Mathematical Models and Methods in Applied Sciences
Vol. 17, No. 8 (2007) 1159–1182

CONVERGENCE OF A *hp*-STREAMLINE DIFFUSION SCHEME FOR VLASOV–FOKKER–PLANCK SYSTEM

M. ASADZADEH

Department of Mathematics,
Chalmers University of Technology and the Göteborg University,
SE-412 96, Göteborg, Sweden
mohammad@math.chalmers.se

A. SOPASAKIS

Department of Mathematics,
University of Massachusetts, Amherst, MA 01003-9305, USA
sopas@math.umass.edu

Received 26 September 2006
Communicated by C. Johnson

We analyze the *hp*-version of the streamline diffusion finite element method for the Vlasov–Fokker–Planck system. For this method we prove the stability estimates and derive sharp *a priori* error bounds in a stabilization parameter $\delta \sim \min(h/p, h^2/\sigma)$, with h denoting the mesh size of the finite element discretization in phase-space-time, p the spectral order of approximation, and σ the transport cross-section.

Keywords: Vlasov–Poisson–Fokker–Planck; *hp*-version; streamline diffusion; stability; convergence.

AMS Subject Classification: 65M60, 82C80

1. Introduction

We study stability and convergence for the *hp*-version streamline diffusion (SD) finite element method for a deterministic Vlasov–Fokker–Planck (VFP) system. During this work we apply some of the *hp*-techniques introduced in Refs. 11–13. The objective is to derive sharp *a priori* *hp*-error bounds for a SD scheme in a L_2-based norm.

The Vlasov-Poisson-Fokker-Planck (VPFP) system arising in the kinetic description of a plasma of Coulomb particles under the influence of a self-consistent internal field and an external force can be formulated as follows: given the initial distribution of particles $f_0(x, v) \geq 0$, in the phase-space variable $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$,

$d = 1, 2, 3$, and the physical parameters $\beta \geq 0$ and $\sigma \geq 0$, find the distribution function $f(x, v, t)$ for $t > 0$, satisfying the nonlinear system of evolution equations

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + \mathrm{div}_v[(E - \beta v)f] = \sigma \Delta_v f, & \text{in } \mathbb{R}^{2d} \times (0, \infty), \\ f(x, v, 0) = f_0(x, v), & \text{for } (x, v) \in \mathbb{R}^{2d}, \\ E(x, t) = \dfrac{\theta}{|\mathcal{S}|^{d-1}} \dfrac{x}{|x|^d} *_x \rho(x, t), & \text{for } (x, t) \in \mathbb{R}^d \times (0, \infty), \\ \rho(x, t) = \displaystyle\int_{\mathbb{R}^d} f(x, v, t)\, dv, & \text{and } \theta = \pm 1, \end{cases} \tag{1.1}$$

where $x \in \mathbb{R}^d$ is the position, $v \in \mathbb{R}^d$ is the velocity, $t > 0$ is the time, $\nabla_x = (\partial/\partial x_1, \ldots, \partial/\partial x_d)$, $\nabla_v = (\partial/\partial v_1, \ldots, \partial/\partial v_d)$, and $\cdot$ is the inner product in $\mathbb{R}^d$. The transport cross-section parameter σ is assumed to be very small and decoupled from $\beta = \mathcal{O}(1)$. $|\mathcal{S}|^{d-1} \sim 1/\omega_d$ is the surface area of the unit disc in $\mathbb{R}^d$, $\rho(x, t)$ is the spatial density, and $*_x$ denotes the convolution in x. E and ρ can be interpreted as the electrical field and charge, respectively.

For a gradient field, i.e. when E is divergence free, and with no viscosity, i.e. for $\beta = 0$, the first equation in (1.1), would become

$$\partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = \sigma \Delta_v f, \tag{1.2}$$

which, with the rest of equations in (1.1), gives rise to a simplified VPFP system. When E is known, we refer to this system as the VFP system. For $\sigma = 0$ and $E(x, t) = -\nabla_x \phi(x, t)$, we obtain the Vlasov–Poisson equation with an internal potential field $\phi(x, t)$ satisfying the Poisson equation

$$\Delta_x \phi(x, t) = -\theta \int_{\mathbb{R}^d} f(x, v, t) dv = -\theta \rho(x, t), \tag{1.3}$$

with the asymptotic boundary condition

$$\begin{cases} \phi(x, t) \to 0, & \text{for } d > 2, \quad \text{as } |x| \to \infty, \\ \phi(x, t) = \mathcal{O}(\log |x|), & \text{for } d = 2, \quad \text{as } |x| \to \infty. \end{cases} \tag{1.4}$$

For $\beta \neq 0$, we have the following modified version of the VPFP equation

$$\partial_t f + v \cdot \nabla_x f - \nabla_x \phi \cdot \nabla_v f = \nabla_v \cdot (\beta v f + \sigma \nabla_v f), \tag{1.5}$$

where ϕ is assumed to be the exact solution for the Poisson equation (1.3).

1.1. *The continuous problem*

The mathematical study of the VPFP/VFP system has been considered by several authors in various settings, see e.g. Refs. 7 and 21. For the linear Fokker–Planck equation:

$$f_t + v \cdot \nabla_x f + \mathbf{E} \cdot \nabla_v f - \sigma \Delta_v f = S, \quad f(x, v, 0) = f_0(x, v), \tag{1.6}$$

where

$$\mathbf{E} = (\mathbf{E}_i(x, v, t))_{i=1}^{d},$$

is a given vector field and $f_0(x, v)$ and $S(x, v, t)$ are given functions; existence, uniqueness, stability and regularity properties of the solution are straightforward generalizations of the one-dimensional classical results due to Baouendi and Grisvard[6] for the degenerate type equations. These generalizations as well as coupling to the nonlinear problem are due to J. L. Lions[16] and require some regularity assumptions on the data: f_0, S and $\mathbf{E}$. Existence, uniqueness and regularity results relevant to the continuous model problem (1.1) can be found, e.g. in Degond,[8] in one- and two-dimensional cases, and in Bouchut[7] and Victory and O'Dwyer[21] in general three-dimensional setting.

As for the numerical studies: several Lagrangian schemes are developed based on particle methods: In Ref. 9, the authors devise and study a deterministic splitting method for approximating VPFP systems, whereby particle methods are used to treat the convective part and the diffusion is simulated by convolving the particle approximation with the field-free Fokker–Planck kernel. In Refs. 10 and 18 finite-difference methods are considered for the one-dimensional VPFP system, with centered differences used to approximate the diffusion in velocity. In Ref. 23, the numerical procedure combines a deterministic particle type computation with a process for periodically reconstructing the distribution function on a fixed grid in one dimension.

In our studies, assuming a continuous Poisson solver for Eq. (1.3), we focus on the numerical convergence analysis of a deterministic model problem for the VFP system in a bounded phase-space-time domain. This is a convection dominated convection-diffusion problem of degenerate type, (full convection, but only small diffusion in v), for which we study the *hp*-version of the *streamline-diffusion* finite element method and derive convergence rates, which are otherwise more involved using, e.g. particle methods; the most common discretization schemes for the Vlasov type equations. More specifically, for the locally regular solution f in the Sobolev class $H^{s_K+1}(K)$, we derive *optimal a priori* error estimates, basically, of order $\mathcal{O}(\delta_K^{s_K+1/2})$ where $\delta_K \sim \min(h_K/p_K, h_K^2/\sigma)$, with h_K and p_K being the local mesh size and the local spectral order, respectively. (see Remark 3 in Sec. 4 and Ref. 13). A corresponding discontinuous Galerkin study as well as numerical implementations are the subject of a forthcoming paper.

In the classical finite element method (h-version) convergence order improvement relies on mesh refinement while keeping the approximation order within the elements at a fixed low value (suitable for problems with highly singular solutions that require small mesh parameter). Some studies on the h-version of the SD finite element method can be found, e.g. in Ref. 14 for advection-diffusion, Navier–Stokes and first-order hyperbolic equations, in Ref. 15 for Euler and Navier–Stokes equations, in Ref. 1 for the Vlasov–Poisson and in Refs. 2 and 3 for the Fokker–Planck and Fermi equations. On the other hand, in the spectral method, the accuracy improvement is accomplished by raising the order of approximation polynomial rather than mesh refinement (advantageous in approximating smooth solutions). However, most realistic problems have local behavior (are locally smooth or locally singular), therefore a more realistic numerical approach would be a combination of mesh refinement in the vicinity of singularities (with lower order polynomial approximations), and higher order polynomial approximations in high regularity regions (with larger, non-refined, mesh parameter). This strategy, which can be viewed as a generalized adaptive approach, is the *hp*-version of the finite element method. For some basic *hp*-finite element studies, see e.g. Refs. 5, 19 and 20.

An outline of this paper is as follows. In Sec. 2 we introduce the notation and approximation spaces necessary for the subsequent development of the theory. In Sec. 3 we derive error estimates for projection operators useful in our final estimates. Our concluding Sec. 4 is devoted to the study of stability estimates and proof of convergence rates for the *hp*-streamline diffusion approximation of the VFP system.

2. Notation and Assumptions

The continuous problem (1.1), as given in Sec. 1, is not appropriate for numerical considerations since it is formulated in a fully unbounded phase-space-time domain, without any asymptotic boundary conditions. Below we restate the problem (1.1) for $\sigma > 0$ and bounded polyhedral domains $\Omega_x \subset \mathbb{R}^d$ and $\Omega_v \subset \mathbb{R}^d$ associated with some boundary conditions. For simplicity we assume that $\Omega := \Omega_x \times \Omega_v$ is a slight deformation of a bounded, canonical, cubic domain $(-x_0, x_0)^d \times (-v_0, v_0)^d$, $d = 1, 2, 3$. We start with a nonhomogeneous, initial-boundary value problem for the VFP system viz,

$$\begin{cases} \partial_t f + G \cdot \nabla f - \sigma \Delta_v f - \operatorname{div}_v(\beta v f) = S, & \text{in } \Omega_T := \Omega \times (0, T), \\ f(x, v, 0) = f_0(x, v), & \text{in } \Omega_0 := \Omega \times \{0\}, \\ f(x, v, t) = w(x, v, t), & \text{in } ([\Gamma_v^- \times \Omega_v] \cup [\Omega_x \times \partial\Omega_v]) \times (0, T], \end{cases} \tag{2.1}$$

and we let $w = 0$ on $\partial\Omega_v$, i.e. we have an elliptic boundary condition in v and a hyperbolic one in x, where for $v \in \Omega_v$, we define, $\Gamma_v^- = \{x \in \partial\Omega_x : \mathbf{n}(x) \cdot v < 0\}$. We also use the following notation:

$$\nabla f := (\nabla_x f, \nabla_v f) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}, \frac{\partial f}{\partial v_1}, \ldots, \frac{\partial f}{\partial v_d} \right), \quad d = 1, 2, 3, \quad \text{and}$$

$$G(f) := (v, -\nabla_x \phi) = \left(v_1, \ldots, v_d, -\frac{\partial \phi}{\partial x_1}, \ldots, -\frac{\partial \phi}{\partial x_d} \right) = (G_1, \ldots, G_{2d}).$$

Here ϕ satisfies

$$-\Delta_x \phi(x,t) = \int_{\Omega_v} f(x,v,t)\,dv, \quad (x,t) \in \Omega_x \times (0,T], \tag{2.2}$$

where $\nabla_x \phi$ is uniformly bounded and $|\nabla_x \phi| \to 0$ as $x \to \partial\Omega_x$. Also note that G is divergent free

$$\operatorname{div} G(f) = \sum_{i=1}^{d} \frac{\partial G_i}{\partial x_i} + \sum_{i=d+1}^{2d} \frac{\partial G_i}{\partial v_{i-d}} = 0, \quad d = 1,2,3. \tag{2.3}$$

Thus, loosely speaking (generalizing the results for the Navier–Stokes equations in two dimensions), one can assume that there is a unique function $\Psi(x,v,t)$ such that $G = \operatorname{rot}_{(x,v)}\Psi$, $\Psi|_\Gamma = 0$, or alternatively

$$-\Delta\Psi(\cdot,\cdot,t) = \aleph(\cdot,\cdot,t) \quad \text{in } \Omega, \quad \Psi = 0, \quad \text{on } \Gamma, \tag{2.4}$$

where $\aleph = \operatorname{rot}_{(x,v)} G$ may be interpreted as the vorticity of the velocity field $G(f)$. For notational simplicity we split the boundary into the in–(out) flow boundaries:

$$\Gamma^{-(+)} = \{(x,v) \in \Gamma := \partial\Omega|\ G\cdot\mathbf{n} < 0(\geq 0)\}, \quad \mathbf{n} = (\mathbf{n}_x, \mathbf{n}_v), \tag{2.5}$$

where $\Gamma := (\partial\Omega_x \times \Omega_v) \cup (\Omega_x \times \partial\Omega_v) \cup (\partial\Omega_x \times \partial\Omega_v)$, $\mathbf{n}_x$ and $\mathbf{n}_v$ are outward unit normals to $\partial\Omega_x$ and $\partial\Omega_v$, respectively, and $G := G(f)$. Note that since $G = (v, -\nabla_x\phi)$ and $|\nabla_x\phi| \to 0$ as $x \to \partial\Omega_x$, thus $G\cdot\mathbf{n} = (v, -\nabla_x\phi)\cdot(\mathbf{n}_x,\mathbf{n}_v) = v\cdot\mathbf{n}_x$, and hence Γ^- "coincides" with Γ_v^- and therefore throughout our estimates, the boundary terms will be taken over Γ^-, except when we explicitly emphasis the role of v, where we shall employ Γ_v^-. Another justification of this is due to the fact that $(\partial\Omega_x \times \partial\Omega_v)$ has a zero measure, and we have assumed that $w = 0$ on $\partial\Omega_v$, therefore there will not be a nonzero contribution from $(x,v) \in (\Omega_x \times \partial\Omega_v)$, and hence the actual (x,v)-support of w is $\Gamma_v^- \equiv \Gamma^-$.

Our discretization scheme concerns the modified problem (2.1), formulated for the bounded domain Ω_T, and NOT! the original VPFP system stated in $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^+$ as in (1.1). In what follows C will denote a general constant independent of the involved parameters on estimates, unless otherwise explicitly specified.

We now denote an approximate solution for (2.1) by $\tilde{f}$ and recall the usual general procedure of a numerical investigation by decomposing the error viz.,

$$f - \tilde{f} = \big(f - \Pi f\big) - \big(\tilde{f} - \Pi f\big) \equiv \eta - \xi,$$

where Π is an appropriate projection/interpolation operator from the space of the continuous solution f into the (finite-dimensional) space of approximate solution $\tilde{f}$. Considering a suitable norm, denoted by $|||\cdot|||$, the process of estimating the error is split into the following two steps: (i) first we use approximation theory results to derive sharp error bounds for the interpolation error $|||\eta|||$, and then (ii) establish

$$|||\xi||| \leq C|||\eta|||, \tag{2.6}$$

which rely on the stability estimates of bounding $|||\tilde{f}|||$ by the $|||\text{data}|||$. The former step has theoretical nature and is related to the character of the projection/interpolation operator Π, whereas the latter depending on the structure of the $|||\cdot|||$-norm, and the numerical approximation techniques, varies in the order of its difficulty.

Below we present some basic assumptions/notation necessary in *hp*-studies for approximating the projection errors, (see, e.g. Ref. 11): assume a partition $\mathcal{P}$ of $\Omega = \Omega_x \times \Omega_v$ into open patches P which are images of canonical two, four or six-dimensional "cubes": $\hat{P} = (-1,1)^{2d} := \hat{I}^{2d},\ d = 1,2,3,\ \hat{I} = (-1,1)$, under smooth bijections F_P:

$$\forall\ P \in \mathcal{P} : P = F_P(\hat{P}).$$

A mesh $\mathcal{T}$ on Ω is constructed by subdividing the patches: For each P, first we subdivide $\hat{P} = (-1,1)^{2d}$, into $2d$-dimensional generalized quadrilateral elements $\hat{\tau}$ ($2d$-dimensional prisms, i.e. generalized triangular elements would work as well) labeled $\hat{\tau}$ which are affine equivalent to $\hat{P}$, we call this mesh $\hat{\mathcal{T}}_P$ (on $\hat{P}$). On each $P \in \mathcal{P}$ we define a mesh $\mathcal{T}_P$ by setting

$$\forall\ P \in \mathcal{P} : \mathcal{T}_P := \{\tau | \tau = F_P(\hat{\tau}),\ \hat{\tau} \in \hat{\mathcal{T}}_P\}.$$

Note that each $\hat{\tau}(\tau)$ is an image of the reference domain $\hat{P}$ under an affine mapping $A_{\hat{\tau}} : \hat{P} \to \hat{\tau}(F_\tau = F_P \circ A_{\hat{\tau}} : \hat{P} \to \tau)$. Now $\mathcal{T} := \cup_{P\in\mathcal{P}} \mathcal{T}_P$ is a mesh on Ω. We also define the function space

$$F_{\mathcal{P}} = \{F_P : P \in \mathcal{P}\},$$

and the polynomial space

$$\mathcal{A}_p = \mathrm{span}\{(\hat{x}, \hat{v})^\alpha : 0 \le \alpha_i \le p,\ 1 \le i \le 2d\},$$

where $(\hat{x}, \hat{v}) \in \hat{P} := \{(\hat{x}, \hat{v}) \in \mathbb{R}^d \times \mathbb{R}^d : |\hat{x}_j| \le 1 \text{ and } |\hat{v}_j| \le 1,\ j = 1, \dots, d\}$.

Now we let $\mathbf{p}$ be a polynomial degree vector in $\mathcal{T}$,

$$\mathbf{p} = \{p_\tau : \tau \in \mathcal{T}\},$$

and define the continuous *hp*-finite element spaces

$$S^{\mathbf{p},k}(\Omega, \mathcal{T}, F_{\mathcal{P}}) := \{f \in H^k(\Omega) : f|_\tau \circ F_\tau \in \mathcal{A}_{p_\tau},\ \tau \in \mathcal{T}\}, \quad k = 0, 1, \dots,$$

for polynomials with degree vector $\mathbf{p}$, and

$$S^{p,k}(\Omega, \mathcal{T}, F_{\mathcal{P}}) := \{f \in S^{\mathbf{p},k}(\Omega, \mathcal{T}, F_{\mathcal{P}}) : \mathbf{p} = (p, p, \dots, p)\},$$

for the uniform polynomial degree $p_\tau = p,\ \forall\ \tau,\ p > 1$.

Finally we denote by $\|f\|_{k,\hat{I}}$ and $|f|_{k,\hat{I}}$ the $H^k(\hat{I})$ norm and seminorm on $\hat{I}$, respectively (we shall suppress $k = 0$, corresponding to the L_2-norm). We also denote by $S^p(\hat{I})$ the set of polynomials of degree p on $\hat{I}$.

Remark 2.1. To invoke the time variable we shall, basically, use the same notation: we assume a partition $\mathcal{Q}$ of $\Omega_T = \Omega \times (0, T)$ into open patches Q which are images of the canonical cube $\hat{Q} = (-1,1)^{2d+1}$ subdivided into elements $\hat{k} := \hat{\tau} \times \hat{\kappa}$, where each $\hat{\kappa}$ is affine equivalent to $\hat{I}$ corresponding to the time interval $(0, T)$. The exception is that the progress in the time direction is performed successively on the slabs

$\Omega_m := \Omega \times (t_m, t_{m+1}),\ m = 0, \ldots, M-1$, with $t_0 = 0$ and $t_M = T$, and may have jump discontinuities across the discrete time levels $t_m,\ m = 1, \ldots, M-1$. A global mesh is now denoted by $\mathcal{K}$.

3. Approximation of the Projection Error

Using the notation of the previous section and mainly the stability estimate (2.6) we now provide estimates for the projection error η, in some suitable norm. For our choice of the norm $|||\cdot|||$, the terms which will be involved in the projection error are, basically, $\|\eta\|$ and $\|\mathcal{D}\eta\|$, where $\mathcal{D} := (\nabla_x, \nabla_v, d/dt)$ denotes the total gradient operator. In this section we estimate these two quantities for our $(2d+1)$-dimensional problem.

To proceed we denote by $\pi_p^i f$ the one-dimensional H^1-projection of f onto the polynomials of degree p in the ith coordinate, where $1 \leq i \leq d$ would correspond to x_i's for the spatial variable, $d+1 \leq i \leq 2d$ to v_i's for the velocity, and $i = 2d+1$ for the time variable. We shall apply the tensor product in $(2d+1)$-dimensions to the following one- and two-dimensional results while we refer to Ref. 20 for the proof.

Proposition 3.1. *Let $f \in H^{k+1}(\hat{I})$ for some $k \geq 0$. Then, for every $p \geq 1$, there exists a projection $\pi_p f \in S^p(\hat{I})$ such that,*

$$\|f' - (\pi_p f)'\|_{\hat{I}}^2 \leq \frac{(p-s)!}{(p+s)!} |f|_{s+1,\hat{I}}^2, \tag{3.1}$$

$$\|f - \pi_p f\|_{\hat{I}}^2 \leq \frac{1}{p(p+1)} \frac{(p-s)!}{(p+s)!} |f|_{s+1,\hat{I}}^2, \tag{3.2}$$

for any $0 \leq s \leq \min(p, k)$. Moreover,

$$\pi_p f(\pm 1) = f(\pm 1). \tag{3.3}$$

In particular for any $f \in H^1(\hat{I})$ we have that,

$$\|(\pi_p f)'\|_{\hat{I}} \leq 2\|f'\|_{\hat{I}}, \quad \|\pi_p f\|_{\hat{I}} \leq \|f\|_{\hat{I}} + \frac{1}{\sqrt{p(p+1)}} \|f'\|_{\hat{I}}. \tag{3.4}$$

Corollary 3.1. *Let $p \geq 1$ and assume that $\psi \in H^{k+1}(\hat{I}^2)$ for some $k \geq 1$. Then for each i, j, $0 \leq i, j \leq 2d+1$, the projections π_p^i and π_p^j satisfy the following estimates*:

$$\|\psi - \pi_p^i \psi\|_{\hat{I}^2}^2 \leq \frac{1}{p(p+1)} \frac{(p-s)!}{(p+s)!} \|\partial_i^{s+1} \psi\|_{\hat{I}^2}^2, \tag{3.5}$$

$$\begin{aligned} \|\pi_p^i(\psi - \pi_p^j \psi)\|_{\hat{I}^2}^2 \leq{}& \frac{2}{p(p+1)} \times \frac{(p-s)!}{(p+s)!} \|\partial_j^{s+1} \psi\|_{\hat{I}^2}^2 \\ &+ \frac{2}{p^2(p+1)^2} \times \frac{(p-(s-1))!}{(p+(s-1))!} \|\partial_i \partial_j^s \psi\|_{\hat{I}^2}^2, \end{aligned} \tag{3.6}$$

where we have identified $\hat{I}_i \times \hat{I}_j$ by $\hat{I}^2$ and π_p^0 by the identity operator.

We just generalize this procedure to arbitrary d (i.e. to $(2d+1)$ dimensions): To this approach we let $\Pi_p = \prod_{i=1}^{2d+1} \pi_p^i$ denote the tensor product projector and recall $\mathcal{D} = (\nabla_x, \nabla_v, d/dt)$, the $2d+1$ total derivative. We also define the binary multi-index $|m|_l \equiv \sum_{n=1}^{l} m_n$, with $m_n = 0$ or 1. Now we can formulate the main result in this section as:

Theorem 3.1. *Let $\hat{Q} := \hat{P} \times \hat{I}$, $p \geq 1$, $f \in H^{k+1}(\hat{Q})$ for some $k \geq 1$, and set $0 \leq s \leq \min(p,k)$. Then we have the following $\|\cdot\| := \|\cdot\|_{L_2(\hat{Q})}$ estimates for $\eta_p := f - \Pi_p f$,*

$$\|\eta_p\|^2 \leq (2d+1) \times \sum_{i=1}^{2d+1} 2^{i-1} \sum_{|m|_{i-1} \leq \min\{i-1, s+1\}} \alpha_p^{|m|_{i-1}+1} \beta_{|m|_{i-1}} \|\partial^{|m|_{i-1}} \partial_i^{s-|m|_{i-1}+1} f\|^2,$$

and its total derivative $\mathcal{D}\eta_p = (\nabla_x, \nabla_v, d/dt)\eta_p$,

$$\|\mathcal{D}\eta_p\|^2 \leq \sum_{i=1}^{2d+1} (2d+1) \times \sum_{j=1}^{2d+1} \sum_{\substack{|m|_{j-1} \leq \min\{j-1, s+1\} \\ m_i = 1}} 2^j \alpha_p^{|m|_{j-1}} \beta_{|m|_{j-1}} \|\partial^{|m|_{j-1}} \partial_j^{s-|m|_{j-1}+1} f\|^2,$$

where $\partial^{|m|_{i-1}} = \partial_1^{m_1} \partial_2^{m_2} \cdots \partial_{i-1}^{m_{i-1}}$, $\alpha_p = \frac{1}{p(p+1)}$ and $\beta_{|m|_k} = \frac{(p-s+|m|_k)!}{(p+s-|m|_k)!}$.

Proof. We may use the telescopic identity

$$f - \Pi_p f = \left(f - \prod_{k=1}^{2d+1} \pi_p^k f \right) = \sum_{i=1}^{2d+1} \left(\prod_{j=0}^{i-1} \pi_p^j \right) (f - \pi_p^i f),$$

to get the estimate

$$\|f - \Pi_p f\|^2 \leq (2d+1) \sum_{i=1}^{2d+1} \left\| \left(\prod_{j=0}^{i-1} \pi_p^j \right) (f - \pi_p^i f) \right\|^2. \tag{3.7}$$

It is easy to show that for $s+1 \geq |m|_n$,

$$\left\| \left(\prod_{j=0}^{n} \pi_p^j \right) (f - \pi_p^{n+1} f) \right\|^2 \leq \sum_{|m|_n \leq n} 2^n \alpha_p^{|m|_n+1} \beta_{|m|_n} \|\partial^{|m|_n} \partial_{n+1}^{s-|m|_n+1} f\|^2. \tag{3.8}$$

Note that for $n = 0$ and 1, (3.8) is just (3.5) and (3.6), respectively. Furthermore, since $\pi_p^0 = \mathrm{id}$, we have by the second inequality in (3.4) and using twice (3.6), (second time with $\psi := \partial_1 f$), that

$$\begin{aligned}
\|\pi_p^1\pi_p^2(f-\pi_p^3 f)\|^2 &\le 2\|\pi_p^2(f-\pi_p^3 f)\|^2 + \frac{2}{p(p+1)}\|\partial_1\pi_p^2(f-\pi_p^3 f)\|^2 \\
&\le 2\frac{2}{p(p+1)}\frac{(p-s)!}{(p+s)!}\|\partial_3^{s+1}f\|^2 \\
&\quad + 2\frac{2}{p^2(p+1)^2}\frac{(p-(s-1))!}{(p+(s-1))!}\|\partial_2\partial_3^s f\|^2 \\
&\quad + \frac{2}{p(p+1)}\left\{\frac{2}{p(p+1)}\frac{(p-(s-1))!}{(p+(s-1))!}\|\partial_1\partial_3^s f\|^2\right. \\
&\quad + \left.\frac{2}{p^2(p+1)^2}\frac{(p-(s-2))!}{(p+(s-2))!}\|\partial_1\partial_2\partial_3^{s-1}f\|^2\right\},
\end{aligned}$$

which gives (3.8) for $n = 2$. For the remaining values of n, i.e. for $3 \le n \le 2d+1$, (3.8) is justified by a similar, however lengthy, "induction-like" procedure which we omit. Recall that, to get non-negative differentiation orders, the parameters should be accordingly related. In the sequel we do not state these relations explicitly.

The first assertion of the theorem now follows from (3.7) and (3.8). To show the second estimate we start by rewriting and subsequently simplifying, via (3.7), the total derivative $\mathcal{D}\eta_p$,

$$\begin{aligned}
\|\mathcal{D}\eta_p\|^2 &= \sum_{i=1}^{2d+1}\left\|\partial_i\left(f-\prod_{k=1}^{2d+1}\pi_p^k f\right)\right\|^2 \\
&\le \sum_{i=1}^{2d+1}(2d+1)\sum_{j=1}^{2d+1}\left\|\partial_i\left(\prod_{l=0}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|^2.
\end{aligned}$$

Below we split the estimate of $\|\partial_i(\prod_{l=0}^{j-1}\pi_p^l)(f-\pi_p^j f)\|_{\hat{Q}}^2$ into the following three possible cases:

Case I: $i \le j-1$. Using the first estimate in (3.4) we have

$$\begin{aligned}
\left\|\partial_i\left(\prod_{l=0}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|^2 &= \left\|\partial_i\pi_p^i\left(\prod_{\substack{l=0\\ l\ne i}}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|^2 \\
&\le 4\left\|\partial_i\left(\prod_{\substack{l=0\\ l\ne i}}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|^2.
\end{aligned}$$

Now since ∂_i is no longer in the direction of any of the remaining projections in the product, we can use the second estimate in (3.4) and (3.8) as

$$\begin{aligned}&\left\|\partial_i\left(\prod_{\substack{l=0\\l\neq i}}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|_{\hat{Q}}^2\\&\quad\leq\sum_{\substack{|m|_{j-1}\leq j-1\\m_i=1}}2^{j-2}\alpha_p^{|m|_{j-1}}\beta_{|m|_{j-1}}\left\|\partial^{|m|_{j-1}}\partial_j^{s-|\bar{m}|_{j-1}+1}f\right\|_{\hat{Q}}^2,\end{aligned}$$

where $|\bar{m}|_{j-1}=|m|_{j-1}$, with $m_k=0$ or 1 for $k\neq i,\ 0\leq k\leq j-1$ and $m_i\equiv 1$. In this way the contribution of ∂_i is included on the right-hand side above. Hence, we have shown the second assertion of the theorem in case I.

Case II: $i=j$. Thus we can write

$$\left\|\partial_i\left(\prod_{l=0}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|_{\hat{Q}}^2=\left\|\partial_j\left(\prod_{l=0}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|_{\hat{Q}}^2=\|\partial_j(\mathcal{F}-\pi_p^j\mathcal{F})\|_{\hat{Q}}^2,$$

where $\mathcal{F}=(\prod_{l=0}^{j-1}\pi_p^l)f$. By (3.1) we have $\|\partial_j(\mathcal{F}-\pi_p^j\mathcal{F})\|_{\hat{Q}}^2\leq\beta_0\|\partial_j^{s+1}\mathcal{F}\|_{\hat{Q}}^2$. This quantity can now be estimated by a (repeated) use of the second estimate in (3.4):

$$\left\|\left(\prod_{l=0}^{j-1}\pi_p^l\right)\varphi\right\|_{\hat{Q}}^2\leq\sum_{|m|_{j-1}\leq j-1}2^{j-1}\alpha_p^{|m|_{j-1}}\|\partial^{|m|_{j-1}}\varphi\|_{\hat{Q}}^2,$$

so that, replacing φ by $\partial_j^{s+1}f$, we obtain the desired result also for the case $i=j$.

Case III: $i>j$. Here we can apply (3.8) directly since ∂_i and the projections in $\|\partial_i(\prod_{l=0}^{j-1}\pi_p^l)(f-\pi_p^j f)\|_{\hat{Q}}^2$, are decoupled and therefore we can derive the estimate

$$\begin{aligned}\left\|\partial_i\left(\prod_{l=0}^{j-1}\pi_p^l\right)(f-\pi_p^j f)\right\|_{\hat{Q}}^2&\leq\sum_{|m|_{j-1}\leq j-1}2^{j-1}\alpha_p^{|m|_{j-1}+1}\beta_{|m|_{j-1}-1}\\&\qquad\times\|\partial_i\partial^{|m|_{j-1}}\partial_j^{s-|m|_{j-1}}f\|_{\hat{Q}}^2.\end{aligned}$$

Summing over i gives the second estimate and completes the proof. □

Remark 3.1. We can write the above estimates in a general setting for a partition $\mathcal{R}$ of a bounded, convex, curved polyhedral domain $D\subset\mathbb{R}^{\mathcal{N}}$: Let $R\in\mathcal{R}$ be an image of the $\mathcal{N}$-dimensional canonical hypercube $\hat{R}:=(-1,1)^{\mathcal{N}}$, with $\mathcal{N}$-dimensional mesh $\mathcal{M}_{\hat{R}}$, under the bijective map $G_R:\ R=G_R(\hat{R})$, and with a corresponding generalized $\mathcal{N}$-dimensional quadrilateral mesh $\mathcal{M}_R$ on R. Then for a global generalized quadrilateral mesh $\mathcal{M}:=\cup_{R\in\mathcal{R}}\mathcal{M}_R$ on D, the projection error estimates are obtained by the change of variables and a simple scaling argument where we assume that the patch $\hat{R}$ is the canonical deformation of R with no significant rescaling. More specifically we assume that there are positive constants c_1 and c_2 such that

$$c_1 \leq h_K/\hat{h}_K \leq c_2, \quad \forall\ K \in \mathcal{M}, \tag{3.9}$$

where $h_K = \mathrm{diam}(K)$, $\hat{h}_K = \mathrm{diam}(\hat{K})$, $K = G_R(\hat{K})$, and $\hat{K} \subset \hat{R}$ is a reference element in the mesh $\mathcal{M}_{\hat{R}}$. All the corresponding notation such as the polynomial degree distribution $\mathbf{r} = \{r_K | K \in \mathcal{M} := \cup_{R\in\mathcal{R}}\mathcal{M}_R\}$, the affine mapping $A_{\hat{K}} : \hat{R} \to \hat{K}$, the patch-map vector $G_{\mathcal{R}} = \{G_R : R \in \mathcal{R}\}$, and the element map $G_K := G_R \circ A_{\hat{K}}$ with $K = G_K(\hat{R})$, as well as the function space $S^{\mathbf{r},k}(D, \mathcal{M}, G_{\mathcal{R}})$ are defined correspondingly as in Sec. 2. However, since in the streamline diffusion method we allow discontinuities in time, we formulate the generalization in fully discontinuous setting using a local version of $S^{\mathbf{r},k}(D, \mathcal{M}, G_{\mathcal{R}})$ with only, elementwise high regularity:

$$S^{\mathbf{r},\mathbf{k}}_{\mathrm{loc}}(D, \mathcal{M}, G_{\mathcal{R}}) := \{f \in S^{\mathbf{r},0}(D, \mathcal{M}, G_{\mathcal{R}}) : f|_K \in H^{k_K+1}(K)\}, \tag{3.10}$$

where $\mathbf{k} := \{k_K : K \in \mathcal{M}\}$. This is a more general setting which is also appropriate in the discontinuous Galerkin studies. Finally, we have the following general result:

Theorem 3.2. *Let $R \in \mathcal{R}$ and the polynomial degree distribution $\mathbf{r}$ be defined as above. $\forall\ K \in \mathcal{M}_R$, let $f|_K \in H^{k_K+1}(K)$ for some $k_K \geq 1$ and define $\Pi f \in S^{\mathbf{r},\mathbf{k}}_{\mathrm{loc}}(D, \mathcal{M}, G_{\mathcal{R}})$ elementwise by $(\Pi f)|_K \circ G_R := \Pi_{r_K}(f|_K \circ G_R)$, $\forall\ K \in \mathcal{M}_R$. Then, for $r_K \geq 1$ and for $0 \leq s_K \leq \min(r_K, k_K)$ we have the following estimates*:

$$\|f - \Pi f\|^2_R \leq C \sum_{K\in\mathcal{M}_R} \left(\frac{h_K}{2}\right)^{2s_K+2} \Phi_1(r_K, s_K)\|\hat{f}\|^2_{s_K+1,\hat{K}},$$

$$\|\mathcal{D}(f - \Pi f)\|^2_R \leq C \sum_{K\in\mathcal{M}_R} \left(\frac{h_K}{2}\right)^{2s_K} \Phi_2(r_K, s_K)\|\hat{f}\|^2_{s_K+1,\hat{K}},$$

where $\hat{f} = f \circ G_R$, $K = G_R(\hat{K})$, $\|\cdot\|_{s_K+1,\hat{K}}$ is the Sobolev norm in $H^{s_K+1}(\hat{K})$ and

$$\Phi_1(p, s) = \mathcal{N} \sum_{i=1}^{\mathcal{N}} 2^{i-1} \sum_{|m|_{i-1}\leq i-1} \alpha_p^{|m|_{i-1}+1}\beta_{|m|_{i-1}},$$

$$\Phi_2(p, s) = \mathcal{N} \sum_{i=1}^{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} 2^j \sum_{\substack{|m|_{j-1}\leq j-1 \\ m_i=1}} \alpha_p^{|m|_{j-1}}\beta_{|m|_{j-1}}.$$

Proof. The proof is based on a scaling argument due to the use of a corresponding affine mapping $A_{\hat{K}}$, this time $A_{\hat{K}} : \hat{R} \to \hat{K}$, on the results of Theorem 3.2 above: A consequence of applying tensor product to the proof of Theorem 3.4 in Ref. 11. $\square$

4. The Streamline Diffusion Method

The SD-method for (2.1) is based on using finite elements over the phase-space-time domain Ω_T. To define this method, following the notation in Sec. 2, let $\mathcal{T}_h = \{\tau\}$ be a finite element subdivision of $\Omega = \Omega_x \times \Omega_v$ into open elements $\tau := F_P(\hat{\tau})$, where P corresponds to an open patch in Ω, further let $0 = t_0 < t_1 < \cdots < t_M = T$ be a partition of the time interval $(0, T)$ into subintervals $I_m := (t_m, t_{m+1})$. For each $m = 0, \ldots, M-1$, we denote by $\mathcal{K}_{h,m} := \{K : \ K = \tau \times I_m, \ \tau \in \mathcal{T}_h\}$ the corresponding subdivision of $\Omega_m := \Omega \times I_m$. Finally let $\mathcal{K}_h = \cup_m \mathcal{K}_{h,m} = \{K\}$ be the subdivision of Ω_T into elements K and with the piecewise constant mesh function h defined by $h(x, v, t) := h_K = \mathrm{diam}(K)$, $(x, v, t) \in K$. We assume that the family of partitions $\{\mathcal{K}_h\}_{h>0}$ is *shape regular*; i.e. for each $K \in \mathcal{K}_h$ there is an inscribed ($(2d+1)$-dimensional) sphere in K such that the ratio of the diameter of this sphere and the diameter of K is bounded from below independent of K and h_K, i.e. there is a positive constant C_0, independent of h, such that

$$C_0 h_K^{2d+1} \le \mathrm{meas}(K), \quad \forall\ K \in \cup_h \mathcal{K}_h. \tag{4.1}$$

Now on each slab Ω_m we define a corresponding finite element space by

$$V_h^{\mathbf{p}_m} = \Big\{ g \in S^{\mathbf{p}_m, \mathbf{k}_m}(\Omega_m, \mathcal{K}_{h,m}) : \ g\Big|_K \in P_{p_K}(K); \ \forall\ K = \tau \times I_m \Big\},$$

where $P_{p_K}(K) := P_{p_K}(\tau) \times P_{p_K}(I_m)$ denotes the set of polynomials in x, v and t of degree at most $p_K \ge 1$ on K and $S^{\mathbf{p}_m, \mathbf{k}_m}$ defined similarly to (3.10), with

$$\mathbf{p}_m := \{p_K | \ K \in \mathcal{K}_{h,m}\}, \ \mathbf{k}_m := \{k_K | \ K \in \mathcal{K}_{h,m}\}.$$

Now we let $\mathbf{q} = (\mathbf{p}_0, \mathbf{p}_2, \ldots, \mathbf{p}_{M-1})$ be the polynomial degree (multi-) vector in the mesh $\mathcal{K}_h$ for the Ω_T, $\mathbf{k} := \{\mathbf{k}_m\}$, and define

$$V_h^{\mathbf{q}} = \prod_{m=0}^{M-1} V_h^{\mathbf{p}_m},$$

to be a finite element space in the whole $\Omega_T = \Omega \times (0, T)$.

To invoke the inhomogeneous boundary condition, for simplicity, first we consider the steady state version of the problem (2.1), viz.

$$\mathcal{B} f := G \cdot \nabla f - \sigma \Delta_v f - \mathrm{div}_v(\beta v f) = S \quad \text{in } \Omega, \quad \text{with } \ f = w \quad \text{on } \Gamma^-, \tag{4.2}$$

where we assume that $S \in L_2$ and $w \in L_2(\Gamma^-)$. The usual (not streamline diffusion) weak formulation of this problem is then to find $f \in H^1$ such that

$$b(G; f, g) = L(g), \quad \forall\ g \in H_0^1(\Omega), \quad \text{with } \ \gamma^- f = w, \tag{4.3}$$

where

$$b(\omega; g, u) = \int_\Omega (\omega \cdot \nabla g \ u + \sigma \nabla_v g \cdot \nabla_v u + \nabla_v \cdot (\beta v g) u) \, dx dv, \tag{4.4}$$

$$L(g) = \int_\Omega f g \, dx \, dv, \quad \gamma^- = \gamma|_{\Gamma^-} \quad \text{and} \quad \gamma : H^1 \to L_2(\Gamma) \ \text{ is the trace operator.}$$

For the existence of a solution, we assume that the given function w is the restriction to Γ^- of the trace of some function $\chi \in H^1$, i.e. $w = \gamma^-\chi$. Setting $\zeta = f - \chi$, we then seek $\zeta \in H_0^1$ satisfying

$$b(G : \zeta, g) = L(g) - b(G; \chi, g), \quad \forall\, g \in H_0^1. \tag{4.5}$$

In this way the right-hand side is a bounded linear functional on H_0^1 and hence it follows from the Lax–Milgram theorem that there exists a unique $\zeta \in H_0^1$ satisfying (4.5). Clearly, $f = \chi + \zeta$ satisfies (4.4) and $\gamma f|_{\Gamma^-} = w$. This solution is unique, for if (4.3) had two solutions f_1 and f_2 with the same data S and w (with w interpreted as the restriction of the trace of f to Γ^-), then their difference $f_1 - f_2 \in H_0^1$ would be a weak solution of the homogeneous version of (4.3) with $S = w = 0$, and hence using the stability estimate

$$|f|_1 \leq C\|S\|,$$

would imply $f_1 - f_2 = 0$, i.e. $f_1 = f_2$. Hence, (4.3) has a unique weak solution. In particular, the solution f is independent of the choice of the extension χ of the boundary value w. Now we define the globally discrete function space

$$\tilde{V}_h^{\mathbf{q}} := \{g \in V_h^{\mathbf{q}} : \; g|_{\Gamma^-} = \tilde{w}\},$$

where $\tilde{w}$ is the nodal interpolant of the inflow boundary function w and both w and $\tilde{w}$ are continuously extended to the boundary Γ with the property that

$$\langle \tilde{w}, u\rangle_{\Gamma^\pm} = \langle w, u\rangle_{\Gamma^\pm}. \tag{4.6}$$

These extensions, in turn, can be considered as restrictions to Γ of functions in H^1 and $V_h^{\mathbf{q}}$, respectively.

Now we return to the SD version and, for further notational convenience, introduce the slab-wise representations:

$$(f, g)_m = (f, g)_{\Omega_m}, \quad \|g\|_m = (g, g)_m^{1/2},$$

and define the inner product and seminorm at the time level t_m by

$$\langle f, g\rangle_m = (f(\cdot,\cdot,t_m), g(\cdot,\cdot,t_m))_\Omega, \quad |g|_m = \langle g, g\rangle_m^{1/2}.$$

We also present the jump term by, $[g] = g^+ - g^-$, where (to include also the case with $\sigma \equiv 0$),

$$\begin{aligned} g^\pm &= \lim_{s\to 0\pm} g(x, v, t+s), && \text{for } (x, v) \in \Omega_x \times \Omega_v, && t \in I, \\ g^\pm &= \lim_{s\to 0\pm} g(x+sv, v, t+s), && \text{for } (x, v) \in (\partial\Omega_x) \times \Omega_v, && t \in I, \end{aligned}$$

and use the following notation for the boundary integrals

$$\begin{aligned} \langle f^\mp, g^\mp\rangle_{\Gamma^\pm} &= \int_{\Gamma^\pm} f^\mp g^\mp \left|(G^h \cdot \mathbf{n})\right| d\nu, && G^h := G(f^h) \equiv G(f^h), \\ \langle f^\mp, g^\mp\rangle_{\lambda_m^\pm} &= \int_{I_m} \langle f^\mp, g^\mp\rangle_{\Gamma^\pm}\, dt, && \lambda_m^\pm := \Gamma^\pm \times I_m, \\ \langle f^\mp, g^\mp\rangle_{\Lambda^\pm} &= \int_0^T \langle f^\mp, g^\mp\rangle_{\Gamma^\pm}\, dt, && \Lambda^\pm := \Gamma^\pm \times (0, T). \end{aligned}$$

4.1. *Stability of the time-dependent SD finite element method*

To choose streamline diffusion test functions we single out the terms in (2.1) which give rise to significant advection: these are f_t, the G-term $G(f^h)\cdot\nabla f$ and, since $\beta = O(1)$, the β-term $\nabla_v\cdot(\beta v f)$. As a result a more appropriate class of test functions, with contributions corresponding to all these terms, may have the form:

$$u + \delta(u_t + G(f^h)\cdot\nabla u + \beta du + \beta v\cdot\nabla_v u), \tag{4.7}$$

where d is the dimension of Ω_v. However, in order to present the analysis in a rather concise form, in the SD test functions we shall not include the terms which result to estimates of comparable order. Here, multiplying the VFP Eq. (2.1) by (4.7) and integrating over Ω_T would produce the following, β-terms:

$$\text{(i)}\;\; \beta d(u_t, u),\;\; \text{(ii)}\;\; \beta^2(\delta v\cdot\nabla_v u, v\cdot\nabla_v u),\;\; \text{(iii)}\;\; (u_t + G(f^h)\cdot\nabla u, \beta v\cdot\nabla_v u). \tag{4.8}$$

As we shall see in (4.12) below, the terms (i) and (iii) also arise in the variational formulation using the simpler test function:

$$u + \delta(u_t + G(f^h)\cdot\nabla u). \tag{4.9}$$

Thus the "actual" new contribution, in using (4.7), is the β-term (ii), i.e.

$$\beta^2\|\delta v\cdot\nabla_v u\|^2. \tag{4.10}$$

This, combined with the contribution of the diffusion term $\sigma\Delta_v f$, give rise to the control of the term

$$\beta^2\|\delta v\cdot\nabla_v u\|^2 + \sigma\|\nabla_v u\|^2 \le \tilde\sigma\|\nabla_v u\|^2 \quad \text{where } \tilde\sigma := \sigma + |\delta|_\infty\beta^2|\Omega_v|^2, \tag{4.11}$$

rather than only $\sigma\|\nabla_v u\|^2$, corresponding to the use of (4.9). Hence basically, the additional contribution to the "artificial" diffusion, which comes from the β-term in (4.7) is of order $\mathcal{O}(\delta)$. To work with (4.7) however, requires somewhat more involved technical details. To be concise, we skip the tedious details and consider the test functions of the form (4.9), bearing in mind that involving β-terms as in (4.7), we would in fact obtain a slightly better estimate controlling (4.11).

In the conventional h version of the SD-method for time-dependent *hyperbolic, or convection dominated* problems, assuming $\tilde f$ to be an approximate solution and using test functions of the form (4.9), where δ is a small parameter (normally $\delta\sim h$), would supply us with a necessary (missing) diffusion term of order δ in the direction of the streamlines: $(1, G(f^h))$. More specifically, in the stability estimates we will be able to control an extra term of the form $\delta\|u_t + G(f^h)\cdot\nabla u\| \sim h\|u_t + G(f^h)\cdot\nabla u\|$. In the hp studies, however, the choice of δ is somewhat more involved and in addition to the equation type it also depends on the choice of the parameters h and p: these are chosen locally (elementwise) in an optimal manner. Therefore, in our estimates, δ would appropriately appear as an elementwise (local) parameter. Below we formulate both global and local time-dependent SD-method for problem (2.1) and continue the analysis of hp-version for the local case. Assuming that $\tilde w$ is the nodal interpolant of w at the inflow boundary nodes satisfying (4.6), the SD-method for (2.1) can now be formulated as follows:

find $f^h \in \tilde{V}_h^{\mathbf{q}} := \{g \in V_h^{\mathbf{q}} : g|_{\Gamma^-} = \tilde{w}\}$ *such that for* $m = 0, \dots, M-1$,

$$\begin{aligned}(f_t^h + G(f^h) \cdot \nabla f^h, u + \delta(u_t + G(f^h) \cdot \nabla u))_m + \sigma(\nabla_v f^h, \nabla_v u)_m \\ + \langle [f^h], u^+ \rangle_m - \sigma(\delta \Delta_v f^h, u_t + G(f^h) \cdot \nabla u)_m \\ - (\nabla_v \cdot (\beta v f^h), u + \delta(u_t + G(f^h) \cdot \nabla u))_m \\ = (S, u + \delta(u_t + G(f^h) \cdot \nabla u))_m \\ + \langle w^+, u^+ \rangle_{\lambda_m^-} + \langle w^-, u^- \rangle_{\lambda_m^+}, \quad \forall\, u \in \tilde{V}_h^{\mathbf{q}},\end{aligned} \tag{4.12}$$

where the expression $(\delta \Delta_v f^h, u_t + G(f^h) \cdot \nabla u)_m$ should be interpreted as a sum of integrals over the elements $K \in \mathcal{K}_{h,m}$ in the slab Ω_m. The problem (4.12) is equivalent to: *find* $f^h \in \tilde{V}_h^{\mathbf{q}}$ *such that*,

$$B_\delta(G(f^h); f^h, u) - J_\delta(f^h, u) = L_\delta(u) \quad \forall\, u \in \tilde{V}_h^{\mathbf{q}}, \tag{4.13}$$

where for a given appropriate function g, the trilinear form B_δ is defined as

$$\begin{aligned}B_\delta(\omega; g, u) = \sum_{m=0}^{M-1} \Big[(g_t + \omega \cdot \nabla g, u + \delta(u_t + G(f^h) \cdot \nabla u))_m \\ + \sigma(\nabla_v g, \nabla_v u)_m - \sigma(\delta \Delta_v g, u_t + G(f^h) \cdot \nabla u)_m \Big] \\ + \sum_{m=1}^{M-1} \langle [g], u^+ \rangle_m + \langle g^+, u^+ \rangle_{\Lambda^-} + \langle g^-, u^- \rangle_{\Lambda^+} + \langle g^+, u^+ \rangle_0,\end{aligned}$$

the bilinear form J_δ by,

$$J_\delta(g, u) = \sum_{m=0}^{M-1} (\nabla_v \cdot (\beta v g), u + \delta(u_t + G(f^h) \cdot \nabla u))_m,$$

and finally the linear form L_δ is given by,

$$\begin{aligned}L_\delta(u) = \sum_{m=0}^{M-1} (S, u + \delta(u_t + G(f^h) \cdot \nabla u))_m \\ + \langle f_0, u^+ \rangle_0 + \langle w^+, u^+ \rangle_{\Lambda^-} + \langle w^-, u^- \rangle_{\Lambda^+}.\end{aligned}$$

Note that B_δ, J_δ and L_δ depend implicitly on f^h through the term $G(f^h)$. In the sequel we relate the cross-section σ to the element size h_K by assuming that $\sigma \leq \min_K h_K,\ K \in \mathcal{K}_h$. Note also that the discrete version of (2.3) takes now the following form:

$$\text{div } G(f^h) = 0. \tag{4.14}$$

Stability and convergence estimates for (4.13) are derived in the triple norm:

$$|||u|||^2 = \frac{1}{2}\Bigg[2\sigma\|\nabla_v u\|^2_{\Omega_T} + |u|^2_M + |u|^2_0 + \sum_{m=1}^{M-1} |[u]|^2_m + 2\|\delta(u_t + G(f^h)\cdot\nabla u)\|^2_{\Omega_T}$$
$$+ \; 3\int_{\partial\Omega\times I} u^2 \big|G^h\cdot \mathbf{n}\big|\, d\nu\, ds\Bigg].$$

Now let $(\cdot,\cdot)_K$ denote the L_2-inner product over K and define the non-negative piecewise constant function δ by

$$\delta|_K = \delta_K, \quad \text{for } K \in \mathcal{K}_h,$$

where δ_K is a non-negative constant on element K. To formulate the local version of (4.13), we replace in the definitions for B_δ, J_δ and L_δ the inner products $(\cdot,\cdot)_m$, over the slab Ω_m by the corresponding sum: $\sum_{K\in\mathcal{K}_{h,m}}(\cdot,\cdot)_K$, and all δ by δ_K. Thus, more specifically we have the problem (4.13), with the trilinear form B_δ defined as:

$$B_\delta(G(f^h); f^h, u) = \sum_{m=0}^{M-1}\sum_{K\in\mathcal{K}_{h,m}} [(f^h_t + G(f^h)\cdot\nabla f^h, u + \delta_K(u_t + G(f^h)\cdot\nabla u))_K$$
$$+\sigma(\nabla_v f^h, \nabla_v u)_K - \delta_K\sigma(\Delta_v f, u_t + G(f^h)\cdot\nabla u)_K]$$
$$+\sum_{m=1}^{M-1} \langle [f^h], u^+\rangle_m + \langle f^{h,+}, u^+\rangle_{\Lambda^-} + \langle f^{h,-}, u^-\rangle_{\Lambda^+} + \langle f^{h,+}, u^+\rangle_0,$$

the bilinear form J_δ as,

$$J_\delta(f^h, u) = \sum_{m=0}^{M-1}\sum_{K\in\mathcal{K}_{h,m}} (\nabla_v\cdot(\beta v f^h), u + \delta_K(u_t + G(f^h)\cdot\nabla u))_K,$$

and the linear form L_δ given by,

$$L_\delta(u) = \sum_{m=0}^{M-1}\sum_{K\in\mathcal{K}_{h,m}} (S, u + \delta_K(u_t + G(f^h)\cdot\nabla u))_K$$
$$+ \langle f_0, u^+\rangle_0 + \langle w^+, u^+\rangle_{\Lambda^-} + \langle w^-, u^-\rangle_{\Lambda^+}.$$

Note that in the h version of the SD approach for the time-dependent problems we interpret $(\cdot,\cdot)_{\Omega_T}$ as $\sum_{m=0}^{M-1}(\cdot,\cdot)_m$ and, assuming discontinuities in the time variable, include jump terms in the time direction. Thus we estimate the sum of the norms over slabs Ω_m as well as the contributions from the jumps over time levels t_m, $m = 1,\ldots,M-1$. Whereas in *hp* version we have, in addition to slab-wise estimates, a further step of identifying $(\cdot,\cdot)_m$ by $\sum_{K\in\mathcal{K}_{h,m}}(\cdot,\cdot)_K$ counting for the local character of the parameters h_K, p_K and δ_K, and consequently replacing some of the terms of the form $(\cdot,\cdot)_m$ and $\|\cdot\|_m$ (e.g. those involving δ_K), by the equivalent ones: $(\cdot,\cdot)_m = \sum_{K\in\mathcal{K}_{h,m}}(\cdot,\cdot)_K$ and $\|\cdot\|_m = \sum_{K\in\mathcal{K}_{h,m}}\|\cdot\|_K$, respectively. Thus in our SD estimates the sum $\sum_{K\in M_R}$ in Theorem 3.2 is identified by $\sum_{m=0}^{M-1}\sum_{K\in K_{h,m}}$ and all the corresponding terms are interpreted accordingly.

Below we prove stability estimates and derive convergence rates for error in $|||\cdot|||$.

Proposition 4.1. *We assume that the mesh $\mathcal{K}_h$ consists of shape-regular elements K and the SD-parameter $\delta_K(:= \delta|_K)$ on K satisfies $0 \le \delta_K \le \min\left(\frac{h_K^2}{\sigma C_I^2}, \frac{h_K}{p_K C_I^2}\right)$, with $C_I = C(C_{\text{inv}}, C_0)$, where C_{inv} is the constant in an inverse estimate and C_0 is as in* (4.1)*. Then the trilinear form $B_\delta(G(f^h), \cdot, \cdot)$ is coercive on $V_h^{\mathbf{q}} \times V_h^{\mathbf{q}}$:*

$$B_\delta(G(f^h); u, u) \ge \frac{1}{2}|||u|||^2, \quad \forall\ u \in V_h^{\mathbf{q}}$$

Further, for any constant $C_1 > 0$ we have for any $u \in V_h^{\mathbf{q}}$,

$$\|u\|_{\Omega_T}^2 \le \left[\frac{1}{C_1}\|u_t + G(f^h)\cdot\nabla u\|_{\Omega_T}^2 + \sum_{m=1}^{M}|u^-|_m^2 + \int_{\partial\Omega\times I} u^2\big|G^h\cdot\mathbf{n}\big|\,d\nu\,ds\right]\delta e^{C_1\delta}.$$

Proof. Starting from our trilinear form,

$$\begin{aligned}
B_\delta(G(f^h); u, u) &= \big(u_t, u\big)_{\Omega_T} + \langle u^+, u^+\rangle_0 - \sigma\sum_{K\in\mathcal{K}_h}\delta_K(\Delta_v u, u_t + G(f^h)\cdot\nabla u)_K \\
&\quad + \sum_{K\in\mathcal{K}_h}\delta_K\|u_t + G(f^h)\cdot\nabla u\|_K^2 + \sigma\|\nabla_v u\|_{\Omega_T}^2 + \sum_{m=1}^{M-1}\langle [u], u^+\rangle_m \\
&\quad + \sum_{m=0}^{M-1}[(G(f^h)\cdot\nabla u,\, u)_m + \langle u^+, u^+\rangle_{\lambda_m^-} + \langle u^-, u^-\rangle_{\lambda_m^+}],
\end{aligned}$$

we work separately on pieces of this form. Integrating by parts,

$$(u_t, u)_{\Omega_T} + \langle u^+, u^+\rangle_0 + \sum_{m=1}^{M-1}\langle [u], u^+\rangle_m = \frac{1}{2}\left[|u|_M^2 + |u|_0^2 + \sum_{m=1}^{M-1}|[u]|^2\right]. \quad (4.15)$$

To estimate the term involving $\delta_K\sigma$ we apply Cauchy–Schwartz and the inverse inequalities, and use the assumption on δ_K, to get

$$\begin{aligned}
&\delta_K\sigma(\Delta_v u, u_t + G(f^h)\cdot\nabla u)_K \\
&\quad \le \frac{1}{2}C_I h_K^{-1}\sqrt{\sigma\delta_K}[\sigma\|\nabla_v u\|_K^2 + \delta_K\|u_t + G(f^h)\cdot\nabla u\|_K^2] \\
&\quad \le \frac{1}{2}[\sigma\|\nabla_v u\|_K^2 + \delta_K\|u_t + G(f^h)\cdot\nabla u\|_K^2],
\end{aligned} \quad (4.16)$$

where, as we mentioned earlier, the constant C_I depends on the constants in the inverse estimate and the shape-regularity constant C_0 of the triangulation $\mathcal{K}_h$. Further using Green's formula and (2.3) we have

$$\begin{aligned}
&(G(f^h)\cdot\nabla u, u)_\Omega + \langle u^+, u^+\rangle_{\Gamma^-} + \langle u^-, u^-\rangle_{\Gamma^+} \\
&\quad = \frac{1}{2}\int_{\partial\Omega} u^2|G(f^h)\cdot\mathbf{n}|\,d\nu + \langle u^+, u^+\rangle_{\Gamma^-} + \langle u^-, u^-\rangle_{\Gamma^+} \\
&\quad = \frac{3}{2}\int_{\partial\Omega} u^2|G(f^h)\cdot\mathbf{n}|\,d\nu.
\end{aligned} \quad (4.17)$$

Now summing (4.16) over K, integrating (4.17) over I_m, summing over m and combining with (4.15) gives the first assertion of the proposition. For the second part we apply (4.17), Grönwall's inequality on $\|u\|^2_{\Omega_T}$ and use Lemma 3.2 of Ref. 1. $\square$

Remark 4.1. Our choice of δ_K is due to the fact that we have a convection dominated problem with a small diffusion term only in v. Further we have not involved the contributions, to the constants, from the parameter p in our inverse estimates. For a genuine convection diffusion problem $\delta_K \sim h_K^2/p^4$, cf. Ref. 12, and the corresponding hp inverse estimate is of order $h_K^{-1}p^2$. It can be shown, however, that both approaches lead to the same final estimates, though the former is much simpler to follow.

Proposition 4.2. *Let $f^h \in \tilde{V}_h^{\mathbf{q}}$ and write $f - f^h = \eta - \xi$, where $\eta = f - \Pi_p f$, $\xi = f^h - \Pi_p f$ and $\Pi_p f \in \tilde{V}_h^{\mathbf{q}}$ is defined as in Sec.* 3. *Further assume that*

$$\|\nabla f\|_\infty + \|G(f)\|_\infty + \|\nabla \eta\|_\infty \le C, \tag{4.18}$$

then we have the following estimate:

$$\begin{aligned}
&\left|B_\delta(G(f); f, \xi) - B_\delta(G(f^h); \Pi_p f, \xi)\right| \\
&\quad \le \frac{1}{8}|||\xi|||^2 + C\int_{\partial\Omega\times I} \eta^2 |G(f^h)\cdot \mathbf{n}|\, d\nu\, ds + C\sum_{K\in\mathcal{K}_h}\Big[h_K^{-1}(\|\eta\|_K^2 + (\|\eta_t\|_K \\
&\qquad + \|\nabla\eta\|_K)^2) + h_K(\|\xi\|_K + \|\eta\|_K)^2\Big] + C(\|\xi\|_{\Omega_T} + \|\eta\|_{\Omega_T})\|\xi\|_{\Omega_T} + \sum_{m=1}^M |\eta_-|_m^2.
\end{aligned}$$

Remark 4.2. For a justification of the assumption (4.18) we could follow the idea in Ref. 15, where by rearranging the nonlinear term a Eulerian type system occurs and (4.18) arises naturally. Also numerical simulations for instance described in Ref. 9, 10, 18 and 23, point out that the nonlinear effects become secondary when the Fokker–Planck diffusion takes over. This phenomenon becomes more pronounced for larger diffusion coefficients. To avoid (4.18) altogether, one can follow the more involved mollifying procedure in Ref. 22.

Proof. Using the definition of η and ξ we write

$$\begin{aligned}
&B_\delta(G(f); f, \xi) - B_\delta(G(f^h); \Pi_p f, \xi) \\
&\quad = B_\delta(G(f^h); \eta, \xi) + B_\delta(G(f); f, \xi) - B_\delta(G(f^h); f, \xi) := T_1 + T_2 - T_3.
\end{aligned}$$

Now we estimate the terms T_1 and $T_2 - T_3$ separately. Starting with T_1, we have

$$\begin{aligned}
T_1 &= B_\delta(G(f^h); \eta, \xi) \\
&= \left(\eta_t, \xi\right)_{\Omega_T} + \langle \eta_+, \xi_+\rangle_0 - \sigma\sum_{K\in\mathcal{K}_h}\delta_K(\Delta_v\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K
\end{aligned}$$

$$+ \sum_{K\in\mathcal{K}_h} \delta_K(\eta_t + G(f^h)\cdot\nabla\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K + \sigma\big(\nabla_v\eta, \nabla_v\xi\big)_{\Omega_T}$$
$$+ \sum_{m=1}^{M-1} \langle[\eta], \xi_+\rangle_m + \sum_{m=0}^{M-1} (G(f^h)\cdot\nabla\eta, \xi)_m + \langle\eta_+, \xi_+\rangle_{\Lambda^-} + \langle\eta_-, \xi_-\rangle_{\Lambda^+}.$$

From the inverse inequality and the assumptions on σ and δ_K we have the estimates:

$$\begin{aligned}\sigma\Big|(\nabla_v\eta, \nabla_v\xi)_K\Big| &\le \sigma\|\nabla_v\eta\|_K\|\nabla_v\xi\|_K \le Ch_K^{-1}\|\eta\|_K\sigma\|\nabla_v\xi\|_K \\ &\le Ch_K^{-1}\|\eta\|_K^2 + \frac{1}{8h_K}\sigma^2\|\nabla_v\xi\|_K^2 \le Ch_K^{-1}\|\eta\|_K^2 + \frac{\sigma}{8}\|\nabla_v\xi\|_K^2 \quad (4.19)\end{aligned}$$

and

$$\begin{aligned}&\delta_K\sigma|(\Delta_v\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K| \\ &\quad\le \delta_K\sigma\|\Delta_v\eta\|_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K \\ &\quad\le C_I\delta_K\sigma h_K^{-2}\|\eta\|_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K \le C_I\delta_K h_K^{-1}\|\eta\|_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K \\ &\quad\le C_I\sqrt{\delta_K h_K^{-1}}\Big[h_K^{-1}\|\eta\|_K^2 + \frac{\delta_K}{8}\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2\Big] \\ &\quad\le C_p\Big[h_K^{-1}\|\eta\|_K^2 + \frac{\delta_K}{8}\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2\Big], \qquad (4.20)\end{aligned}$$

where by assumption on δ_K we have $C_p := C_I\sqrt{\delta_K h_K^{-1}} \le p_K^{-1/2}$. Then integrating by parts on the remaining terms, using (4.14), and a similar argument as in the proof of Proposition 4.1 we get,

$$\begin{aligned}&\sum_{K\in\mathcal{K}_h} (\eta_t + G(f^h)\cdot\nabla\eta, \xi + \delta_K(\xi_t + G(f^h)\cdot\nabla\xi))_K \\ &\qquad + \sum_{m=1}^{M-1} \langle[\eta], \xi\rangle_m + \langle\eta_+, \xi_+\rangle_0 + \langle\eta_+, \xi_+\rangle_{\Lambda^-} + \langle\eta_-, \xi_-\rangle_{\Lambda^+} \\ &\quad = -\big(\eta, \xi_t + G(f^h)\cdot\nabla\xi\big)_{\Omega_T} + \langle\eta_-, \xi_-\rangle_M + C\int_{\partial\Omega\times I} \eta\xi\big|G^h\cdot\mathbf{n}\big|\,d\nu\,ds \\ &\qquad - \sum_{m=1}^{M-1} \langle\eta_-, [\xi]\rangle_m + \sum_{K\in\mathcal{K}_h} \delta_K(\eta_t + G(f^h)\cdot\nabla\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K\end{aligned}$$

which using Cauchy–Schwartz inequality together with (4.19) and (4.20) gives

$$\begin{aligned}|T_1| \le \frac{1}{8}|||\xi|||^2 + C\Big[&\int_{\partial\Omega\times I} \eta^2\big|G^h\cdot\mathbf{n}\big|\,d\nu\,ds + \sum_{K\in\mathcal{K}_h} h_K^{-1}\|\eta\|_K^2 \\ &+ \sum_{m=1}^{M} |\eta_-|_m^2 + \sum_{K\in\mathcal{K}_h} \delta_K\|\eta_t + G(f^h)\cdot\nabla\eta\|_K^2\Big]. \qquad (4.21)\end{aligned}$$

By basic properties on solution of Poisson equation and the definition of G we have

$$\|G(f^h) - G(f)\|_{\Omega_T} \le C\|f - f^h\|_{\Omega_T} \le C(\|\xi\|_{\Omega_T} + \|\eta\|_{\Omega_T}).$$

Using this relation we now bound the last term on the right-hand side of (4.21), (see Ref. 1 for details),

$$\begin{aligned}\|\eta_t + G(f^h)\cdot\nabla\eta\|_{\Omega_T} &\le \|\eta_t\|_{\Omega_T} + \|G(f)\|_\infty\|\nabla\eta\|_{\Omega_T}\\ &\quad + C\|\nabla\eta\|_\infty(\|\xi\|_{\Omega_T} + \|\eta\|_{\Omega_T}).\end{aligned} \tag{4.22}$$

To estimate the term $(T_2 - T_3)$, we follow a similar argument as in Ref. 1 and get

$$\begin{aligned}|T_2 - T_3| &\le C(\|\xi\|_{\Omega_T} + \|\eta\|_{\Omega_T})\|\nabla f\|_\infty\|\xi\|_{\Omega_T}\\ &\quad + C\|\nabla f\|_\infty^2 \sum_{K\in\mathcal{K}_h} h_K(\|\xi\|_K + \|\eta\|_K)^2 + \frac{1}{8}\sum_{K\in\mathcal{K}_h}\delta_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2.\end{aligned} \tag{4.23}$$

Now combining the estimates (4.21)–(4.23), using assumption (4.18) and hiding the term $\frac{1}{8}\sum_{K\in\mathcal{K}_h}\delta_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2$ in $|||\xi|||$, the proof is complete. □

Proposition 4.3. *Under the assumptions of Proposition* 4.2 *we have*

$$\left|J_\delta(f^h,\xi) - J_\delta(f,\xi)\right| \le \frac{1}{8}|||\xi|||^2 + C\|\xi\|_{\Omega_T}^2 + C\sum_{K\in\mathcal{K}_h} h_K^{-1}\|\eta\|_K^2.$$

Proof. Using the definition of ξ and η, we have the identity

$$J_\delta(f^h,\xi) - J_\delta(f,\xi) = J_\delta(\xi,\xi) - J_\delta(\eta,\xi) := J_1 - J_2.$$

Below, we bound the terms J_1 and J_2 separately. For the first term J_1, using integration by parts, boundedness of Ω_v and the fact that $\xi \equiv 0$ on $\partial\Omega\times(0,T)$, ($\xi$ is the difference of two functions in $\tilde{V}_h^{\mathbf{q}}$, which are coinciding at the inflow boundary $\Gamma^-\times(0,T)$ with the nodal interpolant $\tilde{w}$ of w and at the outflow boundary $\Gamma^+\times(0,T)$ with the continuous extension of $\tilde{w}$, or the projection of f on the finite element space), we can easily show that

$$\begin{aligned}|J_1| &= \left|\sum_{K\in\mathcal{K}_h}(\nabla_v\cdot(\beta v\xi), \xi + \delta_K(\xi_t + G(f^h)\cdot\nabla\xi))_K\right|\\ &\le \beta\sum_{K\in\mathcal{K}_h}\left|(d\xi + v\cdot\nabla_v\xi, \xi + \delta_K(\xi_t + G(f^h)\cdot\nabla\xi))_K\right|\\ &\le C\beta d\|\xi\|_{\Omega_T}^2 + \beta\sum_{K\in\mathcal{K}_h}[|v|_{L_\infty(K)}^2\|\nabla_v\xi\|_K^2 + \delta_k^2\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2]\\ &\le C\beta d\|\xi\|_{\Omega_T}^2 + \beta\sum_{K\in\mathcal{K}_h}[h_K^2\|\nabla_v\xi\|_K^2 + \delta_k^2\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2].\end{aligned}$$

The term J_2 is estimated using the integration by parts, boundedness of Ω_v^h, and that $\xi \equiv 0$ on $\partial\Omega \times (0, T)$, viz.

$$\begin{aligned}
|J_2| &= \left| \sum_{K\in\mathcal{K}_h} (\nabla_v \cdot (\beta v \eta), \xi + \delta_K(\xi_t + G(f^h)\cdot\nabla\xi))_K \right| \\
&= \beta \left| \big(d\eta + v\cdot\nabla_v\eta, \xi\big)_{\Omega_T} + \sum_{K\in\mathcal{K}_h} \delta_K(d\eta + v\cdot\nabla_v\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K \right| \\
&= \beta \left| d\big(\eta, \xi\big)_{\Omega_T} - \big(\eta, v\cdot\nabla_v\xi\big)_{\Omega_T} + \sum_{K\in\mathcal{K}_h} \delta_K(d\eta + v\cdot\nabla_v\eta, \xi_t + G(f^h)\cdot\nabla\xi)_K \right| \\
&\le \beta(d+1)\delta^{-1}\|\eta\|_{\Omega_T}^2 + \frac{\beta}{4}\delta(\|\xi\|_{\Omega_T}^2 + |v|_\infty^2\|\nabla_v\xi\|_{\Omega_T}^2) \\
&\quad + \beta \sum_{K\in\mathcal{K}_h} \delta_K(d\|\eta\|_K^2 + |v|_{L_\infty(K)}^2\|\nabla_v\eta\|_K^2 + \frac{1}{2}\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2) \\
&\le C\beta[\delta^{-1}\|\eta\|_{\Omega_T}^2 + \delta\|\xi\|_{\Omega_T}^2 + C_v\delta\|\nabla_v\xi\|_{\Omega_T}^2 \\
&\quad + \sum_{K\in\mathcal{K}_h} \delta_K(\|\eta\|_K^2 + C_v\|\eta\|_{1,K}^2 + \|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2)],
\end{aligned}$$

where $\delta = \max_K \delta_K$. Combining these two estimates, recalling the assumption on β, and δ_K and hiding the terms $\sum_{K\in\mathcal{K}_h} \delta_K\|\xi_t + G(f^h)\cdot\nabla\xi\|_K^2$ and $\sum_{K\in\mathcal{K}_h} \delta_K\|\nabla_v\xi\|_K^2$ in $|||\xi|||^2$ we get the desired result. □

Note that in the above estimate for J_2 we may use the element-size and inverse estimate to write $|v|_{L_\infty(K)}^2\|\nabla_v\eta\|_K^2 \le h_K^2 h_K^{-2}\|\eta\|_K^2$. Thus, in the last step, we can replace $\|\eta\|_{1,K}^2$ by $h_K^2 h_K^{-2}\|\eta\|_K^2 = \|\eta\|_K^2$ and hence get a gradient-free estimate.

We will now derive a stability estimate underlying our main convergence result.

Lemma 4.1. *For ξ and η as above, there exist a constant $C > 0$ such that,*

$$|||\xi|||^2 \le C\left[\int_{\partial\Omega\times I} \eta^2\big|G^h\cdot\mathbf{n}\big|\,d\nu\,ds + \delta^{-1}\|\eta\|_{\Omega_T}^2 + \sum_{m=1}^M |\eta_-|_m^2 + \delta\|\eta\|_{1,Q}^2 + \sum_{m=1}^M |\xi_-|_m^2\delta\right].$$

Proof. The exact solution f satisfies (4.13), i.e.

$$B_\delta(G(f); f, u) - J_\delta(f, u) = L_\delta(u) \quad \forall\, u \in V^{\mathbf{q}}.$$

The coercivity result: Proposition 4.1 yields

$$\begin{aligned}
\frac{1}{2}|||\xi|||^2 &\le B_\delta(G(f^h); f^h - \Pi f, \xi) = L_\delta(\xi) + J_\delta(f^h, \xi) - B_\delta(G(f^h); \Pi f, \xi) \\
&= B_\delta(G(f); f, \xi) - B_\delta(G(f^h); \Pi f, \xi) + J_\delta(f^h, \xi) - J_\delta(f, \xi) \\
&:= \Delta B_\delta + \Delta J_\delta.
\end{aligned}$$

Now we use Propositions 4.2 and 4.3 to bound the terms ΔB_δ and ΔJ_δ. Further using the second result in Proposition 4.1 we estimate $\|\xi\|^2_{\Omega_T}$ and $\|\eta\|^2_{\Omega_T}$ with sufficiently large C_1. Combining all these estimates we obtain the desired result and the proof is complete. □

4.2. *Convergence*

We now put together all of the previously established results and prove our main convergence estimate. Recalling our previous notation $e := f - f^h = f - \Pi_p f + \Pi_p f - f^h := \eta - \xi$, we show that:

Theorem 4.1. *If* $f^h \in \tilde{V}^{\mathbf{q}}$ *satisfies* (4.13) *and* $\delta_K = \min\left(\frac{h_K^2}{\sigma C_I^2}, \frac{h_K}{p_K C_I}\right)$ *for each* $K \in \mathcal{T}$, *then there is a constant* $C > 0$ *such that,*

$$|||f - f^h|||^2 \leq C \sum_{K \in \mathcal{K}_h} h_K^{2s_k+1} \frac{\Phi(p_K, s_K)}{p_K} \|\hat{f}\|^2_{s_K+1,\hat{K}}, \tag{4.24}$$

where $\Phi(p_K, s_K) = \max(\Phi_1(p_K, s_K), \Phi_2(p_K, s_K))$ *as defined in Theorem* 3.2.

Proof. We split the right-hand side of the estimation in Lemma 4.1 and rewrite it concisely as

$$|||\xi|||^2 \leq C(A_1 + A_2 + A_3), \tag{4.25}$$

with

$$A_1 := \sum_{K \in \mathcal{K}_h} \delta_K^{-1} \|\eta\|_K^2 + \delta \|\eta\|^2_{1,\Omega_T},$$

$$A_2 := \int_{\partial\Omega \times I} \eta^2 \left|G^h \cdot \mathbf{n}\right| d\nu\, ds + \sum_{m=1}^{M} |\eta_-|_m^2,$$

$$A_3 := \sum_{m=1}^{M} |\xi_-|_m^2 \sum_{K \in \mathcal{K}_{h,m}} h_K \delta_K.$$

Below we estimate each A_i separately: As for A_1 we have using Theorem 3.2,

$$A_1 \leq \sum_{K \in \mathcal{K}} \left(\frac{h_K}{2}\right)^{2s_K} \Phi(p_K, s_K)(\delta_K^{-1} h_K^2 + \delta_K) \|\hat{f}\|^2_{s_K+1,\hat{K}}. \tag{4.26}$$

To get an estimate for A_2 we use trace estimate combined with an inverse inequality,

$$\|\eta\|^2_{\partial K} \leq C(\|\nabla\eta\|_K \|\eta\|_K + h_K^{-1}\|\eta\|_K^2), \quad \forall\, K \in \mathcal{M}_R, \tag{4.27}$$

which gives,

$$\begin{aligned} A_2 \leq C \sum_{K \in \mathcal{K}} \Bigg[&\left(\frac{h_K}{2}\right)^{s_K} \Phi_2^{1/2}(p_K, s_K) \left(\frac{h_K}{2}\right)^{s_K+1} \Phi_1^{1/2}(p_K, s_K) \\ &+ h_K^{-1} \left(\frac{h_K}{2}\right)^{2s_K+2} \Phi_1(p_K, s_K)\Bigg] \|\hat{f}\|^2_{s_K+1,\hat{K}}, \end{aligned} \tag{4.28}$$

where both terms in A_2 are estimated combining (4.27) with Theorem 3.2 and using $\delta_K = \min\left(\frac{h_K^2}{\sigma C_I^2}, \frac{h_K}{p_K C_I}\right)$. Summing up we can now rewrite (4.25) as,

$$|||\xi|||^2 \le C\left[\sum_{K\in\mathcal{K}} h_K^{2s_k+1}\frac{\Phi(p_K,s_K)}{p_K}\|\hat{f}\|^2_{s_K+1,\hat{K}} + \sum_{m=1}^{M}|\xi_-|_m^2\sum_{K\in\mathcal{K}_{h,m}} h_K\delta_K\right]. \quad (4.29)$$

To proceed, we need to estimate also the A_3 term. For this approach, we use the following discrete Grönwall's type estimate as, e.g. in Ref. 1: If

$$y(\cdot,t_m) \le C + C_1\sum_{j\le m}|y(\cdot,t_j)|\sum_{K\in\mathcal{K}_{h,m}} h_K\delta_K, \quad (4.30)$$

then $y(t_m) \le Ce^{C_1 t} \le Ce^{C_1 T}$. Note that (4.29) also implies,

$$|\xi_-|_m^2 \le C\left[\sum_{K\in\mathcal{K}} h_K^{2s_k+1}\frac{\Phi(p_K,s_K)}{p_K}\|\hat{f}\|^2_{s_K+1,\hat{K}} + \sum_{m=1}^{M}|\xi_-|_m^2\sum_{K\in\mathcal{K}_{h,m}} h_K\delta_K\right], \quad (4.31)$$

which gives using (4.30), (where we interpret the term under $\sum_{\mathcal{K}}$ as a new constant depending on f, $\mathcal{K}$ and $\mathbf{q}$), that

$$|\xi_-|_m^2 \le C\sum_{K\in\mathcal{K}} h_K^{2s_k+1}\frac{\Phi(p_K,s_K)}{p_K}\|\hat{f}\|^2_{s_K+1,\hat{K}}e^{CT}. \quad (4.32)$$

Thus we now also have an estimate for A_3, which together with (4.26), (4.28), $\delta|_K := \delta_K$, gives the desired result. See also Refs. 1 and 15. □

Remark 4.3. One can show that the convergence rate (4.24): $h_K^{2s_k+1}\frac{\Phi(p_K,s_K)}{p_K}$ is indeed of order $\delta_K^{2s_K+1}$. However, this remaining part is basically, similar to the type of estimates derived in Ref. 12 in their full details and therefore are omitted.

Acknowledgments

This work was partially supported by the EU contract ERB FMRX CT97 0157.

References

1. M. Asadzadeh, Streamline diffusion methods for the Vlasov–Poisson equation, *Math. Mod. Numer. Anal.* **24** (1990) 177–196.
2. M. Asadzadeh, Streamline diffusion methods for Fermi and Fokker–Planck equations, *Transp. Th. Statist. Phys.* **26** (1997) 319–340.
3. M. Asadzadeh, A posteriori error estimates for the Fokker–Planck and Fermi pencil beam equations, *Math. Mod. Meth. Appl. Sci.* **10** (2000) 737–769.
4. M. Asadzadeh and P. Kowalczyk, Convergence analysis of the streamline diffusion and discontinuous Galerkin methods for the Vlasov–Fokker–Planck system, *Numer. Meth. Partial Diff. Eqn.* **21** (2005) 472–495.
5. I. Babuska, C. E. Baumann and J. T. Oden, A discontinuous *hp*-finite element method for diffusion problems: 1-D analysis, *Comp. Math.* **37** (1999) 103–122.

6. M. S. Baouendi and P. Grisvard, Sur une équation d'évolution changeant de type, *J. Funct. Anal.* **2** (1968) 352–367.
7. F. Bouchut, Existence and Uniqueness of a global smooth solution for the Vlasov–Poisson–Fokker–Planck system in three dimensions, *J. Funct. Anal.* **111** (1993) 239–258.
8. P. Degond, Global existence of smooth solutions for the Vlasov–Fokker–Planck equation in 1 and 2 space dimensions, *Ann. Sci. Éc. Norm. Sup.* 4^e *sér.* **19** (1986) 519–542.
9. K. J. Havlak and H. D. Victory, Jr., The numerical analysis of random particle methods applied to Vlasov–Poisson–Fokker–Planck kinetic equations, *SIAM J. Numer. Anal.* **33** (1996) 291–317.
10. K. J. Havlak and H. D. Victory, Jr., On deterministic particle methods for solving Vlasov–Poisson–Fokker–Planck kinetic systems, *SIAM J. Numer. Anal.* **35** (1998) 1473–1519.
11. P. Houston, C. Schwab and E. Süli, Stabilized *hp*-finite element methods for first order hyperbolic problems, *SIAM J. Numer. Anal.* **37** (2000) 1618–1643.
12. P. Houston and E. Süli, Stabilized *hp*-finite element approximation of partial differential equations with nonnegative characteristic form, *Computing* **66** (2001) 99–119.
13. P. Houston, C. Schwab and E. Süli, Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems, *SIAM J. Numer. Anal.* **39** (2002) 2133–2163.
14. T. J. Hughes, E. T. Tezduyar and A. Broks, Streamline upwind formulation for advection-diffusion, Navier–Stokes and first order hyperbolic equations, in *Fourth Int. Conf. on Finite Element Methods in Fluids*, Tokyo, July (1982).
15. C. Johnson and J. Saranen, Streamline diffusion methods for the incompressible Euler and Navier–Stokes equations, *Math. Comp.* **47** (1986) 1–18.
16. J. L. Lions, *Equations Différentielles Opérationnelle et Problm̀es aux Limites* (Springer, 1961).
17. J. L. Lions, *Quelques Méthods de Résolution des Problèm aux Limites non Linéaires* (Dunod, 1969).
18. J. Schaeffer, Convergence of a difference scheme for the Vlasov–Poisson–Fokker–Planck system in one dimension, *SIAM J. Numer. Anal.* **35** (1998) 1149–1175.
19. D. Schötzau and C. Schwab, Time discretization of parabolic problems by the *hp*-version of the discontinuous Galerkin finite element methods, *SIAM J. Numer. Anal.* **38** (2000) 837–875.
20. C. Schwab, *p*- and *hp*-finite element methods, in *Theory and Applications in Solid and Fluid Mechanics* (Oxford, 1998).
21. H. D. Victory, Jr. and B. P. O'Dwyer, On classical solutions of Vlasov–Poisson–Fokker–Planck Systems, *Indiana Univ. Math. J.* **39** (1990) 105–156.
22. J. Weckler, On the initial-boundary-value problem for the Vlasov–Poisson system: Existence of weak solutions and stability, *Arch. Rational Mech. Anal.* **130** (1995) 145–161.
23. S. Wollman, Numerical approximation of the Vlasov–Poisson–Fokker–Planck system in one dimension, *J. Comput. Phys.* **202** (2005) 602–644.