Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Consistency and convergence for a family of finite volume discretizations of the Fokker–Planck operator

Martin Heida, Markus Kantner, Artur Stephan

submitted: February 5, 2020

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: martin.heida@wias-berlin.de
markus.kantner@wias-berlin.de
artur.stephan@wias-berlin.de

No. 2684
Berlin 2020

2010 *Mathematics Subject Classification.* 35Q84, 49M25, 65N08.

Key words and phrases. Finite volume, Fokker–Planck, Scharfetter–Gummel, Stolarsky mean, consistency, order of convergence.

M. H. and A. S. are financed by Deutsche Forschungsgemeinschaft (DFG) through Grant CRC 1114 “Scaling Cascades in Complex Systems”, Project C05 *Effective models for materials and interfaces with multiple scales*. The work of M. K. received funding from the DFG under Germany’s Excellence Strategy – EXC2046: MATH+ (Berlin Mathematics Research Center).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Consistency and convergence for a family of finite volume discretizations of the Fokker–Planck operator

Martin Heida, Markus Kantner, Artur Stephan

Abstract

We introduce a family of various finite volume discretization schemes for the Fokker–Planck operator, which are characterized by different weight functions on the edges. This family particularly includes the well-established Scharfetter–Gummel discretization as well as the recently developed square-root approximation (SQRA) scheme. We motivate this family of discretizations both from the numerical and the modeling point of view and provide a uniform consistency and error analysis. Our main results state that the convergence order primarily depends on the quality of the mesh and in second place on the quality of the weights. We show by numerical experiments that for small gradients the choice of the optimal representative of the discretization family is highly non-trivial while for large gradients the Scharfetter–Gummel scheme stands out compared to the others.

1 Introduction

The Fokker–Planck equation (FPE), also known as *Smoluchowski equation* or *Kolmogorov forward equation*, is one of the major equations in theoretical physics and applied mathematics. It describes the time evolution of the probability density function of a particle in an external force field (e.g., fluctuating forces as in Brownian motion). The equation can be generalized to other contexts and observables and has been employed in a broad range of applications, including physical chemistry, protein synthesis, plasma physics and semiconductor device simulation. Thus, there is a huge interest in the development of efficient and robust numerical methods. In the context of finite volume (FV) methods, the central objective is a robust and accurate discretization of the (particle or probability) flux implied by the FPE.

A particularly important discretization scheme for the flux was derived by Scharfetter and Gummel [SG69] in the context of the drift-diffusion model for electronic charge carrier transport in bipolar semiconductor devices [vR50]. The typically exponentially varying carrier densities at p-n junctions lead to unphysical results (spurious oscillations), if the flux is discretized in a naive way using standard finite difference schemes [MW94]. The problem was overcome by considering the flux expression as a one-dimensional boundary value problem along each edge between adjacent mesh nodes. The resulting Scharfetter–Gummel (SG) scheme provides a robust discretization of the flux as it asymptotically approaches the numerically stable discretizations in the drift- (upwind scheme) and diffusion-dominated (central finite difference scheme) limits. The SG-scheme and its several generalizations to more complex physical problem settings are nowadays widely used in semiconductor device simulation [Mar86, FRD$^+$17] and have been extensively studied in the literature [BMP89, EFG06, FKF17, Kan20]. The SG-scheme is also known as *exponential fitting scheme* and was independently discovered by Allan and Southwell [AS55] and Il'in [Il'69] in different contexts.

DOI 10.20347/WIAS.PREPRINT.2684

Recently, an alternative flux discretization method, called *square-root approximation* (SQRA) scheme, has been derived explicitly for high dimensional problems. The original derivation in [LFW13] aims at applications in molecular dynamics and is based on Markov state models. However, it can also be obtained from a maximum entropy path principle [DJSD15] and from discretizing the Jordan–Kinderlehrer–Otto variational formulation of the FPE [Mie13a]. In Section 3.3, we provide a derivation of SQRA scheme, which is motivated from the theory of gradient flows. In contrast to the SG-scheme, the SQRA is very recent and only sparsely investigated. The only contributions on the convergence seem to be [Mie13a] in 1D, [DHWK] (formally, rectangular meshes) and [Hei18] using G-convergence.

The SG and the SQRA schemes both turn out to be special cases of a family of discretization schemes based on weighted Stolarsky means, see Section 3.2. This family is very rich and allows for a general convergence and consistency analysis, which we carry out in Sections 4–5. Interestingly, there seems to be no previous results in the literature.

1.1 The FPE and the SG and SQRA discretization schemes

In this work, we consider the stationary Fokker–Planck equation

$$-\nabla\cdot(\kappa\nabla u)-\nabla\cdot(\kappa u\nabla V)=f, \tag{1.1}$$

which can be equivalently written as

$$\operatorname{div}\mathbf{J}(u,V)=f$$

using the flux $\mathbf{J}(u,V)=-\kappa\left(\nabla u+u\nabla V\right)$, where $\kappa>0$ is a (possibly space-dependent) diffusion coefficient and $V:\Omega\to\mathbb{R}$ is a given potential. The flux $\mathbf{J}$ consists of a diffusive part $\kappa\nabla u$ and a drift part $\kappa u\nabla V$, which compensate for the stationary density $\pi=\mathrm{e}^{-V}$ (also known as the Boltzmann distribution) as $\mathbf{J}(\mathrm{e}^{-V},V)=0$. This reflects the principle of detailed balance in the thermodynamic equilibrium. The right-hand side f describes possible sink or source terms.

The SG and the SQRA schemes of the Fokker–Planck operator $\operatorname{div}\mathbf{J}(u,V)$ that are considered below are given in the form

$$\left(\mathcal{F}_B^{\mathcal{T}}u\right)_i:=-\sum_{j:\,i\sim j}\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(B\left(V_i,V_j\right)u_j-B\left(V_j,V_i\right)u_i\right), \tag{1.2}$$

where $\sum_{j:\,j\sim i}$ indicates a sum over all cells adjacent to the i-th cell of the mesh, m_{ij} is the mass of the interface between the i-th and j-th cell, h_{ij} is the distance between the corresponding nodes and κ_{ij} is the discretized diffusion coefficient κ. We are particularly interested in the two cases

$$B\left(V_i,V_j\right)=B_1\left(V_i-V_j\right):=\frac{V_i-V_j}{\mathrm{e}^{V_i-V_j}-1} \tag{1.3}$$

$$\text{or}\quad B\left(V_i,V_j\right)=B_2\left(V_i-V_j\right):=\mathrm{e}^{-\frac{1}{2}\left(V_i-V_j\right)} \tag{1.4}$$

with either the Bernoulli function B_1 (for SG) or with the SQRA-coefficient B_2. The schemes are derived under the assumption of constant flux, diffusion constant and potential gradient along the respective edges.

In the pure diffusion regime, i.e., for $V_i-V_j\to0$, the Bernoulli function provides $B_1\left(V_i-V_j\right)\to1$, such that the SG scheme approaches a discrete analogue of the diffusive part of the continuous flux: $J_{ij}=\kappa_{ij}(u_i-u_j)/h_{ij}$. In the drift-dominated regime, i.e., for $V_j-V_i\to\pm\infty$, the asymptotics of B_1 recover the upwind scheme

$$J_{i,j}\to-\kappa_{i,j}\frac{V_j-V_i}{h_{i,j}}\begin{cases}u_j & \text{if } V_j>V_i\\ u_i & \text{if } V_j<V_i\end{cases}, \tag{1.5}$$

which is a robust discretization of the drift part of the flux, where the density u is evaluated in the donor cell of the flux. Hence, the Bernoulli function B_1 interpolates between the appropriate discretizations for the drift- and diffusion-dominated limits, which is why the SG scheme is the preferred FV scheme for Fokker–Planck type operators. Indeed, the SQRA scheme is consistent with the diffusive limit, but is less accurate than the SG scheme in the case of strong gradients ∇V.

1.2 The Stolarsky mean approximation schemes

In this work, we investigate the relative L^2-distance between the discrete SQRA and SG solutions on the same mesh and the order of convergence of the SQRA scheme, which was an open problem. It turns out that both methods are members of a broad family of finite volume discretizations that stem from the weighted Stolarsky means

$$S_{\alpha,\beta}(x,y) = \left(\frac{\beta\left(x^\alpha - y^\alpha\right)}{\alpha\left(x^\beta - y^\beta\right)}\right)^{\frac{1}{\alpha-\beta}},$$

see Section 3.2. We benefit from the general structure of these schemes and prove order of convergence on consistent meshes in the sense of the recent work [DPD18]. We will see that the error naturally splits into the consistency error for the discretization of the Laplace operator plus an error which is due to the discretization of the stationary solution π and the Stolarsky mean, see Theorem 5.4.

We will demonstrate below that the Stolarsky discretization schemes for (1.1) read

$$-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(\frac{u_j}{\pi_j}-\frac{u_i}{\pi_i}\right) = f_i, \tag{1.6}$$

where $\pi_i = \mathrm{e}^{-V_i}$, $f_i = \int_{\Omega_i} f$ is the integral of f over the i-th cell and $S_{ij} = S_{\alpha,\beta}(\pi_i,\pi_j)$ is a Stolarsky mean of π_i and π_j. We sometimes refer to the general form (1.6) as discrete FPE.

The Stolarsky means $S_{\alpha,\beta}$ generalize Hölder means and other f-means (see Table 2). An interesting aspect of the above representation is that all these schemes preserve positivity with the discrete linear operator being an M-matrix. Furthermore, with the relative density $U = u/\pi$ we arrive at

$$-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(U_j - U_i\right) = f_i,$$

which is a discretization of the elliptic equation

$$-\nabla\cdot\left(\kappa\pi\nabla U\right) = f,$$

where the discrete Fokker–Planck operator becomes a purely diffusive second order operator in U. Furthermore, if κ is a symmetric strictly positive definite uniformly elliptic matrix, this operator is also symmetric strictly positive definite and uniformly elliptic. In the latter setting, we can thus rule out the occurrence of spurious oscillations in our discretization.

Although we treat the Stolarsky means as an explicit example, note that the main theorems also hold for other smooth means.

1.3 Major contributions of this work

Since we look at the FV discretization of the FPE from a very broad point of view, we summarize our major findings.

- We provide a derivation of the general Stolarsky mean FV discretization in Section 3.2.

- We discuss the gradient structure of the discretization schemes in view of the natural gradient structure of the FPE in Section 3.3.

- We provide order of convergence of the schemes as the fineness of the discretization tends to zero. In particular we show

  - that the order of convergence is mainly determined by two independent parts: the consistency (Def. 2.7) of the mesh (Section 5.1) and an error due to the discretization of π along the edge by the means $S_*(\pi_i, \pi_j)$.
  - that the order of convergence strongly depends on the constant $\alpha + \beta$, where α and β are the Stolarsky coefficients in $S_{ij} = S_{\alpha,\beta}(\pi_i, \pi_j)$ (Corollary 4.3).
  - that the SG coefficients are to be preferred in regions of strong gradient ∇V (Section 5.2).

1.4 Outlook

The results of this work suggest to search for “optimal” parameters α and β in the choice of the Stolarsky mean in order to reduce the error of the approximation as much as possible. However, from an analytical point of view, the quest for such optimal α and β is quite challenging. Moreover, since the optimal choice might vary locally, depending on the local properties of the potential V, we suggest to implement a learning algorithm that provides suitable parameters α and β depending on the local structure of V and the mesh.

1.5 Outline of this work

After some preliminaries regarding notation and a priori estimates in Section 2, we will recall the classical derivation of the SG scheme in Section 3.1 and discuss its formal relation to SQRA. We will then provide a derivation of SQRA from physical principles in Section 3.3, based on the Jordan–Kinderlehrer–Otto [JKO98] formulation of the FPE. In Section 3.2, we show that SG and SQRA are elements of a huge family of discretization schemes (1.6).

Section 5 provides the error analysis and estimates for the consistency and the order of convergence. We distinguish the cases of small and large gradients and have a particular look at cubic meshes.

Finally, we show hat the optimal choice of S_* depends on V and f but is not unique. If $S_{\alpha,\beta}$ denotes one of the Stolarsky means, we will prove in Section 4 that the Stolarsky means satisfying $\alpha + \beta = \mathrm{const}$ show similar quantitative convergence behavior as suggested in Corollary 4.3. Finally, this result is illustrated in Section 6 by numerical simulations.

2 Preliminaries and notation

We collect some concepts and notation, which will frequently be used in this work.

2.1 The Mesh

For a subset $A \subset \mathbb{R}^d$, $\overline{A}$ is the closure of A.

Definition 2.1. Let $\Omega \subset \mathbb{R}^d$ be a polygonal domain. A finite volume mesh of Ω is a triangulation $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$ consisting of a family of control volumes $\mathcal{V} := \{\Omega_i, i = 1, \dots, N\}$ which are convex polytope cells, a family of $(d-1)$-dimensional interfaces

$$\begin{aligned}\mathcal{E} &:= \mathcal{E}_\Omega \cup \mathcal{E}_\partial \\ \mathcal{E}_\Omega &:= \left\{\sigma_{ij} \subset \mathbb{R}^d \,:\, \sigma_{ij} = \partial\Omega_i \cap \partial\Omega_j\right\} \\ \mathcal{E}_\partial &:= \left\{\sigma \subset \mathbb{R}^d \,:\, \sigma = \partial\Omega_i \cap \partial\Omega \text{ is flat}\right\}\end{aligned}$$

and points $\mathcal{P} = \{x_i, i = 1, \dots, N\}$ with $x_i \in \overline{\Omega_i}$ satisfying

(i) $\bigcup_i \overline{\Omega}_i = \overline{\Omega}$

(ii) For every i there exists $\mathcal{E}_i \subset \mathcal{E}$ such that $\overline{\Omega}_i \backslash \Omega_i = \bigcup_{\sigma\in\mathcal{E}_i} \sigma$. Furthermore, $\mathcal{E} = \bigcup_i \mathcal{E}_i$.

(iii) For every i, j either $\overline{\Omega}_i \cap \overline{\Omega}_j = \emptyset$ or $\overline{\Omega}_i \cap \overline{\Omega}_j = \overline{\sigma}$ for $\sigma \in \mathcal{E}_i \cap \mathcal{E}_j$ which will be denoted σ_{ij}.

The mesh is called h-consistent if

(iv) The Family $(x_i)_{i=1\dots N}$ is such that $x_i \neq x_j$ if $i \neq j$ and the straight line D_{ij} going through x_i and x_j is orthogonal to σ_{ij}.

and admissible if

(v) For any boundary interface $\sigma \in \mathcal{E}_\partial \cap \mathcal{E}_i$ it holds $x_i \notin \sigma$ and for $D_{i,\sigma}$ the line through x_i orthogonal to σ it holds that $D_{i,\sigma} \cap \sigma \neq \emptyset$ and let $y_\sigma := D_{i,\sigma} \cap \sigma$.

Property (iv) is assumed in [GHV00] in order to prove a strong form of consistency in the sense of Definition 2.10 below. It is satisfied for example for Voronoi discretizations.

We write m_i for the volume of Ω_i and for $\sigma \in \mathcal{E}$ we denote m_σ its $(d-1)$-dimensional mass. In case $\sigma_{ij} \in \mathcal{E}_i \cap \mathcal{E}_j$ we write $m_{ij} := m_{\sigma_{ij}}$. For the sake of simplicity, we consider $\tilde{\mathcal{P}} := (x_i)_{i=1,\dots,N}$ and $\mathcal{P} := \tilde{\mathcal{P}} \cup \{y_\sigma \,:\, \sigma \in \mathcal{E}_\partial$, according to (v)$\}$. We extend the enumeration of $\tilde{\mathcal{P}}$ to $\mathcal{P} = (x_j)_{j=1,\dots,\tilde{N}}$ and write $i \sim j$ if $x_i, x_j \in \tilde{\mathcal{P}}$ with $\mathcal{E}_i \cap \mathcal{E}_j \neq \emptyset$. Similarly, if $x_i \in \tilde{\mathcal{P}}$ and $x_j = y_\sigma$ for $\sigma \in \mathcal{E}_i$ we write $\sigma_{ij} := \sigma$ and $i \sim j$. Finally, we write $h_{ij} = |x_i - x_j|$.

We further call

$$\mathcal{P}^* := \{u : \mathcal{P} \to \mathbb{R}\}\ , \quad \tilde{\mathcal{P}}^* := \left\{u : \tilde{\mathcal{P}} \to \mathbb{R}\right\}\ , \quad \text{and} \quad \mathcal{E}^* := \{w : \mathcal{E} \to \mathbb{R}\}$$

the discrete functions from $\mathcal{P}$ resp. $\tilde{\mathcal{P}}$ resp. $\mathcal{E}$ to $\mathbb{R}$. For $w \in \mathcal{E}^*$ we write $w_{ij} := w(\sigma)$ if $\sigma_{ij} = \sigma$. Then for fixed i the expression

$$\sum_{j:\, i\sim j} w_{ij} := \sum_{\sigma_{ij}\in\mathcal{E}_i} w_{ij}$$

is the sum over all w_{ij} such that $\mathcal{E}_i \cap \mathcal{E}_j \neq \emptyset$ and

$$\sum_{j\sim i} w_{ij} := \sum_{\sigma_{ij}\in\mathcal{E}} w_{ij} := \sum_{\sigma\in\mathcal{E}} w(\sigma)$$

symbol	meaning	symbol	meaning
u	density	m_i	$\mathrm{vol}(\Omega_i)$
V	real potential on $\Omega \subset \mathbb{R}^d$	h_i	$\mathrm{diam}(\Omega_i)$
κ	diffusion coefficient	σ_{ij}	$\partial\Omega_i \cap \partial\Omega_j$
π	stat. measure $\mathrm{e}^{-V(x)}$ on Ω	m_{ij}	area of σ_{ij}
U	u/π	$\mathbf{h}_{ij}$	$x_i - x_j$
u_i	$u(x_i)$	h_{ij}	$\lvert\mathbf{h}_{ij}\rvert$
π_i	stat. measure $\mathrm{e}^{-V(x_i)}$ on Ω_i	$d_{i,ij}$	$\mathrm{dist}\,(x_i, \sigma_{ij})$
$\bar{f}_i$	$\frac{1}{\lvert\Omega_i\rvert}\int_{\Omega_i} f \mathrm{d}x$	$\mathrm{diam}\mathcal{T}$	diameter, i.e. $\sup_{i\sim j}\lvert x_i - x_j\rvert$
f_i	$m_i \bar{f}_i$	V^*, V_*	$-\infty < V_* \le V \le V^* < \infty$
κ_{ij}	$\frac{h_{ij}\bar\kappa_i\bar\kappa_j}{\bar\kappa_i d_{i,ij}+\bar\kappa_j d_{j,ij}}$	κ^*, κ_*	$0 < \kappa_* \le \kappa \le \kappa^* < \infty$

Table 1: Commonly used notations.

is the sum over all edges.

Moreover, we define the diameter of a triangulation $\mathcal{T}$ as

$$\mathrm{diam}\mathcal{T} = \sup_{i\sim j} |x_i - x_j|.$$

The identity

$$\sum_i \sum_{j:j\sim i} A_{ij} = \sum_{j\sim i} (A_{ij} + A_{ji}) \tag{2.1}$$

will frequently be used throughout this paper, where we often encounter the case $A_{ij} = \alpha_{ij} U_i$ with $\alpha_{ij} = -\alpha_{ji}$:

$$\sum_i \sum_{j:j\sim i} \alpha_{ij} U_i = \sum_{j\sim i} (\alpha_{ij} U_i + \alpha_{ji} U_j) = \sum_{j\sim i} \alpha_{ij} (U_i - U_j) \ . \tag{2.2}$$

Formula (2.1) in particular allows for a discrete integration by parts:

$$\sum_i \sum_{j:j\sim i} (U_j - U_i) U_i = \sum_{j\sim i} ((U_j - U_i) U_i + (U_i - U_j) U_j) = -\sum_{j\sim i} (U_j - U_i)^2 \ . \tag{2.3}$$

On a given mesh $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, we consider the linear discrete operator $\mathcal{L}^{\mathcal{T}}_\kappa : \mathcal{P}^* \to \mathcal{P}^*$, which is defined by a family of non-negative weights $\kappa : \mathcal{E} \to \mathbb{R}$ and acts on functions $u \in \mathcal{P}^*$ via

$$\forall x_i \in \mathcal{P} : \ \left(\mathcal{L}^{\mathcal{T}}_\kappa u\right)_i := \sum_{i\sim j} \kappa_{ij} \frac{m_{ij}}{h_{ij}} (u_j - u_i) \ . \tag{2.4}$$

While (2.4) is very general, it is shown in [GHV00], Lemma 3.3, that the property (iv) of Definition 2.1 comes up with some special consistency properties for the choice of

$$\kappa_{ij} := \frac{\bar\kappa_i \bar\kappa_j}{\bar\kappa_i \frac{d_{i,ij}}{h_{ij}} + \bar\kappa_j \frac{d_{j,ij}}{h_{ij}}} \ , \tag{2.5}$$

where $d_{i,ij}$ and $d_{j,ij}$ are the distances between σ_{ij} and x_i and x_j respectively and averaged diffusion coefficient is defined by $\overline{\kappa}_i = \frac{1}{m_i}\int_{\Omega_i} \kappa(x) \mathrm{d}x$.

Lemma 2.2 (A consistency lemma, [GHV00]). *Let the $\mathcal{T}=(\mathcal{V},\mathcal{E},\mathcal{P})$ satisfy Definition 2.1 (i)–(v) and let $d\in\{2,3\}$ and let h_{ij} be uniformly bounded from above and from below. Then for every $u\in H^2(\Omega)$ it holds*

$$\left|\int_{\sigma_{ij}}\kappa\nabla u\cdot\boldsymbol{\nu}_{ij}-\kappa_{ij}\frac{m_{ij}}{h_{ij}}\left(u(x_j)-u(x_i)\right)\right|\leq Cm_{ij}^{\frac{1}{2}}h_{ij}^{\frac{1}{2}}\left\|u\right\|_{H^2(\Omega_i\cup\Omega_j)}\,.$$

Lemma 2.2 was one of the motivations to provide a more general and powerful concept of consistency in [DPD18], as we will discuss in Section 2.5

2.2 Existence and a priori estimates

From the standard theory of elliptic systems ([Eva98] Chapter 6), we have the following theorem.

Theorem 2.3. *Let Ω be as above and $f\in\mathrm{L}^2(\Omega)$, $\kappa\in C^1\left(\overline{\Omega}:\mathbb{R}^{d\times d}\right)$ such that κ is uniformly bounded, symmetric and elliptic and $V\in\mathrm{C}^2(\overline{\Omega})$. Then there is a unique $u\in\mathrm{H}^2(\Omega)\cap\mathrm{H}^1_0(\Omega)$ solving $-\nabla\cdot(\kappa\nabla u)-\nabla\cdot(\kappa u\nabla V)=f$ in the weak sense.*

In what follows, we frequently use the following transformations in (1.1) and (1.6): we define the relative densities $U=u/\pi$ and $U_i^{\mathcal{T}}=u_i^{\mathcal{T}}/\pi_i$ to find

$$-\nabla\cdot(\kappa\pi\nabla U)=f\,, \tag{2.6}$$

$$\forall i:\qquad-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S(\pi_i,\pi_j)\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)=f_i\,. \tag{2.7}$$

If κ and π are non-degenerate (in the sense of $\pi>c>0$ and $\xi\cdot\kappa\xi>c\left|\xi\right|^2$), the left hand side of (2.7) defines a strongly elliptic operator on the finite volume space $L^2(\mathcal{P})$ and hence there exists a unique solution $U^{\mathcal{T}}$. Concerning the right hand side, using (2.7) as the discretization of (2.6), one natural choice for f_i is $f_i=m_i\bar{f}_i$, where $\bar{f}_i=m_i^{-1}\int_{\Omega_i}f$. We immediately see that the Boltzmann distribution $\pi_i=\exp\left(-V(x_i)\right)=\exp\left(-V_i\right)$ is the stationary solution, i.e., $u_i=\pi_i$ solves (2.7) for $f=0$.

Having shown the existence of solutions to (2.6) and (2.7), we recall the derivation of some natural a priori estimates for both the continuous Fokker–Planck equation and the discretization.

Continuous FPE Let u, resp. $U=u/\pi$, be a solution of the stationary Fokker–Planck equation (2.6) with Dirichlet boundary conditions. Testing with U, we get (assuming homogeneous Dirichlet boundary conditions and exploiting thus the Poincaré inequality), that

$$\begin{aligned}\int_\Omega\kappa\pi|\nabla U|^2=\int_\Omega Uf\quad&\leq C\left(\int_\Omega f^2\right)^{\frac{1}{2}}\left(\int_\Omega\kappa\pi|\nabla U|^2\right)^{\frac{1}{2}}\\ \Rightarrow\int_\Omega\tfrac{1}{\kappa\pi}|\kappa\pi\nabla U|^2&\leq C\int_\Omega f^2\,.\end{aligned} \tag{2.8}$$

Furthermore, the standard theory of elliptic equations (e.g., [Eva98]) yields that $\left\|U\right\|_{H^2(\Omega)}\leq C\left\|f\right\|_{L^2}$, where C depends on the C^1-norm of $\kappa\pi$ and the Poincaré-constant.

Discrete FPE Let $U_i^{\mathcal{T}}$ be a solution of (2.7) with $f_i=m_i\bar{f}_i=\int_{\Omega_i}f\mathrm{d}x$ (as specified in the Tab. 1), i.e.,

$$\forall i:\qquad-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)=m_i\bar{f}_i\,.$$

Then, multiplying with $U_i^{\mathcal{T}}$ we get

$$-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)U_i^{\mathcal{T}}=m_i\bar{f}_iU_i^{\mathcal{T}}.$$

Summing over all $x_i\in\mathcal{P}$ and using (2.3), we conclude with help of the discrete Poincaré inequality (see Theorem 2.5 below)

$$\begin{aligned}\sum_{j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)^2=\sum_i m_i\bar{f}_iU_i^{\mathcal{T}}\leq\sum_i\left((U_i^{\mathcal{T}})^2m_i+\tfrac{1}{\pi_i}\bar{f}_i^2m_i\right)\\ \Rightarrow\sum_{j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)^2\leq C\sum_i m_i\bar{f}_i^2.\end{aligned}$$

Additionally, one gets

$$\sum_{j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}\frac{1}{S_{ij}\kappa_{ij}^2}\left(\kappa_{ij}S_{ij}(U_j^{\mathcal{T}}-U_i^{\mathcal{T}})\right)^2\leq C\sum_i\bar{f}_i^2m_i. \tag{2.9}$$

2.3 Fluxes and L^2-spaces

In order to derive and formulate variational consistence errors for the discrete FPE (2.7), we introduce the discrete fluxes

$$\begin{aligned}J_{ij}^SU^{\mathcal{T}}&:=-\frac{\kappa_{ij}}{h_{ij}}S_{ij}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right),\\ \overline{J}_{ij}U&:=-\frac{1}{m_{ij}}\int_{\sigma_{ij}}\kappa\pi\nabla U\cdot\boldsymbol{\nu}_{ij}\,.\end{aligned}\tag{2.10}$$

In particular, if $S_{ij}=\sqrt{\pi_i\pi_j}$ we get the flux of the SQRA $J_{ij}^{\mathrm{SQRA}}U^{\mathcal{T}}:=-\frac{\kappa_{ij}}{h_{ij}}\sqrt{\pi_i\pi_j}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)$. Note that $\overline{J}_{ij}U$ is the spatial average of $\mathbf{J}(U)=-\kappa\pi\nabla U$ on σ_{ij}. The quantity $J_{ij}^SU^{\mathcal{T}}$ can indeed be considered as a flux in the sense that it will be shown to approximate $\overline{J}_{ij}$, S_{ij} is a discrete approximation of $\pi|_{\sigma_{ij}}$, κ_{ij} is a discrete approximation of $\kappa|_{\sigma_{ij}}$ and $\frac{1}{h_{ij}}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)$ is a discrete version of ∇U.

While former approaches focus on the rate of convergence of $\frac{1}{h_{ij}}\left(u_j^{\mathcal{T}}-u_i^{\mathcal{T}}\right)\to\nabla u$, we additionally follow the approach of [DPD18] applied to U and are interested in the rate of convergence of $J_{ij}^SU^{\mathcal{T}}\to\mathbf{J}(U)$, which is an indirect approach to the original problem as this rate of convergence is directly related to $\frac{1}{h_{ij}}\left(U_j^{\mathcal{T}}-U_i^{\mathcal{T}}\right)\to\nabla U$.

In view of the natural norms for the variational consistency (see (2.17)–(2.19)), we introduce the following

$$\begin{array}{lll}\forall U\in L^2(\Omega): & \|U\|_{L^2(\Omega)}^2:=\int_\Omega U^2\mathrm{d}x & \|U\|_{L^2_\pi(\Omega)}^2:=\int_\Omega\frac{1}{\pi}U^2\mathrm{d}x\\ \forall U\in\mathcal{P}^*: & \|U\|_{L^2(\mathcal{P})}^2:=\sum_{i\in\mathcal{P}}m_iU_i^2 & \|U\|_{L^2_\pi(\mathcal{P})}^2:=\sum_{i\in\mathcal{P}}m_i\frac{1}{\pi_i}U_i^2\\ \forall J\in\mathcal{E}^*: & \|J\|_{L^2(\mathcal{E})}^2:=\sum_{i\sim j}m_{ij}h_{ij}J_{ij}^2 & \|J\|_{L^2_S(\mathcal{E})}^2:=\sum_{i\sim j}m_{ij}h_{ij}\frac{1}{S_{ij}}J_{ij}^2\end{array}\tag{2.11}$$

Let us introduce the discrete flux $J^SU^{\mathcal{T}}\in\mathcal{E}^*$ via $J^SU^{\mathcal{T}}(\sigma_{ij}):=J_{ij}^SU^{\mathcal{T}}$ and similarly also $\frac{1}{\kappa}J^SU^{\mathcal{T}}\in\mathcal{E}^*$ via $J^SU^{\mathcal{T}}(\sigma_{ij}):=\frac{1}{\kappa_{ij}}J_{ij}^SU^{\mathcal{T}}$. With all the above notations, our a priori estimates (2.8) and (2.9) now read

$$\left\|\frac{1}{\sqrt{\kappa}}\mathbf{J}(U)\right\|^2_{L^2_\pi(\Omega)} \leq C\|f\|^2_{L^2_\pi(\Omega)}$$
$$\left\|\frac{1}{\sqrt{\kappa}}J^S U^{\mathcal{T}}\right\|^2_{L^2_S(\mathcal{E})} \leq C\|\bar{f}\|^2_{L^2_\pi(\mathcal{P})}\,.$$

Assuming that the diffusion coefficient is bounded, i.e. $\kappa^* \geq \kappa \geq \kappa_*$, we further get

$$\tfrac{1}{\kappa_*}\|\mathbf{J}(U)\|^2_{L^2_\pi(\Omega)} \leq C\|f\|^2_{L^2_\pi(\Omega)}$$
$$\frac{1}{\kappa^*}\|J^S U^{\mathcal{T}}\|^2_{L^2_S(\mathcal{E})} \leq C\|\bar{f}\|^2_{L^2_\pi(\mathcal{P})}\,.$$

Remark 2.4 (Naturalness of norms)**.** Let us discuss why these norms are natural to consider. The left norms in (2.11) can be interpreted as the Euclidean L^2-norms on Ω, $\mathcal{P}$ and $\mathcal{E}$, while the right norms are the natural norms for the study of the Fokker–Planck equation as they are weighted with the inverse of the Boltzmann distribution π, resp. π_i. Note that assuming V is bounded from above and below, the L^2-norms $\|\cdot\|_{L^2_\pi(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ are equivalent and the same holds true for the two norms in the discrete setting.

Given a discretization $\mathcal{T}$, the linear map

$$C_c\left(\mathbb{R}^d\right) \to \mathbb{R}\,, \qquad f \mapsto \sum_{i\in\mathcal{P}} m_i f(x_i)$$

defines an integral on Ω w.r.t. a discrete measure $\mu_{\mathcal{T}}$ having the property that $\mu_{\mathcal{T}} \to \mathcal{L}^d$ vaguely, where $\mathcal{L}^d$ is the d-dimensional Lebesgue measure. In particular $\mu_{\mathcal{T}}(A) \to \mathcal{L}^d\left(A\right)$ for every bounded measurable set with $\mathcal{L}^d\left(\partial A\right) = 0$. The norm $\|U\|^2_{L^2(\mathcal{P})}$ is simply the L^2-norm based on the measure $\mu_{\mathcal{T}}$.

Similar considerations work also for the norm on $\mathcal{E}^*$. The norm $\|\cdot\|^2_{L^2(\mathcal{E})}$ is given via a measure $\tilde{\mu}_{\mathcal{T}}$ having the property

$$\tilde{\mu}_{\mathcal{T}} : C_c\left(\mathbb{R}^d\right) \to \mathbb{R}\,, \qquad f \mapsto \sum_{i\sim j} m_{ij} h_{ij} f(x_{ij})\,,$$

with the property that $\tilde{\mu}_{\mathcal{T}} \to d\cdot\mathcal{L}^d$ vaguely: every Voronoi cell Ω_i consists of disjoint cones with mass $\frac{1}{d}m_{ij}h_{ij}$, where one has to account for all cones with $j \sim i$. In particular, we obtain $\tilde{\mu}_{\mathcal{T}}(A) \approx d\cdot\mathcal{L}(A)$ for Lipschitz domains – an estimate which then becomes precise in the limit. Without going into details, let us mention that heuristically the prefactor d balances the fact that $J_{ij} \approx \frac{(x_i-x_j)}{|x_i-x_j|}\cdot\nabla U$ which yields for functions $U \in C^1_c\left(\mathbb{R}^d\right)$:

$$\sum_{i\sim j} m_{ij}h_{ij}\left|\frac{(x_i-x_j)}{|x_i-x_j|}\cdot\nabla U\right|^2 \to \int_{\mathbb{R}^d} |\nabla U|^2\,.$$

For the particular case of a rectangular mesh, this is straight forward to verify.

2.4 Poincaré inequalities

In order to derive the a priori estimates in Section 2.2 we need to exploit (discrete) Poincaré inequalities to estimate $\|u\|_{L^2(\Omega)}$ by $\|\nabla u\|_{L^2(\Omega)}$ or $\|u^{\mathcal{T}}\|_{L^2(\mathcal{P})}$ by $\|Du^{\mathcal{T}}\|_{L^2(\mathcal{E})}$, where $\left(Du^{\mathcal{T}}\right)_{ij} = U_j - U_i$. In particular, we use the following theorem.

Theorem 2.5. *Given a mesh $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$ let $h_{\inf} := \inf\{|x-y| : (x,y) \in \mathcal{P}^2\} > 0$ and $h_{\sup} := \sup\{|x-y| : (x,y) \in \mathcal{P}^2\} > 0$ correspondingly. Then for every $u \in L^2(\mathcal{P})$ and for every $\boldsymbol{\eta} \in \mathbb{R}^d$ it holds*

$$\int_\Omega \left|\sum_i u_i \chi_{\Omega_i}(x) - \sum_i u_i \chi_{\Omega_i}(x+\boldsymbol{\eta})\right|^2 dx \le |\boldsymbol{\eta}| \left(\mathrm{diam}\Omega \frac{h_{\sup}}{h_{\inf}} \sum_{i\sim j} \frac{m_{ij}}{h_{ij}} (u_j - u_i)^2\right), \tag{2.12}$$

and particularly

$$\|u\|^2_{L^2(\mathcal{P})} \le (\mathrm{diam}\Omega)^2 \frac{h_{\sup}}{h_{\inf}} \sum_{i\sim j} m_{ij} (u_j - u_i)^2 .$$

Proof. This follows from Lemma A.1 with $C_\# \le \frac{\mathrm{diam}\Omega}{h_0}$ and the choice $|\boldsymbol{\eta}| > \mathrm{diam}\Omega$. □

2.5 Consistency and inf-sup stability

Results such as Lemma 2.2 motivated the authors of the recent paper [DPD18] to define the concepts of consistency and inf-sup stability as discussed in the following. For readability, we will restrict the general framework of [DPD18] to cell-centered finite volume schemes and refer to general concepts only as far as needed.

Definition 2.6 (inf-sup stability)**.** A bilinear form $a_\mathcal{T}$ on $L^2(\mathcal{P})$ for a given mesh $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$ is called inf-sup stable with respect to a norm $\|\cdot\|_{H_\mathcal{T}}$ on a subspace of $H_\mathcal{T} \subset L^2(\mathcal{P})$ if there exists $\gamma > 0$ such that

$$\forall u \in H_\mathcal{T} : \quad \gamma \|u\|_{H_\mathcal{T}} \le \sup_{v \in H_\mathcal{T}} \frac{a_\mathcal{T}(u,v)}{\|v\|_{H_\mathcal{T}}} .$$

Usually, and particularly in our setting, $a_\mathcal{T}$ is the discretization of a continuous bilinear form, say $a(u,v) = \int_\Omega \nabla u \cdot (\kappa \nabla v)$. We are interested in the problem

$$\forall v \in H^1_0(\Omega) : \quad a(u,v) = l(v) , \tag{2.13}$$

where $l : H^1_0(\Omega) \to \mathbb{R}$ is a continuous linear map, and in the convergence of the solutions of the discrete problems

$$\forall v \in L^2(\mathcal{T}) : \quad a_\mathcal{T}(u_\mathcal{T}, v) = l_\mathcal{T}(v) . \tag{2.14}$$

Definition 2.7 (Consistency)**.** Let $B \subset H^1_0(\Omega)$ be a continuously embedded Banach subspace and for given $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$ consider continuous linear operators $\mathcal{R}_\mathcal{T} : B \to L^2(\mathcal{P})$ with uniform bound. Let u be the solution to the linear equation (2.13) and let $l_\mathcal{T} : L^2(\mathcal{P}) \to \mathbb{R}$ be a family of linear functionals. The variational consistency error is the linear form $\mathfrak{E}_\mathcal{T}(u; \cdot) : L^2(\mathcal{P}) \to \mathbb{R}$ where

$$\forall u \in B : \qquad \mathfrak{E}_\mathcal{T}(u; \cdot) := l_\mathcal{T}(\cdot) - a_\mathcal{T}(\mathcal{R}_\mathcal{T} u, \cdot) .$$

Let now a family $(\mathcal{T}, a_\mathcal{T}, l_\mathcal{T})$ with $\mathrm{diam}\mathcal{T} \to 0$ be given and consider the corresponding family of linear discrete problems (2.14). We say that consistency holds if

$$\|\mathfrak{E}_\mathcal{T}(u; \cdot)\|_{H^*_\mathcal{T}} \to 0 \quad \text{as} \quad \mathrm{diam}\mathcal{T} \to 0 .$$

Remark 2.8. A typical situation is the case $d \le 3$, where $H^2(\Omega) \cap H^1_0(\Omega) \hookrightarrow C_0(\Omega)$ continuously. We then might set $B = H^2(\Omega) \cap H^1_0(\Omega)$ and $(\mathcal{R}_\mathcal{T} u)_i := u(x_i)$.

Consistency measures the rate at which $\mathcal{R}_\mathcal{T}u - u_\mathcal{T} \to 0$ and particularly provides a positive answer to the question whether the numerical scheme converges, at least if the solution of (2.13) lies in B. This is formulated in Theorem 10 of [DPD18].

Theorem 2.9 (Theorem 10, [DPD18])**.** *Using the above notation, it holds*

$$\left\|u_\mathcal{T} - \mathcal{R}_\mathcal{T}u\right\|_{H_\mathcal{T}} \le \gamma^{-1}\left\|\mathfrak{E}_\mathcal{T}\left(u;\cdot\right)\right\|_{H^*_\mathcal{T}} \tag{2.15}$$

In our setting, $\|\cdot\|_{H_\mathcal{T}} = \|\cdot\|_{H_{\mathcal{T},\kappa}}$ (see (2.17)) is a norm on the discrete gradients. By the discrete Poincaré inequality, (2.15) also implies an convergence estimate for the discrete solutions itself. The theorem can be understood as a requirement on the regularity of u, resp. the right hand side of (2.13).

The combination of the proofs of Theorems 27 and 33 of [DPD18] shows that the variational consistency error for

$$a\left(u,v\right) = \int_\Omega \nabla u\cdot\kappa\nabla v\,,\qquad a_\mathcal{T}\left(u,v\right) = \sum_{i\sim j}\frac{m_{ij}}{h_{ij}}\left(u_j-u_i\right)\kappa_{ij}\left(v_j-v_i\right)$$

becomes

$$\mathfrak{E}_\mathcal{T}\left(u;v\right) = \sum_{i\sim j}\left(v_j-v_i\right)\left(\int_{\sigma_{ij}}\kappa\nabla u\cdot\boldsymbol{\nu}_{ij} - \frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(\left(\mathcal{R}_\mathcal{T}u\right)_j - \left(\mathcal{R}_\mathcal{T}u\right)_i\right)\right). \tag{2.16}$$

Introducing on $L^2\left(\mathcal{P}\right)$ the $H_\mathcal{T}$-norm given by

$$\left\|u\right\|_{H_{\mathcal{T},\kappa}} := \sum_{i\sim j}\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(u_j-u_i\right)^2, \tag{2.17}$$

we find

$$\left\|\mathfrak{E}_\mathcal{T}\left(u;\cdot\right)\right\|_{H^*_{\mathcal{T},\kappa}} \le \sum_{i\sim j}\frac{h_{ij}}{m_{ij}}\kappa_{ij}^{-1}\left(\int_{\sigma_{ij}}\kappa\nabla u\cdot\boldsymbol{\nu}_{ij} - \frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(\left(\mathcal{R}_\mathcal{T}u\right)_j - \left(\mathcal{R}_\mathcal{T}u\right)_i\right)\right)^2. \tag{2.18}$$

In view of the Poincaré inequality in Theorem 2.5 the norm $\|\cdot\|_{L^2(\mathcal{P})}$ is bounded by $\|\cdot\|_{H_{\mathcal{T},\kappa}}$ in case κ is uniformly bounded away from 0. The right hand side of equation (2.18) gives rise to the definition of a “dual” $H_\mathcal{T}$-norm which we denote

$$\left\|u\right\|_{H^-_{\mathcal{T},\kappa}} := \sum_{i\sim j}\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(u_j-u_i\right)^2. \tag{2.19}$$

With regard to (2.15) and Lemma 2.2, the above considerations motivate the following definition.

Definition 2.10 (φ-consistency)**.** Let $\mathcal{T}_h = \left(\mathcal{V}_h,\mathcal{E}_h,\mathcal{P}_h\right)$ be a family of meshes with $\mathrm{diam}\mathcal{T}_h \to 0$ as $h\to 0$. We say that $\mathcal{T}_h$ is φ-consistent (satisfies φ-consistency) on the subspace $B\subset H^1_0\left(\Omega\right)$ if for every $u\in B$ there exists $C\ge 0$ such that for every $h>0$

$$\sum_{\sigma_{ij}\in\mathcal{E}_h}\frac{h_{ij}}{m_{ij}}\kappa_{ij}^{-1}\left|\int_{\sigma_{ij}}\kappa\nabla u\cdot\boldsymbol{\nu}_{ij} - \kappa_{ij}\frac{m_{ij}}{h_{ij}}\left(\left(\mathcal{R}_{\mathcal{T}_h}u\right)_j - \left(\mathcal{R}_{\mathcal{T}_h}u\right)_i\right)\right|^2 \le C\varphi\left(h\right)^2.$$

Hence, we immediately obtain the following.

Proposition 2.11. *Under the assumptions of Lemma 2.2 and assuming $h_{ij}\le Ch$ for some constant $C>0$ the mesh is φ-consistent with $\varphi(h) = h$. We say that the mesh is h-consistent.*

2.6 Consistency on cubic meshes

For $d \leq 3$, we consider a polygonal domain $\Omega \subset \mathbb{R}^d$ with a cubic mesh where $\Omega_i = x_i + [-h/2, h/2]^d$, $x_i \in h\mathbb{Z} \subset \Omega$,

Then we want to estimate the terms

$$m_{ij}\left|J^S_{ij}\hat{U} - \overline{J}^\star_{ij}U\right| = S_{ij}\left|m_{ij}\kappa_{ij}\frac{\hat{U}_j - \hat{U}_i}{h} - \int_{\sigma_{ij}}\kappa\nabla U\cdot\nu_{ij}\right|$$
$$\leq S_{ij}\left|m_{ij}\kappa_{ij}\frac{\hat{U}_j - \hat{U}_i}{h} - \kappa_{ij}\int_{\sigma_{ij}}\nabla U\cdot\nu_{ij}\right| + S_{ij}\left|\int_{\sigma_{ij}}(\kappa_{ij} - \kappa)\nabla U\cdot\nu_{ij}\right|.$$

In fact the following calculations are quite standard and, therefore, we shorten our considerations.

Now, we want to estimate $\left|m_{ij}\frac{\hat{U}_j - \hat{U}_i}{h} - \int_{\sigma_{ij}}\nabla U\cdot\nu_{ij}\right|$. We have $\hat{U}_j = U(x) + \nabla U\cdot(x_j - x) + O(h^2)$ and $\hat{U}_i = U(x) + \nabla U\cdot(x_i - x) + O(h^2)$. Moreover, we can write $x_i - x = -\frac{h}{2}\nu_{ij} + \tilde{x}$ where $\tilde{x} \perp \nu_{ij}$ and $x_j - x = \frac{h}{2}\nu_{ij} + \tilde{x}$ (the normal ν_{ij} points outside or inside of Ω_i). Hence, we conclude

$$\hat{U}_j = U(x) + \nabla U\cdot(\frac{h}{2}\nu_{ij} + \tilde{x}) + O(h^2)$$
$$\hat{U}_i = U(x) + \nabla U\cdot(-\frac{h}{2}\nu_{ij} + \tilde{x}) + O(h^2).$$

Subtracting both equations, we end up with $\frac{\hat{U}_j - \hat{U}_i}{h} = \nabla U\cdot\nu_{ij} + O(h^2)$, and hence,

$$\left|m_{ij}\frac{\hat{U}_j - \hat{U}_i}{h} - \int_{\sigma_{ij}}\nabla U\cdot\nu_{ij}\right| \leq m_{ij}O(h^2).$$

Theorem 2.12 (Consistency on cubic meshes). *Let $\Omega \subset \mathbb{R}^d$ with $d \leq 3$ be a polygonal domain with a cubic mesh where $\Omega_i = x_i + [-h/2, h/2]^d$, $x_i \in h\mathbb{Z} \subset \Omega$ and let $\kappa \equiv 1$. Then*

$$\|\mathfrak{E}_{\mathcal{T}}(u;\cdot)\|_{H^\star_{\mathcal{T}}} \leq Ch^2.$$

We will consider a general κ in Section 5.3 below.

3 Derivation of the methods and formal comparison

In this section, we recall the original derivation of the Scharfetter–Gummel scheme and then show that both the SG and the SQRA scheme are members of a huge family of discretization schemes. Finally, we provide a physically motivated derivation of the SQRA scheme.

3.1 Motivation of the Scharfetter–Gummel scheme

One dimensional case The Scharfetter–Gummel scheme for the discrete flux on the interval $[0, h]$ is derived under the assumption of constant flux J, force $q = -\mathrm{d}V/\mathrm{d}x$ and diffusion coefficient κ on $[0, h]$. We consider the two-point boundary value problem

$$J = -\kappa\left(\frac{\mathrm{d}u}{\mathrm{d}x} - qu\right) \quad \text{on } [0, h], \qquad u(0) = u_0, \qquad u(h) = u_h, \tag{3.1}$$

where $q : [0, h] \to \mathbb{R}$ describes the constant force inducing the drift current (i.e., the potential V is a linear function on $[0, h]$) and $\kappa > 0$ is a constant, positive diffusion coefficient. The problem has an elementary solution of the form

$$u(x) = \frac{J}{\kappa q} + \left(u_0 - \frac{J}{\kappa q}\right) \mathrm{e}^{qx}.$$

Using the second boundary value $u(h) = u_h$, we get an explicit form for the flux

$$J = \frac{\kappa}{h}\left(u_0 B(-qh) - u_h B(qh)\right), \tag{3.2}$$

where $B(r) = r/(\mathrm{e}^r - 1)$ is the Bernoulli function. Finally, using $q = -(V_h - V_0)/h$, we can equally write

$$J = -\frac{\kappa}{h} \frac{V_h - V_0}{\mathrm{e}^{V_h} - \mathrm{e}^{V_0}} \left(\frac{u_h}{\pi_h} - \frac{u_0}{\pi_0}\right). \tag{3.3}$$

Higher dimensional case In higher dimensions, the flux between two neighboring cells $j \sim i$ is discretized along the same lines as in the one dimensional case (i.e., assumption of constant force, flux and diffusion constant along each edge of the mesh). We project the flux $\mathbf{J}$ on the edge $\mathbf{h}_{ij} := x_j - x_i$

$$\mathbf{h}_{ij} \cdot \mathbf{J} = -\kappa_{ij}\left(\mathbf{h}_{ij} \cdot \nabla u + u \mathbf{h}_{ij} \cdot \nabla V\right),$$

where the assumption of a linear affine potential (inducing the constant force $\mathbf{q} = -\nabla V$) implies that $\mathbf{h}_{ij} \cdot \nabla V = V_i - V_j$. Moreover, we write $u(x) = u(x(s))$ with $x(s) = s x_i + (1 - s) x_j$, where $0 \le s \le 1$ parametrizes the position on the edge. Then, with $\mathrm{d}u/\mathrm{d}s = \mathbf{h}_{ij} \cdot \nabla u$ and $\mathbf{h}_{ij} \cdot \mathbf{J} = h_{ij} J_{ij}$, we arrive at the two-point boundary problem

$$h_{ij} J_{ij} = -\kappa_{ij}\left(\frac{\mathrm{d}u}{\mathrm{d}s} + u(V_i - V_j)\right) \quad \text{on } s \in [0, 1], \qquad u(0) = u_j, \qquad u(1) = u_i,$$

which is equivalent to the one dimensional problem. (3.1). The solution reads

$$J_{ij} = \frac{\kappa_{ij}}{h_{ij}}\left(u_j B(V_i - V_j) - u_i B(-(V_i - V_j))\right),$$

which can also be written as

$$J_{ij} = -\frac{\kappa_{ij}}{h_{ij}} \frac{V_i - V_j}{e^{V_i} - e^{V_j}} \left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j}\right).$$

Remark 3.1. In case of a sufficiently fine discretization that accurately takes into account the structure of V, we can expect that $|V_j - V_i| \ll 1$ is small such that $\frac{V_i - V_j}{\mathrm{e}^{V_i} - \mathrm{e}^{V_j}} \approx \sqrt{\pi_i \pi_j} + O(\pi_i - \pi_j)^2$. We then infer

$$J_{ij} = -\frac{\kappa_{ij}}{h_{ij}} \sqrt{\pi_i \pi_j} \left(\frac{u_j}{\pi_j} - \frac{u_i}{\pi_i}\right),$$

which is the flux discretization according to the SQRA scheme. This becomes more clear in the following sections.

Remark 3.2 (Motivation of discretized diffusion coefficient κ_{ij})**.** Considering inhomogeneous media, where the diffusion coefficient is not necessarily constant, a suitable discretization for the diffusion κ_{ij} is needed. Let us neglect for a moment the drift $\mathbf{q}u$ and assume that we have κ_i on Ω_i around x_i and κ_j on Ω_j around x_j. We compute the flux J_{ij} from x_i to x_j. Let x_0 be the intersection of $\mathbf{h}_{ij}$ and σ_{ij}, and moreover, $d_i = |x_0 - x_i|$ and $d_j = |x_j - x_0|$. The density at x_0 is denoted by u_0. The flux J_i from

x_i to x_0 is then given by $J_i = -\kappa_i \frac{u_0 - u_i}{d_i}$ and the flux J_j from x_0 to x_j is then given by $J_j = -\kappa_j \frac{u_j - u_0}{d_j}$. The flux from x_i to x_j is then given by $J_{ij} = -\kappa_{ij} \frac{u_j - u_i}{d_i + d_j}$. Hence, we have

$$J_{ij} = -\frac{\kappa_{ij}}{d_i + d_j}\left(u_j - u_0 + u_0 - u_i\right) = \frac{\kappa_{ij}}{d_i + d_j}\left(\frac{d_i}{\kappa_i} J_i + \frac{d_j}{\kappa_j} J_j\right).$$

Kirchhoff's law says $J_i = J_j = J_{ij}$, which implies that $1 = \frac{\kappa_{ij}}{d_i + d_j}\left(\frac{d_i}{\kappa_i} + \frac{d_j}{\kappa_j}\right)$ and hence the weighted harmonic mean

$$\kappa_{ij} = \frac{(d_i + d_j)}{\frac{d_i}{\kappa_i} + \frac{d_j}{\kappa_j}} = \frac{1}{\frac{1}{\kappa_i}\frac{d_i}{d_i + d_j} + \frac{1}{\kappa_j}\frac{d_j}{d_i + d_j}}.$$

Note that $1/\kappa$ is the mobility and hence we conclude $\frac{1}{\kappa_{ij}} = \frac{1}{\kappa_i}\frac{d_i}{d_i + d_j} + \frac{1}{\kappa_j}\frac{d_j}{d_i + d_j}$, i.e. the arithmetic mean of the mobilities $1/\kappa_i$ and $1/\kappa_j$.

Interestingly, the harmonic mean is yet another special case of Stolarsky means (see below) for $\alpha = -2$ and $\beta = -1$. Thus classical FV discretizations of classical elliptic problems based on discretizations of $-\Delta$ are another particular case of our general study.

3.2 A family of discretization schemes

Repeating the above calculations from a different point of view reveals some additional structure of the Scharfetter–Gummel scheme and puts it into a broader context.

Taking into account the special structure of the Fokker–Planck equation in (3.1), we solve

$$\frac{1}{\kappa} J = -\left(u'(x) + u(x) V'(x)\right), \qquad u(0) = u_0,\ u(h) = u_h,$$

for a general potential $V : [0, h] \to \mathbb{R}$ not necessarily assumed to be affine. The general solution reads

$$u(x) = -\left(\frac{1}{\kappa} J \int_0^x \mathrm{e}^V + u_0 \mathrm{e}^{V_0}\right) \mathrm{e}^{-V(x)}.$$

The flux can be computed explicitly from the assumption $J = \mathrm{const.}$ and setting $x = h$ in the above formula. This yields

$$J = -\kappa \frac{u_h \mathrm{e}^{V_h} - u_0 \mathrm{e}^{V_0}}{\int_0^h \mathrm{e}^V} = -\kappa \frac{1}{h}\left(\frac{1}{h}\int_0^h \pi^{-1}\right)^{-1}\left(\frac{u_h}{\pi_h} - \frac{u_0}{\pi_0}\right) = -\kappa \pi_{\mathrm{mean}} \frac{1}{h}\left(\frac{u_h}{\pi_h} - \frac{u_0}{\pi_0}\right)$$

for the averaged $\pi_{\mathrm{mean}} = \left(\frac{1}{h}\int_0^h \pi^{-1}\right)^{-1}$, which clearly determines the constant flux along the edge. In particular, assuming that V is affine, i.e. $V(x) = \frac{V_h - V_0}{x_h - x_0}(x - x_0) + V_0$, one easily checks that $\pi_{\mathrm{mean}} = (V_h - V_0)/(\mathrm{e}^{V_h} - \mathrm{e}^{V_0})$, which is the mean corresponding to the Scharfetter–Gummel discretization. However, a potential can also be approximated not by piecewise affine interpolation but in other ways, resulting in different means π_{mean}. We provide an example of such an approximation for the SQRA in the Appendix A.4.

We aim to express π_{mean} by means of the values π_0 and π_h at the boundaries. The choice of this average is non-trivial and determines the quality of the discretization scheme, as we will see below. In the present work, we focus on the (weighted) Stolarsky mean, although there are also other means like general f-means ($M_f(x, y) = f\left(\frac{f^{-1}(x) + f^{-1}(y)}{2}\right)$ for a strictly increasing function f). The Stolarsky mean has the advantage that it is a closed formula for a broad family of popular means and that its derivatives can be computed explicitly.

The weighted Stolarsky mean $S_{\alpha,\beta}$ [Sto75] is given as

$$S_{\alpha,\beta}(x,y) = \left(\frac{\beta(x^\alpha - y^\alpha)}{\alpha(x^\beta - y^\beta)}\right)^{\frac{1}{\alpha-\beta}} , \tag{3.4}$$

whenever these expressions are well defined and continuously extended otherwise. We note the symmetry properties $S_{\alpha,\beta}(x,y) = S_{\alpha,\beta}(y,x) = S_{\beta,\alpha}(x,y)$. Interesting special limit cases are $S_{0,1}(x,y) = \frac{x-y}{\log(x/y)} = \Lambda(x,y)$ (logarithmic mean), $S_{-1,1}(x,y) = \sqrt{xy}$ (geometric mean) and $S_{0,-1}(x,y) = \frac{xy}{\Lambda(x,y)}$ (Scharfetter–Gummel mean). A list of further Stolarsky means is given in Table 2.

An explicit calculation shows that $\partial_x^2 S_{0,-1}(x,x) = -(3x)^{-1}$ and $\partial_x^2 S_{-1,1}(x,x) = -(4x)^{-1}$. For the general Stolarsky mean $S_{\alpha,\beta}$ one obtains (see Appendix A.3)

$$\partial_x^2 S_{\alpha,\beta}(x,x) = \frac{1}{12x}(\alpha+\beta-3) , \tag{3.5}$$

particularly reproducing the above findings for $\partial_x^2 S_{0,-1}$ and $\partial_x^2 S_{-1,1}$. With respect to (1.3)–(1.4), we observe that

$$B_1(V_i - V_j) = \frac{V_i - V_j}{e^{V_i - V_j} - 1} = S_{0,-1}(\pi_i, \pi_j)\,\pi_j^{-1} , \tag{3.6}$$

$$B_2(V_i - V_j) = e^{-\frac{1}{2}(V_i - V_j)} = S_{-1,1}(\pi_i, \pi_j)\,\pi_j^{-1} , \tag{3.7}$$

and (1.2) can be brought into the form (1.6), which we equally write as

$$-\sum_{j:j\sim i} \frac{m_{ij}}{h_{ij}} \kappa_{ij} S_*(\pi_i, \pi_j) \left(\frac{u_j^{\mathcal{T}}}{\pi_j} - \frac{u_i^{\mathcal{T}}}{\pi_i}\right) = f_i , \tag{3.8}$$

where S_* equals either $S_{0,-1}$ or $S_{-1,1}$. For general means $S_{\alpha,\beta}(x,y)$, we have the relation $S_{\alpha,\beta}(x,y) = xS_{\alpha,\beta}(1, y/x)$, such that the weight function for arbitrary parameters α and β reads

$$B_{\alpha,\beta}(x) = S_{\alpha,\beta}(1, e^{-x}) .$$

In particular, it holds for any α and β

$$B_{\alpha,\beta}(-x) = e^x B_{\alpha,\beta}(x) ,$$

which guarantees the consistency of the scheme with the thermodynamic equilibrium.

Interestingly, the derivation of the SQRA in Section 2.2 of [LFW13] relies on the assumption that the flux through a FV-interface has to be proportional to $\left(u_j^{\mathcal{T}}/\pi_j - u_i^{\mathcal{T}}/\pi_i\right)$ with the proportionality factor given by a suitable mean of π_i and π_j. The choice of $S_{-1,1}$ in [LFW13] seems arbitrary, yet it yields very good results [WE17, FKN+19, DHWK].

3.3 The Wasserstein gradient structure of the Fokker–Planck operator and the SQRA method

The choice of S_* turns out to be crucial for the convergence properties. In this section, we look at physical structures which are desirable to be preserved in the discretization procedure. Our considerations are based on the variational structure of the Fokker–Planck equation. Let us note at this point

mean	weight $B(x)$	α	β	$\alpha+\beta$
max	$\begin{cases} \mathrm{e}^{-x}, & x \le 0 \\ 1, & x > 0 \end{cases}$	$+\infty$	1	$+\infty$
quadratic mean	$\sqrt{\frac{1}{2}\left(1+\mathrm{e}^{-2x}\right)}$	4	2	6
arithmetic mean	$\frac{1}{2}\left(1+\mathrm{e}^{-x}\right)$	2	1	3
logarithmic mean	$\frac{1}{x}\left(1-\mathrm{e}^{-x}\right)$	0	1	1
geometric mean (SQRA)	$\mathrm{e}^{-x/2}$	-1	1	0
Scharfetter–Gummel mean	$x/\left(\mathrm{e}^{x}-1\right)$	0	-1	-1
harmonic Mean	$2/\left(\mathrm{e}^{x}+1\right)$	-2	-1	-3
min	$\begin{cases} \mathrm{e}^{x}, & x \le 0 \\ 1, & x > 0 \end{cases}$	$-\infty$	1	$-\infty$

Table 2: Examples for popular mean values expressed as (weighted) Stolarsky means $S_{\alpha,\beta}$ with corresponding weight functions in (1.2) that generalize the Bernoulli function. The geometric mean corresponds to the SQRA, the $S_{0,-1}$-mean to the Scharfetter–Gummel discretization.

that a physically reasonable discretization is not necessarily the best from the rate of convergence point of view. Indeed, this last point will be underlined by numerical simulations in Section 6. However, the physical consideration is helpful to understand the family of Stolarsky discretizations from a further, different point of view.

In [JKO98] it was proved that the Fokker–Planck equation

$$\dot{u} = \nabla \cdot \left(\kappa \nabla u + \kappa u \nabla V\right) \tag{3.9}$$

has the gradient flow formulation $\dot{u} = \partial_\xi \Psi^*(u, -\mathrm{D}E(u))$ where

$$E(u) = \int_\Omega \left(u \log u + Vu - u + 1\right) = \int_\Omega \left(u \log\left(\frac{u}{\pi}\right) - u + 1\right), \qquad \Psi^*(u,\xi) = \frac{1}{2}\int_\Omega \kappa u \left|\nabla \xi\right|^2, \tag{3.10}$$

and $\pi = \mathrm{e}^{-V}$ is the stationary solution of (3.9). Indeed, one easily checks that $\mathrm{D}E(u) = \log u + V = \log\left(\frac{u}{\pi}\right)$ and $\partial_\xi \Psi^*(u,\xi) = -\nabla \cdot (\kappa u \nabla \xi)$ such that it formally holds

$$\partial_\xi \Psi^*(u,\xi)|_{\xi=-\mathrm{D}E(u)} = -\nabla\cdot\left(\kappa u \nabla \xi\right)|_{\xi=-\mathrm{D}E(u)} = \nabla \cdot \left(\kappa u \left(\frac{\nabla u}{u} + \nabla V\right)\right) = \nabla \cdot \left(\kappa \nabla u + \kappa u \nabla V\right) = \dot{u}.$$

Given a particular partial differential equation, the gradient structure might not be unique. For example, the simple parabolic equation $\partial_t u = \Delta u$ can be described by (3.10) with $V = 0$. But at the same time one might choose $E(u) = \int u^2$ with $\Psi^*(\xi) = \int |\nabla \xi|^2$, which plays a role in phase field modeling (see [HMR11] and references therein) or $E(u) = -\int \log u$ with $\Psi^*(\xi) = \int u^2 |\nabla \xi|^2$.

In view of this observation, one might pose the question about “natural” gradient structures of the discretization schemes. This is reasonable if one believes that discretization schemes should incorporate the underlying physical principles. The energy functional is clearly prescribed by (3.10) with the natural discrete equivalent

$$E_{\mathcal{T}}(u) = \sum_i m_i \left(u_i \log\left(\frac{u_i}{\pi_i}\right) - u_i + 1 \right) . \tag{3.11}$$

The discrete linear evolution equation can be expected to be linear. Since we identified the continuous flux to be $\mathbf{J} = -\kappa\pi\nabla U$ with $U = u/\pi$, we expect the form

$$\dot{u}_i m_i = \partial_\xi \Psi^*_{\mathcal{T}}(u, -\mathrm{D}E_{\mathcal{T}}(u)) = \sum_{j:i\sim j} \frac{m_{ij}}{h_{ij}} \kappa_{i,j} \pi_{ij} \left(\frac{u_j}{\pi_j} - \frac{u_i}{\pi_i} \right) \tag{3.12}$$

for some suitably averaged π_{ij}. Equation (3.12) can be understood as a time-reversible (or detailed balanced) Markov process on the finite state space $\mathcal{P}$. Recently, various different gradient structures have been suggested for (3.12): [Mie11, Maa11, EM12, CHLZ12, Mie13b] for a quadratic dissipation as a generalization of the Jordan–Kinderlehrer–Otto approach; and [MPR14, MPPR17], where a dissipation of cosh-type was appeared in the Large deviation rate functional for a hydrodynamic limit of an interacting particle system. All of them can be written in the abstract form

$$\Psi^*_{\mathcal{T}}(u, \xi) = \frac{1}{2} \sum_i \frac{1}{m_i} \sum_{j:i\sim j} \frac{m_{ij}}{h_{ij}} S_{ij} a_{ij}(u, \pi) \psi^* \left(\xi_i - \xi_j \right) , \tag{3.13}$$

where

$$a_{ij}(u, \pi) = \left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j} \right) \partial_\xi \psi^* \left(\log\left(\frac{u_i}{\pi_i}\right) - \log\left(\frac{u_j}{\pi_j}\right) \right)^{-1} . \tag{3.14}$$

In fact, any positive and convex function ψ^* defines a reasonable dissipation functional Ψ^* by (3.13) and (3.14). A special case is when choosing for ψ^* and exponentially fast growing function $\psi^*(r) := \mathsf{C}^*(r) := 2\left(\cosh(r/2) - 1\right)$. Then a_{ij} simplifies to

$$a_{ij}(u, \pi) = \sqrt{\frac{u_i u_j}{\pi_i \pi_j}},$$

and hence, the square root appears. Choosing $S_{ij} = \sqrt{\pi_i \pi_j}$, we end up with a dissipation functional of the form

$$\Psi^*_{\mathcal{T}}(u, \xi) = \sum_i \sum_{j:i\sim j} m_{ij} h_{ij} \sqrt{u_i u_j} \, \frac{1}{h_{ij}^2} \mathsf{C}^* \left(\xi_i - \xi_j \right) . \tag{3.15}$$

There are (at least) three good reasons why choosing this gradient structure, i.e., modeling fluxes in exponential terms: a historical, a mathematical and a physical:

1 Already in Marcelin's PhD thesis from 1915 ([Mar15]) exponential reaction kinetics have been derived, which are still common in chemistry literature.

2 Recently, convergence for families of gradient systems has been derived based on the energy-dissipation principle (the so-called EDP-convergence [Mie16, LMPR17, DFM18]). Vice versa, the above cosh-gradient structure appears as an effective gradient structure applying EDP-convergence to Wasserstein gradient flow problems [LMPR17, FL19].

3 Recalling the gradient structure for the continuous Fokker–Planck equation (3.10), we observe that the dissipation mechanism Ψ^* is totally independent of the particular form of the energy

$\mathcal{E}$, which is determined by the potential V. This is physically understandable, since a change of the energy resulting, e.g., from external fields should not influence the dissipation structure. The same holds for the discretized version (3.15). In fact it was shown in [MS19], that the only discrete gradient structure, where the dissipation does not depend on V resp. $\pi = \mathrm{e}^{-V}$, is the cosh-gradient structure with the SQRA discretization $S_{ij} = S_{-1,1}(\pi_i, \pi_j)$. In particular, this characterizes the SQRA. For convenience, we add a proof for that to the Appendix A.2.

We think that these properties distinguish the SQRA, although in the following the convergence proofs do not really rely on the particular discretization weight S_{ij}.

Remark 3.3 (Convergence of energy and dissipation functional)**.** Let us finally make some comments on the convergence of $E_{\mathcal{T}}$ and $\Psi^*_{\mathcal{T}}$ given in (3.11) and (3.15) to the continuous analogies E and Ψ^*. Γ-convergence can be shown if the fineness of $\mathcal{T}$ tends to 0. For the energies it is clear, since $u \mapsto u \log(u/\pi) - u$ is convex. For the dissipation potentials $\Psi^*_{\mathcal{T}}(u,\xi)$ we observe the following: For smooth functions u and ξ, we have $\frac{1}{h_{ij}^2}\mathsf{C}^*\left(\xi_i - \xi_j\right) \approx \frac{1}{2}\left(\frac{x_i - x_j}{|x_i - x_j|} \cdot \nabla\xi\right)^2 + O(h_{ij}^2)$ and $\sqrt{u_i u_j} \approx u\left(\frac{1}{2}(x_i + x_j)\right)$. The considerations from Section 2.3 then yield $\Psi^*_{\mathcal{T}}(u,\xi) \approx \frac{1}{2}\int_{\boldsymbol{Q}} u \left|\nabla\xi\right|^2$.

For quadratic dissipation, qualitative convergence results in 1-D using the underlying gradient structure are obtained in [DL15] looking at energy-dissipation mechanism, and in [GKMP19] proving convergence of the metric.

4 Comparison of discretization schemes

We mutually compare any two discretization schemes of the form (1.6) in case of Dirichlet boundary conditions. In this case, even though the problem is only defined on $\tilde{\mathcal{P}}$, we can simply sum over all $\mathcal{P}$ once we multiplied with a test function that assumes the value 0 at all $\mathcal{P}\backslash\tilde{\mathcal{P}}$.

Let us recall the formula (2.10) for the fluxes

$$J_{ij}^S U = -\frac{\kappa_{ij}}{h_{ij}} S_{ij}(U_j - U_i).$$

Moreover, let $u_i = U_i \pi_i$ and $\tilde{u}_i = \tilde{U}_i \pi_i$ be the solution of the discrete FPE (1.6) for two different smooth mean coefficients $S_{ij} = S(\pi_i, \pi_j)$ and $\tilde{S}_{ij} = \tilde{S}(\pi_i, \pi_j)$ (e.g. once for Scharfetter–Gummel and once for SQRA) such that

$$\sum_{k:k\sim i} m_{ik} h_{ik} J_{ik}^S U = m_i \bar{f}_i \tag{4.1}$$

$$\sum_{k:k\sim i} m_{ik} h_{ik} J_{ik}^{\tilde{S}} \tilde{U} = m_i \bar{f}_i. \tag{4.2}$$

In order to compare the solutions of (4.1) and (4.2) we take the difference of these two equations and multiply with $E_i = U_i - \tilde{U}_i$. We obtain

$$\begin{aligned} 0 &= \sum_i \sum_{k:k\sim i} m_{ik} h_{ik} \left(J_{ik}^S U - J_{ik}^{\tilde{S}} \tilde{U}\right) E_i \\ &= \sum_i \sum_{k:\,k\sim i} \frac{m_{ik}}{h_{ki}} \kappa_{ij} (S_{ik}(U_i - U_k) - \tilde{S}_{ik}(\tilde{U}_i - \tilde{U}_k)) E_i \end{aligned}$$

Introducing the notation $\alpha_{ik} = \kappa_{ik}\frac{m_{ik}}{h_{ik}}$ and using (2.2) we get

$$
\begin{aligned}
0 &= \sum_{k\sim i} \alpha_{ik}\left(S_{ik}(U_i - U_k) - S_{ik}(\tilde{U}_i - \tilde{U}_k) + \left(S_{ik} - \tilde{S}_{ik}\right)(\tilde{U}_i - \tilde{U}_k)\right)(E_i - E_k)\\
&= \sum_{k\sim i} \alpha_{ik}\left(S_{ik}\left(E_i - E_k\right) + \left(S_{ik} - \tilde{S}_{ik}\right)\left(\tilde{U}_i - \tilde{U}_k\right)\right)\left(E_i - E_k\right).
\end{aligned}
$$

Using the notation $\mathrm{D}_{ik}A = A_k - A_i$ for discrete gradients

$$
\left(\tilde{S}_{ik} - S_{ik}\right)\left(\tilde{U}_i - \tilde{U}_k\right)\left(E_i - E_k\right) \le \frac{1}{2}\left[S_{ik}\left(\mathrm{D}_{ik}E\right)^2 + \frac{(S_{ik} - \tilde{S}_{ik})^2}{S_{ik}}\left(\mathrm{D}_{ik}\tilde{U}\right)^2\right]
$$

we get

$$
\frac{1}{2}\sum_{k\sim i}\alpha_{ik}S_{ik}\left(\mathrm{D}_{ik}E\right)^2 \le \frac{1}{2}\sum_{k\sim i}\frac{(\tilde{S}_{ik} - S_{ik})^2}{S_{ik}\tilde{S}_{ik}}\alpha_{ik}\tilde{S}_{ik}\left(\mathrm{D}_{ik}\tilde{U}\right)^2. \tag{4.3}
$$

In the case of Stolarsky means the constants are more explicit. We have the following expansion of S_{ij}: writing $\pi_{ij} = \frac{1}{2}\left(\pi_i + \pi_j\right)$, $\pi_+ = \pi_- = \frac{1}{2}\left(\pi_i - \pi_j\right)$ and $\pi_i = \pi_0 + \pi_+$ and $\pi_j = \pi_0 - \pi_-$

$$
\begin{aligned}
S_{ij} &= S_{\alpha,\beta}\left(\pi_{ij},\pi_{ij}\right) + \frac{1}{2}\left(\pi_+ - \pi_-\right) + \frac{1}{2}\partial_x^2 S_{\alpha,\beta}\left(\pi_{ij},\pi_{ij}\right)\left(\pi_+ + \pi_-\right)^2 + O\left(\pi_\pm^3\right)\\
&= \pi_{ij} + \frac{\frac{1}{3}\left(\alpha+\beta\right) - 1}{8\,\pi_{ij}}\left(\pi_i - \pi_j\right)^2 + O\left(\pi_i - \pi_j\right)^3.
\end{aligned} \tag{4.4}
$$

In case $\left(\alpha+\beta\right) = \left(\tilde{\alpha}+\tilde{\beta}\right)$, we obtain $S_{ij} - \tilde{S}_{ij} = O\left(\pi_i - \pi_j\right)^3$ and hence this yields the following first comparison result:

Theorem 4.1. *Let $\mathcal{T}$ be a mehs with right hand side $f \in L^2(\mathcal{P})$ and let u and $\tilde{u}$ be a two solution of the discrete FPE for different Stolarsky mean coefficients $S_{ij} = S_{\alpha,\beta}\left(\pi_i,\pi_j\right)$ and $\tilde{S}_{ij} = S_{\tilde{\alpha},\tilde{\beta}}\left(\pi_i,\pi_j\right)$ respectively. Then*

$$
\begin{aligned}
&\frac{1}{2}\sum_{k\sim i}\kappa_{ik}\frac{m_{ik}}{h_{ik}}S_{ik}\left(\mathrm{D}_{ik}E\right)^2\\
&\qquad\le \frac{1}{2}\sum_{k\sim i}\left(\frac{\left(\left(\alpha+\beta\right) - \left(\tilde{\alpha}+\tilde{\beta}\right)\right)^2}{24^2\,\pi_{ij}^2\tilde{S}_{ik}S_{ik}}\left(\pi_i - \pi_j\right)^4 + O\left(\pi_i - \pi_j\right)^5\right)\kappa_{ik}\frac{m_{ik}}{h_{ik}}\left(\mathrm{D}_{ik}\tilde{U}\right)^2
\end{aligned}
$$

In case $\left(\alpha+\beta\right) = \left(\tilde{\alpha}+\tilde{\beta}\right)$ we furthermore find

$$
\frac{1}{2}\sum_{k\sim i}\kappa_{ik}\frac{m_{ik}}{h_{ik}}S_{ik}\left(\mathrm{D}_{ik}E\right)^2 \le \frac{1}{2}\sum_{k\sim i}O\left(\pi_i - \pi_j\right)^6\kappa_{ik}\frac{m_{ik}}{h_{ik}}\left(\mathrm{D}_{ik}\tilde{U}\right)^2.
$$

We aim to refine the above result to an order of convergence result for $J^S U - J^{\tilde{S}}\tilde{U}$.. We introduce the auxiliary smooth mean $\hat{S}_{ik} = \hat{S}(\pi_i,\pi_k)$ and find

$$
\begin{aligned}
\hat{S}_{ik}\left(E_i - E_k\right) &= \hat{S}_{ik}\left(U_i - \tilde{U}_i - \left(U_k - \tilde{U}_k\right)\right)\\
&= S_{ik}(U_i - U_k) - S_{ik}(U_i - U_k) + \tilde{S}_{ik}(\tilde{U}_i - \tilde{U}_k) - \tilde{S}_{ik}(\tilde{U}_i - \tilde{U}_k) + \hat{S}_{ik}\left(U_i - \tilde{U}_i - \left(U_k - \tilde{U}_k\right)\right)\\
&= m_{ik}\alpha_{ik}^{-1}\left(J_{ik}^S U - J_{ik}^{\tilde{S}}\tilde{U}\right) + \left(\hat{S}_{ik} - S_{ik}\right)\left(U_i - U_k\right) - \left(\hat{S}_{ik} - \tilde{S}_{ik}\right)\left(\tilde{U}_i - \tilde{U}_k\right).
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
&\sum_{k\sim i}\alpha_{ik}\left(S_{ik}(U_i-U_k)-\tilde S_{ik}\left(\tilde U_i-\tilde U_k\right)\right)(E_i-E_k)\\
&\quad=\sum_{k\sim i}\frac{h_{ik}}{\kappa_{ik}}m_{ik}\frac{1}{\hat S_{ik}}\left(J_{ik}^{S}U-J_{ik}^{\tilde S}\tilde U\right)^2\\
&\quad+\sum_{k\sim i}m_{ik}\frac{1}{\hat S_{ik}}\left(J_{ik}^{S}U-J_{ik}^{\tilde S}\tilde U\right)\left[\left(\hat S_{ik}-S_{ik}\right)(U_i-U_k)+\left(\hat S_{ik}-\tilde S_{ik}\right)\left(\tilde U_i-\tilde U_k\right)\right],
\end{aligned}
$$

and using Cauchy-Schwartz inequality, we get

$$
\begin{aligned}
\sum_{k\sim i}\alpha_{ik}\left(S_{ik}(U_i-U_k)-\tilde S_{ik}\left(\tilde U_i-\tilde U_k\right)\right)(E_i-E_k)&\le-\frac{1}{2}\sum_{k\sim i}\frac{h_{ik}m_{ik}}{\kappa_{ik}}\frac{1}{\hat S_{ik}}\left(J_{ik}^{S}U-J_{ik}^{\tilde S}\tilde U\right)^2\\
&+\sum_{k\sim i}\frac{m_{ik}\kappa_{ik}}{h_{ik}\hat S_{ik}}\left(\left(\hat S_{ik}-S_{ik}\right)^2(U_i-U_k)^2+\left(\hat S_{ik}-\tilde S_{ik}\right)^2\left(\tilde U_i-\tilde U_k\right)^2\right).
\end{aligned}
$$

Altogether we obtain

$$
\begin{aligned}
\frac{1}{2}\sum_{k\sim i}\frac{h_{ik}m_{ik}}{\kappa_{ik}}\frac{1}{\hat S_{ik}}\left(J_{ik}^{S}U-J_{ik}^{\tilde S}\tilde U\right)^2&\le\sum_{k\sim i}\frac{m_{ik}h_{ik}}{\kappa_{ik}\hat S_{ik}S_{ik}^2}\left(\hat S_{ik}-S_{ik}\right)^2\left(\frac{\kappa_{ik}}{h_{ik}}S_{ik}\left(U_i-U_k\right)\right)^2\\
&+\sum_{k\sim i}\frac{m_{ik}h_{ik}}{\kappa_{ik}\hat S_{ik}\tilde S_{ik}^2}\left(\hat S_{ik}-\tilde S_{ik}\right)^2\left(\frac{\kappa_{ik}}{h_{ik}}\tilde S_{ik}\left(\tilde U_i-\tilde U_k\right)\right)^2.
\end{aligned}
$$

We make once more use of (4.4) writing $C_{\alpha,\beta}:=\frac{\alpha+\beta}{24}$ and exploiting $\pi_i=\pi_{ij}+\pi_{ij}\left(V_i-V_{ij}\right)+O\left(V_i-V_{ij}\right)^2$ with

$$
\begin{aligned}
\pi_i-\pi_j&\approx\pi_{ij}\left(V_i-V_j\right)+O\left(V_i-V_{ij}\right)^2+O\left(V_j-V_{ij}\right)^2\\
S_{ij}&\approx\pi_{ij}+O\left(\pi_i-\pi_j\right).
\end{aligned}
$$

Hence, we conclude the following result.

Theorem 4.2. *Let $\mathcal{T}$ be a mesh with right hand side $f\in L^2(\mathcal{P})$ and let u and $\tilde u$ be two solutions of the discrete FPE for different Stolarsky means S and $\tilde S$. Moreover, let $\hat S$ be any Stolarsky mean and assume that either $\alpha+\beta\neq\hat\alpha+\hat\beta$ or $\tilde\alpha+\tilde\beta\neq\hat\alpha+\hat\beta$. Then the solutions u and $\tilde u$ of the discretized FPE satisfy the symmetrized error estimate up to higher order*

$$
\begin{aligned}
\frac{1}{2}\sum_{k\sim i}\frac{h_{ik}m_{ik}}{\kappa_{ik}}\frac{1}{\hat S_{ik}}\left(J_{ik}^{S}U-J_{ik}^{\tilde S}\tilde U\right)^2&\le\sum_{k\sim i}\frac{m_{ik}h_{ik}}{\kappa_{ik}S_{ik}}\left(C_{\alpha,\beta}-C_{\hat\alpha,\hat\beta}\right)\left(V_i-V_j\right)^2\left(J_{ik}^{S}U\right)^2\\
&+\sum_{k\sim i}\frac{m_{ik}h_{ik}}{\kappa_{ik}\tilde S_{ik}}\left(C_{\tilde\alpha,\tilde\beta}-C_{\hat\alpha,\hat\beta}\right)\left(V_i-V_j\right)^2\left(J_{ik}^{\tilde S}\tilde U\right)^2.
\end{aligned}
$$

More general, for any mean we have

$$
\begin{aligned}
&\frac{1}{2\kappa^*}\|J^SU-J^{\tilde S}\tilde U\|^2_{L^2_{\hat S}(\mathcal{E})}\\
&\qquad\le\frac{1}{\kappa_*}\left\{\sup_{i,k}\frac{\left(\hat S_{ik}-S_{ik}\right)^2}{\hat S_{ik}S_{ik}}\left\|J^SU\right\|^2_{L^2_S(\mathcal{E})}+\sup_{i,k}\frac{\left(\hat S_{ik}-\tilde S_{ik}\right)^2}{\hat S_{ik}\tilde S_{ik}}\left\|J^{\tilde S}\tilde U\right\|^2_{L^2_{\tilde S}(\mathcal{E})}\right\},
\end{aligned}
\tag{4.5}
$$

and in particular for Stolarsky means with $\alpha+\beta=\tilde\alpha+\tilde\beta=\hat\alpha+\hat\beta$ we find the following result:

Corollary 4.3. *Let $\mathcal{T}$ be a mesh with right hand side $f \in L^2(\mathcal{P})$ and let u and $\tilde{u}$ be two solutions of the discrete FPE for different Stolarsky mean coefficients $S_{ij} = S_{\alpha,\beta}(\pi, \pi_j)$ and $\tilde{S}_{ij} = S_{\tilde{\alpha},\tilde{\beta}}(\pi, \pi_j)$ with $\alpha + \beta = \tilde{\alpha} + \tilde{\beta} = \hat{\alpha} + \hat{\beta}$. Then estimate* (4.5) *holds. In particular, we find the refined estimate*

$$\frac{1}{2\kappa^*}\|J^S U - J^{\tilde{S}}\tilde{U}\|^2_{L^2_{\hat{S}}(\mathcal{E})} \leq O\left(\pi_i - \pi_j\right)^6 \left(\left\|J^S U\right\|^2_{L^2_S(\mathcal{E})} + \left\|J^{\tilde{S}}\tilde{U}\right\|^2_{L^2_{\tilde{S}}(\mathcal{E})}\right).$$

In particular, the last result shows that convergence rates are similar up to order 3 for different α, β which satisfy $\alpha + \beta = \mathrm{const}$.

5 Convergence of the discrete FPE

In this section, we derive general estimates for the order of convergence of the Stolarsky FV operators. Throughout this section, we assume that the mesh satisfies the consistency property of Definition 2.10 with a suitable consistency function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and discretization operator $\mathcal{R}_\mathcal{T} : H^1(\Omega) \supset B \to L^2(\mathcal{P})$. The parameters π_i are then given in terms of $\pi_i = \left(\mathcal{R}_\mathcal{T}\pi\right)_i$.

We derive consistency errors for U in Section (5.1) and consistency errors for u in Section (5.2).

5.1 Error Analysis in U

In what follows, we assume that the discrete and the continuous solution satisfy Dirichlet conditions. In view of the continuous and the discrete FPE given in the form (2.6) and (2.7) as well as formula (2.16) we observe that the natural variational consistency error for a given Stolarsky mean S takes the form

$$\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U; v\right) = \sum_{i\sim j}\left(v_j - v_i\right)\left(\int_{\sigma_{ij}} \kappa\pi\nabla U \cdot \boldsymbol{\nu}_{ij} - \kappa_{ij} S_{ij}\frac{m_{ij}}{h_{ij}}\left(\left(\mathcal{R}_\mathcal{T} U\right)_j - \left(\mathcal{R}_\mathcal{T} U\right)_i\right)\right).$$

We recall that an estimate for $\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U; \cdot\right)$ implies an order of convergence estimate by (2.15). Our main result of this section provides a connection between $\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U; \cdot\right)$ and the variational consistency $\mathfrak{E}_\mathcal{T}\left(U; \cdot\right)$ (given by (2.16)) of the second order equation

$$-\nabla \cdot \left(\kappa \nabla U\right) = f$$

with the discretization scheme

$$\forall i : \qquad -\sum_{j:j\sim i} \kappa_{ij}\frac{m_{ij}}{h_{ij}}\left(U_j^\mathcal{T} - U_i^\mathcal{T}\right) = f_i\,.$$

Proposition 5.1. *Let $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$ be a mesh. The variational consistency error $\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U; \cdot\right)$ can be estimated by*

$$\left\|\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U; \cdot\right)\right\|^2_{H^*_{\mathcal{T},\kappa S}} \leq \|\pi\|_\infty \left\|\mathfrak{E}_\mathcal{T}\left(U; \cdot\right)\right\|^2_{H^*_{\mathcal{T},\kappa}} + \sum_{i\sim j}\frac{h_{ij}}{m_{ij}}\kappa_{ij}^{-1} S_{ij}^{-1}\left(\int_{\sigma_{ij}}\left(\pi - S_{ij}\right)\kappa\nabla U \cdot \boldsymbol{\nu}_{ij}\right)^2. \tag{5.1}$$

Proof. For simplicity, we write $\hat{U} := \mathcal{R}_{\mathcal{T}}U$. We observe that

$$\begin{aligned}\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}\left(U;v\right) &= -\sum_{i\sim j}\left(v_j - v_i\right) m_{ij}\left(\overline{J}_{ij}U - J^S_{ij}\hat{U}\right)\\ &= -\sum_{i\sim j}\left(v_j - v_i\right) m_{ij}\left(\left(\overline{J}_{ij}U - \overline{J}^\star_{ij}U\right) + \left(\overline{J}^\star_{ij}U - J^S_{ij}\hat{U}\right)\right),\end{aligned}$$

where

$$\overline{J}^\star_{ij}U := -m_{ij}^{-1}\int_{\sigma_{ij}}\kappa S_{ij}\nabla U\cdot\boldsymbol{\nu}_{ij}\,.$$

satisfies

$$m_{ij}\left|\overline{J}_{ij}U - \overline{J}^\star_{ij}U\right| \leq \left|\int_{\sigma_{ij}}\left(\pi - S_{ij}\right)\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right|. \tag{5.2}$$

Using the fact that

$$m_{ij}\left(\overline{J}^\star_{ij}U - J^S_{ij}\hat{U}\right) = -S_{ij}\left(\int_{\sigma_{ij}}\kappa\nabla U\cdot\boldsymbol{\nu}_{ij} - \kappa_{ij}\frac{m_{ij}}{h_{ij}}\left(\hat{U}_j - \hat{U}_i\right)\right)$$

we obtain

$$\left|\sum_{i\sim j}\left(v_j - v_i\right) m_{ij}\left(\overline{J}^\star_{ij}U - J^S_{ij}\hat{U}\right)\right| \tag{5.3}$$

$$\leq \|v\|_{H_{\mathcal{T},\kappa S}}\left(\sup_{ij} S_{ij}\right)^{\frac{1}{2}}\left(\sum_{i\sim j}\frac{h_{ij}}{m_{ij}}\kappa_{ij}^{-1}\left(\int_{\sigma_{ij}}\kappa\nabla U\cdot\boldsymbol{\nu}_{ij} - \kappa_{ij}\frac{m_{ij}}{h_{ij}}\left(\hat{U}_j - \hat{U}_i\right)\right)^2\right)^{\frac{1}{2}}. \tag{5.4}$$

From (5.2) we conclude

$$\left|\sum_{i\sim j}\left(v_j - v_i\right) m_{ij}\left(\overline{J}_{ij}U - \overline{J}^\star_{ij}U\right)\right| \leq \|v\|_{H_{\mathcal{T},\kappa S}}\left(\sum_{i\sim j}\frac{h_{ij}}{m_{ij}}\kappa_{ij}^{-1}S_{ij}^{-1}\left(\int_{\sigma_{ij}}\left(\pi - S_{ij}\right)\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right)^2\right)^{\frac{1}{2}}. \tag{5.5}$$

Taking together (5.3)–(5.5) we obtain (5.1). □

Lemma 5.2. *Assume there exists a constant $C > 0$ such that for all cells Ω_i, Ω_j with $h_i = \mathrm{diam}\Omega_i$ it holds*

$$\|f\|^2_{L^2(\sigma_{ij})} \leq \frac{1}{h_i}C^2\|f\|^2_{H^1(\Omega_i)}\,. \tag{5.6}$$

Then for C^2-smooth means S

$$\left|\int_{\sigma_{ij}}\left(\pi - S_{ij}\right)\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right| \leq 2C\left(m_{ij}h_i\right)^{\frac{1}{2}}\|\kappa\nabla U\|_{H^1(\Omega_i)}\,. \tag{5.7}$$

Remark 5.3. Note that (5.6) can be easily verified for cubes.

Proof. Observe that

$$\begin{aligned}\int_{\sigma_{ij}}\left|\pi - S_{ij}\right|\left|\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right| &\leq \left(\int_{\sigma_{ij}}\left|\pi - S_{ij}\right|^2\right)^{\frac{1}{2}}\left(\int_{\sigma_{ij}}\left|\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right|^2\right)^{\frac{1}{2}}\\ &\leq c\left(\int_{\sigma_{ij}}\left|\pi - S_{ij}\right|^2\right)^{\frac{1}{2}}\left(\frac{1}{h_i}\|\kappa\nabla U\|^2_{H^1(\Omega_i)}\right)^{\frac{1}{2}}.\end{aligned} \tag{5.8}$$

It remains to study $\frac{1}{m_{ij}}\int_{\sigma_{ij}}|\pi - S_{ij}|^2$ in more detail. We have

$$\pi - S_{ij} = \frac{1}{2}(\pi - \pi_i) + \frac{1}{2}(\pi - \pi_j) + \left(\frac{\pi_i + \pi_j}{2} - S_{ij}\right).$$

The first term can be estimated by $|\pi - \pi_i| \le h_i \cdot \nabla\pi + O(h_i^2)$ and a similar estimate holds for the second term. The last term, assuming that the mean is C^2-smooth, can be estimated by

$$S(\pi_i, \pi_j) - S\left(\frac{\pi_i + \pi_j}{2}, \frac{\pi_i + \pi_j}{2}\right) = \frac{1}{2}(\pi_i - \pi_j)\nabla S \cdot (1, -1)^T + O(|\pi_i - \pi_j|).$$

Using that $\pi_i - \pi_j = \nabla\pi \cdot h_{ij} + O(h_{ij})$ and that $S\left(\frac{\pi_i+\pi_j}{2}, \frac{\pi_i+\pi_j}{2}\right) = \frac{\pi_i+\pi_j}{2}$, we obtain that $|\pi - S_{ij}|^2 \le O\left(h_i^2\right)$. In total we obtain

$$\int_{\sigma_{ij}} |\pi - S_{ij}|\,|\kappa\nabla U \cdot \boldsymbol{\nu}_{ij}| \le 2C\left(m_{ij}h_i^2\right)^{\frac{1}{2}}\left(\frac{1}{h_i}\|\kappa\nabla U\|^2_{H^1(\Omega_i)}\right)^{\frac{1}{2}}.$$

□

Using the above estimates, we can now show the main result of the section.

Theorem 5.4 (Localized order of convergence)**.** *Let the mesh $\mathcal{T}$ be admissible in sense of Definition 2.1 and consistent in sense of Definition 2.10. Let $u \in C_0^2(\Omega)$ be the solution to* (1.1)*. Let $f^{\mathcal{T}} := \mathcal{R}^*_{\mathcal{T}} f$ and let $u^{\mathcal{T}} \in \mathcal{S}^{\mathcal{T}}$ be the solution to* (2.3)*. Moreover, let $\kappa \le \kappa^*$, $b > 0$ and $S \in C^2(\mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0})$. Then it holds it holds*

$$\|u^{\mathcal{T}} - \mathcal{R}_{\mathcal{T}}u\|^2_{H_{\mathcal{T}}} \le C(\kappa_*, \pi, d, \|U\|_{C^2}) \times \left(\varphi(h)^2 + h^2\right).$$

Proof. Inserting estimate (5.7) int to the estimate of the variational consistency, we get

$$\begin{aligned}\|\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}(U;\cdot)\|^2_{H^*_{\mathcal{T},\kappa S}} &\le \|\pi\|_\infty \|\mathfrak{E}_{\mathcal{T}}(U;\cdot)\|^2_{H^*_{\mathcal{T},\kappa}} + C\sum_{i\sim j} h_{ij}\kappa_{ij}^{-1}S_{ij}^{-1}h_i\|\kappa\nabla U\|^2_{H^1(\Omega_i)}\\ &\le \|\pi\|_\infty \|\mathfrak{E}_{\mathcal{T}}(U;\cdot)\|^2_{H^*_{\mathcal{T},\kappa}} + C(\kappa_*, \pi, d)\, h^2 \sum_i \|\kappa\nabla U\|^2_{H^1(\Omega_i)}.\end{aligned}$$

Using (2.15) we obtain an estimate for the discretization error in the form

$$\|u^{\mathcal{T}} - \mathcal{R}_{\mathcal{T}}u\|^2_{H_{\mathcal{T}}} \le \|\pi\|_\infty \|\mathfrak{E}_{\mathcal{T}}(U;\cdot)\|^2_{H^*_{\mathcal{T},\kappa}} + C(\kappa_*, \pi, d, \|U\|_{C^2})\,\mathrm{Size}(\mathcal{T})^2.$$

Using the consistency assumption on the discretization of the pure elliptic problem we obtain the desired estimate. □

5.2 Error Analysis in u

In the following, we will discuss how to derive bounds on the rate of convergence of u instead of U. As a basis for both proofs of this section, we start with the discrete FP operator which we rewrite as

$$-\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}S_{ij}\left(\frac{u_j}{\pi_j} - \frac{u_i}{\pi_i}\right) = -\sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}(u_j - u_i) - \sum_{j:j\sim i}\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(\frac{S_{ij} - \pi_j}{\pi_j}u_j - \frac{S_{ij} - \pi_i}{\pi_i}u_i\right).$$

We have

$$\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}(U;v) - \mathfrak{E}_{\mathcal{T}}(u;v) = \sum_{i\sim j}\left(\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(\frac{S_{ij} - \pi_j}{\pi_j}u_j - \frac{S_{ij} - \pi_i}{\pi_i}u_i\right) - \int_{\sigma_{ij}}\kappa u\nabla V \cdot \boldsymbol{\nu}_{ij}\right)(v_j - v_i),$$

where we want to estimate the right-hand side. For $V_i - V_j = O(h)$ we have

$$\frac{S_{ij}-\pi_j}{\pi_j} = \frac{1}{2}\left(\frac{\pi_i}{\pi_j}-1\right) + O(\pi_i - \pi_j) = \frac{1}{2}(V_j - V_i) + O(\pi_i-\pi_j) + O(V_i - V_j) \tag{5.9}$$

and hence

$$\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}(U;v) - \mathfrak{E}_{\mathcal{T}}(u;v) = \sum_{i\sim j}\left(\frac{m_{ij}}{h_{ij}}\kappa_{ij}\frac{1}{2}(V_j-V_i)(u_i+u_j) - \int_{\sigma_{ij}}\kappa u\nabla V\cdot\boldsymbol{\nu}_{ij} + O(h)\right)(v_j - v_i)\,.$$

Since $\kappa_{ij}\approx\kappa.$, $\frac{u_i+u_j}{2}\approx u$, $\frac{V_j-V_i}{h_{ij}}\approx\nabla V$ it holds $\mathfrak{E}_{\mathcal{T},\mathrm{FPE}}(U;v)\approx\mathfrak{E}_{\mathcal{T}}(u;v)$.

Theorem 5.5. *For smooth potentials $V\in C^2$ it holds $\|\mathfrak{E}_{\mathcal{T}}(u;v)\|_{H^*_{\mathcal{T},\kappa S}} = O(h)$.*

Remark 5.6. The calculation (5.9) is an approximation for small values of $|V_j - V_i|$. In the particular case of large discrete gradients a general approximation of $\frac{S_{ij}-\pi_j}{\pi_j}$ is not at hand. However, in the SG case $S_* = S_{0,-1}$ we observe (compare with (1.5) and (3.6)) introducing $f(x) = \frac{-x-e^x-1}{(e^x-1)x}$ (with $f(x)\to 0$ as $x\to+\infty$ and $f(x)\to 1$ as $x\to-\infty$)

$$\begin{aligned}\frac{1}{h_{ij}}\frac{S_{ij}-\pi_j}{\pi_j} &= \frac{1}{h_{ij}}\frac{V_j - V_i - (\mathrm{e}^{V_i-V_j}-1)}{\mathrm{e}^{V_i-V_j}-1}\\ &= \frac{V_i-V_j}{h_{ij}}f(V_i-V_j)\to\begin{cases}-\nabla V\cdot\boldsymbol{\nu}_{ij} & \text{if } V_i\gg V_j\\ 0 & \text{if } V_j\gg V_i\end{cases}\quad\text{as } h_{ij}\to 0\,.\end{aligned}$$

Hence we observe that the SG method is particularly suited to minimize the error term

$$\frac{m_{ij}}{h_{ij}}\kappa_{ij}\left(\frac{S_{ij}-\pi_j}{\pi_j}u_j - \frac{S_{ij}-\pi_i}{\pi_i}u_i\right) - \int_{\sigma_{ij}}\kappa u\nabla V\cdot\boldsymbol{\nu}_{ij}$$

for large gradients ∇V.

5.3 Qualitative comparison on cubic meshes

In view of Section 2.6 we consider a polygonal domain $\Omega\subset\mathbb{R}^d$ with $d\le 3$ and a cubic mesh where $\Omega_i = x_i + [-h/2,h/2]^d$, $x_i\in h\mathbb{Z}\subset\Omega$ to show that $\left|\int_{\sigma_{ij}}(\pi - S_{ij})\kappa\nabla U\cdot\boldsymbol{\nu}_{ij}\right| = O(h^2)$. In fact the following calculations are quite standard and, therefore, we shorten our considerations. We have for $x\in\sigma_{ij}$

$$\begin{aligned}S_{ij}-\pi(x) &= S(\pi_i,\pi_j) - S(\pi(x),\pi(x)) =\\ &= \nabla S(x)\cdot\begin{pmatrix}\pi_i-\pi(x)\\ \pi_j-\pi(x)\end{pmatrix} + \begin{pmatrix}\pi_i-\pi(x)\\ \pi_j-\pi(x)\end{pmatrix}\cdot\nabla^2 S(x)\cdot\begin{pmatrix}\pi_i-\pi(x)\\ \pi_j-\pi(x)\end{pmatrix} + O(h^3).\end{aligned}$$

Moreover, we have $\pi_i-\pi(x) = \nabla\pi\cdot(x_i-x)$. The gradient of S is given by $(1/2,1/2)^T$ and hence, we $S_{ij}-\pi(x) = \frac{\pi_i+\pi_j-2\pi(x)}{2} + O(h^2)$. We compute the first term in more detail. We have $\pi_j-\pi(x) = \nabla\pi\cdot(x_j-x)$ and $\pi_i-\pi(x) = \nabla\pi\cdot(x_i-x)$ and the sum yields $\pi_i+\pi_j-2\pi(x) = \nabla\pi\cdot(x_i+x_j-2x) = \frac{1}{2}\nabla\pi\cdot\tilde{x}$, where $\tilde{x} = x - \frac{x_i + x_j}{2}$ the coordinate on the cell surface with respect to the middle point $\bar{x} = \frac{x_i + x_j}{2}$. Hence, we get

$$\int_{\sigma_{ij}} (\pi - S_{ij}) \kappa \nabla U \cdot \nu_{ij} = \frac{1}{4} \int_{\sigma_{ij}} \nabla \pi(x) \cdot \tilde{x} \kappa(x) \nabla U(x) \cdot \nu_{ij} \mathrm{d}\sigma(\tilde{x}) + O(h^2).$$

Now we can fix the function $s(x) = \kappa(x) \nabla U(x) \cdot \nu_{ij} \nabla \pi(x)$ with respect to $\bar{x}$. We have $s(x) = s(\bar{x}) + (x - \bar{x}) \nabla s(\bar{x}) + O(h^2)$, which implies (assuming that $U, \pi \in \mathrm{C}^2$ and $\kappa \in \mathrm{C}^1$) that $\int_{\sigma_{ij}} (\pi - S_{ij}) \kappa \nabla U \cdot \nu_{ij} = \frac{1}{4} \int_{\sigma_{ij}} (s(\bar{x}) + (x - \bar{x}) \nabla s(\bar{x})) \cdot \tilde{x} \mathrm{d}\sigma(\tilde{x}) + O(h^2) = \frac{1}{4} \int_{\sigma_{ij}} s(\bar{x}) \cdot \tilde{x} \mathrm{d}\sigma(\tilde{x}) + O(h^2)$. But the first vanishes, since the interface σ_{ij} is symmetric w.r.t. the mid point $\bar{x}$ and we are integrating along $\tilde{x}$. Hence, we have $\left| \int_{\sigma_{ij}} (\pi - S_{ij}) \kappa \nabla U \cdot \boldsymbol{\nu}_{ij} \right| = O(h^2)$.

Hence, iterating the above argument twice for κ and π and exploiting in the first step Theorem 2.12 we proved the following.

Theorem 5.7. *Let $d \leq 3$. On a polygonal domain $\Omega \subset \mathbb{R}^d$ with a cubic mesh where $\Omega_i = x_i + [-h/2, h/2]^d$, $x_i \in h\mathbb{Z} \subset \Omega$, it holds*

$$\| \mathfrak{E}_{\mathcal{T},\mathrm{FPE}} (U; \cdot) \|_{H^*_{\mathcal{T},\kappa S}} \leq C h^2.$$

6 Numerical simulation and convergence analysis

In this section, we provide a numerical convergence analysis of the flux discretization schemes based on weighted Stolarsky means described in the previous sections. For the sake of simplicity, we restrict ourselves to one-dimensional examples, for which already non-trivial results can be observed.

Example 6.1. We consider the potential $V(x) = 2 \sin(2\pi x)$ and the right hand side $f(x) = x(1 - x)$ on $x = (0, 1)$. We assume the diffusion constant $\kappa = 1$ and Dirichlet boundary conditions $u(0) = 0$ and $u(1) = 1$. The Stolarsky mean discretizations are compared point-wise with a numerically computed reference solution u_{ref} (and J_{ref}) that was obtained by the shooting method (using a fourth order Runge–Kutta scheme) in combination with Brent's root finding algorithm [Bre71] on a very fine grid with 136474 nodes ($h \approx 7.3 \times 10^{-6}$).

The convergence results are summarized in Fig. 1. In Figure 1 (a), the logarithmic error $\log_{10}(\|u - u_{\text{ref}}\|_{L_2})$ is shown in the (α, β)-plane of the Stolarsky mean parameters for an equidistant mesh with $2^{10} + 1 = 1025$ nodes. First, we note that the accuracy for a mean $S_{\alpha,\beta}$ is indeed practically invariant along $\alpha + \beta = \mathrm{const}$, which is consistent with our analytical result in Section 4. In this particular example, we observe optimal accuracy at about $\alpha + \beta \approx 4.2$. This coincides with the convergence results under mesh refinement shown in Figure 1 (b), where the fastest convergence is obtained for the scheme involving the $S_{3.2,1}$-mean. The other considered schemes, however, show as well a quadratic convergence behavior with a slightly larger constant. Interestingly, for the same example, we find that the optimal mean for an accurate approximation of the flux J is on $\alpha + \beta = -3$, see Figure 1 (c). This is further evidences in Figure 1 (d), where the harmonic mean $S_{-1,-2}$ converges significantly faster than the other schemes. Obviously, in the present example, the minimal attainable error for both u and J can not be achieved by the same discretization scheme.

Example 6.2. We consider the potential $V(x) = 5(x + 1)x$. The right hand side function, the diffusion constant and the boundary conditions are the same as in Example 6.1. The problem has an exact

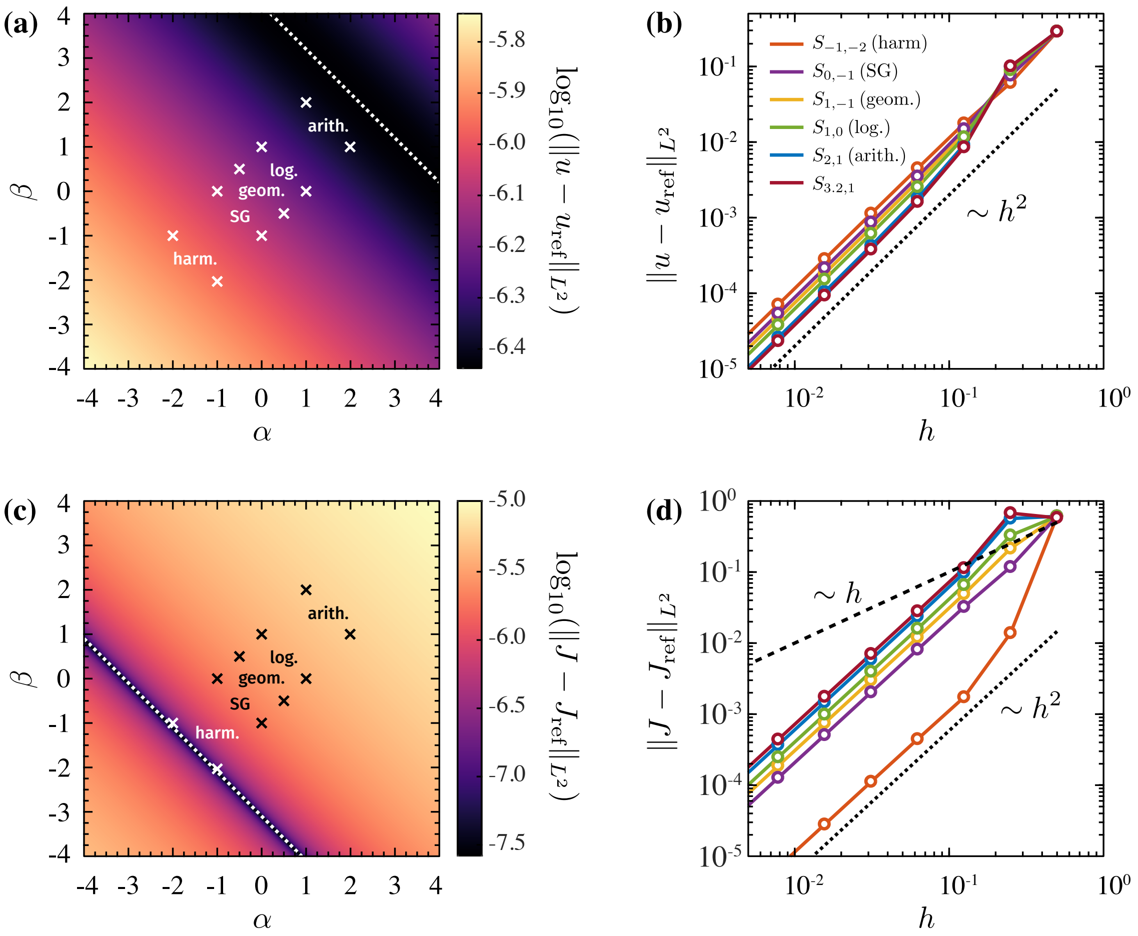

Fig. 1. Numerical results for Example 6.1. **(a)** Discretization error $\log_{10}(\|u - u_{\text{ref}}\|_{L_2})$ in the (α, β)-plane on an equidistant mesh with $2^{10} + 1$ nodes. The error is color-coded. Several special means (see Tab. 2) are highlighted by crosses. Notice the symmetry $S_{\alpha,\beta}(x, y) = S_{\beta,\alpha}(x, y)$. **(b)** Quadratic convergence of the discrete solution to the exact reference solution u_{exact} under mesh refinement in the L_2-norm. See the inset for a legend and color-coding of the considered means $S_{\alpha,\beta}$. In the present example, the best numerical result for u is achieved by $S_{3.2,1}$. **(c)** Logarithmic error of the numerically computed flux density $\log_{10}(\|J - J_{\text{ref}}\|_{L_2})$ in the (α, β)-plane on the same mesh as in (a). **(d)** Convergence of the numerically computed flux density to J_{ref}. In contrast to the convergence of u shown in (b), here the harmonic average $S_{-1,-2}$ yield the highest accuracy.

solution involving the imaginary error function (which is related to the Dawson function), that has been obtained using Wolfram Mathematica [WR17].

The numerical results are show in Figure 2. The discretization errors of both the density u and the flux J shown in Figure 2 (a) and (c) exhibit a sharp minimum on $\alpha + \beta = -1$. This includes the Scharfetter–Gummel mean $S_{0,-1}$, which converges fastest to the exact reference solutions for u and J, as shown in 2 (b) and (d). The SQRA scheme, with geometric mean $S_{\alpha,-\alpha}$, is found to be second best in the present example.

The numerical results are in line with our previous statements from Remark 5.6: In the case of strong gradients ∇V, the Scharfetter–Gummel scheme provides the most accurate flux discretization, in particular, the SG mean $S_{0,-1}$ is the only Stolarsky mean that recovers the upwind scheme (1.5). Away from that drift-dominated regime, the situation is less clear and other averages $S_{\alpha,\beta}$ can be superior, see for instance Example 6.1.

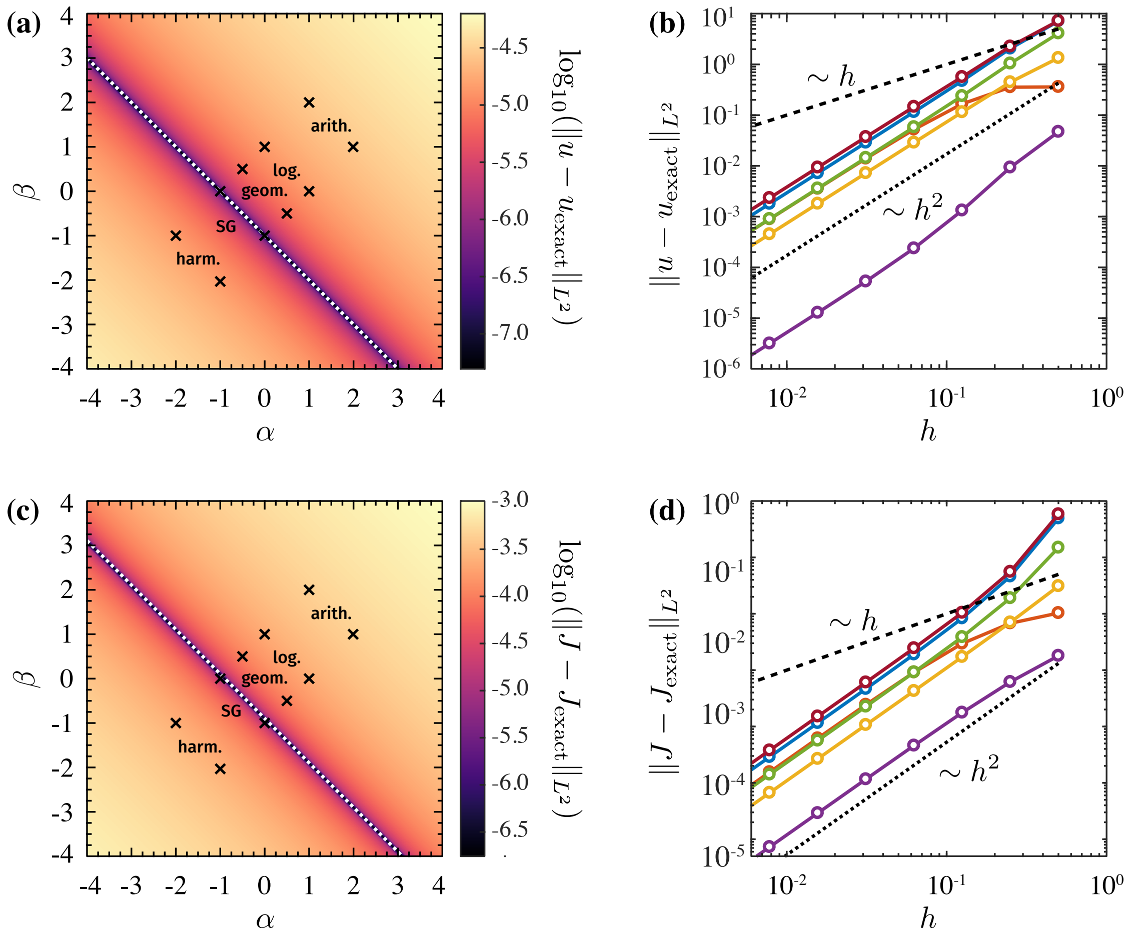

Fig. 2. Discretization errors and convergence behavior of the numerically computed u and J in Example 6.2 using the Stolarsky mean schemes. The errors in **(a)** and **(c)** are color-coded. The coloring of the means in **(b)** and **(d)** is the same as in Figure 1 (b). The plots clearly show a superior performance of the Scharfetter–Gummel scheme, which corresponds to the Stolarsky mean $S_{0,-1}$ for the approximation of both the density u and the flux J.

A Appendix

A.1 A General Poincaré Inequality

We derive a general Poincaré inequality on meshes. The idea behind the proof seems to go back to Hummel [Hum99] and has been adapted in a series of works e.g. [Hei18, HKP17]. Let $e_0 = 0$ and $(e_i)_{i=1,\dots,n}$ be the canonical basis of $\mathbb{R}^n$. Define:

$$D^{d-1} := \{\nu \in \mathbb{S}^{d-1} \mid \exists m \in \{1,\cdots,d\} : \nu \cdot e_i = 0 \;\forall\; i \in \{0,1,\cdots,m-1\} \text{ and } \nu \cdot e_m > 0\}\,.$$

Every $\nu \in \mathbb{S}^{d-1}$ satisfies $\nu \cdot e_i \neq 0$ for at least one e_i. Thus, for every $\nu \in \mathbb{S}^{d-1}$ it holds $\nu \in D^{d-1}$ if and only if $-\nu \notin D^{d-1}$.

We denote $\Gamma = \bigcup_{\sigma \in \mathcal{E}_\Omega} \sigma$ and say that $x \in \Gamma$ is a Lipschitz point if Γ is a Lipschitz graph in a neighborhood of x. The set of Lipschitz-Points is called $\Gamma_L \subset \Gamma$ and we note that for the $(d-1)$-dimensional Hausdorff-measure of $\Gamma \backslash \Gamma_L$ it holds $\mathcal{H}^{d-1}(\Gamma \backslash \Gamma_L) = 0$.

For $x \in \Gamma_L$, we denote $\nu_x \in D^{d-1}$ the normal vector to Γ in x.. Let

$$\mathcal{C}^1_0(\Omega;\Gamma) := \left\{u \in C(\Omega \backslash \Gamma) \;:\; u|_{\partial\Omega} \equiv 0\,,\; \forall i\, \exists v_i \in C^1\left(\overline{\Omega_i}\right) : u|_{\Omega_i} = v_i\right\}$$

and for $u \in \mathcal{C}^1_{K,0}(\Omega)$ define in Lipschitz points $x \in \Gamma_L$

$$u_\pm(x) := \lim_{h\to 0} \left(u\left(x \pm h\nu_x \right) \right) \,, \quad [\![u]\!](x) := u_+(x) - u_-(x) \,.$$

For two points $x, y \in \mathbb{R}^n$ denote (x, y) the closed straight line segment connecting x and y and for $\xi \in (x, y) \cap \Gamma_L$ denote

$$[\![u]\!]_{x,y}(\xi) := \lim_{h\to 0} \left(u\left(\xi + h(y - x) \right) - u\left(\xi - h(y - x) \right) \right)$$

the jump of the function u at ξ in direction $(y - x)$, i.e. $[\![u]\!]_{x,y}(\xi) \in \pm [\![u]\!]\,(\xi)$. We can extend $[\![u]\!]$ to Γ by $[\![u]\!]\,(x) = 0$ for $x \in \Gamma\backslash\Gamma_L$ and define

$$\|u\|_{H^1(\Omega;\Gamma)} := \left(\int_{\Omega\backslash\Gamma} |\nabla u|^2 + \int_\Gamma [\![u]\!]^2 \right)^{\frac{1}{2}} \,,$$
$$H^1_0\left(\Omega;\Gamma\right) := \overline{\mathcal{C}^1_0\left(\Omega;\Gamma\right)}^{\|\cdot\|_{H^1(\Omega;\Gamma)}} \,.$$

Then we find the following result:

Lemma A.1 (Semi-discrete Poincaré inequality)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. The space $H^1_0\left(\Omega;\Gamma\right)$ is linear and closed for every $s \in [0, \frac{1}{2})$ and there exists a positive constant $C_s > 0$ such that the following holds: Suppose there exists a constant $C_\# > 0$ such that for almost all $(x, y) \in \Omega^2$ it holds $\#\left((x, y) \cap \Gamma\right) \le C_\#$.. Then for every $u \in H^1_0\left(\Omega;\Gamma\right)$ it holds*

$$\|u\|^2_{H^s(\Omega)} \le C_s \left(C_\# \int_\Gamma [\![u]\!]^2 + \|\nabla u\|^2_{L^2(\Omega\backslash\Gamma)} \right) . \tag{A.1}$$

Furthermore, for every $u \in H^1\left(\Omega;\Gamma\right)$ and every $\boldsymbol{\eta} \in \mathbb{R}^d$ it holds

$$\int_\Omega |u(x) - u(x + \boldsymbol{\eta})|^2 \,dx \le |\boldsymbol{\eta}| \left(C_\# \int_\Gamma [\![u]\!]^2 + \|\nabla u\|^2_{L^2(\Omega\backslash\Gamma)} \right) . \tag{A.2}$$

Proof. In what follows, given $u \in \mathcal{C}^1_0(\Omega;\Gamma)$, we write $\widehat{\nabla u}(x) := \nabla u(x)$ if $x \in \Omega\backslash\Gamma$ and $\widehat{\nabla u}(x) = 0$ else. For $y \in \mathbb{R}^d$ we denote $(x, y) = \{x + s\,(y - x) \,:\, s \in [0, 1]\}$. Using $2ab < a^2 + b^2$, we infer for $u \in \mathcal{C}^1_0(\Omega;\Gamma)$ and $x, y \in \overline{\Omega}\backslash\Gamma$ such that $(x, y) \cap \Gamma$ is finite the inequality

$$\begin{aligned} |u(x) - u(y)|^2 &\le \left(\sum_{\xi\in(x,y)\cap\Gamma} [\![u]\!]_{x,y}(\xi) + \int_0^1 \widehat{\nabla u}\left(x + s(y - x)\right) \cdot (x - y)\, ds \right)^2 \\ &< |x - y|^2 \int_0^1 \left|\widehat{\nabla u}\left(x + s(y - x)\right)\right|^2 ds + \left(\sum_{\xi\in(x,y)\cap\Gamma} [\![u]\!]_{x,y}(\xi) \right)^2 \end{aligned}$$

Since $[\![u]\!]_{x,y} = [\![u]\!]$ we compute

$$\left(\sum_{\xi\in(x,y)\cap\Gamma} [\![u]\!]_{x,y}(\xi) \right)^2 \le \#\left((x, y) \cap \Gamma\right) \sum_{\xi\in(x,y)\cap\Gamma} [\![u]\!]^2(\xi)$$

and obtain

$$\begin{aligned} |u(x) - u(y)|^2 < |x - y|^2 \int_0^1 \left|\widehat{\nabla u}\left(x + s(y - x)\right)\right|^2 ds \\ + \#\left((x, y) \cap \Gamma\right) \sum_{\xi\in(x,y)\cap\Gamma} [\![u]\!]^2(\xi) \,. \end{aligned} \tag{A.3}$$

We fix $\eta > 0$ and consider the orthonormal basis $(e_i)_{i=1,\dots,d}$ of $\mathbb{R}^d$. The determinant of the first fundamental form of Γ is bigger than 1 almost everywhere. Hence we can observe that

$$\begin{aligned}\int_\Omega \sum_{\xi\in(x,x+\eta e_1)\cap\Gamma} [\![u]\!]^2(\xi)\,\mathrm{d}x &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}^{d-1}} \sum_{\xi\in(x,x+\eta e_1)\cap\Gamma} [\![u]\!]^2(\xi)\,\mathrm{d}x_2\dots\mathrm{d}x_d\right)\mathrm{d}x_1\\ &\le \int_{\mathbb{R}}\int_{\Gamma\cap((x_1,x_1+\eta)\times\mathbb{R}^{d-1})} [\![u]\!]^2(x)\,\mathrm{d}\sigma\,\mathrm{d}x_1\\ &\le \eta\int_\Gamma [\![u]\!]^2(x)\,\mathrm{d}x\,,\end{aligned}$$

where we used that the surface elements are bigger than 1. Furthermore, we have

$$\eta^2\int_0^1 \left|\widehat{\nabla u}\,(x+s\eta e_1)\right|^2 ds = \eta\int_0^\eta \left|\widehat{\nabla u}\,(x+se_1)\right|^2 ds\,.$$

Replacing e_1 in the above calculations with any unit vector e, we obtain from integration of (A.3) with $y = x+\boldsymbol{\eta}$, $\boldsymbol{\eta} = \eta e$, over Ω that

$$\int_\Omega |u(x)-u(x+\boldsymbol{\eta})|^2\,dx \le |\boldsymbol{\eta}|\left(C_\#\int_\Gamma [\![u]\!]^2 + \|\nabla u\|^2_{L^2(\Omega\setminus\Gamma)}\right).$$

Dividing by $|\boldsymbol{\eta}|$ and integrating over $\boldsymbol{\eta}\in\mathbb{R}^d$, we obtain that for every $s\in[0,\frac12)$ there exists a positive constant $C_s > 0$ independent from u and K such that

$$\|u\|^2_{H^s(\Omega)} \le C_s\left(C_\#\int_\Gamma [\![u]\!]^2 + \|\nabla u\|^2_{L^2(\Omega\setminus\Gamma)}\right). \tag{A.4}$$

Hence, by approximation, the last two estimates hold for all $u\in H^1_0(\Omega;\Gamma)$.. □

A.2 Physical relevance of the geometric mean

Theorem A.2. *Let $S_{ij} = S_*(\pi_i,\pi_j)$ be a Stolarsky mean and let ψ^* be a symmetric strictly convex function with $\psi^*(0) = 0$. If $\partial_\pi(S_{ij}a_{ij}) = 0$ then $S_{ij} = \sqrt{\pi_i\pi_j}$ and ψ^* is proportional to C^*.*

Proof of Theorem A.2. The case $S_{ij} = \sqrt{\pi_i\pi_j}$ and $\psi^*(\xi) = \cosh\xi - 1$ was explained in detail in [Hei18].

In the general case, symmetry of ψ^* in $\xi_i-\xi_j$ implies $\psi^*(\xi_i-\xi_j) = \psi^*(|\xi_i-\xi_j|)$. We make use of the fact that the original $\mathsf{C}^*(\xi) = \cosh\xi - 1$ is a bijection on $[0,\infty)$ and suppose that hence $\psi^*(\xi_i-\xi_j) = \theta(\mathsf{C}^*(\xi_i-\xi_j))$. This implies particularly that

$$0 \le x\,\partial_x(\theta(\mathsf{C}^*(x))) = x\,\partial_\xi\theta(\mathsf{C}^*(x))\,\partial_x\mathsf{C}^*(x)\,.$$

Furhtermore, the symmetry of ψ^* implies by the last inequality that $\partial_\xi\theta(\mathsf{C}^*(x)) > 0$. Inserting this information in (3.13) and (3.14) we observe that

$$S_{ij}\left(\frac{u_i}{\pi_i}-\frac{u_j}{\pi_j}\right)\partial_\xi\theta\left(\mathsf{C}^*\left(\ln\left(\frac{u_i}{\pi_i}\right)-\ln\left(\frac{u_j}{\pi_j}\right)\right)\right)^{-1}\sinh\left(\ln\left(\frac{u_i}{\pi_i}\right)-\ln\left(\frac{u_j}{\pi_j}\right)\right)^{-1}$$

has to be independent from π_i and π_j. From the above case $S_{ij} = \sqrt{\pi_i\pi_j}$, we know that

$$\sqrt{\pi_i\pi_j}\left(\frac{u_i}{\pi_i}-\frac{u_j}{\pi_j}\right)\sinh\left(\ln\left(\frac{u_i}{\pi_i}\right)-\ln\left(\frac{u_j}{\pi_j}\right)\right)^{-1}$$

is constant in π_i and π_j. Hence it remains to show that

$$f(\pi_i,\pi_j) := S_{ij}\sqrt{\pi_i\pi_j}^{-1}\partial_\xi\psi\left(\frac{u_i}{u_j}\frac{\pi_j}{\pi_i}+\frac{u_j}{u_i}\frac{\pi_i}{\pi_j}\right)^{-1}$$

is independent from π_i and π_j if and only if $\partial_\xi\psi = \mathrm{const}$ and $S_{ij} = \sqrt{\pi_i\pi_j}$.

Assume first that $S_{ij}\sqrt{\pi_i\pi_j}^{-1} = \mathrm{const}$. Then for $p = \frac{\pi_i}{\pi_j}$ we obtain that

$$\partial_p\left(\partial_\xi\theta\left(\frac{u_i}{u_j}p^{-1}+\frac{u_j}{u_i}p\right)^{-1}\right) = 0$$

has to hold. This implies that $\partial_\xi\psi = \mathrm{const}$.

If $S_{ij}\sqrt{\pi_i\pi_j}^{-1} \neq \mathrm{const}$, we use the definition of the weighted Stolarsky means given in (3.4) and note that

$$S_{ij} := S(\pi_i,\pi_j) = \left(\frac{\beta(\pi_i^\alpha-\pi_j^\alpha)}{\alpha(\pi_i^\beta-\pi_j^\beta)}\right)^{\frac{1}{\alpha-\beta}} = \pi_j\left(\frac{\beta(p^\alpha-1)}{\alpha(p^\beta-1)}\right)^{\frac{1}{\alpha-\beta}},$$

where again $p = \frac{\pi_i}{\pi_j}$. Hence we obtain that

$$\begin{aligned} f(\pi_i,\pi_j) = \tilde f(p) &:= \sqrt{\frac{1}{p}}\left(\frac{\beta(p^\alpha-1)}{\alpha(p^\beta-1)}\right)^{\frac{1}{\alpha-\beta}}\partial_\xi\theta\left(\frac{u_i}{u_j}p^{-1}+\frac{u_j}{u_i}p\right)^{-1} \\ &= \left(\frac{\beta\left(p^{\frac{\alpha}{2}}-p^{-\frac{\alpha}{2}}\right)}{\alpha\left(p^{\frac{\beta}{2}}-p^{-\frac{\beta}{2}}\right)}\right)^{\frac{1}{\alpha-\beta}}\partial_\xi\theta\left(\frac{u_i}{u_j}p^{-1}+\frac{u_j}{u_i}p\right)^{-1}\end{aligned}$$

has to be independent of π_i and π_j. But then, $\tilde f$ is independent of p. Now, we define $a = \frac{u_j}{u_i}$ and observe that

$$\tilde f\left(\frac{1}{a^2p}\right) = \left(\frac{\beta\left((a^2p)^{-\frac{\alpha}{2}}-(a^2p)^{\frac{\alpha}{2}}\right)}{\alpha\left((a^2p)^{-\frac{\beta}{2}}-(a^2p)^{\frac{\beta}{2}}\right)}\right)^{\frac{1}{\alpha-\beta}}\partial_\xi\theta\left(\frac{u_i}{u_j}p^{-1}+\frac{u_j}{u_i}p\right)^{-1}.$$

We assume for $\alpha \neq \beta$. The case $\alpha = \beta$ can follows by continuity. For any p it should holds $\tilde f\left(\frac{1}{a^2p}\right) = \tilde f(p)$, which implies

$$\left(\frac{\beta\left(p^{\frac{\alpha}{2}}-p^{-\frac{\alpha}{2}}\right)}{\alpha\left(p^{\frac{\beta}{2}}-p^{-\frac{\beta}{2}}\right)}\right)^{\frac{1}{\alpha-\beta}} = \left(\frac{\beta\left((a^2p)^{-\frac{\alpha}{2}}-(a^2p)^{\frac{\alpha}{2}}\right)}{\alpha\left((a^2p)^{-\frac{\beta}{2}}-(a^2p)^{\frac{\beta}{2}}\right)}\right)^{\frac{1}{\alpha-\beta}},$$

or equivalently, after introducing $q^2 = p$,

$$\left(a^\alpha-a^\beta\right)q^{\alpha+\beta}+\left(a^\beta-a^{-\alpha}\right)q^{\beta-\alpha}+\left(a^{-\beta}-a^\alpha\right)q^{\alpha-\beta}+\left(a^{-\alpha}-a^{-\beta}\right)q^{-\beta-\alpha} = 0.$$

Since $\alpha \neq \beta$, one of the terms $q^{\pm\alpha\pm\beta}$ grows faster than the other. Hence we conclude that $a^\alpha = a^{\pm\beta}$ which means, $a = 1$, a contradiction. □

A.3 Properties of the Stolarsky mean

Lemma A.3. *For every of the above Stolarsky means $S_*(x,y)$ it holds*

$$\partial_x S_*(x,x) = \partial_y S_*\left(x,x\right) = \frac{1}{2} \text{ and } \partial_x^2 S_*\left(x,x\right) = \partial_y^2 S_*\left(x,x\right) = -\partial_{xy}^2 S_*\left(x,x\right) = -\partial_{yx}^2 S_*\left(x,x\right)\,.$$

Proof. Since $S_*(x,x) = x$ and S_* is symmetric in x and y, we find from differentiating $\partial_x S_* = \partial_y S_* = \frac{1}{2}$. From the last equality, we find $\partial_x S_*(x,x) - \partial_y S_*(x,x) = 0$ as well as $\partial_x S_*(x,x) + \partial_y S_*(x,x) = 1$ and differentiation yields

$$\partial_x^2 S_*\left(x,x\right) - \partial_y^2 S_*\left(x,x\right) - \partial_{xy}^2 S_*\left(x,x\right) + \partial_{yx}^2 S_*\left(x,x\right) = 0\,, \tag{A.5}$$

$$\partial_x^2 S_*\left(x,x\right) + \partial_y^2 S_*\left(x,x\right) + \partial_{xy}^2 S_*\left(x,x\right) + \partial_{yx}^2 S_*\left(x,x\right) = 0\,. \tag{A.6}$$

Since $-\partial_{xy}^2 S_*\left(x,x\right) + \partial_{yx}^2 S_*\left(x,x\right) = 0$, equation (A.5) yields $\partial_x^2 S_*\left(x,x\right) = \partial_y^2 S_*\left(x,x\right)$. Inserting the last two relations into (A.6) yields $\partial_{xy}^2 S_*\left(x,x\right) = \partial_{yx}^2 S_*\left(x,x\right) = -\partial_x^2 S_*\left(x,x\right)$. □

Lemma A.4. *It holds* (3.5)$\partial_x^2 S_{\alpha,\beta}\left(\pi,\pi\right) = \frac{1}{12\pi}\left(\alpha+\beta-3\right)$.

Proof. We know from Lemma A.3 that $\partial_x S_{\alpha,\beta}\left(x,x\right) = \frac{1}{2}$ and $\partial_x^2 S_{\alpha,\beta}\left(x,x\right) = -\partial_y \partial_x S_{\alpha,\beta}\left(x,x\right)$. Hence we find

$$\partial_x S_{\alpha,\beta}\left(x+h,x-h\right) - \frac{1}{2} = \begin{pmatrix} h \\ -h \end{pmatrix}\begin{pmatrix} \partial_x^2 S_{\alpha,\beta}\left(x,x\right) \\ \partial_y \partial_x S_{\alpha,\beta}\left(x,x\right) \end{pmatrix} = 2h\partial_x^2 S_{\alpha,\beta}\left(x,x\right)\,.$$

We make use of the explicit form

$$\partial_x S_{\alpha,\beta}\left(x,y\right) = \left(\frac{\beta}{\alpha}\right)^{\frac{1}{\alpha-\beta}} \frac{\left(x^\alpha - y^\alpha\right)^{\frac{1}{\alpha-\beta}-1}}{\left(x^\beta - y^\beta\right)^{\frac{1}{\alpha-\beta}-1}} \frac{\alpha\left(x^\beta - y^\beta\right)x^\alpha - \beta\left(x^\alpha - y^\alpha\right)x^\beta}{\left(\alpha-\beta\right)\ x\ \left(x^\beta - y^\beta\right)^2}$$

for $x \neq y$. We insert $x = x+h$ and $y = x-h$ and make use of the following expansions

$$\begin{aligned}
\left(\left(x+h\right)^\alpha - \left(x-h\right)^\alpha\right)^c &= \left(\alpha h x^{\alpha-1}\right)^c\left(2^c + O\left(h^2\right)\right) \\
\beta\left(\left(x+h\right)^\alpha - \left(x-h\right)^\alpha\right)\left(x+h\right)^\beta &= 2\alpha\beta h x^{\alpha+\beta-1} + 2\alpha\beta^2 h^2 x^{\alpha+\beta-2} \\
&\quad + \frac{1}{3}\alpha\beta h^3\left(\alpha^2 - 3\alpha + 3\beta^2 - 3\beta + 2\right) + O\left(h^4\right) \\
\alpha\left(\left(x+h\right)^\beta - \left(x-h\right)^\beta\right)\left(x+h\right)^\alpha &= 2\alpha\beta h x^{\alpha+\beta-1} + 2\alpha^2\beta h^2 x^{\alpha+\beta-2} \\
&\quad + \frac{1}{3}\alpha\beta h^3\left(\beta^2 - 3\beta + 3\alpha^2 - 3\alpha + 2\right) + O\left(h^4\right) \\
\left(x+h\right)\ \left(\left(x+h\right)^\beta - \left(x-h\right)^\beta\right)^2 &= 4\beta^2 h^2 x^{2\beta-1} + 4\beta^2 h^3 x^{2\beta-2} + O\left(h^4\right)
\end{aligned}$$

$$\begin{aligned}
&\alpha\left(\left(x+h\right)^\beta - \left(x-h\right)^\beta\right)\left(x+h\right)^\alpha - \beta\left(\left(x+h\right)^\alpha - \left(x-h\right)^\alpha\right)\left(x+h\right)^\beta \\
&\qquad = 2\alpha\beta\left(\alpha-\beta\right)h^2 x^{\alpha+\beta-2} + \frac{\alpha\beta}{3}h^3 x^{\alpha+\beta-3}\left(2\alpha^2 - 2\beta^2\right) + O\left(h^4\right)
\end{aligned}$$

to obtain

$$\frac{\beta\left(x^{\alpha}-y^{\alpha}\right)x^{\beta}-\alpha\left(x^{\beta}-y^{\beta}\right)x^{\alpha}}{\left(\alpha-\beta\right)\ x\ \left(x^{\beta}-y^{\beta}\right)^{2}}=\frac{\alpha\left(x^{\alpha+\beta-2}+h\frac{1}{3}x^{\alpha+\beta-3}\left(\alpha+\beta\right)+O\left(h^{2}\right)\right)}{2\beta\left(x^{2\beta-1}+hx^{2\beta-2}+O\left(h^{2}\right)\right)}$$

and

$$\frac{\left(x^{\alpha}-y^{\alpha}\right)^{\frac{1}{\alpha-\beta}-1}}{\left(x^{\beta}-y^{\beta}\right)^{\frac{1}{\alpha-\beta}-1}}\approx\left(\frac{\alpha}{\beta}\right)^{\frac{1}{\alpha-\beta}-1}\left(\frac{x^{\alpha-1}\left(1+O\left(h^{2}\right)\right)}{x^{\beta-1}\left(1+O\left(h^{2}\right)\right)}\right)^{\frac{1}{\alpha-\beta}-1}.$$

Together with

$$\begin{aligned}\frac{a+bh}{c+dh}&=\frac{a}{c}+\frac{bc-ad}{c^{2}}h+O\left(h^{2}\right)\\ \left(\frac{1+ah^{2}}{1+bh^{2}}\right)^{c}&=1+ch^{2}(a-b)+O\left(h^{4}\right)\end{aligned}$$

we find

$$\begin{aligned}\partial_{x}S_{\alpha,\beta}\left(x+h,x-h\right)&=\left(\frac{\left(1+O\left(h^{2}\right)\right)}{\left(1+O\left(h^{2}\right)\right)}\right)^{\frac{1}{\alpha-\beta}-1}\left(\frac{\left(1+h\frac{1}{3}x^{-1}\left(\alpha+\beta\right)+O\left(h^{2}\right)\right)}{2\left(1+hx^{-1}+O\left(h^{2}\right)\right)}\right)\\&=\left(\frac{1}{2}+\frac{\frac{2}{3}\left(\alpha+\beta\right)-2}{4x}h\right)+O\left(h^{2}\right)\end{aligned}$$

and hence (3.5). □

A.4 Approximation of potentials to get the SQRA mean

The aim of this section is to provide a class of potentials which are easy to handle and which generate the SQRA-mean $S_{-1,1}(\pi_0,\pi_h)$ by $\pi_{\text{mean}}=\left(\frac{1}{h}\int_0^h\pi^{-1}\right)^{-1}$. Clearly, choosing the constant potential $V(x):=V_c:=-\log S_{-1,1}(\pi_0,\pi_h)$ we obtain right mean. Although this works for any means, this has two drawbacks

1 The potential jumps and hence the gradient is somewhere infinite, which means that at these points the force on the particles is infinitely high which is not physical.

2 Approximating a general function by piecewise constants, on each interval the accuracy is only of order h. However, approximating a function by affine interpolation the accuracy is of order h^2 on each interval (see below for the calculation).

So we want to get a potential which may be used as a good approximation (i.e. approximating of order h^2), is physical (i.e. continuous) and generates the SQRA-mean. Note, that most considerations below also work for other Stolarsky means. For simplicity we focus on the SQRA mean $S_{-1,1}$.

A.4.1 Approximation order for linear approximation

Let us first realize that a linear interpolation provides an approximation of order h^2. Let $V:[0,h]\to\mathbb{R}$ be a general C^2-potential. We define with $V(0)=V_0$ and $V(h)=V_h$

$$\tilde{V}(x)=V_0+\frac{V_h-V_0}{h}x.$$

Then one easily checks that

$$V(x)=V_0+\partial_x V(0)x+\frac{1}{2}\partial_x^2 V(0)x^2+O(h^3)$$

and hence,

$$V(x)-\tilde{V}(x)=\left(\partial_x V(0)-\frac{V_h-V_0}{h}\right)x+\frac{1}{2}\partial_x^2 V(0)x^2+O(h^3).$$

Clearly, we also have

$$V_h=V_0+\partial_x V(0)h+\frac{1}{2}\partial_x^2 V(0)h^2+O(h^3)$$

which yields

$$V(x)-\tilde{V}(x)=-\frac{1}{2}\partial_x^2 V(0)hx+\frac{1}{2}\partial_x^2 V(0)x^2+O(h^3)=\frac{1}{2}\partial_x^2 V(0)(x-h)x+O(h^3)=O(h^2).$$

A.4.2 Definition of potentials $\hat{V}$ which generate the SQRA mean

We consider a piecewise linear potential of the form

$$\hat{V}(x)=\begin{cases}\frac{V_c-V_0}{x_1}x+V_0 & ,x\in[0,x_1]\\ V_c & ,x\in[x_1,x_2]\\ \frac{V_h-V_c}{h-x_2}(x-x_2)+V_c & ,x\in[x_2,h]\end{cases}.$$

where $x_1,x_2\in[0,h]$ are firstly arbitrary and $V_c=-\log S_{-1,1}(\pi_0,\pi_h)=\frac{1}{2}(V_h+V_0)$. The potential is clearly continuous. Then

$$\frac{1}{h}\int_0^h \mathrm{e}^{\hat{V}(x)}\mathrm{d}x=\frac{x_1}{h}\frac{\mathrm{e}^{V_c}-\mathrm{e}^{V_0}}{V_c-V_0}+\frac{x_2-x_1}{h}\mathrm{e}^{V_c}+\frac{h-x_2}{h}\frac{\mathrm{e}^{V_h}-\mathrm{e}^{V_c}}{V_h-V_c}.$$

Introducing the ratios $\alpha=\frac{x_1}{h}$ and $\beta=\frac{h-x_2}{h}$ (which are in $[0,1/2]$) , we want to solve $\frac{1}{h}\int_0^h \mathrm{e}^{\hat{V}(x)}\mathrm{d}x=\mathrm{e}^{\frac{1}{2}(V_h+V_0)}$. Indeed, introducing the difference of the potentials $\bar{V}=V_h-V_0$, we obtain

$$\lambda=\frac{\alpha}{\beta}=\frac{\mathrm{e}^{\bar{V}/2}-\bar{V}/2-1}{\mathrm{e}^{-\bar{V}/2}+\bar{V}/2-1}\approx 1+\frac{1}{3}\bar{V}+\frac{1}{18}\bar{V}^2.$$

Hence, any value α,β satisfying this ratio generates a potential with the SQRA-mean.

A.4.3 Proof that the potential approximates an arbitrary potential of order h^2

Since the linear potentials approximates a general potential of order h^2 it suffices to approximate the linear potential $\tilde{V}$ by $\hat{V}$. We show that there are α,β satisfying $\frac{\alpha}{\beta}=\lambda$, such that $\|\hat{V}-\tilde{V}\|_{C([x_i,x_{i+1}])}=$

$O(h^2)$. The difference of $\hat{V}$ and $\tilde{V}$ is the largest at $x = x_1$ or $x = x_2$. We estimate both differences. We have

$$\tilde{V}(x_1) = V_0 + \frac{V_h - V_0}{h} x_1 = V_0 + \alpha \bar{V}, \;\; \tilde{V}(x_2) = V_0 + \frac{V_h - V_0}{h} x_2 = V_0 + (1-\beta)\bar{V}.$$

Hence we have to estimate

$$\Delta_1 := |V_0 - V_c + \alpha \bar{V}|, \;\; \Delta_2 := |V_0 - V_c + (1-\beta)\bar{V}|.$$

In the case of SQRA, one possible choice for α, β is given by $\alpha + \beta = 1$. Then $\Delta_1 = \Delta_2 = |V_0 - V_c + \alpha \bar{V}| = |V_0 - V_c + \frac{\lambda}{1+\lambda}\bar{V}| = \frac{1}{1+\lambda}|(1+\lambda)(V_0 - V_c) + \lambda \bar{V}|$. We have $V_0 - V_c = -\bar{V}/2$, and hence

$$\Delta_1 = \Delta_2 = \frac{1}{1+\lambda}\frac{\bar{V}}{2}|\lambda - 1|.$$

One can check that $\lambda \approx 1 + \bar{V}/3$ and hence, $\Delta_1 + \Delta_2 \approx \frac{V^2}{6} \approx O(h^2)$.

References

[AS55] D. N. de G. Allan and R. V. Southwell. Relaxation methods applied to determine the motion in two dimensions of a viscous fluid past a fixed cylinder. *Q. J. Mech. Appl. Math.*, 8(2):129–145, 1955.

[BMP89] Franco Brezzi, Luisa Donatella Marini, and Paola Pietra. Numerical simulation of semiconductor devices. *Comput. Methods Appl. Mech. Eng.*, 75(1-3):493–514, 1989.

[Bre71] Richard P. Brent. An algorithm with guaranteed convergence for finding a zero of a function. *Comput. J.*, 14(4):422–425, 1971.

[CHLZ12] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. 203(3):969–1008, 2012.

[DFM18] Patrick Dondl, Thomas Frenzel, and Alexander Mielke. A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM Control Optim. Calc. Var.*, 2018. To appear. WIAS preprint 2459.

[DHWK] L. Donati, M. Heida, M. Weber, and B. Keller. Estimation of the initesimal generator by square-root approximation. *In preparation.*

[DJSD15] Purushottam D Dixit, Abhinav Jain, Gerhard Stock, and Ken A Dill. Inferring transition rates of networks from populations in continuous-time markov processes. *Journal of chemical theory and computation*, 11(11):5464–5472, 2015.

[DL15] Karoline Disser and Matthias Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Networks Heterg. Media*, 10(2):233–253, 2015.

[DPD18] Daniele A Di Pietro and Jérôme Droniou. A third strang lemma and an aubin–nitsche trick for schemes in fully discrete formulation. *Calcolo*, 55(3):40, 2018.

[EFG06] R. Eymard, J. Fuhrmann, and K. Gärtner. A finite volume scheme for nonlinear parabolic equations derived from one-dimensional local dirichlet problems. *Numer. Math.*, 102(3):463–495, 2006.

[EM12] Matthias Erbar and Jan Maas. Ricci curvature of finite Markov chains via convexity of the entropy. 206(3):997–1038, 2012.

[Eva98] L.C. Evans. *Partial Differential Equations*. AMS, 1998.

[FKF17] Patricio Farrell, Thomas Koprucki, and Jürgen Fuhrmann. Computational and analytical comparison of flux discretizations for the semiconductor device equations beyond Boltzmann statistics. *Journal of Computational Physics*, 346:497–513, 2017.

[FKN$^+$19] Konstantin Fackeldey, Péter Koltai, Peter Névir, Henning Rust, Axel Schild, and Marcus Weber. From metastable to coherent sets – time-discretization schemes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(1):012101, 2019.

[FL19] Thomas Frenzel and Matthias Liero. Effective diffusion in thin structures via generalized gradient systems and EDP-convergence. *WIAS Preprint 2601*, 2019.

[FRD$^+$17] Patricio Farrell, Nella Rotundo, Duy Hai Doan, Markus Kantner, Jürgen Fuhrmann, and Thomas Koprucki. Drift-Diffusion Models. In Joachim Piprek, editor, *Handbook of Optoelectronic Device Modeling and Simulation: Lasers, Modulators, Photodetectors, Solar Cells, and Numerical Methods*, volume 2, chapter 50, pages 731–771. CRC Press, Taylor & Francis Group, Boca Raton, 2017.

[GHV00] Thierry Gallouët, Raphaele Herbin, and Marie Hélene Vignal. Error estimates on the approximate finite volume solution of convection diffusion equations with general boundary conditions. *SIAM Journal on Numerical Analysis*, 37(6):1935–1972, 2000.

[GKMP19] Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Homogenisation of one-dimensional discrete optimal transport. *arXiv:1905.05757*, 2019.

[Hei18] Martin Heida. Convergences of the squareroot approximation scheme to the Fokker–Planck operator. *Mathematical Models and Methods in Applied Sciences*, 28(13):2599–2635, 2018.

[HKP17] Martin Heida, Ralf Kornhuber, and Joscha Podlesny. Fractal homogenization of multiscale interface problems. *arXiv preprint arXiv:1712.01172*, 2017.

[HMR11] M. Heida, J. Màlek, and K.R. Rajagopal. On the development and generalizations of Allen-Cahn and Stefan equations within a thermodynmic framework. *to be submitted to Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, 2011.

[Hum99] H.K. Hummel. *Homogenization of Periodic and Random Multidimensional Microstructures*. PhD thesis, Technische Universität Bergakademie Freiberg, 1999.

[Il'69] A. M. Il'in. Differencing scheme for a differential equation with a small parameter affecting the highest derivative. *Mathematical notes of the Academy of Sciences of the USSR*, 6(2):237–248, 1969. Translated from Mat. Zametki, Vol. 6, No. 2, pp. 237–248 (1969).

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[Kan20] Markus Kantner. Generalized Scharfetter–Gummel schemes for electro-thermal transport in degenerate semiconductors using the Kelvin formula for the Seebeck coefficient. *Journal of Computational Physics*, 402:109091, 2020.

[LFW13] Han Cheng Lie, Konstantin Fackeldey, and Marcus Weber. A square root approximation of transition rates for a markov state model. *SIAM Journal on Matrix Analysis and Applications*, 34:738–756, 2013.

[LMPR17] Matthias Liero, Alexander Mielke, Mark A. Peletier, and D. R. Michiel Renger. On microscopic origins of generalized gradient structures. *Discr. Cont. Dynam. Systems Ser. S*, 10(1):1–35, 2017.

[Maa11] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261:2250–2292, 2011.

[Mar15] René Marcelin. Contribution a l'étude de la cinétique physico-chimique. *Annales de Physique*, III:120–231, 1915.

[Mar86] P. A. Markowich. *The stationary Semiconductor device equations*. Springer, Vienna, 1986.

[Mie11] Alexander Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24:1329–1346, 2011.

[Mie13a] Alexander Mielke. Geodesic convexity of the relative entropy in reversible markov chains. *Calculus of Variations and Partial Differential Equations*, 48(1):1–31, 2013.

[Mie13b] Alexander Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Part. Diff. Eqns.*, 48(1):1–31, 2013.

[Mie16] Alexander Mielke. On evolutionary Γ-convergence for gradient systems (Ch. 3). In A. Muntean, J. Rademacher, and A. Zagaris, editors, *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, Lecture Notes in Applied Math. Mechanics Vol. 3, pages 187–249. Springer, 2016. Proc. of Summer School in Twente University, June 2012.

[MPPR17] Alexander Mielke, Robert I. A. Patterson, Mark A. Peletier, and D. R. Michiel Renger. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. *SIAM J. Appl. Math.*, 77(4):1562–1585, 2017.

[MPR14] Alexander Mielke, Mark A. Peletier, and D. R. Michiel Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

[MS19] Alexander Mielke and Artur Stephan. Coarse-graining via edp-convergence for linear fast-slow reaction systems. *WIAS preprint 2643*, 2019.

[MW94] J J H Miller and Song Wang. An analysis of the Scharfetter–Gummel box method for the stationary semiconductor device equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 28(2):123–140, 1994.

[SG69] D.L. Scharfetter and H.K. Gummel. Large-signal analysis of a silicon read diode oscillator. *IEEE Trans. Electron Devices*, 16(1):64–77, 1969.

[Sto75] Kenneth B. Stolarsky. Generalizations of the logarithmic mean. *Mathematics Magazine*, 48(2):87–92, 1975.

[vR50] W. W. van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell Syst. Tech. J.*, 29(4):560–607, Oct 1950.

[WE17] Marcus Weber and Natalia Ernst. A fuzzy-set theoretical framework for computing exit rates of rare events in potential-driven diffusion processes. *arXiv preprint arXiv:1708.00679*, 2017.

[WR17] Inc. Wolfram Research. Mathematica, 2017.