_\nu w_\nu}{c_{H_0^1} \Big(\sum_{\nu'\in\tilde\Lambda} w_{\nu'}^2 \Big)^{\frac{1}{2}}} \\
&= \frac{1}{c_{H_0^1}} \Big(\sum_{\nu'\in\tilde\Lambda} v_{\nu'}^2 \Big)^{\frac{1}{2}},
\end{aligned}$$

where we additionally use the identification of the dual of ℓ_2 with itself in the last line. Proceeding in the same way and using the inverse Riesz inequality in line four yields the other bound and thus (iii). □

The proof of Theorem 3.7 demonstrates that the L_2-orthonormality of the one-dimensional basis is crucial for ensuring that the Riesz constants of the product basis remain independent of the spatial dimension d. As shown in [DSS09], the Riesz constants scale exponentially in d for spline-wavelets with Dirichlet boundary conditions, unless the basis is L_2-orthonormal. Hence, choosing a basis that is not orthonormal leads to exponential growth in the condition number of the bi-infinite matrix $\mathbf{B}$ and thus to exponential growth in the computational costs of the adaptive method. The L_2-orthonormal assumption was also made for the adaptive low-rank methods for elliptic problems in [BD16].

As already mentioned in Section 2.1, we will use L_2-orthonormal spline wavelets with sufficiently many vanishing moments as a one-dimensional basis. Furthermore, we assume that the support size is exponentially decreasing with the wavelet level, that is,

$$|\operatorname{supp} \theta_{\nu_t}| \lesssim 2^{-|\nu_t|}, \quad |\operatorname{supp} \varphi_{\nu_1}| \lesssim 2^{-|\nu_1|}$$

uniformly in $\nu_t \in \vee_t$ and $\nu_1 \in \vee_1$. These properties will be important for the compressibility of the corresponding matrices.

Next, we introduce scaling sequences $\bar{s}^{\mathcal{X}}$ and $\bar{s}^{\mathcal{Y}}$ for the Riesz basis $\mathcal{X}$ and $\mathcal{Y}$ from (3.9) and (3.10) that are suitable for our purposes. We choose $\bar{s}^{\mathcal{Y}}_{\nu_t,\nu_x} = \|\Psi_{\nu_x}\|_V^{-1}$, which will lead to equality in (3.10). This results in

$$\bar{\mathbf{D}}_{\mathcal{Y}} = \mathbf{I}_t \otimes \bar{\mathbf{D}} \tag{3.24}$$

with

$$\bar{\mathbf{D}} = \big(\|\Psi_{\nu_x}\|^{-1}_{H_0^1(\Omega)} \delta_{\nu_x,\nu_x'} \big)_{\nu_x,\nu_x'\in\vee_x} \tag{3.25}$$

and $\mathbf{I}_t = (\delta_{\nu_t,\nu_t'})_{\nu_t,\nu_t'\in\vee_t}$ is the identity matrix on $\ell_2(\vee_t)$ for the time index set $\vee_t$. We will not directly work with the matrices $\bar{\mathbf{D}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ since their entries are not separable and thus the rank of the operator can be infinite. Nevertheless, the scaling matrices can be well approximated by finite sums of low-rank matrices. The matrix $\bar{\mathbf{D}}$ has already been investigated in this regard in [BD16]. This exponential sum approximation of $\bar{\mathbf{D}}$ will be later analyzed in more detail in Section 4.2.

Using Theorem 3.7 (ii), we obtain

$$\begin{aligned}\left(\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + \|\theta_{\nu_t}\|^2_{H^1(I)}\|\Psi_{\nu_x}\|^2_{H^{-1}(\Omega)}\right)^{\frac{1}{2}} &\eqsim \left(\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + \|\theta_{\nu_t}\|^2_{H^1(I)}\|\Psi_{\nu_x}\|^{-2}_{H^1_0(\Omega)}\right)^{\frac{1}{2}}\\ &\eqsim \frac{\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + \|\theta_{\nu_t}\|_{H^1(I)}}{\|\Psi_{\nu_x}\|_{H^1_0(\Omega)}}\end{aligned}$$

with constants independent of d. Thus, for

$$\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} = \frac{\|\Psi_{\nu_x}\|_{H^1_0(\Omega)}}{\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + \|\theta_{\nu_t}\|_{H^1(I)}} \tag{3.26}$$

the relation (3.9) holds true with constants independent of d. The choice of this specific scaling sequence is also motivated by the low-rank approximability of the corresponding scaling matrix $\bar{\mathbf{D}}_{\mathcal{X}}$ defined in (3.17). This is analyzed further in Section 4.2.

In this work, we only consider time-independent second-order elliptic operators $A(t) = A: V \to V'$. Then correspondingly, the bilinear form $a(t;\cdot,\cdot)$ defined in (3.2) is time-independent as well. Hence we write $a(t;v,w) = a(v,w)$. Assuming additionally also an L_2-orthonormal temporal basis, the operator $\mathbf{T}$ in (3.18) has the form

$$\mathbf{T} = \mathbf{I}_t \otimes \mathbf{T}_x + \mathbf{T}_t \otimes \mathbf{I}_x, \tag{3.27}$$

where

$$\begin{aligned}\mathbf{T}_t &= \Big(\int_I \theta'_{\nu_t}(t)\theta_{\nu_t'}(t)\,\mathrm{d}t\Big)_{\nu_t',\nu_t\in\vee_t}\\ \mathbf{T}_x &= \big(|I|a(\Psi_{\nu_x},\Psi_{\nu_x'})\big)_{\nu_x',\nu_x\in\vee_x}\end{aligned} \tag{3.28}$$

and $T = |I|$ is the length of the time interval. We define

$$\begin{aligned}\bar{\mathbf{B}}_x &= \bar{\mathbf{D}}_{\mathcal{Y}}(\mathbf{I}_t \otimes \mathbf{T}_x)\bar{\mathbf{D}}_{\mathcal{X}},\\ \bar{\mathbf{B}}_t &= \bar{\mathbf{D}}_{\mathcal{Y}}(\mathbf{T}_t \otimes \mathbf{I}_x)\bar{\mathbf{D}}_{\mathcal{X}},\end{aligned} \tag{3.29}$$

which yields $\bar{\mathbf{B}}_1 = \bar{\mathbf{B}}_x + \bar{\mathbf{B}}_t$.

In particular, we will consider the heat equation, i.e., $A = -\Delta$. In this case, we have

$$\mathbf{T}_x = \Big(|I|\int_\Omega \nabla\Psi_{\nu_x}\cdot\nabla\Psi_{\nu_x'}\Big)_{\nu_x',\nu_x\in\vee_x}$$

The more general case of a second-order elliptic operator with constant coefficients M, q, and c in (3.21) can be treated precisely as in [BD16].

3.2 Alternative Variational Formulations

For the high-dimensional spatial case, this work focuses primarily on the variational formulation introduced in Section 3.1. However, we still mention some alternative

formulations and their usability for the adaptive method.

3.2.1 Alternative Handling of Initial Condition

Another space-time variational formulation was investigated in [CS11]. In contrast to the variational formulation in Section 3.1, we multiply (3.3) by smooth test functions which vanish at time $t = T$ and apply integration by parts to the first term. Then we arrive at the following variational formulation: find $u \in L_2(I;V)$ such that

$$b(u,v) = f(v) \quad \text{for all } v \in L_2(I;V) \cap \{w \in H^1(I;V') : w(T) = 0\},$$

where

$$\begin{aligned} b(v,w) &= \int_I -\langle v(t), \partial_t w(t)\rangle_H + a(t; v(t), w(t))\,\mathrm{d}t, \\ f(w) &= \int_I \langle g(t), w(t)\rangle_H \,\mathrm{d}t + \langle u_0, w(0)\rangle_H. \end{aligned}$$

This variational formulation was also studied in the special case of the heat equation in [BJ89], where $a(t;\phi,\psi) = \int_\Omega \nabla\phi \cdot \nabla\psi$.

Theorem 3.8 ([CS11]). *The operator $B \in \mathcal{L}(L_2(I;V), (L_2(I;V) \cap \{w \in H^1(I;V') : w(T) = 0\})')$ defined by $(Bv)(w) = b(v,w)$ is boundedly invertible.*

This theorem can be proven in the same way as Theorem 3.3 using the Banach-Nečas-Babuška theorem.

With slight adjustments, we can treat this formulation in the same way as the formulation (3.7) with the methods in Chapter 4. In contrast to the previous variational formulation, the initial condition is incorporated differently, thereby avoiding some technical difficulties (see Section 4.3.4). However, the variational formulation (3.7) yields stronger convergence of the approximate solution, namely in the $\mathcal{X} = L_2(I;V) \cap H^1(I;V')$ norm.

For details on more general settings, in particular different smoothness of the spaces, we refer to [CS11, Theorem 2.4].

3.2.2 Saddle Point Formulation

The two previous variational formulations have different trial and test spaces. Finally, we mention a variational formulation in which the trial and test spaces coincide. For this purpose we follow [And16; SW21] and define

$$A_s = \tfrac{1}{2}(A + A'), \quad A_a = \tfrac{1}{2}(A - A'),$$

which are the symmetric and antisymmetric parts of the operator A defined in (3.2). Furthermore, let

$$C = \partial_t + A_a.$$

In the next step, we use the Brézis-Ekeland-Nayroles variational principle, where we refer to [BE76] and [And12, Section 3.2.4] for details. By applying this principle, we obtain the variational formulation: find $u \in \mathcal{X}$ such that

$$\tilde{b}(u,v) = \tilde{f}(v) \quad \text{for all } v \in \mathcal{X}, \tag{3.30}$$

where

$$\begin{aligned}\tilde{b}(v,w) &= \int_I \langle A_s^{-1}(t)C(t)v(t), C(t)w(t)\rangle_{V\times V'} + \langle A_s(t)v(t), w(t)\rangle_{V'\times V}\,\mathrm{d}t \\ &\quad + \langle v(T), w(T)\rangle_H, \\ \tilde{f}(w) &= \int_I \langle g(t), A_s^{-1}(t)C(t)w(t)\rangle_{V'\times V}\,\mathrm{d}t + \langle u_0, w(0)\rangle_H.\end{aligned}$$

Using the Lax-Milgram theorem, we can show that there exists a unique solution for (3.30). For the ellipticity of the variational formulation, we refer to [And12].

Following [SW21], one can reformulate the variational formulation as a saddle point problem by using the Schur complement. A possible numerical approach for this formulation would be to use a fixed-point iteration for saddle point problems, such as the Uzawa iteration.

Remark 3.9. The method described in Chapter 4 can in general also be adapted to these two variational formulations. For the variational formulation in Section 3.2.1, we have to adjust the bases for the trial and test spaces accordingly and adjust the right-hand side. For the variational formulation in Section 3.2.2 we have to define an adaptive Uzawa iteration similarly as the space-time adaptive Richardson iteration in Section 4.4.

Chapter 4

Adaptive Low-Rank Method

In this chapter, we introduce an adaptive low-rank method for finding an approximation of the solution of the space-time variational formulation in Section 3.1. In particular, we design our method to fulfill several properties. The method should guarantee convergence in the space-time norm $\|\cdot\|_{\mathcal{X}}$ defined in (3.6). Furthermore, we want to avoid that the method suffers from the curse of dimensionality, meaning that, under reasonable assumptions on the problem, the computational complexity should not grow exponentially with the number of spatial dimensions d. To achieve this, we derive estimates of the computational complexity that show we can circumvent the curse of dimensionality under natural assumptions.

The starting point of the method is the representation of the solution of the parabolic PDE by the bi-infinite matrix equation in (3.15). According to the Riesz basis property, we have

$$\|u - \tilde{u}\| \eqsim \|\mathbf{u} - \tilde{\mathbf{u}}\|_{\ell_2(\vee)}, \tag{4.1}$$

where $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are the coefficient vectors of u and $\tilde{u}$ with respect to the Riesz basis of $\mathcal{X}$. Moreover, by construction of the Riesz basis in Section 3.1.2, the constants in (4.1) are independent of the spatial dimension d. Based on this equivalence, we obtain an approximation of u directly from an approximation of $\mathbf{u}$. To find a suitable approximation of $\mathbf{u}$, we use an inexact version of the Richardson iteration for the bi-infinite matrix equation (3.15). The Richardson iteration for solving a linear system of equations $\tilde{\mathbf{A}}\mathbf{u} = \tilde{\mathbf{f}}$ for a symmetric positive definite $\tilde{\mathbf{A}}$ is defined by

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \omega(\tilde{\mathbf{A}}\mathbf{u}_k - \tilde{\mathbf{f}}).$$

For a suitable choice of the relaxation parameter $\omega > 0$ we obtain

$$\|\mathbf{u} - \mathbf{u}_{k+1}\| \leq \rho\|\mathbf{u} - \mathbf{u}_k\|$$

independent of k. However, for a bi-infinite matrix, $\tilde{\mathbf{A}}\mathbf{u}_k$ is generally not finitely supported, even if $\mathbf{u}_k$ itself is finitely supported. Thus we cannot perform the iteration exactly. To overcome this issue, we replace the application of the matrix $\tilde{\mathbf{A}}$ by an approximate application with controlled accuracy, which yields a finitely supported approximation. By a suitable choice of this accuracy, we can still guarantee the convergence of the method.

Additionally, we represent the iterates $\mathbf{u}_k$ in low-rank form. However, the scaling matrices $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ in the representation (3.16) of the bi-infinite matrix $\bar{\mathbf{B}}$ are not compatible with the low-rank structure. Hence, we replace these scaling matrices with equivalent ones suited to the Richardson iteration in low-rank representation.

Moreover, during the iteration, the support and the ranks of the iterates $\mathbf{u}_k$ grow. The computational complexity of the method depends on the support and the ranks

of the iterates. Hence, as intermediate steps, we apply routines that reduce the support and the ranks of the iterates while controlling the approximation error.

The structure of this chapter is as follows. In Section 4.1, we discuss the low-rank representation of the coefficient vectors and introduce methods to reduce the ranks and the supports of low-rank tensors in a controlled way. Next, in Section 4.2 we introduce the equivalent scaling operators. In Section 4.3, we present the adaptive application of the bi-infinite matrix. Then in Section 4.4, we state the inexact Richardson iteration. We analyze the convergence of the method and prove that we can circumvent the curse of dimensionality under certain assumptions. Finally, in Section 4.5, we present numerical results for the proposed method.

4.1 Adaptive Methods for Parabolic PDEs

One of the most important questions is how to handle the time variable in the low-rank method. For example, in the works of [AT15; Boi+19], the time variable is treated as a separate tensor mode. However, this might entail disadvantages related to the approximability of the solution in this format. Additionally, it has some algorithmic issues concerning the low-rank preconditioning. We will comment further on this in Theorem 4.25.

The approach in this work, following [BF24], differs from the earlier methods that treat the time variable as a separate tensor mode. Instead, we use a sparse expansion for the time variable and apply low-rank approximations only in the spatial variable. Let $\mathbf{e}_{\nu_t} \in \ell_2(\vee_t)$ be the Kronecker vector, i.e., $(\mathbf{e}_{\nu_t})_{\mu_t} = \delta_{\nu_t,\mu_t}$ for $\mu_t \in \vee_t$. A given $\mathbf{u} \in \ell_2(\vee)$ with $\vee = \vee_t \times \vee_x$ can be expressed as

$$\mathbf{u} = \sum_{\nu_t \in \vee_t} \mathbf{e}_{\nu_t} \otimes \mathbf{u}_{\nu_t} \quad \text{with } \mathbf{u}_{\nu_t} \in \ell_2(\vee_x), \text{ for all } \nu_t \in \vee_t. \tag{4.2}$$

Furthermore, the *spatial components* $\mathbf{u}_{\nu_t}$ can be represented in the hierarchical tensor format described in Section 2.4.3, independently for each $\nu_t \in \vee_t$, but with a shared dimension tree $\mathbb{T}_d$. The use of the same dimension tree facilitates the addition of different $\mathbf{u}_{\nu_t}$ as described in Section 2.4.5.1, which will be necessary due to the application of the operator $\mathbf{T}_t$ defined in (3.28). Correspondingly, we denote for $\mathbf{u}$ by

$$\operatorname{rank}_{\nu_t}(\mathbf{u}) = \operatorname{rank}(\mathbf{u}_{\nu_t}) \tag{4.3}$$

the ranks of the spatial components $\mathbf{u}_{\nu_t}$ for $\nu_t \in \vee_t$ with rank defined in (2.16). To simplify the notation, we write

$$\operatorname{rank}_\infty\big((r_{\nu_t})_{\nu_t \in \vee_t}\big) = \big\|(|r_{\nu_t}|_\infty)_{\nu_t \in \vee_t}\big\|_{\ell_\infty}$$

for $(r_{\nu_t})_{\nu_t \in \vee_t} \in \mathcal{R}^{\vee_t}$ with $\mathcal{R}$ defined in Theorem 2.14 and

$$\operatorname{rank}_\infty(\mathbf{u}) = \operatorname{rank}_\infty\big((\operatorname{rank}_{\nu_t}(\mathbf{u}))_{\nu_t \in \vee_t}\big).$$

This quantity measures the maximum of the maximum ranks over all spatial components. The maximal rank of each spatial component is important for the computational complexity of the HSVD in (2.20), the storage cost in Theorem 2.17 and for performing operations in hierarchical tensor format in Section 2.4.5. To extend this to the representation (4.2), we are thus interested in the maximum over all spatial components. Let $\mathbf{v} \in \ell_2(\vee)$ be finitely supported and let the spatial components $\mathbf{v}_{\nu_t}$ be hierarchical tensors of size $n_{\nu_t,1} \times \cdots \times n_{\nu_t,d}$. Furthermore, let n_t be the temporal

support, which is equivalent to the number of nonvanishing spatial components. Then according to Theorem 2.17, the total storage required for $\mathbf{v}$ is bounded by

$$\sum_{\nu_t \in \vee_t} \Big((d-1) |\mathrm{rank}_{\nu_t}(\mathbf{v})|_\infty^3 + \sum_{i=1}^{d} n_{\nu_t,i} |\mathrm{rank}_{\nu_t}(\mathbf{v})|_\infty \Big) \\ \leq (d-1) n_t \, \mathrm{rank}_\infty(\mathbf{v})^3 + \mathrm{rank}_\infty(\mathbf{v}) \sum_{i=1}^{d} \sum_{\nu_t \in \vee_t} n_{\nu_t,i}.$$

Hence, we can bound the total storage costs in terms of the maximum rank, the sum of the one-dimensional support sizes and the temporal support size.

A suitable approximation of the solution $\mathbf{u}$ is given in the same format as in (4.2), but with a finite sum representing the sparsity in time and as small as possible ranks representing the low-rank approximations in the spatial variables. The required storage of this approximation, as described above, depends not only on the ranks but also on the size of the one-dimensional supports in the spatial components. Sparsity in these supports is also important. In the remaining part of this section, we will describe methods to gain approximation with near-optimal ranks and near-optimal sparsity in the one-dimensional supports of the spatial components.

Finally, we mention the following relation between the ℓ_2-norm of $\mathbf{u}$ and the ℓ_2-norm of its spatial components: based on the representation (4.2), we have the property

$$\|\mathbf{u}\|_{\ell_2(\vee)}^2 = \sum_{\nu_t \in \vee_t} \|\mathbf{u}_{\nu_t}\|_{\ell_2(\vee_x)}^2. \tag{4.4}$$

4.1.1 Near-Optimal Low-Rank Recompression

The storage and computational costs of the method depend on the ranks of the spatial components. In this section, we state the method from [BF24, Section 3.2.1], which yields a near-optimal low-rank approximation for a given target accuracy $\eta > 0$. Here, near-optimal means that the approximation is optimal up to a constant factor compared to the best approximation. In particular, we want to specify the dependence of the spatial dimension d on this constant factor.

As mentioned in Theorem 2.22, the truncation in HSVD yields near-best approximations for the hierarchical tensor format. In the elliptic case using a single hierarchical tensor representation from [BD16], the approximation quality is quantified by the maximum hierarchical rank. For the parabolic case, as in [BF24], the approximation error is quantified by the ℓ_∞-norm of the maximal ranks of each temporal basis index, as in (4.3).

Definition 4.1. Let $\mathbf{v} \in \ell_2(\vee)$ with spatial components $(\mathbf{v}_{\nu_t})_{\nu_t \in \vee_t}$ in hierarchical tensor format according to Theorem 2.19 and $(\mathsf{r}_{\nu_t})_{\nu_t \in \vee_t} \in \mathcal{R}^{\vee_t}$ with $\mathsf{r}_{\nu_t} = (r_{\nu_t,e})_{e \in \mathbb{E}_d} \in \mathcal{R} = (\mathbb{N}_0 \cup \{\infty\})^{\mathbb{E}_d}$, where $\mathbb{E}_d$ are the effective edges of the dimension tree $\mathbb{T}_d$. Then we define the truncation error bound by

$$\lambda\big((\mathsf{r}_{\nu_t})_{\nu_t \in \vee_t}; \mathbf{v}\big) = \Big(\sum_{\nu_t \in \vee_t} \sum_{e \in \mathbb{E}_d} \sum_{k > r_{\nu_t,e}} \big(\sigma_k^{[e]}(\mathbf{v}_{\nu_t})\big)^2 \Big)^{\frac{1}{2}}.$$

For a given $\eta > 0$, let $\mathsf{r}(\mathbf{v}, \eta) \in \mathcal{R}^{\vee_t}$ be defined such that the truncation error bound is smaller than η, i.e.,

$$\lambda(\mathsf{r}(\mathbf{v}, \eta); \mathbf{v}) \leq \eta,$$

and that the maximal rank over all time indices $\mathrm{rank}_\infty\big(\mathsf{r}(\mathbf{v}, \eta)\big)$ is minimal.

We mention here that $\mathrm{r}(\mathbf{v},\eta)$ is in general not unique.

Let $\mathbf{v} \in \ell_2(\vee)$ be in the form (4.2) and let $\eta > 0$. We define

$$\hat{\mathrm{P}}_\eta(\mathbf{v}) = \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \mathrm{P}_{\mathbf{v}_{\nu_\mathrm{t}},\mathrm{r}_{\nu_\mathrm{t}}(\mathbf{v},\eta)}\mathbf{v}_{\nu_\mathrm{t}},$$

where $\mathrm{P}_{\mathbf{v}_\mathrm{x},\mathrm{r}_\mathrm{x}}$ is the projection defined in (2.22), which applied to $\mathbf{v}_\mathrm{x} \in \ell_2(\vee_\mathrm{x})$ realizes the HSVD with ranks at most $\mathrm{r}_\mathrm{x} \in \mathcal{R}$. Then by definition

$$\|\mathbf{v} - \hat{\mathrm{P}}_\eta(\mathbf{v})\| \le \lambda(\mathrm{r}(\mathbf{v},\eta);\mathbf{v}) \le \eta, \quad \operatorname{rank}_{\nu_\mathrm{t}}\big(\hat{\mathrm{P}}_\eta(\mathbf{v})\big) = \mathrm{r}_{\nu_\mathrm{t}}(\mathbf{v},\eta), \quad \nu_\mathrm{t} \in \vee_\mathrm{t}. \tag{4.5}$$

Next, we introduce the class of representations with bounded ranks $(\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}}$ by

$$\mathcal{F}\big((\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}}\big) = \Big\{ \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \mathbf{v}_{\nu_\mathrm{t}} : \mathbf{v}_{\nu_\mathrm{t}} \in \mathcal{H}(\mathrm{r}_{\nu_\mathrm{t}}) \text{ for all } \nu_\mathrm{t} \in \vee_\mathrm{t} \Big\}$$

with $\mathcal{H}$ defined in (2.15). The quasi-optimality property of the truncation by HSVD in Theorem 2.22 holds for these representations as well with the same constant.

Proposition 4.2 ([BF24, Proposition 3.3])**.** *Let $\mathbf{v} \in \ell_2(\vee)$. Then for $(\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \in \mathcal{R}^{\vee_\mathrm{t}}$, one has*

$$\lambda\big((\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}};\mathbf{v}\big) \le \kappa_\mathrm{P} \inf_{\mathbf{w}\in\mathcal{F}((\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}})} \|\mathbf{w}-\mathbf{v}\|$$

with $\kappa_\mathrm{P} = \sqrt{2d-3}$.

Proof. For any $\mathbf{v}_\mathrm{x} \in \ell_2(\vee_\mathrm{x})$ and $\hat{\mathrm{r}} = (\hat{r}_e)_{e\in\mathbb{E}_d} \in \mathcal{R}$, we have by Theorem 2.22 the estimate

$$\sum_{e\in\mathbb{E}_d}\sum_{k>\hat{r}_e} \big|\sigma_k^{[e]}(\mathbf{v}_\mathrm{x})\big|^2 \le \kappa_\mathrm{P}^2 \inf_{\mathbf{w}\in\mathcal{H}(\hat{\mathrm{r}})} \|\mathbf{w}-\mathbf{v}_\mathrm{x}\|^2.$$

Now let $(\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \subset \mathcal{R}$ and $\mathbf{v} \in \ell_2(\vee)$ in the form (4.2). By applying the previous estimate to each temporal index $\nu_\mathrm{t} \in \vee_\mathrm{t}$, we obtain

$$\begin{aligned}
\lambda^2\big((\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}};\mathbf{v}\big) &\le \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \kappa_\mathrm{P}^2 \inf_{\mathbf{w}_{\nu_\mathrm{t}}\in\mathcal{H}(\mathrm{r}_{\nu_\mathrm{t}})} \|\mathbf{w}_{\nu_\mathrm{t}} - \mathbf{v}_{\nu_\mathrm{t}}\|^2_{\ell_2(\vee_\mathrm{x})} \\
&= \kappa_\mathrm{P}^2 \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \inf_{\mathbf{w}_{\nu_\mathrm{t}}\in\mathcal{H}(\mathrm{r}_{\nu_\mathrm{t}})} \|\mathbf{e}_{\nu_\mathrm{t}} \otimes \mathbf{w}_{\nu_\mathrm{t}} - \mathbf{e}_{\nu_\mathrm{t}} \otimes \mathbf{v}_{\nu_\mathrm{t}}\|^2_{\ell_2(\vee)} \\
&= \kappa_\mathrm{P}^2 \inf_{\mathbf{w}\in\mathcal{F}((\mathrm{r}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}})} \|\mathbf{w}-\mathbf{v}\|^2_{\ell_2(\vee)},
\end{aligned}$$

where we used that the entries on each sum are disjoint and that $\vee = \vee_\mathrm{t} \times \vee_\mathrm{x}$. □

Correspondingly, we define the best approximation error with maximum rank $r \in \mathbb{N}_0$ by

$$\sigma_r^{\mathcal{F}}(\mathbf{v}) = \inf\Big\{ \|\mathbf{v}-\mathbf{w}\| : \mathbf{w} \in \mathcal{F}\big((r_{\nu_\mathrm{t}})_{\nu_\mathrm{t}}\big) \text{ with } (r_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \in \mathcal{R}^{\vee_\mathrm{t}}, \operatorname{rank}_\infty\big((r_{\nu_\mathrm{t}})_{\nu_\mathrm{t}}\big) \le r \Big\}. \tag{4.6}$$

Furthermore, we introduce the best approximation ranks $\bar{\mathrm{r}}(\mathbf{v},\eta) \in \mathcal{R}^{\vee_\mathrm{t}}$ which depend on the sequence $\mathbf{v} \in \ell_2(\vee)$ as well as on the target accuracy $\eta > 0$, by

$$\operatorname{rank}_\infty\big(\bar{\mathrm{r}}(\mathbf{v},\eta)\big) = \min\big\{ r \in \mathbb{N}_0 : \sigma_r^{\mathcal{F}}(\mathbf{v}) \le \eta \big\}. \tag{4.7}$$

Again, the choice of the best approximation ranks may not be unique, since we are only interested in the maximum rank.

The following lemma relates the ranks of the truncation to the best attainable ranks within a particular ball around the solution. It is an analogue of the single hierarchical tensor case in [BD15, Lemma 2].

Lemma 4.3 ([BF24, Lemma 3.4]). *Let $\alpha, \eta > 0$ and let $\mathbf{u}, \mathbf{v} \in \ell_2(\vee)$ satisfy $\|\mathbf{u} - \mathbf{v}\| \leq \eta$. Then we have*

$$\|\mathbf{u} - \hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})\| \leq (1 + \kappa_{\mathrm{P}}(1+\alpha))\eta \tag{4.8}$$

with $\kappa_{\mathrm{P}} = \sqrt{2d-3}$, while

$$\operatorname{rank}_\infty\big(\hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})\big) \leq \operatorname{rank}_\infty\big(\bar{\mathsf{r}}(\mathbf{u}, \alpha\eta)\big). \tag{4.9}$$

Proof. Using the triangle inequality as well as (4.5) and $\|\mathbf{u} - \mathbf{v}\| \leq \eta$, we obtain

$$\begin{aligned}\|\mathbf{u} - \hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})\| &\leq \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v} - \hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})\| \\ &\leq (1 + \kappa_{\mathrm{P}}(1+\alpha))\eta,\end{aligned}$$

which yields (4.8). Moreover, let $\tilde{\mathrm{P}}_{\mathrm{x},\mathbf{u}_\mathrm{x},\mathsf{r}_\mathrm{x}} \colon \ell_2(\vee_\mathrm{x}) \to \mathcal{H}(r_\mathrm{x})$ be a linear projection such that $\tilde{\mathrm{P}}_{\mathrm{x},\mathbf{u}_\mathrm{x},\mathsf{r}_\mathrm{x}}\mathbf{u}_\mathrm{x}$ is a best approximation in $\mathcal{H}(\mathsf{r}_\mathrm{x})$ as in [BD15, Lemma 1], i.e.,

$$\|\mathbf{u}_\mathrm{x} - \tilde{\mathrm{P}}_{\mathrm{x},\mathbf{u}_\mathrm{x},\mathsf{r}_\mathrm{x}}\mathbf{u}_\mathrm{x}\|_{\ell_2(\vee_\mathrm{x})} = \min_{\mathbf{w}\in\mathcal{H}(\mathsf{r}_\mathrm{x})} \|\mathbf{u}_\mathrm{x} - \mathbf{w}\|_{\ell_2(\vee_\mathrm{x})}.$$

Furthermore, let $\mathbf{E}_{\nu_\mathrm{t}} = \mathbf{e}_{\nu_\mathrm{t}}^\top \mathbf{e}_{\nu_\mathrm{t}}$ be the bi-infinite matrix with one at the diagonal entry ν_t and zero elsewhere.

Then for $\mathbf{u} \in \ell_2(\vee)$ and $\mathsf{r} \in \mathcal{R}^{\vee_\mathrm{t}}$, the operator $\tilde{\mathrm{P}}_{\mathbf{u},\mathsf{r}} \colon \ell_2(\vee) \to \mathcal{F}(\mathsf{r})$ of the form

$$\tilde{\mathrm{P}}_{\mathbf{u},\mathsf{r}} = \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \mathbf{E}_{\nu_\mathrm{t}} \otimes \tilde{\mathrm{P}}_{\mathrm{x},\mathbf{u}_{\nu_\mathrm{t}},\mathsf{r}_{\nu_\mathrm{t}}}$$

has the property

$$\|\mathbf{u} - \tilde{\mathrm{P}}_{\mathbf{u},\mathsf{r}}\mathbf{u}\|_{\ell_2(\vee)} = \min_{\mathbf{w}\in\mathcal{F}(\mathsf{r})} \|\mathbf{u} - \mathbf{w}\|_{\ell_2(\vee)}.$$

Furthermore, $\tilde{\mathrm{P}}_{\mathbf{u},\mathsf{r}}$ is a linear projection. Let $\mathbf{I} \colon \ell_2(\vee) \to \ell_2(\vee)$ be the identity operator. According to Theorem 4.2, we obtain

$$\begin{aligned}\lambda(\bar{\mathsf{r}}(\mathbf{u}, \alpha\eta); \mathbf{v}) &\leq \kappa_P \inf_{\mathbf{w}\in\mathcal{F}(\bar{\mathsf{r}}(\mathbf{u},\alpha\eta))} \|\mathbf{v} - \mathbf{w}\| \\ &\leq \kappa_P \|\mathbf{v} - \tilde{\mathrm{P}}_{\mathbf{u},\bar{\mathsf{r}}(\mathbf{u},\alpha\eta)}\mathbf{v}\| \\ &\leq \kappa_P \big(\|(\mathbf{I} - \tilde{\mathrm{P}}_{\mathbf{u},\bar{\mathsf{r}}(\mathbf{u},\alpha\eta)})(\mathbf{v} - \mathbf{u})\| + \|\mathbf{u} - \tilde{\mathrm{P}}_{\mathbf{u},\bar{\mathsf{r}}(\mathbf{u},\alpha\eta)}\mathbf{u}\|\big) \\ &\leq \kappa_P(1+\alpha)\eta.\end{aligned} \tag{4.10}$$

Moreover, by (4.5), we have

$$\operatorname{rank}_{\nu_\mathrm{t}}\big(\hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})\big) = \mathsf{r}_{\nu_\mathrm{t}}(\mathbf{v}, \kappa_{\mathrm{P}}(1+\alpha)\eta), \quad \nu_\mathrm{t} \in \vee_\mathrm{t}.$$

By definition, $\mathsf{r}(\mathbf{v}, \kappa_{\mathrm{P}}(1+\alpha)\eta)$ is chosen such that $\operatorname{rank}_\infty\big(\mathsf{r}(\mathbf{v}, \kappa_{\mathrm{P}}(1+\alpha)\eta)\big)$ is the minimal value such that the bound $\lambda(\mathsf{r}(\mathbf{v}, \kappa_{\mathrm{P}}(1+\alpha)\eta), \mathbf{v}) \leq \kappa_{\mathrm{P}}(1+\alpha)\eta$ is fulfilled. Combining this minimality condition with (4.10) yields (4.9). □

Next, we introduce approximation classes, again in the same way as for single low-rank approximations in [BD15, Definition 4].

Definition 4.4. We call a strictly increasing, positive sequence $\gamma = \big(\gamma(n)\big)_{n\in\mathbb{N}_0}$ with $\gamma(0) = 1$ and $\gamma(n) \to \infty$ as $n \to \infty$ a *growth sequence*. Given a growth sequence γ, we define

$$\|\mathbf{v}\|_{\mathcal{A}(\gamma)} = \sup_{r\in\mathbb{N}_0} \gamma(r)\sigma_r^{\mathcal{F}}(\mathbf{v}), \quad \mathcal{A}(\gamma) = \{\mathbf{v} \in \ell_2(\vee) : \|\mathbf{v}\|_{\mathcal{A}(\gamma)} < \infty\}$$

with $\sigma_r^{\mathcal{F}}(\mathbf{v})$ defined in (4.6). By $\gamma^{-1}\colon \mathbb{R}^+ \to \mathbb{N}_0$, we denote the left-continuous inverse of γ, meaning

$$\gamma^{-1}(y) = \min\{n \in \mathbb{N}_0 : \gamma(n) \geq y\}.$$

Furthermore, we call the growth sequence γ *admissible* if

$$\rho_\gamma = \sup_{n\in\mathbb{N}} \frac{\gamma(n)}{\gamma(n-1)} < \infty.$$

Each sequence with at most exponential growth is an admissible growth sequence.

Remark 4.5. Let $\mathbf{v} \in \mathcal{A}(\gamma)$, $\eta > 0$ and r be minimal such that $\sigma_r^{\mathcal{F}}(\mathbf{v}) \leq \eta$ with $\sigma_r^{\mathcal{F}}$ from (4.6). Then we have

$$\|\mathbf{v}\|_{\mathcal{A}(\gamma)} \geq \gamma(r-1)\sigma_{r-1}^{\mathcal{F}}(\mathbf{v}) > \gamma(r-1)\eta.$$

Using that γ is a growth sequence and the definition of the left-continuous inverse, we can estimate the maximum rank needed to achieve accuracy η by

$$r \leq \gamma^{-1}(\|\mathbf{v}\|_{\mathcal{A}(\gamma)}\eta^{-1}).$$

Furthermore, if we have a rank bound of the form

$$\tilde{r} \leq \gamma^{-1}(C\|\mathbf{v}\|_{\mathcal{A}(\gamma)}\eta^{-1}) \tag{4.11}$$

for some constant $C > 0$. Then using that γ is admissible, we obtain

$$\tilde{r} \leq \gamma^{-1}(\|\mathbf{v}\|_{\mathcal{A}(\gamma)}\eta^{-1}) + k$$

if k is chosen such that $\rho_\gamma^k \geq C$. Hence, bounds of the form (4.11) are near-optimal.

The following result is an extension of Theorem 4.3, where we additionally assume that $\mathbf{u} \in \mathcal{A}(\gamma)$. For the proof, we follow the idea of [BD15, Theorem 6], where this statement was proven for the single hierarchical tensor case and a slightly modified definition of $\|\cdot\|_{\mathcal{A}(\gamma)}$.

Theorem 4.6 ([BF24, Theorem 3.6])**.** *Let $\kappa_{\mathrm{P}} = \sqrt{2d-3}$ and $\alpha > 0$. Assume that $\mathbf{u} \in \mathcal{A}(\gamma)$ with a growth sequence γ and that $\mathbf{v} \in \ell_2(\vee)$ satisfies $\|\mathbf{u} - \mathbf{v}\| \leq \eta$ for some $\eta > 0$. Then defining $\mathbf{w}_\eta = \hat{\mathrm{P}}_{\kappa_{\mathrm{P}}(1+\alpha)\eta}(\mathbf{v})$, we have*

$$\|\mathbf{u} - \mathbf{w}_\eta\| \leq (1 + \kappa_{\mathrm{P}}(1+\alpha))\eta, \tag{4.12}$$

$$\operatorname{rank}_\infty(\mathbf{w}_\eta) \leq \gamma^{-1}\big(\|\mathbf{u}\|_{\mathcal{A}(\gamma)}\alpha^{-1}\eta^{-1}\big), \tag{4.13}$$

and

$$\|\mathbf{w}_\eta\|_{\mathcal{A}(\gamma)} \leq \big(\alpha^{-1}(1 + \kappa_{\mathrm{P}}(1+\alpha)) + 1\big)\|\mathbf{u}\|_{\mathcal{A}(\gamma)}. \tag{4.14}$$

Proof. The bound (4.12) was already shown in Theorem 4.3. Based on (4.9), we have

$$\operatorname{rank}_\infty(\mathbf{w}_\eta) \leq \operatorname{rank}_\infty\big(\bar{\mathrm{r}}(\mathbf{u}, \alpha\eta)\big).$$

Hence, for $\operatorname{rank}_\infty\big(\bar{r}(\mathbf{u},\alpha\eta)\big) = 0$, the statement (4.13) is trivial. Next, we consider $\tilde{r} = \operatorname{rank}_\infty\big(\bar{r}(\mathbf{u},\alpha\eta)\big) \geq 1$. Based on the definition of $\|\cdot\|_{\mathcal{A}(\gamma)}$ and on the minimality condition (4.6), we have

$$\|\mathbf{u}\|_{\mathcal{A}(\gamma)} \geq \gamma(\tilde{r}-1)\sigma^{\mathcal{F}}_{\tilde{r}-1}(\mathbf{u}) > \gamma(\tilde{r}-1)\alpha\eta.$$

By the definition of the left-continuous inverse, this yields (4.13).

Let $r < \operatorname{rank}_\infty(\bar{r}(\mathbf{u},\alpha\eta))$. Then we have

$$\begin{aligned}\sigma^{\mathcal{F}}_r(\mathbf{w}_\eta) &\leq \inf_{\substack{\mathrm{r}\in\mathcal{R}^{\vee_\mathrm{t}}\\ \operatorname{rank}_\infty(\mathrm{r})\leq r}} \|\mathbf{w}_\eta - \tilde{\mathrm{P}}_{\mathbf{u},\mathrm{r}}\mathbf{u}\| \\ &\leq \|\mathbf{w}_\eta - \mathbf{u}\| + \inf_{\substack{\mathrm{r}\in\mathcal{R}^{\vee_\mathrm{t}}\\ \operatorname{rank}_\infty(\mathrm{r})\leq r}} \|\mathbf{u} - \tilde{\mathrm{P}}_{\mathbf{u},\mathrm{r}}\mathbf{u}\| \\ &\leq (1+\kappa_\mathrm{P}(1+\alpha))\eta + \sigma^{\mathcal{F}}_r(\mathbf{u}),\end{aligned}$$

where we have used (4.12) and the definition of $\sigma^{\mathcal{F}}_r(\mathbf{u})$ from (4.6) in the last line. Based on the minimality condition (4.7) on $\bar{r}(\mathbf{u},\alpha\eta)$ we know that $\sigma^{\mathcal{F}}_r(\mathbf{u}) > \alpha\eta$. Thus

$$\begin{aligned}\gamma(r)\sigma^{\mathcal{F}}_r(\mathbf{w}_\eta) &\leq \gamma(r)\alpha^{-1}(1+\kappa_\mathrm{P}(1+\alpha))\alpha\eta + \gamma(r)\sigma^{\mathcal{F}}_r(\mathbf{u}) \\ &\leq (\alpha^{-1}(1+\kappa_\mathrm{P}(1+\alpha))+1)\gamma(r)\sigma^{\mathcal{F}}_r(\mathbf{u}) \\ &\leq (\alpha^{-1}(1+\kappa_\mathrm{P}(1+\alpha))+1)\|\mathbf{u}\|_{\mathcal{A}(\gamma)}.\end{aligned}$$

Furthermore, by (4.9) of Theorem 4.3, we have $\sigma^{\mathcal{F}}_r(\mathbf{w}_\eta) = 0$ for $r \geq \operatorname{rank}_\infty(\bar{r}(\mathbf{u},\alpha\eta))$. Combining this with the definition of $\|\cdot\|_{\mathcal{A}(\gamma)}$ yields (4.14).

□

4.1.2 Contractions and Coarsening

As seen in Theorem 2.17, the storage of the hierarchical low-rank approximation depends not only on the ranks, but also on the size of the mode frames, which in our case is equivalent to the size of the one-dimensional supports of the spatial components. In this section, we describe the method from [BF24, Section 3.2.2] on how to choose a finite number of basis functions for a suitable approximation, which is an adaptation of the single hierarchical tensor case in [BD15; BD16].

In the following, we always assume that we have the same index set $\vee_1$ for each spatial dimension and correspondingly $\vee_\mathrm{x} = \bigtimes_{i=1}^d \vee_1$ and $\vee = \vee_\mathrm{t} \times \vee_\mathrm{x}$.

The basic idea is to use the concept of best N-term approximation from Section 2.2 applied to *tensor contractions*. These tensor contractions measure for a fixed dimension $i \in \{1,\dots,d\}$ the influence of a one-dimensional basis function corresponding to an index $\nu_i \in \vee_1$ by summing over all indices $(\mu_1,\dots,\mu_d) \in \vee_x$ where $\mu_i = \nu_i$. This procedure was first used in [BD15; BD16] for the single hierarchical tensor case for elliptic problems and later extended to the parabolic case in [BF24]. In the latter, this concept is applied for a given $\mathbf{u} \in \ell_2(\vee)$ in (4.2) independently to each spatial component $\mathbf{u}_{\nu_\mathrm{t}}$ for $\nu_\mathrm{t} \in \vee_\mathrm{t}$, which yields a set of contractions over all time indices, all one-dimensional spatial indices, and all spatial dimensions.

Definition 4.7. For $\mathbf{v}_\mathrm{x} \in \ell_2(\vee_\mathrm{x})$, we define the spatial contractions

$$\pi^{(i)}(\mathbf{v}_\mathrm{x}) = \big(\pi^{(i)}_\mu(\mathbf{v}_\mathrm{x})\big)_{\mu\in\vee_1} \in \ell_2(\vee_1)$$

by the ℓ_2 norm of the entries where the index $\mu \in \vee_1$ is active in the i-th component, which is

$$\pi_\mu^{(i)}(\mathbf{v}_\mathrm{x}) = \left(\sum_{\substack{\nu \in \vee_\mathrm{x} \\ \nu_i = \mu}} (\mathbf{v}_\mathrm{x})_\nu^2 \right)^{\frac{1}{2}} = \left(\sum_{\nu_1 \in \vee_1} \cdots \sum_{\nu_{i-1} \in \vee_1} \sum_{\nu_{i+1} \in \vee_1} \cdots \sum_{\nu_d \in \vee_1} (\mathbf{v}_\mathrm{x})_{\nu_1,\ldots,\nu_{i-1},\mu,\nu_{i+1},\ldots,\nu_d}^2 \right)^{\frac{1}{2}}. \tag{4.15}$$

Furthermore, let $\mathbf{v} \in \ell_2(\vee)$ with spatial components $(\mathbf{v}_{\nu_\mathrm{t}})_{\nu_\mathrm{t} \in \vee_\mathrm{t}}$. Then we define the spatio-temporal contractions as

$$\pi^{(\mathrm{t},i)}(\mathbf{v}) = \left(\pi_\mu^{(i)}(\mathbf{v}_{\nu_\mathrm{t}}) \right)_{(\nu_\mathrm{t},\mu) \in \vee_\mathrm{t} \times \vee_1}, \quad i = 1, \ldots, d, \tag{4.16}$$

as well as the temporal contractions

$$\pi^{(\mathrm{t})}(\mathbf{v}) = \left(\pi_{\nu_\mathrm{t}}^{(\mathrm{t})}(\mathbf{v}) \right)_{\nu_\mathrm{t} \in \vee_\mathrm{t}} = \left(\|\mathbf{v}_{\nu_\mathrm{t}}\| \right)_{\nu_\mathrm{t} \in \vee_\mathrm{t}}.$$

However, calculating the contractions using (4.15) is not feasible for our approach, since it requires a summation over a $d-1$-dimensional index set. This would again lead to a curse of dimensionality. Nevertheless, an observation from [BD15] shows that this can be circumvented by using orthogonality properties of the HSVD representation. This can also be applied in this context, since spatio-temporal contractions are the spatial contractions of the spatial components. For the proof, we follow mainly the lines of [BD15, Proposition 2].

Proposition 4.8 ([BF24, Proposition 3.10]). *Let $\mathbf{v} \in \ell_2(\vee)$ with corresponding spatial components $(\mathbf{v}_{\nu_\mathrm{t}})_{\nu_\mathrm{t} \in \vee_\mathrm{t}}$.*

(i) *We have $\|\mathbf{v}_{\nu_\mathrm{t}}\| = \|\pi^{(i)}(\mathbf{v}_{\nu_\mathrm{t}})\|$ for $i = 1, \ldots, d$ and $\nu_\mathrm{t} \in \vee_\mathrm{t}$.*

(ii) *We have $\|\mathbf{v}\| = \left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|$ for $i = 1, \ldots, d$.*

(iii) *Let $\Lambda_{\nu_\mathrm{t}}^{(i)} \subseteq \vee_1$ for each $i = 1, \ldots, d$ and $\nu_\mathrm{t} \in \vee_\mathrm{t}$. Moreover, let*

$$\Lambda = \bigcup_{\nu_\mathrm{t} \in \vee_\mathrm{t}} \{\nu_\mathrm{t}\} \times \Lambda_{\nu_\mathrm{t}}^{(1)} \times \cdots \times \Lambda_{\nu_\mathrm{t}}^{(d)},$$

then we have

$$\|\mathbf{v} - \mathrm{R}_\Lambda \mathbf{v}\| \leq \left(\sum_{\nu_\mathrm{t} \in \vee_\mathrm{t}} \sum_{i=1}^{d} \sum_{\nu_1 \in \vee_1 \setminus \Lambda_{\nu_\mathrm{t}}^{(i)}} \left| \pi_{\nu_\mathrm{t},\nu_1}^{(\mathrm{t},i)}(\mathbf{v}) \right|^2 \right)^{\frac{1}{2}} \tag{4.17}$$

with the restriction operator R_Λ defined in (2.5).

(iv) *For $i = 1, \ldots, d$, let in addition $\mathbf{U}_{\nu_\mathrm{t}}^{(i)}$ be the mode frames and $\sigma_k^{(i)}(\mathbf{v}_{\nu_\mathrm{t}})$ be the sequence of singular values of an HSVD of the spatial components $\mathbf{v}_{\nu_\mathrm{t}}$ according to Theorem 2.19. Then*

$$\pi_{\nu_\mathrm{t},\nu_1}^{(\mathrm{t},i)}(\mathbf{v}) = \left(\sum_k \left| \left(\mathbf{U}_{\nu_\mathrm{t}}^{(i)}\right)_{\nu_1,k} \right|^2 \left| \sigma_k^{(i)}(\mathbf{v}_{\nu_\mathrm{t}}) \right|^2 \right)^{\frac{1}{2}}.$$

Proof. Property (i) is a direct consequence of the definition of the contractions. Combining this with (4.4) yields property (ii). Property (iv) is a direct consequence of Theorem 2.19 with the orthonormality of the mode frames and the definition of the

singular values, as well as (2.18). For more details, we refer to [BD15, Proposition 2]. Let $\Lambda_{\nu_t} = \Lambda_{\nu_t}^{(1)} \times \cdots \times \Lambda_{\nu_t}^{(d)}$. Then we have

$$\begin{aligned}
\|\mathbf{v} - \mathrm{R}_\Lambda \mathbf{v}\|^2 &= \Big\| \sum_{\nu_t \in \vee_t} \mathbf{e}_{\nu_t} \otimes \big(\mathbf{v}_{\nu_t} - \mathrm{R}_{\Lambda_{\nu_t}} \mathbf{v}_{\nu_t}\big) \Big\|^2 \\
&= \sum_{\nu_t \in \vee_t} \|\mathbf{v}_{\nu_t} - \mathrm{R}_{\Lambda_{\nu_t}} \mathbf{v}_{\nu_t}\|^2 \\
&\leq \sum_{\nu_t \in \vee_t} \big\|\mathbf{v}_{\nu_t} - \mathrm{R}_{\Lambda_{\nu_t}^{(1)} \times \vee_1 \times \cdots \times \vee_1} \mathbf{v}_{\nu_t}\big\|^2 + \cdots + \big\|\mathbf{v}_{\nu_t} - \mathrm{R}_{\vee_1 \times \cdots \times \vee_1 \times \Lambda_{\nu_t}^{(d)}} \mathbf{v}_{\nu_t}\big\|^2 \\
&= \sum_{\nu_t \in \vee_t} \sum_{i=1}^{d} \sum_{\nu_1 \in \vee_1 \setminus \Lambda_{\nu_t}^{(i)}} \big|\pi_{\nu_t,\nu_1}^{(\mathrm{t},i)}(\mathbf{v})\big|^2,
\end{aligned}$$

which yields (iii). □

Property (iv) of Theorem 4.8 offers us a possibility to efficiently compute the contractions $\pi^{(\mathrm{t},i)}(\mathbf{v})$ for $\mathbf{v} \in \ell_2(\vee)$. First, we compute the HSVD of all spatial components. Then we sum up at most $\operatorname{rank}_\infty(\mathbf{v})$ summands to calculate $\pi_{\nu_t,\nu_1}^{(\mathrm{t},i)}(\mathbf{v})$ for each $\nu_t \in \vee_t$ and $\nu_1 \in \vee_1$. Hence, no summation over high-dimensional index sets is needed. Moreover, since the calculation of an HSVD is needed for the truncation procedure anyway, the overall computational costs are not influenced.

The estimate (4.17) can alternatively be written as

$$\|\mathbf{v} - \mathrm{R}_\Lambda \mathbf{v}\| \leq \left(\sum_{i=1}^{d} \big\|\pi^{(\mathrm{t},i)}(\mathbf{v}) - \mathrm{R}_{\Lambda^{(i)}} \pi^{(\mathrm{t},i)}(\mathbf{v})\big\|^2 \right)^{\frac{1}{2}},$$

where $\Lambda^{(i)}$ is defined by

$$\Lambda^{(i)} = \bigcup_{\nu_t \in \vee_t} \big\{(\nu_t, \nu_1) : \nu_1 \in \Lambda_{\nu_t}^{(i)}\big\}.$$

As a direct consequence, we have

$$\|\mathbf{v} - \mathrm{R}_\Lambda \mathbf{v}\| \leq \sum_{i=1}^{d} \big\|\pi^{(\mathrm{t},i)}(\mathbf{v}) - \mathrm{R}_{\Lambda^{(i)}} \pi^{(\mathrm{t},i)}(\mathbf{v})\big\|. \tag{4.18}$$

We have the following subadditivity properties for the contractions, which are a direct consequence of the triangle inequality.

Proposition 4.9. *Let $N \in \mathbb{N}$ and $i \in \{1, \ldots, d\}$ as well as $\nu_t \in \vee_t$ and $\nu_1 \in \vee$. Then we have*

$$\begin{aligned}
\pi_{\nu_1}^{(i)}\Big(\sum_{n=1}^{N} \mathbf{v}_n\Big) &\leq \sum_{n=1}^{N} \pi_{\nu_1}^{(i)}(\mathbf{v}_n) && \text{for } \mathbf{v}_1, \ldots, \mathbf{v}_n \in \ell_2(\vee_x), \\
\pi_{\nu_t,\nu_1}^{(\mathrm{t},i)}\Big(\sum_{n=1}^{N} \mathbf{v}_n\Big) &\leq \sum_{n=1}^{N} \pi_{\nu_t,\nu_1}^{(\mathrm{t},i)}(\mathbf{v}_n) && \text{for } \mathbf{v}_1, \ldots, \mathbf{v}_n \in \ell_2(\vee), \\
\pi_{\nu_t}^{(\mathrm{t})}\Big(\sum_{n=1}^{N} \mathbf{v}_n\Big) &\leq \sum_{n=1}^{N} \pi_{\nu_t}^{(\mathrm{t})}(\mathbf{v}_n) && \text{for } \mathbf{v}_1, \ldots, \mathbf{v}_n \in \ell_2(\vee).
\end{aligned}$$

Next, we define the coarsening procedure. As in the single hierarchical tensor case [BD15], the coarsening procedure is based on the best N-term approximation of the set of all contractions over all spatial dimensions. This means that we do not consider the contractions of each spatial dimension separately, but rather choose

the N largest contractions from all dimensions together. We consider the set of all spatio-temporal contractions over all dimensions

$$\left\{\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_1}(\mathbf{v}) : \nu_1 \in \vee_1, \nu_\mathrm{t} \in \vee_\mathrm{t}, i = 1,\dots,d\right\}.$$

We determine indices $i_j \in \{1,\dots,d\}$, $\nu_{\mathrm{t},j} \in \vee_\mathrm{t}$ and $\nu_{1,j} \in \vee_1$ such that we have a non-decreasing sequence

$$\pi^{(\mathrm{t},i_1)}_{\nu_{\mathrm{t},1},\nu_{1,1}}(\mathbf{v}) \geq \pi^{(\mathrm{t},i_2)}_{\nu_{\mathrm{t},2},\nu_{1,2}}(\mathbf{v}) \geq \ldots \geq \pi^{(\mathrm{t},i_j)}_{\nu_{\mathrm{t},j},\nu_{1,j}}(\mathbf{v}) \geq \ldots.$$

From this ordering, we keep the N largest terms and build from the corresponding indices the index sets for each temporal index and each dimension by

$$\Lambda^{(i)}_{\nu_\mathrm{t}}(\mathbf{v},N) = \left\{\nu_{1,j} : i_j = i, \nu_{\mathrm{t},j} = \nu_\mathrm{t}, j = 1,\dots,N\right\} \subset \vee_1 .$$

For each time index, we build the Cartesian product sets

$$\Lambda_{\nu_\mathrm{t}}(\mathbf{v},N) = \bigtimes_{i=1}^{d} \Lambda^{(i)}_{\nu_\mathrm{t}}(\mathbf{v},N) \subset \vee_\mathrm{x}$$

for all time indices $\nu_\mathrm{t} \in \vee_\mathrm{t}$ and finally the set

$$\Lambda(\mathbf{v},N) = \bigcup_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \left(\{\nu_\mathrm{t}\} \times \Lambda_{\nu_\mathrm{t}}(\mathbf{v},N)\right) \subset \vee. \tag{4.19}$$

Hence, for each time index, we have a Cartesian product structure over the spatial variables, but the spatial index sets may differ between time indices. This structure fits perfectly into our framework, since we can have different hierarchical tensor representations for each temporal index $\nu_\mathrm{t} \in \vee_\mathrm{t}$.

Due to the product structure of the index set, it can happen that $\Lambda_{\nu_\mathrm{t}}(\mathbf{v},N) = \varnothing$ even if $\Lambda^{(i)}_{\nu_\mathrm{t}}(\mathbf{v},N) \neq \varnothing$ for multiple $i = 1,\dots,d$. Nevertheless, by construction

$$\sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \sum_{i=1}^{d} \#\Lambda^{(i)}_{\nu_\mathrm{t}}(\mathbf{v},N) \leq N$$

and

$$\sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \sum_{i=1}^{d} \sum_{\nu_1\in(\vee_1\setminus\Lambda^{(i)}_{\nu_\mathrm{t}}(\mathbf{v},N))} |\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_1}(\mathbf{v})|^2 = \min_{\hat\Lambda}\left(\sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \sum_{i=1}^{d} \sum_{\nu_1\in(\vee_1\setminus\hat\Lambda^{(i)}_{\nu_\mathrm{t}})} |\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_1}(\mathbf{v})|^2\right), \tag{4.20}$$

where the minimum is taken over all sets $\hat\Lambda$, which are of the form

$$\hat\Lambda = \bigcup_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \{\nu_\mathrm{t}\} \times \bigtimes_{i=1}^{d} \hat\Lambda^{(i)}_{\nu_\mathrm{t}}, \quad \text{where} \quad \sum_{\nu_\mathrm{t}\in\vee_\mathrm{t}} \sum_{i=1}^{d} \#\hat\Lambda^{(i)}_{\nu_\mathrm{t}} \leq N . \tag{4.21}$$

We are primarily interested in the approximation properties of the coarsening routine. For this purpose, we compare the approximation error to the best approximation error that is achievable using N one-dimensional basis functions. We define the index set $\bar\Lambda(\mathbf{v},N)$ such that it fulfills (4.21) as well as

$$\|\mathbf{v} - \mathrm{R}_{\bar\Lambda(\mathbf{v},N)}\mathbf{v}\| = \min_{\sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}))\leq N} \|\mathbf{v} - \mathbf{w}\|. \tag{4.22}$$

The following proposition shows that the coarsening procedure is quasi-optimal.

Proposition 4.10 ([BF24, Proposition 3.12]). *For any $\mathbf{v} \in \ell_2(\vee)$ with spatial components $(\mathbf{v}_{\nu_t})_{\nu_t \in \vee_t}$ we have*

$$\|\mathbf{v} - \mathrm{R}_{\Lambda(\mathbf{v},N)}\mathbf{v}\| \leq \mu_N(\mathbf{v}) \tag{4.23}$$

with $\Lambda(\mathbf{v}, N)$ defined in (4.19), where

$$\mu_N(\mathbf{v}) = \left(\sum_{\nu_t \in \vee_t} \sum_{i=1}^{d} \sum_{\nu_1 \in (\vee_1 \setminus \Lambda^{(i)}_{\nu_t}(\mathbf{v},N))} |\pi^{(t,i)}_{\nu_t,\nu_1}(\mathbf{v})|^2 \right)^{\frac{1}{2}}.$$

Furthermore

$$\mu_N(\mathbf{v}) \leq \kappa_C \|\mathbf{v} - \mathrm{R}_{\bar{\Lambda}(\mathbf{v},N)}\mathbf{v}\| \tag{4.24}$$

with $\kappa_C = \sqrt{d}$, where $\bar{\Lambda}(\mathbf{v}, N)$ is defined by (4.22).

Proof. The estimate (4.23) is an immediate consequence of the property (4.17) from Theorem 4.8. Let $\hat{\Lambda}$ satisfy (4.21). Then using (4.20), we obtain

$$\begin{aligned}
(\mu_N(\mathbf{v}))^2 &= \sum_{\nu_t \in \vee_t} \sum_{i=1}^{d} \sum_{\nu_1 \in (\vee_1 \setminus \hat{\Lambda}^{(i)}_{\nu_t})} |\pi^{(t,i)}_{\nu_t,\nu_1}(\mathbf{v})|^2 \\
&= \sum_{\nu_t \in \vee_t} \left\| \mathbf{v}_{\nu_t} - \mathrm{R}_{\hat{\Lambda}^{(1)}_{\nu_t} \times \vee_1 \times \cdots \times \vee_1} \mathbf{v}_{\nu_t} \right\|^2 + \cdots + \left\| \mathbf{v}_{\nu_t} - \mathrm{R}_{\vee_1 \times \cdots \times \vee_1 \times \hat{\Lambda}^{(d)}_{\nu_t}} \mathbf{v}_{\nu_t} \right\|^2 \\
&\leq \sum_{\nu_t \in \vee_t} d \left\| \mathbf{v}_{\nu_t} - \mathrm{R}_{\hat{\Lambda}_{\nu_t}} \mathbf{v}_{\nu_t} \right\|^2 \\
&= d \|\mathbf{v} - \mathrm{R}_{\hat{\Lambda}}\mathbf{v}\|^2,
\end{aligned}$$

where we used (4.4) in the last line. Since $\hat{\Lambda}$ was arbitrary, the statement holds in particular for $\hat{\Lambda} = \bar{\Lambda}(\mathbf{v}, N)$, which yields (4.24). □

Remark 4.11. Ideally, we want a coarsening algorithm with a runtime that depends linearly on the sum of the one-dimensional supports over all time indices

$$n = \sum_{i=1}^{d} \#\mathrm{supp}(\pi^{(t,i)}(\mathbf{v})).$$

Exact sorting algorithms, such as quicksort, have a runtime of $\mathcal{O}(n \log n)$ to sort n elements. To circumvent this, the sorting can be replaced by a *quasi-sorting* using binary binning. In this approach, we do not assign an exact order to the entries. Instead, we arrange the entries into specific bins that represent certain value ranges. Quasi-sorting have also used in the coarsening procedure for the one-dimensional case. For details, we refer to [Bar05, Section 1.2] and [DSS09, Section 4.2].

In the adaptive procedure, we want to apply the coarsening routine such that the approximation has a given accuracy $\eta > 0$. For this purpose, we define

$$N(\mathbf{v}, \eta) = \min\{N : \mu_N(\mathbf{v}) \leq \eta\} \tag{4.25}$$

and the corresponding coarsening operator by

$$\hat{\mathrm{C}}_\eta(\mathbf{v}) = \mathrm{R}_{\Lambda(\mathbf{v},N(\mathbf{v},\eta))}\mathbf{v}. \tag{4.26}$$

Due to (4.23) in Theorem 4.10, this choice ensures

$$\|\mathbf{v} - \hat{\mathrm{C}}_\eta(\mathbf{v})\| \leq \eta. \tag{4.27}$$

A coarsening routine was also used in the context of adaptive wavelet methods for low-dimensional problems in [CDD01; CDD02], where the routine was directly applied to the coefficient vector. This is not suitable for our case, since calculating all coefficients would lead to the curse of dimensionality. Furthermore, due to our low-rank structure, it is not possible to remove individual coefficients because of the product structure of the representation.

4.1.3 Combination of Tensor Recompression and Coarsening

Finally, we investigate the combination of low-rank recompression and basis coarsening. To this end, we consider the following result from [BF24], which is an extension of the single hierarchical tensor case [BD15, Theorem 7] with slightly modified constants. It shows that the combination of these two operators yields quasi-optimal ranks and provides quasi-optimal mode frame sizes over all time indices.

Theorem 4.12 ([BF24, Theorem 3.13])**.** *Let $\mathbf{u}, \mathbf{v} \in \ell_2(\vee)$ with $\|\mathbf{u} - \mathbf{v}\| \leq \eta$. We further assume $\mathbf{u} \in \mathcal{A}(\gamma)$, where γ is an growth sequence and $\pi^{(\mathrm{t},i)}(\mathbf{u}) \in \mathcal{A}^s(\vee_\mathrm{t} \times \vee_1)$ for $i = 1, \dots, d$ with $\mathcal{A}^s$ defined in (2.6). As before, let $\kappa_\mathrm{P} = \sqrt{2d-3}$ and $\kappa_\mathrm{C} = \sqrt{d}$. Then for*

$$\mathbf{w}_\eta = \hat{\mathrm{C}}_{\kappa_\mathrm{C}(\kappa_\mathrm{P}+1)(1+\alpha)\eta}\left(\hat{\mathrm{P}}_{\kappa_\mathrm{P}(1+\alpha)\eta}(\mathbf{v})\right),$$

we have

$$\|\mathbf{u} - \mathbf{w}_\eta\| \leq \big(1 + \kappa_\mathrm{P}(1+\alpha) + \kappa_\mathrm{C}(\kappa_\mathrm{P}+1)(1+\alpha)\big)\eta \tag{4.28}$$

as well as

$$\begin{aligned} \operatorname{rank}_\infty(\mathbf{w}_\eta) &\leq \gamma^{-1}\big((\alpha\eta)^{-1}\|\mathbf{u}\|_{\mathcal{A}(\gamma)}\big), \\ \|\mathbf{w}_\eta\|_{\mathcal{A}(\gamma)} &\leq C_1\|\mathbf{u}\|_{\mathcal{A}(\gamma)}, \end{aligned} \tag{4.29}$$

where $C_1 = \alpha^{-1}(1 + \kappa_\mathrm{P}(1+\alpha)) + 1$ and

$$\begin{aligned} \sum_{i=1}^d \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)) &\leq d\eta^{-\frac{1}{s}}\alpha^{-\frac{1}{s}}\Big(\sum_{i=1}^d \big\|\pi^{(\mathrm{t},i)}(\mathbf{u})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}, \\ \sum_{i=1}^d \big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\big\|_{\mathcal{A}^s} &\leq C_2 \sum_{i=1}^d \big\|\pi^{(\mathrm{t},i)}(\mathbf{u})\big\|_{\mathcal{A}^s} \end{aligned} \tag{4.30}$$

with $C_2 = 2^s(1+2^s) + 2^{2s}\alpha^{-1}(1 + \kappa_\mathrm{P}(1+\alpha) + \kappa_\mathrm{C}(\kappa_\mathrm{P}+1)(1+\alpha))d^{\max\{1,s\}}$.

Proof. We set $\hat{\mathbf{w}} = \hat{\mathrm{P}}_{\kappa_\mathrm{P}(1+\alpha)\eta}(\mathbf{v})$. By the triangle inequality, we have

$$\begin{aligned} \|\mathbf{u} - \mathbf{w}_\eta\| &\leq \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v} - \hat{\mathrm{P}}_{\kappa_\mathrm{P}(1+\alpha)\eta}(\mathbf{v})\| + \|\hat{\mathbf{w}} - \hat{\mathrm{C}}_{\kappa_\mathrm{C}(\kappa_\mathrm{P}+1)(1+\alpha)\eta}(\hat{\mathbf{w}})\| \\ &\leq \eta + \kappa_\mathrm{P}(1+\alpha)\eta + \kappa_\mathrm{C}(\kappa_\mathrm{P}+1)(1+\alpha)\eta, \end{aligned}$$

where we used (4.8) from Theorem 4.3 as well as (4.27) in the second line, which yields (4.28).

The statements in (4.29) are a direct consequence of Theorem 4.6. The additional application of the coarsening operator can at most reduce the ranks and does not affect these estimates.

The main part is the proof of the estimates (4.30). Let $N \in \mathbb{N}$ be minimal such that $\|\mathbf{u} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\mathbf{u}\| \le \alpha\eta$. Then

$$\begin{aligned}\|\hat{\mathbf{w}} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\hat{\mathbf{w}}\| &\le \|(\mathbf{I} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)})(\mathbf{u} - \hat{\mathbf{w}})\| + \|\mathbf{u} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\mathbf{u}\| \\ &\le \|\mathbf{u} - \hat{\mathbf{w}}\| + \|\mathbf{u} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\mathbf{u}\| \\ &\le (1 + \kappa_{\mathrm{P}}(1+\alpha) + \alpha)\eta,\end{aligned}$$

where we used Theorem 4.3 to bound the first term on the right-hand side as well as $\|\mathbf{I} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\| \le 1$. Hence, by (4.24), we have

$$\begin{aligned}\mu_N(\hat{\mathbf{w}}) &\le \kappa_{\mathrm{C}}\|\hat{\mathbf{w}} - \mathrm{R}_{\bar{\Lambda}(\hat{\mathbf{w}},N)}\hat{\mathbf{w}}\| \\ &\le \kappa_{\mathrm{C}}\|\hat{\mathbf{w}} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N)}\hat{\mathbf{w}}\| \\ &\le \kappa_{\mathrm{C}}(1+\alpha)(\kappa_{\mathrm{P}}+1)\eta.\end{aligned} \tag{4.31}$$

Moreover, by definition, we have $\mathbf{w}_\eta = \hat{\mathrm{C}}_{\kappa_{\mathrm{C}}(\kappa_{\mathrm{P}}+1)(1+\alpha)\eta}(\hat{\mathbf{w}})$. Using the definition (4.26) of the coarsening operator combined with the minimality condition (4.25), we have

$$\sum_{i=1}^{d} \#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right) \le N.$$

Without loss of generality, we may assume that $N \ge d$, since the one-dimensional support must be at least one for each dimension, or the output will be zero. Keeping the optimality (4.22) condition of the best coarsening operator in mind, we have

$$\begin{aligned}\alpha\eta &< \|\mathbf{u} - \mathrm{R}_{\bar{\Lambda}(\mathbf{u},N-1)}\mathbf{u}\| \\ &\le \|\mathbf{u} - \mathrm{R}_{\Lambda(\mathbf{u},N-1)}\mathbf{u}\| \\ &\le \inf_{\sum_{i=1}^d \#\Lambda^{(t,i)} \le N-1} \left(\sum_{i=1}^{d} \left\|\pi^{(\mathrm{t},i)}(\mathbf{u}) - \mathrm{R}_{\Lambda^{(t,i)}}\pi^{(\mathrm{t},i)}(\mathbf{u})\right\|^2\right)^{\frac{1}{2}} \\ &\le \sum_{i=1}^{d} \inf_{\#\Lambda^{(t,i)} \le (N-1)/d} \left\|\pi^{(\mathrm{t},i)}(\mathbf{u}) - \mathrm{R}_{\Lambda^{(t,i)}}\pi^{(\mathrm{t},i)}(\mathbf{u})\right\| \\ &\le \left(1 + \left\lfloor \frac{N-1}{d} \right\rfloor\right)^{-s} \sum_{i=1}^{d} \left\|\pi^{(\mathrm{t},i)}(\mathbf{u})\right\|_{\mathcal{A}^s} \\ &\le \left(\frac{N}{d}\right)^{-s} \sum_{i=1}^{d} \left\|\pi^{(\mathrm{t},i)}(\mathbf{u})\right\|_{\mathcal{A}^s},\end{aligned}$$

where we used (4.18) in the fourth line and Theorem 2.8 (iv) in the fifth line. We note that by (4.31), the coarsening operator $\hat{\mathrm{C}}_{\kappa_{\mathrm{C}}(\kappa_{\mathrm{P}}+1)(1+\alpha)\eta}$ retains at most N terms. We combine this fact with the previous estimate, which results in

$$\sum_{i=1}^{d} \#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right) \le N \le d\alpha^{-\frac{1}{s}}\eta^{-\frac{1}{s}} \left(\sum_{i=1}^{d} \left\|\pi^{(\mathrm{t},i)}(\mathbf{u})\right\|_{\mathcal{A}^s}\right)^{\frac{1}{s}}. \tag{4.32}$$

Hence, the first statement in (4.30) is shown. Now let $\hat{N}_i = \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta))$ for $i = 1, \dots, d$, and let $\hat{N} = \hat{N}_1 + \cdots + \hat{N}_d \le N$. We may assume without loss of generality

$\hat{N}_i > 0$ for $i = 1, \dots, d$. Resolve (4.32) for η and inserting into (4.28), we observe

$$\|\mathbf{u} - \mathbf{w}_\eta\| \leq \hat{N}^{-s} C(\alpha) d^s \sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}, \tag{4.33}$$

where $C(\alpha) = \alpha^{-1}(1 + \kappa_{\mathrm{P}}(1+\alpha) + \kappa_{\mathrm{C}}(\kappa_{\mathrm{P}}+1)(1+\alpha))$.

Let $\mathbf{u}_{\hat{N}_i}^{(i)}$ be the best $\hat{N}_i$-term approximation to $\pi^{(\mathrm{t},i)}(\mathbf{u})$, then by using the properties of Theorem 2.8 we observe

$$\begin{aligned}
\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\|_{\mathcal{A}^s} &\leq 2^s \left(\|\mathbf{u}_{\hat{N}_i}^{(i)}\|_{\mathcal{A}^s} + \|\mathbf{u}_{\hat{N}_i}^{(i)} - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\|_{\mathcal{A}^s} \right) \\
&\leq 2^s \left(\|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s} + (2\hat{N}_i)^s \|\mathbf{u}_{\hat{N}_i}^{(i)} - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\| \right) \\
&\leq 2^s \big(\|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s} + (2\hat{N}_i)^s \big(\|\mathbf{u}_{\hat{N}_i}^{(i)} - \pi^{(\mathrm{t},i)}(\mathbf{u})\| + \|\pi^{(\mathrm{t},i)}(\mathbf{u}) - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\| \big) \big) \\
&\leq 2^s \left((1+2^s) \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s} + (2\hat{N}_i)^s \|\pi^{(\mathrm{t},i)}(\mathbf{u}) - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\| \right),
\end{aligned}$$

where we used that $\#\operatorname{supp}\left(\mathbf{u}_{\hat{N}_i}^{(i)} - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right) \leq 2\hat{N}_i$ in the second line. As a consequence of the Cauchy-Schwarz inequality, we have the componentwise estimate

$$|\pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_{\mathrm{x}}}(\mathbf{u}) - \pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_{\mathrm{x}}}(\mathbf{w}_\eta)| \leq \pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_{\mathrm{x}}}(\mathbf{u} - \mathbf{w}_\eta),$$

which yields

$$\|\pi^{(\mathrm{t},i)}(\mathbf{u}) - \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\| \leq \|\pi^{(\mathrm{t},i)}(\mathbf{u} - \mathbf{w}_\eta)\| = \|\mathbf{u} - \mathbf{w}_\eta\|.$$

Combining this fact with (4.33), we obtain

$$\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\|_{\mathcal{A}^s} \leq 2^s(1+2^s)\|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s} + 2^s C(\alpha) d^s \hat{N}^{-s} (2\hat{N}_i)^s \sum_{k=1}^{d} \|\pi^{(\mathrm{t},k)}(\mathbf{u})\|_{\mathcal{A}^s}.$$

Considering the case $s < 1$ and $s \geq 1$ separately, we observe

$$\hat{N}^{-s} \sum_{i=1}^{d} (\hat{N}_i)^s \leq d^{\max(0,1-s)},$$

where we use the Hölder inequality for the case $s < 1$. Summing over $i = 1, \dots, d$, we arrive at the second statement in (4.30). □

Remark 4.13. The recompression routine requires an HSVD of its inputs. Moreover, the coarsening routine also needs an HSVD to efficiently compute the contractions in the sense of Theorem 4.8 (iv). For each finitely supported $\mathbf{v} \in \ell_2(\vee)$, using the bound (2.20) for the costs of the HSVD for each spatial component $\mathbf{v}_{\nu_{\mathrm{t}}} \in \ell_2(\vee_{\mathrm{x}})$, $\nu_{\mathrm{t}} \in \vee_{\mathrm{t}}$, the total number of operations required for computing the decomposition for $\mathbf{v}$ is by bounded by

$$\begin{aligned}
\operatorname{flops}(\operatorname{hsvd}(\mathbf{v})) &\lesssim \sum_{\nu_{\mathrm{t}} \in \vee_{\mathrm{t}}} \left(d|\operatorname{rank}_{\nu_{\mathrm{t}}}(\mathbf{v})|_\infty^4 + |\operatorname{rank}_{\nu_{\mathrm{t}}}(\mathbf{v})|_\infty^2 \sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(i)}(\mathbf{v}_{\nu_{\mathrm{t}}})) \right) \\
&\leq d \operatorname{rank}_\infty^4(\mathbf{v}) \#\operatorname{supp}(\pi^{(\mathrm{t})}(\mathbf{v})) + \operatorname{rank}_\infty^2(\mathbf{v}) \sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{v})).
\end{aligned}$$

When we compute the HSVD for redundant hierarchical tensor representations with representation ranks greater than $\operatorname{rank}_{\nu_t}(\mathbf{v})$ for $\nu_t \in \vee_t$, the latter need to be replaced by the respective representation ranks in the above bound.

The HSVD calculation thus dominates the costs of the recompression and coarsening routines when we use a quasi-sorting algorithm as described in Theorem 4.11.

4.2 Low-Rank Preconditioning

As seen in Section 3.1.1, the solution $u \in \mathcal{X}$ of the space-time variational formulation is given by the solution $\mathbf{u}$ of the bi-infinite matrix equation $\bar{\mathbf{B}}\mathbf{u} = \bar{\mathbf{f}}$ from (3.14), where $\mathbf{u}$ are the coefficients of the basis representation of u. As mentioned in the beginning of the chapter, we will consider an iterative solver for this bi-infinite matrix equation, which yields approximations $(\mathbf{u}_k)_k$ with $\mathbf{u}_k \to \mathbf{u}$ as $k \to \infty$. One main goal is to avoid the curse of dimensionality, i.e., to circumvent the exponential growth of computational cost with the number of spatial dimensions d. As we will see in Section 4.4, the computational complexity of the iterative solver depends on the condition number of the bi-infinite matrix. In this regard, it is crucial that our bi-infinite matrix $\bar{\mathbf{B}}$ is bounded invertible with a condition number independent of d. As seen in Section 3.1.2, the diagonal scaling matrices $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ defined in (3.17) play an important role in this regard.

However, there is one issue with the scaling matrices. For the calculation of the iterates $\mathbf{u}_k$ of the iterative solver, we need to be able to apply appropriate approximations of the operator $\bar{\mathbf{B}}$. In the example of the heat equation, the operator $\mathbf{T}_x$ from (3.27) has a low-rank structure, due to its representation (2.24) and Theorem 2.20. Thus, $\mathbf{T}$ from (3.27) has a low-rank structure as well. However, the scaling matrices do not have a finite rank representation due to the non-separable structure of the diagonal entries of $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$. Hence the operator $\bar{\mathbf{B}}$ has unbounded ranks. Since the iterates $\mathbf{u}_k$ are based on approximate application of the operator $\bar{\mathbf{B}}$, we are not able to perform the operations in hierarchical tensor format according to Section 2.4.5 anymore since this would lead to a summation of infinitely many terms. To solve this issue, we replace the scaling operators by $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ that are equivalent to $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ in the sense that

$$\|\mathbf{D}_{\mathcal{X}}\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\|, \|\mathbf{D}_{\mathcal{X}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\| < \infty, \quad \|\mathbf{D}_{\mathcal{Y}}\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}\|, \|\mathbf{D}_{\mathcal{Y}}^{-1}\bar{\mathbf{D}}_{\mathcal{Y}}\| < \infty \tag{4.34}$$

and can be approximated by separable operators in an efficient, quantifiable way.

If we substitute the scaling matrices, we get the modified system

$$\mathbf{B}\mathbf{u} = \mathbf{f},$$

where

$$\begin{aligned} \mathbf{B} &= \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathcal{Y}}\mathbf{T}\mathbf{D}_{\mathcal{X}} \\ \mathbf{T}_0\mathbf{D}_{\mathcal{X}} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathcal{Y}} & 0 \\ 0 & \mathbf{I}_x \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{T}_0 \end{pmatrix} \mathbf{D}_{\mathcal{X}}, \\ \mathbf{f} &= \begin{pmatrix} \mathbf{D}_{\mathcal{Y}}\mathbf{g}_1 \\ \mathbf{g}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathcal{Y}} & 0 \\ 0 & \mathbf{I}_x \end{pmatrix} \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{pmatrix}, \end{aligned} \tag{4.35}$$

with $\mathbf{T}$ and $\mathbf{T}_0$ from (3.18) as well as $\mathbf{g}_1$ and $\mathbf{g}_2$ from (3.20). Based on the definitions of $\mathbf{B}$, $\mathbf{f}$ and $\bar{\mathbf{B}}$, $\bar{\mathbf{f}}$ in (3.16) and (3.19), we have that $\mathbf{u}$ fulfills $\mathbf{B}\mathbf{u} = \mathbf{f}$ if and only if

$$\bar{\mathbf{B}}(\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X}}\mathbf{u}) = \bar{\mathbf{f}}.$$

Thus, the solution of the modified system is directly related to the solution of the original system. Moreover, we have

$$\|\mathbf{B}\| \le \left\|\mathbf{D}_{\mathcal{X}}\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\right\|\left\|\mathbf{D}_{\mathcal{Y}}\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}\right\|\|\bar{\mathbf{B}}\|, \quad \left\|\mathbf{B}^{-1}\right\| \le \left\|\mathbf{D}_{\mathcal{X}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\right\|\left\|\mathbf{D}_{\mathcal{Y}}^{-1}\bar{\mathbf{D}}_{\mathcal{Y}}\right\|\left\|\bar{\mathbf{B}}^{-1}\right\|.$$

As mentioned before, for the computational complexity of the iterative solver it is crucial that the condition number is independent of the number of spatial dimensions d. Thus it is important that the operator norms in (4.34) are bounded independently of d. Furthermore, for practical reasons, the constants should not be too large.

In the remainder of this section, we give a construction of matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ which fulfil (4.34) and can be approximated well by low-rank matrices. For both matrices, we use exponential sum approximation. For the scaling matrix $\mathbf{D}_{\mathcal{X}}$, we follow [BF24, Section 4.1]. For $\mathbf{D}_{\mathcal{Y}}$, we follow the idea of [BD16] and combine it with the structural properties (3.24).

4.2.1 Scaling Matrix $\mathbf{D}_{\mathcal{X}}$

For the construction of the scaling matrix we separate again between the temporal and spatial parts. Let $\mathbf{E}_{\nu_t} = \mathbf{e}_{\nu_t}\mathbf{e}_{\nu_t}^{\top}$, which is the diagonal matrix of zeros except for a one at the diagonal entry ν_t. We aim to construct a scaling matrix of the form

$$\mathbf{D}_{\mathcal{X}} = \sum_{\nu_t \in \mathcal{V}_t} \mathbf{E}_{\nu_t} \otimes \mathbf{D}_{\mathcal{X},\nu_t}.$$

If we apply a scaling matrix of this form to a vector $\mathbf{v} \in \ell_2(\vee)$ in the format (4.2), we get

$$\mathbf{D}_{\mathcal{X}}\mathbf{v} = \sum_{\nu_t \in \mathcal{V}_t} \mathbf{e}_{\nu_t} \otimes (\mathbf{D}_{\mathcal{X},\nu_t}\mathbf{v}_{\nu_t}).$$

Since we are interested in the maximum rank over all spatial components, we can treat each scaling matrix $\mathbf{D}_{\mathcal{X},\nu_t}$ independently.

For the construction of $\mathbf{D}_{\mathcal{X},\nu_t}$, we use an exponential sum approximation as in [BF24]. We need relative error bounds for approximating these scaling matrices. To this end, we represent the entries of the scaling matrix $\mathbf{D}_{\mathcal{X},\nu_t}$ as integrals with separable integrands using the Laplace transform. Then we transform these integrals to the whole real line and subsequently approximate the integral by a quadrature rule with an infinite sum. As final step, we truncate the infinite sum to a finite sum, which yields an approximation of the scaling matrix by a separable operator.

For the Laplace transforms that we obtain in our case, the *Dawson function* plays an important role. This function is defined by

$$F_D(x) = \frac{\sqrt{\pi}}{2}e^{-x^2}\operatorname{erfi}(x) = e^{-x^2}\int_0^x e^{y^2}\,\mathrm{d}y, \tag{4.36}$$

where erfi is the imaginary error function. A visualization of the Dawson function in the interval $[0,5]$ is given in Figure 4.1. The Dawson function, based on the imaginary error function, is also defined for complex-valued arguments. Before we proceed with the construction of the scaling matrices, we state some properties of the Dawson function that will be helpful in the following.

Proposition 4.14. *The Dawson function F_D defined in (4.36) fulfills*

(i) $F_D'(x) = 1 - 2xF_D(x)$ *for all* $x \ge 0$,

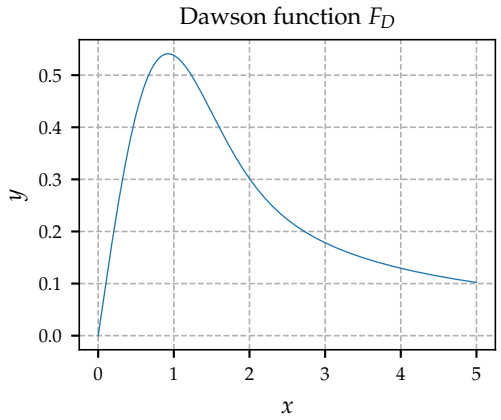


FIGURE 4.1: The Dawson function F_D defined in (4.36) in the interval $[0,5]$.

(ii) $0 \le F_D(x) \le x$ *for all* $x \ge 0$,

(iii) $2F_D(x) \ge x$ *for all* $0 \le x \le 1$,

(iv) $xF_D(x) \le 1$ *for all* $x \ge 0$,

(v) $-1 \le F_D'(x) \le 0$ *for all* $x \ge 1$,

(vi) $F_D'(x) \ge 0$ *for all* $0 \le x \le \sqrt{\ln(2)}$,

(vii) $F_D(x) \ge \frac{2x^2-1}{4x(x^2-1)}$ *for all* $x > \frac{9}{4}$.

(viii) *Let* $a, c > 0$ *and* $b \in \mathbb{R}$ *with* $|b| < a$. *Then*

$$\left|1 - 2c(a+bi)F_D\big(c(a+bi)\big)\right| \le c_{\mathrm{daw}}(a,b)$$

with

$$c_{\mathrm{daw}}(a,b) = 1 + \sqrt{\frac{\pi(a^2+b^2)}{2e(a^2-b^2)}} + \frac{2\sqrt{a^2+b^2}}{a} x_{\min} F_D(x_{\min}),$$

where $x_{\min}$ *is the global minimizer of* F_D' *(see [Was25]).*

The proof of this proposition is based directly on the definition of the Dawson function. For completeness, we give the proof in Chapter A.

For the construction of the scaling matrix, we start with an inverse Laplace transform of the entries for a fixed time index. By (3.26), the entries are of the form

$$\frac{\sqrt{s}}{s+a}, \quad \text{where} \quad s = \|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)}, a = \|\theta_{\nu_t}\|_{H^1(I)}. \tag{4.37}$$

In the next step, we consider the inverse Laplace transform of this function for fixed a, which is a key ingredient for obtaining a separable representation.

Lemma 4.15. *Let* $a > 0$. *Then the inverse Laplace transform of* $s \mapsto \frac{\sqrt{s}}{s+a}$ *is given by*

$$\frac{\sqrt{s}}{s+a} = \int_0^\infty \frac{1}{\sqrt{\pi y}}\Big(1 - 2\sqrt{ya}F_D\big(\sqrt{ya}\big)\Big)e^{-ys}\,\mathrm{d}y. \tag{4.38}$$

The proof is based on elementary calculations and can be found in Chapter A.

As seen in (4.37), we insert the squared H_0^1-norm of the d-dimensional spatial basis functions for s. Based on (3.23), the squared H_0^1-norm is the sum of the squared one-dimensional H_0^1-norms for product basis functions if the one-dimensional basis functions are L_2-orthonormal. Since this sum appears in the exponential term on the right-hand side of (4.38), it will decompose into a product of one-dimensional terms, resulting in a separable structure. Hence, by the inverse Laplace transform, we arrive at a separable representation of the integrand, but the expression is an integral over this separable structure. The main idea in the following steps is first to approximate the integral by an infinite sum using a quadrature rule and then to truncate it to a finite sum.

First, we consider approximating the integral by an infinite sum. To this end, we use the following definition and approximation error bound from [Ste93].

Definition 4.16. For $\zeta > 0$ and $0 < \varepsilon < 1$ let

$$\begin{aligned} D_\zeta &= \{z \in \mathbb{C} : |\operatorname{Im} z| < \zeta\}, \\ D_\zeta(\varepsilon) &= \{z \in \mathbb{C} : |\operatorname{Re} z| < \varepsilon^{-1}, |\operatorname{Im} z| < \zeta(1-\varepsilon)\}. \end{aligned}$$

Furthermore, for g analytic in D_ζ let

$$N_1(g, D_\zeta) = \lim_{\varepsilon \to 0} \int_{\partial D_\zeta(\varepsilon)} |g(z)|\,|\mathrm{d}z|.$$

Theorem 4.17 ([Ste93, Theorem 3.2.1])**.** *Let g be analytic in D_ζ with $N_1(g, D_\zeta) < \infty$ and $h > 0$, then*

$$\left| \int_{\mathbb{R}} g(x)\,\mathrm{d}x - h \sum_{k \in \mathbb{Z}} g(kh) \right| \leq \frac{e^{-\pi\zeta/h}}{2\sinh(\pi\zeta/h)} N_1(g, D_\zeta).$$

To be able to apply Theorem 4.17, we need to transform the integral in (4.38) to the whole real line. If we apply the substitution $ya = e^x$, we arrive at

$$\begin{aligned} \frac{\sqrt{s}}{s+a} &= \frac{1}{\sqrt{a\pi}} \int_{\mathbb{R}} \left(1 - 2e^{\frac{x}{2}} F_D\left(e^{\frac{x}{2}}\right)\right) e^{\frac{x}{2} - \frac{s}{a}e^x}\,\mathrm{d}x \\ &= \frac{2}{\sqrt{a\pi}} \int_{\mathbb{R}} \tfrac{\mathrm{d}}{\mathrm{d}x}(F_D(e^{x/2}))e^{-\frac{s}{a}e^x}\,\mathrm{d}x. \end{aligned}$$

Finally, we truncate the arising sum to arrive at a finite rank approximation. To do this, it is important to investigate the monotonicity of the integrand.

Lemma 4.18. *Let $c > 0$ and*

$$f_c(x) = \left(1 - 2e^{\frac{x}{2}} F_D\left(e^{\frac{x}{2}}\right)\right) e^{\frac{x}{2} - ce^x}. \tag{4.39}$$

Then $f_c(x)$ is monotonically increasing for $x > 2\ln(\frac{7}{4})$. Furthermore, there exists $z = z(c) < 0$ such that $f_c(x)$ is monotonically increasing in $(-\infty, z)$ and monotonically decreasing in $(z, \ln(\ln(2)))$.

Proof. We have

$$f_c'(x) = e^{\frac{x}{2} - ce^x}\left(F_D'(e^{\frac{x}{2}})\big(1 - e^x(c+1)\big) - \tfrac{1}{2}\right).$$

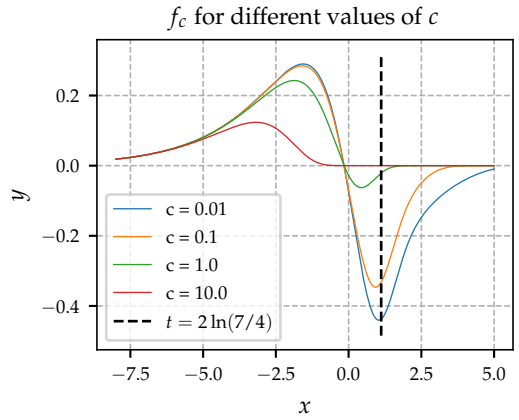


FIGURE 4.2: The function f_c defined in (4.39) for different values of c.

The first factor is positive for all $x \in \mathbb{R}$. Moreover, by Theorem 4.14 (vii), for all $c > 0$ and $x > 2\ln(\frac{7}{4})$ we obtain

$$\begin{aligned} F_D'(e^{\frac{x}{2}})\big(1 - e^x(c+1)\big) &= \Big(2e^{\frac{x}{2}} F_D\big(e^{\frac{x}{2}}\big) - 1\Big)(e^x(c+1) - 1) \\ &\geq \frac{1}{2e^x - 2}(e^x - 1) \\ &= \tfrac{1}{2}. \end{aligned}$$

Hence, $f_c(x)$ is monotonically increasing for $x > 2\ln(\frac{7}{4})$. Moreover, F_D' is monotonically decreasing in $[0,1]$ and $F_D'(0) = 1$. Combining this with $f_c'(\ln(\ln(2))) < 0$ yields the second statement. □

A visualization of the function f_c for different values of c, illustrating the monotonicity properties from Theorem 4.18, is given in Figure 4.2.

We combine these results in the following theorem, which is a slightly modified version of [BF24, Theorem 4.1].

Theorem 4.19. *For each $a > 0$, let*

$$\alpha_a(x) = \tfrac{1}{a}e^x, \qquad w_a(x) = \tfrac{1}{\sqrt{\pi a}}\big(1 - 2e^{x/2}F_D(e^{x/2})\big)e^{x/2} = \tfrac{1}{\sqrt{\pi a}}F_D'(e^{x/2})e^{x/2}.$$

For an arbitrary but fixed $\delta \in (0,1)$, let h_a be chosen such that

$$0 < h_a \leq \frac{2\pi^2}{3\ln\big(44\frac{1+a}{\delta}\big)} \tag{4.40}$$

and set

$$n_a^+ = n_a^+(\delta) = \left\lceil h_a^{-1} \max\left\{ \ln\Big(a\ln\big(\tfrac{8}{\sqrt{\pi}\delta}\big)\Big), 2\ln\big(\tfrac{7}{4}\big)\right\}\right\rceil. \tag{4.41}$$

Then with

$$\Phi_{a,n}(s) = \sum_{k=-n}^{n_a^+} h_a w_a(kh_a) e^{-\alpha_a(kh_a)s}, \quad \Phi_{a,\infty}(s) = \lim_{n\to\infty} \Phi_{a,n}(s),$$

one has

$$\left|\frac{\sqrt{s}}{s+a} - \Phi_{a,\infty}(s)\right| \le \frac{\delta}{2}\frac{\sqrt{s}}{s+a} \quad \text{for all } s \in [1,\infty). \tag{4.42}$$

Moreover, let $\eta > 0$, $K > 1$ *and*

$$n \ge \min\Big\{n \in \mathbb{N}_0 : n \ge h_a^{-1}|\ln(\ln(2))| \wedge \max\big\{f_a(1,h_a n), f_a(K,h_a n)\big\} \le \min\{\tfrac{\delta}{2},\eta\}\Big\} \tag{4.43}$$

with

$$f_a(s,y) = (1+h_a)\Big(2\sqrt{\tfrac{a}{s\pi}}e^{-\frac{s}{a}e^{-y}}F_D\big(e^{-\frac{y}{2}}\big) + \operatorname{erf}\Big(\sqrt{\tfrac{s}{a}}e^{-\frac{y}{2}}\Big)\Big).$$

Then one has additionally

$$\left|\frac{\sqrt{s}}{s+a} - \Phi_{a,n}(s)\right| \le \delta\frac{\sqrt{s}}{s+a} \quad \text{for all } s \in [1,K], \tag{4.44}$$

$$|\Phi_{a,\infty}(s) - \Phi_{a,n}(s)| \le \eta\frac{\sqrt{s}}{s+a} \quad \text{for all } s \in [1,K]. \tag{4.45}$$

Remark 4.20. In Figure 4.3, the exponential sum approximation $\Phi_{a,n}$ from Theorem 4.19 is visualized for $a = 1, 10$ and different values of K and δ. Additionally, the corresponding relative errors corresponding to (4.44) are shown. The results indicate that the theoretical bounds from Theorem 4.19 are not sharp. In particular, the actual errors are much smaller than the theoretical bounds. Hence, for the practical realization, the number of terms in the exponential sum can be chosen smaller than suggested by Theorem 4.19.

Proof of Theorem 4.19. We follow mainly the lines of [BF24]. The function

$$g(z) = \frac{1}{\sqrt{a\pi}}\left(1 - 2e^{\frac{z}{2}}F_D(e^{\frac{z}{2}})\right)e^{\frac{z}{2}-\frac{s}{a}e^{z}}$$

is holomorphic in the strip $D_{\pi/2} = \{x+iy : x \in \mathbb{R}, |y| < \frac{\pi}{2}\}$. Based on Theorem 4.14 (viii), we have

$$\left|1 - 2e^{\frac{z}{2}}F_D(e^{\frac{z}{2}})\right| \le c_{\mathrm{daw}}(\cos(\tfrac{y}{2}),\sin(\tfrac{y}{2})), \quad \text{for } z = x+yi \in D_{\pi/2}.$$

Additionally, we have

$$\left|e^{\frac{x+yi}{2}-\frac{s}{a}e^{x+yi}}\right| = e^{\mathrm{Re}(\frac{x+yi}{2}-\frac{s}{a}e^{x+yi})} = e^{\frac{x}{2}-\frac{s}{a}e^{x}\cos(y)},$$

which yields

$$\lim_{x\to\pm\infty} g(x+iy) = 0, \quad |y| < \tfrac{\pi}{2}.$$

Using this, we arrive at

$$N_1(g, D_\zeta) = \int_{\mathbb{R}} |g(x+i\zeta)|\,\mathrm{d}x + \int_{\mathbb{R}} |g(x-i\zeta)|\,\mathrm{d}x$$

for $|\zeta| < \pi/2$. In particular, for the special choice of $\zeta = \pi/3$, we have

$$|N_1(g, D_{\pi/3})| \le \frac{2c_{\mathrm{daw}}\big(\frac{\sqrt{3}}{2},\frac{1}{2}\big)}{\sqrt{a\pi}}\int_{\mathbb{R}} e^{\frac{x}{2}-\frac{s}{2a}e^{x}}\,\mathrm{d}x = \frac{2\sqrt{2}c_{\mathrm{daw}}\big(\frac{\sqrt{3}}{2},\frac{1}{2}\big)}{\sqrt{s}},$$

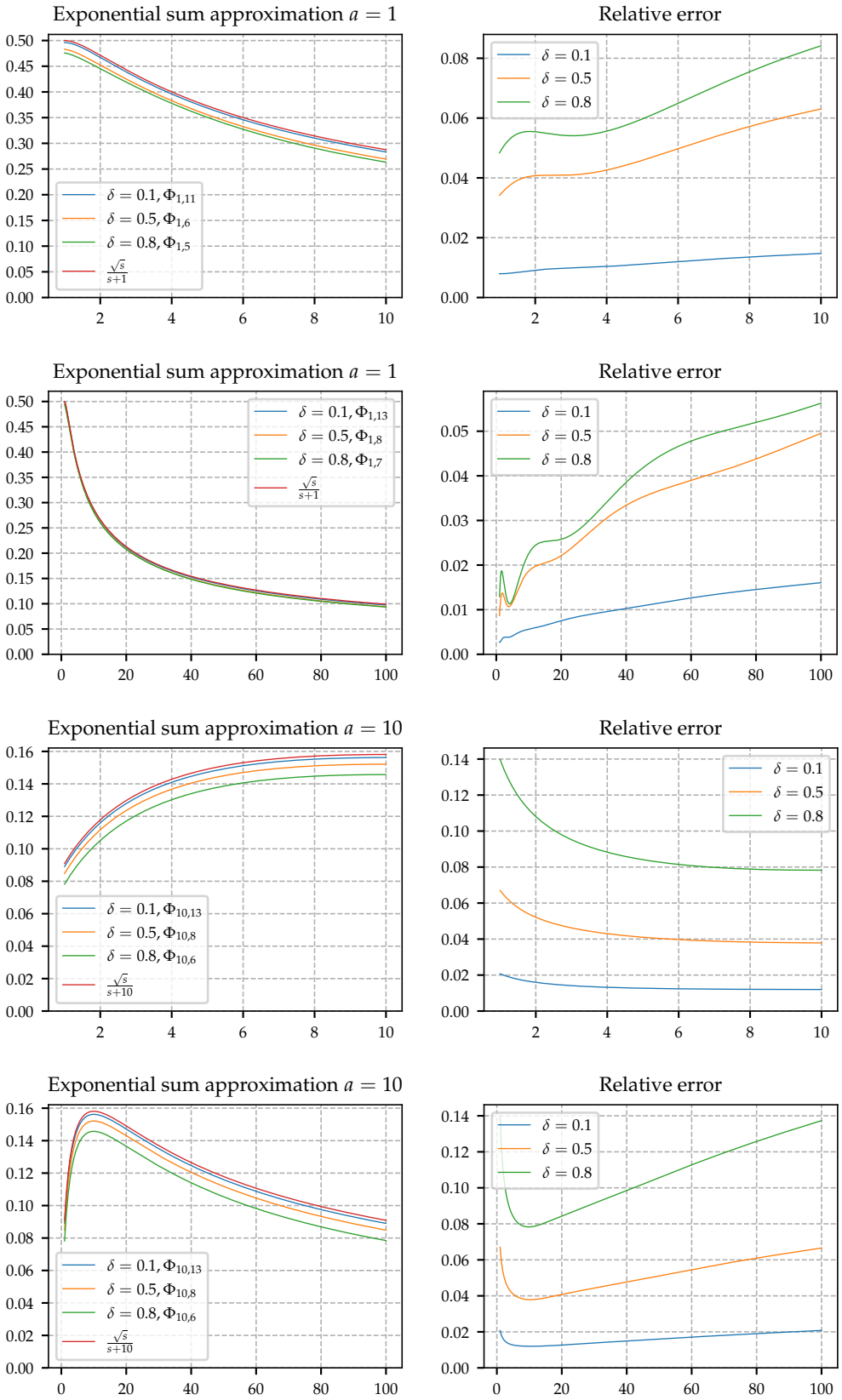


FIGURE 4.3: On the left side exponential sum approximation for different values of δ, a and K and the exact function are visualized. On the right side, the corresponding relative errors $|\frac{\sqrt{s}}{s+a} - \Phi_{a,n}(s)|(\frac{\sqrt{s}}{s+a})^{-1}$ are shown.

where we used Theorem 4.14 (viii) and that $\sqrt{\pi}\,\operatorname{erf}(\sqrt{c}\,e^{x/2})/\sqrt{c}$ is the antiderivative of $e^{\frac{x}{2}-ce^x}$ for $c>0$. We apply Theorem 4.17 next, which yields

$$\begin{aligned}\Big|\frac{\sqrt{s}}{s+a}-\sum_{k\in\mathbb{Z}}h_a w_a(kh_a)e^{-\alpha_a(kh_a)s}\Big| &\le \frac{e^{-\frac{\pi^2}{3h_a}}}{2\sinh(\frac{\pi^2}{3h_a})}2\sqrt{2}c_{\mathrm{daw}}\big(\tfrac{\sqrt{3}}{2},\tfrac{1}{2}\big)\frac{1}{\sqrt{s}}\\ &\le 11e^{-\frac{2\pi^2}{3h_a}}\frac{1}{\sqrt{s}}\\ &\le 11(1+a)e^{-\frac{2\pi^2}{3h_a}}\frac{\sqrt{s}}{s+a}\\ &\le \frac{\delta}{4}\frac{\sqrt{s}}{s+a},\end{aligned}\tag{4.46}$$

where we used $2\sqrt{2}c_{\mathrm{daw}}(\frac{\sqrt{3}}{2},\frac{1}{2})\le 11$ in the second line, $s\ge 1$ in the third line and the choice of h_a in (4.40) in the last line.

Let $c=s/a$. Based on the definition of w_a and Theorem 4.14 (v), we have $w_a(x)\le 0$ for $x\ge 0$. Combining this with the choice of n_a^+ and the monotonicity properties from Theorem 4.18, we arrive at

$$\begin{aligned}\Big|\sum_{k>n_a^+}h_a w_a(kh_a)e^{-\alpha_a(kh_a)s}\Big| &= -\sum_{k>n_a^+}h_a w_a(kh_a)\,e^{-\alpha_a(kh_a)s}\\ &\le -\int_{n_a^+}^{\infty}h_a w_a(xh_a)\,e^{-\alpha_a(xh_a)s}\,\mathrm{d}x\\ &= -\frac{1}{\sqrt{\pi a}}\int_{n_a^+h_a}^{\infty}\Big(1-2e^{\frac{x}{2}}F_D\big(e^{\frac{x}{2}}\big)\Big)e^{\frac{x}{2}-ce^x}\,\mathrm{d}x\\ &= -\frac{1}{\sqrt{\pi a}}\left[\frac{2}{c+1}e^{-ce^x}F_D(e^{x/2})+\sqrt{\pi}\frac{\sqrt{c}}{c+1}\operatorname{erf}\big(\sqrt{c}e^{x/2}\big)\right]_{n_a^+h_a}^{\infty}\\ &= \frac{1}{\sqrt{\pi}}\frac{\sqrt{s}}{s+a}\left(\frac{2}{\sqrt{c}}e^{-ce^{n_a^+h_a}}F_D\big(e^{\frac{n_a^+h_a}{2}}\big)-\sqrt{\pi}\operatorname{erfc}\Big(\sqrt{c}e^{\frac{n_a^+h_a}{2}}\Big)\right),\end{aligned}$$

where we used the definition $\operatorname{erfc}(z)=1-\operatorname{erf}(z)$ of the complementary error function as well as

$$\frac{1}{\sqrt{a}}\frac{\sqrt{c}}{c+1}=\frac{\sqrt{s}}{s+a}.$$

Furthermore, for the complementary error function, we have, by [AS13, Eq. 7.1.13], the estimate

$$\operatorname{erfc}(x)\ge\frac{2}{\sqrt{\pi}}e^{-x^2}\frac{1}{x+\sqrt{x^2+2}}\ge\frac{2}{\sqrt{\pi}}e^{-x^2}\frac{x}{2x^2+1},\qquad x>0.$$

By the choice of n_a^+ in (4.41), we have $e^{n_a^+ h_a} \geq \frac{1}{c}$. Combining all of this yields

$$\begin{aligned}
&\Big|\sum_{k>n_a^+} h_a w_a(kh_a) e^{-\alpha_a(kh_a)s}\Big| \\
&\quad\leq \frac{2}{\sqrt{\pi}} \frac{\sqrt{s}}{s+a} e^{-ce^{n_a^+ h_a}} \left(\frac{1}{\sqrt{c}} F_D\Big(e^{\frac{n_a^+ h_a}{2}}\Big) - \frac{\sqrt{c} e^{\frac{n_a^+ h_a}{2}}}{2ce^{n_a^+ h_a}+1} \right) \\
&\quad= \frac{\sqrt{s}}{s+a} \frac{2}{\sqrt{\pi}} e^{-ce^{n_a^+ h_a}} \frac{e^{\frac{n_a^+ h_a}{2}} F_D\Big(e^{\frac{n_a^+ h_a}{2}}\Big) - ce^{n_a^+ h_a}\Big(1 - 2e^{\frac{n_a^+ h_a}{2}} F_D\Big(e^{\frac{n_a^+ h_a}{2}}\Big)\Big)}{(2ce^{n_a^+ h_a}+1)\sqrt{c} e^{\frac{n_a^+ h_a}{2}}} \\
&\quad\leq \frac{\sqrt{s}}{s+a} \frac{2}{\sqrt{\pi}} e^{-ce^{n_a^+ h_a}} \\
&\quad< \frac{\delta}{4} \frac{\sqrt{s}}{s+a},
\end{aligned}$$

where we used $yF_D(y) \leq 1$ and $-F_D'(y) \leq 1$ for $y \geq 1$ from Theorem 4.14 in the third line as well as $s \geq 1$ and the choice of n_a^+ in the last line. Combining this result with (4.46) as well as the triangle inequality yields (4.42).

In the final step, we truncate the sum in the lower range. In contrast to before, we cannot use the monotonicity directly, since by Theorem 4.18 the function

$$u_a(x) = w_a(x) e^{-\alpha_a(x)s}$$

is not strictly monotone for $x < 0$. Let N be chosen according to Theorem 4.18 such that $-Nh_a < \ln(\ln(2))$ and that u_a is monotonically decreasing in $(-Nh_a, \ln(\ln(2)))$ and monotonically increasing in $(-\infty, -(N+1)h_a)$. Moreover, based on the choice of $n \geq h_a^{-1}|\ln(\ln(2))|$, we know by Theorem 4.14 (vi) that $u_a(kh_a) > 0$ for all $k < -n$. Hence,

$$\begin{aligned}
\Big|\sum_{k<-n} h_a u_a(kh_a)\Big| &= \sum_{k>n} h_a u_a(-kh_a) \\
&= \sum_{k=n+1}^{N-1} h_a u_a(-kh_a) + \sum_{k=N+2}^{\infty} h_a u_a(-kh_a) \\
&\quad + h_a(u_a(-Nh_a) + u_a(-(N+1)h_a)) \\
&\leq \int_{(n+1)h_a}^{Nh_a} u_a(-x)\,\mathrm{d}x + \int_{(N+1)h_a}^{\infty} u_a(-x)\,\mathrm{d}x \\
&\quad + h_a(u_a(-Nh_a) + u_a(-h_a(N+1))) \\
&\leq \int_{(n+1)h_a}^{\infty} u_a(-x)\,\mathrm{d}x + h_a \max\{u_a(-Nh_a), u_a(-(N+1)h_a)\},
\end{aligned}$$

where we used in the last line that by monotonicity

$$u_a(-x) \geq \min\{u_a(-Nh_a), u_a(-(N+1)h_a)\}, \quad x \in [-(N+1)h_a, -Nh_a].$$

Setting $c = s/a$ as before, we observe

$$\begin{aligned}\int_y^\infty u_a(-x)\,\mathrm{d}x &= -\frac{1}{\sqrt{a\pi}}\left[\frac{2}{c+1}e^{-ce^{-x}}F_D\big(e^{-\frac{x}{2}}\big)+\sqrt{\pi}\frac{\sqrt{c}}{c+1}\operatorname{erf}\big(\sqrt{c}e^{-\frac{x}{2}}\big)\right]_y^\infty\\ &= \frac{1}{\sqrt{a\pi}}\left(\frac{2}{c+1}e^{-ce^{-y}}F_D\big(e^{-\frac{y}{2}}\big)+\sqrt{\pi}\frac{\sqrt{c}}{c+1}\operatorname{erf}\big(\sqrt{c}e^{-\frac{y}{2}}\big)\right)\\ &\geq \frac{1}{\sqrt{a\pi}}\frac{2c+1}{c+1}e^{-\frac{y}{2}}e^{-ce^{-y}}\\ &\geq u_a(-y),\end{aligned}$$

where we use that $2F_D(x) \geq x$ for $x \in [0,1]$ by Theorem 4.14 and $\sqrt{\pi}\operatorname{erf}(z) \geq 2ze^{-z^2}$ in the third line as well as $F_D'(x) \leq 1$ in the last line. Combining this with the previous estimate yields

$$\begin{aligned}\Big|\sum_{k<-n} h_a u_a(kh_a)\Big| &\leq (1+h_a)\int_{(n+1)h_a}^\infty u_a(-x)\,\mathrm{d}x\\ &= \frac{(1+h_a)}{\sqrt{a\pi}}\left(\frac{2}{c+1}e^{-ce^{-nh_a}}F_D\big(e^{-\frac{nh_a}{2}}\big)+\sqrt{\pi}\frac{\sqrt{c}}{c+1}\operatorname{erf}\big(\sqrt{c}e^{-\frac{nh_a}{2}}\big)\right)\\ &= (1+h_a)\frac{\sqrt{s}}{s+a}\left(\operatorname{erf}\left(\frac{\sqrt{s}}{\sqrt{a}}e^{-\frac{nh_a}{2}}\right)+\frac{2\sqrt{a}}{\sqrt{s}\sqrt{\pi}}e^{-\frac{s}{a}e^{-nh_a}}F_D\left(e^{-\frac{nh_a}{2}}\right)\right)\\ &= \frac{\sqrt{s}}{s+a}f_a(s,nh_a).\end{aligned} \tag{4.47}$$

Based on the choice of n, we know by Theorem 4.14 (i) and (vi) that

$$F\big(e^{-\frac{nh_a}{2}}\big)e^{-\frac{nh_a}{2}} < \frac{1}{2}.$$

Moreover, we have

$$\frac{\mathrm{d}f_a}{\mathrm{d}s}(s,nh_a) = (1+h_a)e^{-\frac{s}{a}e^{-nh_a}}e^{-\frac{nh_a}{2}}\frac{1}{\sqrt{sa\pi}}\left(1-2e^{-\frac{nh_a}{2}}F_D\big(e^{-\frac{nh_a}{2}}\big)-\frac{a}{s}e^{\frac{nh_a}{2}}F_D\big(e^{-\frac{nh_a}{2}}\big)\right).$$

Hence, f_a is monotonically decreasing with respect to the first variable in some interval $s \in (0,z]$ and monotonically increasing in $s \in (z,\infty)$. Based on this fact, the maximum value of the function $f_a(s,nh_a)$ for $s \in [1,K]$ has to be located at one of the boundary values. Thus, by the choice of n in (4.43) and the estimate in (4.47), we arrive at

$$|\Phi_{a,\infty}(s) - \Phi_{a,n}(s)| \leq \min\{\eta, \tfrac{\delta}{2}\}\frac{\sqrt{s}}{s+a},$$

which shows (4.45). Taking into account (4.42) and the triangle inequality we obtain (4.44). □

Remark 4.21. We note here that the bound of h_a in (4.40) to achieve a relative error δ might not be sharp. The bound can in general be improved, e.g., by more optimized choices of ζ in the proof of Theorem 4.19. The same holds true for the choice of n_a^+.

Based on Theorem 4.19, we can define the corresponding scaling matrices. Let

$$s_{\min} = \big(\min_{v\in V_x}\|\Psi_v\|_V\big)^{-1} \tag{4.48}$$

and

$$a_{\nu_t} = s_{\min}^2 \|\theta_{\nu_t}\|_{H^1(I)}, \quad \nu_t \in \vee_t. \tag{4.49}$$

Furthermore, we fix a value $\delta \in (0,1)$ and choose $h_{a_{\nu_t}}$ and $n^+_{a_{\nu_t}}$ for each $\nu_t \in \vee_t$ as in (4.40) and (4.41). Especially for $h_{a_{\nu_t}}$, we choose the upper bound. For shorthand notation, we write $h_{\nu_t} = h_{a_{\nu_t}}$. We define

$$s^{\mathcal{X}}_{n,\nu_t,\nu_x} = s_{\min} \Phi_{a_{\nu_t},n}\big(s_{\min}^2 \|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)}\big)$$

for $n \in \mathbb{N}_0$. For $\mathsf{n} = (n_{\nu_t})_{\nu_t \in \vee_t} \subset \mathbb{N}_0$ and $\mathbf{v} \in \ell_2(\vee)$ in the form (4.2) with spatial components $(\mathbf{v}_{\nu_t})_{\nu_t \in \vee_t}$ we define

$$\mathbf{D}_{\mathcal{X},\mathsf{n}} \mathbf{v} = \sum_{\nu_t \in \vee_t} \mathbf{e}_{\nu_t} \otimes \big(\mathbf{D}_{\mathcal{X},\nu_t,n_{\nu_t}} \mathbf{v}_{\nu_t}\big)$$

with

$$\mathbf{D}_{\mathcal{X},\nu_t,n_{\nu_t}} \mathbf{v}_{\nu_t} = \Big(s^{\mathcal{X}}_{n_{\nu_t},\nu_t,\nu_x} v_{\nu_t,\nu_x}\Big)_{\nu_x \in \vee_x}.$$

As a direct consequence the bi-infinite scaling matrix $\mathbf{D}_{\mathcal{X},\nu_t,n_{\nu_t}}$ can be represented as a sum of $1 + n^+_{a_{\nu_t}}(\delta) + n_{\nu_t}$ separable terms, since by (3.23) we have

$$e^{-y\|\Psi_\nu\|^2_{H^1_0(\Omega)}} = e^{-y\sum_{i=1}^d \|\psi_{\nu_i}\|^2_{H^1_0(0,1)}} = \prod_{i=1}^d e^{-y\|\psi_{\nu_i}\|^2_{H^1_0(0,1)}}, \quad y \in \mathbb{R}. \tag{4.50}$$

The reference scaling matrix $\mathbf{D}_{\mathcal{X}}$ is defined in the same fashion by

$$\mathbf{D}_{\mathcal{X}} \mathbf{v} = \sum_{\nu_t \in \vee_t} \mathbf{e}_{\nu_t} \otimes \big(\mathbf{D}_{\mathcal{X},\nu_t} \mathbf{v}_{\nu_t}\big)$$

with

$$\mathbf{D}_{\mathcal{X},\nu_t} \mathbf{v}_{\nu_t} = \Big(s^{\mathcal{X}}_{\nu_t,\nu_x} v_{\nu_t,\nu_x}\Big)_{\nu_x \in \vee_x},$$

where

$$s^{\mathcal{X}}_{\nu_t,\nu_x} = s_{\min} \Phi_{a_{\nu_t},\infty}\big(s_{\min}^2 \|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)}\big).$$

Based on the definition of $\Phi_{a,n}$ and $\Phi_{a,\infty}$ we have

$$s^{\mathcal{X}}_{\nu_t,\nu_x} = \lim_{n\to\infty} s^{\mathcal{X}}_{n,\nu_t,\nu_x}.$$

In the next step, we state the approximation properties (4.44) and (4.45) in terms of the scaling matrices. For this purpose, we define the index set

$$\Lambda_K = \big\{\nu_x \in \vee_x : s_{\min}^2 \|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} \le K\big\} \tag{4.51}$$

and

$$M_{\mathcal{X},\nu_t}(\eta; K) = \min\Big\{n \in \mathbb{N} : n \ge \tfrac{|\ln(\ln(2))|}{h_{\nu_t}} \wedge \max\big\{f_{a_{\nu_t}}(1, h_{\nu_t} n), f_{a_{\nu_t}}(K, h_{\nu_t} n)\big\} \le \min\{\tfrac{\delta}{2}, \eta\}\Big\}$$

as well as

$$M_{\mathcal{X},0,\nu_t}(K) = M_{\mathcal{X},\nu_t}(\tfrac{\delta}{2}, K),$$

where a_{ν_t} is given by (4.49). Then for all $\nu_t \in \vee_t$, $\eta > 0$, $K > 1$ and $\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}$ defined in (3.26), we obtain for $n \geq M_{\nu_t}(\eta; K)$ and $\nu_x \in \Lambda_K$ the estimate

$$\begin{aligned}\left|\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} - s^{\mathcal{X}}_{n,\nu_t,\nu_x}\right| &= s_{\min}\left|\frac{s_{\min}\|\Psi_{\nu_x}\|_{H^1_0(\Omega)}}{s^2_{\min}\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + s^2_{\min}\|\theta_{\nu_t}\|^2_{H^1(I)}} - \Phi_{a_{\nu_t},n}\left(s^2_{\min}\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)}\right)\right| \\ &\leq s_{\min}\delta\frac{s_{\min}\|\Psi_{\nu_x}\|_{H^1_0(\Omega)}}{s^2_{\min}\|\Psi_{\nu_x}\|^2_{H^1_0(\Omega)} + s^2_{\min}\|\theta_{\nu_t}\|^2_{H^1(I)}} \\ &= \delta\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x},\end{aligned}$$

where we used (4.44) in the second line. This statement is equivalent to

$$\left|\left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\right)^{-1}\left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x} - s^{\mathcal{X}}_{n,\nu_t,\nu_x}\right)\right| \leq \delta.$$

In the same manner, by using (4.45) we get

$$\left|\left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\right)^{-1}\left(s^{\mathcal{X}}_{\nu_t,\nu_x} - s^{\mathcal{X}}_{n,\nu_t,\nu_x}\right)\right| \leq \eta.$$

In the iterative solver, we apply the scaling matrices to vectors with finite support. The spatial components of the vector will in general have different supports. Hence the number of terms does not only depend on the temporal index, but also on the corresponding support of the spatial component. We define

$$\Lambda_{\mathsf{K}} = \bigcup_{\nu_t \in \vee_t} \{\nu_t\} \times \Lambda_{K_{\nu_t}} \tag{4.52}$$

for $\mathsf{K} \in \mathbb{R}^{\vee_t}$, where $\Lambda_{K_{\nu_t}}$ is defined in (4.51), and set

$$\begin{aligned}\mathsf{M}_{\mathcal{X}}(\eta;\mathsf{K}) &= \left(M_{\mathcal{X},\nu_t}(\eta; K_{\nu_t})\right)_{\nu_t \in \vee_t}, \\ \mathsf{M}_{\mathcal{X},0}(\mathsf{K}) &= \left(M_{\mathcal{X},0,\nu_t}(K_{\nu_t})\right)_{\nu_t \in \vee_t}.\end{aligned} \tag{4.53}$$

Then for $\eta > 0$ and $\mathsf{K} \in \mathbb{R}^{\vee_t}$ with $\mathsf{K}_{\nu_t} > 1$ for all $\nu_t \in \vee_t$, we have

$$\begin{aligned}\|\bar{\mathbf{D}}^{-1}_{\mathcal{X}}(\bar{\mathbf{D}}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}}}\| &\leq \delta, \\ \|\bar{\mathbf{D}}^{-1}_{\mathcal{X}}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}}}\| &\leq \eta\end{aligned} \tag{4.54}$$

for $\mathsf{n} \geq \mathsf{M}_{\mathcal{X}}(\eta;\mathsf{K})$ with the restriction operator $\mathrm{R}_{\Lambda_{\mathsf{K}}}$ defined in (2.5).

As a consequence of the definition of $\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}$ from (3.26) and (4.42), we have

$$1 - \tfrac{\delta}{2} \leq \left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\right)^{-1} s^{\mathcal{X}}_{\nu_t,\nu_x} \leq 1 + \tfrac{\delta}{2}, \quad (\nu_t,\nu_x) \in \vee.$$

Furthermore, by the choice of $\mathsf{n} \geq \mathsf{M}_{\mathcal{X}}(\eta;\mathsf{K})$ and Theorem 4.14 (vi), for each $\nu_t \in \vee_t$ we have $w_{a_{\nu_t}}(kh_{\nu_t}) > 0$ for all $k \leq n_{\nu_t}$ with $w_{a_{\nu_t}}$ defined in Theorem 4.19. Hence

$$s^{\mathcal{X}}_{n_{\nu_t},\nu_t,\nu_x} \leq s^{\mathcal{X}}_{\nu_t,\nu_x} \tag{4.55}$$

for each $\nu_t \in \vee_t$ and thus

$$\left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\right)^{-1} s^{\mathcal{X}}_{n_{\nu_t},\nu_t,\nu_x} \leq 1 + \tfrac{\delta}{2}.$$

We can restate (4.55) in terms of the scaling matrices, which yields

$$\|\mathbf{D}_{\mathcal{X},\mathsf{n}}\mathbf{D}^{-1}_{\mathcal{X}}\mathrm{R}_{\Lambda_{\mathsf{K}}}\| \leq 1 \tag{4.56}$$

for all $\mathsf{n} \geq M_{\mathcal{X},\nu_t}(\mathsf{K})$. Under the additional restriction $n_{\nu_t} \geq M_{\mathcal{X},0,\nu_t}(K_{\nu_t})$ and $\nu_x \in \Lambda_{K_{\nu_t}}$, we have by (4.44) the estimate

$$\left(\bar{s}^{\mathcal{X}}_{\nu_t,\nu_x}\right)^{-1} s^{\mathcal{X}}_{n_{\nu_t},\nu_t,\nu_x} \geq 1-\delta.$$

Again, we restate this in terms of the scaling matrices.

Remark 4.22. The first statement in (4.34) is fulfilled with

$$\begin{aligned} \|\mathbf{D}_{\mathcal{X}}\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\| &\leq 1+\tfrac{\delta}{2}, \\ \|\bar{\mathbf{D}}_{\mathcal{X}}\mathbf{D}_{\mathcal{X}}^{-1}\| &\leq (1-\tfrac{\delta}{2})^{-1}. \end{aligned} \tag{4.57}$$

The bound on the operator norms especially depend only on the choice of δ and are independent of the spatial dimension d. Furthermore, for $\mathsf{n} = (n_{\nu_t})_{\nu_t \in \vee_t}$ with $n_{\nu_t} \geq h_{\nu_t}^{-1}|\ln(\ln(2))|$ for all $\nu_t \in \vee_t$, we have

$$\|\mathbf{D}_{\mathcal{X},\mathsf{n}}\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\| \leq 1+\tfrac{\delta}{2}.$$

Let $\mathsf{K} \in \mathbb{R}^{\vee_t}$ with $\mathsf{K}_{\nu_t} > 1$ for all $\nu_t \in \vee_t$, then we have

$$\|\mathbf{D}_{\mathcal{X},\mathsf{n}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\mathrm{R}_{\Lambda_{\mathsf{K}}}\| \leq (1-\delta)^{-1}$$

for $\mathsf{n} \geq \mathsf{M}_{\mathcal{X},0}(\mathsf{K})$, where the inequality is to understood element-wise.

For the analysis of our method, we need an upper bound for $M_{\mathcal{X},\nu_t}(K;\eta)$ in terms of $K > 1$, $\eta > 0$, and $\nu_t \in \vee_t$.

Lemma 4.23 ([BF24, Lemma 4.3])**.** *Let $0 < \eta < 1, K \in \mathbb{R}$ and $\nu_t \in \vee_t$. Then with $h_{\nu_t} = h_{a_{\nu_t}}$ defined in (4.40), we have*

$$\begin{aligned} &M_{\mathcal{X},\nu_t}(\eta;K) \\ &\quad \leq 1 + 2h_{\nu_t}^{-1}\Big(C + |\ln(\min\{\tfrac{\delta}{2},\eta\})| + \tfrac{1}{2}\ln(a_{\nu_t}) + \max\{0, \tfrac{1}{2}\ln(K) - \ln(a_{\nu_t})\}\Big), \end{aligned} \tag{4.58}$$

where $C = \ln(\frac{12}{\sqrt{\pi}})$.

Proof. We define

$$g_{a_{\nu_t}}(y) = \tfrac{12}{\sqrt{\pi}} e^{-\frac{y}{2}} \sqrt{a_{\nu_t}} \max\left\{1, \tfrac{\sqrt{K}}{a_{\nu_t}}\right\}.$$

Using Theorem 4.14 (ii) and $\operatorname{erf}(x) \leq \frac{2}{\sqrt{\pi}}x$ for all $x > 0$ as well as $h_{\nu_t} \leq 2$ and $c + c^{-1} \leq 2\max\{c, c^{-1}\}$ yields

$$f_{a_{\nu_t}}(s,y) \leq g_{a_{\nu_t}}(y) \quad \text{for all } y \in \mathbb{R}, s \in [1,K].$$

Let $n \in \mathbb{N}$ be minimal such that $g_{a_{\nu_t}}(nh_{\nu_t}) \leq \eta$. Then we obtain

$$n - 1 \leq 2h_{\nu_t}^{-1}\left(C + |\ln(\min\{\tfrac{\delta}{2},\eta\})| + \tfrac{1}{2}\ln(a_{\nu_t}) + \max\{0, \tfrac{1}{2}\ln(K) - \ln(a_{\nu_t})\}\right)$$

with $C = \ln(\frac{12}{\sqrt{\pi}})$. Due to $f_{a_{\nu_t}}(s, nh_{\nu_t}) \leq g_{a_{\nu_t}}(nh_{\nu_t}) \leq \eta$ for all $s \in [1,K]$, we have $M_{\mathcal{X},\nu_t}(\eta;K) \leq n$. From this, we conclude (4.58). □

Hence, the upper bound depends on the time index via the corresponding norm of the basis function, which enters in a_{ν_t}. However, it also depends on the spatial support of the corresponding spatial component to which we apply the scaling matrix.

Remark 4.24. The step size h_{ν_t} is not fixed for each time index. Moreover, it depends on the norm of the corresponding temporal wavelet basis function, which enters in a_{ν_t}. If we set h_{ν_t} to the upper bound in (4.40), we get $h_{\nu_t} \eqsim (1+|\nu_t|)^{-1}$, where $|\nu_t|$ is the corresponding level of the wavelet basis function θ_{ν_t}. Hence, $n^+_{a_{\nu_t}}$ grows by (4.41) quadratically in $|\nu_t|$. Furthermore, for the lower bound of n, we have $n \gtrsim h^{-1}_{\nu_t} \eqsim (1+|\nu_t|)$, i.e., n is at least of the magnitude of the level of the temporal wavelet basis function.

Remark 4.25. In this approach, the scaling matrix $\mathbf{D}_{\mathcal{X}} = \sum_{\nu_t \in \mathcal{V}_t} \mathbf{E}_{\nu_t} \otimes \mathbf{D}_{\mathcal{X},\nu_t}$ is approximated by a sum of separable terms independently for each time index ν_t. Based on our separable representation (4.2), each temporal component $\mathbf{D}_{\mathcal{X},\nu_t}$ of the scaling matrix acts on exactly one low-rank approximation. This fact is heavily used by fixing the norm of the temporal basis function that enters in a_{ν_t}. If, instead of using the separable approach, we treat time as one of the tensor components, we would need to find a direct low-rank approximation to the scaling matrix $\mathbf{D}_{\mathcal{X}}$, which seems substantially more difficult. We are not aware of any error-controlled constructions as in our approach. However, there are heuristic approaches to construct low-rank approximations of a rescaling equivalent to $\mathbf{D}_{\mathcal{X}}$ in the discretized case in [AT15].

4.2.2 Scaling Matrix $\mathbf{D}_{\mathcal{Y}}$

The scaling operator $\bar{\mathbf{D}}_{\mathcal{Y}}$ has the form

$$\bar{\mathbf{D}}_{\mathcal{Y}} = \mathbf{I}_t \otimes \bar{\mathbf{D}} = \sum_{\nu_t \in \mathcal{V}_t} \mathbf{E}_{\nu_t} \otimes \bar{\mathbf{D}}$$

with $\bar{\mathbf{D}}$ defined in (3.25) and $\mathbf{E}_{\nu_t} = \mathbf{e}_{\nu_t}\mathbf{e}_{\nu_t}^\top$. An equivalent scaling matrix $\mathbf{D}$ for $\bar{\mathbf{D}}$ which fulfills

$$\|\mathbf{D}\bar{\mathbf{D}}^{-1}\|, \|\mathbf{D}^{-1}\bar{\mathbf{D}}\| < \infty,$$

with bounds independent of the number of spatial dimensions d and which can be approximated by separable operators in an efficient quantifiable way, was already constructed in [BD16, Section 4.1]. In this section, we give the main results and refer to [BD16] for the proofs. Additionally, we present the corresponding adaptation of the application of $\mathbf{D}_{\mathcal{Y}}$, as in [BF24, Section 4.2]. As for the scaling matrix $\mathbf{D}_{\mathcal{X}}$, we can exploit the fact that the spatial support may differ across the spatial components. Hence, we can approximate $\mathbf{D}$ differently for each component.

The procedure is the same as for $\mathbf{D}_{\mathcal{X}}$ in the previous section. We start by the inverse Laplace transform of the function $t \mapsto t^{-1/2}$ which gives us the representation (c.f. [AS13, Eq.29.3.4])

$$\frac{1}{\sqrt{s}} = \int_0^\infty \frac{1}{\sqrt{\pi t}} e^{-st}\,\mathrm{d}t.$$

Applying a change of variables, we obtain

$$\frac{1}{\sqrt{s}} = \frac{2}{\sqrt{\pi}} \int_{\mathbb{R}} \frac{e^{-s\ln^2(1+e^x)}}{1+e^{-x}}\,\mathrm{d}x.$$

As before we are interested in exponential sum approximation, which yields relative error bounds. For exponential sum approximations of the function $t \mapsto t^{-1/2}$ with absolute error bounds, we refer to [Hac19, Section 9.7.2.3]. Proceeding as before, using the trapezoidal rule from Theorem 4.17 and truncating the sum, we arrive at the following statement.

Theorem 4.26 ([BD16, Theorem 9]). *For an arbitrary but fixed $\delta \in (0,1)$, let h be chosen such that*

$$0 \le h \le \frac{\pi^2}{5|\ln(\delta/4)| + 4}$$

and set

$$n^+ = n^+(\delta) = \left\lceil h^{-1} \max\left\{ \tfrac{4}{\sqrt{\pi}} \sqrt{\left|\ln\left(\tfrac{\delta}{4}\right)\right|} \right\} \right\rceil.$$

Moreover, let $\alpha(x) = \ln^2(1+e^x)$ and $w(x) = 2\pi^{-1/2}(1+e^{-x})^{-1}$. Then defining

$$\Phi_n(s) = \sum_{k=-n}^{n^+} hw(kh) e^{-\alpha(kh)s}, \quad \Phi_\infty(s) = \lim_{n\to\infty} \Phi_n(s),$$

one has

$$\left| \frac{1}{\sqrt{s}} - \Phi_n(s) \right| \le \frac{\delta}{2} \frac{1}{\sqrt{s}}, \quad s \in [1,\infty).$$

Furthermore, let $\eta > 0$, $K > 1$ and

$$n \ge \left\lceil h^{-1} \Big(\ln\left(\tfrac{2}{\sqrt{\pi}}\right) + \left|\ln\left(\min\left\{ \tfrac{\delta}{2}, \eta \right\}\right)\right| + \tfrac{1}{2}\ln(K) \Big) \right\rceil.$$

Then one has additionally

$$\left| \frac{1}{\sqrt{s}} - \Phi_n(s) \right| \le \delta \frac{1}{\sqrt{s}} \quad \text{for all } s \in [1,K],$$
$$|\Phi_\infty(s) - \Phi_n(s)| \le \eta \frac{1}{\sqrt{s}} \quad \text{for all } s \in [1,K].$$

The proof of this theorem is similar to the proof of Theorem 4.19. For details, we refer to [BD16].

We define the scaling matrices $\mathbf{D}$ and $\mathbf{D}_n$ by

$$\mathbf{D}_n \mathbf{v} = \left(s_{n,\nu}^V \mathbf{v}_\nu \right)_{\nu \in \vee_{\mathrm{x}}}, \quad \text{where} \quad s_{n,\nu}^V = s_{\min} \Phi_n\left(s_{\min}^2 \|\Psi_\nu\|_V^2 \right),$$
$$\mathbf{D} \mathbf{v} = \left(s_\nu^V \mathbf{v}_\nu \right)_{\nu \in \vee_{\mathrm{x}}}, \quad \text{where} \quad s_\nu^V = s_{\min} \Phi_\infty\left(s_{\min}^2 \|\Psi_\nu\|_V^2 \right)$$

for $\mathbf{v} \in \ell_2(\vee_{\mathrm{x}})$ and $s_{\min}$ defined in (4.48). As before using (4.50), the bi-infinite scaling matrix $\mathbf{D}_n$ is a sum of $1 + n^+(\delta) + n$ separable terms.

In the remainder of this section, we state the relevant results about the scaling matrices. For $\eta > 0$, $h \in \mathbb{R}$, $0 < \delta < 1$ and $K \in \mathbb{R}$, let

$$\begin{aligned} M_{\mathcal{Y}}(\eta; K) &= \left\lceil h^{-1} \Big(\ln\left(\tfrac{2}{\sqrt{\pi}}\right) + \left|\ln\left(\min\left\{ \tfrac{\delta}{2}, \eta \right\}\right)\right| + \tfrac{1}{2}\ln(K) \Big) \right\rceil, \\ M_{\mathcal{Y},0}(K) &= M_{\mathcal{Y}}\left(\tfrac{\delta}{2}, K\right). \end{aligned} \tag{4.59}$$

Let $n \ge M_{\mathcal{Y}}(\eta; K)$ and Λ_K as in (4.51). Then we have

$$\begin{aligned} \|\bar{\mathbf{D}}^{-1}(\bar{\mathbf{D}} - \mathbf{D}_n) \mathrm{R}_{\Lambda_K}\| &\le \delta, \\ \|\bar{\mathbf{D}}^{-1}(\mathbf{D} - \mathbf{D}_n) \mathrm{R}_{\Lambda_K}\| &\le \eta. \end{aligned} \tag{4.60}$$

Remark 4.27 ([BD16, Remark 11]). For the diagonal operators $\bar{\mathbf{D}}, \mathbf{D}$ and $\mathbf{D}_n$ with $n \in \mathbb{N}_0$, we have

$$\begin{aligned} \|\bar{\mathbf{D}}^{-1}\mathbf{D}_n\|, \|\bar{\mathbf{D}}^{-1}\mathbf{D}\| &\le 1 + \tfrac{\delta}{2}, \\ \|\mathbf{D}^{-1}\bar{\mathbf{D}}\| &\le (1 - \tfrac{\delta}{2})^{-1}. \end{aligned} \tag{4.61}$$

Moreover, for any $K > 1$ and $n \geq M_{\mathcal{Y},0}(K)$, we have

$$\|\bar{\mathbf{D}}\mathbf{D}_n^{-1}\mathrm{R}_{\Lambda_K}\| \leq (1-\delta)^{-1}.$$

Due to the representation (3.24) of $\bar{\mathbf{D}}_{\mathcal{Y}}$, it is straightforward to define an equivalent scaling for $\bar{\mathbf{D}}_{\mathcal{Y}}$. We set

$$\mathbf{D}_{\mathcal{Y}} = \mathbf{I}_{\mathsf{t}} \otimes \mathbf{D} = \sum_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}} \mathbf{E}_{\nu_{\mathsf{t}}} \otimes \mathbf{D}, \tag{4.62}$$

which still has unbounded ranks. For the versions with finite ranks, we stay as general as possible and allow for a different number of terms in the approximation for each temporal index. Hence, for $\mathsf{n} = (n_{\nu_{\mathsf{t}}})_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}} \subset \mathbb{N}_0$, we define

$$\mathbf{D}_{\mathcal{Y},\mathsf{n}} = \sum_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}} \mathbf{E}_{\nu_{\mathsf{t}}} \otimes \mathbf{D}_{n_{\nu_{\mathsf{t}}}}.$$

By definition of w in Theorem 4.26 we have $\Phi_n(s) \leq \Phi_\infty(s)$ for all $s \in \mathbb{R}$. Hence, we obtain

$$\|\mathbf{D}_{\mathcal{Y},\mathsf{n}}\mathbf{D}_{\mathcal{Y}}^{-1}\| \leq 1 \tag{4.63}$$

for all $\mathsf{n} \in \mathbb{N}^{\vee_{\mathsf{t}}}$. For $\mathsf{K} = (K_{\nu_{\mathsf{t}}})_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}} \in \mathbb{R}^{\vee_{\mathsf{t}}}$ we define as before

$$\begin{aligned}\mathsf{M}_{\mathcal{Y}}(\eta;\mathsf{K}) &= (M_{\mathcal{Y}}(\eta;K_{\nu_{\mathsf{t}}}))_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}},\\ \mathsf{M}_{\mathcal{Y},0}(\mathsf{K}) &= (M_{\mathcal{Y},0}(K_{\nu_{\mathsf{t}}}))_{\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}}.\end{aligned}$$

Then if $\mathsf{n} \geq \mathsf{M}_{\mathcal{Y}}(\eta;\mathsf{K})$, by (4.60), we have

$$\begin{aligned}\|\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}(\bar{\mathbf{D}}_{\mathcal{Y}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}}}\| &\leq \delta,\\ \|\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}(\mathbf{D}_{\mathcal{Y}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}}}\| &\leq \eta\end{aligned} \tag{4.64}$$

with Λ_{K} defined in (4.52).

Remark 4.28. The second statement in (4.34) is fulfilled with

$$\begin{aligned}\|\mathbf{D}_{\mathcal{Y}}\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}\| &\leq 1 + \tfrac{\delta}{2},\\ \|\bar{\mathbf{D}}_{\mathcal{Y}}\mathbf{D}_{\mathcal{Y}}^{-1}\| &\leq \left(1 - \tfrac{\delta}{2}\right)^{-1}.\end{aligned} \tag{4.65}$$

The bounds on the operator norms depend only on the choice of δ and are independent of the number of spatial dimensions d. Moreover, for $\mathsf{n} \in \mathbb{N}_0^{\vee_{\mathsf{t}}}$ we have

$$\|\mathbf{D}_{\mathcal{Y},\mathsf{n}}\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}\| \leq 1 + \tfrac{\delta}{2}.$$

Let $\mathsf{K} \in \mathbb{R}^{\vee_{\mathsf{t}}}$ with $\mathsf{K}_{\nu_{\mathsf{t}}} > 1$ for all $\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}$, then for $\mathsf{n} \geq \mathsf{M}_{\mathcal{Y},0}(\mathsf{K})$, we have

$$\|\mathbf{D}_{\mathcal{Y},\mathsf{n}}^{-1}\bar{\mathbf{D}}_{\mathcal{Y}}\mathrm{R}_{\Lambda_{\mathsf{K}}}\| \leq (1-\delta)^{-1}.$$

Hence, by comparing Theorem 4.22 and Theorem 4.28, we see that we get similar properties for both scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ with respect to the equivalent scaling matrices $\bar{\mathbf{D}}_{\mathcal{X}}$ and $\bar{\mathbf{D}}_{\mathcal{Y}}$ as well as for their finite-rank approximations.

4.2.3 Relation to Low-Dimensional Scaling Matrices

In the next section, we analyze the adaptive approximation of rescaled low-rank operators. Here we make use of the compressibility properties of the one-dimensional components as in Theorem 2.30. These compressibility properties are expressed in

terms of the one-dimensional scaling matrices. Therefore, we need the following relation between the high-dimensional scaling matrices $\mathbf{D}$, $\mathbf{D}_{\mathcal{X}}$, and $\mathbf{D}_{\mathcal{Y}}$ and the low-dimensional scaling matrices $\bar{\mathbf{D}}_1^\tau : \mathbb{R}^{\vee_1} \to \mathbb{R}^{\vee_1}$ and $\bar{\mathbf{D}}_t^\tau : \mathbb{R}^{\vee_t} \to \mathbb{R}^{\vee_t}$ which are defined by

$$\begin{aligned} \bar{\mathbf{D}}_1^\tau \mathbf{v} &= \left(\|\psi_\nu\|_{H_0^1(0,1)}^{-\tau} v_\nu \right)_{\nu \in \vee_1}, \quad \bar{\mathbf{D}}_1 \mathbf{v} = \bar{\mathbf{D}}_1^1 \mathbf{v}, \quad \text{for } \mathbf{v} \in \ell_2(\vee_1), \\ \bar{\mathbf{D}}_t^\tau \mathbf{v} &= \left(\|\theta_{\nu_t}\|_{H^1(I)}^{-\tau} v_{\nu_t} \right)_{\nu_t \in \vee_t}, \quad \bar{\mathbf{D}}_t \mathbf{v} = \bar{\mathbf{D}}_t^1 \mathbf{v}, \quad \text{for } \mathbf{v} \in \ell_2(\vee_t) \end{aligned} \tag{4.66}$$

with the one-dimensional wavelet basis functions from Section 3.1.2.

Lemma 4.29 ([BF24, Lemma 5.4]). *For* $\mathbf{M}_1 \in \mathbb{R}^{\vee_1 \times \vee_1}$, $\mathbf{M}_t \in \mathbb{R}^{\vee_t \times \vee_t}$ *and* $\nu_t \in \vee_t$,

$$\|\mathbf{D}[\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{M}_1 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1]\mathbf{D}_{\mathcal{X},\nu_t}\| \leq C_\delta \|\bar{\mathbf{D}}_1 \mathbf{M}_1 \bar{\mathbf{D}}_1\|, \tag{4.67}$$

$$\|\mathbf{D}_{\mathcal{Y}}[\mathbf{M}_t \otimes \mathbf{I}_x]\mathbf{D}_{\mathcal{X}}\| \leq C_\delta \left\|\mathbf{M}_t \bar{\mathbf{D}}_t\right\|, \tag{4.68}$$

where $\mathbf{I}_1 = (\delta_{\nu,\mu})_{\nu,\mu \in \vee_1}$ *is the identity matrix on* $\ell_2(\vee_1)$ *and* $C_\delta = (1 + \frac{\delta}{2})^2$.

Proof. By definition of $\bar{\mathbf{D}}_{\mathcal{X}}$ from (3.26) and $\bar{\mathbf{D}}$ from (3.25),

$$\begin{aligned} \|[\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \bar{\mathbf{D}}_1^{-1} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1]\bar{\mathbf{D}}_{\mathcal{X},\nu_t}\| &\leq 1, \\ \|\bar{\mathbf{D}}[\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \bar{\mathbf{D}}_1^{-1} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1]\| &\leq 1. \end{aligned} \tag{4.69}$$

Combining this with (4.57) and (4.61) yields

$$\begin{aligned} &\|\mathbf{D}[\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{M}_1 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1]\mathbf{D}_{\mathcal{X},\nu_t}\| \\ &\quad \leq C_\delta^2 \left\|\bar{\mathbf{D}}[\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{M}_1 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1]\bar{\mathbf{D}}_{\mathcal{X},\nu_t}\right\| \\ &\quad \leq C_\delta^2 \left\|\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \bar{\mathbf{D}}_1 \mathbf{M}_1 \bar{\mathbf{D}}_1 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1\right\| \end{aligned}$$

and thus (4.67). Due to the particular structure of the operator $\mathbf{D}_{\mathcal{Y}}$ from (4.62),

$$\mathbf{D}_{\mathcal{Y}}(\mathbf{M}_t \otimes \mathbf{I}_x) = (\mathbf{I}_t \otimes \mathbf{D})(\mathbf{M}_t \otimes \mathbf{I}_x) = (\mathbf{M}_t \otimes \mathbf{I}_x)(\mathbf{I}_t \otimes \mathbf{D}) = (\mathbf{M}_t \otimes \mathbf{I}_x)\mathbf{D}_{\mathcal{Y}}.$$

Furthermore, by definition of $\bar{\mathbf{D}}_{\mathcal{X}}$ from (3.26) and $\bar{\mathbf{D}}_{\mathcal{Y}}$ from (3.24), we obtain

$$\|(\bar{\mathbf{D}}_t^{-1} \otimes \mathbf{I}_x)\bar{\mathbf{D}}_{\mathcal{Y}}\bar{\mathbf{D}}_{\mathcal{X}}\| \leq 1.$$

Combining this fact with (4.57) and (4.65) yields (4.68). □

4.3 Adaptive Approximation of Rescaled Low-Rank Operators

In this section, we discuss the application of low-rank operators rescaled by the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ of the previous section. We are particularly interested in the operator $\mathbf{T}$ defined in (3.27). We assume that $\mathbf{T}_x$ has a low-rank representation as in (2.21) with rank R. Thus, we have

$$\mathbf{T}_x = \sum_{\mathsf{n} \in \mathsf{K}_d(R)} c_{\mathsf{n}} \bigotimes_{i=1}^{d} \mathbf{T}_{n_i}^{(i)}, \tag{4.70}$$

with component tensor $(c_{\mathsf{n}})_{\mathsf{n} \in \mathsf{K}_d(R)}$ in hierarchical format with ranks at most R.

Example 4.30. For the heat equation where the spatial operator $\mathbf{T}_\mathrm{x}$ corresponds to the representation of the Laplacian, the operator is, as already seen in (2.24), of the form

$$\mathbf{T}_\mathrm{x} = \mathbf{T}_2 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 + \cdots + \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{T}_2. \tag{4.71}$$

This representation is a special case of that one in Theorem 2.20, as it uses identical one-dimensional operators in each dimension. Hence, $\mathbf{T}_\mathrm{x}$ can be written in the form (4.70) with

$$\mathbf{T}_1^{(i)} = \mathbf{I}_1, \quad \mathbf{T}_2^{(i)} = \mathbf{T}_2 \quad \text{for } i = 1, \dots, d$$

and the corresponding transfer tensors form Theorem 2.20. The corresponding rank of the component tensor is $R = 2$.

When applying the rescaled operators $\mathbf{B}_1$ and $\mathbf{B}_2$ defined in (4.35) to $\mathbf{v} \in \ell_2(\vee)$, the output generally possesses infinite support, even if $\mathbf{v}$ itself is finitely supported. We aim to develop an efficient scheme that approximately applies the operators $\mathbf{B}_1$ and $\mathbf{B}_2$ to finitely supported $\mathbf{v} \in \ell_2(\vee)$ with finite hierarchical ranks. Specifically, for a given error tolerance $\eta > 0$, we seek an approximation $\mathbf{w}_\eta$ such that $\|\mathbf{B}_1\mathbf{v} - \mathbf{w}_\eta\| \leq \eta$. This approximation should have quasi-optimal bounds on hierarchical ranks and lower-dimensional support sizes $\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta))$ with respect to η. Similarly, we want an approximate solution for $\mathbf{B}_2\mathbf{v}$. Note that $\mathbf{B}_2$ maps into $\ell_2(\vee_\mathrm{x})$ instead of $\ell_2(\vee)$.

We start with the adaptive application of the operator

$$\mathbf{B}_1 = \mathbf{B}_\mathrm{t} + \mathbf{B}_\mathrm{x},$$

where

$$\mathbf{B}_\mathrm{t} = \mathbf{D}_\mathcal{X}(\mathbf{T}_\mathrm{t} \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}, \quad \mathbf{B}_\mathrm{x} = \mathbf{D}_\mathcal{Y}(\mathbf{I}_\mathrm{t} \otimes \mathbf{T}_\mathrm{x})\mathbf{D}_\mathcal{X}, \tag{4.72}$$

which is the analogue definition to (3.29) with replaced scaling matrices. In the following, we first consider the operators $\mathbf{B}_\mathrm{x}$ and $\mathbf{B}_\mathrm{t}$ separately and subsequently combine the results for the adaptive application of $\mathbf{B}_1$. In the final part of this section, we discuss the adaptive application of $\mathbf{B}_2$.

4.3.1 Spatial Operator

We start with the adaptive application of the operator $\mathbf{B}_\mathrm{x}$ from (4.72), where $\mathbf{T}_\mathrm{x}$ is in the form (4.70). Specifically, we are interested in the heat equation, i.e. $\mathbf{T}_\mathrm{x}$ of the form (4.71) representing the Laplacian. We assume that the operators $\bar{\mathbf{D}}_1\mathbf{T}_{n_i}^{(i)}\bar{\mathbf{D}}_1$ with $\bar{\mathbf{D}}_1$ defined in (4.66) are s^*-compressible in the sense of Theorem 2.23, i.e., for each $s < s^*$ there exists a sequence $(\mathbf{T}_{n_i,j}^{(i)})_{j\in\mathbb{N}_0}$ with

$$\big\|\bar{\mathbf{D}}_1\big(\mathbf{T}_{n_i}^{(i)} - \mathbf{T}_{n_i,j}^{(i)}\big)\bar{\mathbf{D}}_1\big\| \leq \beta_j\big(\bar{\mathbf{D}}_1\mathbf{T}_{n_i}^{(i)}\bar{\mathbf{D}}_1\big)2^{-sj}. \tag{4.73}$$

For the adaptive application of $\mathbf{B}_\mathrm{x}$ to $\mathbf{v} \in \ell_2(\vee)$, we adopt the core methodology of lower-dimensional problems as in [CDD01]: a priori knowledge about the bi-infinite matrix $\mathbf{B}_\mathrm{x}$, namely its compressibility properties, is combined with a posteriori knowledge of the vector $\mathbf{v}$, namely the magnitude of its entries. In the high-dimensional case, computing all entries of $\mathbf{v}$ is infeasible and suffers from the curse of dimensionality. Instead, as in the high-dimensional elliptic case [BD16], we use the information about the contractions of $\mathbf{v}$, as here in the parabolic setting, the spatio-temporal contractions (4.16). Moreover, we use the compressibility of the one-dimensional operators $\mathbf{T}_{n_i}^{(i)}$ in the low-rank format.

We construct approximations $\mathbf{w}_J$ for $\mathbf{B}_\mathrm{x}\mathbf{v}$ for a refinement parameter $J \in \mathbb{N}$, where a higher J represents a finer approximation. Next, we express the approximations in terms of a given error tolerance $\eta > 0$. We define the approximations for operators $\mathbf{B}_\mathrm{x}$ as in (4.72) with general $\mathbf{T}_\mathrm{x}$ in the form (4.70). As in the one-dimensional case [CDD01], we sort the indices by the magnitude of the values, here by the magnitude of the spatio-temporal contractions. To this end, let $\bar{\Lambda}_j^{(i)}(\mathbf{v})$ be the support of the best 2^j-term approximation of $\pi^{(\mathrm{t},i)}(\mathbf{v})$. Furthermore, let

$$\bar{\Lambda}_{\nu_\mathrm{t},j}^{(i)}(\mathbf{v}) = \big\{\nu_1 : (\nu_\mathrm{t},\nu_1) \in \bar{\Lambda}_j^{(i)}(\mathbf{v})\big\}, \quad j = 0,\dots,J, \tag{4.74}$$

be the corresponding one-dimensional index sets for each temporal index $\nu_\mathrm{t} \in \vee_\mathrm{t}$. Since we have spatial index sets for each time index separately, we can proceed as in the elliptic case [BD16] and distribute them across disjoint sets. For each time index $\nu_\mathrm{t} \in \vee_\mathrm{t}$, let $\bar{\Lambda}_{\nu_\mathrm{t},-1}^{(i)}(\mathbf{v}) = \varnothing$ and

$$\Lambda_{\nu_\mathrm{t},[p]}^{(i)}(\mathbf{v}) = \begin{cases} \bar{\Lambda}_{\nu_\mathrm{t},p}^{(i)}(\mathbf{v}) \setminus \bar{\Lambda}_{\nu_\mathrm{t},p-1}^{(i)}(\mathbf{v}), & p = 0,\dots,J, \\ \vee_1 \setminus \bar{\Lambda}_{\nu_\mathrm{t},J}^{(i)}(\mathbf{v}), & p = J+1, \\ \varnothing, & p > J+1. \end{cases} \tag{4.75}$$

From this, we reconstruct the index set containing all time indices

$$\Lambda_{[p]}^{(\mathrm{t},i)}(\mathbf{v}) = \big\{(\nu_\mathrm{t},\nu_1) : \nu_\mathrm{t} \in \vee_\mathrm{t}, \nu_1 \in \Lambda_{\nu_\mathrm{t},[p]}^{(i)}(\mathbf{v})\big\}. \tag{4.76}$$

Next, we give a bound on the norm of the restricted contractions.

Lemma 4.31. *Let $\mathbf{v} \in \ell_2(\vee)$ with $\pi^{(\mathrm{t},i)}(\mathbf{v}) \in \mathcal{A}^s$ for some $s > 0$. Then with the restriction operator defined in (2.5), we have*

$$\Big\|\mathrm{R}_{\Lambda_{[p]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{v})\Big\| \le 2^{-s(p-1)}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}$$

for $p = 0,\dots,J+1$ and $i = 1,\dots,d$.

Proof. By construction, we have $\Lambda_{[0]}^{(\mathrm{t},i)} = \bar{\Lambda}_0^{(i)}(\mathbf{v})$. Hence, the statement for $p = 0$ follows by Theorem 2.8 (v) combined with $\#\Lambda_{[0]}^{(\mathrm{t},i)} = 1$. For $p \ge 1$, we have

$$\begin{aligned} \Big\|\mathrm{R}_{\Lambda_{[p]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{v})\Big\| &\le \Big\|\pi^{(\mathrm{t},i)}(\mathbf{v}) - \mathrm{R}_{\bar{\Lambda}_{p-1}^{(i)}}\pi^{(\mathrm{t},i)}(\mathbf{v})\Big\| \\ &\le (2^{p-1}+1)^{-s}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s} \\ &\le 2^{-s(p-1)}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}, \end{aligned}$$

where we used Theorem 2.8 (iv) in the second line. □

Additionally, we define the approximation of the one-dimensional operators

$$\tilde{\mathbf{T}}_{n_i,[p]}^{(i)} = \begin{cases} \mathbf{T}_{n_i,J-p}^{(i)}, & p = 0,\dots,J, \\ 0, & p > J \end{cases} \tag{4.77}$$

from fine to coarse approximations, where $(\mathbf{T}_{n_i,j}^{(i)})_{j\in\mathbb{N}_0}$ is the approximation sequence of the compressibility assumption according to (4.73). Based on this, we define for

each time index $\nu_t \in \vee_t$ and $\mathbf{v} \in \ell_2(\vee)$ the approximate one-dimensional operator by

$$\tilde{\mathbf{T}}^{(i)}_{\nu_t,n_i,\mathbf{v}} = \sum_{p\in\mathbb{N}_0} \tilde{\mathbf{T}}^{(i)}_{n_i,[p]} \mathrm{R}_{\Lambda^{(i)}_{\nu_t,[p]}(\mathbf{v})}. \tag{4.78}$$

Since we use the a posteriori information through $\Lambda^{(t,i)}_{[p]}(\mathbf{v})$ for the construction, the adaptive approximation operator depends on the input $\mathbf{v} \in \ell_2(\vee)$. To make it as general as possible, we define the approximate application of the operator, defined by an input $\mathbf{v}$, to a vector $\mathbf{w} \in \ell_2(\vee)$ in the form (4.2) by

$$\begin{aligned}\tilde{\mathbf{T}}_{x,J,\mathbf{v}}\mathbf{w} &= \sum_{\nu_t\in\vee_t} \mathbf{e}_{\nu_t} \otimes \left(\tilde{\mathbf{T}}_{x,\nu_t,J,\mathbf{v}}\mathbf{w}_{\nu_t}\right) \\ &= \sum_{\nu_t\in\vee_t} \mathbf{e}_{\nu_t} \otimes \left(\sum_{\mathbf{n}\in\mathrm{K}_d(R)} c_{\mathbf{n}} \bigotimes_{i=1}^{d} \tilde{\mathbf{T}}^{(i)}_{\nu_t,n_i,\mathbf{v}}\right)\mathbf{w}_{\nu_t}.\end{aligned} \tag{4.79}$$

As a direct consequence, one can see that we approximate the operator $\mathbf{T}_x$ differently for each time index, again by using the information of the contractions of $\mathbf{v}$. Based on this, we define the approximation

$$\tilde{\mathbf{B}}_{x,J}(\mathbf{v}) = \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{x,J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{v}. \tag{4.80}$$

To avoid technicalities, we give the proof of the operator approximation error estimates for the Laplacian, that is, for $\mathbf{T}_x$ of the form (4.71).

Remark 4.32. The following analysis of the Laplacian as a second-order elliptic operator can be easily extended to more general second-order elliptic operators. One example are time-independent second-order elliptic operators $A \equiv A(t)$ in (3.21) with constant coefficients in the diffusion matrix $\mathbf{M}$ in (3.21), as in the elliptic case in [BD16, Section 6.2]. In this case, $\mathbf{B}_x$ again has a structure with the identity in time as in (4.72). Hence, the spatial part of the operator can be treated independently for each temporal index. The analysis can be done analogously to the analysis below by using the results from the elliptic case [BD16, Theorem 34, Lemmas 35 and 36]. In this case, the rank r of the operator $\mathbf{T}_x$ in (4.70) would change. For example, for a tridiagonal diffusion matrix, we obtain rank $r = 5$ [BD16, Example 5].

For notational simplicity, we define the one-dimensional operator $\mathbf{C}\colon \ell_2(\vee_1) \to \ell_2(\vee_1)$ by

$$\mathbf{C}_2 = \bar{\mathbf{D}}_1\mathbf{T}_2\bar{\mathbf{D}}_1.$$

Based on Theorem 2.30, the operator $\mathbf{C}_2$ is s^*-compressible for a suitable choice of the wavelet basis. Under these conditions, we can analyze the adaptive application of the spatial operator.

Lemma 4.33 ([BF24, Lemma 5.5]). *Let $\mathbf{B}_x = \mathbf{D}_{\mathcal{X}}(\mathbf{I}_t \otimes \mathbf{T}_x)\mathbf{D}_{\mathcal{Y}}$ with $\mathbf{T}_x$ from (4.71). Assume that (4.73) holds for $\mathbf{T}_2$ for any $s < s^*$ with the corresponding sequence $(\alpha_j(\mathbf{C}_2))_{j\in\mathbb{N}_0}$ and $(\beta_j(\mathbf{C}_2))_{j\in\mathbb{N}_0}$ of the compressibility definition. Additionally, let $\mathbf{v}, \mathbf{w} \in \ell_2(\vee)$ in the form (4.2) with $\pi^{(t,i)}(\mathbf{v}) \in \mathcal{A}^s$ for $i = 1,\dots,d$. Then for each $J \in \mathbb{N}_0$ and $\tilde{\mathbf{B}}_{x,J}$ defined in (4.80) with the $\mathbf{v}$-dependent partitions (4.75), we have the a posteriori bound*

$$\|\mathbf{B}_x\mathbf{w} - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{x,J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{w}\| \le e_{x,J,\mathbf{v}}(\mathbf{w}), \tag{4.81}$$

with

$$e_{\mathrm{x},J,\mathbf{v}}(\mathbf{w}) = C_\delta \sum_{i=1}^{d} \Big(\sum_{p=0}^{J} 2^{-s(J-p)} \beta_{J-p}(\mathbf{C}_2) \big\| \mathrm{R}_{\Lambda^{(\mathrm{t},i)}_{[p]}(\mathbf{v})} \pi^{(\mathrm{t},i)}(\mathbf{w}) \big\| + \|\mathbf{C}_2\| \big\| \mathrm{R}_{\Lambda^{(\mathrm{t},i)}_{[J+1]}(\mathbf{v})} \pi^{(\mathrm{t},i)}(\mathbf{w}) \big\| \Big), \tag{4.82}$$

C_δ *as in Theorem 4.29, as well as the a priori estimate*

$$\|\mathbf{B}_\mathrm{x}\mathbf{v} - \tilde{\mathbf{B}}_{\mathrm{x},J}(\mathbf{v})\| \le C_\delta 2^{-sJ} \Big(\|\mathbf{C}_2\| + 2^s \, \|\beta(\mathbf{C}_2)\|_{\ell_1} \Big) \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}(\mathbf{v}) \big\|_{\mathcal{A}^s}. \tag{4.83}$$

Moreover, for $i = 1, \dots, d$, *we have the support bound*

$$\#\mathrm{supp}\Big(\pi^{(\mathrm{t},i)} \big(\tilde{\mathbf{B}}_{\mathrm{x},J}(\mathbf{v}) \big) \Big) \le 2 \|\alpha(\mathbf{C}_2)\|_{\ell_1} 2^J. \tag{4.84}$$

Proof. Using the structure of $\mathbf{B}_\mathrm{x}$ and $\mathbf{w}$ as well as the definition of $\tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}$ and the scaling matrices $\mathbf{D}_\mathcal{X}$ and $\mathbf{D}_\mathcal{Y}$, we have

$$\begin{aligned} \|\mathbf{B}_\mathrm{x}\mathbf{w} - \mathbf{D}_\mathcal{Y}\tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}\mathbf{D}_\mathcal{X}\mathbf{w}\| &= \Big\| \mathbf{D}_\mathcal{Y}(\mathbf{I}_\mathrm{t} \otimes \mathbf{T}_\mathrm{x} - \tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}) \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} (\mathbf{e}_{\nu_\mathrm{t}} \otimes (\mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}})) \Big\| \\ &= \Big\| \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \big(\mathbf{D}(\mathbf{T}_\mathrm{x} - \tilde{\mathbf{T}}_{\mathrm{x},\nu_\mathrm{t},J,\mathbf{v}}) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}} \big) \Big\|. \end{aligned}$$

Furthermore, using the definition of $\tilde{\mathbf{T}}_{\mathrm{x},\nu_\mathrm{t},J,\mathbf{v}}$ in (4.79) and the triangle inequality as well as (4.78), we obtain

$$\begin{aligned} &\|\mathbf{B}_\mathrm{x}\mathbf{w} - \mathbf{D}_\mathcal{Y}\tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}\mathbf{D}_\mathcal{X}\mathbf{w}\| \\ &\quad \le \sum_{i=1}^{d} \Big\| \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \big(\mathbf{D}\big(\mathbf{I}_1 \otimes \cdots \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{\nu_\mathrm{t},2,\mathbf{v}}) \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}} \big) \Big\| \\ &\quad \le \sum_{i=1}^{d} \sum_{p \in \mathbb{N}_0} \Big\| \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \big(\mathbf{D}\big(\mathbf{I}_1 \otimes \cdots \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p],\mathbf{v}}) \mathrm{R}_{\Lambda^{(i)}_{\nu_\mathrm{t},[p]}(\mathbf{v})} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}} \big) \Big\| \end{aligned} \tag{4.85}$$

with the natural interpretation of the terms for $i = 1$ and $i = d$. Using that diagonal operators commute as well as (4.4) and (4.67) from Theorem 4.29, we obtain

$$\begin{aligned} &\Big\| \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \mathbf{e}_{\nu_\mathrm{t}} \otimes \big(\mathbf{D}\big(\mathbf{I}_1 \otimes \cdots \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p],\mathbf{v}}) \mathrm{R}_{\Lambda^{(i)}_{\nu_\mathrm{t},[p]}(\mathbf{v})} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}} \big) \Big\|^2 \\ &\quad \le \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \Big\| \mathbf{D}\big(\mathbf{I}_1 \otimes \cdots \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p],\mathbf{v}}) \mathrm{R}_{\Lambda^{(i)}_{\nu_\mathrm{t},[p]}(\mathbf{v})} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \mathbf{w}_{\nu_\mathrm{t}} \Big\|^2 \\ &\quad \le \sum_{\nu_\mathrm{t} \in \mathcal{V}_\mathrm{t}} \Big\| \mathbf{D}\big(\mathbf{I}_1 \otimes \cdots \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p],\mathbf{v}}) \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \Big\|^2 \Big\| \mathrm{R}_{\Lambda^{(i)}_{\nu_\mathrm{t},[p]}(\mathbf{v})} \pi^{(i)}(\mathbf{w}_{\nu_\mathrm{t}}) \Big\|^2 \\ &\quad \le C_\delta^2 \Big\| \bar{\mathbf{D}}_1 (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p]}) \bar{\mathbf{D}}_1 \Big\|^2 \Big\| \mathrm{R}_{\Lambda^{(\mathrm{t},i)}_{[p]}(\mathbf{v})} \pi^{(\mathrm{t},i)}(\mathbf{w}) \Big\|^2. \end{aligned}$$

By definition, we have $\Lambda^{(\mathrm{t},i)}_{[p]}(\mathbf{v}) = \varnothing$ for $p \ge J+2$ and $\tilde{\mathbf{T}}^{(i)}_{2,[J+1]} = 0$. Furthermore, by definition of $\tilde{\mathbf{T}}^{(i)}_{2,[p]}$ in (4.77) and the s^*-compressibility of $\mathbf{C}_2$ in (4.73), we have

$$\Big\| \bar{\mathbf{D}}_1 (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p]}) \bar{\mathbf{D}}_1 \Big\| = \Big\| \bar{\mathbf{D}}_1 (\mathbf{T}_2 - \mathbf{T}_{2,J-p}) \bar{\mathbf{D}}_1 \Big\| \le \beta_{J-p}(\mathbf{C}_2) 2^{-s(J-p)} \quad \text{for } p = 0, \dots, J.$$

Inserting the previous estimates into (4.85) yields (4.81).

The a priori estimate (4.83) is a direct consequence of the a posteriori estimate (4.81) combined with Theorem 4.31.

Since the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ are diagonal operators, their application does not change the support. As in the one-dimensional case, one can see that by the definition of $\tilde{\mathbf{T}}^{(i)}_{n_i,[p]}$ in (4.78), we apply in each term operators with $\alpha_{J-j}2^{J-j}$ entries in each column and row to vectors with support size at most 2^j. Hence, the support size is bounded by

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{B}}_{\mathrm{x},J}(\mathbf{v}))\right) \leq 2(\alpha_J(\mathbf{C}_2)2^J2^0 + \cdots + \alpha_0(\mathbf{C}_2)2^02^J) \leq 4\|\alpha(\mathbf{C}_2)\|_{\ell_1}2^J, \tag{4.86}$$

which yields (4.84). The additional factor two is based on the rank of the representation of the Laplacian. □

The support bound can be further improved using the structure of the operator $\mathbf{T}_{\mathrm{x}}$ as in Theorem 4.30 with $\mathbf{T}^{(i)}_1 = \mathbf{I}_1$ being the identity operator. Using the properties of the identity, would lead to the bound

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{B}}_{\mathrm{x},J}(\mathbf{v}))\right) \leq (2\|\alpha(\mathbf{C}_2)\|_{\ell_1}+1)2^J.$$

However, the bound (4.84) indicates the generalizability to more general second-order elliptic operators, since the factor two corresponds to the rank of the operator $\mathbf{T}_{\mathrm{x}}$.

Due to the identity in the time variable, the operator $\mathbf{B}_{\mathrm{x}}$ does not involve any interactions between the different spatial components $(\mathbf{v}_{\nu_{\mathrm{t}}})_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}}$ of the input $\mathbf{v}$. Hence, the estimates are similar to the elliptic case [BD16, Lemma 36].

The described approximate application of the spatial operator $\mathbf{B}_{\mathrm{x}}$ is not yet practically realizable, as it still involves the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$, which are of infinite rank. In Section 4.3.3, we discuss the necessary modifications if we use the rank-truncated scaling matrices.

4.3.2 Temporal Operator

In this section, we discuss the adaptive approximate application of the temporal operator $\mathbf{B}_{\mathrm{t}}$ defined in (4.72). By applying the operator without scaling matrices, i.e. $\mathbf{T}_{\mathrm{t}} \otimes \mathbf{I}_{\mathrm{x}}$, to a given $\mathbf{v} \in \ell_2(\vee)$ in the form (4.2), we obtain

$$(\mathbf{T}_{\mathrm{t}} \otimes \mathbf{I}_{\mathrm{x}})\mathbf{v} = \sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}} (\mathbf{T}_{\mathrm{t}}\mathbf{e}_{\nu_{\mathrm{t}}}) \otimes \mathbf{v}_{\nu_{\mathrm{t}}} = \sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}} \mathbf{e}_{\nu_{\mathrm{t}}} \otimes \Big(\sum_{\mu_{\mathrm{t}}\in\vee_{\mathrm{t}}} (\mathbf{T}_{\mathrm{t}})_{\mu_{\mathrm{t}},\nu_{\mathrm{t}}}\mathbf{v}_{\mu_{\mathrm{t}}}\Big). \tag{4.87}$$

Hence, in contrast to the spatial operator without scaling matrices, the output of the operator $\mathbf{T}_{\mathrm{x}} \otimes \mathbf{I}_{\mathrm{x}}$ will in general have spatial components with infinite ranks if $\#\operatorname{supp}(\pi^{(\mathrm{t})}(\mathbf{v})) = \infty$, even if all spatial components of the input $\mathbf{v}$ have finite ranks. This fact is based on the interactions between the different spatial components caused by the operator $\mathbf{T}_{\mathrm{t}}$. Moreover, even if the input $\mathbf{v}$ has finite support, the output of the operator $\mathbf{w} = (\mathbf{T}_{\mathrm{t}} \otimes \mathbf{I}_{\mathrm{x}})\mathbf{v}$ might in general have spatial components $\mathbf{w}_{\nu_{\mathrm{t}}}$ with

$$\operatorname{rank}_\infty(\mathbf{w}_{\nu_{\mathrm{t}}}) = \operatorname{rank}_\infty(\mathbf{v})\#\operatorname{supp}\left(\pi^{(\mathrm{t})}(\mathbf{v})\right),$$

which is a potential huge rank increase. This increase is based on the representation (4.87) and on the rank bounds of the sum of different low-rank tensors in Section 2.4.5.1. Moreover, in general $\mathbf{w}$ will be infinitely supported, even if $\mathbf{v}$ is finitely supported.

Due to the potential strong rank increase, we require stronger compressibility assumptions for $\mathbf{T}_\mathrm{t}$. By Theorem 2.31 we know that

$$\mathbf{C}_\mathrm{t} = \mathbf{T}_\mathrm{t}\bar{\mathbf{D}}_\mathrm{t}$$

is super-compressible in the sense of Theorem 2.25 and Theorem 2.27 for a suitable choice of the temporal wavelet basis. For the super-compressible operator $\mathbf{T}_\mathrm{t}$, we thus assume without loss of generality that there exists a sequence $(\mathbf{T}_{\mathrm{t},j})_{j\in\mathbb{N}_0}$ such that $\|(\mathbf{T}_\mathrm{t} - \mathbf{T}_{\mathrm{t},j})\bar{\mathbf{D}}_\mathrm{t}\| \le \beta_j 2^{-sj}$ with exponent $s > 0$ as in the compressibility of the spatial operator $\mathbf{C}_2$ in (4.73). We recall that by Theorem 2.27, the operator $\mathbf{T}_{\mathrm{t},J}$ has at most cJ nonzero entries in each row and column for each $J \in \mathbb{N}_0$.

As in the spatial case, we construct the adaptive approximation by combining the a priori knowledge of the operator $\mathbf{B}_\mathrm{t}$ with the a posteriori knowledge of the input $\mathbf{v}$ by using the contractions $\pi^{(\mathrm{t},i)}(\mathbf{v})$. We define

$$\begin{aligned}\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v}) &= \bigcup_{\nu_\mathrm{t}\in\mathcal{V}_\mathrm{t}} \{\nu_\mathrm{t}\} \times \bar{\Lambda}_{\nu_\mathrm{t},j}^{(1)}(\mathbf{v}) \times \cdots \times \bar{\Lambda}_{\nu_\mathrm{t},j}^{(d)}(\mathbf{v}),\\ \mathbf{v}_j &= \mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{v},\end{aligned} \tag{4.88}$$

which is the combination of the best 2^j-term approximations of the individual spatial dimensions by the index sets $\bar{\Lambda}_{\nu_\mathrm{t},j}^{(i)}(\mathbf{v})$ defined in (4.74). For brevity, we define $\bar{\Lambda}_{-1}^{(\mathrm{t})}(\mathbf{v}) = \varnothing$. As an immediate consequence of the definition, we have

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{v}_j)\right) \le 2^j, \quad j \in \mathbb{N}_0. \tag{4.89}$$

In general, the lower-dimensional support can be smaller than 2^j, since it is possible that $\bar{\Lambda}_{\nu_\mathrm{t},j}^{(k)}(\mathbf{v}) = \varnothing$ even if $\bar{\Lambda}_{\nu_\mathrm{t},j}^{(i)}(\mathbf{v}) \neq \varnothing$, resulting in $(\mathbf{v}_j)_{\nu_\mathrm{t}} = 0$.

For a given refinement parameter $J \in \mathbb{N}_0$ and $\mathbf{v}_j$ from (4.88), we define the approximation

$$\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}) = \mathbf{D}_\mathcal{Y}(\mathbf{T}_{\mathrm{t},J} \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}\mathbf{v}_0 + \sum_{j=1}^{J} \mathbf{D}_\mathcal{Y}(\mathbf{T}_{\mathrm{t},J-j} \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}(\mathbf{v}_j - \mathbf{v}_{j-1}),$$

where the operators $\mathbf{T}_{\mathrm{t},j}$ are as in Theorem 2.27. This representation aligns with the scheme for the spatial operator, since we apply finer approximations of the operator to the more influential parts of the input $\mathbf{v}$. For making use of the super-compressibility, we reformulate the approximation by reordering the terms to

$$\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}) = \sum_{j=0}^{J-1} \mathbf{D}_\mathcal{Y}((\mathbf{T}_{\mathrm{t},J-j} - \mathbf{T}_{\mathrm{t},J-j-1}) \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}\mathbf{v}_j + \mathbf{D}_\mathcal{Y}(\mathbf{T}_{\mathrm{t},0} \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}\mathbf{v}_J. \tag{4.90}$$

Here we apply consecutive differences of the operator approximations $\mathbf{T}_{\mathrm{t},j+1} - \mathbf{T}_{\mathrm{t},j}$, which by the super-compressibility have only a constant number of nonzero entries in each row and column.

The approximation $\tilde{\mathbf{B}}_{\mathrm{t},J}$ is defined with respect to the scaling matrices $\mathbf{D}_\mathcal{X}$ and $\mathbf{D}_\mathcal{Y}$. In the practical realization, we replace them by their rank-truncated versions

from Section 4.2. Hence, we define the approximate application of the operator defined by $\mathbf{v}$ to a general vector $\mathbf{w} \in \ell_2(\vee)$. To this end, let

$$\begin{aligned}\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}} &= \sum_{j=0}^{J-1}\big((\mathbf{T}_{\mathrm{t},J-j} - \mathbf{T}_{\mathrm{t},J-j-1}) \otimes \mathbf{I}_{\mathrm{x}}\big)\mathrm{R}_{\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v})} + (\mathbf{T}_{\mathrm{t},0} \otimes \mathbf{I}_{\mathrm{x}})\mathrm{R}_{\bar\Lambda_J^{(\mathrm{t})}(\mathbf{v})} \\ &= (\mathbf{T}_{\mathrm{t},J} \otimes \mathbf{I}_{\mathrm{x}})\mathrm{R}_{\bar\Lambda_0^{(\mathrm{t})}(\mathbf{v})} + \sum_{j=1}^{J}(\mathbf{T}_{\mathrm{t},J-j} \otimes \mathbf{I}_{\mathrm{x}})\Big(\mathrm{R}_{\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v})} - \mathrm{R}_{\bar\Lambda_{j-1}^{(\mathrm{t})}(\mathbf{v})}\Big),\end{aligned} \tag{4.91}$$

and note that

$$\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}) = \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{v}.$$

By construction, we have $\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v}) \subseteq \bar\Lambda_{j+1}^{(\mathrm{t})}(\mathbf{v})$ for each $j \in \mathbb{N}_0$. Hence, for each $\mathbf{w} \in \ell_2(\vee)$ we have

$$\begin{aligned}\big\|\mathrm{R}_{\bar\Lambda_{j+1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\big\| &= \big\|\mathrm{R}_{\bar\Lambda_{j+1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v})}\big(\mathrm{R}_{\bar\Lambda_{j+1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\big)\big\| \\ &\le \sum_{i=1}^{d}\big\|\mathrm{R}_{\bar\Lambda_{j+1}^{(i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w}) - \mathrm{R}_{\bar\Lambda_j^{(i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\big\| \\ &\le \sum_{i=1}^{d}\big\|\mathrm{R}_{\Lambda_{[j+1]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\big\|,\end{aligned} \tag{4.92}$$

where we used (4.18) and the definition of $\Lambda_{[j]}^{(\mathrm{t},i)}(\mathbf{v})$ in (4.76).

Remark 4.34. As an alternative, instead of using the index sets $\bar\Lambda_j^{(\mathrm{t})}(\mathbf{v})$ in (4.88), one can use the index sets from the coarsening routine $\Lambda(\mathbf{v}, N)$ in (4.19) with $N = 2^j$. In the present setting, this would lead to essentially the same results.

Lemma 4.35 ([BF24, Lemma 5.10])**.** *Let $\mathbf{B}_{\mathrm{t}} = \mathbf{D}_{\mathcal{Y}}(\mathbf{T}_{\mathrm{t}} \otimes \mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}$ and assume that $\mathbf{T}_{\mathrm{t}}$ is super-compressible in the sense of Theorem 2.27. Given any $\mathbf{v}, \mathbf{w} \in \ell_2(\vee)$ and $J \in \mathbb{N}_0$. Then whenever $\pi^{(\mathrm{t},i)}(\mathbf{v}) \in \mathcal{A}^s$ for some $0 < s < s^*$ and $i = 1, \dots, d$, one has the a posteriori error bound*

$$\|\mathbf{B}_{\mathrm{t}}\mathbf{w} - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{w}\| \le e_{\mathrm{t},J,\mathbf{v}}(\mathbf{w}), \tag{4.93}$$

where

$$\begin{aligned}e_{\mathrm{t},J,\mathbf{v}}(\mathbf{w}) = C_\delta \sum_{i=1}^{d}\Big(\sum_{j=0}^{J} 2^{-s(J-j)}\beta(\mathbf{C}_t)_{J-j}\big\|\mathrm{R}_{\Lambda_{[j]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\big\| \\ + \|\mathbf{C}_{\mathrm{t}}\|\big\|\mathrm{R}_{\Lambda_{[J+1]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\big\|\Big)\end{aligned} \tag{4.94}$$

as well as the a priori error bound

$$\|\mathbf{B}_{\mathrm{t}}\mathbf{v} - \tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v})\| \le C_\delta 2^{-sJ}(2^s\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\mathbf{C}_{\mathrm{t}}\|)\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}, \tag{4.95}$$

where $\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v})$ is defined by (4.90). Moreover, one has the support estimate

$$\#\operatorname{supp}\Big(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}))\Big) \le c(\mathbf{C}_t)2^J, \quad i = 1, \dots, d. \tag{4.96}$$

Proof. Using the identity

$$\mathbf{B}_\mathrm{t}\mathbf{w} = \sum_{j=0}^{J} \mathbf{B}_\mathrm{t}\Big(\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\Big) + \mathbf{B}_\mathrm{t}\Big(\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\Big)$$

combined with the triangle inequality and the definition of $\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}}$ in (4.91), we obtain

$$\begin{aligned}
\|\mathbf{B}_\mathrm{t}\mathbf{w} - \mathbf{D}_\mathcal{Y}\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}}\mathbf{D}_\mathcal{X}\mathbf{w}\| &\le \sum_{j=0}^{J} \|\mathbf{D}_\mathcal{Y}((\mathbf{T}_{\mathrm{t},J-j} - \mathbf{T}_\mathrm{t}) \otimes \mathbf{I}_\mathrm{x})\mathbf{D}_\mathcal{X}\| \left\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| \\
&\quad + \|\mathbf{B}_\mathrm{t}\| \left\|\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| \\
&\le \sum_{j=0}^{J} C_\delta \|(\mathbf{T}_{\mathrm{t},J-j} - \mathbf{T}_\mathrm{t})\hat{\mathbf{D}}_\mathrm{t}\| \left\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| \\
&\quad + C_\delta \|\mathbf{C}_\mathrm{t}\| \left\|\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\|,
\end{aligned}$$

where we used (4.68) from Theorem 4.29 in the last step. Moreover, using the super-compressibility of $\mathbf{T}_\mathrm{t}$ and the estimate (4.92) yields (4.93).

The a priori estimate (4.95) follows as for the bound (4.83) of the spatial operator by further estimating the a posteriori estimate (4.93) using Theorem 4.31.

As noted above, $(\mathbf{T}_{\mathrm{t},j+1} - \mathbf{T}_{\mathrm{t},j})$ has only $c(\mathbf{C}_t)$ nonzero entries in each column and by choice $\mathbf{T}_{\mathrm{t},0} = 0$. In view of (4.89) and (4.90), we have

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}))\right) \le \sum_{j=0}^{J-1} c(\mathbf{C}_t)2^j \le c(\mathbf{C}_t)2^J$$

since the diagonal operators $\mathbf{D}_\mathcal{X}$ and $\mathbf{D}_\mathcal{Y}$ do not change the support. □

Again, as for the spatial operator, the approximate application of the temporal operator $\mathbf{B}_\mathrm{t}$ is not yet practically realizable since it still involves the scaling matrices $\mathbf{D}_\mathcal{X}$ and $\mathbf{D}_\mathcal{Y}$, which are of infinite rank.

4.3.3 Combination of Operators

In this section, we investigate the combination of the spatial and temporal operator approximations. In particular, we rephrase the results in terms of a given error tolerance $\eta > 0$ instead of refinement levels J. Furthermore, we replace the scaling matrices $\mathbf{D}_\mathcal{X}$ and $\mathbf{D}_\mathcal{Y}$ by the rank-truncated scaling matrices $\mathbf{D}_{\mathcal{X},\mathrm{n}_1}$ and $\mathbf{D}_{\mathcal{Y},\mathrm{n}_2}$ to arrive at a practical realizable versions.

First, we state the results for the combination of the two operators $\mathbf{B}_1 = \mathbf{B}_\mathrm{x} + \mathbf{B}_\mathrm{t}$ from (4.35) in terms of a refinement level, which is a direct consequence of Theorems 4.33 and 4.35.

Corollary 4.36. *Under the assumptions of Theorem 4.33 on* $\mathbf{T}_\mathrm{x}$ *and Theorem 4.35 on* $\mathbf{T}_\mathrm{t}$, *let* $\mathbf{v} \in \ell_2(\vee)$ *as well as*

$$\tilde{\mathbf{B}}_{1,J}(\mathbf{v}) = \tilde{\mathbf{B}}_{\mathrm{t},J}(\mathbf{v}) + \tilde{\mathbf{B}}_{\mathrm{x},J}(\mathbf{v})$$

and

$$\tilde{\mathbf{T}}_{J,\mathbf{v}} = \tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}} + \tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}. \tag{4.97}$$

Then for $\mathbf{w} \in \ell_2(\vee)$ *and*

$$e_{J,\mathbf{v}}(\mathbf{w}) = e_{\mathrm{x},J,\mathbf{v}}(\mathbf{w}) + e_{\mathrm{t},J,\mathbf{v}}(\mathbf{w}), \tag{4.98}$$

with $e_{\mathrm{x},J,\mathbf{v}}(\mathbf{w})$ *defined in* (4.82) *and* $e_{\mathrm{t},J,\mathbf{v}}(\mathbf{w})$ *in* (4.94)*, one has the a posteriori bound*

$$\|\mathbf{B}_1\mathbf{w} - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{w}\| \le e_{J,\mathbf{v}}(\mathbf{w}).$$

Furthermore, we have the a priori error bound

$$\|\mathbf{B}_1\mathbf{v} - \tilde{\mathbf{B}}_{1,J}(\mathbf{v})\| \le C_\delta C_{\mathrm{t},1} 2^{-sJ} \sum_{i=1}^{d} \big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}$$

with $C_{\mathrm{t},1} = 2^s(\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\beta(\mathbf{C}_2)\|_{\ell_1}) + \|\mathbf{C}_\mathrm{t}\| + \|\mathbf{C}_2\|$*, and*

$$\#\mathrm{supp}\big(\pi^{(\mathrm{t},i)}\big(\tilde{\mathbf{B}}_{1,J}(\mathbf{v})\big)\big) \le \hat{\alpha}(\mathbf{B}_1)2^J \tag{4.99}$$

with $\hat{\alpha}(\mathbf{B}_1) = c(\mathbf{C}_t) + 4\|\alpha(\mathbf{C}_2)\|_{\ell_1}$.

For notational simplicity, we write $e_J(\mathbf{v}) = e_{J,\mathbf{v}}(\mathbf{v})$.

Remark 4.37. For practical purposes, we are interested in the adaptive application of the operator $\mathbf{B}_1$ to finitely supported vectors $\mathbf{v} \in \ell_2(\vee)$. In this case, there exists a $p(\mathbf{v}) \in \mathbb{N}_0$ such that $\Lambda_{[p]}^{(\mathrm{t},i)}(\mathbf{v}) = \varnothing$ for $i = 1,\dots,d$ for all $p > p(\mathbf{v})$. By definition of $e_{J,\mathbf{v}}$ in (4.98), for $J \ge p(\mathbf{v})$ we have

$$\begin{aligned} e_J(\mathbf{v}) &\le C_\delta \sum_{i=1}^{d}\sum_{j=0}^{J} 2^{-s(J-j)}(\beta(\mathbf{C}_t)_{J-j} + \beta_{J-p}(\mathbf{C}_2))\big\|\mathrm{R}_{\Lambda_{[j]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{v})\big\| \\ &\le C_\delta 2^{-s(J-p(\mathbf{v}))} \sum_{i=1}^{d}\sum_{j=0}^{p(\mathbf{v})} 2^{-s(p(\mathbf{v})-j)}\tilde{\beta}\big\|\mathrm{R}_{\Lambda_{[j]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{v})\big\| \end{aligned}$$

with $\tilde{\beta} = \max_{j\in\mathbb{N}_0}(\beta(\mathbf{C}_t)_j + \beta_j(\mathbf{C}_2)) < \infty$. Since the upper bound monotonically tends to zero as J increases, for each $\eta > 0$ there exists a minimal $J \in \mathbb{N}_0$ such that $e_J(\mathbf{v}) \le \eta$.

Theorem 4.36 is still not suitable for practical purposes since the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ have unbounded ranks. Hence, we do not only have to replace $\mathbf{T}$ by some approximation $\tilde{\mathbf{T}}$, but also the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$ by the scaling matrices $\mathbf{D}_{\mathcal{X},\mathsf{n}_1}$ and $\mathbf{D}_{\mathcal{Y},\mathsf{n}_2}$ which have bounded ranks for each time index. The following lemma gives us an estimate of the error when replacing $\mathbf{T}$ by its approximate version, as well as the scaling matrices by their truncated versions.

Lemma 4.38 ([BF24, Lemma 5.13])**.** *Let* $\mathbf{B}_1 = \mathbf{D}_{\mathcal{Y}}\mathbf{T}\mathbf{D}_{\mathcal{X}}$ *with* $\mathbf{T}$ *be defined as in* (3.27)*,* $\tilde{\mathbf{T}} \in \mathbb{R}^{\vee\times\vee}$ *and*

$$\mathbf{G} = \mathbf{B}_1 - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}} = \mathbf{D}_{\mathcal{Y}}(\mathbf{T} - \tilde{\mathbf{T}})\mathbf{D}_{\mathcal{X}}.$$

Let $\mathbf{v} \in \ell_2(\vee)$ *and* $\mathsf{K}_1, \mathsf{K}_2 \in \mathbb{R}^{\vee_\mathrm{t}}$ *be chosen such that*

$$\mathrm{supp}(\mathbf{v}) \subseteq \Lambda_{\mathsf{K}_1}, \quad \mathrm{supp}(\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}}\mathbf{v}) \subseteq \Lambda_{\mathsf{K}_2},$$

where Λ_{K_1} *and* Λ_{K_2} *are defined by* (4.52)*. Then whenever* $\mathsf{n}_1 \ge \mathsf{M}_{\mathcal{X}}(\eta, \mathsf{K}_1)$*,* $\mathsf{n}_2 \ge \mathsf{M}_{\mathcal{Y}}(\eta, \mathsf{K}_2)$*, we have*

$$\begin{aligned} \big\|(\mathbf{B}_1 - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\big\| &\le \|\mathbf{G}\mathbf{v}\| + \|\mathbf{G}(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\| \\ &\quad + \eta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{G}(\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\| + 2\eta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\|\|\mathbf{v}\|, \end{aligned} \tag{4.100}$$

where $\mathbf{I} = (\delta_{\nu,\mu})_{\nu,\mu\in\vee}$ *is the identity operator on* $\ell_2(\vee)$.

Proof. We follow the lines of the proof of [BD16, Lemma 32], where a similar statement was proven in the context of elliptic problems, but with different scaling matrices.

By definition of $\mathbf{G}$ and the triangle inequality, we have

$$\begin{aligned}\left\|(\mathbf{B}_1 - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| &\le \|\mathbf{G}\mathbf{v}\| + \left\|(\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| \\ &\le \|\mathbf{G}\mathbf{v}\| + \left\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| \\ &\quad + \left\|(\mathbf{D}_{\mathcal{Y}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2})\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{v}\right\|.\end{aligned} \tag{4.101}$$

To arrive at the estimate (4.100), we further have to estimate the last two terms on the right-hand side. Using $\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}} = \mathbf{B}_1 - \mathbf{G}$, we have

$$\begin{aligned}\left\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| &= \left\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}}(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| \\ &\le \left\|\mathbf{B}_1(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| + \left\|\mathbf{G}(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\|\end{aligned}$$

and

$$\begin{aligned}\left\|\mathbf{B}_1(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| &\le \|\mathbf{B}_1\|\|\mathbf{D}_{\mathcal{X}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\|\|\bar{\mathbf{D}}_{\mathcal{X}}^{-1}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}_1}}\|\|\mathbf{v}\| \\ &\le (1 - \tfrac{\delta}{2})^{-1}\eta\|\mathbf{B}_1\|\|\mathbf{v}\|,\end{aligned}$$

where we used (4.54) and (4.57) in the final step as well as $\mathbf{v} = \mathrm{R}_{\Lambda_{\mathsf{K}_1}}\mathbf{v}$ and the definition of n_1.

For the last term on the right-hand side in (4.101), we can proceed similarly and use that diagonal operators commute. We obtain

$$\begin{aligned}\left\|(\mathbf{D}_{\mathcal{Y}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2})\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{v}\right\| &\le \|\mathbf{D}_{\mathcal{Y}}^{-1}\bar{\mathbf{D}}_{\mathcal{Y}}\|\|\bar{\mathbf{D}}_{\mathcal{Y}}^{-1}(\mathbf{D}_{\mathcal{Y}} - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2})\mathrm{R}_{\Lambda_{\mathsf{K}_2}}\|\left\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{v}\right\| \\ &\le (1 - \tfrac{\delta}{2})^{-1}\eta\left\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X}}(\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| \\ &\le (1 - \tfrac{\delta}{2})^{-1}\eta\Big(\left\|\mathbf{G}(\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| + \|\mathbf{B}_1\|\|\mathbf{v}\|\Big),\end{aligned}$$

where we used (4.64) and (4.65) as well as the definition of Λ_{K_2} and n_2 in the second line and (4.56) in the last line. Inserting these estimates into (4.101) yields (4.100). □

Remark 4.39. In the context of interest, we assume $\tilde{\mathbf{T}}$ to be an approximation of $\mathbf{T}$. Hence, under the assumptions of Theorem 4.38 with the additional natural assumption

$$\|\mathbf{G}\| = \|\mathbf{D}_{\mathcal{Y}}(\mathbf{T} - \tilde{\mathbf{T}})\mathbf{D}_{\mathcal{X}}\| \le \|\mathbf{D}_{\mathcal{Y}}\mathbf{T}\mathbf{D}_{\mathcal{X}}\| = \|\mathbf{B}_1\|,$$

we get the modified result

$$\left\|(\mathbf{B}_1 - \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\right\| \le \|\mathbf{G}\mathbf{v}\| + 4\eta(1 - \tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\|\|\mathbf{v}\|.$$

The proof of this estimate is straightforward by further estimating the second and third terms on the right-hand side of (4.100) in the same manner as the other terms.

Next, we state a result about the adaptive approximation of the operator $\mathbf{B}_1$ with the rank-truncated scaling matrices in terms of a given error tolerance $\eta > 0$. To this end, we define the corresponding refinement level for the given tolerance

$$J(\eta;\mathbf{v}) = \min\big\{J \in \mathbb{N}_0 : e_J(\mathbf{v}) \le \tfrac{\eta}{4}\big\}, \tag{4.102}$$

which exists by Theorem 4.37. Furthermore, we define

$$c(\eta;\mathbf{v}) = \frac{\eta}{4\|\mathbf{B}_1\|\|\mathbf{v}\|}(1 - \tfrac{\delta}{2})$$

and the quantities related to the number of terms in the rank truncated scaling matrices by

$$\mathsf{m}_{\mathcal{X}}(\eta;\mathbf{v}) = \mathsf{M}_{\mathcal{X}}\big(c(\eta;\mathbf{v});\mathsf{K}(J(\eta;\mathbf{v});\mathbf{v})\big),$$
$$\mathsf{m}_{\mathcal{Y}}(\eta;\mathbf{v}) = \mathsf{M}_{\mathcal{Y}}\big(c(\eta;\mathbf{v});\mathsf{K}(J(\eta;\mathbf{v});\mathbf{v})\big)$$

with $\mathsf{m}_{\mathcal{X}}$, $\mathsf{m}_{\mathcal{Y}}$ defined in (4.53) and (4.59) respectively, and where

$$\mathsf{K}(J;\mathbf{v}) = \min\{\mathsf{K} > 0 : \operatorname{supp}(\mathbf{v}) \cup \operatorname{supp}(\tilde{\mathbf{T}}_{J,\mathbf{v}}\mathbf{v}) \subseteq \Lambda_{\mathsf{K}}\} \tag{4.103}$$

with Λ_{K} defined in (4.52). Again, the minimum is to be understood element-wise for each K_{ν_t} with $\nu_t \in \vee_t$.

Theorem 4.40 ([BF24, Theorem 5.14]). *Let $0 \neq \mathbf{v} \in \ell_2(\vee)$ and $0 < \eta < 2\|\mathbf{B}_1\|\|\mathbf{v}\|$. We fix*

$$J = J(\eta;\mathbf{v}), \quad \mathsf{n}_1 = \mathsf{m}_{\mathcal{X}}(\eta;\mathbf{v}), \quad \mathsf{n}_2 = \mathsf{m}_{\mathcal{Y}}(\eta;\mathbf{v}).$$

Then with $\mathbf{w}_\eta = \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}_{J,\mathbf{v}}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{v}$ and $\tilde{\mathbf{T}}_{J,\mathbf{v}}$ from (4.97), we have

$$\|\mathbf{B}_1\mathbf{v} - \mathbf{w}_\eta\| \leq \eta\,. \tag{4.104}$$

Proof. To simplify notation, we write

$$\tilde{\mathbf{T}}_J = \tilde{\mathbf{T}}_{J,\mathbf{v}}, \quad \tilde{\mathbf{v}} = (\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}, \quad \tilde{\mathbf{d}} = \mathbf{v} - \tilde{\mathbf{v}} = (\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}$$

as well as

$$\mathsf{K} = \mathsf{K}(J(\eta;\mathbf{v}),\mathbf{v}), \quad \zeta = c(\eta;\mathbf{v}).$$

By definition of e_J, we have for diagonal operators $\tilde{\mathbf{D}}$ with $\|\tilde{\mathbf{D}}\| \leq \tilde{c}$ that

$$e_{J,\mathbf{v}}(\tilde{\mathbf{D}}\mathbf{v}) \leq \tilde{c}e_J(\mathbf{v}) = \tilde{c}e_{J,\mathbf{v}}(\mathbf{v}). \tag{4.105}$$

Furthermore, by the choice of n_1 as well as (4.54) and (4.57), we have

$$\|\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathrm{R}_{\Lambda_{\mathsf{K}}}\| \leq \|\mathbf{D}_{\mathcal{X}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\|\|\bar{\mathbf{D}}_{\mathcal{X}}^{-1}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathrm{R}_{\Lambda_{\mathsf{K}}}\| \leq (1-\tfrac{\delta}{2})^{-1}\zeta. \tag{4.106}$$

Applying Theorem 4.38 with $\tilde{\mathbf{T}} = \tilde{\mathbf{T}}_{J,\mathbf{v}}$ yields

$$\begin{aligned}
\|\mathbf{B}_1\mathbf{v} - \mathbf{w}_\eta\| &\leq \|\mathbf{B}_1\mathbf{v} - \tilde{\mathbf{B}}_{1,J}(\mathbf{v})\| + \|\mathbf{B}_1\tilde{\mathbf{d}} - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_J\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{d}}\| \\
&\quad + \zeta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\tilde{\mathbf{v}} - \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_J\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{v}}\| + 2\zeta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\|\|\mathbf{v}\| \\
&\leq e_J(\mathbf{v}) + e_{J,\mathbf{v}}(\tilde{\mathbf{d}}) + \zeta(1-\tfrac{\delta}{2})^{-1}e_{J,\mathbf{v}}(\tilde{\mathbf{v}}) + 2\zeta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\|\|\mathbf{v}\| \\
&\leq e_J(\mathbf{v})\left(1 + 2\zeta(1-\tfrac{\delta}{2})^{-1}\right) + 2\zeta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_1\|\|\mathbf{v}\| \\
&\leq 2e_J(\mathbf{v}) + \tfrac{\eta}{2},
\end{aligned}$$

where we used (4.105) in combination with (4.56) and (4.106) in the fourth line as well as the choice of ζ and $\eta \leq 2\|\mathbf{B}_1\|\|\mathbf{v}\|$ in the final step. By the choice of J according to (4.102), we arrive at (4.104). □

Remark 4.41. The restriction $\eta \leq 2\|\mathbf{B}_1\|\|\mathbf{v}\|$ in Theorem 4.40 is not a real restriction, since for $\eta \geq \|\mathbf{B}_1\|\|\mathbf{v}\|$ we have $\|\mathbf{B}_1\mathbf{v}\| \leq \eta$. Hence, we can approximate $\mathbf{B}_1\mathbf{v}$ by the zero sequence.

Remark 4.42. For the proof of Theorem 4.40, it is not necessary to use the same K for both scaling matrices. For the practical realization, it is beneficial to choose

$$\begin{aligned}\mathsf{K}_{\mathcal{X}}(J;\mathbf{v}) &= \min\{\mathsf{K}>0 : \operatorname{supp}(\mathbf{v}_J)\subseteq\Lambda_{\mathsf{K}}\},\\ \mathsf{K}_{\mathcal{Y}}(J;\mathbf{v}) &= \min\{\mathsf{K}>0 : \operatorname{supp}(\tilde{\mathbf{T}}_{J,\mathbf{v}}\mathbf{v}_J)\subseteq\Lambda_{\mathsf{K}}\},\end{aligned}$$

with $\mathbf{v}_J$ defined in (4.88), which results in smaller values for n_1 or n_2 in many cases, leading to lower ranks. However, for the analysis of the method, we cannot benefit from this and keep the simpler notation using the same K for both scaling matrices.

In the next step, we provide estimates for the support size, the ranks, and the $\mathcal{A}^s$-norm of the approximation. We reuse the results from Theorems 4.33 and 4.35, where the results are stated in terms of the refinement level J, and reformulate them accordingly with respect to the given tolerance η. Furthermore, for the ranks, we have to take into account the ranks of the truncated scaling matrices.

Lemma 4.43 ([BF24, Lemma 5.15])**.** *Under the assumptions of Theorem 4.40, we have*

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right) \le 2^{1+\frac{2}{s}}\hat{\alpha}(\mathbf{B}_1)(C_\delta C_{\mathrm{t},1})^{\frac{1}{s}}\eta^{-\frac{1}{s}}\left(\sum_{i=1}^{d}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)^{\frac{1}{s}} \tag{4.107}$$

with C_δ from Theorem 4.29 and $C_{\mathrm{t},1}$, $\hat{\alpha}(\mathbf{B}_1)$ from Theorem 4.36, as well as

$$\left\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right\|_{\mathcal{A}^s} \le 2^s\hat{C}_{\mathrm{x}}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s} + 2^s\hat{C}_{\mathrm{t}}\sum_{j=1}^{d}\|\pi^{(\mathrm{t},j)}(\mathbf{v})\|_{\mathcal{A}^s}, \tag{4.108}$$

where

$$\begin{aligned}\hat{C}_{\mathrm{x}} &= 2^s C_\delta\Big(2(d-1)\|\mathbf{C}_2\| + 2^{3s+2}(2^s-1)^{-1}\|\alpha(\mathbf{C}_2)\|_{\ell_1}^s(\|\mathbf{C}_2\| + 2^s\|\beta(\mathbf{C}_2)\|)\Big),\\ \hat{C}_{\mathrm{t}} &= 2^{1+4s}C_\delta c(\mathbf{C}_t)^s(2^s\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\mathbf{C}_{\mathrm{t}}\|).\end{aligned}$$

Moreover, we have the rank estimate

$$\operatorname{rank}_\infty(\mathbf{w}_\eta) \le \|\hat{\mathsf{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty}\|\hat{\mathsf{m}}_{\mathcal{Y}}(\eta;\mathbf{v})\|_{\ell_\infty}(c(\mathbf{C}_t)J(\eta;\mathbf{v})+2)\operatorname{rank}_\infty(\mathbf{v}), \tag{4.109}$$

where

$$\begin{aligned}\hat{\mathsf{m}}_{\mathcal{X},\nu_{\mathrm{t}}}(\eta;\mathbf{v}) &= 1 + n^+_{a_{\nu_{\mathrm{t}}}}(\delta) + \mathsf{m}_{\mathcal{X},\nu_{\mathrm{t}}}(\eta;\mathbf{v}), \quad \nu_{\mathrm{t}}\in\vee_{\mathrm{t}},\\ \hat{\mathsf{m}}_{\mathcal{Y},\nu_{\mathrm{t}}}(\eta;\mathbf{v}) &= 1 + n^+(\delta) + \mathsf{m}_{\mathcal{Y},\nu_{\mathrm{t}}}(\eta;\mathbf{v}), \quad \nu_{\mathrm{t}}\in\vee_{\mathrm{t}}.\end{aligned} \tag{4.110}$$

Proof. We start with the estimate of the support size. The support is not affected by changes in the scaling matrices, since diagonal operators do not change the support. We define the quantity

$$\bar{J}(\eta;\mathbf{v}) = \arg\min\left\{J\in\mathbb{N}_0 : C_\delta C_{\mathrm{t},1}2^{-sJ}\sum_{i=1}^{d}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s} \le \tfrac{\eta}{4}\right\},$$

similarly as $J(\eta;\mathbf{v})$ in (4.102), but using the a priori estimate from Theorem 4.36 instead of the a posteriori estimate. As a direct consequence of

$$e_J(\mathbf{v}) \le C_\delta C_{\mathrm{t},1}2^{-sJ}\sum_{i=1}^{d}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s},$$

we have $J(\eta;\mathbf{v}) \le \bar{J}(\eta;\mathbf{v})$. By the minimality of $\bar{J}(\eta;\mathbf{v})$, we have

$$C_\delta C_{\mathrm{t},1} 2^{-s(\bar{J}(\eta;\mathbf{v})-1)} \sum_{i=1}^{d} \left\| \pi^{(\mathrm{t},i)}(\mathbf{v}) \right\|_{\mathcal{A}^s} \ge \frac{\eta}{4},$$

from which we infer

$$2^{J(\eta;\mathbf{v})} \le 2^{\bar{J}(\eta;\mathbf{v})} \le 2^{1+\frac{2}{s}} (C_\delta C_{\mathrm{t},1})^{\frac{1}{s}} \eta^{-\frac{1}{s}} \left(\sum_{i=1}^{d} \left\| \pi^{(\mathrm{t},i)}(\mathbf{v}) \right\|_{\mathcal{A}^s} \right)^{\frac{1}{s}}. \tag{4.111}$$

Inserting this into (4.99) from Theorem 4.36 yields (4.107).

The factor of rank increase is bounded by the product of the bounds on the hierarchical ranks of each of the factors $\mathbf{D}_{\mathcal{Y},\mathsf{n}_2}$, $\tilde{\mathbf{T}}_{J,\mathbf{v}}$ and $\mathbf{D}_{\mathcal{X},\mathsf{n}_1}$. Hence, the factors $\|\hat{\mathsf{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty}$ and $\|\hat{\mathsf{m}}_{\mathcal{Y}}(\eta;\mathbf{v})\|_{\ell_\infty}$ in (4.109) results from the rank truncated scaling matrices $\mathbf{D}_{\mathcal{X},\mathsf{n}_1}$ and $\mathbf{D}_{\mathcal{Y},\mathsf{n}_2}$. The last factor corresponds to the representation

$$\mathbf{T} = \mathbf{T}_\mathrm{t} \otimes \mathbf{I}_\mathrm{x} + \mathbf{I}_\mathrm{t} \otimes \mathbf{T}_\mathrm{x}.$$

The rank of the spatial operator $\mathbf{T}_\mathrm{x}$ is $R = 2$. For the temporal operator we get by the representation (4.91) and by Theorem 2.27 an increase of $c(\mathbf{C}_t)J(\eta;\mathbf{v})$ since each row and column of $(\mathbf{T}_{\mathrm{t},j+1} - \mathbf{T}_{\mathrm{t},j})$ has at most $c(\mathbf{C}_t)$ nonzero entries per row and column. Hence, in total we get for $\tilde{\mathbf{T}}_{J,\mathbf{v}}$ the factor $c(\mathbf{C}_t)J(\eta;\mathbf{v}) + 2$, which yields (4.109).

In the final step we prove the approximability estimate (4.108). Using (4.63) yields

$$\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_i}(\mathbf{w}_\eta) = \pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_i}(\mathbf{D}_{\mathcal{Y},\mathsf{n}_2} \tilde{\mathbf{T}}_{J,\mathbf{v}} \mathbf{D}_{\mathcal{X},\mathsf{n}_1} \mathbf{v}) \le \pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_i}(\mathbf{D}_{\mathcal{Y}} \tilde{\mathbf{T}}_{J,\mathbf{v}} \mathbf{D}_{\mathcal{X}} \tilde{\mathbf{v}})$$

with $\tilde{\mathbf{v}} = (\mathbf{D}_{\mathcal{X}}^{-1} \mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}$. We look separately at the temporal operator $\mathbf{T}_\mathrm{t} \otimes \mathbf{I}_\mathrm{x}$ and spatial operator $\mathbf{I}_\mathrm{t} \otimes \mathbf{T}_\mathrm{x}$. Then the result is an immediate consequence of the modified triangle inequality in Theorem 2.8 (ii). Without loss of generality, we consider only the case $i = 1$. Based on the symmetry of $\mathbf{T}_\mathrm{x}$, the other cases can be proved in the same manner. We start with the spatial operator $\mathbf{I}_\mathrm{t} \otimes \mathbf{T}_\mathrm{x}$. Due to $\tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}$ being a tensor product with the identity in time and the diagonal structure of the scaling matrices $\mathbf{D}_{\mathcal{X}}$ and $\mathbf{D}_{\mathcal{Y}}$, we obtain

$$\begin{aligned} \pi^{(\mathrm{t},1)}_{\nu_\mathrm{t},\nu_1}(\mathbf{D}_{\mathcal{Y}} \tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}} \mathbf{D}_{\mathcal{X}} \tilde{\mathbf{v}}) &= \pi^{(1)}_{\nu_1}(\mathbf{D} \tilde{\mathbf{T}}_{\mathrm{x},\nu_\mathrm{t},J,\mathbf{v}} \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \tilde{\mathbf{v}}_{\nu_\mathrm{t}}) \\ &\le D_{1,\nu_\mathrm{t},\nu_1} + D_{2,\nu_\mathrm{t},\nu_1}, \end{aligned}$$

with

$$\begin{aligned} D_{1,\nu_\mathrm{t},\nu_1} &= \pi^{(1)}_{\nu_1}\Big(\mathbf{D}\Big(\mathbf{I}_1 \otimes \sum_{\mathsf{n} \in \mathsf{K}_d(1,2,\dots,2)} c_\mathsf{n} \bigotimes_{j=2}^{d} \tilde{\mathbf{T}}^{(j)}_{\nu_\mathrm{t},n_i,\mathbf{v}}\Big) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \tilde{\mathbf{v}}_{\nu_\mathrm{t}}\Big), \\ D_{2,\nu_\mathrm{t},\nu_1} &= \pi^{(1)}_{\nu_1}\Big(\mathbf{D}(\tilde{\mathbf{T}}^{(1)}_{\nu_\mathrm{t},2,\mathbf{v}} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1) \mathbf{D}_{\mathcal{X},\nu_\mathrm{t}} \tilde{\mathbf{v}}_{\nu_\mathrm{t}}\Big), \end{aligned}$$

where we used the structure of the Laplacian $\mathbf{T}_\mathrm{x}$ from Theorem 4.30 and the subadditivity from Theorem 4.9. To estimate $D_{1,\nu_\mathrm{t},\nu_1}$, we introduce the notation

$$\check{\mathbf{D}}_i = \left(\Big(\sum_{j \ne i} \|\psi_{\nu_j}\|^2_{H^1_0(0,1)} \Big)^{-\frac{1}{2}} \delta_{\nu,\nu'} \right)_{\nu,\nu' \in \mathcal{V}_\mathrm{x}}, \quad i = 1,\dots,d,$$

which is the diagonal scaling matrix where we omit the i-th component from the scaling. We observe that

$$\|\mathbf{D}\check{\mathbf{D}}_i^{-1}\| \le \|\mathbf{D}\bar{\mathbf{D}}^{-1}\| \|\bar{\mathbf{D}}\check{\mathbf{D}}_i^{-1}\| \le 1 + \tfrac{\delta}{2},$$

where we used (3.25) and (4.61). Moreover, we have

$$\|\check{\mathbf{D}}_i^{-1}\mathbf{D}_{\mathcal{X},\nu_t}\| \le \|\check{\mathbf{D}}_i^{-1}\bar{\mathbf{D}}\|\|\bar{\mathbf{D}}^{-1}\bar{\mathbf{D}}_{\mathcal{X},\nu_t}\|\|\bar{\mathbf{D}}_{\mathcal{X},\nu_t}^{-1}\mathbf{D}_{\mathcal{X},\nu_t}\| \le \|\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X}}\| \le 1+\tfrac{\delta}{2},$$

where we used (4.57) and that by definition of $\bar{\mathbf{D}}_{\mathcal{X}}$ we have $\|\bar{\mathbf{D}}^{-1}\bar{\mathbf{D}}_{\mathcal{X},\nu_t}\| \le 1$.

Since we are considering the Laplacian, we know that $c_{\mathsf{n}} = 1$ if n is a permutation of $(2,1,\dots,1)$ and $c_{\mathsf{n}} = 0$ otherwise. Hence, using that $\mathbf{D}$ and $\check{\mathbf{D}}_1$ are diagonal matrices, we obtain

$$\begin{aligned}
D_{1,\nu_t,\nu_1} &\le (1+\tfrac{\delta}{2})\pi_{\nu_1}^{(1)}\Big(\check{\mathbf{D}}_1\Big(\mathbf{I}_1 \otimes \sum_{\mathsf{n}\in \mathsf{K}_d(1,2,\dots,2)} c_{\mathsf{n}} \bigotimes_{i=2}^{d} \tilde{\mathbf{T}}^{(i)}_{\nu_t,n_i,\mathbf{v}}\Big)\check{\mathbf{D}}_1(\check{\mathbf{D}}_1^{-1}\mathbf{D}_{\mathcal{X},\nu_t})\tilde{\mathbf{v}}_{\nu_t}\Big)\\
&\le (1+\tfrac{\delta}{2}) \sum_{\mathsf{n}\in \mathsf{K}_d(1,2,\dots,2)} |c_{\mathsf{n}}| \Big\|\check{\mathbf{D}}_1\Big(\mathbf{I}_1 \otimes \bigotimes_{i=2}^{d} \tilde{\mathbf{T}}^{(i)}_{\nu_t,n_i,\mathbf{v}}\Big)\check{\mathbf{D}}_1\Big\| \pi_{\nu_1}^{(1)}(\check{\mathbf{D}}_1^{-1}\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{v}}_{\nu_t})\\
&\le (1+\tfrac{\delta}{2}) \sum_{i=2}^{d} \|\bar{\mathbf{D}}_1\tilde{\mathbf{T}}^{(i)}_{\nu_t,2,\mathbf{v}}\bar{\mathbf{D}}_1\| \pi_{\nu_1}^{(1)}((\mathbf{I}_1 \otimes \check{\mathbf{D}}_1^{-1})\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{v}}_{\nu_t})\\
&\le 2(d-1)\|\mathbf{C}_2\|(1+\tfrac{\delta}{2})^2\pi_{\nu_t,\nu_1}^{(\mathrm{t},1)}(\tilde{\mathbf{v}})\\
&\le C_\delta 2(d-1)\|\mathbf{C}_2\|\pi_{\nu_t,\nu_1}^{(\mathrm{t},1)}(\mathbf{v}),
\end{aligned}$$

where we used a slightly modified version of (4.69) in the third line with $\bar{\mathbf{D}}$ replaced by $\check{\mathbf{D}}_1$, that $\tilde{\mathbf{T}}^{(i)}_{\nu_t,2,\mathbf{v}}$ is an approximation of $\mathbf{T}_2$ in the fourth line as well as (4.56) in the last line.

Next, we examine D_{2,ν_t,ν_1}. Using $\|\mathbf{D}\bar{\mathbf{D}}_1^{-1}\| \le 1+\frac{\delta}{2}$, we have

$$D_{2,\nu_t,\nu_1} \le (1+\tfrac{\delta}{2})\pi_{\nu_1}^{(1)}\Big(\Big((\bar{\mathbf{D}}_1\tilde{\mathbf{T}}^{(1)}_{\nu_t,2,\mathbf{v}}\bar{\mathbf{D}}_1)\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1\Big)(\bar{\mathbf{D}}_1^{-1}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1)\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{v}}_{\nu_t}\Big).$$

We set

$$\hat{\mathbf{v}}_{\nu_t} = (\bar{\mathbf{D}}_1^{-1}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1)\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{v}}_{\nu_t}, \quad \nu_t \in \vee_t$$

and

$$\begin{aligned}
\tilde{\mathbf{C}}^{(1)}_{\nu_t,2,\mathbf{v}} &= \bar{\mathbf{D}}_1\tilde{\mathbf{T}}^{(1)}_{\nu_t,2}\bar{\mathbf{D}}_1,\\
\mathbf{C}_{2,j} &= \bar{\mathbf{D}}_1\mathbf{T}_{2,j}\bar{\mathbf{D}}_1, \quad j\in\mathbb{N}_0.
\end{aligned}$$

Furthermore, we define the disjoint sets

$$\begin{aligned}
\hat{\Lambda}_{\nu_t,[0]} &= \operatorname{supp}\operatorname{range}\mathbf{C}_{2,0}\mathrm{R}_{\Lambda^{(1)}_{\nu_t,[0]}},\\
\hat{\Lambda}_{\nu_t,[q]} &= \Big(\bigcup_{j+l=q}\operatorname{supp}\operatorname{range}\mathbf{C}_{2,j}\mathrm{R}_{\Lambda^{(1)}_{\nu_t,[l]}}\Big)\setminus\Big(\bigcup_{i<q}\hat{\Lambda}_{\nu_t,[i]}\Big), \quad q>0,
\end{aligned}$$

where $\Lambda^{(1)}_{\nu_t,[j]} = \Lambda^{(1)}_{\nu_t,[j]}(\mathbf{v})$ is defined in (4.75). Moreover, for notational brevity, we set $\Lambda^{(\mathrm{t},1)}_{[j]} = \Lambda^{(\mathrm{t},1)}_{[j]}(\mathbf{v})$. Naturally, for each $k\in\mathbb{N}_0$ we define the union over all time indices by

$$\hat{\Lambda}_{[k]} = \bigcup_{\nu_t\in\vee_t}\{\nu_t\}\times\hat{\Lambda}_{\nu_t,[k]}.$$

Thus, for

$$\hat{\mathbf{w}} = \sum_{\nu_t\in\vee_t}\mathbf{e}_{\nu_t}\otimes\Big(\tilde{\mathbf{C}}^{(1)}_{\nu_t,2}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1\Big)\hat{\mathbf{v}}_{\nu_t}$$

we have

$$\left\|\mathrm{R}_{\hat{\Lambda}_{[q]}}\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\right\|^2 = \sum_{\nu_\mathrm{t}\in\mathcal{V}_\mathrm{t}}\left\|\mathrm{R}_{\hat{\Lambda}_{\nu_\mathrm{t},[q]}}\pi^{(1)}((\tilde{\mathbf{C}}^{(1)}_{2,\nu_\mathrm{t}}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1)\hat{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2. \tag{4.112}$$

For each term on the right-hand side can be proceeded as in the proof for the elliptic case in [BD15, Theorem 8] to obtain

$$\begin{aligned}\left\|\mathrm{R}_{\hat{\Lambda}_{\nu_\mathrm{t},[q]}}\pi^{(1)}((\tilde{\mathbf{C}}^{(1)}_{2,\nu_\mathrm{t}}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1)\hat{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2 \le 3\tilde{C}_1\Bigg(&\sum_{l=0}^{q-1}\gamma_{q,l}2^{-2s(q-l-1)}\left\|\mathrm{R}_{\Lambda^{(1)}_{\nu_\mathrm{t},[l]}}\pi^{(1)}(\hat{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2\\ &+\sum_{l=q}^{J}\beta_{J-l}(\mathbf{C}_2)2^{-2s(J-l)}\left\|\mathrm{R}_{\Lambda^{(1)}_{\nu_\mathrm{t},[l]}}\pi^{(1)}(\hat{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2\\ &+\|\mathbf{C}_2\|\left\|\mathrm{R}_{\bigcup_{j\ge q}\Lambda^{(1)}_{\nu_\mathrm{t},[j]}}\pi^{(1)}(\hat{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2\Bigg),\end{aligned} \tag{4.113}$$

where $\gamma_{q,l} = \beta_{J-l}(\mathbf{C}_2) + \beta_{q-l-1}(\mathbf{C}_2)$ as well as $\tilde{C}_1 = \|\mathbf{C}_2\| + 2^s\|\beta(\mathbf{C}_2)\|$. Additionally, we have

$$\|(\bar{\mathbf{D}}_1^{-1}\otimes\mathbf{I}_1\otimes\cdots\otimes\mathbf{I}_1)\mathbf{D}_{\mathcal{X},\nu_\mathrm{t}}\| \le 1+\tfrac{\delta}{2}$$

and thus

$$\pi^{(1)}_{\nu_1}(\hat{\mathbf{v}}_{\nu_\mathrm{t}}) \le (1+\tfrac{\delta}{2})\pi^{(1)}_{\nu_1}(\tilde{\mathbf{v}}_{\nu_\mathrm{t}}).$$

Moreover, using the definition of $\tilde{\mathbf{v}} = (\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathrm{n}_1})\mathbf{v}$ as well as (4.56), we obtain

$$\sum_{\nu_\mathrm{t}\in\mathcal{V}_\mathrm{t}}\left\|\mathrm{R}_{\Lambda^{(1)}_{\nu_\mathrm{t},[q]}}\pi^{(1)}(\tilde{\mathbf{v}}_{\nu_\mathrm{t}})\right\|^2 = \left\|\mathrm{R}_{\Lambda^{(\mathrm{t},1)}_{[q]}}\pi^{(\mathrm{t},1)}(\tilde{\mathbf{v}})\right\|^2 \le \left\|\mathrm{R}_{\Lambda^{(\mathrm{t},1)}_{[q]}}\pi^{(\mathrm{t},1)}(\mathbf{v})\right\|^2.$$

Combining this with (4.112) and (4.113), we arrive at

$$\begin{aligned}\left\|\mathrm{R}_{\hat{\Lambda}_{[q]}}\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\right\|^2 \le 3(1+\tfrac{\delta}{2})^2\tilde{C}_1\Bigg(&\sum_{l=0}^{q-1}\gamma_{q,l}2^{-2s(q-l-1)}\left\|\mathrm{R}_{\Lambda^{(\mathrm{t},1)}_{[l]}}\pi^{(\mathrm{t},1)}(\mathbf{v})\right\|^2\\ &+\sum_{l=q}^{J}\beta_{J-l}(\mathbf{C}_2)2^{-2s(J-l)}\left\|\mathrm{R}_{\Lambda^{(\mathrm{t},1)}_{[l]}}\pi^{(\mathrm{t},1)}(\mathbf{v})\right\|^2\\ &+\|\mathbf{C}_2\|\left\|\mathrm{R}_{\bigcup_{j\ge q}\Lambda^{(\mathrm{t},1)}_{[j]}}\pi^{(\mathrm{t},1)}(\mathbf{v})\right\|^2\Bigg).\end{aligned}$$

Then Theorem 4.31 yields

$$\|\mathrm{R}_{\hat{\Lambda}_{[q]}}\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\| \le 2^{-sq}2^{2(s+1)}(1+\tfrac{\delta}{2})\tilde{C}_1\|\pi^{(\mathrm{t},1)}(\mathbf{v})\|_{\mathcal{A}^s}$$

for $q = 0,\dots,J$. By construction, the index sets $\hat{\Lambda}_{[q]}$ are disjoint and by the same argument as in (4.86), we observe that $\#\hat{\Lambda}_{[q]} \le \|\alpha(\mathbf{C}_2)\|_{\ell_1}2^q$. By the definition of the $\mathcal{A}^s$-norm, we have

$$\begin{aligned}\|\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\|_{\mathcal{A}^s} &\le \sup_{q\in\mathbb{N}_0}\Big(1+\sum_{j\le q}|\hat{\Lambda}_{[j]}|\Big)^s\|\mathrm{R}_{\bigcup_{j\ge q}\hat{\Lambda}_{[j]}}\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\|\\ &\le \sup_{q\in\mathbb{N}_0}\Big(1+\sum_{j\le q}|\hat{\Lambda}_{[j]}|\Big)^s\sum_{j\ge q}\|\mathrm{R}_{\hat{\Lambda}_{[j]}}\pi^{(\mathrm{t},1)}(\hat{\mathbf{w}})\|\\ &\le \frac{2^{3s+2}}{2^s-1}(1+\tfrac{\delta}{2})\|\alpha(\mathbf{C}_2)\|^s_{\ell_1}\tilde{C}_1\|\pi^{(\mathrm{t},1)}(\tilde{\mathbf{v}})\|_{\mathcal{A}^s}.\end{aligned}$$

Combining the two estimates for the spatial operator with Theorem 2.8 and (4.56) results in

$$\|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x},J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{v}}\|_{\mathcal{A}^s} \leq 2^s\hat{C}_{\mathrm{x}}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}.$$

In the final step, we consider the temporal operator $\tilde{\mathbf{B}}_{\mathrm{t},J}$ with $J = J(\eta;\mathbf{v})$. Here, we use restrictions of the index sets in all variables simultaneously, and thus we need to proceed differently from the case of the spatial operator. The basic idea is to estimate the $\mathcal{A}^s$-norm of $\pi^{(\mathrm{t},i)}(\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t},J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{v}})$ for a fixed dimension $i \in \{1,\dots,d\}$ by the sum of the $\mathcal{A}^s$-norm of the contractions of $\mathbf{v}$ over all dimensions. The idea of the proof is to work directly with the definition of the $\mathcal{A}^s$-norm from Theorem 2.7 and construct suitable sets to estimate the norm. To this end, for $q = 0,\dots,J$ we define

$$\begin{aligned}\tilde{\mathbf{w}}_q &= \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t},q,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{v}},\\ \hat{\Lambda}_q^{(i)} &= \operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_q)\right).\end{aligned}$$

By (4.96) from Theorem 4.35 we have

$$\#\hat{\Lambda}_q^{(i)} \leq 2c2^q.$$

For $N > \#\hat{\Lambda}_J^{(i)}$ we have

$$\inf_{\#\Lambda=N}\|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J) - \mathrm{R}_\Lambda\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\| \leq \|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J) - \mathrm{R}_{\hat{\Lambda}_J^{(i)}}\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\| = 0.$$

Now let $0 < N \leq \#\hat{\Lambda}_J^{(i)}$. Then there exists a $q < J$ such that $\#\hat{\Lambda}_q^{(i)} < N \leq \#\hat{\Lambda}_{q+1}^{(i)}$. We observe that

$$\begin{aligned}\left\|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J) - \mathrm{R}_{\hat{\Lambda}_q^{(i)}}\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\right\|^2 &= \sum_{(\nu_\mathrm{t},\nu_i)\notin\hat{\Lambda}_q^{(i)}}\left(\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_i}(\tilde{\mathbf{w}}_J)\right)^2\\ &= \sum_{(\nu_\mathrm{t},\nu_i)\notin\hat{\Lambda}_q^{(i)}}\left(\pi^{(\mathrm{t},i)}_{\nu_\mathrm{t},\nu_i}(\tilde{\mathbf{w}}_J - \tilde{\mathbf{w}}_q)\right)^2\\ &\leq \|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J - \tilde{\mathbf{w}}_q)\|^2\\ &= \|\tilde{\mathbf{w}}_J - \tilde{\mathbf{w}}_q\|^2,\end{aligned}$$

where we have used the definition of $\hat{\Lambda}_q^{(i)}$ in the second line. From this we infer

$$\begin{aligned}&(N+1)^s\inf_{\#\Lambda=N}\|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J) - \mathrm{R}_\Lambda\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\|\\ &\quad\leq (2c(\mathbf{C}_t)2^{q+1}+1)^s\|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J) - \mathrm{R}_{\hat{\Lambda}_q^{(i)}}\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\|\\ &\quad\leq (2c(\mathbf{C}_t)2^{q+1}+1)^s\|\tilde{\mathbf{w}}_J - \tilde{\mathbf{w}}_q\|\\ &\quad\leq 2^{3s}c(\mathbf{C}_t)^s2^{sq}\left(\|\mathbf{B}_\mathrm{t}\tilde{\mathbf{v}} - \tilde{\mathbf{w}}_J\| + \|\mathbf{B}_\mathrm{t}\tilde{\mathbf{v}} - \tilde{\mathbf{w}}_q\|\right)\\ &\quad\leq 2^{3s}c(\mathbf{C}_t)^s2^{sq}(e_{\mathrm{t},J,\mathbf{v}}(\tilde{\mathbf{v}}) + e_{\mathrm{t},q,\mathbf{v}}(\tilde{\mathbf{v}}))\end{aligned}$$

with $e_{\mathrm{t},j,\mathbf{v}}(\tilde{\mathbf{v}})$ defined in (4.94). Using (4.56) and Theorem 4.31 yields

$$\|\mathrm{R}_{\Lambda^{(\mathrm{t},i)}_{[p]}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\tilde{\mathbf{v}})\| \leq \|\mathrm{R}_{\Lambda^{(\mathrm{t},i)}_{[p]}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{v})\| \leq 2^{-s(p-1)}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s},$$

and thus

$$e_{\mathrm{t},p,\mathbf{v}}(\tilde{\mathbf{v}}) \leq C_\delta 2^{-s(p-1)}(2^s\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\mathbf{C}_\mathrm{t}\|)\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}$$

for all $p = 0,\dots,J$. Since N was chosen arbitrary, by definition of the $\mathcal{A}^s$-norm we arrive at

$$\|\pi^{(\mathrm{t},i)}(\tilde{\mathbf{w}}_J)\|_{\mathcal{A}^s} \leq 2C_\delta c(\mathbf{C}_t)^s 2^{4s}(2^s\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\mathbf{C}_\mathrm{t}\|)\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}.$$

Using Theorem 2.8 (ii) for the spatial and temporal part yields (4.108). □

Remark 4.44. The estimate for the $\mathcal{A}^s$-norm in (4.108) in Theorem 4.43 depends on the number of spatial dimensions d in two different ways. For the spatial part, the dependence is direct due to the factor $\hat{C}_\mathrm{x}$, which is linear in d. The temporal part, however, depends on d only indirectly via the sum over the contractions of all dimensions. In most cases, we are not particularly interested in the approximability of the contractions of $\mathbf{w}_\eta$ in a specific dimension, but more in the sum over all dimensions. Under the assumptions of Theorem 4.43, we have

$$\begin{aligned}\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\|_{\mathcal{A}^s} &\leq 2^s\hat{C}_\mathrm{x}\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s} + d2^s\hat{C}_\mathrm{t}\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}\\ &\leq C(\delta,s,\mathbf{B}_1)d\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}.\end{aligned}$$

Hence, both spatial and temporal parts depend linearly on the number of spatial dimensions d.

Remark 4.45. The bounds in Theorem 4.43 are similar to the bounds obtained in [BD15] for the elliptic case, but for rank estimate (4.109) we have an additional factor $J(\eta;\mathbf{v})$. However, as we will show in Section 4.4, this factor does not influence the principal behaviour of the computational complexity of the method, which means we can still circumvent the curse of dimensionality under natural assumptions.

In the remainder of this section, we analyze the computational complexity of the adaptive approximate application of the operator $\mathbf{B}_1$ and the behaviour of the quantities $\|\hat{\mathrm{m}}_\mathcal{X}(\mathbf{v};\eta)\|_{\ell_\infty}$ and $\|\hat{\mathrm{m}}_\mathcal{Y}(\mathbf{v};\eta)\|_{\ell_\infty}$ arising in the rank estimates (4.109).

Lemma 4.46 ([BF24, Lemma 5.17])**.** *Under the assumptions of Theorem 4.40, the number* $\mathrm{flops}(\mathbf{w}_\eta)$ *of floating point operations to compute* $\mathbf{w}_\eta$ *is bounded by*

$$\begin{aligned}\mathrm{flops}(\mathbf{w}_\eta) \lesssim{}& (8 + c(\mathbf{C}_t)^4(J(\eta;\mathbf{v}))^4)\hat{m}^3(\eta;\mathbf{v})\,\mathrm{rank}_\infty^3(\mathbf{v})\sum_{i=1}^{d}\#\mathrm{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{v})\right)\\ &+ d\eta^{-\frac{1}{s}}\,\mathrm{rank}_\infty(\mathbf{v})J(\eta;\mathbf{v})\hat{m}(\eta;\mathbf{v})c(\mathbf{C}_t)\hat{\alpha}(\mathbf{B}_1)C_{\mathrm{t},1}^{\frac{1}{s}}\left(\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}\right)^{\frac{1}{s}},\end{aligned} \tag{4.114}$$

where $\hat{m}(\eta;\mathbf{v}) = \|\hat{\mathrm{m}}_\mathcal{X}(\eta;\mathbf{v})\|_{\ell_\infty}\|\hat{\mathrm{m}}_\mathcal{Y}(\eta;\mathbf{v})\|_{\ell_\infty}$ *with* $\hat{\mathrm{m}}_\mathcal{X}$ *and* $\hat{\mathrm{m}}_\mathcal{Y}$ *defined in* (4.110)*, and where the hidden constant is independent of* $\eta, \mathbf{B}_1, \mathbf{v}$ *and* d*.*

Proof. Since the adaptive application is the sum of the spatial and temporal parts, we can estimate the number of operations for each part separately. To this end, let $\mathbf{w}_\eta = \mathbf{w}_{\eta,\mathrm{x}} + \mathbf{w}_{\eta,\mathrm{t}}$ with $\mathbf{w}_{\eta,\mathrm{x}}$ and $\mathbf{w}_{\eta,\mathrm{t}}$ being the output of the spatial and temporal

operator, respectively. We start with the spatial operator $\mathbf{B}_\mathrm{x}$, where we can proceed similarly to the elliptic case [BD15; BD16] due to the identity in time.

The output $\mathbf{w}_\eta$ is a sum of hierarchical tensors arising from the application of operators in hierarchical tensor format. As described in Section 2.4.5, the new transfer tensors are given by concatenating the transfer tensor of each term, which again are Kronecker products of the transfer tensors of the operator and of the input tensor. Hence, we can calculate each entry of the new transfer tensors in $\mathcal{O}(1)$ operations. Having in mind that we have $d-1$ transfer tensors per hierarchical tensor, the number of operations to calculate all transfer tensors of $\mathbf{w}_{\eta,\mathrm{x}}$ is bounded by

$$\begin{aligned}(d-1)\#\operatorname{supp}\big(\pi^{(\mathrm{t})}(\mathbf{w}_{\eta,\mathrm{x}})\big)\operatorname{rank}_\infty^3(\mathbf{w}_{\eta,\mathrm{x}}) &\leq 8\hat{m}^3(\eta;\mathbf{v})\operatorname{rank}_\infty^3(\mathbf{v})d\#\operatorname{supp}\big(\pi^{(\mathrm{t})}(\mathbf{v})\big)\\ &\leq 8\hat{m}^3(\eta;\mathbf{v})\operatorname{rank}_\infty^3(\mathbf{v})\sum_{i=1}^d\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{v})\big),\end{aligned}$$

where we used that the rank of the Laplace operator $R = 2$ and the structure of $\mathbf{B}_\mathrm{x}$ with the identity in time. The entries in $\mathbf{T}_2$, and thus in $\mathbf{T}_{2,j}$, can be calculated in $\mathcal{O}(1)$ operations, since we can integrate piecewise polynomials of fixed degree exactly in a constant number of operations using quadrature. Hence, we can calculate each entry in the mode frames in constant time as well. From this, we infer that the number of operations to calculate the mode frames is bounded by a multiple of the total number of entries in the mode frames, which can be estimated by

$$\begin{aligned}&\operatorname{rank}_\infty(\mathbf{w}_{\eta,\mathrm{x}})\sum_{i=1}^d\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{w}_{\eta,\mathrm{x}})\big)\\ &\qquad\lesssim d\eta^{-\frac{1}{s}}\operatorname{rank}_\infty(\mathbf{v})\hat{m}(\mathbf{v},\eta)\hat{\alpha}(\mathbf{B}_1)(4C_\delta C_{\mathrm{t},1})^{\frac{1}{s}}\Big(\sum_{i=1}^d\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\end{aligned}$$

where we used (4.107) and (4.109) from Theorem 4.43.

Next, we examine the adaptive application of the temporal operator. By definition of $\mathbf{v}_i$ in (4.88), we need $d2^i\operatorname{rank}_\infty(\mathbf{v})$ operations to calculate the mode frames of $\mathbf{v}_i$ and at most $\#\operatorname{supp}(\pi^{(\mathrm{t})}(\mathbf{v}))d\operatorname{rank}_\infty^3(\mathbf{v})$ operations to calculate the transfer tensors.

As before, we divide the analysis of the costs of the application step into the calculation of the mode frames and of the transfer tensors. Due to the structure of $\mathbf{B}_\mathrm{t}$ with the identity in the spatial variables, the corresponding spatial components in low-rank format are rescaled, but remain otherwise unchanged. The number of operations required to compute the mode frames of $\mathbf{w}_{\eta,\mathrm{t}}$ can be estimated by the number of entries in the mode frames of $\mathbf{w}_{\eta,\mathrm{t}}$, because each entry results only from multiplication with the corresponding diagonal entries of the scaling matrices which has constant costs. Therefore, the number of operations for the mode frames is bounded by a fixed multiple

$$\begin{aligned}&\operatorname{rank}_\infty(\mathbf{w}_{\eta,\mathrm{t}})\sum_{i=1}^d\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{w}_{\eta,\mathrm{t}})\big)\\ &\qquad\lesssim d\eta^{-\frac{1}{s}}\operatorname{rank}_\infty(\mathbf{v})\hat{m}(\mathbf{v},\eta)J(\eta;\mathbf{v})c(\mathbf{C}_t)\hat{\alpha}(\mathbf{B}_1)(4C_\delta C_{\mathrm{t},1})^{\frac{1}{s}}\Big(\sum_{i=1}^d\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\end{aligned}$$

where we again used the results from Theorem 4.43. In the final step, we consider the transfer tensors of $\mathbf{w}_{\eta,\mathrm{t}}$. Using the super-compressibility of $\mathbf{T}_\mathrm{t}$ from Theorem 2.27

we obtain

$$\#\operatorname{supp}(\pi^{(\mathrm{t})}(\mathbf{w}_{\eta,\mathrm{t}})) \le cJ(\eta;\mathbf{v})\#\operatorname{supp}(\pi^{(\mathrm{t})}(\mathbf{v})).$$

The application of the operator is a multiplication of the low-rank representations by a certain value, which is done by the application of the scaling matrices as well. The result is then stored at the corresponding position, thereby increasing the rank. Additionally, we know by the same arguments as in Theorem 4.43 that the maximum rank grows at most by the factor $c(\mathbf{C}_t)J(\eta;\mathbf{v})\hat{m}(\eta;\eta)$. Again, using that we have $d-1$ transfer tensors per active temporal index, the number of operations to calculate the transfer tensors of $\mathbf{w}_{\eta,\mathrm{t}}$ is bounded by a fixed multiple of

$$\begin{aligned}
&\#\operatorname{supp}\big(\pi^{(\mathrm{t})}(\mathbf{w}_{\eta,\mathrm{t}})\big)\operatorname{rank}_\infty^3(\mathbf{w}_{\eta,\mathrm{t}})(d-1)\\
&\quad\le c(\mathbf{C}_t)J(\eta;\mathbf{v})\#\operatorname{supp}\big(\pi^{(\mathrm{t})}(\mathbf{v})\big)(d-1)c(\mathbf{C}_t)^3(J(\eta;\mathbf{v}))^3\hat{m}^3(\eta;\mathbf{v})\operatorname{rank}_\infty^3(\mathbf{v})\\
&\quad\le c(\mathbf{C}_t)^4(J(\eta;\mathbf{v}))^4\hat{m}^3(\eta;\mathbf{v})\operatorname{rank}_\infty^3(\mathbf{v})\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{v})\big).
\end{aligned}$$

Combining all these estimates yields (4.114). □

Next, we investigate the quantities $\|\hat{\mathrm{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty}$ and $\|\hat{\mathrm{m}}_{\mathcal{Y}}(\eta;\mathbf{v})\|_{\ell_\infty}$ to get a better understanding of the bound on the ranks of the adaptive application in (4.109). The goal is to bound these quantities in terms of the error tolerance η and of the maximum active one-dimensional wavelet level of the input $\mathbf{v}$. To estimate the latter, we need additional regularity and make use of the level decay property defined in Theorem 2.28.

In what follows, for finitely supported $\mathbf{v} \in \ell_2(\vee)$ with corresponding spatial components $(\mathbf{v}_{\nu_\mathrm{t}})_{\nu_\mathrm{t}\in\vee_\mathrm{t}}$ we denote by

$$L_{\mathrm{x}}(\mathbf{v}) = \max\{|\nu_i| : \nu = (\nu_1,\dots,\nu_d) \in \operatorname{supp}(\mathbf{v}_{\nu_\mathrm{t}}), \nu_\mathrm{t} \in \vee_\mathrm{t}, i = 1,\dots,d\}$$

the maximal active one-dimensional spatial wavelet level and by

$$L_{\mathrm{t}}(\mathbf{v}) = \max\{|\nu_\mathrm{t}| : \nu_\mathrm{t} \in \vee_\mathrm{t}, \mathbf{v}_{\nu_\mathrm{t}} \neq 0\}$$

the maximal active temporal wavelet level. Let

$$L(\mathbf{v}) = \max\{L_{\mathrm{x}}(\mathbf{v}), L_{\mathrm{t}}(\mathbf{v})\} \tag{4.115}$$

be the overall maximal one-dimensional active wavelet level of $\mathbf{v}$. In what follows, we follow the arguments of [BD16, Section 6.2] and [BF24, Section 5.3]. By the definition of the quantities (4.110) and in view of (4.58) and (4.59), we need to find an upper bound for $\ln(\|\mathsf{K}\|_{\ell_\infty})$ with $\mathsf{K} = \mathsf{K}(J(\eta;\mathbf{v});\mathbf{v})$ defined in (4.103).

Lemma 4.47. *Let $\mathbf{v}_{\mathrm{x}} \in \ell_2(\vee_{\mathrm{x}})$ and $K > 0$ with*

$$L_{\mathrm{x}}(\mathbf{v}_{\mathrm{x}}) \le \frac{1}{2}\log_2(K) - \log_2(cs_{\min}^{(1)}), \tag{4.116}$$

where

$$s_{\min}^{(1)} = \Big(\min_{\nu_1\in\vee_1}\|\psi_{\nu_1}\|_{H_0^1}\Big)^{-1} \tag{4.117}$$

and $c = \max_{\nu_1\in\vee_1} 2^{-|\nu_1|}\|\psi_{\nu_1}\|_{H_0^1}$. Then we have $\operatorname{supp}(\mathbf{v}_{\mathrm{x}}) \subset \Lambda_K$ with Λ_K defined in (4.51).

Proof. By definition of the $H_0^1(\Omega)$-norm and the $L_2(0,1)$-orthonormality of the one-dimensional wavelets, for $\nu \in \vee_{\mathrm{x}}$ we have

$$s_{\min}\|\Psi_\nu\|_{H_0^1(\Omega)} \leq s_{\min}\sqrt{d}\max_{i=1,\dots,d}\|\psi_{\nu_i}\|_{H_0^1(0,1)} \leq s_{\min}\sqrt{d}c2^{\max_i|\nu_i|} \leq s_{\min}^{(1)}c2^{\max_i|\nu_i|},$$

where we used $s_{\min}^{-1} = (s_{\min}^{(1)})^{-1}\sqrt{d}$ with $s_{\min}$ from (4.48). By choosing K to satisfy (4.116), we have for all $\nu \in \operatorname{supp}(\mathbf{v}_{\mathrm{x}})$ that

$$s_{\min}^2\|\Psi_\nu\|_{H_0^1(\Omega)}^2 \leq (s_{\min}^{(1)})^2c^2 2^{2\max_i|\nu_i|} \leq K. \qquad \square$$

Lemma 4.48. *Let the operators $\mathbf{T}_2$ and $\mathbf{T}_{\mathrm{t}}$ have the level decay property from Theorem 2.28. Then under the assumptions of Theorem 4.40, we have the estimates*

$$\|\hat{\mathsf{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty} \leq \tilde{C}_1(\delta,s,\mathbf{B}_1)\left[1 + L(\mathbf{v}) + |\ln(\eta)| + \ln\left(\sum_{i=1}^d \left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right]^2, \tag{4.118}$$

$$\|\hat{\mathsf{m}}_{\mathcal{Y}}(\eta;\mathbf{v})\|_{\ell_\infty} \leq \tilde{C}_2(\delta,s,\mathbf{B}_1)\left[1 + L_{\mathrm{x}}(\mathbf{v}) + |\ln\eta| + \ln\left(\sum_{i=1}^d \left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right], \tag{4.119}$$

with constants $\tilde{C}_1(\delta,s,\mathbf{B}_1), \tilde{C}_2(\delta,s,\mathbf{B}_1) > 0$ independent of η and $\mathbf{v}$.

Proof. By the minimality property of K in (4.103) and by Theorem 4.47, for all $\nu_{\mathrm{t}} \in \vee_{\mathrm{t}}$ we have

$$\begin{aligned}\tfrac{1}{2}\log_2(\mathsf{K}_{\nu_{\mathrm{t}}}) &\leq L_{\mathrm{x}}((\tilde{\mathbf{T}}_{J(\eta;\mathbf{v}),\mathbf{v}}\tilde{\mathbf{v}})_{\nu_{\mathrm{t}}}) + \log_2(cs_{\min}^{(1)}) \\ &\leq L_{\mathrm{x}}(\tilde{\mathbf{T}}_{J(\eta;\mathbf{v}),\mathbf{v}}\tilde{\mathbf{v}}) + \log_2(cs_{\min}^{(1)}) \\ &\leq L_{\mathrm{x}}(\tilde{\mathbf{v}}) + C_1(\mathbf{B}_1,s)J(\eta;\mathbf{v}) + \log_2(cs_{\min}^{(1)}),\end{aligned}$$

where we used the level decay property from Theorem 2.28 in the last step. The application of the scaling matrix does not influence the support because of the diagonal structure. Hence, the scaling matrices do not affect the maximum active spatial wavelet level either.

By the choice of η in Theorem 4.40, we have $c(\eta;\mathbf{v}) \leq 1$. Hence,

$$\begin{aligned}\left|\ln\left(\min\left\{\tfrac{\delta}{2}, c(\eta;\mathbf{v})\right\}\right)\right| &= -\ln\left(\min\left\{\tfrac{\delta}{2}, c(\eta;\mathbf{v})\right\}\right) \\ &= \ln\left(\max\left\{\tfrac{2}{\delta}, \tfrac{4\|\mathbf{B}_1\|}{\eta}(1-\tfrac{\delta}{2})^{-1}\|\mathbf{v}\|\right\}\right) \\ &\leq \ln\left(\tfrac{2}{\delta}\tfrac{4\|\mathbf{B}_1\|}{\eta}(1-\tfrac{\delta}{2})^{-1}\max\{1,\|\mathbf{v}\|\}\right) \\ &\leq C_2(\mathbf{B}_1,\delta) + |\ln\eta| + \max\{0,\ln(\|\mathbf{v}\|)\}\end{aligned} \tag{4.120}$$

and $\|\mathbf{v}\| = \left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\| \leq \left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}$ for all $i = 1,\dots,d$. Furthermore, by (4.111), we have

$$J(\eta;\mathbf{v}) \leq \frac{1}{s}\left(|\ln\eta| + \ln\left(C_3(\mathbf{B}_1,s,\delta)\sum_{i=1}^d\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right). \tag{4.121}$$

Combining all of this yields (4.119).

Now we consider $\mathsf{m}_{\mathcal{X}}$, which depends not only on the spatial supports of $\mathbf{v}$ at the respective temporal basis indices, but also on the norm of each temporal basis index. Based on Theorem 4.24, for $a_{\nu_{\mathrm{t}}}$ defined in (4.49), we have $\ln(a_{\nu_{\mathrm{t}}}) \eqsim 1 + |\nu_{\mathrm{t}}|$ as well as $h_{\nu_{\mathrm{t}}}^{-1} \eqsim (1+|\nu_{\mathrm{t}}|)$. Now let ν_{t} be an active temporal index of $\mathbf{v}$. Based on Theorem 4.23,

we do a case separation and proceed as before. If $K_{\nu_t} > a_{\nu_t}^2$, we have

$$m_{\mathcal{X},\nu_t}(\eta;\mathbf{v}) \le C_4(\delta,s,\mathbf{B}_1)(1+|\nu_t|)\left[1+L_x(\mathbf{v})+|\ln\eta|+\ln\left(\sum_{i=1}^{d}\left\|\pi^{(t,i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right],$$

where we used $|\nu_t| \ge 0$ for all $\nu_t \in \vee_t$. In the case $K_{\nu_t} \le a_{\nu_t}^2$, we have

$$m_{\mathcal{X},\nu_t}(\eta;\mathbf{v}) \le C_5(\delta,s,\mathbf{B}_1)(1+|\nu_t|)\left[1+|\nu_t|+|\ln\eta|+\ln\left(\sum_{i=1}^{d}\left\|\pi^{(t,i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right].$$

Based on the level decay property, we can bound the maximal temporal level by

$$L_t(\tilde{\mathbf{T}}_{J(\eta;\mathbf{v}),\mathbf{v}}\tilde{\mathbf{v}}) \le L_t(\mathbf{v}) + C_1(\mathbf{B}_1,s)J(\eta;\mathbf{v}).$$

Due to the diagonal structure of the scaling matrices, $m_{\mathcal{X},\nu_t}$ can be chosen as zero for $|\nu_t| > L_t(\tilde{\mathbf{T}}_{J(\eta;\mathbf{v})}\mathbf{v})$. As in the other case, we use the level decay property, which results in

$$\|m_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty} \le C_6(\delta,s,\mathbf{B}_1)\left[1+L(\mathbf{v})+|\ln(\eta)|+\ln\left(\sum_{i=1}^{d}\left\|\pi^{(t,i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right]^2.$$

In contrast to the scaling matrix $\mathbf{D}_\mathcal{Y}$, the upper bound $n^+_{a_{\nu_t}}(\delta)$ depends on the temporal wavelet as well. By the definition of $n^+_{a_{\nu_t}}$ in (4.41) and Theorem 4.24, we obtain that $n^+_{a_{\nu_t}}$ grows quadratically with respect to the maximal temporal wavelet level. From this, we infer (4.118). □

The estimate (4.119) is similar to the one obtained in [BD16] for elliptic problems, which is quite natural since the scaling matrix does not depend directly on the temporal basis function.

Remark 4.49. Based on the ordering of the scaling matrix in $\mathbf{B}_1$, the rank-truncated approximations of the scaling matrix $\mathbf{D}_\mathcal{X}$ are applied to $\mathbf{v}$ directly. Hence, in this case, we could directly use $L_t(\mathbf{v})$ instead of $L_t(\tilde{\mathbf{T}}_{J(\eta;\mathbf{v}),\mathbf{v}}\mathbf{v})$. However, as considered in further detail in Section 4.4, we do not only apply the operator $\mathbf{B}_1$, but also its transpose $\mathbf{B}_1^\top = \mathbf{D}_\mathcal{X}\mathbf{T}^\top\mathbf{D}_\mathcal{Y}$. Because the scaling matrices are diagonal, they coincide with their transposes, but the order of the matrices is reversed. The above estimates thus cover the two extreme cases.

4.3.4 Initial Value Operator

In this section, we discuss the adaptive approximate application of the initial value operator $\mathbf{B}_2 = \mathbf{T}_0\mathbf{D}_\mathcal{X}$ defined in (4.35) and of its transpose $\mathbf{B}_2^\top = \mathbf{D}_\mathcal{X}\mathbf{T}_0^\top$. Similar to the temporal operator $\mathbf{T}_t$ the operator $\mathbf{T}_0$ is of the form

$$\mathbf{T}_0 = \mathbf{T}_{0,t} \otimes \mathbf{I}_x \tag{4.122}$$

with the identity in the spatial variable. However, $\mathbf{T}_{0,t}$ maps from $\ell_2(\vee_t)$ to $\mathbb{R}$ with

$$(\mathbf{T}_{0,t})_{\nu_t} = \theta_{\nu_t}(0), \quad \nu_t \in \vee_t.$$

As we will discuss in more detail in Section 4.4, we are not interested in the adaptive application of $\mathbf{B}_2$ itself. Instead, we focus on the adaptive application of $\mathbf{A}_2 = \mathbf{B}_2^\top\mathbf{B}_2$ and $\mathbf{B}_2^\top$, since we will solve a least-squares formulation. There are two

possibilities to realize the adaptive application of $\mathbf{A}_2$. In [SS09; BF24] the adaptive application of $\mathbf{B}_2$ and $\mathbf{B}_2^\top$ are considered separately. Then these two adaptive applications are applied one after the other with properly chosen tolerances. To construct a suitable approximation of $\mathbf{B}_2$, one can proceed similarly to [SS09; BF24] using the decay of the entries. Based on the choice of temporal wavelets, for each $\nu_t \in \vee_t$ there exists a $t_0 \in [0,T]$ with $\theta_{\nu_t}(t_0) = 0$. Hence, we obtain

$$(\theta_{\nu_t}(0))^2 = \left| \int_0^{t_0} \frac{\mathrm{d}}{\mathrm{d}t}\big(\theta_{\nu_t}(t)^2\big)\,\mathrm{d}t \right| \leq 2\int_0^{t_0} |\theta_{\nu_t}(t)||\theta'_{\nu_t}(t)|\,\mathrm{d}t \leq 2\|\theta_{\nu_t}\|_{L_2(0,T)}\|\theta_{\nu_t}\|_{H^1(0,T)}.$$

By this trace theorem-like estimate, the orthonormality of the wavelet basis and the connection between the level and the H^1-norm, we have

$$|\theta_{\nu_t}(0)| \leq \sqrt{2}\sqrt{\|\theta_{\nu_t}\|_{L_2}\|\theta_{\nu_t}\|_{H^1}} = \sqrt{2}\|\theta_{\nu_t}\|_{H^1}^{\frac{1}{2}} \eqsim 2^{\frac{1}{2}|\nu_t|}. \tag{4.123}$$

Hence, based on the definition of $\bar{\mathbf{D}}_{\mathcal{X}}$ from (3.26) and the definition of $\mathbf{T}_0$, we have for $\bar{\mathbf{B}}_2 = \mathbf{T}_0\bar{\mathbf{D}}_{\mathcal{X}}$ the entry-wise estimate

$$\begin{aligned}
|(\bar{\mathbf{B}}_2)_{\nu_x,(\nu_t,\nu_x)}| &= \frac{|\theta_{\nu_t}(0)|\|\Psi_{\nu_x}\|_{H_0^1}}{\|\Psi_{\nu_x}\|_{H_0^1}^2 + \|\theta_{\nu_t}\|_{H^1}} \\
&\lesssim \frac{\|\theta_{\nu_t}\|_{H^1}^{\frac{1}{2}}\|\Psi_{\nu_x}\|_{H_0^1}}{\|\Psi_{\nu_x}\|_{H_0^1}^2 + \|\theta_{\nu_t}\|_{H^1}} \\
&\leq \min\left\{\|\theta_{\nu_t}\|_{H^1}^{-\frac{1}{2}}\|\Psi_{\nu_x}\|_{H_0^1}, \|\theta_{\nu_t}\|_{H^1}^{\frac{1}{2}}\|\Psi_{\nu_x}\|_{H_0^1}^{-1}\right\} \\
&\eqsim 2^{-\left|\frac{1}{2}|\nu_t| - \log_2\left(\|\Psi_{\nu_x}\|_{H_0^1}\right)\right|}
\end{aligned} \tag{4.124}$$

for $\nu_t \in \vee_t$ with $\theta_{\nu_t}(0) \neq 0$ and $\nu_x \in \vee_x$. The decay of the entries depends strongly on the coupling between the temporal and spatial wavelets. Using the relation

$$\|\Psi_{\nu_x}\|_{H_0^1}^2 = \sum_{i=1}^d \|\psi_{\nu_i}\|_{H_0^1}^2 \eqsim \sum_{i=1}^d 2^{2|\nu_i|} \leq d\,2^{2\max_i|\nu_i|}$$

of the maximal one-dimensional wavelet level and the H_0^1-norm, we similarly get

$$|(\bar{\mathbf{B}}_2)_{\nu_x,(\nu_t,\nu_x)}| \lesssim \sqrt{d}\,2^{-\left|\frac{1}{2}|\nu_t| - \max_i|\nu_i|\right|}. \tag{4.125}$$

Based on the two estimates (4.124) and (4.125), we see that the construction of approximation sequences of $\mathbf{B}_2$ and $\mathbf{B}_2^\top$ relies on the coupling between temporal and spatial wavelets. Specifically, an example of such an approximation for $\mathbf{B}_2$ is provided in [SS09], while a corresponding construction for high-dimensional cases is detailed in [BF24].

In this work, we follow the second possibility, which is to construct an adaptive application of $\mathbf{A}_2 = \mathbf{B}_2^\top\mathbf{B}_2 \colon \ell_2(\vee) \to \ell_2(\vee)$ directly. For this operator, we can avoid the coupling between temporal and spatial wavelets and proceed as in Section 4.3.2. After this construction, we give an adaptive application of $\mathbf{B}_2^\top$, where we have to take into account this coupling.

4.3.4.1 Adaptive Approximation of $\mathbf{A}_2$

Unlike the entries of $\mathbf{D}_{\mathcal{X}}$, which are defined as infinite sums, the entries of the scaling matrix $\bar{\mathbf{D}}_{\mathcal{X}}$ have a closed-form. Therefore we first analyze the approximability of $\bar{\mathbf{A}}_2 = \bar{\mathbf{B}}_2^\top \bar{\mathbf{B}}_2$. Using (4.57), we can then directly infer the approximability of $\mathbf{A}_2$ from the approximability of $\bar{\mathbf{A}}_2$. By definition, we have

$$\bar{\mathbf{A}}_2 = \bar{\mathbf{D}}_{\mathcal{X}} \mathbf{S}_0 \bar{\mathbf{D}}_{\mathcal{X}} = \bar{\mathbf{D}}_{\mathcal{X}} (\mathbf{S}_{0,\mathrm{t}} \otimes \mathbf{I}_{\mathrm{x}}) \bar{\mathbf{D}}_{\mathcal{X}}$$

with

$$\mathbf{S}_0 = \mathbf{T}_0^\top \mathbf{T}_0, \quad \mathbf{S}_{0,\mathrm{t}} = \mathbf{T}_{0,\mathrm{t}}^\top \mathbf{T}_{0,\mathrm{t}}, \tag{4.126}$$

which means

$$(\mathbf{S}_{0,\mathrm{t}})_{\nu_\mathrm{t},\mu_\mathrm{t}} = \theta_{\nu_\mathrm{t}}(0)\theta_{\mu_\mathrm{t}}(0), \quad \nu_\mathrm{t}, \mu_\mathrm{t} \in \vee_\mathrm{t}.$$

By definition, we have $(\bar{\mathbf{A}}_2)_{(\mu_\mathrm{t},\nu_\mathrm{x}),(\nu_\mathrm{t},\mu_\mathrm{x})} = 0$ for $\nu_\mathrm{x} \neq \mu_\mathrm{x}$. Moreover, proceeding as in (4.124), we obtain

$$\begin{aligned}
|(\bar{\mathbf{A}}_2)_{(\mu_\mathrm{t},\nu_\mathrm{x}),(\nu_\mathrm{t},\nu_\mathrm{x})}| &= \frac{|\theta_{\nu_\mathrm{t}}(0)|\|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}}{\|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}^2 + \|\theta_{\nu_\mathrm{t}}\|_{H^1}} \frac{|\theta_{\mu_\mathrm{t}}(0)|\|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}}{\|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}^2 + \|\theta_{\mu_\mathrm{t}}\|_{H^1}} \\
&\lesssim \frac{\|\theta_{\nu_\mathrm{t}}\|_{H^1}^{\frac{1}{2}} \|\theta_{\mu_\mathrm{t}}\|_{H^1}^{\frac{1}{2}} \|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}^2}{\|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}^4 + \|\Psi_{\nu_\mathrm{x}}\|_{H_0^1}^2 (\|\theta_{\nu_\mathrm{t}}\|_{H^1} + \|\theta_{\mu_\mathrm{t}}\|_{H^1}) + \|\theta_{\nu_\mathrm{t}}\|_{H^1}\|\theta_{\mu_\mathrm{t}}\|_{H^1}} \\
&\leq \frac{\|\theta_{\nu_\mathrm{t}}\|_{H^1}^{\frac{1}{2}} \|\theta_{\mu_\mathrm{t}}\|_{H^1}^{\frac{1}{2}}}{\|\theta_{\nu_\mathrm{t}}\|_{H^1} + \|\theta_{\mu_\mathrm{t}}\|_{H^1}} \\
&\eqsim 2^{-\frac{||\nu_\mathrm{t}|-|\mu_\mathrm{t}||}{2}}.
\end{aligned} \tag{4.127}$$

Unlike $\bar{\mathbf{B}}_2$, for $\bar{\mathbf{A}}_2$, the decay of the entries only depend on the difference of the temporal wavelet levels, avoiding the coupling with the spatial wavelets. Next, we define an approximation sequence of $\bar{\mathbf{A}}_2$ based on this decay of the entries. To this end, for $J \in \mathbb{N}_0$ let $\bar{\mathbf{S}}_{0,\mathrm{t},j} \in \mathbb{R}^{\vee_\mathrm{t} \times \vee_\mathrm{t}}$ be defined by

$$(\bar{\mathbf{S}}_{0,\mathrm{t},j})_{\nu_\mathrm{t},\mu_\mathrm{t}} = \begin{cases} (\mathbf{S}_{0,\mathrm{t}})_{\nu_\mathrm{t},\mu_\mathrm{t}}, & ||\nu_\mathrm{t}| - |\mu_\mathrm{t}|| \leq j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.128}$$

Moreover, we define

$$\bar{\mathbf{S}}_{0,j} = \bar{\mathbf{S}}_{0,\mathrm{t},j} \otimes \mathbf{I}_{\mathrm{x}}.$$

This means that we set any entry of $\mathbf{S}_{0,j}$ to zero for which the difference in the temporal wavelet levels is more than j. Since our wavelets are compactly supported, only a fixed number $\bar{c}$ of them per level are nonzero at $t = 0$. Thus for each row of $\bar{\mathbf{S}}_{0,\mathrm{t},j}$ there are at most $(2j+1)\bar{c}$ nonzero entries. Due to the structure with the identity in the spatial variables, the same is true for $\bar{\mathbf{S}}_{0,j}$.

Let $(\nu_t, \nu_x) \in \vee$. Then by (4.127), we obtain

$$\begin{aligned}\sum_{(\mu_t,\mu_x)\in\vee} \left|(\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \bar{\mathbf{S}}_{0,j})\bar{\mathbf{D}}_{\mathcal{X}})_{(\mu_t,\mu_x),(\nu_t,\nu_x)}\right| &= \sum_{\substack{\mu_t\in\vee_t \\ ||\nu_t|-|\mu_t||>j}} \left|(\bar{\mathbf{A}}_2)_{(\mu_t,\nu_x),(\nu_t,\nu_x)}\right| \\ &\lesssim \sum_{\substack{\mu_t\in\vee_t \\ ||\nu_t|-|\mu_t||>j \\ \theta_{\mu_t}(0)\neq 0}} 2^{-\frac{||\nu_t|-|\mu_t||}{2}} \\ &\lesssim \sum_{k=j+1}^{\infty} \bar{c}2^{-\frac{k}{2}} \\ &\lesssim \bar{c}2^{-\frac{j}{2}}.\end{aligned}$$

By construction, $\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \bar{\mathbf{S}}_{0,j})\bar{\mathbf{D}}_{\mathcal{X}}$ is symmetric. Hence, we also have

$$\sum_{(\mu_t,\mu_x)\in\vee} |(\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \bar{\mathbf{S}}_{0,j})\bar{\mathbf{D}}_{\mathcal{X}})_{(\nu_t,\nu_x),(\mu_t,\mu_x)}| \lesssim \bar{c}2^{-\frac{j}{2}}.$$

Using Theorem 2.29, we conclude

$$\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \bar{\mathbf{S}}_{0,j})\bar{\mathbf{D}}_{\mathcal{X}}\| \lesssim \bar{c}2^{-\frac{j}{2}},$$

which yields the super-compressibility of $\bar{\mathbf{A}}_2$.

Remark 4.50. As mentioned in Theorem 2.27, by taking suitable sequences in the super-compressibility, we can match any rate. To this end, for given $s > 0$ let $(\mathbf{S}_{0,j})_{j\in\mathbb{N}_0}$ be chosen such that

$$\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \mathbf{S}_{0,j})\bar{\mathbf{D}}_{\mathcal{X}}\| \lesssim \beta_j^{(0)}2^{-js}$$

with $(\beta_j^{(0)})_{j\in\mathbb{N}_0} \in \ell_1(\mathbb{N}_0)$ and that only $c^{(0)}j$ many entries are nonzero in each row and column of $\mathbf{S}_{0,j}$ for some constant $c^{(0)} > 0$ independent of j. Furthermore, we assume that the $\mathbf{S}_{0,j}$ are chosen nested, i.e., only $c^{(0)}$ entries are nonzero in each row and column of $\mathbf{S}_{0,j} - \mathbf{S}_{0,j-1}$.

Now, we can proceed similarly as for the temporal operator $\mathbf{B}_t$ in Section 4.3.2. For a given refinement parameter $J \in \mathbb{N}_0$, we define the approximation

$$\begin{aligned}\tilde{\mathbf{A}}_{2,J}(\mathbf{v}) &= \mathbf{D}_{\mathcal{X}}\mathbf{S}_{0,J}\mathbf{D}_{\mathcal{X}}\mathbf{v}_0 + \sum_{j=1}^{J} \mathbf{D}_{\mathcal{X}}\mathbf{S}_{0,J-j}\mathbf{D}_{\mathcal{X}}(\mathbf{v}_j - \mathbf{v}_{j-1}) \\ &= \sum_{j=0}^{J-1} \mathbf{D}_{\mathcal{X}}(\mathbf{S}_{0,J-j} - \mathbf{S}_{0,J-j-1})\mathbf{D}_{\mathcal{X}}\mathbf{v}_j + \mathbf{D}_{\mathcal{X}}\mathbf{S}_{0,0}\mathbf{D}_{\mathcal{X}}\mathbf{v}_J\end{aligned} \tag{4.129}$$

with $\mathbf{v}_j$ defined in (4.88).

Again, as in the spatial and temporal case, we also want to be able to apply the approximation of the operator defined by the input $\mathbf{v}$ to a general vector $\mathbf{w} \in \ell_2(\vee)$. To this end, we define the operator

$$\begin{aligned}\tilde{\mathbf{S}}_{0,J,\mathbf{v}} &= \sum_{j=0}^{J-1} \big(\mathbf{S}_{0,J-j} - \mathbf{S}_{0,J-j-1}\big)\mathrm{R}_{\bar{\Lambda}_j^{(t)}(\mathbf{v})} + \mathbf{S}_{0,0}\mathrm{R}_{\bar{\Lambda}_J^{(t)}(\mathbf{v})} \\ &= \mathbf{S}_{0,J}\mathrm{R}_{\bar{\Lambda}_0^{(t)}(\mathbf{v})} + \sum_{j=1}^{J} \mathbf{S}_{0,J-j}\Big(\mathrm{R}_{\bar{\Lambda}_j^{(t)}(\mathbf{v})} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(t)}(\mathbf{v})}\Big),\end{aligned} \tag{4.130}$$

where $\bar{\Lambda}_j^{(\mathrm{t})}$ are defined in (4.88). We note that

$$\tilde{\mathbf{A}}_{2,J}(\mathbf{v}) = \mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_{0,J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{v}.$$

Next, we state the results for the adaptive application of $\mathbf{A}_2$ with respect to the refinement level J.

Lemma 4.51. *Let $\mathbf{A}_2 = \mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}$ and assume that the operator $\mathbf{S}_0$ defined in (4.126) is super-compressible in the sense of Theorem 4.50 with corresponding sequence $(\beta_j^{(0)})_{j\in\mathbb{N}_0}$ and $c^{(0)}$. Let $\mathbf{v},\mathbf{w} \in \ell_2(\vee)$ and $J \in \mathbb{N}_0$. Then whenever $\pi^{(\mathrm{t},i)}(\mathbf{v}) \in \mathcal{A}^s$ for some $0 < s < s^*$ and $i = 1,\dots,d$, one has the a posteriori error bound*

$$\|\mathbf{A}_2\mathbf{w} - \mathbf{D}_{\mathcal{X}}\mathbf{S}_{0,J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{w}\| \le e_{0,J,\mathbf{v}}(\mathbf{w}), \tag{4.131}$$

where

$$e_{0,J,\mathbf{v}}(\mathbf{w}) = \sum_{i=1}^{d} C_\delta \sum_{j=0}^{J} 2^{-s(J-j)}\beta_{J-j}^{(0)}\left\|\mathrm{R}_{\Lambda_{[j]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\right\| + \|\mathbf{A}_2\|\left\|\mathrm{R}_{\Lambda_{[J+1]}^{(\mathrm{t},i)}(\mathbf{v})}\pi^{(\mathrm{t},i)}(\mathbf{w})\right\|$$

as well as the a priori error bound

$$\|\mathbf{A}_2\mathbf{v} - \tilde{\mathbf{A}}_{2,J}(\mathbf{v})\| \le 2^{-sJ}(C_\delta 2^s\|\beta^{(0)}\|_{\ell_1} + \|\mathbf{A}_2\|)\sum_{i=1}^{d}\|\pi^{(\mathrm{t},i)}(\mathbf{v})\|_{\mathcal{A}^s}, \tag{4.132}$$

where $\tilde{\mathbf{A}}_{2,J}(\mathbf{v})$ is defined by (4.129). Moreover, one has the support estimate

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{2,J}(\mathbf{v}))\right) \le c^{(0)}2^J, \quad i = 1,\dots,d.$$

Proof. The proof of this lemma is similar to the proof of Theorem 4.35. Hence, we follow the same arguments.

Using the identity

$$\mathbf{A}_2\mathbf{w} = \sum_{j=0}^{J}\mathbf{A}_2\Big(\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\Big) + \mathbf{A}_2\Big(\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\Big)$$

combined with the triangle inequality and the definition of $\tilde{\mathbf{S}}_{0,J,\mathbf{v}}$ in (4.130), we obtain

$$\begin{aligned}
&\|\mathbf{A}_2\mathbf{w} - \mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_{0,J,\mathbf{v}}\mathbf{D}_{\mathcal{X}}\mathbf{w}\| \\
&\quad\le \sum_{j=0}^{J}\|\mathbf{D}_{\mathcal{X}}(\mathbf{S}_0 - \mathbf{S}_{0,J-j})\mathbf{D}_{\mathcal{X}}\|\left\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| + \|\mathbf{A}_2\|\left\|\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| \\
&\quad\le \sum_{j=0}^{J} C_\delta\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{S}_0 - \mathbf{S}_{0,J-j})\bar{\mathbf{D}}_{\mathcal{X}}\|\left\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{v})}\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\| + \|\mathbf{A}_2\|\left\|\mathbf{w} - \mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{v})}\mathbf{w}\right\|,
\end{aligned}$$

where we used (4.57) in the last step. Using the super-compressibility of $\mathbf{S}_0$ and the estimate (4.92) yields (4.131). The a priori estimate (4.132) is a direct consequence of the a posteriori estimate (4.131) combined with Theorem 4.31.

By assumption $(\mathbf{S}_{0,j+1} - \mathbf{S}_{0,j})$ has only $c^{(0)}$ nonzero entries in each column and by choice $\mathbf{S}_{0,0} = 0$. In view of (4.129) and (4.89), we have

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{2,J}(\mathbf{v}))\right) \le \sum_{j=0}^{J-1} c^{(0)}2^j \le c^{(0)}2^J,$$

since the diagonal operator $\mathbf{D}_{\mathcal{X}}$ does not change the support. □

For notational simplicity, we write $e_{J,0}(\mathbf{v}) = e_{J,0,\mathbf{v}}(\mathbf{v})$. As in Section 4.3.3, we next state the results in terms of a given tolerance $\eta > 0$ instead of the refinement level J and go from the full scaling matrix $\mathbf{D}_{\mathcal{X}}$ to its rank-truncated version $\mathbf{D}_{\mathcal{X},\mathsf{n}}$. To do this, we state the following lemma as an analogue of Theorem 4.38.

Lemma 4.52. *Let* $\mathbf{A}_2 = \mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}$ *with* $\mathbf{S}_0$ *defined in* (4.126), $\tilde{\mathbf{S}}_0 \in \mathbb{R}^{\vee\times\vee}$ *and*

$$\mathbf{G} = \mathbf{A}_2 - \mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_0\mathbf{D}_{\mathcal{X}} = \mathbf{D}_{\mathcal{X}}(\mathbf{S}_0 - \tilde{\mathbf{S}}_0)\mathbf{D}_{\mathcal{X}}.$$

Let $\mathbf{v} \in \ell_2(\vee)$ *and* $\mathsf{K}_1, \mathsf{K}_2 \in \mathbb{R}^{\vee_{\mathsf{t}}}$ *be such that*

$$\operatorname{supp}(\mathbf{v}) \subseteq \Lambda_{\mathsf{K}_1}, \quad \operatorname{supp}(\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_0\mathbf{D}_{\mathcal{X}}\mathbf{v}) \subseteq \Lambda_{\mathsf{K}_2},$$

where Λ_{K_1} *and* Λ_{K_2} *are defined by* (4.52). *Then whenever* $\mathsf{n}_1 \geq \mathsf{M}_{\mathcal{X}}(\eta, \mathsf{K}_1)$, $\mathsf{n}_2 \geq \mathsf{M}_{\mathcal{Y}}(\eta, \mathsf{K}_2)$, *we have*

$$\begin{aligned}\big\|(\mathbf{A}_2 - \mathbf{D}_{\mathcal{X},\mathsf{n}_2}\tilde{\mathbf{S}}_0\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\big\| \leq \|\mathbf{G}\mathbf{v}\| + \|\mathbf{G}(\mathbf{I} - \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\| \\ + \eta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{G}(\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1})\mathbf{v}\| + 2\eta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{A}_2\|\|\mathbf{v}\|.\end{aligned}$$

The proof of this lemma is analogous to the proof of Theorem 4.38, since we have the same properties for the scaling matrix $\mathbf{D}_{\mathcal{X}}$ as for $\mathbf{D}_{\mathcal{Y}}$. Hence, we omit the proof here.

Next, we state a result on the adaptive approximation of $\mathbf{A}_2$ based on a given tolerance $\eta > 0$ using the rank-truncated scaling matrices. As for the operator $\mathbf{B}_1$ in Section 4.3.3, we define the corresponding refinement level for the given tolerance by

$$J_0(\eta;\mathbf{v}) = \min\big\{J \in \mathbb{N}_0 : e_{J,0}(\mathbf{v}) \leq \tfrac{\eta}{4}\big\}, \tag{4.133}$$

which exists by a similar argument as in Theorem 4.37. We define the quantities

$$c_0(\eta;\mathbf{v}) = \frac{\eta}{4\|\mathbf{A}_2\|\|\mathbf{v}\|}(1-\tfrac{\delta}{2})$$

and

$$\mathsf{m}_{\mathcal{X},0}(\eta;\mathbf{v}) = \mathsf{M}_{\mathcal{X}}\big(c_0(\eta;\mathbf{v}); \mathsf{K}_0(J_0(\eta;\mathbf{v});\mathbf{v})\big),$$

where

$$\mathsf{K}_0(J;\mathbf{v}) = \min\{\mathsf{K} > 0 : \operatorname{supp}(\mathbf{v}) \cup \operatorname{supp}(\tilde{\mathbf{S}}_{0,J,\mathbf{v}}\mathbf{v}) \subseteq \Lambda_{\mathsf{K}}\}.$$

Again, the minimum is to be understood element-wise for each $K_{\nu_{\mathsf{t}}}$ with $\nu_{\mathsf{t}} \in \vee_{\mathsf{t}}$.

Theorem 4.53. *Let* $0 \neq \mathbf{v} \in \ell_2(\vee)$ *and* $0 < \eta < 2\|\mathbf{A}_2\|\|\mathbf{v}\|$. *We fix*

$$J = J_0(\eta;\mathbf{v}), \quad \mathsf{n} = \mathsf{m}_{\mathcal{X},0}(\eta;\mathbf{v}).$$

Then with $\mathbf{w}_\eta = \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{S}}_{0,J,\mathbf{v}}\mathbf{D}_{\mathcal{X},\mathsf{n}}\mathbf{v}$, *we have*

$$\|\mathbf{A}_2\mathbf{v} - \mathbf{w}_\eta\| \leq \eta\,.$$

This theorem is an analogue of Theorem 4.40, and the proof can be done in the same way using Theorem 4.52 instead of Theorem 4.38. Hence, we omit the proof here. Theorems 4.41 and 4.42 apply also in this case.

Next, we give estimates on the support size, the ranks and the $\mathcal{A}^s$-norms of the approximations.

Lemma 4.54. *Under the assumptions of Theorem 4.53, we have*

$$\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\big) \leq 2^{1+\frac{2}{s}}\tilde{C}_0^{\frac{1}{s}}\eta^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}$$

with $\tilde{C}_0 = C_\delta 2^s\|\beta^{(0)}\|_{\ell_1} + \|\mathbf{A}_2\|$, *as well as*

$$\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\big\|_{\mathcal{A}^s} \leq \hat{C}_0\sum_{j=1}^{d}\|\pi^{(\mathrm{t},j)}(\mathbf{v})\|_{\mathcal{A}^s},$$

where $\hat{C}_0 = 2^{1+4s}(c^{(0)})^s(C_\delta 2^s\|\beta(\mathbf{C}_t)\|_{\ell_1} + \|\mathbf{A}_2\|)$. *Moreover, we have the rank estimate*

$$\operatorname{rank}_\infty(\mathbf{w}_\eta) \leq \|\hat{\mathrm{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty}^2 c^{(0)} J_0(\eta;\mathbf{v})\operatorname{rank}_\infty(\mathbf{v}), \tag{4.134}$$

where

$$\hat{\mathrm{m}}_{\mathcal{X},0,\nu_\mathrm{t}}(\eta;\mathbf{v}) = 1 + n^+_{a_{\nu_\mathrm{t}}}(\delta) + \mathrm{m}_{\mathcal{X},0,\nu_\mathrm{t}}(\eta;\mathbf{v}), \quad \nu_\mathrm{t} \in \vee_\mathrm{t}.$$

The proof of this lemma is a simplified version of the proof of Theorem 4.43, since we only have to consider the super-compressible part here and do not have to deal with the additional spatial operator. For the proof, we can follow the same arguments and use the statements of Theorem 4.51, where the statements are given with respect to the refinement level J. Hence, we omit the proof here.

In view of Theorem 4.44, we are in most situations not interested in the compressibility of the spatio-temporal contractions of a specific dimension, but rather in the sum over all dimensions. As a direct consequence of Theorem 4.54, we have

$$\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\big\|_{\mathcal{A}^s} \leq d\hat{C}_0\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}.$$

In the following lemma, we investigate the computational complexity of the adaptive application of the operator $\mathbf{A}_2$.

Lemma 4.55. *Under the assumptions of Theorem 4.53, the number* $\operatorname{flops}(\mathbf{w}_\eta)$ *of floating point operations to compute* $\mathbf{w}_\eta$ *is bounded by*

$$\begin{aligned}\operatorname{flops}(\mathbf{w}_\eta) \lesssim\ & (c^{(0)})^4(J_0(\eta;\mathbf{v}))^4\hat{m}_0^3(\eta;\mathbf{v})\operatorname{rank}_\infty^3(\mathbf{v})\sum_{i=1}^{d}\#\operatorname{supp}\Big(\pi^{(\mathrm{t},i)}(\mathbf{v})\Big)\\ &+ d\eta^{-\frac{1}{s}}\operatorname{rank}_\infty(\mathbf{v})J_0(\eta;\mathbf{v})\hat{m}_0(\eta;\mathbf{v})c^{(0)}\tilde{C}_0^{\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{v})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\end{aligned}$$

where $\hat{m}_0(\eta;\mathbf{v}) = \|\hat{\mathrm{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty}^2$, *and where the hidden constant is independent of* $\eta, \mathbf{A}_2, \mathbf{v}$ *and* d.

Again, this statement is a simplified version of the statement in Theorem 4.46, and we can prove it in the same way. In this case, we only need to consider the super-compressible part. Hence, we omit the proof here.

We end this section with an estimate of $\|\hat{\mathrm{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty}$, which is an estimate of the rank increase caused by the rank-truncated scaling matrix $\mathbf{D}_{\mathcal{X},\mathrm{n}}$. By construction of $(\bar{\mathbf{S}}_{0,j})_{j\in\mathbb{N}}$ with $\bar{\mathbf{S}}_{0,j} = \bar{\mathbf{S}}_{0,j,\mathrm{t}} \otimes \mathbf{I}_\mathrm{x}$, where $\bar{\mathbf{S}}_{0,j,\mathrm{t}}$ is defined in (4.128), we can directly infer that $\mathbf{A}_2$ has the level decay property from Theorem 2.28. The parameter γ in the level decay property will depend on the parameter s from the compressibility.

Furthermore, to bound $\|\hat{\mathfrak{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty}$ we can proceed as for $\|\hat{\mathfrak{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty}$ in Theorem 4.48.

Lemma 4.56. *Under the assumptions of Theorem 4.53, we have the estimate*

$$\|\hat{\mathfrak{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty} \leq C(\delta,s,\mathbf{A}_2)\left[1+L(\mathbf{v})+|\ln(\eta)|+\ln\left(\sum_{i=1}^{d}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right]^2$$

with the maximal one-dimensional wavelet level $L(\mathbf{v})$ defined in (4.115).

This estimate can be proven as for $\|\hat{\mathfrak{m}}_{\mathcal{X}}(\eta;\mathbf{v})\|_{\ell_\infty}$ in Theorem 4.48 using the corresponding estimates for $\mathbf{A}_2$, J_0 and K_0 instead.

As one can see in the rank estimate (4.134) from Theorem 4.54, the increase caused by the adaptive application of $\mathbf{A}_2$ does not only depend on $\|\hat{\mathfrak{m}}_{\mathcal{X},0}(\eta;\mathbf{v})\|_{\ell_\infty}$ but also on the refinement level $J_0(\eta;\mathbf{v})$. Proceeding as in Section 4.3.3 by combining the minimality of $J_0(\eta;\mathbf{v})$ in (4.133) with the a priori estimate (4.132) from Theorem 4.51 we obtain the bound

$$J_0(\eta;\mathbf{v}) \leq \frac{1}{s}\left(|\ln\eta|+\ln\left(C(\mathbf{A}_2,s,\delta)\sum_{i=1}^{d}\left\|\pi^{(\mathrm{t},i)}(\mathbf{v})\right\|_{\mathcal{A}^s}\right)\right). \tag{4.135}$$

4.3.4.2 Adaptive Approximation of $\mathbf{B}_2^\top$

As final part, we consider the adaptive application of the adjoint $\mathbf{B}_2^\top$ of the initial-value operator. In [BF24, Section 5.4], a possible way for an adaptive application of this operator was already discussed. There, the input was additionally decomposed into terms with fixed maximal spatial levels to make use of the relation (4.125).

Here, we follow a different approach. Instead of applying different approximations of $\mathbf{B}_2^\top$ with different precision at once, we apply a fixed approximation of the operator. Nevertheless, as for the other operators, we again combine a priori information of the operator with a posteriori information of the input.

For a finitely supported input $\mathbf{v} \in \ell_2(\vee)$ and a refinement level $j \in \mathbb{N}_0$ we define the approximation

$$(\mathbf{T}_{0,j,\mathbf{v}})_{\nu_\mathrm{x},(\nu_\mathrm{t},\nu_\mathrm{x})} = \begin{cases} (\mathbf{T}_0)_{\nu_\mathrm{x},(\nu_\mathrm{t},\nu_\mathrm{x})}, & |\nu_\mathrm{t}| \leq 2L_\mathrm{x}(\mathbf{v})+j, \\ 0, & |\nu_\mathrm{t}| > 2L_\mathrm{x}(\mathbf{v})+j. \end{cases}$$

Let $\nu \in \vee_\mathrm{x}$ with $\max_i|\nu_i| \leq L_\mathrm{x}(\mathbf{v})$. Then using the identity in the spatial variable as well as (4.125), we obtain

$$\begin{aligned}
\sum_{(\mu_\mathrm{t},\mu_\mathrm{x})\in\vee}\left|(\bar{\mathbf{D}}\mathcal{X}(\mathbf{T}_0^\top-\mathbf{T}_{0,j,\mathbf{v}}^\top))_{(\mu_\mathrm{t},\mu_\mathrm{x}),\nu_\mathrm{x}}\right| &= \sum_{\substack{\mu_\mathrm{t}\in\vee_\mathrm{t}\\ |\nu_\mathrm{t}|>2L_\mathrm{x}(\mathbf{v})+j}}\left|(\bar{\mathbf{B}}_2^\top)_{(\mu_\mathrm{t},\nu_\mathrm{x}),\nu_\mathrm{x}}\right| \\
&\lesssim \sqrt{d}\sum_{\substack{\mu_\mathrm{t}\in\vee_\mathrm{t}\\ |\nu_\mathrm{t}|>2L_\mathrm{x}(\mathbf{v})+j\\ \theta_{\mu_\mathrm{t}}(0)\neq 0}} 2^{-\left|\frac{1}{2}|\nu_\mathrm{t}|-\max_i|\nu_i|\right|} \\
&\lesssim \sqrt{d}\sum_{k=j+1}^{\infty}\bar{c}2^{-\frac{k}{2}} \\
&\lesssim \sqrt{d}\bar{c}2^{-\frac{j}{2}},
\end{aligned}$$

where $\bar{c}$ is the number of nonvanishing wavelets per level at $t = 0$. Moreover, for $(\nu_t, \nu_x) \in \vee$ with $\max_i |\nu_{x,i}| \le L_x(\mathbf{v})$, we have

$$\sum_{\mu_x \in \vee_x} \left|(\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{T}_0^\top - \mathbf{T}_{0,j,\mathbf{v}}^\top))_{(\nu_t,\nu_x),\mu_x}\right| = \begin{cases} 0, & |\nu_t| \le 2L_x(\mathbf{v}) + j, \\ \left|(\bar{\mathbf{B}}_2^\top)_{(\nu_t,\nu_x),\nu_x}\right|, & |\nu_t| > 2L_x(\mathbf{v}) + j, \end{cases} \lesssim \sqrt{d}2^{-\frac{j}{2}},$$

where we again used (4.125). Let

$$\Lambda_x(\mathbf{v}) = \left\{\nu \in \vee_x : \max_i |\nu_i| \le L_x(\mathbf{v})\right\}$$

and let $\mathbf{I}_{\Lambda_x(\mathbf{v})}$ be the identity operator restricted on $\Lambda_x(\mathbf{v})$, meaning that $(\mathbf{I}_{\Lambda_x(\mathbf{v})})_{\nu_x,\nu_x}$ is one if $\nu_x \in \Lambda_x(\mathbf{v})$ and zero otherwise. Then by Theorem 2.29 we conclude

$$\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{T}_0^\top - \mathbf{T}_{0,j,\mathbf{v}}^\top)\mathbf{I}_{\Lambda_x(\mathbf{v})}\| \lesssim \sqrt{d}2^{-\frac{j}{2}}. \tag{4.136}$$

We define

$$\tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v}) = \mathbf{D}_{\mathcal{X}}\mathbf{T}_{0,J,\mathbf{v}}^\top\mathbf{v} = \sum_{\substack{\nu_t \in \vee_t \\ |\nu_t| \le 2L_x(\mathbf{v})+j}} \mathbf{e}_{\nu_t} \otimes \theta_{\nu_t}(0)\mathbf{D}_{\mathcal{X},\nu_t}\mathbf{v}, \tag{4.137}$$

where we used the definition of $\mathbf{T}_{0,j,\mathbf{v}}$ and $\mathbf{T}_0$ in the last step. As before, we first analyze the approximation error with respect to the refinement level J.

Lemma 4.57. *Let $\mathbf{B}_2$ be defined by (4.35). Given any finitely supported $\mathbf{v} \in \ell_2(\vee_x)$ and $J \in \mathbb{N}_0$, let $\tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v})$ be defined by (4.137). Then whenever $\pi^{(i)}(\mathbf{v}) \in \mathcal{A}^s$ for some $0 < s < s^*$ and $i = 1, \dots, d$, one has*

$$\|\mathbf{B}_2^\top\mathbf{v} - \tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v})\| \le e_{0,J}^{\mathrm{ad}}(\mathbf{v})$$

with

$$e_{0,J}^{\mathrm{ad}}(\mathbf{v}) = CC_\delta^{\frac{1}{2}}d^{-\frac{1}{2}}2^{-\frac{J}{2}}\sum_{i=1}^d \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s},$$

where C depend on the number $\bar{c}$ of nonvanishing temporal basis functions per level at $t = 0$. Moreover, one has the support estimate

$$\#\operatorname{supp}\left(\pi^{(t,i)}(\tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v}))\right) \le \bar{c}(2L_x(\mathbf{v}) + J)\#\operatorname{supp}\left(\pi^{(i)}(\mathbf{v})\right), \quad i = 1, \dots, d, \tag{4.138}$$

where $L_x(\mathbf{v})$ is the maximal one-dimensional wavelet level of $\mathbf{v}$.

Proof. For each $\nu_t \in \vee_t$ with $\theta_{\nu_t}(0) \neq 0$, we have

$$\#\operatorname{supp}\left(\pi^{(i)}(\theta_{\nu_t}(0)\mathbf{D}_{\mathcal{X},\nu_t}\mathbf{v})\right) = \#\operatorname{supp}\left(\pi^{(i)}(\mathbf{v})\right)$$

since $\mathbf{D}_{\mathcal{X},\nu_t}$ is a diagonal operator. Hence, (4.138) is a direct consequence of (4.137). Furthermore,

$$\begin{aligned} \|\mathbf{B}_2^\top\mathbf{v} - \tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v})\| &\le \|\mathbf{D}_{\mathcal{X}}\bar{\mathbf{D}}_{\mathcal{X}}^{-1}\|\left\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{T}_0^\top - \mathbf{T}_{0,j,\mathbf{v}}^\top)\mathbf{I}_{\Lambda_x(\mathbf{v})}\right\|\|\mathbf{v}\| \\ &\le C_\delta^{\frac{1}{2}}Cd^{-\frac{1}{2}}2^{-\frac{J}{2}}\sum_{i=1}^d \|\pi^{(i)}(\mathbf{v})\| \\ &\le CC_\delta^{\frac{1}{2}}d^{-\frac{1}{2}}2^{-\frac{J}{2}}\sum_{i=1}^d \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}, \end{aligned} \tag{4.139}$$

where we used Theorem 4.8 (i), (4.57) and (4.136) in the second line as well as Theorem 2.8 (i) in the last line. □

Since $\mathbf{D}_{\mathcal{X}}$ has unbounded ranks, we need to switch again to the rank-truncated scaling matrix $\mathbf{D}_{\mathcal{X},\mathsf{n}}$. In this case, we only apply a scaling matrix from the left side. Hence, we get the following result, which is conceptually analogous to Theorems 4.38 and 4.52.

Lemma 4.58 ([BF24, Lemma 5.27]). *Let $\mathbf{T}_0$ be defined as in (3.18), let $\tilde{\mathbf{T}} \in \mathbb{R}^{\vee\times\vee_{\mathsf{x}}}$, $\mathbf{v} \in \ell_2(\vee_{\mathsf{x}})$, $\mathsf{K} \in \mathbb{N}_0^{\vee_{\mathsf{t}}}$ be such that $\operatorname{supp}(\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{T}}\mathbf{v}) \subset \Lambda_{\mathsf{K}}$ and let $\mathbf{G} = \mathbf{D}_{\mathcal{X}}(\mathbf{T}_0^\top - \tilde{\mathbf{T}})$. Then whenever $\mathsf{n} \geq \mathsf{M}_{\mathcal{X}}(\eta, \mathsf{K})$, we have*

$$\|(\mathbf{D}_{\mathcal{X}}\mathbf{T}_0^\top - \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{T}})\mathbf{v}\| \leq (1+\eta(1-\tfrac{\delta}{2})^{-1})\|\mathbf{G}\mathbf{v}\| + \eta(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_2\|\|\mathbf{v}\|.$$

Proof. The proof of this lemma is a simplified version of the proof of Theorem 4.38, since we only apply a truncated scaling matrix from the left rather than from both sides. We have

$$\begin{aligned}\|(\mathbf{D}_{\mathcal{X}}\mathbf{T}_0^\top - \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{T}})\mathbf{v}\| &\leq \|\mathbf{G}\mathbf{v}\| + \|(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}})\tilde{\mathbf{T}}\mathbf{v}\| \\ &\leq \|\mathbf{G}\mathbf{v}\| + \|\mathbf{D}_{\mathcal{X}}^{-1}\bar{\mathbf{D}}_{\mathcal{X}}\|\|\bar{\mathbf{D}}_{\mathcal{X}}^{-1}(\mathbf{D}_{\mathcal{X}} - \mathbf{D}_{\mathcal{X},\mathsf{n}})\mathrm{R}_{\Lambda_{\mathsf{K}}}\|\|\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{T}}\mathbf{v}\| \\ &\leq \|\mathbf{G}\mathbf{v}\| + \eta(1-\tfrac{\delta}{2})^{-1}(\|\mathbf{G}\mathbf{v}\| + \|\mathbf{D}_{\mathcal{X}}\mathbf{T}_0^\top\mathbf{v}\|),\end{aligned}$$

where we used (4.57) as well as (4.54) in the last step. Using $\mathbf{B}_2 = \mathbf{D}_{\mathcal{X}}\mathbf{T}_0^\top$ concludes the proof. □

In view of the above lemma, we choose the approximation parameters similarly as for the other operators. For a given tolerance $\eta > 0$ and a finitely supported input $\mathbf{v} \in \ell_2(\vee_{\mathsf{x}})$, we define

$$J_0^{\mathrm{ad}}(\eta;\mathbf{v}) = \min\big\{J \in \mathbb{N}_0 : e_{0,J}^{\mathrm{ad}}(\mathbf{v}) \leq \tfrac{\eta}{4}\big\}, \quad c_0^{\mathrm{ad}}(\eta;\mathbf{v}) = \frac{\eta}{2\|\mathbf{B}_2\|\|\mathbf{v}\|}(1-\tfrac{\delta}{2})$$

and

$$\mathsf{m}_{\mathcal{X},0}^{\mathrm{ad}}(\eta;\mathbf{v}) = \mathsf{M}_{\mathcal{X}}\big(c_0^{\mathrm{ad}}(\eta;\mathbf{v}); K_0^{\mathrm{ad}}\big(J_0^{\mathrm{ad}}(\eta;\mathbf{v});\mathbf{v}\big)\big),$$

where

$$\mathsf{K}_0^{\mathrm{ad}}(J;\mathbf{v}) = \min\big\{\mathsf{K} > 0 : \operatorname{supp}(\mathbf{T}_{0,J,\mathbf{v}}^\top) \subseteq \Lambda_{\mathsf{K}}\big\}.$$

Theorem 4.59. *Let $\mathbf{v} \in \ell_2(\vee_{\mathsf{x}})$ and $0 < \eta < 2\|\mathbf{B}_2\|\|\mathbf{v}\|$. We fix $J = J_0^{\mathrm{ad}}(\eta;\mathbf{v})$ and $\mathsf{n} = \mathsf{m}_{\mathcal{X},0}^{\mathrm{ad}}(\eta;\mathbf{v})$. Then for $\mathbf{w}_\eta = \mathbf{D}_{\mathcal{X},\mathsf{n}}\mathbf{T}_{0,J,\mathbf{v}}^\top\mathbf{v}$, we have*

$$\|\mathbf{B}_2^\top\mathbf{v} - \mathbf{w}_\eta\| \leq \eta.$$

Proof. Using Theorems 4.57 and 4.58, we obtain

$$\begin{aligned}\|\mathbf{B}_2^\top\mathbf{v} - \mathbf{w}_\eta\| &\leq (1 + c_0^{\mathrm{ad}}(\eta;\mathbf{v})(1-\tfrac{\delta}{2})^{-1})\|\mathbf{B}_2^\top\mathbf{v} - \tilde{\mathbf{B}}_{2,J}^{\mathrm{ad}}(\mathbf{v})\| + c_0^{\mathrm{ad}}(\eta;\mathbf{v})(1-\tfrac{\delta}{2})^{-1}\|\mathbf{B}_2\|\|\mathbf{v}\| \\ &\leq 2e_{0,J}^{\mathrm{ad}}(\mathbf{v}) + \tfrac{\eta}{2} \\ &\leq \eta,\end{aligned}$$

where we used the assumption $\eta < 2\|\mathbf{B}_2\|\|\mathbf{v}\|$ in the second line. □

We mention here that by the same argument as in Theorem 4.41, the assumption $\eta < 2\|\mathbf{B}_2\|\|\mathbf{v}\|$ does not pose a limitation, since for larger η we can set $\mathbf{w}_\eta = 0$.

Remark 4.60. For the practical realization, one can also choose

$$J_0^{\mathrm{ad}}(\eta;\mathbf{v}) = \min\left\{J \in \mathbb{N}_0 : C\sqrt{C_\delta d}\,2^{-\frac{J}{2}}\|\mathbf{v}\| \le \tfrac{\eta}{4}\right\}$$

with C from Theorem 4.57. Based on (4.139), the statement of Theorem 4.59 is still valid for this choice of J_0^{ad}. Moreover, by making this choice, we avoid computing the $\mathcal{A}^s$-norms of the contractions of $\mathbf{v}$.

Lemma 4.61. *Under the assumptions of Theorem 4.59, we have*

$$\#\operatorname{supp}\left(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right) \le C_1\big(2L_{\mathrm{x}}(\mathbf{v}) + J_0^{\mathrm{ad}}(\eta;\mathbf{v})\big)\#\operatorname{supp}\left(\pi^{(i)}(\mathbf{v})\right), \tag{4.140}$$

and

$$\left\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right\|_{\mathcal{A}^s} \le C_2\big(2L_{\mathrm{x}}(\mathbf{v}) + J_0^{\mathrm{ad}}(\eta;\mathbf{v})\big)^s\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}, \tag{4.141}$$

as well as

$$J_0^{\mathrm{ad}}(\eta;\mathbf{v}) \le C_3\Big(|\ln(\eta)| + \ln\Big(\sum_{i=1}^d\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}\Big)\Big), \tag{4.142}$$

where the constants $C_1, C_2, C_3 > 0$ depend on δ, s and $\bar{c}$. Moreover, we have the rank estimate

$$\operatorname{rank}_\infty(\mathbf{w}_\eta) \le \left\|\hat{\mathrm{m}}^{\mathrm{ad}}_{\mathcal{X},0}(\eta;\mathbf{v})\right\|_{\ell_\infty}|\operatorname{rank}(\mathbf{v})|_\infty, \tag{4.143}$$

where

$$\hat{\mathrm{m}}^{\mathrm{ad}}_{\mathcal{X},0,\nu_{\mathrm{t}}}(\eta;\mathbf{v}) = 1 + n^+_{a_{\nu_{\mathrm{t}}}}(\delta) + \mathrm{m}^{\mathrm{ad}}_{\mathcal{X},0,\nu_{\mathrm{t}}}(\eta;\mathbf{v}), \quad \nu_{\mathrm{t}} \in \vee_{\mathrm{t}}.$$

Proof. As seen in Section 2.4.5.1, the multiplication with scalar values does not affect the hierarchical ranks. Hence, the rank estimate (4.143) is an immediate consequence of

$$\mathbf{T}^{\top}_{0,J,\mathbf{v}}\mathbf{v} = \sum_{\substack{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}\\ |\nu_{\mathrm{t}}|\le 2L_{\mathrm{x}}(\mathbf{v})+j}} \mathbf{e}_{\nu_{\mathrm{t}}}\otimes\theta_{\nu_{\mathrm{t}}}(0)\mathbf{v}$$

combined with the ranks of the rank-truncated scaling matrix $\mathbf{D}_{\mathcal{X},\mathsf{n}}$.

By the minimality condition of $J = J_0^{\mathrm{ad}}(\eta;\mathbf{v})$ we have

$$CC_\delta^{\frac12}d^{-\frac12}2^{-\frac{J-1}{2}}\sum_{i=1}^d\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s} = e^{\mathrm{ad}}_{0,J-1}(\mathbf{v}) \ge \frac{\eta}{4}.$$

Hence, we have

$$J \le \log_2\Big(4\sqrt{2}CC_\delta^{\frac12}d^{-\frac12}\eta^{-1}\sum_{i=1}^d\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}\Big) \le \tilde{C}\Big(|\ln(\eta)| + \ln\Big(\sum_{i=1}^d\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}\Big)\Big),$$

which yields (4.142). The estimate (4.140) is a direct consequence of (4.138) from Theorem 4.57, since the diagonal operator $\mathbf{D}_{\mathcal{X},\mathsf{n}}$ does not change the support.

We finally show the estimate (4.141). By definition of $\mathbf{w}_\eta$ and $\mathbf{T}_{0,J,\mathbf{v}}$, we know that

$$N_{\mathrm{t}} = \#\operatorname{supp}\left(\pi^{(\mathrm{t})}(\mathbf{w}_\eta)\right) \le \bar{c}(2L_{\mathrm{x}}(\mathbf{v}) + J).$$

Now let $i \in \{1,\dots,d\}$ and $N \in \mathbb{N}$ be arbitrary, and we set $N_{\mathbf{v}} = \lfloor\frac{N}{N_{\mathrm{t}}}\rfloor$. Since $\mathbf{v}$ is finitely supported, there exists an index set Λ_i with

$$\left\|\pi^{(i)}(\mathbf{v}) - \mathrm{R}_{\Lambda_i}\pi^{(i)}(\mathbf{v})\right\| = \min_{\#\hat{\Lambda}\le N_{\mathbf{v}}}\left\|\pi^{(i)}(\mathbf{v}) - \mathrm{R}_{\hat{\Lambda}}\pi^{(i)}(\mathbf{v})\right\|. \tag{4.144}$$

We define the corresponding index set

$$\Lambda = \{\nu_{\mathrm{t}} : |\nu_{\mathrm{t}}| \le 2L_x(\mathbf{v}) + J, \theta_{\nu_{\mathrm{t}}}(0) \neq 0\} \times \Lambda_i.$$

By construction, we have $\#\Lambda \le N_{\mathrm{t}} N_{\mathbf{v}} \le N$. Using (4.124), for a fixed $\nu_{\mathrm{x}} \in \vee_{\mathrm{x}}$, we have

$$\begin{aligned}\sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}} \left|(\mathbf{B}_2^\top)^2_{(\nu_{\mathrm{t}},\nu_{\mathrm{x}}),\nu_{\mathrm{x}}}\right| &\lesssim \sum_{\substack{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}\\ \theta_{\nu_{\mathrm{t}}}(0)\neq 0}} 2^{-\left||\nu_{\mathrm{t}}|-\log_2\left(\|\Psi_{\nu_{\mathrm{x}}}\|_{H^1_0}\right)\right|}\\ &\le 2\tilde{c}\sum_{k=0}^{\infty} 2^{-j}\\ &= 4\tilde{c}.\end{aligned} \tag{4.145}$$

Furthermore, by construction of Λ and the identity in the spatial variable of $\mathbf{B}_2$, we have

$$\begin{aligned}\left\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta) - \mathrm{R}_\Lambda \pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right\|^2 &= \sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}}\sum_{\nu_i\notin\Lambda_i} \pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_i}(\mathbf{w}_\eta)^2\\ &\le (1-\tfrac{\delta}{2})^{-2}\sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}}\sum_{\nu_i\notin\Lambda_i} \pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_i}(\bar{\mathbf{D}}\mathcal{X}\mathbf{T}^\top_{0,J,\mathbf{v}}\mathbf{v})^2\\ &= (1-\tfrac{\delta}{2})^{-2}\sum_{|\nu_{\mathrm{t}}|\le 2L_{\mathrm{x}}+J}\sum_{\nu_i\notin\Lambda_i} \pi^{(\mathrm{t},i)}_{\nu_{\mathrm{t}},\nu_i}(\bar{\mathbf{B}}_2^\top\mathbf{v})^2\\ &\le (1-\tfrac{\delta}{2})^{-2}\sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}}\sum_{\nu_i\notin\Lambda_i}\sum_{\substack{\mu_{\mathrm{x}}\in\vee_{\mathrm{x}}\\ \mu_i=\nu_i}} (\bar{\mathbf{B}}_2^\top\mathbf{v})^2_{\nu_{\mathrm{t}},\mu_{\mathrm{x}}}\\ &= (1-\tfrac{\delta}{2})^{-2}\sum_{\nu_i\notin\Lambda_i}\sum_{\substack{\mu_{\mathrm{x}}\in\vee_{\mathrm{x}}\\ \mu_i=\nu_i}} (\mathbf{v}_{\mu_{\mathrm{x}}})^2\sum_{\nu_{\mathrm{t}}\in\vee_{\mathrm{t}}}(\bar{\mathbf{B}}_2^\top)^2_{(\nu_{\mathrm{t}},\mu_{\mathrm{x}}),\mu_{\mathrm{x}}}\\ &\le C(1-\tfrac{\delta}{2})^{-2}\sum_{\nu_i\notin\Lambda_i}\pi^{(i)}_{\nu_i}(\mathbf{v})^2\\ &= C(1-\tfrac{\delta}{2})^{-2}\|\pi^{(i)}(\mathbf{v}) - \mathrm{R}_{\Lambda_i}\pi^{(i)}(\mathbf{v})\|^2,\end{aligned}$$

where we have used (4.56) and (4.57) in the second line, the definition of $\mathbf{T}^\top_{0,J,\mathbf{v}}$ in the third line, the definition of $\bar{\mathbf{B}}_2$ with the identity in the spatial variable in the fifth line and (4.145) in the sixth line. Hence, we obtain

$$\begin{aligned}(N+1)^s \inf_{\#\hat{\Lambda}\le N}\left\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta) - \mathrm{R}_{\hat{\Lambda}}\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right\| &\le 2^s N^s \left\|\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta) - \mathrm{R}_\Lambda\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\right\|\\ &\le \frac{2^s\sqrt{C}}{1-\frac{\delta}{2}} N_{\mathrm{t}}^s (N_{\mathbf{v}}+1)^s \|\pi^{(i)}(\mathbf{v}) - \mathrm{R}_{\Lambda_i}\pi^{(i)}(\mathbf{v})\|\\ &\le \frac{2^s\bar{c}^s\sqrt{C}}{1-\frac{\delta}{2}} x(2L_{\mathrm{x}}(\mathbf{v})+J)^s \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s},\end{aligned}$$

where we have used (4.144) and the definition of the $\mathcal{A}^s$ norm in the final step. Since the right-hand side is independent of N, we obtain (4.141). □

In the next step, we analyze the computational complexity of the adaptive application of $\mathbf{B}_2^\top$.

Lemma 4.62. *Under the assumptions of Theorem 4.59, for the number* $\operatorname{flops}(\mathbf{w}_\eta)$ *of floating point operations required to compute* $\mathbf{w}_\eta$ *we have the bound*

$$\begin{aligned}\operatorname{flops}(\mathbf{w}_\eta) \lesssim (2L_{\mathrm{x}}(\mathbf{v}) + J_0^{\mathrm{ad}}(\eta;\mathbf{v}))|\operatorname{rank}(\mathbf{v})|_\infty \big\|\hat{\mathrm{m}}^{\mathrm{ad}}_{\mathcal{X},0}(\eta;\mathbf{v})\big\|_{\ell_\infty} \sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(i)}(\mathbf{v})\big)\\ + d(2L_{\mathrm{x}}(\mathbf{v}) + J_0^{\mathrm{ad}}(\eta;\mathbf{v}))|\operatorname{rank}(\mathbf{v})|_\infty^3 \big\|\hat{\mathrm{m}}^{\mathrm{ad}}_{\mathcal{X},0}(\eta;\mathbf{v})\big\|^3_{\ell_\infty}, \quad (4.146)\end{aligned}$$

where the hidden constant is independent of $\eta, \mathbf{v}$ *and* d.

Proof. The entries in the core tensors and mode frames can each be computed in constant time, since they are determined solely by applying the entries of the diagonal scaling matrix and the temporal wavelet values at $t = 0$. Hence, we can estimate the computational costs by the number of entries in the hierarchical tensor representation of $\mathbf{w}_\eta$. Consequently, we obtain

$$\operatorname{flops}(\mathbf{w}_\eta) \lesssim \operatorname{rank}_\infty(\mathbf{w}_\eta) \sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{w}_\eta)\big) + d\#\operatorname{supp}\big(\pi^{(\mathrm{t})}(\mathbf{w}_\eta)\big)\operatorname{rank}^3_\infty(\mathbf{w}_\eta).$$

Combining this estimate with Theorem 4.61 and the definition of the operator $\mathbf{T}^\top_{0,J,\mathbf{v}}$ yields (4.146). □

Lemma 4.63. *Under the assumptions of Theorem 4.59, we have the estimate*

$$\big\|\hat{\mathrm{m}}^{\mathrm{ad}}_{\mathcal{X},0}(\eta;\mathbf{v})\big\|_{\ell_\infty} \le C(\delta,\mathbf{B}_2)\Big(1 + L_{\mathrm{x}}(\mathbf{v}) + |\ln(\eta)| + \ln\Big(\sum_{i=1}^{d}\|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}\Big)\Big)^2.$$

Proof. Let $\mathsf{K} = K_0^{\mathrm{ad}}(J_0^{\mathrm{ad}}(\eta;\mathbf{v});\mathbf{v})$. By the minimality property of K and Theorem 4.47, we obtain the estimate

$$\begin{aligned}\tfrac{1}{2}\log_2(\mathsf{K}_{\nu_\mathrm{t}}) &\le L_{\mathrm{x}}((\mathbf{T}^\top_{0,J_0^{\mathrm{ad}}(\eta;\mathbf{v}),\mathbf{v}}\mathbf{v}) + \log_2(cs^{(1)}_{\min})\\ &\le L_{\mathrm{x}}(\mathbf{v}) + \log_2(cs^{(1)}_{\min}),\end{aligned}$$

with $s^{(1)}_{\min}$ defined in (4.117), where we have used that $\mathbf{T}_0$ is of the form (4.122). Proceeding as in Theorems 4.47 and 4.48, we obtain

$$\mathrm{m}^{\mathrm{ad}}_{\mathcal{X},0,\nu_\mathrm{t}}(\eta;\mathbf{v}) \le C_2(\delta,s,\mathbf{B}_2)\big(1 + \max\{L_{\mathrm{x}}(\mathbf{v}),|\nu_\mathrm{t}|\} + |\ln\eta| + \max\{0,\ln(\|\mathbf{v}\|)\}\big)^2.$$

Furthermore, by construction of the adaptive approximate application, we only need to consider the temporal indices ν_t with $|\nu_\mathrm{t}| \le 2L_{\mathrm{x}}(\mathbf{v}) + J_0^{\mathrm{ad}}(\eta;\mathbf{v})$. Combining this with $\|\mathbf{v}\| = \|\pi^{(i)}(\mathbf{v})\| \le \|\pi^{(i)}(\mathbf{v})\|_{\mathcal{A}^s}$ and the estimate for $J_0^{\mathrm{ad}}(\eta;\mathbf{v})$ in Theorem 4.61 concludes the proof. □

Remark 4.64. The estimates for the quantities of the adaptive application of $\mathbf{B}_2^\top$ are mostly comparable to the ones for the adaptive application routine described in [BF24, Section 5.4]. In the approach of this work, we obtain a better estimate for the arising ranks, since we only apply one fixed approximation of the operator. However, we have slightly different estimates for the support in (4.140) and for the approximability in (4.141). In comparison to the estimates in [BF24, Lemma 5.29], the support depends on the support of the input $\mathbf{v}$ but only logarithmically on the tolerance η. However, the approximability in the $\mathcal{A}^s$-norm in (4.141) depends additionally on the tolerance η through $J_0^{\mathrm{ad}}(\eta;\mathbf{v})$.

4.4 Space-Time Adaptive Richardson Iteration

In this section, we analyze an iterative scheme based on the Richardson iteration to approximate the solution of the bi-infinite matrix equation from the space-time variational formulation of the parabolic PDE in Section 3.1. As already mentioned at the beginning of the chapter, the Richardson iteration for solving a linear equation $\tilde{\mathbf{A}}\mathbf{u} = \tilde{\mathbf{f}}$ for a symmetric positive definite $\tilde{\mathbf{A}}$ is given by

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \omega(\tilde{\mathbf{A}}\mathbf{u}_k - \tilde{\mathbf{f}}).$$

For a suitable choice of the relaxation parameter $\omega > 0$ such that

$$\|\mathbf{I} - \omega\tilde{\mathbf{A}}\| \leq \rho \in (0,1),$$

we obtain the error reduction

$$\|\mathbf{u} - \mathbf{u}_{k+1}\| \leq \rho\|\mathbf{u} - \mathbf{u}_k\| \tag{4.147}$$

independently of k. However, for bi-infinite matrix equations, $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{A}}\mathbf{u}_k$ are generally not finitely supported, even if $\mathbf{u}_k$ itself is. To address this, we use an approximate Richardson iteration. Here, the operator is applied adaptively as described in Section 4.3, and the right-hand side is approximated adaptively as well. Moreover, we compress and coarsen the iterates, as described in Section 4.1, to keep ranks and supports small. We combine all these approximations with suitable tolerances to ensure convergence, as in (4.147).

For this method, we analyze its convergence and computational complexity. Especially for the latter, we are interested in the dependence on the desired accuracy and the number of spatial dimensions d. The goal of this section is to show that the proposed method can circumvent the curse of dimensionality under suitable assumptions, meaning that the computational complexity does not grow exponentially with respect to d.

4.4.1 A Convergent Solver

For the construction of the method, we mainly follow [BF24], which again is a modified version of the construction for elliptic problems in [BD15; BD16]. The following construction differs from the one in [BF24] based on the different treatment of the initial value operator in Section 4.3.4. The algorithm, stated in Algorithm 4.1, is based on [BD15, Algorithm 1]. Differences arise from the modified implementations of the used subroutines APPLY, COARSEN, RECOMPRESS, and RHS. The algorithm consists of an outer and an inner loop. For each outer loop k, we start with an approximation $\mathbf{u}_k$ of the solution $\mathbf{u}$ with $\|\mathbf{u} - \mathbf{u}_k\| \leq 2^{-k}\varepsilon_0$. Then we perform a fixed number of inexact Richardson iteration steps in the inner loop to reduce the error to $\kappa_1 2^{-(k+1)}\varepsilon_0$. Then we apply the RECOMPRESS and COARSEN routines with suitable tolerances to obtain quasi-optimal ranks and supports. The tolerances of the two routines are chosen such that

$$\|\mathbf{u} - \mathbf{u}_{k+1}\| \leq (\kappa_1 + \kappa_2 + \kappa_3)2^{-(k+1)}\varepsilon_0 \leq 2^{-(k+1)}\varepsilon_0.$$

A visualization of the iteration is given in Figure 4.4. Here, the black arrows indicate the inner iteration steps $\mathbf{w}_{k,j}$ and the red arrows indicate the outer iteration steps for $\mathbf{u}_k$ by applying recompression and coarsening. The node sizes reflect the rank and support size of the corresponding tensor, which rise during the inner iteration and drop after each outer iteration step.

Algorithm 4.1: solve$(\mathbf{A}, \mathbf{f}, \varepsilon) = \mathbf{u}_\varepsilon$

Input: $\omega > 0$ and $\rho \in (0,1)$ s.t. $\|\mathbf{I} - \omega\mathbf{A}\| \le \rho$, $c_\mathbf{A} \ge \|\mathbf{A}^{-1}\|, \varepsilon_0 \ge c_\mathbf{A}\|\mathbf{B}^\top \mathbf{f}\|$, $\kappa_1, \kappa_2, \kappa_3 \in (0,1)$ with $\kappa_1 + \kappa_2 + \kappa_3 \le 1$ and $\beta_1 \ge 0, \beta_2 > 0$.

Output: $\mathbf{u}_\varepsilon$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \le \varepsilon$.

begin
  $\mathbf{u}_0 = 0$
  $k = 0$
  $I = \min\left\{j : \rho^j(1 + (\omega + \beta_1 + \beta_2)j) \le \frac{1}{2}\kappa_1\right\}$
  while $2^{-k}\varepsilon_0 > \varepsilon$ **do**
    $\mathbf{w}_{k,0} = \mathbf{u}_k$
    $j = 0$
    repeat
      $\eta_{k,j} = \rho^{j+1}2^{-k}\varepsilon_0$
      $\mathbf{r}_{k,j} = \text{APPLY}\left(\mathbf{w}_{k,j}, \frac{1}{2}\eta_{k,j}\right) - \text{RHS}\left(\frac{1}{2}\eta_{k,j}\right)$
      $\mathbf{w}_{k,j+1} = \text{COARSEN}\left(\text{RECOMPRESS}\left(\mathbf{w}_{k,j} - \omega\mathbf{r}_{k,j}, \beta_1\eta_{k,j}\right), \beta_2\eta_{k,j}\right)$
      $j \mathrel{+}= 1$
    until $j \ge I$ or $c_\mathbf{A}\rho\|\mathbf{r}_{k,j-1}\| + (c_\mathbf{A}\rho + \omega + \beta_1 + \beta_2)\eta_{k,j-1} \le \kappa_1 2^{-(k+1)}\varepsilon_0$;
    $\mathbf{u}_{k+1} = \text{COARSEN}\left(\text{RECOMPRESS}\left(\mathbf{w}_{k,j}, \kappa_2 2^{-(k+1)}\varepsilon_0\right), \kappa_3 2^{-(k+1)}\varepsilon_0\right)$
    $k \mathrel{+}= 1$
  end
  $\mathbf{u}_\varepsilon = \mathbf{u}_k$
end

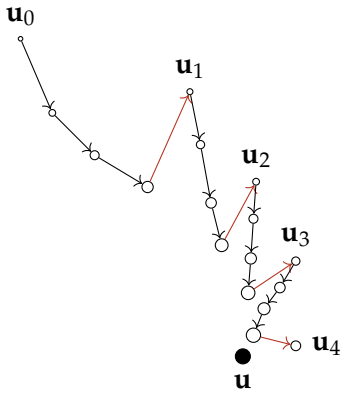


FIGURE 4.4: Visualization of the inexact Richardson iteration in Algorithm 4.1. The black arrows visualize the inner iteration steps. The red arrows indicate the recompression and coarsening step at the end of each outer iteration.

The realization of the subroutines COARSEN and RECOMPRESS are based on the routines described in Section 4.1 and fulfil

$$\begin{aligned} \|\mathbf{v} - \text{RECOMPRESS}(\mathbf{v}, \eta)\| &\le \eta, \\ \|\mathbf{v} - \text{COARSEN}(\mathbf{v}, \eta)\| &\le \eta. \end{aligned} \tag{4.148}$$

The APPLY routine is the adaptive approximation of the operator described in Section 4.3. Since the operator $\mathbf{B}$ given in (4.35) is non-symmetric with unequal domain

and codomain, we consider, as in [SS09], the equivalent least squares formulation

$$\mathbf{B}^\top \mathbf{B}\mathbf{u} = \mathbf{B}^\top \mathbf{f}. \tag{4.149}$$

Setting

$$\mathbf{A} = \mathbf{B}^\top \mathbf{B} = \mathbf{B}_1^\top \mathbf{B}_1 + \mathbf{B}_2^\top \mathbf{B}_2 = \mathbf{A}_1 + \mathbf{A}_2,$$

the operator $\mathbf{A}$ is symmetric, bounded and elliptic on $\ell_2(\vee)$, since $\|\mathbf{A}\| = \|\mathbf{B}\|^2$ and

$$\langle \mathbf{A}\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{B}\mathbf{v}, \mathbf{B}\mathbf{v}\rangle \geq \|\mathbf{B}^{-1}\|^{-2}\|\mathbf{v}\|^2.$$

Next, we define the adaptive application of the operator $\mathbf{A}$. Here, we handle the two parts $\mathbf{A}_1$ and $\mathbf{A}_2$ separately, since in the previous section we derived an adaptive application scheme for $\mathbf{A}_2$ directly. However, for $\mathbf{A}_1$ we only have an adaptive application scheme for the two parts $\mathbf{B}_1$ and $\mathbf{B}_1^\top$ separately. We can define the adaptive application of $\mathbf{A}_1$ to an input $\mathbf{w} \in \ell_2(\vee)$ as in [SS09; CDD02] by

$$\operatorname{APPLY}_{\mathbf{A}_1}(\mathbf{v}, \eta) = \operatorname{APPLY}_{\mathbf{B}_1^\top}\big(\operatorname{APPLY}_{\mathbf{B}_1}\big(\mathbf{v}, \tfrac{\eta}{2\|\mathbf{B}_1\|}\big), \tfrac{\eta}{2}\big), \tag{4.150}$$

where $\operatorname{APPLY}_{\mathbf{B}_1}(\mathbf{v}, \eta)$ is the realization of $\mathbf{w}_\eta$ of Theorem 4.40 for the operator $\mathbf{B}_1$ and correspondingly $\operatorname{APPLY}_{\mathbf{B}_1^\top}(\mathbf{v}, \eta)$ for the operator $\mathbf{B}_1^\top$, which we can do in the same way since the adjoint of an s^*-compressible operator is s^*-compressible and the adjoint of a super-compressible operator is super-compressible. As a direct consequence of the triangle inequality, with $\mathbf{z} = \operatorname{APPLY}_{\mathbf{B}_1}\big(\mathbf{v}, \frac{\eta}{2\|\mathbf{B}_1\|}\big)$ we have

$$\begin{aligned}
&\|\mathbf{A}_1\mathbf{v} - \operatorname{APPLY}_{\mathbf{A}_1}(\mathbf{v}, \eta)\| \\
&\quad \leq \big\|\mathbf{B}_1^\top\big(\mathbf{B}_1\mathbf{v} - \operatorname{APPLY}_{\mathbf{B}_1}(\mathbf{v}, \tfrac{\eta}{2\|\mathbf{B}_1\|})\big)\big\| + \big\|\mathbf{B}_1^\top\mathbf{z} - \operatorname{APPLY}_{\mathbf{B}_1^\top}\big(\mathbf{z}, \tfrac{\eta}{2}\big)\big\| \\
&\quad \leq \|\mathbf{B}_1\|\tfrac{\eta}{2\|\mathbf{B}_1\|} + \tfrac{\eta}{2} \\
&\quad = \eta.
\end{aligned} \tag{4.151}$$

Furthermore, let $\operatorname{APPLY}_{\mathbf{A}_2}(\mathbf{v}, \eta)$ denote the realization of $\mathbf{w}_\eta$ from Theorem 4.53 for the operator $\mathbf{A}_2$. With these two ingredients, we can define the adaptive application of $\mathbf{A}$ with tolerance $\eta > 0$ by

$$\operatorname{APPLY}(\mathbf{v}, \eta) = \operatorname{APPLY}_{\mathbf{A}}(\mathbf{v}, \eta) = \operatorname{APPLY}_{\mathbf{A}_1}(\mathbf{v}, \eta_1) + \operatorname{APPLY}_{\mathbf{A}_2}(\mathbf{v}, \eta_2), \tag{4.152}$$

where $\eta_1 + \eta_2 = \eta$. By the triangle inequality and the previous observations, we obtain

$$\|\mathbf{A}\mathbf{v} - \operatorname{APPLY}(\mathbf{v}, \eta)\| \leq \eta_1 + \eta_2 = \eta. \tag{4.153}$$

Remark 4.65. In Algorithm 4.1, we use the APPLY routine for the operator $\mathbf{A}$ defined above in each step with different tolerances $\eta_{k,j}$. As described in (4.152), we have to split the tolerances $\eta_{k,j}$ into two parts $\eta_{k,j} = \eta_{k,j}^{(1)} + \eta_{k,j}^{(2)}$ for the two parts $\mathbf{A}_1$ and $\mathbf{A}_2$ of the operator. In general, we could use different splittings for each call of the APPLY routine. Nevertheless, for the analysis we assume that the ratio of the tolerances is fixed, i.e. $\eta_{k,j}^{(1)} = \lambda\eta_{k,j}$ and $\eta_{k,j}^{(2)} = (1-\lambda)\eta_{k,j}$ for some fixed $\lambda \in (0,1)$.

The routine for approximating the right-hand side

$$\mathbf{B}^\top \mathbf{f} = \mathbf{B}_1^\top \mathbf{f}_1 + \mathbf{B}_2^\top \mathbf{f}_2$$

is given by

$$\begin{aligned}\mathrm{RHS}(\eta) &= \mathrm{RHS}_{\mathbf{B}^\top \mathbf{f}}(\eta) \\ &= \mathrm{APPLY}_{\mathbf{B}_1^\top}\Big(\mathrm{RHS}_{\mathbf{f}_1}\Big(\tfrac{\eta_1}{2\|\mathbf{B}_1\|}\Big), \tfrac{\eta_1}{2}\Big) + \mathrm{APPLY}_{\mathbf{B}_2^\top}\Big(\mathrm{RHS}_{\mathbf{f}_2}\Big(\tfrac{\eta_2}{2\|\mathbf{B}_2\|}\Big), \tfrac{\eta_2}{2}\Big)\end{aligned} \tag{4.154}$$

with $\eta_1 + \eta_2 = \eta$, where $\mathrm{RHS}_{\mathbf{f}_1}(\eta)$ and $\mathrm{RHS}_{\mathbf{f}_2}(\eta)$ are routines fulfilling

$$\begin{aligned}\|\mathbf{f}_1 - \mathrm{RHS}_{\mathbf{f}_1}(\eta)\|_{\ell_2(\vee)} &\le \eta, \\ \|\mathbf{f}_2 - \mathrm{RHS}_{\mathbf{f}_2}(\eta)\|_{\ell_2(\vee_{\mathrm{x}})} &\le \eta.\end{aligned} \tag{4.155}$$

Following the same arguments as in (4.151), we obtain

$$\begin{aligned}\Big\|\mathbf{B}_1^\top \mathbf{f}_1 - \mathrm{APPLY}_{\mathbf{B}_1^\top}\Big(\mathrm{RHS}_{\mathbf{f}_1}\Big(\tfrac{\eta_1}{2\|\mathbf{B}_1\|}\Big), \tfrac{\eta_1}{2}\Big)\Big\| &\le \eta_1, \\ \Big\|\mathbf{B}_2^\top \mathbf{f}_2 - \mathrm{APPLY}_{\mathbf{B}_2^\top}\Big(\mathrm{RHS}_{\mathbf{f}_2}\Big(\tfrac{\eta_2}{2\|\mathbf{B}_2\|}\Big), \tfrac{\eta_2}{2}\Big)\Big\| &\le \eta_2,\end{aligned}$$

which yields

$$\|\mathbf{B}^\top \mathbf{f} - \mathrm{RHS}(\eta)\| \le \eta_1 + \eta_2 = \eta. \tag{4.156}$$

In the following statement, we show the convergence of the adaptive Richardson iteration in Algorithm 4.1, which is a special case of [BD15, Proposition 5].

Proposition 4.66. *Let the conditions of Algorithm 4.1 be fulfilled. Then the intermediate results $\mathbf{u}_k$ satisfy $\|\mathbf{u}_k - \mathbf{u}\| \le 2^{-k}\varepsilon_0$, and in particular $\|\mathbf{u}_\varepsilon - \mathbf{u}\| \le \varepsilon$.*

Proof. The proof can be done in the same way as in [BD15, Proposition 5] since the proof is not based on the practical realization of the four subroutines COARSEN, RECOMPRESS, RHS, and APPLY, but only on their error properties.

We show the statement by induction. By definition of ε_0 and $c_{\mathbf{A}}$, the statement is true for $k = 0$. The idea for the induction step is to show that after the end of the inner loop, we have the error bound

$$\|\mathbf{w}_{k,j} - \mathbf{u}\| \le \kappa_1 2^{-(k+1)}\varepsilon_0.$$

Then the result is a direct consequence of the properties of the RECOMPRESS and COARSEN routines as well as $\kappa_1 + \kappa_2 + \kappa_3 \le 1$.

By definition of $\mathbf{r}_{k,j}$ and the triangle inequality, we have

$$\|\mathbf{A}\mathbf{w}_{k,j} - \mathbf{B}^\top \mathbf{f} - \mathbf{r}_{k,j}\| \le \big\|\mathbf{A}\mathbf{w}_{k,j} - \mathrm{APPLY}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}\big)\big\| + \big\|\mathbf{B}^\top \mathbf{f} - \mathrm{RHS}\big(\tfrac{1}{2}\eta_{k,j}\big)\big\| \le \eta_{k,j}.$$

Hence, we obtain

$$\begin{aligned}\|\mathbf{w}_{k,j+1} - \mathbf{u}\| &= \|\mathrm{COARSEN}\left(\mathrm{RECOMPRESS}\left(\mathbf{w}_{k,j} - \omega\mathbf{r}_{k,j}, \beta_1\eta_{k,j}\right), \beta_2\eta_{k,j}\right) - \mathbf{u}\| \\ &\le \|\mathbf{w}_{k,j} - \omega\mathbf{r}_{k,j} - \mathbf{u}\| + (\beta_1 + \beta_2)\eta_{k,j} \\ &\le \|(\mathbf{I} - \omega\mathbf{A})(\mathbf{w}_{k,j} - \mathbf{u})\| + \omega\|\mathbf{A}\mathbf{w}_{k,j} - \mathbf{B}^\top \mathbf{f} - \mathbf{r}_{k,j}\| + (\beta_1 + \beta_2)\eta_{k,j} \\ &\le \rho\|\mathbf{w}_{k,j} - \mathbf{u}\| + (\omega + \beta_1 + \beta_2)\eta_{k,j},\end{aligned}$$

where we used (4.148) in the second line as well as the estimates (4.153) and (4.156) in the last line. If the inner loop terminates by the first condition $j \ge I$, we obtain by

recursively applying the previous estimate that

$$\begin{aligned}\|\mathbf{w}_{k,I} - \mathbf{u}\| &\leq \rho^I \|\mathbf{w}_{k,0} - \mathbf{u}\| + (\omega + \beta_1 + \beta_2) \sum_{j=0}^{I-1} \rho^{I-1-j} \eta_{k,j} \\ &\leq \rho^I 2^{-k} \varepsilon_0 + (\omega + \beta_1 + \beta_2) I \rho^I 2^{-k} \varepsilon_0 \\ &\leq \kappa_1 2^{-(k+1)} \varepsilon_0,\end{aligned}$$

where we used the induction hypothesis and the definition of $\eta_{k,j} = \rho^{j+1} 2^{-k} \varepsilon_0$ in the second line and the choice of I in line 4 of Algorithm 4.1 in the last line.

If the inner loop terminates by the second condition, we have

$$\begin{aligned}\|\mathbf{w}_{k,j} - \mathbf{u}\| &\leq \rho \|\mathbf{w}_{k,j-1} - \mathbf{u}\| + (\omega + \beta_1 + \beta_2) \eta_{k,j-1} \\ &\leq c_{\mathbf{A}} \rho \|\mathbf{A}\mathbf{w}_{k,j-1} - \mathbf{B}^\top \mathbf{f} - \mathbf{r}_{k,j-1} + \mathbf{r}_{k,j-1}\| + (\omega + \beta_1 + \beta_2) \eta_{k,j-1} \\ &\leq c_{\mathbf{A}} \rho \|\mathbf{r}_{k,j-1}\| + (c_{\mathbf{A}} \rho + \omega + \beta_1 + \beta_2) \eta_{k,j-1} \\ &\leq \kappa_1 2^{-(k+1)} \varepsilon_0,\end{aligned}$$

where we used the termination condition for the inner loop in the last line. We have thus shown the first statement by induction.

The outer loop terminates for $2^{-k} \varepsilon_0 < \varepsilon$. Hence,

$$\|\mathbf{u}_\varepsilon - \mathbf{u}\| = \|\mathbf{u}_k - \mathbf{u}\| \leq 2^{-k} \varepsilon_0 \leq \varepsilon,$$

which concludes the proof. □

A similar proof for the one-dimensional case without low-rank structure and therefore without RECOMPRESS can be found in [CDD02].

4.4.2 Computational Complexity

As seen in Theorem 4.66, quite general assumptions guarantee the convergence of the adaptive approximate Richardson iteration. However, to show that the method can circumvent the curse of dimensionality, we need further assumptions on the solution $\mathbf{u}$, the data $\mathbf{f}$, and the components of the bi-infinite matrix $\mathbf{B}$. Furthermore, we have to specify the parameters $\kappa_1, \kappa_2, \kappa_3$ in Algorithm 4.1.

The following assumptions are similar to the ones in [BF24] with minor modifications based on the different treatment of the initial value operator in Section 4.3.4. The assumptions below are stated for the operators corresponding to the heat equation. In particular, we state assumptions for the one-dimensional parts of the spatial operator $\mathbf{T}_x$ from (2.24) related to the Laplacian. However, the assumptions can be adapted to more general parabolic problems, for example, the second-order elliptic operators discussed in Theorem 4.32.

Assumptions 1. *Let $s^*, \tau > 0$ be fixed. We require the following properties for the matrices related to $\mathbf{B}$ and the right-hand side $\mathbf{f}$ defined in (4.35).*

(i) *The one-dimensional operator $\mathbf{C}_2 = \bar{\mathbf{D}}_1 \mathbf{T}_2 \bar{\mathbf{D}}_1$ is s^*-compressible according to Theorem 2.23 with the level decay property from Theorem 2.28 with the corresponding approximations $\mathbf{C}_{2,j} = \bar{\mathbf{D}}_1 \mathbf{T}_{2,j} \bar{\mathbf{D}}_1$. Additionally, the operator has* Sobolev stability of order τ, *that is,*

$$\|\bar{\mathbf{D}}_1^{-\tau} (\mathbf{C}_2 - \mathbf{C}_{2,j}) \bar{\mathbf{D}}_1^{\tau}\| < C_{\tau,2} \beta_j(\mathbf{C}_2)$$

for some constant $C_{\tau,2} > 0$ and all $j \in \mathbb{N}_0$.

(ii) *The one-dimensional operator* $\mathbf{C}_\mathrm{t} = \mathbf{B}_\mathrm{t}\bar{\mathbf{D}}_\mathrm{t}$ *is super-compressible according to Theorem 2.25 with the level decay property and approximations* $\mathbf{C}_{\mathrm{t},j} = \mathbf{T}_{\mathrm{t},j}\bar{\mathbf{D}}_\mathrm{t}$. *Additionally, the operator has Sobolev stability of order* τ, *that is,*

$$\|\bar{\mathbf{D}}_\mathrm{t}^{-\tau}(\mathbf{C}_\mathrm{t} - \mathbf{C}_{\mathrm{t},j})\bar{\mathbf{D}}_\mathrm{t}^{\tau}\| < C_{\tau,\mathrm{t}}\beta_j(\mathbf{C}_\mathrm{t})$$

for some constant $C_{\tau,\mathrm{t}} > 0$ *and all* $j \in \mathbb{N}_0$.

(iii) *The high-dimensional operator* $\mathbf{A}_2 = \mathbf{D}_\mathcal{X}\mathbf{S}_0\mathbf{D}_\mathcal{X}$ *is super-compressible with the level decay property and approximations* $\mathbf{A}_{2,j} = \mathbf{D}_\mathcal{X}\mathbf{S}_{0,j}\mathbf{D}_\mathcal{X}$. *Additionally, the operator has Sobolev stability of order* τ, *that is,*

$$\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_\mathrm{x})(\mathbf{A}_2 - \mathbf{A}_{2,j})(\bar{\mathbf{D}}_\mathrm{t}^{\tau}\otimes\mathbf{I}_\mathrm{x})\| < C_{\tau,0}\beta_j^{(0)}$$

for some constant $C_{\tau,0} > 0$ *and all* $j \in \mathbb{N}_0$.

(iv) *The number of operations required for evaluating each entry in the approximations* $\mathbf{T}_{2,j}$ *and* $\mathbf{T}_{\mathrm{t},j}$ *is uniformly bounded.*

(v) *We have an estimate* $c_\mathbf{A} \geq \|\mathbf{A}^{-1}\|$ *with* $c_\mathbf{A} \lesssim \|\mathbf{A}^{-1}\|$, *and the initial error estimate* ε_0 *satisfies* $\varepsilon_0 \eqsim \|\mathbf{A}^{-1}\|\|\mathbf{B}^\top\mathbf{f}\|$.

(vi) *The contractions of* $\mathbf{f}_1$ *and* $\mathbf{f}_2$ *are compressible, that is,* $\pi^{(\mathrm{t},i)}(\mathbf{f}_1) \in \mathcal{A}^s(\vee_\mathrm{t}\times\vee_1)$ *and* $\pi^{(i)}(\mathbf{f}_2) \in \mathcal{A}^s(\vee_1)$, $i = 1,\dots,d$, *for any* s *with* $0 < s < s^*$.

(vii) *For* $\mathbf{T}_2, \mathbf{T}_\mathrm{t}, \mathbf{S}_0$ *and* $\mathbf{f} = [\mathbf{D}_\mathcal{Y}\mathbf{g}_1, \mathbf{g}_2]$, *we assume*

$$\begin{aligned}
\|\bar{\mathbf{D}}_1^{1-\tau}\mathbf{T}_2\bar{\mathbf{D}}_1^{1+\tau}\| &< \infty, & \|\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\mathbf{T}_\mathrm{t}\bar{\mathbf{D}}_\mathrm{t}^{1+\tau}\| &< \infty,\\
\|\bar{\mathbf{D}}_\mathrm{t}^{1-\tau}\mathbf{T}_\mathrm{t}^\top\bar{\mathbf{D}}_\mathrm{t}^{\tau}\| &< \infty, & \|\bar{\mathbf{D}}^{-2\tau}\mathbf{g}_2\| &< \infty,\\
\|\bar{\mathbf{D}}_\mathcal{Y}^{1-\tau}\mathbf{g}_1\| &< \infty, & \|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_\mathrm{x})\bar{\mathbf{D}}_\mathcal{Y}\mathbf{g}_1\| &< \infty,
\end{aligned}$$

and

$$\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_\mathrm{x})\bar{\mathbf{D}}_\mathcal{X}\mathbf{S}_0\bar{\mathbf{D}}_\mathcal{X}(\bar{\mathbf{D}}_\mathrm{t}^{\tau}\otimes\mathbf{I}_\mathrm{x})\| < \infty \tag{4.157}$$

for some $\tau > 0$ *to which we refer to as* excess regularity of order τ.

The compressibility and the level decay properties are satisfied for a suitable choice of wavelet bases by Section 2.5 and Section 4.3.4. Similar to the elliptic case [BD16, Remark 19], the Sobolev stability of order τ is satisfied if the wavelet basis is sufficiently regular to be a Riesz basis also for $H^{1+s}(\Omega)$ for some $s > \tau > 0$ with appropriate renormalization, and if g, h from (3.3) fulfill $g \in L_2(0,T;H^{-1+\tau}(\Omega))$ and $h \in H^{2\tau}(\Omega)$.

Furthermore, by choosing spline wavelets, such as the multiwavelets from Section 2.1.3, we can calculate the entries of $\mathbf{T}_2$ defined in (2.25) and $\mathbf{T}_\mathrm{t}$ in (3.28) exactly in constant time using quadrature rules, since the entries are products of piecewise polynomials.

Remark 4.67. The bound (4.157) in Assumption 1 (vii) is in particular true for $\tau < \frac{1}{2}$. Following the same arguments as in (4.127) we have

$$\big((\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_\mathrm{x})\bar{\mathbf{D}}_\mathcal{X}\mathbf{S}_0\bar{\mathbf{D}}_\mathcal{X}(\bar{\mathbf{D}}_\mathrm{t}^{\tau}\otimes\mathbf{I}_\mathrm{x})\big)_{(\mu_\mathrm{t},\nu_\mathrm{x}),(\nu_\mathrm{t},\nu_\mathrm{x})} \lesssim 2^{-\left(\frac{1}{2}-\tau\right)\||\nu_\mathrm{t}|-|\mu_\mathrm{t}|\|}.$$

Combining this with Theorem 2.29 and the fact that only a fixed number of wavelets do not vanish at $t = 0$ per level, we obtain the desired result. In the same manner, for

fixed $0 < \tau < \frac{1}{2}$ we can show Assumption 1 (iii) by choosing the sequence $\beta^{(0)}$ in an appropriate way.

In general, the construction of the routine $\mathrm{RHS}_{\mathbf{f}}$ for approximating $\mathbf{f}$ will highly depend on the source term g and the initial condition h in the parabolic PDE (3.3) as well as on the choice of the wavelet bases. Hence, we collect some properties that this routine should fulfil.

Assumptions 2. *The procedure* $\mathrm{RHS}_{\mathbf{f}}(\eta) = [\mathrm{RHS}_{\mathbf{f}_1}(\eta_1)\ \mathrm{RHS}_{\mathbf{f}_2}(\eta_2)]^\top$ *with* $\eta_1^2 + \eta_2^2 = \eta^2$ *is assumed to have the following properties.*

(i) *There exists an approximation* $\mathbf{f}_{1,\eta} = \mathrm{RHS}_{\mathbf{f}_1}(\eta)$ *such that* $\|\mathbf{f}_1 - \mathbf{f}_{1,\eta}\|_{\ell_2(\vee)} \le \eta$ *and*

$$\begin{aligned}
\big\|\pi^{(\mathrm{t},i)}(\mathbf{f}_{1,\eta})\big\|_{\mathcal{A}^s} &\le C^{\mathrm{sparse}}\big\|\pi^{(\mathrm{t},i)}(\mathbf{f}_1)\big\|_{\mathcal{A}^s},\\
\sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\mathbf{f}_{1,\eta})) &\le C^{\mathrm{supp}}\, d\,\eta^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{f}_1)\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\\
\mathrm{rank}_\infty(\mathbf{f}_{1,\eta}) &\le C_{\mathbf{f}}^{\mathrm{rank}}\,(1+|\ln\eta|)^{b_{\mathbf{f}}},\\
L(\mathbf{f}_{1,\eta}) &\le C^{\mathrm{lvl}}(|\ln\eta| + \ln d),
\end{aligned}$$

as well as an approximation $\mathbf{f}_{2,\eta} = \mathrm{RHS}_{\mathbf{f}_2}(\eta)$ *such that* $\|\mathbf{f}_2 - \mathbf{f}_{2,\eta}\|_{\ell_2(\vee_{\mathrm{x}})} \le \eta$ *and*

$$\begin{aligned}
\|\pi^{(i)}(\mathbf{f}_{2,\eta})\|_{\mathcal{A}^s} &\le C^{\mathrm{sparse}}\|\pi^{(i)}(\mathbf{f}_2)\|_{\mathcal{A}^s},\\
\sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(i)}(\mathbf{f}_{2,\eta})) &\le C^{\mathrm{supp}}\, d\,\eta^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\|\pi^{(i)}(\mathbf{f}_2)\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\\
|\mathrm{rank}(\mathbf{f}_{2,\eta})|_\infty &\le C_{\mathbf{f}}^{\mathrm{rank}}\,(1+|\ln\eta|)^{b_{\mathbf{f}}},\\
L_{\mathrm{x}}(\mathbf{f}_{2,\eta}) &\le C^{\mathrm{lvl}}(|\ln\eta| + \ln d),
\end{aligned}$$

where the constants $C^{\mathrm{sparse}}, C^{\mathrm{supp}}, C_{\mathbf{f}}^{\mathrm{rank}}, C^{\mathrm{lvl}} > 0$, $b_{\mathbf{f}} \ge 1$ *are independent of* η, *and* C^{sparse}, $C^{\mathrm{supp}}, C^{\mathrm{lvl}}$ *are independent of* $\mathbf{f}$.

(ii) *The number of operations required for evaluating* $\mathrm{RHS}_{\mathbf{f}}(\eta)$ *is bounded by*

$$\mathrm{flops}(\mathbf{f}_\eta) \le C_{\mathbf{f}}^{\mathrm{ops}}(d)\big[(1+|\ln\eta|)^{3b_{\mathbf{f}}} + (1+|\ln\eta|)^{b_{\mathbf{f}}}\eta^{-\frac{1}{s}}\big]$$

with a constant $C_{\mathbf{f}}^{\mathrm{ops}}(d)$ *independent of* η.

(iii) *The routine* $\mathrm{RHS}_{\mathbf{f}}$ *preserves the excess regularity of the problem, that is, there exists* $C_{\mathbf{f}}^{\mathrm{reg}} > 0$ *independent of* η *such that*

$$\begin{aligned}
\big\|\bar{\mathbf{D}}_{\mathcal{Y}}^{-\tau}\mathbf{f}_{1,\eta}\big\| &\le C_{\mathbf{f}}^{\mathrm{reg}}\big\|\bar{\mathbf{D}}_{\mathcal{Y}}^{-\tau}\mathbf{f}_1\big\|,\\
\big\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{f}_{1,\eta}\big\| &\le C_{\mathbf{f}}^{\mathrm{reg}}\big\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{f}_1\big\|,\\
\big\|\bar{\mathbf{D}}^{-2\tau}\mathbf{f}_{2,\eta}\big\| &\le C_{\mathbf{f}}^{\mathrm{reg}}\big\|\bar{\mathbf{D}}^{-2\tau}\mathbf{f}_2\big\|.
\end{aligned}$$

In the numerical experiments in Section 4.5, we consider finitely supported right-hand sides. In this case, the construction of a routine RHS that fulfills these assumptions is trivial. For the elliptic case, in [BD16, Appendix B], the approximation of more general right-hand sides that satisfy similar assumptions is discussed.

Under Assumptions 1 and 2, we are especially interested in whether our algorithm produces low-rank sparse approximations. To observe these low-rank sparse approximations, our solution $\mathbf{u}$ needs to have the corresponding approximability

properties. However, these assumptions are not explicitly used in the adaptive solver in Algorithm 4.1. This property is also referred to as universality. Regarding the approximability, we state the following assumptions on the solution $\mathbf{u}$.

Assumptions 3. *Concerning the approximability of the solution* $\mathbf{u}$*, we assume:*

(i) $\mathbf{u} \in \mathcal{A}(\gamma_{\mathbf{u}})$ *with* $\gamma_{\mathbf{u}}(n) = e^{c_{\mathbf{u}} n^{1/b_{\mathbf{u}}}}$ *for some* $c_{\mathbf{u}} > 0$, $b_{\mathbf{u}} \geq 1$ *and* $\mathcal{A}(\gamma_{\mathbf{u}})$ *defined in Theorem 4.4.*

(ii) $\pi^{(\mathrm{t},i)}(\mathbf{u}) \in \mathcal{A}^s$ *for* $i = 1, \dots, d$*, for any* s *with* $0 < s < s^*$ *with* $\mathcal{A}^s$ *defined in Theorem 2.7.*

One of the main goals is to analyze the computational complexity of the algorithm with respect to the number of spatial dimensions d. In particular, we want to show that we can avoid the curse of dimensionality, meaning exponential growth in computational cost with respect to d. To do this, it is essential to state how the important quantities depend on the number of spatial dimensions.

Assumptions 4. *For the dependence with respect to the number of spatial dimensions* d*, we assume the following.*

(i) *The following are independent of* d*: the constants* $c_{\mathbf{u}}$, $b_{\mathbf{u}}$, C^{sparse}, C^{supp}, $C^{\mathrm{rank}}_{\mathbf{f}}$, $C^{\mathrm{lvl}}_{\mathbf{f}}$; *the excess regularity index* τ*, and* $C^{\mathrm{reg}}_{\mathbf{f}}$.

(ii) *The following quantities remain bounded independently of* d*:* $\|\mathbf{B}\|$, $\|\mathbf{B}^{-1}\|$*, as well as* $\|\mathbf{D}_{\mathcal{Y}}^{1-\tau}\mathbf{g}_1\|$, $\|\mathbf{D}^{-\tau}\mathbf{g}_2\|$*; the quantities* $\|\mathbf{u}\|_{\mathcal{A}(\gamma_{\mathbf{u}})}$ *and* $\|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}$ *in Assumption 3 and* $\|\pi^{(\mathrm{t},i)}(\mathbf{f}_1)\|_{\mathcal{A}^s}$, $\|\pi^{(i)}(\mathbf{f}_2)\|_{\mathcal{A}^s}$ *in Assumption 2 (i), each for* $i = 1, \dots, d$.

(iii) *In addition, we assume that* $C^{\mathrm{ops}}_{\mathbf{f}}(d)$ *as in Assumption 3 (ii) grows at most polynomially as* $d \to \infty$.

The independence of $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ of the spatial dimension d was already discussed in Section 3.1.2 for a suitable choice of the wavelet bases.

Remark 4.68. Since $\|\mathbf{A}\| = \|\mathbf{B}\|^2$ and $\|\mathbf{A}^{-1}\| = \|\mathbf{B}^{-1}\|^2$, we can infer the d-independence of the norms of $\mathbf{A}$ and $\mathbf{A}^{-1}$ directly from the d-independence of the norms of $\mathbf{B}$ and $\mathbf{B}^{-1}$. If we choose $\omega = 2(\|\mathbf{A}\| + \|\mathbf{A}^{-1}\|^{-1})^{-1}$ in Algorithm 4.1, we obtain

$$\|\mathbf{I} - \omega\mathbf{A}\| = \frac{\mathrm{cond}(\mathbf{A}) - 1}{\mathrm{cond}(\mathbf{A}) + 1} < 1$$

with $\mathrm{cond}(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$. Hence, both ω and the reduction parameter ρ are independent of the number of spatial dimensions d as well.

After stating the necessary assumptions, we can now state the main theorem. This theorem is split into two parts. First, we state the properties of the approximation with respect to support sizes, ranks, and compressibility. In the second part, we provide bounds on the computational complexity of the algorithm under additional assumptions.

Theorem 4.69 ([BF24, Theorem 6.7])**.** *Suppose that Assumptions 1 and 2 hold and that Assumption 3 is valid for the solution* $\mathbf{u}$ *of* $\mathbf{A}\mathbf{u} = \mathbf{B}^\top\mathbf{f}$*. Let* $\varepsilon < \varepsilon_0$ *with* ε_0 *chosen as in Assumption 1 (v), let* $\alpha > 0$ *and let* $\kappa_{\mathrm{P}} = \sqrt{2d-3}, \kappa_{\mathrm{C}} = \sqrt{d}$ *with* $d \geq 3$*. Moreover, let the constants* $\kappa_1, \kappa_2, \kappa_3$ *in Algorithm 4.1 be chosen as*

$$\begin{aligned} \kappa_1 &= \big(1 + (1+\alpha)(\kappa_{\mathrm{P}} + \kappa_{\mathrm{C}} + \kappa_{\mathrm{P}}\kappa_{\mathrm{C}})\big)^{-1}, \\ \kappa_2 &= (1+\alpha)\kappa_{\mathrm{P}}\kappa_1, \\ \kappa_3 &= \kappa_{\mathrm{C}}(1+\kappa_{\mathrm{P}})(1+\alpha)\kappa_1, \end{aligned} \tag{4.158}$$

and let $\beta_1 \geq 0$, $\beta_2 > 0$ be arbitrary but fixed. Then the approximate solution $\mathbf{u}_\varepsilon$ computed by Algorithm 4.1 satisfies

$$\operatorname{rank}_\infty(\mathbf{u}_\varepsilon) \leq \gamma_{\mathbf{u}}^{-1}\big(2(\alpha\kappa_1)^{-1}\,\|\mathbf{u}\|_{\mathcal{A}(\gamma_{\mathbf{u}})}\,\varepsilon^{-1}\big), \tag{4.159}$$

$$\sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{u}_\varepsilon)) \lesssim d^{1+s^{-1}} \Big(\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}} \varepsilon^{-\frac{1}{s}}, \tag{4.160}$$

as well as

$$\|\mathbf{u}_\varepsilon\|_{\mathcal{A}(\gamma_{\mathbf{u}})} \lesssim \sqrt{d}\,\|\mathbf{u}\|_{\mathcal{A}(\gamma_{\mathbf{u}})}, \tag{4.161}$$

$$\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u}_\varepsilon)\|_{\mathcal{A}^s} \lesssim d^{1+\max\{1,s\}} \sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}. \tag{4.162}$$

The multiplicative constant in (4.161) depends only on α, those in (4.160) and (4.162) depend only on α and s.

If, in addition, Assumption 4 hold, then for the number of required operations to calculate $\mathbf{u}_\varepsilon$ by Algorithm 4.1, we have the estimate

$$\operatorname{flops}(\mathbf{u}_\varepsilon) \leq C d^{c_1+c_2\ln d}(\ln d)^{c_3+c_4\ln d}(1+|\ln\varepsilon|)^{c_5+c_6\ln d}\,\varepsilon^{-\frac{1}{s}}, \tag{4.163}$$

where $C, c_1, c_2, c_3, c_4, c_5$ and c_6 are constants independent of ε and d.

As we will see in the proof of the theorem, the exponents stated in this work are slightly better than the exponents observed in [BF24, Theorem 6.7], which is due to the different treatment of the initial value operator. We will comment further on the optimality of the exponents in Theorem 4.80.

The proof of the bound of the quantities of $\mathbf{u}_\varepsilon$ is mainly based on Theorem 4.12. For the proof of the bound on the computational complexity, we need some additional preparation.

Since we want to solve the normal equation (4.149), we write for notational simplicity

$$\mathbf{z}_{k,j} = \mathrm{APPLY}_{\mathbf{B}_1}\left(\mathbf{w}_{k,j}, \frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\right) \tag{4.164}$$

for the intermediate results in the adaptive application of $\mathbf{B}_1^\top\mathbf{B}_1$, where $\eta_{k,j}^{(1)} = \lambda\eta_{k,j}$ with some $\lambda \in (0,1)$. The remainder of this section is dedicated to the proof of this theorem. We divide the proof into multiple smaller lemmas.

4.4.2.1 Maximum Level and Support of the Iterates

The ranks of the output of the APPLY routine depend by Theorems 4.43 and 4.54 and Theorems 4.48 and 4.56 on the maximal wavelet levels present in the current iterate. Therefore, bounding the maximal spatial and temporal level is necessary. The following result establishes these bounds after the COARSEN routine is applied. This routine is used after each step in the perturbed Richardson iteration since $\beta_2 > 0$ in Algorithm 4.1. These bounds depend on the excess regularity assumption and on the support sizes of the spatial-temporal contractions.

Lemma 4.70 ([BF24, Lemma 6.8]). *For given $\mathbf{v} \in \ell_2(\vee)$, we assume*

$$N = \sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathbf{v})\big) < \infty,$$

and that for some $\tau > 0$, one has $\|(\bar{\mathbf{D}}_t^{-\tau} \otimes \mathbf{I}_1)\pi^{(t,i)}(\mathbf{v})\| < \infty$ and $\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{v})\| < \infty$ for all $i = 1, \dots, d$. Let $\eta > 0$, $\mathbf{v}_\eta = \mathrm{COARSEN}(\mathbf{v}, \eta)$ and

$$C_{\vee_t} = \sup_{\nu_t \in \vee_t} \|\theta_{\nu_t}\|_{H^1}^{-\tau} 2^{\tau|\nu_t|}, \quad C_{\vee_1} = \sup_{\nu \in \vee_1} \|\psi_\nu\|_{H_0^1}^{-\tau} 2^{\tau|\nu|}.$$

Then for all $i = 1, \dots, d$ and $(\nu_t, \nu_1) \in \operatorname{supp}(\pi^{(t,i)}(\mathbf{v}_\eta))$, we have

$$|\nu_t| \le \tau^{-1} \log_2 \left(\eta^{-1} C_{\vee_t} \|(\bar{\mathbf{D}}_t^{-\tau} \otimes \mathbf{I}_1)\pi^{(t,i)}(\mathbf{v})\| \sqrt{N}\right),$$
$$|\nu_1| \le \tau^{-1} \log_2 \left(\eta^{-1} C_{\vee_1} \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{v})\| \sqrt{N}\right).$$

Proof. The proof is an adaptation of the elliptic case in [BD16, Lemma 37] and follows the same lines using spatio-temporal contractions instead.

Suppose that $i \in \{1, \dots, d\}$ and $(\nu_t, \nu_1) \in \operatorname{supp}(\pi^{(t,i)}(\mathbf{v}_\eta))$ with

$$|\nu_i| > \tau^{-1} \log_2 \left(\eta^{-1} C_{\vee_1} \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{v})\| \sqrt{N}\right).$$

Then by the definition of $\bar{\mathbf{D}}_1$, we have

$$\begin{aligned} |\pi^{(t,i)}_{\nu_t,\nu_i}(\mathbf{v})| &\le \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{v})\| \|\psi_{\nu_i}\|_{H_0^1}^{-\tau} \\ &\le \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{v})\| C_{\vee_1} 2^{-\tau|\nu_i|} \\ &< \frac{\eta}{\sqrt{N}}. \end{aligned}$$

We set

$$\hat{\Lambda} = \left\{(j, \mu_t, \mu_1) : j = 1, \dots, d, (\mu_t, \mu_1) \notin \operatorname{supp}\left(\pi^{(t,j)}(\mathbf{v}_\eta)\right)\right\}.$$

By the definition of COARSEN in Section 4.1.2, we have $\pi^{(t,j)}_{\mu_t,\mu_j}(\mathbf{v}) \le \pi^{(t,i)}_{\nu_t,\nu_i}(\mathbf{v})$ for all $(j, \mu_t, \mu_j) \in \hat{\Lambda}$. Hence,

$$\sum_{(j,\mu_t,\mu_j) \in \hat{\Lambda} \cup \{(i,\nu_t,\nu_i)\}} |\pi^{(t,j)}_{\mu_t,\mu_j}(\mathbf{v})|^2 \le N \frac{\eta^2}{N} = \eta^2.$$

This is a contradiction to the minimality condition (4.20) of COARSEN. The proof for the temporal levels is analogous. □

This lemma is applied to line 11 of Algorithm 4.1. Controlling the maximal spatial and temporal levels of iterates $\mathbf{w}_{k,j}$ is essential; thus, having $\beta_2 > 0$ is crucial for the coarsening step, while the recompress step remains optional, i.e., $\beta_1 = 0$ can be chosen.

Remark 4.71. Theorem 4.70 provides a way to construct the routine $\mathrm{RHS}_{\mathbf{f}_1}$ so that it fulfills the level bounds in Assumption 2, assuming the remaining conditions are already met. If we modify the routine by first applying RHS with a smaller tolerance, and then applying COARSEN with an adapted tolerance, the level bounds will also be fulfilled. This holds by considering the excess regularity from Assumption 1 (vii). For the routine $\mathrm{RHS}_{\mathbf{f}_2}$, we can proceed similarly by using [BD16, Lemma 37] from the elliptic problem.

To obtain suitable bounds for the maximal levels of the iterates $\mathbf{w}_{k,j}$, we have to estimate $\|(\bar{\mathbf{D}}_t^{-\tau} \otimes \mathbf{I}_1)\pi^{(t,i)}(\mathbf{w}_{k,j})\|$ and $\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j})\|$. The bounds on the

ranks in Theorems 4.48 and 4.56 inherently depend on the lower-dimensional support and the specified tolerance. Next, we state a result about the stability of the contractions under the application of the COARSEN and RECOMPRESS routines.

Lemma 4.72 ([BF24, Lemma 6.9]). *For any $\tau, \eta > 0$ and $i = 1, \dots, d$, we have*

$$\begin{aligned}
\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_1)\pi^{(\mathrm{t},i)}(\textsc{coarsen}(\mathbf{v}, \eta))\| &\le \|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{v})\|, \\
\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\textsc{coarsen}(\mathbf{v}, \eta))\| &\le \|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{v})\|, \\
\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_1)\pi^{(\mathrm{t},i)}(\textsc{recompress}(\mathbf{v}, \eta))\| &\le \|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{v})\|, \\
\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\textsc{recompress}(\mathbf{v}, \eta))\| &\le \|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{v})\|,
\end{aligned}$$

for any $\mathbf{v} \in \ell_2(\vee)$.

This lemma is an adaptation of [BD16, Lemma 38] of the elliptic case to the spatio-temporal setting. The proof is an immediate consequence of the stability of the contractions under application of the RECOMPRESS and COARSEN routines.

Before we examine the evolution of the quantities $\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|$ and $\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|$ during the iteration, we need the following result about the regularity of $\mathbf{B}_2^{\top}$.

Lemma 4.73 ([BF24, Lemma 6.4]). *For all $0 < \tau < \frac{1}{2}$ we have*

$$\left\|\left(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_{\mathrm{x}}\right)\mathbf{B}_2^{\top}\bar{\mathbf{D}}^{2\tau}\right\| \le \sqrt{2}(1 + \tfrac{\delta}{2}).$$

Proof. By the definition of the matrices and the trace theorem (4.123), we obtain

$$\begin{aligned}
\left|\left((\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_{\mathrm{x}})\bar{\mathbf{D}}_{\mathcal{X}}\mathbf{T}_0^{\top}\bar{\mathbf{D}}^{2\tau}\right)_{(\nu_{\mathrm{t}},\nu_{\mathrm{x}}),\nu_{\mathrm{x}}}\right| &= \|\theta_{\nu_{\mathrm{t}}}\|_{H^1}^{\tau} \frac{\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}}{\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^2 + \|\theta_{\nu_{\mathrm{t}}}\|_{H^1}} |\theta_{\nu_{\mathrm{t}}}(0)| \|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^{-2\tau} \\
&\le \sqrt{2} \frac{\|\theta_{\nu_{\mathrm{t}}}\|_{H^1}^{\frac{1}{2}(1+2\tau)} (\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^2)^{\frac{1}{2}(1-2\tau)}}{\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^2 + \|\theta_{\nu_{\mathrm{t}}}\|_{H^1}} \\
&\le \sqrt{2} \frac{\frac{1}{2}(1+2\tau)\|\theta_{\nu_{\mathrm{t}}}\|_{H^1} + \frac{1}{2}(1-2\tau)\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^2}{\|\Psi_{\nu_{\mathrm{x}}}\|_{H_0^1}^2 + \|\theta_{\nu_{\mathrm{t}}}\|_{H^1}} \\
&\le \sqrt{2},
\end{aligned}$$

where we used Young's inequality in the third step. Combining this estimate with (4.57) yields the desired result. □

Based on the excess regularity assumptions in Assumption 1 (vii), we introduce

$$\zeta_{\mathbf{f}} = \max\left\{\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_{\mathrm{x}})\bar{\mathbf{D}}_{\mathcal{Y}}\mathbf{g}_1\right\|, \left\|\bar{\mathbf{D}}_{\mathcal{Y}}^{1-\tau}\mathbf{g}_1\right\|, \left\|\bar{\mathbf{D}}^{-2\tau}\mathbf{g}_2\right\|\right\} < \infty, \tag{4.165}$$

$$\begin{aligned}
\zeta_{\mathbf{B}} = \max\Big\{ &\left\|\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\mathbf{C}_{\mathrm{t}}\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\right\|, \left\|\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\mathbf{C}_{\mathrm{t}}^{\top}\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\right\|, \\
&\left\|\bar{\mathbf{D}}_1^{-\tau}\mathbf{C}_2\bar{\mathbf{D}}_1^{\tau}\right\|, \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau} \otimes \mathbf{I}_{\mathrm{x}})\bar{\mathbf{A}}_2(\bar{\mathbf{D}}_{\mathrm{t}}^{\tau} \otimes \mathbf{I}_{\mathrm{x}})\right\| \Big\} < \infty.
\end{aligned} \tag{4.166}$$

Recall that the operator $\mathbf{C}_2$ is symmetric. Hence, we also have

$$\left\|\bar{\mathbf{D}}_1^{-\tau}\mathbf{C}_2^{\top}\bar{\mathbf{D}}_1^{\tau}\right\| = \left\|\bar{\mathbf{D}}_1^{-\tau}\mathbf{C}_2\bar{\mathbf{D}}_1^{\tau}\right\| \le \zeta_{\mathbf{B}}.$$

The next statement is a modified version of [BF24, Proposition 6.10]. The differences arise from the different treatment of the initial-value operator.

Proposition 4.74. *Assume the conditions of Theorem 4.69 hold. Then the iterates* $\mathbf{w}_{k,j}$ *of Algorithm 4.1 satisfy*

$$\max\Big\{\big\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|,\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|\Big\}\le\frac{\gamma^{kI+j}-1}{\gamma-1}\bar{C}_{\mathbf{f}},\tag{4.167}$$

with

$$\begin{aligned}\gamma&=1+\omega\max\left\{\gamma_{\mathrm{t},\mathbf{B}_1}^2+\gamma_{\mathrm{t},\mathbf{A}_2},\gamma_{\mathrm{x},\mathbf{B}_1}^2+\gamma_{\mathrm{x},\mathbf{A}_2}\right\},\\ \bar{C}_{\mathbf{f}}&=\omega C_{\mathbf{f}}^{\mathrm{reg}}\zeta_{\mathbf{f}}\max\left\{\gamma_{\mathrm{t},\mathbf{B}_1}+\gamma_{\mathrm{t},\mathbf{B}_2},\gamma_{\mathrm{x},\mathbf{B}_1}+\gamma_{\mathrm{x},\mathbf{B}_2}\right\},\end{aligned}\tag{4.168}$$

where

$$\begin{aligned}\gamma_{\mathrm{x},\mathbf{B}_1}&=C_\delta(2(d-1)\|\mathbf{C}_2\|+2\|\mathbf{C}_{\mathrm{t}}\|+\zeta_{\mathbf{B}}+C_{\tau,2}\|\beta(\mathbf{C}_2)\|_{\ell_1}),&\gamma_{\mathrm{x},\mathbf{B}_2}&=2(s_{\min}^{(1)})^{\tau}\|\mathbf{B}_2\|,\\ \gamma_{\mathrm{t},\mathbf{B}_1}&=C_\delta(2\|\bar{\mathbf{B}}_{\mathrm{x}}\|+\zeta_{\mathbf{B}}+C_{\tau,\mathrm{t}}\|\beta(\mathbf{C}_{\mathrm{t}})\|_{\ell_1}),&\gamma_{\mathrm{t},\mathbf{B}_2}&=\sqrt{2}(1+\tfrac{\delta}{2}),\\ \gamma_{\mathrm{x},\mathbf{A}_2}&=2\|\mathbf{A}_2\|,&\gamma_{\mathrm{t},\mathbf{A}_2}&=C_\delta(\zeta_{\mathbf{B}}+C_{\tau,0}\|\beta^{(0)}\|_{\ell_1}),\end{aligned}$$

with $s_{\min}^{(1)}$ *from* (4.117).

Proof. The proof follows the ideas of [BF24, Proposition 6.10] and [BD15, Proposition 39].

We start with the proof of the bound for $\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|$. In the outer loop k, each inner loop step j is of the form

$$\mathbf{w}_{k,j+1}=\mathrm{COARSEN}\left(\mathrm{RECOMPRESS}\big((\mathbf{I}-\omega\tilde{\mathbf{A}}_{k,j})\mathbf{w}_{k,j}+\omega\mathbf{h}_{k,j},\beta_1\eta_{k,j}\big),\beta_2\eta_{k,j}\right)$$

with $\tilde{\mathbf{A}}_{k,j}\mathbf{w}_{k,j}=\mathrm{APPLY}(\mathbf{w}_{k,j},\frac{1}{2}\eta_{k,j})$ and $\mathbf{h}_{k,j}=\mathrm{RHS}(\frac{1}{2}\eta_{k,j})$. Using the stability of the contractions under the application of the RECOMPRESS and COARSEN routines from Theorem 4.72 and the triangle inequality, we obtain

$$\begin{aligned}\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\big\|&\le\Big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}\big((\mathbf{I}-\omega\tilde{\mathbf{A}}_{k,j})\mathbf{w}_{k,j}+\omega\mathbf{h}_{k,j}\big)\Big\|\\ &\le\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|\\ &\quad+\omega\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}\mathbf{w}_{k,j})\big\|\\ &\quad+\omega\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\big\|.\end{aligned}\tag{4.169}$$

In the next steps, we need to derive further estimates for the last two terms. By the definition of $\tilde{\mathbf{A}}_{k,j}$ and the triangle inequality, we have

$$\begin{aligned}&\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}\mathbf{w}_{k,j})\big\|\\ &\qquad\le\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(1)}\mathbf{w}_{k,j})\big\|+\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\big\|,\end{aligned}\tag{4.170}$$

where

$$\begin{aligned}\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(1)}\mathbf{w}_{k,j})\|&=\Big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{B}_1^{\mathsf{T}}}\Big(\mathbf{z}_{k,j},\frac{\eta_{k,j}^{(1)}}{4}\Big)\Big)\Big\|\\ &=\Big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{B}_1^{\mathsf{T}}}\Big(\mathrm{APPLY}_{\mathbf{B}_1}\Big(\mathbf{w}_{k,j},\frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\Big),\frac{\eta_{k,j}^{(1)}}{4}\Big)\Big)\Big\|\end{aligned}$$

and

$$\big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\big\|=\Big\|(\mathbf{I}_{\mathrm{t}}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{A}_2}\Big(\mathbf{w}_{k,j},\frac{\eta_{k,j}^{(2)}}{2}\Big)\Big)\Big\|$$

with $\eta_{k,j}^{(1)} + \eta_{k,j}^{(2)} = \eta_{k,j}$. We start with the part involving $\tilde{\mathbf{A}}_{k,j}^{(1)}$. The principal idea is to do a stepwise estimate

$$\begin{aligned}\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}\Big(\mathrm{APPLY}_{\mathbf{B}_1^\top}\Big(\mathbf{z}_{k,j}, \tfrac{\eta_{k,j}^{(1)}}{4}\Big)\Big)\right\| &\leq \gamma_{\mathsf{x},\mathbf{B}_1}\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}(\mathbf{z}_{k,j})\right\|, \\ \left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}(\mathbf{z}_{k,j})\right\| &\leq \gamma_{\mathsf{x},\mathbf{B}_1}\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}(\mathbf{w}_{k,j})\right\|.\end{aligned} \tag{4.171}$$

These two estimates can be proven exactly in the same manner since $\mathbf{B}_1$ and $\mathbf{B}_1^\top$ have the same properties. Hence, we restrict ourselves to the proof of the second estimate. For notational simplicity, we define $\bar{\mathbf{D}}_{\mathsf{x},i} \in \mathbb{R}^{\mathsf{V}_\mathsf{x}\times\mathsf{V}_\mathsf{x}}$ by

$$\bar{\mathbf{D}}_{\mathsf{x},i} = \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \bar{\mathbf{D}}_1 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1,$$

where $\bar{\mathbf{D}}_1$ is at the i-th position. By construction of the adaptive application routine in Theorem 4.40, we have $\mathbf{z}_{k,j} = \mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}_{k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}$ with

$$\eta = \frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}, \quad \tilde{\mathbf{T}}_{k,j} = \tilde{\mathbf{T}}_{J(\eta;\mathbf{w}_{k,j}),\mathbf{w}_{k,j}}, \quad \mathsf{n}_1 = \mathsf{m}_\mathcal{X}(\eta;\mathbf{w}_{k,j}), \quad \mathsf{n}_2 = \mathsf{m}_\mathcal{Y}(\eta;\mathbf{w}_{k,j}).$$

Using that $\bar{\mathbf{D}}_1$ is a diagonal matrix, we obtain

$$\begin{aligned}\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}(\mathbf{z}_{k,j})\right\| &= \left\|\pi^{(\mathsf{t},i)}\big((\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}_{k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}\big)\right\| \\ &= \left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y},\mathsf{n}_2}\tilde{\mathbf{T}}_{k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}\right\| \\ &\leq \left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}\right\|,\end{aligned}$$

where in the last step we use that diagonal operators commute as well as (4.63). By the construction in (4.97), we have $\tilde{\mathbf{T}}_{k,j} = \tilde{\mathbf{T}}_\mathsf{t} + \tilde{\mathbf{T}}_\mathsf{x}$ with $\tilde{\mathbf{T}}_\mathsf{t} = \tilde{\mathbf{T}}_{\mathsf{t},J(\eta;\mathbf{w}_{k,j}),\mathbf{w}_{k,j}}$ and $\tilde{\mathbf{T}}_\mathsf{x} = \tilde{\mathbf{T}}_{\mathsf{x},J(\eta;\mathbf{w}_{k,j}),\mathbf{w}_{k,j}}$. Hence, we obtain

$$\begin{aligned}&\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathsf{t},i)}\big(\mathbf{z}_{k,j}\big)\right\| \\ &\qquad\leq \left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| + \left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{x}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\|\end{aligned} \tag{4.172}$$

with $\tilde{\mathbf{w}}_{k,j} = \mathbf{D}_\mathcal{X}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}$. We start by estimating the first term on the right-hand side. By the definition of $\tilde{\mathbf{T}}_\mathsf{t}$ in (4.91) with only restrictions in the spatial variables, we have

$$(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}} = \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}}(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau}),$$

where we use the fact that diagonal operators commute. Hence, we obtain

$$\begin{aligned}\left\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| &\leq \|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}}\|\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{D}_\mathcal{X}^{-1}\mathbf{D}_{\mathcal{X},\mathsf{n}_1}\mathbf{w}_{k,j}\| \\ &\leq \|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_\mathsf{t}\mathbf{D}_{\mathcal{X}}\|\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{w}_{k,j}\| \\ &\leq 2\|\mathbf{D}_{\mathcal{Y}}\mathbf{T}_\mathsf{t}\mathbf{D}_{\mathcal{X}}\|\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{w}_{k,j}\| \\ &\leq 2C_\delta\|\mathbf{C}_\mathsf{t}\|\|(\mathbf{I}_\mathsf{t} \otimes \bar{\mathbf{D}}_{\mathsf{x},i}^{-\tau})\mathbf{w}_{k,j}\|,\end{aligned}$$

where we use that diagonal operators commute and (4.56) in the second line, that $\tilde{\mathbf{T}}_\mathsf{t}$ is an approximation of $\mathbf{T}_\mathsf{t}$ in the third line and (4.68) from Theorem 4.29 in the last line. Next, we consider the second term on the right-hand side of (4.172). Here we make use of the structure of the Laplacian corresponding to $\mathbf{T}_\mathsf{x}$. By definition, we

have

$$\begin{aligned}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{w} &= \sum_{\nu_{\mathrm{t}}\in\mathcal{V}_{\mathrm{t}}} \mathbf{e}_{\nu_{\mathrm{t}}} \otimes (\tilde{\mathbf{T}}_{\mathrm{x},\nu_{\mathrm{t}}}\mathbf{w}_{\nu_{\mathrm{t}}}) \\ &= \sum_{\nu_{\mathrm{t}}\in\mathcal{V}_{\mathrm{t}}} \mathbf{e}_{\nu_{\mathrm{t}}} \otimes \Big(\sum_{i=1}^{d} \tilde{\mathbf{T}}^{(i)}_{\mathrm{x},\nu_{\mathrm{t}}}\mathbf{w}_{\nu_{\mathrm{t}}}\Big)\end{aligned}$$

with

$$\tilde{\mathbf{T}}^{(i)}_{\mathrm{x},\nu_{\mathrm{t}}} = \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \tilde{\mathbf{T}}^{(i)}_{\nu_{\mathrm{t}},2,\mathbf{w}_{k,j}} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1.$$

By construction, we have

$$\tilde{\mathbf{T}}_{\mathrm{x}} = \sum_{i=1}^{d} \tilde{\mathbf{T}}^{(i)}_{\mathrm{x}}, \quad \text{where} \quad \tilde{\mathbf{T}}^{(i)}_{\mathrm{x}} = \sum_{\nu_{\mathrm{t}}\in\mathcal{V}_{\mathrm{t}}} \mathbf{E}_{\nu_{\mathrm{t}}} \otimes \tilde{\mathbf{T}}^{(i)}_{\mathrm{x},\nu_{\mathrm{t}}}.$$

Using this representation, we obtain

$$\big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| \le \sum_{j=1}^{d}\big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}^{(j)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\|.$$

For $j \neq i$, we have

$$(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}^{(j)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}} = \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}^{(j)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})$$

by the definition of $\tilde{\mathbf{T}}^{(j)}_{\mathrm{x}}$ and $\bar{\mathbf{D}}_{\mathrm{x},i}$, since $\tilde{\mathbf{T}}^{(j)}_{\mathrm{x},\nu_{\mathrm{t}}}$ has a identity in the i-th position and $\bar{\mathbf{D}}_{\mathrm{x},i}$ acts only on the j-th position. Proceeding as for the temporal part, we arrive at

$$\big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}^{(j)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| \le 2C_\delta\|\mathbf{C}_2\|\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{w}_{k,j}\|, \quad j \neq i.$$

As the final step, we need to consider the case $j = i$. We have

$$\begin{aligned}&\big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}^{(i)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| \\ &\qquad \le \big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\mathbf{T}^{(i)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| + \big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}(\mathbf{T}^{(i)}_{\mathrm{x}} - \tilde{\mathbf{T}}^{(i)}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\|\end{aligned}$$

with

$$\mathbf{T}^{(i)}_{\mathrm{x}} = \mathbf{I}_{\mathrm{t}} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes \mathbf{T}_2 \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1,$$

where $\mathbf{T}_2$ is at the i-th position. We obtain

$$\begin{aligned}\big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{D}_{\mathcal{Y}}\mathbf{T}^{(i)}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| &\le \big\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i}\mathbf{D})\mathbf{T}^{(i)}_{\mathrm{x}}(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{1+\tau}_{\mathrm{x},i})\big\|\big\|\mathbf{D}_{\mathcal{X}}(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-1-\tau}_{\mathrm{x},i})\tilde{\mathbf{w}}_{k,j}\big\| \\ &\le \big\|\bar{\mathbf{D}}^{-1}_{\mathrm{x},i}\mathbf{D}\big\|\big\|\bar{\mathbf{D}}^{1-\tau}_{1}\mathbf{T}_2\bar{\mathbf{D}}^{1+\tau}_{1}\big\|\big\|\mathbf{D}_{\mathcal{X}}(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-1}_{\mathrm{x},i})\big\|\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\tilde{\mathbf{w}}_{k,j}\| \\ &\le C_\delta\zeta_{\mathbf{B}}\big\|\bar{\mathbf{D}}^{-1}_{\mathrm{x},i}\bar{\mathbf{D}}\big\|\big\|\bar{\mathbf{D}}_{\mathcal{X}}(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-1}_{\mathrm{x},i})\big\|\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\tilde{\mathbf{w}}_{k,j}\| \\ &\le C_\delta\zeta_{\mathbf{B}}\|(\mathbf{I}_{\mathrm{t}} \otimes \bar{\mathbf{D}}^{-\tau}_{\mathrm{x},i})\mathbf{w}_{k,j}\|,\end{aligned}$$

where we used that diagonal operators commute and the definition of $\bar{\mathbf{D}}_{\mathrm{x},i}$ in the second line, (4.166), (4.57) and (4.65) in the third line, as well as the definition of $\bar{\mathbf{D}}_{\mathrm{x},i}$, $\bar{\mathbf{D}}$ and $\bar{\mathbf{D}}_{\mathcal{X}}$ and (4.56) in the last line. Using the identity in the time variable of the

spatial operator, we obtain for $\tilde{\mathbf{w}}_{k,j} = \sum_{\nu_t \in \mathbb{V}_t} \mathbf{e}_{\nu_t} \otimes \tilde{\mathbf{w}}_{\nu_t,k,j}$ the estimate

$$\begin{aligned}
&\big\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{D}_{\mathcal{Y}}(\mathbf{T}_x^{(i)} - \tilde{\mathbf{T}}_x^{(i)})\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\|^2 \\
&\quad = \sum_{\nu_t \in \mathbb{V}_t} \big\|\bar{\mathbf{D}}_{x,i}^{-\tau}\mathbf{D}(\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{\nu_t,2,\mathbf{w}_{k,j}}) \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1)\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{w}}_{\nu_t,k,j}\big\|^2 \\
&\quad \le \sum_{\nu_t \in \mathbb{V}_t} \Big(\sum_{p \in \mathbb{N}_0} \big\|\bar{\mathbf{D}}_{x,i}^{-\tau}\mathbf{D}(\mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1 \otimes (\mathbf{T}_2 - \tilde{\mathbf{T}}^{(i)}_{2,[p]})\mathrm{R}_{\Lambda^{(i)}_{\nu_t,p}(\mathbf{w}_{k,j})} \otimes \mathbf{I}_1 \otimes \cdots \otimes \mathbf{I}_1)\mathbf{D}_{\mathcal{X},\nu_t}\tilde{\mathbf{w}}_{\nu_t,k,j}\big\| \Big)^2 \\
&\quad \le \sum_{\nu_t \in \mathbb{V}_t} \Big(\sum_{j \in \mathbb{N}_0} \|\bar{\mathbf{D}}_{x,i}^{-1}\mathbf{D}\|\|\bar{\mathbf{D}}_1^{1-\tau}(\mathbf{T}_2 - \mathbf{T}_{2,j})\bar{\mathbf{D}}_1^{1+\tau}\|\|\mathbf{D}_{\mathcal{X},\nu_t}\bar{\mathbf{D}}_{x,i}^{-1}\|\|\bar{\mathbf{D}}_{x,i}^{-\tau}\tilde{\mathbf{w}}_{\nu_t,k,j}\| \Big)^2 \\
&\quad \le C_\delta^2 \sum_{\nu_t \in \mathbb{V}_t} \|\bar{\mathbf{D}}_{x,i}^{-\tau}\tilde{\mathbf{w}}_{\nu_t,k,j}\|^2 C_{\tau,2}^2 \|\beta(\mathbf{C}_2)\|_{\ell_1}^2 \\
&\quad \le C_\delta^2 C_{\tau,2}^2 \|\beta(\mathbf{C}_2)\|_{\ell_1}^2 \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{w}_{k,j}\|^2,
\end{aligned}$$

where we used (4.78) in the third line, $\|\mathrm{R}_{\Lambda^{(i)}_{\nu_t,p}(\mathbf{w}_{k,j})}\| \le 1$ in the fourth line, (4.57) (4.65) and Assumption 1 (i) in the fifth line as well as (4.56) in the last line. Combining all the estimates above yields

$$\begin{aligned}
&\big\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_x\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\big\| \\
&\qquad\qquad \le C_\delta(2(d-1)\|\mathbf{C}_2\| + \zeta_{\mathbf{B}} + C_{\tau,2}\|\beta(\mathbf{C}_2)\|_{\ell_1})\big\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j})\big\|.
\end{aligned}$$

Inserting this into (4.172), we arrive at

$$\big\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}\big(\mathbf{z}_{k,j}\big)\big\| \le \gamma_{x,\mathbf{B}_1}\big\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j})\big\|.$$

Next, we estimate the second term on the right-hand side in (4.170). Here, we can proceed similarly to the temporal operator $\tilde{\mathbf{T}}_t$, since we have an identity in the spatial variables. We have $\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j} = \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{S}}_{0,k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}}\mathbf{w}_{k,j}$ with

$$\eta = \frac{\eta_{k,j}^{(2)}}{4}, \quad \tilde{\mathbf{S}}_{0,k,j} = \tilde{\mathbf{S}}_{0,J_0(\eta;\mathbf{w}_{k,j}),\mathbf{w}_{k,j}}, \quad \mathsf{n} = \mathsf{m}_{\mathcal{X},0}(\eta;\mathbf{w}_{k,j}).$$

By construction of $\tilde{\mathbf{S}}_{0,k,j}$ in (4.130) with only restrictions in the spatial variables, we have

$$\begin{aligned}
\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\| &= \|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{S}}_{0,k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}}\mathbf{w}_{k,j}\| \\
&= \|\mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{S}}_{0,k,j}\mathbf{D}_{\mathcal{X},\mathsf{n}}(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{w}_{k,j}\| \\
&\le \|\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_{0,k,j}\mathbf{D}_{\mathcal{X}}\|\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_{x,i}^{-\tau})\mathbf{w}_{k,j}\| \\
&\le 2\|\mathbf{A}_2\|\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j})\|,
\end{aligned}$$

where we used (4.56) in the third line and that $\tilde{\mathbf{S}}_{0,k,j}$ is an approximation of $\mathbf{S}_0$ in the last line. Inserting the estimates into (4.170), by (4.169) we obtain

$$\begin{aligned}
&\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j+1})\| \\
&\qquad\qquad \le \gamma\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{w}_{k,j})\| + \omega\|(\mathbf{I}_t \otimes \bar{\mathbf{D}}_1^{-\tau})\pi^{(t,i)}(\mathbf{h}_{k,j})\|. \quad (4.173)
\end{aligned}$$

In the next step, we estimate the second term on the right-hand side above. Then the assertion follows from the recursive application of this estimate. By definition of the

routine RHS in (4.154) and the triangle inequality we have

$$\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\| \le \|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(1)})\| + \|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(2)})\|, \quad (4.174)$$

where $\mathbf{h}_{k,j}^{(1)}$ and $\mathbf{h}_{k,j}^{(2)}$ are defined by the adaptive application of $\mathbf{B}_1^\top$ and $\mathbf{B}_2^\top$, respectively, as

$$\mathbf{h}_{k,j}^{(1)} = \mathrm{APPLY}_{\mathbf{B}_1^\top}\left(\mathbf{f}_{k,j}^{(1)}, \tfrac{\eta_{k,j}^{(1)}}{4}\right), \quad \mathbf{h}_{k,j}^{(2)} = \mathrm{APPLY}_{\mathbf{B}_2^\top}\left(\mathbf{f}_{k,j}^{(2)}, \tfrac{\eta_{k,j}^{(2)}}{4}\right) \quad (4.175)$$

with $\eta_{k,j}^{(1)} + \eta_{k,j}^{(2)} = \eta_{k,j}$, where

$$\mathbf{f}_{k,j}^{(1)} = \mathrm{RHS}_{\mathbf{f}_1}\left(\tfrac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\right), \quad \mathbf{f}_{k,j}^{(2)} = \mathrm{RHS}_{\mathbf{f}_2}\left(\tfrac{\eta_{k,j}^{(2)}}{4\|\mathbf{B}_2\|}\right).$$

For the first term on the right-hand side of (4.174), we can use the results from above and observe, similarly to (4.171), the estimate

$$\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(1)})\| \le \gamma_{\mathrm{x},\mathbf{B}_1}\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{f}_{k,j}^{(1)})\|.$$

Combining this with Assumption 2 (iii) and (4.165), we obtain

$$\begin{aligned}\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(1)})\| &\le \gamma_{\mathrm{x},\mathbf{B}_1}\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_{\mathrm{x},i}^{-\tau})\mathbf{f}_{k,j}^{(1)}\| \\ &\le \gamma_{\mathrm{x},\mathbf{B}_1}\|\bar{\mathbf{D}}_\mathcal{Y}^{-\tau}\mathbf{f}_{k,j}^{(1)}\| \\ &\le \gamma_{\mathrm{x},\mathbf{B}_1}C_\mathbf{f}^\mathrm{reg}\|\bar{\mathbf{D}}_\mathcal{Y}^{-\tau}\mathbf{f}_1\| \\ &\le \gamma_{\mathrm{x},\mathbf{B}_1}C_\mathbf{f}^\mathrm{reg}\zeta_\mathbf{f}.\end{aligned}$$

By definition of the adaptive application of $\mathbf{B}_2^\top$, we have $\mathbf{h}_{k,j}^{(2)} = \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad}\mathbf{f}_{k,j}^{(2)}$, where

$$\eta = \frac{\eta_{k,j}^{(2)}}{4}, \quad \tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad} = \mathbf{T}_{0,J_0^\mathrm{ad}(\eta;\mathbf{f}_{k,j}^{(2)}),\mathbf{f}_{k,j}^{(2)}}^\top, \quad \mathsf{n} = \mathsf{m}_{\mathcal{X},0}^\mathrm{ad}(\eta;\mathbf{f}_{k,j}^{(2)}).$$

By construction of $\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad}$ in Section 4.3.4.2, we have

$$(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_{\mathrm{x},i}^{-\tau})\mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad} = \mathbf{D}_{\mathcal{X},\mathsf{n}}\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad}\bar{\mathbf{D}}_{\mathrm{x},i}^{-\tau}.$$

Hence, we obtain

$$\begin{aligned}\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(2)})\| &\le \|\mathbf{D}_\mathcal{X}\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad}\|\|\bar{\mathbf{D}}_{\mathrm{x},i}^{-\tau}\mathbf{f}_{k,j}^{(2)}\| \\ &\le 2\|\mathbf{B}_2^\top\|(s_\mathrm{min}^{(1)})^\tau\|\bar{\mathbf{D}}^{-2\tau}\mathbf{f}_{k,j}^{(2)}\| \\ &\le 2(s_\mathrm{min}^{(1)})^\tau\|\mathbf{B}_2\|C_\mathbf{f}^\mathrm{reg}\|\bar{\mathbf{D}}^{-2\tau}\mathbf{f}_2\| \\ &\le \gamma_{\mathrm{x},\mathbf{B}_2}C_\mathbf{f}^\mathrm{reg}\zeta_\mathbf{f},\end{aligned}$$

where we use (4.56) in the first line as well as $\|\bar{\mathbf{D}}_{\mathrm{x},i}^{\tau}\| \le s_\mathrm{min}^{(1)}$ with $s_\mathrm{min}^{(1)}$ from (4.117) and that $\mathbf{D}_\mathcal{X}\tilde{\mathbf{T}}_{0,k,j}^\mathrm{ad}$ is an approximation of $\mathbf{B}_2^\top$ in the second line.

Inserting these estimates into (4.174) yields

$$\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\| \le \omega^{-1}\bar{C}_\mathbf{f}.$$

Inserting this into (4.173), we arrive at

$$\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\right\| \le \gamma\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\| + \bar{C}_\mathbf{f}.$$

Applying this estimate recursively over j yields

$$\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\| \le \sum_{l=0}^{j-1}\gamma^l\bar{C}_\mathbf{f} + \gamma^j\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,0})\right\|.$$

By Theorem 4.72, the application of COARSEN and RECOMPRESS at the end of the outer loop does not influence the regularity estimates. Since we have at most I inner iteration steps, we have

$$\begin{aligned}\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\| &\le \sum_{l=0}^{j-1}\gamma^l\bar{C}_\mathbf{f} + \gamma^j\Big(\sum_{l=0}^{I-1}\gamma^l\bar{C}_\mathbf{f} + \gamma^I\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k-1,0})\right\|\Big)\\ &\le \bar{C}_\mathbf{f}\sum_{l=0}^{j+I-1}\gamma^l + \gamma^{j+I}\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k-1,0})\right\|.\end{aligned}$$

By definition of Algorithm 4.1 we have $\mathbf{w}_{0,0} = \mathbf{u}_0 = 0$. Applying the previous estimate recursively over k yields

$$\left\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\| \le \bar{C}_\mathbf{f}\sum_{l=0}^{j+kI-1}\gamma^l = \frac{\gamma^{kI+j}-1}{\gamma-1}\bar{C}_\mathbf{f}.$$

The estimate for $\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|$ follows by similar arguments. For completeness, the proof is given in Chapter A. □

To further estimate the right-hand side of (4.167) in Theorem 4.74, we need a bound for the maximal number of inner iteration steps I in Algorithm 4.1.

Lemma 4.75. *Under the assumptions of Theorem 4.69, there exists a d-independent constant c such that*

$$I \le c\ln(d),$$

where I is the maximal number of inner iteration steps of Algorithm 4.1.

Proof. We assume without loss of generality that $I \ge 1$. By the minimality condition of I in line 4 of Algorithm 4.1 we have

$$\rho^I I(1+\omega+\beta_1+\beta_2) \ge \rho^I(1+(\omega+\beta_1+\beta_2)(I-1)) \ge \frac{\rho}{2}\kappa_1.$$

By definition of κ_1 in (4.158) and Theorem 4.68, there exists a constant c_1 independent of d such that

$$c_1 d \ge 2\rho^{-1}(1+\omega+\beta_1+\beta_2)\kappa_1^{-1} \ge e^{\ln(\rho^{-1})I}I^{-1} \ge \frac{e\ln(\rho^{-1})}{2}e^{\frac{\ln(\rho^{-1})}{2}I},$$

where we used $e^{ax}x^{-1} \ge \frac{ea}{2}e^{\frac{a}{2}x}$ for $a > 0$ in the last step. Hence, we obtain

$$\tfrac{\ln(\rho^{-1})}{2}I \le \ln(\tfrac{2c_1}{e\ln(\rho^{-1})}) + \ln(d)$$

Choosing c appropriately yields the assertion. □

In the next step, we estimate the approximability of the iterates $\mathbf{w}_{k,j}$ in the $\mathcal{A}^s$-norm as well as the support size of the iterates. The following result is similar to the result in [BF24, Lemma 6.11], but with a better dependence on the additional logarithmic factors. This difference stems from the different treatment of the initial-value operator. Due to this different treatment, there is less coupling between the different quantities. In [BF24, Lemma 5.29], for the realization of the adaptive approximation of the initial value operator, the support as well as the compressibility of the contractions additionally depend on the maximal wavelet level. Therefore, these quantities must be estimated together and cannot be considered separately, as in the following result. The following result is analogous to the findings for the elliptic case in [BD16, Section 6.5]. In particular, we are interested in the dependence of these quantities on the number of spatial dimensions d and on the current tolerance $\eta_{k,j}$. To simplify the notation, we define

$$\hat{C}_{\mathbf{u},\mathbf{f}} = \max_{i=1,\dots,d} \max\Big\{1, \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}^{\frac{1}{s}}, \|\pi^{(\mathrm{t},i)}(\mathbf{f}_1)\|_{\mathcal{A}^s}^{\frac{1}{s}}, \|\pi^{(i)}(\mathbf{f}_2)\|_{\mathcal{A}^s}^{\frac{s+1}{s}}\Big\},$$

which is independent of d by Assumption 4.

Lemma 4.76. *Let $\mathbf{w}_{k,j}$ be defined by Algorithm 4.1 and let $\varepsilon_0 \le 1$. Under the assumptions of Theorem 4.69, one has*

$$\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|_{\mathcal{A}^s} \le \bar{C}_1(\tilde{C}d^2)^j d^{p_0}\big(|\ln \eta_{k,j}| + (\ln d)\big)^s \hat{C}_{\mathbf{u},\mathbf{f}}^s, \tag{4.176}$$

$$\sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})) \le \bar{C}_2(\tilde{C}d^2)^{\frac{j}{s}} d^{p_1}\big(|\ln \eta_{k,j}| + (\ln d)\big)\eta_{k,j}^{-\frac{1}{s}} \hat{C}_{\mathbf{u},\mathbf{f}}, \tag{4.177}$$

where $p_0 = 2 + \max\{1, s\}$, $p_1 = \max\{2 + 3s^{-1}, 1 + 4s^{-1}\}$ and $\tilde{C}, \bar{C}_1, \bar{C}_2$ are d-independent constants.

Proof. The statements are obviously true for $\mathbf{w}_{0,0} = \mathbf{u}_0 = 0$. Using the definition of Algorithm 4.1 for $\mathbf{w}_{k,0} = \mathbf{u}_k$ with $k \ge 1$, we have by (4.30) in Theorem 4.12 the estimate

$$\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,0})\|_{\mathcal{A}^s} \le C_1 d^{1+\max\{1,s\}} \sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s} \le C_1 d^{p_0} \hat{C}_{\mathbf{u},\mathbf{f}}^s,$$

and

$$\sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,0})) \le C_2 d(2^{-k}\varepsilon_0)^{-\frac{1}{s}} \Big(\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{u})\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}} \le C_3 d^{p_1} \eta_{k,0}^{-\frac{1}{s}} \hat{C}_{\mathbf{u},\mathbf{f}},$$

where we used the definition of κ_{C} and κ_{P}. The idea of the proof is to show the estimate inductively over the inner iteration index j for fixed outer iteration index k. We start with the proof of (4.176). Based on the stability of the contractions under RECOMPRESS and COARSEN, we have

$$\sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\|_{\mathcal{A}^s} \le 2^s \sum_{i=1}^{d} \|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|_{\mathcal{A}^s} + \omega \|\pi^{(\mathrm{t},i)}(\mathbf{r}_{k,j})\|_{\mathcal{A}^s}, \tag{4.178}$$

where we used Theorem 2.8 (ii). Using again this property and the definition of $\mathbf{r}_{k,j}$ in Algorithm 4.1, we obtain

$$\sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}(\mathbf{r}_{k,j}) \big\|_{\mathcal{A}^s} \leq 2^s \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}\big)\big) \big\|_{\mathcal{A}^s} + \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{RHS}\big(\tfrac{1}{2}\eta_{k,j}\big)\big) \big\|_{\mathcal{A}^s}. \quad (4.179)$$

In the next steps, we estimate the two terms on the right-hand side separately. Using Theorem 2.8 (ii) as well as (4.152), we have

$$\sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}\big)\big) \big\|_{\mathcal{A}^s} \leq 2^s \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_1}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}^{(1)}\big)\big) \big\|_{\mathcal{A}^s} + \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_2}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}^{(2)}\big)\big) \big\|_{\mathcal{A}^s},$$

where $\eta_{k,j}^{(1)} = \lambda\eta_{k,j}$ and $\eta_{k,j}^{(2)} = (1-\lambda)\eta_{k,j}$ for some $\lambda \in (0,1)$. Using the definition of $\mathrm{APPLY}_{\mathbf{A}_1}$ in (4.150) as well as Theorem 4.43 or Theorem 4.44, respectively, we obtain

$$\sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_1}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}^{(1)}\big)\big) \big\|_{\mathcal{A}^s} \lesssim d^2 \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j}) \big\|_{\mathcal{A}^s},$$

where the hidden constant is independent of d and $\eta_{k,j}$. Moreover, by Theorem 4.54, we have

$$\sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_2}\big(\mathbf{w}_{k,j}, \tfrac{1}{2}\eta_{k,j}^{(2)}\big)\big) \big\|_{\mathcal{A}^s} \lesssim d \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j}) \big\|_{\mathcal{A}^s}.$$

Next, we bound the second term on the right-hand side of (4.179). For this purpose, we define

$$\mathbf{f}_{k,j}^{(1)} = \mathrm{RHS}_{\mathbf{f}_1}\left(\frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\right), \quad \mathbf{f}_{k,j}^{(2)} = \mathrm{RHS}_{\mathbf{f}_2}\left(\frac{\eta_{k,j}^{(2)}}{4\|\mathbf{B}_2\|}\right).$$

Then by definition of RHS in (4.154) and Theorem 2.8 (ii), we have

$$\sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{RHS}\big(\tfrac{1}{2}\eta_{k,j}\big)\big) \big\|_{\mathcal{A}^s} \leq 2^s \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{B}_1^\top}\big(\mathbf{f}_{k,j}^{(1)}, \tfrac{1}{4}\eta_{k,j}^{(1)}\big)\big) \big\|_{\mathcal{A}^s} + \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{B}_2^\top}\big(\mathbf{f}_{k,j}^{(2)}, \tfrac{1}{4}\eta_{k,j}^{(2)}\big)\big) \big\|_{\mathcal{A}^s}. \quad (4.180)$$

Again using Theorem 4.43 or Theorem 4.44, respectively, yields

$$\begin{aligned} \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{B}_1^\top}\big(\mathbf{f}_{k,j}^{(1)}, \tfrac{1}{4}\eta_{k,j}^{(1)}\big)\big) \big\|_{\mathcal{A}^s} &\lesssim d \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}\big(\mathbf{f}_{k,j}^{(1)}\big) \big\|_{\mathcal{A}^s} \\ &\lesssim d \sum_{i=1}^{d} \big\| \pi^{(\mathrm{t},i)}(\mathbf{f}_1) \big\|_{\mathcal{A}^s} \\ &\leq d^2 \hat{C}_{\mathbf{u},\mathbf{f}}^s, \end{aligned}$$

where we used Assumptions 2 and 4 in the second step. For the second term in (4.180), we first estimate

$$\begin{aligned} 2L_{\mathrm{x}}\big(\mathbf{f}_{k,j}^{(2)}\big) + J_0^{\mathrm{ad}}\big(\tfrac{1}{4}\eta_{k,j}^{(2)}; \mathbf{f}_{k,j}^{(2)}\big) &\lesssim L_{\mathrm{x}}\big(\mathbf{f}_{k,j}^{(2)}\big) + |\ln(\tfrac{1}{4}\eta_{k,j}^{(2)})| + \ln\Big(\sum_{i=1}^{d}\big\|\pi^{(i)}\big(\mathbf{f}_{k,j}^{(2)}\big)\big\|\Big) \\ &\lesssim \ln(d) + |\ln(\eta_{k,j})| + \ln\Big(\max_{i=1,\dots,d}\big\|\pi^{(i)}\big(\mathbf{f}_{k,j}^{(2)}\big)\big\|\Big) \\ &\lesssim (\ln(d) + |\ln(\eta_{k,j})|)\max\Big\{1, \max_{i=1,\dots,d}\big\|\pi^{(i)}\big(\mathbf{f}_{k,j}^{(2)}\big)\big\|\Big\} \\ &\leq (\ln(d) + |\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}}, \end{aligned} \tag{4.181}$$

where we have used (4.142) from Theorem 4.61 in the first line and Assumptions 2 and 4 as well as $\ln(d) > 1$ in the second line, as well as the definition of $\hat{C}_{\mathbf{u},\mathbf{f}}$ in the last line. Combining this estimate with (4.141) yields

$$\begin{aligned} \sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{B}_2^\top}\big(\mathbf{f}_{k,j}^{(2)}, \tfrac{1}{4}\eta_{k,j}^{(2)}\big)\big)\big\|_{\mathcal{A}^s} &\lesssim (\ln(d) + |\ln(\eta_{k,j})|)^s \hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}s}\sum_{i=1}^{d}\big\|\pi^{(i)}\big(\mathbf{f}_{k,j}^{(2)}\big)\big\| \\ &\lesssim d(\ln(d) + |\ln(\eta_{k,j})|)^s \hat{C}_{\mathbf{u},\mathbf{f}}^{s}. \end{aligned}$$

Inserting all of these estimates into (4.178), we arrive at

$$\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\big\|_{\mathcal{A}^s} \leq C_3 d^2 \sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s} + C_3 d^2(\ln(d) + |\ln(\eta_{k,j})|)^s \hat{C}_{\mathbf{u},\mathbf{f}}^{s}$$

for a suitable constant C_3 independent of d and $\eta_{k,j}$. Next, we show that for all $j \geq 0$,

$$\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s} \leq C_4 d^{p_0}\hat{C}_{\mathbf{u},\mathbf{f}}^{s}(\ln(d) + |\ln(\eta_{k,j})|)^s (C_3 d^2 + 1)^j$$

with $C_4 = \max\{C_1, C_3\}$. The statement is true for $j = 0$ as shown at the beginning of the proof. Using the recursion above, we have

$$\begin{aligned} \sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\big\|_{\mathcal{A}^s} &\leq C_4 d^{p_0}\hat{C}_{\mathbf{u},\mathbf{f}}^{s}(\ln(d) + |\ln(\eta_{k,j})|)^s ((C_3 d^2 + 1)^j (C_3 d^2) + 1) \\ &\leq C_4 d^{p_0}\hat{C}_{\mathbf{u},\mathbf{f}}^{s}(\ln(d) + |\ln(\eta_{k,j+1})|)^s (C_3 d^2 + 1)^{j+1}, \end{aligned}$$

where we used that $\varepsilon_0 \leq 1$ and thus $\eta_{k,j} \leq 1$. Setting $\tilde{C} = C_3 + 1$ and $\tilde{C}_1 = C_4$ yields the assertion (4.176).

Next, we consider the support estimate. We define the intermediate result before coarsening and recompression as

$$\bar{\mathbf{w}}_{k,j+1} = \mathbf{w}_{k,j} - \omega \mathbf{r}_{k,j}.$$

Applying the routines COARSEN and RECOMPRESS can only decrease the support. Hence, we obtain

$$\begin{aligned} \sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})) &\leq \sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\bar{\mathbf{w}}_{k,j+1})) \\ &\leq \sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})) + \sum_{i=1}^{d}\#\mathrm{supp}(\pi^{(\mathrm{t},i)}(\mathbf{r}_{k,j})). \end{aligned} \tag{4.182}$$

Furthermore, we have

$$\begin{aligned}\#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{r}_{k,j}))&\\ \leq \#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j},\tfrac{\eta_{k,j}}{2}\big)\big)\big)&+\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{RHS}\big(\tfrac{\eta_{k,j}}{2}\big)\big)\big)\end{aligned}\tag{4.183}$$

for each $i = 1,\dots,d$. As before, we estimate the two terms on the right-hand side separately. Using (4.107) from Theorem 4.43 as well as Theorem 4.44, we obtain

$$\begin{aligned}\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_1}\big(\mathbf{w}_{k,j},\tfrac{\eta_{k,j}^{(1)}}{2}\big)\big)\big) &\lesssim d\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\\ &\lesssim d^{1+\frac{1}{s}}\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}},\end{aligned}$$

where $\mathbf{z}_{k,j}$ is defined in (4.164). Additionally, using Theorem 4.54, we have

$$\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{A}_2}\big(\mathbf{w}_{k,j},\tfrac{\eta_{k,j}^{(2)}}{2}\big)\big)\big)\lesssim d\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}.$$

Combining the previous two estimates with the definition of APPLY in (4.152) yields

$$\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j},\tfrac{\eta_{k,j}}{2}\big)\big)\big)\lesssim d^{1+\frac{1}{s}}\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}.$$

Inserting this estimate into (4.176) and using the fact that $p_1 = 1 + s^{-1} + s^{-1}p_0$, we obtain

$$\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j},\tfrac{\eta_{k,j}}{2}\big)\big)\big)\leq C_5 d^{p_1}\eta_{k,j}^{-\frac{1}{s}}(\tilde{C}d^2)^{\frac{j}{s}}(\ln(d)+|\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}}\tag{4.184}$$

for some d-independent constant C_5. Next, we estimate the support of the contractions arising from the routine RHS. Using again (4.107) from Theorem 4.43, we obtain

$$\begin{aligned}\sum_{i=1}^{d}\#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}\big(\mathrm{APPLY}_{\mathbf{B}_1^\top}\big(\mathbf{f}_{k,j}^{(1)},\tfrac{1}{4}\eta_{k,j}^{(1)}\big)\big)\big)&\lesssim d\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}\big(\mathbf{f}_{k,j}^{(1)}\big)\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}\\ &\lesssim d\eta_{k,j}^{-\frac{1}{s}}\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{f}_1)\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}}\\ &\leq d^{1+\frac{1}{s}}\eta_{k,j}^{-\frac{1}{s}}\hat{C}_{\mathbf{u},\mathbf{f}},\end{aligned}$$

where we used Assumptions 2 and 4 in the second line and the definition of $\hat{C}_{\mathbf{u},\mathbf{f}}$ in the last line. Moreover, using (4.140) and (4.181), we obtain

$$\begin{aligned}
&\sum_{i=1}^{d} \#\operatorname{supp}\Big(\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{B}_2^\top}\Big(\mathbf{f}_{k,j}^{(2)}, \tfrac{\eta_{k,j}^{(2)}}{4}\Big)\Big)\Big) \\
&\qquad \lesssim (\ln(d) + |\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}} \sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(i)}(\mathbf{f}_{k,j}^{(2)})\big) \\
&\qquad \lesssim d\eta_{k,j}^{-\frac{1}{s}}(\ln(d) + |\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}} \Big(\sum_{i=1}^{d} \|\pi^{(i)}(\mathbf{f}_2)\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}} \\
&\qquad \lesssim d^{1+s^{-1}}\eta_{k,j}^{-\frac{1}{s}}(\ln(d) + |\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}},
\end{aligned}$$

where we used Assumptions 2 and 4 in the second line and the definition of $\hat{C}_{\mathbf{u},\mathbf{f}}$ in the last line. Combining the last two estimates with the definition of the routine RHS yields

$$\sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\mathrm{RHS}(\tfrac{\eta_{k,j}}{2}))\big) \le C_6 d^{1+s^{-1}}\eta_{k,j}^{-\frac{1}{s}}(\ln(d) + |\ln(\eta_{k,j})|)\hat{C}_{\mathbf{u},\mathbf{f}} \tag{4.185}$$

for some d-independent constant C_6.

We now proceed by induction over j. Let statement (4.177) be true for a fixed j with $\hat{C}_2 = \max\{C_5, C_6\}$ and $\tilde{C} \ge d^{-2}3^s$, which we can choose without loss of generality. By inserting (4.183) as well as (4.184) and (4.185) into (4.182), we obtain

$$\begin{aligned}
\sum_{i=1}^{d} \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\bar{\mathbf{w}}_{k,j+1})) &\le 3\bar{C}_2(\tilde{C}d^2)^{\frac{j}{s}} d^{p_1}\big(|\ln \eta_{k,j}| + (\ln d)\big)\eta_{k,j}^{-\frac{1}{s}}\hat{C}_{\mathbf{u},\mathbf{f}} \\
&\le \bar{C}_2(\tilde{C}d^2)^{\frac{j+1}{s}} d^{p_1}\big(|\ln \eta_{k,j+1}| + (\ln d)\big)\eta_{k,j+1}^{-\frac{1}{s}}\hat{C}_{\mathbf{u},\mathbf{f}},
\end{aligned}$$

where we again used that $\eta_{k,j} \le 1$. Since the statement is true for $j = 0$ as shown at the beginning of the proof, this concludes the induction and thus the proof. □

Remark 4.77. We assume $\varepsilon_0 \le 1$ in Theorem 4.76 for notionally brevity to ensure that $\eta_{k,j} \le 1$ for all k, j. The assumption is not restrictive, as we are interested in the regime, where the tolerances are small, since the computational cost of the algorithm is dominated by the iterates with small tolerances. If $\varepsilon_0 > 1$, we can proceed similarly and use for $\eta_{k,j} > 1$ the estimate $|\ln(\eta_{k,j})| \le |\ln(\varepsilon_0)|$, where we note that ε_0 is independent of d by definition of ε_0 in Algorithm 4.1 and Assumption 4. For simplicity of the proofs, we only consider the case $\varepsilon_0 \le 1$ in the following.

In the proof of this lemma, we have shown that the support estimate (4.177) is not only true for the iterates $\mathbf{w}_{k,j+1}$ but also for the intermediate quantity

$$\bar{\mathbf{w}}_{k,j+1} = \mathbf{w}_{k,j} - \omega \mathbf{r}_{k,j}$$

before the application of COARSEN and RECOMPRESS. Having in mind Theorem 4.70, this is important for the estimation of the maximal wavelet level of the iterates, which we consider next.

Lemma 4.78. *Let $\mathbf{w}_{k,j}$ be defined by Algorithm 4.1 and let $\varepsilon_0 \le 1$. Under the assumptions of Theorem 4.69, one has*

$$L(\mathbf{w}_{k,j}) \le C((\ln d)^2|\ln \eta_{k,j}| + (\ln d)^3) \tag{4.186}$$

for some d-independent constant C.

Proof. For the proof we follow the lines of [BF24, Lemma 6.11] and [BD16, Section 6.5]. The main idea is to combine Theorem 4.70, Theorem 4.74, and Theorem 4.76 in a suitable way.

The statement is obviously true for $\mathbf{w}_{0,0} = \mathbf{u}_0 = 0$. In the first step, we show that the statement is true for $\mathbf{w}_{k,j}$ for all $j \geq 1$. Using $\beta_2 > 0$ in Algorithm 4.1 we have by Theorem 4.70 and Theorem 4.74 the estimate

$$\begin{aligned} L(\mathbf{w}_{k,j+1}) &\leq \tau^{-1} \log_2 \Big(C_1 (\beta_2 \eta_{k,j})^{-1} \gamma^{kI+j+1} \Big(\sum_{i=1}^{d} \#\operatorname{supp}\big(\pi^{(\mathrm{t},i)}(\bar{\mathbf{w}}_{k,j+1})\big)\Big)^{\frac{1}{2}} \Big) \\ &\leq \tau^{-1} \log_2 \Big(C_2 \eta_{k,j}^{-1} \gamma^{kI+j+1} \bar{C}_{\mathbf{f}} (\tilde{C} d^2)^{\frac{j+1}{2s}} d^{\frac{p_1}{2}} (|\ln \eta_{k,j+1}| + \ln d)^{\frac{1}{2}} \eta_{k,j+1}^{-\frac{1}{2s}} \hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{1}{2}} \Big), \end{aligned} \tag{4.187}$$

where we additionally used (4.177) from Theorem 4.76 in the second step. Since γ from (4.168) depend on d only by $\gamma_{\mathrm{x},\mathbf{B}_1}$, there exists a d-independent constant C_3 such that

$$\gamma \leq C_3 d^2.$$

Furthermore, by Theorem 4.75, we have

$$j + 1 \leq I \lesssim \ln(d).$$

Moreover, by definition of $\eta_{k,j}$ in Algorithm 4.1, we have

$$k = \log_2(\eta_{k,j}^{-1} \rho^{j+1} \varepsilon_0) \leq |\log_2(\eta_{k,j})| + (j+1)|\log_2(\rho)| + |\log_2(\varepsilon_0)| \lesssim (|\ln \eta_{k,j}| + \ln d),$$

where we used Assumption 4. Combining everything, we have

$$\log_2(\gamma^{kI+j+1}) = (kI + j + 1) \log_2(\gamma) \leq C_4((\ln d)^3 + (\ln d)^2 |\ln(\eta_{k,j})|)$$

for some d-independent constant C_4. Inserting this into (4.187) and using logarithmic laws yields

$$L(\mathbf{w}_{k,j+1}) \leq C_5((\ln d)^3 + (\ln d)^2 |\ln(\eta_{k,j+1})|)$$

for some d-independent constant C_5, where we used that $\varepsilon_0 \leq 1$ and thus $\eta_{k,j} \leq 1$.

In the final step, we show that the statement is true for $\mathbf{w}_{k,0}$ with $k \geq 1$ as well. Here we can proceed exactly as before, using the fact that we apply COARSEN at the end of each outer loop. Furthermore, by definition we have

$$|\ln \eta_{k,I}| = \Big|\ln\Big(\frac{\rho^I}{2}\Big) + \ln(\eta_{k+1,0})\Big| \lesssim (\ln(d) + |\ln \eta_{k+1,0}|).$$

Then proceeding as before using Theorem 4.70, combined with the estimate above yields the assertion. □

4.4.2.2 Hierarchical Ranks of Iterates

In the following part, we estimate the hierarchical ranks of the iterates $\mathbf{w}_{k,j}$ in Algorithm 4.1. To do this, we make use of the rank bounds in Theorems 4.43, 4.54 and 4.61. In the bounds for the ranks, there appear the quantities

$$J\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)}; \mathbf{w}_{k,j}\big), \quad J\big(\tfrac{1}{4}\eta_{k,j}^{(1)}, \mathbf{z}_{k,j}\big), \quad J_0\big(\tfrac{1}{2}\eta_{k,j}^{(2)}; \mathbf{w}_{k,j}\big)$$

and the quantities related to scaling matrices

$$\big\|\hat{\mathrm{m}}_{\mathcal{X}}\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty},\quad \big\|\hat{\mathrm{m}}_{\mathcal{Y}}\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty},\quad \big\|\hat{\mathrm{m}}_{\mathcal{X},0}\big(\tfrac{1}{2}\eta_{k,j}^{(2)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty},$$

as well as

$$\big\|\hat{\mathrm{m}}_{\mathcal{X}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)};\mathbf{z}_{k,j}\big)\big\|_{\ell_\infty},\quad \big\|\hat{\mathrm{m}}_{\mathcal{Y}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)};\mathbf{z}_{k,j}\big)\big\|_{\ell_\infty}.$$

The quantities above can be bounded by the maximal wavelet level of the iterates and the $\mathcal{A}^s$-norm of the contractions as shown in Theorems 4.48 and 4.56 as well as in (4.121) and (4.135).

Lemma 4.79. *Let $\mathbf{w}_{k,j}$ be defined by Algorithm 4.1 and let $\varepsilon_0 \le 1$. Under the assumptions of Theorem 4.69, one has*

$$\operatorname{rank}_\infty(\mathbf{w}_{k,j}) \le C d^{p_2} (\ln d)^{14j} \big(|\ln \eta_{k,j}| + (\ln d)\big)^{8j} (\ln(d) + |\ln \eta_{k,j}|)^b \tag{4.188}$$

with $b = \max\{b_{\mathbf{u}}, b_{\mathbf{f}}\}$ with $b_{\mathbf{u}}$ and $b_{\mathbf{f}}$ from Assumptions 2 and 3, respectively, and where p_2, C are d-independent constants.

Proof. The proof follows the ideas of [BF24, Lemma 6.12] and [BD16, Section 6.5]. Clearly, the statement is true for $\mathbf{w}_{0,0} = \mathbf{u}_0 = 0$. In the first step, we show that the statement is true for all $\mathbf{w}_{k,0}$ with $k \ge 1$. By Assumption 3 (i) and the definition of the left-continuous inverse, we have

$$\gamma_{\mathbf{u}}^{-1}(m) = \Big\lceil \big(c_{\mathbf{u}}^{-1} \ln(m)\big)^{b_{\mathbf{u}}} \Big\rceil. \tag{4.189}$$

Furthermore, as seen in the proof of Theorem 4.66, at the end of the inner iteration, we have

$$\|\mathbf{w}_{k,j} - \mathbf{u}\| \le \kappa_1 2^{-(k+1)} \varepsilon_0.$$

By the definition of $\mathbf{w}_{k+1,0} = \mathbf{u}_{k+1}$ in Algorithm 4.1, we are able to apply Theorem 4.12, which yields

$$\begin{aligned}
\operatorname{rank}_\infty(\mathbf{w}_{k+1,0}) &\le \gamma_{\mathbf{u}}^{-1}\Big(\|\mathbf{u}\|_{\mathcal{A}(\gamma_{\mathbf{u}})}\big(\alpha\kappa_1 2^{-k+1}\varepsilon_0\big)^{-1}\Big) \\
&\le \Big\lceil \big(c_{\mathbf{u}}^{-1} \ln\big(\|\mathbf{u}\|_{\mathcal{A}(\gamma_{\mathbf{u}})}\rho\alpha^{-1}\kappa_1^{-1}\eta_{k+1,0}^{-1}\big)\big)^{b_{\mathbf{u}}} \Big\rceil \\
&\le C_1 (\ln(d) + |\ln(\eta_{k+1,0})|)^{b_{\mathbf{u}}},
\end{aligned}$$

for some d-independent constant C_1, where we used (4.189) and the definition of $\eta_{k+1,0}$ in the second line as well as $\kappa_1^{-1} \lesssim d$ and Assumption 4 for the d-independence of the other constants.

In the next step, we show that the statement is true for all $\mathbf{w}_{k,j}$ with $j \ge 1$. First, we mention that by (4.176) from Theorem 4.76 we have

$$\ln\Big(\sum_{i=1}^{d} \big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s}\Big) \lesssim (\ln d)^2 + |\ln \eta_{k,j}| \tag{4.190}$$

for a sufficiently large constant, where we used $j \le I$ combined with Theorem 4.75 as well as $\ln(d) \ge 1$. Inserting this into (4.121) and (4.135) yields

$$J\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big) \lesssim (\ln d)^2 + |\ln \eta_{k,j}|, \tag{4.191}$$

$$J_0\big(\tfrac{1}{2}\eta_{k,j}^{(2)};\mathbf{w}_{k,j}\big) \lesssim (\ln d)^2 + |\ln \eta_{k,j}|. \tag{4.192}$$

Moreover, combining Theorem 4.44 with the definition of $\mathbf{z}_{k,j}$ in (4.164) as well as (4.190), we correspondingly have

$$\ln\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\big\|_{\mathcal{A}^s}\Big) \lesssim (\ln d)^2 + |\ln \eta_{k,j}|$$

and thus

$$J\big(\tfrac{1}{4}\eta_{k,j}^{(1)}, \mathbf{z}_{k,j}\big) \lesssim (\ln d)^2 + |\ln \eta_{k,j}|. \tag{4.193}$$

By inserting (4.186) from Theorem 4.78 as well as (4.190) into (4.118) and (4.119) from Theorem 4.48 yields

$$\begin{aligned}\big\|\hat{\mathrm{m}}_{\mathcal{X}}\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty} &\lesssim \big((\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3\big)^2,\\ \big\|\hat{\mathrm{m}}_{\mathcal{Y}}\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty} &\lesssim (\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3,\end{aligned} \tag{4.194}$$

where we again use $\ln(d) \geq 1$. Furthermore, by Theorem 4.56, we obtain

$$\big\|\hat{\mathrm{m}}_{\mathcal{X},0}\big(\tfrac{1}{2}\eta_{k,j}^{(2)};\mathbf{w}_{k,j}\big)\big\|_{\ell_\infty} \lesssim \big((\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3\big)^2. \tag{4.195}$$

Additionally, we have to estimate the quantities with respect to the intermediate results $\mathbf{z}_{k,j}$. By the level decay property, we obtain, as in Theorem 4.48, the estimate

$$L(\mathbf{z}_{k,j}) \leq L(\mathbf{w}_{k,j}) + C_2 J\big(\tfrac{1}{4\|\mathbf{B}_1\|}\eta_{k,j}^{(1)};\mathbf{w}_{k,j}\big)$$

for some suitable constant C_2 independent of d and $\eta_{k,j}$. Combining this result with Theorem 4.48 as well as Theorem 4.44 and (4.191), we obtain

$$\begin{aligned}\big\|\hat{\mathrm{m}}_{\mathcal{X}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)};\mathbf{z}_{k,j}\big)\big\|_{\ell_\infty} &\lesssim \Big(1 + L(\mathbf{z}_{k,j}) + \big|\ln\big(\tfrac{1}{4}\eta_{k,j}^{(1)}\big)\big| + \ln\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\big\|_{\mathcal{A}^s}\Big)\Big)^2\\ &\lesssim \Big(L(\mathbf{w}_{k,j}) + |\ln(\eta_{k,j})| + (\ln d)^2 + \ln\Big(\sum_{i=1}^{d}\big\|\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|_{\mathcal{A}^s}\Big)\Big)^2\\ &\lesssim \big((\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3\big)^2,\end{aligned} \tag{4.196}$$

where we again used (4.186) from Theorem 4.78 and (4.190) in the last line. Proceeding similarly for $\hat{\mathrm{m}}_{\mathcal{Y}}$ yields

$$\big\|\hat{\mathrm{m}}_{\mathcal{Y}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)};\mathbf{z}_{k,j}\big)\big\|_{\ell_\infty} \lesssim (\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3. \tag{4.197}$$

As for the other quantities, we prove the rank bound (4.188) iteratively. We know by Section 2.4.5.1 that the ranks of a sum can be bounded by the sum of the ranks of the individual terms. Hence, we have

$$\begin{aligned}\operatorname{rank}_\infty(\mathbf{w}_{k,j+1}) \leq \operatorname{rank}_\infty(\mathbf{w}_{k,j}) &+ \operatorname{rank}_\infty\big(\mathrm{APPLY}\big(\mathbf{w}_{k,j}, \tfrac{\eta_{k,j}}{2}\big)\big)\\ &+ \operatorname{rank}_\infty\big(\mathrm{RHS}\big(\tfrac{\eta_{k,j}}{2}\big)\big).\end{aligned} \tag{4.198}$$

Combining the estimations above with the rank estimates in Theorems 4.43 and 4.54 and with the definition of APPLY in (4.152) yields

$$\begin{aligned}&\operatorname{rank}_\infty\left(\text{APPLY}(\mathbf{w}_{k,j}, \tfrac{\eta_{k,j}}{2})\right)\\ &\quad\le C_2\left((\ln d)^2|\ln\eta_{k,j}| + (\ln d)^3\right)^6\left((\ln d)^2 + |\ln\eta_{k,j}|\right)^2\operatorname{rank}_\infty(\mathbf{w}_{k,j})\\ &\quad\le C_2(\ln d)^{14}(\ln(d) + |\ln\eta_{k,j}|)^8\operatorname{rank}_\infty(\mathbf{w}_{k,j})\end{aligned} \tag{4.199}$$

for some suitable d-independent constant C_2. In the next step, we estimate the ranks arising from the routine RHS. We mention again that the ranks of a sum are bounded by the sum of the ranks of the individual terms and consider both terms on the right-hand side of the definition of RHS in (4.154) separately. For this purpose, we define as before

$$\mathbf{f}_{k,j}^{(1)} = \text{RHS}_{\mathbf{f}_1}\Big(\tfrac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\Big), \quad \mathbf{f}_{k,j}^{(2)} = \text{RHS}_{\mathbf{f}_2}\Big(\tfrac{\eta_{k,j}^{(2)}}{4\|\mathbf{B}_2\|}\Big).$$

Combining the estimate (4.121) with Assumptions 2 and 4 we obtain

$$\begin{aligned}J\big(\tfrac{1}{4}\eta_{k,j}^{(1)}, \mathbf{f}_{k,j}^{(1)}\big) &\lesssim 1 + |\ln\eta_{k,j}| + \ln\Big(\sum_{i=1}^d \big\|\pi^{(\mathrm{t},i)}\big(\mathbf{f}_{k,j}^{(1)}\big)\big\|_{\mathcal{A}^s}\Big)\\ &\lesssim \ln(d) + |\ln\eta_{k,j}|.\end{aligned} \tag{4.200}$$

Furthermore, using Theorem 4.48 and combining it with Assumptions 2 and 4, we obtain

$$\begin{aligned}\big\|\hat{\mathrm{m}}_{\mathcal{X}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)}; \mathbf{f}_{k,j}^{(1)}\big)\big\|_{\ell_\infty} &\lesssim (\ln(d) + |\ln\eta_{k,j}|)^2,\\ \big\|\hat{\mathrm{m}}_{\mathcal{Y}}\big(\tfrac{1}{4}\eta_{k,j}^{(1)}; \mathbf{f}_{k,j}^{(1)}\big)\big\|_{\ell_\infty} &\lesssim \ln(d) + |\ln\eta_{k,j}|.\end{aligned} \tag{4.201}$$

Now applying (4.109) together with Assumptions 2 and 4, yields

$$\operatorname{rank}_\infty\left(\text{APPLY}_{\mathbf{B}_1^\top}\big(\mathbf{f}_{k,j}^{(1)}, \tfrac{1}{4}\eta_{k,j}^{(1)}\big)\right) \lesssim (\ln(d) + |\ln\eta_{k,j}|)^4(1 + |\ln\eta_{k,j}|)^{b_\mathbf{f}}.$$

For the second term on the right-hand side, we have analogously by Theorem 4.63 as well as Assumptions 2 and 4 the estimate

$$\big\|\hat{\mathrm{m}}_{\mathcal{X},0}^{\mathrm{ad}}\big(\tfrac{1}{4}\eta_{k,j}^{(2)}; \mathbf{f}_{k,j}^{(2)}\big)\big\|_{\ell_\infty} \lesssim (\ln(d) + |\ln\eta_{k,j}|)^2. \tag{4.202}$$

Hence, using (4.143) from Theorem 4.62, we have

$$\begin{aligned}\operatorname{rank}_\infty\left(\text{APPLY}_{\mathbf{B}_2^\top}\big(\mathbf{f}_{k,j}^{(2)}, \tfrac{1}{4}\eta_{k,j}^{(2)}\big)\right) &\lesssim (\ln(d) + |\ln\eta_{k,j}|)^2\big|\operatorname{rank}\big(\mathbf{f}_{k,j}^{(2)}\big)\big|_\infty\\ &\lesssim (\ln(d) + |\ln\eta_{k,j}|)^2(1 + |\ln\eta_{k,j}|)^{b_\mathbf{f}}.\end{aligned}$$

Combining the two estimates for the right-hand side of RHS, we obtain

$$\operatorname{rank}_\infty\left(\text{RHS}\big(\tfrac{\eta_{k,j}}{2}\big)\right) \le C_3(\ln(d) + |\ln\eta_{k,j}|)^4(1 + |\ln\eta_{k,j}|)^{b_\mathbf{f}} \tag{4.203}$$

for some suitable d-independent constant C_3.

In the next step, we show inductively that

$$\operatorname{rank}_\infty(\mathbf{w}_{k,j}) \le C_1C_4^j(\ln d)^{14j}\big(|\ln\eta_{k,j}| + (\ln d)\big)^{8j}(\ln(d) + |\ln\eta_{k,j}|)^b$$

with $C_4 = 1 + C_2 + C_1^{-1}C_3$ and $b = \max\{b_\mathbf{u}, b_\mathbf{f}\}$ with $b_\mathbf{u}$ and $b_\mathbf{f}$ from Assumptions 2 and 3, respectively. The statement is true for $j = 0$ as shown at the beginning of

the proof. Inserting the estimates (4.199) and (4.203) into (4.198) shows that the statement is true for $j+1$ if it is true for j, since the application of COARSEN and RECOMPRESS can only decrease the ranks. Hence, the above estimation holds for all $j \geq 0$. In the final step, we use that by Theorem 4.75 we have $j \leq I \leq c\ln(d)$ for some d-independent constant c. Furthermore, we have

$$C_4^j \leq C_4^{c\ln(d)} = d^{c\ln(C_4)}.$$

Setting $C = C_1$ and $p_2 = c\ln(C_4)$ concludes the proof. □

4.4.2.3 Complexity of Algorithm 4.1

Using the derived bounds on the quantities of the iterates, we can now prove a bound on the computational complexity of the algorithm.

Proof of Theorem 4.69. By Theorem 4.77, we consider without loss of generality only the case $\varepsilon_0 \leq 1$. Let $\mathbf{u}_\varepsilon = \mathbf{u}_{k+1}$ for some $k \in \mathbb{N}_0$. Then by Theorem 4.66 and the termination condition of Algorithm 4.1, we have

$$\|\mathbf{u} - \mathbf{u}_{k+1}\| \leq 2^{-k+1}\varepsilon_0 \leq \varepsilon \leq 2^{-k}\varepsilon_0.$$

Furthermore, there exists a $j \leq I$ with

$$\mathbf{u}_{k+1} = \mathrm{COARSEN}\big(\mathrm{RECOMPRESS}\big(\mathbf{w}_{k,j}, \kappa_2 2^{-(k+1)}\varepsilon_0\big), \kappa_3 2^{-(k+1)}\varepsilon_0\big).$$

As one can see in the proof of Theorem 4.66, we have

$$\|\mathbf{u} - \mathbf{w}_{k,j}\| \leq \kappa_1 2^{-k+1}\varepsilon_0 \leq \kappa_1\varepsilon.$$

By definition of κ_2 and κ_3 we are now able to apply Theorem 4.12 with $\eta = \kappa_1 2^{-(k+1)}\varepsilon_0$. Hence, the statements (4.161) and (4.162) are a direct consequence of (4.29) and Assumption 3. Furthermore, (4.159) and (4.160) are a direct consequence of (4.30) and Assumption 3.

For the computational complexity, we proceed as follows. First, we estimate the costs for each inner iteration step. Then we estimate the costs for each outer-loop step and subsequently for the total algorithm.

The computational cost for each inner iteration step consists of three parts: the costs of RHS, APPLY, and the combined costs of COARSEN and RECOMPRESS. We start with the costs for RHS. These costs again consist, by the definition of RHS in (4.154), of three parts. The first part is $\mathrm{flops}(\mathbf{f}_{k,j})$ for computing $\mathbf{f}_{k,j} = \big(\mathbf{f}_{k,j}^{(1)}, \mathbf{f}_{k,j}^{(2)}\big)^\top$, where

$$\mathbf{f}_{k,j}^{(1)} = \mathrm{RHS}_{\mathbf{f}_1}\Big(\tfrac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\Big), \quad \mathbf{f}_{k,j}^{(2)} = \mathrm{RHS}_{\mathbf{f}_2}\Big(\tfrac{\eta_{k,j}^{(2)}}{4\|\mathbf{B}_2\|}\Big).$$

Using Assumption 2 (ii) as well as Assumption 4 (iii), we obtain

$$\mathrm{flops}(\mathbf{f}_{k,j}) \lesssim d^q\big((1+|\ln\eta_{k,j}|)^{3b_\mathbf{f}} + (1+|\ln\eta_{k,j}|)^{b_\mathbf{f}}\eta_{k,j}^{-\frac{1}{s}}\big)$$

for some suitable $q > 0$. Furthermore, using Theorem 4.46 as well as (4.200) and (4.201), we obtain

$$\begin{aligned}\operatorname{flops}\Big(\operatorname{APPLY}_{\mathbf{B}_1^\top}\Big(\mathbf{f}_{k,j}^{(1)}, \tfrac{\eta_{k,j}^{(1)}}{4}\Big)\Big) &\lesssim (\ln(d) + |\ln \eta_{k,j}|)^{13} \operatorname{rank}_\infty^3\big(\mathbf{f}_{k,j}^{(1)}\big) \sum_{i=1}^d \#\operatorname{supp}(\pi^{(\mathrm{t},i)}(\mathbf{f}_{k,j}^{(1)})) \\ &\quad + d\eta_{k,j}^{-\frac{1}{s}} (\ln(d) + |\ln \eta_{k,j}|)^4 \operatorname{rank}_\infty\big(\mathbf{f}_{k,j}^{(1)}\big)\Big(\sum_{i=1}^d \big\|\pi^{(\mathrm{t},i)}\big(\mathbf{f}_{k,j}^{(1)}\big)\big\|_{\mathcal{A}^s}\Big)^{\frac{1}{s}} \\ &\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{1+\frac{1}{s}} \eta_{k,j}^{-\frac{1}{s}} (\ln(d) + |\ln \eta_{k,j}|)^{13} (1 + |\ln \eta_{k,j}|)^{3b_{\mathbf{f}}},\end{aligned}$$

where we used Assumptions 2 and 4 in the last step. Moreover, using Theorem 4.62 as well as (4.181) and (4.202) we have

$$\begin{aligned}&\operatorname{flops}\Big(\operatorname{APPLY}_{\mathbf{B}_2^\top}\Big(\mathbf{f}_{k,j}^{(2)}, \tfrac{\eta_{k,j}^{(2)}}{4}\Big)\Big) \\ &\quad\lesssim \hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}} (\ln(d) + |\ln(\eta_{k,j})|)^3 \big|\operatorname{rank}\big(\mathbf{f}_{k,j}^{(2)}\big)\big|_\infty \sum_{i=1}^d \#\operatorname{supp}\big(\pi^{(i)}\big(\mathbf{f}_{k,j}^{(2)}\big)\big) \\ &\qquad + d\hat{C}_{\mathbf{u},\mathbf{f}}^{\frac{s}{s+1}} (\ln(d) + |\ln(\eta_{k,j})|)^7 \big|\operatorname{rank}\big(\mathbf{f}_{k,j}^{(2)}\big)\big|_\infty^3 \\ &\quad\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{1+\frac{1}{s}} \eta_{k,j}^{-\frac{1}{s}} (\ln(d) + |\ln(\eta_{k,j})|)^7 (1 + |\ln \eta_{k,j}|)^{3b_{\mathbf{f}}}.\end{aligned}$$

Combining the three estimates yields

$$\operatorname{flops}\big(\operatorname{RHS}\big(\tfrac{\eta_{k,j}}{2}\big)\big) \le C_1 \hat{C}_{\mathbf{u},\mathbf{f}} d^q d^{1+\frac{1}{s}} \eta_{k,j}^{-\frac{1}{s}} (\ln(d) + |\ln(\eta_{k,j})|)^7 (1 + |\ln \eta_{k,j}|)^{3b_{\mathbf{f}}}.$$

In the next step we estimate the costs for $\operatorname{APPLY}\big(\mathbf{w}_{k,j}, \frac{\eta_{k,j}}{2}\big)$. Again, these costs are split into three parts: the adaptive application of $\mathbf{B}_1$, $\mathbf{B}_1^\top$, and $\mathbf{A}_2$. Using Theorems 4.46 and 4.76 as well as (4.191) and (4.194), we obtain

$$\begin{aligned}&\operatorname{flops}\Big(\operatorname{APPLY}_{\mathbf{B}_1}\Big(\mathbf{w}_{k,j}, \frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}\Big)\Big) \\ &\qquad\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j}^{-\frac{1}{s}} (\tilde{C}d^2)^{\frac{j}{s}} \ln(d)^{22} ((\ln d) + |\ln \eta_{k,j}|)^{14} \operatorname{rank}_\infty^3(\mathbf{w}_{k,j})\end{aligned}$$

with $p_1 \ge 1 + p_0 s^{-1}$ from Theorem 4.76. Similarly, by additional using Theorem 4.43 as well as (4.193), (4.196) and (4.197), we have

$$\begin{aligned}\operatorname{flops}\Big(\operatorname{APPLY}_{\mathbf{B}_1^\top}\Big(\mathbf{z}_{k,j}, \tfrac{\eta_{k,j}^{(1)}}{4}\Big)\Big) &\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j}^{-\frac{1}{s}} (\tilde{C}d^2)^{\frac{j}{s}} \ln(d)^{22} ((\ln d) + |\ln \eta_{k,j}|)^{14} \operatorname{rank}_\infty^3(\mathbf{z}_{k,j}) \\ &\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j}^{-\frac{1}{s}} (\tilde{C}d^2)^{\frac{j}{s}} \ln(d)^{43} ((\ln d) + |\ln \eta_{k,j}|)^{26} \operatorname{rank}_\infty^3(\mathbf{w}_{k,j}),\end{aligned}$$

where we used (4.109) from Theorem 4.43 combined with (4.181) and (4.202) in the second step. Moreover, using Theorem 4.55 combined with Theorem 4.76 as well as (4.192) and (4.195), we obtain

$$\begin{aligned}&\operatorname{flops}\Big(\operatorname{APPLY}_{\mathbf{A}_2}\Big(\mathbf{w}_{k,j}, \frac{\eta_{k,j}^{(2)}}{2}\Big)\Big) \\ &\qquad\lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j}^{-\frac{1}{s}} (\tilde{C}d^2)^{\frac{j}{s}} \ln(d)^{28} ((\ln d) + |\ln \eta_{k,j}|)^{17} \operatorname{rank}_\infty^3(\mathbf{w}_{k,j}).\end{aligned}$$

Combining these three estimates with the definition of APPLY in (4.152) yields

$$\operatorname{flops}\left(\operatorname{APPLY}\left(\mathbf{w}_{k,j}, \tfrac{\eta_{k,j}}{2}\right)\right) \lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j}^{-\frac{1}{s}} (\tilde{C} d^2)^{\frac{j}{s}} \ln(d)^{43} ((\ln d) + |\ln \eta_{k,j}|)^{26} \operatorname{rank}_\infty^3(\mathbf{w}_{k,j}).$$

As already mentioned in Theorem 4.13, the computational complexity of COARSEN and RECOMPRESS are dominated by the computational complexity of the required calculation of the HSVD. As one can see in the proof of Theorem 4.76, the estimates (4.176) and (4.177) are valid for $\bar{\mathbf{w}}_{k,j}$ as well. Hence, the computational costs for COARSEN and RECOMPRESS can be estimated by

$$\operatorname{flops}(\mathbf{w}_{k,j+1}) \lesssim \hat{C}_{\mathbf{u},\mathbf{f}} d^{p_1} \eta_{k,j+1}^{-\frac{1}{s}} (\tilde{C} d^2)^{\frac{j+1}{s}} ((\ln d) + |\ln \eta_{k,j+1}|) \operatorname{rank}_\infty^4(\bar{\mathbf{w}}_{k,j+1}).$$

In the next step, we combine these estimates with the rank bound of the iterates from Theorem 4.79. Here, we mention that by the proof of Theorem 4.79, the statement (4.188) is also true for $\bar{\mathbf{w}}_{k,j}$. Hence, inserting (4.188) into the estimates above yields that the computational costs for each inner iteration step can be bounded by a fixed multiple independent of d and $\eta_{k,j}$ by

$$\hat{C}_{\mathbf{u},\mathbf{f}} d^{p_3} \eta_{k,j+1}^{-\frac{1}{s}} (\tilde{C} d^2)^{\frac{j+1}{s}} (\ln d)^{56(j+1)} ((\ln d) + |\ln \eta_{k,j+1}|)^{32(j+1)+4b}$$

with $p_3 = p_2 + \max\{p_1, 1 + s^{-1} + q\}$ and p_2 from Theorem 4.79, where we used that without loss of generality $\eta_{k,j} \leq 1$. Using this property again, as well as

$$\sum_{j=0}^{I-1} \eta_{k,j+1}^{-\frac{1}{s}} \lesssim \eta_{k,I}^{-\frac{1}{s}}$$

and $j + 1 \leq I$, we obtain that the computational costs for the l-th outer iteration step can be bounded by a fixed multiple independent of d and $\eta_{k,I}$ by

$$\hat{C}_{\mathbf{u},\mathbf{f}} d^{p_3} \eta_{l,I}^{-\frac{1}{s}} (\tilde{C} d^2)^{\frac{I}{s}} (\ln d)^{56I} ((\ln d) + |\ln \eta_{l,I}|)^{32I+4b}.$$

Using that

$$\sum_{l=0}^{k} \eta_{l,I}^{-\frac{1}{s}} \lesssim \eta_{k+1,I}^{-\frac{1}{s}},$$

the total computational costs to calculate $\mathbf{u}_\varepsilon = \mathbf{u}_{k+1}$ is bounded by a fixed multiple independent of d and $\eta_{k,I}$ of

$$\hat{C}_{\mathbf{u},\mathbf{f}} d^{p_3} \eta_{k+1,I}^{-\frac{1}{s}} (\tilde{C} d^2)^{\frac{I}{s}} (\ln d)^{56I} ((\ln d) + |\ln \eta_{k,I}|)^{32I+4b}.$$

In the final step, we express the estimate in terms of ε. For this purpose, we mention that by the definition of $\eta_{k+1,I}$ and by Theorem 4.75, we have

$$\eta_{k+1,I}^{-\frac{1}{s}} = 2^{\frac{1}{s}} \eta_{k,I}^{-\frac{1}{s}} = 2^{\frac{1}{s}} \rho^{-\frac{1}{s}} \rho^{-\frac{1}{s}I} (2^{-k} \varepsilon_0)^{-\frac{1}{s}} \leq 2^{\frac{1}{s}} \rho^{-\frac{1}{s}} \varepsilon^{-\frac{1}{s}} \rho^{-\frac{1}{s} c(\ln d)} = 2^{\frac{1}{s}} \rho^{-\frac{1}{s}} \varepsilon^{-\frac{1}{s}} d^{cs^{-1}|\ln \rho|}.$$

Hence, there exists a constant C_1 independent of d such that

$$(\ln d) + |\ln \eta_{k,I}| \leq C_1 ((\ln d) + |\ln \varepsilon|).$$

Again using Theorem 4.75, we arrive at

$$\operatorname{flops}(\mathbf{u}_\varepsilon) \lesssim d^a\, d^{2cs^{-1}\ln d}(\ln d)^{88c\ln d}(\ln d)^{4b}(1+|\ln\varepsilon|)^{32c(\ln d)+4b}\,\varepsilon^{-\frac{1}{s}}$$

with

$$a = p_3 + cs^{-1}(|\ln\rho| + \ln(\tilde{C} + 32\ln(C_1))).$$

Combining this with Assumption 4 yields (4.163). □

Remark 4.80. At this point, we mention that the estimates are not optimized and, in particular, not sharp. Hence, the exponents in the computational complexity estimates could probably be improved by a more detailed analysis. However, since the primary purpose of this work is to show that we can avoid the curse of dimensionality, the observed estimates are sufficient for our purpose. As we will see in Section 4.5, the numerical results indicate that the practical performance of the proposed method is better than the theoretical estimates suggest. In particular, we observe polynomial scaling of complexity with respect to the number of spatial dimensions d.

4.5 Numerical Experiments

In this section, we present numerical results to illustrate the performance of the proposed method for the heat equation with two different test cases and in different spatial dimensions d.

In our implementation, we use hierarchical tensor representations with a linear dimension tree as in (2.19) and Figure 2.6. As a one-dimensional wavelet basis, we use L_2-orthonormal, continuously differentiable, piecewise polynomial Donovan-Geronimo-Hardin multiwavelets [DGH99] of polynomial degree six. These wavelets satisfy our necessary assumptions. In particular, the L_2-orthonormality is crucial, as stated in Section 3.1.2, to ensure that the condition number of the bi-infinite matrix is independent of the number of spatial dimensions. Moreover, as seen in Section 2.5, with appropriate rescaling, we obtain s^*-compressibility of the one-dimensional operator $\mathbf{T}_2$ defined in (2.25) and super-compressibility of $\mathbf{T}_t$.

In particular, the goal of this section is to compare the results of Theorem 4.69 with numerical simulations. To do this, we consider the high-dimensional heat equation

$$\partial_t u - \Delta u = g \quad \text{in } \Omega = (0,1)^d, \qquad u|_{t=0} = h, \tag{4.204}$$

with homogeneous Dirichlet boundary conditions on $\partial\Omega$ for different spatial dimensions d. In particular, we consider two different versions: one with $\|g\|_{L_2} = 1$ and a vanishing initial condition $h = 0$, and one with a vanishing source term $g = 0$ and $\|h\|_{L_2} = 1$. Due to this renormalization of the right-hand side, it is sufficient to consider absolute residual norms.

We mention here that we use a slightly modified version of the routine APPLY as described in [BD16, Section 7.2]. Since, in general, the rank-truncated scaling matrix can have many terms, the intermediate output of APPLY may have a high rank, but with many small singular values. Hence, we apply the RECOMPRESS routine systematically to the intermediate results within APPLY to reduce the ranks, but with relatively small tolerances. The tolerances are chosen in such a way that the error bound

$$\|\mathbf{A}\mathbf{v} - \text{APPLY}(\mathbf{v};\eta)\| \le \eta$$

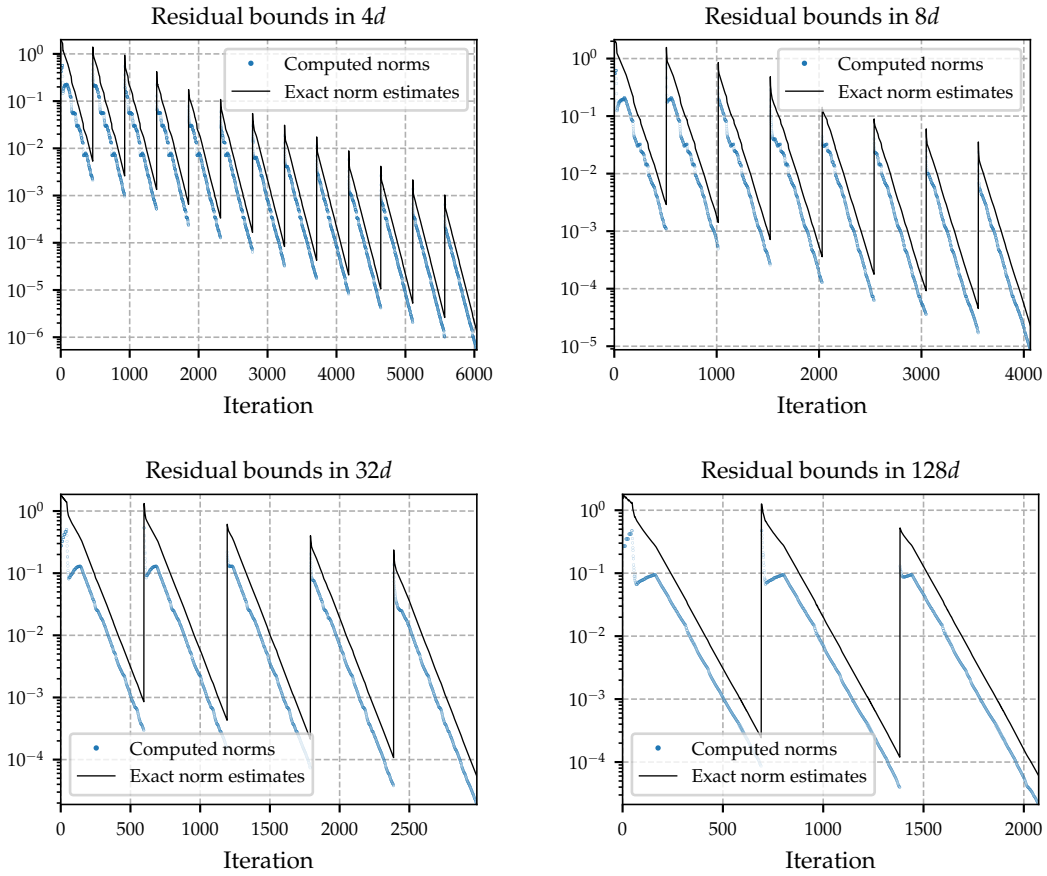


FIGURE 4.5: Residual norms of computed error estimates and error bounds in dependence on the iteration number for the heat equation with source term, for $d = 4, 8, 32, 128$.

is still valid. In particular, we proceed in this way for the subroutines of APPLY in (4.152).

Furthermore, for simplicity, we choose g and h in the two test cases to be finite sums of basis functions. By proceeding in this way, the Assumption 2 is obviously satisfied. For test cases with infinitely supported functions g or h, one can adapt the concepts of the elliptic case in [BD16, Appendix B].

4.5.1 Vanishing Initial Condition

As the first test case, we consider the one with vanishing initial condition and non-vanishing source term. For simplicity, we choose the source term

$$g(t,x) = \theta_{\nu_t}(t) \prod_{i=1}^{d} \psi_{\nu_1}(x_i),$$

for one chosen $\nu_t \in \vee_t$ and $\nu_1 \in \vee_1$. The corresponding coefficient vector $\mathbf{f}_1$ is one at index $(\nu_t, \nu_1, \dots, \nu_1)$ and zero elsewhere, and $\mathbf{f}_2 = 0$. Hence, $\operatorname{rank}_\infty(\mathbf{f}_1) = 1$ and $\#\operatorname{supp}(\pi^{(t,i)}(\mathbf{f}_1)) = 1$ for each $i = 1, \dots, d$. The parameters in Algorithm 4.1 are chosen according to the convergence analysis.

Figure 4.5 illustrates the computed residual norms $\|\mathbf{r}_{k,j}\|$ for $\mathbf{r}_{k,j}$ defined in Algorithm 4.1 as well as the upper bounds $\|\mathbf{r}_{k,j}\| + \eta_{k,j}$ for the exact residual norms

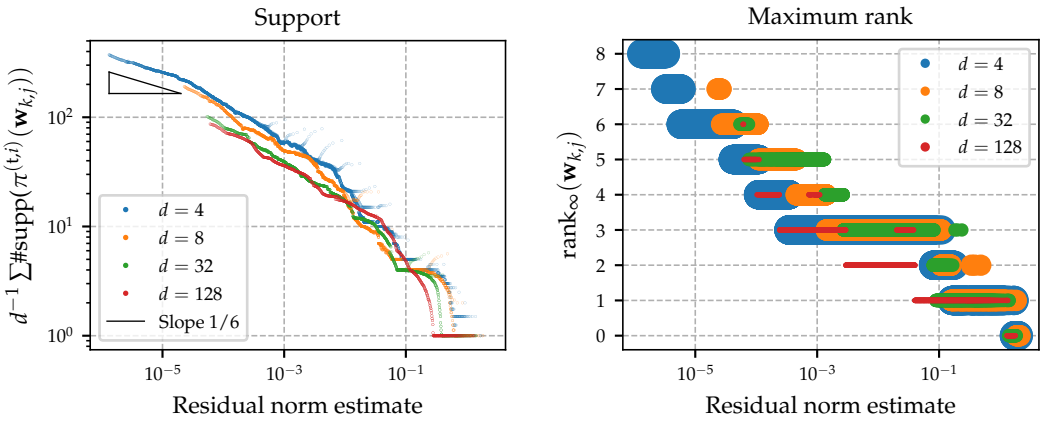


FIGURE 4.6: Averaged sum of the one-dimensional support (left) and maximum rank per iteration versus the current residual norm estimate (right) for the heat equation with source term, for $d = 4, 8, 32, 128$.

$\|\mathbf{A}\mathbf{w}_{k,j} - \mathbf{B}^\top \mathbf{f}\|$ for different spatial dimensions $d \in \{4, 8, 32, 128\}$. Since the operator $\mathbf{A}$ is bounded and elliptic, we know that

$$\|\mathbf{A}\mathbf{w}_{k,j} - \mathbf{B}^\top \mathbf{f}\| \eqsim \|\mathbf{u} - \mathbf{w}_{k,j}\|,$$

where by the choice of our basis, the constants in the equivalence are by Section 3.1 independent of the spatial dimension d. Furthermore, by the Riesz basis property, we have

$$\|u - w_{k,j}\|_{\mathcal{X}} \eqsim \|\mathbf{u} - \mathbf{w}_{k,j}\|,$$

again with constants independent of d, where $w_{k,j}$ is the function corresponding to the coefficient tensor $\mathbf{w}_{k,j}$. Hence, this method allows for control of the error in the space-time norm. As one can see in Figure 4.5, the method behaves as expected with an increase of the residuals after each outer iteration step based on the application of the COARSEN and RECOMPRESS routines in line 14 of Algorithm 4.1. As shown in Theorem 4.75, the number of inner iteration steps increases for an increasing number of spatial dimensions d. Even for high spatial dimensions up to $d = 128$, the method produces approximations with residual norms of order 10^{-4}. Due to the longer runtime for higher spatial dimensions, we can perform fewer outer iteration steps for comparable time: 13 for $d = 4$ and only 3 for $d = 128$.

Figure 4.6 illustrates the averaged sum of the one-dimensional support sizes as well as the maximum ranks in dependence on the current residual norm estimate for each iterate $\mathbf{w}_{k,j}$. There is a large pre-asymptotic range, but for small residual norm estimates in the case $d = 4$, the expected rate of $s = \frac{1}{6}$ according to Theorem 4.69 and the choice of the wavelet basis with six vanishing moments is observed. Moreover, we observe that the sum of the one-dimensional support sizes grows linearly in d for fixed residual norm estimates, which is better than the theoretical estimate in Theorem 4.69. Furthermore, Figure 4.6 shows polynomial growth of the maximum ranks in a semi-logarithmic plot, which indicates poly-logarithmic growth with respect to the residual norm estimates.

Next, we investigate the individual spatial components $(\mathbf{w}_{k,j})_{\nu_t}$ for each time index $\nu_t \in \vee_t$ of the iterates $\mathbf{w}_{k,j}$ in Algorithm 4.1. One of the main features of our approach is that it allows for different ranks and spatial supports at each time index.

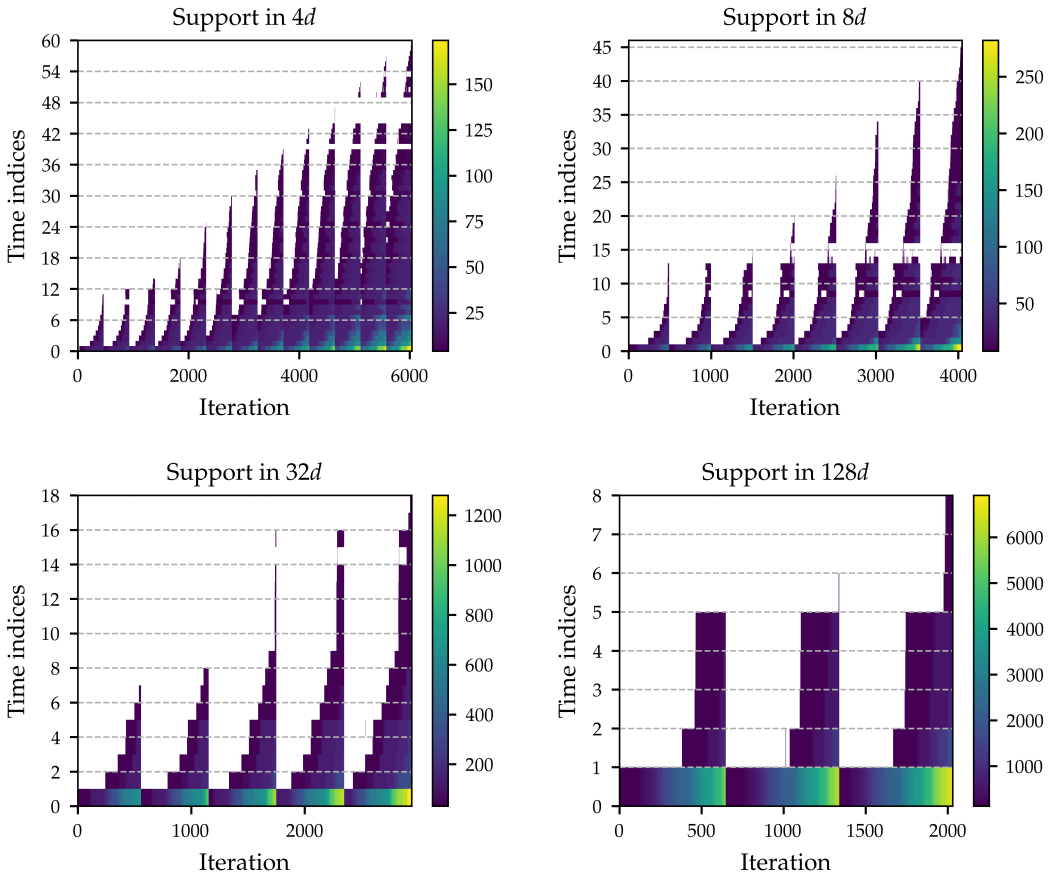


FIGURE 4.7: Sum of one-dimensional support for each time index (4.205) in dependence on the iteration number for the heat equation with source term, for $d = 4, 8, 32, 128$.

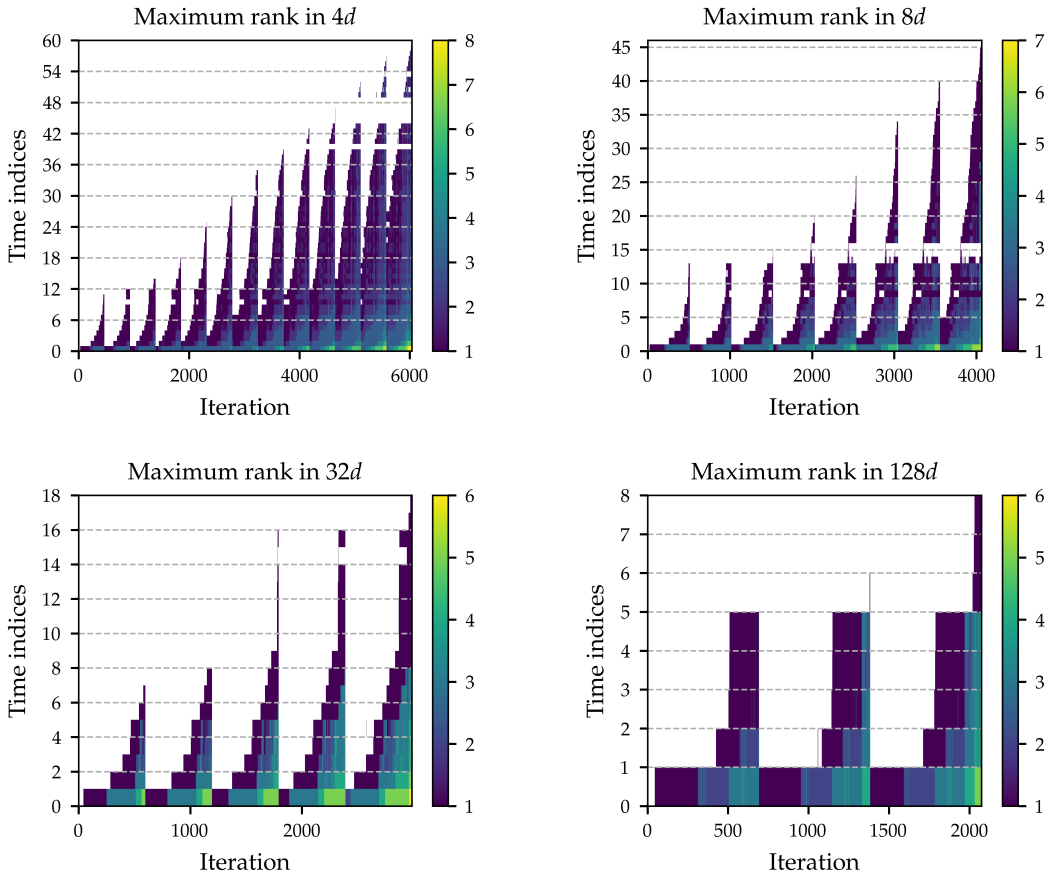


FIGURE 4.8: Residual norms of computed error estimates and error bounds in dependence on the iteration number for the heat equation with source term, for $d = 4, 8, 32, 128$.

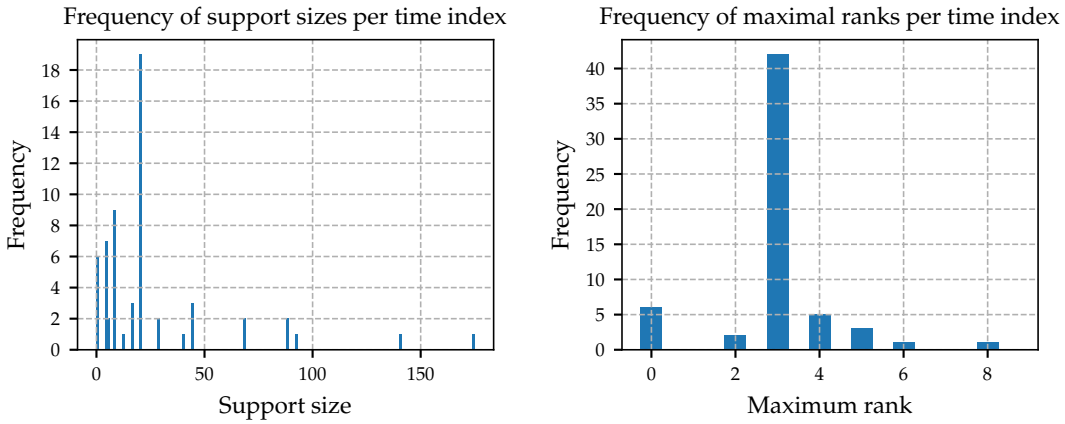


FIGURE 4.9: Frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for the heat equation with source term for $d = 4$.

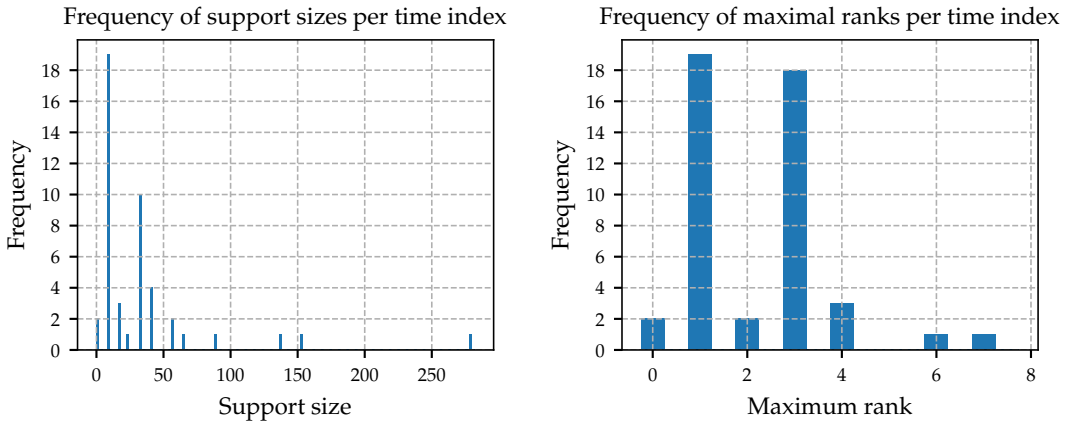


FIGURE 4.10: Frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for the heat equation with source term for $d = 8$.

In Figure 4.7, the sum of the one-dimensional support sizes for each time index, i.e.,

$$\sum_{i=1}^{d} \#\mathrm{supp}(\pi^{(i)}((\mathbf{w}_{k,j})_{\nu_t})), \quad \nu_t \in \vee_t, \tag{4.205}$$

is illustrated over the course of the iteration for different spatial dimensions d. Figure 4.8 shows the corresponding maximal rank for each spatial component. Analogously to Figure 4.5, we observe a substantial decrease in the support sizes after each outer iteration step. The support sizes remain relatively small for most time indices, and only a few require larger support sizes. This behaviour demonstrates one of the significant advantages of our method over using time as a tensor component. In this case, one could not exploit this sparsity pattern.

This observation is further supported by Figure 4.9, where the frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for $d = 4$ are illustrated. Here, only 13 out of 60 time indices have a support size larger than 20 in comparison to the maximal support size of 174. Analogously, only 10 out of 60 time indices have a maximal rank larger than

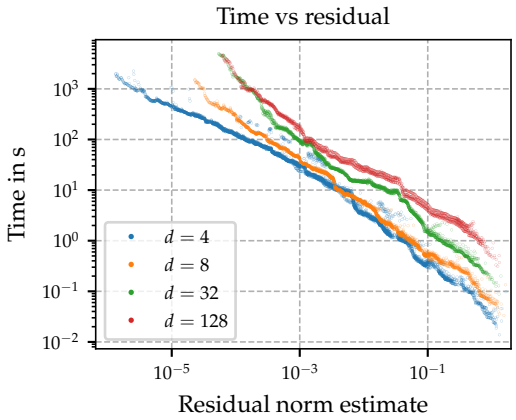


FIGURE 4.11: Time in seconds required for each inner iteration step for the heat equation with source term for $d = 4, 8, 32, 128$ in dependence on the achieved residual estimate.

three in comparison to the overall maximal rank of eight. However, the reduction of the ranks is not as pronounced as for the support size. The same observations hold for the test cases with higher spatial dimensions. Figure 4.10 illustrates the corresponding frequencies for $d = 8$. For dimensions $d = 32$ and $d = 128$ the number of active time indices is significantly smaller. Hence, we omit the corresponding figures for the frequencies.

Figure 4.11 shows the time required for each inner iteration step in dependence on the achieved residual estimate for different spatial dimensions d. Here, we use time in seconds as a measure of computational complexity. We observe algebraic behaviour with respect to the residual norm estimates, but, as in Figure 4.6, we are still in the pre-asymptotic regime. Based on the theoretical results from Theorem 4.69, we would expect a rate of $s = \frac{1}{6}$ besides additional poly-logarithmic terms. The dependence of the computational costs with respect to the spatial dimension d in Figure 4.11 is polynomial, which is in accordance with Theorem 4.69. However, since we are still in the pre-asymptotic behaviour, we cannot predict the exact exponent of this polynomial growth.

4.5.2 Vanishing Source Term

Next, we consider a test case with vanishing source term and nonvanishing initial condition, meaning $g = 0$ and $h \neq 0$ in (4.204). Again, for simplicity, we choose an initial condition that is a product of one basis function per dimension, i.e.,

$$h(x) = \prod_{i=1}^{d} \psi_{\nu_1}(x_i), \quad \nu_1 \in \vee_1.$$

In this test case, we observe that the temporal support of the iterates grows faster than in the test case with vanishing initial condition from Section 4.5.1, leading to higher computational costs. Hence, we restrict ourselves to spatial dimensions $d \in \{4, 8, 16\}$ in this case. Moreover, we choose larger tolerances for the routines APPLY, COARSEN, and RECOMPRESS in Algorithm 4.1 than suggested by the convergence analysis, to reduce computational costs. In particular in line 10 of Algorithm 4.1 we replace $\eta_{k,j}$ by $10\eta_{k,j}$ and in line 11 we replace $\eta_{k,j}$ by $2\eta_{k,j}$. Since the error estimates

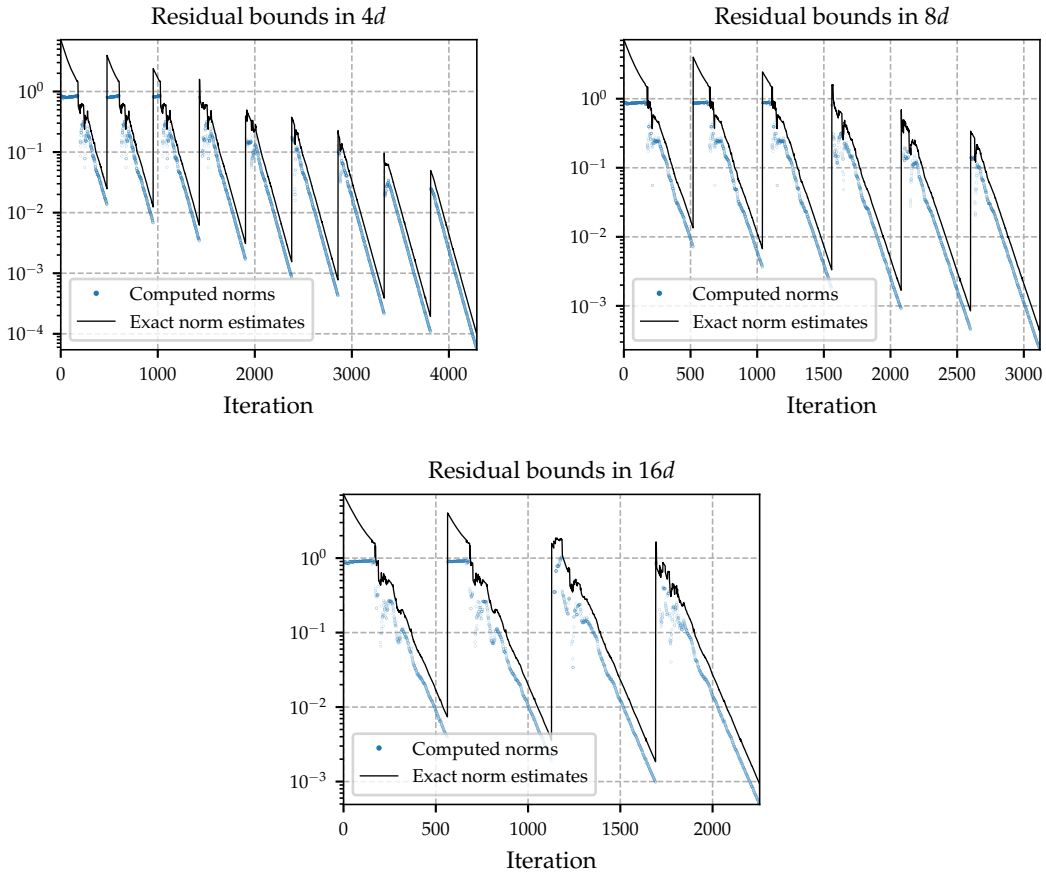


FIGURE 4.12: Residual norms of computed error estimates and error bounds in dependence on the iteration number for the heat equation with initial condition, for $d = 4, 8, 16$.

are not sharp in most cases, we do not observe any impact on the convergence of the method.

For this test case, we illustrate the same quantities as for the experiment with vanishing initial condition in Section 4.5.1. We start with the computed approximate residual norms $\|\mathbf{r}_{k,j}\|$ and the upper bounds $\|\mathbf{r}_{k,j}\| + \eta_{k,j}$ for the exact residual norms $\|\mathbf{A}\mathbf{w}_{k,j} - \mathbf{B}^\top \mathbf{f}\|$ for different spatial dimensions $d \in \{4, 8, 16\}$ in Figure 4.12. The principal behaviour is similar to the previous test case, where we observe an increase in residuals after each outer loop step and a decrease within each inner loop. At the beginning of the inner iterations, the residual estimates are slightly unstable based on the higher tolerances for the subroutines. However, for smaller residual norms, which is the relevant regime, the residual estimates behave as expected.

Figure 4.13 illustrates the averaged sum of the one-dimensional support sizes as well as the maximum ranks in dependence on the current residual norm estimate for each iterate $\mathbf{w}_{k,j}$. In contrast to the other test case, we do not leave the pre-asymptotic regime for the sum of the one-dimensional supports, not even for spatial dimension $d = 4$ as in Figure 4.6. However, we still observe polynomial growth of the support sizes with respect to the residual norm estimates and the spatial dimension d. Moreover, the maximal rank is bounded by four for all iterates, which is moderately small. In particular, it is much smaller than the maximal sum of the one-dimensional support sizes over all spatial components, which is approximately 200, as one can see in Figure 4.14. This behaviour indicates that we benefit substantially from the

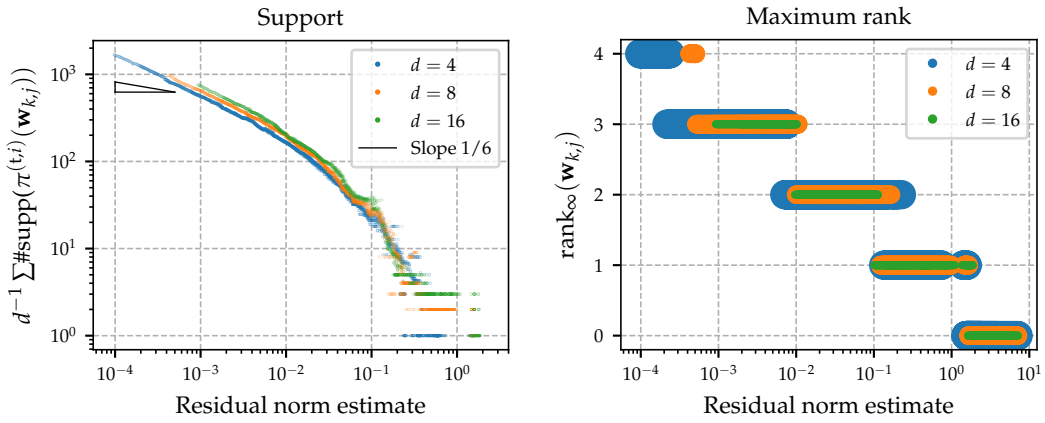


FIGURE 4.13: Averaged sum of the one-dimensional support and maximum rank per iteration versus the current residual norm estimate for the heat equation with initial condition, for $d = 4, 8, 16$.

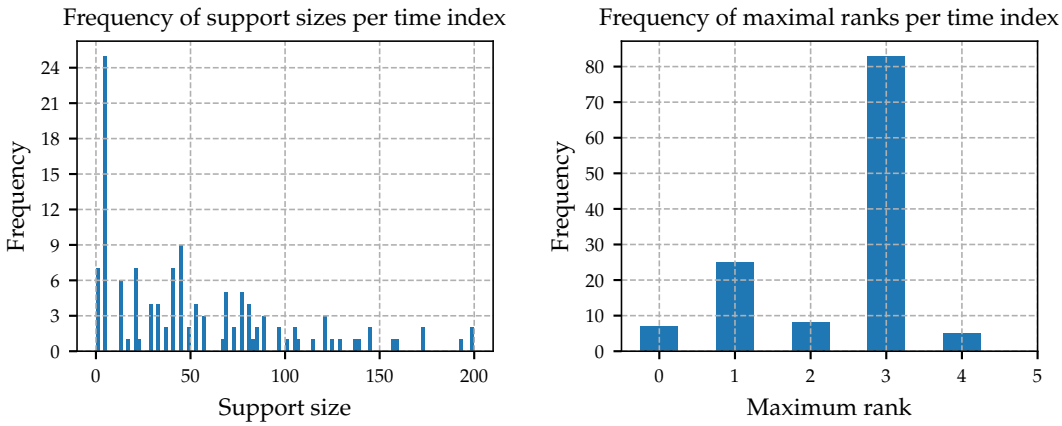


FIGURE 4.14: Frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for the heat equation with initial condition for $d = 4$.

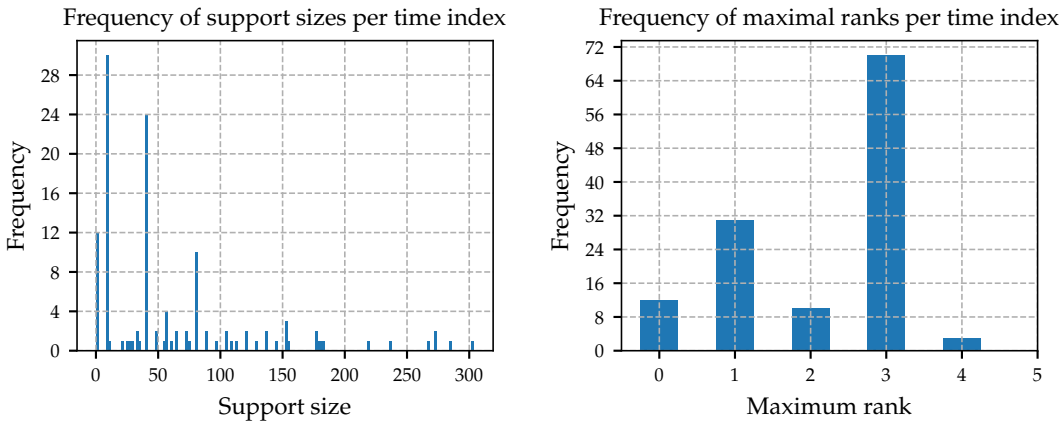


FIGURE 4.15: Frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for the heat equation with initial condition for $d = 8$.

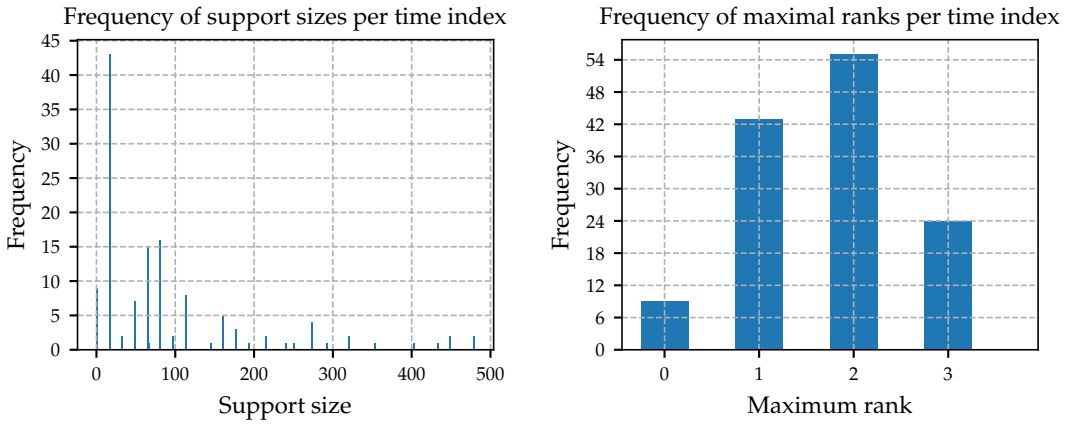


FIGURE 4.16: Frequencies of the sum of one-dimensional support sizes and maximum ranks over all time indices at the final iteration step for the heat equation with initial condition for $d = 16$.

low-rank structure of the solution. Furthermore, as for the other test case, the results indicate poly-logarithmic growth of the maximal ranks with respect to the residual norm estimates.

In Figures 4.17 and 4.18, we investigate the support sizes and the maximum ranks of the respective spatial components $(\mathbf{w}_{k,j})_{\nu_t}$ for each time index $\nu_t \in \vee_t$ of the iterates $\mathbf{w}_{k,j}$ in Algorithm 4.1. In this test case, we observe a significantly faster increase of the temporal support, caused by the adaptive application of the initial-value operator $\mathbf{B}_2^\top$ onto $\mathbf{f}_2$, which is zero in the other test case due to the vanishing initial condition. However, we observe in Figure 4.14 that most time indices still require only small support sizes. For $d = 4$ at the final iteration step, we even have $\#\mathrm{supp}((\mathbf{w}_{k,j})_{\nu_t}) = 1$ for 25 of the spatial components. The same observation holds for the final iteration steps for $d = 8$ and $d = 16$, which are illustrated in Figures 4.15 and 4.16.

Figure 4.19 shows the time required for each inner iteration step in dependence on the achieved residual estimate for different spatial dimensions d. In this test case, we have higher computational costs per iteration than in the previous test case in Figure 4.11. The higher runtime is reasonable due to the higher support sizes as seen in Figure 4.17 in comparison to Figure 4.7. For comparable residual norm estimates, the sum of the one-dimensional supports is approximately ten times greater, which explains the higher costs due to the dependence of costs on support. Due to the higher computational costs, we could perform fewer iterations in this test case. Although we are still in the pre-asymptotic regime, we observe a polynomial dependence on the residual norm estimate and the spatial dimension d.

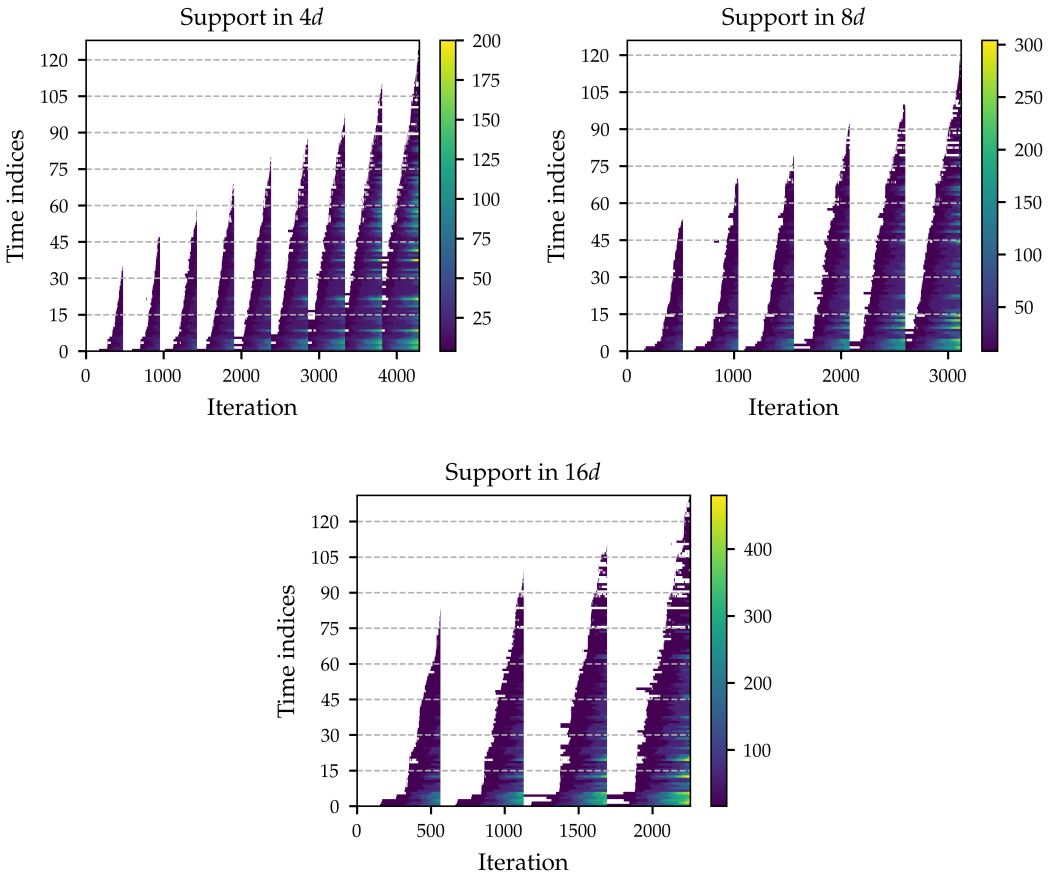


FIGURE 4.17: Sum of one-dimensional support for each time index in dependence on the iteration number for the heat equation with initial condition, for $d = 4, 8, 16$.

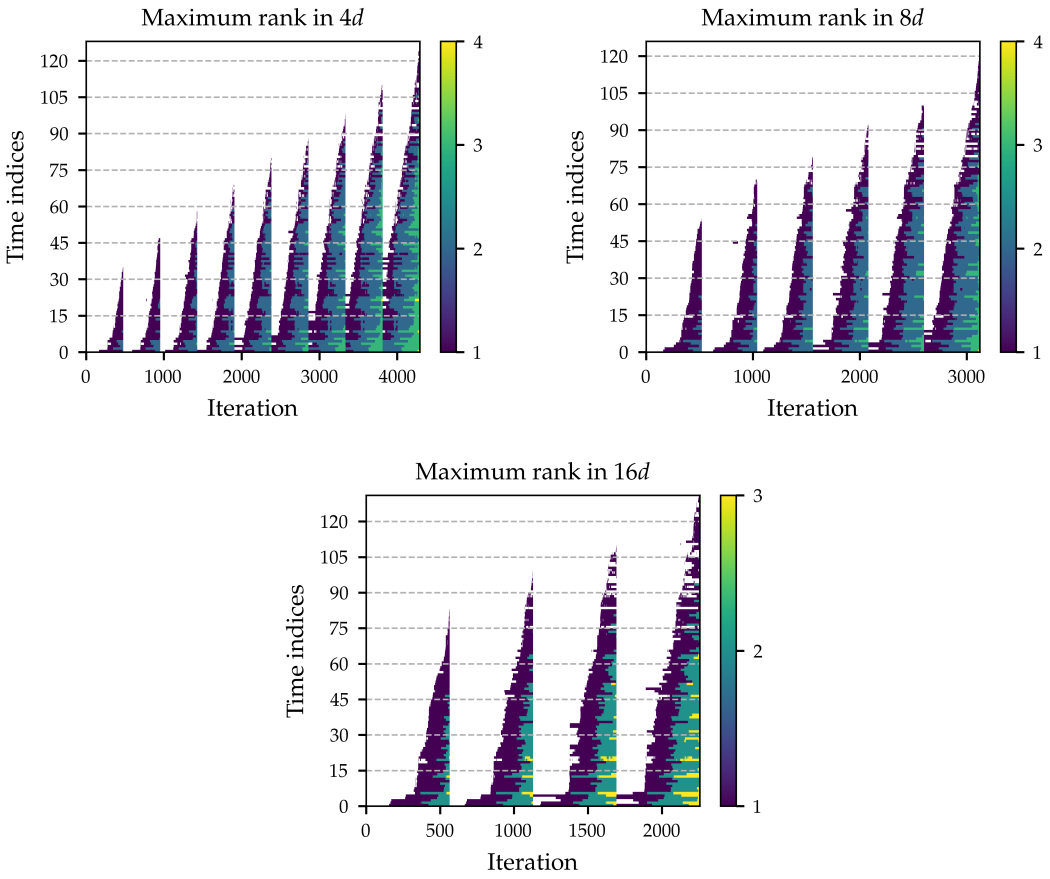


FIGURE 4.18: Residual norms of computed error estimates and error bounds in dependence on the iteration number for the heat equation with initial condition, for $d = 4, 8, 16$.

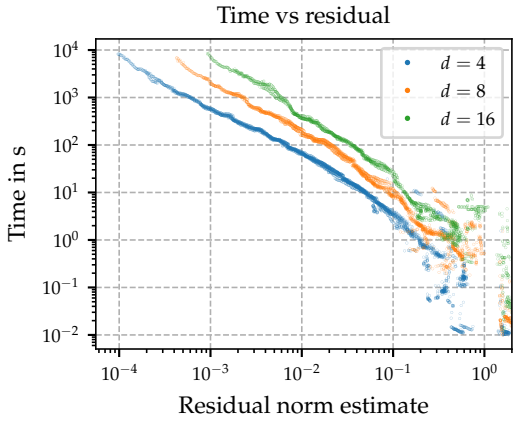


FIGURE 4.19: Time in seconds required for each inner iteration step for the heat equation with initial condition for $d = 4, 8, 16$ in dependence on the achieved residual estimate.

Chapter 5

Magnetic Particle Imaging

Magnetic Particle Imaging (MPI) is a tracer-based tomographic imaging method introduced in [GW05]. The idea behind MPI is to image the spatially dependent concentration of superparamagnetic iron oxide nanoparticles. Unlike other imaging techniques with high temporal resolution, these particles have the advantage of not emitting harmful radiation, which is beneficial for in vivo applications. Furthermore, we can achieve high spatial resolution below 1 mm in the order of milliseconds using this imaging method.

There are several possibilities for clinical applications of MPI. Examples are vascular imaging, cancer imaging, stem cell tracking, and hyperthermia treatment monitoring. For an overview and comparisons to other imaging modalities, we refer to [Yan+22].

In the following, we give a brief introduction to the physical principles of MPI and the challenges of mathematical models for the processes underlying MPI. For a full introduction to MPI and the explanation of the physical details, we refer to [KB12].

Superparamagnetic particles have no remanence magnetization, which means that they lose their magnetization when the external magnetic field is removed. Moreover, these particles have a nonlinear magnetization response and a maximal magnetization, called saturation magnetization. The fundamental mechanism of MPI relies on this nonlinear magnetization response of the nanoparticles to an external magnetic field. For the physical modeling of this response, different models have been proposed in the literature, ranging from simple models assuming thermal equilibrium of the particles to more complex models based on stochastic differential equations (SDEs). For an overview of different mathematical models for this response, we refer to [Klu18].

The external time-dependent magnetic field $\vec{H}\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ is applied to the spatial region of interest $\Omega \subset \mathbb{R}^3$. The magnetic field is of the form

$$\vec{H}(x,t) = \vec{H}_S(x) + \vec{H}_D(t) \tag{5.1}$$

with a static selection field $\vec{H}_S\colon \mathbb{R}^3 \to \mathbb{R}^3$ and a dynamic drive field $\vec{H}_D\colon \mathbb{R} \to \mathbb{R}^3$. The selection field creates a field-free point, i.e., a point $x \in \mathbb{R}^3$, where $\vec{H}_S(x) = 0$. The dynamic drive field moves this field-free point through the imaging volume along a predefined trajectory. The superparamagnetic particles respond to the applied magnetic field with a magnetization $M\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$. Figure 5.1 illustrates a cross-section of the static selection field, the dynamic drive field, and the resulting magnetic field for a fixed point of time. The drive field is of the form

$$\vec{H}_D(t) = \begin{pmatrix} A_1 \sin(2\pi\nu_x t) \\ A_2 \sin(2\pi\nu_y t) \\ A_3 \sin(2\pi\nu_z t) \end{pmatrix} \tag{5.2}$$

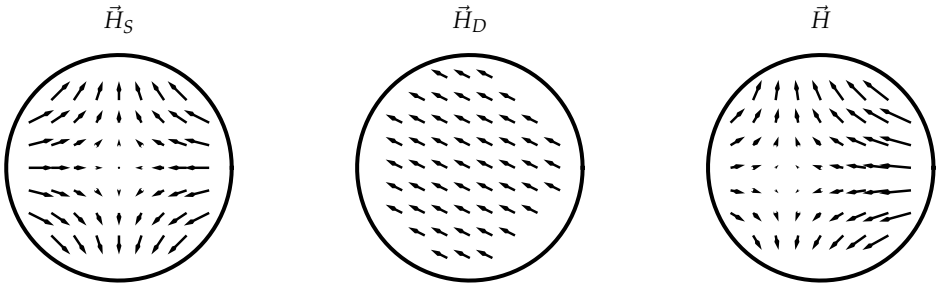


FIGURE 5.1: Visualization of a cross section of the magnetic fields in MPI for a fixed point of time: (left) selection field $\vec{H}_S$ creating a field-free point at the center, (middle) dynamic drive field $\vec{H}_D$ moving the field-free point in a circular trajectory, (right) resulting magnetic field $\vec{H} = \vec{H}_S + \vec{H}_D$.

with amplitudes $A_1, A_2, A_3 \geq 0$ and excitation frequencies $\nu_x, \nu_y, \nu_z > 0$. Following [Wei+09], the excitation frequencies are chosen to be similar but slightly different, i.e., $\nu_x \approx \nu_y \approx \nu_z$, which results in a Lissajous trajectory. For more details on the drive field trajectory design, we refer to [KB12, Section 2.5.6]. The static field is of the form

$$\vec{H}_S(x) = \begin{pmatrix} G_x & 0 & 0 \\ 0 & G_y & 0 \\ 0 & 0 & G_z \end{pmatrix} x$$

with gradient strengths $G_x, G_y, G_z \in \mathbb{R}$. Based on Gauss's law for magnetism, the divergence of the magnetic field has to be zero, which implies $G_x + G_y + G_z = 0$. A standard setup used in MPI scanners is $G_y = -2G_x = -2G_z$, which results in

$$\vec{H}_S(x) = G \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} x \tag{5.3}$$

with gradient strength $G \in \mathbb{R}$. For more details on the design of the selection field, we refer to [KB12, Section 2.4].

As the field-free point moves over the object, the magnetization of the particles near the field-free point changes rapidly. According to the induction principle, this change of magnetization induces a voltage u, which is measured by multiple receive coils. This induced voltage is given by

$$u(t) = -\mu_0 \int_\Omega \mathbf{p}^{\mathrm{R}}(x) \cdot \partial_t M(x,t) \, \mathrm{d}x, \tag{5.4}$$

where Ω is the imaging domain, μ_0 is the magnetic permeability, and $\mathbf{p}^{\mathrm{R}} \colon \mathbb{R}^3 \to \mathbb{R}^3$ is the sensitivity profile of the receive coil. The goal of MPI is to reconstruct the spatially dependent particle concentration from the measured induced voltage.

We define the mean magnetic moment $\bar{m} \colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ as

$$\bar{m}(x,t) = \frac{1}{N} \sum_{i=1}^{N} m_i(x,t),$$

where $m_i(x,t) \in \mathbb{R}^3$ is the magnetic moment of the i-th particle at position x and

time t and N is the number of particles. Based on the nonmagnetic coating of the particles, the particle-particle interactions are suppressed and negligible. Hence, the magnetization $M(x,t) = c(x)\bar{m}(x,t)$ is the product of the particle concentration $c\colon \mathbb{R}^3 \to \mathbb{R}$ and the mean magnetic moment $\bar{m}$. We can thus rewrite the induced voltage defined in (5.4) as

$$u(t) = \int_\Omega s(x,t)c(x)\,\mathrm{d}x,$$

where the system function $s\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ is given by

$$s(x,t) = -\mu_0 \mathbf{p}^{\mathrm{R}}(x) \cdot \partial_t \bar{m}(x,t). \tag{5.5}$$

The inverse or imaging problem of MPI is the reconstruction of the particle concentration c from measurements of the induced voltage u. This is an ill-posed inverse problem that requires regularization techniques. For more details on the ill-posedness of the imaging problem, we refer to [KJL18] and for regularization methods to [KB12, Section 5.4].

To solve the inverse problem, we need accurate approximations of the system function s in (5.5). Currently, standard practice involves experimentally measuring the system function. This means that we identify $\bar{m}$ or s from given tuples of known concentrations c and measured voltages u by moving a probe through the field of view. However, this approach is time-consuming, has high memory requirements, and is sensitive to measurement errors, which typically appear based on background noise. Moreover, the main issue is that these measurements have to be repeated every time we vary the parameters of the scanner, such as the amplitude and frequencies, or the parameters of the particle, such as core size, viscosity, anisotropy constant, and so forth. For more details on the measurement-based approach, we refer to [KB12, Section 4.8.1].

In this work, we follow an alternative approach. The idea is to mathematically model the mean magnetic moment $\bar{m}$ and use relation (5.5) for the calculation of the system function. This approach allows us to easily vary the physical parameters of the particles and the scanner without the need for any measurements. The choice of an accurate mathematical model for the mean magnetic moment $\bar{m}$ is crucial for the quality of the resulting system function and thus for the image quality of the reconstructed particle concentration.

Remark 5.1. As mentioned above, the linear dependence of the magnetization M on the particle concentration c is based on the assumption of negligible particle-particle interactions. As investigated in [Loe+16], this assumption might not be valid for all tracers. Nevertheless, in this work, we will always assume negligible particle-particle interactions.

The outline of this chapter is as follows: We start with an overview of mathematical models for the mean magnetic moment $\bar{m}$. Here, we focus on the more accurate models using both Brown and Néel relaxation mechanisms. This results in a coupled SDE model describing the mean magnetic moment. Then based on this SDE model, we get the corresponding Fokker-Planck equation for the probability density function of the magnetization and particle direction.

Parameter	Symbol	Value
Physical constants		
Magnetic permeability	μ_0	$4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$
Boltzmann constant	k_B	$1.38064852 \times 10^{-23}\,\mathrm{J\,K^{-1}}$
Scanner (sinusoidal)		
Excitation frequency	$\nu = (\nu_x, \nu_y, \nu_z)$	$\approx 25000\,\mathrm{Hz}$
Excitation amplitudes	$A = (A_1, A_2, A_3)$	$\approx 0.02\,\mathrm{T}\mu_0^{-1}$
Gradient strength	G	$7\,\mathrm{T\,m^{-1}}\mu_0^{-1}$
Particle		
Temperature	T_{emp}	293 K
Saturation magnetization	M_S	$474000\,\mathrm{J\,m^{-3}\,T^{-1}}$
Particle core diameter	D_C	20nm
Particle core volume	V_C	$1/6\pi D_C^3$
Particle diameter	D_H	25nm
Particle volume	V_H	$1/6\pi D_H^3$
Dynamic viscosity of fluid	η	$1.0 \times 10^{-5}\,\mathrm{Pa\,s}$
Anisotropy constant	K_{anis}	$625\,\mathrm{J\,m^{-3}}$
Gyromagnetic ratio	γ	$1.75 \times 10^{11}\,\mathrm{rad\,s^{-1}\,T^{-1}}$
Damping parameter	α	0.1

TABLE 5.1: List of physical parameters in MPI with exemplary values for a typical MPI scanner and particles from [Klu18; KJL18; AKK20].

5.1 Mathematical Models

There exist various mathematical models for the description of the mean magnetic moment $\bar{m}$ that result from different physical assumptions. Stronger assumptions lead to simpler models, which are computationally cheaper to solve numerically but generally less accurate.

A strong assumption is that the particles are always in thermal equilibrium, meaning that the nanoparticles align instantaneously with the magnetic field. In this case, there is no time delay between the change of the magnetic field and the response of the particles. This leads to the so-called *equilibrium model*, i.e.,

$$\bar{m}(x,t) = m_0 \mathcal{L}(\beta \|\vec{H}(x,t)\|_2) \frac{\vec{H}(x,t)}{\|\vec{H}(x,t)\|_2},$$

where $\beta, m_0 > 0$ depend on the properties of the nanoparticles and $\mathcal{L}\colon \mathbb{R} \to \mathbb{R}$ is the Langevin function

$$\mathcal{L}(z) = \coth(z) - \frac{1}{z}.$$

For details on the equilibrium model, we refer to [KB12, Section 2.2.2].

However, in practice, the particles do not respond instantaneously to changes in the magnetic field. As seen in [KM17; KSK19], model-based reconstructions using the equilibrium model lead to worse image quality compared to measured system functions. Nevertheless, the equilibrium model is computationally cheap and is capable of describing the basic principles of MPI.

A list of relevant physical parameters with exemplary values for a typical MPI scanner and particles from [Klu18; KJL18; AKK20] is given in Table 5.1. The physical parameters of the scanner are the excitation frequencies ν_x, ν_y, ν_z and amplitudes A_1, A_2, A_3 of the drive field as well as the gradient strength G of the selection

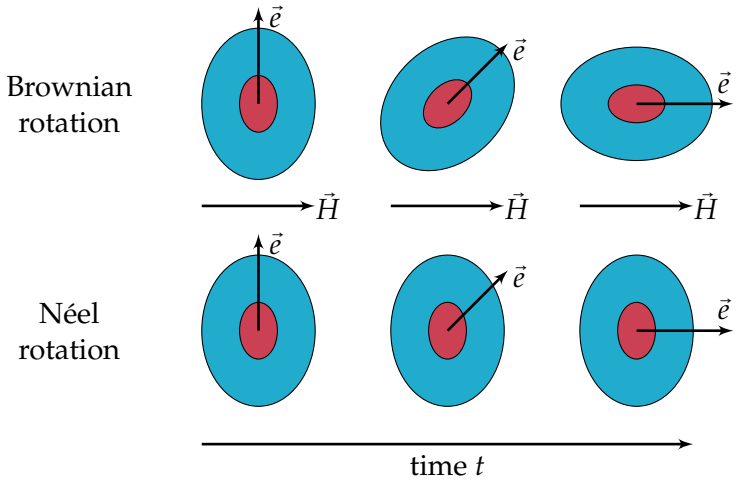


FIGURE 5.2: Illustration of Brown and Néel rotation of a magnetic nanoparticle. For a clearer illustration of the mechanisms, the particles are represented as ellipses.

field. As illustrated in Figure 5.2, the particle consists of a magnetic core (red) and a nonmagnetic hydrodynamic coating (blue), which are represented by ellipses for a clearer visualization of the rotation mechanisms. We have to distinguish between the core diameter D_C and the diameter of the entire particle D_H, which is the sum of the core diameter and two times the thickness of the coating. More physical parameters of the particles are, for example, the saturation magnetization M_S, the dynamic viscosity η of the surrounding fluid, and the constant K_{anis} expressing the anisotropy of the particle.

5.1.1 Relaxation Mechanism

For accurate modeling of the mean magnetic moment $\bar{m}$, we have to take into account the *relaxation time* of the particles, which is the time delay between the change of the magnetic field and the response of the particles. There are two different mechanisms, namely Brownian and Néel rotation, which were first considered in the context of MPI in [Wei+10].

The Brownian rotation describes the physical rotation of the entire nanoparticle, whereas Néel rotation describes the internal rotation of the magnetic moment inside the particle without physical rotation of the particle itself. The two rotation mechanisms are illustrated in Figure 5.2, where $\vec{e}$ represents the direction of the magnetic moment.

The relaxation time for Brownian rotation is given by

$$\tau_B = \frac{3\eta V_H}{k_B T_{\text{emp}}}, \tag{5.6}$$

where η is the viscosity of the surrounding fluid, V_H is the volume of the particle, k_B is the Boltzmann constant, and T_{emp} is the temperature. The relaxation time for Néel rotation is given by

$$\tau = \tau_N \exp\left(\tfrac{K_{\text{anis}} V_C}{k_B T_{\text{emp}}}\right),$$

where

$$\tau_N = \frac{V_C M_S}{k_B T_{\text{emp}} \gamma} \frac{1+\alpha^2}{2\alpha}, \tag{5.7}$$

and V_C is the magnetic core volume of the particle, M_S is the saturation magnetization, γ is the gyromagnetic ratio, K_{anis} is the anisotropy constant of the particle, and α is a dimensionless damping parameter.

Hence, the relaxation times depend on the physical properties of the particles as well as on the surrounding fluid and temperature. Moreover, for the relation of the two time constants τ_N and τ_B related to the relaxation times, the ratio of the magnetic coating to the core size is of special interest.

In different physical states, one of the two relaxation mechanisms can dominate, and thus the other one can be neglected. If the anisotropy constant is sufficiently large, the magnetic moment vector is aligned with the easy axis, which is the direction in which spontaneous magnetization is most likely to occur. In this case, the Brownian rotation dominates. This is typically the case for larger particles. Incorporating the Brownian rotation as well as the concept of thermal noise, we arrive at a Langevin equation for the mean magnetic moment [Klu18]. The Langevin equation is a SDE.

In contrast, for immobilized particles, we only need to consider Néel relaxation. Considering the magnetic field, the anisotropy of the particle, and the thermal noise, we can again derive a Langevin equation for the mean magnetic moment, which in this case is based on the Landau-Lifshitz-Gilbert equation. For more details, we refer to [Klu18]. However, in many practical applications none of the mechanisms dominates, and both relaxation mechanisms are relevant and have to be taken into account.

5.1.2 Coupled Brown-Néel Model

In this work, we focus on accurate modeling of the mean magnetic moment $\bar{m}$. Hence, we consider both relaxation mechanisms simultaneously and consider the coupled Brown-Néel model.

In what follows, we denote by $\vec{n} \in \mathbb{S}$ the normalized direction of the shell, where

$$\mathbb{S} = \{x \in \mathbb{R}^3 : \|x\|_2 = 1\}$$

is the unit sphere in $\mathbb{R}^3$. Moreover, we denote by $\vec{e} \in \mathbb{S}$ the normalized direction of the magnetic moment.

The coupled model describes the dynamics of $\vec{n}$ and $\vec{e}$ for fixed scanner and particle parameters as well as for a fixed spatial position $x \in \mathbb{R}^3$. Hence, we can interpret the spatial location as a parameter, which we emphasize by the notation $\vec{H}_x \colon \mathbb{R} \to \mathbb{R}^3$ with $\vec{H}_x(t) = \vec{H}(x,t)$, where $\vec{H}$ is the magnetic field defined in (5.1).

In this work, we will not discuss the physical derivation of the coupled model and refer to [Wei18] for details. The dynamics of the particle direction $\vec{n}$ and the magnetic moment direction $\vec{e}$ are described by an SDE, or more precisely, the Langevin

equation. In the notation of [Wei18], the equation is given by

$$\begin{aligned}\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} &= \frac{K_{\text{anis}}V_c}{3\eta V_H}(\vec{n}\cdot\vec{e})[(\vec{n}\times\vec{e})\times\vec{n}] + \frac{1}{\sqrt{\tau_B}}\vec{Y}\times\vec{n},\\ \frac{\mathrm{d}\vec{e}}{\mathrm{d}t} &= \frac{1}{2\tau_N}\frac{M_S V_C}{\alpha k_B T_{\text{emp}}}\Big(\Big[\vec{e}\times\Big(\vec{H}_x + \tfrac{2K_{\text{anis}}}{M_S}(\vec{e}\cdot\vec{n})\vec{n}\Big)\Big] + \Big[\vec{e}\times\Big(\vec{H}_x + \tfrac{2K_{\text{anis}}}{M_S}(\vec{e}\cdot\vec{n})\vec{n}\Big)\Big]\times\vec{e}\Big)\\ &\quad + \frac{1}{\sqrt{(1+\alpha^2)\tau_N}}\big(\vec{e}\times\vec{X} + (\vec{e}\times\vec{X})\times\vec{e}\big).\end{aligned} \tag{5.8}$$

This equation contains two independent three-dimensional standard white noise processes $\vec{Y}$ and $\vec{X}$, which are based on the Brownian and Néel rotation, respectively. White noise processes are stochastic processes with a mean of zero and zero autocorrelation for all nonzero time differences, which means that

$$\begin{aligned}\mathbb{E}(Z_i(t)) &= 0, \quad i &&= 1,\dots,6,\\ \mathbb{E}(Z_i(t_1)Z_j(t_2)) &= \delta_{ij}\mathbf{1}_{\{0\}}(t_1-t_2), \quad i,j &&= 1,\dots,6,\ t_1,t_2>0\end{aligned}$$

for the indicator function $\mathbf{1}_{\{0\}}\colon\mathbb{R}\to\{0,1\}$ with $\mathbf{1}_{\{0\}}(0) = 1$ and $\mathbf{1}_{\{0\}}(t) = 0$ for $t \neq 0$ and

$$Z(t) = \begin{pmatrix}\vec{Y}(t)\\ \vec{X}(t)\end{pmatrix} \in \mathbb{R}^6.$$

Hence, the white noise processes $\vec{X}$ and $\vec{Y}$ can be interpreted as distributional derivatives of standard Wiener processes. For details on this, we refer to [Øks03; Gli11]. The interpretation of the SDE (5.8) is in the Stratonovich sense, which is the prevailing option for SDEs on manifolds due to its invariance under diffeomorphisms. For details, we refer to [Hsu02, Section 1.2]. For the meaning of the physical parameters in (5.8), we refer to Table 5.1.

In notation according to [Pav14] using the Wiener process instead of the distributional derivatives, we write

$$\mathrm{d}\begin{pmatrix}\vec{n}_t\\ \vec{e}_t\end{pmatrix} = a(\vec{n}_t,\vec{e}_t,t)\mathrm{d}t + \hat{\sigma}(\vec{n}_t,\vec{e}_t)\circ\mathrm{d}W_t,$$

where the equation is interpreted as a stochastic integral equation in the Stratonovich sense and W_t is a six-dimensional Wiener process. Moreover, the drift term is given by

$$a(\vec{n},\vec{e},t) = \begin{pmatrix}a_1(\vec{n},\vec{e})\\ a_2(\vec{n},\vec{e},t)\end{pmatrix} \tag{5.9}$$

with

$$\begin{aligned}a_1(\vec{n},\vec{e}) &= \tfrac{K_{\text{anis}}V_c}{3\eta V_H}(\vec{n}\cdot\vec{e})[(\vec{n}\times\vec{e})\times\vec{n}],\\ a_2(\vec{n},\vec{e},t) &= \tfrac{1}{2\tau_N}\tfrac{M_S V_C}{\alpha k_B T_{\text{emp}}}\Big(\Big[\vec{e}\times\Big(\vec{H}_x(t) + \tfrac{2K_{\text{anis}}}{M_S}(\vec{e}\cdot\vec{n})\vec{n}\Big)\Big]\\ &\qquad + \Big[\vec{e}\times\Big(\vec{H}_x(t) + \tfrac{2K_{\text{anis}}}{M_S}(\vec{e}\cdot\vec{n})\vec{n}\Big)\Big]\times\vec{e}\Big).\end{aligned} \tag{5.10}$$

The diffusion matrix is given in the block form

$$\hat{\sigma}(\vec{n},\vec{e}) = \begin{pmatrix}\frac{1}{\sqrt{\tau_B}}\Lambda(\vec{n})^\top & 0\\ 0 & \frac{1}{\sqrt{(1+\alpha^2)\tau_N}}\Big(\Lambda(\vec{e}) + \Lambda(\vec{e})^\top\Lambda(\vec{e})\Big)\end{pmatrix},$$

where the skew-symmetric matrix $\Lambda(\vec{m}) \in \mathbb{R}^{3\times 3}$ is defined as

$$\Lambda(\vec{m}) = \begin{pmatrix} 0 & -m_3 & m_2 \\ m_3 & 0 & -m_1 \\ -m_2 & m_1 & 0 \end{pmatrix}.$$

By definition, $\Lambda(\vec{m})$ satisfies

$$\Lambda(\vec{m})\vec{u} = \vec{m} \times \vec{u}, \quad \Lambda(\vec{m})^\top \vec{u} = \vec{u} \times \vec{m} \quad \vec{u}, \vec{m} \in \mathbb{R}^3.$$

For details on the interpretation in the Stratonovich sense, the Stratonovich stochastic integral, Langevin equations, and stochastic processes, we refer to [Pav14].

In this model, the directions of the particle $\vec{n}$ and the magnetic moment $\vec{e}$ are coupled via the drift term a defined in (5.9). The coupling is based on the fact that a change in particle direction $\vec{n}$ or direction of magnetization $\vec{e}$ causes a change in energy. This, in turn, leads to a torque of the outer or inner angular momentum, respectively. For more details on the physical background of the coupling, we refer to [Wei18].

5.2 Fokker-Planck Equation

Under certain assumptions, the solution of a Stratonovich-type SDE is directly related to the solution of a Fokker-Planck equation for the probability density function of the stochastic variable. For details on this relation, we refer to [Pav14, Section 3.4]. The quantity of interest for MPI is the mean magnetic moment $\bar{m}$, which can be computed as an expectation value over the probability density function of the magnetic moment: let $p: \mathbb{S} \times \mathbb{S} \times I \to \mathbb{R}$ be the probability density function of the particle direction $\vec{n}$ and magnetic moment $\vec{e}$ at time $t \in I$, then

$$\bar{m}(t) = \int_{\mathbb{S}\times\mathbb{S}} \vec{e}\, p(\vec{n}, \vec{e}, t)\, \mathrm{d}(\vec{n}, \vec{e}). \tag{5.11}$$

The corresponding Fokker-Planck equation for f is a deterministic parabolic PDE on the Cartesian product $\mathbb{S} \times \mathbb{S}$ of two unit spheres. Since the PDE is posed on the surfaces of the spheres, we have to use the corresponding surface differential operators. Before we state the Fokker-Planck equation, we define these operators following [DE13].

Definition 5.2. The surface gradient of a function $f: \mathbb{S} \to \mathbb{R}$ on the sphere $\mathbb{S}$ at point $\vec{n}$ is defined by the projection of the standard gradient onto the tangent space of the sphere, i.e.,

$$\nabla_{\mathbb{S}} f = (I - \vec{n}\vec{n}^\top)\nabla \tilde{f} = \nabla \tilde{f} - (\vec{n} \cdot \nabla \tilde{f})\vec{n},$$

where $\tilde{f}$ is a smooth extension of f to a neighborhood of $\mathbb{S}$ such that $\tilde{f}|_{\mathbb{S}} = f$. Note that for notational simplicity, we will now write f for the extension as well. Furthermore, for a vector field $F : \mathbb{S} \to \mathbb{R}^3$, the surface divergence is defined as

$$\mathrm{div}_{\mathbb{S}}(F) = \mathrm{trace}(D_{\mathbb{S}} F) = (\nabla_{\mathbb{S}} F_1)_1 + (\nabla_{\mathbb{S}} F_2)_2 + (\nabla_{\mathbb{S}} F_3)_3,$$

where $D_{\mathbb{S}} F$ is the Jacobian with respect to gradient $\nabla_{\mathbb{S}}$. Moreover, we define the Laplace–Beltrami operator of a function $f: \mathbb{S} \to \mathbb{R}$ by

$$\Delta_{\mathbb{S}} f = \mathrm{div}_{\mathbb{S}}(\nabla_{\mathbb{S}} f).$$

As a direct consequence of Theorem 5.2, we have the following properties of the surface differential operators on the sphere.

Corollary 5.3. *For $f:\mathbb{S}\to\mathbb{R}$ and $F:\mathbb{S}\to\mathbb{R}^3$, we have*

$$\begin{aligned}\vec{n}\cdot\nabla_{\mathbb{S}}f &= \vec{n}\cdot\nabla f-(\vec{n}\cdot\nabla f)(\vec{n}\cdot\vec{n}) = \vec{n}\cdot\nabla f-\vec{n}\cdot\nabla f = 0,\\ \vec{n}\times\nabla_{\mathbb{S}}f &= \vec{n}\times\nabla f-(\vec{n}\cdot\nabla f)\vec{n}\times\vec{n} = \vec{n}\times\nabla f,\\ \operatorname{div}_{\mathbb{S}}(fF) &= f\operatorname{div}_{\mathbb{S}}(F)+F\cdot\nabla_{\mathbb{S}}f.\end{aligned}$$

For a function $f:\mathbb{S}\times\mathbb{S}\to\mathbb{R}$ defined on the product of two spheres, we denote by $\nabla_{\mathbb{S}_1}f$ and $\nabla_{\mathbb{S}_2}f$ the surface gradients with respect to the first and second sphere, respectively. Analogously, we define the surface divergence operators $\operatorname{div}_{\mathbb{S}_1}$ and $\operatorname{div}_{\mathbb{S}_2}$.

In the final part of this section, we state the Fokker-Planck equation for the coupled SDE system (5.8). The derivation in polar coordinates can be found in [Wei18]. For related SDE systems on the sphere, using only Brownian or Néel relaxation, the Fokker-Planck equations in Cartesian coordinates were derived in [Klu18]. For more details on the general derivation in Cartesian coordinates of the Fokker-Planck equation related to Langevin equations, we refer to [Ris96; Pav14] and, in particular, to [Gli11, Section 6.3] for Langevin equations on manifolds.

Proceeding as in [Klu18], we arrive at the Fokker-Planck equation for the coupled SDE system (5.8) of the form

$$\partial_t p = -\operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2}\big(ap-\mathbf{M}\,\nabla_{\mathbb{S}_1\times\mathbb{S}_2}p\big),\quad \text{in } \mathbb{S}\times\mathbb{S}\times[0,T] \tag{5.12}$$

with a defined in (5.9), final time $T>0$, and where

$$\nabla_{\mathbb{S}_1\times\mathbb{S}_2}p(\vec{n},\vec{e},t) = \begin{pmatrix}\nabla_{\mathbb{S}_1}p(\vec{n},\vec{e},t)\\ \nabla_{\mathbb{S}_2}p(\vec{n},\vec{e},t)\end{pmatrix}$$

as well as

$$\operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2}\begin{pmatrix}F_1\\F_2\end{pmatrix} = \operatorname{div}_{\mathbb{S}_1}(F_1)+\operatorname{div}_{\mathbb{S}_2}(F_2)$$

for $F_1,F_2:\mathbb{S}\times\mathbb{S}\times\mathbb{R}\to\mathbb{R}^3$. Moreover, the diffusion tensor is given by

$$\mathbf{M} = \begin{pmatrix}\frac{1}{2\tau_B}\mathbf{I}_3 & 0\\ 0 & \frac{1}{(1+\alpha^2)\tau_N}\mathbf{I}_3\end{pmatrix}, \tag{5.13}$$

where $\mathbf{I}_3\in\mathbb{R}^{3\times3}$ is the identity matrix. Furthermore, we have the initial condition $p(\vec{n},\vec{e},0)=p_0(\vec{n},\vec{e})$, where p_0 is the density of the initial distribution of $(\vec{n}_0,\vec{e}_0)$.

The diffusion matrix $\mathbf{M}$ depends on the time constants τ_B and τ_N defined in (5.6) and (5.7), respectively, and thus on the physical parameters of the nanoparticles and the surrounding fluid. In particular, we have

$$\mathbf{M} = \mathbf{M}(\eta, V_H, V_C, M_S, \gamma, \alpha, T_{\text{emp}}).$$

Moreover, the convection vector a depends on these physical parameters, as well as on the parameters of the applied magnetic field $\vec{H}$ and the spatial location $x\in\mathbb{R}^3$. Hence, we have

$$a(\vec{n},\vec{e},t) = a(\vec{n},\vec{e},t;x_1,x_2,x_3,\eta,K_{\text{anis}},V_H,V_C,M_S,\gamma,\alpha,T_{\text{emp}},\nu_x,\nu_y,\nu_z,A_x,A_y,A_z,G).$$

For notational simplicity, we denote by

$$\mu = (x_1, x_2, x_3, \eta, K_{\text{anis}}, V_H, V_C, M_S, \gamma, \alpha, T, \nu_x, \nu_y, \nu_z, A_x, A_y, A_z, G) \tag{5.14}$$

the collection of all parameters and write $\mathbf{M} = \mathbf{M}(\mu)$ and $a = a(\cdot, \cdot, \cdot; \mu)$.

As mentioned at the beginning of the chapter, we are interested in the mean magnetic moment $\bar{m}$ for different parameter μ. This calculation can be realized by approximating the solution of the Fokker-Planck equation (5.12) for different parameters μ and evaluating the expectation formula (5.11). To emphasize the parameter dependence, we write $p(\vec{n}, \vec{e}, t; \mu)$ for the solution of (5.12) with diffusion matrix $\mathbf{M} = \mathbf{M}(\mu)$ and convection term $a = a(\cdot, \cdot, \cdot; \mu)$. Correspondingly, we write $\bar{m}(t; \mu)$ for the parameter-dependent mean magnetic moment.

For the numerical approximation of the solution of (5.12) for a fixed parameter μ, we use a method that discretizes the angular variables $\vec{n}$ and $\vec{e}$ as well as the temporal variable t simultaneously. Hence, we rewrite (5.12) in angular-time divergence form

$$\begin{aligned} 0 &= \operatorname{div}_{\mathbb{S}_1 \times \mathbb{S}_2 \times \mathbb{R}} \left(\begin{pmatrix} a_1 \\ a_2 \\ 1 \end{pmatrix} p - \begin{pmatrix} \frac{1}{2\tau_B} \nabla_{\mathbb{S}_1} p \\ \frac{1}{(1+\alpha^2)\tau_N} \nabla_{\mathbb{S}_2} p \\ 0 \end{pmatrix} \right) && \text{in } \mathbb{S} \times \mathbb{S} \times [0, T], \\ p_0(\vec{n}, \vec{e}) &= p(\vec{n}, \vec{e}, 0) && \text{for all } \vec{n}, \vec{e} \in \mathbb{S}. \end{aligned} \tag{5.15}$$

One of the challenges when approximating the solution of (5.15) is that, for typical physical parameters in MPI, the problem is convection-dominated. This means that the diffusion coefficients in (5.13) are small compared to the magnitude of the convection vector a. Hence, we will use a variational formulation that is well-suited for convection-dominated problems. We will comment on this in more detail in Section 6.2.

Chapter 6

Reduced Basis Method for Convection-Dominated Parametric Fokker-Planck Equation

As described in Chapter 5, in MPI we are interested in the mean magnetic moment $\bar{m}(\cdot, \mu)$ for various parameter combinations $\mu \in \mathcal{P}$, where $\mathcal{P}$ denotes the set of parameters of interest. In particular, we are interested in the mean magnetic moment for different spatial locations $x \in \mathbb{R}^3$, which is essential for the spatial reconstruction in MPI. The spatial location x is one of the many parameters in μ. As seen in Section 5.2, the mean magnetic moment for a fixed parameter combination is the expectation (5.11) of the solution of the Fokker-Planck equation (5.12). However, solving the Fokker-Planck equation numerically for each parameter combination $\mu \in \mathcal{P}$ is extremely costly if the parameter space is large. Assuming each parameter in (5.14) can take 10 different values, we would have to solve the Fokker-Planck equation numerically $\#\mathcal{P} = 10^{18}$ times. Moreover, we aim for approximations of the mean magnetic moment with guarantees on the solution quality.

In order to efficiently compute the mean magnetic moment for many parameter combinations while maintaining accurate solutions, we use a model order reduction technique, namely the reduced basis (RB) method. The principal idea of the RB method is as follows: first we compute approximations $p^{(i)} \approx p(\cdot; \mu^{(i)})$ of the solution of the Fokker-Planck equation in an offline phase for selected parameter combinations $\mu^{(i)} \in \mathcal{P}$ and $i = 1, \dots, n$ using a suitable variational formulation. Then, in the online phase, we express the solution for a new parameter $\mu \in \mathcal{P}$ as a linear combination of the precomputed solutions

$$p(\cdot; \mu) \approx p_{\mathrm{RB}}(\cdot; \mu) = \sum_{i=1}^{n} c_i(\mu) p^{(i)}. \tag{6.1}$$

Using this representation, we can directly compute the mean magnetic moment using the linearity of the expectation (5.11). Hence, we obtain

$$\bar{m}(t; \mu) \approx \bar{m}_{\mathrm{RB}}(t; \mu) = \sum_{i=1}^{n} c_i(\mu) \bar{m}^{(i)}(t),$$

where $\bar{m}^{(i)}(t)$ is the mean magnetic moment corresponding to $p^{(i)}$. In what follows, we write $p(\mu) = p(\cdot; \mu)$ for notational simplicity.

To be able to calculate appropriate approximations $p^{(i)} \approx p(\mu^{(i)})$ in the offline phase, we need a suitable variational formulation for the convection-dominated Fokker-Planck equation in Section 5.2. The computation of each approximation

for fixed parameter can be already very expensive, since we have to solve a five-dimensional PDE. Nevertheless, using specific properties of the variational formulation, we can derive a low-dimensional system of equations of size $n \times n$ for the coefficients $c_i(\mu)$ in (6.1). Then the computational costs of calculating the coefficients $c_i(\mu)$ for a new parameter only depend on the number n of precomputed solutions, not on the dimension of the underlying variational formulation. This means that the expensive offline phase for solving the higher-dimensional PDEs does not influence the computational costs of the online phase. Hence, if n is small compared to the dimension of the full model, we can compute approximations of the mean magnetic moment efficiently for numerous parameter combinations.

We use a variational formulation for the Fokker-Planck equation in MPI, applied to the angular-time divergence form in (5.15), which yields a simultaneous discretization in time and angular variables. Consequently, unlike traditional time-stepping methods, this space-time approach computes the solution for the entire time interval at once rather than at discrete, sequential time steps.

In this section, we mainly follow the ideas proposed in [Fal+25] for the parametric Fokker-Planck equation related to MPI. This chapter is structured as follows. In Section 6.1, we give a brief overview of the RB method and its main ingredients. In Section 6.2, we present a variational formulation that is well-suited for convection-dominated problems and fits into the framework of the RB method. In Section 6.3 we subsequently collect some properties of finite elements on surfaces. In particular, we present Raviart-Thomas elements of arbitrary order on spherical triangles. Finally, in Section 6.4, we present numerical results for the parametric Fokker-Planck equation with fixed parameters as well as results for the RB method.

6.1 Reduced Basis Method

In this section, we give a brief overview of the RB method, where we mainly follow the notation of [QMN16; HRS16]. In particular, we focus on its main ingredients that are necessary for the application of this method to the parametric Fokker-Planck equation in Section 5.2. However, we retain the notation as general as possible. To this end, we denote by $u(\mu)$ the solution of the parametrized PDE for a parameter combination $\mu \in \mathcal{P}$, where $\mathcal{P}$ is the set of parameters.

The RB method is a model order reduction technique for parametrized PDEs, which is divided into two parts. First, in the *offline phase*, we compute accurate approximations $u^{(i)} \approx u(\mu^{(i)})$ of the solution for selected parameter combinations $\mu^{(i)} \in \mathcal{P}$ using a suitable numerical method. These approximations $u^{(i)}$ are called *snapshots* and form the *reduced basis* $\{u^{(1)}, \dots, u^{(n)}\}$.

Then in the *online phase*, we use the RB to compute approximations of $u(\mu)$ for new parameter combinations $\mu \in \mathcal{P}$. The main idea is to express the approximation as a linear combination of the precomputed snapshots as in (6.1). The RB method relies on the assumption that the solution manifold

$$\mathcal{M} = \{u(\mu) : \mu \in \mathcal{P}\}, \tag{6.2}$$

which is the set of solutions for all parameter combinations, can be well approximated by a low-dimensional subspace. Under certain assumptions, the representation in this subspace leads to a low-dimensional system of equations for the coefficients $\mathbf{c}(\mu)$ in (6.1), which can therefore be solved efficiently. We can interpret this approach as a low-rank approximation of the solution manifold. Here we separate between the domain and the parameter space, as can be seen in (6.1), where the

snapshots $p^{(i)}$ depend on the domain variables and the coefficients $c_i(\mu)$ depend on the parameters.

For the calculation of snapshots in the offline phase and for the derivation of the low-dimensional system of equations in the online phase, we require a suitable variational formulation of the underlying parametrized PDE. We first describe the RB method in abstract form independently of the specific choice of the variational formulation. A concrete choice of the variational formulation for the parametric Fokker-Planck equation (5.15) related to MPI will be stated in Section 6.2.

We assume the solution $u(\mu)$ of the parametrized PDE to be in some function space W and that we have a finite-dimensional finite element space $W_h \subset W$. Moreover, we assume that there is a parametrized bilinear form $B(\cdot,\cdot;\mu)\colon W \times W \to \mathbb{R}$ and a parametrized linear functional $l(\cdot;\mu)\colon W \to \mathbb{R}$. For a fixed parameter μ, the variational formulation is: find $u_h(\mu) \in W_h$ such that

$$B(u_h(\mu), v_h; \mu) = l(v_h; \mu) \quad \text{for all } v_h \in W_h. \tag{6.3}$$

Furthermore, we assume that we have a norm related to the bilinear form. Since the variational formulation is parameter-dependent, the norm will also generally depend on the parameter μ. We denote this parameter-dependent norm by $\|\cdot\|_\mu$. Moreover, we assume the bilinear form to be uniformly bounded with respect to this norm, i.e., it exists a constant $\gamma > 0$ independent of μ and for each $\mu \in \mathcal{P}$ there exist constants $\gamma(\mu) \leq \gamma < \infty$ such that

$$|B(w_h, v_h; \mu)| \leq \gamma(\mu)\|w_h\|_\mu \|v_h\|_\mu \quad \text{for all } w_h, v_h \in W_h.$$

In many cases, the bilinear form is additionally assumed to be coercive. However, the bilinear form associated to the variational formulation we consider in Section 6.2 is not coercive. Instead, we assume the variational formulation to be inf-sup stable, i.e., there exists a constant $\beta > 0$ such that

$$\beta(\mu) = \inf_{w_h \in W_h \setminus \{0\}} \sup_{v_h \in W_h \setminus \{0\}} \frac{B(w_h, v_h; \mu)}{\|w_h\|_\mu \|v_h\|_\mu} \geq \beta \quad \text{for all } \mu \in \mathcal{P}. \tag{6.4}$$

If we already have an RB approximation $u_{\mathrm{RB}}(\mu) \approx u(\mu)$, then we can estimate the accuracy of the RB approximation using the triangle inequality by

$$\|u(\mu) - u_{\mathrm{RB}}(\mu)\|_\mu \leq \|u(\mu) - u_h(\mu)\|_\mu + \|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_\mu, \tag{6.5}$$

which is the sum of the error of the finite element approximation and the RB approximation with respect to the finite element solution. Based on the properties of the variational formulation, we assume that we can control the finite element error. Thus, the focus of the RB method is to control the second term in (6.5).

6.1.1 Offline Phase

The goal of the offline phase is to construct a RB space $W_{\mathrm{RB}} \subset W_h$ with

$$W_{\mathrm{RB}} = \operatorname{span}\left\{u^{(1)}, \ldots, u^{(n)}\right\},$$

such that for each parameter $\mu \in \mathcal{P}$ there exists a RB approximation $u_{\mathrm{RB}}(\mu) \in W_{\mathrm{RB}}$ which is a good approximation of the finite element solution $u_h(\mu)$, meaning that the error $\|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_\mu$ is suitably small.

Algorithm 6.1: Greedy Algorithm

Input: tol > 0, $\mu_1 \in \mathcal{P}_\delta$ with $\mathcal{P}_\delta$ finite subset of $\mathcal{P}$.
Output: Reduced basis space W_{RB}.

begin
  $n = 1$
  repeat
    Compute snapshot $u^{(n)} = u_h(\mu_n)$
    $W_{\mathrm{RB}} = \mathrm{span}\{u^{(1)}, \dots, u^{(n)}\}$
    for $\mu \in \mathcal{P}_\delta$ **do**
      Compute RB approximation $u_{\mathrm{RB}}(\mu)$
      Evaluate error estimator $\eta(\mu)$
    end
    $n = n + 1$
    $\mu_n = \arg\max_{\mu \in \mathcal{P}_\delta} \eta(\mu)$
  until $\eta(\mu_n) \leq$ tol;
end

Compared to the online phase, the offline phase is allowed to be computationally expensive, since we only need to calculate the RB once. The main challenge is selecting suitable parameter combinations $\mu^{(i)}$ for computing the snapshots $u^{(i)}$. The snapshots themselves can be calculated using the variational formulation.

There are multiple strategies for selecting parameter combinations for the snapshots. In this work, we use the *greedy algorithm*, which iteratively selects parameter combinations based on an error estimator $\eta: \mathcal{P} \to \mathbb{R}$ for the RB approximation. This error estimator estimates the second term on the right-hand side of (6.5), i.e., we assume

$$\|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_\mu \leq \eta(\mu) \quad \text{for all } \mu \in \mathcal{P}. \tag{6.6}$$

Remark 6.1. The error estimator in (6.6) can also be chosen to estimate the error in a parameter-independent norm $\|\cdot\|_W$, i.e.,

$$\|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_W \leq \eta(\mu) \quad \text{for all } \mu \in \mathcal{P}.$$

Alternatively, instead of estimating the approximation error, we can estimate the error of a quantity of interest. For example, in the context of the parametric Fokker-Planck equation in Section 5.2, we could estimate the corresponding error of the mean magnetic moment in a suitable norm.

The greedy algorithm is summarized in Algorithm 6.1. The algorithm is an iterative procedure, where in each iteration we increase the RB space by one basis function, called snapshot. Assume the RB space is already constructed by n snapshots. Then the basic idea is to select the parameter combination μ_{n+1} for which the error estimator $\eta(\mu)$ is maximal over a finite training set $\mathcal{P}_\delta \subset \mathcal{P}$. The evaluation of the error estimator usually requires the computation of the RB approximation $u_{\mathrm{RB}}(\mu)$ for all $\mu \in \mathcal{P}_\delta$. Hence, we calculate the RB approximation for each parameter, evaluate the error estimator, and select the parameter μ_{n+1} with the maximal error estimator. Then we compute the snapshot $u^{(n+1)}$ for the new parameter combination μ_{n+1} and extend the RB space. We repeat this procedure and thus add basis elements until the error estimator is below a given tolerance tol > 0 for all parameter combinations in

the training set, i.e.,

$$\eta(\mu) \leq \text{tol} \quad \text{for all } \mu \in \mathcal{P}_\delta.$$

To ensure that finding the maximal value of the error estimator is computationally tractable, we restrict the evaluation to a finite subset of parameters $\mathcal{P}_\delta$, which is called *training set*. This set should be chosen such that it represents the whole parameter space, which means that the solution manifold of the training set should be close to the solution manifold of the full parameter space $\mathcal{P}$. However, for large discrete subsets, it is essential that we can compute the RB approximation and evaluate the error estimator efficiently.

The approximation properties of the greedy algorithm can be analyzed under certain assumptions using the theory of Kolmogorov n-widths. For details, we refer to [Buf+12; HRS16].

6.1.2 Online Phase

The goal of the online phase is the efficient computation of approximations of the solutions of the PDE for new parameter combinations $\mu \in \mathcal{P}$ in the RB space

$$W_{\text{RB}} = \text{span}\{u^{(1)}, \dots, u^{(n)}\}.$$

This means that we determine coefficients $\mathbf{c}(\mu) \in \mathbb{R}^n$ such that

$$u_{\text{RB}}(\mu) = \sum_{i=1}^{n} c_i(\mu) u^{(i)} \approx u(\mu)$$

is a suitable approximation. In general, we assume that $n \ll N$, where $N = \dim(W_h)$ is the dimension of the finite element space, meaning the solution manifold in (6.2) can be well approximated by the low-dimensional subspace W_{RB}. Ideally, the computational costs for computing the coefficients $\mathbf{c}(\mu)$ depend only on the number of snapshots n and not on the dimension of the finite element space N. To achieve this, we need certain assumptions on the bilinear form $B(\cdot,\cdot;\mu)$ and the linear functional $l(\cdot;\mu)$.

Definition 6.2. The bilinear form $B(\cdot,\cdot;\mu)$ and the linear functional $l(\cdot;\mu)$ have an *affine decomposition* if there exist parameter-independent bilinear forms $B_q : W \times W \to \mathbb{R}$ for $q = 1, \dots, Q_B$ and parameter-independent linear functionals $l_q : W \to \mathbb{R}$ for $q = 1, \dots, Q_l$ as well as parameter functions $\theta_q^{(B)} : \mathcal{P} \to \mathbb{R}$ for $q = 1, \dots, Q_B$ and $\theta_q^{(l)} : \mathcal{P} \to \mathbb{R}$ for $q = 1, \dots, Q_l$ with $Q_B, Q_l \in \mathbb{N}$ such that

$$B(w, v; \mu) = \sum_{q=1}^{Q_B} \theta_q^{(B)}(\mu) B_q(w, v) \qquad \text{for all } w, v \in W,$$

$$l(v; \mu) = \sum_{q=1}^{Q_l} \theta_q^{(l)}(\mu) l_q(v) \qquad \text{for all } v \in W.$$

Let $\{\phi_1, \dots, \phi_N\}$ be a basis of the finite element space W_h. Then by the same arguments as in (3.1.1), the solution $u_h(\mu) = \sum_{i=1}^{N} u_i(\mu)\phi_i$ of the variational formulation (6.3) is given by the solution of the linear system

$$\mathbf{A}_h(\mu)\mathbf{u}(\mu) = \mathbf{f}_h(\mu),$$

where $\mathbf{u}(\mu) = (u_i(\mu))_{i=1,\dots,N} \in \mathbb{R}^N$ and

$$\begin{aligned} \mathbf{A}_h(\mu) &= (B(\phi_j, \phi_i; \mu))_{i,j=1,\dots,N}, \\ \mathbf{f}_h(\mu) &= (l(\phi_i; \mu))_{i=1,\dots,N}. \end{aligned} \tag{6.7}$$

Using the affine decomposition of the bilinear form, we obtain

$$\mathbf{A}_h(\mu) = \sum_{q=1}^{Q_B} \theta_q^{(B)}(\mu) \mathbf{A}_h^{(q)}, \tag{6.8}$$

where

$$\mathbf{A}_h^{(q)} = (B_q(\phi_j, \phi_i))_{i,j=1,\dots,N}, \quad q = 1, \dots, Q_B.$$

Similarly, we have

$$\mathbf{f}_h(\mu) = \sum_{q=1}^{Q_l} \theta_q^{(l)}(\mu) \mathbf{f}_h^{(q)}, \tag{6.9}$$

where

$$\mathbf{f}_h^{(q)} = (l_q(\phi_i))_{i=1,\dots,N}, \quad q = 1, \dots, Q_l.$$

Since the bilinear forms B_q and the linear forms l_q are parameter-independent, the matrices $\mathbf{A}_h^{(q)}$ and the vectors $\mathbf{f}_h^{(q)}$ are parameter-independent as well. Hence, we also have an affine decomposition of the system matrix $\mathbf{A}_h(\mu)$ and the right-hand side $\mathbf{f}_h(\mu)$. Moreover, the matrices $\mathbf{A}_h^{(q)}$ and the vectors $\mathbf{f}_h^{(q)}$ can be precomputed.

In the case of a coercive bilinear form, the RB approximation can be computed using a Galerkin projection onto the RB space W_{RB}. However, in our case, the bilinear form is only inf-sup stable (6.4) with respect to the space W_h. In general, the inf-sup stability is not inherited from W_h to the RB space W_{RB}. As an alternative, we consider a least-squares approach as described in [QMN16, Section 3.3.2]. Therefore, we assume that there exists a symmetric positive-definite matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{N \times N}$ such that

$$\|\tilde{u}\|_W \eqsim \|\mathbf{A}_h(\mu)\tilde{\mathbf{u}}\|_{\tilde{\mathbf{A}}^{-1}} \quad \text{for all } \tilde{u} = \sum_{i=1}^{N} \tilde{u}_i \phi_i \in W_h \text{ with } \tilde{\mathbf{u}} = (\tilde{u}_i)_{i=1,\dots,N}, \tag{6.10}$$

where $\|\cdot\|_{\tilde{\mathbf{A}}^{-1}}$ is the norm induced by the matrix $\tilde{\mathbf{A}}^{-1}$, i.e.,

$$\|\mathbf{v}\|_{\tilde{\mathbf{A}}^{-1}}^2 = \langle \mathbf{v}, \tilde{\mathbf{A}}^{-1}\mathbf{v} \rangle \quad \text{for } \mathbf{v} \in \mathbb{R}^N.$$

In many applications, the matrix $\tilde{\mathbf{A}}$ is chosen as the discretization of the inner product of the function space W, i.e.,

$$\|\tilde{u}\|_W^2 = \langle \tilde{u}, \tilde{u} \rangle_W = \langle \tilde{\mathbf{u}}, \tilde{\mathbf{A}}\tilde{\mathbf{u}} \rangle = \|\tilde{\mathbf{u}}\|_{\tilde{\mathbf{A}}}^2 \quad \text{for all } \tilde{u} = \sum_{i=1}^{N} \tilde{u}_i \phi_i \text{ with } \tilde{\mathbf{u}} = (\tilde{u}_i)_{i=1,\dots,N}.$$

Moreover, we have

$$\|\mathbf{v}\|_{\tilde{\mathbf{A}}^{-1}} = \|\tilde{\mathbf{A}}^{-1}\mathbf{v}\|_{\tilde{\mathbf{A}}} \quad \text{for } \mathbf{v} \in \mathbb{R}^N$$

and in particular

$$\|\mathbf{A}_h(\mu)\tilde{\mathbf{u}}\|_{\tilde{\mathbf{A}}^{-1}} = \|\tilde{\mathbf{A}}^{-1}\mathbf{A}_h(\mu)\tilde{\mathbf{u}}\|_{\tilde{\mathbf{A}}}.$$

Hence, the constants in (6.10) depend on $\tilde{\mathbf{A}}^{-1}\mathbf{A}_h(\mu)$ and thus $\tilde{\mathbf{A}}$ can be interpreted as a preconditioner for the matrix $\mathbf{A}_h(\mu)$.

Considering the solution $u_h(\mu) \in W_h$ of the variational formulation and an approximation $\tilde{u} \in W_h$ with corresponding coefficient vector $\tilde{\mathbf{u}} \in \mathbb{R}^N$, we have

$$\|u_h(\mu) - \tilde{u}\|_\mu \approx \|\mathbf{A}_h(\mu)\tilde{\mathbf{u}} - \mathbf{f}_h(\mu)\|_{\tilde{\mathbf{A}}^{-1}}. \tag{6.11}$$

Based on this equivalence, we define the RB approximation

$$u_{\mathrm{RB}}(\mu) = \sum_{i=1}^{n} c_i(\mu) u^{(i)} \in W_{\mathrm{RB}},$$

determined by the coefficients $\mathbf{c}(\mu)$ as the solution of the least-squares problem

$$\mathbf{c}(\mu) = \arg\min_{\tilde{\mathbf{c}} \in \mathbb{R}^n} \left\| \mathbf{A}_h(\mu) \sum_{i=1}^{n} \tilde{c}_i(\mu) u^{(i)} - \mathbf{f}_h(\mu) \right\|_{\tilde{\mathbf{A}}^{-1}}. \tag{6.12}$$

Setting

$$\mathbf{U} = \begin{pmatrix} \vdots & & \vdots \\ \mathbf{u}^{(1)} & \cdots & \mathbf{u}^{(n)} \\ \vdots & & \vdots \end{pmatrix} \in \mathbb{R}^{N \times n}, \tag{6.13}$$

where $\mathbf{u}^{(i)} \in \mathbb{R}^N$ is the coefficient vector of the snapshot $u^{(i)}$, the least-squares problem (6.12) can be rewritten as

$$\mathbf{c}(\mu) = \arg\min_{\tilde{\mathbf{c}} \in \mathbb{R}^n} \left\| \mathbf{A}_h(\mu)\mathbf{U}\tilde{\mathbf{c}} - \mathbf{f}_h(\mu) \right\|_{\tilde{\mathbf{A}}^{-1}}.$$

Using the definition of the norm $\|\cdot\|_{\tilde{\mathbf{A}}^{-1}}$, we obtain that the solution of the least-squares problem is given by the solution of the normal equation

$$(\mathbf{A}_h(\mu)\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{A}_h(\mu)\mathbf{U}\mathbf{c}(\mu) = (\mathbf{A}_h(\mu)\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{f}_h(\mu).$$

We set

$$\begin{aligned} \mathbf{A}_{\mathrm{RB}}(\mu) &= (\mathbf{A}_h(\mu)\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{A}_h(\mu)\mathbf{U}, \\ \mathbf{f}_{\mathrm{RB}}(\mu) &= (\mathbf{A}_h(\mu)\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{f}_h(\mu). \end{aligned} \tag{6.14}$$

Using the affine representation (6.8) of the matrix $\mathbf{A}_h(\mu)$, we obtain

$$\mathbf{A}_{\mathrm{RB}}(\mu) = \sum_{p,q=1}^{Q_B} \theta_p^{(B)}(\mu)\theta_q^{(B)}(\mu)(\mathbf{A}_h^{(p)}\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{A}_h^{(q)}\mathbf{U}.$$

Hence, we define the parameter-independent matrices

$$\mathbf{G}^{(p,q)} = (\mathbf{A}_h^{(p)}\mathbf{U})^\top \tilde{\mathbf{A}}^{-1} \mathbf{A}_h^{(q)}\mathbf{U} \in \mathbb{R}^{n \times n} \quad \text{for } p, q = 1, \ldots, Q_B. \tag{6.15}$$

Since $(\mathbf{A}_h^{(q)}\mathbf{U}) \in \mathbb{R}^{N \times n}$, the computation of the matrices $\mathbf{G}^{(p,q)}$ requires the solution of n linear systems of size $N \times N$. However, the matrices $\mathbf{G}^{(p,q)}$ are parameter-independent. Thus, they need to be calculated only once and can be precomputed in the offline phase. Hence, we have the affine representation

$$\mathbf{A}_{\mathrm{RB}}(\mu) = \sum_{p,q=1}^{Q_B} \theta_p^{(B)}(\mu)\theta_q^{(B)}(\mu)\mathbf{G}^{(p,q)}. \tag{6.16}$$

Algorithm 6.2: Online Phase

Input: $\mu \in \mathcal{P}_\delta$,
$\mathbf{G}^{(p,q)}$ for $p,q = 1,\dots,Q_B$,
$\mathbf{h}^{(p,q)}$ for $p = 1,\dots,Q_B$, $q = 1,\dots,Q_l$.

Output: $u_{\mathrm{RB}}(\mu)$.

begin

Compute $\theta_p^{(B)}(\mu)$ for $p = 1,\dots,Q_B$

Compute $\theta_q^{(l)}(\mu)$ for $q = 1,\dots,Q_l$

Assemble the matrix and right-hand side

$$\mathbf{A}_{\mathrm{RB}}(\mu) = \sum_{p,q=1}^{Q_B} \theta_p^{(B)}(\mu)\theta_q^{(B)}(\mu)\mathbf{G}^{(p,q)}, \quad \mathbf{f}_{\mathrm{RB}}(\mu) = \sum_{p=1}^{Q_B}\sum_{q=1}^{Q_l} \theta_p^{(B)}(\mu)\theta_q^{(l)}(\mu)\mathbf{h}^{(p,q)}$$

Solve the linear system $\mathbf{A}_{\mathrm{RB}}(\mu)\mathbf{c}(\mu) = \mathbf{f}_{\mathrm{RB}}(\mu)$

$u_{\mathrm{RB}}(\mu) = \sum_{i=1}^{n} c_i(\mu)u^{(i)}$

end

Similarly, using the affine representations (6.8) and (6.9), we obtain

$$\mathbf{f}_{\mathrm{RB}}(\mu) = \sum_{p=1}^{Q_B}\sum_{q=1}^{Q_l} \theta_p^{(B)}(\mu)\theta_q^{(l)}(\mu)(\mathbf{A}_h^{(p)}\mathbf{U})^\top \tilde{\mathbf{A}}^{-1}\mathbf{f}_h^{(q)}.$$

As before, we define the parameter-independent vectors

$$\mathbf{h}^{(p,q)} = (\mathbf{A}_h^{(p)}\mathbf{U})^\top \tilde{\mathbf{A}}^{-1}\mathbf{f}_h^{(q)} = (\tilde{\mathbf{A}}^{-1}\mathbf{A}_h^{(p)}\mathbf{U})^\top \mathbf{f}_h^{(q)} \in \mathbb{R}^n \quad \text{for } p = 1,\dots,Q_B,\ q = 1,\dots,Q_l,$$

where we used the symmetry of the matrix $\tilde{\mathbf{A}}$. The vectors $\mathbf{h}^{(p,q)}$ are also parameter-independent and can be precomputed in the offline phase. Moreover, the matrices $\tilde{\mathbf{A}}^{-1}\mathbf{A}_h^{(p)}\mathbf{U} \in \mathbb{R}^{N\times n}$ have already been computed during the calculation of the matrices $\mathbf{G}^{(p,q)}$ and can be reused here. Hence, we have the affine representation of the right-hand side of the normal equation

$$\mathbf{f}_{\mathrm{RB}}(\mu) = \sum_{p=1}^{Q_B}\sum_{q=1}^{Q_l} \theta_p^{(B)}(\mu)\theta_q^{(l)}(\mu)\mathbf{h}^{(p,q)}.$$

The online phase is summarized in Algorithm 6.2. We start by evaluating the parameter-dependent functions $\theta_p^{(B)}(\mu)$ and $\theta_q^{(l)}(\mu)$ for $p = 1,\dots,Q_B$ and $q = 1,\dots,Q_l$ for a given parameter $\mu \in \mathcal{P}$. Since the matrices $\mathbf{G}^{(p,q)}$ and the vectors $\mathbf{h}^{(p,q)}$ are precomputed, we can assemble the RB matrix $\mathbf{A}_{\mathrm{RB}}(\mu) \in \mathbb{R}^{n\times n}$ by summing Q_B^2 matrices of size $n \times n$ and the right-hand side $\mathbf{f}_{\mathrm{RB}}(\mu) \in \mathbb{R}^n$ by summing $Q_B Q_l$ vectors of size n. Finally, we solve the RB system of equations $\mathbf{A}_{\mathrm{RB}}(\mu)\mathbf{c}(\mu) = \mathbf{f}_{\mathrm{RB}}(\mu)$ of size $n \times n$. Then the RB approximation is given by

$$u_{\mathrm{RB}}(\mu) = \sum_{i=1}^{n} c_i(\mu)u^{(i)}. \tag{6.17}$$

Hence, the computational costs to compute the coefficients $\mathbf{c}(\mu)$ are independent of

the dimension of the finite element space N and depend only on the number of snapshots n as well as the numbers Q_B and Q_l of the terms in the affine decompositions.

In many practical applications, we are not interested in the solution of the PDE itself, but in specific quantities of interest, e.g., $Q(u(\mu))$, where $Q\colon W \to V$ is a linear mapping into some normed space V.

Remark 6.3. In the context of the parametric Fokker-Planck equation related to MPI in Section 5.2, we are interested in the mean magnetic moment defined in (5.11). This quantity of interest is given by

$$Q(p) = \int_{\mathbb{S}\times\mathbb{S}} \vec{e}\, p(\vec{n},\vec{e},\cdot)\,\mathrm{d}(\vec{n},\vec{e}) \in L_2(0,T).$$

In this case, we can calculate the coefficient vector $\mathbf{c}(\mu)$ as before, but we do not need to compute the actual RB approximation $u_{\mathrm{RB}}(\mu)$ in (6.17). Instead, we can precompute the quantities of interest for the snapshots in the offline phase, i.e., we compute $Q^{(i)} = Q(u^{(i)})$ for $i = 1,\dots,n$. Then using the linearity of Q, we obtain

$$Q(u_{\mathrm{RB}}(\mu)) = \sum_{i=1}^{n} c_i(\mu) Q^{(i)}.$$

6.2 Hybrid Discontinuous Galerkin Finite Element Method

In this section, we present a finite element method for the Fokker-Planck equation (5.12) that fits within the RB framework presented in Section 6.1. As mentioned in Section 5.2, the Fokker-Planck equation is convection-dominated for typical parameters in MPI. For this reason, we need a finite element method that is suitable for this type of problem. In this work, we use the hybrid discontinuous Galerkin (HDG) method described in [ES10]. However, in principle, the finite element method can be replaced by any other suitable method that fits into the RB framework.

In previous works regarding Fokker-Planck equations related to MPI, different numerical schemes have been considered. For a simplified Fokker-Planck equation using only Brown or Néel relaxation mechanisms, a finite volume method was used for the discretization of the spherical variable in [AKK20], which yields an ordinary differential equation system in time. This ordinary differential equation can then be solved by an time-stepping scheme. Another approach is spherical harmonics, which was considered in [Wei18] for the coupled model and in [AKK20] for the simplified model. However, both of these methods are generally not well-suited for convection-dominated problems. Moreover, in our approach we aim for high-order methods.

To introduce the HDG method, we first present the method in a simplified setting and in the notation of [ES10]. Then we state the necessary modifications to apply this method to the Fokker-Planck equation (5.12) and to the RB framework in Section 6.1.

We consider a stationary convection-diffusion equation of the form

$$\begin{aligned} \operatorname{div}(-\varepsilon\nabla u + \beta u) &= f && \text{in } \Omega, \\ u &= g_D && \text{on } \partial\Omega, \end{aligned} \tag{6.18}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded open domain, $\varepsilon\colon \Omega \to \mathbb{R}$ is a nonnegative function and $\beta\colon \Omega \to \mathbb{R}^d$ is a vector field. We introduce the diffusive flux $\sigma = -\varepsilon\nabla u$ and rewrite

(6.18) in mixed form as

$$\begin{aligned} \sigma + \varepsilon \nabla u = 0, \quad \operatorname{div}(\sigma + \beta u) &= f && \text{in } \Omega, \\ u &= g_D && \text{on } \partial\Omega. \end{aligned} \tag{6.19}$$

6.2.1 Basic Assumptions and Notation

Before we state the variational formulation, we introduce some basic notation and assumptions from [ES10]. In particular, not all of them are satisfied in the Fokker-Planck equation (5.15) related to MPI. However, some of these assumptions are for notational simplicity and can be relaxed. We will comment further on this in Section 6.2.4.

Let $\mathcal{T}_h$ be a shape-regular partition of the domain Ω into simplices T with mesh size $h > 0$. Moreover, let $\mathcal{E}_h$ be the set of all facets of the elements $T \in \mathcal{T}_h$. We assume that each element $T \in \mathcal{T}_h$ is generated by a regular mapping $F_T \colon \hat{T} \to T$, where $\hat{T}$ is the reference simplex in $\mathbb{R}^d$. Similarly, we assume that each facet $E \in \mathcal{E}_h$ is generated by a regular mapping $F_E \colon \hat{E} \to E$, where $\hat{E}$ is the reference facet in $\mathbb{R}^{d-1}$. Additionally, we denote by $\partial\mathcal{T}_h$ the set of all element boundaries ∂T associated with the corresponding simplex. Hence, the same facet can appear multiple times if two elements share it.

We assume homogeneous Dirichlet boundary conditions, i.e., $g_D = 0$. Regarding the coefficients, we assume that ε and β are constant on each element $T \in \mathcal{T}_h$ and that the normal components of the convection field β are continuous across element interfaces, i.e., for each facet $E \in \mathcal{E}_h$ shared by two elements $T_1, T_2 \in \mathcal{T}_h$, we have $\beta|_{T_1} \cdot n_{T_1} + \beta|_{T_2} \cdot n_{T_2} = 0$, where n_{T_i} is the outward unit normal vector of T_i on E for $i = 1, 2$. This assumption implies $\operatorname{div}(\beta) = 0$. Furthermore, it induces a splitting of the element boundaries into inflow and outflow boundaries

$$\partial T^{\text{out}} = \{x \in \partial T : \beta_T \cdot n_T > 0\}, \quad \partial T^{\text{in}} = \partial T \setminus \partial T^{\text{out}}, \quad T \in \mathcal{T}_h. \tag{6.20}$$

Correspondingly, we define by $\partial\mathcal{T}^{\text{out}}$ and $\partial\mathcal{T}^{\text{in}}$, the sets of all element inflow and outflow boundaries, respectively, again associated with the corresponding simplices. In the same manner, we define the global inflow and outflow boundaries of the domain Ω and denote them by $\partial\Omega^{\text{out}}$ and $\partial\Omega^{\text{in}}$, respectively.

Next, we introduce the broken Sobolev space with respect to the mesh $\mathcal{T}_h$ by

$$H^s(\mathcal{T}_h) = \{v \in L_2(\Omega) : v|_T \in H^s(T), T \in \mathcal{T}_h\}, \quad s \geq 0.$$

Moreover, for $v \in H^{s+1}(\mathcal{T}_h)$, we denote by $\nabla v \in (H^s(\mathcal{T}_h))^d$ the piecewise gradient defined by $(\nabla v)|_T = \nabla(v|_T)$ for all $T \in \mathcal{T}_h$. For functions $u, v \in (L_2(\Omega))^n$, we have the standard L_2-inner product

$$\langle u, v \rangle_{\mathcal{T}_h} = \sum_{T \in \mathcal{T}_h} \langle u, v \rangle_T, \quad \text{where} \quad \langle u, v \rangle_T = \int_T u \cdot v \, \mathrm{d}x.$$

The function spaces for the element interfaces are defined by

$$\begin{aligned} L_2(\mathcal{E}_h) &= \{\xi \colon \xi \in L_2(E), \text{for all } E \in \mathcal{E}_h\}, \\ L_2(\partial\mathcal{T}_h) &= \{v \colon v \in L_2(\partial T), \text{for all } T \in \mathcal{T}_h\}. \end{aligned}$$

Functions in $L_2(\partial\mathcal{T}_h)$ are potentially double-valued on the element interfaces if two elements share the corresponding facet. Functions in $L_2(\mathcal{E}_h)$ are single-valued on each facet. However, they can be identified with functions in $L_2(\partial\mathcal{T}_h)$ by duplicating

the values on each element interface. Based on this identification, we have $L_2(\mathcal{E}_h) \subset L_2(\partial\mathcal{T}_h)$. Correspondingly, we define for $\lambda, \xi \in L_2(\mathcal{T}_h)$ the inner product

$$\langle \lambda, \xi \rangle_{\partial\mathcal{T}_h} = \sum_{T \in \mathcal{T}_h} \langle \lambda, \xi \rangle_{\partial T}, \quad \text{with} \quad \langle \lambda, \xi \rangle_{\partial T} = \int_{\partial T} \lambda \xi \, \mathrm{d}s.$$

Next, we define the finite element spaces associated to the function spaces

$$\Sigma = (H^1(\mathcal{T}_h))^d, \quad V = H^1(\mathcal{T}_h), \quad M = \{\xi \in L^2(\mathcal{E}_H) : \xi = 0 \text{ on } \partial\Omega\}.$$

To this end, let $k \in \mathbb{N}_0$ be a fixed polynomial degree. We denote by $\mathbb{P}_k(\hat{T})$ and $\mathbb{P}_k(\hat{E})$ the space of polynomials of degree at most k on the reference simplex $\hat{T}$ and the reference facet $\hat{E}$, respectively. Moreover, let $\mathbb{RT}_{k,d}$ be the Raviart-Thomas space of order k in dimension d on the reference simplex $\hat{T}$, which is a vector-valued space of polynomials. We will give more details on the Raviart-Thomas space in Section 6.3.2. Based on these polynomial spaces and the mappings from the reference elements, we define the function spaces by

$$\begin{aligned}
\Sigma_h &= \left\{ \tau_h \in (L_2(\Omega))^d : \tau_h|_T = \tfrac{1}{\det(DF_T)} DF_T(\hat{\tau} \circ F_T^{-1}), \hat{\tau} \in \mathbb{RT}_{k,d}(\hat{T}), T \in \mathcal{T}_h \right\}, \\
V_h &= \left\{ v_h \in L_2(\Omega) : v_h|_T = \hat{v} \circ F_T^{-1}, \hat{v} \in \mathbb{P}_k(\hat{T}), T \in \mathcal{T}_h \right\}, \\
M_h &= \left\{ \xi_h \in L_2(\mathcal{E}_h) : \xi_h = 0 \text{ on } \partial\Omega, \xi_h|_E = \hat{\xi} \circ F_E^{-1}, \hat{\xi} \in \mathbb{P}_k(\hat{E}), E \in \mathcal{E}_h \right\},
\end{aligned}$$

where DF_T denotes the Jacobian matrix of the mapping F_T. Additionally, we set

$$\begin{aligned}
W &= \Sigma \times V \times M, \\
W_h &= \Sigma_h \times V_h \times M_h.
\end{aligned}$$

In the case of an affine mapping F_T and F_E, the spaces Σ_h, V_h, and M_h consist of piecewise polynomials of degree k on each element $T \in \mathcal{T}_h$ and facet $E \in \mathcal{E}_h$, respectively. However, even if the mappings are not affine but differentiable, we have $W_h \subset W$.

6.2.2 Variational Formulation

In this section, we state the HDG variational formulation from [ES10] for the stationary convection-diffusion equation (6.18) using the mixed formulation (6.19). The basic idea is to combine formulations for elliptic and hyperbolic problems that recover the original behaviour in the respective limits. The two subproblems have to be considered in such a way that they can be coupled later.

6.2.2.1 Pure Diffusion Subproblem

We start with the pure diffusion subproblem [ES10, Section 2.2], i.e., we set $\beta = 0$ in (6.19) and assume that $\varepsilon > 0$. The corresponding mixed variational formulation is given by: find $(\sigma, u) \in H(\mathrm{div}, \Omega) \times L_2(\Omega)$ such that

$$\begin{aligned}
\langle \varepsilon^{-1}\sigma, \tau \rangle_{\mathcal{T}_h} - \langle u, \mathrm{div}(\tau) \rangle_{\mathcal{T}_h} &= 0, && \text{for all } \tau \in H(\mathrm{div}, \Omega), \\
\langle \mathrm{div}(\sigma), v \rangle_{\mathcal{T}_h} &= \langle f, v \rangle_{\mathcal{T}_h}, && \text{for all } v \in L_2(\Omega).
\end{aligned}$$

Using a conforming discretization of this formulation leads to a saddle point problem. To avoid this, we use hybridization, where we refer to [AB85; BF91] for details.

The idea of hybridization is to allow the discrete fluxes σ_h, τ_h on $\mathcal{T}_h$ to be discontinuous piecewise polynomial functions instead of being functions in $H(\mathrm{div}, \Omega)$. To ensure the continuity of the normal fluxes, we introduce additional unknowns $\lambda \in M$ on the element interfaces and require σ_h to fullfil

$$\langle \sigma_h \cdot n, \lambda \rangle_{\partial \mathcal{T}_h} = 0 \quad \text{for all } \lambda \in M.$$

By the definition of M, this condition directly implies the continuity of the normal fluxes across element interfaces. In the next step, this condition is relaxed in the variational formulation by replacing M with the discrete space M_h. However, in the limit as $h \to 0$, the continuity is recovered. The hybridized discretized formulation reads as follows: find $(\sigma_h, u_h, \lambda_h) \in W_h$ such that

$$\begin{aligned} \langle \varepsilon^{-1} \sigma_h, \tau_h \rangle_{\mathcal{T}_h} - \langle u_h, \mathrm{div} \tau_h \rangle_{\mathcal{T}_h} + \langle \lambda_h, \tau_h \cdot n \rangle_{\partial \mathcal{T}_h} &= 0 && \text{for all } \tau_h \in \Sigma_h, \\ \langle \mathrm{div}(\sigma_h), v_h \rangle_{\mathcal{T}_h} &= \langle f, v_h \rangle_{\mathcal{T}_h} && \text{for all } v_h \in V_h, \\ \langle \sigma_h \cdot n, \xi_h \rangle_{\partial \mathcal{T}_h} &= 0 && \text{for all } \xi_h \in M_h. \end{aligned}$$

Using integration by parts, we obtain the hybrid mixed finite element method: find $(\sigma_h, u_h, \lambda_h) \in W_h$ such that

$$B_D((\sigma_h, u_h, \lambda_h), (\tau_h, v_h, \xi_h)) = l_D((\tau_h, v_h, \xi_h)) \quad \text{for all } (\tau_h, v_h, \xi_h) \in W_h,$$

where the bilinear form $B_D \colon W \times W \to \mathbb{R}$ is defined by

$$\begin{aligned} B_D((\sigma_h, u_h, \lambda_h), (\tau_h, v_h, \xi_h)) = \langle \varepsilon^{-1} \sigma_h, \tau_h \rangle_{\mathcal{T}_h} + \langle \nabla u_h, \tau_h \rangle_{\mathcal{T}_h} + \langle \lambda_h - u_h, \tau_h \cdot n \rangle_{\partial \mathcal{T}_h} \\ + \langle \sigma_h, \nabla v_h \rangle_{\mathcal{T}_h} + \langle \sigma_h \cdot n, \xi_h - v_h \rangle_{\partial \mathcal{T}_h}, \end{aligned} \tag{6.21}$$

and the linear functional $l_D \colon W \to \mathbb{R}$ by

$$l_D((\tau_h, v_h, \xi_h)) = -\langle f, v_h \rangle_{\mathcal{T}_h}. \tag{6.22}$$

6.2.2.2 Pure Convection Subproblem

Next, we consider the pure convection subproblem [ES10, Section 2.3], i.e., we set $\varepsilon = 0$ in (6.18), which together with the assumption $g_D = 0$ in Section 6.2.1 leads to the hyperbolic problem

$$\begin{aligned} \mathrm{div}(\beta u) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega^{\mathrm{in}}. \end{aligned}$$

For the hyperbolic problem, we use a discontinuous Galerkin formulation with upwind stabilization, which was first introduced and analyzed in [RH73; LR74].

For each element $E \in \mathcal{E}_h$, we denote by $T^+ = T$ the upwind element, i.e., the element attached to E such that $\beta \cdot n_T \geq 0$. If $\beta \cdot n_T > 0$, the choice is unique based on the continuity assumption of the normal components of β across element interfaces. If $\beta \cdot n_T = 0$, we can make an arbitrary choice. Correspondingly, for a function $u \in L_2(\Omega)$, we denote by $u^+ = u|_{\partial T^+}$ the upwind value and set $u^+|_{\partial\Omega^{\mathrm{in}}} = 0$ to incorporate the boundary conditions.

Then the discrete discontinuous Galerkin formulation with upwind stabilization reads as follows: find $u_h \in H^1(\mathcal{T}_h)$ such that

$$\langle \mathrm{div}(\beta u), v \rangle_{\mathcal{T}_h} + \langle \beta \cdot n(u^+ - u), v \rangle_{\partial \mathcal{T}^{\mathrm{in}}} = \langle f, v \rangle_{\mathcal{T}_h}, \quad \text{for all } v \in H^1(\mathcal{T}_h).$$

Again, using integration by parts as well as $u = u^+$ on $\partial\mathcal{T}^{\text{out}}$, we obtain that

$$\langle u, \beta \cdot \nabla v\rangle_{\mathcal{T}_h} - \langle \beta \cdot n u^+, v\rangle_{\partial\mathcal{T}^{\text{in}}} - \langle \beta \cdot n u, v\rangle_{\partial\mathcal{T}^{\text{out}}} = -\langle f, v\rangle_{\mathcal{T}_h}, \quad \text{for all } v \in H^1(\mathcal{T}_h).$$

Next, we introduce an additional variable on the element interfaces to make the variational formulation suitable with the hybrid formulation in Section 6.2.2.1. To this end, let $\lambda = u^+$, and we define

$$\{\lambda/u\} = \begin{cases} \lambda & E \subset \partial T^{\text{in}}, \\ u & E \subset \partial T^{\text{out}}, \end{cases}, \quad T \in \mathcal{T}_h.$$

Using the finite element spaces, we arrive at the HDG formulation for the pure convection problem: find $(u_h, \lambda_h) \in V_h \times M_h$ such that

$$B_C((u_h, \lambda_h), (v_h, \xi_h)) = l_C((v_h, \xi_h)), \quad \text{for all } (v_h, \xi_h) \in V_h \times M_h,$$

where the bilinear form $B_C\colon (V \times M) \times (V \times M) \to \mathbb{R}$ is defined by

$$B_C((u_h, \lambda_h), (v_h, \xi_h)) = \langle u_h, \beta \cdot \nabla v_h\rangle_{\mathcal{T}_h} + \langle \beta \cdot n\{\lambda_h/u_h\}, \xi_h - v_h\rangle_{\partial\mathcal{T}_h} \tag{6.23}$$

and the linear functional $l_C\colon (V \times M) \to \mathbb{R}$ by

$$l_C((v_h, \xi_h)) = -\langle f, v_h\rangle_{\mathcal{T}_h}. \tag{6.24}$$

This formulation is algebraically equivalent to the standard discontinuous Galerkin formulation with upwind stabilization. Hence, the introduction of the additional variable, which is necessary for the coupling with the diffusion subproblem, does not alter the fundamental properties of the formulation. For details, we refer to [ES10].

Remark 6.4. The upwind elements T^+ depend on the convection field β. Hence, for parameter-dependent problems with a parameter-dependent convection field, the upwind elements can change with varying parameters. If this is the case, the variational formulation will no longer be affine in the parameters. We discuss this in greater detail in Section 6.2.4.2.

6.2.2.3 Combined Formulation

Finally, we consider the full convection-diffusion problem (6.19) and combine the two subproblems from Sections 6.2.2.1 and 6.2.2.2. The two variational formulations for the subproblems are designed so that they can be coupled directly. The combined HDG formulation reads as follows: find $(\sigma_h, u_h, \lambda_h) \in W_h$ such that

$$B((\sigma_h, u_h, \lambda_h), (\tau_h, v_h, \xi_h)) = f((\tau_h, v_h, \xi_h)), \quad \text{for all } (\tau_h, v_h, \xi_h) \in W_h, \tag{6.25}$$

where the bilinear form $B\colon W \times W \to \mathbb{R}$ is defined by

$$\begin{aligned} &B((\sigma_h, u_h, \lambda_h), (\tau_h, v_h, \xi_h)) \\ &\qquad = B_D((\sigma_h, u_h, \lambda_h), (\tau_h, v_h, \xi_h)) + B_C((u_h, \lambda_h), (v_h, \xi_h)), \end{aligned} \tag{6.26}$$

with B_D from (6.21) and B_C from (6.23), and the linear functional $l\colon W \to \mathbb{R}$ is defined by

$$l((\tau_h, v_h, \xi_h)) = -(f, v_h)_{\mathcal{T}_h}.$$

The right-hand side is divided into two parts for the diffusion and convection subproblem, respectively. Combining both parts and using the definition of l_D from (6.22) and l_C from (6.24) leads to the above definition of l.

The variational formulation is consistent in the following sense.

Proposition 6.5 ([ES10, Theorem 2.5])**.** *Let u denote the solution of (6.19) with sufficient regularity and let $\sigma = -\varepsilon\nabla u$. Moreover, let $\lambda = u|_{\mathcal{E}_h}$. Then $(\sigma, u, \lambda) \in W$ satisfies*

$$B((\sigma, u, \lambda), (\tau_h, v_h, \xi_h)) = l((\tau_h, v_h, \xi_h)), \quad \text{for all } (\tau_h, v_h, \xi_h) \in W_h.$$

For the proof and more details on the sufficient regularity of u, we refer to [ES10].

Moreover, the method is locally and globally conservative, meaning it prevents nonphysical behaviour, such as an increase in mass or energy. For details, we refer to [ES10, Proposition 2.6].

6.2.3 Mesh-Dependent Norm and Error Estimate

In the next step, we introduce a mesh-dependent norm. Using this norm, we can show boundedness and an inf-sup condition of the bilinear form. Then we state an a priori error estimate for the HDG method. In this work, we use the mesh-dependent energy norms introduced in [ES10].

Definition 6.6. For $(\sigma, u, \lambda), (\tau, v, \xi) \in W$, we define the mesh-dependent inner products

$$\begin{aligned}
\left\langle (\sigma, u, \lambda), (\tau, v, \xi) \right\rangle_D &= \left\langle \varepsilon^{-1}\sigma, \tau \right\rangle_{\mathcal{T}_h} + \left\langle \varepsilon\nabla u, \nabla v \right\rangle_{\mathcal{T}_h} + \left\langle \varepsilon h^{-1}(\lambda - u), \xi - v \right\rangle_{\partial\mathcal{T}_h}, \\
\left\langle (u, \lambda), (v, \xi) \right\rangle_C &= \left\langle |\beta|^{-1} h\beta \cdot \nabla u, \beta \cdot \nabla v \right\rangle_{\mathcal{T}_h} + \left\langle |\beta \cdot n|(\lambda - u), \xi - v \right\rangle_{\partial\mathcal{T}_h}, \\
\left\langle (\sigma, u, \lambda), (\tau, v, \xi) \right\rangle_W &= \left\langle (\sigma, u, \lambda), (\tau, v, \xi) \right\rangle_D + \left\langle (u, \lambda), (v, \xi) \right\rangle_C.
\end{aligned} \tag{6.27}$$

Here, $|\beta|$ and $|\beta \cdot n|$ denote appropriate bounds for the convection field β on single elements and for $\beta \cdot n$ on single facets, respectively. The corresponding mesh-dependent norms are defined by

$$\begin{aligned}
\|(\sigma, u, \lambda)\|_D &= \left(\left\langle (\sigma, u, \lambda), (\sigma, u, \lambda) \right\rangle_D \right)^{\frac{1}{2}}, \\
\|(u, \lambda)\|_C &= \left(\left\langle (u, \lambda), (u, \lambda) \right\rangle_C \right)^{\frac{1}{2}}, \\
\|(\sigma, u, \lambda)\|_W &= \left(\left\langle (\sigma, u, \lambda), (\sigma, u, \lambda) \right\rangle_W \right)^{\frac{1}{2}}.
\end{aligned}$$

The boundedness of the bilinear forms B_D, B_C, and B cannot be shown directly with respect to the above norms. Hence, we introduce the following stronger norms.

Definition 6.7. For $(\sigma, u, \lambda) \in W$ let

$$\begin{aligned}
\|(\sigma, u, \lambda)\|_{D,*} &= \left(\|(\sigma, u, \lambda)\|_D^2 + \left\langle \varepsilon^{-1} h\tau \cdot n, \tau \cdot n \right\rangle_{\partial\mathcal{T}_h} \right)^{\frac{1}{2}}, \\
\|(u, \lambda)\|_{C,*} &= \left(\left\langle h^{-1}|\beta| u, u \right\rangle_{\mathcal{T}_h} + \left\langle |\beta \cdot n| \{\lambda/u\}, \{\lambda/u\} \right\rangle_{\partial\mathcal{T}_h} \right)^{\frac{1}{2}}, \\
\|(\sigma, u, \lambda)\|_{W,*} &= \left(\|(\sigma, u, \lambda)\|_{D,*}^2 + \|(u, \lambda)\|_{C,*}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Using the norms from Theorems 6.6 and 6.7, we obtain the following boundedness results.

Proposition 6.8 ([ES10, Corollary 3.8]). *We have*

$$B((\sigma,u,\lambda),(\tau_h,v_h,\xi_h)) \le C\|(\sigma,u,\lambda)\|_{W,*}\|(\tau_h,v_h,\xi_h)\|_W$$

for all $(\sigma,u,\lambda) \in W$ *and* $(\tau_h,v_h,\xi_h) \in W_h$ *with a constant* $C > 0$ *independent of the mesh size* h.

The idea of the proof is to show the boundedness of the bilinear forms B_D and B_C separately, i.e.,

$$\begin{aligned} B_D((\sigma,u,\lambda),(\tau_h,v_h,\xi_h)) &\le C_D\|(\sigma,u,\lambda)\|_{D,*}\|(\tau_h,v_h,\xi_h)\|_D, \\ B_C((u,\lambda),(v_h,\xi_h)) &\le C_C\|(u,\lambda)\|_{C,*}\|(v_h,\xi_h)\|_C. \end{aligned} \tag{6.28}$$

Using the definition of B in (6.26) and the definitions of the norms in Theorems 6.6 and 6.7 yields the desired result. For details on the proof of the bounds in (6.28), we refer to [ES10, Proposition 3.4 and Proposition 3.6].

Proposition 6.9 ([ES10, Proposition 3.7]). *We have*

$$\sup_{0\neq(\tau_h,v_h,\xi_h)\in W_h} \frac{B((\sigma_h,u_h,\lambda_h),(\tau_h,v_h,\xi_h))}{\|(\tau_h,v_h,\xi_h)\|_W} \ge c\|(\sigma_h,u_h,\lambda_h)\|_W$$

for all $(\sigma_h,u_h,\lambda_h) \in W_h$ *with a constant* $c > 0$ *independent of the mesh size* h.

This property is equivalent to an inf-sup condition of the bilinear form $B\colon W_h \times W_h \to \mathbb{R}$. The statement can be shown by testing with appropriate functions in W_h and using the stability properties of the individual bilinear forms B_D and B_C in [ES10, Proposition 3.2 and Proposition 3.5].

As the final point of this section, we state an a priori error estimate for the HDG method.

Theorem 6.10 ([ES10, Theorem 3.14]). *Let* $(\sigma_h,u_h,\lambda_h) \in W_h$ *be the solution of the variational formulation* (6.25). *Moreover, for some* $\delta > 0$, *let* $u \in H^1(\Omega) \cap H^{\frac{3}{2}+\delta}(\mathcal{T}_h)$ *be the solution of* (6.19) *and* $\sigma = -\varepsilon\nabla u$. *Then we have*

$$\|(\sigma-\sigma_h,u-u_h,u-\lambda_h)\|_W \le Ch^s\left|(\sqrt{\varepsilon}+h^{\frac{1}{2}}\sqrt{|\beta|})u\right|_{H^{s+1}(\mathcal{T}_h)}$$

for all $\frac{1}{2} < s \le k$, *where* $C > 0$ *is a constant independent of the mesh size* h, k *is the polynomial degree of the finite element spaces, and* $|\cdot|_{H^{s+1}(\mathcal{T}_h)}$ *denotes the broken Sobolev seminorm with respect to the triangulation* $\mathcal{T}_h$.

The proof of this theorem is based on the boundedness and inf-sup condition of the bilinear form B in Theorems 6.8 and 6.9 as well as an interpolation estimates. For details, we refer to [ES10].

By (6.5) the error of the reduced basis approximation can be estimated by the sum of the error of the HDG approximation and the error of the reduced basis approximation in the HDG space. In this regard, Theorem 6.10 is crucial for the solution quality of the RB method, since it provides a bound for the error of the HDG finite element approximation.

6.2.4 Magnetic Particle Imaging Fokker-Planck Equation

In the previous sections, we have described the RB method and the HDG discretization in general settings under certain assumptions. However, not all of these assumptions are satisfied for the Fokker-Planck equation related to MPI in (5.12). In

this section, we consider the necessary modifications to apply the HDG method in the context of the Fokker-Planck equation and of the RB framework in Section 6.1.

We divide the section into two parts. First, we state the necessary modifications of the HDG formulation to make it applicable to the Fokker-Planck equation (5.15). We then modify the formulation to satisfy the RB method requirements in Section 6.1.

6.2.4.1 Fokker-Planck Equation

One assumption of the HDG method in Section 6.2 is a piecewise constant convection field. However, in the Fokker-Planck equation for MPI in (5.12), the convection field

$$\beta(\vec{n},\vec{e},t) = \begin{pmatrix} a_1(\vec{n},\vec{e}) \\ a_2(\vec{n},\vec{e},t) \\ 1 \end{pmatrix} \tag{6.29}$$

with a_1 and a_2 from (5.10) is not piecewise constant. Moreover, β is not divergence-free, i.e., $\operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2\times\mathbb{R}}(\beta) \neq 0$. However, neither assumption is necessary for formulating the variational formulation. These assumptions are used in [ES10] for the convergence and stability analysis of the method. Nevertheless, convergence of discontinuous Galerkin methods can also be shown under more general assumptions, including non-divergence-free convection fields. For details, we refer to [BDS17] and the references therein.

Due to the non-piecewise constant convection field, a face E of the partition $\mathcal{T}_h$ may not be strictly an inflow or outflow boundary. Since the inflow and outflow boundaries in (6.20) are considered in a pointwise sense, these boundaries do not align with the faces of the simplices anymore. Nevertheless, we can still use the same formulation with a pointwise interpretation of the inflow and outflow boundaries. An alternative would be to consider the inflow and outflow boundaries in an integral sense, according to the mass flux in finite volume schemes. To this end, we define

$$\partial T^{\text{out}} = \left\{ x \in E \subset \partial T : E \in \mathcal{E}_h, \int_E \beta_T \cdot n_T \, \mathrm{d}s > 0 \right\}, \quad \partial T^{\text{in}} = \partial T \setminus \partial T^{\text{out}}, \quad T \in \mathcal{T}_h.$$

Hence, a facet is considered as an outflow boundary if more mass is leaving the element than entering it.

Moreover, as stated in Section 6.2.1, we assume homogeneous Dirichlet boundary conditions in the HDG method for the domain $\Omega = \mathbb{S}\times\mathbb{S}\times[0,T]$. Based on the definition of β in (6.29) and the properties of the sphere, specifically $\partial\mathbb{S} = \varnothing$, we have $\partial\Omega^{\text{in}} = \mathbb{S}\times\mathbb{S}\times\{0\}$. This means that the inflow at the boundary of the domain is given by the initial condition of the Fokker-Planck equation. However, in the Fokker-Planck equation for MPI in (5.12), we have inhomogeneous Dirichlet boundary conditions at the inflow boundary of the domain $\partial\Omega^{\text{in}} = \mathbb{S}\times\mathbb{S}\times\{0\}$ due to the initial condition p_0. To tackle this issue, we homogenize the boundary conditions by introducing an appropriate lifting function that incorporates the inhomogeneous Dirichlet boundary conditions. We define $\tilde{p}\colon \mathbb{S}\times\mathbb{S}\times[0,T]$ as

$$\tilde{p}(\vec{n},\vec{e},t) = p(\vec{n},\vec{e},t) - p_0(\vec{n},\vec{e}),$$

where p is the solution of the Fokker-Planck equation (5.15) with initial condition p_0 and T is the final time. Then we obtain the homogenized Fokker-Planck equation

$$\operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2\times\mathbb{R}}\left(\begin{pmatrix}a_1\\a_2\\1\end{pmatrix}\tilde{p}-\begin{pmatrix}\frac{1}{2\tau_B}\nabla_{\mathbb{S}_1}\tilde{p}\\\frac{1}{(1+\alpha^2)\tau_N}\nabla_{\mathbb{S}_2}\tilde{p}\\0\end{pmatrix}\right)=f,\quad \tilde{p}(\vec{n},\vec{e},0)=0 \tag{6.30}$$

for $\tilde{p}$ with

$$\begin{aligned} f(\vec{n},\vec{e},t) &= \operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2\times\mathbb{R}}\left(\begin{pmatrix}a_1\\a_2\\1\end{pmatrix}p_0-\begin{pmatrix}\frac{1}{2\tau_B}\nabla_{\mathbb{S}_1}p_0\\\frac{1}{(1+\alpha^2)\tau_N}\nabla_{\mathbb{S}_2}p_0\\0\end{pmatrix}\right)(\vec{n},\vec{e},t)\\ &= \operatorname{div}_{\mathbb{S}_1}\big(a_1p_0-\tfrac{1}{2\tau_B}\nabla_{\mathbb{S}_1}p_0\big)(\vec{n},\vec{e})+\operatorname{div}_{\mathbb{S}_2}\big(a_2p_0-\tfrac{1}{(1+\alpha^2)\tau_N}\nabla_{\mathbb{S}_2}p_0\big)(\vec{n},\vec{e},t), \end{aligned} \tag{6.31}$$

where we used that p_0 does not dependent on the time. Instead of solving the Fokker-Planck equation (5.15), we solve the homogenized problem in (6.30) with homogeneous Dirichlet boundary conditions, i.e., homogenous initial condition. Using the expectation formula (5.11), we can directly calculate the mean magnetic moment from the solution $\tilde{p}$ of (6.30) by using the linearity of the expectation,

$$\begin{aligned} \bar{m}(t) &= \int_{\mathbb{S}\times\mathbb{S}}\vec{e}p(\vec{n},\vec{e},t)\,\mathrm{d}(\vec{n},\vec{e})\\ &= \int_{\mathbb{S}\times\mathbb{S}}\vec{e}\tilde{p}(\vec{n},\vec{e},t)\,\mathrm{d}(\vec{n},\vec{e})+\int_{\mathbb{S}\times\mathbb{S}}\vec{e}p_0(\vec{n},\vec{e})\,\mathrm{d}(\vec{n},\vec{e}). \end{aligned}$$

The initial condition is parameter-independent. Thus, the second integral on the right-hand side can be precomputed.

Moreover, we need to rewrite the system as mixed formulation similar to (6.19) in order to apply the HDG method. Hence, let

$$\varepsilon_1=\frac{1}{2\tau_B},\quad \varepsilon_2=\frac{1}{(1+\alpha^2)\tau_N}.$$

We introduce the diffusive fluxes $\sigma_1=-\varepsilon_1\nabla_{\mathbb{S}_1}\tilde{p}$ and $\sigma_2=-\varepsilon_2\nabla_{\mathbb{S}_2}\tilde{p}$. Then we can rewrite (6.30) in the mixed form

$$\begin{aligned} \begin{pmatrix}\sigma_1\\\sigma_2\end{pmatrix}+\begin{pmatrix}\varepsilon_1\nabla_{\mathbb{S}_1}\tilde{p}\\\varepsilon_2\nabla_{\mathbb{S}_2}\tilde{p}\end{pmatrix}=0,\quad \operatorname{div}_{\mathbb{S}_1\times\mathbb{S}_2\times\mathbb{R}}\left(\begin{pmatrix}\sigma_1\\\sigma_2\\0\end{pmatrix}+\beta\tilde{p}\right)&=f \qquad \text{in } \Omega,\\ \tilde{p}&=0 \qquad \text{on } \partial\Omega^{\text{in}}, \end{aligned}$$

with β defined in (6.29) and $\Omega=\mathbb{S}\times\mathbb{S}\times[0,T]$. Alternatively, we can write the system in the form

$$\begin{aligned} \sigma_1+\varepsilon_1\nabla_{\mathbb{S}_1}\tilde{p}=0,\quad \sigma_2+\varepsilon_2\nabla_{\mathbb{S}_2}\tilde{p}&=0, && \text{in } \Omega,\\ \operatorname{div}_{\mathbb{S}_1}(\sigma_1+a_1\tilde{p})+\operatorname{div}_{\mathbb{S}_2}(\sigma_2+a_2\tilde{p})+\operatorname{div}_{\mathbb{R}}(\tilde{p})&=f, && \text{in } \Omega,\\ \tilde{p}&=0, && \text{on } \partial\Omega^{\text{in}}, \end{aligned} \tag{6.32}$$

where we write $\operatorname{div}_{\mathbb{R}}(\tilde{p})=\partial_t\tilde{p}$ for notational consistency.

Next, we have to consider the product structure of the domain $\Omega=\mathbb{S}\times\mathbb{S}\times[0,T]$. To this end, we construct a mesh $\mathcal{P}_h$ that is a product of meshes of the two spheres

$\mathbb{S}$ and $[0,T]$ with mesh sizes $\mathsf{h} = (h_1, h_2, h_\mathsf{t})$. Let $\mathcal{T}_{h_1}^{(1)}$ and $\mathcal{T}_{h_2}^{(2)}$ be the triangulations of the two spheres $\mathbb{S}$ with mesh sizes h_1 and h_2, respectively. We will discuss the concrete implementation of triangulation for spheres later in Section 6.3.1. Moreover, let $\mathcal{I}_{h_\mathsf{t}}$ be a partition of $[0,T]$ into intervals with mesh size h_t. Then we define the product mesh

$$\mathcal{P}_\mathsf{h} = \mathcal{T}_{h_1}^{(1)} \times \mathcal{T}_{h_2}^{(2)} \times \mathcal{I}_{h_\mathsf{t}}.$$

Based on the product structure, we have

$$\partial\mathcal{P}_\mathsf{h} = \partial\mathcal{P}_\mathsf{h}^{(1)} \cup \partial\mathcal{P}_\mathsf{h}^{(2)} \cup \partial\mathcal{P}_\mathsf{h}^{(\mathsf{t})},$$

with

$$\begin{aligned}
\partial\mathcal{P}_\mathsf{h}^{(1)} &= \partial\mathcal{T}_{h_1}^{(1)} \times \mathcal{T}_{h_2}^{(2)} \times \mathcal{I}_{h_\mathsf{t}},\\
\partial\mathcal{P}_\mathsf{h}^{(2)} &= \mathcal{T}_{h_1}^{(1)} \times \partial\mathcal{T}_{h_2}^{(2)} \times \mathcal{I}_{h_\mathsf{t}},\\
\partial\mathcal{P}_\mathsf{h}^{(\mathsf{t})} &= \mathcal{T}_{h_1}^{(1)} \times \mathcal{T}_{h_2}^{(2)} \times \partial\mathcal{I}_{h_\mathsf{t}},
\end{aligned}$$

where we again identify each facet with the corresponding element $P \in \mathcal{P}_\mathsf{h}$. Correspondingly, let $\mathcal{E}_\mathsf{h}$ denote the set of facets of the product mesh $\mathcal{P}_\mathsf{h}$, which is given by

$$\mathcal{E}_\mathsf{h} = (\mathcal{E}_{h_1}^{(1)} \times \mathcal{T}_{h_2}^{(2)} \times \mathcal{I}_{h_\mathsf{t}}) \cup (\mathcal{T}_{h_1}^{(1)} \times \mathcal{E}_{h_2}^{(2)} \times \mathcal{I}_{h_\mathsf{t}}) \cup (\mathcal{T}_{h_1}^{(1)} \times \mathcal{T}_{h_2}^{(2)} \times \mathcal{V}_{h_\mathsf{t}}),$$

where $\mathcal{E}_{h_1}^{(1)}$ and $\mathcal{E}_{h_2}^{(2)}$ denote the sets of facets of the triangulation $\mathcal{T}_{h_1}^{(1)}$ and $\mathcal{T}_{h_2}^{(2)}$, respectively, and $\mathcal{V}_{h_\mathsf{t}}$ denotes the set of endpoints of the intervals in $\mathcal{I}_{h_\mathsf{t}}$. Moreover, based on the product structure, we consider the function spaces

$$\begin{aligned}
\Sigma &= \Sigma_1 \times \Sigma_2\\
&= \left(\left(H^1\left(\mathcal{T}_{h_1}^{(1)}\right)\right)^3 \otimes L_2(\mathbb{S}_2) \otimes L_2([0,T])\right) \times \left(L_2(\mathbb{S}_1) \otimes \left(H^1\left(\mathcal{T}_{h_2}^{(2)}\right)\right)^3 \otimes L_2([0,T])\right),\\
V &= H^1(\mathcal{P}_\mathsf{h}),\\
M &= \left\{\xi \in L_2(\mathcal{E}_\mathsf{h}) : \xi = 0 \text{ on } \partial\Omega^{\text{in}}\right\}.
\end{aligned}$$

Then we combine these function spaces by

$$\begin{aligned}
W_i &= \Sigma_i \times V \times M, \quad i = 1,2,\\
W &= \Sigma \times V \times M.
\end{aligned}$$

We introduce the spaces Σ_i and W_i in order to define the bilinear forms for the diffusion part of the two spheres separately.

Next, we define the corresponding finite element spaces. Here, we allow for different polynomial degrees in the different subdomains. To this end let $\mathsf{k} = (k_1, k_2, k_\mathsf{t})$ be the vector of polynomial degrees. For the construction of the finite element spaces, we use the tensor-product structure of the product mesh $\mathcal{P}_\mathsf{h}$. Hence, for $i = 1,2$ we define

$$\begin{aligned}
\Sigma_{h_i}^{(i)} &= \left\{\tau_h \in (L_2(\mathbb{S}))^3 : \tau_h|_T = \tfrac{1}{\det(DF_T)} DF_T(\hat{\tau} \circ F_T^{-1}), \hat{\tau} \in \mathbb{RT}_{k_i,2}(\hat{T}), T \in \mathcal{T}_{h_i}^{(i)}\right\},\\
V_{h_i}^{(i)} &= \left\{v_h \in L_2(\mathbb{S}) : v_h|_T = \hat{v} \circ F_T^{-1}, \hat{v} \in \mathbb{P}_{k_i}(\hat{T}), T \in \mathcal{T}_{h_i}^{(i)}\right\},\\
M_{h_i}^{(i)} &= \left\{\xi_h \in L_2(\mathcal{E}_{h_i}^{(i)}) : \xi_h|_E = \hat{\xi} \circ F_E^{-1}, \hat{\xi} \in \mathbb{P}_{k_i}(\hat{E}), E \in \mathcal{E}_{h_i}^{(i)}\right\},
\end{aligned} \tag{6.33}$$

with the Raviart-Thomas space $\mathbb{RT}_{k_i,2}(\hat{T})$ and the space of polynomials $\mathbb{P}_{k_i}(\hat{T})$ of

order k_i on the reference triangle $\hat{T}$ defined in Section 6.2.1. Here, F_T is the mapping from the reference triangle $\hat{T}$ in $\mathbb{R}^2$ to a spherical triangle in $\mathbb{S}$. The interpretation of $\det(DF_T)$ for $DF_T \in \mathbb{R}^{3\times 2}$ will be discussed in Section 6.3.2.2. Moreover, let $\hat{I} = (0,1)$ be the reference interval, F_I the affine mapping from $\hat{I}$ to an interval $I \in \mathcal{I}_{h_t}$ and $\mathbb{P}_{k_t}(\hat{I})$ be the space of polynomials of degree k_t on $\hat{I}$. Then we define

$$V_{h_t}^{(t)} = \left\{ v_h \in L_2([0,T]) : v_h|_I = \hat{v} \circ F_I^{-1}, \hat{v} \in \mathbb{P}_{k_t}(\hat{I}), I \in \mathcal{I}_{h_t} \right\}.$$

Based on these spaces, we define the finite element spaces for the product mesh $\mathcal{P}_h$ by

$$\begin{aligned} \Sigma_h &= \Sigma_{h,1} \times \Sigma_{h,2} = \left(\Sigma_{h_1}^{(1)} \otimes V_{h_2}^{(2)} \otimes V_{h_t}^{(t)} \right) \times \left(V_{h_1}^{(1)} \otimes \Sigma_{h_2}^{(2)} \otimes V_{h_t}^{(t)} \right), \\ V_h &= V_{h_1}^{(1)} \otimes V_{h_2}^{(2)} \otimes V_{h_t}^{(t)}, \end{aligned} \tag{6.34}$$

and

$$\begin{aligned} M_h &= \begin{cases} \xi \in L_2(\mathcal{E}_h) : \xi|_{T_1 \times T_2 \times I} \in M_{h_1}^{(1)} \otimes V_{h_2}^{(2)} \otimes V_{h_t}^{(t)}, & E_1 \in \mathcal{E}_{h_1}^{(1)}, T_2 \in \mathcal{T}_{h_2}^{(2)}, I \in \mathcal{I}_{h_t}, \\ \xi|_{T_1 \times E_2 \times I} \in V_{h_1}^{(1)} \otimes M_{h_2}^{(2)} \otimes V_{h_t}^{(t)}, & T_1 \in \mathcal{T}_{h_1}^{(1)}, E_2 \in \mathcal{E}_{h_2}^{(2)}, I \in \mathcal{I}_{h_t}, \\ \xi|_{T_1 \times T_2 \times \{t\}} \in V_{h_1}^{(1)} \otimes V_{h_2}^{(2)}, & T_1 \in \mathcal{T}_{h_1}^{(1)}, T_2 \in \mathcal{T}_{h_2}^{(2)}, 0 \neq t \in \mathcal{V}_{h_t}, \end{cases} \\ &\cong (M_{h_1}^{(1)} \otimes V_{h_2}^{(2)} \otimes V_{h_t}^{(t)}) \oplus (V_{h_1}^{(1)} \otimes M_{h_2}^{(2)} \otimes V_{h_t}^{(t)}) \oplus (V_{h_1}^{(1)} \otimes V_{h_2}^{(2)} \otimes \mathbb{R}^{\mathcal{I}_{h_t}}). \end{aligned}$$

Correspondingly, we set

$$\begin{aligned} W_{h,i} &= \Sigma_{h,i} \times V_h \times M_h, \quad i = 1,2, \\ W_h &= \Sigma_h \times V_h \times M_h. \end{aligned} \tag{6.35}$$

In the following, we define the corresponding HDG variational formulation based on these spaces. For the surface divergence of the two spheres in (6.32), we use the combined formulation, and for the time derivative, we use the pure convection formulation from Section 6.2.2.2. Hence, we define the bilinear form $B_D^{(i)} : W_i \times W_i \to \mathbb{R}$ for $i = 1,2$ by

$$\begin{aligned} &B_D^{(i)}((\sigma,u,\lambda),(\tau,v,\xi)) \\ &\quad = \langle \varepsilon_i^{-1}\sigma, \tau \rangle_{\mathcal{P}_h} + \langle \nabla_{\mathbb{S}_i} u, \tau \rangle_{\mathcal{P}_h} + \langle \lambda - u, \tau \cdot n \rangle_{\partial \mathcal{P}_h^{(i)}} + \langle \sigma, \nabla_{\mathbb{S}_i} v \rangle_{\mathcal{P}_h} + \langle \sigma \cdot n, \xi - v \rangle_{\partial \mathcal{P}_h^{(i)}}. \end{aligned}$$

Then we define the bilinear form $B_D : W \times W \to \mathbb{R}$ by

$$\begin{aligned} &B_D((\sigma,u,\lambda),(\tau,v,\xi)) \\ &\quad = B_D^{(1)}\left((\sigma^{(1)},u,\lambda), (\tau^{(1)},v,\xi) \right) + B_D^{(2)}\left((\sigma^{(2)},u,\lambda), (\tau^{(2)},v,\xi) \right), \end{aligned} \tag{6.36}$$

where

$$\sigma = \begin{pmatrix} \sigma^{(1)} \\ \sigma^{(2)} \end{pmatrix}, \quad \tau = \begin{pmatrix} \tau^{(1)} \\ \tau^{(2)} \end{pmatrix}.$$

In the next step, we state the convection parts of the variational formulation. Considering the convection part of the variational formulation in (6.23) separately for the two spheres and the time derivative and then combining them, we arrive at the

bilinear form $B_C\colon (V \times M) \times (V \times M) \to \mathbb{R}$ defined by

$$\begin{aligned} B_C((u,\lambda),(v,\xi)) &= \langle u, \beta \cdot \nabla_{S_1 \times S_2 \times \mathbb{R}} v\rangle_{\mathcal{P}_h} + \langle \beta \cdot n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h} \\ &= \langle u, a_1 \cdot \nabla_{S_1} v\rangle_{\mathcal{P}_h} + \langle a_1 \cdot n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h^{(1)}} \\ &\quad + \langle u, a_2 \cdot \nabla_{S_2} v\rangle_{\mathcal{P}_h} + \langle a_2 \cdot n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h^{(2)}} \\ &\quad + \langle u, \partial_t v\rangle_{\mathcal{P}_h} + \langle n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h^{(t)}}. \end{aligned} \tag{6.37}$$

The total bilinear form $B\colon W \times W \to \mathbb{R}$ is then given by (6.26). The linear functional $l\colon W \to \mathbb{R}$ is defined as in Section 6.2.2.3 by

$$l((\tau, v, \xi)) = -\langle f, v\rangle_{\mathcal{P}_h}. \tag{6.38}$$

6.2.4.2 Reduced Basis

We now modify the HDG formulation from Section 6.2.4 for compatibility with the RB method as in Section 6.1. In particular, we need to define an error estimator and we need to ensure that the bilinear form and the linear functional are affine with respect to the parameters.

Remark 6.11. The parameters in the diffusion coefficients $\varepsilon_1, \varepsilon_2$ and in the convection fields a_1 and a_2 enter as factors of the whole expression or as factors of specific terms. The only exception are the frequencies ν_x, ν_y and ν_z in the drive field $\vec{H}_D$ in (5.2). However, we only have a discrete set of frequencies that are used in the drive field, which are determined by the hardware of the MPI scanner. Since the set of different frequencies is typically small, we can treat the frequencies as fixed parameters and consider separate RB for each combination of frequencies (ν_x, ν_y, ν_z).

We start by decomposing the diffusion and convection terms into affine functions. We can subsequently use these decompositions to obtain the affine decompositions of the bilinear form and the linear functional. Clearly, we have the affine decompositions of the diffusion terms

$$\varepsilon_1 = \frac{1}{2\tau_B} = \theta_{\varepsilon_1}(\mu), \quad \varepsilon_2 = \frac{1}{(1+\alpha^2)\tau_N} = \theta_{\varepsilon_2}(\mu)$$

given by (5.13) with the definition of τ_B and τ_N in (5.6) and (5.7), respectively. Moreover, for a_1 defined in (5.10), we have

$$a_1(\vec{n}, \vec{e}; \mu) = \theta_{a_1}(\mu)\tilde{a}_1(\vec{n}, \vec{e}) \tag{6.39}$$

with

$$\theta_{a_1}(\mu) = \frac{KV_c}{3\eta V_H}, \quad \tilde{a}_1(\vec{n}, \vec{e}) = (\vec{n} \cdot \vec{e})[(\vec{n} \times \vec{e}) \times \vec{n}].$$

The affine decomposition of the convection field a_2 consists of more terms based on the applied magnetic field $\vec{H}$. Using the definitions of the selection and the drive field in (5.2) and (5.3), respectively, we have

$$\vec{H}_S(x;\mu) = Gx_1\vec{H}_S^{(1)} + Gx_2\vec{H}_S^{(2)} + Gx_3\vec{H}_S^{(3)} = -G\frac{x_1}{2}e_1 + Gx_2e_2 - G\frac{x_3}{2}e_3$$

with unit vectors e_i in $\mathbb{R}^3$ as well as

$$\begin{aligned}\vec{H}_D(t;\mu) &= A_1\vec{H}_D^{(1)}(t) + A_2\vec{H}_D^{(2)}(t) + A_3\vec{H}_D^{(3)}(t)\\ &= A_1\begin{pmatrix}\sin(2\pi\nu_x t)\\0\\0\end{pmatrix} + A_2\begin{pmatrix}0\\\sin(2\pi\nu_y t)\\0\end{pmatrix} + A_3\begin{pmatrix}0\\0\\\sin(2\pi\nu_z t)\end{pmatrix}.\end{aligned}$$

Hence, using (5.1) and $\vec{H}(x,t) = \vec{H}_x(t)$, we obtain

$$\vec{H}_x(t) = \sum_{i=1}^{3}\theta_D^{(i)}(\mu)\vec{H}_D^{(i)}(t) + \sum_{i=1}^{3}\theta_S^{(i)}(\mu)\vec{H}_S^{(i)},$$

where

$$\theta_D^{(i)}(\mu) = A_i, \quad \theta_S^{(i)}(\mu) = Gx_i, \quad i = 1,2,3.$$

Combining this with the definition of a_2 in (5.10) yields the affine decomposition

$$a_2(\vec{n},\vec{e},t;\mu) = \sum_{i=1}^{7}\theta_{a_2}^{(i)}(\mu)\tilde{a}_2^{(i)}(\vec{n},\vec{e},t) \tag{6.40}$$

with

$$\begin{aligned}\theta_{a_2}^{(i)}(\mu) &= \frac{1}{2\tau_N}\frac{M_S V_C}{\alpha k_B T}\theta_D^{(i)}(\mu), \qquad i = 1,2,3,\\ \theta_{a_2}^{(i+3)}(\mu) &= \frac{1}{2\tau_N}\frac{M_S V_C}{\alpha k_B T}\theta_S^{(i)}(\mu), \qquad i = 1,2,3,\\ \theta_{a_2}^{(7)}(\mu) &= \frac{KV_C}{\tau_N\alpha k_B T},\end{aligned}$$

and

$$\begin{aligned}\tilde{a}_2^{(i)}(\vec{n},\vec{e},t) &= \vec{e}\times\vec{H}_D^{(i)}(t) + \left(\vec{e}\times\vec{H}_D^{(i)}(t)\right)\times\vec{e}, \qquad i = 1,2,3,\\ \tilde{a}_2^{(i+3)}(\vec{n},\vec{e},t) &= \vec{e}\times\vec{H}_S^{(i)} + \left(\vec{e}\times\vec{H}_S^{(i)}\right)\times\vec{e}, \qquad i = 1,2,3,\\ \tilde{a}_2^{(7)}(\vec{n},\vec{e},t) &= (\vec{n}\cdot\vec{e})(\vec{e}\times\vec{n} + (\vec{e}\times\vec{n})\times\vec{e}).\end{aligned}$$

Using the affine decompositions of the diffusion and convection terms combined with (6.31) and (6.38), we obtain

$$\begin{aligned}l((\tau,v,\xi);\mu) &= -\langle f(\mu),v\rangle_{\mathcal{P}_h}\\ &= \sum_{i=1}^{2}\langle\varepsilon_i(\mu)\Delta_{S_i}p_0,v\rangle_{\mathcal{P}_h} - \langle\mathrm{div}_{S_i}(a_i(\mu)p_0),v\rangle_{\mathcal{P}_h}\\ &= \sum_{q=1}^{10}\theta_q^{(l)}(\mu)l_q((\tau,v,\xi))\end{aligned}$$

with suitable linear functionals $l_q\colon W\to\mathbb{R}$ and coefficient functions $\theta_q^{(l)}\colon\mathcal{P}\to\mathbb{R}$ for $q = 1,\dots,10$. We arrive at 10 terms, since we have one term for each diffusion coefficient and one term for a_1, as well as 7 terms for a_2.

Next, we consider the affine decomposition of the bilinear form. As already mentioned in Theorem 6.4, the upwind elements can be parameter-dependent based on the convection field. However, based on the definition of β in (6.29) and the affine decomposition of a_1 in (6.39), the upwind elements of $\partial\mathcal{P}_h^{(1)}$ and $\partial\mathcal{P}_h^{(t)}$ are parameter-independent since $\theta_{a_1}(\mu) > 0$ for all $\mu\in\mathcal{P}$.

Nevertheless, based on the affine decomposition of a_2 in (6.40), the upwind elements of $\partial\mathcal{P}_h^{(2)}$ are generally parameter-dependent. There are different options to tackle this problem. One option is the empirical interpolation method, first introduced and used in [Bar+04; Gre+07]. The idea of the empirical interpolation method is to approximate the non-affine bilinear form by an affine decomposition. However, this method can lead to having many terms in the affine decomposition in order to get a good approximation. For more details on the empirical interpolation method, we refer to [HRS16, Section 5.2]. In this work, we follow another approach and replace the upwind flux on $\partial\mathcal{P}_h^{(2)}$ with a Lax-Friedrichs flux, which has also been used in the context of RB methods in [HO08] and in the context of HDG in [SWM12]. Instead of taking the upwind value, the Lax-Friedrichs flux averages the values from neighboring elements and adds a stabilization term. Hence, we replace the term

$$\langle a_2 \cdot n\{\lambda_h/u_h\}, \xi_h - v_h\rangle_{\partial\mathcal{P}_h^{(2)}}$$

in (6.37) by

$$\langle s(\lambda_h - u_h), v_h - \xi_h\rangle_{\partial\mathcal{P}_h^{(2)}} - \langle a_2 \cdot n\lambda_h, v_h\rangle_{\partial\mathcal{P}_h^{(2)}}$$

with a suitable stabilization function $s\colon \mathbb{S}\times\mathbb{S}\times[0,T]\to\mathbb{R}$ satisfying

$$s(\vec{n},\vec{e},t) \geq \|a_2(\vec{n},\vec{e},t)\| \quad \text{for all } \vec{n},\vec{e}\in\mathbb{S}, t\in[0,T].$$

In the context of the RB method, it is important to choose a stabilization function that is affine in the parameters. To this end, we choose

$$s(\vec{n},\vec{e},t;\mu) = \sum_{i=1}^{7} |\theta_{a_2}^{(i)}(\mu)| \|\tilde{a}_2^{(i)}(\vec{n},\vec{e},t)\|_2$$

correspondingly to (6.40). Hence, we arrive at the modified bilinear form for the convection part $\tilde{B}_C\colon (V\times M)\times(V\times M)\to\mathbb{R}$ defined by

$$\begin{aligned}\tilde{B}_C((u,\lambda),(v,\xi)) = {}& \langle u, a_1\cdot\nabla_{\mathbb{S}_1} v\rangle_{\mathcal{P}_h} + \langle a_1\cdot n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h^{(1)}} \\ & + \langle u, a_2\cdot\nabla_{\mathbb{S}_2} v\rangle_{\mathcal{P}_h} + \langle s(\lambda-u), v-\xi\rangle_{\partial\mathcal{P}_h^{(2)}} - \langle a_2\cdot n\lambda, v\rangle_{\partial\mathcal{P}_h^{(2)}} \\ & + \langle u, \partial_t v\rangle_{\mathcal{P}_h} + \langle n\{\lambda/u\}, \xi - v\rangle_{\partial\mathcal{P}_h^{(t)}}.\end{aligned} \tag{6.41}$$

Since the flux in time direction is constant, we also keep the upwind flux on $\partial\mathcal{P}_h^{(t)}$ similar to the flux on the first sphere.

With the above adaptations we arrive at an adapted version of the HDG variational formulation that can be applied to the Fokker-Planck equation for MPI in (5.12): find $(\sigma_h, u_h, \lambda_h)\in W_h$ such that

$$B((\sigma_h,u_h,\lambda_h),(\tau_h,v_h,\xi_h)) = l((\tau_h,v_h,\xi_h)) \quad \text{for all } (\tau_h,v_h,\xi_h)\in W_h,$$

where $B\colon W\times W\to\mathbb{R}$ is defined by

$$\begin{aligned}&B((\sigma_h,u_h,\lambda_h),(\tau_h,v_h,\xi_h)) \\ &\qquad = B_D((\sigma_h,u_h,\lambda_h),(\tau_h,v_h,\xi_h)) + \tilde{B}_C((u_h,\lambda_h),(v_h,\xi_h))\end{aligned} \tag{6.42}$$

with B_D from (6.36) and $\tilde{B}_C$ from (6.41) as well as $l\colon W\to\mathbb{R}$ from (6.38). For notational simplicity, we have neglected the parameter-dependence in the bilinear form and the

linear functional in the formulation above. As for the linear functional, one can show that there exists an affine decomposition of the bilinear form

$$B((\sigma,u,\lambda),(\tau,v,\xi);\mu) = \sum_{q=1}^{Q_B} \theta_q^{(B)}(\mu) B_q((\sigma,u,\lambda),(\tau,v,\xi))$$

for suitable bilinear forms $B_q : W \times W \to \mathbb{R}$ and coefficient functions $\theta_q^{(B)} : \mathcal{P} \to \mathbb{R}$ with $Q_B \in \mathbb{N}$. The concrete derivation of the bilinear forms B_q and the coefficient functions $\theta_q^{(B)}$ is straightforward but lengthy and similar to the derivation of the linear functional above. Hence, we omit the details here.

Finally, we define the error estimator for the RB method as described in Section 6.1. Let $\{\Phi_1,\dots,\Phi_N\}$ be a basis of the finite element space W_h and let $\mathbf{A}_h(\mu) \in \mathbb{R}^{N\times N}$ be the matrix representation of the bilinear form $B(\cdot,\cdot;\mu)$ with respect to this basis as in (6.7). Moreover, let $\mathbf{f}(\mu) \in \mathbb{R}^N$ be the vector representation of the linear functional $l(\cdot;\mu)$ with respect to the same basis. We assume that we have computed the snapshots $(\sigma^{(i)},u^{(i)},\lambda^{(i)})$ for $i = 1,\dots,n$ with corresponding coefficient vector $\mathbf{w}^{(i)}$ in the basis representation, i.e.,

$$(\sigma^{(i)},u^{(i)},\lambda^{(i)}) = \sum_{j=1}^{N} w_j^{(i)} \Phi_j.$$

Moreover, let $\mathbf{U}$ be defined correspondingly to (6.13) with the coefficient vectors $\mathbf{w}^{(i)}$ as columns. Additionally, let $\tilde{\mathbf{c}}(\mu) \in \mathbb{R}^n$ be the coefficient vector of the RB solution with respect to the RB, i.e., the solution of

$$\mathbf{A}_{\mathrm{RB}}(\mu)\mathbf{c}(\mu) = \mathbf{f}_{\mathrm{RB}}(\mu)$$

with $\mathbf{A}_{\mathrm{RB}}(\mu)$ and $\mathbf{f}_{\mathrm{RB}}(\mu)$ defined in (6.14). Let $\tilde{\mathbf{A}}$ be a suitable parameter-independent preconditioner for $\mathbf{A}_h(\mu)$. In this work, we choose $\tilde{\mathbf{A}}$ as the matrix representation of the inner product $\langle\cdot,\cdot\rangle_{W,\mu^*} = \langle\cdot,\cdot\rangle_W$ on W defined in (6.27) for a fixed parameter $\mu^* \in \mathcal{P}$ and with respect to the basis $\{\Phi_1,\dots,\Phi_N\}$. Hence, $\tilde{\mathbf{A}}$ is defined such that

$$\big\langle(\sigma,u,\lambda),(\tau,v,\xi)\big\rangle_{W,\mu^*} = \tilde{\mathbf{w}}^\top \tilde{\mathbf{A}} \mathbf{w}$$

for all $(\sigma,u,\lambda),(\tau,v,\xi) \in W_\mathsf{h}$, where

$$(\sigma,u,\lambda) = \sum_{j=1}^{N} \tilde{w}_j \Phi_j, \quad (\tau,v,\xi) = \sum_{j=1}^{N} w_j \Phi_j.$$

The inner product $\langle\cdot,\cdot\rangle_W$ from (6.27) extended to the MPI problem is in general parameter-dependent via the diffusion coefficients $\varepsilon_1(\mu^*)$ and $\varepsilon_2(\mu^*)$ as well as the convection fields $a_1(\mu^*)$ and $a_2(\mu^*)$. For this reason, we choose the reference parameter μ^* for the definition of the inner product and thus for the matrix $\tilde{\mathbf{A}}$.

With this choice of $\tilde{\mathbf{A}}$, we define the error estimator $\eta : \mathcal{P} \to \mathbb{R}$ by

$$\eta(\mu) = \left\| \mathbf{A}_h(\mu)\mathbf{U}\mathbf{c}(\mu) - \mathbf{f}_h(\mu) \right\|_{\tilde{\mathbf{A}}^{-1}}, \tag{6.43}$$

which is the residual measured in the norm induced by $\tilde{\mathbf{A}}^{-1}$. The relation of the error estimator and the error of the RB solution compared to the finite element solution, was discussed in Section 6.1.2, and in particular in (6.11). Moreover, we can bound

the error of the RB solution compared to finite element solution by the residual in the dual norm using the inf-sup property in Theorem 6.9. For more details, we refer to [QMN16, Section 3.6.1].

As mentioned in Section 6.1.2, we can efficiently evaluate the coefficient vector $\mathbf{c}(\mu)$ if we precalculate the parameter-independent matrices of the decomposition of $\mathbf{A}_{\mathrm{RB}}(\mu)$ in (6.16). Moreover, by the same concepts, we can efficiently evaluate the $\|\cdot\|_{\tilde{\mathbf{A}}^{-1}}$ norms in (6.43).

By the choice of the error estimator (6.43), inequality (6.6) holds with an additional constant $C > 0$ that depends on the choice of the preconditioner $\tilde{\mathbf{A}}$ related to (6.10). As mentioned in Theorem 6.1, we can also measure the error in a parameter-independent norm and choose $\|\cdot\|_W = \|\cdot\|_{W,\mu^*}$, where μ^* is the reference parameter corresponding to the preconditioner $\tilde{\mathbf{A}}$.

6.3 Finite Elements on Surfaces

The Fokker-Planck equation (5.15) related to MPI is posed on a surface consisting of two copies of the unit sphere $\mathbb{S} \subset \mathbb{R}^3$ with the corresponding surface gradient and surface divergence defined in Theorem 5.2. To apply the HDG method from Section 6.2, we need to define certain quantities related to finite elements on surfaces. The finite element spaces in (6.33) are defined on the surface, where the spaces $V_{h_i}^{(i)}$ are defined by a suitable transformation of a polynomial spaces from a reference triangle and the spaces $\Sigma_{h_i}^{(i)}$ by a suitable transformation of a Raviart-Thomas space from the reference triangle. In particular, to make these finite element spaces more precise, we need to characterize the triangulation on a sphere and these transformations from the reference triangle in $\mathbb{R}^2$.

Especially, we introduce the Raviart-Thomas basis functions of arbitrary degree on spherical triangles. To do this, we first recall the definition of the Raviart-Thomas basis functions on planar simplices. Then introduce the corresponding functions on spherical triangles using the Piola transformation.

Due to the product structure of the finite element spaces in (6.34), we can restrict the discussion to finite elements on a single sphere.

6.3.1 Triangulation of the Sphere

For finite element methods on surfaces, there are two common approaches. The first one is approximating the surface with planar elements; the second is working with surface elements. Figure 6.1 visualizes both approaches for the special case of the unit sphere. In the first approach, the sphere is approximated by a polyhedral surface consisting of planar triangles. In this case, an additional geometric error arises from this approximation. For details on finite element methods using planar approximations of surfaces, we refer to [DE13]. An advantage of this approach is the simple implementation, since the mapping from the reference triangle to a planar triangle is affine. However, for planar triangles, the normal vectors of two adjacent triangles are not opposite to each other, meaning $n_{T_1} \neq -n_{T_2}$ for two triangles T_1 and T_2 sharing an edge. This can lead to issues when using upwind schemes, as inflow can occur in both triangles. Hence, in this work, we follow the second approach and use spherical triangles at the price of non-affine mappings.

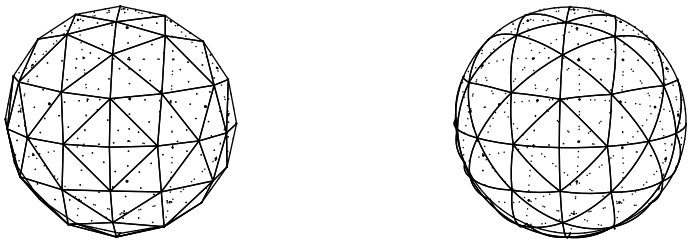


FIGURE 6.1: Planar triangulation (left) and spherical triangulation (right) of the unit sphere.

6.3.1.1 Transformation of Triangular Elements Onto a Sphere

In this section, we introduce the mapping from a reference triangle to a spherical triangle on the unit sphere. To this end, we denote by $\hat{T} \subset \mathbb{R}^2$ the reference triangle, which is

$$\hat{T} = \left\{ x \in \mathbb{R}^2 : x_1, x_2 > 0,\ x_1 + x_2 < 1 \right\}.$$

Furthermore, let $T \subset \mathbb{R}^3$ be a planar triangle identified by the vertices $y^{(1)}, y^{(2)}, y^{(3)} \in \mathbb{R}^3$. Hence

$$T = \left\{ y \in \mathbb{R}^3 : y = \lambda_1 y^{(1)} + \lambda_2 y^{(2)} + \lambda_3 y^{(3)},\ \lambda_i > 0,\ \lambda_1 + \lambda_2 + \lambda_3 = 1 \right\},$$

which means T is the interior of the convex hull of these three points. Then the corresponding spherical triangle is defined by the projection of the planar triangle onto the unit sphere, i.e.,

$$\tilde{T} = \left\{ z \in \mathbb{R}^3 : z = \frac{y}{\|y\|},\ y \in T \right\}.$$

We define the mapping from the reference triangle to the spherical triangle as the composition of two mappings. First, we map the reference triangle to the planar triangle by an affine mapping. Then we project the planar triangle onto the sphere. We define the affine mapping $F_{\hat{T}\to T}\colon \hat{T} \to T$ from the reference triangle to the planar triangle by

$$F_{\hat{T}\to T}(x) = \mathbf{A}_T x + y^{(1)},$$

where

$$\mathbf{A}_T = \begin{pmatrix} \vdots & \vdots \\ y^{(2)} - y^{(1)} & y^{(3)} - y^{(1)} \\ \vdots & \vdots \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

The corresponding inverse mapping $F^{-1}_{\hat{T}\to T}\colon T \to \hat{T}$ is given by

$$F^{-1}_{\hat{T}\to T}(y) = \mathbf{B}_T(y - y^{(1)}),$$

where

$$\mathbf{B}_T = (\mathbf{A}_T^\top \mathbf{A}_T)^{-1} \mathbf{A}_T^\top \in \mathbb{R}^{2\times 3}$$

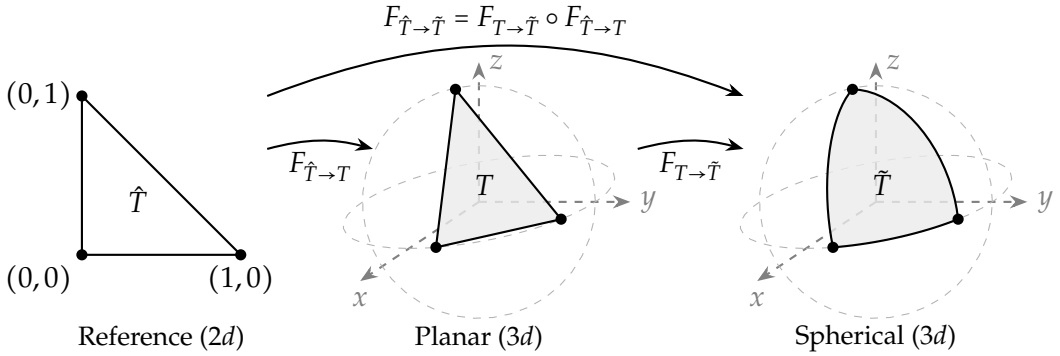


FIGURE 6.2: Mapping $F_{\hat{T}\to\tilde{T}}$ from the reference triangle $\hat{T}$ to the spherical triangle $\tilde{T}$. First, we map from the reference triangle $\hat{T}$ to the planar triangle T using the affine mapping $F_{\hat{T}\to T}$. Then we project the planar triangle T onto the sphere using the mapping $F_{T\to\tilde{T}}$.

is the Moore–Penrose inverse. Additionally, we define the mapping $F_{T\to\tilde{T}}\colon T \to \tilde{T}$ from the planar triangle to the spherical triangle by

$$F_{T\to\tilde{T}}(y) = \frac{y}{\|y\|}.$$

The mapping $F_{\hat{T}\to\tilde{T}}\colon \hat{T} \to \tilde{T}$ from the reference triangle to the spherical triangle is then given by the composition of the two mappings,

$$F_{\hat{T}\to\tilde{T}} = F_{T\to\tilde{T}} \circ F_{\hat{T}\to T}. \tag{6.44}$$

A visualization of the mapping is given in Figure 6.2.

Finally, we derive an expression for the inverse mapping from the reference triangle to the spherical triangle. For this purpose, we determine the inverse of the mapping from a planar triangle to a spherical triangle $F^{-1}_{T\to\tilde{T}}$. Then we have

$$F^{-1}_{\hat{T}\to\tilde{T}} = F^{-1}_{\hat{T}\to T} \circ F^{-1}_{T\to\tilde{T}}.$$

The planar triangle is part of a plane, i.e.,

$$T \subset \left\{ y \in \mathbb{R}^3 : \nu_T \cdot y = d_T \right\},$$

where

$$\nu_T = \left(y^{(2)} - y^{(1)}\right) \times \left(y^{(3)} - y^{(1)}\right)$$

is the normal vector of the plane and

$$d_T = y^{(1)} \cdot \nu_T = \det\left(y^{(1)}, y^{(2)} - y^{(1)}, y^{(3)} - y^{(1)}\right) = \det\left(y^{(1)}, y^{(2)}, y^{(3)}\right).$$

Then the inverse mapping from the spherical triangle to the planar triangle is given by the projection onto this plane

$$F^{-1}_{T\to\tilde{T}}(y) = \frac{\det(y^{(1)}, y^{(2)}, y^{(3)})}{\nu_T \cdot y} y.$$

The finite element spaces $V^{(i)}_{h_i}$ defined in (6.33) of the HDG method to solve the

MPI Fokker-Planck equation in Section 6.2.4 are thus defined using the mapping $F_{\hat{T}\to\tilde{T}}$ for $i = 1, 2$.

Remark 6.12. Since the mapping $F_{\hat{T}\to\tilde{T}}$ from the reference triangle to the spherical triangle is not affine, the finite element spaces using this mapping are no longer polynomial. However, for sufficiently fine triangulations of the sphere, we have $\|F_{\hat{T}\to T}(x)\| \approx 1$ for all $x \in \hat{T}$. Thus, we obtain

$$F_{\hat{T}\to\tilde{T}}(x) = \frac{F_{\hat{T}\to T}(x)}{\|F_{\hat{T}\to T}(x)\|} \approx F_{\hat{T}\to T}(x), \quad x \in \hat{T}.$$

Hence, the mapping is close to an affine mapping and we can expect that the finite element spaces defined by using the mapping $F_{\hat{T}\to\tilde{T}}$ are close to polynomial spaces. This property is helpful for the computation of the integrals in the HDG method, since we can use quadrature rules for polynomials to compute the integrals approximately.

6.3.1.2 Spherical Edges

Next, we consider the edges of a spherical triangle. In the context of the HDG method from Section 6.2, we are interested in the transformation from the reference edge to a spherical edge and in the normal vector on the spherical edge. In particular, the finite element spaces $M_{h_i}^{(i)}$ in (6.33) are defined via this transformation.

First, we define the mapping from a reference edge $\hat{E} = [0, 1]$ to a spherical edge. Let $E, \tilde{E}$ be the edges of the planar triangle T and spherical triangle $\tilde{T}$ respectively. The edges are defined by the vertices $y^{(1)}, y^{(2)} \in \mathbb{R}^3$, i.e.,

$$E = \left\{y \in \mathbb{R}^3 : y = (1-t)y^{(1)} + ty^{(2)},\ t \in [0,1]\right\}, \quad \tilde{E} = \left\{z \in \mathbb{R}^3 : z = \tfrac{y}{\|y\|},\ y \in E\right\}.$$

Correspondingly, we define the mappings

$$\begin{aligned} F_{\hat{E}\to E}&: \hat{E} \to E, \quad t \mapsto y^{(1)} + t\left(y^{(2)} - y^{(1)}\right), \\ F_{E\to\tilde{E}}&: E \to \tilde{E}, \quad y \mapsto \frac{y}{\|y\|}, \\ F_{\hat{E}\to\tilde{E}}&: \hat{E} \to \tilde{E}, \quad t \mapsto \left(F_{E\to\tilde{E}} \circ F_{\hat{E}\to E}\right)(t). \end{aligned}$$

The inverse mapping of $F_{\hat{E}\to\tilde{E}}$ is again given by the concatenation of the inverses of the other two mappings, i.e.,

$$F_{\hat{E}\to\tilde{E}}^{-1} = F_{\hat{E}\to E}^{-1} \circ F_{E\to\tilde{E}}^{-1},$$

where $F_{E\to\tilde{E}}^{-1}$ is the projection from the spherical edge to the planar edge and $F_{\hat{E}\to E}^{-1}$ is the mapping from the planar edge to the reference edge.

Using the product rule and properties of the cross product, we obtain

$$F'_{\hat{E}\to\tilde{E}}(t) = \frac{\left(F_{\hat{E}\to\tilde{E}}(t) \times \left(y^{(2)} - y^{(1)}\right)\right) \times F_{\hat{E}\to\tilde{E}}(t)}{\left\|F_{\hat{E}\to E}(t)\right\|}$$

as well as

$$\left\|F'_{\hat{E}\to\tilde{E}}(t)\right\| = \frac{\left\|y^{(1)} \times y^{(2)}\right\|}{\left\|F_{\hat{E}\to E}(t)\right\|^2}.$$

Since E is a line segment and the unit ball is convex, we have $\|F_{\hat{E}\to E}(t)\| \leq 1$ for all $t \in [0,1]$.

Next, we consider the normal vector of a spherical edge. The normal vector at a certain point on the spherical edge is orthogonal to the edge and lies in the tangent space of the sphere, i.e., it is orthogonal to the point itself. Hence, the normalized normal vector of a spherical edge $\tilde{E}$ is given by

$$n_T(t) = \frac{F'_{\hat{E}\to\tilde{E}}(t) \times F_{\hat{E}\to\tilde{E}}(t)}{\|F'_{\hat{E}\to\tilde{E}}(t) \times F_{\hat{E}\to\tilde{E}}(t)\|} = \frac{y^{(1)} \times y^{(2)}}{\|y^{(1)} \times y^{(2)}\|}.$$

In particular, the normal vector is constant along the spherical edge.

6.3.2 Raviart-Thomas Functions

The goal of this section is to define Raviart-Thomas elements for spherical triangles. For this purpose, we proceed as follows. First, we recall the definition of Raviart-Thomas elements on reference simplices. Then we discuss the Piola transformation for mappings onto manifolds in $\mathbb{R}^3$, and subsequently combine these two concepts to define Raviart-Thomas elements on spherical triangles.

6.3.2.1 Definition for Simplicial Elements

In this section, we keep the notation as general as possible and define the Raviart-Thomas elements for simplicial elements in $\mathbb{R}^d$, where we follow the notation of [EG21, Chapter 14]. We start with some basic definitions. First, we define the set of multi-indices

$$\mathcal{A}_{k,d} = \left\{\alpha \in \mathbb{N}_0^d : |\alpha| = \alpha_1 + \cdots + \alpha_d \leq k\right\}$$

and the subset

$$\mathcal{A}^H_{k,d} = \left\{\alpha \in \mathbb{N}_0^d : |\alpha| = k\right\}.$$

Moreover, for $\alpha \in \mathbb{N}_0^d$ and $x \in \mathbb{R}^d$, we use the standard notation

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}.$$

Definition 6.13. For $k \in \mathbb{N}$ and $d \geq 2$ let

$$\mathbb{P}_{k,d} = \left\{p \in C(\mathbb{R}^d, \mathbb{R}) : p(\mathbf{x}) = \sum_{\alpha \in \mathcal{A}_{k,d}} a_\alpha x^\alpha\right\}$$

be the set of polynomials of degree k and let

$$\mathbb{P}^H_{k,d} = \left\{p \in C(\mathbb{R}^d, \mathbb{R}) : p(\mathbf{x}) = \sum_{\alpha \in \mathcal{A}^H_{k,d}} a_\alpha x^\alpha\right\}$$

be the set of homogeneous polynomials of degree k. The Raviart-Thomas space $\mathbb{RT}_{k,d}$ of $\mathbb{R}^d$-valued polynomials is then defined by

$$\mathbb{RT}_{k,d} = (\mathbb{P}_{k,d})^d \oplus x\mathbb{P}^H_{k,d}.$$

As a direct consequence of the definition, we have

$$\operatorname{div}(\mathbb{RT}_{k,d}) = \{\operatorname{div}(f) : f \in \mathbb{RT}_{k,d}\} = \mathbb{P}_{k,d}.$$

Now let K be a simplex in $\mathbb{R}^d$ with vertices $\{v^{(i)}\}_{i=1,\dots,d}$. We denote the set of all faces of this simplex by

$$\mathcal{F}_K = \{F_1, \dots, F_{d+1}\},$$

where face F_i is the face opposite to the vertex $v^{(i)}$. Each face $F \in \mathcal{F}_K$ is oriented with a corresponding normal vector ν_F, where ν_F is scaled such that $|\nu_F| = |F|$. Next, for facet $F \in \mathcal{F}_K$ let $T_F \colon \hat{S}^{d-1} \to F$ be an affine mapping from the $(d-1)$-dimensional unit simplex $\hat{S}^{d-1}$ to F.

The Raviart-Thomas finite element basis functions are defined by properties on the moments of the normal component of the flux across the faces and moments of the normal component of the flux in the interior of the simplex. To this end, we define the corresponding linear forms.

Definition 6.14. Let $\{\zeta_m\}_{m=1,\dots,m_1}$ be a basis of $\mathbb{P}_{k,d-1}(\hat{S}^{d-1})$. Then we define the linear form $\sigma^{\mathrm{f}}_{F,m} \colon \mathbb{RT}_{k,d} \to \mathbb{R}$ for each $m \in \{1, \dots, m_1\}$ and face $F \in \mathcal{F}_K$ by

$$\sigma^{\mathrm{f}}_{F,m}(v) = \frac{1}{|F|} \int_F (v \cdot \nu_F)(\zeta_m \circ T_F^{-1}) \,\mathrm{d}s.$$

Moreover, let $\{\psi_m\}_{m=1,\dots,m_2}$ be a basis of $\mathbb{P}_{k-1,d}(K)$. Then we define the linear form $\sigma^{\mathrm{c}}_{j,m} \colon \mathbb{RT}_{k,d} \to \mathbb{R}$ for each $m \in \{1, \dots, m_2\}$ and $j = 1, \dots, d$ by

$$\sigma^{\mathrm{c}}_{j,m}(v) = \frac{1}{|K|} \int_K (v \cdot \nu_{F_j}) \psi_m \,\mathrm{d}s.$$

Finally, we denote by

$$\Sigma = \left\{\sigma^{\mathrm{f}}_{F,m} : F \in \mathcal{F}_K, m = 1, \dots, m_1\right\} \cup \left\{\sigma^{\mathrm{c}}_{j,m} : j = 1, \dots, d, m = 1, \dots, m_2\right\}$$

the set of all such linear forms.

Let $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ be the set of linear forms according to Theorem 6.14. Then the corresponding Raviart-Thomas finite element basis functions $\Phi_1, \dots, \Phi_n \in \mathbb{RT}_{k,d}$ are defined by the property

$$\sigma_i(\Phi_j) = \delta_{i,j}. \tag{6.45}$$

For a proof that $\{\Phi_i\}_{i=1,\dots,n}$ is indeed a basis of $\mathbb{RT}_{k,d}$, we refer to [EG21, Proposition 14.15]. The corresponding linear forms can be interpreted as degrees of freedom for the Raviart-Thomas elements.

In the final step, we will discuss the calculation of the Raviart-Thomas finite element basis functions. We assume that the set of linear forms Σ is given. Moreover, let $\{p_l\}$ be a basis of $\mathbb{RT}_{k,d}$ and $\mathbf{A} \in \mathbb{R}^{n \times n}$ be defined by

$$A_{i,j} = \sigma_i(p_j).$$

Then the Raviart-Thomas finite element basis functions Φ_i are given by

$$\Phi_i(x) = \sum \lambda_l p_l(x),$$

where $\lambda \in \mathbb{R}^n$ is the solution of $\mathbf{A}\lambda = \mathbf{e}_i$ and $\mathbf{e}_i$ is the i-th unit vector in $\mathbb{R}^d$.

Remark 6.15. The Raviart-Thomas finite element basis functions depend on the choice of the basis functions for the linear forms in Theorem 6.14. For the implementation of the finite element method, it can be useful to choose the basis functions according to the polynomial basis used for the finite element spaces on the simplex

K and its faces. In this case, some integrals appearing in the bilinear form of the HDG method can be calculated analytically.

6.3.2.2 Piola Transformation

In the previous section, we have defined the Raviart-Thomas elements on reference simplices in $\mathbb{R}^d$. In the context of the MPI Fokker-Planck equation, we are interested in Raviart-Thomas elements on spherical triangles, i.e., on subsets of two-dimensional manifolds in $\mathbb{R}^3$. The Raviart-Thomas elements shall have the same degrees of freedom as the Raviart-Thomas elements on the reference triangle, meaning that the linear forms defined in Theorem 6.14 are preserved under the mapping from the reference triangle to the spherical triangle. To do this, we introduce the *Piola transformation*. For notational simplicity, we keep the notation as general as possible and discuss the Piola transformation for mappings onto general two-dimensional manifolds M in $\mathbb{R}^3$. Therefore, we follow [Rog+13; Cas+16].

Assume we have a regular bijective mapping $\varphi\colon \hat{T} \to K$, where K is a subset of a two-dimensional manifold M in $\mathbb{R}^3$ and $\hat{T} \subset \mathbb{R}^2$ is the reference triangle. In particular, we are interested in the case where $\varphi = F_{\hat{T}\to\tilde{T}}$ is the mapping from the reference triangle $\hat{T}$ to a spherical triangle $\tilde{T}$ defined in Section 6.3.1.1. We define the Jacobian matrix of the mapping φ by

$$\mathbf{J} = D\varphi = \begin{pmatrix} \frac{\mathrm{d}\varphi_1}{\mathrm{d}x_1} & \frac{\mathrm{d}\varphi_1}{\mathrm{d}x_2} \\ \frac{\mathrm{d}\varphi_2}{\mathrm{d}x_1} & \frac{\mathrm{d}\varphi_2}{\mathrm{d}x_2} \\ \frac{\mathrm{d}\varphi_3}{\mathrm{d}x_1} & \frac{\mathrm{d}\varphi_3}{\mathrm{d}x_2} \end{pmatrix}.$$

Since M is a two-dimensional manifold, the Jacobian matrix $\mathbf{J}$ has full column rank two. Furthermore, let the pseudo-determinant be defined by

$$\det(\mathbf{J}) = \pm \| \nabla_{x_1} \varphi \times \nabla_{x_2} \varphi \|,$$

which corresponds to the area of the parallelogram spanned by the columns of $\mathbf{J}$, where the sign depends on the orientation of K with respect to the manifold. The Moore-Penrose inverse of $\mathbf{J}$ is given by

$$\mathbf{J}^\dagger = \left(\mathbf{J}^\top \mathbf{J}\right)^{-1} \mathbf{J}^\top.$$

Next, we consider the relation between the standard gradient in $\mathbb{R}^2$ and the surface gradient of the manifold with respect to the mapping φ. For a general definition of the surface gradient for manifolds with local parametrization, we refer to [DE13]. For the special case of the unit sphere, a definition of the surface gradient is given in Theorem 5.2. Let $\hat{f}\colon \mathbb{R}^2 \to \mathbb{R}$ and let $f = \hat{f} \circ \varphi^{-1}$. Then we have

$$(\nabla_M f) \circ \varphi = (\mathbf{J}^\dagger)^\top \nabla \hat{f}.$$

For a proof of this relation, we refer to [Cas+16].

The degree of freedoms defined by the linear forms in Theorem 6.14 are not preserved directly by the mapping φ, i.e., the condition (6.45) is generally not satisfied for the basis functions $\hat{\Phi}_j \circ \varphi^{-1}$, where $\hat{\Phi}_j$ are the Raviart-Thomas finite element basis

functions on the reference triangle $\hat{T}$. To overcome this, we use the Piola transformation, which is defined by

$$\Phi = \left(\frac{1}{\det(\mathbf{J})}\mathbf{J}\hat{\Phi}\right) \circ \varphi^{-1}$$

for a vector field $\hat{\Phi}\colon \mathbb{R}^2 \to \mathbb{R}^2$. Based on this transformation, we have the property

$$\operatorname{div}_M(\Phi) = \left(\frac{1}{\det(\mathbf{J})}\operatorname{div}(\hat{\Phi})\right) \circ \varphi^{-1}.$$

Moreover, let $\hat{\psi}\colon \hat{T} \to \mathbb{R}$ be a sufficiently regular function and $\psi = \hat{\psi} \circ \varphi^{-1}$. Then we have

$$\begin{aligned}
\int_K \langle \Phi, \nabla_M \psi \rangle \,\mathrm{d}z &= \int_{\hat{T}} \langle \hat{\Phi}, \nabla \hat{\psi} \rangle \,\mathrm{d}x, \\
\int_K \operatorname{div}_M(\Phi)\psi \,\mathrm{d}z &= \int_{\hat{T}} \operatorname{div}(\hat{\Phi})\hat{\psi} \,\mathrm{d}x, \\
\int_{\partial K} \langle \Phi, \mathbf{n} \rangle \psi \,\mathrm{d}\sigma &= \int_{\partial \hat{T}} \langle \hat{\Phi}, \hat{\mathbf{n}} \rangle \hat{\psi} \,\mathrm{d}s.
\end{aligned}$$

For details of these properties, we refer to [Cas+16]. Based on these properties, the degrees of freedom defined by the linear forms in Theorem 6.14 are preserved. Hence, the Raviart-Thomas finite element space in (6.33) is defined using the Piola transformation of the Raviart-Thomas elements on the reference triangle $\hat{T}$.

Finally, we consider the special case of a spherical triangle $\tilde{T}$ defined by the vertices $y^{(1)}, y^{(2)}, y^{(3)} \in \mathbb{S}$. To this end, let $\varphi = F_{\hat{T}\to\tilde{T}}$ be the mapping from the reference triangle $\hat{T}$ to the spherical triangle $\tilde{T}$ defined in (6.44). Then we have

$$\mathbf{J}(x) = \left(\frac{\left(F_{\hat{T}\to\tilde{T}}(x) \times \left(y^{(2)} - y^{(1)}\right)\right) \times F_{\hat{T}\to\tilde{T}}(x)}{\|F_{\hat{T}\to T}(x)\|} \quad \frac{\left(F_{\hat{T}\to\tilde{T}}(x) \times \left(y^{(3)} - y^{(1)}\right)\right) \times F_{\hat{T}\to\tilde{T}}(x)}{\|F_{\hat{T}\to T}(x)\|} \right)$$

as well as

$$\det(\mathbf{J}(x)) = \|F_{\hat{T}\to T}(x)\|^{-3} \det\left(y^{(1)}, y^{(2)}, y^{(3)}\right).$$

6.4 Numerical Experiments

In the final part of this chapter, we present numerical results for the HDG method applied to the MPI Fokker-Planck equation discussed in Section 6.2.4. We consider the case of fixed parameters as well as the RB method introduced in Section 6.1.

The direct visualization of the full approximation of the Fokker-Planck equation is infeasible due to the dimensionality of the domain $\Omega = \mathbb{S} \times \mathbb{S} \times [0, T]$. Therefore, throughout this section, we visualize the quantity of interest of the solution, that is, the mean magnetic moment $\bar{m} \in (L_2(0,T))^3$ given by the expectation formula (5.11).

We assume that the triangulation of the sphere is a uniform refinement of the initial triangulation. The initial triangulation is the projection of an octahedron onto the sphere. A refinement splits each triangle into four smaller triangles by connecting its midpoints. Figure 6.3 visualizes the initial triangulation and the first two refinement levels. We denote by j_1 and j_2 the refinement levels of the triangulations of the first and second spheres, respectively. Moreover, we denote by $n = \#\mathcal{I}_{h_t}$ the number of time intervals, where $h_t = T/n$ with final time T.

The direct computation of the finite element solutions of the HDG method for the MPI Fokker-Planck equation is already computationally expensive, given the five-dimensional domain $\Omega = \mathbb{S} \times \mathbb{S} \times [0, T]$. However, this direct calculation is only required during the offline phase of the RB method. In this work, we focus solely

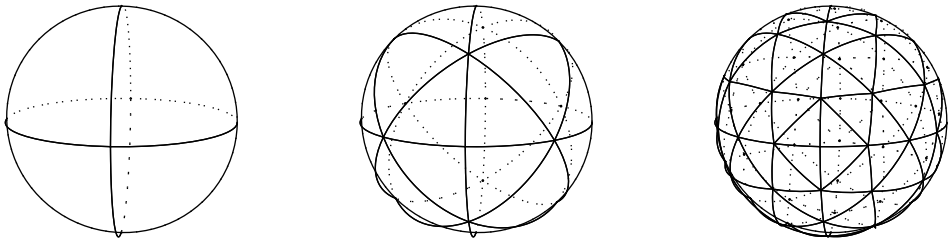


FIGURE 6.3: Different refinement levels for the triangulation of the sphere. The left figure corresponds to the initial triangulation, the middle figure corresponds to the first refinement level, and the right figure corresponds to the second refinement level.

on efficiency improvements utilizing the online phase. Optimizing the offline finite element computations is a separate challenge, which we discuss further in Chapter 7.

6.4.1 Solutions for Fixed Parameters

First, we consider the HDG method for the MPI Fokker-Planck equation for fixed parameter combinations. Before we present the results, we investigate the structure of the stiffness matrix $\mathbf{A}_h(\mu)$ defined in (6.7). In particular, the structure depends on the choice of the basis $\{\Phi_1, \dots, \Phi_N\}$ of the finite element space W_h defined in (6.35) as well as on the ordering of the basis functions. Keeping in mind the structure of the bilinear form B defined in (6.42), there are no direct interactions between the degrees of freedom defined on elements on different time intervals, i.e, the degrees of freedom defined on $T_1^{(1)} \times T_1^{(2)} \times I_1$ and the degrees of freedom defined on $T_2^{(1)} \times T_2^{(2)} \times I_2$ for $T_1^{(1)}, T_2^{(1)} \in \mathcal{T}_{h_1}^{(1)}$, $T_1^{(2)}, T_2^{(2)} \in \mathcal{T}_{h_2}^{(2)}$ and $I_1, I_2 \in \mathcal{I}_{h_t}$ if $I_1 \neq I_2$. Next, we consider interactions between degrees of freedom on elements and on faces. We consider two adjacent time intervals $I_1 = [t_0, t_1]$, $I_2 = [t_1, t_2]$. Hence, based on the use of upwind fluxes in the time direction, the degrees of freedom defined on the facets $T_1^{(1)} \times T_1^{(2)} \times \{t_1\}$ only depend on the degrees of freedom defined on $T_1^{(1)} \times T_1^{(2)} \times I_1$ but not on the degrees of freedom defined on $T_1^{(1)} \times T_1^{(2)} \times I_2$. Moreover, the degrees of freedom on $T_1^{(1)} \times T_1^{(2)} \times I_2$ depend on the degrees of freedom defined on $T_1^{(1)} \times T_1^{(2)} \times \{t_1\}$ but not on the degrees of freedom defined on $T_1^{(1)} \times T_1^{(2)} \times \{t_2\}$. A visualization of the flux in time direction in a simplified setting with only one triangle is given in Figure 6.4. Based on this visualization, we can see the dependencies of the degrees of freedom in the temporal direction. This behaviour carries over to the case of multiple triangles and non-planar triangles due to the product structure.

Hence, if we order the degrees of freedom first by the time interval and then by the spherical elements, we obtain a block lower-bidiagonal structure of the stiffness

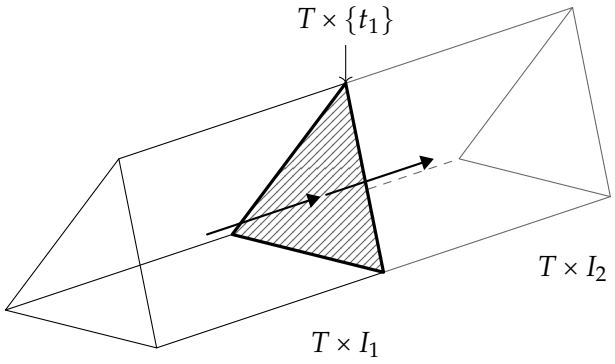


FIGURE 6.4: Schematic illustration of the flux in time direction in a simplified setting with only one triangle, which additionally is planar.

matrix $\mathbf{A}_h(\mu)$. The matrix is of the form

$$\mathbf{A}_h(\mu) = \begin{pmatrix} \mathbf{A}_{h,1}(\mu) & 0 & 0 & \dots & 0 \\ \mathbf{U} & \mathbf{A}_{h,2}(\mu) & 0 & \dots & 0 \\ 0 & \mathbf{U} & \mathbf{A}_{h,3}(\mu) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{U} & \mathbf{A}_{h,n}(\mu) \end{pmatrix}, \tag{6.46}$$

where n is the number of time intervals, $\mathbf{A}_{h,i}(\mu)$ is the stiffness matrix corresponding to the time interval I_i, and $\mathbf{U}$ is a matrix that corresponds to the upwind flux in time direction. Since the flux in the time direction is constant, the matrix $\mathbf{U}$ is independent of the time interval and the parameter.

Based on the block structure (6.46) of the stiffness matrix, we can solve the system of equations

$$\mathbf{A}_h(\mu)\mathbf{u} = \mathbf{f}_h(\mu)$$

by a forward substitution, similar to the forward substitution for solving linear system with a triangular matrix. Hence, we can solve the system of equations by solving n smaller systems.

Next, we present numerical results for the HDG method for the MPI Fokker-Planck equation for fixed parameter combinations. In particular, we fix all parameters except one and visualize the mean magnetic moment $\bar{m}(t)$ for different values of this remaining parameter. In these experiments, we use the parameters from [Azi+25].

For simplicity, we assume that the spatial location is fixed and chosen such that $\vec{H}_S(x) = 0$, where $\vec{H}_S$ is the static selection field in (5.1). Hence, we can neglect the spatial dependence here. Moreover, we consider a scanner with excitation only in x-direction, i.e., $\vec{H}_D(t) = (A_1 \sin(2\pi\nu_x t), 0, 0)^\top$, where A_1 is the excitation amplitude and ν_x is the excitation frequency. The remaining parameters are summarized in Table 6.1.

In the following, we present the results for the mean magnetic moment $\bar{m}(t)$ for different values of the saturation magnetization M_S, the particle core diameter D_C, and the anisotropy constant K_{anis}. We vary one of the three parameters at a time and choose the other two according to Table 6.1. The parameters of the HDG method

Parameter	Symbol	Value
Scanner		
Excitation frequency	ν_x	25000 Hz
Excitation amplitudes	A_1	0.02 Tμ_0^{-1}
Particle		
Temperature	T_{emp}	300 K
Saturation magnetization	M_S	300000 J m^{-3} T^{-1}
Particle core diameter	D_C	30 nm
Particle diameter	D_H	D_C + 5 nm
Dynamic viscosity of fluid	η	1.0×10^{-3} Pa s
Anisotropy constant	K_{anis}	3000 J m^{-3}
Gyromagnetic ratio	γ	1.75×10^{11} rad s^{-1} T^{-1}
Damping parameter	α	1.0

TABLE 6.1: List of physical parameters for numerical experiments with fixed parameter, taken from [Azi+25].

Parameter	Refinement			Degree			Dof
	j_1	j_2	n	k_1	k_2	k_t	
Saturation magnetization	0	3	400	0	2	2	$2.67 \cdot 10^8$
Particle core diameter	0	3	400	0	2	2	$2.67 \cdot 10^8$
Anisotropy constant	0	3	400	0	2	2	$2.67 \cdot 10^8$
	0	4	100	0	2	2	$2.67 \cdot 10^8$

TABLE 6.2: Parameter of the HDG method for the numerical experiments with fixed parameters: j_1 and j_2 denote the refinement levels for the triangulation of the two spheres, n denotes the number of time intervals, and $\mathsf{k} = (k_1, k_2, k_t)$ denotes the polynomial degrees. The last column shows the total number of degrees of freedom (dof) for the finite element space.

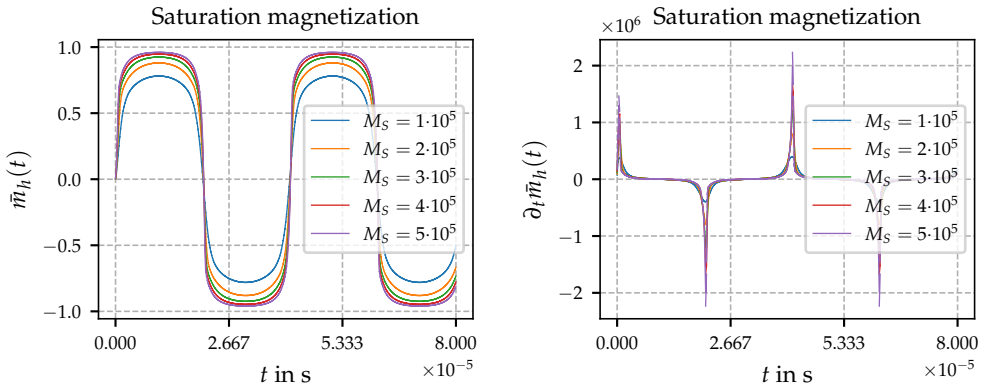


FIGURE 6.5: Approximation of the mean magnetic moment $\bar{m}(t)$ (left) and piecewise time derivative of the mean magnetic moment $\partial_t \bar{m}(t)$ for different values of the saturation magnetization M_S.

for these experiments are summarized in Table 6.2, which are the refinement level of the triangulation of the spheres, the number of time intervals, and the polynomial degrees of the finite element space for each component. Moreover, we state the total number of degrees of freedom for the entire finite element space for each experiment.

We begin by investigating the saturation magnetization M_S. The mean magnetic moment $\bar{m}(t)$ is visualized in the left part of Figure 6.5 for different values of the saturation magnetization $M_S \in \{i \cdot 10^5 \mathrm{J\,m^{-3}\,T^{-1}} : i = 1, \dots, 5\}$. Since we use a discontinuous Galerkin method, the approximation of the Fokker-Planck equation is discontinuous in time, and thus the corresponding mean magnetic moment is discontinuous. In the right part of Figure 6.5, we visualize the piecewise time derivative of the mean magnetic moment $\partial_t \bar{m}(t)$ for the same values of the saturation magnetization M_S. As expected, the amplitude of the mean magnetic moment $\bar{m}(t)$ increases with increasing saturation magnetization M_S.

In this experiment, we choose the final time to be $T = 2/\nu_x$. Moreover, we use the refinement levels $j_1 = 0$ and $j_2 = 3$ for the triangulation of the spheres, which corresponds to 8 and 512 triangles, respectively, and $n = 400$ time intervals as well as the polynomial degrees $\mathsf{k} = (0, 2, 2)$. This choice of parameters leads to a total number of degrees of freedom of approximately 267 million. The computation of the solution of the system of equations in a reasonable time is only possible due to the block structure of the stiffness matrix $\mathbf{A}_h(\mu)$ and the use of a forward substitution. Solving the system of equations in Julia [Bez+17] (via LAPACK [And+99]) required approximately 19.5 hours. Benchmarks were conducted on a system equipped with 640 GB of RAM and dual 24-core AMD EPYC Genoa 9254 processors.

Numerical experiments indicate that a finer triangulation of the second sphere, as chosen in this experiment, is required to obtain a stable approximation of the mean magnetic moment $\bar{m}(t)$ for the chosen parameters. This behaviour is reasonable, since the magnetic field enters the convection term of the second sphere.

Next, we investigate how the particle core diameter D_C influences the mean magnetic moment $\bar{m}(t)$. Again, we use the HDG method for the approximation of the Fokker-Planck equation with refinement levels $j_1 = 0$ and $j_2 = 3$ for the triangulation of the spheres, $n = 400$ time intervals, and polynomial degrees $\mathsf{k} = (0, 2, 2)$. The remaining parameters are stated in Table 6.1. The mean magnetic moment $\bar{m}(t)$ and the piecewise time derivative of the mean magnetic moment $\partial_t \bar{m}(t)$ are visualized in Figure 6.6 for different values of the particle core diameter $D_C \in \{15, 20, 25, 30, 35\}$nm.

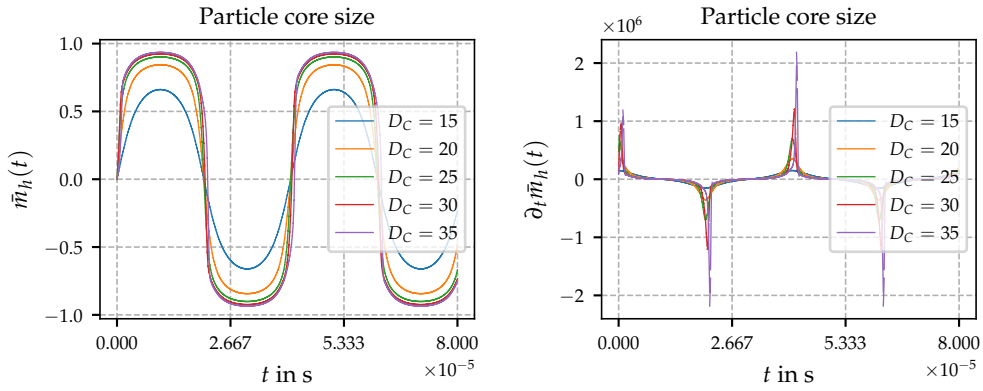


FIGURE 6.6: Approximation of the mean magnetic moment $\bar{m}(t)$ (left) and piecewise time derivative of the mean magnetic moment $\partial_t \bar{m}(t)$ for different values of the particle core diameter D_C.

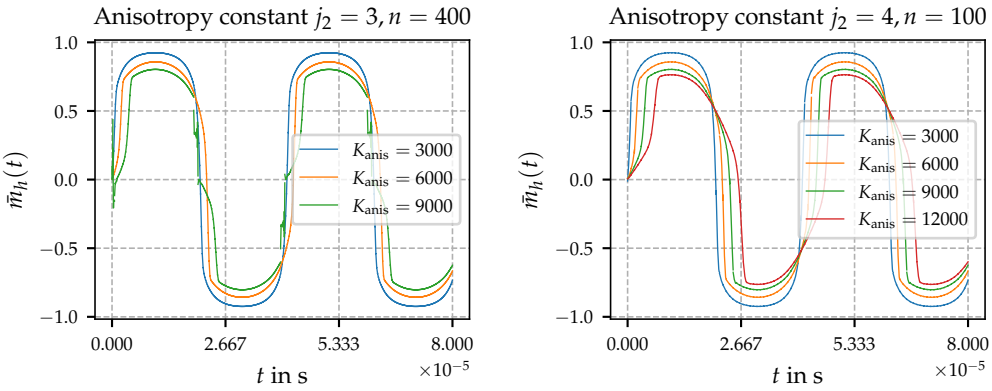


FIGURE 6.7: Approximation of the mean magnetic moment $\bar{m}(t)$ for different values of the anisotropy constant K_{anis} and different discretizations.

The behaviour of the mean magnetic moment $\bar{m}(t)$ for different values of the particle core size is comparable to the behaviour observed in [Azi+25].

Finally, we examine the influence of the anisotropy constant K_{anis} on the mean magnetic moment $\bar{m}(t)$. Again, we use the polynomial degrees $\mathsf{k} = (0,2,2)$, the refinement level $j_1 = 0$ for the first sphere, and the final time $T = 2/\nu_x$. In Figure 6.7, we visualize the mean magnetic moment $\bar{m}(t)$ for different values of the anisotropy constant K_{anis}, where we use different refinement levels j_2 and number of time intervals n in the left and right figure, respectively. The results indicate that a finer triangulation of the second sphere is required to obtain stable approximations of the mean magnetic moment $\bar{m}(t)$ for larger values of the anisotropy constant K_{anis}. A higher value of the anisotropy constant K_{anis} leads to a stronger coupling between the two spheres in the advection term (5.10) of the Fokker-Planck equation (5.12). Even for the results with higher refinement level of the second sphere, the discontinuous approach of the method can be seen upon closer inspection. Therefore, we omit the visualization of the corresponding piecewise derivatives for this case. Since we use an additional refinement level for the second sphere in the current experiments, but only a quarter of the time intervals, the number of degrees of freedom remains the same. Nevertheless, the runtime for solving the system of equations is slightly

higher, since we have less but larger blocks in the representation (6.46) of the stiffness matrix $\mathbf{A}_h(\mu)$.

The results for the anisotropy constant K_{anis} indicate that the choice of the discretization parameters, in particular the refinement level of the second sphere, can have a significant influence on the stability of the approximation of the mean magnetic moment $\bar{m}(t)$ for certain parameter combinations. One possible reason for this could be the use of the Lax-Friedrichs flux in the second sphere, since this flux can be more diffusive than the upwind flux, and we have a convection-dominated problem. Hence, it is important to choose the discretization parameters carefully in order to obtain stable approximations of the mean magnetic moment $\bar{m}(t)$.

6.4.2 Reduced Basis Method

To conclude, we present numerical results for the RB method applied to the MPI Fokker-Planck equation. In particular, we consider two test cases: one with excitation only in the x-direction and one with excitation in all three directions.

As stated in Section 6.1, the affine decomposition of the RB matrix $\mathbf{A}_{\text{RB}}(\mu)$ in (6.16) is crucial for an efficient computation of the RB approximation for a new parameter combination $\mu \in \mathcal{P}$ in the online phase, but also for the efficient selection of the next parameter combination in the greedy algorithm during the offline phase. However, for the computation of the parameter-independent matrices $\mathbf{G}^{(p,q)}$ of the affine decomposition of $\mathbf{A}_{\text{RB}}(\mu)$ in (6.15), we need to solve multiple linear systems of equations of the matrix $\tilde{\mathbf{A}}$, which is the matrix representation of the inner product as described in Section 6.2.4.2. Based on the definition of the inner product in (6.27), the matrix $\tilde{\mathbf{A}}$, in contrast to the stiffness matrix $\mathbf{A}_h(\mu)$, does not have a block lower-bidiagonal matrix, but a block tridiagonal matrix. Hence, the calculation of the solution of the linear systems of equations is more expensive than the solution of the linear systems of equations with the stiffness matrix $\mathbf{A}_h(\mu)$, since we cannot use forward substitution anymore. Therefore, we restrict ourselves to a coarser discretization of the finite element spaces in the HDG method in this section. Based on the possible stability issues for coarser discretization as observed in the previous section, we choose the parameter ranges in such a way that we obtain stable approximations of the mean magnetic moment $\bar{m}(t)$ for the chosen parameters and discretization.

Moreover, in this work, we do not consider all parameters in the parameter space of the RB method. We fix certain parameters and vary only a subset to span the parameter space $\mathcal{P}$. The fixed and variable parameters of the RB numerical experiments are summarized in Table 6.3. Moreover, the corresponding parameter ranges for the individual variable parameters are given in the last column of Table 6.3. The parameter space $\mathcal{P}$ is then given by the Cartesian product of the parameter ranges of the individual variable parameters.

In the experiments in this chapter, we use the refinement levels $j_1 = 0$ and $j_2 = 1$ for the triangulation of the spheres, $n = 100$ time intervals, polynomial degrees $\mathsf{k} = (0, 1, 2)$ and final time $T = \frac{1}{25000}$ for the HDG method in the offline phase.

6.4.2.1 Excitation in One Direction

We start with the test case where excitation only appears in the x-direction, i.e., we set $A_2 = A_3 = 0$. Hence, we have four variable parameters in this case. For the training set $\mathcal{P}_\delta$, we use a uniform grid with 20 grid points in each parameter direction, which leads to a total of $20^4 = 1.6 \cdot 10^5$ parameter combinations in the training set $\mathcal{P}_\delta$

Fixed Parameter	**Symbol**	**Value/Range**
Excitation frequency	ν_x	$2.5 \cdot 10^6/102$ Hz
Excitation frequency	ν_y	$2.5 \cdot 10^6/96$ Hz
Excitation frequency	ν_z	$2.5 \cdot 10^6/99$ Hz
Saturation magnetization	M_S	$474000\,\mathrm{J\,m^{-3}\,T^{-1}}$
Dynamic viscosity of fluid	η	1.0×10^{-3} Pa s
Gyromagnetic ratio	γ	$1.75 \times 10^{11}\,\mathrm{rad\,s^{-1}\,T^{-1}}$
Damping parameter	α	0.1
Anisotropy constant	K_{anis}	$3000\,\mathrm{J\,m^{-3}}$
Variable Parameter	**Symbol**	**Range**
Excitation amplitude	A_1	$[8, 16]\,\mathrm{mT}\mu_0^{-1}$
Excitation amplitude	A_2	$[8, 16]\,\mathrm{mT}\mu_0^{-1}$
Excitation amplitude	A_3	$[8, 16]\,\mathrm{mT}\mu_0^{-1}$
Temperature	T_{emp}	$[283.15, 303.15]$ K
Particle core diameter	D_C	$[15, 30]$ nm
Particle diameter	D_H	$D_C + [0, 15]$ nm

TABLE 6.3: List of physical parameters for numerical experiments in the RB method. The parameters are divided into fixed parameters and variable parameters, spanning the parameter space $\mathcal{P}$ for the RB method.

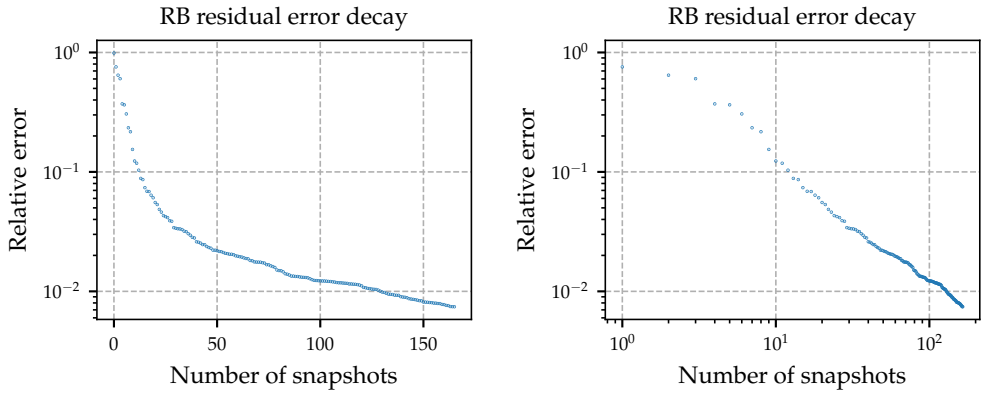


FIGURE 6.8: Relative error decay of the error estimator during the greedy algorithm with respect to the number of snapshots for the test case with excitation in one direction.

for the greedy algorithm. Figure 6.8 visualizes the relative error decay of the maximal value of the error estimator during the greedy algorithm with respect to the number of snapshots. We observe rapid decay of the error estimator in the first iterations of the greedy algorithm, indicating that the RB space captures the underlying low-dimensional structure of the solution manifold quickly. After a certain number of snapshots, the decay of the error estimator slows but continues to decrease.

Next, we discuss the quality of the RB solution for a new parameter combination $\mu \in \mathcal{P}$ in the online phase, using an RB space with 131 snapshots, which corresponds to the required number of snapshots to archive a relative error of the error estimator of 10^{-2}. The quantity of interest of the RB solution for a random parameter in the parameter space $\mathcal{P}$ is visualized in the left column of Figure 6.9, i.e., the mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$. In the right column of Figure 6.9, we visualize the difference between the quantity of interest of the

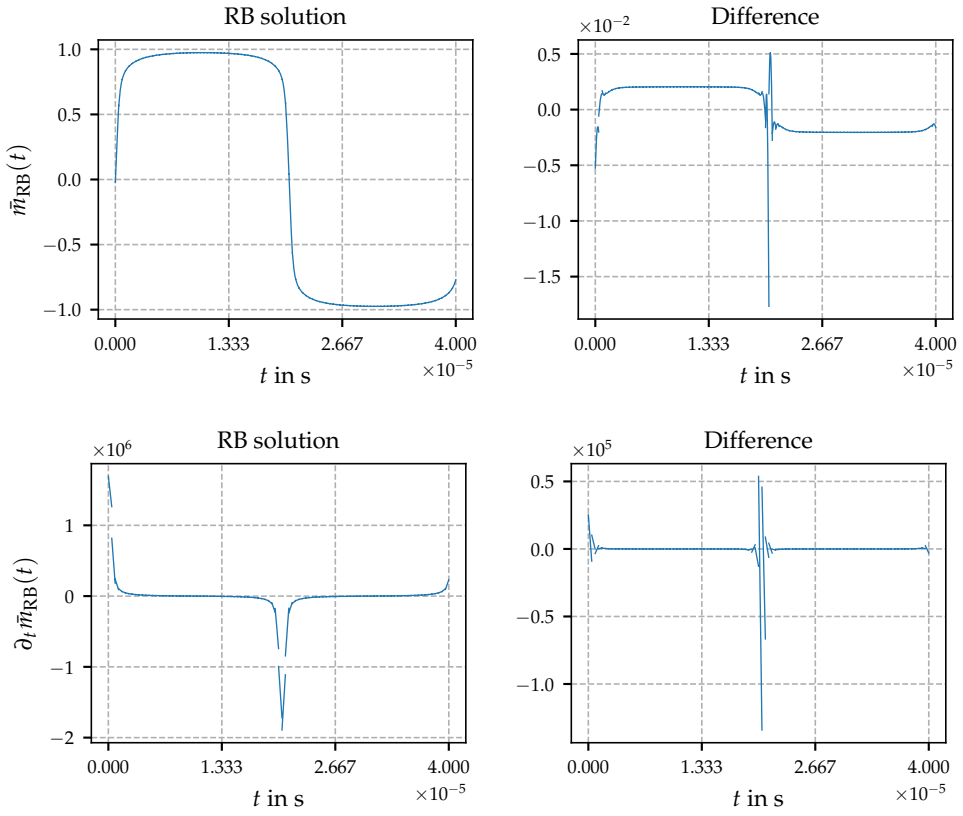


FIGURE 6.9: Mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$ for a new parameter combination $\mu \in \mathcal{P}$ with excitation in one direction computed by the RB method in the left column. In the right column, the difference between the RB solution and the finite element solution in terms of the mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$ is visualized.

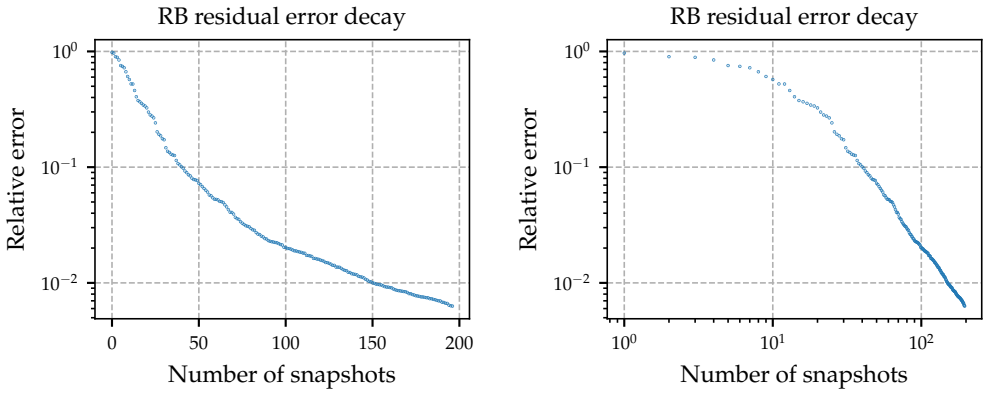


FIGURE 6.10: Relative error decay of the error estimator during the greedy algorithm with respect to the number of snapshots for the test case with excitation in three directions.

RB solution and the finite element solution. The relative L_2-error of the mean magnetic moment $\bar{m}(t)$ given by the RB solution compared to the finite element solution is approximately $2.2 \cdot 10^{-3}$ and the relative piecewise H^1-error is approximately $2.8 \cdot 10^{-2}$. The finite element solution requires the solution of a system of equations with approximately $2.1 \cdot 10^6$ degrees of freedom. In contrast, the RB solution requires only the solution of a system with 131 unknown. The runtime for the calculation of the finite element solution with the same hardware and software configuration as in Section 6.4.1 is approximately 3 minutes, while the runtime for the RB solution is approximately 0.2 milliseconds. Hence, we can achieve a speed-up of approximately $9 \cdot 10^5$, while still maintaining a good accuracy of the RB solution compared to the finite element solution.

However, we are actually interested in the approximation properties with respect to the exact solution of the Fokker-Planck equation. As seen in (6.5), the accuracy can be quantified as the sum of the RB error and the discretization error of the HDG method. As mentioned before, the use of finer discretization for the HDG method is not feasible here without additional effort due to the computational costs for the calculation of the parameter-independent matrices $\mathbf{G}^{(p,q)}$ in the affine decomposition of the RB matrix $\mathbf{A}_{\mathrm{RB}}(\mu)$. Nevertheless, the results indicate that the RB method achieves good accuracy compared to the finite element solution while significantly reducing the computational costs.

6.4.2.2 Excitation in Three Directions

Next, we consider the test case with excitation in all three directions. In this case, we have six variable parameters. For the training set $\mathcal{P}_\delta$, we again use a uniform grid, but with 10 grid points in each parameter direction, which leads to a total of 10^6 parameter combinations in the training set $\mathcal{P}_\delta$ for the greedy algorithm. Figure 6.10 shows the relative error decay of the maximal value of the error estimator during the greedy algorithm with respect to the number of snapshots. We observe a behaviour similar to that in the test case with excitation in one direction, where the error estimator decays rapidly in the first iterations, followed by slower decay as the number of snapshots increases. In this case, we need more snapshots to achieve a relative error of the error estimator of 10^{-2}, namely 152 compared to 131 snapshots in the test

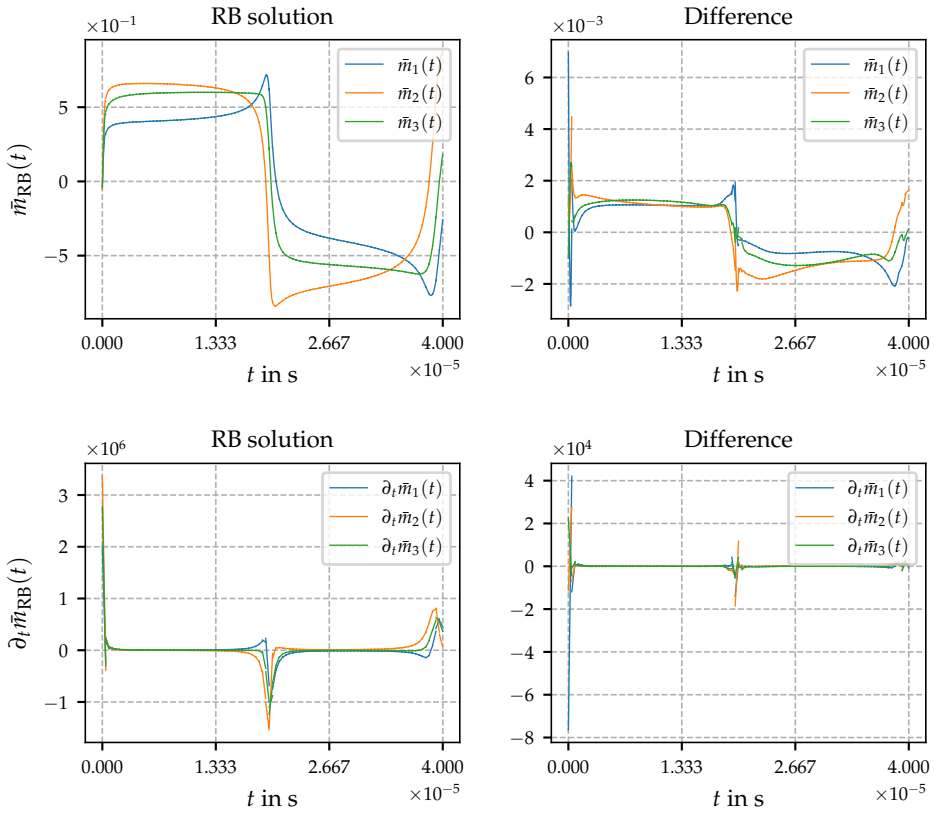


FIGURE 6.11: Mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$ for a new parameter combination $\mu \in \mathcal{P}$ with excitation in three direction computed by the RB method in the left column. In the right column, the difference between the RB solution and the finite element solution in terms of the mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$ is visualized.

case with excitation in one direction. This demonstrates that the solution manifold is more complex in this case due to the higher number of variable parameters.

In Figure 6.11, we again visualize the mean magnetic moment $\bar{m}(t)$ and its piecewise time derivative $\partial_t \bar{m}(t)$ for a new parameter combination $\mu \in \mathcal{P}$ given by the RB method. Moreover, we visualize the difference between the mean magnetic moment and its piecewise time derivative, compared with the mean magnetic moment from the finite element solution. The relative L_2-error of the mean magnetic moment $\bar{m}(t)$ from the RB method compared to the finite element method is approximately $2.0 \cdot 10^{-3}$ and the relative piecewise H^1-error is approximately $1.1 \cdot 10^{-2}$. The finite element solution requires the solution of a system of equations with approximately $2.1 \cdot 10^6$ degrees of freedom. In contrast, the RB solution requires only the solution of a system of 152 equations. The runtime for the solution of the system of equations for the finite element solution is again approximately 3 minutes. In contrast, the runtime for solving the system of equations for the RB solution is approximately 0.5 milliseconds. Hence, we can achieve a speed-up of approximately $3.6 \cdot 10^5$.

As final numerical experiment, we investigate the error decay of the greedy algorithm for a test case with fewer variables. Specifically, we fix the temperature and the particle diameter. For the training set $\mathcal{P}_\delta$, we use a uniform grid with 50 grid points in each parameter direction. Figure 6.12 visualizes the relative error decay of the maximal value of the error estimator during the greedy algorithm with respect

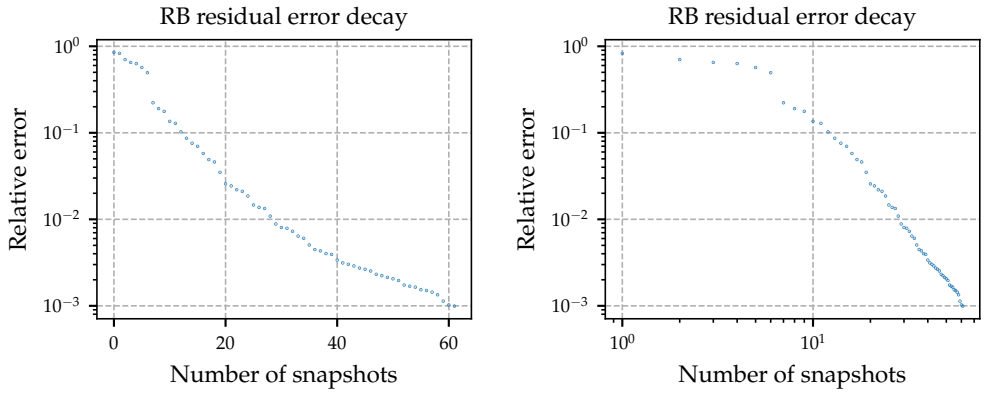


FIGURE 6.12: Relative error decay of the error estimator during the greedy algorithm with respect to the number of snapshots for the test case with excitation in three directions and parameters A_1, A_2, A_3 and D_H.

to the number of snapshots for this test case. We observe a faster decay of the error estimator compared to the test case that includes the temperature and particle diameter as variable parameters. This indicates that the approximation of the solution manifold is more complex when the temperature and particle diameter are included as variable parameters.

As demonstrated in Section 6.4.1, the HDG methods require higher refinement levels and polynomial degrees to yield stable approximations for certain parameters. Consequently, exploring a broader parameter space and improving accuracy within the RB method will demand more sophisticated techniques for the Fokker-Planck equation in the offline phase, a point we will discuss further in Section 7.2. Despite these limitations, this section successfully establishes the preliminary feasibility of the RB method for this space-time formulation. The results indicate that the RB approach achieves comparable accuracy to the finite element solution while significantly reducing online computational costs. Hence, these results are a promising starting point for model based image reconstruction methods for MPI.

Chapter 7

Conclusion and Outlook

In this work, we examined high-dimensional parabolic PDEs across two settings: high-dimensional spatial domains and parametric configurations. For both problems, we have developed efficient numerical methods. For both problems, we employed similar methodologies founded on space-time variational formulations and low-rank approximations.

In the following, we summarize the main contributions of this work and outline possible future research directions. These directions are discussed separately below for the two types of high-dimensional PDEs.

7.1 High-Dimensional Spatial Domain

For parabolic PDEs on high-dimensional spatial domains, we proposed a space-time adaptive low-rank method in Chapter 4 based on the adaptive low-rank method [BD16] for elliptic problems. Our approach combines a sparse wavelet expansion in time with adaptive low-rank tensor approximations in the spatial domain. A key contribution of this approach is the flexibility to adapt spatial ranks independently for each temporal basis function. The numerical method is a perturbed version of the Richardson iteration for the bi-infinite matrix-vector equation arising from the space-time variational formulation in Section 3.1. The condition number of the bi-infinite matrix is bounded independently of the spatial dimension d by an appropriate choice of Riesz bases, which requires the application of scaling matrices. To maintain the low-rank structure during the iteration, we constructed separable representations of these scaling matrices in Section 4.2 using exponential sum approximations. Since we have a bi-infinite matrix, we cannot apply the operator directly in the Richardson iteration. Instead, we apply the operator approximately with controlled accuracy, which is crucial for the convergence of the method. The application of the operator while controlling the support sizes, the ranks, and related quantities was analyzed in Section 4.3.

We showed in Section 4.4 that the method converges in the underlying space-time norm. Moreover, we proved under natural regularity assumptions on the solution and the operator that the method circumvents the curse of dimensionality. Numerical experiments on the high-dimensional heat equation in Section 4.5 validated these theoretical results, demonstrating the effectiveness of the method up to spatial dimension $d = 128$.

As stated in Theorem 4.32, the analysis of the method, in particular the adaptive application of the operator, can be extended to more general time-independent elliptic operators by combining the ideas of this work with the adaptive low-rank method for elliptic problems from [BD16]. In particular, we can investigate convection-diffusion problems. For convection-dominated problems, an interesting test would be to examine the distribution of ranks across the different time indices.

Another interesting direction is the extension to time-dependent elliptic operators $A(t)$ in (3.21). Then the corresponding operator $\mathbf{B}_X$ in (4.72) would no longer have the identity structure in time. An open question is whether the method can be extended by replacing the identity in time with a compressible operator. In particular, one would need to investigate the adaptive application of such an operator and the compressibility properties required for it.

Moreover, we can use the same approach for the variational formulations with alternative handling of the initial condition in Section 3.2.1 and the saddle point formulation in Section 3.2.2. In particular, the variational formulation in Section 3.2.1 is closely related to the one used in this work. The main difference is that the initial condition is not incorporated into the operator but is instead treated on the right-hand side and in the test space. However, much of the methodology in this work can be reused, for example, the adaptive application of the operator. Nevertheless, we have to adapt the basis of the test space accordingly. The variational formulation in Section 3.2.2 can be reformulated as a saddle-point problem. Iterative solvers for saddle point problems, such as the Uzawa iteration, have already been used in the context of adaptive wavelet methods in [DDU02]. Hence, one interesting question is whether we can propose an adaptive low-rank Uzawa iteration with properties similar to those of the perturbed Richardson iteration in Section 4.4.

In [AU20], an adaptive wavelet Galerkin method using hierarchical Tucker format was proposed for elliptic PDEs. This method is based on an adaptively refined Galerkin discretization instead of a Richardson iteration. In principle, this method can also be applied to the space-time variational formulation with some modifications. For low-dimensional problems, one major advantage of the adaptive wavelet Galerkin method over the perturbed Richardson iteration is that we can omit the coarsening step if we do not increase the index size in the refinement sets too much [GHS07]. However, for high-dimensional problems, this is not possible due to the product structure of the basis functions, which are crucial for the low-rank structure. Nevertheless, it would be interesting to compare the adaptive wavelet Galerkin low-rank method with the adaptive low-rank method based on the perturbed Richardson iteration.

7.2 Parametric Fokker-Planck Equation

In Chapter 5, we described the importance of accurate modeling in MPI using Brown and Néel relaxation mechanisms. This modeling leads to a representation of the mean magnetic moment as the expectation of the solution of a parametric Fokker-Planck equation. Accurate solutions of this equation for many parameters would enable model-based image reconstruction in MPI as an alternative to system-matrix calibration.

To be able to compute accurate solutions for many parameters, we introduced the general concept of RB methods for parametric PDEs in Section 6.1. Moreover, we presented the HDG method in Section 6.2, a finite element method well-suited for convection-dominated problems. In particular, we showed in Section 6.2.4 that these two methods can be applied to the parametric Fokker-Planck equation arising in MPI under minor modifications.

Since the parametric Fokker-Planck equation is posed on a manifold, we recalled some properties of finite element methods on surfaces in Section 6.3. Then we presented numerical results for the parametric Fokker-Planck equation related to MPI

in Section 6.4. We showed that we can compute accurate solutions for fixed parameters. Moreover, preliminary results for the RB method with coarser discretization showed that we can achieve good accuracy for many parameters with a small number of snapshots. This is a promising result for the application of the RB method to the parametric Fokker-Planck equation arising in MPI, as it significantly reduces computational costs up to a factor of $9 \cdot 10^5$, enabling model-based image reconstruction and parameter studies.

As seen in the numerical results, the computational costs for fixed parameter combinations grow rapidly with finer discretizations due to the 5-dimensional domain. One direction for future research is to improve the efficiency of the HDG method for the Fokker-Planck equation. A possible approach is to make use of the product structure of the domain and apply combination techniques [Gri92]. Another possibility to improve efficiency in the offline phase of the RB method is to compute multiple snapshots in parallel, as in [RUV24].

For the specific Fokker-Planck equation arising in MPI, not all assumptions of the HDG method are fulfilled. Moreover, we made minor modifications to the HDG method to make it compatible with the RB method, such as using Lax-Friedrichs fluxes. Hence, a rigorous analysis of the HDG method for the Fokker-Planck equation arising in MPI is still missing. In particular, it would be interesting to investigate the stability of the method. This would also give insights for the numerical results in Section 6.4.1, where we observed some stability issues for varying the anisotropy constant.

Moreover, as stated in Section 6.1, the RB method can be combined with every suitable variational formulation. Hence, a comparison of different variational formulations for the parametric Fokker-Planck equation in the context of RB methods would be interesting. A possible alternative is the robust Petrov-Galerkin formulation in [BS14], which is also well-suited for convection-dominated problems.

Furthermore, it would be interesting to extend the RB method to include all parameters of the parametric Fokker-Planck equation. In particular, the spatial dependence due to the static selection field $\vec{H}_S$, which is treated as a parameter, was not included in the numerical tests of the RB method yet. If all parameters are included, the modeling can also be used for parameter studies and for optimal experimental design in MPI.

Appendix A

Supplementary Proofs

In this appendix, we provide supplementary proofs of some statements that were omitted in the main text.

Proof of Theorem 2.29. Let $\mathbf{v} \in \ell_2(\hat{\mathbb{V}})$. Then by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\|\mathbf{M}\mathbf{v}\|^2 &= \sum_{\nu\in\mathbb{V}} \Big(\sum_{\mu\in\hat{\mathbb{V}}} m_{\nu,\mu} x_\mu \Big)^2 \\
&\leq \sum_{\nu\in\mathbb{V}} \Big(\sum_{\mu\in\hat{\mathbb{V}}} |m_{\nu,\mu}| \hat{\omega}_\mu \Big) \Big(\sum_{\mu\in\hat{\mathbb{V}}} |m_{\nu,\mu}| \hat{\omega}_\mu^{-1} x_\mu^2 \Big) \\
&\leq \sum_{\nu\in\mathbb{V}} \hat{c}\omega_\nu \sum_{\mu\in\hat{\mathbb{V}}} |m_{\nu,\mu}| \hat{\omega}_\mu^{-1} x_\mu^2 \\
&= \hat{c} \sum_{\mu\in\hat{\mathbb{V}}} \hat{\omega}_\mu^{-1} x_\mu^2 \sum_{\nu\in\mathbb{V}} |m_{\nu,\mu}| \omega_\nu \\
&\leq \hat{c} \sum_{\mu\in\hat{\mathbb{V}}} \hat{\omega}_\mu^{-1} x_\mu^2 c\hat{\omega}_\mu \\
&= c\hat{c}\|\mathbf{v}\|^2.
\end{aligned}$$

The statement then follows from the definition of the operator norm. □

Proof of Theorem 4.14. Statement (i) is a direct consequence of the product rule and the fundamental theorem of calculus.

The statement $0 \leq F_D(x)$ is a direct consequence of the definition of the Dawson function in (4.36). Combining this with (i) yields $F_D'(x) \leq 1$ for all $x \geq 0$. Since $F_D(0) = 0$, we obtain $F_D(x) \leq x$ for all $x \geq 0$.

Next, we prove (iii). We define $g(x) = 2F_D(x) - x$. Then by (i), we have

$$g'(x) = 1 - 4xF_D(x).$$

One can easily check that there exists a $y \in [0,1]$ with $g'(x) \geq 0$ for all $x \in [0,y]$ and $g'(x) \leq 0$ for all $x \in [y,1]$. Since $g(0) = 0$ as well as $g(1) = 2F_D(1) - 1 > 0$, we obtain $g(x) \geq 0$ for all $x \in [0,1]$, which proves the claim.

Now we consider (iv). By (i),we have

$$xF_D(x) = \tfrac{1}{2}(1 - F_D'(x)) \leq \tfrac{1}{2}(1 - F_D(x_{\min})) \leq 1,$$

where $x_{\min}$ is by [Was25] the global minimizer of F_D'.

Next, we prove (v). Let $x \geq 1$, then

$$\begin{aligned} 2xF_D(x) &= 2xe^{-x^2}\int_0^1 e^{t^2}\,\mathrm{d}t + 2xe^{-x^2}\int_1^x e^{t^2}\,\mathrm{d}t \\ &\geq xe^{-x^2}e + e^{-x^2}\int_1^x 2te^{t^2}\,\mathrm{d}t \\ &= xe^{-x^2}e + 1 - e^{-x^2}e \\ &\geq 1. \end{aligned}$$

Combining this with (i) as well as (iv) yields (v).

One can easily check by monotonicity that

$$\int_0^x e^{t^2}\,\mathrm{d}t \leq \frac{e^{x^2}-1}{x}, \quad x > 0.$$

Hence, combining this with (i), we obtain

$$F_D'(x) = 1 - 2xF_D(x) \geq 1 - 2(1-e^{-x^2}) \geq 0, \quad x \leq \sqrt{\log(2)}.$$

The lower bound is a direct consequence of (i) and (iv).

Statement (vii) can be easily shown by considering

$$h(x) = 4x(x^2-1)F_D(x) - (2x^2-1)$$

and showing $h(x) \to 0$ for $x \to \infty$ and $h'(x) < 0$ for $x > \frac{9}{4}$.

The proof of (viii) will be done only for the plus case. The case with the minus sign can be shown in the same way. Since e^{s^2} is holomorphic, the integral is path-independent. Hence, for $z = c(a+bi)$, we have

$$\begin{aligned} \int_0^z e^{\xi^2}\,\mathrm{d}\xi &= \int_0^{\operatorname{Im} z} e^{(iy)^2}\,\mathrm{d}y + \int_0^{\operatorname{Re} z} e^{(x+i\operatorname{Im}(z))^2}\,\mathrm{d}x \\ &= \frac{\sqrt{\pi}}{2}\operatorname{erf}(cb) + e^{-(cb)^2}\int_0^{ac} e^{x^2}e^{2icbx}\,\mathrm{d}x. \end{aligned}$$

Using this, as well as the triangle inequality, yields

$$\begin{aligned} |F_D(z)| &= \left|e^{-z^2}\right|\left|\int_0^z e^{\xi^2}\,\mathrm{d}\xi\right| \\ &\leq e^{-c^2(a^2-b^2)}\frac{\sqrt{\pi}}{2}\operatorname{erf}(cb) + F_D(ac). \end{aligned}$$

As a consequence,

$$\begin{aligned} |1-2zF_D(z)| &\leq 1 + 2|z||F_D(z)| \\ &\leq 1 + \sqrt{\pi(a^2+b^2)}ce^{-c^2(a^2-b^2)}\operatorname{erf}(cb) + 2c\sqrt{a^2+b^2}F_D(ac) \\ &\leq 1 + \sqrt{\tfrac{\pi(a^2+b^2)}{2e(a^2-b^2)}} + \tfrac{2\sqrt{a^2+b^2}}{a}x_{\min}F_D(x_{\min}), \end{aligned}$$

where we used in the last line that $F_D'(x) = 1 - 2xF_D(x)$ has a global minimum at $x_{\min}$ as well as $\operatorname{erf}(x) \leq 1$. □

Proof of Theorem 4.15. Combining Theorem 4.14 with the chain rule, we obtain

$$\frac{1}{\sqrt{\pi}\sqrt{y}}\Big(1-2\sqrt{ya}F_D\big(\sqrt{ya}\big)\Big) = \frac{2}{\sqrt{a\pi}}\frac{\mathrm{d}}{\mathrm{d}y}(F_D(\sqrt{ya}))$$

Hence, by integration by parts, we have

$$\begin{aligned}
\int_0^\infty \frac{1}{\sqrt{\pi y}}\Big(1-2\sqrt{ya}F_D\big(\sqrt{ya}\big)\Big)e^{-ys}\,\mathrm{d}y &= \frac{2}{\sqrt{a\pi}}\Big[F_D(\sqrt{ya})e^{-ys}\Big]_0^\infty \\
&\quad + \frac{2}{\sqrt{a\pi}}s\int_0^\infty F_D(\sqrt{ya})e^{-ys}\,\mathrm{d}y \\
&= \frac{2}{\sqrt{a\pi}}s\int_0^\infty e^{-ys}e^{-ya}\int_0^{\sqrt{ya}} e^{t^2}\,\mathrm{d}t\,\mathrm{d}y \\
&= \frac{2}{\sqrt{a\pi}}s\int_0^\infty \int_{\frac{t^2}{a}}^\infty e^{t^2}e^{-y(s+a)}\,\mathrm{d}y\,\mathrm{d}t \\
&= \frac{2}{\sqrt{a\pi}}s\int_0^\infty \frac{1}{s+a}e^{-\frac{s}{a}t^2}\,\mathrm{d}t \\
&= \frac{\sqrt{s}}{s+a},
\end{aligned}$$

where we used Fubini's theorem in the third line and the integration formula for the Gaussian in the last line. □

Proof of Theorem 4.74. The proof for $\|(\mathbf{I}_\mathrm{t}\otimes\bar{\mathbf{D}}_1^{-\tau})\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|$ was already given in Section 4.4.2.1. In this part we proof the estimate (4.167) for $(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_i)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})$.

We can proceed as in (4.169) to obtain

$$\begin{aligned}
\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\big\| &\le \big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\| \\
&\quad + \omega\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}\mathbf{w}_{k,j})\big\| \\
&\quad + \omega\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\big\|.
\end{aligned}$$

Again, we estimate the last two terms on the right-hand side separately. For the second term, similar to (4.170), we have

$$\begin{aligned}
&\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}\mathbf{w}_{k,j})\big\| \\
&\qquad\le \big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\tilde{\mathbf{A}}_{k,j}^{(1)}\mathbf{w}_{k,j}\big)\big\| + \big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j}\big)\big\|, \quad \text{(A.1)}
\end{aligned}$$

As in the other case, we again estimate these two terms separately. For the first term, the idea is to do a stepwise estimate as in (4.171), meaning

$$\begin{aligned}
\Big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{B}_1^\top}\Big(\mathbf{z}_{k,j},\tfrac{\eta_{k,j}^{(1)}}{4}\Big)\Big)\Big\| &\le \gamma_{\mathrm{t},\mathbf{B}_1}\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\big\|, \\
\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\big\| &\le \gamma_{\mathrm{t},\mathbf{B}_1}\big\|(\bar{\mathbf{D}}_\mathrm{t}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\big\|.
\end{aligned}$$

Since these two estimates can be shown in the same way, we only show the second one in detail. Using the definition of $\mathbf{z}_{k,j}$ and proceeding as in the other case, we

arrive at a result similar to (4.172) at

$$\begin{aligned}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{z}_{k,j}\big)\right\| \\ \leq \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| + \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| \quad \text{(A.2)}\end{aligned}$$

with $\tilde{\mathbf{w}}_{k,j} = \mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathrm{n}_1}\mathbf{w}_{k,j}$. This time, we start by estimating the second term on the right-hand side. By the definition of $\tilde{\mathbf{T}}_{\mathrm{x}}$ in Section 4.3.1, we have only restrictions in the temporal variable. Hence, we obtain

$$(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}} = \mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}}),$$

where we also used the fact that diagonal operators commute. Hence, we arrive at

$$\begin{aligned}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| &\leq \|\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\|\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathrm{n}_1}\mathbf{w}_{k,j}\right\| \\ &\leq 2\|\mathbf{D}_{\mathcal{Y}}\mathbf{T}_{\mathrm{x}}\mathbf{D}_{\mathcal{X}}\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{w}_{k,j}\| \\ &\leq 2C_\delta\|\bar{\mathbf{B}}_{\mathrm{x}}\|\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\|,\end{aligned}$$

where we used that diagonal operators commute and (4.56) in the second line, that $\tilde{\mathbf{T}}_{\mathrm{x}}$ is an approximation of $\mathbf{T}_{\mathrm{x}}$ in the third line as well as (4.57) and (4.65) in the last line. Next, we consider the second term in (A.2). Let $J = J(\eta;\mathbf{w}_{k,j})$ for $\eta = \frac{\eta_{k,j}^{(1)}}{4\|\mathbf{B}_1\|}$ and $\bar{\Lambda}_{-1}^{(\mathrm{t})} = \varnothing$. By definition of $\tilde{\mathbf{T}}_{\mathrm{t}}$ in (4.91), we have

$$\begin{aligned}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}\tilde{\mathbf{T}}_{\mathrm{t}}\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{w}}_{k,j}\right\| &\leq \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}(\mathbf{T}_{\mathrm{t}}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}}\tilde{\mathbf{w}}_{k,j}\right\| \\ &\quad + \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}((\mathbf{T}_{\mathrm{t}}\otimes\mathbf{I}_{\mathrm{x}})-\tilde{\mathbf{T}}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}}\tilde{\mathbf{w}}_{k,j}\right\|\end{aligned}$$

since $\tilde{\mathbf{T}}_{\mathrm{t}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}} = \tilde{\mathbf{T}}_{\mathrm{t}}$ and diagonal operators commute. For the first term on the right-hand side, we have

$$\begin{aligned}&\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}(\mathbf{T}_{\mathrm{t}}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\right\| \\ &\quad = C_\delta\left\|\mathbf{D}_{\mathcal{Y}}(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\mathbf{T}_{\mathrm{t}}\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\tilde{\mathbf{w}}_{k,j}\right\| \\ &\quad \leq C_\delta\|\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\mathbf{T}_{\mathrm{t}}\bar{\mathbf{D}}_{\mathrm{t}}^{1+\tau}\|\left\|\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\right\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\tilde{\mathbf{w}}_{k,j}\| \\ &\quad \leq C_\delta\zeta_{\mathbf{B}}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\|,\end{aligned}$$

where we use (4.68) from Theorem 4.29 in the second line as well as (4.166) and (4.56) in the last line. Furthermore, for the second term, we have

$$((\mathbf{T}_{\mathrm{t}}\otimes\mathbf{I}_{\mathrm{x}})-\tilde{\mathbf{T}}_{\mathrm{x}})\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})} = \sum_{j=0}^{J}((\mathbf{T}_{\mathrm{t}}-\mathbf{T}_{\mathrm{t},J-j})\otimes\mathbf{I}_{\mathrm{x}})\Big(\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{w}_{k,j})}-\mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{w}_{k,j})}\Big).$$

Using the triangle inequality, we thus obtain

$$\begin{aligned}
&\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{Y}}((\mathbf{T}_{\mathrm{t}}\otimes\mathbf{I}_{\mathrm{x}})-\tilde{\mathbf{T}}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}}\tilde{\mathbf{w}}_{k,j}\right\| \\
&\quad\le\sum_{j=0}^{J}\left\|\mathbf{D}_{\mathcal{Y}}(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}(\mathbf{T}_{\mathrm{t}}-\mathbf{T}_{\mathrm{t},J-j})\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\right\|\left\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{w}_{k,j})}-\mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{w}_{k,j})}\right\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\tilde{\mathbf{w}}_{k,j}\| \\
&\quad\le\sum_{j=0}^{J}C_\delta\left\|\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}(\mathbf{T}_{\mathrm{t}}-\mathbf{T}_{\mathrm{t},J-j})\bar{\mathbf{D}}_{\mathrm{t}}^{1+\tau}\right\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{w}_{k,j}\| \\
&\quad\le\sum_{j=0}^{J}C_\delta C_{\tau,\mathrm{t}}\beta_{J-j}(\mathbf{C}_{\mathrm{t}})\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\| \\
&\quad\le C_\delta C_{\tau,\mathrm{t}}\|\beta(\mathbf{C}_{\mathrm{t}})\|_{\ell_1}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\|,
\end{aligned}$$

where we used (4.68) and (4.56) in the third line and Assumption 1 (ii) in the fourth line.

Combining all these estimates, we arrive at

$$\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{z}_{k,j})\right\|\le\gamma_{\mathrm{t},\mathbf{B}_1}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\|$$

In the next step, we will estimate the second term on the right-hand side in (A.1), i.e.,

$$\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\right\|=\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\Big(\mathrm{APPLY}_{\mathbf{A}_2}\Big(\mathbf{w}_{k,j},\frac{\eta_{k,j}^{(2)}}{2}\Big)\Big)\right\|$$

For this part, we can proceed similarly to $\mathbf{T}_{\mathrm{t}}$ using the corresponding assumption for higher-regularity of $\mathbf{A}_2$ from Assumption 1. For $J=J_0(\eta;\mathbf{w}_{k,j})$ with $\eta=\frac{\eta_{k,j}^{(2)}}{4}$, we have by construction of $\tilde{\mathbf{S}}_{0,k,j}$ in (4.130) and by (4.56) the estimate

$$\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\right\|\le\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{S}}_{0,k,j}\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\right\|$$

with $\tilde{\mathbf{w}}_{k,j}=\mathbf{D}_{\mathcal{X}}^{-1}\mathbf{D}_{\mathcal{X},\mathrm{n}}\mathbf{w}_{k,j}$. By application of the triangle inequality, we arrive at

$$\begin{aligned}
\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j})\right\|&\le\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\right\| \\
&\quad+\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}(\mathbf{S}_0-\tilde{\mathbf{S}}_{0,k,j})\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\right\|.
\end{aligned}$$

Again, we treat both terms on the right-hand side independently. For the first term on the right-hand side, we have

$$\begin{aligned}
&\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\right\| \\
&\quad\le\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}(\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\|\left\|\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\right\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\tilde{\mathbf{w}}_{k,j}\| \\
&\quad\le C_\delta\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathbf{S}_0\mathbf{D}_{\mathcal{X}}(\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{w}_{k,j}\| \\
&\quad\le C_\delta\zeta_{\mathbf{B}}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\right\|,
\end{aligned}$$

where we used (4.57), (4.65) and (4.56) in the third line as well as (4.166) in the last line. For the second term, we have

$$(\mathbf{S}_0 - \tilde{\mathbf{S}}_{0,k,j})\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})} = \sum_{j=0}^{J}(\mathbf{S}_0 - \mathbf{S}_{0,J-j})\Big(\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{w}_{k,j})} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{w}_{k,j})}\Big).$$

Combining this with the triangle inequality, we obtain

$$\begin{aligned}
&\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}\mathcal{X}(\mathbf{S}_0 - \tilde{\mathbf{S}}_{0,k,j})\mathbf{D}\mathcal{X}\mathrm{R}_{\bar{\Lambda}_J^{(\mathrm{t})}(\mathbf{w}_{k,j})}\tilde{\mathbf{w}}_{k,j}\| \\
&\quad\le \sum_{j=0}^{J}\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}\mathcal{X}(\mathbf{S}_0 - \mathbf{S}_{0,J-j})\mathbf{D}\mathcal{X}(\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\|\Big\|\mathrm{R}_{\bar{\Lambda}_j^{(\mathrm{t})}(\mathbf{w}_{k,j})} - \mathrm{R}_{\bar{\Lambda}_{j-1}^{(\mathrm{t})}(\mathbf{w}_{k,j})}\Big\| \\
&\qquad\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\tilde{\mathbf{w}}_{k,j}\| \\
&\quad\le \sum_{j=0}^{J} C_\delta\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\bar{\mathbf{D}}\mathcal{X}(\mathbf{S}_0 - \mathbf{S}_{0,J-j})\bar{\mathbf{D}}\mathcal{X}(\bar{\mathbf{D}}_{\mathrm{t}}^{\tau}\otimes\mathbf{I}_{\mathrm{x}})\|\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{w}_{k,j}\| \\
&\quad\le \sum_{j=0}^{J} C_\delta C_{\tau,0}\beta_{J-j}^{(0)}\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\| \\
&\quad\le C_\delta C_{\tau,0}\|\beta^{(0)}\|_{\ell_1}\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|,
\end{aligned}$$

where we used (4.57), (4.65) and (4.56) in the third line as well as Assumption 1 (iii) in the fourth line. Combining these two estimates, we arrive at

$$\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\tilde{\mathbf{A}}_{k,j}^{(2)}\mathbf{w}_{k,j}\big)\| \le \gamma_{\mathrm{t},\mathbf{A}_2}\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\|.$$

Inserting this into (A.1), we obtain

$$\begin{aligned}
&\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j+1})\| \\
&\qquad\le \gamma\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{w}_{k,j})\| + \omega\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\|. \quad \text{(A.3)}
\end{aligned}$$

Hence, it remains open to estimate $\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\|$. As in (4.174) we can split this term into

$$\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j})\| \le \|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}^{(1)}\big)\| + \|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}^{(2)}\big)\|, \quad \text{(A.4)}$$

where $\mathbf{h}_{k,j}^{(1)}$ and $\mathbf{h}_{k,j}^{(2)}$ are defined in (4.175). Since $\mathbf{h}_{k,j}^{(1)}$ is defined by the adaptive application of $\mathbf{B}_1^\top$ we can proceed as before and obtain

$$\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}^{(1)}\big)\| \le \gamma_{\mathrm{t},\mathbf{B}_1}\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{f}_{k,j}^{(1)}\big)\|.$$

Combining this with Assumption 2 (iii) and (4.165) as well as (4.65), we arrive at

$$\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}^{(1)}\big)\| \le (1+\tfrac{\delta}{2})\gamma_{\mathrm{t},\mathbf{B}_1}C_{\mathbf{f}}^{\mathrm{reg}}\zeta_{\mathbf{f}}.$$

Finally, we need to estimate $\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}(\mathbf{h}_{k,j}^{(2)})\|$. By definition of the adaptive application of $\mathbf{B}_2^\top$, we have $\mathbf{h}_{k,j}^{(2)} = \mathbf{D}\mathcal{X}_{,\mathrm{n}}\tilde{\mathbf{T}}_{0,k,j}^{\mathrm{ad}}\mathbf{f}_{k,j}^{(2)}$ where

$$\eta = \tfrac{\eta_{k,j}^{(2)}}{4}, \quad \tilde{\mathbf{T}}_{0,k,j}^{\mathrm{ad}} = \mathbf{T}_{0,J_0^{\mathrm{ad}}(\eta;\mathbf{f}_{k,j}^{(2)}),\mathbf{f}_{k,j}^{(2)}}^{\top}.$$

Then we obtain

$$\begin{aligned}\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}^{(2)}\big)\right\| &= \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{h}_{k,j}^{(2)}\right\| \\ &= \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X},\mathrm{n}}\tilde{\mathbf{T}}_{0,k,j}^{\mathrm{ad}}\mathbf{f}_{k,j}^{(2)}\right\| \\ &\leq \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\tilde{\mathbf{T}}_{0,k,j}^{\mathrm{ad}}\mathbf{f}_{k,j}^{(2)}\right\| \\ &\leq \left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_{\mathrm{x}})\mathbf{D}_{\mathcal{X}}\mathbf{T}_0^{\top}\mathbf{f}_{k,j}^{(2)}\right\| \\ &\leq \sqrt{2}(1+\tfrac{\delta}{2})\left\|\bar{\mathbf{D}}^{-2\tau}\mathbf{f}_{k,j}^{(2)}\right\| \\ &\leq \sqrt{2}(1+\tfrac{\delta}{2})C_{\mathbf{f}}^{\mathrm{reg}}\zeta_{\mathbf{f}},\end{aligned}$$

where we used (4.56) in the third line, the definition of $\tilde{\mathbf{T}}_{0,k,j}^{\mathrm{ad}}$ from (4.137) as restriction in the fourth line, Theorem 4.73 in fifth line as well as Assumption 2 (iii) in the last line.

Inserting these estimates into (A.4) yields

$$\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{h}_{k,j}\big)\right\| \leq \omega^{-1}\bar{C}_{\mathbf{f}}.$$

Further inserting this into (A.3), we arrive at

$$\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{w}_{k,j+1}\big)\right\| \leq \gamma\left\|(\bar{\mathbf{D}}_{\mathrm{t}}^{-\tau}\otimes\mathbf{I}_1)\pi^{(\mathrm{t},i)}\big(\mathbf{w}_{k,j}\big)\right\| + \bar{C}_{\mathbf{f}}.$$

From here on, we can proceed exactly as in the other case, using the above recursion formula. This concludes the proof. □

List of Abbreviations

HDG	**H**ybrid **D**iscontinuous **G**alerkin
HSVD	**H**ierachical **S**ingular **V**alue **D**ecomposition
MPI	**M**agnetic **P**article **I**maging
PDE	**P**artial **D**ifferential **E**quation
RB	**R**educed **B**asis
SDE	**S**tochastic **D**ifferential **E**quation
SVD	**S**ingular **V**alue **D**ecomposition

List of Figures

2.1 DGH multiwavelets . . . 11
2.2 Canonical polyadic format . . . 17
2.3 Tucker format . . . 19
2.4 Binary dimension tree in $4d$ . . . 20
2.5 Hierarchical tensor representation in $4d$ . . . 21
2.6 Linear dimension trees . . . 22

4.1 Dawson function . . . 59
4.2 Function f_c related to Dawson function . . . 61
4.3 Exponential-sum approximation . . . 63
4.4 Visualization of inexact Richardson iteration . . . 108
4.5 Residual bounds for heat equation with source term . . . 137
4.6 Support and rank versus residual norm estimate for heat equation with source term . . . 138
4.7 Support size for heat equation with source term . . . 139
4.8 Ranks per time index for heat equation with source term . . . 140
4.9 Support and rank frequencies for heat equation with source term in $d = 4$ . . . 141
4.10 Support and rank frequencies for heat equation with source term in $d = 8$ . . . 141
4.11 Time per Richardson step for heat equation with source term . . . 142
4.12 Residual bounds for heat equation with initial condition . . . 143
4.13 Support and rank versus residual norm estimate for heat equation with initial condition . . . 144
4.14 Support and rank frequencies for heat equation with initial condition in $d = 4$ . . . 144
4.15 Support and rank frequencies for heat equation with initial condition in $d = 8$ . . . 144
4.16 Support and rank frequencies for heat equation with initial condition in $d = 16$ . . . 145
4.17 Support size for heat equation with initial condition . . . 146
4.18 Ranks per time index for heat equation with initial condition . . . 147
4.19 Time per Richardson step for heat equation with initial condition . . . 147

5.1 Magnetic field . . . 150
5.2 Brown and Néel rotation . . . 153

6.1 Planar and spherical triangulation of the sphere . . . 183
6.2 Mapping reference to spherical triangle . . . 184
6.3 Refinement levels for triangulation of the sphere . . . 190
6.4 Illustration flux in time direction . . . 191
6.5 Mean magnetic moment for different values of saturation magnetization 193
6.6 Mean magnetic moment for different values of particle core diameter 194

6.7 Mean magnetic moment for different values of anisotropy constant . 194
6.8 Relative error decay during greedy algorithm with excitation in one direction 196
6.9 RB solution and difference to finite element solution with excitation in one direction 197
6.10 Relative error decay during greedy algorithm with excitation in three directions 198
6.11 RB solution and difference to finite element solution with excitation in three directions 199
6.12 Relative error decay during greedy algorithm with excitation in three directions with fewer parameters 200

List of Tables

5.1 Physical parameters in MPI 152

6.1 Physical parameters for numerical experiments with fixed parameter 192
6.2 Parameter of the HDG method for the numerical experiments with fixed parameters 192
6.3 Physical parameters for RB numerical experiments 196

Bibliography

[AB85] Douglas. N. Arnold and Franco Brezzi. "Mixed and Nonconforming Finite Element Methods: Implementation, Postprocessing and Error Estimates". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 19.1 (1985), pp. 7–32. ISSN: 0764-583X, 1290-3841. DOI: 10.1051/m2an/1985190100071.

[AKK20] Hannes Albers, Tobias Kluth, and Tobias Knopp. *A Simulation Framework for Particle Magnetization Dynamics of Large Ensembles of Single Domain Particles: Numerical Treatment of Brown/Néel Dynamics and Parameter Identification Problems in Magnetic Particle Imaging*. 2020. DOI: 10.48550/ARXIV.2010.07772.

[And12] Roman Andreev. "Stability of Space-Time Petrov-Galerkin Discretizations for Parabolic Evolution Equations". PhD thesis. ETH Zurich, 2012. DOI: 10.3929/ETHZ-A-007563932. HDL: 20.500.11850/59269.

[And16] Roman Andreev. "Wavelet-In-Time Multigrid-In-Space Preconditioning of Parabolic Evolution Equations". In: *SIAM Journal on Scientific Computing* 38.1 (Jan. 2016), A216–A242. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/140998639.

[And+99] Eward Anderson et al. *LAPACK Users' Guide*. Third. Society for Industrial and Applied Mathematics, Jan. 1999. ISBN: 978-0-89871-447-0 978-0-89871-960-4. DOI: 10.1137/1.9780898719604.

[AS13] Milton Abramowitz and Irene A. Stegun, eds. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. 9. Dover print.; [Nachdr. der Ausg. von 1972]. Dover Books on Mathematics. New York, NY: Dover Publ, 2013. ISBN: 978-0-486-61272-0.

[AT15] Roman Andreev and Christine Tobler. "Multilevel Preconditioning and Low-rank Tensor Iteration for Space–Time Simultaneous Discretizations of Parabolic PDEs". In: *Numerical Linear Algebra with Applications* 22.2 (Mar. 2015), pp. 317–337. ISSN: 1070-5325, 1099-1506. DOI: 10.1002/nla.1951.

[AU20] Mazen Ali and Karsten Urban. "HT-AWGM: A Hierarchical Tucker–Adaptive Wavelet Galerkin Method for High-Dimensional Elliptic Problems". In: *Advances in Computational Mathematics* 46.4 (Aug. 2020), p. 59. ISSN: 1019-7168, 1572-9044. DOI: 10.1007/s10444-020-09797-9.

[Azi+25] Ebrahim Azizi, Shahriar Mostufa, Bahareh Rezaei, Shumin Xin, Jiajia Sun, Zongqian Shi, Jenifer Gómez-Pastora, and Kai Wu. "Applying Stochastic Langevin Function with Coupled Brownian–Néel Relaxations to Study the Dynamic Magnetizations of Nanoparticle Tracers in Magnetic Particle Imaging". In: *Journal of Physics D: Applied Physics* 58.13 (Mar. 2025), p. 135002. ISSN: 0022-3727, 1361-6463. DOI: 10.1088/1361-6463/adafbd.

[Bac12] Markus Bachmayr. "Adaptive Low-Rank Wavelet Methods and Applications to Two-Electron Schrödinger Equations". PhD thesis. RWTH Aachen, 2012.

[Bac+21] Markus Bachmayr, Henrik Eisenmann, Emil Kieri, and André Uschmajew. "Existence of Dynamical Low-Rank Approximations to Parabolic Problems". In: *Mathematics of Computation* 90.330 (Mar. 2021), pp. 1799–1830. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/mcom/3626.

[Bac23] Markus Bachmayr. "Low-Rank Tensor Methods for Partial Differential Equations". In: *Acta Numerica* 32 (May 2023), pp. 1–121. ISSN: 0962-4929, 1474-0508. DOI: 10.1017/s0962492922000125.

[Bar+04] Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T. Patera. "An 'Empirical Interpolation' Method: Application to Efficient Reduced-Basis Discretization of Partial Differential Equations". In: *Comptes Rendus. Mathématique* 339.9 (Oct. 2004), pp. 667–672. ISSN: 1631-073X, 1778-3569. DOI: 10.1016/j.crma.2004.08.006.

[Bar05] Arne Barinka. "Fast Computation Tools for Adaptive Wavelet Schemes". PhD thesis. RWTH Aachen, 2005.

[BD15] Markus Bachmayr and Wolfgang Dahmen. "Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations". In: *Foundations of Computational Mathematics* 15.4 (Aug. 2015), pp. 839–898. ISSN: 1615-3375, 1615-3383. DOI: 10.1007/s10208-013-9187-3.

[BD16] Markus Bachmayr and Wolfgang Dahmen. "Adaptive Low-Rank Methods: Problems on Sobolev Spaces". In: *SIAM Journal on Numerical Analysis* 54.2 (Jan. 2016), pp. 744–796. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/140978223.

[BDS17] Dirk Broersen, Wolfgang Dahmen, and Rob Stevenson. "On the Stability of DPG Formulations of Transport Equations". In: *Mathematics of Computation* 87.311 (Sept. 2017), pp. 1051–1082. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/mcom/3242.

[BDS25] Markus Bachmayr, Matthieu Dolbeault, and Polina Sachsenmaier. *Iterative Thresholding Low-Rank Time Integration*. 2025. DOI: 10.48550/ARXIV.2507.15848.

[BE76] Haïm Brezis and Ivar Ekeland. "Un Principe Variationnel Associ6~ Certaines 6quations Paraboliques". In: *CR Acad. Sci. Paris, S6r. A–B* 282 (1976).

[BEU25] Markus Bachmayr, Henrik Eisenmann, and André Uschmajew. "Dynamical Low-Rank Tensor Approximations to High-Dimensional Parabolic Problems: Existence and Convergence of Spatial Discretizations". In: *Numerische Mathematik* 157.3 (June 2025), pp. 781–822. ISSN: 0029-599X, 0945-3245. DOI: 10.1007/s00211-025-01465-8.

[Bez+17] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. "Julia: A Fresh Approach to Numerical Computing". In: *SIAM Review* 59.1 (Jan. 2017), pp. 65–98. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/141000671.

[BF24] Markus Bachmayr and Manfred Faldum. "A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic Partial Differential Equations". In: *Journal of Complexity* 82 (June 2024), p. 101839. ISSN: 0885064X. DOI: 10.1016/j.jco.2024.101839.

[BF91] Franco Brezzi and Michel Fortin, eds. *Mixed and Hybrid Finite Element Methods*. Vol. 15. Springer Series in Computational Mathematics. New York, NY: Springer New York, 1991. ISBN: 978-1-4612-7824-5 978-1-4612-3172-1. DOI: 10.1007/978-1-4612-3172-1.

[BG04] Hans-Joachim Bungartz and Michael Griebel. "Sparse Grids". In: *Acta Numerica* 13 (May 2004), pp. 147–269. ISSN: 0962-4929, 1474-0508. DOI: 10.1017/S0962492904000182.

[BJ89] Ivo Babuska and Tadeusz Janik. "The H-p Version of the Finite Element Method for Parabolic Equations. Part I. The P-version in Time". In: *Numerical Methods for Partial Differential Equations* 5.4 (Dec. 1989), pp. 363–399. ISSN: 0749-159X, 1098-2426. DOI: 10.1002/num.1690050407.

[Boi+19] Thomas Boiveau, Virginie Ehrlacher, Alexandre Ern, and Anthony Nouy. "Low-Rank Approximation of Linear Parabolic Equations by Space-Time Tensor Galerkin Methods". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 53.2 (Mar. 2019), pp. 635–658. ISSN: 0764-583X, 1290-3841. DOI: 10.1051/m2an/2018073.

[BS14] Dirk Broersen and Rob Stevenson. "A Robust Petrov–Galerkin Discretisation of Convection–Diffusion Equations". In: *Computers & Mathematics with Applications* 68.11 (Dec. 2014), pp. 1605–1618. ISSN: 08981221. DOI: 10.1016/j.camwa.2014.06.019.

[BS17] Markus Bachmayr and Reinhold Schneider. "Iterative Methods Based on Soft Thresholding of Hierarchical Tensors". In: *Foundations of Computational Mathematics* 17.4 (Aug. 2017), pp. 1037–1083. ISSN: 1615-3375, 1615-3383. DOI: 10.1007/s10208-016-9314-z.

[Buf+12] Annalisa Buffa, Yvon Maday, Anthony Patera, Christophe Prud'homme, and Gabriel Turinici. "*A Priori* Convergence of the Greedy Algorithm for the Parametrized Reduced Basis Method". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 46.3 (May 2012), pp. 595–603. ISSN: 0764-583X, 1290-3841. DOI: 10.1051/m2an/2011056.

[Cas+16] Douglas A. Castro, Philippe R.B. Devloo, Agnaldo M. Farias, Sônia M. Gomes, and Omar Durán. "Hierarchical High Order Finite Element Bases for H(Div) Spaces Based on Curved Meshes for Two-Dimensional Regions or Manifolds". In: *Journal of Computational and Applied Mathematics* 301 (Aug. 2016), pp. 241–258. ISSN: 03770427. DOI: 10.1016/j.cam.2016.01.053.

[CDD01] Albert Cohen, Wolfgang Dahmen, and Ronald DeVore. "Adaptive Wavelet Methods for Elliptic Operator Equations: Convergence Rates". In: *Mathematics of Computation* 70.233 (Jan. 2001), pp. 27–75. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/S0025-5718-00-01252-7.

[CDD02] Albert Cohen, Wolfgang Dahmen, and Ronald DeVore. "Adaptive Wavelet Methods II—Beyond the Elliptic Case". In: *Foundations of Computational Mathematics* 2.3 (Aug. 2002), pp. 203–202. ISSN: 16153375. DOI: 10.1007/s102080010027.

[CDV93] Albert Cohen, Ingrid Daubechies, and Pierre Vial. "Wavelets on the Interval and Fast Wavelet Transforms". In: *Applied and Computational Harmonic Analysis* 1.1 (Dec. 1993), pp. 54–81. ISSN: 10635203. DOI: 10.1006/acha.1993.1005.

[Coh03] Albert Cohen. *Numerical Analysis of Wavelet Methods*. 1st ed. Studies in Mathematics and Its Applications v. 32. Amsterdam Boston: Elsevier, 2003. ISBN: 978-0-444-51124-9.

[Con20] Dajana Conte. "Dynamical Low-Rank Approximation to the Solution of Parabolic Differential Equations". In: *Applied Numerical Mathematics* 156 (Oct. 2020), pp. 377–384. ISSN: 01689274. DOI: 10.1016/j.apnum.2020.05.011.

[CQ92] Charles K. Chui and Ewald Quak. "Wavelets on a Bounded Interval". In: *Numerical Methods in Approximation Theory, Vol. 9*. Ed. by Dietrich Braess and Larry L. Schumaker. Basel: Birkhäuser Basel, 1992, pp. 53–75. ISBN: 978-3-0348-9702-0 978-3-0348-8619-2. DOI: 10.1007/978-3-0348-8619-2_4.

[CS11] Nabi Chegini and Rob Stevenson. "Adaptive Wavelet Schemes for Parabolic Problems: Sparse Matrices and Numerical Results". In: *SIAM Journal on Numerical Analysis* 49.1 (Jan. 2011), pp. 182–212. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/100800555.

[CVK16] Heyrim Cho, Daniele Venturi, and George E. Karniadakis. "Numerical Methods for High-Dimensional Probability Density Function Equations". In: *Journal of Computational Physics* 305 (Jan. 2016), pp. 817–837. ISSN: 00219991. DOI: 10.1016/j.jcp.2015.10.030.

[Dah01] Wolfgang Dahmen. "Wavelet Methods for PDEs — Some Recent Developments". In: *Journal of Computational and Applied Mathematics* 128.1-2 (Mar. 2001), pp. 133–185. ISSN: 03770427. DOI: 10.1016/S0377-0427(00)00511-2.

[Dau88] Ingrid Daubechies. "Orthonormal Bases of Compactly Supported Wavelets". In: *Communications on Pure and Applied Mathematics* 41.7 (Oct. 1988), pp. 909–996. ISSN: 0010-3640, 1097-0312. DOI: 10.1002/cpa.3160410705.

[DDU02] Stephan Dahlke, Wolfgang Dahmen, and Karsten Urban. "Adaptive Wavelet Methods for Saddle Point Problems—Optimal Convergence Rates". In: *SIAM Journal on Numerical Analysis* 40.4 (Jan. 2002), pp. 1230–1262. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/S003614290139233X.

[DE13] Gerhard Dziuk and Charles M. Elliott. "Finite Element Methods for Surface PDEs". In: *Acta Numerica* 22 (May 2013), pp. 289–396. ISSN: 0962-4929, 1474-0508. DOI: 10.1017/S0962492913000056.

[DGH96] George C. Donovan, Jeffrey S. Geronimo, and Douglas P. Hardin. "Intertwining Multiresolution Analyses and the Construction of Piecewise-Polynomial Wavelets". In: *SIAM Journal on Mathematical Analysis* 27.6 (Nov. 1996), pp. 1791–1815. ISSN: 0036-1410, 1095-7154. DOI: 10.1137/S0036141094276160.

[DGH99] George. C. Donovan, Jeffrey S. Geronimo, and Douglas P. Hardin. "Orthogonal Polynomials and the Construction of Piecewise Polynomial Smooth Wavelets". In: *SIAM Journal on Mathematical Analysis* 30.5 (Jan. 1999), pp. 1029–1056. ISSN: 0036-1410, 1095-7154. DOI: 10.1137/S0036141096313112.

[DKO12] Sergey. V. Dolgov, Boris N. Khoromskij, and Ivan. V. Oseledets. "Fast Solution of Parabolic Problems in the Tensor Train/Quantized Tensor Train Format with Initial Application to the Fokker–Planck Equation". In: *SIAM Journal on Scientific Computing* 34.6 (Jan. 2012), A3016–A3038. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/120864210.

[DKU99] Wolfgang Dahmen, Angela Kunoth, and Karsten Urban. "Biorthogonal Spline Wavelets on the Interval—Stability and Moment Conditions". In: *Applied and Computational Harmonic Analysis* 6.2 (Mar. 1999), pp. 132–196. ISSN: 10635203. DOI: 10.1006/acha.1998.0247.

[DL00] Robert Dautray and Jacques-Louis Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 5: Evolution Problems I*. Ed. by Ian N. Sneddon. Trans. by Alan Craig. Berlin, [Germany] ; Heidelberg, [Germany]: Springer, 2000. ISBN: 978-3-540-66101-6 978-3-642-58090-1.

[DSS09] Tammo Jan Dijkema, Christoph Schwab, and Rob Stevenson. "An Adaptive Wavelet Method for Solving High-Dimensional Elliptic PDEs". In: *Constructive Approximation* 30.3 (Dec. 2009), pp. 423–455. ISSN: 0176-4276, 1432-0940. DOI: 10.1007/s00365-009-9064-0.

[EG21] Alexandre Ern and Jean-Luc Guermond. *Finite Elements I: Approximation and Interpolation*. Vol. 72. Texts in Applied Mathematics. Cham: Springer International Publishing, 2021. ISBN: 978-3-030-56340-0 978-3-030-56341-7. DOI: 10.1007/978-3-030-56341-7.

[ES10] Herbert Egger and Joachim Schoberl. "A Hybrid Mixed Discontinuous Galerkin Finite-Element Method for Convection-Diffusion Problems". In: *IMA Journal of Numerical Analysis* 30.4 (Oct. 2010), pp. 1206–1234. ISSN: 0272-4979, 1464-3642. DOI: 10.1093/imanum/drn083.

[Fal25] Manfred Faldum. "A Space-Time Adaptive Low-Rank Method for High-Dimensional Parabolic PDEs". In: *Oberwolfach Reports*. Ed. by Markus Bachmayr, Albert Cohen, Angela Kunoth, and Olga Mula. Vol. 37. Computation and Learning in High Dimensions. 2025, pp. 2021–2023. DOI: 10.4171/OWR/2025/37.

[Fal+25] Manfred Faldum, Markus Bachmayr, Volkmar Schulz, and Franziska Schrank. "Efficient Solvers for Coupled Brown-Néel Fokker-Planck Equations". In: *International Journal on Magnetic Particle Imaging IJMPI* 11.1 Suppl 1 (Mar. 2025). DOI: 10.18416/IJMPI.2025.2503011.

[GH11] Lars Grasedyck and Wolfgang Hackbusch. "An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples". In: *Computational Methods in Applied Mathematics* 11.3 (2011), pp. 291–304. ISSN: 1609-9389, 1609-4840. DOI: 10.2478/cmam-2011-0016.

[GHM94] Jeffrey S. Geronimo, Douglas P. Hardin, and Peter.R. Massopust. "Fractal Functions and Wavelet Expansions Based on Several Scaling Functions". In: *Journal of Approximation Theory* 78.3 (Sept. 1994), pp. 373–401. ISSN: 00219045. DOI: 10.1006/jath.1994.1085.

[GHS07] Tsogtgerel Gantumur, Helmut Harbrecht, and Rob Stevenson. "An Optimal Adaptive Wavelet Method without Coarsening of the Iterands". In: *Mathematics of Computation* 76.258 (Sept. 2007), pp. 615–629. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/S0025-5718-06-01917-X.

[Gli11] Yuri E. Gliklikh. *Global and Stochastic Analysis with Applications to Mathematical Physics*. Theoretical and Mathematical Physics. London: Springer London, 2011. ISBN: 978-0-85729-162-2 978-0-85729-163-9. DOI: 10.1007/978-0-85729-163-9.

[GMU17] Silke Glas, Antonia Mayerhofer, and Karsten Urban. "Two Ways to Treat Time in Reduced Basis Methods". In: *Model Reduction of Parametrized Systems*. Ed. by Peter Benner, Mario Ohlberger, Anthony Patera, Gianluigi Rozza, and Karsten Urban. Vol. 17. Cham: Springer International Publishing, 2017, pp. 1–16. ISBN: 978-3-319-58785-1 978-3-319-58786-8. DOI: 10.1007/978-3-319-58786-8_1.

[Gra10] Lars Grasedyck. "Hierarchical Singular Value Decomposition of Tensors". In: *SIAM Journal on Matrix Analysis and Applications* 31.4 (Jan. 2010), pp. 2029–2054. ISSN: 0895-4798, 1095-7162. DOI: 10.1137/090764189.

[Gre+07] Martin A. Grepl, Yvon Maday, Ngoc C. Nguyen, and Anthony T. Patera. "Efficient Reduced-Basis Treatment of Nonaffine and Nonlinear Partial Differential Equations". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 41.3 (May 2007), pp. 575–605. ISSN: 0764-583X, 1290-3841. DOI: 10.1051/m2an:2007031.

[Gri92] Michael Griebel. "The Combination Technique for the Sparse Grid Solution of PDE's on Multiprocessor Machines". In: *Parallel Processing Letters* 02.01 (Mar. 1992), pp. 61–70. ISSN: 0129-6264, 1793-642X. DOI: 10.1142/S0129626492000180.

[GW05] Bernhard Gleich and Jürgen Weizenecker. "Tomographic Imaging Using the Nonlinear Response of Magnetic Particles". In: *Nature* 435.7046 (June 2005), pp. 1214–1217. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature03808.

[Hac19] Wolfgang Hackbusch. *Tensor Spaces and Numerical Tensor Calculus*. Vol. 56. Springer Series in Computational Mathematics. Cham: Springer International Publishing, 2019. ISBN: 978-3-030-35553-1 978-3-030-35554-8. DOI: 10.1007/978-3-030-35554-8.

[HK09] Wolfgang Hackbusch and Stefan Kühn. "A New Scheme for the Tensor Representation". In: *Journal of Fourier Analysis and Applications* 15.5 (Oct. 2009), pp. 706–722. ISSN: 1069-5869, 1531-5851. DOI: 10.1007/s00041-009-9094-9.

[HO08] Bernard Haasdonk and Mario Ohlberger. "Reduced Basis Method for Finite Volume Approximations of Parametrized Linear Evolution Equations". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 42.2 (Mar. 2008), pp. 277–302. ISSN: 0764-583X, 1290-3841. DOI: 10.1051/m2an:2008001.

[HRS16] Jan S. Hesthaven, Gianluigi Rozza, and Benjamin Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. SpringerBriefs in Mathematics. Cham: Springer International Publishing, 2016. ISBN: 978-3-319-22469-5 978-3-319-22470-1. DOI: 10.1007/978-3-319-22470-1.

[Hsu02] Elton P. Hsu. *Stochastic Analysis on Manifolds*. Graduate Studies in Mathematics Volume 38. Providence, Rhode Island: American Mathematical Society, 2002. ISBN: 978-0-8218-0802-3 978-1-4704-2090-1.

[KB12] Tobias Knopp and Thorsten M. Buzug. *Magnetic Particle Imaging: An Introduction to Imaging Principles and Scanner Instrumentation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. ISBN: 978-3-642-04198-3 978-3-642-04199-0. DOI: 10.1007/978-3-642-04199-0.

[KJL18] Tobias Kluth, Bangti Jin, and Guanglian Li. "On the Degree of Ill-Posedness of Multi-Dimensional Magnetic Particle Imaging". In: *Inverse Problems* 34.9 (Sept. 2018), p. 095006. ISSN: 0266-5611, 1361-6420. DOI: 10.1088/1361-6420/aad015.

[KL07] Othmar Koch and Christian Lubich. "Dynamical Low-Rank Approximation". In: *SIAM Journal on Matrix Analysis and Applications* 29.2 (Jan. 2007), pp. 434–454. ISSN: 0895-4798, 1095-7162. DOI: 10.1137/050639703.

[KL10] Othmar Koch and Christian Lubich. "Dynamical Tensor Approximation". In: *SIAM Journal on Matrix Analysis and Applications* 31.5 (Jan. 2010), pp. 2360–2375. ISSN: 0895-4798, 1095-7162. DOI: 10.1137/09076578X.

[Klu18] Tobias Kluth. "Mathematical Models for Magnetic Particle Imaging". In: *Inverse Problems* 34.8 (Aug. 2018), p. 083001. ISSN: 0266-5611, 1361-6420. DOI: 10.1088/1361-6420/aac535.

[KM17] Tobias Kluth and Peter Maass. "Model Uncertainty in Magnetic Particle Imaging: Nonlinear Problem Formulation and Model-Based Sparse Reconstruction". In: *International Journal on Magnetic Particle Imaging* Vol.3 (2017), 10 pages. ISSN: 2365-9033. DOI: 10.18416/IJMPI.2017.1707004.

[KSK19] Tobias Kluth, Patryk Szwargulski, and Tobias Knopp. "Towards Accurate Modeling of the Multidimensional Magnetic Particle Imaging Physics". In: *New Journal of Physics* 21.10 (Oct. 2019), p. 103032. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ab4938.

[KSU15] Sebastian Kestler, Kristina Steih, and Karsten Urban. "An Efficient Space-Time Adaptive Wavelet Galerkin Method for Time-Periodic Parabolic Partial Differential Equations". In: *Mathematics of Computation* 85.299 (Aug. 2015), pp. 1309–1333. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/mcom/3009.

[LC85] William Allan Light and Elliott Ward Cheney. *Approximation Theory in Tensor Product Spaces*. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1985. ISBN: 978-3-540-16057-1 978-3-540-39741-0. DOI: 10.1007/bfb0075391.

[Loe+16] Norbert Loewa, Patricia Radon, Olaf Kosch, and Frank Wiekhorst. "Concentration Dependent MPI Tracer Performance". In: *International Journal on Magnetic Particle Imaging* Vol.2 (2016), 5 pages. ISSN: 2365-9033. DOI: 10.18416/IJMPI.2016.1601001.

[LR74] Pierre Lasaint and Pierre-Arnaud Raviart. "On a Finite Element Method for Solving the Neutron Transport Equation". In: *Mathematical Aspects of Finite Elements in Partial Differential Equations*. Elsevier, 1974, pp. 89–123. ISBN: 978-0-12-208350-1. DOI: 10.1016/B978-0-12-208350-1.50008-X.

[Mal09] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Elsevier, 2009. ISBN: 978-0-12-374370-1. DOI: 10.1016/B978-0-12-374370-1.X0001-8.

[Mey91] Yves Meyer. "Ondelettes Sur l'intervalle". In: *Revista Matemática Iberoamericana* 7.2 (Aug. 1991), pp. 115–133. ISSN: 0213-2230, 2235-0616. DOI: 10.4171/rmi/107.

[Øks03] Bernt Øksendal. *Stochastic Differential Equations*. Universitext. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. ISBN: 978-3-540-04758-2 978-3-642-14394-6. DOI: 10.1007/978-3-642-14394-6.

[Ose11] Ivan V. Oseledets. "Tensor-Train Decomposition". In: *SIAM Journal on Scientific Computing* 33.5 (Jan. 2011), pp. 2295–2317. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/090752286.

[Pav14] Grigorios A. Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. Vol. 60. Texts in Applied Mathematics. New York, NY: Springer New York, 2014. ISBN: 978-1-4939-1322-0 978-1-4939-1323-7. DOI: 10.1007/978-1-4939-1323-7.

[QMN16] Alfio Quarteroni, Andrea Manzoni, and Federico Negri. *Reduced Basis Methods for Partial Differential Equations*. Vol. 92. UNITEXT. Cham: Springer International Publishing, 2016. ISBN: 978-3-319-15430-5 978-3-319-15431-2. DOI: 10.1007/978-3-319-15431-2.

[RH73] William Reed and Thomas Hill. *Triangular Mesh Methods for the Neutron Transport Equation*. Tech. rep. Los Alamos, New Mexico, 1973.

[Ris96] Hannes Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Ed. by Hermann Haken. Vol. 18. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996. ISBN: 978-3-540-61530-9 978-3-642-61544-3. DOI: 10.1007/978-3-642-61544-3.

[Rog+13] Marie E. Rognes, Daseul A. Ham, Christopher J. Cotter, and Andrew T. T. McRae. "Automating the Solution of PDEs on the Sphere and Other Manifolds in FEniCS 1.2". In: *Geoscientific Model Development* 6.6 (Dec. 2013), pp. 2099–2119. ISSN: 1991-9603. DOI: 10.5194/gmd-6-2099-2013.

[RS19] Nikolaos Rekatsinas and Rob Stevenson. "An Optimal Adaptive Tensor Product Wavelet Solver of a Space-Time FOSLS Formulation of Parabolic Evolution Problems". In: *Advances in Computational Mathematics* 45.2 (Apr. 2019), pp. 1031–1066. ISSN: 1019-7168, 1572-9044. DOI: 10.1007/s10444-018-9644-2.

[RUV24] Niklas Reich, Karsten Urban, and Jürgen Vorloeper. *A Parallel Batch Greedy Algorithm in Reduced Basis Methods: Convergence Rates and Numerical Results*. 2024. DOI: 10.48550/ARXIV.2407.11631.

[SS09] Christoph Schwab and Rob Stevenson. "Space-Time Adaptive Wavelet Methods for Parabolic Evolution Problems". In: *Mathematics of Computation* 78.267 (Sept. 2009), pp. 1293–1318. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/s0025-5718-08-02205-9.

[Ste04] Rob Stevenson. "On the Compressibility of Operators in Wavelet Coordinates". In: *SIAM Journal on Mathematical Analysis* 35.5 (Jan. 2004), pp. 1110–1132. ISSN: 0036-1410, 1095-7154. DOI: 10.1137/S0036141002411520.

[Ste09] Rob Stevenson. "Adaptive Wavelet Methods for Solving Operator Equations: An Overview". In: *Multiscale, Nonlinear and Adaptive Approximation*. Ed. by Ronald DeVore and Angela Kunoth. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 543–597. ISBN: 978-3-642-03412-1 978-3-642-03413-8. DOI: 10.1007/978-3-642-03413-8_13.

[Ste93] Frank Stenger. *Numerical Methods Based on Sinc and Analytic Functions*. Vol. 20. Springer Series in Computational Mathematics. New York, NY: Springer New York, 1993. ISBN: 978-1-4612-7637-1 978-1-4612-2706-9. DOI: 10.1007/978-1-4612-2706-9.

[SU09] Winfried Sickel and Tino Ullrich. "Tensor Products of Sobolev–Besov Spaces and Applications to Approximation from the Hyperbolic Cross". In: *Journal of Approximation Theory* 161.2 (Dec. 2009), pp. 748–786. ISSN: 0021-9045. DOI: 10.1016/j.jat.2009.01.001.

[SVVW22] Rob Stevenson, Raymond Van Venetië, and Jan Westerdiep. "A Wavelet-in-Time, Finite Element-in-Space Adaptive Method for Parabolic Evolution Equations". In: *Advances in Computational Mathematics* 48.3 (June 2022), p. 17. ISSN: 1019-7168, 1572-9044. DOI: 10.1007/s10444-022-09930-w.

[SW21] Rob Stevenson and Jan Westerdiep. "Stability of Galerkin Discretizations of a Mixed Space–Time Variational Formulation of Parabolic Evolution Equations". In: *IMA Journal of Numerical Analysis* 41.1 (Jan. 2021), pp. 28–47. ISSN: 0272-4979, 1464-3642. DOI: 10.1093/imanum/drz069.

[SWM12] Jochen Schuetz, Michael Woopen, and Georg May. "A Hybridized DG/Mixed Scheme for Nonlinear Advection-Diffusion Systems, Including the Compressible Navier-Stokes Equations". In: *50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*. Nashville, Tennessee: American Institute of Aeronautics and Astronautics, Jan. 2012. ISBN: 978-1-60086-936-5. DOI: 10.2514/6.2012-729.

[Tob12] Christine Tobler. "Low-Rank Tensor Methods for Linear Systems and Eigenvalue Problems". PhD thesis. ETH Zurich, 2012, 1 Band. DOI: 10.3929/ETHZ-A-007587832. HDL: 20.500.11850/60368.

[TV16] Francesca Tantardini and Andreas Veeser. "The L^2-Projection and Quasi-Optimality of Galerkin Methods for Parabolic Equations". In: *SIAM Journal on Numerical Analysis* 54.1 (Jan. 2016), pp. 317–340. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/140996811.

[UP12] Karsten Urban and Anthony T. Patera. "A New Error Bound for Reduced Basis Approximation of Parabolic Partial Differential Equations". In: *Comptes Rendus. Mathématique* 350.3-4 (Feb. 2012), pp. 203–207. ISSN: 1631-073X, 1778-3569. DOI: 10.1016/j.crma.2012.01.026.

[UP13] Karsten Urban and Anthony Patera. "An Improved Error Bound for Reduced Basis Approximation of Linear Parabolic Problems". In: *Mathematics of Computation* 83.288 (Oct. 2013), pp. 1599–1615. ISSN: 0025-5718, 1088-6842. DOI: 10.1090/S0025-5718-2013-02782-2.

[Vid03] Guifré Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". In: *Physical Review Letters* 91.14 (Oct. 2003), p. 147902. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.91.147902.

[Was25] Eric Wasserstein. "Entry A133843". In: *The on-line encyclopedia of integer sequences* (2025).

[Wei+09] Jürgen Weizenecker, Bernhard Gleich, Jürgen Rahmer, Hannes Dahnke, and Jörn Borgert. "Three-Dimensional Real-Time *in Vivo* Magnetic Particle Imaging". In: *Physics in Medicine and Biology* 54.5 (Mar. 2009), pp. L1–L10. ISSN: 0031-9155, 1361-6560. DOI: 10.1088/0031-9155/54/5/L01.

[Wei+10] Jürgen Weizenecker, Bernhard Gleich, Jürgen Rahmer, and Jörn Borgert. "Particle Dynamics of Mono-Domain Particles in Magnetic Particle Imaging". In: *Magnetic Nanoparticles*. Institute of Medical Engineering, University of Lübeck, Germany: WORLD SCIENTIFIC, Aug. 2010, pp. 3–15. ISBN: 978-981-4324-67-0 978-981-4324-68-7. DOI: 10.1142/9789814324687_0001.

[Wei18] Jürgen Weizenecker. "The Fokker–Planck Equation for Coupled Brown–Néel-rotation". In: *Physics in Medicine & Biology* 63.3 (Jan. 2018), p. 035004. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aaa186.

[Yan14] Masayuki Yano. "A Space-Time Petrov–Galerkin Certified Reduced Basis Method: Application to the Boussinesq Equations". In: *SIAM Journal on Scientific Computing* 36.1 (Jan. 2014), A232–A266. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/120903300.

[Yan+22] Xue Yang, Guoqing Shao, Yanyan Zhang, Wei Wang, Yu Qi, Shuai Han, and Hongjun Li. "Applications of Magnetic Particle Imaging in Biomedicine: Advancements and Prospects". In: *Frontiers in Physiology* 13 (July 2022), p. 898426. ISSN: 1664-042X. DOI: 10.3389/fphys.2022.898426.