ESAIM: M2AN 43 (2009) 445–485
DOI: 10.1051/m2an:2008051

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

SPECTRAL GALERKIN APPROXIMATION OF FOKKER-PLANCK EQUATIONS WITH UNBOUNDED DRIFT

David J. Knezevic[1] and Endre Süli[1]

Abstract. This paper is concerned with the analysis and implementation of spectral Galerkin methods for a class of Fokker-Planck equations that arises from the kinetic theory of dilute polymers. A relevant feature of the class of equations under consideration from the viewpoint of mathematical analysis and numerical approximation is the presence of an unbounded drift coefficient, involving a smooth convex potential U that is equal to $+\infty$ along the boundary ∂D of the computational domain D. Using a symmetrization of the differential operator based on the Maxwellian M corresponding to U, which vanishes along ∂D, we remove the unbounded drift coefficient at the expense of introducing a degeneracy, through M, in the principal part of the operator. The general class of admissible potentials considered includes the FENE (finitely extendible nonlinear elastic) model. We show the existence of weak solutions to the initial-boundary-value problem, and develop a fully-discrete spectral Galerkin method for such degenerate Fokker-Planck equations that exhibits optimal-order convergence in the Maxwellian-weighted H^1 norm on D. In the case of the FENE model, we also discuss variants of these analytical results when the Fokker-Planck equation is subjected to an alternative class of transformations proposed by Chauvière and Lozinski; these map the original Fokker-Planck operator with an unbounded drift coefficient into Fokker-Planck operators with unbounded drift and reaction coefficients, that have improved coercivity properties in comparison with the original operator. The analytical results are illustrated by numerical experiments for the FENE model in two space dimensions.

Mathematics Subject Classification. 65M70, 65M12, 35K20, 82C31, 82D60.

Received January 29, 2008. Revised July 17, 2008.
Published online December 17, 2008.

1. Introduction

This paper is concerned with the numerical approximation of the Fokker-Planck equation

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_x \cdot (\underset{\sim}{u}(\underset{\sim}{x},t)\,\psi) + \underset{\sim}{\nabla}_q \cdot \left((\underset{\approx}{\nabla}_x \underset{\sim}{u})\underset{\sim}{q}\,\psi \right) = \varepsilon\,\Delta_x \psi + \frac{1}{2\,\lambda}\,\underset{\sim}{\nabla}_q \cdot \left(\underset{\sim}{\nabla}_q \psi + \underset{\sim}{F}(\underset{\sim}{q})\,\psi \right), \tag{1.1}$$

that arises from the kinetic theory of dilute polymers [12,13]; see also [4,5,7] and references therein. Here, ε and λ are two positive parameters, referred to as *centre-of-mass diffusion coefficient* and *relaxation time*, respectively, $\Omega \subset \mathbb{R}^d$ is the flow-domain of the polymer and $D \subset \mathbb{R}^d$ is the set of admissible orientation vectors of polymer chains. Typically $D = B(\underset{\sim}{0}; \sqrt{b}\,)$, where $b > 0$ is a nondimensional parameter that measures the

Keywords and phrases. Spectral methods, Fokker-Planck equations, transport-diffusion problems, FENE.

[1] OUCL, University of Oxford, Parks Road, Oxford, OX1 3QD, UK. david.knezevic@balliol.ox.ac.uk; davek@comlab.ox.ac.uk; endre.suli@comlab.ox.ac.uk

Article published by EDP Sciences

maximum possible extension of polymer chains, and $B(\underset{\sim}{0};s)$ is the open ball with radius s centred at the origin $\underset{\sim}{0}$ in $\mathbb{R}^d$, $d \in \{2,3\}$. Henceforth, unless otherwise stated, D will denote $B(\underset{\sim}{0};\sqrt{b}\,)$.

Equation (1.1) governs the evolution, over a nonempty, bounded and closed time interval $[0,T]$, of the probability density function $\psi : (\underset{\sim}{x},\underset{\sim}{q},t) \in \Omega \times D \times [0,T] \mapsto \psi(\underset{\sim}{x},\underset{\sim}{q},t)$ of a $2d$-component stochastic process that models random fluctuations of polymer molecules in a solvent due to thermal agitation. The solvent is an incompressible Newtonian fluid with velocity $\underset{\sim}{u}$ whose motion is governed by the Navier-Stokes equation forced by the divergence of the non-Newtonian extra stress tensor, defined as the integral of $\underset{\sim}{F}(\underset{\sim}{q}) \otimes \underset{\sim}{q}\psi(\underset{\sim}{x},\underset{\sim}{q},t)$ over D. In the simplest models of this kind, elastic effects are incorporated by modelling the polymer chains as dumbbells, *i.e.*, as pairs of massless beads connected by an elastic spring, with spring force $\underset{\sim}{F} : D \to \mathbb{R}^d$ defined by a *spring potential* $U : \mathbb{R}_{\geq 0} \to \mathbb{R}$ through

$$\underset{\sim}{F}(\underset{\sim}{q}) := U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)\,\underset{\sim}{q}, \qquad \underset{\sim}{q} \in D. \tag{1.2}$$

We adopt the following structural hypotheses.

Hypothesis A. The spring potential $U \in \mathrm{C}^1([0,\frac{b}{2}))$ is a non-negative monotonic increasing function, with $U(0)=0$, $\lim_{s\to b/2_-} U(s) = +\infty$, $\lim_{s\to b/2_-}(\frac{b}{2}-s)U'(s) < \infty$.

Hypothesis A is consistent with the physical requirement that, in order to faithfully model *finite* stretching of polymer chains, the spring force $\underset{\sim}{F}(\underset{\sim}{q})$ should have infinite intensity when the maximum admissible elongation $|\underset{\sim}{q}| = \sqrt{b}$ is reached; *i.e.*, the function $\underset{\sim}{q} \mapsto U'(\frac{1}{2}|\underset{\sim}{q}|^2)$ should tend to $+\infty$ as $\mathfrak{d}(\underset{\sim}{q}) := \mathrm{dist}(\underset{\sim}{q},\partial D) = \sqrt{b}-|\underset{\sim}{q}| \to 0_+$.

Given a spring potential U, consider the associated (normalized) *Maxwellian* M defined by

$$\underset{\sim}{q} \mapsto M(\underset{\sim}{q}) := \frac{1}{C(b)} \exp\left(-U(\tfrac{1}{2}|\underset{\sim}{q}|^2)\right) \in \mathrm{L}^1(D), \qquad \text{where} \quad C(b) := \int_D \exp\left(-U(\tfrac{1}{2}|\underset{\sim}{q}|^2)\right) \mathrm{d}\underset{\sim}{q}.$$

Since, by Hypothesis A, $U(\frac{1}{2}|\underset{\sim}{q}|^2) \to +\infty$ as $\mathfrak{d}(\underset{\sim}{q}) \to 0_+$, we have that $M(\underset{\sim}{q}) \to 0_+$ as $\mathfrak{d}(\underset{\sim}{q}) \to 0_+$.

Hypothesis B. $\sqrt{M} \in \mathrm{H}^1_0(D)$, and M is a *weight function of type 3* on D in the sense of Triebel [33], p. 247, Definition 3.2.1.3c; *i.e.*, there exist positive constants c_1, c_2 and λ, and a positive monotonic increasing function τ, defined on the interval $(0,\lambda)$, such that $c_1\,\tau(\mathfrak{d}(\underset{\sim}{q})) \leq M(\underset{\sim}{q}) \leq c_2\,\tau(\mathfrak{d}(\underset{\sim}{q}))$ for all $\underset{\sim}{q} \in D$ satisfying $\mathfrak{d}(\underset{\sim}{q}) < \lambda$.

Example 1.1. Consider the function U defined by

$$U(s) := -f(s)\ln\left(1-\frac{2s}{b}\right), \qquad s \in [0,\tfrac{b}{2}), \qquad \text{with } b > 2,$$

where $f \in \mathrm{C}^1[0,\frac{b}{2}]$ is a nondecreasing function, positive on $(0,\frac{b}{2}]$, with $f(\frac{b}{2}) > 1$; then U and the associated Maxwellian M satisfy Hypotheses A and B, respectively.

Hypotheses A and B will be assumed throughout the paper. In Section 2 we shall also invoke the following additional assumption, which is not required elsewhere.

Hypothesis C. With U as in Hypothesis A and $L(s) := \ln(1-\frac{2s}{b})$, there exist ω and γ in $\mathbb{R}$ such that the mapping $\underset{\sim}{q} \mapsto (U+\gamma L)(\frac{1}{2}|\underset{\sim}{q}|^2)$ is ω-*convex* on D in the following sense: there exists $c_0 \in \mathbb{R}_{>0}$ such that, for each $\underset{\sim}{q} \in D$, the Hessian

$$H(\underset{\sim}{q}) := \left(\frac{\partial^2}{\partial q_i\,\partial q_j}\,(U+\gamma L)(\tfrac{1}{2}|\underset{\sim}{q}|^2)\right)$$

of $\underset{\sim}{q} \mapsto (U+\gamma L)(\frac{1}{2}|\underset{\sim}{q}|^2)$ satisfies $H(\underset{\sim}{q}) \geq c_0(1-|\underset{\sim}{q}|^2/b)^{\omega}\,\mathrm{Id}$, where Id is the $d\times d$ identity matrix.

Example 1.2. In the case of the FENE (finitely extendible nonlinear elastic) polymer model

$$U(s) := -\frac{b}{2}\ln\left(1-\frac{2s}{b}\right), \qquad U'(s) = \frac{1}{1-\frac{2s}{b}}, \qquad s \in [0,\tfrac{b}{2}), \qquad \text{with } b > 2.$$

It will be shown in Section 2 that the function $\underset{\sim}{q} \in D = B(\underset{\sim}{0}; \sqrt{b}) \mapsto (U + \gamma L)(\frac{1}{2}|q|^2)$ is ω-convex with $\omega = -1$ (or, briefly, (-1)-convex) and $c_0 = (b - 2\gamma)/b$, for all $\gamma \in [0, 1]$. Thus the FENE spring potential U satisfies Hypothesis C with $\omega = -1$ and any $\gamma \in [0, 1]$. The associated normalized Maxwellian is

$$M(\underset{\sim}{q}) = \frac{1}{C(b)} \left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{\frac{b}{2}}, \qquad \underset{\sim}{q} \in D = B(\underset{\sim}{0}; \sqrt{b}\,).$$

Clearly, there exist positive constants c_1 and c_2 such that $c_1 \le M(\underset{\sim}{q})/[\mathfrak{d}(\underset{\sim}{q})]^{b/2} \le c_2$ for all $\underset{\sim}{q} \in D$ (*i.e.*, $\lambda = \sqrt{b}$); hence M is a weight function of type 3 on D. Also, thanks to the assumption $b > 2$, $\sqrt{M} \in \mathrm{H}^1_0(D)$. For $b \gg 1$, M decays to 0 very rapidly as $\underset{\sim}{q}$ approaches ∂D. In numerical simulations typically $b \in [10, 100]$.

Following Kolmogorov [24], the Fokker-Planck equation can be recast as follows:

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_x \cdot (\,\underset{\sim}{u}(\underset{\sim}{x}, t)\psi) + \underset{\sim}{\nabla}_q \cdot \left(\underset{\approx}{\kappa}(\underset{\sim}{x}, t)\,\underset{\sim}{q}\,\psi\right) = \varepsilon\,\Delta_x \psi + \frac{1}{2\,\lambda}\,\underset{\sim}{\nabla}_q \cdot \left(M(\underset{\sim}{q})\underset{\sim}{\nabla}_q \left(\frac{\psi}{M}\right)\right),$$

where $\underset{\approx}{\kappa}(\underset{\sim}{x}, t) := (\underset{\sim}{\nabla}_x\,\underset{\sim}{u})$. The probability density ψ is a function of $2d+1$ independent variables: $\underset{\sim}{x} \in \mathbb{R}^d$, $\underset{\sim}{q} \in \mathbb{R}^d$ and $t \in \mathbb{R}_{\ge 0}$. Since the dependence of the coefficients in the equation on $\underset{\sim}{x}$ and $\underset{\sim}{q}$ is separated/factorized, an efficient approach to the numerical solution of this equation in $2d + 1$ variables is based on operator-splitting with respect to $(\underset{\sim}{q}, t)$ and $(\underset{\sim}{x}, t)$; see Chauvière and Lozinski [17,18,27]. Thereby, the resulting time-dependent transport-diffusion equation with respect to $(\underset{\sim}{x}, t)$ is completely standard, $\psi_t + \underset{\sim}{\nabla}_x \cdot (\underset{\sim}{u}(\underset{\sim}{x}, t)\psi) = \varepsilon\Delta_x\psi$, while the transport-diffusion equation with respect to $(\underset{\sim}{q}, t)$ is

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_q \cdot (\underset{\approx}{\kappa}\,\underset{\sim}{q}\psi) = \frac{1}{2\lambda}\underset{\sim}{\nabla}_q \cdot \left(M(\underset{\sim}{q})\underset{\sim}{\nabla}_q \left(\frac{\psi}{M}\right)\right), \qquad (\underset{\sim}{q}, t) \in D \times (0, T]. \tag{1.3}$$

Equation (1.3) is supplemented with the following initial and boundary conditions:

$$\psi(\underset{\sim}{q}, 0) = \psi_0(\underset{\sim}{q}), \qquad \text{for all } \underset{\sim}{q} \in D, \tag{1.4}$$

$$\psi(\underset{\sim}{q}, t) = o\left(\sqrt{M(\underset{\sim}{q})}\,\right), \qquad \text{as } \mathfrak{d}(\underset{\sim}{q}) \to 0_+, \ \text{ for all } t \in (0, T]. \tag{1.5}$$

Here, the initial datum ψ_0 is such that $\psi_0 \ge 0$ and $\int_D \psi_0(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q} = 1$.

The central difficulty, from both the analytical and the computational point of view, is now the presence in (1.3) of the degenerate Maxwellian $M(\underset{\sim}{q})$, with $\lim_{\mathfrak{d}(\underset{\sim}{q}) \to 0_+} M(\underset{\sim}{q}) = 0$. Thus we shall ignore the coupling between the Fokker-Planck equation and the Navier-Stokes system, suppress the dependence of the probability density function ψ on the variable $\underset{\sim}{x}$, assume that the $d \times d$ tensor $\underset{\approx}{\kappa} = \underset{\sim}{\nabla}_x\,\underset{\sim}{u}$ is independent of $\underset{\sim}{x}$, belongs to $(\mathrm{C}[0, T])^{d\times d}$ and is such that $\mathrm{tr}(\underset{\approx}{\kappa})(t) = 0$ for all $t \in [0, T]$, and we focus our attention on the numerical solution of (1.3), (1.4), (1.5). For theoretical results concerning the existence of weak solutions to coupled Navier-Stokes-Fokker-Planck systems and a detailed survey of related literature we refer to [4,6,7] and [26].

Most numerical methods developed for the Fokker-Planck equation have been based on the 'original' form,

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_q \cdot \left(\underset{\approx}{\kappa}\underset{\sim}{q}\,\psi\right) = \frac{1}{2\,\lambda}\,\underset{\sim}{\nabla}_q \cdot \left(\underset{\sim}{\nabla}_q\psi + \underset{\sim}{F}(\underset{\sim}{q})\,\psi\right), \tag{1.6}$$

see, for example, [17,18,27] or [1,2]. From the theoretical viewpoint at least, the advantage of (1.3) over (1.6), is that on transformation into weak form the diffusion operator becomes symmetric (see (1.7)), which facilitates the analysis of the Fokker-Planck equation for a general class of Maxwellians. Notwithstanding this potential theoretical advantage, the computational benefits, or otherwise, of discretizing (1.3) rather then (1.6) remain to be understood.

The aim of this paper is therefore two-fold:

(a) Our principal objective is to develop the mathematical and numerical analysis of equation (1.3) for a general class of Maxwellians. The discretization of the equation is based on a spectral Galerkin method in the spatial variable $\underset{\sim}{q}$ coupled with backward Euler time-stepping. One can, of course, consider more accurate time discretization schemes, such as an nth-order backward differentiation formula, BDFn, $n \in \{2, \ldots, 6\}$, for example. High-order time discretization of the problem is, however, a secondary consideration to the central theme of the paper, and we do not discuss it here.

(b) In the special case of the FENE model, we shall show how the results under (a) can be adapted to the case of an alternative discretization proposed in [17,18,27], which applies a transformation, different from Kolmogorov's symmetrizing transformation considered under (a), to the 'original' form (1.6) of the Fokker-Planck equation. The transformed equation is then approximated in the same way as in (a), using a spectral Galerkin method in space and a backward Euler discretization in time.

Since the analytical arguments under (b) are almost identical to those under (a), for the sake of brevity we shall focus our attention on (a), but we shall systematically indicate the key adjustments that need to be made in order to obtain the corresponding results under (b). We begin by defining the relevant function spaces.

Let

$$\mathfrak{H} := \left\{ \varphi \in \mathrm{L}^2_{\mathrm{loc}}(D) \,:\, \int_D \left(\frac{\varphi}{\sqrt{M}} \right)^2 \mathrm{d}\underset{\sim}{q} < \infty \right\}, \quad \mathfrak{K} := \left\{ \varphi \in \mathfrak{H} \,:\, \int_D \left(\left(\frac{\varphi}{\sqrt{M}} \right)^2 + \left| \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\frac{\varphi}{M} \right) \right|^2 \right) \mathrm{d}\underset{\sim}{q} < \infty \right\}$$

and define $\mathfrak{K}_0$ as the closure of $\sqrt{M}\mathrm{C}_0^\infty(D)$ in the norm of $\mathfrak{K}$. Taking our test functions as φ/M with $\varphi \in \mathfrak{K}_0$, we obtain the following weak formulation of the initial-boundary-value problem (1.3).

Given $\psi_0 \in \mathfrak{H}$, find $\psi \in \mathrm{L}^\infty(0,T;\mathfrak{H}) \cap \mathrm{L}^2(0,T;\mathfrak{K}_0)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_D \frac{\psi\varphi}{M}\, \mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}\, \underset{\sim}{q}) \frac{\psi}{\sqrt{M}} \cdot \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\frac{\varphi}{M} \right) \mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda} \int_D \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\frac{\psi}{M} \right) \cdot \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\frac{\varphi}{M} \right) \mathrm{d}\underset{\sim}{q} = 0 \quad \forall \varphi \in \mathfrak{K}_0, \tag{1.7}$$

in the sense of distributions on $(0,T)$, and $\psi(\cdot,0) = \psi_0(\cdot)$.

Now, by introducing the notation

$$\hat{\varphi} := \frac{\varphi}{\sqrt{M}} \qquad \text{and} \qquad \underset{\sim}{\nabla}_M \hat{\varphi} := \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\frac{\hat{\varphi}}{\sqrt{M}} \right)$$

we can reformulate (1.7) on observing that, by the definition of $\mathfrak{K}$, we have $\varphi \in \mathfrak{K}_0$ if, and only if, $\hat{\varphi} \in \mathrm{H}^1_0(D;M)$, where $\mathrm{H}^1_0(D;M)$ is the closure of $\mathrm{C}_0^\infty(D)$ in the norm of $\mathrm{H}^1(D;M)$, and

$$\mathrm{H}^1(D;M) := \left\{ \zeta \in \mathrm{L}^2(D) : \|\zeta\|^2_{\mathrm{H}^1(D;M)} := \int_D \left(|\zeta|^2 + |\underset{\sim}{\nabla}_M \zeta|^2 \right) \mathrm{d}\underset{\sim}{q} < \infty \right\}.$$

When applied to an element of $\mathrm{H}^1_0(D;M)$ the norm $\|\cdot\|_{\mathrm{H}^1(D;M)}$ will be written $\|\cdot\|_{\mathrm{H}^1_0(D;M)}$. As a matter of fact, we shall show below that $\mathrm{C}_0^\infty(D)$ is dense in $\mathrm{H}^1(D;M)$ and therefore, perhaps somewhat counter-intuitively, $\mathrm{H}^1_0(D;M) = \mathrm{H}^1(D;M)$, and also $\mathfrak{K}_0 = \mathfrak{K}$.

Remark 1.3. We note in passing that the substitution $\hat{\varphi} = \varphi/\sqrt{M}$ also appears in the recent paper by Du *et al.* [20], though the operator $\underset{\sim}{\nabla}_M$ does not.

In the case of the FENE Maxwellian (*cf.* Ex. 1.2), Chauvière and Lozinski [17,18,27] used a spectral method to approximate $\psi/M^{2s/b}$ instead of $\psi/\sqrt{M}$, where s is a parameter that was chosen on the basis of numerical experiments. Clearly, the two expressions coincide when $s = b/4$; on the other hand, the values $s = 2$ and $s = 2.5$ were recommended in [17,18,27] on computational grounds for $d = 2$ and $d = 3$, respectively. More will be said in Sections 3, 4 and 6 about the analytical implications of using, in the special case of the FENE model,

the substitution $\hat{\psi} := \psi/M^{2s/b}$ instead of the substitution $\hat{\psi} := \psi/\sqrt{M}$. In particular, we shall show that both substitutions result in unconditionally stable and convergent numerical methods, although in the case of the Chauvière and Lozinski type substitution it will be necessary to assume for this purpose that $b \geq 4s^2/(2s-1)$ with $s > 1/2$, while the Kolmogorov symmetrization will be seen to result in a stable and optimally convergent scheme for all $b > 2$. In Section 7 we shall perform quantitative comparisons of the two approaches through numerical experiments.

With the notational conventions defined above, (1.7) has the following form.

Given $\hat{\psi}_0 := \psi_0/\sqrt{M} \in \mathrm{L}^2(D)$, find $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}\,\underset{\sim}{q})\hat{\psi}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_M\hat{\psi}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} = 0 \qquad \forall\hat{\varphi}\in\mathrm{H}^1_0(D;M), \tag{1.8}$$

in the sense of distributions on $(0,T)$, and $\hat{\psi}(\cdot,0) = \hat{\psi}_0(\cdot)$.

The function space $\mathrm{H}^1_0(D;M)$ may appear exotic. We shall see however in Section 2 that this is not so: it will be shown in Section 2.2 that, under Hypotheses A and B, $\mathrm{H}^1(D;M) = \mathrm{H}^1_0(D;M)$ and $\mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D;M)$. We shall further illuminate the structure of Maxwellian-weighted spaces by applying the Brascamp-Lieb inequality with a probability measure whose Radon-Nikodým derivative is the Maxwellian. The connection between $\mathrm{H}^1_0(D;M)$ and $\mathrm{H}^1_0(D)$ will prove helpful in the development of Galerkin methods for (1.8), since the construction of finite-dimensional subspaces of $\mathrm{H}^1_0(D)$ and the analysis of their approximation properties are well-understood.

In Section 3 we shall revisit the weak formulation (1.8) of the initial boundary value problem. We shall construct a backward Euler semidiscretization of the weak formulation and show that this has a unique solution. We shall then use a compactness argument to establish the existence of weak solutions to the initial-boundary-value problem. We also show the uniqueness of the weak solution. In the process, we shall prove the unconditional stability of the temporal semidiscretization in the $\ell^\infty(0,T;\mathrm{L}^2(D))$ and $\ell^2(0,T;\mathrm{H}^1_0(D;M))$ norms. Our arguments do not invoke compact embedding of (Maxwellian-)weighted Sobolev spaces, and no growth/decay conditions (such as a Muckenhaupt condition) need to be imposed on the Maxwellian M beyond the conditions on U and M stated in Hypotheses A and B above. Elliptic and parabolic operators with unbounded drift coefficients, albeit in nonconservative form, have been considered recently by Cerrai, Da Prato, Lunardi and others (see, for example, [16,19]); the technique herein, based on semidiscretization in time and passage to the limit using a weak compactness argument, is different from the semigroup theoretic approach used in those papers. We also show how, in the case of the FENE model with $b \geq 4s^2/(2s-1)$ and $s > 1/2$, our results can be carried across, independent of the spatial dimension d, to a weak formulation that results from using the alternative substitution $\hat{\psi} := \psi/M^{2s/b}$; the cases of $s = 2$ and $s = 2.5$ correspond to the methods proposed by Chauvière and Lozinski [17,18,27] for $d = 2$ and $d = 3$, respectively.

In Section 4 we develop the fully-discrete method and, using the stability results from Section 3, we derive a bound on the global error in terms of the approximation error in a suitably defined spectral projection operator.

In Section 5 we give the precise definition of our projection operator: its nonstandard form stems from a *decomposition lemma*, Lemma 5.2, for elements of the Sobolev space $\mathrm{H}^1(D)$ in polar co-ordinates. The result can be seen as a Sobolev space variant of the Malgrange preparation theorem [22].

We complete our convergence analysis in Section 6 by showing that, under Hypotheses A and B, the method exhibits optimal-order convergence in the Maxwellian-weighted norm $\|\cdot\|_{\ell^2(0,T;\mathrm{H}^1_0(D;M))}$ with respect to the spatial and temporal discretization parameters.

Section 7 is devoted to numerical experiments that illustrate the performance of the method. Since the case of two space dimensions ($d = 2$) is sufficiently representative, for ease of presentation in Sections 5, 6 and 7 we have confined ourselves to this case; all of our results in Sections 5 and 6 have obvious extensions to three space dimensions. The stability bounds and existence and uniqueness results presented in Sections 3 and 4 are valid in any number of space dimensions.

2. The Brascamp-Lieb inequality

Suppose that D is a convex open set, $D \subset \mathbb{R}^d$ (*e.g.*, $D = B(\underset{\sim}{0}; \sqrt{b}\,)$, $b > 0$). Consider a probability measure μ supported on D with density $\exp(-V(\underset{\sim}{q}))$, $\underset{\sim}{q} \in D$, with respect to the Lebesgue measure $\mathrm{d}\underset{\sim}{q}$ on $\mathbb{R}^d$, where V is a convex function on D; μ is usually referred to as *Boltzmann measure*. In particular,

$$\mu(B) = \int_B \mathrm{d}\mu = \int_B \exp(-V(\underset{\sim}{q}))\,\mathrm{d}\underset{\sim}{q},$$

for any μ-measurable set $B \subset D$, with $\mu(D) = 1$. The following geometric functional inequality comes from the paper of Bobkov and Ledoux [14].

Theorem 2.1 (Brascamp-Lieb inequality)**.** *Assume that V is a twice continuously differentiable and convex function on a convex open set $D \subset \mathbb{R}^d$, such that, for each $\underset{\sim}{q} \in D$, the Hessian*

$$H(\underset{\sim}{q}) := \left(\frac{\partial^2 V(\underset{\sim}{q})}{\partial q_i\,\partial q_j}\right)$$

is positive definite. Then, for any sufficiently smooth function f,

$$\mathrm{Var}_\mu(f) := \mathbb{E}_\mu[(f - \mathbb{E}_\mu[f])^2] \le \int_D \langle H^{-1}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_q f, \underset{\sim}{\nabla}_q f\rangle\,\mathrm{d}\mu, \qquad \textit{where} \quad \mathbb{E}_\mu[f] = \int_D f\,\mathrm{d}\mu.$$

In terms of simpler notation, the Brascamp-Lieb inequality can be restated as follows:

$$\int_D \left[f(\underset{\sim}{q}) - \int_D f(\underset{\sim}{p})\,\mathrm{e}^{-V(\underset{\sim}{p})}\,\mathrm{d}\underset{\sim}{p}\right]^2 \mathrm{e}^{-V(\underset{\sim}{q})}\,\mathrm{d}\underset{\sim}{q} \le \int_D \langle H^{-1}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_q f, \underset{\sim}{\nabla}_q f\rangle\,\mathrm{e}^{-V(\underset{\sim}{q})}\,\mathrm{d}\underset{\sim}{q},$$

for any sufficiently smooth function f.

Corollary 2.2. *Assume that V is a twice continuously differentiable and ω-convex function on $D = B(\underset{\sim}{0}; \sqrt{b}\,)$, in the sense that there exist $c_0 > 0$ and $\omega \in \mathbb{R}$ such that the Hessian H of V satisfies $H(q) \ge c_0(1 - |\underset{\sim}{q}|^2/b)^\omega\,\mathrm{Id}$ for each $\underset{\sim}{q} \in D$. Then, for any sufficiently smooth function f,*

$$\int_D \left[f(\underset{\sim}{q}) - \int_D f(\underset{\sim}{p})\,\mathrm{e}^{-V(\underset{\sim}{p})}\,\mathrm{d}\underset{\sim}{p}\right]^2 \mathrm{e}^{-V(\underset{\sim}{q})}\,\mathrm{d}\underset{\sim}{q} \le \frac{1}{c_0}\int_D \left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-\omega} |\underset{\sim}{\nabla}_q f(\underset{\sim}{q})|^2 \mathrm{e}^{-V(\underset{\sim}{q})}\,\mathrm{d}\underset{\sim}{q}.$$

Proof. Under the hypotheses of the corollary

$$\underset{\sim}{\xi}^{\mathrm{T}} H(\underset{\sim}{q})\underset{\sim}{\xi} \ge c_0\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{\omega} |\underset{\sim}{\xi}|^2 \qquad \forall \underset{\sim}{\xi} \in \mathbb{R}^d, \quad |\underset{\sim}{q}| < \sqrt{b}.$$

Thus, in particular, V is a convex function on D, with a positive definite Hessian at each point in D. Since $H(\underset{\sim}{q})$, for $|\underset{\sim}{q}| < \sqrt{b}$, is a symmetric positive definite matrix, we deduce that

$$\langle H(\underset{\sim}{q})^{-1}\underset{\sim}{\eta}, \underset{\sim}{\eta}\rangle = \underset{\sim}{\eta}^{\mathrm{T}} H(\underset{\sim}{q})^{-1}\underset{\sim}{\eta} \le \frac{1}{c_0}\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-\omega} |\underset{\sim}{\eta}|^2 \qquad \forall \underset{\sim}{\eta} \in \mathbb{R}^d, \quad |\underset{\sim}{q}| < \sqrt{b}.$$

Hence, on taking $\underset{\sim}{\eta} = \underset{\sim}{\nabla}_q f$, the desired bound follows directly from the Brascamp-Lieb inequality. □

2.1. **Application to the FENE potential**

Let $D = B(\underset{\sim}{0}; \sqrt{b}\,)$ where $b > 2$, and define

$$U_\beta(s) := -\frac{\beta}{2}\ln\left(1 - \frac{2s}{b}\right), \qquad C(\beta) := \int_D \mathrm{e}^{-U_\beta(\frac{1}{2}|\underset{\sim}{q}|^2)}\,\mathrm{d}\underset{\sim}{q},$$

with $0 \le s < \frac{b}{2}$, $\beta := b - 2\gamma$, $0 \le \gamma \le 1$. The FENE potential corresponds to $\beta = b$ (*i.e.*, to $\gamma = 0$). Further, let

$$V(\underset{\sim}{q}) := U_\beta\left(\tfrac{1}{2}|\underset{\sim}{q}|^2\right) + \ln C(\beta), \qquad M_\beta(\underset{\sim}{q}) := \frac{1}{C(\beta)}\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{\frac{b}{2}} = \frac{C(b)}{C(\beta)}M_b(\underset{\sim}{q}).$$

Note that the exponent in M_β is $b/2$, not $\beta/2$, so M_β is not the normalized Maxwellian of U_β, except when $\beta = b$. Note further that, for all $\underset{\sim}{\xi} \in \mathbb{R}^d$ and all $\underset{\sim}{q} \in D$,

$$\sum_{i=1,\,j=1}^{d} \xi_i\xi_j \frac{\partial^2 V(\underset{\sim}{q})}{\partial q_i \partial q_j} = \left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-2}\left\{\frac{\beta}{b}|\underset{\sim}{\xi}|^2\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right) + \frac{2\beta}{b^2}\langle\underset{\sim}{\xi}, \underset{\sim}{q}\rangle^2\right\} \ge \frac{\beta}{b}\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}|\underset{\sim}{\xi}|^2.$$

Hence $\underset{\sim}{q} \mapsto V(\underset{\sim}{q})$ is (-1)-convex on D, with $c_0 = \beta/b$. The same is true of $\underset{\sim}{q} \mapsto U_\beta(\frac{1}{2}|\underset{\sim}{q}|^2) = V(\underset{\sim}{q}) - \ln C(\beta)$. Since $U_\beta = U_b + \gamma L$, it follows that $\underset{\sim}{q} \mapsto (U_b + \gamma L)(\frac{1}{2}|\underset{\sim}{q}|^2)$ is (-1)-convex on D, with $c_0 = (b - 2\gamma)/b$, for all $\gamma \in [0, 1]$; hence Hypothesis C holds with $\omega = -1$. Applying Corollary 2.2 with $\omega = -1$ and $c_0 = \beta/b$ yields

$$\int_D \left[f(\underset{\sim}{q}) - \int_D f(\underset{\sim}{p})M_\beta(\underset{\sim}{p})\left(1 - \frac{|\underset{\sim}{p}|^2}{b}\right)^{-\gamma}\mathrm{d}\underset{\sim}{p}\right]^2 M_\beta(\underset{\sim}{q})\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-\gamma}\mathrm{d}\underset{\sim}{q} \le \frac{b}{\beta}\int_D |\underset{\sim}{\nabla}_q f(\underset{\sim}{q})|^2 M_\beta(\underset{\sim}{q})\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{1-\gamma}\mathrm{d}\underset{\sim}{q},$$

where $\gamma \in [0, 1]$. We shall now consider the two extreme cases: $\gamma = 0$ and $\gamma = 1$.

2.1.1. *Case 1*

Let $\gamma = 0$ (whereupon $\beta = b$). Then, by writing $M := M_b$ and taking $f = \hat{\psi}/\sqrt{M}$, we get

$$\int_D \left[\hat{\psi} - \sqrt{M(\underset{\sim}{q})}\int_D \hat{\psi}(\underset{\sim}{p})\sqrt{M(\underset{\sim}{p})}\,\mathrm{d}\underset{\sim}{p}\right]^2 \mathrm{d}\underset{\sim}{q} + \frac{1}{b}\int_D |\underset{\sim}{q}|^2 |\underset{\sim}{\nabla}_M\hat{\psi}(\underset{\sim}{q})|^2\,\mathrm{d}\underset{\sim}{q} \le \int_D |\underset{\sim}{\nabla}_M\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}.$$

This implies the following Poincaré inequality, on noting that $\mathrm{Ker}(\underset{\sim}{\nabla}_M) = \{\lambda\sqrt{M} : \lambda \in \mathbb{R}\}$:

$$\inf_{c \in \mathrm{Ker}(\underset{\sim}{\nabla}_M)} \int_D |\hat{\psi} - c\,|^2\,\mathrm{d}\underset{\sim}{q} \le \int_D |\underset{\sim}{\nabla}_M\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}. \tag{2.1}$$

2.1.2. *Case 2*

Let $\gamma = 1$, take $f = \hat{\psi}/\sqrt{M}$ and note that M_β and $M := M_b$ only differ by the multiplicative factor $C(b)/C(\beta)$, where $\beta = b - 2$ with $b > 2$. Then,

$$\int_D \left[\hat{\psi}(\underset{\sim}{q}) - \frac{C(b)\sqrt{M(\underset{\sim}{q})}}{C(b-2)}\int_D \hat{\psi}(\underset{\sim}{p})\sqrt{M}(\underset{\sim}{p})\left(1 - \frac{|\underset{\sim}{p}|^2}{b}\right)^{-1}\mathrm{d}\underset{\sim}{p}\right]^2\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}\mathrm{d}\underset{\sim}{q} \le \frac{b}{b-2}\int_D |\underset{\sim}{\nabla}_M\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}. \tag{2.2}$$

Hence, we obtain the following Poincaré-Hardy inequality:

$$\inf_{c \in \mathrm{Ker}(\underset{\sim}{\nabla}_M)} \int_D \frac{|\hat{\psi} - c\,|^2}{1 - \frac{|\underset{\sim}{q}|^2}{b}}\,\mathrm{d}\underset{\sim}{q} \le \frac{b}{b-2}\int_D |\underset{\sim}{\nabla}_M\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}. \tag{2.3}$$

This can be seen as a refinement of the Poincaré inequality (2.1) in the sense that the left-hand side of (2.3) is an upper bound on the left-hand side of (2.1) (at the expense of increasing the multiplicative constant on the right-hand side of (2.1) from 1 to $b/(b-2)$, $b>2$). The inequalities (2.1) and (2.3) hold, in particular, for any $\hat{\psi} \in \sqrt{M}\,\mathrm{C}^\infty(\overline{D})$. Next, we shall show by a density argument that they are also valid for all $\hat{\psi} \in \mathrm{H}^1(D;M)$.

2.2. Density results for the space $\mathbf{H}^1(D;M)$

Since the density results below are not specific to the FENE model, we shall state them more generally, for any potential U and associated Maxwellian M that satisfy Hypotheses A and B, respectively. We shall then state additional results that hold when the list of assumptions is supplemented by Hypothesis C. Recall that the FENE model satisfies Hypotheses A, B and C (*cf.* Ex. 1.2).

(a) Suppose that the Maxwellian M satisfies Hypothesis B; M is then a weight-function of Type 3 in the sense of Triebel. According to [33], Theorem 3.2.2a, the weighted Sobolev space $\mathrm{H}^1_M(D) = \{v \in \mathrm{L}^2_M(D) : \underset{\sim}{\nabla}_q v \in (\mathrm{L}^2_M(D))^d\}$ is a Hilbert space with respect to the norm $\|\cdot\|_{\mathrm{H}^1_M(D)}$ defined by

$$\|v\|_{\mathrm{H}^1_M(D)} := \left(\|v\|^2_{\mathrm{L}^2_M(D)} + \|\underset{\sim}{\nabla}_q v\|^2_{\mathrm{L}^2_M(D)}\right)^{\frac{1}{2}},$$

and $\mathrm{L}^2_M(D) = (1/\sqrt{M})\,\mathrm{L}^2(D)$ is a Hilbert space with norm $\|\cdot\|_{\mathrm{L}^2_M(D)}$ defined by $\|v\|_{\mathrm{L}^2_M(D)} := \|\sqrt{M}v\|$, where $\|\cdot\|$ denotes the $\mathrm{L}^2(D)$ norm induced by the $\mathrm{L}^2(D)$ inner product $(\cdot,\cdot)$. By [33], Theorem 3.2.2c, $\mathrm{C}^\infty(\overline{D})$ is dense in both $\mathrm{H}^1_M(D)$ and $\mathrm{L}^2_M(D)$; see also Chapter I, Section 7, in Kufner [25], or one of [10,11]. Thus, $\sqrt{M}\,\mathrm{C}^\infty(\overline{D})$ is dense in the Hilbert spaces $\mathrm{H}^1(D;M)$ and $\mathrm{L}^2(D)$, whereby $\mathrm{H}^1(D;M)$ is dense in $\mathrm{L}^2(D)$.

(b) Now suppose that U satisfies Hypothesis A and the associated Maxwellian M satisfies Hypothesis B. It follows from Hardy's inequality (see, for example, [3,28]) that

$$\int_D \left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right)^{-2} |\hat{\psi}(\underset{\sim}{q})|^2 \,\mathrm{d}\underset{\sim}{q} \le 4b\|\underset{\sim}{\nabla}_q \hat{\psi}\|^2 \qquad \forall \hat{\psi} \in \mathrm{H}^1_0(D). \tag{2.4}$$

Since $\underset{\sim}{\nabla}_M \hat{\psi} = \underset{\sim}{\nabla}_q \hat{\psi} + \frac{1}{2}\underset{\sim}{q} U'\left(\frac{1}{2}|\underset{\sim}{q}|^2\right)\hat{\psi}$ and Hypothesis A implies the existence of $C_1 \in \mathbb{R}_{>0}$ (for the FENE model $C_1 = 1$) such that $(1-|\underset{\sim}{q}|^2/b)^2 |U'(\frac{1}{2}|\underset{\sim}{q}|^2)|^2 \le C_1^2$ for all $\underset{\sim}{q} \in D$, it follows that

$$\|\underset{\sim}{\nabla}_M \hat{\psi}\| \le (1 + C_1 b)\|\underset{\sim}{\nabla}_q \hat{\psi}\| \qquad \forall \hat{\psi} \in \mathrm{H}^1_0(D). \tag{2.5}$$

Thus, (2.5) now implies that $\mathrm{H}^1_0(D) \subset \mathrm{H}^1(D;M)$.

Let us show that $\mathrm{H}^1(D;M) = \mathrm{H}^1_0(D;M)$. As $\sqrt{M}\mathrm{C}^\infty(\overline{D}) \subset \mathrm{H}^1_0(D) \subset \mathrm{H}^1(D;M)$ and $\sqrt{M}\mathrm{C}^\infty(\overline{D})$ is dense in $\mathrm{H}^1(D;M)$ (*cf.* (a) above), we deduce that $\mathrm{H}^1_0(D)$ is dense in $\mathrm{H}^1(D;M)$. Since $\mathrm{C}^\infty_0(D)$ is dense in $\mathrm{H}^1_0(D)$, it then follows from (2.5) that $\mathrm{C}^\infty_0(D)$ is also dense in $\mathrm{H}^1(D;M)$. By definition, $\mathrm{H}^1_0(D;M)$ is the closure of $\mathrm{C}^\infty_0(D)$ in $\mathrm{H}^1(D;M)$; thus we deduce that $\mathrm{H}^1(D;M) = \mathrm{H}^1_0(D;M)$, and therefore $\mathfrak{K} = \mathfrak{K}_0$. As $\mathrm{H}^1(D;M)$ is continuously and densely embedded into $\mathrm{L}^2(D)$, it follows that $\mathrm{H}^1_0(D;M)$ is continuously and densely embedded into $\mathrm{L}^2(D)$.

(c) As the FENE potential U and Maxwellian M satisfy Hypotheses A and B, respectively, the density of $\sqrt{M}\,\mathrm{C}^\infty(\overline{D})$ in $\mathrm{H}^1(D;M)$ implies that (2.1) holds for all $\hat{\psi} \in \mathrm{H}^1(D;M)$. To show that, in the case of the FENE potential, (2.3) holds for all $\hat{\psi} \in \mathrm{H}^1(D;M)$, note that (2.2) holds, with the outer integral on the left-hand side of (2.2) replaced by an integral over $D_\varepsilon := \{\underset{\sim}{q} \in D : |\underset{\sim}{q}|^2 < b(1-\varepsilon)\}$, $\varepsilon \in (0,1)$, for any $\hat{\psi} \in \mathrm{C}^\infty_0(D)$, and hence, by the density of $\mathrm{C}^\infty_0(D)$ in $\mathrm{H}^1(D;M)$, for any $\hat{\psi} \in \mathrm{H}^1(D;M)$. Letting $\varepsilon \to 0_+$, Lebesgue's monotone convergence theorem implies that (2.2) holds for all $\hat{\psi} \in \mathrm{H}^1(D;M)$; hence (2.3) holds for all $\hat{\psi} \in \mathrm{H}^1(D;M)$. More generally, let U satisfy Hypotheses A and C with $\omega = -1$ (and $\gamma = 0$ or $\gamma = 1$), and let the associated Maxwellian satisfy Hypothesis B. If $\gamma = 0$, then (2.1) holds for all $\hat{\psi} \in \mathrm{H}^1(D;M)$, as in the case of the FENE

model, and if $\gamma = 1$, then (2.3) holds for all $\hat{\psi} \in \mathrm{H}^1(D, M)$, with $b/(b-2)$ replaced by $1/c_0$ (and $c_0 > 0$ as in Hypothesis C).

3. Backward Euler semidiscretization: existence and uniqueness of weak solutions

As was noted in the Introduction, by setting $\hat{\psi}(\cdot, t) := \psi(\cdot, t)/\sqrt{M}$ for $t \in [0, T]$ and $\hat{\varphi} := \varphi/\sqrt{M}$ in (1.7) and writing $\hat{\psi}_0 := \psi_0/\sqrt{M}$, we arrive at the following weak formulation of the initial-boundary-value problem (1.3), (1.4), (1.5):

Given $\hat{\psi}_0 \in \mathrm{L}^2(D)$, find $\hat{\psi} \in \mathrm{L}^\infty(0, T; \mathrm{L}^2(D)) \cap \mathrm{L}^2(0, T; \mathrm{H}^1_0(D; M))$ such that (1.8) holds in the sense of distributions on $(0, T)$, and $\hat{\psi}(\cdot, 0) = \hat{\psi}_0(\cdot)$.

The function ψ, representing a weak solution to the problem (1.7), is then recovered from $\hat{\psi}$ through the substitution $\psi := \sqrt{M}\,\hat{\psi}$. Thus, instead of constructing a Galerkin approximation to ψ, our aim is to construct a Galerkin approximation to $\hat{\psi}$ from a finite-dimensional subspace of the function space $\mathrm{H}^1_0(D; M)$; we shall then produce an approximation to ψ by multiplying the approximation to $\hat{\psi}$ by $\sqrt{M}$. First, however, we shall construct a time-semidiscretization of (1.8) and use a compactness argument to show the existence of weak solutions; we shall then also show the uniqueness of weak solutions.

Let $N_T \geq 1$ be an integer, $\Delta t = T/N_T$, and $t^n = n\Delta t$, for $n = 0, 1, \ldots, N_T$. Discretizing (1.8) in time using the backward Euler method yields the following semidiscrete numerical scheme.

Given $\hat{\psi}^0 := \hat{\psi}_0 = \psi_0/\sqrt{M} \in \mathrm{L}^2(D)$, find $\hat{\psi}^{n+1} \in \mathrm{H}^1_0(D; M)$, $n = 0, \ldots, N_T - 1$, such that

$$\int_D \frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\Delta t}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}^{n+1}) \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_M \hat{\psi}^{n+1} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q} = 0 \quad \forall \hat{\varphi} \in \mathrm{H}^1_0(D; M). \tag{3.1}$$

3.1. Well-posedness of the semidiscrete problem (3.1) and passage to the limit $\Delta t \to 0_+$

Let us first show that for any Δt, sufficiently small, problem (3.1) has a unique solution. To this end, we consider the bilinear form $\mathfrak{B}(\cdot, \cdot)$ defined on $\mathrm{H}^1_0(D; M) \times \mathrm{H}^1_0(D; M)$ by

$$\mathfrak{B}(\hat{\psi}, \hat{\varphi}) := \frac{1}{\Delta t}\int_D \hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}) \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_M \hat{\psi} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q},$$

and, for $\hat{\psi}^n \in \mathrm{L}^2(D)$ fixed, we define the linear functional $\ell(\hat{\psi}^n; \cdot)$ on $\mathrm{H}^1_0(D; M)$ by

$$\ell(\hat{\psi}^n; \hat{\varphi}) := \frac{1}{\Delta t}\int_D \hat{\psi}^n\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}.$$

Clearly,

$$\mathfrak{B}(\hat{\psi}, \hat{\psi}) \geq \frac{1}{\Delta t}\left(1 - \Delta t\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}\right)\int_D |\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q} + \frac{1}{4\lambda}\int_D |\underset{\sim}{\nabla}_M \hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q},$$

and hence, on assuming that $\Delta t\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)} < 1$ and letting $c_{\Delta t} := \frac{1}{\Delta t}\left(1 - \Delta t\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}\right)$, we deduce that

$$\mathfrak{B}(\hat{\psi}, \hat{\psi}) \geq \min\left(c_{\Delta t}, \frac{1}{4\lambda}\right)\|\hat{\psi}\|^2_{\mathrm{H}^1_0(D;M)}. \tag{3.2}$$

Also, by a simple application of the Cauchy-Schwarz inequality, $\mathfrak{B}(\cdot, \cdot)$ is a bounded bilinear functional on $\mathrm{H}^1_0(D; M) \times \mathrm{H}^1_0(D; M)$ and, for any $\hat{\psi}^n \in \mathrm{L}^2(D)$, $\ell(\hat{\psi}^n; \cdot)$ is a bounded linear functional on $\mathrm{H}^1_0(D; M)$.

Since $\mathrm{H}^1_0(D;M)$ is a Hilbert space with norm $\|\cdot\|_{\mathrm{H}^1_0(D;M)}$, the Lax-Milgram theorem implies the existence of a unique solution $\hat{\psi}^{n+1} \in \mathrm{H}^1_0(D;M)$ such that

$$\mathfrak{B}(\hat{\psi}^{n+1},\hat{\varphi}) = \ell(\hat{\psi}^n;\hat{\varphi}) \qquad \forall \hat{\varphi} \in \mathrm{H}^1_0(D;M), \qquad n=0,1,\ldots,N_T-1. \tag{3.3}$$

As $\hat{\psi}^0 \in \mathrm{L}^2(D)$, we have thus shown that, for any $\Delta t = T/N_T$ such that $\Delta t \lambda b \|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)} < 1$, the problem (3.1) has a unique solution $\{\hat{\psi}^n \in \mathrm{H}^1_0(D;M) \,:\, n=1,\ldots,N_T\}$.

For the purposes of the convergence analysis that will be carried out below, we consider an extended version of the scheme (3.1) with a nonzero right-hand side:

$$\begin{aligned}\int_D \frac{\hat{\psi}^{n+1}-\hat{\psi}^n}{\Delta t}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}^{n+1})\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \tfrac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} = \\ \int_D \mu^{n+1}\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \int_D \underset{\sim}{\nu}^{n+1}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} \qquad \forall\hat{\varphi}\in\mathrm{H}^1_0(D;M), \quad n=0,\ldots,N_T-1,\end{aligned} \tag{3.4}$$

where $\mu^{n+1} \in \mathrm{L}^2(D)$ and $\underset{\sim}{\nu}^{n+1} \in (\mathrm{L}^2(D))^d$ for all $n \geq 0$. We have the following stability result for (3.4).

Lemma 3.1 (the first stability inequality)**.** *Let $\Delta t = T/N_T$, $N_T \geq 1$, $\underset{\approx}{\kappa} \in (\mathrm{C}[0,T])^{d\times d}$, $\hat{\psi}^0 \in \mathrm{L}^2(D)$, and define $c_0 := 1 + 4\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}$. If Δt is such that $0 < c_0\Delta t \leq 1/2$, then we have, for all m such that $1 \leq m \leq N_T$,*

$$\|\hat{\psi}^m\|^2 + \sum_{n=0}^{m-1}\Delta t\left\|\frac{\hat{\psi}^{n+1}-\hat{\psi}^n}{\sqrt{\Delta t}}\right\|^2 + \sum_{n=0}^{m-1}\frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \leq \mathrm{e}^{2c_0 m\Delta t}\left\{\|\hat{\psi}^0\|^2 + \sum_{n=0}^{m-1}2\Delta t\left(\|\mu^{n+1}\|^2 + 4\lambda\|\underset{\sim}{\nu}^{n+1}\|^2\right)\right\}.$$

Proof. Let $0 \leq n \leq N_T - 1$. Setting $\hat{\varphi} = \hat{\psi}^{n+1}$, we write the first term in (3.4) as

$$\int_D \frac{\hat{\psi}^{n+1}-\hat{\psi}^n}{\Delta t}\,\hat{\psi}^{n+1}\,\mathrm{d}\underset{\sim}{q} = \frac{1}{2\Delta t}\left(\|\hat{\psi}^{n+1}\|^2 - \|\hat{\psi}^n\|^2\right) + \frac{1}{2\Delta t}\|\hat{\psi}^{n+1}-\hat{\psi}^n\|^2$$

using the identity $(\alpha-\beta)\alpha = \frac{1}{2}(\alpha^2-\beta^2) + \frac{1}{2}(\alpha-\beta)^2$.

Applying the Cauchy-Schwarz inequality to the transport term in (3.4), we have

$$\int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}^{n+1})\cdot\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\,\mathrm{d}\underset{\sim}{q} \leq \sqrt{b}\,|\underset{\approx}{\kappa}^{n+1}|\,\|\hat{\psi}^{n+1}\|\,\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|.$$

Combining these results and applying the Cauchy-Schwarz inequality to the right-hand side terms in (3.4) gives

$$\begin{aligned}\|\hat{\psi}^{n+1}\|^2 + \|\hat{\psi}^{n+1}-\hat{\psi}^n\|^2 + \frac{\Delta t}{\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \;&\leq\; \|\hat{\psi}^n\|^2 + 2\Delta t\sqrt{b}\,|\underset{\approx}{\kappa}^{n+1}|\|\hat{\psi}^{n+1}\|\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\| \\ &\quad + 2\Delta t\|\mu^{n+1}\|\|\hat{\psi}^{n+1}\| + 2\Delta t\|\underset{\sim}{\nu}^{n+1}\|\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\| \\ &=:\; \|\hat{\psi}^n\|^2 + \mathrm{T}_1 + \mathrm{T}_2 + \mathrm{T}_3.\end{aligned}$$

Using Cauchy's inequality $2\alpha\beta \leq \varepsilon\alpha^2 + \varepsilon^{-1}\beta^2$ with $\varepsilon > 0$ on each of T_1 and T_3, we deduce that

$$\mathrm{T}_1 \leq \varepsilon\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 + \frac{1}{\varepsilon}\Delta t^2 b|\underset{\approx}{\kappa}^{n+1}|^2\|\hat{\psi}^{n+1}\|^2, \qquad \mathrm{T}_3 \leq \varepsilon\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 + \frac{1}{\varepsilon}\Delta t^2\|\underset{\sim}{\nu}^{n+1}\|^2.$$

Choosing $\varepsilon = \Delta t/(4\lambda)$ then gives

$$\|\hat{\psi}^{n+1}\|^2 + \|\hat{\psi}^{n+1}-\hat{\psi}^n\|^2 + \frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \leq \|\hat{\psi}^n\|^2 + 4\Delta t\lambda b|\underset{\approx}{\kappa}^{n+1}|^2\|\hat{\psi}^{n+1}\|^2 + 4\Delta t\lambda\|\underset{\sim}{\nu}^{n+1}\|^2 + \mathrm{T}_2.$$

Similarly, we have $\mathrm{T}_2 \leq \Delta t\|\hat{\psi}^{n+1}\|^2 + \Delta t\|\mu^{n+1}\|^2$, and therefore, on defining $c_0 := 1 + 4\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}$, we get

$$(1 - c_0\Delta t)\|\hat{\psi}^{n+1}\|^2 + \|\hat{\psi}^{n+1} - \hat{\psi}^n\|^2 + \frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \leq \|\hat{\psi}^n\|^2 + \Delta t\|\mu^{n+1}\|^2 + 4\Delta t\lambda\|\underset{\sim}{\nu}^{n+1}\|^2.$$

As $c_0\Delta t \leq \frac{1}{2}$, dividing through by $(1 - c_0\Delta t)$ and using the fact that $1 \leq \frac{1}{1-c_0\Delta t} \leq 2$, we have

$$\begin{aligned}\|\hat{\psi}^{n+1}\|^2 + \|\hat{\psi}^{n+1} - \hat{\psi}^n\|^2 + \frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 &\leq \frac{1}{1 - c_0\Delta t}\left(\|\hat{\psi}^n\|^2 + \Delta t\|\mu^{n+1}\|^2 + 4\Delta t\lambda\|\underset{\sim}{\nu}^{n+1}\|^2\right)\\ &\leq (1 + 2c_0\Delta t)\|\hat{\psi}^n\|^2 + 2\Delta t\left(\|\mu^{n+1}\|^2 + 4\lambda\|\underset{\sim}{\nu}^{n+1}\|^2\right). \quad (3.5)\end{aligned}$$

Summing over $n = 0, \ldots, m-1$ in (3.5) we obtain

$$\begin{aligned}\|\hat{\psi}^m\|^2 + \sum_{n=0}^{m-1}\Delta t\left\|\frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\sqrt{\Delta t}}\right\|^2 + \sum_{n=0}^{m-1}\frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \leq\\ \left\{\|\hat{\psi}^0\|^2 + \sum_{n=0}^{m-1}2\Delta t\left(\|\mu^{n+1}\|^2 + 4\lambda\|\underset{\sim}{\nu}^{n+1}\|^2\right)\right\} + 2c_0\sum_{n=0}^{m-1}\Delta t\|\hat{\psi}^n\|^2, \quad (3.6)\end{aligned}$$

for all $m \in \{1, \ldots, N_T\}$. By induction (or by a discrete Gronwall lemma) we deduce that

$$\begin{aligned}\|\hat{\psi}^m\|^2 + \sum_{n=0}^{m-1}\Delta t\left\|\frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\sqrt{\Delta t}}\right\|^2 + \sum_{n=0}^{m-1}\frac{\Delta t}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^{n+1}\|^2 \leq\\ \mathrm{e}^{2c_0 m\Delta t}\left\{\|\hat{\psi}^0\|^2 + \sum_{n=0}^{m-1}2\Delta t\left(\|\mu^{n+1}\|^2 + 4\lambda\|\underset{\sim}{\nu}^{n+1}\|^2\right)\right\}, \qquad 1 \leq m \leq N_T,\end{aligned}$$

and that completes the proof. □

We shall now use this stability result to show the existence of weak solutions *via* a weak compactness argument. We shall also show the uniqueness of the weak solution.

Theorem 3.2. *Suppose that* $\hat{\psi}_0 \in \mathrm{L}^2(D)$ *and that* $\underset{\approx}{\kappa} \in (\mathrm{C}[0,T])^{d\times d}$. *Then, there exists a unique function* $\hat{\psi}$ *in* $\mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)) \cap \mathrm{C}([0,T];\mathrm{L}^2(D))$, *such that*

$$(\hat{\psi}(\cdot,0) - \hat{\psi}_0, \hat{w}) = 0 \qquad \forall \hat{w} \in \mathrm{L}^2(D)$$

and

$$\begin{aligned}-(\hat{\psi}_0, \hat{\varphi}(\cdot,0)) - \int_0^T\int_D \hat{\psi}\frac{\partial\hat{\varphi}}{\partial t}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t - \int_0^T\int_D (\underset{\approx}{\kappa}\,\underset{\sim}{q}\,\hat{\psi})\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t + \frac{1}{2\lambda}\int_0^T\int_D \underset{\sim}{\nabla}_M\hat{\psi}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t = 0\\ \forall\hat{\varphi} \in \mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M)), \quad \hat{\varphi}(\cdot,T) = 0. \quad (3.7)\end{aligned}$$

The function $\psi = \sqrt{M}\hat{\psi}$ *will be called the weak solution of the initial-boundary-value problem* (1.3), (1.4), (1.5).

Proof. **Step 1.** Let us denote by $\hat{\psi}^{\Delta t} \in \mathrm{C}([0,T];\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ the continuous piecewise linear interpolant, with respect to $t \in [0,T]$, of the semidiscrete solution $\{\hat{\psi}^n : n = 0, \ldots, N_T\}$ to (3.1), defined by

$$\hat{\psi}^{\Delta t}(\cdot,t)|_{[t^n,t^{n+1}]} := \frac{t - t^n}{\Delta t}\hat{\psi}^{n+1} + \frac{t^{n+1} - t}{\Delta t}\hat{\psi}^n, \qquad t \in [t^n, t^{n+1}], \quad n = 0, \ldots, N_T - 1, \quad (3.8)$$

and let

$$\hat{\psi}^{\Delta t,+}(\cdot,t) := \hat{\psi}^{n+1}(\cdot), \qquad \hat{\psi}^{\Delta t,-}(\cdot,t) := \hat{\psi}^{n}(\cdot), \qquad t \in [t^n, t^{n+1}], \quad n = 0, \ldots, N_T - 1. \tag{3.9}$$

We shall denote by $\hat{\psi}^{\Delta t,(\pm)}$ any one of the functions $\hat{\psi}^{\Delta t}$, $\hat{\psi}^{\Delta t,+}$, $\hat{\psi}^{\Delta t,-}$ defined above; $(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t}$ will denote the sequence of functions $\hat{\psi}^{\Delta t,(\pm)}$, indexed by $\Delta t = T/N_T \to 0_+$, for T fixed, as $N_T \to \infty$.

Using analogous notation for $\underset{\approx}{\kappa}$, equation (3.1), with $\hat{\varphi} \in \mathrm{H}^1_0(D;M)$ replaced by $\hat{\varphi}(t,\cdot) \in \mathrm{H}^1_0(D;M)$ for $t \in (0,T]$ where $\hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$, and summed over $n = 0, \ldots, N_T - 1$, yields

$$\int_0^T \int_D \frac{\partial \hat{\psi}^{\Delta t}}{\partial t}\, \hat{\varphi}\, \mathrm{d}\underset{\sim}{q}\, \mathrm{d}t - \int_0^T \int_D (\underset{\approx}{\kappa}^{\Delta t,+}\, \underset{\sim}{q}\, \hat{\psi}^{\Delta t,+}) \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q}\, \mathrm{d}t + \frac{1}{2\lambda} \int_0^T \int_D \underset{\sim}{\nabla}_M \hat{\psi}^{\Delta t,+} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q}\, \mathrm{d}t = 0 \qquad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)). \tag{3.10}$$

It follows from Lemma 3.1 with $\mu = 0$ and $\underset{\sim}{\nu} = \underset{\sim}{0}$ that

$$(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t} \quad \text{is bounded in} \quad \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)), \tag{3.11}$$

$$(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t} \quad \text{is bounded in} \quad \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)), \tag{3.12}$$

$$\left(\frac{\hat{\psi}^{\Delta t,+} - \hat{\psi}^{\Delta t,-}}{\sqrt{\Delta t}}\right)_{\Delta t} \quad \text{is bounded in} \quad \mathrm{L}^2(0,T;\mathrm{L}^2(D)). \tag{3.13}$$

Now, (3.11) and (3.12) imply that we can extract a subsequence from $(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t}$, which for the sake of notational simplicity we still denote by $(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t}$, such that, as $\Delta t \to 0_+$,

$$(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t} \quad \text{weak-}* \text{ converges in} \quad \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)), \tag{3.14}$$

$$(\hat{\psi}^{\Delta t,(\pm)})_{\Delta t} \quad \text{weakly converges in} \quad \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)). \tag{3.15}$$

Specifically, (3.14) implies the existence of $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D))$ such that

$$\int_0^T (\hat{\psi}^{\Delta t}(t) - \hat{\psi}(t), \hat{\varphi}(t))\, \mathrm{d}t \to 0 \quad \text{as} \quad \Delta t \to 0_+ \qquad \forall \hat{\varphi} \in \mathrm{L}^1(0,T;\mathrm{L}^2(D)). \tag{3.16}$$

On the other hand (3.15) implies the existence of $\hat{\psi}^*$ such that

$$\int_0^T \langle \hat{\psi}^{\Delta t}(t) - \hat{\psi}^*(t), \hat{\varphi}(t)\rangle\, \mathrm{d}t \to 0 \quad \text{as} \quad \Delta t \to 0_+ \qquad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)'), \tag{3.17}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between the Hilbert space $\mathrm{H}^1_0(D;M)$ and its dual space $\mathrm{H}^1_0(D;M)'$.

Identifying, by means of the Riesz representation theorem, $\mathrm{L}^2(D)$ with $\mathrm{L}^2(D)'$, we deduce that $\mathrm{H}^1_0(D;M) \subset \mathrm{L}^2(D) = \mathrm{L}^2(D)' \subset \mathrm{H}^1_0(D;M)'$, so that each space is dense in the next one in the chain, with continuous and injective embedding (*cf.* Sect. 2.2). Hence, $\langle \hat{\psi}, \hat{\varphi}\rangle = (\hat{\psi}, \hat{\varphi})$ for all $\hat{\psi} \in \mathrm{H}^1_0(D;M)$ and all $\hat{\varphi} \in \mathrm{L}^2(D)$. Returning to (3.17), we then deduce that

$$\int_0^T (\hat{\psi}^{\Delta t}(t) - \hat{\psi}^*(t), \hat{\varphi}(t))\, \mathrm{d}t = \int_0^T \langle \hat{\psi}^{\Delta t}(t) - \hat{\psi}^*(t), \hat{\varphi}(t)\rangle\, \mathrm{d}t \to 0 \qquad \text{as } \Delta t \to 0_+ \quad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{L}^2(D)).$$

Subtracting this from (3.16) yields

$$\int_0^T (\hat{\psi}(t) - \hat{\psi}^*(t), \hat{\varphi}(t))\, \mathrm{d}t = 0 \qquad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{L}^2(D)),$$

and therefore $\hat{\psi} = \hat{\psi}^* \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$.

It remains to show that the weak-$*$ limits $\hat{\psi}^{\pm}$ of the sequences $(\hat{\psi}^{\Delta t,\pm})_{\Delta t}$ in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(D))$ are also equal to $\hat{\psi}$. We shall show below that $\hat{\psi}^+ = \hat{\psi}^-$. Once we have done so, recalling from the definitions of $\hat{\psi}^{\Delta t}$ and $\hat{\psi}^{\Delta t,\pm}$ that

$$\hat{\psi}^{\Delta t}(\cdot,t) - \hat{\psi}^{\pm}(\cdot,t) = \frac{t-t^n}{\Delta t}(\hat{\psi}^{\Delta t,+}(\cdot,t) - \hat{\psi}^{+}(\cdot,t)) + \frac{t^{n+1}-t}{\Delta t}(\hat{\psi}^{\Delta t,-}(\cdot,t) - \hat{\psi}^{-}(\cdot,t))$$

for all $t \in [t^n, t^{n+1}]$ and $n = 0,\ldots,N_T-1$, and passing to the weak-$*$ limit in $\mathrm{L}^\infty(0,T;\mathrm{L}^2(D))$ as $\Delta t \to 0_+$, will imply that $\hat{\psi} = \hat{\psi}^{\pm}$.

To show that $\hat{\psi}^+ = \hat{\psi}^-$, observe that

$$\left|\int_0^T (\hat{\psi}^{\Delta t,+} - \hat{\psi}^{\Delta t,-}, \hat{\varphi})\,\mathrm{d}t\right| \le \left\|\frac{\hat{\psi}^{\Delta t,+} - \hat{\psi}^{\Delta t,-}}{\sqrt{\Delta t}}\right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))} \sqrt{\Delta t}\,\|\hat{\varphi}\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))},$$

for any $\hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{L}^2(D)) \subset \mathrm{L}^1(0,T;\mathrm{L}^2(D))$. Since by (3.13) the first factor on the right-hand side is bounded, independent of Δt, on passing to the limit $\Delta t \to 0_+$, it follows that

$$\lim_{\Delta t \to 0_+} \int_0^T (\hat{\psi}^{\Delta t,+} - \hat{\psi}^{\Delta t,-}, \hat{\varphi})\,\mathrm{d}t = 0 \qquad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{L}^2(D)).$$

Therefore,

$$\int_0^T (\hat{\psi}^{+} - \hat{\psi}^{-}, \hat{\varphi})\,\mathrm{d}t = 0 \qquad \forall \hat{\varphi} \in \mathrm{L}^2(0,T;\mathrm{L}^2(D)).$$

This, in turn, implies that $\hat{\psi}^+ = \hat{\psi}^-$. Thereby, as has been argued above, $\hat{\psi} = \hat{\psi}^+ = \hat{\psi}^-$.

Step 2. Next we pass to the limit $\Delta t \to 0_+$ in (3.10). Integrating by parts in the first term appearing on the left-hand side of equation (3.10), with $\hat{\varphi} \in \mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M)) \hookrightarrow \mathrm{C}([0,T];\mathrm{H}^1_0(D;M))$, we deduce that

$$(\hat{\psi}^{\Delta t}(\cdot,T),\hat{\varphi}(\cdot,T)) - (\hat{\psi}^{\Delta t}(\cdot,0),\hat{\varphi}(\cdot,0)) - \int_0^T\int_D \hat{\psi}^{\Delta t}\frac{\partial \hat{\varphi}}{\partial t}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t - \int_0^T\int_D (\underset{\approx}{\kappa}^{\Delta t,+}\,\underset{\sim}{q}\,\hat{\psi}^{\Delta t,+}) \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t$$
$$+ \frac{1}{2\lambda}\int_0^T\int_D \underset{\sim}{\nabla}_M \hat{\psi}^{\Delta t,+} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t = 0 \qquad \forall \hat{\varphi} \in \mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M)). \tag{3.18}$$

As $\hat{\psi}^{\Delta t}(\cdot,0) := \hat{\psi}_0(\cdot)$ and the sequence $(\underset{\approx}{\kappa}^{\Delta t,+})_{\Delta t}$ converges (strongly) in $(\mathrm{L}^\infty(0,T))^{d\times d}$ to $\underset{\approx}{\kappa}$, passing to the limit $\Delta t \to 0_+$ in (3.18) we deduce that the associated limiting function $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ satisfies (3.7). In particular, on choosing $\hat{\varphi} = \hat{\zeta}\cdot\hat{w} \in \mathrm{C}^\infty_0(0,T)\otimes \mathrm{H}^1_0(D;M)$ in (3.7), where $\hat{\zeta} \in \mathrm{C}^\infty_0(0,T)$ and $\hat{w} \in \mathrm{H}^1_0(D;M)$ are arbitrary, it follows from (3.7) that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\hat{\psi},\hat{w}) - (\underset{\approx}{\kappa}\,\underset{\sim}{q}\hat{\psi}, \underset{\sim}{\nabla}_M \hat{w}) + \frac{1}{2\lambda}(\underset{\sim}{\nabla}_M \hat{\psi}, \underset{\sim}{\nabla}_M \hat{w}) = 0 \qquad \forall \hat{w} \in \mathrm{H}^1_0(D;M), \tag{3.19}$$

in the sense of distributions on $(0,T)$, with $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$. Hence, the limiting function $\hat{\psi}$ satisfies (1.8), as required.

Step 3. It remains to show that $\hat{\psi}$ also satisfies the required initial condition. We proceed as follows. Since, for $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$, the second and third term on the left-hand side of (3.19) belong to $\mathrm{L}^2(0,T)$ for every $\hat{w} \in \mathrm{H}^1_0(D;M)$, the same is true of the first term on the left-hand side of (3.19). Therefore, $t \in [0,T] \mapsto (\hat{\psi}(\cdot,t),\hat{w})$ belongs to $\mathrm{H}^1(0,T)$ for all $\hat{w} \in \mathrm{H}^1_0(D;M)$. By the Sobolev embedding $\mathrm{H}^1(0,T) \hookrightarrow \mathrm{C}[0,T]$ we deduce that, for every $\hat{w} \in \mathrm{H}^1_0(D;M)$, $t \in [0,T] \mapsto (\hat{\psi}(\cdot,t),\hat{w})$ is a.e. equal to a function that is defined and

continuous on $[0,T]$; *i.e.*, $\hat{\psi} \in \mathrm{C}_{\mathrm{weak}}([0,T];\mathrm{L}^2(D))$, the set of all weakly continuous functions from $[0,T]$ into $\mathrm{L}^2(D)$.

Thus it makes sense to multiply (3.19) by $\hat{\zeta} \in \mathrm{H}^1(0,T)$, such that $\hat{\zeta}(T)=0$, integrate over $[0,T]$ and integrate by parts with respect to t in the first term to deduce, on writing $\hat{\varphi} = \hat{\zeta}\cdot\hat{w}$, that

$$-(\hat{\psi}(\cdot,0),\hat{\varphi}(\cdot,0)) - \int_0^T\int_D \hat{\psi}\frac{\partial\hat{\varphi}}{\partial t}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t - \int_0^T\int_D (\underset{\approx}{\kappa}\,\underset{\sim}{q}\,\hat{\psi})\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t + \frac{1}{2\lambda}\int_0^T\int_D \underset{\sim}{\nabla}_M\hat{\psi}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t = 0$$
$$\forall\hat{\varphi}\in\mathrm{H}^1(0,T)\otimes\mathrm{H}^1_0(D;M),\quad \hat{\varphi}(\cdot,T)=0. \tag{3.20}$$

Applying (3.7) with $\hat{\varphi}\in\mathrm{H}^1(0,T)\otimes\mathrm{H}^1_0(D;M)\subset\mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M))$ and comparing with (3.20) it follows that $(\hat{\psi}(\cdot,0)-\hat{\psi}_0,\hat{\varphi}(\cdot,0))=0$ for all $\hat{\varphi}\in\mathrm{H}^1(0,T)\otimes\mathrm{H}^1_0(D;M)$, $\hat{\varphi}(\cdot,T)=0$, and therefore, since $\mathrm{H}^1_0(D;M)$ is dense in $\mathrm{L}^2(D)$, it follows that $(\hat{\psi}(\cdot,0)-\hat{\psi}_0,\hat{w})=0$ for all $\hat{w}\in\mathrm{L}^2(D)$. We shall prove in Step 4 that, in fact, $\hat{\psi}\in\mathrm{C}([0,T];\mathrm{L}^2(D))$, which will then show that the function $\hat{\psi}$ satisfies the initial condition $\hat{\psi}(\cdot,0)=\hat{\psi}_0$ (and therefore $\psi=\sqrt{M}\hat{\psi}$ satisfies the corresponding initial condition $\psi(\cdot,0)=\psi_0$ $(=\sqrt{M}\hat{\psi}_0)$).

Step 4. Let us show that $\psi=\sqrt{M}\hat{\psi}$, with $\hat{\psi}\in\mathrm{L}^\infty(0,T;\mathrm{L}^2(D))\cap\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ defined by (3.7), is the *unique* weak solution to the initial-boundary-value problem. We begin by observing that, for any $\hat{\psi}\in\mathrm{L}^\infty(0,T;\mathrm{L}^2(D))\cap\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$,

$$\left|\int_0^T\left\{(\underset{\approx}{\kappa}\,\underset{\sim}{q}\hat{\psi},\underset{\sim}{\nabla}_M\hat{\varphi}) - \frac{1}{2\lambda}(\underset{\sim}{\nabla}_M\hat{\psi},\underset{\sim}{\nabla}_M\hat{\varphi})\right\}\mathrm{d}t\right| \le C\|\hat{\psi}\|_{\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))}\|\underset{\sim}{\nabla}_M\hat{\varphi}\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}$$

for all $\hat{\varphi}\in\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$, where $C:=\left(b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}+1/(4\lambda^2)\right)^{\frac{1}{2}}$. On bounding $\|\underset{\sim}{\nabla}_M\hat{\varphi}\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}$ by $\|\hat{\varphi}\|_{\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))}$, we deduce the existence of $G\in\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)')$ such that, by (3.7),

$$-(\hat{\psi}_0,\hat{\varphi}(\cdot,0)) - \int_0^T\int_D\hat{\psi}\frac{\partial\hat{\varphi}}{\partial t}\,\mathrm{d}\underset{\sim}{q}\,\mathrm{d}t = \int_0^T\langle G,\hat{\varphi}\rangle\,\mathrm{d}t \qquad \forall\hat{\varphi}\in\mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M)),\quad \hat{\varphi}(\cdot,T)=0.$$

Hence,

$$-\int_0^T\left\langle\hat{\psi},\frac{\partial\hat{\varphi}}{\partial t}\right\rangle\mathrm{d}t = \int_0^T\langle G,\hat{\varphi}\rangle\,\mathrm{d}t \qquad \forall\hat{\varphi}\in\mathrm{C}^\infty_0(0,T;\mathrm{H}^1_0(D;M)).$$

By virtue of Lemma 1.1 in Chapter 3, Section 1.1 of Temam [32] with $X=\mathrm{H}^1_0(D;M)'$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\psi},\hat{w}\rangle = \langle G,\hat{w}\rangle \qquad \forall\hat{w}\in\mathrm{H}^1_0(D;M),$$

in the sense of distributions on $(0,T)$, and $\hat{\psi}$ is almost everywhere equal to a continuous function from $[0,T]$ into $\mathrm{H}^1_0(D;M)'$. In fact, since $\hat{\psi}\in\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ and

$$\frac{\partial\hat{\psi}}{\partial t} = G\in\mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)'),$$

it follows from Lemma 1.2 in Chapter 3, Section 1.2 of Temam [32] (with $V=\mathrm{H}^1_0(D;M)$, $H=\mathrm{L}^2(D)$ and $V'=\mathrm{H}^1_0(D;M)'$) that $\hat{\psi}$ is a.e. equal to a continuous function from $[0,T]$ into $\mathrm{L}^2(D)$ and the following identity holds in the sense of distributions of $(0,T)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\hat{\psi}\|^2 = 2\left\langle\frac{\partial\hat{\psi}}{\partial t},\hat{\psi}\right\rangle.$$

Now, suppose that $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M))$ is a weak solution of the initial-boundary-value problem, defined by (3.7). Then, for any $s \in (0,T]$,

$$\int_0^s \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\hat{\psi}\|^2\,\mathrm{d}t = \int_0^s \left\langle \frac{\partial\hat{\psi}}{\partial t}, \hat{\psi}\right\rangle \mathrm{d}t = \int_0^s \langle G, \hat{\psi}\rangle\,\mathrm{d}t = \int_0^s \left\{ ((\underset{\approx}{\kappa}\,\underset{\sim}{q})\hat{\psi}, \underset{\sim}{\nabla}_M\hat{\psi}) - \frac{1}{2\lambda}(\underset{\sim}{\nabla}_M\hat{\psi}, \underset{\sim}{\nabla}_M\hat{\psi})\right\}\mathrm{d}t.$$

Therefore,

$$\begin{aligned}\frac{1}{2}\left(\|\hat{\psi}(s)\|^2 - \|\hat{\psi}_0\|^2\right) + \frac{1}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}\|^2_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))} &= \int_0^s ((\underset{\approx}{\kappa}\,\underset{\sim}{q})\hat{\psi}, \underset{\sim}{\nabla}_M\hat{\psi})\mathrm{d}t \\ &\le \sqrt{b}\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)}\|\hat{\psi}\|_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))}\|\underset{\sim}{\nabla}_M\hat{\psi}\|_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))} \text{ for a.e. } s \in (0,T].\end{aligned}$$

This implies that

$$\|\hat{\psi}(s)\|^2 + \frac{1}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}\|^2_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))} \le \|\hat{\psi}_0\|^2 + 2\lambda\, b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}\|\hat{\psi}\|^2_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))} \qquad \text{for a.e. } s \in (0,T].$$

Thus, by Gronwall's lemma, any weak solution $\hat{\psi}$ to (3.7) satisfies the following energy inequality

$$\|\hat{\psi}(s)\|^2_{\mathrm{L}^\infty(0,s;\mathrm{L}^2(D))} + \frac{1}{2\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}\|^2_{\mathrm{L}^2(0,s;\mathrm{L}^2(D))} \le \|\hat{\psi}_0\|^2 \exp\left(2s\lambda\, b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}\right) \qquad \text{for a.e. } s \in (0,T].$$

Note, in particular, that if $\hat{\psi}_0 = 0$, then $\hat{\psi}(\cdot,s) = 0$ in $\mathrm{L}^2(D)$ for a.e. $s \in (0,T]$, which in turn implies the uniqueness of a weak solution. □

Next we shall show that $\psi = \sqrt{M}\hat{\psi}$ has the usual properties of a probability density function: if ψ_0 is non-negative and has unit integral over D, then the same is true of $\psi(\cdot,t)$ for all $t \in [0,T]$.

Lemma 3.3. *Let $\psi_0 \in \mathfrak{H}$ and $\psi = \sqrt{M}\hat{\psi}$ where $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D;M)) \cap \mathrm{C}([0,T];\mathrm{L}^2(D))$ is the weak solution to (3.7) subject to the initial condition $\hat{\psi}_0 = \psi_0/\sqrt{M}$ (*i.e.*, the function ψ is the weak solution of the initial-boundary-value problem (1.3), (1.4), (1.5)). Then,*

$$\int_D \psi(\underset{\sim}{q},t)\,\mathrm{d}\underset{\sim}{q} = \int_D \psi_0(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q} \qquad \forall t \in [0,T).$$

Furthermore if $\psi_0 \ge 0$ a.e. on D, then $\psi(\cdot,t) \ge 0$ a.e. on D for all $t \in [0,T]$.

Proof. Fix any $t \in (0,T)$, and let $\varepsilon \in (0, T-t]$. Consider the function $\hat{\varphi}_\varepsilon$ defined by

$$\hat{\varphi}_\varepsilon(\underset{\sim}{q},s) := \begin{cases} \sqrt{M} & \text{for } s \in [0,t], \\ \sqrt{M}(t+\varepsilon-s)/\varepsilon & \text{for } s \in [t,t+\varepsilon), \\ 0 & \text{for } s \in [t+\varepsilon,T]. \end{cases}$$

Clearly, $\hat{\varphi}_\varepsilon \in \mathrm{H}^1(0,T;\mathrm{H}^1_0(D;M))$ and $\hat{\varphi}_\varepsilon(\cdot,T) = 0$. Taking $\hat{\varphi}_\varepsilon$ as test function in (3.7) we obtain

$$-(\hat{\psi}_0, \sqrt{M}\,) + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} (\hat{\psi}(\cdot,s), \sqrt{M}\,)\,\mathrm{d}s = 0.$$

Passing to the limit $\varepsilon \to 0_+$ yields $-(\hat{\psi}_0, \sqrt{M}\,) + (\hat{\psi}(\cdot,t), \sqrt{M}\,) = 0$, whereby $(\psi(\cdot,t),1) = (\psi_0,1)$, as required, for all $t \in (0,T)$; for $t = 0$ the equality holds trivially.

Now, suppose that $\psi_0 \in \mathfrak{H}$ and $\psi_0 \geq 0$; then, $\hat{\psi}_0 \in \mathrm{L}^2(D)$ and $\hat{\psi}_0 \geq 0$. For Δt as in Lemma 3.1, consider the sequence of functions $(\hat{\psi}^n)_{n=0}^{N_T} \subset \mathrm{H}_0^1(D;M)$ defined by (3.3). By Lemma 3.5 below (with $L = 0$ and $[x]_\pm := (x \pm |x|)/2$ for $x \in \mathbb{R}$), we have that $([\hat{\psi}^n]_-)_{n=0}^{N_T} \subset \mathrm{H}_0^1(D;M)$. It follows from (3.3) that

$$\mathfrak{B}([\hat{\psi}^{n+1}]_-\,,\,[\hat{\psi}^{n+1}]_-) = \mathfrak{B}(\hat{\psi}^{n+1}\,,\,[\hat{\psi}^{n+1}]_-) = \ell(\hat{\psi}^n;[\hat{\psi}^{n+1}]_-).$$

Suppose, for induction, that $\hat{\psi}^n \geq 0$; this is certainly true for $n = 0$, since $\hat{\psi}^0 = \hat{\psi}_0 \geq 0$. Hence,

$$\ell(\hat{\psi}^n;[\hat{\psi}^{n+1}]_-) = \frac{1}{\Delta t}\int_D \hat{\psi}^n(\underset{\sim}{q})[\hat{\psi}^{n+1}(\underset{\sim}{q})]_-\,\mathrm{d}\underset{\sim}{q} \leq 0.$$

Therefore, $\mathfrak{B}([\hat{\psi}^{n+1}]_-\,,\,[\hat{\psi}^{n+1}]_-) \leq 0$; thus, (3.2) implies that $\|[\hat{\psi}^{n+1}]_-\|_{\mathrm{H}_0^1(D;M)} \leq 0$, whereby $[\hat{\psi}^{n+1}]_- = 0$ and hence $\hat{\psi}^{n+1} \geq 0$. By induction, $\hat{\psi}^n \geq 0$ for all $n = 0, 1, \ldots, N_T$. Therefore, each of the functions $\hat{\psi}^{\Delta t}$, $\hat{\psi}^+$ and $\hat{\psi}^-$, defined in the proof of Theorem 3.2, is non-negative on $D \times [0,T]$. Hence the limiting function $\hat{\psi}$ of the sequence(s), as $\Delta t \to 0_+$, is also non-negative on $D \times [0,T]$. □

Remark 3.4. We note in passing that if $\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}(t)\,\underset{\sim}{q} \leq 0$ for all $t \in [0,T]$ then, by considering the expression, $\mathfrak{B}([\hat{\psi}^{n+1} - L\sqrt{M}\,]_+\,,\,[\hat{\psi}^{n+1} - L\sqrt{M}\,]_+)$ one can show by induction, as in the proof above, with

$$L = \mathrm{ess.sup}_{\underset{\sim}{q}\in D}\ \hat{\psi}_0(\underset{\sim}{q})/\sqrt{M(\underset{\sim}{q})},$$

that $\mathfrak{B}([\hat{\psi}^{n+1} - L\sqrt{M}\,]_+\,,\,[\hat{\psi}^{n+1} - L\sqrt{M}\,]_+) = 0$ for all $n = 0, 1, \ldots, N_T - 1$. Consequently, by inequality (3.2), $[\hat{\psi}^{n+1} - L\sqrt{M}\,]_+ = 0$; *i.e.*, $\hat{\psi}^{n+1} \leq L\sqrt{M}$. This then implies, on passage to the limit $\Delta t \to 0_+$, that

$$\mathrm{ess.sup}_{(\underset{\sim}{q},t)\in D\times[0,T]}\ \hat{\psi}(\underset{\sim}{q},t)/\sqrt{M(\underset{\sim}{q})} \leq \mathrm{ess.sup}_{\underset{\sim}{q}\in D}\ \hat{\psi}_0(\underset{\sim}{q})/\sqrt{M(\underset{\sim}{q})}.$$

Hence,

$$\mathrm{ess.sup}_{(\underset{\sim}{q},t)\in D\times[0,T]}\ \psi(\underset{\sim}{q},t)/M(\underset{\sim}{q}) \leq \mathrm{ess.sup}_{\underset{\sim}{q}\in D}\ \psi_0(\underset{\sim}{q})/M(\underset{\sim}{q}),$$

which can be thought of as a maximum principle for the initial-boundary-value problem[1].

Lemma 3.5. *Suppose that* $\hat{\varphi} \in \mathrm{H}_0^1(D;M)$ *and* $L \geq 0$. *Then,*

$$\nabla_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ = \begin{cases} \nabla_M(\,\hat{\varphi} - L\sqrt{M}\,) = \nabla_M\hat{\varphi} & \text{if } \hat{\varphi} > L\sqrt{M}, \\ 0 & \text{if } \hat{\varphi} \leq L\sqrt{M}; \end{cases} \tag{3.21}$$

and

$$\nabla_M[\,\hat{\varphi} - L\sqrt{M}\,]_- = \begin{cases} \nabla_M(\,\hat{\varphi} - L\sqrt{M}\,) = \nabla_M\hat{\varphi} & \text{if } \hat{\varphi} < L\sqrt{M}, \\ 0 & \text{if } \hat{\varphi} \geq L\sqrt{M}. \end{cases} \tag{3.22}$$

Furthermore, $[\,\hat{\varphi} - L\sqrt{M}\,]_+$ *and* $[\,\hat{\varphi} - L\sqrt{M}\,]_-$ *belong to* $\mathrm{H}_0^1(D;M)$.

Proof. We shall prove (3.21); the proof of (3.22) is analogous, *mutatis mutandis*. We begin by noting that since $L \geq 0$ and $\sqrt{M} > 0$ on D,

$$|[\,\hat{\varphi} - L\sqrt{M}\,]_+| \leq |\hat{\varphi}|. \tag{3.23}$$

Following [7], for any $\varepsilon > 0$, we define the following regularization of $[\,\cdot\,]_+$:

$$p_{+,\varepsilon}(s) := \begin{cases} (s^2 + \varepsilon^2)^{\frac{1}{2}} - \varepsilon & \text{if } s > 0, \\ 0 & \text{if } s \leq 0. \end{cases}$$

[1] If $\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}(t)\,\underset{\sim}{q} \leq 0$ for all $\underset{\sim}{q} \in \mathbb{R}^d$ and $t \in [0,T]$, and $\mathrm{tr}(\underset{\approx}{\kappa}(t)) = 0$ for all $t \in [0,T]$, then $\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}(t)\underset{\sim}{q} = 0$ for all $\underset{\sim}{q} \in \mathbb{R}^d$ and $t \in [0,T]$.

Clearly, $0 \le p_{+,\varepsilon}(s) \le [s]_+$ for all $s \in \mathbb{R}$. Observe that

$$\underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ = \underset{\sim}{\nabla}_q[\,\hat{\varphi} - L\sqrt{M}\,]_+ + \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)[\,\hat{\varphi} - L\sqrt{M}\,]_+$$

in the sense of d-component distributions on D. Let $\underset{\sim}{\eta} \in (\mathrm{C}_0^\infty(D))^d$ be fixed. Thus,

$$\begin{aligned}
\langle \underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ \,,\, \underset{\sim}{\eta}\rangle &= \langle \underset{\sim}{\nabla}_q[\,\hat{\varphi} - L\sqrt{M}\,]_+ + \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)[\,\hat{\varphi} - L\sqrt{M}\,]_+ \,,\, \underset{\sim}{\eta}\rangle \\
&= -\langle [\,\hat{\varphi} - L\sqrt{M}\,]_+ \,,\, \underset{\sim}{\nabla}_q \cdot \underset{\sim}{\eta}\rangle + \langle \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)[\,\hat{\varphi} - L\sqrt{M}\,]_+ \,,\, \underset{\sim}{\eta}\rangle \\
&= -\int_D [\,\hat{\varphi} - L\sqrt{M}\,]_+ (\underset{\sim}{\nabla}_q \cdot \underset{\sim}{\eta})\,\mathrm{d}\underset{\sim}{q} + \int_D \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)[\,\hat{\varphi} - L\sqrt{M}\,]_+ \cdot \underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q}.
\end{aligned}$$

Let χ_S denote the characteristic function of a set $S \subset D$. Since $\underset{\sim}{\eta}$ has compact support in D, by Lebesgue's dominated convergence theorem we deduce that

$$\begin{aligned}
\langle \underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ \,,\, \underset{\sim}{\eta}\rangle &= -\lim_{\varepsilon\to 0_+}\int_D p_{+,\varepsilon}(\hat{\varphi} - L\sqrt{M}\,)(\underset{\sim}{\nabla}_q \cdot \underset{\sim}{\eta})\,\mathrm{d}\underset{\sim}{q} + \lim_{\varepsilon\to 0_+}\int_D \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)p_{+,\varepsilon}(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} \\
&= \lim_{\varepsilon\to 0_+}\int_D p'_{+,\varepsilon}(\hat{\varphi} - L\sqrt{M}\,)\underset{\sim}{\nabla}_q(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} + \lim_{\varepsilon\to 0_+}\int_D \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)p_{+,\varepsilon}(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} \\
&= \int_D \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_q(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} + \int_D \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} \\
&= \int_D \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\left\{\underset{\sim}{\nabla}_q(\hat{\varphi} - L\sqrt{M}\,) + \tfrac{1}{2}\underset{\sim}{q}U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)(\hat{\varphi} - L\sqrt{M}\,)\right\}\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} \\
&= \int_D \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M(\hat{\varphi} - L\sqrt{M}\,)\cdot\underset{\sim}{\eta}\,\mathrm{d}\underset{\sim}{q} = \langle \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M(\hat{\varphi} - L\sqrt{M}\,)\,,\,\underset{\sim}{\eta}\rangle\cdot
\end{aligned}$$

Since these equalities hold for all $\underset{\sim}{\eta} \in (\mathrm{C}_0^\infty(D))^d$, it follows that $\underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ = \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M(\hat{\varphi} - L\sqrt{M}\,)$. As $\sqrt{M} \in \mathrm{Ker}(\underset{\sim}{\nabla}_M)$, we deduce that $\chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M(\hat{\varphi} - L\sqrt{M}\,) = \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M\hat{\varphi}$, and that proves (3.21). Now, since $\underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ = \chi_{\hat{\varphi}>L\sqrt{M}}(\underset{\sim}{q})\,\underset{\sim}{\nabla}_M\hat{\varphi}$, and the right-hand side in this equality belongs to $\mathrm{L}^2(D)$ (recall that $\hat{\varphi} \in \mathrm{H}_0^1(D;M)$ by hypothesis), it follows that $\underset{\sim}{\nabla}_M[\,\hat{\varphi} - L\sqrt{M}\,]_+ \in \mathrm{L}^2(D)$. Hence, and by (3.23), $[\,\hat{\varphi} - L\sqrt{M}\,]_+ \in \mathrm{H}_0^1(D;M)$, as required. □

By the next lemma, if $\underset{\approx}{\kappa} \in (\mathrm{H}^1(0,T))^{d\times d}$ and $\hat{\psi}_0 \in \mathrm{H}_0^1(D;M)$, then we have stability in stronger norms.

Lemma 3.6 (the second stability inequality)**.** *Let $\Delta t = T/N_T$, $N_T \ge 1$, $\underset{\approx}{\kappa} \in (\mathrm{H}^1(0,T))^{d\times d}$, $\hat{\psi}^0 \in \mathrm{H}_0^1(D;M)$, and define $c_0 := 1 + 4\lambda b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}$. If Δt is such that $0 < c_0\Delta t \le 1/2$, then, for all m such that $1 \le m \le N_T$,*

$$\begin{aligned}
\Delta t\sum_{n=0}^{m-1}\left\|\frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\Delta t}\right\|^2 &+ \frac{1}{4\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^m\|^2 + \frac{1}{2\lambda}\sum_{n=0}^{m-1}\Delta t\left\|\underset{\sim}{\nabla}_M\frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\sqrt{\Delta t}}\right\|^2 \le \\
&\mathrm{e}^{2c_1 m\Delta t}\left\{2\Delta t\sum_{n=0}^{m-1}\|\mu^{n+1}\|^2 + 12\lambda\max_{1\le n\le m}\|\underset{\sim}{\nu}^n\|^2 + \Delta t\sum_{n=1}^{m-1}\left\|\frac{\underset{\sim}{\nu}^{n+1} - \underset{\sim}{\nu}^n}{\Delta t}\right\|^2\right. \\
&\left. + \frac{1}{\lambda}\|\underset{\sim}{\nabla}_M\hat{\psi}^0\|^2 + \left(b\,\|\underset{\approx}{\kappa}t\|^2_{\mathrm{L}^2(0,T)} + 12\lambda b\,\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)}\right)\mathfrak{S}(\hat{\psi}^0,\mu,\underset{\sim}{\nu},m\Delta t)\right\},
\end{aligned}$$

where $\mathfrak{S}(\hat{\psi}^0,\mu,\underset{\sim}{\nu},m,\Delta t)$ is the right-hand side of the inequality from Lemma 3.1 and $c_1 = 4\lambda(1 + b\,\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,T)})$.

Proof. The proof is similar to that of Lemma 3.1, except one uses the test function $\hat{\varphi} = (\hat{\psi}^{n+1} - \hat{\psi}^n)/\Delta t$. □

It follows from Lemma 3.6, by an identical argument as in the proof of Theorem 3.2, that the weak solution $\hat{\psi}$ of (3.7) belongs to $\mathrm{H}^1(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^\infty(0,T;\mathrm{H}^1_0(D;M))$, provided that $\underset{\approx}{\kappa} \in (\mathrm{H}^1(0,T))^{d\times d}$ and $\hat{\psi}_0 \in \mathrm{H}^1_0(D;M)$.

The stability result in Lemma 3.1 will be useful in Section 4, but for now we note that setting $\mu = 0$ and $\underset{\sim}{\nu} = \underset{\sim}{0}$ in Lemmas 3.1 and 3.6 demonstrates the unconditional stability of the time semidiscretization in various norms. We also note that, evidently, any fully-discrete method based on the semidiscrete scheme (3.1) and conforming Galerkin discretization in $\underset{\sim}{q}$ using a finite-dimensional subspace $\mathcal{P}_N$ of $\mathrm{H}^1_0(D;M)$ will be unconditionally stable in the norms appearing on the left-hand sides of the bounds in Lemmas 3.1 and 3.6.

3.2. **Well-posedness of a Chauvière-Lozinski type transformed FENE model**

In this section we show that, in the case of the FENE model, the weak formulation resulting from the substitution $\hat{\psi} := \psi/M^{2s/b}$ with $b \geq 4s^2/(2s-1)$ and $s > 1/2$ also leads to a well-posed problem and a stable semidiscretization in any number of space dimensions. The minimum value of the function $s \in (0,\infty) \mapsto 4s^2/(2s-1)$ is attained at $s = 1$, yielding the maximum range of b values, $b \geq 4$. This transformation was proposed by Chauvière and Lozinski [17,18,27] in the special cases $s = 2$ and $s = 2.5$, where these values were chosen on the basis of numerical experiments in two and three space dimensions, respectively. For the sake of brevity, we shall confine ourselves to establishing an energy estimate analogous to our first stability inequality in Lemma 3.1. A weak compactness argument identical to the one above then shows the existence of a unique (corresponding) weak solution. The discussion in this section is restricted to the FENE model; however our arguments can be extended to more general models by adopting additional structural hypotheses on the potential U (see, for example, Sect. 2.3 in [7]).

Inserting $\psi(\underset{\sim}{q}) = [M(\underset{\sim}{q})]^{2s/b}\hat{\psi}(\underset{\sim}{q})$ into our model problem (1.3), where now M is the FENE Maxwellian, yields, on noting that $\mathrm{tr}(\underset{\approx}{\kappa})(t) = 0$ for all $t \in [0,T]$,

$$\frac{\partial\hat{\psi}}{\partial t} - \frac{1}{2\lambda}\Delta_q\hat{\psi} = \frac{1}{2\lambda}\left[\left(1-\frac{4s}{b}\right)\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}\underset{\sim}{q} - 2\lambda(\underset{\approx}{\kappa}\,\underset{\sim}{q})\right]\cdot\underset{\sim}{\nabla}_q\hat{\psi}$$
$$+\frac{1}{2\lambda}\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2}\left[d\left(1-\frac{2s}{b}\right)\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)+\frac{2(s-1)(2s-b)}{b^2}|\underset{\sim}{q}|^2+\frac{4s\lambda}{b}(\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}\,\underset{\sim}{q})\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)\right]\hat{\psi}. \quad (3.24)$$

Denoting by $\underset{\sim}{A}(\underset{\sim}{q},t)$ the expression in the first square bracket on the right-hand side of (3.24) and by $B(\underset{\sim}{q},t)$ the expression in the second square bracket, multiplying (3.24) by any $\hat{\varphi} \in \mathrm{H}^1_0(D)$, integrating the resulting expression over D, and integrating by parts in the second term on the left-hand side, yields the following weak formulation.

Given $\hat{\psi}_0 = \psi_0/M^{2s/b} \in \mathrm{L}^2(D)$, find $\hat{\psi} \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(D)) \cap \mathrm{L}^2(0,T;\mathrm{H}^1_0(D))$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_q\hat{\psi}\cdot\underset{\sim}{\nabla}_q\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} = \frac{1}{2\lambda}\int_D(\underset{\sim}{A}(\underset{\sim}{q},t)\cdot\underset{\sim}{\nabla}_q\hat{\psi})\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2}B(\underset{\sim}{q},t)\,\hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}, \quad (3.25)$$

for all $\hat{\varphi} \in \mathrm{H}^1_0(D)$, in the sense of distributions on $(0,T)$, and with $\hat{\psi}(\cdot,0) = \hat{\psi}_0$.

The backward Euler semidiscretization of this weak formulation is as follows.
Given $\hat{\psi}^0 := \hat{\psi}_0 = \psi_0/M^{2s/b} \in \mathrm{L}^2(D)$, find $\hat{\psi}^{n+1} \in \mathrm{H}^1_0(D)$, $n = 0, 1, \ldots, N_T - 1$, such that

$$\int_D \frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\Delta t}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_q\hat{\psi}^{n+1}\cdot\underset{\sim}{\nabla}_q\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} =$$
$$\frac{1}{2\lambda}\int_D (\underset{\sim}{A}(\underset{\sim}{q},t^{n+1})\cdot\underset{\sim}{\nabla}_q\hat{\psi}^{n+1})\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2} B(\underset{\sim}{q},t^{n+1})\,\hat{\psi}^{n+1}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} \quad \forall\hat{\varphi}\in\mathrm{H}^1_0(D). \tag{3.26}$$

We begin by showing that, for Δt sufficiently small and all $b \geq 4s^2/(2s-1)$ and $s > 1/2$, this problem has a unique solution. To this end, for $t \in [0,T]$ fixed, we consider the bilinear form defined on $\mathrm{H}^1_0(D)\times\mathrm{H}^1_0(D)$ by

$$\mathfrak{C}(\hat{\psi},\hat{\varphi}) := \frac{1}{\Delta t}\int_D \hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda}\int_D \underset{\sim}{\nabla}_q\hat{\psi}\cdot\underset{\sim}{\nabla}_q\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \frac{1}{2\lambda}\int_D (\underset{\sim}{A}(\underset{\sim}{q},t)\cdot\underset{\sim}{\nabla}_q\hat{\psi})\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} - \frac{1}{2\lambda}\int_D \left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2} B(\underset{\sim}{q},t)\hat{\psi}\,\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}.$$

Now, taking $\hat{\varphi} = \hat{\psi} \in \mathrm{C}^\infty_0(D)$, integration by parts in the third integral in the definition of $\mathfrak{C}$, and then merging the resulting integral with the fourth integral in the definition of $\mathfrak{C}$, yields

$$\begin{aligned}\mathfrak{C}(\hat{\psi},\hat{\psi}) \;=\; & \frac{1}{\Delta t}\|\hat{\psi}\|^2 + \frac{1}{2\lambda}\|\underset{\sim}{\nabla}_q\hat{\psi}\|^2 + \frac{1}{2\lambda}\left(2s-1-\frac{4s^2}{b}\right)\int_D \frac{|\underset{\sim}{q}|^2}{b}\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2}|\hat{\psi}|^2\,\mathrm{d}q \\ & - \frac{1}{4\lambda}\int_D \left[d + \frac{8s\lambda}{b}(\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}\,\underset{\sim}{q})\right]\left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}|\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}.\end{aligned}$$

Assuming that $b \geq 4s^2/(2s-1)$ with $s > 1/2$, and recalling that $|\underset{\sim}{q}| < \sqrt{b}$ for $\underset{\sim}{q} \in D$, we then have that

$$\mathfrak{C}(\hat{\psi},\hat{\psi}) \geq \frac{1}{\Delta t}\|\hat{\psi}\|^2 + \frac{1}{2\lambda}\|\underset{\sim}{\nabla}_q\hat{\psi}\|^2 - \frac{1}{4\lambda}(d + 8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)})\int_D \left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}|\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}.$$

Let us note that for, any $\beta > 0$,

$$\int_D \left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-1}|\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q} \leq \frac{1}{4\beta}\int_D |\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q} + \beta\int_D \left(1-\frac{|\underset{\sim}{q}|^2}{b}\right)^{-2}|\hat{\psi}|^2\,\mathrm{d}\underset{\sim}{q}. \tag{3.27}$$

Hence, by (2.4) and fixing β as the unique solution of the equation $4b\left(d + 8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)}\right)\beta = 1$, we have that

$$\mathfrak{C}(\hat{\psi},\hat{\psi}) \geq \frac{1}{\Delta t}\left(1 - \frac{b\Delta t}{4\lambda}(d + 8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)})^2\right)\|\psi\|^2 + \frac{1}{4\lambda}\|\underset{\sim}{\nabla}_q\hat{\psi}\|^2 \qquad \forall\hat{\psi}\in\mathrm{C}^\infty_0(D).$$

Recalling that $\mathrm{C}^\infty_0(D)$ is dense in $\mathrm{H}^1_0(D)$ and, by [10,11], also in the $(1-|\underset{\sim}{q}|^2/b)^{-2}$-weighted L^2 space, $\mathrm{L}^2_{M^{-4/b}}(D)$, we thus deduce that, for any $\Delta t < 4\lambda/(b(d+8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)})^2)$, the bilinear form $\mathfrak{C}$ is coercive on $\mathrm{H}^1_0(D)\times\mathrm{H}^1_0(D)$. The existence of a unique solution $\{\hat{\psi}^n\}_{n=0}^{N_T}$ to the semidiscretization (3.26) then follows from the Lax-Milgram theorem, as in the previous section. Using the above coercivity argument, the proof of stability of (3.26), stated in Lemma 3.7 below, is completely analogous to the proof of Lemma 3.1 and is therefore omitted.

Lemma 3.7 (stability inequality)**.** *Let $\Delta t = T/N_T$, $N_T \geq 1$, $\underset{\approx}{\kappa} \in (\mathrm{C}[0,T])^{d\times d}$, $\hat{\psi}^0 \in \mathrm{L}^2(D)$, $b \geq 4s^2/(2s-1)$ with $s > 1/2$, and define $c_0 := b(d + 8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)})^2/(2\lambda)$. If Δt is such that $0 < c_0\Delta t \leq 1/2$, then we have,*

for all m such that $1 \le m \le N_T$,

$$\|\hat{\psi}^m\|^2 + \sum_{n=0}^{m-1} \Delta t \left\| \frac{\hat{\psi}^{n+1} - \hat{\psi}^n}{\sqrt{\Delta t}} \right\|^2 + \sum_{n=0}^{m-1} \frac{\Delta t}{2\lambda} \|\underset{\sim}{\nabla}_q \hat{\psi}^{n+1}\|^2 \le \mathrm{e}^{2c_0 m \Delta t} \|\hat{\psi}^0\|^2.$$

The existence of a unique weak solution to (3.25) now follows from Lemma 3.7 by a weak compactness argument, in the same way as in the previous section in the case of a general Maxwellian. In particular (3.11) and (3.13) still hold with $\hat{\psi}^{\Delta t,(\pm)}$ defined by (3.8) and (3.9), but now using the sequence $\{\hat{\psi}^n : n = 0, \dots, N_T\}$ generated by (3.26); (3.12) also holds, with $\mathrm{H}^1_0(D;M)$ replaced by $\mathrm{H}^1_0(D)$. The rest of the argument is then identical to that in the proof of Theorem 3.2, using (3.27), Hardy's inequality (2.4) and the fact that $\mathrm{H}^1_0(D) \subset \mathrm{L}^2(D) = \mathrm{L}^2(D)' \subset \mathrm{H}^1_0(D)'$, where each space is dense in the next one in the chain, with continuous and injective embedding.

4. The fully-discrete method

We now return to the semidiscrete method (3.1) based on the symmetrized version of the Fokker-Planck equation and describe the construction of a fully-discrete numerical method that stems from this semidiscretization. At the end of the section we shall comment on the extension of our results to a fully-discrete method based on the semidiscretization (3.26) of the Chauvière-Lozinski-transformed Fokker-Planck equation (3.24) for the FENE model.

Let $\mathcal{P}_N(D)$ be a finite-dimensional subspace of $\mathrm{H}^1_0(D;M)$, to be chosen below, and let $\hat{\psi}^n_N \in \mathcal{P}_N(D)$ be the solution at time level n of our fully-discrete Galerkin method:

$$\int_D \frac{\hat{\psi}^{n+1}_N - \hat{\psi}^n_N}{\Delta t}\, \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\, \underset{\sim}{q}\, \hat{\psi}^{n+1}_N) \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda} \int_D \underset{\sim}{\nabla}_M \hat{\psi}^{n+1}_N \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} = 0$$
$$\forall \hat{\varphi} \in \mathcal{P}_N(D), \quad n = 0, \dots, N_T - 1, \tag{4.1}$$

$$\hat{\psi}^0_N(\cdot) := \text{the } \mathrm{L}^2(D) \text{ orthogonal projection of } \hat{\psi}_0(\cdot) = \hat{\psi}(\cdot, 0) \text{ onto } \mathcal{P}_N(D). \tag{4.2}$$

Remark 4.1. If the linear space $\mathcal{P}_N(D)$ is selected so that $\sqrt{M} \in \mathcal{P}_N(D)$, then, since $\sqrt{M} \in \mathrm{Ker}(\underset{\sim}{\nabla}_M)$, it follows on taking $\hat{\varphi} = \sqrt{M}$ in (4.1) that

$$\int_D \sqrt{M(\underset{\sim}{q})}\, \hat{\psi}^n_N(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q} = \int_D \sqrt{M(\underset{\sim}{q})}\, \hat{\psi}^0_N(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q}, \qquad n = 1, \dots, N_T,$$

whereby, on letting $\psi^n_N := \sqrt{M}\hat{\psi}^n_N$, we have that

$$\int_D \psi^n_N(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q} = \int_D \psi^0_N(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q}, \qquad n = 1, \dots, N_T.$$

The function ψ^n_N represents an approximation to the probability density function $\psi = \sqrt{M}\hat{\psi}$ at $t = t^n$. Since, by Lemma 3.3, $\int_D \psi(\underset{\sim}{q}, t)\, \mathrm{d}\underset{\sim}{q} = \int_D \psi_0(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q} = 1$ for all $t \ge 0$, we deduce, by choosing $\mathcal{P}_N(D)$ so that $\sqrt{M} \in \mathcal{P}_N(D)$, that this integral identity is preserved under discretization. The integral $\int_D \psi(\underset{\sim}{q}, t)\, \mathrm{d}\underset{\sim}{q}$ will sometimes be referred to as the *volume* of ψ.

Our objective is to derive a bound on the global error $e^n_N := \hat{\psi}(\cdot, t^n) - \hat{\psi}^n_N$. Clearly,

$$e^n_N = (\hat{\psi}(\cdot, t^n) - \hat{\Pi}_N \hat{\psi}(\cdot, t^n)) + (\hat{\Pi}_N \hat{\psi}(\cdot, t^n) - \hat{\psi}^n_N) =: \eta^n + \xi^n,$$

where $\hat{\Pi}_N \hat{\psi}(\cdot, t^n) \in \mathcal{P}_N(D)$ is a certain projection of $\hat{\psi}(\cdot, t^n)$ onto $\mathcal{P}_N(D)$ that will be defined below. For the moment, the specific choices of $\mathcal{P}_N \subset \mathrm{H}^1_0(D;M)$ and $\hat{\Pi}_N$ are irrelevant.

We begin by bounding norms of ξ in terms of suitable norms of η. Substituting ξ into (4.1), setting $\hat{\varphi} = \xi^{n+1}$, and noting that $\xi^n = \hat{\psi}(\cdot, t^n) - \hat{\psi}_N^n - \eta^n$, we have

$$\int_D \frac{\xi^{n+1} - \xi^n}{\Delta t}\, \xi^{n+1}\, \mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\, \underset{\sim}{q}\, \xi^{n+1}) \cdot \underset{\sim}{\nabla}_M \xi^{n+1}\, \mathrm{d}\underset{\sim}{q} + \frac{1}{2\lambda} \int_D \underset{\sim}{\nabla}_M \xi^{n+1} \cdot \underset{\sim}{\nabla}_M \xi^{n+1}\, \mathrm{d}\underset{\sim}{q} = \int_D \mu^{n+1}\, \xi^{n+1}\, \mathrm{d}\underset{\sim}{q} + \int_D \underset{\sim}{\nu}^{n+1} \cdot \underset{\sim}{\nabla}_M \xi^{n+1}\, \mathrm{d}\underset{\sim}{q}, \quad (4.3)$$

for $n = 0, \ldots, N_T - 1$, where

$$\mu^{n+1} := \left(\frac{\hat{\psi}(\cdot, t^{n+1}) - \hat{\psi}(\cdot, t^n)}{\Delta t} - \frac{\partial \hat{\psi}}{\partial t}(\cdot, t^{n+1}) \right) - \frac{\eta^{n+1} - \eta^n}{\Delta t}, \quad (4.4)$$

$$\underset{\sim}{\nu}^{n+1} := \underset{\approx}{\kappa}^{n+1} \underset{\sim}{q} \eta^{n+1} - \frac{1}{2\lambda} \underset{\sim}{\nabla}_M \eta^{n+1}. \quad (4.5)$$

Since $\mathcal{P}_N(D) \subset \mathrm{H}_0^1(D; M)$, (4.3) is in the form of (3.4); hence, applying Lemma 3.1, we obtain

$$\|\xi^m\|^2 + \frac{1}{2\lambda} \sum_{n=0}^{m-1} \Delta t \|\underset{\sim}{\nabla}_M \xi^{n+1}\|^2 \le \mathrm{e}^{2c_0 m \Delta t} \left\{ \|\xi^0\|^2 + \sum_{n=0}^{m-1} 2\Delta t \left(\|\mu^{n+1}\|^2 + 4\lambda \|\underset{\sim}{\nu}^{n+1}\|^2 \right) \right\}, \quad (4.6)$$

for $m = 1, \ldots, N_T$. Let us first consider the term $\|\xi^0\|$ on the right-hand side of (4.6). Since $\hat{\psi}_N^0$ is the $\mathrm{L}^2(D)$ orthogonal projection of $\hat{\psi}(\cdot, 0) = \hat{\psi}_0$ onto $\mathcal{P}_N(D)$, we have $(\xi^0, \hat{\varphi}_N) = -(\eta^0, \hat{\varphi}_N)$ for all $\hat{\varphi}_N \in \mathcal{P}_N(D)$. Setting $\hat{\varphi}_N = \xi^0$ here and applying the Cauchy-Schwarz inequality on the right-hand side yields $\|\xi^0\| \le \|\eta^0\|$.

By the triangle inequality we have the following bound on $\|\underset{\sim}{\nu}^{n+1}\|$:

$$\|\underset{\sim}{\nu}^{n+1}\| \le \sqrt{b}\, |\underset{\approx}{\kappa}^{n+1}|\, \|\eta^{n+1}\| + \frac{1}{2\lambda} \|\underset{\sim}{\nabla}_M \eta^{n+1}\|, \qquad n = 0, \ldots, N_T - 1.$$

Hence for the third term on the right-hand-side of (4.6), we have

$$\begin{aligned} \sum_{n=0}^{m-1} 8\lambda \Delta t \|\underset{\sim}{\nu}^{n+1}\|^2 &\le \sum_{n=0}^{m-1} \Delta t \left(16\lambda b |\underset{\approx}{\kappa}^{n+1}|^2 \|\eta^{n+1}\|^2 + \frac{4}{\lambda} \|\underset{\sim}{\nabla}_M \eta^{n+1}\|^2 \right) \\ &\le 4c_2 \sum_{n=0}^{m-1} \Delta t \|\eta^{n+1}\|_{\mathrm{H}_0^1(D;M)}^2 = 4c_2 \|\eta\|_{\ell^2(0,t^m;\mathrm{H}_0^1(D;M))}^2, \end{aligned}$$

for $m = 1, \ldots, N_T$, where $c_2 := \max\left(1/\lambda\,,\, 4\lambda b \|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)}^2\right)$.

It remains to bound $\|\mu^{m+1}\|$. We begin by observing that

$$\|\mu^{m+1}\| \le \left\| \frac{\hat{\psi}(\cdot, t^{n+1}) - \hat{\psi}(\cdot, t^n)}{\Delta t} - \frac{\partial \hat{\psi}}{\partial t}(\cdot, t^{n+1}) \right\| + \left\| \frac{\eta^{n+1} - \eta^n}{\Delta t} \right\| =: I + II.$$

Bounding both I and II by Taylor's theorem with integral remainder yields

$$I^2 \le \Delta t \int_{t^n}^{t^{n+1}} \left\| \frac{\partial^2 \hat{\psi}}{\partial t^2}(\cdot, t) \right\|^2 \mathrm{d}t \qquad \text{and} \qquad II^2 \le \int_D \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \left| \frac{\partial \eta}{\partial t}(\underset{\sim}{q}, t) \right|^2 \mathrm{d}t\, \mathrm{d}\underset{\sim}{q} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \left\| \frac{\partial \eta}{\partial t}(\cdot, t) \right\|^2 \mathrm{d}t.$$

Therefore, we now have that

$$\begin{aligned}\sum_{n=0}^{m-1} 2\Delta t\|\mu^{n+1}\|^2 &\leq 4\sum_{n=0}^{m-1}\Delta t^2\int_{t^n}^{t^{n+1}}\left\|\frac{\partial^2\hat{\psi}}{\partial t^2}(\cdot,t)\right\|^2\mathrm{d}t + 4\sum_{n=0}^{m-1}\int_{t^n}^{t^{n+1}}\left\|\frac{\partial\eta}{\partial t}(\cdot,t)\right\|^2\mathrm{d}t\\ &= 4\Delta t^2\left\|\frac{\partial^2\hat{\psi}}{\partial t^2}\right\|^2_{\mathrm{L}^2(0,t^m;\mathrm{L}^2(D))} + 4\left\|\frac{\partial\eta}{\partial t}\right\|^2_{\mathrm{L}^2(0,t^m;\mathrm{L}^2(D))}.\end{aligned}$$

Combining the bounds on the three terms on the right-hand side of (4.6) we deduce that

$$\begin{aligned}&\|\xi^m\|^2 + \frac{1}{2\lambda}\sum_{n=0}^{m-1}\Delta t\|\underset{\sim}{\nabla}_M\xi^{n+1}\|^2 \leq\\ &\quad \mathrm{e}^{2c_0m\Delta t}\left(\|\eta^0\|^2 + 4c_2\|\eta\|^2_{\ell^2(0,t^m;\mathrm{H}^1_0(D;M))} + 4\left\|\frac{\partial\eta}{\partial t}\right\|^2_{\mathrm{L}^2(0,t^m;\mathrm{L}^2(D))} + 4\Delta t^2\left\|\frac{\partial^2\hat{\psi}}{\partial t^2}\right\|^2_{\mathrm{L}^2(0,t^m;\mathrm{L}^2(D))}\right). \qquad (4.7)\end{aligned}$$

It remains to bound the first three terms in the bracket on the right-hand side of (4.7). To do so we need to make a specific choice of the finite-dimensional space $\mathcal{P}_N(D)$ from which approximations to $\hat{\psi}\in\mathrm{H}^1_0(D;M)$ are sought, and we also need to specify the projector $\hat{\Pi}_N$. These issues will be discussed in the next section. We shall then return, in Section 6, to (4.7) and complete the convergence analysis of the numerical method.

Remark 4.2. In the case of the FENE model with $b\geq 4s^2/(2s-1)$ and $s>1/2$ a bound analogous to (4.7) can be shown to hold for the fully-discrete version of the semidiscretization (3.26) based on a Chauvière and Lozinski type transformation, with suitable fixed positive constants c_0 and c_2, except that $\mathcal{P}_N(D)$ is then taken to be a finite-dimensional subspace of $\mathrm{H}^1_0(D)$, $\underset{\sim}{\nabla}_M\xi^{n+1}$ on the left-hand side of the bound (4.7) is replaced by $\underset{\sim}{\nabla}_q\xi^{n+1}$, and the norm $\|\cdot\|_{\ell^2(0,t^m;\mathrm{H}^1_0(D;M))}$ on the right-hand side of (4.7) is replaced by $\|\cdot\|_{\ell^2(0,t^m;\mathrm{H}^1_0(D))}$. The main steps of the proof are identical to those above: the Cauchy-Schwarz inequality and inequalities (2.4) and (3.27) are used in the course of bounding the terms on the right-hand side of an error identity analogous to (4.3) relating the sequence $\{\xi^m\}_{m=0}^{N_T}$ to the sequence $\{\eta^m\}_{m=0}^{N_T}$, while the terms on the left-hand side of the error identity are bounded below as in the proof the stability inequality stated in Lemma 3.7.

We note in particular that the fully-discrete version of the semidiscretization (3.26) based on a Chauvière and Lozinski type transformation $\hat{\psi}=\psi/M^{2s/b}$ and the finite-dimensional Galerkin subspace $\mathcal{P}_N(D)\subset\mathrm{H}^1_0(D)$ is unconditionally stable in the sense that the sequence of numerical solutions $\{\hat{\psi}_N^n\}_{n=0}^{N_T}$ generated by the fully-discrete scheme satisfies the stability inequality stated in Lemma 3.7, with $\Delta t=T/N_T$, $N_T\geq 1$, $\underset{\approx}{\kappa}\in(\mathrm{C}[0,T])^{d\times d}$, $\hat{\psi}_N^0\in\mathcal{P}_N(D)$, $b\geq 4s^2/(2s-1)$, $s>1/2$, $c_0:=b(d+8s\lambda\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)})^2/(2\lambda)$, $0<c_0\Delta t\leq 1/2$, and ψ^m, ψ^{m-1} and ψ^0 replaced by ψ_N^m, ψ_N^{m-1} and ψ_N^0, respectively, without any conditions relating Δt to N. The proof of this is identical to that of Lemma 3.7, *mutatis mutandis*. We thus deduce that for $b\gg 1$ a time-step limitation of the form $\Delta t=\mathcal{O}(b^{-1})$ is needed in order to ensure that $0<c_0\Delta t\leq 1/2$, and thereby the stability of the method. In this respect the scheme behaves identically to the fully-discrete numerical method (4.1), (4.2), based on the symmetrized form of the Fokker-Planck equation (*cf.* the conditions of Lem. 3.1, for example).

5. Approximation results

We showed in Section 2.2(b) that, under Hypotheses A and B stated in the Introduction, $\mathrm{H}^1_0(D)\subset\mathrm{H}^1(D;M)=\mathrm{H}^1_0(D;M)$. Therefore, any finite-dimensional space $\mathcal{P}_N(D)\subset\mathrm{H}^1_0(D)$ is, trivially, also contained in $\mathrm{H}^1_0(D;M)$. Our aim now is to make a specific choice of $\mathcal{P}_N(D)$ and to explore the approximation properties of our chosen space.

Remark 5.1. We noted in Remark 4.1 that if, in addition, $\sqrt{M} \in \mathcal{P}_N(D)$, then

$$\int_D \psi_N^n(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q} = \int_D \psi_N^0(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}.$$

Since, by Hypothesis B, $\sqrt{M} \in \mathrm{H}_0^1(D)$, one can ensure that this integral identity holds by including $\sqrt{M}$ in the finite-dimensional space $\mathcal{P}_N(D)$.

Our definition of $\mathcal{P}_N(D)$ and the choice of the projector $\hat{\Pi}_N : \mathrm{H}_0^1(D;M) \to \mathcal{P}_N(D)$ depend on the number d of space dimensions. Since the case of $d = 2$ is sufficiently representative, for the sake of brevity and ease of presentation we shall confine ourselves to two space dimensions, that is, when D is a disc of radius $\sqrt{b}$ in $\mathbb{R}^2$.

Let D_0 denote the slit disc $D_0 := D \setminus \{(q_1, 0) : 0 \le q_1 < \sqrt{b}\}$. It is natural to transform D_0 into the rectangle $(r,\theta) \in R := (0,1) \times (0, 2\pi)$ in a polar co-ordinate system, using the (bijective) change of variables $\underset{\sim}{q} = (q_1, q_2) = (\sqrt{b}\, r\cos\theta, \sqrt{b}\, r \sin\theta) \in D_0$ where $(r,\theta) \in R$. Given $f \in \mathrm{H}^1(D)$, we define $\tilde{f}$ on R by

$$\tilde{f}(r,\theta) = f(q_1,q_2), \qquad \underset{\sim}{q} = (q_1,q_2) \in D_0, \quad (r,\theta) \in R, \quad q_1 = \sqrt{b}\, r\cos\theta, \quad q_2 = \sqrt{b}\, r\sin\theta. \tag{5.1}$$

Thus,

$$\|f\|^2_{\mathrm{H}^1(D)} = \|f\|^2_{\mathrm{H}^1(D_0)} = \int_0^1 r \int_0^{2\pi} \left(b|\tilde{f}|^2 + |\mathrm{D}_r \tilde{f}|^2 + \left| \frac{\mathrm{D}_\theta \tilde{f}}{r} \right|^2 \right) \mathrm{d}\theta\, \mathrm{d}r.$$

Motivated by this identity and writing, here and henceforth, $\tilde{w}(r) := r$ for our weight-function on the interval $(0,1)$, we define the space

$$\tilde{\mathrm{H}}^1_{\tilde{w}}(R) := \{\tilde{f} \in \mathrm{L}^2_{\mathrm{loc}}(0,1;\mathrm{H}^1_p(0,2\pi)) : \tilde{f} \in \mathrm{L}^2_{\tilde{w}}(R), \quad \mathrm{D}_r\tilde{f} \in \mathrm{L}^2_{\tilde{w}}(R) \quad \text{and} \quad \frac{1}{r}\mathrm{D}_\theta \tilde{f} \in \mathrm{L}^2_{\tilde{w}}(R)\}, \tag{5.2}$$

equipped with the norm $\|\cdot\|_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)}$ defined by

$$\|\tilde{f}\|^2_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)} := \int_0^1 \tilde{w}(r) \int_0^{2\pi} \left(|\tilde{f}|^2 + |\mathrm{D}_r\tilde{f}|^2 + \left| \frac{\mathrm{D}_\theta \tilde{f}}{r} \right|^2 \right) \mathrm{d}\theta\, \mathrm{d}r, \tag{5.3}$$

where $\mathrm{L}^2_{\tilde{w}}(R)$ is the $\tilde{w}$-weighted space of square-integrable functions on R, with norm $\|\cdot\|_{\mathrm{L}^2_{\tilde{w}}(R)}$ defined by

$$\|\tilde{f}\|^2_{\mathrm{L}^2_{\tilde{w}}(R)} := \int_0^1 \tilde{w}(r) \int_0^{2\pi} |\tilde{f}(r,\theta)|^2\, \mathrm{d}\theta\, \mathrm{d}r = \int_R |\tilde{f}(r,\theta)|^2\, r\, \mathrm{d}r\, \mathrm{d}\theta,$$

and, for a non-negative integer t, the periodic Sobolev space $\mathrm{H}^t_p(0,2\pi)$ is given by

$$\mathrm{H}^t_p(0,2\pi) := \{\tilde{f} \in \mathrm{H}^t_{\mathrm{loc}}(\mathbb{R}) : \tilde{f}(\theta + 2\pi) = \tilde{f}(\theta) \quad \forall \theta \in \mathbb{R}\}.$$

Let $\tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$ denote the subspace of $\tilde{\mathrm{H}}^1_{\tilde{w}}(R)$ consisting of all functions $\tilde{f}$ such that the trace $\tilde{f}(1,\cdot) = 0$.

For non-negative integers s, t we define the weighted space $\mathrm{H}^{s,t}_{\tilde{w}}(R) := \mathrm{H}^s_{\tilde{w}}(0,1;\mathrm{H}^t_p(0,2\pi))$, equipped with the norm $\|\cdot\|_{\mathrm{H}^{s,t}_{\tilde{w}}(R)}$ given by:

$$\|\tilde{f}\|^2_{\mathrm{H}^{s,t}_{\tilde{w}}(R)} := \sum_{0\le i\le s,\, 0\le j\le t} \int_0^1 \tilde{w}(r) \int_0^{2\pi} |\mathrm{D}_r^i \mathrm{D}_\theta^j\, \tilde{f}(r,\theta)|^2\, \mathrm{d}\theta\, \mathrm{d}r.$$

Similarly, for integers $s \geq 1$ and $t \geq 0$, we define $\mathrm{H}^{s,t}_{\tilde{w},0}(R) := \mathrm{H}^s_{\tilde{w},0}(0,1;\mathrm{H}^t_p(0,2\pi))$, where $\mathrm{H}^s_{\tilde{w},0}(0,1) := \mathrm{H}^s_{\tilde{w}}(0,1) \cap \mathrm{H}^1_{\tilde{w},0}(0,1)$. Here, $\mathrm{H}^1_{\tilde{w},0}(0,1)$ denotes the set of all $\tilde{u} \in \mathrm{H}^1_{\tilde{w}}(0,1)$ such that $\tilde{u}(1) = 0$, endowed with the following inner product and norm:

$$(\tilde{u},\tilde{v})_{\mathrm{H}^1_{\tilde{w},0}(0,1)} := \int_0^1 \tilde{w}(r)\,\mathrm{D}_r\tilde{u}\,\mathrm{D}_r\tilde{v}\,\mathrm{d}r \qquad \text{and} \qquad \|\tilde{u}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} := \{(\tilde{u},\tilde{u})_{\mathrm{H}^1_{\tilde{w},0}(0,1)}\}^{\frac{1}{2}}.$$

Note that $\tilde{w}$ is a Jacobi weight function (*i.e.*, of the form $(1-s)^\alpha(1+s)^\beta, s \in (-1,1)$ with $\alpha,\beta > -1$) when transformed to $(-1,1)$.

We now introduce the projection operators that we will use. Due to the Cartesian product structure of the set R it is natural to define distinct projection operators in the r and θ co-ordinate directions. In the θ-direction we use the orthogonal projection in the $\mathrm{L}^2(0,2\pi)$ inner product (*i.e.*, truncation of the Fourier series) denoted, for $N \geq 1$, by $P^F_N : \mathrm{L}^2(0,2\pi) \to \mathbb{S}_N(0,2\pi)$ where $\mathbb{S}_N(0,2\pi)$ is the space of all trigonometric polynomials in $\theta \in [0,2\pi]$ of degree N or less. We denote by $\mathbb{S}_{N_\theta,0}(0,2\pi)$ the orthogonal complement in $\mathbb{S}_{N_\theta}(0,2\pi)$, with respect to the $\mathrm{L}^2(0,2\pi)$ inner product, of the one-dimensional subspace spanned by constant functions.

The appropriate choice of projector in the r-direction is less immediate. We define, for $N \geq 1$, the operator $P^J_N : \mathrm{H}^1_{\tilde{w},0}(0,1) \to \mathbb{P}_{N,0}(0,1)$ as the orthogonal projection in the $\mathrm{H}^1_{\tilde{w},0}(0,1)$ inner product, where $\mathbb{P}_{N,0}(0,1)$ is the space of all algebraic polynomials in $r \in [0,1]$, of degree N or less, that vanish at $r = 1$.

It is tempting to define a two-dimensional projector onto $\mathbb{S}_N(0,2\pi) \otimes \mathbb{P}_{N,0}(0,1)$ as the tensor product of the projectors P^F_N and P^J_N. Unfortunately, this choice is inadequate due to the presence of the singular factor $1/r$ in the weighted Sobolev norm $\|\cdot\|_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)}$, and a different definition is required. In order to motivate our choice of the two-dimensional projector below, we state the following result that can be seen as a variant of the Malgrange preparation theorem [22].

Lemma 5.2 (decomposition lemma)**.** *Let $\tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w}}(R)$ and, for $\varepsilon \in (0,1)$, define $R_\varepsilon := (\varepsilon,1)\times(0,2\pi)$. There exist $\tilde{g}_1 \in \mathrm{H}^1_{\tilde{w}}(0,1)$ and $\tilde{g}_2 \in \mathrm{H}^{0,1}_{\tilde{w}}(R)$, with $\tilde{g}_2 \in \mathrm{H}^1(R_\varepsilon)$ for each $\varepsilon \in (0,1)$ and $r\tilde{g}_2 \in \tilde{\mathrm{H}}^1_{\tilde{w}}(R)$, such that*

$$\tilde{g}(r,\theta) = \tilde{g}_1(r) + r\tilde{g}_2(r,\theta) \quad \textit{for a.e. } (r,\theta) \in R \qquad \textit{and} \qquad \tilde{g}_1(r) := \frac{1}{2\pi}(g(r,\cdot),1)_{\mathrm{L}^2(0,2\pi)}.$$

This is the unique such decomposition of $\tilde{g}$. If $\tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$, then $\tilde{g}_1 \in \mathrm{H}^1_{\tilde{w},0}(0,1)$ and $r\tilde{g}_2 \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$, with $\tilde{g}_2(1,\cdot) = 0$ in the sense of the trace theorem on $\mathrm{H}^1(R_\varepsilon)$, $\varepsilon \in (0,1)$.

Proof. Let $\tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w}}(R)$; then, by virtue of Fubini's theorem, $\tilde{g}(r,\cdot) \in \mathrm{H}^1_p(0,2\pi)$ for a.e. $r \in (0,1)$. Let us define, for $r \in (0,1)$, the Fourier coefficients of $\tilde{g}(r,\cdot)$ by

$$\tilde{\gamma}_n(r) := \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} \tilde{g}(r,\theta)\exp(-\mathrm{i}n\theta)\,\mathrm{d}\theta, \qquad n = 0,1,\ldots$$

According to Parseval's identity,

$$\|\tilde{g}\|^2_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)} = \sum_{n\in\mathbb{Z}}\int_0^1 \left(|\tilde{\gamma}_n(r)|^2 + |\tilde{\gamma}_n'(r)|^2 + n^2\left|\frac{\tilde{\gamma}_n(r)}{r}\right|^2\right) r\,\mathrm{d}r < \infty,$$

whereby, in particular, $\tilde{\gamma}_0 \in \mathrm{H}^1_{\tilde{w}}(0,1)$ and

$$\tilde{\gamma}_n \in \mathrm{H}^1(0,1;r^{-1},r) := \left\{\tilde{f} \in \mathrm{H}^1_{\mathrm{loc}}(0,1) \,:\, \int_0^1 \left(r^{-1}|\tilde{f}(r)|^2 + r|\tilde{f}'(r)|^2\right)\,\mathrm{d}r < \infty\right\} \qquad \forall n \in \mathbb{Z}\setminus\{0\}.$$

For any $\varepsilon \in (0,1)$ and $n \in \mathbb{Z}\setminus\{0\}$, $\tilde\gamma_n \in \mathrm{H}^1(\varepsilon,1)$, and hence $\tilde\gamma_n \in \mathrm{C}(0,1]$. Also, for $0 < r_1 < r_2 < 1$,

$$\begin{aligned}\tilde\gamma_n(r_2)^2 - \tilde\gamma_n(r_1)^2 &= \int_{r_1}^{r_2} \frac{\mathrm{d}}{\mathrm{d}s}(\tilde\gamma_n(s)^2)\,\mathrm{d}s = 2\int_{r_1}^{r_2} \frac{\tilde\gamma_n(s)}{\sqrt{s}}\sqrt{s}\,\tilde\gamma_n'(s)\,\mathrm{d}s \\ &\le 2\left(\int_{r_1}^{r_2} s^{-1}|\tilde\gamma_n(s)|^2\,\mathrm{d}s\right)^{\frac12}\left(\int_{r_1}^{r_2} s|\tilde\gamma_n'(s)|^2\,\mathrm{d}s\right)^{\frac12},\end{aligned}$$

which is finite by the definition of $\mathrm{H}^1(0,1;r^{-1},r)$, and hence the left-most integral above is finite also. Since the integral is a continuous function of its limits, it follows that $\tilde\gamma_n^2 \in \mathrm{C}[0,1]$, and hence that $|\tilde\gamma_n| = \sqrt{\tilde\gamma_n^2} \in \mathrm{C}[0,1]$. Therefore, we have that (for $n \in \mathbb{Z}\setminus\{0\}$) $|\tilde\gamma_n| \in \mathrm{C}[0,1]$ and $\tilde\gamma_n \in \mathrm{C}(0,1]$, and it follows straightforwardly that $\tilde\gamma_n \in \mathrm{C}[0,1]$, $n \in \mathbb{Z}\setminus\{0\}$. However, Parseval's identity above then implies that, necessarily, $\tilde\gamma_n(0) = 0$ for all $n \in \mathbb{Z}\setminus\{0\}$.

Now, let us define $\tilde G_n(r) := \tilde\gamma_n(r)/r$ for $n \in \mathbb{Z}\setminus\{0\}$, $r \in (0,1]$ and $\tilde E_n(\theta) := (\exp(\mathrm{i}n\theta))/\sqrt{2\pi}$, $n \in \mathbb{Z}$, $\theta \in [0,2\pi]$. By Parseval's identity, again, $\sqrt{r^2+n^2}\,\tilde G_n \in \mathrm{L}^2_{\tilde w}(0,1)$, $n \in \mathbb{Z}\setminus\{0\}$. With these definitions, we have the following Fourier series expansion of $\tilde g$:

$$\tilde g = \frac{1}{\sqrt{2\pi}}\tilde\gamma_0 + r\sum_{n\in\mathbb{Z}\setminus\{0\}} \tilde G_n \tilde E_n,$$

with equality in the sense of $\tilde{\mathrm{H}}^1_{\tilde w}(R)$. We define $\tilde g_1 := \tilde\gamma_0/\sqrt{2\pi}$ and $\tilde g_2 = \sum_{n\in\mathbb{Z}\setminus\{0\}} \tilde G_n\tilde E_n$ to deduce the stated decomposition $\tilde g(r,\theta) = \tilde g_1(r) + r\tilde g_2(r,\theta)$, and we note that $\tilde g_1 = \frac{1}{2\pi}(\tilde g,1)_{\mathrm{L}^2(0,2\pi)} \in \mathrm{H}^1_{\tilde w}(0,1)$ and $\tilde g_2 \in \mathrm{H}^{0,1}_{\tilde w}(R)$; moreover, trivially, $r\tilde g_2 = \tilde g - \tilde g_1 \in \tilde{\mathrm{H}}^1_{\tilde w}(R)$. Also, since $\tilde g \in \tilde{\mathrm{H}}^1_{\tilde w}(R)$ it follows that $\tilde g \in \mathrm{H}^1(R_\varepsilon)$ and $\tilde g_1 \in \mathrm{H}^1(\varepsilon,1)$ for any $\varepsilon \in (0,1)$. Hence, $\tilde g_2 = (\tilde g - \tilde g_1)/r \in \mathrm{H}^1(R_\varepsilon)$ for any $\varepsilon \in (0,1)$.

For $\tilde g_1 = \tilde\gamma_0/\sqrt{2\pi}$ fixed, as in the statement of the lemma, the uniqueness of $\tilde g_2$ follows easily by *reductio ad absurdum*: suppose that $\tilde h_2$ is another function, with the same regularity properties as $\tilde g_2$, and such that $\tilde g = \tilde g_1 + r\tilde h_2$. Then, $r(\tilde h_2 - \tilde g_2) = 0$ a.e. on R, and therefore $\tilde h_2 = \tilde g_2$ a.e. on R.

The final statement of the lemma follows directly from the definitions of $\tilde\gamma_n$, $n \in \mathbb{Z}$ and the definitions of $\tilde g_1$ and $\tilde g_2$ *via* the $\tilde\gamma_n$, $n \in \mathbb{Z}$. □

Suppose that $\tilde g \in \tilde{\mathrm{H}}^1_{\tilde w,0}(R)$. On applying Lemma 5.2 we deduce that $\tilde g$ has the (unique) decomposition

$$\tilde g(r,\theta) = \tilde g_1(r) + r\tilde g_2(r,\theta), \tag{5.4}$$

where $\tilde g_1 := \frac{1}{2\pi}(\tilde g,1)_{\mathrm{L}^2(0,2\pi)} \in \mathrm{H}^1_{\tilde w,0}(0,1)$, $\tilde g_2 \in \mathrm{H}^{0,1}_{\tilde w}(R)$ and $\tilde g_2(1,\cdot) = 0$. Note also that $(g_2(r,\cdot),1)_{\mathrm{L}^2(0,2\pi)} = 0$ for a.e. $r \in (0,1)$. We shall assume in addition that $\tilde g_2(\cdot,\theta) \in \mathrm{H}^1_{\tilde w,0}(0,1)$ for a.e. $\theta \in (0,2\pi)$; by virtue of Fubini's theorem, a convenient sufficient condition for this is that $\tilde g_2 \in \mathrm{H}^{1,0}_{\tilde w,0}(R)$, for example. We then define

$$\tilde P_N^J \tilde g(\cdot,\theta) := P_N^J \tilde g_1(\cdot) + rP_N^J \tilde g_2(\cdot,\theta), \qquad \theta \in (0,2\pi),$$

where $P_N^J : \mathrm{H}^1_{\tilde w,0}(0,1) \to \mathbb{P}_{N,0}(0,1)$ is the orthogonal projector defined above.

There are a number of approximation results available in the literature related to projectors in Jacobi-weighted inner products (see for example [9] or [15]). Since the setting here is specific, we shall establish the required approximation properties of the univariate projector P_N^J from first principles. The approximation properties of $\tilde P_N^J$ and of our two-dimensional projector $P_N^F\tilde P_N^J$ will then follow. The relevant results are stated in the next two lemmas.

Lemma 5.3. *Suppose that $\tilde g \in \mathrm{H}^k_{\tilde w,0}(0,1)$ with $k \ge 1$; then,*

$$\|\tilde g - P_N^J\tilde g\|_{\mathrm{H}^1_{\tilde w}(0,1)} \le cN^{1-k}\|\tilde g\|_{\mathrm{H}^k_{\tilde w}(0,1)} \tag{5.5}$$

and

$$\|\tilde{g} - P_N^J \tilde{g}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)} \leq cN^{-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}. \tag{5.6}$$

Proof. Let us first prove (5.5). Note that by Pythagoras' theorem,

$$\|\tilde{g} - P_N^J \tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} = \left(\|\tilde{g}\|^2_{\mathrm{H}^1_{\tilde{w},0}(0,1)} - \|P_N^J \tilde{g}\|^2_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \right)^{\frac{1}{2}} \leq \|\tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \leq \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}.$$

If $k = 1$, the right-most term in this chain is equal to $1 \cdot N^{1-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}$, while if $k \geq 2$ and $1 \leq N < k-1$, then it is bounded by $(k-1)^{k-1} N^{1-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}$.

Finally, if $k \geq 2$ and $N \geq \max(2, k-1)$, then we recall that, by the definition of P_N^J,

$$\|\tilde{g} - P_N^J \tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \leq \|\tilde{g} - \tilde{v}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \qquad \forall \tilde{v} \in \mathbb{P}_{N,0}(0,1).$$

Select, in particular,

$$\tilde{v}(r) = -\int_r^1 Q_{N-1}^J \mathrm{D}_s \tilde{g}(s)\, \mathrm{d}s, \qquad r \in [0,1],$$

where Q_{N-1}^J is the orthogonal projector in $\mathrm{L}^2_{\tilde{w}}(0,1)$ onto $\mathbb{P}_{N-1}(0,1)$, the set of all algebraic polynomials of degree $N-1$ or less on the interval $[0,1]$. Thus,

$$\|\tilde{g} - P_N^J \tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \leq \|\mathrm{D}_r \tilde{g} - \mathrm{D}_r \tilde{v}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)} = \|\mathrm{D}_r \tilde{g} - Q_{N-1}^J(\mathrm{D}_r \tilde{g})\|_{\mathrm{L}^2_{\tilde{w}}(0,1)} \leq c\,(N-1)^{1-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)},$$

where the last bound (scaled from the standard interval $(-1,1)$ to $(0,1)$) comes from Section 5.7.1 of Canuto *et al.* [15], and is valid for $N \geq \max(2, k-1)$, $k \geq 2$. Hence, after bounding $(N-1)^{1-k}$ by $2^{k-1} N^{1-k}$ (recall that $N \geq 2$ by hypothesis), we deduce that

$$\|\tilde{g} - P_N^J \tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \leq c\, 2^{k-1} N^{1-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}.$$

Now choosing $\hat{c} = \max\{(k-1)^{k-1}, c\, 2^{k-1}\}$ for $k \geq 1$, with the convention that $0^0 := 1$, we have that

$$\|\tilde{g} - P_N^J \tilde{v}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \leq \hat{c} N^{1-k} \|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}$$

for all $N \geq 1$ (regardless of whether or not $N \geq k-1$). Since by the Friedrichs inequality

$$\|\tilde{v}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)} \leq \frac{1}{2} \|\mathrm{D}_r \tilde{v}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)} \qquad \forall \tilde{v} \in \mathrm{H}^1_{\tilde{w},0}(0,1) \tag{5.7}$$

$\|\cdot\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)}$ and $\|\cdot\|_{\mathrm{H}^1_{\tilde{w}}(0,1)}$ are equivalent norms on $\mathrm{H}^1_{\tilde{w},0}(0,1)$, we deduce (5.5) for any $N \geq 1$.

The proof of (5.6) is based on a duality argument. Let $e := \tilde{g} - P_N^J \tilde{g}$ and note that, by the hypotheses of the lemma on $\tilde{g}$, we have $e \in \mathrm{L}^2_{\tilde{w}}(0,1)$. Consider the mixed Neumann-Dirichlet boundary-value problem:

$$-\mathrm{D}_r(r \mathrm{D}_r z_e(r)) = r\, e(r), \quad r \in (0,1), \qquad \lim_{r \to 0_+} r \mathrm{D}_r z_e(r) = 0, \quad z_e(1) = 0. \tag{5.8}$$

By (5.7) and the Lax-Milgram theorem, this has a unique weak solution $z_e \in \mathrm{H}^1_{\tilde{w},0}(0,1)$ satisfying

$$(z_e, v)_{\mathrm{H}^1_{\tilde{w},0}(0,1)} = (e, v)_{\mathrm{L}^2_{\tilde{w}}(0,1)} \qquad \forall v \in \mathrm{H}^1_{\tilde{w},0}(0,1), \qquad \text{and, by (5.7)}, \qquad \|z_e\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)} \leq \frac{5}{16} \|e\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)}. \tag{5.9}$$

We shall show that in fact $\mathrm{D}_r^2 z_e \in \mathrm{L}^2_{\tilde{w}}(0,1)$, and thereby $z_e \in \mathrm{H}^2_{\tilde{w},0}(0,1)$. To this end, note that

$$\mathrm{D}_r z_e(r) = -\frac{1}{r} \int_0^r s\, e(s)\, \mathrm{d}s, \qquad r \in (0,1].$$

Hence, $\mathrm{D}_r z_e \in \mathrm{C}(0,1]$ and, on recalling that $e \in \mathrm{L}^2_{\tilde{w}}(0,1)$, the Cauchy-Schwarz inequality yields

$$|\mathrm{D}_r z_e(r)|^2 \le \frac{1}{2}\int_0^r s|e(s)|^2\,\mathrm{d}s, \qquad r \in (0,1]. \tag{5.10}$$

This inequality implies that $\lim_{r\to 0+}\mathrm{D}_r z_e(r) = 0$ and that, for any $\varepsilon \in (0,1)$,

$$\int_\varepsilon^1 \frac{1}{r}\,|\mathrm{D}_r z_e(r)|^2\,\mathrm{d}r \le \frac{1}{2\varepsilon}\int_0^1 s|e(s)|^2\,\mathrm{d}s.$$

Thus, $\sqrt{r}(r^{-1}\mathrm{D}_r z_e) \in \mathrm{L}^2(\varepsilon,1)$; hence, by (5.8), $\sqrt{r}\,\mathrm{D}_r^2 z_e = -\sqrt{r}\,(e + r^{-1}\mathrm{D}_r z_e) \in \mathrm{L}^2(\varepsilon,1)$. Multiplying this equality by $\sqrt{r}\,\mathrm{D}_r^2 z_e$ and integrating over the interval $(\varepsilon,1)$ gives

$$\int_\varepsilon^1 r\,|\mathrm{D}_r^2 z_e(r)|^2\,\mathrm{d}r + \int_\varepsilon^1 \mathrm{D}_r z_e(r)\,\mathrm{D}_r^2 z_e(r)\,\mathrm{d}r = -\int_\varepsilon^1 r\,e(r)\,\mathrm{D}_r^2 z_e(r)\,\mathrm{d}r.$$

Hence, by computing explicitly the second integral on the left-hand side and applying Cauchy's inequality $|\alpha\beta| \le \frac{1}{2}(\alpha^2+\beta^2)$ on the right-hand side, we obtain

$$\int_\varepsilon^1 r\,|\mathrm{D}_r^2 z_e(r)|^2\,\mathrm{d}r + |\mathrm{D}_r z_e(1)|^2 \le \int_\varepsilon^1 r\,|e(r)|^2\,\mathrm{d}r + |\mathrm{D}_r z_e(\varepsilon)|^2.$$

Passing to the limit $\varepsilon \to 0_+$ and omitting the second term on the left-hand side gives that $\mathrm{D}_r^2 z_e \in \mathrm{L}^2_{\tilde{w}}(0,1)$ and

$$\int_0^1 r\,|\mathrm{D}_r^2 z_e(r)|^2\,\mathrm{d}r \le \int_0^1 r\,|e(r)|^2\,\mathrm{d}r.$$

Combining this with our earlier bound from (5.9), we have that $\|z_e\|^2_{\mathrm{H}^2_{\tilde{w}}(0,1)} \le \frac{21}{16}\|e\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)}$.

We are now ready to embark on the analysis of the projection error in the $\mathrm{L}^2_{\tilde{w}}(0,1)$ norm. Recalling that $e = \tilde{g} - P_N^J\tilde{g} \in \mathrm{H}^1_{\tilde{w},0}(0,1)$, we deduce from the weak formulation (5.9), the definition of the orthogonal projector P_N^J, the Cauchy-Schwarz inequality, (5.5) and the $\mathrm{H}^2_{\tilde{w}}(0,1)$ norm bound just derived that

$$\begin{aligned}
\|\tilde{g} - P_N^J\tilde{g}\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} &= (e, \tilde{g} - P_N^J\tilde{g})_{\mathrm{L}^2_{\tilde{w}}(0,1)} = (z_e, \tilde{g} - P_N^J\tilde{g})_{\mathrm{H}^1_{\tilde{w},0}(0,1)} = (\tilde{g} - P_N^J\tilde{g}, z_e - P_N^J z_e)_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \\
&\le \|\tilde{g} - P_N^J\tilde{g}\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)}\|z_e - P_N^J z_e\|_{\mathrm{H}^1_{\tilde{w},0}(0,1)} \le cN^{1-k}\|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}\cdot N^{-1}\|z_e\|_{\mathrm{H}^2_{\tilde{w}}(0,1)} \\
&\le cN^{-k}\|\tilde{g}\|_{\mathrm{H}^k_{\tilde{w}}(0,1)}\|\tilde{g} - P_N^J\tilde{g}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)}, \qquad k \ge 1.
\end{aligned}$$

Dividing the left-most and the right-most term in this chain by $\|\tilde{g} - P_N^J\tilde{g}\|_{\mathrm{L}^2_{\tilde{w}}(0,1)}$ gives (5.6). □

Next, for $\tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$, with decomposition given in (5.4), we define the projection operator $\tilde{\Pi}_N : \tilde{\mathrm{H}}^1_{\tilde{w},0}(R) \to \mathcal{P}_N(R)$ as:

$$(\tilde{\Pi}_N\tilde{g})(r,\theta) := (P^F_{N_\theta}\tilde{P}^J_{N_r}\tilde{g})(r,\theta) = (\tilde{P}^J_{N_r}P^F_{N_\theta}\tilde{g})(r,\theta),$$

where the finite-dimensional space $\mathcal{P}_N(R)$ is defined as

$$\mathcal{P}_N(R) := \mathbb{P}_{N_r,0}(0,1) \oplus (r\mathbb{P}_{N_r,0}(0,1) \otimes \mathbb{S}_{N_\theta,0}(0,2\pi)).$$

The structure of this space reflects the decomposition (5.4). Note that the constant functions have been factored out of the space $\mathbb{S}_{N_\theta}(0,2\pi)$ in the definition of $\mathcal{P}_N(R)$; this is appropriate because, as observed above, $(g_2(r,\cdot),1)_{\mathrm{L}^2(0,2\pi)} = 0$. The lemma below establishes optimal order approximation results for this projector.

Lemma 5.4. *Let $\tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$, with decomposition $\tilde{g}(r,\theta) = \tilde{g}_1(r) + r\tilde{g}_2(r,\theta)$, where $\tilde{g}_1 = \frac{1}{2\pi}(\tilde{g},1)_{\mathrm{L}^2(0,2\pi)} \in \mathrm{H}^1_{\tilde{w},0}(0,1)$, $\tilde{g}_2 \in \mathrm{H}^{0,1}_{\tilde{w}}(R)$, $\tilde{g}_2(1,\cdot) = 0$, and assume, in addition, that $\tilde{g}_2(\cdot,\theta) \in \mathrm{H}^1_{\tilde{w},0}(0,1)$ for a.e. $\theta \in (0,2\pi)$. If $\tilde{g}_1 \in \mathrm{H}^{k+1}_{\tilde{w}}(0,1)$ and $\tilde{g}_2 \in \mathrm{H}^{k+1,0}_{\tilde{w}}(R) \cap \mathrm{H}^{k,1}_{\tilde{w}}(R) \cap \mathrm{H}^{0,l+1}_{\tilde{w}}(R) \cap \mathrm{H}^{1,l}_{\tilde{w}}(R)$ for some $k,l \geq 1$, then*

$$\begin{aligned}\|\tilde{g} - \tilde{\Pi}_N\tilde{g}\|_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)} &\leq C_1 N_r^{-k}\left(\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k+1,0}_{\tilde{w}}(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k,1}_{\tilde{w}}(R)}\right)^{\frac{1}{2}} \\ &\quad + C_2 N_\theta^{-l}\left(\|\tilde{g}_2\|^2_{\mathrm{H}^{0,l+1}_{\tilde{w}}(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{1,l}_{\tilde{w}}(R)}\right)^{\frac{1}{2}}. \end{aligned} \tag{5.11}$$

If $\tilde{g}_1 \in \mathrm{H}^k_{\tilde{w}}(0,1)$ and $\tilde{g}_2 \in \mathrm{H}^{k,0}_{\tilde{w}}(R) \cap \mathrm{H}^{0,l}_{\tilde{w}}(R)$ for some $k,l \geq 1$, then

$$\|\tilde{g} - \tilde{\Pi}_N\tilde{g}\|_{\mathrm{L}^2_{\tilde{w}}(R)} \leq C_1 N_r^{-k}\left(\|\tilde{g}_1\|^2_{\mathrm{H}^k_{\tilde{w}}(0,1)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k,0}_{\tilde{w}}(R)}\right)^{\frac{1}{2}} + C_2 N_\theta^{-l}\|\tilde{g}_2\|_{\mathrm{H}^{0,l}_{\tilde{w}}(R)}. \tag{5.12}$$

Proof. The left-hand side in (5.11) is given by:

$$\begin{aligned}\|\tilde{g} - \tilde{\Pi}_N\tilde{g}\|^2_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)} &= \int_0^1 \tilde{w}(r)\int_0^{2\pi}\left\{(\tilde{g} - \tilde{\Pi}_N\tilde{g})^2 + (\mathrm{D}_r\tilde{g} - \mathrm{D}_r(\tilde{\Pi}_N\tilde{g}))^2\right\}\mathrm{d}\theta\,\mathrm{d}r \\ &\quad + \int_0^1 r^{-1}\int_0^{2\pi}(\mathrm{D}_\theta\tilde{g} - \mathrm{D}_\theta(\tilde{\Pi}_N\tilde{g}))^2\,\mathrm{d}\theta\,\mathrm{d}r =: I + II.\end{aligned}$$

Let us first consider term I; we treat the two terms in the, inner, θ-integral in I separately. First, using the L^2-error bound for Fourier projection, as well as the fact that $\|P^F_{N_\theta}\|_{\mathcal{L}(\mathrm{L}^2_p(0,2\pi),\mathrm{L}^2_p(0,2\pi))} \leq 1$, we obtain

$$\begin{aligned}\|\tilde{g}(r,\cdot) - \tilde{\Pi}_N\tilde{g}(r,\cdot)\|^2_{\mathrm{L}^2(0,2\pi)} &\leq \left(\|\tilde{g}(r,\cdot) - P^F_{N_\theta}\tilde{g}(r,\cdot)\|_{\mathrm{L}^2(0,2\pi)} + \|P^F_{N_\theta}(\tilde{g}(r,\cdot) - \tilde{P}^J_{N_r}\tilde{g}(r,\cdot))\|_{\mathrm{L}^2(0,2\pi)}\right)^2 \\ &\leq \left(C_3 N_\theta^{-l}\|\mathrm{D}^l_\theta\tilde{g}(r,\cdot)\|_{\mathrm{L}^2(0,2\pi)} + \|\tilde{g}(r,\cdot) - \tilde{P}^J_{N_r}\tilde{g}(r,\cdot)\|_{\mathrm{L}^2(0,2\pi)}\right)^2 \\ &\leq 2C_3^2 N_\theta^{-2l}\|\mathrm{D}^l_\theta\tilde{g}_2(r,\cdot)\|^2_{\mathrm{L}^2(0,2\pi)} + 2\|\tilde{g}(r,\cdot) - \tilde{P}^J_{N_r}\tilde{g}(r,\cdot)\|^2_{\mathrm{L}^2(0,2\pi)},\end{aligned}$$

where $\mathrm{D}^l_\theta\tilde{g} = r\mathrm{D}^l_\theta\tilde{g}_2$ and $0 \leq r \leq 1$ have been used in the last line. Similarly,

$$\begin{aligned}\|\mathrm{D}_r\tilde{g}(r,\cdot) - \mathrm{D}_r(\tilde{\Pi}_N\tilde{g}(r,\cdot))\|^2_{\mathrm{L}^2(0,2\pi)} &\leq 4C_3^2 N_\theta^{-2l}\left(\|\mathrm{D}^l_\theta\tilde{g}_2(r,\cdot)\|^2_{\mathrm{L}^2(0,2\pi)} + \|\mathrm{D}_r\mathrm{D}^l_\theta\tilde{g}_2(r,\cdot)\|^2_{\mathrm{L}^2(0,2\pi)}\right) \\ &\quad + 2\|\mathrm{D}_r\tilde{g}(r,\cdot) - \mathrm{D}_r(\tilde{P}^J_{N_r}\tilde{g}(r,\cdot))\|^2_{\mathrm{L}^2(0,2\pi)}.\end{aligned}$$

Therefore,

$$I \leq 6\,C_3^2 N_\theta^{-2l}\int_0^{2\pi}\left(\|\mathrm{D}^l_\theta\tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} + \|\mathrm{D}_r\mathrm{D}^l_\theta\tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)}\right)\mathrm{d}\theta + 2\int_0^{2\pi}\|\tilde{g}(\cdot,\theta) - \tilde{P}^J_{N_r}\tilde{g}(\cdot,\theta)\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)}\mathrm{d}\theta.$$

Now we bound the final term on the right-hand side of the last inequality using the univariate bound (5.5):

$$\begin{aligned}\|\tilde{g}(\cdot,\theta) - \tilde{P}^J_{N_r}\tilde{g}(\cdot,\theta)\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)} &\leq 2\|\tilde{g}_1 - P^J_{N_r}\tilde{g}_1\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)} + 2\|r(\tilde{g}_2(\cdot,\theta) - P^J_{N_r}\tilde{g}_2(\cdot,\theta))\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)} \\ &\leq C^2 N_r^{-2k}\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} \\ &\quad + 2\int_0^1 \tilde{w}(r)\left\{(2+r^2)(\tilde{g}_2(r,\theta) - P^J_{N_r}\tilde{g}_2(r,\theta))^2 + 2r^2(\mathrm{D}_r(\tilde{g}_2(r,\theta) - P^J_{N_r}\tilde{g}_2(r,\theta)))^2\right\}\mathrm{d}r \\ &\leq C^2 N_r^{-2k}\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + 6\|\tilde{g}_2(\cdot,\theta) - P^J_{N_r}\tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)} \\ &\leq C_4^2 N_r^{-2k}\left(\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + \|\tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)}\right).\end{aligned}$$

Therefore,

$$\begin{aligned} I &\leq 6\,C_3^2 N_\theta^{-2l} \int_0^{2\pi} \left(\|\mathrm{D}_\theta^l \tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} + \|\mathrm{D}_r \mathrm{D}_\theta^l \tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} \right) \mathrm{d}\theta \\ &\quad + 2\,C_4^2 N_r^{-2k} \int_0^{2\pi} \left(\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + \|\tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} \right) \mathrm{d}\theta, \end{aligned} \tag{5.13}$$

which is an optimal-order bound on I.

Next we consider II. Since θ-differentiation commutes with the projectors $P^J_{N_r}$ and $P^F_{N_\theta}$, we have

$$\begin{aligned} II &\leq 2 \int_0^1 r^{-1} \int_0^{2\pi} |\mathrm{D}_\theta \tilde{g}(r,\theta) - P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}(r,\theta)|^2 \mathrm{d}\theta\, \mathrm{d}r \\ &\quad + 2 \int_0^1 r^{-1} \int_0^{2\pi} |P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}(r,\theta) - \tilde{P}^J_{N_r}(P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}(r,\theta))|^2 \mathrm{d}\theta\, \mathrm{d}r. \end{aligned}$$

Therefore,

$$\begin{aligned} II &\leq 2 \int_0^1 r^{-1} \int_0^{2\pi} \left| r\mathrm{D}_\theta \tilde{g}_2(r,\theta) - r P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta) \right|^2 \mathrm{d}\theta\, \mathrm{d}r \\ &\quad + 2 \int_0^{2\pi} \int_0^1 r^{-1} |r P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta) - \tilde{P}^J_{N_r}(r P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta))|^2 \,\mathrm{d}r\, \mathrm{d}\theta \\ &\leq C_5^2 N_\theta^{-2l} \int_0^1 \tilde{w}(r) \int_0^{2\pi} |\mathrm{D}_\theta^{l+1} \tilde{g}_2(r,\theta)|^2 \,\mathrm{d}\theta\, \mathrm{d}r \\ &\quad + 2 \int_0^{2\pi} \int_0^1 \tilde{w}(r) |P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta) - P^J_{N_r}(P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta))|^2 \,\mathrm{d}r\, \mathrm{d}\theta \\ &\leq C_5^2 N_\theta^{-2l} \int_0^{2\pi} \|\mathrm{D}_\theta^{l+1} \tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} \,\mathrm{d}\theta + C_6^2 N_r^{-2k} \int_0^{2\pi} \|P^F_{N_\theta} \mathrm{D}_\theta \tilde{g}_2(r,\theta)\|^2_{\mathrm{H}^k_{\tilde{w}}(0,1)} \,\mathrm{d}\theta. \end{aligned}$$

We have used the fact that $\tilde{P}^J_{N_r}(r\tilde{g}_2) = r P^J_{N_r}(\tilde{g}_2)$ as well as the $\mathrm{L}^2_w(0,1)$ norm error bound for $P^J_{N_r}$ stated in (5.6). For the second integral in the last line in the bound on II we have

$$\sum_{j=0}^{k} \int_0^1 \tilde{w}(r) \|P^F_{N_\theta} \mathrm{D}_r^j \mathrm{D}_\theta \tilde{g}_2(\cdot,r)\|^2_{\mathrm{L}^2(0,2\pi)} \mathrm{d}r \leq \sum_{j=0}^{k} \int_0^1 \tilde{w}(r) \|\mathrm{D}_r^j \mathrm{D}_\theta \tilde{g}_2(\cdot,r)\|^2_{\mathrm{L}^2(0,2\pi)} \mathrm{d}r.$$

Therefore,

$$II \leq C_5^2 N_\theta^{-2l} \int_0^{2\pi} \|\mathrm{D}_\theta^{l+1} \tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} \,\mathrm{d}\theta + C_6^2 N_r^{-2k} \int_0^{2\pi} \|\mathrm{D}_\theta \tilde{g}_2(\cdot,\theta)\|^2_{\mathrm{H}^k_{\tilde{w}}(0,1)} \mathrm{d}\theta.$$

Combining the bounds for I and II with suitable constants C_1 and C_2, we obtain

$$\begin{aligned} \|\tilde{g} - P^F_{N_\theta} \tilde{P}^J_{N_r} \tilde{g}\|_{\tilde{\mathrm{H}}^1_{\tilde{w}}(R)} &\leq C_1 N_r^{-k} \left\{ \int_0^{2\pi} (\|\tilde{g}_1\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k+1}_{\tilde{w}}(0,1)} + \|\mathrm{D}_\theta \tilde{g}_2\|^2_{\mathrm{H}^k_{\tilde{w}}(0,1)}) \,\mathrm{d}\theta \right\}^{\frac{1}{2}} \\ &\quad + C_2 N_\theta^{-l} \left\{ \int_0^{2\pi} (\|\mathrm{D}_\theta^{l+1} \tilde{g}_2\|^2_{\mathrm{L}^2_{\tilde{w}}(0,1)} + \|\mathrm{D}_\theta^l \tilde{g}_2\|^2_{\mathrm{H}^1_{\tilde{w}}(0,1)}) \,\mathrm{d}\theta \right\}^{\frac{1}{2}}, \end{aligned} \tag{5.14}$$

which is (5.11). The proof of the $\mathrm{L}^2_{\tilde{w}}(R)$ norm bound (5.12) is very similar: its main ingredients are, in fact, contained in the argument above. For the sake of brevity we omit the details. □

The bounds (5.11) and (5.12) can now be straightforwardly mapped from R to D_0. We define $\mathcal{P}_N(D)$ as $\mathcal{P}_N(R)$ mapped from R to D_0 using the polar coordinate transformation (5.1), and we suppose that $\hat{\psi} \in \mathcal{H}^{k+1,l+1}(D)$, with $k, l \geq 1$, where

$$\begin{aligned}\mathcal{H}^{k,l}(D) \quad := \quad & \{g \in \mathrm{H}^1_0(D) \,:\, \tilde{g} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R) \text{ has a decomposition } \tilde{g}(r,\theta) = \tilde{g}_1(r) + r\tilde{g}_2(r,\theta), \\ & \text{with } \tilde{g}_1 = \tfrac{1}{2\pi}(\tilde{g},1)_{\mathrm{L}^2(0,2\pi)} \in \mathrm{H}^k_{\tilde{w},0}(0,1) \text{ and } \tilde{g}_2 \in \mathrm{H}^{k,0}_{\tilde{w},0}(R) \cap \mathrm{H}^{k-1,1}_{\tilde{w}}(R) \cap \mathrm{H}^{0,l}_{\tilde{w}}(R) \cap \mathrm{H}^{1,l-1}_{\tilde{w}}(R)\},\end{aligned}$$

equipped with the norm $\|g\|_{\mathcal{H}^{k,l}(D)} := \left(\|g\|^2_{\mathcal{H}^k_r(D)} + \|g\|^2_{\mathcal{H}^l_\theta(D)}\right)^{\frac{1}{2}}$ where, for $\tilde{g} = \tilde{g}_1 + r\tilde{g}_2 \in \mathcal{H}^{k,l}(D)$,

$$\|g\|_{\mathcal{H}^k_r(D)} := \left(\|\tilde{g}_1\|^2_{\mathrm{H}^k_{\tilde{w}}(0,1)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k,0}_{\tilde{w}}(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k-1,1}_{\tilde{w}}(R)}\right)^{\frac{1}{2}} \quad \text{and} \quad \|g\|_{\mathcal{H}^l_\theta(D)} := \left(\|\tilde{g}_2\|^2_{\mathrm{H}^{0,l}_{\tilde{w}}(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{1,l-1}_{\tilde{w}}(R)}\right)^{\frac{1}{2}}.$$

We define

$$\hat{\Pi}_N : \mathcal{H}^{1,1}(D) \to \mathcal{P}_N(D) \quad \text{by} \quad (\hat{\Pi}_N g)(q_1,q_2) = (\tilde{\Pi}_N \tilde{g})(r,\theta), \qquad g \in \mathcal{H}^{1,1}(D).$$

Thus, recalling (2.5) and noting that $\mathcal{H}^{k,l}(D) \subset \mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D;M)$, $k,l \geq 1$, we deduce from (5.11) that

$$\|\hat{\psi} - \hat{\Pi}_N\hat{\psi}\|_{\mathrm{H}^1_0(D;M)} \leq C_1 N_r^{-k} \|\hat{\psi}\|_{\mathcal{H}^{k+1}_r(D)} \; + \; C_2 N_\theta^{-l} \|\hat{\psi}\|_{\mathcal{H}^{l+1}_\theta(D)} \tag{5.15}$$

for all $\hat{\psi} \in \mathcal{H}^{k+1,l+1}(D)$, with $k,l \geq 1$. Similarly, we obtain from (5.12) that

$$\|\hat{\psi} - \hat{\Pi}_N\hat{\psi}\|_{\mathrm{L}^2(D)} \leq C_1 N_r^{-k} \|\hat{\psi}\|_{\mathcal{H}^k_r(D)} \; + \; C_2 N_\theta^{-l} \|\hat{\psi}\|_{\mathcal{H}^l_\theta(D)} \tag{5.16}$$

for all $\hat{\psi} \in \mathcal{H}^{k,l}(D)$, with $k,l \geq 1$.

6. Convergence analysis of the numerical method

In this section we complete the convergence analysis of the fully-discrete numerical method (4.1), (4.2), based on the symmetrized form of the Fokker-Planck equation. At the end of the section we shall comment on the extension of our results to a fully-discrete method that stems from the alternative semidiscretization (3.26) in the case of the FENE model.

We see from (4.7) that in order to obtain bounds on the norms of ξ appearing on the left-hand side of (4.7) we need to bound the following terms:

$$\|\eta^0\|, \quad \|\eta\|_{\ell^2(0,T;\mathrm{H}^1_0(D;M))} \quad \text{and} \quad \left\|\frac{\partial \eta}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}.$$

It follows from (5.15), (5.16) and the definition of $\eta := \hat{\psi} - \hat{\Pi}_N\hat{\psi}$ that

$$\|\eta^0\| \leq \|\hat{\psi}_0 - \hat{\Pi}_N\hat{\psi}_0\| \leq C_1 N_r^{-k} \|\hat{\psi}_0\|_{\mathcal{H}^k_r(D)} \; + \; C_2 N_\theta^{-l} \|\hat{\psi}_0\|_{\mathcal{H}^l_\theta(D)},$$

$$\|\eta\|_{\ell^2(0,T;\mathrm{H}^1_0(D;M))} \leq C_1 N_r^{-k} \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}^{k+1}_r(D))} \; + \; C_2 N_\theta^{-l} \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}^{l+1}_\theta(D))},$$

$$\left\|\frac{\partial \eta}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))} \leq C_1 N_r^{-k} \left\|\frac{\partial \hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathcal{H}^k_r(D))} + \; C_2 N_\theta^{-l} \left\|\frac{\partial \hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathcal{H}^l_\theta(D))},$$

with $k, l \geq 1$, provided that $\hat{\psi}$ is such that the right-hand sides of these inequalities are finite. Substituting these three bounds into the right-hand side of (4.7) we deduce, with $m\Delta t \leq T$, $m = 0, 1, \ldots, N_T$, that

$$\|\xi\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} + \|\underset{\sim}{\nabla}_M \xi\|_{\ell^2(0,T;\mathrm{L}^2(D))} \leq C_1 N_r^{-k} \left(\|\hat{\psi}_0\|_{\mathcal{H}_r^k(D)} + \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_r^{k+1}(D))} + \left\| \frac{\partial \hat{\psi}}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_r^k(D))} \right)$$
$$+ C_2 N_\theta^{-l} \left(\|\hat{\psi}_0\|_{\mathcal{H}_\theta^l(D)} + \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_\theta^{l+1}(D))} + \left\| \frac{\partial \hat{\psi}}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_\theta^l(D))} \right) + C_3 \Delta t \left\| \frac{\partial^2 \hat{\psi}}{\partial t^2} \right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}. \quad (6.1)$$

Note, also, that

$$\|\eta\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} \leq C_1 N_r^{-k} \|\hat{\psi}\|_{\ell^\infty(0,T;\mathcal{H}_r^k(D))} + C_2 N_\theta^{-l} \|\hat{\psi}\|_{\ell^\infty(0,T;\mathcal{H}_\theta^l(D))}, \quad (6.2)$$
$$\|\underset{\sim}{\nabla}_M \eta\|_{\ell^2(0,T;\mathrm{L}^2(D))} \leq C_1 N_r^{-k} \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_r^{k+1}(D))} + C_2 N_\theta^{-l} \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_\theta^{l+1}(D))}. \quad (6.3)$$

Now, by the triangle inequality,

$$\|\hat{\psi} - \hat{\psi}_N\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} + \|\underset{\sim}{\nabla}_M(\hat{\psi} - \hat{\psi}_N)\|_{\ell^2(0,T;\mathrm{L}^2(D))} \leq$$
$$\|\xi\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} + \|\underset{\sim}{\nabla}_M \xi\|_{\ell^2(0,T;\mathrm{L}^2(D))} + \|\eta\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} + \|\underset{\sim}{\nabla}_M \eta\|_{\ell^2(0,T;\mathrm{L}^2(D))},$$

whereby (6.1), (6.2) and (6.3) give

$$\|\hat{\psi} - \hat{\psi}_N\|_{\ell^\infty(0,T;\mathrm{L}^2(D))} + \|\underset{\sim}{\nabla}_M(\hat{\psi} - \hat{\psi}_N)\|_{\ell^2(0,T;\mathrm{L}^2(D))} \leq$$
$$C_1 N_r^{-k} \left(\|\hat{\psi}\|_{\ell^\infty(0,T;\mathcal{H}_r^k(D))} + \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_r^{k+1}(D))} + \left\| \frac{\partial \hat{\psi}}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_r^k(D))} \right)$$
$$+ C_2 N_\theta^{-l} \left(\|\hat{\psi}\|_{\ell^\infty(0,T;\mathcal{H}_\theta^l(D))} + \|\hat{\psi}\|_{\ell^2(0,T;\mathcal{H}_\theta^{l+1}(D))} + \left\| \frac{\partial \hat{\psi}}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_\theta^l(D))} \right) + C_3 \Delta t \left\| \frac{\partial^2 \hat{\psi}}{\partial t^2} \right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}.$$

We recall that $\psi = \sqrt{M}\hat{\psi}$, and we define $\psi_N^n := \sqrt{M}\hat{\psi}_N^n$. Consequently,

$$\|\psi - \psi_N\|_{\ell^\infty(0,T;\mathfrak{H})} + \|\psi - \psi_N\|_{\ell^2(0,T;\mathfrak{K})} \leq$$
$$C_1 N_r^{-k} \left(\left\| \frac{\psi}{\sqrt{M}} \right\|_{\ell^\infty(0,T;\mathcal{H}_r^k(D))} + \left\| \frac{\psi}{\sqrt{M}} \right\|_{\ell^2(0,T;\mathcal{H}_r^{k+1}(D))} + \left\| \frac{1}{\sqrt{M}} \frac{\partial \psi}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_r^k(D))} \right)$$
$$+ C_2 N_\theta^{-l} \left(\left\| \frac{\psi}{\sqrt{M}} \right\|_{\ell^\infty(0,T;\mathcal{H}_\theta^l(D))} + \left\| \frac{\psi}{\sqrt{M}} \right\|_{\ell^2(0,T;\mathcal{H}_\theta^{l+1}(D))} + \left\| \frac{1}{\sqrt{M}} \frac{\partial \psi}{\partial t} \right\|_{\mathrm{L}^2(0,T;\mathcal{H}_\theta^l(D))} \right)$$
$$+ C_3 \Delta t \left\| \frac{1}{\sqrt{M}} \frac{\partial^2 \psi}{\partial t^2} \right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(D))}, \quad (6.4)$$

with $k, l \geq 1$, provided that ψ is such that right-hand side is finite.

That completes the convergence analysis of the method in the case of $d = 2$. For $d = 3$ the argument is identical, and rests on a three-dimensional analogue of Lemma 5.2. We omit the details.

Starting from the second stability inequality stated in Lemma 3.6 and proceeding in an identical manner as above, one can derive analogous error bounds in the $h^1(0, T; \mathfrak{H})$ and $\ell^\infty(0, T; \mathfrak{K})$ norms.

Remark 6.1. In the case of the FENE Maxwellian, $\sqrt{M} \in \mathcal{P}_N(D)$ if, and only if, there exists a positive integer m such that $b = 4m$ and $N_r \geq 2m$. In order to ensure that, more generally, $\sqrt{M} \in \mathcal{P}_N(D)$ regardless of the specific choice of b and the value of N_r, one can simply enrich $\mathcal{P}_N(D)$ by adding $\sqrt{M}$ as an extra basis function. However, in general the polynomials in $\mathcal{P}_N(D)$ approximate $\sqrt{M}$ very closely, so this leads to a highly ill-conditioned basis. A better solution is to add the component of $\sqrt{M}$ orthogonal to $\mathcal{P}_N(D)$ (in the $\mathrm{L}^2(D)$ inner product, for example) to the basis, rather than $\sqrt{M}$ itself. This is implemented in Section 7 for a numerical example in which b is not divisible by 4 and is shown to work well in that case.

Remark 6.2. We make a second comment regarding the FENE model. Starting from the variant of the inequality (4.7) alluded to in Remark 3.7 in connection with the fully-discrete spectral method based on the semidiscretization (3.26) with $b \geq 4s^2/(2s-1)$ and $s > 1/2$, one can derive an optimal-order error bound analogous to (6.4). The core of the argument is identical to the one above, and is therefore omitted.

7. Implementation of the numerical method

Numerical methods for solving the Fokker-Planck equation arising from the FENE dumbbell model for dilute polymeric fluids have been the focus of some attention recently; Du *et al.* [20] developed a finite difference scheme that preserved the unit integral property and the positivity of ψ, Chauvière and Lozinski [17,18,27] developed a spectral method for this problem and Ammar *et al.* [1,2] proposed a reduced-basis method for solving the Fokker-Planck equation for FENE dumbbell chains. For a survey of, alternative, stochastic techniques for the numerical simulation of polymeric liquids we refer to the monograph of Öttinger [30] and the article of Jourdain *et al.* [23], for example.

In this section we discuss the implementation of two spectral Galerkin methods based on the formulation (4.1), (4.2). We also present computational results, both to verify the convergence results derived in previous sections and to demonstrate the effectiveness of these numerical methods in practice. Finally, we compare the two spectral Galerkin methods based on the formulation (4.1), (4.2) with the method of Chauvière and Lozinski based on the "original" form (1.6) of the Fokker-Planck equation (or, more precisely, its transformed version (3.24) resulting from the substitution (7.9), with $s = 2$).

Following Section 5 we restrict our attention to the case $d = 2$ and suppose that $\hat{\psi} \in \mathrm{H}^1_0(D)$. Hence, $\tilde{\psi} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$, where $\tilde{\psi}(r,\theta) := \hat{\psi}(q_1,q_2)$ with $q_1 = \sqrt{b}\, r\cos\theta$, $q_2 = \sqrt{b}\, r\sin\theta$. Using the decomposition (5.4), $\tilde{\psi}$ can be written in polar co-ordinates as follows:

$$\tilde{\psi}(r,\theta) = \tilde{\psi}_1(r) + r\tilde{\psi}_2(r,\theta), \qquad (r,\theta) \in R = (0,1)\times(0,2\pi), \tag{7.1}$$

where, as in Section 5, r has been scaled from $(0,\sqrt{b})$ to $(0,1)$, and $\tilde{\psi}_1 := \frac{1}{2\pi}(\tilde{\psi},1)_{\mathrm{L}^2(0,2\pi)}$. In the context of spectral methods in polar co-ordinates, (7.1) is referred to by Shen as the *essential pole condition* [31]. This condition is a "first-order" form of the following full pole-condition [21]: in order that a function

$$\tilde{\psi}(r,\theta) = \sum_{n\in\mathbb{Z}} \tilde{\gamma}_n(r)\tilde{E}_n(\theta), \qquad \text{where} \quad \tilde{E}_n(\theta) := \frac{1}{\sqrt{2\pi}}\exp(\mathrm{i}n\theta),$$

is infinitely differentiable when transformed from polar to Cartesian co-ordinates, it is necessary that, for each $n \in \mathbb{Z}\setminus\{0\}$,

$$\tilde{\gamma}_n(r) = \mathcal{O}(r^{|n|}) \qquad \text{as } r \to 0_+. \tag{7.2}$$

That (7.1) is a "first-order" form of the full pole condition is easily seen by writing $\tilde{\gamma}_n(r) = r^{|n|}\tilde{G}_n(r)$, with $\tilde{G}_n(r) = \mathcal{O}(1)$ as $r \to 0_+$; hence,

$$\tilde{\psi}(r,\theta) = \frac{1}{\sqrt{2\pi}}\,\tilde{\gamma}_0(r) + r\sum_{n\in\mathbb{Z}\setminus\{0\}}^{\infty} r^{|n|-1}\tilde{G}_n(r)\tilde{E}_n(\theta) =: \tilde{\psi}_1(r) + r\tilde{\psi}_2(r,\theta),$$

with $\tilde{\psi}_1(r) = \tilde{\gamma}_0(r)/\sqrt{2\pi} = \frac{1}{2\pi}(\tilde{\psi},1)_{\mathrm{L}^2(0,2\pi)}$, as required.

The full pole condition (7.2) is consistent with the result established in the proof of Lemma 5.2 stating that the expansion coefficients $\tilde{\gamma}_n$, $n \in \mathbb{Z} \setminus \{0\}$, of a function in $\tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$ satisfy $\tilde{\gamma}_n(r) = o(1)$ as $r \to 0_+$, although the conditions (7.2) are clearly much more restrictive.

In order to fit into the framework of the numerical analysis in Sections 5 and 6, each element of $\mathcal{P}_N(R)$ should satisfy (7.1) to ensure that $\mathcal{P}_N(D)$ is contained in $\mathrm{H}^1_0(D)$. The discrete space $\mathcal{P}_N(R)$, introduced in Section 5, satisfies this property. In this section we define a spectral Galerkin method for the Fokker-Planck equation based on a particular basis (denoted $\mathcal{A}$) for $\mathcal{P}_N(R)$ that satisfies the same decomposition.

For the purpose of comparison, we also introduce a second basis, $\mathcal{B}$, in which each function satisfies the full pole condition, (7.2). Thus, on mapping $\mathcal{B}$ from R to D we obtain a basis for a finite-dimensional subspace of $\mathrm{C}^\infty(\overline{D}) \cap \mathrm{C}_0(\overline{D}) \subset \mathrm{H}^1_0(D)$. The reason for considering this second basis is that typical solutions of the FENE Fokker-Planck equation are smooth on D, and therefore it is likely that in practice a Galerkin method based on $\mathcal{B}$ will be more accurate than a method based on $\mathcal{A}$: mapping the basis $\mathcal{A}$ from R to D yields a finite-dimensional subspace of $\mathrm{H}^1_0(D)$ only, which contains functions that are not smooth at the origin in D. We note, however, that the span of $\mathcal{B}$ does not coincide with $\mathcal{P}_N(R)$, and therefore the approximation properties of $\mathcal{B}$ are not covered by the results in Section 5 that led to the error bounds in Section 6. Hence, the numerical results for basis $\mathcal{A}$ are intended to verify the analysis developed in the previous sections, while basis $\mathcal{B}$ is introduced to indicate the gain in performance that can be obtained by satisfying (7.2). By requiring more regularity from the basis than it being a finite-dimensional subspace of $\mathrm{H}^1_0(D)$ one could modify the arguments in Section 5 to derive convergence estimates based on a pole condition of higher order than (5.4), but this would make the derivation of the approximation results more laborious (*e.g.*, the projector $\tilde{P}^J_N$ would have to obey (7.2) rather than (7.1)). Before introducing bases $\mathcal{A}$ and $\mathcal{B}$, we make the following observation.

Remark 7.1. Let $\hat{\psi}$ be the weak solution of (1.8) corresponding to a given initial condition $\hat{\psi}_0$, define $\hat{\psi}^*(\underset{\sim}{q}, t) := \hat{\psi}(-\underset{\sim}{q}, t)$ and suppose that $\hat{\psi}_0$ is invariant under the change of independent variable $\underset{\sim}{q} \mapsto -\underset{\sim}{q}$, *i.e.*, $\hat{\psi}_0(\underset{\sim}{q}) = \hat{\psi}_0(-\underset{\sim}{q})$ for a.e. $\underset{\sim}{q} \in D$. On noting that $M(\underset{\sim}{q}) = M(-\underset{\sim}{q})$, $\underset{\sim}{q} \in D$, it follows that the weak formulation (1.8) is also invariant under this change of variable; hence $\hat{\psi}$ and $\hat{\psi}^*$ are weak solutions to the same initial boundary-value problem. It follows by uniqueness of the weak solution established in Section 3 that $\hat{\psi}(\underset{\sim}{q}, t) \equiv \hat{\psi}^*(\underset{\sim}{q}, t)$, *i.e.*, $\hat{\psi}(\underset{\sim}{q}, t) = \hat{\psi}(-\underset{\sim}{q}, t)$ for a.e. $\underset{\sim}{q} \in D$ and a.e. $t \in [0, T]$. This evenness of $\hat{\psi}$ in the D domain with respect to $\underset{\sim}{q}$ translates into π-periodicity of $\tilde{\psi}$ in the R domain with respect to θ. An identical statement applies to the numerical solution $(\hat{\psi}^n_N)_{n=0}^{N_T}$ defined by (4.1), (4.2), provided $\mathcal{P}_N(D) \subset \mathrm{H}^1_0(D)$ is such that whenever a function $\underset{\sim}{q} \mapsto v(\underset{\sim}{q})$ belongs to $\mathcal{P}_N(D)$ its even reflection $\underset{\sim}{q} \mapsto v(-\underset{\sim}{q})$ also belongs to $\mathcal{P}_N(D)$: if $\hat{\psi}_0(\underset{\sim}{q}) = \hat{\psi}_0(-\underset{\sim}{q})$ for a.e. $\underset{\sim}{q} \in D$, uniqueness of the $\mathrm{L}^2(D)$ projection of $\hat{\psi}^0$ onto $\mathcal{P}_N(D)$ implies that $\hat{\psi}^0_N(\underset{\sim}{q}) = \hat{\psi}^0_N(-q)$ for a.e. $\underset{\sim}{q} \in D$. Uniqueness of the numerical solution then yields $\hat{\psi}^n_N(\underset{\sim}{q}) = \hat{\psi}^n_N(-\underset{\sim}{q})$ for a.e. $\underset{\sim}{q} \in D$ and all $n = 0, \ldots, N_T$.

The above remark demonstrates that (1.8) captures an important symmetry property of the dumbbell model for polymeric fluids: the configuration probability density function ψ is required to be symmetric about the origin in D because the beads of a dumbbell are indistinguishable. As long as $\hat{\psi}_0$ and $\mathcal{P}_N(D)$ are invariant under the change of independent variable $\underset{\sim}{q} \mapsto -\underset{\sim}{q}$ described in Remark 7.1, the numerical solution will inherit the symmetry of the analytical solution implied by the symmetry of the initial condition. A consequence of this observation is that we should require the basis functions in $\mathcal{A}$ and $\mathcal{B}$ to obey the same symmetry condition; following [18], this is achieved in the definitions below by only including even trigonometric modes in θ. Strictly speaking therefore $\mathcal{A}$ is chosen to be a basis for the linear subspace of $\mathcal{P}_N(R)$ consisting of all π-periodic functions. Note, however, that if the solution were 2π-periodic, then one could simply include odd trigonometric modes as well. We are now ready to define the bases $\mathcal{A}$ and $\mathcal{B}$.

Basis $\mathcal{A}$. Let $\mathcal{A} := \mathcal{A}_1 \cup \mathcal{A}_2$ where:

$$\mathcal{A}_1 := \{(1-r)P_k(r) : k = 0, \ldots, N_r - 1\},$$
$$\mathcal{A}_2 := \{r(1-r)P_k(r)\Phi_{il}(\theta) : k = 0, \ldots, N_r - 1; \quad i = 0, 1; \quad l = 1, \ldots, N_\theta\}.$$

P_k is a polynomial of degree k in $r \in [0,1]$ and $\Phi_{il}(\theta) = (1-i)\cos(2l\theta) + i\sin(2l\theta)$, $\theta \in [0,\pi]$. We denote by P_k the kth Chebyshev polynomial scaled from $[-1,1]$ to $[0,1]$. The numerical method is not particularly sensitive to this choice of polynomial, however, and other choices work well also. Notice that the polynomials in $\mathcal{A}_1$ and $\mathcal{A}_2$ both contain the factor $(1-r)$ in order to impose the homogeneous Dirichlet boundary condition on ∂D, and functions in $\mathcal{A}_2$ contain an extra factor of r to enforce the essential pole condition. Basis $\mathcal{A}$ is chosen so as to mimic the decomposition (7.1) of the analytical solution $\tilde{\psi} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$ in polar co-ordinates: the role of $\mathrm{span}(\mathcal{A}_1)$ is to approximate $\tilde{\psi}_1$ while $\mathrm{span}(\mathcal{A}_2)$ is meant to approximate $r\tilde{\psi}_2$.

Basis $\mathcal{B}$. This is, effectively, the basis proposed by Matsushima and Marcus [29] and Verkley [34], except that, as above, we ensure that the functions are zero at $r=1$ and that they are π-periodic in θ:

$$\mathcal{B} = \{W_{lk}(r)\Phi_{il}(\theta) : k = 0,\dots,N_r-1; \quad i = 0,1; \quad l = i,\dots,N_\theta\}, \tag{7.3}$$

where $W_{lk}(r) = r^{2l}(1-r^2)J_k^{(0,2l)}(2r^2-1)$ and $J_k^{(\alpha,\beta)}(x)$ is the Jacobi polynomial on $[-1,1]$ of degree k with respect to the weight $(1-x)^\alpha(1+x)^\beta$ (Φ_{il} is the same as in $\mathcal{A}$). Each element of $\mathcal{B}$ satisfies (7.2).

$\mathcal{A}$ and $\mathcal{B}$ both have cardinality $N := N_r\,(2N_\theta+1)$. Expressing trial and test functions in terms of either $\mathcal{A}$ or $\mathcal{B}$, it is now straightforward to determine the discretization matrices corresponding to the integrals

$$\int_D \hat{\psi}_N^{n+1}\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}, \qquad \int_D \underset{\sim}{\nabla}_M\hat{\psi}_N^{n+1}\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q}, \qquad \int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}_N^{n+1})\cdot\underset{\sim}{\nabla}_M\hat{\varphi}\,\mathrm{d}\underset{\sim}{q} \tag{7.4}$$

from (4.1). These matrices are labeled $\mathbf{M}$, $\mathbf{S}$ and $\mathbf{C}^{n+1}$ for mass, stiffness and convection respectively.

Using the ansatz $\tilde{\psi}_N^{n+1}(r,\theta) = \sum_{v=1}^N \tilde{\Psi}_v^{n+1}X_v(r,\theta)$ for trial functions, where X_v is a basis function (from either $\mathcal{A}$ or $\mathcal{B}$) for $1 \le v \le N$, denoting test functions as X_u for $1 \le u \le N$ and mapping (7.4) from D to R yields:

$$\mathbf{M}_{uv} = \int_0^1\int_0^\pi b\,r\,X_v(r,\theta)X_u(r,\theta)\,\mathrm{d}r\,\mathrm{d}\theta, \tag{7.5}$$

$$\mathbf{S}_{uv} = \int_0^1\int_0^\pi \left\{ r\,\frac{\partial X_v}{\partial r}\frac{\partial X_u}{\partial r} + \frac{1}{r}\frac{\partial X_v}{\partial\theta}\frac{\partial X_u}{\partial\theta} + \frac{b}{2}\frac{r^2}{1-r^2}\frac{\partial}{\partial r}(X_uX_v) + \frac{b^2}{4}\frac{r^3}{(1-r^2)^2}X_vX_u \right\}\mathrm{d}r\,\mathrm{d}\theta, \tag{7.6}$$

$$\begin{aligned}\mathbf{C}^{n+1}_{uv} &= \int_0^1\int_0^\pi br\,X_v\frac{\partial X_u}{\partial\theta}\,(-\kappa_{11}^{n+1}\sin 2\theta - \kappa_{12}^{n+1}\sin^2\theta + \kappa_{21}\cos^2\theta)\,\mathrm{d}r\,\mathrm{d}\theta \\ &\quad + \int_0^1\int_0^\pi \left(b\,r^2\,X_v\frac{\partial X_u}{\partial r} + \frac{b^2}{2}\frac{r^3}{1-r^2}X_vX_u\right)\left(\kappa_{11}^{n+1}\cos 2\theta + \frac{1}{2}(\kappa_{12}^{n+1}+\kappa_{21}^{n+1})\sin 2\theta\right)\mathrm{d}r\,\mathrm{d}\theta.\end{aligned} \tag{7.7}$$

Note that if the X_u, X_v do not satisfy (7.1), then the entries of $\mathbf{S}$ may be undefined.

With these discretization matrices in hand, the numerical solution is computed by solving the following linear system for the coefficient vector $\tilde{\mathbf{\Psi}}^{n+1} := (\tilde{\Psi}_1^{n+1},\dots,\tilde{\Psi}_N^{n+1})^{\mathrm{T}} \in \mathbb{R}^N$, $n = 0,1,\dots,N_T-1$:

$$\left(\mathbf{M} + \Delta t\left(\frac{1}{2\lambda}\mathbf{S} - \mathbf{C}^{n+1}\right)\right)\tilde{\mathbf{\Psi}}^{n+1} = \mathbf{M}\tilde{\mathbf{\Psi}}^n, \tag{7.8}$$

with $\tilde{\Psi}^0$ defined by the initial datum. Then, the numerical approximation to the probability density function itself is obtained as $\psi_N^{n+1}(\underset{\sim}{q}) = \sqrt{M(\underset{\sim}{q})}\,\tilde{\psi}_N^{n+1}(r,\theta)$, where $r = |\underset{\sim}{q}|/\sqrt{b}$ and $\tilde{\psi}_N^{n+1}(r,\theta) = \sum_{v=1}^N \tilde{\Psi}_v^{n+1}X_v(r,\theta)$.

For ease of evaluation, the integrals in (7.5), (7.6) and (7.7) can be factorized into products of 1-dimensional integrals over r and θ. We evaluate the θ-integrals exactly using trigonometric identities, and, noting that the r-integrands are all polynomials, we use Gauss quadrature to evaluate the r-integrals to machine precision. $\mathbf{M}$ and $\mathbf{S}$ are constant matrices, which can be pre-computed and reused, but if $\underset{\approx}{\kappa}$ is time-varying, we must reassemble $\mathbf{C}^{n+1}$ at every time-step. However, it is straightforward to factor out the dependence of $\mathbf{C}^{n+1}$ on $\underset{\approx}{\kappa}$

so that the integrals that determine $\mathbf{C}^{n+1}$ need not be evaluated more than once. We use LU-decomposition to solve (7.8), which is appropriate because the spectral discretization matrices are generally of moderate size.

We now present some numerical results. For simplicity, in the computations considered below we always use the normalized Maxwellian (which satisfies the symmetry property required in Remark 7.1 and also has unit volume) as the initial condition, so that $\hat{\psi}_0(\underset{\sim}{q}) = \sqrt{M(\underset{\sim}{q})}$. Also, most of the results presented in this section are for computations in which b was chosen to be divisible by 4 so that the spaces span($\mathcal{A}$) and span($\mathcal{B}$) naturally contain $\sqrt{M}$, as in Remark 6.1. However, the basis enrichment technique described in Remark 6.1 was implemented to obtain the results in Table 3 (in which $b = 10$) and, as discussed below, it worked well for that problem.

Henceforth, the two numerical methods that use basis $\mathcal{A}$ and basis $\mathcal{B}$, respectively, will be referred to as method $\mathcal{A}$ and method $\mathcal{B}$.

First of all we present results from solving the Fokker-Planck equation with parameters $b = 16$, $\lambda = 1.2$ and $\kappa_{11} = -\kappa_{22} = 1.1$, $\kappa_{12} = 0.9$, $\kappa_{21} = -0.6$ and with $\Delta t = 0.05$. These parameters were chosen somewhat arbitrarily, but the intention here is to visualize a typical evolution of ψ_N towards steady state, and to provide an initial qualitative comparison of methods $\mathcal{A}$ and $\mathcal{B}$ (quantitative convergence results will be presented below). By taking $(N_r, N_\theta) = (26, 20)$ with basis $\mathcal{A}$ and $(N_r, N_\theta) = (21, 15)$ with basis $\mathcal{B}$, the solutions from the two methods were indistinguishable to the eye and appear to be fully resolved. As foreshadowed above, $\mathcal{A}$ required more degrees-of-freedom than $\mathcal{B}$ to resolve the solution to comparable accuracy in this case because, as can be seen in Figure 1, ψ_N is smooth at the origin in Cartesian co-ordinates whereas the basis functions in $\mathcal{A}$ are not necessarily smooth there. Nevertheless, a clear advantage of basis $\mathcal{A}$ over basis $\mathcal{B}$ is that it is built by relying on the essential pole condition only, as manifested by the decomposition in Lemma 5.2, which only requires the most basic smoothness hypothesis, that $\tilde{\psi} \in \tilde{\mathrm{H}}^1_{\tilde{w},0}(R)$ (implied by the assumption that the weak solution $\hat{\psi} \in \mathrm{H}^1_0(D; M)$ belongs to $\mathrm{H}^1_0(D)$). A related important observation is that, as long as $\hat{\psi} \in \mathrm{H}^1_0(D)$, the error bound (6.4), the definition of $\mathcal{H}^{k,l}(D)$ and the convergence rate delivered by basis $\mathcal{A}$ depend on the smoothness of $\tilde{\psi}$ on R, not on the smoothness of $\hat{\psi}$ on D (see also [31], p. 1585); this feature can be advantageous: for example, the error bound (6.4), resulting from the Cartesian product structure of R, indicates how potential anisotropic smoothness of $\hat{\psi}$ in the radial and azimuthal directions can be exploited by admitting different, unrelated, polynomial degrees N_r and N_θ in the radial and azimuthal directions, respectively.

Figure 1 shows snapshots of ψ_N at $t = 0$, $t = 1$, $t = 2$ and $t = 3$, and ψ_N is close to steady state at $t = 3$.

To provide a quantitative study of the spatial accuracy of the numerical methods defined in this section, we use the fact that when $\underset{\approx}{\kappa}$ is a symmetric tensor the exact steady-state solution of the Fokker-Planck equation is given by $\psi_{\mathrm{exact}}(\underset{\sim}{q}) := CM(\underset{\sim}{q}) \exp(\lambda \underset{\sim}{q}^{\mathrm{T}} \underset{\approx}{\kappa} \underset{\sim}{q})$ where C is a normalization constant chosen so that $\int_D \psi_{\mathrm{exact}}(\underset{\sim}{q}) \mathrm{d}\underset{\sim}{q} = 1$; see [13]. We now consider a particular case, referred to as *extensional flow*, in which $\underset{\approx}{\kappa} = \mathrm{diag}(\delta, -\delta)$. This generally provides a good test case for numerical methods for the Fokker-Planck equation because it yields particularly sharp solution profiles that are challenging to resolve, and also the exact steady-state solution is available for comparison. In order to compare the convergence rates of methods $\mathcal{A}$ and $\mathcal{B}$, we solved two distinct extensional flow problems for: (i) $(b, \lambda, \delta) = (12, 1, 1)$ and (ii) $(b, \lambda, \delta) = (20, 1, 2)$, with a range of choices of (N_r, N_θ). In order to compare to the known exact steady-state solution, we took 2000 time-steps (with $\Delta t = 0.05$ and $T = 100$) in each case so that the final numerical solution is a very close approximation to the steady-state solution. This allows us to compare the spatial convergence rates of the two numerical methods without worrying about temporal discretization error. Tables 1 and 2 show the relative errors (in the $\mathrm{L}^2(D)$ and $\mathrm{H}^1(D; M)$ norms) between the exact and the computed steady-state solutions for extensional flows (i) and (ii), respectively.

We can see from the data in the tables that methods $\mathcal{A}$ and $\mathcal{B}$ converge rapidly for both problem (i) and problem (ii) and that for each choice of (N_r, N_θ), basis $\mathcal{B}$ outperforms basis $\mathcal{A}$ – again this is because the solution profiles are smooth at the origin in Cartesian co-ordinates, see Figure 2. Nevertheless, the rapid convergence of method $\mathcal{A}$ is consistent with the spectral error estimates established in Section 6 (recall that these error estimates do not apply to method $\mathcal{B}$ because span($\mathcal{B}$) is not the same as $\mathcal{P}_N(R)$ analyzed in Sect. 5). It is also

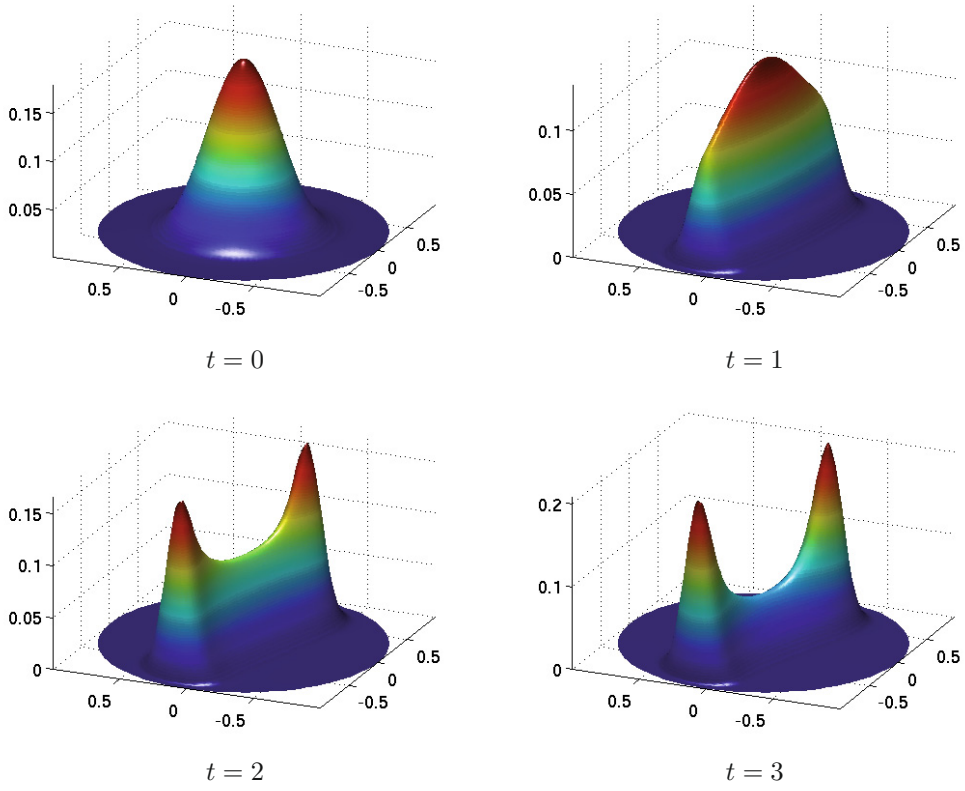


FIGURE 1. Snapshots of ψ_N at $t = 0$, $t = 1$, $t = 2$ and $t = 3$ illustrating evolution towards steady state. In this case, we have $\Delta t = 0.05$, $b = 16$, $\lambda = 1.2$ and $\kappa_{11} = -\kappa_{22} = 1.1$, $\kappa_{12} = 0.9$, $\kappa_{21} = -0.6$. This computation was performed using basis $\mathcal{A}$ and basis $\mathcal{B}$ with $(N_r, N_\theta) = (26, 20)$ and $(N_r, N_\theta) = (21, 15)$, respectively. The solutions were fully resolved in each of these two cases.

clear that problem (ii) is more challenging to resolve than problem (i); with both $\mathcal{A}$ and $\mathcal{B}$, more basis functions are required to attain a given accuracy for problem (ii) than for problem (i). Note that the greater difficulty of resolving extensional flow (ii) is encoded in the convergence estimates in Section 6 because the constants in these estimates depend exponentially on b, δ (*via* $\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)}$) and T. Moreover, the factor $\mathrm{e}^{2c_0 m\Delta t}$ on the right-hand side in Lemma 3.1 permits exponential growth in time of the norm of $\hat{\psi}_N$, and this is reflected in the first row of Table 2 in which the solutions computed with $(N_r, N_\theta) = (10, 10)$ for extensional flow (ii) resulted in numerical overflow[2]. Note that this overflow behaviour was only observed in the case of under-resolved computations that led to numerical solutions containing numerical oscillations *i.e.* it was not observed in rows 2, 3 and 4 of Table 2; note also that Chauvière and Lozinski's method behaves in the same way for under-resolved solutions, as shown in Table 3.

The (fully resolved) solutions corresponding to extensional flow problems (i) and (ii) are shown in Figure 2, and in each case both ψ_N and $\tilde{\psi}_N$ are plotted. It is clear that the solution profiles corresponding to (ii) are much more severe, and therefore it is not surprising that more modes were required in this case. The quantity of interest in these computations is ψ_N, but $\tilde{\psi}_N$ is also plotted to emphasize the numerical difficulties that are

[2]When $\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}(t)\underset{\sim}{q} = 0$ for all $t \in [0, T]$, Lemma 3.1, with $\mu = 0$ and $\underset{\sim}{\nu} = \underset{\sim}{0}$, can be sharpened. The inequality holds with $c_0 = 0$, showing that the expression on the left-hand side of the inequality is bounded by $\|\hat{\psi}^0\|^2$, uniformly in T, b and $\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T)}$.

TABLE 1. Relative errors in the $\mathrm{L}^2(D)$ and $\mathrm{H}^1(D;M)$ norms (*i.e.* $\|\hat{\psi}_N - \hat{\psi}_{\text{exact}}\|/\|\hat{\psi}_{\text{exact}}\|$ and $\|\hat{\psi}_N - \hat{\psi}_{\text{exact}}\|_{\mathrm{H}^1(D;M)}/\|\hat{\psi}_{\text{exact}}\|_{\mathrm{H}^1(D;M)}$, respectively) for extensional flow (i) at steady-state, *i.e.* $b = 12$, $\lambda = 1$ and $\delta = 1$. $\hat{\psi}_N$ is an approximation to the steady-state solution obtained by taking 2000 time-steps with $\Delta t = 0.05$, and $\hat{\psi}_{\text{exact}}$ is the exact steady-state solution which is known in this case because $\underset{\approx}{\kappa}$ is symmetric.

	Relative $\mathrm{L}^2(D)$ error		Relative $\mathrm{H}^1(D;M)$ error	
(N_r, N_θ)	Basis $\mathcal{A}$	Basis $\mathcal{B}$	Basis $\mathcal{A}$	Basis $\mathcal{B}$
(10,10)	3.63×10^{-2}	4.61×10^{-3}	7.90×10^{-2}	8.82×10^{-3}
(15,15)	3.36×10^{-3}	9.19×10^{-6}	8.58×10^{-3}	2.33×10^{-5}
(20,20)	5.13×10^{-5}	4.63×10^{-9}	1.64×10^{-4}	1.52×10^{-8}
(25,25)	2.94×10^{-7}	1.74×10^{-12}	1.13×10^{-6}	6.94×10^{-12}
(30,30)	8.31×10^{-10}	1.70×10^{-13}	3.77×10^{-9}	1.70×10^{-13}

TABLE 2. Relative errors in the $\mathrm{L}^2(D)$ and $\mathrm{H}^1(D;M)$ norms for extensional flow (ii) at steady-state, *i.e.* $(b, \lambda, \delta) = (20, 1, 2)$. The time-stepping strategy to compute the approximate steady-state solution was the same as in Table 1. The hyphens in the first row indicate that we obtained numerical overflow in those computations.

	Relative $\mathrm{L}^2(D)$ error		Relative $\mathrm{H}^1(D;M)$ error	
(N_r, N_θ)	Basis $\mathcal{A}$	Basis $\mathcal{B}$	Basis $\mathcal{A}$	Basis $\mathcal{B}$
(10,10)	–	–	–	–
(20,20)	3.91×10^{-2}	1.72×10^{-3}	4.88×10^{-2}	2.54×10^{-3}
(30,30)	1.50×10^{-3}	2.97×10^{-6}	2.61×10^{-3}	4.49×10^{-6}
(40,40)	2.54×10^{-5}	5.97×10^{-9}	4.55×10^{-5}	5.94×10^{-9}

encountered as b and δ are increased. In the plots corresponding to (i), the peaks in $\tilde{\psi}_N$ are higher than in ψ_N, but only by a factor of about 20. For (ii) on the other hand, the peaks in $\tilde{\psi}_N$ are higher by a factor of roughly 5000. The causes of this behaviour are two-fold: with $\delta = 2$ the flow has stronger extensional character and therefore the solution peaks are expected to be more concentrated and also, the larger value of b means that $\sqrt{M}$ is more strongly degenerate near ∂D so that $\hat{\psi}_N = \psi_N/\sqrt{M}$ takes larger values near the boundary. This second point can be seen as a drawback, for $b \gg 1$, of the fully-discrete numerical method (4.1), (4.2), based on the symmetrized form of the Fokker-Planck equation. Presumably Chauvière and Lozinski [18] fixed their value of s ($s = 2$ for $d = 2$ and $s = 2.5$ for $d = 3$) in the transformation

$$\hat{\psi}(\underset{\sim}{q}) := \psi(\underset{\sim}{q})/[M(\underset{\sim}{q})]^{2s/b} = \psi(\underset{\sim}{q})/(1 - |\underset{\sim}{q}|^2/b)^s \tag{7.9}$$

so as to avoid a similar effect; indeed, they presented some numerical results for $b = 200$. Values of b this large do not appear to be feasible with the fully-discrete method (4.1), (4.2), based on the substitution $\hat{\psi}_N = \psi_N/\sqrt{M}$.

As has been noted in Remark 4.2, there is in fact no difference between the stability properties of the method based on (4.1), (4.2) and of a Chauvière and Lozinski type method. However, if $b \gg 1$, for a typical ψ we have that $\|\psi/\sqrt{M}\|_{\mathrm{L}^\infty(D)} = \|\psi/(1 - |\underset{\sim}{q}|^2/b)^{b/4}\|_{\mathrm{L}^\infty(D)} \gg \|\psi/(1 - |\underset{\sim}{q}|^2/b)^2\|_{\mathrm{L}^\infty(D)}$. Hence, compared to a Chauvière and Lozinski type method with the recommended choice of $s = 2$ for $d = 2$, the maximum value of the numerical approximation $\hat{\psi}_N$ to the function $\hat{\psi}$ defined by the scheme (4.1), (4.2) can be much larger when $b \gg 1$, and can thereby require greater computational effort to resolve to a given accuracy. The computational results that we consider in this section are therefore restricted to moderate values of b.

With these precursors, we now compare the accuracy of methods $\mathcal{A}$ and $\mathcal{B}$ to that of the spectral method of Chauvière and Lozinski discussed in [18]. In Table 2 of that paper, the authors presented convergence data for

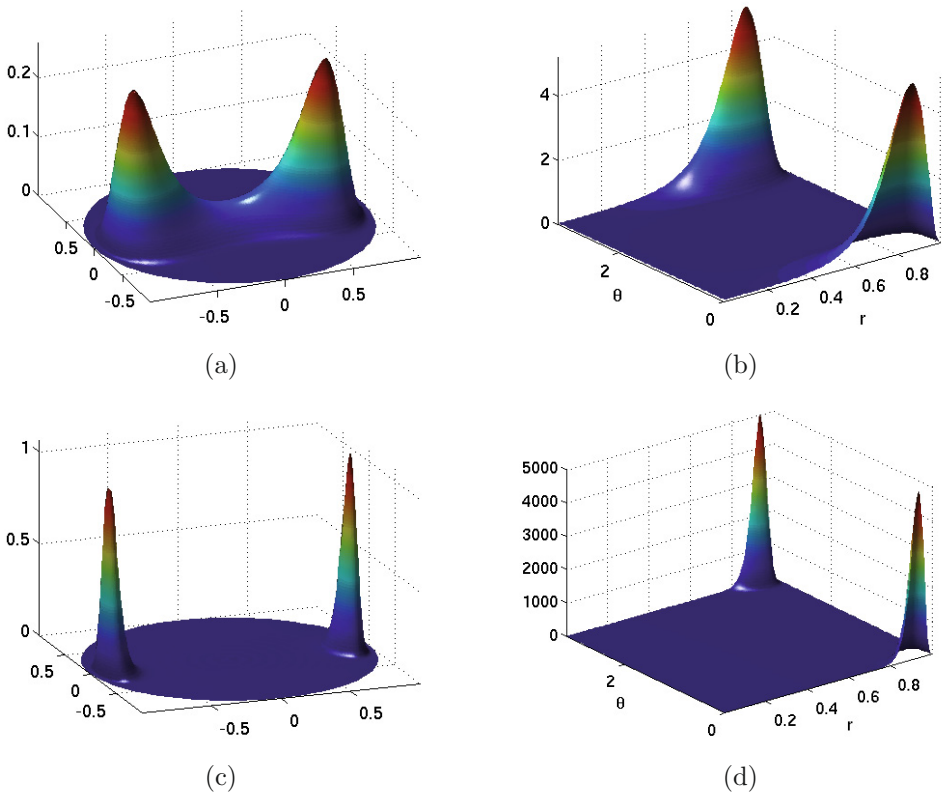


FIGURE 2. Numerical approximations to the steady state solution for extensional flow problems (i) and (ii) using $(N_r, N_\theta) = (30, 30)$ and $(N_r, N_\theta) = (40, 40)$, respectively. Plots (a) and (b) show ψ_N and $\tilde{\psi}_N$ respectively, at steady state for problem (i) and (c), (d) show ψ_N and $\tilde{\psi}_N$ for (ii). The purpose of plots (b) and (d) is to demonstrate that $\tilde{\psi}_N$ usually has a much steeper solution profile than ψ_N and this effect is amplified if either δ or b (or both) are increased.

the $(1,1)$-component of the *polymeric extra-stress* tensor, $\underset{\approx}{\tau} = (\tau_{ij})$[3], computed for an extensional flow at steady state for the parameters $(b, \lambda, \delta) = (10, 1, 5)$. The tensor $\underset{\approx}{\tau}$ is defined as follows:

$$\underset{\approx}{\tau}(t) = \int_D \underset{\sim}{F}(\underset{\sim}{q}) \otimes \underset{\sim}{q} \psi(\underset{\sim}{q}, t)\, \mathrm{d}\underset{\sim}{q}, \tag{7.10}$$

where $\underset{\sim}{F}$ is the FENE spring force. Table 3 reproduces Chauvière and Lozinski's results and compares them to the corresponding results for methods $\mathcal{A}$ and $\mathcal{B}$. Note that in this problem b is not divisible by 4. Therefore, in order to ensure that the volume of ψ_N is conserved with methods $\mathcal{A}$ and $\mathcal{B}$, we added the component of $\sqrt{M}$ orthogonal to span($\mathcal{A}$) (resp. span($\mathcal{B}$)) to the bases to obtain an enriched discrete space that contains $\sqrt{M}$ (*cf.* Rem. 6.1)[4]. This ensured that the volume of ψ_N was conserved to machine precision (except in the cases that rounding error polluted the results, these are indicated by hyphens in the table).

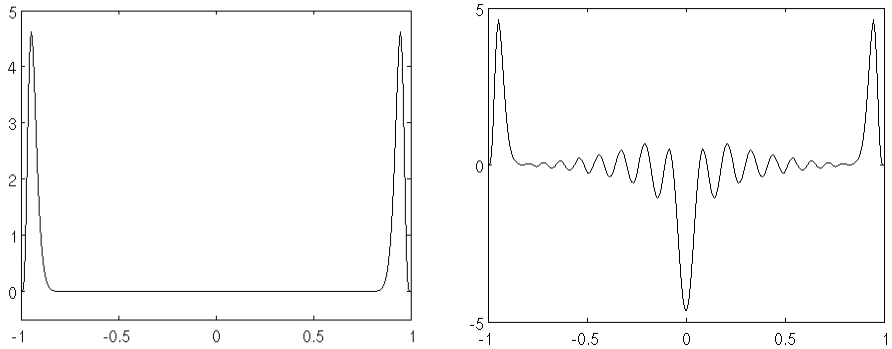
The data in Table 3 show that for this problem method $\mathcal{B}$ converges at a comparable rate to the method of Chauvière and Lozinski, whereas $\mathcal{A}$ appears to converge more slowly. Note that the reason why method $\mathcal{B}$ and

[3]In the context of polymeric fluids, $\underset{\approx}{\tau}$ represents the contribution of the polymer molecules to the macroscopic stress field.

[4]Orthogonalization was performed in the $\mathrm{L}^2(D)$ inner product.

TABLE 3. Comparison of the relative errors in τ_{11} for extensional flow with $(b, \lambda, \delta) = (10, 1, 5)$. The three schemes compared are methods $\mathcal{A}$ and $\mathcal{B}$ and the spectral method of Chauvière and Lozinski. The data for the method of Chauvière and Lozinski is taken from Table 2 in [18].

	Relative error of τ_{11}		
(N_r, N_θ)	Basis $\mathcal{A}$	Basis $\mathcal{B}$	Chauvière and Lozinski
(11,5)	–	–	–
(13,6)	–	4.8×10^{-2}	0.35
(21,10)	1.8×10^{-3}	2.0×10^{-2}	2.0×10^{-2}
(31,15)	2.1×10^{-4}	1.4×10^{-4}	1.4×10^{-4}
(41,20)	1.3×10^{-5}	8.7×10^{-7}	2.1×10^{-7}



(a) (b)

FIGURE 3. Cross-sections of the solution of the extensional flow problem with $b = 12$, $\lambda = 1$ and $\delta = 5$ at steady state, obtained using method $\mathcal{B}$. The fully-resolved solution in (a) was obtained using $(N_r, N_\theta) = (41, 20)$, and the under-resolved solution in (b) was obtained with $(N_r, N_\theta) = (26, 20)$.

Chauvière and Lozinski's method converge at a similar rate (at least in this case where b is relatively low) is that both methods involve ansatzes that impose extra regularity at the origin in Cartesian co-ordinates; basis $\mathcal{B}$ satisfies the pole condition (7.2), and Chauvière and Lozinski use a transformation that enforces $\frac{\partial \psi}{\partial r}\Big|_{r=0} = 0$, which, when combined with π-periodicity in θ, has a similar effect.

Remark 7.2. It was proved in Lemma 3.3 that the weak solution of the initial-boundary-value problem (1.3), (1.4), (1.5) is non-negative a.e. on D. This property is not guaranteed to hold for the numerical solution. However, our numerical experiments consistently show that if there are sufficiently many modes in the approximation space to accurately resolve the solution then this non-negativity property is preserved under discretization. This is illustrated in Figure 3 in which two cross-sections of the numerical solution for the $(b, \lambda, \delta) = (12, 1, 5)$ extensional flow are shown: the numerical solution on the left is fully resolved, while the one on the right is under-resolved. In the under-resolved case there are oscillations and clearly $\psi_N \geq 0$ is not satisfied throughout D, whereas the non-negativity property is accurately captured in the fully resolved case.

8. CONCLUSIONS

The Fokker-Planck equation (1.1) has been the subject of active research recently, as a component of bead-spring type Navier-Stokes-Fokker-Planck models for dilute polymeric fluids. We focused our attention on Fokker-Planck equations with unbounded drift, such as those that arise from modelling polymer molecules as FENE dumbbells, where the spring potential $\underset{\sim}{q} \in D \mapsto U(\underset{\sim}{q}) \in \mathbb{R}_{\geq 0}$ appearing in the Fokker-Planck equation tends

to $+\infty$ as $\underset{\sim}{q}$ approaches ∂D, where D is a ball in $\mathbb{R}^d$. The purpose of this paper has been to develop a rigorous foundation for the numerical approximation of such Fokker-Planck equations. We symmetrized the principal part of the differential operator by introducing the Maxwellian M associated with U, and applied the transformation $\hat{\psi} = \psi/\sqrt{M}$. The resulting weak formulation (1.8) facilitated the development of a number of analytical results in Sections 3 and 4, including existence and uniqueness of weak solutions of the semidiscretized equation (3.1) and, on passing to the limit $\Delta t \to 0_+$, of (1.8) also. Using the approximation results derived in Section 5, optimal-order convergence of the fully-discrete spectral Galerkin method (4.1), (4.2) was established for the case of $d = 2$; an analogous procedure could be carried out for $d = 3$. This analysis was performed for spring potentials that satisfy Hypotheses A and B; see Example 1.1. The FENE potential is a special case of this family and also satisfies a third structural hypothesis, Hypothesis C; for such potentials further results can be deduced *via* the Brascamp-Lieb inequality (*cf.* Sect. 2). For example, by virtue of (2.3), not only does the method converge in the $\mathrm{L}^2(D)$ norm but also in the norm of the weighted factor space $\mathrm{L}^2_{M^{-2/b}}(D)/\mathrm{Ker}(\underset{\sim}{\nabla}_M)$.

In the case of the FENE model we indicated the extension of our analysis to a class of numerical methods based on another change of variable, proposed by Chauvière and Lozinski; here, instead of a Kolmogorov-type symmetrization, a different transformation, (7.9), is applied to the Fokker-Planck equation. We showed that, at the analytical level at least, the two approaches lead to methods with very similar stability and accuracy properties. Section 7 addressed issues related to the implementation of numerical methods for the FENE Fokker-Planck equation. Numerical results were presented for two distinct implementations, methods $\mathcal{A}$ and $\mathcal{B}$, and these methods were also compared to the spectral method discussed in the paper of Chauvière and Lozinski [18] on the basis of numerical results reported therein. We showed that methods $\mathcal{A}$ and $\mathcal{B}$ work well for values of b up to about 20, and are comparable to the method formulated in [18] in terms of computational efficiency in this parameter range, with method $\mathcal{B}$ being more accurate than method $\mathcal{A}$, and of a very similar accuracy as the method in [18].

A spectral method is natural in the context of this problem because the boundary of the domain D is smooth and D can be easily transformed into a rectangular domain R. One could, however, also conceive of a finite-element method directly in Cartesian co-ordinates, without mapping D to R, and in this case much of the analysis of this paper would carry over. By choosing a finite-element space $\mathcal{P}_N(D) \subset \mathrm{H}^1_0(D)(\subset \mathrm{H}^1_0(D;M))$ and recalling (2.5) and the approximation results of Bernardi on d-dimensional exact triangulations of D (*cf.* Thm. 5.1 in [8]), one could easily deduce optimal error bounds from (4.7) (as well as its analogue based on a Chauvière and Lozinski type transformation). We note, however, that in order to guarantee the unit-volume property by selecting $\mathcal{P}_N(D)$ so that $\sqrt{M} \in \mathcal{P}_N(D)$ (*cf.* Rem. 4.1), in addition to choosing b to be a multiple of 4 as in the spectral methods above, one would now need to work with *piecewise* polynomials of degree $b/2$ at least.

The goal of our future work is to extend the numerical methods and analytical results herein to the coupled Navier-Stokes-Fokker-Planck model, building on the recent paper [6] where convergence to weak solutions of coupled Navier-Stokes-Fokker-Planck systems has been shown for a general class of Galerkin schemes (without convergence rates) in the special case when the velocity field $\underset{\sim}{u}$ is corotational (*i.e.*, $\underset{\sim}{q}^{\mathrm{T}}\underset{\approx}{\kappa}\underset{\sim}{q} = 0$, with $\underset{\approx}{\kappa} = \underset{\approx}{\nabla}_x \underset{\sim}{u}$).

References

[1] A. Ammar, B. Mokdad, F. Chinesta and R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *J. Non-Newtonian Fluid Mech.* **139** (2006) 153–176.

[2] A. Ammar, B. Mokdad, F. Chinesta and R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids. Part II: Transient simulation using space-time separated representations. *J. Non-Newtonian Fluid Mech.* **144** (2007) 98–121.

[3] F.G. Avkhadiev and K.-J. Wirths, Unified Poincaré and Hardy inequalities with sharp constants for convex domains. *ZAMM Z. Angew. Math. Mech.* **87** (2007) 632–642.

[4] J.W. Barrett and E. Süli, Existence of global weak solutions to kinetic models of dilute polymers. *Multiscale Model. Simul.* **6** (2007) 506–546.

[5] J.W. Barrett and E. Süli, Existence of global weak solutions to dumbbell models for dilute polymers with microscopic cut-off. *Math. Mod. Meth. Appl. Sci.* **18** (2008) 935–971.

[6] J.W. Barrett and E. Süli, Numerical approximation of corotational dumbbell models for dilute polymers. *IMA J. Numer. Anal.* (2008) online. Available at http://imajna.oxfordjournals.org/cgi/content/abstract/drn022.
[7] J.W. Barrett, C. Schwab and E. Süli, Existence of global weak solutions for some polymeric flow models. *Math. Mod. Meth. Appl. Sci.* **15** (2005) 939–983.
[8] C. Bernardi, Optimal finite-element interpolation on curved domains. *SIAM J. Numer. Anal.* **26** (1989) 1212–1240.
[9] C. Bernardi and Y. Maday, Spectral methods, in *Handbook of Numerical Analysis* **V**, P. Ciarlet and J. Lions Eds., Elsevier (1997).
[10] O.V. Besov and A. Kufner, The density of smooth functions in weight spaces. *Czechoslova. Math. J.* **18** (1968) 178–188.
[11] O.V. Besov, J. Kadlec and A. Kufner, Certain properties of weight classes. *Dokl. Akad. Nauk SSSR* **171** (1966) 514–516.
[12] R.B. Bird, C.F. Curtiss, R.C. Armstrong and O. Hassager, *Dynamics of Polymeric Liquids, Vol.* **1**, *Fluid Mechanics*. Second edition, John Wiley and Sons (1987).
[13] R.B. Bird, C.F. Curtiss, R.C. Armstrong and O. Hassager, *Dynamics of Polymeric Liquids, Vol.* **2**, *Kinetic Theory*. Second edition, John Wiley and Sons (1987).
[14] S. Bobkov and M. Ledoux, From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.* **10** (2000) 1028–1052.
[15] C. Canuto, A. Quarteroni, M.Y. Hussaini and T.A. Zang, *Spectral Methods: Fundamentals in Single Domains*. Springer (2006).
[16] S. Cerrai, *Second Order PDE's in Finite and Infinite Dimensions, A Probabilistic Approach*, *Lecture Notes in Mathematics* **1762**. Springer (2001).
[17] C. Chauvière and A. Lozinski, Simulation of complex viscoelastic flows using Fokker-Planck equation: 3D FENE model. *J. Non-Newtonian Fluid Mech.* **122** (2004) 201–214.
[18] C. Chauvière and A. Lozinski, Simulation of dilute polymer solutions using a Fokker-Planck equation. *Comput. Fluids* **33** (2004) 687–696.
[19] G. Da Prato and A. Lunardi, On a class of elliptic operators with unbounded coefficients in convex domains. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* **15** (2004) 315–326.
[20] Q. Du, C. Liu and P. Yu, FENE dumbbell model and its several linear and nonlinear closure approximations. *Multiscale Model. Simul.* **4** (2005) 709–731.
[21] H. Eisen, W. Heinrichs and K. Witsch, Spectral collocation methods and polar coordinate singularities. *J. Comput. Phys.* **96** (1991) 241–257.
[22] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*. Springer (1973).
[23] B. Jourdain, T. Lelièvre and C. Le Bris, Numerical analysis of micro-macro simulations of polymeric fluid flows: A simple case. *Math. Mod. Meth. Appl. Sci.* **12** (2002) 1205–1243.
[24] A.N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104** (1931).
[25] A. Kufner, *Weighted Sobolev Spaces*, *Teubner-Texte zur Mathematik*. Teubner (1980).
[26] T. Li and P.-W. Zhang, Mathematical analysis of multi-scale models of complex fluids. *Commun. Math. Sci.* **5** (2007) 1–51.
[27] A. Lozinski and C. Chauvière, A fast solver for Fokker-Planck equation applied to viscoelastic flows calculation: 2D FENE model. *J. Comput. Phys.* **189** (2003) 607–625.
[28] M. Marcus, V.J. Mizel and Y. Pinchover, On the best constant for Hardy's inequality in $\mathbf{R}^n$. *Trans. Amer. Math. Soc.* **350** (1998) 3237–3255.
[29] T. Matsushima and P.S. Marcus, A spectral method for polar coordinates. *J. Comput. Phys.* **120** (1995) 365–374.
[30] H.C. Öttinger, *Stochastic Processes in Polymeric Fluids*. Springer (1996).
[31] J. Shen, Efficient spectral Galerkin methods III: Polar and cylindrical geometries. *SIAM J. Sci. Comput.* **18** (1997) 1583–1604.
[32] R. Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*. Third edition, North-Holland, Amsterdam (1984).
[33] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*. Second edition, Johan Ambrosius Barth, Heidelberg (1995).
[34] W.T.M. Verkley, A spectral model for two-dimensional incompressible fluid flow in a circular basin I. Mathematical formulation. *J. Comput. Phys.* **136** (1997) 100–114.