IEEE Open Journal of the
Industrial Electronics Society

Received 28 May 2024; accepted 27 June 2024. Date of publication 4 July 2024;
date of current version 18 July 2024. The review of this article was arranged by Associate Editor Jorge Solsona.

Digital Object Identifier 10.1109/OJIES.2024.3423405

Signal Recovery in Power Systems by Correlated Gaussian Processes

MARCEL ZIMMER [1], DANIELE CARTA [1] (Member, IEEE), THIEMO PESCH [1], AND ANDREA BENIGNI [1,2,3] (Senior Member, IEEE)

[1]ICE-1: Energy Systems Engineering, Forschungszentrum Jülich, 52428 Jülich, Germany
[2]RWTH Aachen University, 52056 Aachen, Germany
[3]JARA-Energy, 52425 Jülich, Germany

CORRESPONDING AUTHOR: MARCEL ZIMMER (e-mail: m.zimmer@fz-juelich.de).

This work was supported by the Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE).

ABSTRACT This article proposes the application of correlated Gaussian processes (Corr-GPs) for the recovery of missing intervals in power systems signals. Based on only local power system topology, the presented algorithm combines cross-channel information of the considered signals with a universal, non-parametric probabilistic machine learning regression to recover missing data. Starting from the theoretical background, the proposed approach is presented and contextualized in the framework of signal recovery for power systems. Then, by making use of real data collected from the Living Lab Energy Campus—a real-life laboratory established at Forschungszentrum Jülich—we demonstrate the use of the proposed approach for recovering distribution grid signals. We evaluate the performances of Corr-GP compared with those of other state-of-the-art techniques. In addition to outperformance in terms of recovery accuracy, it is explained when and how the accuracy of the reconstructed signal is independent of the missing interval length. Finally, detailed insights about key characteristics of the proposed approach that generate practical benefits for system operators are provided. A self-aware failing indication allowing system operators a direct evaluation of the recovered data and enabling further improvement of the proposed approach is presented, as well as recommendations for field implementation.

INDEX TERMS Gaussian processes (GPs), monitoring, power system monitoring.

I. INTRODUCTION

A. MOTIVATION

Nowadays, power systems are monitored with an increasing number of measurement devices, aimed at providing the system operator with fundamental information about the operating conditions of the system. From smart meters and power quality meters [1], largely applied in low-voltage networks, to phasor measurement units (PMUs) [2], mostly used in transmission and distribution systems, every day huge datasets are generated in the field. In control centers, this large amount of data is provided as input to routines for monitoring (e.g., state estimation [2] or the estimation of line parameters [3]), automation (e.g., harmonic analysis [4]), and control (e.g., for voltage stability [5]) that strongly rely on these measurements. Within this framework, it is evident that the successful operation of these algorithms depends not only on the number and accuracy of measurements but also on the data quality [6]. In particular, bad data is the term commonly used when referring to corrupted data and data loss [7]. Such issues can originate from malfunctioning of the devices in the monitoring chain (instrument transformers and measurement devices), issues with the communication network, or failures when storing the data. An example with reference to a PMU-based measurement system is reported in [8], while in [9], the problem of outliers in a SCADA-based monitoring environment is addressed. Bad data can also originate from a tough environment and thus can play a key role in the monitoring of health conditions of components where, as underlined in [10], data considered incorrect in other fields are indicators of faults in machinery. Moreover, it is worth mentioning that bad data can also originate from malicious cyberattacks, which are causing an increasing concern [11]. Since the issues mentioned above

are getting more and more common, being able to recover a missing section of a signal would help the system operator overcome the issues due to bad data and ensure the proper operation of network managing routines.

B. RELATED WORK

Signal recovery methods applied to large-scale monitoring systems can roughly be classified into single-channel and multichannel algorithms [12]. Classical single-channel methods range from expectation–maximization methods [13] to mean methods [12]. As linear interpolation (*Lin-Int*) is a common representative of those methods, we use it to evaluate the method proposed in this work. Modern single-channel approaches include both regression and matrix completion algorithms. Konstantinopoulos et al. [14] addressed the problem of data losses in a monitoring framework composed of PMUs utilizing temporal online algorithm for PMU data processing (OLAP-t). Such an algorithm allows the recovery of signals exhibiting temporal patterns subjected to prolonged data outages. Alternatively, based on the method of multipliers in the alternating direction, in [15], the recovery of missing PMU data has been addressed by exploiting the low-rank property of the corresponding measurement matrix. Addressing the matrix completion algorithms, we evaluate the proposed approach against a compressive sensing (CS)-based recovery technique, which is presented in [16]. One such technique, which has been shown to provide the best performance in case of varying signals, uses a discrete cosine transform to reformulate the recovery problem on a sparse basis, such that a norm-1 minimization algorithm can be applied to recover not-null components. Due to the lack of information within larger missing intervals, single-channel methods have a natural limitation.

A straightforward cross-channel approach utilizes related signals, referred to as *Rep-phase* in this article, such as other phases of a multiphase system. Although this approach is highly dependent on the balance of the power system and is not universally applicable, any dedicated algorithm has to outperform this trivial method. More elaborated multichannel algorithms include the power system's graph structure, extended matrix completion algorithms, as well as machine learning techniques. Similarly to [14] and [15], but combined with Kirchhoff's laws and three-phase circuit relations as regularization terms, in [17], the recovered values are given by using cross-correlations of synchrophasor measurements to complete a low-rank tensor. Wu et al. [18] considered the spatial and temporal graphs to define the correlations between system components and quantities, before applying a regression neural network for the recovery of missing data in large power systems. Similarly, in [17], cross-correlations of synchrophasor measurements are recovered by the completion of a low-rank tensor using Kirchhoff's laws and three-phase circuit relations as regularization terms. In [19], a jointly low-rank tensor completion is applied to recover data from logistic systems. In [15], the recovery of missing PMU data has been addressed by exploiting the low-rank property of the corresponding measurement matrix. In [20], a network expectation–maximization algorithm is proposed to recover data in spatiotemporal applications for small missing intervals. Using neural networks including graph structures, missing power system synchrophasors caused by communication delays are recovered as reported in [21]. In [22], a generative adversarial network is used for missing data reconstruction.

To summarize, single-channel recovery algorithms capture typical signal behaviors and are mainly capable of dealing with short missing data intervals. Multichannel approaches can incorporate simultaneously appearing information but might be limited by unknown system information in practical applications. For example, in [22], the entire power system has to be known first, a simulation model has to be built, and a decent amount of scenarios, e.g., 12 000 in the application in [22], has to be simulated to generate training data. In other words, the need for Big Data in artificial intelligence might be a limitation in a practical setting. Moreover, a changing system topology, which once in a while appears in real power systems, cannot be tracked without dedicated effort. On top, the inherent strong dependency of artificial intelligence on the specific dataset, i.e., on a specific regularization scheme, makes universal application difficult. Consequently, there is a need for a multichannel approach based only on local system topology that is capable of recovering large missing intervals from a relatively small amount of known data.

The proposed approach is based on the technique of multioutput Gaussian processes (GPs). With applications in different domains including distinct domain-specific peculiarities, these types of models have been applied in [23] and [24]. In [23], a multioutput GP is utilized for model building of crowdsourced traffic data. With a focus on real-time information processing, the application of data collection from a weather sensor network is presented in [24].

C. CONTRIBUTION

In this article, a correlated Gaussian process (Corr-GP) framework is formulated for the application within the field of power systems' data recovery. First, we give a construction of single and multichannel models for use in power systems, utilized for signal recovery. We then emphasize cross-channel information transfer that encodes local power system's topology. In particular, we show that the amount of needed data to recover a signal is comparable to the desired signal to be recovered. Thus, the proposed approach can handle large periods of missing data, while getting along with comparably small amounts of data. We analyze and demonstrate the proposed signal recovery algorithm's self-aware failing indication, yielding trustworthy recovery results. More precisely, we show that the proposed approach allows for a direct and inherent evaluation of performance, which can be utilized as a possibly automated reliable recovery procedure applied to large databases. We demonstrate that the results provided by the proposed approach are independent of the length of the missing window if sufficient local cross-channel information

is available. Moreover, supported by a detailed mathematical analysis, we show that the proposed approach is almost unconstrained within the desired application. In particular, we explicitly point out that the valid assumption for any sensor reading of Gaussian (measurement) noise is sufficient for the proposed approach to handle any continuous power system signal. The entire evaluation and analysis of the above statements are performed using real data collected from the field. We apply the proposed framework to a real-live laboratory in operation at Forschungszentrum Jülich (FZJ) called the Living Lab Energy Campus (LLEC).

The rest of this article is organized as follows. Section II presents the theoretical background. In Section III, we present the proposed methods and provide a theoretical discussion. Section IV describes the case study. The analysis and results are presented in Section V. Finally, Section VI concludes this article.

II. THEORETICAL BACKGROUND

In this section, the basic mathematical formulation of GPs is presented.

A GP is a collection of random variables, any finite number of which have a joint Gaussian distribution [25]. In line with the given definition, we may assume the training data $(X, Y) = \{(X, Y)|(x, y) \in \mathbb{R}^N \times \mathbb{R}, N \in \mathbb{N}\}$, the test data X^*, as well as the joint set to have a Gaussian distribution, respectively. In the application of signal recovery, Y will be the available signal values with corresponding time instances X, and X^* denote the time instances of the signal to be recovered. Using Bayes' theorem [25], the joint process can be reduced to the test set $X^* = \{x^*|x^* \in \mathbb{R}^N\}$, yielding the desired prediction at X^*, by conditioning on the training set. Accordingly, it is possible to define a prior and a posterior of the process. Formally, this procedure is computed according to the following equations.

Let X be a compact subset of $\mathbb{R}^n$. We call $\mu_{\text{prior}} : X \to \mathbb{R}$ a prior mean function. For the sake of simplicity, we assume that $\mu_{\text{prior}} = 0$ throughout as justified in Section III-A for the given case study. The positive semidefinite function $\kappa : X \times X' \to \mathbb{R}$ is called a prior covariance function. Synonymously used and evident from the following definition, we refer to κ as the kernel function, kernel matrix, or simply kernel as we have

$$K\left(X, X'\right)_{ij} := \left(\kappa\left(x_i, x_j\right)\right) \quad \forall x_i \in X \quad \forall x_j \in X'. \tag{1}$$

Noise is taken into account by adding $\sigma_n \in \mathbb{R}$ to the diagonal, i.e., $K_n(X, X) = K(X, X) + \mathbf{1}\sigma_n$, where $\mathbf{1}$ denotes the identity matrix. Although called prior noise, in practice the value of σ_n accounts for multiple effects. First, as the label suggests, additive, independent, identically distributed white noise is included in the model. Second, a small positive number is added to the diagonal of $K_n(X, X)$, fixing numerical issues concerning matrix inversion via Cholesky decomposition [26]. On top, it also allows to take into account aspects not to be captured explicitly, which may be treated as noise.

Any κ is specified by a set of hyperparameters determined by training the GP. This, is due to the fact that the GP is nonparametric itself, and thus, the hyperparameters defining (1) might be trained. More precisely, any given process is trained by minimizing the objective function, called negative log marginal likelihood, which can be computed, as shown in [25], and reads as follows:

$$\frac{1}{2} Y^T K_n(X, X)^{-1} Y + \frac{1}{2} \ln |K_n(X, X)| + \frac{N}{2} \ln(2\pi) \tag{2}$$

where N denotes the number of training data points with respect to the hyperparameters defining the selected kernel. The three parts of (2) can be interpreted as follows. The first part is responsible for fitting the model to the data while the second part suppresses model complexity. The third part is a constant, which originates from the computation of the marginal likelihood [25].

The training is performed according to a gradient-based optimizer [25], whereby an limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm is used in the given implementation [26]. With optimally determined kernel, the posterior mean μ_{post} and the posterior covariance function σ_{post} can be computed as follows:

$$\mu_{\text{post}} = K(X^*, X) K_n(X, X)^{-1} Y \tag{3}$$

$$\Sigma_{\text{post}} = K(X^*, X^*) - K(X^*, X) K_n(X, X)^{-1} K(X, X^*). \tag{4}$$

III. PROPOSED APPROACH

In this section, the application of GPs as a signal-recovering method in power systems is proposed. First, we discuss theoretical insights, which allow for stating key findings and characteristics of the proposed method. Second, we show how to build GP models for power systems applications.

A. GP CHARACTERISTICS FOR SIGNAL RECOVERY

By definition of GPs, the Gaussian assumption holds at a function space level, not on the data itself. This is rooted in the fact that any finite number of function evaluations have a joint Gaussian distribution, not only individual measurements themselves. Thus, a generic signal of interest interpreted as a time series does not have to explicitly obtain a Gaussian distribution, e.g., with respect to temporal structures. Nevertheless, it is worth underlining that the Gaussian assumption on individual measurement values can still be valid for conventional measurement devices. Based on models lying densely in the space of continuous functions, the considered GPs are universal approximators [27], meaning any continuous function can be approximated by a GP. Thus, no constraints exist within the proposed application.

In the framework of signal recovery, we refer to the available or known data as the training set. The temporal component of the data to be recovered is called the test set. By definition, the prior, as well as the posterior, is Gaussian distributed in the respective sense. Thus, each posterior comes with a mean and a covariance, μ_{post} and Σ_{post}, respectively. In particular, we commonly refer to the posterior mean values, μ_{post}, as the recovered signal. The posterior variance, σ_{post}, is given by the square root of the diagonal of Σ_{post}. Strictly

speaking, σ_{post} cannot be interpreted as an error interval in the usual sense, and thus cannot be associated with the model's accuracy. However, as being a major concern of this work, it provides a measure of the model's precision, as well as, a self-contained failing indication. Thus, on top of recovering missing data, the suggested models provide a reliable, in a sense self-aware technique, conveniently usable by system operators.

With regard to the prior mean function, it should be noted that a nonzero mean function corresponds to a coupling of the nonparametric GP described above with a parametric model [25]. In the given power system application, due to the continuous power demand fluctuations, the signals of interest are nondeterministic. To this end, imposing a prebuild structure on the data by a parametric model is not applicable due to the required flexibility and the considered time scale. Thus, we choose $\mu_{\text{prior}} = 0$.

B. CORR-GPS FOR SIGNAL RECOVERY

1) SINGLE-CHANNEL MODEL—VOLTAGE AND CURRENT MAGNITUDES

The kernel is the defining property of the GP models under consideration. We start by selecting a single-channel kernel for the modeling of power system signals. Since sums and products of kernels also yield valid kernels [25], it is worth highlighting that there are infinite ways to construct GPs for individual power system signals. As we are dealing with time-depending signals, we use $t, t' \in T \subset \mathbb{R}$ to indicate time variables instead of generic $x, x' \in X$ as used in Section II in the following.

A roughly varying voltage or current rms signal, collected from the low-voltage side of a transformer with a reporting rate of 1 min, as investigated in this work, is best described by a one-time continuous differentiable kernel, e.g., the $\nu = 3/2$ Matérn kernel. Such a kernel is formulated as follows:

$$\mathrm{M}^{3/2}(t, t') = \sigma^2 \left(\frac{l + \sqrt{3}|t - t'|}{l}\right) \exp\left(\frac{-\sqrt{3}|t - t'|}{l}\right) \quad (5)$$

where the hyperparameter $l \in \mathbb{R}^+$ defines a timescale of the given process and $\sigma \in \mathbb{R}^+$ accounts for a scaling in magnitudes' direction.

Furthermore, in order to take slowly varying trends into account, a smooth kernel, such as the squared exponential, denoted as radial basis function (RBF), can be considered

$$\mathrm{RBF}(t, t') = \sigma^2 \exp\left(-\frac{|t - t'|^2}{2\,l^2}\right). \quad (6)$$

It is worth to note two important aspects. First, due to the universal property, for data characterized by a high time resolution, the specific choice appears to be less important as the presented kernels yield a universal approximation. On the contrary, in the case of coarser datasets, the signals of interest tend to appear smoother, which implies the usage of smooth kernels. Second, as described in the next section, the key feature of the presented approach to power system signals is the coupling of different single-channel signals into a mutual multichannel signal. As this appears to be the dominant effect of the proposed approach, the single-channel kernel selection becomes less significant in Corr-GPs.

2) MULTICHANNEL MODEL—CROSS-CORRELATED VOLTAGE AND CURRENT MAGNITUDES

Intrinsically induced by their local topological nature, power system signals display a certain degree of similarity depending on the level of balance present in the given system. The proposed approach takes these characteristic cross-correlations between certain signals explicitly into account. The key mathematical operation accomplishing this task is the tensor product $\otimes$, a multiplication structure yielding to GP models with multiple outputs [25]. More specifically, as a mathematical operation, the tensor product provides a separable kernel, i.e., a composition of a kernel describing both single-channel characteristics and a cross-correlation kernel [28]. In other words, by construction, the proposed method accounts for the balance induced by the power system topology while being flexible with respect to the type of signal considered. We refer to such models as Corr-GPs, highlighting the main characteristics of the proposed approach. This type of model has been applied in different domains, such as transport analysis [23] and anomaly detection [29].

At the level of matrices, the tensor product is established in terms of the Kronecker product, also denoted by $\otimes$. In particular, given a kernel matrix $K(T, T')$ and a symmetric positive semidefinite $d \times d$-matrix $B^{[d]}$, a d-dimensional model can be constructed by

$$B^{[d]} \otimes K\left(T, T'\right) = \begin{pmatrix} b_{11}K(T, T') & \dots & b_{1d}K(T, T') \\ \vdots & \ddots & \vdots \\ b_{d1}K(T, T') & \dots & b_{dd}K(T, T') \end{pmatrix} \quad (7)$$

where the superscript $[\cdot]$ denotes the dimension of the associated square matrix. The dimension d of the coregionalization matrix, $B^{[d]}$, corresponds to the number of coupled signals. As being the key component of the proposed approach, we want to emphasize the interpretation of $B^{[d]}$ in terms of power system signal recovery. First, it should be noted that the entries $\{b_{ij}|i, j = 1, \dots, d\}$ of $B^{[d]}$ are trainable hyperparameters of the given model. Diagonal entries $\{b_{ij}|i = j; i, j = 1, \dots, d\}$ describe the self-contribution of the signal to be recovered to the proposed model while off-diagonal terms $\{b_{ij}|i \neq j; i, j = 1, \dots, d\}$ account for cross-correlations among related signals induced by the local system's topology. The latter is the key component of the proposed approach to incorporate simultaneous local power system information, improving the accuracy of recovery results. Emphasizing the consequences of this technicality regarding the given application, we give the following comments.

First, note that the possibility of including simultaneously appearing information, of arbitrary size, about the power system allows the proposed approach to handle large missing

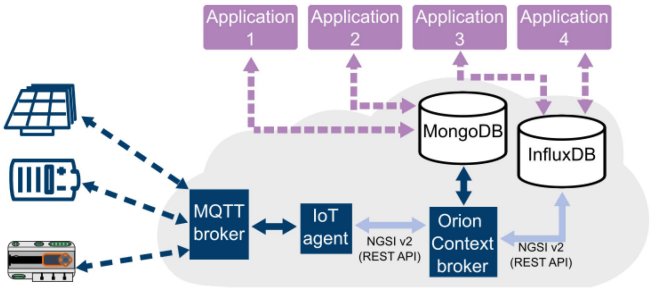

FIGURE 1. Simplified scheme of FZJ ICT FIWARE-based platform.

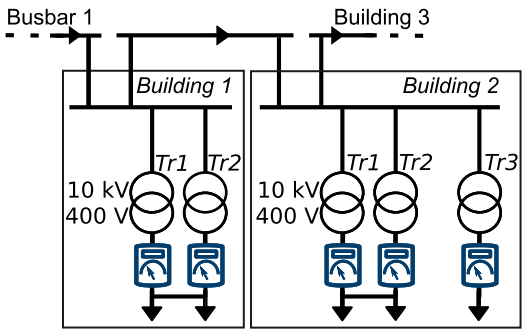

FIGURE 2. Overview of the considered buildings.

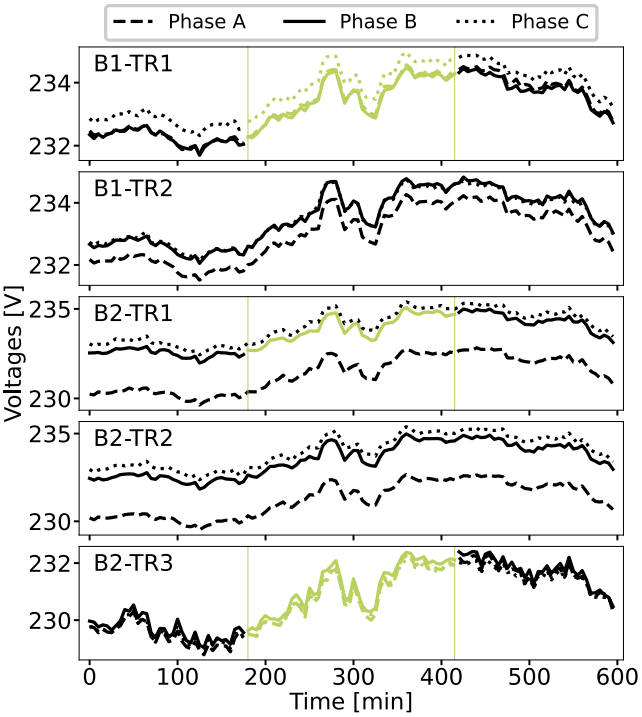

FIGURE 3. Reference signals—voltages.

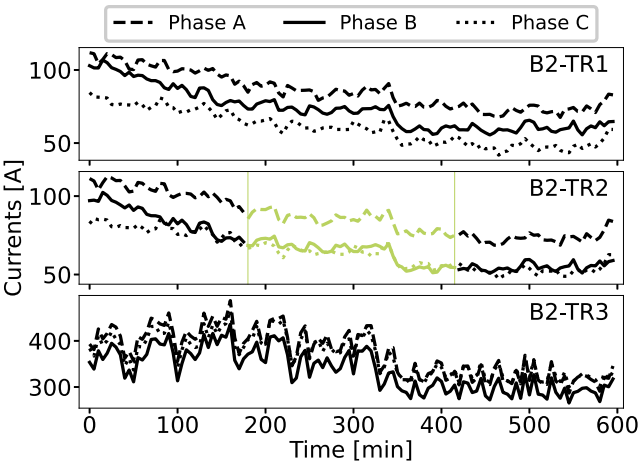

FIGURE 4. Reference signals—currents.

periods. Practically constrained by the availability of sufficient simultaneous system information, theoretically, the proposed approach can recover missing intervals of arbitrary size. Second, the different interpretation of elements of $B^{[d]}$ induces a particular choice of the dataset from which the missing interval can be recovered. To train diagonal elements $\{b_{11}, b_{22}, \dots\}$, the available data on both sides of the missing period are selected. We considered the missing window to be in between two known intervals of almost the same size, as visualized for this case study in Figs. 3 and 4. It should be noted that this is done for a matter of convenience and by no means mandatory to the proposed approach. In fact, having known data before and after the missing period is recommended as it conditions the proposed approach on both ends of the missing window. For off-diagonal entries $\{b_{ij} | i \neq j; i, j = 1, \dots, d\}$, correlated signals are chosen, which are selected by the local power system's topology.

With this at hand, we may define two types of models, one being the generalization of the other. Let K_{basis} be a kernel. The intrinsic coregionalization model (ICM) is defined as follows:

$$K_{\text{ICM}}(T, T') = B^{[d]} \otimes K_{\text{basis}}\left(T, T'\right). \tag{8}$$

This type of model is used in the case of a single class of correlated signals, such as, for example, balanced three-phase current magnitudes.

For multiple classes, e.g., several types of similarities within a set of signals, more sophisticated models are needed. To this end, several ICMs could be added in order to obtain a linear coregionalization model (LCM)

$$K_{\text{LCM}}(T, T') = \sum_{q} B_q^{[d]} \otimes K_{\text{basis},q}\left(T, T'\right) \tag{9}$$

where q denotes the number of considered ICMs. For example, this might be used for modeling q transformers.

Beyond the separable kernel subclass of multioutput GPs, nonseparable kernels as described in [28] do exist. Concerning less evident correlations, this class of kernels might be worth investigating. Nevertheless, given the presented application, such models are out of the scope of this article.

C. MULTICHANNEL RECOVERED SIGNALS

For j signals $(s_1^*(t_1^*), \ldots s_j^*(t_j^*))$ to be recovered from i correlated signals $(s_1(t_1), \ldots s_i(t_i))$, the proposed approach reads as follows:

$$\begin{pmatrix} s_1^*(t_1^*) \\ \vdots \\ s_j^*(t_j^*) \end{pmatrix} = \sum_q \begin{pmatrix} b_{11}^q K^q(t_1^*, t_1) & \ldots & b_{1\,d}^q K^q(t_1^*, t_i) \\ \vdots & \ddots & \vdots \\ b_{d1}^q K^q(t_j^*, t_1) & \ldots & b_{dd}^q K^q(t_j^*, t_i) \cdot \end{pmatrix}$$
$$\begin{pmatrix} b_{11}^q K^q(t_1,t_1)+\sigma_{s_1}^2 & \ldots & b_{1d}^q K^q(t_1,t_i) \\ \vdots & \ddots & \vdots \\ b_{d1}^q K^q(t_i,t_1) & \ldots & b_{dd}^q K^q(t_i,t_i)+\sigma_{s_i}^2 \end{pmatrix}^{-1} \begin{pmatrix} s_1(t_1) \\ \vdots \\ s_i(t_i) \end{pmatrix} \tag{10}$$

where q runs over the corresponding similarity classes and $d = i + j$. For the base kernel $K^q(t, t')$, we propose to choose a Matérn kernel (5). For larger models with $q > 2$, an additional kernel squared exponential kernel (6) should be included.

IV. CASE STUDY—LLEC AT FZJ

The power grid of the LLEC of FZJ serves as a test bench in this article. In particular, it is worth noticing that the proposed data-recovery solution is being developed for its implementation on the FIWARE-based information and communication technology (ICT) platform of LLEC.

A. ICT PLATFORM

The ICT platform of FZJ is based on the open-source Internet-of-Things (IoT) platform FIWARE [30]. A schematic representation of its implementation in the LLEC framework is shown in Fig. 1. The platform allows gathering data from various components on the field (e.g., photovoltaic systems, batteries, monitoring devices, and so on) in real time, via an MQTT broker. Then, via the IoT agent, data are converted into the NGSIv2 data format and sent via representational state transfer (REST) application programming interface (API) to the FIWARE Orion Context Broker. Here, only the most updated data are stored in the MongoDB database, whereas the historical data are stored in the InfluxDB database. Since this large-scale framework is used to store data from different components and sensors, the internal structure of the databases consists of multiple subdatabases customized to properly store data from specific components. It is worth noting that different devices and sensors are typically characterized by different reporting rates. For example, temperature sensors installed in the rooms of the FZJ campus provide change-of-value measurements, i.e., no fixed reporting rate can be given. Instead, the power system data considered in this work, rms voltages and currents are collected from the energy meters installed on the field, are provided with a fixed reporting rate of 5 min. All these data are first saved independently in the MongoDB, overwriting the data stored in the previous reporting interval, and then saved in the InfluxDB, concatenated to the corresponding set of historical data. In addition to the specific quantities for each device/sensor, the metadata structure is also composed of more general information, such as timestamps, unique identification number of the device/sensor, location, and so on. External applications and services can read/write data directly from the databases. In particular, the ICT platform is modular, which allows different applications to run independently, suppressing malfunctioning due to chain effects.

An example of such an application is constituted by the bad data detection and handling tool presented in [7]. This multilevel application aims to detect bad data and provide reconstructed/recovered data. The signal recovery method presented in this article has been developed for its application in the recovery module of the abovementioned application, to recover large sets of missing data. Nevertheless, since this work aims to present the Corr-GP-based method, and underline its capabilities as a standalone tool, the proposed method has been tested outside the routine presented in [7]. In particular, the datasets needed for the tests have been identified by the authors, according to the settings of the considered test cases, and the corresponding data have been collected directly from the historical database, i.e., InfluxDB.

B. POWER SYSTEM

Over the campus, the power distribution is carried out on the 10-kV level via two main busbars. In total, 136 three-phase medium-voltage/low-voltage (MV/LV) transformers are used to supply the low-voltage loads on the 400-V level. On each transformer, the rms values of voltages and currents on each phase are monitored by means of measurement devices in compliance with class 1 of the standard IEC 61036 [31], as well as the total active and reactive powers. The collected data are stored in internal databases, introduced in the previous section, for monitoring, control, and planning purposes. Nevertheless, as underlined in [7], it is often the case that a large amount of data are missing or corrupted from this database. Thus, in this article, we will consider the recovery of missing data and voltage or current rms magnitudes in particular, by means of the proposed Corr-GP.

For the tests, two different buildings have been considered, as shown in Fig. 2.

The corresponding different internal configurations can be described as follows.

1) *Building 1* is characterized by two MV/LV transformers (in the following labeled as B1-TR1 and B1-TR2) in parallel.
2) *Building 2*, directly connected to Building 1, being in the same ring, is characterized by three MV/LV transformers (in the following labeled as B2-TR1, B2-TR2, and B2-TR3), two of which are connected in parallel.

For these two buildings, a 10-h interval of voltages and currents without missing data has been identified in the historical database. Then, in order to be able to evaluate and discuss key characteristics of the proposed Corr-GP, different missing intervals have been simulated during the tests. The considered voltage and current signals are reported in Figs. 3 and 4. Black curves denote training data, while green sections underline the

TABLE 1. Summary of Missing and Available Data or Each Test Case

Case	# Missing-Channels	# Available-Channels	# Missing Data	# Available Data
1-a,b	1	2	48, 1:48	312, 359:312
2	3	12	144	1656
3	3	6	144	936

maximum interval of the considered missing data, i.e., the sets under test. In particular, during the tests, the total duration of these intervals will vary from 5 min to a total period of 4 h. Different patterns are used to represent the different phases. In order to properly evaluate the performance of the proposed method and to underline its associated coupling features, three different test cases have been considered.

In Case 1, the recovery of the voltage magnitude on phase B of the first three-phase MV/LV transformer in Building 2, i.e., B2-Tr1-B, will be investigated by taking into account the magnitude of the voltages on the two other phases of the same transformer, i.e., B2-TR1-A and B2-TR2-C. First, case 1-a analyzes the performance with respect to a missing 4-h interval compared with the state-of-the-art methods, i.e., single-output Gaussian process (S-GP), CS [16], Lin-Int, and the naive approach of substitution by data collected from phases A and C, denoted by Rep-A and Rep-C, respectively. Second, case 1-b analyzes the effect of the missing interval length and a comparison is made with the stated alternatives.

In Case 2, in addition to Case 1, the voltage magnitudes collected from all transformers present in Buildings 1 and 2 are provided as input to the proposed method. More specifically, all of the three phases from the first transformer in Building 1, B1-Tr1, and the last transformer in Building 2, B2-Tr3, are considered missing for 4 h. In case 2-a, scalability and the ability to identify relevant cross-correlations of Corr-GPs are discussed. In case 2-b, a self-failing indication is presented.

Finally, to show the applicability of Corr-GPs to signals with different characteristics, the recovery of the current magnitudes of Building 2 is considered in Case 3. In particular, the current profile of B2-Tr2 will be recovered on the basis of the current magnitudes collected from B2-Tr1 and B2-Tr3.

To further clarify the considered experimental setup, a summary of missing and available data for each considered test case is reported in Table 1, in terms of rms samples. In the table, also the number of channels (i.e., signals) considered in the test cases is indicated. Thus, with reference to Case 1-a, a total of 360 samples are considered from three channels (i.e., 120 samples per signal), 48 of which are considered missing. It is worth noting that, although the proposed approach is rooted in the realm of probabilistic machine learning, the amount of available data (training data) to recover missing data (test data) is comparable in size.

V. TESTS AND RESULTS

In this section, the validity of the proposed approach will be investigated by discussing its performance with reference to the test cases described in the previous section. All of the

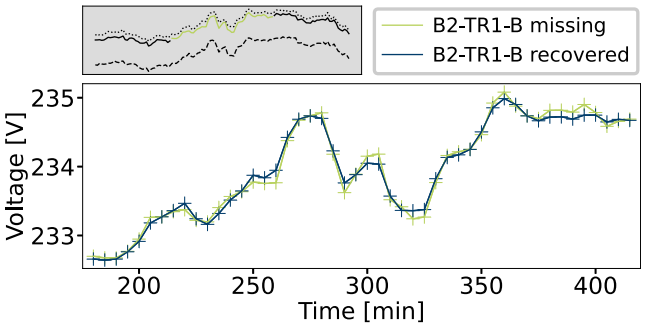

FIGURE 5. Case 1-a—True and recovered voltage profiles of the three-phase transformer B2-Tr1, with Corr-GP.

TABLE 2. Comparison of Different Methods

	Corr-GP	Rep-A	Rep-C	Lin-Int	CS	S-GP
NRMSE [%]	0.031	0.995	0.1651	0.230	0.333	0.322

results are evaluated with respect to the normalized root mean squared error, stated as follows:

$$\mathrm{NRMSE}(y_{\mathrm{true}}, y_{\mathrm{rec}}) = \frac{1}{\bar{y}_{\mathrm{true}}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{\mathrm{true},i} - y_{\mathrm{rec},i})^2} \tag{11}$$

where y_{true} and y_{rec} denote the true and the recovered values, respectively, and $\bar{y}_{\mathrm{true}}$ denotes the arithmetic mean of $y_{\mathrm{true},i}$ over the N samples. Below, we abbreviate (11) by normalized root mean squared error (NRMSE).

A. CASE 1-A: ACCURACY, PRECISION, AND BASIC CHARACTERISTICS

The basic characteristics and the performance of Corr-GPs as a recovery tool will be presented in the following considering signals derived from the collection of voltage magnitudes of the first three-phase transformer in Building 2, B2-Tr1. The explicit model reads as follows:

$$K_{\mathrm{Case1}}(T, T') = B^{[3]} \otimes M^{3/2}(T, T'). \tag{12}$$

In Fig. 5, the recovery results for a missing 4-h interval are presented.

The color-contrasting upper subplot recaps the available data of 10 h taken into account. Focusing on the missing interval, the green crosses denote the signals considered missing while blue crosses describe the recovered values by the Corr-GP. It has to be noticed that in the figure it is not possible to appreciate the blue error bars representing the posterior variances $2\sigma_{\mathrm{post}}$, since their value is in the order of magnitude of 0.08 V. This is in line with the interpretation of posterior variances of Corr-GPs provided in Section III-A.

The comparison to alternative methods is reported in Table 2, in terms of the NRMSE, and displayed in Fig. 6. Contextualizing recap and color coding of Fig. 6 is used the same way as before. From Figs. 5 and 6 and Table 2, the first two main results can be deduced. In particular, the

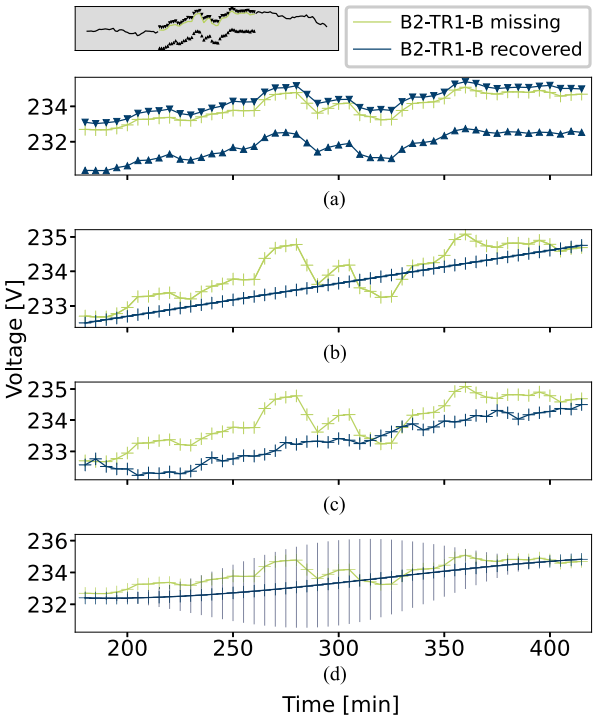

FIGURE 6. Case 1-a—Comparison between recovered voltage profiles of phase B of transformer B2-Tr1. (a) Rep-A / Rep-C. (b) Lin-Int. (c) CS. (d) S-GP $\pm\sigma$.

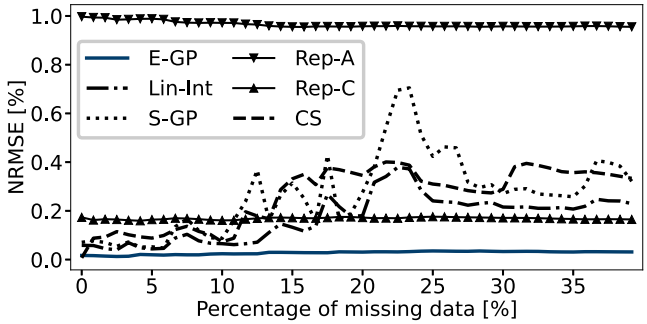

FIGURE 7. Case 1-b—Comparison of NRMSE of different methods with respect to the length of the missing signal.

overlapping green and blue curves in Fig. 5, with an NRMSE of 0.031%, underline the accuracy of the Corr-GP recovery. Compared to the alternative methods, an improvement by a factor between 5.3 and 32.0 can be archived by using the proposed Corr-GP.

With the naive approach Rep-A and Rep-C, which can also be seen as a model correlating different phases, the accuracy of this rudimentary method is unpredictable in practice as it is fully relying on the network balance. While Rep-C is less accurate than the Corr-GP by a factor of 5.3, Rep-A deviates by a factor of 32.0. Lin-Int, CS, and S-GPs with Matérn kernel (5) are less accurate by factors of 7.4, 10.7, and 10.4, respectively.

Furthermore, being in a range too small to be visualized in Fig. 5, the posterior covariance indicates a high precision of the given Corr-GP. This is in particular evident when compared to the S-GP, shown in Fig. 6(d). In both cases, the same (base) kernel (5) is used. This comparison clearly indicates the value of taking cross-correlations between different phases into account. While the single-phase GP, defined by (5), is not able to accurately recover the missing signal due to missing information during the recovery period, the use of a separable kernel, given by (7), enables the inclusion of simultaneously occurring information. Contrarily, if no cross-correlation information is available, the proposed Corr-GP approach reduces to an uncoupled multioutput model, i.e., the correlation matrices have zero off-diagonal terms. As this is equivalent to multiple S-GP, in the absence of sufficient correlations, the accuracy of Corr-GP reduces to the given single-output GP alternatives.

B. CASE 1-B: MISSING INTERVAL INDEPENDENT RECOVERY ACCURACY

Besides the high accuracy and precision of Corr-GPs, a main characteristic is the independence of the recovery performances with respect to the missing interval length as long as a good level of correlation exists between the signal to be reconstructed and other signals in the systems. In practical applications, this property is met for sufficient large power systems by local topological system's characteristic. Using the same dataset as above, this property is shown by varying the missing interval length from 5 min (1 sample) to 4 h (48 samples), with an increasing step of 5 min. Individual NRMSEs of the methods discussed above with respect to the percentage of missing values in reference to the total interval length are reported in Fig. 7. The blue line is used for the results of the Corr-GP while the remaining methods are reported in black. From the figure, it is possible to deduce that the performance of Corr-GP is almost independent of the percentage amount of missing data. Due to the lack of further network unbalances in the observed period, such stability also holds for the replacement by the other phases Rep-A and Rep-C. As for the other methods considered in the comparison, it is possible to observe that the corresponding NRMSE is sensitive to the increase in the portion of missing data. In other words, these results provide an experimental validation, within the given case study, of the independence of the proposed algorithm's accuracy with respect to the missing period size, theoretically discussed in Section III-B2.

Furthermore, the presence of peaks in the curves of these methods underlines how their performances are also affected by the different characteristics of the available and the missing signals. These considerations are not only shown in the given test case but are also clear from a theoretic perspective. As these methods do not take into account any information within the missing time frame, they cannot be expected to provide precise results for larger intervals.

TABLE 3. Case 2—NRMSE of Corr-GP for B1-Tr1 and B2-Tr3

	NRMSE (%)					
	B1-Tr1			B2-Tr3		
	Ph. A	*Ph. B*	*Ph. C*	*Ph. A*	*Ph. B*	*Ph. C*
$K^{\text{ICM}}_{\text{Case2}}(T,T')$	0.023	0.022	0.024	0.685	0.350	0.186
$K^{\text{LCM}}_{\text{Case2}}(T,T')$	0.016	0.012	0.023	0.153	0.153	0.143

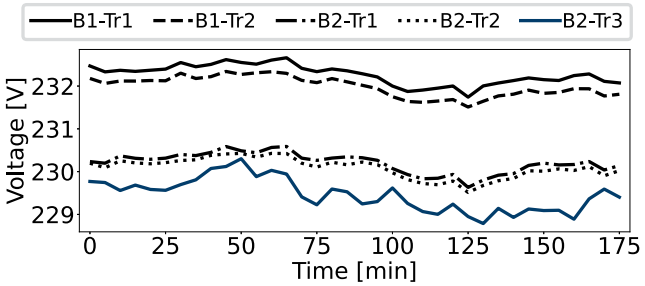

FIGURE 8. Phase A for three hours of available signals of all five transformers.

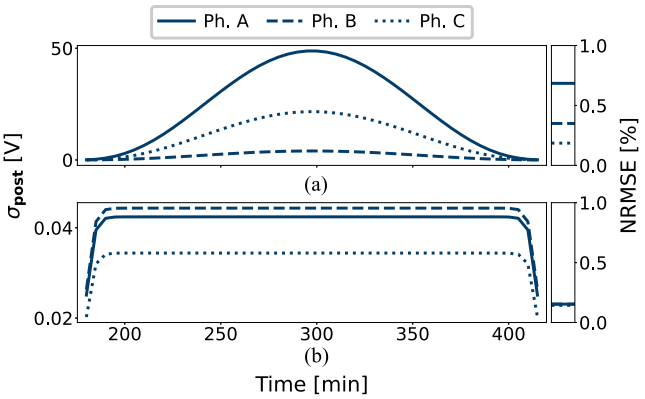

FIGURE 9. Posterior variances (left-hand side) for transformer B2-Tr3, by (a) ICM and (b) LCM, with corresponding NRMSE (right-hand side).

C. CASE 2-A: CROSS-CORRELATION IDENTIFICATION

In the second case study, the amount of information provided as input to the Corr-GP method is increased, by considering the voltage magnitudes of all five transformers present in both Building 1 and Building 2. Instead of recovering the missing window only from one phase of a single transformer, here, three phases from two different transformers located in different buildings, B1-Tr1 and B2-Tr3, will be considered as missing. The comparison of the results achieved with the previous ICM model for five three-phase transformers, formulated as follows:

$$K^{\text{ICM}}_{\text{Case2}}\left(T,T'\right) = B^{[15]} \otimes M^{3/2}\left(T,T'\right) \tag{13}$$

and the generalized LCM model, formulated as follows:

$$K^{\text{LCM}}_{\text{Case2}}\left(T,T'\right) = B^{[15]}_1 \otimes M^{3/2}_1\left(T,T'\right) \tag{14a}$$

$$+\, B^{[15]}_2 \otimes M^{3/2}_2\left(T,T'\right) \tag{14b}$$

are reported in terms of the NRMSE, in Table 3.

While the results for transformer B1-Tr1 are comparable for both considered Corr-GP's and similar to those discussed in Case 1, different considerations should be made for the accuracy of B2-Tr3. We elaborate on this and discuss a key feature of Corr-GP's in the rest of this section.

In general, signals collected across different buildings are expected to have fewer similarities than signals collected within a single transformer. As an example, we report the current magnitudes of phase A for the first three hours of the considered interval for all five transformers in Fig. 8.

From Fig. 8, it is possible to appreciate that transformers B1-Tr1, B1-Tr2, B2-Tr1, and B2-Tr2, denoted in black, are quite balanced while transformer B2-Tr3, denoted in blue, slightly deviates. Thus, taking two classes of signal types into account by using an LCM instead of an ICM model, leads to the better performance of the Corr-GP for transformer B2-Tr3, reported in Table 3. The overall better performance of the LCM Corr-GP model (14) for the recovery of B1-Tr1 compared with those of B2-Tr3 is also evident, as B2-Tr3 is the only transformer displaying this isolated behavior. In fact, there are less cross-correlations available to recover voltage magnitudes of B2-Tr3. In summary, the Corr-GP learns during training which cross-correlation to take into account. If less cross-information is available in this regard, the corresponding signals are treated separately, giving results that are similarly accurate to the single-phase methods provided in Case 1.

D. CASE 2-B: SELF-AWARE FAILING INDICATION

So far, the discussion mainly focused on the Corr-GP's performance. In this section, an additional feature of Corr-GP for practical, and possibly automated, usage by system operators is underlined. In Fig. 9, posterior variances are shown for both models under test with respect to the overall 4-h missing interval as well as the corresponding accuracy results.

As shown in Fig. 9(a), the ICM approach is characterized by a significant increase of σ_{post}-values when compared with the LCM model depicted in Fig. 9(b). Moreover, a dependence between σ_{post} (left-hand side of Fig. 9) and the NRMSE (right-hand side of Fig. 9) is evident. Thus, not only the proper Corr-GP model, in this case the LCM, comes with a high accuracy, but the less fitting Corr-GP approach can be detected by a high posterior variance. Contrarily to the common usage as an error estimate of the resulting values, the posterior variance of GPs displays an error estimate of the selected model. It is worth recalling that, in line with the interpretation discussed in Section III-A, the posterior variance describes the uncertainty of the Corr-GP associated with each recovered sample, where the latter is the corresponding Corr-GP mean value. This means that the higher the posterior variance, the less the selected Corr-GP model fits the recovery task of a desired accuracy. Thus, the notion of posterior variances yields a self-failing indication allowing a direct evaluation of the performance of the given model. Consequently, an automated iterative procedure could be performed by system operators, by gradually increasing model complexity. In cases without

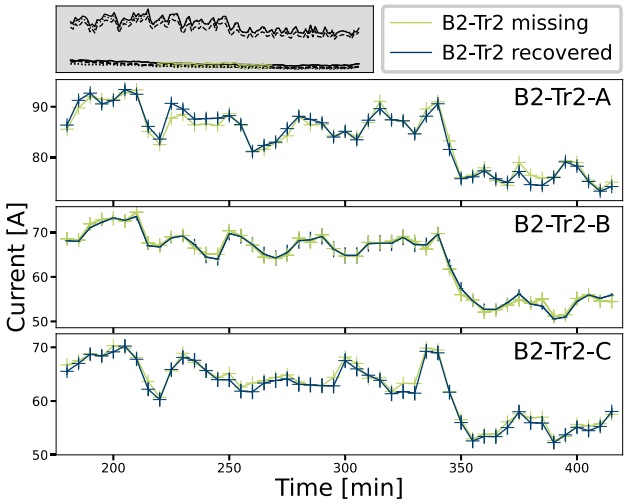

FIGURE 10. Case 3—True and recovered current profiles for the three phases of transformer B2-Tr2.

detailed prior knowledge of the similarity classes provided by the topology and load conditions of the power system, or within an automated operation, a trial and error approach could be performed. First, a basic ICM model of the proposed approach corresponding to a single similarity class including posterior variances is computed. If, based on the posterior variance, the desired model accuracy is not sufficient, an additional ICM term is appended and the extended Corr-GP is retrained. This procedure might be iterated until an acceptable value for the posterior variance is reached. In other words, the iterations terminate if the number of added ICM terms corresponds to the number of similarity classes. However, in practical applications, the number of possible iterations is bounded by computational resources, thus, limiting a fully automated operation in cases where no cross-channel information is available. In cases where detailed information can be provided by system operators or a dedicated pre-analysis of the given signals is performed, the trial and error procedure might be neglected. In such a case, the posterior variance serves as a consistency check.

E. CASE 3: MODELING OF MULTI CROSS-CORRELATED SIGNALS

Due to their different characteristics, current signals cannot be modeled as the voltage ones. Evident from Figs. 4 and 10, in Case 3 the current magnitudes are considered, which are characterized by less homogeneity and cross-correlation compared to the voltages. Taking into account the 10-h current profiles of all phases of B2-Tr1 and B2-Tr-3, as well as the preceding and following 3-h intervals of B2-Tr2, the recovered 4-h interval profiles are displayed in blue in Fig. 10. In particular, missing current magnitudes to be recovered are denoted in green.

By a closer inspection of Fig. 4, it is possible to observe that there exist three similarity classes in the given data set, yielding (15b). Moreover, to also have a correlation term across distinct classes, we take long-term correlations into account, established by (15a). Thus, the proposed Corr-GP model reads as follows:

$$K_{\text{Case3}}\left(X, X'\right) = B_0^{[9]} \otimes \text{RBF}\left(T, T'\right) \tag{15a}$$

$$+ \sum_{q=1}^{3} B_q^{[9]} \otimes M_q\left(T, T'\right) \tag{15b}$$

and gives accuracies with respect to the NRMSE of 1.2%, 1.0%, and 1.3% for phases A, B and C, respectively. Contrarily, in this scenario Lin-Int results in 4.6%, 7.5%, and 5.5%. It should be emphasized that the findings of the previous Section V-D allow an automated determination of the right LCM model.

From Fig. 4, it is possible to observe that the current profiles of B2-Tr1 and B2-Tr2 are very similar, while the currents flowing in the third transformer are almost four times higher, and also characterized by a more volatile profile. Considering the results shown in Fig. 10, we can thus observe that the proposed recovery method is able to distinguish between the characteristics of the different coupled signals and to give more weight to the information coming from the more similar inputs.

F. LIMITATIONS, FIELD IMPLEMENTATION, AND FUTURE PROSPECT

As shown in the previous analysis, the proposed approach is neither limited by the missing window length nor the type of correlation under consideration. However, it is worth underlying that the absence of any correlated signal makes the proposed approach not applicable. Thus, regarding an automated application of the proposed approach, the relevant domain is within large-scale monitoring systems, in which the existence of correlated signals can be guaranteed in practical scenarios. It is currently planned to implement Corr-GPs as part of the recovery module running in the modular tool for bad data detection and recovery of the ICT platform of FZJ [7]. Nevertheless, before applying it in a large scale, automated environment, certain aspects and limitations of Corr-GP have to be considered.

One aspect to consider is the computational complexity associated with the matrix inversion by Cholesky decomposition in (2)–(4). In [24], an updated Cholesky decomposition is proposed to accelerate the calculation of the matrix inverse. However, this means a continuous operation, i.e., the inclusion of every data point into the proposed approach. In the recovery of large missing windows, this is not beneficial since the outage of sensors, communication problems within the database, or maintenance work do not frequently occur but are limited to individual occasions. A possible reduction of the computational burden can be obtained with the use of a minimal dataset in each recovery task, i.e., the ratio between known and missing data has to be minimized. Within a large-scale application, this requires a statistical analysis to ensure reliable applicability. Another possibility is the limitation of

hyperparameters to a relevant domain. Although retraining cannot be avoided completely as timescales and magnitudes might change during online operation, the number of optimizer restarts might be decreased. Another possibility is an analytical preanalysis of the given signals to predetermine the number of needed similarity classes. This might reduce the number of possible model iterations, and could thus, reduce computational costs.

The implementation can be easily performed using GPy [26] by means of APIs as described in [7]. However, depending on the actual computational load, scaling might be limited due to computational resources. In this case, a tailored GPU-based computational environment including a specialized Corr-GP implementation might be needed.

VI. CONCLUSION

In this work, power systems' data recovery by Corr-GPs has been proposed, discussed, and evaluated. During the tests that were performed considering real voltage and current signals collected from the campus network of Forschungszentrum Jülich, the proposed Corr-GP was capable of outperforming the alternative state-of-the-art methods. Moreover, it has been shown that, contrary to the considered alternatives, the accurate performance is mainly independent of the missing interval length. Besides the approximately one order of magnitude higher accuracy when cross-correlated information is available, Corr-GPs perform comparably accurate to the stated alternatives if no such correlated information is available. On top, the proposed method yields a self-failing indication in terms of posterior variances, thus recovery precision, allowing system operators a direct evaluation and further improvement if needed. Finally, key aspects, such as limitations and computational complexity regarding the field implementation, have been discussed.

REFERENCES

[1] V. L. Srinivas and J. Wu, "Topology and parameter identification of distribution network using smart meter and μPMU measurements," *IEEE Trans. Instrum. Meas.*, vol. 71, May 2022, Art. no. 9004114, doi: 10.1109/TIM.2022.3175043.

[2] W. Song, J. He, J. Lin, H. Ye, X. Ling, and C. Lu, "Bias analysis of PMU-Based state estimation and its linear Bayesian improvement," *IEEE Trans. Ind. Informat.*, vol. 20, no. 2, pp. 1607–1617, Feb. 2024.

[3] P. A. Pegoraro, C. Sitzia, A. V. Solinas, S. Sulis, D. Carta, and A. Benigni, "Compensation of systematic errors for improved PMU-based fault detection and location in three-phase distribution grids," *IEEE Trans. Instrum. Meas.*, early access, May 13, 2024, doi: 10.1109/TIM.2024.3400340.

[4] Y. Wang, H. Ma, X. Xiao, Y. Wang, Y. Zhang, and H. Wang, "Harmonic state estimation for distribution networks based on multi-measurement data," *IEEE Trans. Power Del.*, vol. 38, no. 4, pp. 2311–2325, Aug. 2023.

[5] Y. Li, S. Zhang, Y. Li, J. Cao, and S. Jia, "PMU measurements-based short-term voltage stability assessment of power systems via deep transfer learning," *IEEE Trans. Instrum. Meas.*, vol. 72, Sep. 2023, Art. no. 2526111, doi: 10.1109/TIM.2023.3311065.

[6] A. Gholami et al., "Detection and classification of anomalies in power distribution system using outlier filtered weighted least square," *IEEE Trans. Ind. Informat.*, vol. 20, no. 5, pp. 7513–7523, May 2024.

[7] D. Carta, M. Zimmer, T. Pesch, and A. Benigni, "Bad data detection and handling in ICT platforms for energy systems," in *Proc. IEEE 12th Int. Workshop Appl. Meas. Power Syst.*, 2022, pp. 1–6.

[8] Y. Liu et al., "Robust event classification using imperfect real-world PMU data," *IEEE Internet Things J.*, vol. 10, no. 9, pp. 7429–7438, May 2023.

[9] V. S. Patel, S. Chakrabarti, and A. Sharma, "A weighted deep neural network for processing measurements for state estimation," *IEEE Trans. Ind. Informat.*, vol. 20, no. 3, pp. 3530–3538, Mar. 2024.

[10] X. Xu, Y. Lei, and Z. Li, "An incorrect data detection method for Big Data cleaning of machinery condition monitoring," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 2326–2336, Mar. 2020.

[11] L. Liang and S. Liu, "Event-triggered distributed attack detection and fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 72, Oct. 2023, Art. no. 3520111, doi: 10.1109/TIM.2022.3216674.

[12] J. Du, M. Hu, and W. Zhang, "Missing data problem in the monitoring system: A review," *IEEE Sensors J.*, vol. 20, no. 23, pp. 13984–13998, Dec. 2020.

[13] D. Wang, S. Zhang, M. Gan, and J. Qiu, "A novel EM identification method for Hammerstein systems with missing output data," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2500–2508, Apr. 2020.

[14] S. Konstantinopoulos, G. M. De Mijolla, J. H. Chow, H. Lev-Ari, and M. Wang, "Synchrophasor missing data recovery via data-driven filtering," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4321–4330, Sep. 2020.

[15] M. Liao, D. Shi, Z. Yu, Z. Yi, Z. Wang, and Y. Xiang, "An alternating direction method of multipliers based approach for PMU data recovery," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4554–4565, Jul. 2019.

[16] D. Carta and A. Benigni, "Performance evaluation of a missing data recovery approach based on compressive sensing," in *Proc. IEEE 11th Int. Workshop Appl. Meas. Power Syst.*, 2021, pp. 1–6.

[17] A. Ghasemkhani, I. Niazazari, Y. Liu, H. Livani, V. A. Centeno, and L. Yang, "A regularized tensor completion approach for PMU data recovery," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1519–1528, Mar. 2021.

[18] T. Wu, Y.-J. A. Zhang, Y. Liu, W. C. Lau, and H. Xu, "Missing data recovery in large power systems using network embedding," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 680–691, Jan. 2021.

[19] Y. Gao, L. T. Yang, J. Yang, D. Zheng, and Y. Zhao, "Jointly low-rank tensor completion for estimating missing spatiotemporal values in logistics systems," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1814–1822, Feb. 2023.

[20] A. Grover and B. Lall, "A recursive method for estimating missing data in spatio-temporal applications," *IEEE Trans. Ind. Informat.*, vol. 18, no. 4, pp. 2714–2723, Apr. 2022.

[21] J. J. Q. Yu, A. Y. S. Lam, D. J. Hill, Y. Hou, and V. O. K. Li, "Delay aware power system synchrophasor recovery and prediction framework," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3732–3742, Jul. 2019.

[22] J. Fang, L. Zheng, and C. Liu, "A novel method for missing data reconstruction in smart grid using generative adversarial networks," *IEEE Trans. Ind. Informat.*, vol. 20, no. 3, pp. 4408–4417, Mar. 2024.

[23] F. Rodrigues, K. Henrickson, and F. C. Pereira, "Multi-output Gaussian processes for crowdsourced traffic data imputation," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 594–603, Feb. 2019.

[24] M. Osborne, S. Roberts, A. Rogers, S. Ramchurn, and N. Jennings, "Towards real-time information processing of sensor network data using computationally efficient multi-output Gaussian processes," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, 2008, pp. 109–120.

[25] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning* (Adaptive Computation and Machine Learning Series). Cambridge, MA, USA: MIT Press, Jan. 2006.

[26] GPy, "GPy: A Gaussian process framework in Python," [Online]. Available: since 2012 http://github.com/SheffieldML/GPy

[27] C. A. Micchelli, Y. Xu, and H. Zhang, "Universal kernels," *J. Mach. Learn. Res.*, vol. 7, pp. 2651–2667, Dec. 2006.

[28] M. A. Álvarez, L. Rosasco, and N. D. Lawrence, "Kernels for vector-valued functions: A review," *Found. Trends in Mach. Learn.*, vol. 4, no. 3, pp. 195–266, 2012.

[29] W. Cho, Y. Kim, and J. Park, "Hierarchical anomaly detection using a multioutput Gaussian process," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 261–272, Jan. 2020.

[30] "FIWARE - open APIs for open minds," Feb. 2021. [Online]. Available: https://www.fiware.org/

[31] IEC, "Alternating current static watt-hour meters for active energy (classes 1 and 2)," Standard IEC 61036, 2000.

MARCEL ZIMMER received the B.Sc. and M.Sc. degrees in physics from the University of Cologne, Cologne, Germany, in 2014 and 2018, respectively, and the B.Sc. degree in mathematics from the University of Cologne, in 2018. He is currently working toward the Ph.D. degree in mechanical engineering with the Institute for "Institute of Climate and Energy Systems: Energy Systems Engineering, ICE-1," Forschungszentrum Jülich, Jülich, Germany.

DANIELE CARTA (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and the Ph.D. degree in industrial engineering from the University of Cagliari, Cagliari, Italy, in 2013, 2016, and 2020, respectively.

He is currently the Group Leader of the "Control Solution" group with the "Institute of Climate and Energy Systems: Energy Systems Engineering, ICE-1," Forschungszentrum Jülich, Jülich, Germany.

THIEMO PESCH received the diploma in electrical engineering and information technology as well as in economic studies and the Ph.D. degree in mechanical engineering from RWTH Aachen University, Aachen, Germany, in 2008, 2009, and 2019, respectively.

He is currently the Head of the Departments "Energy Grids" and "High Performance Computing" with the "Institute of Climate and Energy Systems: Energy Systems Engineering, ICE-1," Forschungszentrum Jülich, Jülich, Germany.

ANDREA BENIGNI (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from Politecnico di Milano, Milano, Italy, in 2005 and 2008, respectively, and the Ph.D. degree from RWTH-Aachen University, Aachen, Germany, in 2013, all in electrical engineering.

From 2014 to 2019, he was an Assistant Professor with the Department of Electrical Engineering, University of South Carolina, Columbia, SC, USA. Since 2019, he has been a Full Professor with RWTH-Aachen and Director of the "Institute of Climate and Energy Systems: Energy Systems Engineering, ICE-1," Forschungszentrum Jülich, Jülich, Germany.