Journal of Computational Electronics (2025) 24:171
https://doi.org/10.1007/s10825-025-02412-4

Thermodynamically consistent stabilization of the drift-diffusion model for arbitrary band structures and carrier statistics

Tobias Linn[1] · **Max Renner**[1] · **Christoph Jungemann**[1]

Received: 2 July 2025 / Accepted: 19 August 2025

Abstract

To this day, the drift-diffusion model remains the most widely applied semiconductor simulation tool. This is due to its unrivaled numerical robustness when it is discretized with the finite volume method and the Scharfetter–Gummel stabilization. Unfortunately, this stabilization is only valid for nondegenerate carrier statistics. Several extensions of the Scharfetter–Gummel scheme to degenerate semiconductors have been proposed; however, they either rely on additional approximations or lack the stability for a full-scale device simulation. In this paper, we address this issue and present a generalization of the Scharfetter–Gummel scheme using no further approximations. Our scheme works for arbitrary band structures and coarse grids and is guaranteed to be thermodynamically consistent. Similar to Scharfetter–Gummel, it leads to a diagonally dominant Jacobian (M-matrix) for the discrete continuity equation preserving its excellent stability properties. An implementation of the algorithm is available online via Zenodo under the MIT license. It has already been used in a 2D device simulation at 4K where it exhibited excellent stability at a negligible runtime penalty.

Keywords Drift-diffusion model · Generalized Scharfetter–Gummel · Fermi–Dirac statistics · Arbitrary band structures

1 Introduction

Since the 1960s, the simulation of semiconductors has played a central role in the development of electronic devices. To this day, the drift-diffusion (DD) model remains the most important and widely used approach for modeling charge transport in semiconductor materials. A significant part of the model's enduring success can be attributed to the well-known Scharfetter–Gummel (SG) stabilization scheme [1]. It provides a simple analytical formula for the discretized current density along an edge that in 1D leads to a nodally exact solution under the assumption of nondegenerate Maxwell–Boltzmann carrier statistics, a constant mobility and no generation or recombination. It also ensures nonnegative carrier densities regardless of the grid spacing and, moreover, guarantees thermodynamic consistency, i.e., the discretized current is exactly zero at thermal equilibrium.

Unfortunately, the assumption of Maxwell–Boltzmann carrier statistics makes the scheme invalid in the degenerate case, e.g., for high doping concentrations, at cryogenic temperatures or for organic semiconductors. When Fermi–Dirac statistics is used instead, an implicit integral equation for the edge current arises [2] which cannot be solved analytically [3]. Hence, it has to be solved either by approximations or by numerical means.

Approximation-based schemes include the diffusion-enhanced scheme [4, 5] and the inverse activity scheme [6]. While these approaches preserve thermodynamic consistency and can be used for actual device simulation [7], they require very fine grids or suffer from reduced accuracy far from equilibrium. The generalized SG scheme presented in [2] uses an analytic approximation of the Fermi–Dirac integral which reduces the problem to an implicit algebraic equation, solvable by fixed-point iteration or Newton's method. More recently, in [3, 8] the integral was computed without further approximations using numerical quadrature rules, e.g., Gauss–Legendre quadrature. While this approach appears to be very promising, we find several problems with its implementation, in particular when using nonadaptive quadrature rules. Furthermore,

✉ Tobias Linn
tl@ithe.rwth-aachen.de

Max Renner
mr@ithe.rwth-aachen.de

Christoph Jungemann
cj@ithe.rwth-aachen.de

[1] Institute of Electromagnetic Theory, RWTH Aachen University, Kackertstr. 15-17, 52072 Aachen, Germany

Published online: 09 September 2025

thermodynamic consistency is only given up to the accuracy of the numerical quadrature and not guaranteed.

In this paper, we will address these issues and develop an improved version of [3, 8] using the adaptive Tanh–Sinh quadrature and improved Taylor series expansions around critical points. The scheme uses no further approximations as compared to the original SG scheme and works for arbitrary band structures, coarse grids and at an accuracy up to machine precision. A Fortran implementation of our method is publically available at [10]. The function takes the same input and output as the SG expression (including the relevant derivatives) and can therefore easily replace the SG scheme in any device simulator. This implementation has already been used in steady-state and transient 2D device simulations at 4K [9] and is thus proven to be robust enough even for the most extreme conditions. As the calculation can be effectively parallelized, the runtime penalty in [9] was less than 6 %.

The remaining paper is structured as follows. In Sect. 2, the DD model and the implicit integral equation for the current density are presented. Subsequently, the quadrature scheme used for the solution method is outlined in Sect. 3. The algorithm to calculate the edge current is developed in Sect. 4. Lastly, we demonstrate the capability of our method and compare it to existing schemes from the literature in Sect. 5.

2 Drift-diffusion model

In the DD model, the current density is given as

$$\boldsymbol{J} = qen\mu\boldsymbol{\nabla}\Phi, \tag{1}$$

where q is the sign of the carrier charge considered (e.g., $q = -1$ for electrons), e is the positive elementary charge, n is the carrier density, μ is the mobility, and Φ is the quasi-Fermi potential. The carrier density is given by

$$n = n_0\mathcal{F}(\eta), \tag{2}$$

where n_0 is a reference density (e.g., $n_0 = N_\mathrm{C}$, the effective density of states for electrons in the conduction band of a semiconductor) and $\mathcal{F}(\eta)$ is the cumulative distribution function, with the normalized chemical potential

$$\eta = \frac{\varphi - \varphi_0 - \Phi}{V_\mathrm{T}}. \tag{3}$$

φ is the electrostatic potential and $V_\mathrm{T} = \frac{k_\mathrm{B}T}{e}$ is the thermal voltage for the temperature T, where k_B is the Boltzmann constant. φ_0 is a reference potential (e.g., $\varphi_0 = \frac{E_\mathrm{C}}{e}$ for electrons in the conduction band of a semiconductor, with the conduction band edge energy E_C)

The cumulative distribution function is given by

$$\mathcal{F}(\eta) = \int_{-\infty}^{\infty} Z(t)f(t-\eta)\mathrm{d}t, \tag{4}$$

where $Z(t)$ is the normalized density of states and $f(\epsilon)$ is the distribution density with the normalized energy ϵ. For example, assuming 3D bulk semiconductors with parabolic bands and nondegenerate statistics (Maxwell–Boltzmann) we get

$$Z^\mathrm{parabolic}(t) = \frac{2}{\sqrt{\pi}}\sqrt{\max(t,0)} \tag{5}$$

and

$$f^\mathrm{MB}(\epsilon) = e^{-\epsilon}. \tag{6}$$

This results in $\mathcal{F}(\eta) = \exp(\eta)$. In the degenerate case, we need to use Fermi–Dirac statistics instead with

$$f^\mathrm{FD}(\epsilon) = \frac{1}{1+e^{\epsilon}}. \tag{7}$$

In this case, $\mathcal{F}(\eta)$ is given by $\mathcal{F}_{1/2}(\eta)$, the Fermi–Dirac integral of order 1/2, for which efficient computation methods are available (e.g., [11]). For other band structures, $\mathcal{F}(\eta)$ usually has to be calculated numerically using Eq. (4).

In the SG discretization, it is assumed that the electric field, the mobility and the current density are constant along an edge of the grid. Using these assumptions, a differential equation for η can be derived

$$\frac{\partial\eta}{\partial x} = \Delta\varphi - \frac{j}{\mathcal{F}(\eta)}, \tag{8}$$

where $x \in [0,1]$ is the dimensionless edge coordinate, $\Delta\varphi = \frac{\varphi_\mathrm{R}-\varphi_\mathrm{L}}{V_\mathrm{T}}$ is the normalized potential drop along the edge, and $j = \frac{l}{qe\mu V_\mathrm{T} n_0}J_\mathrm{edge}$ is the normalized edge current density with the edge length l, where J_edge is the component of $\boldsymbol{J}$ along the edge. Together with the boundary conditions $\eta(0) = \eta_\mathrm{L}$ and $\eta(1) = \eta_\mathrm{R}$, Eq. (8) is transformed into an integral equation [2]

$$\int_{\eta_\mathrm{L}}^{\eta_\mathrm{R}} \underbrace{\frac{1}{\Delta\varphi - \frac{j}{\mathcal{F}(\eta)}}}_{u(\eta)}\mathrm{d}\eta = 1. \tag{9}$$

This equation relates the unknown current j to the three parameters $\Delta\varphi$, η_L and η_R.

Assuming Maxwell–Boltzmann statistics, the equation can be solved analytically, resulting in the well-known expression derived by Scharfetter and Gummel [1]

$$j^\mathrm{SG} = B(-\Delta\varphi)\mathcal{F}_\mathrm{L} - B(\Delta\varphi)\mathcal{F}_\mathrm{R}, \tag{10}$$

with the positive Bernoulli function $B(\Delta\varphi) = \frac{\Delta\varphi}{\exp(\Delta\varphi)-1}$ and the normalized densities $\mathcal{F}_\mathrm{R} = \mathcal{F}(\eta_\mathrm{R})$ and $\mathcal{F}_\mathrm{L} = \mathcal{F}(\eta_\mathrm{L})$.

For Fermi–Dirac statistics and arbitrary band structures, it is not possible to find an analytical expression for $j(\Delta\varphi, \eta_L, \eta_R)$. Equation (9) therefore has to be either approximated or solved by numerical means. As stated before, in this work we choose the numerical approach.

A reliable numerical solution scheme of Eq. (9) is complicated because the integrand $u(\eta)$ can drastically change its behavior depending on j and $\Delta\varphi$. If j and $\Delta\varphi$ have different signs, $u(\eta)$ is quite well behaved (see Fig. 1). In contrast, if they have the same sign, a pole exists at $\eta_0 = \mathcal{F}^{-1}(j/\Delta\varphi)$ (see Fig. 2). Even though in [12] it was shown that a possible pole will never be located inside of the integration interval for the correct solution j, it can nevertheless get arbitrarily close to one of the integration bounds.

This leads to problems when the integral is calculated by generic numerical quadrature rules, especially for nonadaptive ones as in [3, 8]. For instance, for Gauss–Legendre quadrature it is assumed that the integrand can be approximated well by a polynomial, which is, however, not true close to a singularity. As the pole can move arbitrarily close to the integration bounds, the error of such a nonadaptive quadrature scheme can then become arbitrarily large. Another issue is that for large integration intervals we would either need more quadrature points or alternatively concentrate the existing points at the interesting regions of the integrand. This cannot be accomplished by a nonadaptive scheme.

In order to develop a more robust solution method for Eq. (9), we start with some general remarks. First, since $j(\Delta\varphi, \eta_L, \eta_R) = -j(-\Delta\varphi, \eta_R, \eta_L)$, we can limit the discussion to $\Delta\varphi \geq 0$. This is equivalent to flipping the edge direction.

Furthermore, depending on $\Delta\varphi$ and $\Delta\eta = \eta_R - \eta_L$, we can identify four special cases with an analytical solution. First, $\Delta\varphi = \Delta\eta$ results in $j = 0$ and a linear $\eta(x) = \eta_L + \Delta\eta x$. Second, $\Delta\eta = 0$ implies that $\eta(x) = \text{const} = \eta_L$ and in turn $j = \Delta\varphi\mathcal{F}(\eta_L)$. Third, the purely diffusive case with $\Delta\varphi = 0$ gives $j = -\int_{\eta_L}^{\eta_R} \mathcal{F}(\eta)\mathrm{d}\eta$ [8]. The fourth special case is $\Delta\varphi = \Delta\eta = 0$ with $j = 0$ and $\eta(x) = \text{const}$. In Table 1, all special cases are summarized and labeled.

The solution method developed in the following chapters consists of an adaptive quadrature scheme which is suitable even for near-singular integrands and Taylor series expansions around each of the special cases in Table 1. Before the solution procedure is presented in detail, the quadrature scheme used for the numerical solutions of the integrals is outlined in the next section.

Table 1 Special cases where analytical solutions exist

1	$\Delta\varphi \neq 0$	2	$\Delta\varphi = 0$
1a	$\Delta\eta = \Delta\varphi$	2a	$\Delta\eta \neq 0$
1b	$\Delta\eta = 0$	2b	$\Delta\eta = 0$

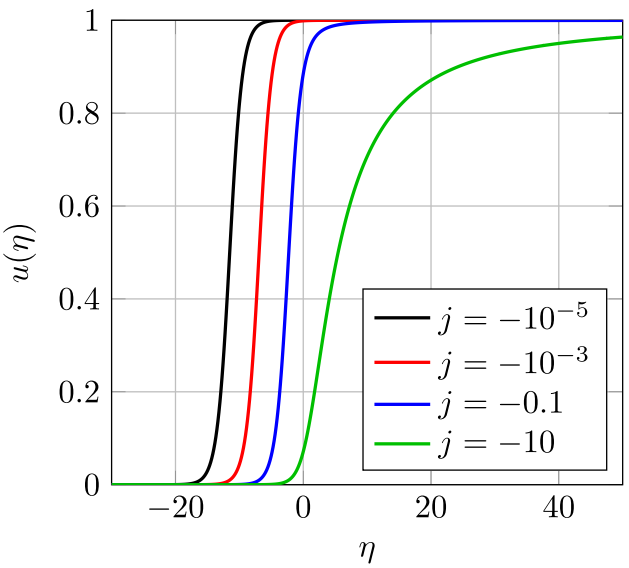

Fig. 1 Integrand for $\mathcal{F}(\eta) = \mathcal{F}_{1/2}(\eta)$ and for various negative values of j and $\Delta\varphi = 1$

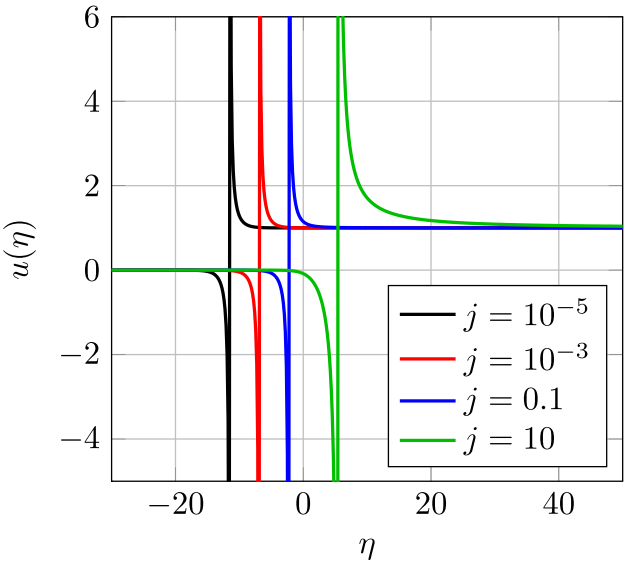

Fig. 2 Integrand for $\mathcal{F}(\eta) = \mathcal{F}_{1/2}(\eta)$ and various positive values of j and $\Delta\varphi = 1$

3 Double-exponential quadrature

A simple yet powerful adaptive quadrature method is the Tanh–Sinh quadrature [13, 14], which can also deal with singularities close to the integration bounds. The idea is to make a variable transformation

$$\eta(t) = \overline{\eta} + \frac{\Delta\eta}{2}\tanh(\sinh(t)) \tag{11}$$

with $\overline{\eta} = \frac{\eta_\mathrm{L}+\eta_\mathrm{R}}{2}$. This pushes possible singularities at the integral boundaries into $\pm\infty$, making this technique particularly suitable for near-singular integrands. The transformed integral is then computed using the trapezoidal rule, which has exponential convergence for integrands that decay at a double-exponential rate [15]

$$\begin{aligned}\int_{\eta_\mathrm{L}}^{\eta_\mathrm{R}} u(\eta)\mathrm{d}\eta &= \int_{-\infty}^{\infty} \frac{\Delta\eta}{2}\frac{\cosh(t)}{\cosh^2(\sinh(t))}u(\eta(t))\mathrm{d}t \\ &\approx \sum_{k=-\infty}^{\infty} w_k u(\eta_k)\end{aligned} \tag{12}$$

with

$$w_k = \frac{h\Delta\eta}{2}\frac{\cosh(kh)}{\cosh^2(\sinh(kh))} \tag{13}$$

and

$$\eta_k = \overline{\eta} + \frac{\Delta\eta}{2}\tanh(\sinh(kh)). \tag{14}$$

The infinite sum is cut off when the summands become too small and the calculation is repeated by reducing the step size $h = 1, \frac{1}{2}, \frac{1}{4}, \ldots$ until sufficient accuracy (close to machine precision) has been achieved.

If one of the integration bounds is infinite, the closely related Exp–Sinh quadrature [13] can be used instead with

$$\eta = \eta_1 + \exp(\sinh(t)). \tag{15}$$

Similarly, if both integration bounds are infinite, we use Sinh–Sinh quadrature [13] with

$$\eta = \eta_1 + \sinh(\sinh(t)). \tag{16}$$

These infinite versions of the quadrature routine are used for the calculation of $\mathcal{F}(\eta)$ (Eq. (4)).

4 Solution procedure

4.1 Calculation of $\mathcal{F}(\eta)$

Before the normalized edge current j can be calculated, we first need a way to compute the cumulative distribution function $\mathcal{F}(\eta)$. In contrast with parabolic bands where $\mathcal{F}(\eta) = \exp(\eta)$ (Maxwell–Boltzmann statistics) or $\mathcal{F}(\eta) = \mathcal{F}_{1/2}(\eta)$ (Fermi–Dirac statistics), for general band structures there is usually no efficient implementation available. In that case, $\mathcal{F}(\eta)$ has to be calculated numerically using Eq. (4). To that end, we use the infinite versions of the double-exponential quadrature (Eq. (15) or Eq. (16)) from the previous chapter. To improve performance in the subsequent calculations, a lookup table is created where $\mathcal{F}(\eta)$ is sampled for a wide range of values for η. To keep the size of the table small, we use cubic Hermite spline interpolation in between the sample points, taking the first derivative at each point into account. Each of the intervals is adaptively refined based on the interpolation error in its center, where we aim for a relative error below 10^{-13}. We choose between logarithmic (i.e., the logarithm of $\mathcal{F}(\eta)$ is interpolated) and direct interpolation and use whichever is more accurate within each interval.

4.2 Solution in the general case

Once we have an efficient way to compute $\mathcal{F}(\eta)$, we can continue with the numerical solution of Eq. (9). We start with the general case where none of the conditions in Table 1 apply. We rewrite the equation as a residual equation

$$r(j) = \int_{\eta_\mathrm{L}}^{\eta_\mathrm{R}} \frac{\mathcal{F}(\eta)}{\mathcal{F}(\eta)\Delta\varphi - j}\mathrm{d}\eta - 1 \overset{!}{=} 0. \tag{17}$$

The integral is calculated using Tanh–Sinh quadrature. In Fig. 3, the effect of the transformation from Eq. (11) is demonstrated. For the chosen parameters, the original integrand has a pole at η_0 located just to the left of the integration interval. The transformed integrand on the other hand is much more well behaved and decays rapidly to 0 for large $|t|$.

The resulting nonlinear equation is solved for j using Newton's method, where the $(i+1)$-th approximation of j is given by

$$j_{i+1} = j_i - \frac{r(j_i)}{r'(j_i)}. \tag{18}$$

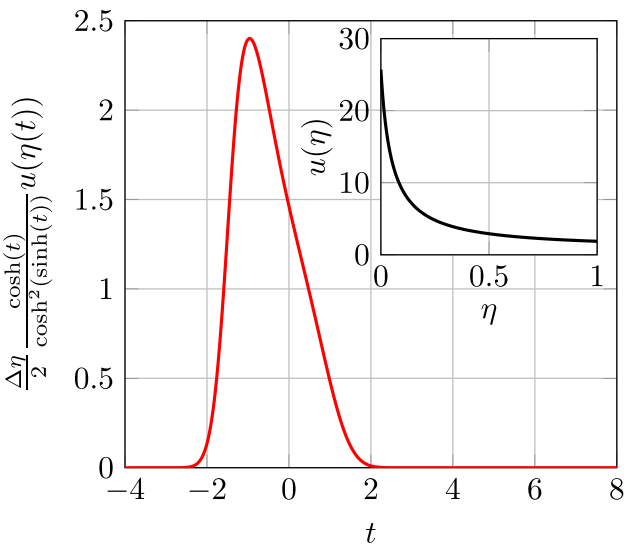

Fig. 3 Original (inset) and transformed integrand for $\Delta\varphi = 1$, $j \approx 0.735$, $\eta_0 = -0.05$, $\eta_\mathrm{L} = 0$ and $\eta_\mathrm{R} = 1$

The calculation of both r and its derivative with respect to j is straightforward, with

$$r(j) = \sum_k w_k \frac{\mathcal{F}(\eta_k)}{\mathcal{F}(\eta_k)\Delta\varphi - j} - 1 \tag{19}$$

and

$$r'(j) = \sum_k w_k \frac{\mathcal{F}(\eta_k)}{\left(\mathcal{F}(\eta_k)\Delta\varphi - j\right)^2}. \tag{20}$$

Similar to [8], we use the diffusion-enhanced Scharfetter–Gummel scheme [4, 5] as an initial guess

$$j_0 = g\left(B\left(-\frac{\Delta\varphi}{g}\right)\mathcal{F}_\mathrm{L} - B\left(\frac{\Delta\varphi}{g}\right)\mathcal{F}_\mathrm{R}\right) \tag{21}$$

with the enhancement factor

$$g = \frac{\Delta\eta}{\log(\mathcal{F}_\mathrm{R}) - \log(\mathcal{F}_\mathrm{L})}. \tag{22}$$

To increase stability, we combine Newton's method with the bisection method, i.e., we keep track of an upper and lower bound between which the solution must be located. If the Newton update would lead to a value outside of that range, we disregard it and use the interval midpoint instead. At every step of the iteration, the bounds are updated accordingly. This procedure ensures that the iteration always converges to the correct solution. To get an initial range, we use the mean value theorem and the monotonicity of $\mathcal{F}$:

$$\mathcal{F}(\eta^{\min})|\Delta\varphi - \Delta\eta| \le |j| \le \mathcal{F}(\eta^{\max})|\Delta\varphi - \Delta\eta| \tag{23}$$

with $\eta^{\min} = \min(\eta_\mathrm{L}, \eta_\mathrm{R})$ and $\eta^{\max} = \max(\eta_\mathrm{L}, \eta_\mathrm{R})$.

Depending on η_L, η_R and $\Delta\varphi$, we can further reduce this range, improving the speed of convergence in cases where bisection is needed.

(i) For $\Delta\eta < 0$, we know that $\frac{\partial\eta}{\partial x} < 0$; therefore, the lower bound of j can be updated with

$$j > \Delta\varphi\mathcal{F}(\eta_\mathrm{L}). \tag{24}$$

(ii) Similarly, $0 < \Delta\eta < \Delta\varphi$ implies that $\frac{\partial\eta}{\partial x} > 0$, thus

$$j < \Delta\varphi\mathcal{F}(\eta_\mathrm{L}). \tag{25}$$

(iii) Lastly, for $\Delta\eta > \Delta\varphi$ we know that $j < 0$ and $\frac{\partial\eta}{\partial x} > \Delta\varphi$. Furthermore,

$$\frac{\partial^2\eta}{\partial x^2} = j\frac{\mathcal{F}'(\eta)}{\mathcal{F}(\eta)^2}\left(\Delta\varphi - \frac{j}{\mathcal{F}(\eta)}\right) < 0 \tag{26}$$

since $\mathcal{F}'(\eta) > 0$. We conclude that

$$\begin{aligned}\eta(x) &= \eta_\mathrm{R} + \int_1^x \frac{\partial\eta}{\partial x'}\mathrm{d}x' \\ &\le \eta_\mathrm{R} + \int_1^x \left.\frac{\partial\eta}{\partial x'}\right|_{x'=0} \mathrm{d}x' \\ &= \eta_\mathrm{R} + \left(\Delta\varphi - \frac{j}{\mathcal{F}(\eta_\mathrm{L})}\right)(x-1) \\ &\le \eta_\mathrm{R} + \Delta\varphi(x-1)\,.\end{aligned} \tag{27}$$

Thus, we get an estimate for $\frac{\partial\eta}{\partial x}$ which does not depend on η

$$\frac{\partial\eta}{\partial x} \ge \Delta\varphi - \frac{j}{\mathcal{F}(\eta_\mathrm{R} + \Delta\varphi(x-1))}. \tag{28}$$

Integrating over x then yields an update for the lower bound of j with

$$j \ge \frac{\Delta\varphi(\Delta\varphi - \Delta\eta)}{\int_{\eta_\mathrm{R}-\Delta\varphi}^{\eta_\mathrm{R}} (\mathcal{F}(\eta))^{-1}\mathrm{d}\eta}, \tag{29}$$

where we can calculate the integral similar to before using Tanh–Sinh quadrature.

The iteration is stopped once accuracy close to machine precision has been achieved using the criterion

$$|\Delta j| \le \epsilon_r|j| + \epsilon_a, \tag{30}$$

where we use a relative tolerance of $\epsilon_r = 2 \cdot 10^{-14}$ and an absolute tolerance of $\epsilon_a = 10^{-16}$.

Once the solution has been found, we use implicit differentiation to calculate derivatives of j with respect to η_L, η_R and $\Delta\varphi$, e.g.,

$$\frac{\partial j}{\partial\eta_\mathrm{L}} = -\frac{\frac{\partial r}{\partial\eta_\mathrm{L}}}{\frac{\partial r}{\partial j}}. \tag{31}$$

Note that we do not use the Leibniz integral rule here, instead we directly calculate the needed derivatives of r using the explicit quadrature formula (see Eq. (19)). This involves calculating the derivatives of the quadrature nodes η_k and weights w_k.

4.3 Solution in special cases

If either $\Delta\varphi$ or $\Delta\eta$ or both are small, the general method will become numerically unstable [8]. Additionally, thermodynamic consistency requires that j is exactly zero for $\Delta\varphi = \Delta\eta$. To address this, we develop series expansions for each of the special cases given in Table 1. These expansions have the additional advantage that the solution is obtained non-iteratively, reducing the runtime significantly.

4.3.1 Case 1a

For $\Delta\varphi \approx \Delta\eta$, the current is $j \approx 0$. We replace the residual by the two-dimensional Taylor series around $j = 0, \Delta\varphi = \Delta\eta$

$$r(j, \Delta\varphi) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{r_{k,l}}{k!\, l!} j^k (\Delta\varphi - \Delta\eta)^l \overset{!}{=} 0, \tag{32}$$

with

$$r_{0,0} = 0 \quad \text{and} \quad r_{k,l} = (-1)^l (k+l)!\, \Delta\eta^{-k-l-1} J_k, \tag{33}$$

where

$$I_k = \int_{\eta_\mathrm{L}}^{\eta_\mathrm{R}} \mathcal{F}(\eta)^k \mathrm{d}\eta \quad \text{and} \quad J_k = I_{-k}. \tag{34}$$

Except for the trivial value $I_0 = \Delta\eta$, these quantities are calculated using Tanh–Sinh quadrature.

Subsequently, the current density is written as an expansion in $\Delta\varphi - \Delta\eta$

$$j = \sum_{i=1}^{\infty} j_i (\Delta\varphi - \Delta\eta)^i. \tag{35}$$

The coefficients are evaluated by implicit differentiation of Eq. (32) (e.g., $j_1 = -r_{0,1}/r_{1,0}$) and are given up to fifth order in Table 2. Note that using Eq. (35) guarantees thermodynamic consistency exactly, i.e., $j = 0$ for $\Delta\varphi = \Delta\eta$.

4.3.2 Case 1b

For $\Delta\eta \approx 0$, the current is $j \approx \Delta\varphi \mathcal{F}(\overline{\eta})$. We write

$$\mathcal{F}(\eta) = \sum_{l=0}^{\infty} \frac{\mathcal{F}_l}{l!} \left(\eta - \overline{\eta}\right)^l \quad \text{with} \quad \mathcal{F}_l = \mathcal{F}^{(l)}(\overline{\eta}). \tag{36}$$

The current is expanded in terms of $\Delta\eta$ with

$$j = \sum_{i=0}^{\infty} j_i \Delta\eta^i \quad \text{and} \quad j_0 = \Delta\varphi \mathcal{F}_0. \tag{37}$$

Inserting both expressions into Eq. (9) and applying the transformation $\eta = \overline{\eta} + t\Delta\eta$ yields

$$\int_{-1/2}^{1/2} \underbrace{\frac{\sum_{i=0}^{\infty} \frac{\mathcal{F}_i}{i!} t^i \Delta\eta^i}{\sum_{i=1}^{\infty} \left(\frac{\mathcal{F}_i}{i!} t^i \Delta\varphi - j_i\right) \Delta\eta^{i-1}}}_{a(t)} dt \overset{!}{=} 1 \tag{38}$$

Next, the transformed integrand is also expanded in terms of $\Delta\eta$ with

$$a(t) = \sum_{k=0}^{\infty} a_k(t) \Delta\eta^k. \tag{39}$$

The first few coefficients are given by

$$\begin{aligned}
a_0 &= -\frac{\mathcal{F}_0}{j_1 - \mathcal{F}_1 t \Delta\varphi} \\
a_1 &= \frac{2\mathcal{F}_0 j_2 - 2\mathcal{F}_1 j_1 t + 2\mathcal{F}_1^2 t^2 \Delta\varphi - \mathcal{F}_0 \mathcal{F}_2 t^2 \Delta\varphi}{2\left(j_1 - \mathcal{F}_1 t \Delta\varphi\right)^2} \\
a_2 &= \frac{\mathcal{F}_0 j_3 - \mathcal{F}_0 \mathcal{F}_3 t^3 \frac{\Delta\varphi}{6}}{\left(j_1 - \mathcal{F}_1 t \Delta\varphi\right)^2} - \frac{\mathcal{F}_0 \left(j_2 - \mathcal{F}_2 t^2 \frac{\Delta\varphi}{2}\right)^2}{\left(j_1 - \mathcal{F}_1 t \Delta\varphi\right)^3} \\
&\quad - \frac{\mathcal{F}_2 t^2}{2\left(j_1 - \mathcal{F}_1 t \Delta\varphi\right)} + \frac{\mathcal{F}_1 t \left(j_2 - \mathcal{F}_2 t^2 \Delta\varphi\right)}{\left(j_1 - \mathcal{F}_1 t \Delta\varphi\right)^2}.
\end{aligned} \tag{40}$$

Integrating the series expansion and solving Eq. (38) separately for each exponent of $\Delta\eta$ finally results in the coefficients given in Table 3.

4.3.3 Case 2a

For $\Delta\varphi \approx 0$, the current is $j \approx -I_1$. Again, the residual is replaced by a two-dimensional Taylor series around $j = -I_1, \Delta\varphi = 0$

$$r(j, \Delta\varphi) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{r_{k,l}}{k!\, l!} \left(j + I_1\right)^k \Delta\varphi^l, \tag{41}$$

with

$$r_{0,0} = 0 \quad \text{and} \quad r_{k,l} = (-1)^l (k+l)! \frac{I_{l+1}}{I_1^{k+l+1}}. \tag{42}$$

The current is then given by

$$j = \sum_{i=0}^{\infty} j_i \Delta\varphi^i, \tag{43}$$

where the coefficients are given in Table 4. Compared to [8] where only the first term of the series was used, our expression is more accurate.

4.3.4 Case 2b

For $\Delta\varphi \approx 0$ and $\Delta\eta \approx 0$, we start with the result from Case 2a and use the same expansion for $\mathcal{F}(\eta)$ as in Case 1b (Eq. (36)). Subsequently, the integrals I_k and the coefficients j_i from Table 4 are expanded in terms of $\Delta\eta$. Ignoring higher-order terms, we get

$$j = (\Delta\varphi - \Delta\eta)\Bigg(\mathcal{F}_0 - \frac{\mathcal{F}_1^2}{12\mathcal{F}_0}\Delta\eta\Delta\varphi + \frac{\mathcal{F}_2}{24}\Delta\eta^2 + \mathcal{O}(\Delta\eta\Delta\varphi^3) + \mathcal{O}(\Delta\eta^2\Delta\varphi^2) + \mathcal{O}(\Delta\eta^3\Delta\varphi) + \mathcal{O}(\Delta\eta^4)\Bigg), \tag{44}$$

which corresponds to the expression developed in [16]. Note that there is no term $\mathcal{O}(\Delta\varphi^n)$ inside the parentheses; nevertheless, in contrast with [8] we do not use this formula for Case 1b, as terms proportional to $\Delta\eta\Delta\varphi^k$ exist for arbitrarily large k. Unless $\Delta\eta$ is exactly zero, it is therefore questionable to use this expression for large $\Delta\varphi$. However, note that for small $\Delta\varphi$ the expansion is consistent with Eq. (37).

4.4 Algorithm

To solve the discretized continuity equation, we need the edge current density J_{edge} as a function of $n_{\text{L/R}}$ and $\varphi_{\text{L/R}}$. The first step is to compute $\eta_{\text{L/R}} = \mathcal{F}^{-1}(\frac{n_{\text{L/R}}}{n_0})$ and $\Delta\varphi = \frac{\varphi_{\text{R}}-\varphi_{\text{L}}}{V_{\text{T}}}$. Next, j is calculated using the procedure outlined in Algorithm 1, where $\epsilon_{\text{c}} = 10^{-3}$ is used as a numerical tolerance. See [10] for the actual implementation as well as some example code. Finally, j is denormalized using $J_{\text{edge}} = \frac{qe\mu V_{\text{T}} n_0}{l} j$. To assemble the Jacobian of the discretized system, in addition the derivatives of J_{edge} with respect to the nodal densities and potentials are needed. These can be easily computed from the derivatives of j w.r.t. η_{L}, η_{R} and $\Delta\varphi$ by using the chain rule.

Algorithm 1 Summary of solution method

```
if Δφ < 0 then
    Flip edge direction and repeat.
else if Δφ > ε_c then
    if |Δη − Δφ| ≤ ε_c Δφ then
        Use series expansion Eq. (35).
    else if |Δη| ≤ ε_c then
        Use series expansion Eq. (37).
    else
        Use Newton's method to solve Eq. (17).
    end if
else
    if |Δη| > ε_c then
        Use series expansion Eq. (43).
    else
        Use series expansion Eq. (44).
    end if
end if
```

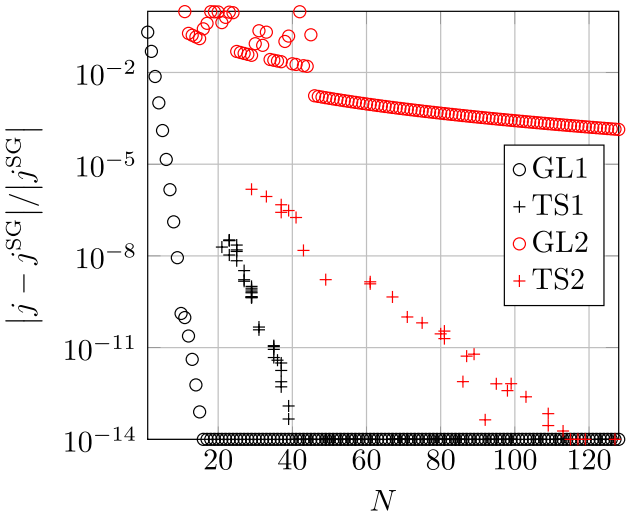

Fig. 4 Relative error of the current density versus number of quadrature points for parabolic bands with Maxwell–Boltzmann statistics using $\eta_1 = -10$, $\eta_2 = -5$ and $\Delta\varphi = 3$ (case 1) and $\Delta\varphi = -10$ (case 2). We compare Gauss–Legendre quadrature (GL1 and GL2) to the adaptive Tanh–Sinh quadrature (TS1 and TS2)

5 Results

5.1 Comparison

We start by comparing our work with the existing scheme developed in [3, 8], where Gauss–Legendre and similar quadrature rules were used to calculate the integral. To get a reference, we limit this comparison to parabolic bands in conjunction with Maxwell–Boltzmann statistics, since for this case the exact solution is known (see Eq. (10)). We measure the performance of the respective scheme by comparing the relative error of the current to the number of quadrature points used in the integration. As it is difficult to directly control the number of points in the adaptive scheme, we simply modify the relative tolerance in the integration algorithm instead. To increase the variation in the number of points and show the trend more clearly, for each value we also randomly choose an initial step size for the trapezoidal rule between 0.5 and 1.0 (other than for this comparison it is simply set to 1.0).

The results are shown in Fig. 4 for two different cases. While in the first case Gauss–Legendre quadrature performs excellently and gives a highly accurate result close to machine precision with less than 20 points, in the second case the relative error is still above 10^{-4} even for 128 points. The cause for this is the existence of a pole close to the integration interval, and the situation can be made arbitrarily worse by changing $\Delta\varphi$ to even larger negative values.

In contrast with this, using the simple adaptive Tanh–Sinh quadrature rule the pole is handled without problems and

the number of required points for the same precision is only moderately increased. Even though it does not perform as good as Gauss–Legendre quadrature in the first case and needs roughly twice as many points for a similar accuracy, overall it is far more robust as it works well even in difficult cases.

As an additional test, we assess the stability properties of both schemes. We recall that in order to achieve a stable solution in a device simulation $\frac{\partial j}{\partial \mathcal{F}_\mathrm{L}} > 0$ and $\frac{\partial j}{\partial \mathcal{F}_\mathrm{R}} < 0$ is necessary, such that the resulting Jacobian of the discretized system is an M-matrix [17, 18]. This is clearly fulfilled for the Scharfetter–Gummel scheme (Eq. (10)) which is the main reason for its excellent stability properties. In [12], it was shown that this is also the case for the generalized Scharfetter–Gummel scheme, under the assumption that the integral equation is solved exactly.

In order to test this property for the two numerical implementations, we compare the solution that both schemes give by varying the chemical potential at one of the edge endpoints (see Fig. 5). The Tanh–Sinh solution fits perfectly to the analytical solution and is monotonous, leading to the same excellent stability properties as the Scharfetter–Gummel scheme. Meanwhile, the solution obtained using Gauss–Legendre quadrature is not only inaccurate for small η_R, it also violates the monotonicity of j. This essentially disqualifies it from being used in a full-scale device simulation, as the Jacobian would no longer be guaranteed to be an M-matrix. While increasing the number of quadrature points improves the situation to some extent, the inability of the nonadaptive quadrature to handle arbitrary parameters is also highlighted, as the number of required points is clearly dependent on η_R (as well as the other parameters) and cannot be simply set to a fixed value.

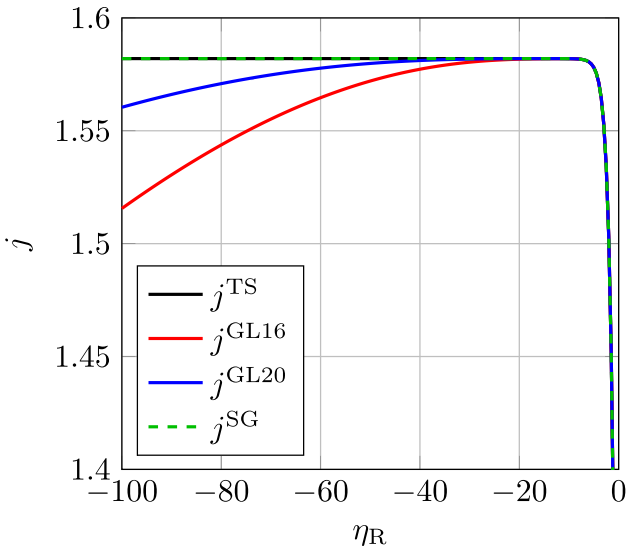

Fig. 5 Current density versus chemical potential at the right-hand side of the edge for parabolic bands with Maxwell–Boltzmann statistics for $\eta_\mathrm{L} = 0$ and $\Delta\varphi = 1$. Comparison between Tanh–Sinh (TS) quadrature, Gauss–Legendre with 16 and 20 points (GL16, GL20) and the analytical solution given by the Scharfetter–Gummel (SG) scheme

5.2 Band tailing model

To demonstrate the ability of our scheme to work with arbitrary band structures, we use a simple model for band tailing [19, 20]. An exponential band tail is added to the parabolic density of states by introducing a parameter t_0 below which exponential behavior dominates. For simplicity, we model $Z(t)$ to be continuously differentiable, so t_0 is the only free parameter [21]:

$$Z(t) = \begin{cases} \frac{2}{\sqrt{\pi}}\sqrt{t_0}e^{\frac{t-t_0}{2t_0}} & \text{for} \quad t < t_0 \\ \frac{2}{\sqrt{\pi}}\sqrt{t} & \text{for} \quad t \geq t_0. \end{cases} \tag{45}$$

Though rather simplistic, in conjunction with Fermi–Dirac statistics such a model can be used to simulate the band tailing effect present in MOSFETs at cryogenic temperatures.

In Fig. 6, the resulting density of states is shown for different values of t_0. For $t_0 \to 0$, we get the original square root behavior from Eq. (5).

$\mathcal{F}(\eta)$ is calculated from Eq. (4) using Sinh–Sinh quadrature. As described in Sect. 4.1, a lookup table is created to speed up subsequent calculations. The resulting table size is small at approximately 200 KiB depending on the value of t_0. In Fig. 7, the cumulative distribution function is shown. While for large η the value of t_0 seems to have no visible influence, it reduces the slope of the exponential Boltzmann tail significantly.

In Fig. 8, the current density versus the potential drop is shown for $\Delta\eta = 10$ and different values of t_0. Since larger values of t_0 result in significantly higher densities at small η, $|j|$ is also increased. However, it is clearly visible that independent on the value of t_0 the current is zero for $\Delta\varphi = \Delta\eta$, which demonstrates the exact thermodynamic consistency of our numerical scheme. For comparison, the diffusion-enhanced approximation is also plotted. While this approximation agrees well with the exact solution for positive j, noticeable deviations appear for negative j. Nevertheless, it remains thermodynamically consistent, as the current still vanishes at $\Delta\varphi = \Delta\eta$.

5.3 Organic semiconductors

In organic semiconductors, charge transport is better described by hopping between individual sites instead of using continuous energy bands. By assuming a Gaussian

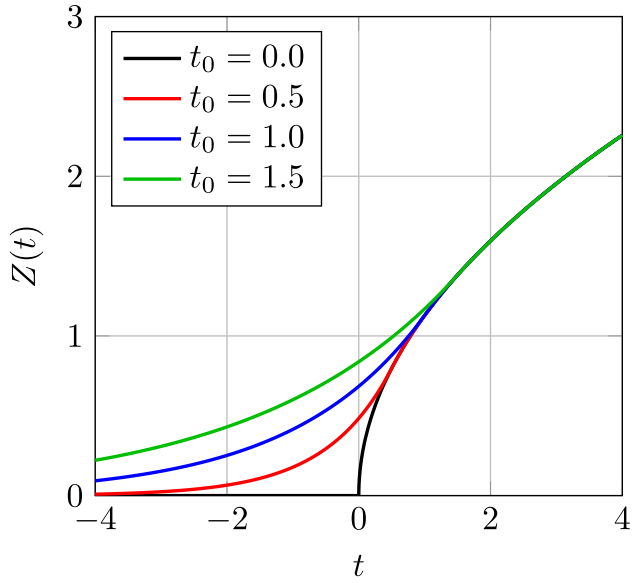

Fig. 6 Normalized density of states for parabolic bands including band tailing

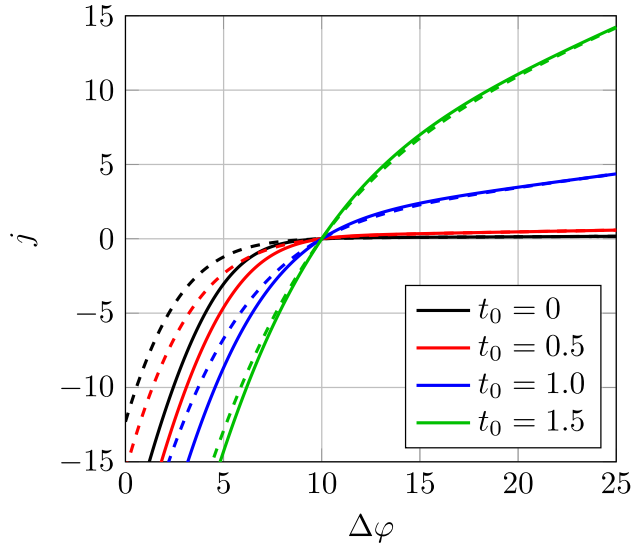

Fig. 8 Current versus potential drop for $\eta_L = -5$ and $\eta_R = 5$ for parabolic bands with band tailing and Fermi–Dirac statistics. Dashed lines indicate the diffusion-enhanced approximation (Eq. (21))

density of states [22–24] and Fermi–Dirac statistics, the individual energy sites can be lumped together and transport can be modeled using a DD model. The form of Z is then given by

$$Z(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{t^2}{2\sigma^2}) \tag{46}$$

where σ is a material parameter, with typical values of about $4 - 5$ [25].

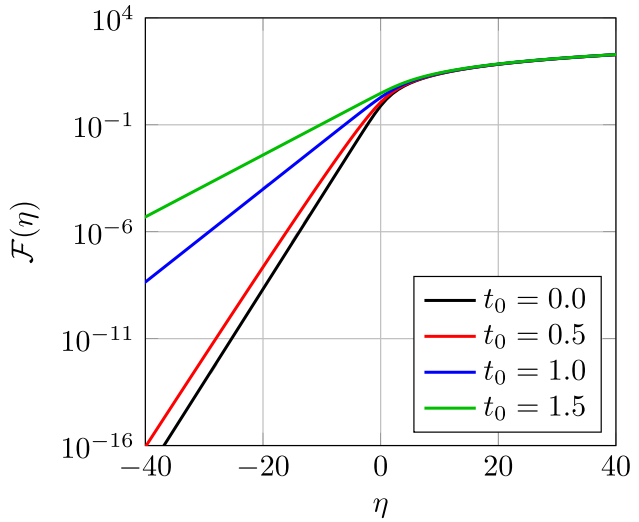

Fig. 7 Cumulative distribution function for parabolic bands including band tailing and Fermi–Dirac statistics

In Fig. 9, the cumulative distribution function is shown for various values of σ. It is calculated like before using Sinh–Sinh quadrature, and a lookup table is created for each value of σ. As opposed to before, $\mathcal{F}(\eta)$ is limited and does never increase above the value of 1, regardless of how large η gets, since $Z(t)$ is absolutely integrable.

In Fig. 10, the current density versus the chemical potential difference is plotted for a fixed value of $\Delta\varphi$. As before, it vanishes for $\Delta\eta = \Delta\varphi$, demonstrating thermodynamic consistency of the numerical scheme once more.

6 Conclusion

In this paper, we presented a generalization of the SG stabilization scheme to degenerate semiconductors improving an existing scheme from the literature. Compared to before, our method uses a more robust numerical quadrature method and new and improved series expansions for several special cases. The approach works for arbitrary band structures and uses no further approximations as compared to the original SG scheme. We demonstrated the superiority of the new scheme by performing multiple tests and also showed that the previous scheme can lead to problems in the case of full-scale device simulations. As opposed to before, thermodynamic consistency is also guaranteed exactly and not subjected to numerical noise.

While the approach is somewhat complex, especially compared to the original SG scheme, the CPU penalty in actual device simulations is rather small. Solving for the

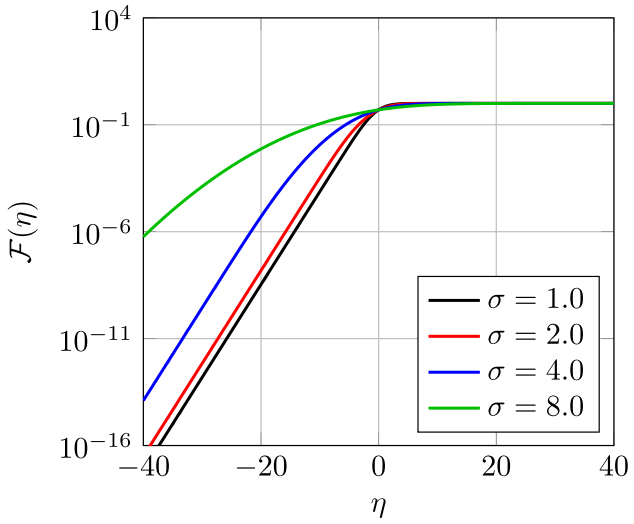

Fig. 9 Gauss–Fermi integral for different values of σ

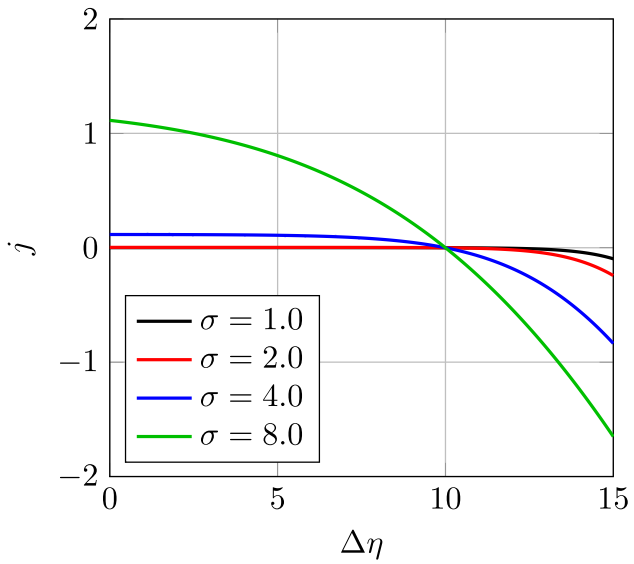

Fig. 10 Current versus chemical potential drop with $\eta_L = -10$ and $\Delta\varphi = 10$ for Gauss–Fermi distribution function

edge currents can be efficiently parallelized, such that the effect on runtime is almost negligible.

Appendix A Series expansion coefficients

See Tables 2, 3 and 4.

Table 2 Coefficients for $\Delta\varphi \approx \Delta\eta$ (Eq. (35))

j_1	I_0/J_1
j_2	$(J_1^2 - J_2 I_0)/J_1^3$
j_3	$(2J_2^2 I_0 - J_1^2 J_2 - J_1 J_3 I_0)/J_1^5$
j_4	$(2J_1^2 J_2^2 - J_1^3 J_3 - 5J_2^3 I_0 - J_1^2 J_4 I_0 + 5J_1 J_2 J_3 I_0)/J_1^7$
j_5	$(14 I_0 J_2^4 - J_1^4 J_4 - 5J_1^2 J_2^3 - I_0 J_1^3 J_5 + 5J_1^3 J_2 J_3$ $+3I_0 J_1^2 J_3^2 - 21 I_0 J_1 J_2^2 J_3 + 6 I_0 J_1^2 J_2 J_4)/J_1^9$

Table 3 Coefficients for $\Delta\eta \approx 0$ (Eq. (37))

j_0	$\Delta\varphi\mathcal{F}_0$
j_1	$-\frac{C}{2}\left(\exp\left(\frac{C}{\mathcal{F}_0}\right)+1\right)/\left(\exp\left(\frac{C}{\mathcal{F}_0}\right)-1\right)$
j_2	$\frac{C^2\mathcal{F}_0(A+\mathcal{F}_1^2)+2Aj_1(C^2-4j_1(\mathcal{F}_0+j_1))}{8C\mathcal{F}_0^3\mathcal{F}_1}$
j_3	$\Big(192A^2 j_1^5 + 48\mathcal{F}_0 j_1^4(8A^2 - B\mathcal{F}_0^2 + 2A\mathcal{F}_0\mathcal{F}_2)$ $-24j_1^3(2B\mathcal{F}_0^4 + A^2(3C^2 - 8\mathcal{F}_0^2) - 4A\mathcal{F}_0^3\mathcal{F}_2)$ $-12C^2\mathcal{F}_0 j_1^2(10A^2 - B\mathcal{F}_0^2 + 3A\mathcal{F}_0\mathcal{F}_2)$ $-4C^2\mathcal{F}_0^2 j_1(12A^2 - \mathcal{F}_0^2(3B + \mathcal{F}_1\mathcal{F}_3) + 6\mathcal{F}_0\mathcal{F}_2 A)$ $+3C^4(A^2(\mathcal{F}_0 + 2j_1) + A\mathcal{F}_0\mathcal{F}_1^2)\Big)/\left(96C^2\mathcal{F}_0^4\mathcal{F}_1^2\right)$
A	$\mathcal{F}_1^2 - \mathcal{F}_0\mathcal{F}_2$
B	$\mathcal{F}_2^2 - \mathcal{F}_1\mathcal{F}_3$
C	$\mathcal{F}_1\Delta\varphi$

Table 4 Coefficients for $\Delta\varphi \approx 0$ (Eq. (43))

j_0	$-I_1$
j_1	I_2/I_1
j_2	$(I_2^2 - I_1 I_3)/I_1^3$
j_3	$(I_1^2 I_4 + 2I_2^3 - 3I_1 I_2 I_3)/I_1^5$
j_4	$(5I_2^4 - I_1^3 I_5 + 2I_1^2 I_3^2 - 10 I_1 I_2^2 I_3 + 4I_1^2 I_2 I_4)/I_1^7$

Author contributions T.L. and M.R. developed the algorithm. T.L. wrote the program code and original manuscript draft and prepared the figures. M.R. tested the program code. Supervision was done by C.J. All authors reviewed the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants or other support was received during the preparation of this manuscript.

Data availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors have no relevant financial or nonfinancial interests to disclose.

References

1. Scharfetter, D.L., Gummel, H.K.: Large-signal analysis of a silicon Read diode oscillator. IEEE Trans. Electron Devices **16**(1), 64–77 (1969)
2. Koprucki, T., Gärtner, K.: Discretization scheme for drift-diffusion equations with strong diffusion enhancement. Opt. Quant. Electron. **45**(7), 791–796 (2013)
3. Farrell, P., Patriarca, M., Fuhrmann, J., Koprucki, T.: Comparison of thermodynamically consistent charge carrier flux discretizations for Fermi-Dirac and Gauss-Fermi statistics. Opt. Quant. Electron. **50**(2), 101 (2018)
4. Bessemoulin-Chatard, M.: A finite volume scheme for convection-diffusion equations with nonlinear diffusion derived from the Scharfetter–Gummel scheme. Numer. Math. **121**, 637–670 (2012)
5. Koprucki, T., Rotundo, N., Farrell, P., Doan, D.H., Fuhrmann, J.: On thermodynamic consistency of a Scharfetter–Gummel scheme based on a modified thermal voltage for drift-diffusion equations with diffusion enhancement. Opt. Quant. Electron. **47**, 1327–1332 (2015)
6. Fuhrmann, J.: Comparison and numerical treatment of generalised Nernst-Planck models. Comput. Phys. Commun. **196**, 166–178 (2015)
7. Stanojević, Z., González-Medina, J.M., Schanovsky, F., Karner, M.: Quasi-Fermi-based charge transport scheme for device simulation in cryogenic, wide bandgap, and high-voltage applications. IEEE Trans. Electron Devices **70**(2), 708–713 (2022)
8. Patriarca, M., Farrell, P., Fuhrmann, J., Koprucki, T.: Highly accurate quadrature-based Scharfetter–Gummel schemes for charge transport in degenerate semiconductors. Comput. Phys. Commun. **235**, 40–49 (2019)
9. Renner, M., Linn, T., Jungemann, C.: Stabilization of the drift-diffusion model for arbitrary carrier statistics. In: 2025 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (2025). In press
10. Linn, T.: TobiasLinnITHE/generalized_SG. https://doi.org/10.5281/zenodo.15781506
11. Fukushima, T.: Precise and fast computation of Fermi-Dirac integral of integer and half integer order by piecewise minimax rational approximation. Appl. Math. Comput. **259**, 708–729 (2015)
12. Gärtner, K.: Existence of bounded discrete steady state solutions of the van Roosbroeck system with monotone Fermi-Dirac statistic functions. J. Comput. Electron. **14**, 773–787 (2015)
13. Takahasi, H., Mori, M.: Double exponential formulas for numerical integration. Public. Res. Inst. Math. Sci. **9**(3), 721–741 (1974)
14. Evans, G., Forbes, R., Hyslop, J.: The tanh transformation for singular integrals. Int. J. Comput. Math. **15**(1–4), 339–358 (1984)
15. Takahashi, H.: Error estimation in the numerical integration of analytic functions. Report Comput. Center **3**, 41–108 (1970)
16. Farrell, P., Koprucki, T., Fuhrmann, J.: Computational and analytical comparison of flux discretizations for the semiconductor device equations beyond Boltzmann statistics. J. Comput. Phys. **346**, 497–513 (2017)
17. Varga, R.S.: Matrix Iterative Analysis. Series in Automatic Computation. Prentice-Hall, Englewood Cliffs, New Jersey (1962)
18. Meinerzhagen, B., Jungemann, C.: Spherical harmonics expansion and multi-scale modeling. In: Rudan, M., Brunetti, R., Reggiani, S., *et al.* (eds.) Springer Handbook of Semiconductor Devices, pp. 1413–1450. Springer, Cham, Switzerland (2023). Chap. 39
19. Bohuslavskyi, H., Jansen, A.G.M., Barraud, S., Barral, V., Cassé, M., Guevel, L., Jehl, X., Hutin, L., Bertrand, B., Billiot, G., Pillonnet, G., Arnaud, F., Galy, P., Franceschi, S., Vinet, M., Sanquer, M.: Cryogenic subthreshold swing saturation in FD-SOI MOSFETs described with band broadening. IEEE Electron Device Lett. **40**(5), 784–787 (2019)
20. Beckers, A., Jazaeri, F., Enz, C.: Theoretical limit of low temperature subthreshold swing in field-effect transistors. IEEE Electron Device Lett. **41**(2), 276–279 (2020)
21. Jungemann, C., Richstein, B., Linn, T., Knoch, J.: Device modeling for admittance spectroscopy of PMOSFETs at cryogenic temperatures. IEEE Trans. Electron Devices **71**(4), 2322–2328 (2024)
22. Bässler, H.: Charge transport in disordered organic photoconductors. a Monte Carlo simulation study. Physica Status Solidi B (Basic Research);(Germany) **175**(1) (1993)
23. Novikov, S.V., Dunlap, D.H., Kenkre, V.M., Parris, P.E., Vannikov, A.V.: Essential role of correlations in governing charge transport in disordered organic materials. Phys. Rev. Lett. **81**(20), 4472 (1998)
24. Hulea, I., Brom, H., Houtepen, A., Vanmaekelbergh, D., Kelly, J., Meulenkamp, E.: Wide energy-window view on the density of states and hole mobility in poly (p-phenylene vinylene). Phys. Rev. Lett. **93**(16), 166601 (2004)
25. Martens, H., Blom, P., Schoo, H.: Comparative study of hole transport in poly (p-phenylene vinylene) derivatives. Phys. Rev. B **61**(11), 7489 (2000)