

# Hypocoercivity-compatible finite element methods for the long-time computation of Kolmogorov's equation

Emmanuil H. Georgoulis <sup>\*†‡</sup>

December 18, 2020

## Abstract

This work is concerned with the development of a family of Galerkin finite element methods for the classical Kolmogorov equation. Kolmogorov's equation serves as a sufficiently rich, for our purposes, model problem for kinetic-type equations and is characterised by diffusion in *one* of the two (or three) spatial directions only. Nonetheless, its solution admits typically decay properties to some long time equilibrium, depending on closure by suitable boundary/decay-at-infinity conditions. A key attribute of the proposed family of methods is that they also admit similar decay properties at the (semi)discrete level for very general families of triangulations. The method construction uses ideas by the general theory of *hypocoercivity* developed by Villani [24], along with judicious choice of numerical flux functions. These developments turn out to be sufficient to imply that the proposed finite element methods admit a priori error bounds with constants *independent of the final time*, despite Kolmogorov equation's degenerate diffusion nature. Thus, the new methods provably allow for robust error analysis for final times tending to infinity. The extension to three spatial dimensions is also briefly discussed.

## 1 Introduction

Degenerate parabolic problems in kinetic modelling are often characterised by the explicit presence of diffusion/dissipation in *some* of the spatial directions only, yet may still admit decay properties to some long time equilibrium. The development of Galerkin finite element methods for such problems satisfying provably also such decay properties is, remarkably, an unexplored area. This is despite the potential attraction that general, possibly highly non-uniform or adaptive, grid approximations can bring to kinetic simulations, especially those concerning expansive long-time simulation phenomena. Aiming to contribute in this direction, we shall develop and analyse a new class of Galerkin finite element methods for a simple kinetic equation, equipped with suitable closures by known boundary/periodicity conditions from the literature. To this end, we shall be concerned with the classical Kolmogorov partial differential equation (PDE)

$$\mathcal{L}u \equiv u_t - u_{xx} + xu_y = f, \quad \text{in } (0, \infty) \times \Omega, \quad (1)$$

---

<sup>\*</sup>School of Mathematics and Actuarial Science, University of Leicester, Leicester LE1 7RH, United Kingdom.  
Email: [Emmanuil.Georgoulis@le.ac.uk](mailto:Emmanuil.Georgoulis@le.ac.uk)

<sup>†</sup>Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Zografou 15780, Greece

<sup>‡</sup>Institute of Applied and Computational Mathematics, Foundation of Research and Technology – Hellas, Vassilika Vouton 700 13 Heraklion, Crete, Greece

with  $\Omega \subset \mathbb{R}^2$ , for suitably smooth forcing  $f \in L_2(\mathbb{R}_+; H^1(\Omega))$ , and Cauchy initial data  $u(0, \cdot) = u_0 \in L_2(\Omega)$ . The notational conventions used above are somewhat non-standard in kinetic theory, whereby  $\mathcal{L}$  may be written as  $\mathcal{L}f := f_t - f_{vv} + vf_x$  with  $v$  denoting the particle velocity variable,  $x$  the displacement/position and  $f$  the respective probability density function. Kolmogorov showed already in 1934 that (1) admits a smooth fundamental solution outside the pole and, hence, it is hypoelliptic, i.e., it admits a smooth solution in  $(0, \infty) \times \mathbb{R}^2$  for any smooth initial and forcing data  $u_0, f$ , respectively [17]. This may be a surprising assertion at first sight, since there is no explicit dissipation built into the PDE in the  $y$ -variable. The advection field  $(0, x)^T$ , however, is “appropriately non-constant” so that it suffices to propagate the built-in dissipation in the  $x$ -variable onto the entire spatial domain. In 1967, Hörmander, in the celebrated work [15], gave a sufficient condition for hypoellipticity for a wide class of 2nd order operators with non-negative characteristic form, which are nowadays often referred to as *Hörmander sum-of-squares operators*; interestingly, a motivating example for the developments in [15] has, indeed, been Kolmogorov’s equation (1). Hence, although (1) may appear to be a rather special equation at first sight, its significance as a model problem is paramount, for it encompasses a number of pertinent structural properties of large classes of kinetic models; we refer, e.g., to [24] for very instructive expositions.

The proof of trend to equilibrium for the directly related inhomogeneous Fokker-Planck equation has been given by Hérau & Nier in [14] upon realising that certain entropies admitting mixed derivatives ( $u_{xy}$  or, respectively,  $f_{vx}$ ) give rise to full gradients in certain weighted Sobolev spaces, from which Poincaré-Friedrichs inequalities (also known as spectral gap properties in this context) are sufficient to prove decay to an equilibrium distribution. Related ideas have also been used by Eckmann & Hairer in [9] to study the spectral properties of certain hypoelliptic operators involving degenerate diffusions and by Mouhot & Neumann [19] in the study of kinetic models with integral-type collision operators, among others. The idea of using entropies involving mixed derivatives was elevated to a general framework in proving decay to equilibrium for kinetic equations by Villani [24] via the introduction of the concept of *hypocoercivity*. Roughly speaking, hypocoercivity is the property of certain degenerate elliptic or parabolic differential operators to yield dissipation of the solution also in the directions where *no* diffusion is explicitly present. Astonishingly, sufficient conditions for hypocoercivity are given by the rank spanned by Hörmander vector fields and their commutators with respect to the vector field related to 1st order term [24]. Thus, sufficient conditions for hypocoercivity are closely related, but *not* identical, to the classical Hörmander’s rank condition for hypoellipticity.

To fix ideas, denoting by  $(\cdot, \cdot)_{L_2}$  the inner product of  $L_2$ , over some  $\Omega \subset \mathbb{R}^2$ , and respective norm  $\|\cdot\|_{L_2} := \sqrt{(\cdot, \cdot)_{L_2}}$  and, assuming suitable boundary conditions, the classical energy-type analysis for (1) yields

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|_{L_2}^2 + \|u_x\|_{L_2}^2 = (f, u)_{L_2}, \quad (2)$$

for almost all  $t \in (0, t_f]$ , for some final time  $t_f \in \mathbb{R}_+$ , upon observing that the skew-symmetric term  $(xu_y, u)$  vanishes. A combination of Cauchy-Schwarz’ and Young’s inequalities on the last term on the right-hand side of (2) thus yields

$$\int_0^{t_f} (f, u)_{L_2} dt \leq \frac{1}{2\delta} \|f\|_{L_2(0, t_f; L_2)}^2 + \frac{\delta}{2} \|u\|_{L_2(0, t_f; L_2)}^2, \quad (3)$$

for any  $\delta > 0$ . Therefore, Grönwall’s Lemma implies

$$\|u(t_f)\|_{L_2}^2 \leq e^{\delta t_f} (\delta^{-1} \|f\|_{L_2(0, t_f; L_2)}^2 + \|u_0\|_{L_2}^2). \quad (4)$$

Since the right-hand side of (4) is always greater than or equal to  $\|u_0\|_{L_2}^2$  for every  $f \in L_2(0, \infty; L_2)$ , we conclude that the basic energy estimate may be severely pessimistic as  $t_f \rightarrow \infty$

for any fixed  $\delta > 0$ : for any finite final time  $t_f > 0$ , the above implies that the right-hand side of (4) will be positive away from zero for any non-trivial initial condition. Notice that when  $f = 0$ , we can take  $\delta \rightarrow 0$ , yielding the stability estimate  $\|u(t_f)\|_{L_2} \leq \|u_0\|_{L_2}$ . This is at odds with the fast decay to 0 observed for this problem by the solution of Kolmogorov's equation when equipped with suitable conditions, e.g., periodic boundary conditions or decay-to-infinity conditions when  $\Omega = \mathbb{R}^2$ . Evidently the key culprit in the above analysis is the absence of control in a time-integrated norm on the left-hand side of (4). The situation does not appear to improve if different versions of Grönwall-type inequalities are applied instead.

On the other hand, for the non-degenerate parabolic equation

$$\tilde{\mathcal{L}}u \equiv u_t - u_{xx} - u_{yy} + xu_y = f, \quad \text{in } (0, \infty) \times \Omega, \quad (5)$$

with the same forcing and initial conditions, we can easily deduce the energy identity

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|_{L_2}^2 + \|\nabla u\|_{L_2}^2 = (f, u)_{L_2}. \quad (6)$$

Assuming now the validity of a Poincaré-Friedrichs inequality of the form  $\|u\|_{L_2}^2 \leq C_{PF} \|\nabla u\|_{L_2}^2$ , and working as before, we can arrive at the estimate

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|_{L_2}^2 + C_{PF}^{-1} \|u\|_{L_2}^2 \leq (f, u)_{L_2}, \quad (7)$$

from which, Grönwall's Lemma implies

$$\|u(t_f)\|_{L_2}^2 \leq e^{-C_{PF}^{-1}t_f} \|u_0\|_{L_2}^2 + C_{PF} \int_0^{t_f} e^{-C_{PF}^{-1}(t_f-t)} \|f(t)\|_{L_2}^2 dt, \quad (8)$$

whose right-hand side clearly decays exponentially as  $t_f \rightarrow \infty$  when  $f = 0$ , for instance. This is in sharp contrast with (4).

A key idea in the context of hypocoercivity, on the other hand, is to arrive at a modified energy setting whereby Poincaré-Friedrichs inequality is valid for the *original* problem (1). In our specific setting of Kolmogorov's equation, upon considering carefully constructed families of Hilbertian norms  $\sqrt{(\cdot, \cdot)_{\text{hc}}} =: \|\cdot\|_{\text{hc}} := (\|\cdot\|_{L_2}^2 + \|\sqrt{A} \nabla \cdot\|_{L_2}^2)^{1/2}$ , where  $A \in \mathbb{R}_{\text{symm}}^{2 \times 2}$  are suitable positive semidefinite matrices, it is possible to arrive at an energy estimate of the form

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|_{\text{hc}}^2 + c(\|\nabla u\|_{L_2}^2 + \|u_x\|_{\text{hc}}^2) \leq (f, u)_{\text{hc}}, \quad (9)$$

for suitable/compatible boundary conditions, with  $c$  a positive constant. Crucially, if  $u$  is such that it satisfies a Poincaré-Friedrichs inequality, we can arrive to an estimate of the form (8) for the original, degenerate Kolmogorov's equations. As a result, in the homogeneous case  $f = 0$  (8) implies exponential decay as  $t_f \rightarrow \infty$ .

Obviously, (4) is not a satisfactory state of affairs not only at the PDE level, but also under the prism of the numerical analysis for these problems. Indeed, almost all provably convergent, general numerical methods which can be potentially applied to this class of PDEs use standard energy-type arguments as in (4) for their error analysis, resulting to possibly severely pessimistic a priori and/or a posteriori error bounds for large final times  $t_f$ . Available Galerkin-type numerical methods for (degenerate) second order evolution problems of diffusion type typically also suffer in this respect; we non-exhaustively refer to [1, 16, 10, 4] as representative works presenting different approaches. A numerical approach accommodating instead the functional setting of (9), thereby taking advantage of hypocoercivity properties is, therefore, desirable.

To the best of our knowledge, there exist only few, yet quite inspiring, exceptions which employ hypocoercivity ideas in the context of numerical methods. In particular, the work of Porretta & Zuazua [21] proves the hypocoercivity of classical finite difference operators discretising Kolmogorov's equation for  $f = 0$  on uniform spatial grids over  $[0, \infty) \times \mathbb{R}^2$ . The recent manuscript by Dujardin, Hérau, & Lafitte [8] takes the approach of [21] further by establishing discrete versions of weighted Poincaré inequalities for difference operators, thereby showing decay to equilibrium for finite difference approximations of the inhomogeneous two-dimensional Fokker-Planck equation, which is, of course, closely related to Kolmogorov's equation. Both these approaches apply directly the abstract semigroup result of Villani [24, Theorem 18] to the finite difference operators used by proving the necessary commutator properties required for its validity. With regard to practical works, Foster, Lohéac & Tran [12] discuss the development of a Lagrangian-type splitting method based on a carefully constructed similarity transformation and linear finite elements over quasiuniform meshes, although no proof of decay to equilibrium is given for the numerical method itself. Also, Bessemoulin-Chatard & Filbet [2] present design principles for the construction of equilibrium-preserving finite volume methods for nonlinear degenerate parabolic problems. Recently, Bessemoulin-Chatard, Herda & Rey [3] presented an asymptotic-preserving in the diffusive limit finite volume scheme for a class of one-dimensional kinetic equations, using the hypocoercivity framework of Dolbeault, Mouhot & Schmeiser [6].

For practical computations, one is typically forced to confine the spatial computational domain into an open and bounded one, say  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$ . If the objective is to compute equilibrium states for the whole of  $\mathbb{R}^d$ , the truncation of the computational domain and the imposition of boundary conditions should be performed carefully, for it may alter the rate of decay to equilibrium compared to the non-confined problem [12]. If one is interested instead in computations on bounded domains as such, then the modelling often dictates the imposition of additional boundary conditions/constraints, such as solution periodicity or far field decay. Indeed, as hypocoercivity is manifested via certain properties of commutators of directional derivatives appearing in the differential operator, integration by parts is an essential tool. As a result when confined on bounded domains the very nature of boundary conditions enforced becomes essential. As we shall see below, the imposition of boundary conditions in a fashion allowing concepts of hypocoercivity to be ported to bounded spatial domains is a challenge both in the PDE and the numerical analysis contexts. Indeed, to the best of our knowledge *no* exhaustive study of hypocoercivity-inducing boundary conditions for problems posed on bounded spatial domains has been carried out in the respective PDE literature.

This work is concerned with the development of a Galerkin-type finite element method which is compatible with hypocoercivity concepts and, therefore, allows for the proof of error bounds whose error constants *do not* grow as the final time  $t_f \rightarrow \infty$ . This is in sharp contrast to estimates derived via standard energy arguments as discussed above; cf. (4). In particular, as we shall show below, there will be *no dependence* of the error constant on the final time  $t_f$  in the setting of (9). This, in turn, will highlight theoretically any potential advantages of the numerical methods presented below for long time computations. The construction of the finite element method is based on a variational interpretation of the modified entropies involving mixed derivatives of Hérau & Nier [14] and Villani [24], which are able to reveal the implicit diffusion properties in the  $y$ -variable also. The derived modified variational problem is based on the PDE and higher order laws stemming from the PDE and involving up to 4th order partial derivatives, when viewed formally in strong form. The finite element method is constructed via a Galerkin projection onto a suitable  $C^0$ -finite element space subordinate to a, possibly, highly non-uniform triangulation. The proposed method is, nevertheless, non-conforming due to the 4th order symbol of the variational problem. To recover, therefore, the consistency of the finite element method without compromising on the hypocoercivity properties, the variational

form has to be enriched by carefully constructed consistent numerical fluxes over the skeleton of the triangulation. The proposed finite element method is proven to decay to equilibrium when  $f = 0$ , when we equip the PDE with additional natural boundary conditions; this is, to the best of our knowledge, the first manifestation of hypocoercivity for a Galerkin scheme in the literature. Moreover, we prove a priori error bounds in the natural norm determined by the hypocoercivity result.

We view the developments presented below as proof of concept that Galerkin finite element methods over general non-uniform grids can be constructed in a compatible fashion to retain hypocoercivity properties at discrete level. This, in turn, leads to a priori error bounds which are *robust* with respect to their dependence on the final time  $t_f$ . Kolmogorov's equation has the role of a sufficiently rich model problem to highlight a new finite element methodology for hypocoercive PDEs on bounded computational domains with view to retaining the long-time behaviour of the exact initial/boundary value problems. Apart from the mathematical challenge, the development of numerical methods for kinetic equations on general, possibly non-quasi-uniform meshes has important practical ramifications. Indeed, modern applications of kinetic equations involve the movement of large particles for which Lagrangian schemes appear to be preferential by practitioners which greatly benefits from the ability to discretise over unstructured, possibly moving, meshes.

The remainder of this work is structured as follows. In Section 2, we introduce a variational formulation motivated by design concepts of hypocoercivity [24], which is shown to be coercive with respect to an  $H^1$ -equivalent norm in Section 3. A discrete bilinear form using continuous, yet non-conforming, elements with appropriately constructed numerical fluxes ensuring the coercivity in the discrete setting also is proposed in Section 4 and is shown to admit completely analogous decay to equilibrium properties as the respective continuous problem in the non-external forcing scenario. The error analysis of the proposed finite element method is given in Section 5 along with a discussion on its properties. Finally, we present a series of numerical experiments highlighting the predicted convergence rates by the theory; this is the content of Section 6.

To simplify notation, we shall abbreviate the  $L_2(\omega)$ -inner product and  $L_2(\omega)$ -norm for a Lebesgue-measurable subset  $\omega \subset \mathbb{R}^d$  as  $(\cdot, \cdot)_\omega$  and  $\|\cdot\|_\omega$ , respectively. Moreover, when  $\omega = \Omega$ , we shall further compress the notation to  $(\cdot, \cdot) \equiv (\cdot, \cdot)_\Omega$  and  $\|\cdot\| \equiv \|\cdot\|_\Omega$ . We shall also make use of the standard notation  $H^k(\omega)$  for Hilbertian Sobolev spaces,  $k \in \mathbb{R}$ .

## 2 A special weak formulation

The problem of closing a degenerate parabolic PDE with suitable boundary conditions is well understood via the classical theory of linear second order equations with non-negative characteristic form [11, 20]. In particular, with  $\mathbf{n}(\cdot) := (n_1(\cdot), n_2(\cdot))^T$  denoting the unit outward normal vector at almost every point of  $\partial\Omega$ , which is assumed to be piecewise smooth and Lipschitz, we first define the elliptic portion of the boundary

$$\partial_0\Omega := \{(x, y) \in \partial\Omega : n_1(x, y) \neq 0\}.$$

On the non-elliptic portion of the boundary  $\partial\Omega \setminus \partial_0\Omega$ , we define the inflow and outflow boundaries:

$$\partial_-\Omega := \{(x, y) \in \partial\Omega \setminus \partial_0\Omega : xn_2(x, y) < 0\}, \quad \partial_+\Omega := \{(x, y) \in \partial\Omega \setminus \partial_0\Omega : xn_2(x, y) \geq 0\}.$$

(For notational brevity, we have included any characteristic portions  $\{(x, y) \in \partial\Omega \setminus \partial_0\Omega : xn_2(x, y) = 0\}$  of the boundary into the outflow part, since their treatment is identical in what

follows.) Introducing the notation  $Lu \equiv -u_{xx} + xu_y$ , we consider the initial/boundary-value problem:

$$\begin{aligned} u_t + Lu &= u_t - u_{xx} + xu_y = f, & \text{in } (0, t_f] \times \Omega, \\ u &= u_0, & \text{on } \{0\} \times \Omega, \\ u &= 0, & \text{on } (0, t_f] \times \partial_{-,0}\Omega, \end{aligned} \quad (10)$$

with  $\partial_{-,0}\Omega := \partial_-\Omega \cup \partial_0\Omega$ , for  $t_f > 0$  and for  $f \in H^1(\Omega)$ , noting the non-standard regularity assumption for  $f$ . The well-posedness of the above problem is assured upon assuming that  $\partial_{-,0}\Omega$  has positive one-dimensional Hausdorff measure [20]. It is also possible to impose Neumann-type boundary conditions on parts of  $\partial_0\Omega$ ; this is not done here in the interest of simplicity of the presentation only. For a three-dimensional counterpart of the above model problem we refer to Remark 3.1.

Let  $A := \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$  be a symmetric and non-negative definite matrix with  $\alpha, \beta, \gamma$  non-negative parameters, and set

$$\delta := 1 - \frac{\beta}{\sqrt{\alpha\gamma}}; \quad (11)$$

the fact that  $A$  is non-negative definite implies that  $0 \leq \delta \leq 1$ . Assuming sufficient regularity for the exact solution  $u$ , so that the following calculations are well defined, (1) implies  $\nabla u_t + \nabla Lu = \nabla f$ , which, tested against  $A\nabla v$  for any  $v \in H^1(\Omega)$ , results in

$$(\nabla u_t, A\nabla v) + (\nabla Lu, A\nabla v) = (\nabla f, A\nabla v);$$

here and below, we denote  $(\cdot, \cdot) \equiv (\cdot, \cdot)_{L_2(\Omega)}$  and  $\|\cdot\| \equiv \|\cdot\|_{L_2(\Omega)}$  for brevity. Resorting to (1) once more, after an integration by parts, we deduce the variational form

$$(u_t, v) + (\nabla u_t, A\nabla v) + (u_x, v_x) + (xu_y, v) + (\nabla Lu, A\nabla v) = (f, v) + (\nabla f, A\nabla v), \quad (12)$$

for all  $v \in H_{-,0}^1(\Omega) := \{v \in H^1(\Omega) : v|_{\partial_{-,0}\Omega} = 0\}$ . The latter will be the basis of constructing a finite element method with the sought-after properties. To highlight the idea behind this construction in the context of hypocoercivity, we shall now discuss the properties of the spatial part of the operator on the left-hand side of (12), viz.,

$$a(u, v) := (u_x, v_x) + (xu_y, v) + (\nabla Lu, A\nabla v). \quad (13)$$

### 3 The hypocoercivity of $a(\cdot, \cdot)$

We begin by somewhat modifying the approach presented in Villani [24] arriving, nevertheless, to an effectively similar result. For accessibility, we shall first revert to (1), rather than (10), i.e., we have  $\Omega = \mathbb{R}^2$ , assuming that  $|u(x, y)| \rightarrow 0$  as  $|(x, y)| \rightarrow \infty$ ; the case  $\Omega \subset \mathbb{R}^2$  will be discussed below.

Setting  $v = u$  into (13) gives

$$a(u, u) = \|u_x\|^2 + (xu_y, u) + (-u_{xxx} + (xu_y)_x, \alpha u_x + \beta u_y) + (-u_{xxy} + xu_{yy}, \beta u_x + \gamma u_y),$$

which, upon integration by parts, already yields

$$\begin{aligned} a(u, u) &= \|u_x\|^2 + (u_{xx}, \alpha u_{xx} + \beta u_{yx}) + ((xu_y)_x, \alpha u_x + \beta u_y) \\ &\quad + (u_{xy}, \beta u_{xx} + \gamma u_{yx}) - (xu_y, \beta u_{xy} + \gamma u_{yy}) \\ &= \|u_x\|^2 + \alpha \|u_{xx}\|^2 + \beta \|u_{yy}\|^2 + \gamma \|u_{xy}\|^2 + \alpha (u_y, u_x) + 2\beta (u_{xy}, u_{xx}), \end{aligned}$$

noting that  $(xu_{xy}, u_x) = 0 = (xu_y, u_{yy})$  from the divergence theorem and the decay properties of  $u$  as  $|(x, y)| \rightarrow \infty$ . Young's and Cauchy-Schwarz' inequalities now imply

$$a(u, u) \geq (1 - \epsilon)\|u_x\|^2 + (\alpha - \sqrt{\beta})\|u_{xx}\|^2 + \left(\beta - \frac{\alpha^2}{4\epsilon}\right)\|u_y\|^2 + \left(\gamma - \beta^{\frac{3}{2}}\right)\|u_{xy}\|^2.$$

Selecting  $\alpha > 0$ ,  $\beta = 4\alpha^2/9$ ,  $\gamma = \alpha^3/3$ , and  $\epsilon = 3/4$ , for instance, we deduce

$$a(u, u) \geq \frac{1}{4}\|u_x\|^2 + \frac{\alpha}{3}\|u_{xx}\|^2 + \frac{\alpha^2}{9}\|u_y\|^2 + \frac{\alpha^3}{27}\|u_{xy}\|^2. \quad (14)$$

Another instructive choice is  $\alpha > 0$ ,  $\beta = \alpha^2$ ,  $\gamma = \alpha^3$  and  $\epsilon = 1/2$ , giving

$$2a(u, u) \geq \|u_x\|^2 + \alpha^2\|u_y\|^2. \quad (15)$$

Thus, remarkably, using the non-standard test function  $v - \nabla \cdot (A \nabla v)$  for well-chosen  $A$ , results into  $L$  providing coercivity with respect to  $u_y$  also! This is a manifestation of the concept of hypocoercivity. As briefly discussed above, the mechanics of this idea can be traced back to the celebrated work of Hörmander [15] for sufficient conditions for hypoellipticity for second order linear operators: the first order term is “appropriately non-constant” with respect to the independent variables  $t, x, y$ , allowing for control also in the commutator vector field  $[\partial_x, x\partial_y] = \partial_y$ , using the usual notation  $[A, B] := AB - BA$  for a commutator of an algebraic structure. Informally speaking, the first order vector field in conjunction with the direction of  $\partial_x$  is able to “propagate” the  $x$ -variable dissipation of  $L$  to the whole of  $\mathbb{R}^2$ , resulting to the property of hypocoercivity [24]. Notice that the choice of  $\alpha, \beta, \gamma$  yielding (14) corresponds to  $v - \nabla \cdot (A \nabla v)$  being an elliptic differential operator, while the respective choice for (15) corresponds to the parabolic operator  $v - \nabla \cdot (A \nabla v)$ . At this point, we prefer to remain general in the choice of  $A$ , very much following the point of view taken by Villani [24]. Some concrete choices of  $A$  will be needed below to fully define the ensuing numerical method.

The key second ingredient in proving solution behaviour as  $t_f \rightarrow \infty$  is the availability of a Poincaré-Friedrichs/spectral gap inequality. Such inequalities typically hold for functions in standard Lebesgue spaces on bounded domains or for periodic ones, along with some boundary or moment conditions. For problems posed on whole Euclidean spaces, hypocoercivity is typically manifested in appropriately weighted Lebesgue spaces [24]. In the present work, we are primarily interested in the mechanics of porting hypocoercivity into piecewise smooth Galerkin-type solutions; as such we shall assume that we have available compatible closure/boundary conditions for hypocoercivity to hold at the PDE level.

Boundary/closure conditions in kinetic PDE models posed on bounded domains  $\Omega \subset \mathbb{R}^d$  are designed to convey various physical properties in the vicinity of boundaries, e.g., reflection, diffusion, periodicity, etc. In the case of the classical Kolmogorov equation considered here, the  $x$ -variable is typically associated with particle velocity, while the  $y$ -variable models position/displacement. Performing the above calculations on a bounded domain  $\Omega$  instead, integrations by parts yield

$$\begin{aligned} a(u, u) &= \|u_x\|^2 + \frac{1}{2}\|\sqrt{xn_2}u\|_{\partial_+\Omega}^2 + \|\sqrt{A}\nabla u_x\|^2 + \beta\|u_y\|^2 + \alpha(u_y, u_x) \\ &\quad + \left(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u\right)_{\partial\Omega}. \end{aligned} \quad (16)$$

To arrive again at (14) (or (15)), it is sufficient to ensure that

$$\left(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u\right)_{\partial\Omega} \geq 0. \quad (17)$$

Alternatively, it may be possible to control part of  $(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u)_{\partial\Omega}$  by the non-negative terms on the right-hand side of (16), with the remaining part to satisfy a non-negativity condition.

We now give some specific examples. For illustration, we set  $\Omega := (x^-, x^+) \times (y^-, y^+)$ . Then,

$$\partial_0\Omega = \{(x^-, y), (x^+, y) : y^- \leq y < y^+\}, \quad \partial_{\pm}\Omega = \{(x, y^{\pm}) : x^- < x \leq x^+\}.$$

Further, we assume  $u \in H^{5/2+\varepsilon}(\Omega)$ -smoothness of the solution,  $\varepsilon > 0$  so that the last term on the right-hand side of (16) is well-defined; this is a reasonable assumption in this context due to the hypoellipticity of Kolmogorov's equation [15].

As a first example, we remove the homogeneous boundary conditions on  $\partial_{-,0}\Omega$  and, instead, we consider the family of sufficiently smooth functions on the torus  $\mathbb{T}$ . The continuity of partial derivatives implies that the last term on the right-hand side of (16), together with any additional boundary terms arising by the removal of the homogeneous boundary conditions, vanish.

As a second example, we restrict the family of solutions to periodic functions in the diffusive  $x$ -direction, and we further assume that  $u$  has vanishing inflow flux, viz.,  $u_y = 0$ . This gives

$$(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u)_{\partial_0\Omega} = 0, \quad (18)$$

since  $xn_2 = 0$  on  $\partial_0\Omega$ . Moreover, since  $n_1 = 0$  and  $u_y = 0$  on  $\partial_{\pm}\Omega$  and, observing that  $u_x = 0$  on  $\partial_{-}\Omega$  (as  $u_x$  is the tangential derivative on  $\partial_{-}\Omega$  of the *constant* initial condition), we deduce

$$(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u)_{\partial_{-}\Omega} = \frac{1}{2}(xn_2\nabla u, A\nabla u)_{\partial_{-}\Omega} = 0.$$

Therefore, we deduce

$$(\frac{1}{2}xn_2\nabla u - n_1\nabla u_x, A\nabla u)_{\partial\Omega} = (\frac{xn_2}{2}\nabla u, A\nabla u)_{\partial_{\pm}\Omega} = \frac{1}{2}\|\sqrt{|xn_2|}A\nabla u\|_{\partial_{+}\Omega}^2.$$

We refer, e.g., to [22, 8] for further discussion on additional boundary conditions. Different combinations of assumptions are also possible to yield (17), such as vanishing second derivative conditions across the elliptic boundary, etc.

**Remark 3.1.** *To highlight the potential generality of the approach, we briefly consider the three-dimensional version of Kolmogorov's equation: for  $\Omega \subset \mathbb{R}^3$ , find  $u : (0, t_f] \times \Omega \rightarrow \mathbb{R}$ , such that*

$$\begin{aligned} u_t + L_3u &= u_t - u_{xx} + xu_y + yu_z = f, & \text{in } (0, t_f] \times \Omega, \\ u &= u_0, & \text{on } \{0\} \times \Omega, \\ u &= 0, & \text{on } (0, t_f] \times \partial_{-,0}\Omega, \end{aligned} \quad (19)$$

for  $t_f > 0$  and  $f \in H^1(\Omega)$ . Remarkably, (19) is smoothing, despite possessing explicit diffusion in one spatial dimension only. Indeed, we have, respectively,

$$[\partial_x, \mathbf{b} \cdot \nabla] = \partial_y, \quad [\partial_y, \mathbf{b} \cdot \nabla] = \partial_z,$$

thereby, Hörmander's rank condition is satisfied [15], implying that (19) is, in fact, hypoelliptic! Moreover, since full rank is achieved via commutators involving the skew-symmetric 1st order part  $\mathbf{b} \cdot \nabla u$ , the PDE in (19) satisfies also the commutator hypotheses of [24, Theorem 24]. The developments discussed in this work can be transferred to this problem also with minor modifications only.

Returning to the case of a general polygonal domain  $\Omega \subset \mathbb{R}^2$ , we shall assume the validity of the additional boundary conditions:

$$\left(\frac{1}{2}x_{n_2}\nabla u - n_1\nabla u_x\right)|_{\partial\Omega\setminus\partial_+\Omega} = 0 \quad \text{and} \quad n_1\nabla u_x|_{\partial_+\Omega} = 0, \quad (20)$$

for the remainder of this work, which trivially imply (17). We do so, instead of prescribing particular additional boundary conditions; we shall clearly state if the validity of (20) is assumed or not for the proof of various results below. We stress that, under minor modifications only, the theory presented below can still be valid for other suitable assumptions, e.g., for periodic solutions also by careful inspection of the proofs. The combination of suitable boundary/closure conditions, whose nature ensures the availability of Poincaré-Friedrichs inequalities, is key in the solution behaviour as  $t_f \rightarrow \infty$ . Nonetheless, in the present work, our focus is on the derivation of Galerkin discretizations of the bilinear form  $a(\cdot, \cdot)$ ; as such, we adopt for simplicity the convenient setting (20) for hypocoercivity-inducing boundary conditions.

## 4 A finite element method

Motivated by the above discussion, we return to the variational representation (12) of problem (10), (20) on a bounded domain  $\Omega \subset \mathbb{R}^2$ . For simplicity of the presentation we, henceforth, assume that  $\Omega$  is polygonal also; extension to curved domains is possible via isoparametric versions of the finite element spaces discussed below. As is standard in discretisation by a Galerkin type finite element method, we introduce a family of triangulations of  $\Omega$ , say  $\mathcal{T}$ , consisting of mutually disjoint open triangular elements  $T \in \mathcal{T}$ , whose closures cover  $\bar{\Omega}$  exactly. Let also  $h : \cup_{T \in \mathcal{T}} T \rightarrow \mathbb{R}_+$  be the local meshsize function defined elementwise by  $h|_T := h_T := \text{diam}(T)$ . For simplicity, we further assume that  $\mathcal{T}$  is shape-regular, in the sense that the radius  $\rho_T$  of the largest inscribed circle of each  $T \in \mathcal{T}$  is bounded from below with respect to each element's diameter  $h|_T$ , uniformly as  $\|h\|_{L^\infty(\Omega)} \rightarrow 0$  under mesh refinement. Also, we assume that  $\mathcal{T}$  is locally quasi-uniform in the sense that the diameters of adjacent elements are uniformly bounded from above and below. Finally, let  $\Gamma := \cup_{T \in \mathcal{T}} \partial T$  denote the mesh skeleton, and  $\Gamma_{\text{int}} := \Gamma \setminus \partial\Omega$ .

We define the standard, continuous, finite element space subordinate to  $\mathcal{T}$  by

$$V_h \equiv V_h^p := \{V \in H_{-,0}^1(\Omega) : V|_T \in \mathbb{P}_p(T), T \in \mathcal{T}\},$$

with  $\mathbb{P}_p(\omega)$ ,  $\omega \subset \mathbb{R}^2$ , denoting the space of polynomials of total degree at most  $p$  over  $\omega$ ,  $p = 2, 3, \dots$ . Further, we define the *jump*  $\llbracket w \rrbracket$  of  $w$  across the common element interface  $e := \partial T^+ \cap \partial T^-$  of two adjacent elements  $T^+, T^-$ , by  $\llbracket w \rrbracket|_e := w|_{T^+} \mathbf{n}|_{T^+} + w|_{T^-} \mathbf{n}|_{T^-}$  with  $\mathbf{n}|_T$  denoting the unit outward normal vector to each point of  $\partial T$ ; correspondingly, we also define the *average*  $\{\!\{w\}\!\}_e := (w|_{T^+} + w|_{T^-})/2$  of  $w$  across  $e$ . Finally, for boundary faces  $e \subset \partial T \cap \partial\Omega$ , we set  $\llbracket w \rrbracket|_e := w|_T \mathbf{n}|_T$  and  $\{\!\{w\}\!\}_e := w|_T$ , respectively. The above jump and average definitions are trivially extended to vector-valued functions by component-wise application.

We shall also make use of the *broken Sobolev spaces*  $H^r(\Omega, \mathcal{T}) := \{v|_T \in H^r(T), T \in \mathcal{T}\}$ , with respective Hilbertian norm  $\|w\|_{H^r(\Omega, \mathcal{T})} := \left(\sum_{T \in \mathcal{T}} \|\nabla^r w\|_T^2 + \|\{\!\{h\}\!\}^{1/2-r} \llbracket w \rrbracket\|_{\Gamma_{\text{int}}}^2\right)^{1/2}$ ,  $r \geq 1$ . Further, we define the broken gradient  $\nabla_{\mathcal{T}}$  to be the element-wise gradient operator with  $\nabla_{\mathcal{T}} w|_T = \nabla w|_T$ ,  $T \in \mathcal{T}$ , for  $w \in H^1(\Omega, \mathcal{T})$ .

Starting from (12), we consider the spatially discrete finite element method: find  $U \equiv$

$U(t) \in V_h$ ,  $t \in (0, t_f]$ , such that

$$\begin{aligned} & (U_t, V) + (U_x, V_x) + (xU_y, V) \\ & + (\nabla U_t, A\nabla V) + (A\nabla_\tau U_x, \nabla_\tau V_x) + (\nabla_\tau(xU_y), A\nabla V) + s_h(U, V) \\ & = (f, V) + (\nabla f, A\nabla V), \end{aligned} \quad (21)$$

for all  $V \in V_h$ , whereby

$$\begin{aligned} s_h(U, V) := & - \int_{\Gamma_{\text{int}} \cup \partial_- \Omega} (0, x)^T \cdot (\alpha \llbracket U_x \rrbracket \llbracket V_x \rrbracket + \beta \llbracket U_x \rrbracket \llbracket V_y \rrbracket + \beta \llbracket U_y \rrbracket \llbracket V_x \rrbracket + \gamma \llbracket U_y \rrbracket \llbracket V_y \rrbracket) \, ds \\ & + \int_{\Gamma_{\text{int}}} \frac{|xn_2|}{2} (\kappa \llbracket U_x \rrbracket \cdot \llbracket V_x \rrbracket + \lambda \llbracket U_y \rrbracket \cdot \llbracket V_y \rrbracket) \, ds \\ & - \int_{\Gamma_{\text{int}} \cup \partial_0 \Omega} \left( \llbracket A\nabla U_x \rrbracket \cdot [\nabla V]_1 + \llbracket A\nabla V_x \rrbracket \cdot [\nabla U]_1 - \tau [\nabla U]_1 \cdot A[\nabla V]_1 \right) \, ds, \end{aligned}$$

with  $[v]_1|_e := (v|_e)^+ n_1^+ + (v|_e)^- n_1^-$ , for each internal face  $e$  and  $n_1$  denoting the first component of the unit normal vector to  $e$  and  $[v]_1|_e := (v|_e)^+ n_1^+$ , for  $e \subset \partial\Omega$ , for some  $\tau : \Gamma \rightarrow \mathbb{R}$  non-negative function to be determined precisely below and user-defined parameters  $\kappa, \lambda \geq 0$ ; we also set  $U(0) := \Pi u_0 \in V_h$ , with  $\Pi : L_2(\Omega) \rightarrow V_h$  denoting the orthogonal  $L_2$ -projection onto the finite element space. For brevity, we shall adopt the notation:

$$s_{1,h}(w, v) := \int_{\Gamma_{\text{int}}} \frac{|xn_2|}{2} (\kappa \llbracket w_x \rrbracket \cdot \llbracket v_x \rrbracket + \lambda \llbracket w_y \rrbracket \cdot \llbracket v_y \rrbracket) \, ds.$$

The term  $s_{1,h}(w, v)$  introduces numerical diffusion in the  $(0, x)^T$ -direction when  $\kappa, \lambda > 0$ . Such additional numerical diffusion may be desirable in the presence of rough initial conditions. When additional numerical diffusion is not advantageous, we can choose  $\kappa, \lambda = 0$  giving  $s_{1,h}(\cdot, \cdot) \equiv 0$ .

As we shall see below, the stabilisation  $s_h(\cdot, \cdot)$  is constructed so that the method (21) is *consistent* with (12) upon assuming the additional boundary conditions (20).

**Lemma 4.1.** *Let  $\tau : \Gamma \rightarrow \mathbb{R}$  with  $\tau = C_\tau p^2 / \llbracket h \rrbracket$ , for  $C_\tau > 0$  large enough but independent of  $h$  and of  $p$ . Then, for any  $0 < \alpha < 1/2$ , there exist  $0 < \gamma < \beta < 1$  and a positive constant  $c_0$ , independent of the approximate solution  $U \in V_h$  and of  $t_f$ , such that for any  $0 < \zeta < 1$ , we have*

$$\begin{aligned} & \|U(t_f)\|^2 + \|\sqrt{A}\nabla U(t_f)\|^2 + (1 - \zeta) \int_0^{t_f} (c_0 \|U\|^2 + \|\sqrt{A}\nabla U\|^2) \, dt \\ & + \int_0^{t_f} (\|\sqrt{xn_2}U\|_{\partial_+ \Omega}^2 + \|\sqrt{xn_2}A\nabla U\|_{\partial_+ \Omega}^2 \\ & \quad + \|\sqrt{A}\nabla_\tau U_x\|^2 + s_{1,h}(U, U) + \|\sqrt{\tau}A[\nabla U]_1\|_{\Gamma_{\text{int}} \cup \partial_0 \Omega}^2) \, dt \\ & \leq \frac{4}{\zeta} \int_0^{t_f} (c_0^{-1} \|f\|^2 + \|\sqrt{A}\nabla f\|^2) \, dt + \|\Pi u_0\|^2 + \|\sqrt{A}\nabla \Pi u_0\|^2; \end{aligned} \quad (22)$$

when  $f \equiv 0$ , we can trivially select  $\zeta = 0$  in the proof, so that the first term on the right-hand side of (22) vanishes.

*Proof.* Setting  $V = U$  in (21), standard arguments give

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} (\|U\|^2 + \|\sqrt{A}\nabla U\|^2) + \|U_x\|^2 + \frac{1}{2} \|\sqrt{xn_2}U\|_{\partial_+ \Omega}^2 \\ & \quad + \|\sqrt{A}\nabla_\tau U_x\|^2 + (\nabla_\tau(xU_y), A\nabla U) + s_h(U, U) \\ & = (f, U) + (\nabla f, A\nabla U), \end{aligned} \quad (23)$$

upon observing that  $(xU_y, U) = \|\sqrt{xn_2}U\|_{\partial_+\Omega}^2/2$  as  $U = 0$  on  $\partial_-\Omega$ . We calculate

$$\begin{aligned} (\nabla(xU_y), A\nabla U)_T &= \alpha(U_y, U_x)_T + \beta\|U_y\|_T^2 + \alpha(U_x, xU_{yx})_T + \beta(U_y, xU_{yx})_T \\ &\quad + \beta(xU_{yy}, U_x)_T + \gamma(xU_{yy}, U_y)_T. \end{aligned}$$

Observe now the identities

$$\begin{aligned} \alpha(U_x, xU_{yx})_T &= \frac{\alpha}{2} \int_{\partial T} xn_2 U_x^2 ds, \quad \gamma(U_y, xU_{yy})_T = \frac{\gamma}{2} \int_{\partial T} xn_2 U_y^2 ds, \\ \beta(xU_{yy}, U_x)_T &= -\beta(xU_y, U_{xy})_T + \beta \int_{\partial T} xn_2 U_x U_y ds. \end{aligned}$$

Then, summing over all  $T \in \mathcal{T}$ , we deduce

$$(\nabla_{\mathcal{T}}(xU_y), A\nabla U) = \alpha(U_y, U_x) + \beta\|U_y\|^2 + \frac{1}{2} \int_{\Gamma} (0, x)^T \cdot [\alpha U_x^2 + 2\beta U_x U_y + \gamma U_y^2] ds. \quad (24)$$

Using (24) into (23), noting that  $\llbracket WV \rrbracket = \llbracket W \rrbracket \llbracket V \rrbracket + \llbracket V \rrbracket \llbracket W \rrbracket$  on  $\Gamma_{\text{int}}$  for  $V, W \in V_h$ , and integrating with respect to  $t \in [0, t_f]$ , therefore, gives

$$\begin{aligned} &\|U(t_f)\|^2 + \|\sqrt{A}\nabla U(t_f)\|^2 + 2 \int_0^{t_f} \|\sqrt{A}\nabla_{\mathcal{T}} U_x\|^2 dt + \int_0^{t_f} \|\sqrt{xn_2}U\|_{\partial_+\Omega}^2 dt \\ &+ \int_0^{t_f} \int_{\partial_+\Omega} xn_2(\alpha U_x^2 + 2\beta U_x U_y + \gamma U_y^2) ds dt + 2 \int_0^{t_f} (\|U_x\|^2 + \alpha(U_y, U_x) + \beta\|U_y\|^2) dt \\ &- 4 \int_0^{t_f} \int_{\Gamma_{\text{int}} \cup \partial_0\Omega} \llbracket A\nabla U_x \rrbracket [\nabla U]_1 ds dt + 2 \int_0^{t_f} (\|\sqrt{\tau A}[\nabla U]_1\|_{\Gamma_{\text{int}} \cup \partial_0\Omega}^2 + s_{1,h}(U, U)) dt \\ &= 2 \int_0^{t_f} ((f, U) + (\nabla f, A\nabla U)) dt + \|\Pi u_0\|^2 + \|\sqrt{A}\nabla \Pi u_0\|^2. \end{aligned} \quad (25)$$

Focusing on the fifth and sixth terms on the left-hand side of (25), recalling (11), the Cauchy-Schwarz and Young inequalities, along with elementary manipulations, imply

$$\int_{\partial_+\Omega} xn_2(\alpha U_x^2 + 2\beta U_x U_y + \gamma U_y^2) ds = \int_{\partial_+\Omega} xn_2 |\sqrt{A}\nabla U|^2 ds \geq 0, \quad (26)$$

and

$$2(\|U_x\|^2 + \alpha(U_y, U_x) + \beta\|U_y\|^2) \geq 2(1 - \epsilon)\|U_x\|^2 + (2\beta - \frac{\alpha^2}{2\epsilon})\|U_y\|^2 = \|\sqrt{B}\nabla U\|^2, \quad (27)$$

respectively, upon defining the diagonal matrix  $B = \text{diag}(2(1 - \epsilon), 2\beta - \alpha^2/(2\epsilon))$ , for any  $0 < \epsilon < 1$ . We, then, have

$$\|\sqrt{B}\nabla U\|^2 = \|\sqrt{B - A}\nabla U\|^2 + \|\sqrt{A}\nabla U\|^2 \geq \lambda_{B-A}^{\min} \|\nabla U\|^2 + \|\sqrt{A}\nabla U\|^2,$$

where  $\lambda_{B-A}^{\min}$  denotes the smallest eigenvalue of  $B - A$ . There exist values  $0 < \gamma < \beta$  such that  $\lambda_{B-A}^{\min} > 0$ . For instance, setting  $\beta = \alpha^2$ ,  $\gamma = \alpha^3$ , and  $\epsilon = 1/2$ , we have  $\det(B - A) = \alpha^2(1 - 2\alpha)$ . A numerical investigation reveals that  $\max_{0 \leq \alpha \leq 1/2} \lambda_{B-A}^{\min} \approx 0.054429$  attained for a value  $\alpha \approx 0.35060$ , noting that  $B - A$  has positive determinant for  $0 < \alpha < 1/2$ .

Since  $U = 0$  on  $\partial_{-,0}\Omega$ , which was assumed to be of positive one-dimensional Hausdorff measure, the validity of a Poincaré-Friedrichs/spectral gap inequality  $\|U\|^2 \leq C_{PF} \|\nabla U\|^2$  with

$C_{PF} \equiv C_{PF}(\Omega)$  positive constant, independent of  $U$ , is assured, thereby allowing us to conclude the lower bound

$$\|\sqrt{B}\nabla U\|^2 \geq C_{PF}^{-1} \lambda_{B-A}^{\min} \|U\|^2 + \|\sqrt{A}\nabla U\|^2. \quad (28)$$

Thus, with the choices  $\alpha \approx 0.35060$ ,  $\beta = \alpha^2$ ,  $\gamma = \alpha^3$ , and  $\epsilon = 1/2$ , we can select  $c_0 = 0.054429/C_{PF}$ . As we shall see below,  $c_0$  effectively represents the exponent of the exponential decay to equilibrium; cf. Theorem 4.3 below.

Further, the symmetry of  $A$  along with the Cauchy-Schwarz inequality and a standard inverse estimate yield, respectively,

$$\begin{aligned} \int_{\Gamma_{\text{int}} \cup \partial_0 \Omega} \{A \nabla U_x\} \cdot [\nabla U]_1 \, ds &\leq \|\varphi^{-1/2} \{\sqrt{A} \nabla U_x\}\|_{\Gamma_{\text{int}} \cup \partial_0 \Omega} \|\sqrt{\varphi A} [\nabla U]_1\|_{\Gamma_{\text{int}} \cup \partial_0 \Omega} \\ &\leq \|\sqrt{C_{\text{inv}} p^2 / h \varphi} \sqrt{A} \nabla_{\mathcal{T}} U_x\| \|\sqrt{\varphi A} [\nabla U]_1\|_{\Gamma_{\text{int}} \cup \partial_0 \Omega} \\ &\leq \frac{1}{4} \|\sqrt{A} \nabla_{\mathcal{T}} U_x\|^2 + \|\sqrt{\tau A} [\nabla U]_1\|_{\Gamma_{\text{int}} \cup \partial_0 \Omega}^2, \end{aligned} \quad (29)$$

using the standard inverse inequality  $\|v\|_{\partial T}^2 \leq C_{\text{inv}} p^2 / h_T \|v\|_T^2$  on (29), and selecting  $\varphi = \tau \geq 4C_{\text{inv}} p^2 / \{h\}$  and resorting to the local quasiuniformity of the mesh. Inserting the last two bounds into (25) and, using the Cauchy-Schwarz inequality, setting  $c_0 := C_{PF}^{-1} \lambda_{B-A}^{\min}$ , we deduce the result for any  $0 < \zeta < 1$ ; when  $f \equiv 0$  we can trivially select  $\zeta = 0$ .  $\square$

We note that the specific choice of  $A$  as a function of  $\alpha$  provided in the proof above is *not* unique. As such, we have avoided providing exact values of  $\alpha, \beta, \gamma$  in the statement of Lemma 4.1. Nonetheless, preliminary numerical investigation shows that the choice of  $A$  does not affect significantly the performance of the method, provided that  $A$  is chosen to ensure (28) away from limiting values [7].

**Remark 4.2.** *It is possible to optimise further the constant  $c_0$  by working as follows: starting from  $\|\sqrt{B}\nabla U\|^2$ , we have*

$$\|\sqrt{B}\nabla U\|^2 = \|\sqrt{B - \nu A} \nabla U\|^2 + \nu \|\sqrt{A} \nabla U\|^2 \geq \lambda_{B-\nu A}^{\min} \|\nabla U\|^2 + \nu \|\sqrt{A} \nabla U\|^2,$$

for suitable  $\nu > 0$ . This way it is possible to arrive at a different, possibly larger, value for  $\lambda_{B-\nu A}^{\min}$  and, consequently, for  $c_0$ .

The last proof highlights that the validity of the Poincaré-Friedrichs inequality is of central importance here, and in general in the study of trend to equilibrium for kinetic equations [22, 14, 24]. In the present context of bounded spatial domain  $\Omega$ , we have taken the viewpoint of using standard (non-weighted) integral norms. Thus the Lebesgue measure of  $\Omega$  is present in the constant  $c_0$ . The extension of the above framework to general Fokker-Planck equations on function spaces weighted by equilibrium distributions is an important question and will be discussed elsewhere. The availability of such a Poincaré-Friedrichs inequality (also known as spectral gap property in kinetic theory) facilitates the crucial feature of the proposed method: the absence of an exponential term of the form  $\exp(ct_f)$  multiplying the data terms in the above stability estimate. Moreover, (21) also immediately implies the decay of numerical solutions for the respective homogeneous problem.

**Theorem 4.3** (Decay via hypocoercivity). *Under the hypotheses of Lemma 4.1, the finite element method (21) for the homogeneous problem (10) with  $f \equiv 0$  satisfies*

$$\|U(t_f)\|^2 + \|\sqrt{A}\nabla U(t_f)\|^2 \leq e^{-\min\{1, c_0\}t_f} \left( \|\Pi u_0\|^2 + \|\sqrt{A}\nabla \Pi u_0\|^2 \right),$$

for all  $t_f > 0$ .

*Proof.* Starting from (23) and using the estimates from the proof of Lemma 4.1, we arrive at the estimate

$$\frac{d}{dt}(\|U\|^2 + \|\sqrt{A}\nabla U\|^2) + \min\{1, c_0\}(\|U\|^2 + \|\sqrt{A}\nabla U\|^2) \leq 0,$$

from which an application of Grönwall's Lemma already implies the result.  $\square$

## 5 Error analysis

We continue by proving a priori error bounds for (21). For brevity, we shall denote by  $\mathcal{B}(\cdot, \cdot) : H^{5/2+\varepsilon}(\Omega, \mathcal{T}) \times H^{5/2+\varepsilon}(\Omega, \mathcal{T}) \rightarrow \mathbb{R}$ ,  $\varepsilon > 0$ , the bilinear form given by

$$\mathcal{B}(w, v) := (w_x, v_x) + (xw_y, v) + (A\nabla_{\mathcal{T}}w_x, \nabla_{\mathcal{T}}v_x) + (\nabla_{\mathcal{T}}(xw_y), A\nabla v) + s_h(w, v).$$

Implicitly in the proof of Lemma 4.1, we proved also the following (hypo)coercivity result.

**Lemma 5.1** ((Hypo)coercivity). *Let*

$$\begin{aligned} \|w\| := & \left( \|\sqrt{B}\nabla w\|^2 + \|\sqrt{xn_2}w\|_{\partial_+\Omega}^2 + \|\sqrt{xn_2}A\nabla w\|_{\partial_+\Omega}^2 \right. \\ & \left. + \|\sqrt{A}\nabla_{\mathcal{T}}w_x\|^2 + \|\sqrt{\tau}A[\nabla w]_1\|_{\Gamma_{\text{int}} \cup \partial_0\Omega}^2 + s_{1,h}(w, w) \right)^{1/2}, \end{aligned}$$

with  $B$  as in the proof of Lemma 4.1. Then,  $\|\cdot\|$  is a norm on  $H_{-,0}^1(\Omega) \cap H^2(\Omega, \mathcal{T})$ , and we have

$$\mathcal{B}(w, w) \geq \frac{1}{2}\|w\|^2,$$

for all  $w \in V_h$ .  $\square$

For the remaining of this work, unless explicitly stated, we assume that  $\epsilon$  is away from 1 and that  $\beta \sim \alpha^2$ , so that  $B = c_B \text{diag}(1, \alpha^2)$  for a given constant  $0 < c_B \leq 1$ . Thus, we can take  $C_0 = (2c_B)^{-1}$ . Note that this also covers the case of  $A$  being singular, i.e., the operator  $\nabla \cdot A\nabla(\cdot)$  being parabolic. Next, we establish the consistency of the bilinear form  $\mathcal{B}$ .

**Lemma 5.2.** *Assume that for the solution  $u$  of (10), (20) we have  $u(t, \cdot) \in H_{-,0}^1(\Omega) \cap H^{5/2+\varepsilon}(\Omega, \mathcal{T})$ ,  $\varepsilon > 0$ , for almost all  $t \in (0, t_f]$ . Then, for almost all  $t \in (0, t_f]$  and for all  $V \in V_h$ , we have*

$$(u_t, V) + (\nabla u_t, A\nabla V) + \mathcal{B}(u, V) = (f, V) + (\nabla f, A\nabla V). \quad (30)$$

*Proof.* The proof follows by integration by parts and by the smoothness of  $u$ , noting that from (20), we have  $(\frac{1}{2}xn_2\nabla u - n_1\nabla_{\mathcal{T}}u_x, A\nabla V)_{\partial\Omega \setminus \partial_+\Omega} = 0$  and  $(n_1\nabla_{\mathcal{T}}u_x, A\nabla v)_{\partial_+\Omega} = 0$ .  $\square$

Let  $\pi : H_{-,0}^1(\Omega) \rightarrow V_h$ , a projection operator onto the finite element space to be defined precisely below, and set

$$\rho := u - \pi u, \quad \vartheta := \pi u - U.$$

Employing (30) and (21) with  $V = \vartheta$ , gives the error equation

$$\frac{1}{2} \frac{d}{dt} (\|\vartheta(s)\|^2 + \|\sqrt{A}\nabla \vartheta(s)\|^2) + \mathcal{B}(\vartheta, \vartheta) = -(\rho_t, \vartheta) - (\nabla \rho_t, A\nabla \vartheta) - \mathcal{B}(\rho, \vartheta). \quad (31)$$

We shall now construct a  $\pi$ , so that  $\mathcal{B}(\rho, \vartheta) = 0$ .

**Lemma 5.3** ((Hypo)elliptic projection). *Assume that  $B = c_B \text{diag}(1, \alpha^2)$ , for a given constant  $0 < c_B \leq 1$ . For every  $v \in H_{-,0}^1(\Omega) \cap H^{5/2+\varepsilon}(\Omega, \mathcal{T})$ ,  $\varepsilon > 0$ , there exists a unique  $\pi v \in V_h$  such that*

$$\mathcal{B}(\pi v, V) = \mathcal{B}(v, V), \quad (32)$$

for all  $V \in V_h$ . Moreover, assuming further that  $v \in H^{k_T}(T)$  for  $T \in \mathcal{T}$ ,  $k_T \geq 3$ , we have the approximation estimate:

$$\|v - \pi v\|^2 \leq C(A, \kappa, \lambda) \sum_{T \in \mathcal{T}} \frac{h_T^{2s_T}}{p^{2(k_T-3)}} x_T |u|_{H^{k_T}(T)}^2, \quad (33)$$

with  $s_T := \min\{p+1, k_T\} - 2$  and  $x_T := \max_{x \in T} \{1, |x|^2\}$ ,  $T \in \mathcal{T}$ .

*Proof.* Uniqueness (and, therefore, existence due to linearity) follows from the coercivity of  $\mathcal{B}$  shown in Lemma 5.1. Setting now  $\eta := v - \mathcal{P}v$  and  $\xi := \mathcal{P}v - \pi v$ , with  $\mathcal{P}$  denoting a projection operator from  $H_{-,0}^1(\Omega)$  onto  $V_h$  with optimal approximation properties with respect to the local meshsize  $h$ , we have

$$\begin{aligned} \frac{1}{2} \|\xi\|^2 &\leq \mathcal{B}(\xi, \xi) = -\mathcal{B}(\eta, \xi) \\ &= -(\eta_x, \xi_x) - (\eta_y, x\xi) - (\sqrt{A} \nabla_{\mathcal{T}} \eta_x, \sqrt{A} \nabla_{\mathcal{T}} \xi_x) \\ &\quad - (\nabla_{\mathcal{T}}(x\eta_y), A \nabla \xi) - s_h(\eta, \xi). \end{aligned} \quad (34)$$

Now, integration by parts yields

$$-(\eta_y, x\xi) = (x\eta, \xi_y) - \int_{\partial_+ \Omega} x n_2 \eta \xi \, ds, \quad (35)$$

and

$$\begin{aligned} &(\nabla_{\mathcal{T}}(x\eta_y), A \nabla \xi) \\ &= \sum_{T \in \mathcal{T}} \left( (\eta_y + x\eta_{xy}, \alpha \xi_x + \beta \xi_y)_T + (x\eta_{yy}, \beta \xi_x + \gamma \xi_y)_T \right) \\ &= \alpha(\eta_y, \xi_x) + \beta(\eta_y, \xi_y) - \sum_{T \in \mathcal{T}} \left( (\eta_x, x(\alpha \xi_x + \beta \xi_y))_T + (\eta_y, x(\beta \xi_x + \gamma \xi_y))_T \right) \\ &\quad + \sum_{T \in \mathcal{T}} \int_{\partial T} x n_2 (\alpha \eta_x \xi_x + \beta \eta_x \xi_y + \beta \eta_y \xi_x + \gamma \eta_y \xi_y) \, ds. \end{aligned} \quad (36)$$

Observe that the last term on the right-hand side of (36) is equal to

$$\begin{aligned} &\int_{\Gamma_{\text{int}} \cup \partial_- \Omega} (0, x)^T \cdot (\alpha \llbracket \eta_x \rrbracket \{\xi_x\} + \beta \llbracket \eta_x \rrbracket \{\xi_y\} + \beta \llbracket \eta_y \rrbracket \{\xi_x\} + \gamma \llbracket \eta_y \rrbracket \{\xi_y\}) \, ds \\ &+ \int_{\Gamma_{\text{int}}} (0, x)^T \cdot (\alpha \{\eta_x\} \llbracket \xi_x \rrbracket + \beta \{\eta_x\} \llbracket \xi_y \rrbracket + \beta \{\eta_y\} \llbracket \xi_x \rrbracket + \gamma \{\eta_y\} \llbracket \xi_y \rrbracket) \, ds \\ &+ \int_{\partial_+ \Omega} x n_2 A \nabla \eta \cdot \nabla \xi \, ds, \end{aligned} \quad (37)$$

and note that the first term of (37) cancels with the first term of  $s_h(\eta, \xi)$ . Using now (35), (36) and (37) into (34), along with standard Cauchy-Schwarz' and Young's inequality arguments,

and working as in (29) gives

$$\begin{aligned}
& C^{-1} |||\xi|||^2 \\
& \leq \|\eta_x\|^2 + \beta^{-1}\|x\eta\|^2 + \|\sqrt{xn_2}\eta\|_{\partial_+\Omega}^2 + \|\sqrt{xn_2}A\nabla\eta\|_{\partial_+\Omega}^2 + \|\sqrt{A}\nabla_{\mathcal{T}}\eta_x\|^2 \\
& \quad + (\alpha^2 + \beta)\|\eta_y\|^2 + \|\sqrt{\tau A}[\nabla\eta]_1\|_{\Gamma_{\text{int}}}^2 + \kappa\|\sqrt{|xn_2|}[\eta_x]\|_{\Gamma_{\text{int}}}^2 + \lambda\|\sqrt{|xn_2|}[\eta_y]\|_{\Gamma_{\text{int}}}^2 \\
& \quad + \sum_{T \in \mathcal{T}} \left( \|\mu^{-1}x\eta_x\|_T(\alpha\|\mu(\xi_x)_y\|_T + \beta\|\mu(\xi_y)_y\|_T) + \|\mu^{-1}x\eta_y\|_T(\beta\|\mu(\xi_x)_y\|_T + \gamma\|\mu(\xi_y)_y\|_T) \right) \\
& \quad + \|\sqrt{A/\tau}\{\nabla\eta_x\}\|_{\Gamma_{\text{int}}}^2 + (\alpha^2\kappa^{-1} + \beta^2\lambda^{-1})\|\sqrt{xn_2}\{\eta_x\}\|_{\Gamma_{\text{int}}}^2 + (\beta^2\kappa^{-1} + \gamma^2\lambda^{-1})\|\sqrt{xn_2}\{\eta_y\}\|_{\Gamma_{\text{int}}}^2,
\end{aligned}$$

for any  $\mu, \kappa, \lambda > 0$ , for some  $C > 1$ , independent of the, relevant to the argument, parameters. Alternatively, when  $\kappa, \lambda = 0$ , we work as follows: using the inverse estimate  $\|\xi_x\|_{\partial T}^2 \leq Cp^2/h_T\|\xi_x\|_T^2$  and, similarly for  $\xi_y$ , we further estimate the second term on the right-hand side of (37) to arrive instead at

$$\begin{aligned}
& C^{-1} |||\xi|||^2 \\
& \leq \|\eta_x\|^2 + \beta^{-1}\|x\eta\|^2 + \|\sqrt{xn_2}\eta\|_{\partial_+\Omega}^2 + \|\sqrt{A}\nabla_{\mathcal{T}}\eta_x\|^2 + (\alpha^2 + \beta)\|\eta_y\|^2 + \|\sqrt{\tau A}[\nabla\eta]_1\|_{\Gamma}^2 \\
& \quad + \sum_{T \in \mathcal{T}} \left( \|\mu^{-1}x\eta_x\|_T(\alpha\|\mu(\xi_x)_y\|_T + \beta\|\mu(\xi_y)_y\|_T) + \|\mu^{-1}x\eta_y\|_T(\beta\|\mu(\xi_x)_y\|_T + \gamma\|\mu(\xi_y)_y\|_T) \right) \\
& \quad + C|AB^{-1/2}|_2(\|xn_2h^{-1/2}\{\eta_x\}\|_{\Gamma_{\text{int}}}^2 + \|xn_2h^{-1/2}\{\eta_y\}\|_{\Gamma_{\text{int}}}^2) + \frac{1}{4}\|\sqrt{B}\nabla\xi\|^2,
\end{aligned}$$

with  $|\cdot|_2$  denoting the matrix-2-norm. Applying the inverse estimate  $\|\nabla v\|_T^2 \leq Cp^4/h_T^2\|v\|_T^2$ , for  $v \in \mathbb{P}_p(T)$  on the terms containing  $\mu\xi$  in the last two alternative estimates, selecting  $\mu = h/p^2$ , and performing standard manipulations, we arrive at the combined estimate

$$\begin{aligned}
C^{-1} |||\xi|||^2 & \leq \|\eta_x\|^2 + \beta^{-1}\|x\eta\|^2 + \|\sqrt{xn_2}\eta\|_{\partial_+\Omega}^2 + \|\sqrt{A}\nabla_{\mathcal{T}}\eta_x\|^2 + \|\sqrt{\tau A}[\nabla\eta]_1\|_{\Gamma}^2 \\
& \quad + (\alpha^2 + \beta)(\|\mu^{-1}x\eta_x\|^2 + \|\eta_y\|^2) + (\beta^2 + \frac{\gamma^2}{\beta})\|\mu^{-1}x\eta_y\|^2 \\
& \quad + \min \left\{ (\alpha^2\kappa^{-1} + \beta^2\lambda^{-1})\|\sqrt{xn_2}\{\eta_x\}\|_{\Gamma_{\text{int}}}^2 + (\beta^2\kappa^{-1} + \gamma^2\lambda^{-1})\|\sqrt{xn_2}\{\eta_y\}\|_{\Gamma_{\text{int}}}^2, \right. \\
& \quad \left. C|AB^{-1/2}|_2(\|xn_2h^{-1/2}\{\eta_x\}\|_{\Gamma_{\text{int}}}^2 + \|xn_2h^{-1/2}\{\eta_y\}\|_{\Gamma_{\text{int}}}^2) \right\} \\
& \quad + \kappa\|\sqrt{|xn_2|}[\eta_x]\|_{\Gamma_{\text{int}}}^2 + \lambda\|\sqrt{|xn_2|}[\eta_y]\|_{\Gamma_{\text{int}}}^2,
\end{aligned}$$

which holds for any  $\kappa, \lambda \geq 0$ . Choosing  $\mathcal{P}$  to be an  $hp$ -optimal projection operator onto the finite element space  $V_h$ , we can assume approximation estimates of the form

$$\|\nabla^m \eta\|_T \leq Ch_T^{\min\{p+1, k\}-m} p^{m-k} |u|_{H^k(T)}$$

for  $k > m$ , with  $u \in H^k(T)$ ,  $T \in \mathcal{T}$  (see, e.g., [23, 4] for examples of such operators). These, together with the trace estimate  $\|w\|_{\partial T}^2 \leq C\|w\|_T\|w\|_{H^1(T)}$  and the triangle inequality  $|||\rho||| \leq |||\eta||| + |||\xi|||$ , already imply (33).  $\square$

We now show the (potentially super-)approximation for the left-hand side of (31). Such results are typical to finite element method for parabolic problems. We shall show that this is the case also for (21) discretising the, degenerate, Kolmogorov equation.

**Lemma 5.4.** Assume that the exact solution  $u$  to (10), (20) satisfies  $u \in H^1(0, t_f; H_{-,0}^1(\Omega) \cap H^{5/2+\varepsilon}(\Omega, \mathcal{T}))$ . Then, under the hypotheses of Lemma 4.1, we have the following bound

$$\begin{aligned} \|\vartheta(t)\|^2 + \|\sqrt{A}\nabla\vartheta(t)\|^2 + \frac{1}{2} \int_0^t \|\vartheta\|^2 ds &\leq \|\vartheta(0)\|^2 + \|\sqrt{A}\nabla\vartheta(0)\|^2 \\ &+ 4 \max\{1, c_0^{-2}\} \int_0^t \|\sqrt{B}\nabla\rho_t\|^2 ds, \end{aligned} \quad (38)$$

for any  $t \in (0, t_f]$ , for some positive constant  $C$ , depending only on the shape-regularity and the local quasi-uniformity of the mesh.

*Proof.* Starting from (31), upon integration in time between 0 and  $t \in (0, t_f]$ , multiplying by 2 and using standard arguments, gives

$$\begin{aligned} \|\vartheta(t)\|^2 + \|\sqrt{A}\nabla\vartheta(t)\|^2 + \int_0^t \|\vartheta\|^2 ds &\leq \|\vartheta(0)\|^2 + \|\sqrt{A}\nabla\vartheta(0)\|^2 \\ &+ 2 \int_0^t (c_0^{-1} \|\rho_t\|^2 + \|\sqrt{A}\nabla\rho_t\|^2) ds \\ &+ \frac{1}{2} \int_0^t (c_0 \|\vartheta\|^2 + \|\sqrt{A}\nabla\vartheta\|^2) ds, \end{aligned} \quad (39)$$

recalling the notation  $c_0 = C_{PF}^{-1} \lambda_{B-A}^{\min}$ . Using now (28), we deduce the result.  $\square$

The above developments yield an a priori error bound in the norm appearing on the left-hand side of (38).

**Theorem 5.5.** Assume that  $u_0, u, u_t \in H_{-,0}^1(\Omega) \cap H^{k_T}(T)$  for  $k_T \geq 3$ ,  $T \in \mathcal{T}$ , the latter two for almost all  $t \in (0, t_f]$ , with  $u$  being the exact solution to (10), (20). Then, under the assumptions of Lemmata 4.1 and 5.3, the error  $e := u - U$  of the finite element method (21) satisfies the bound:

$$\|e(t)\|^2 + \|\sqrt{A}\nabla e(t)\|^2 + \int_0^t \|e\|^2 ds \leq C(A, \kappa, \lambda) \sum_{T \in \mathcal{T}} \mathcal{E}_T(u_0, t, u), \quad (40)$$

where

$$\mathcal{E}_T(u_0, t, u) := \frac{h_T^{2s_T} x_T}{p^{2(k_T-3)}} \left( \max\{1, c_0^{-1}\} |u_0|_{H^{k_T}(T)}^2 + |u_t|_{L_2(0,t;H^{k_T}(T))}^2 + |u|_{L_2(0,t;H^{k_T}(T))}^2 \right),$$

with  $s_T := \min\{p+1, k_T\} - 2$ , and the constant  $C(A, \kappa, \lambda) > 0$  is bounded away from  $+\infty$  for  $0 < \alpha \leq 1/2$  and all  $\kappa, \lambda \geq 0$  and independent of the mesh parameters and of  $u$ . Moreover,  $C(A, \kappa, \lambda)$  is independent of the final time  $t_f$ .

*Proof.* Working as for (28), we have

$$\begin{aligned} \|\vartheta(0)\|^2 + \|\sqrt{A}\nabla\vartheta(0)\|^2 &\leq \max\{1, c_0^{-1}\} \|\sqrt{B}\nabla\vartheta(0)\|^2 \\ &\leq 2 \max\{1, c_0^{-1}\} (\|\sqrt{B}\nabla\rho(0)\|^2 + \|\sqrt{B}\nabla(u_0 - \Pi u_0)\|^2). \end{aligned}$$

Similarly, we also have

$$\|\rho(0)\|^2 + \|\sqrt{A}\nabla\rho(0)\|^2 \leq \max\{1, c_0^{-1}\} \|\sqrt{B}\nabla\rho(0)\|^2.$$

The last two estimates can be further bounded from above by (33) and  $hp$ -approximation bounds for the orthogonal  $L_2$ -projection [18, 5], respectively. Also, from (33), we have

$$\int_0^t (\|\sqrt{B}\nabla\rho_t\|^2 + \|\rho\|^2) \, ds \leq C \sum_{T \in \mathcal{T}} \frac{h_T^{2s_T} x_T}{p^{2(k_T-3)}} (|u_t|_{L_2(0,t;H^{k_T}(T))}^2 + |u|_{L_2(0,t;H^{k_T}(T))}^2). \quad (41)$$

Combining the above bounds with the triangle inequality recalling that  $u - u_h = \rho + \vartheta$ , the result already follows.  $\square$

The crucial property of the above error estimate is the *independence* of the respective constant  $C(A, \kappa, \lambda)$  from the final time  $t_f$ ; this is a direct consequence of the hypocoercivity-compatible discretisation (21).

Assumption  $u_0 \in H_{-,0}^1(\Omega)$  is a natural compatibility condition in this context. Since a typical setting in kinetic simulation concerns initial profiles with compact support within a computational domain  $\Omega$ ,  $u_0 \in H_{-,0}^1(\Omega)$  is of significant practical relevance also. Nonetheless, we shall now discuss the dependence of the error committed by the numerical method as  $t_f \rightarrow \infty$  and we shall see that the effect of the initial condition error  $u_0 - \Pi u_0$  diminishes exponentially as  $t_f \rightarrow \infty$ . Thus, possibly incompatible initial conditions  $u_0 \in H^1(\Omega) \setminus H_{-,0}^1(\Omega)$  will have an exponentially diminishing effect in the accuracy of the method with respect to  $t_f$ .

**Theorem 5.6.** *Assume that  $u, u_t \in H_{-,0}^1(\Omega) \cap H^{k_T}(T)$  for  $k_T \geq 3$ ,  $T \in \mathcal{T}$  for almost all  $t \in (0, t_f]$ , with  $u$  being the exact solution to (10), (20), and  $u_0 \in H^1(\Omega)$ . Under the hypotheses of Lemmata 4.1 and 5.3 and of Theorem 5.5, the error  $e := u - U$  of the finite element method (21) satisfies the bound:*

$$\begin{aligned} \|e(t_f)\|^2 + \|\sqrt{A}\nabla e(t_f)\|^2 &\leq e^{-\min\{1, c_0\}t_f} (\|\pi u_0 - \Pi u_0\|^2 + \|\sqrt{A}\nabla(\pi u_0 - \Pi u_0)\|^2) \\ &\quad + \frac{C(A, \kappa, \lambda)}{\min\{1, c_0\}} \sum_{T \in \mathcal{T}} \frac{h_T^{2s_T} x_T}{p^{2(k_T-3)}} \mathcal{H}_T(t, u), \end{aligned} \quad (42)$$

where

$$\mathcal{H}_T(t, u) := |u(t_f)|_{H^{k_T}(T)}^2 + \int_0^{t_f} e^{-\min\{1, c_0\}(t_f-t)} |u_t|_{H^{k_T}(T)}^2 \, dt,$$

for  $s_T := \min\{p+1, k_T\} - 2$ , with  $C(A, \kappa, \lambda)$  independent of  $t_f$ , of  $u$ , and of the mesh parameters.

*Proof.* Starting from (31), we use Lemmata 5.1 and 5.3 and, subsequently, apply (28) to arrive at

$$\frac{d}{dt} (\|\vartheta(t)\|^2 + \|\sqrt{A}\nabla\vartheta(t)\|^2) + \min\{1, c_0\} (\|\vartheta\|^2 + \|\sqrt{A}\nabla\vartheta\|^2) \leq c_0^{-1} \|\rho_t\|^2 + \|\sqrt{A}\nabla\rho_t\|^2,$$

upon multiplication by 2. Grönwall's Lemma, thus, implies

$$\begin{aligned} \|\vartheta(t_f)\|^2 + \|\sqrt{A}\nabla\vartheta(t_f)\|^2 &\leq e^{-\min\{1, c_0\}t_f} (\|\vartheta(0)\|^2 + \|\sqrt{A}\nabla\vartheta(0)\|^2) \\ &\quad + \int_0^{t_f} e^{-\min\{1, c_0\}(t_f-t)} (c_0^{-1} \|\rho_t\|^2 + \|\sqrt{A}\nabla\rho_t\|^2) \, dt. \end{aligned} \quad (43)$$

As before, we also have

$$\|\rho_t\|^2 + \|\sqrt{A}\nabla\rho_t\|^2 \leq \max\{1, c_0^{-1}\} \|\sqrt{B}\nabla\rho_t\|^2 \leq \frac{C(A, \kappa, \lambda)}{\min\{1, c_0\}} \sum_{T \in \mathcal{T}} \frac{h_T^{2s_T} x_T}{p^{2(k_T-3)}} |u_t|_{H^{k_T}(T)}^2.$$

The last bound combined with (43) and the completely analogous estimate

$$\|\rho(t_f)\|^2 + \|\sqrt{A}\nabla\rho(t_f)\|^2 \leq \max\{1, c_0^{-1}\} \|\sqrt{B}\nabla\rho(t_f)\|^2,$$

already imply the result.  $\square$

Therefore, (21) admits completely analogous long time properties compared to the PDE problem (10), (20).

**Remark 5.7.** *The first term on the right-hand side of (43) vanishes upon altering the finite element method (21) at  $t = 0$  from  $U(0) = \Pi u_0$  to  $U(0) = \pi u_0$ . In practical terms, this results in a computational overhead of a stiffness matrix solve as opposed to a mass matrix one.*

The a priori error estimate derived in Theorem 5.5 is optimal with respect to mesh-size  $h$ , upon observing that the bilinear form  $\mathcal{B}$  is the weak form of a formally 4th order differential operator. At the same time, the highest order terms involving  $A$  of  $\mathcal{B}(w, v)$  play the role of “stabilisation” in this non-standard Galerkin context, resulting in increase of the spatial operator order from second to fourth. The estimate (40) is slightly suboptimal with respect to the polynomial degree  $p$ , as is typical in error analyses of  $hp$ -version interior penalty procedures, such as (21), involving inverse estimates [13]. Correspondingly, the a priori error bound in Theorem 5.6 is formally suboptimal with respect to the mesh size  $h$  by one order. For instance, for  $p = 2$  and  $k_T = 3$ ,  $T \in \mathcal{T}$ , we have first order convergence with respect to the mesh-size  $h$ . This is because the errors are measured in weaker norms than the ones appearing on the right-hand side of (42). We view this as a reasonable price to pay for the long time robustness of the method when  $t_f \rightarrow \infty$ . Indeed, in many practical scenarios, such as long time computations of decay to equilibrium distributions, the favourable exponential dependence with respect to  $t_f$  may be preferable to a slight suboptimality in the rate of convergence with respect to the mesh size  $h$ . Moreover, given the typically high smoothness of the exact solution  $u$ , the use of high order finite element spaces may diminish further the practical significance of this slight  $h$ -suboptimality.

Nevertheless, to address the aforementioned slight  $h$ -suboptimality in Theorem 5.6, one may be tempted to consider scaling of the stabilisation terms involving  $A$  by appropriate powers of the mesh-size  $h$ . This will reduce the stiffness matrix scaling to one of a second order operator; this is a standard practice in classical Galerkin contexts, e.g., for streamline upwinded Petrov-Galerkin methods for convection-dominated problems. In the present context, however, such scaling would introduce new challenges, most important of which is that the spectral gap constant  $c_0$  would be proportional to a negative power of  $h$ . Although such dependence may not turn out to be catastrophic in certain scenarios, e.g., when  $u_t$  decays very fast (cf., (42) in Theorem 5.6), this is a challenging question that requires further research.

Another possibility to retrieve optimality in the rate of convergence for the method (21) in weaker norms is the incorporation of Aubin-Nitsche type duality techniques. Using the  $h$ -optimal error control of the (hypo)elliptic projection error  $\|u - \pi u\|$ , along with regularity estimates, it is conceivable that the order of convergence with respect to  $h$  and  $p$  in Theorem 5.6 can be improved, since the  $H^1$ -seminorm is weaker than  $\|\cdot\|$ . Although this is classical in parabolic problems, it appears to be a challenge in the present context of hypocoercive operators. This is due to the non-closedness of the hypocoercivity property with respect to the adjoint operation in the present variational context (12). This constitutes an interesting direction for future research.

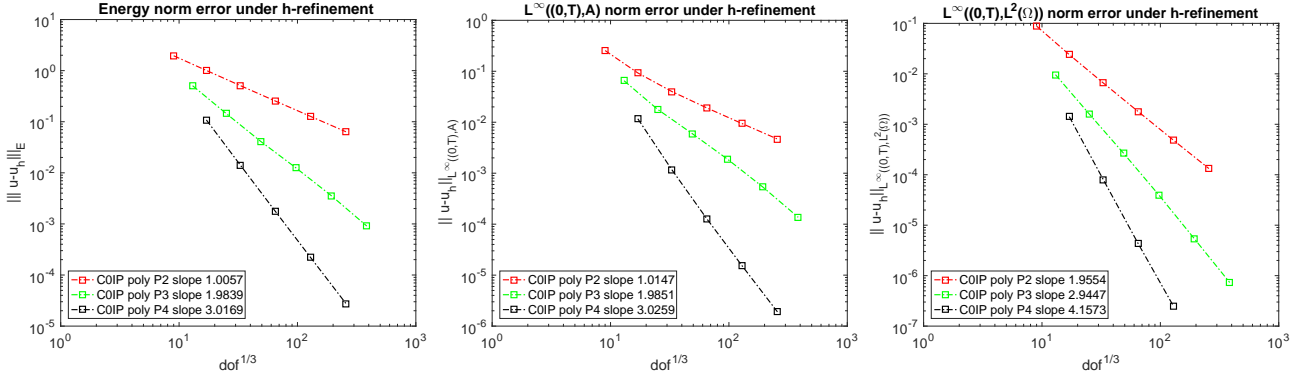


Figure 1: Example 1. Convergence history under  $h$ -refinement for the errors  $|||e|||$  (left),  $||\sqrt{A}\nabla e||_{L^\infty(0,t_f;L_2(\Omega))}$  (centre),  $||e||_{L^\infty(0,t_f;L_2(\Omega))}$  (right), for  $p = 2, 3, 4$ , against the square root of total space-time degrees of freedom (Total DoF) $^{1/2}$ .

## 6 Numerical experiments

We now present some basic numerical experiments aiming to verify the convergence rates predicted by the theory. For the (not studied theoretically) time discretization, we have implemented a discontinuous Galerkin timestepping method of degree  $p - 2$  with  $p$  being the local polynomial degree of the spatial discretization; this choice formally balances the expected rates of convergence, while maintaining a practical computational cost. For the experiments below we have made the following parameter choices:  $t_f = 0.5$ ,  $\Omega = (0, 1)^2$ ,  $\kappa = \lambda = 0$ ,  $\alpha = 0.35$ ,  $\beta = \alpha^2$ ,  $\gamma = \alpha^3$ , and  $C_\tau = 10$ . We note that altering  $\alpha$  within the permitted range does not appear to have essential impact on the errors observed.

### 6.1 Example 1

We first consider the problem (10) and  $u_0, f$  so that the exact solution is given by the smooth time-independent periodic function  $u(t, x, y) := \sin^2(\pi x) \sin^2(\pi y)$ . The purpose of this example is to observe the convergence of the spatial discretization.

In Figure 1, the convergence rates under  $h$ -refinement are presented for the errors  $|||e|||$  (left),  $||\sqrt{A}\nabla e||_{L^\infty(0,t_f;L_2(\Omega))}$  (centre),  $||e||_{L^\infty(0,t_f;L_2(\Omega))}$  (right), and for local polynomial degrees  $p = 2, 3, 4$  against the square root of total space-time degrees of freedom (Total DoF) $^{1/2}$ . The space-time subdivision is constructed by tensorizing a sequence of spatial triangular meshes of 32, 128, 512, 2,048, 8,192, and 32,768 elements with the whole time interval  $[0, t_f]$ ; there is no error in time, so one time step is sufficient. For the errors  $|||e|||$  and  $||\sqrt{A}\nabla e||$ , we observe convergence rates  $O(h^{p-1})$  as expected by the theory, while the error  $||e||$  converges at the higher rate  $O(h^p)$  although the theory currently only predicts  $O(h^{p-1})$ .

### 6.2 Example 2

We now consider the problem (10) and  $u_0, f$  so that the exact solution is given by the smooth function  $u(t, x, y) := (2 - t)^{-1} \exp(-(x - 1/2)^2 - (y - 1/2)^2) \sin(\pi(x - t)) \sin^2(\pi y)$ . Note that (20) is satisfied. The purpose of this experiment is to observe the convergence in the practical setting of the discontinuous Galerkin timestepping of degree  $p - 2$  in conjunction with  $p$ -th polynomial degree spatial discretization (21).

The space-time subdivision is constructed by tensorizing a sequence of spatial triangular meshes of 32, 128, 512, 2,048, and 8,192 elements with respective subdivisions of  $[0, t_f]$  into

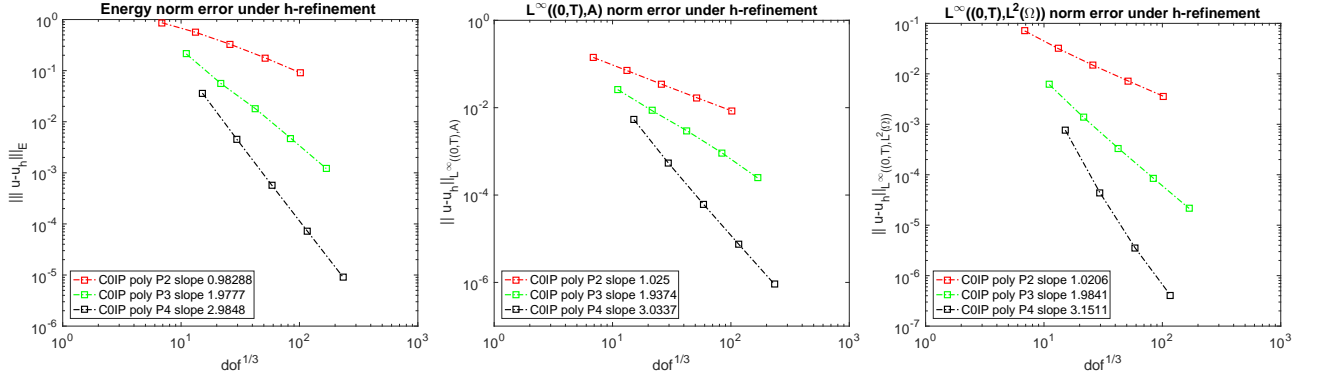


Figure 2: Example 2. Convergence history under  $h$ -refinement for the errors  $|||e|||$  (left),  $\|\sqrt{A}\nabla e\|_{L^\infty(0,t_f;L_2(\Omega))}$  (centre),  $\|e\|_{L^\infty(0,t_f;L_2(\Omega))}$  (right), for  $p = 2, 3, 4$ , against the cubic root of total space-time degrees of freedom (Total DoF) $^{1/3}$ .

4, 8, 16, 32, and 64 time steps; notice that this choice corresponds to balanced spatial and temporal discretization parameters. In Figure 2, the convergence rates under  $h$ -refinement are presented for the errors  $|||e|||$  (left),  $\|\sqrt{A}\nabla e\|_{L^\infty(0,t_f;L_2(\Omega))}$  (centre),  $\|e\|_{L^\infty(0,t_f;L_2(\Omega))}$  (right), and for local polynomial degrees  $p = 2, 3, 4$  against the cubic root of total space-time degrees of freedom (Total DoF) $^{1/3}$ . We observe convergence rates  $O(h^{p-1})$  for all errors. We expect that the reduced order of convergence for the error  $\|e\|_{L^\infty(0,t_f;L_2(\Omega))}$  compared to Example 1 is due to the lower order temporal discretization.

We finally note that the behaviour of the proposed method in practical scenarios with regard to the dependence on the final time  $t_f$  is a rather subtle and interesting topic in its own right and it will be discussed in detail elsewhere [7].

## Acknowledgements

I am sincerely grateful to Zhaonan (Peter) Dong (Cardiff, UK) for helpful comments on an earlier version of this work and for providing a computer code used for the numerical experiments. The financial support of The Leverhulme Trust (research project grant no. RPG-2015-306) is also gratefully acknowledged. Also, this research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant” (Project Number: 3270).

## References

- [1] G. AKRIVIS, M. CROUZEIX, AND V. THOMÉE, *Numerical methods for ultraparabolic equations*, Calcolo, 31 (1994), pp. 179–190 (1996).
- [2] M. BESSEMOULIN-CHATARD AND F. FILBET, *A finite volume scheme for nonlinear degenerate parabolic equations*, SIAM J. Sci. Comput., 34 (2012), pp. B559–B583.
- [3] M. BESSEMOULIN-CHATARD, M. HERDA, AND T. REY, *Hypoocoercivity and diffusion limit of a finite volume scheme for linear kinetic equations*, Math. Comp., 89 (2020), pp. 1093–1133.

- [4] A. CANGIANI, Z. DONG, E. H. GEORGOULIS, AND P. HOUSTON, *hp-version discontinuous Galerkin methods on polygonal and polyhedral meshes*, SpringerBriefs in Mathematics, Springer, Cham, 2017.
- [5] A. CHERNOV, *Optimal convergence estimates for the trace of the polynomial  $L^2$ -projection operator on a simplex*, Math. Comp., 81 (2012), pp. 765–787.
- [6] J. DOLBEAULT, C. MOUHOT, AND C. SCHMEISER, *Hypocoercivity for linear kinetic equations conserving mass*, Trans. Amer. Math. Soc., 367 (2015), pp. 3807–3828.
- [7] Z. DONG AND E. H. GEORGOULIS, In preparation.
- [8] G. DUJARDIN, F. HÉRAU, AND P. LAFITTE, *Coercivity, hypocoercivity, exponential time decay and simulations for discrete Fokker-Planck equations*, Numer. Math., 144 (2020), pp. 615–697.
- [9] J.-P. ECKMANN AND M. HAIRER, *Spectral properties of hypoelliptic operators*, Comm. Math. Phys., 235 (2003), pp. 233–253.
- [10] A. ERN AND J.-L. GUERMOND, *Discontinuous Galerkin methods for Friedrichs’ systems. II. Second-order elliptic PDEs*, SIAM J. Numer. Anal., 44 (2006), pp. 2363–2388.
- [11] G. FICHERA, *Sulle equazioni differenziali lineari ellittico-paraboliche del secondo ordine*, Atti Accad. Naz. Lincei. Mem. Cl. Sci. Fis. Mat. Nat. Sez. I. (8), 5 (1956), pp. 1–30.
- [12] E. L. FOSTER, J. LOHÉAC, AND M.-B. TRAN, *A structure preserving scheme for the Kolmogorov-Fokker-Planck equation*, J. Comput. Phys., 330 (2017), pp. 319–339.
- [13] E. H. GEORGOULIS, E. HALL, AND J. M. MELENK, *On the suboptimality of the  $p$ -version interior penalty discontinuous Galerkin method*, J. Sci. Comput., 42 (2010), pp. 54–67.
- [14] F. HÉRAU AND F. NIER, *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential*, Arch. Ration. Mech. Anal., 171 (2004), pp. 151–218.
- [15] L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.
- [16] P. HOUSTON, C. SCHWAB, AND E. SÜLI, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.
- [17] A. KOLMOGOROFF, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2), 35 (1934), pp. 116–117.
- [18] H. LI AND J. SHEN, *Optimal error estimates in Jacobi-weighted Sobolev spaces for polynomial approximations on the triangle*, Math. Comp., 79 (2010), pp. 1621–1646.
- [19] C. MOUHOT AND L. NEUMANN, *Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus*, Nonlinearity, 19 (2006), pp. 969–998.
- [20] O. A. OLEĬNIK AND E. V. RADKEVIČ, *Second order equations with nonnegative characteristic form*, Plenum Press, New York-London, 1973. Translated from the Russian by Paul C. Fife.

- [21] A. PORRETTA AND E. ZUAZUA, *Numerical hypocoercivity for the Kolmogorov equation*, Math. Comp., 86 (2017), pp. 97–119.
- [22] H. RISKEN, *The Fokker-Planck equation*, vol. 18 of Springer Series in Synergetics, Springer-Verlag, Berlin, second ed., 1989. Methods of solution and applications.
- [23] C. SCHWAB, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.
- [24] C. VILLANI, *Hypocoercivity*, Mem. Amer. Math. Soc., 202 (2009), pp. iv+141.