ESAIM: M2AN 43 (2009) 1117–1156
DOI: 10.1051/m2an/2009034

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

A HETEROGENEOUS ALTERNATING-DIRECTION METHOD FOR A MICRO-MACRO DILUTE POLYMERIC FLUID MODEL

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

Abstract. We examine a heterogeneous alternating-direction method for the approximate solution of the FENE Fokker–Planck equation from polymer fluid dynamics and we use this method to solve a coupled (macro-micro) Navier–Stokes–Fokker–Planck system for dilute polymeric fluids. In this context the Fokker–Planck equation is posed on a high-dimensional domain and is therefore challenging from a computational point of view. The heterogeneous alternating-direction scheme combines a spectral Galerkin method for the Fokker–Planck equation in *configuration space* with a finite element method in *physical space* to obtain a scheme for the high-dimensional Fokker–Planck equation. Alternating-direction methods have been considered previously in the literature for this problem (*e.g.* in the work of Lozinski, Chauvière and collaborators [*J. Non-Newtonian Fluid Mech.* **122** (2004) 201–214; *Comput. Fluids* **33** (2004) 687–696; *CRM Proc. Lect. Notes* **41** (2007) 73–89; Ph.D. Thesis (2003); *J. Computat. Phys.* **189** (2003) 607–625]), but this approach has not previously been subject to rigorous numerical analysis. The numerical methods we develop are fully-practical, and we present a range of numerical results demonstrating their accuracy and efficiency. We also examine an advantageous superconvergence property related to the polymeric extra-stress tensor. The heterogeneous alternating-direction method is well suited to implementation on a parallel computer, and we exploit this fact to make large-scale computations feasible.

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

Received October 23, 2008. Revised March 17, 2009.
Published online August 1st, 2009.

1. Introduction

Simulating dilute polymeric fluids by solving a Navier–Stokes–Fokker–Planck system of partial differential equations (also known as a *micro-macro model*) is well known to be a challenging problem in computational rheology. It is worth highlighting at the outset that there is an extensive literature on numerical methods for this problem, but most of the previous work uses either a fully macroscopic approach in order to circumvent the multiscale nature of the Navier–Stokes–Fokker–Planck system (see the text [31] for an overview of this field) or a stochastic approach in which the micro-macro system is treated using Monte-Carlo-type methods (*cf.* [30]). Our approach is rather different; we focus on using deterministic numerical methods (specifically, finite element and spectral methods) to directly solve this multiscale model. We refer to this as the *deterministic multiscale approach*. This approach was used successfully by Chauvière and Lozinski [10,11,27], however, those authors

Keywords and phrases. Multiscale modelling, kinetic models, dilute polymers, alternating-direction methods, spectral methods, finite element methods, high-performance computing.

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

Article published by EDP Sciences

did not focus on analysis of their numerical methods. The aim of the current work, therefore, is to present rigorous numerical analysis of fully-practical numerical methods for simulating dilute polymeric fluids and to demonstrate the effectiveness of these numerical methods in practice.

The starting point for mathematical modelling of polymeric fluids is generally to represent the fluid by a simplified approximation; in the present case, by a suspension of noninteracting dumbbells (two masses connected by a spring with force law $\underset{\sim}{F}$) in a Newtonian solvent. The spring force law has a corresponding potential, $U : \mathbb{R}_{\geq 0} \to \mathbb{R}$, such that $\underset{\sim}{F}(\underset{\sim}{q}) = U'(\frac{1}{2}|\underset{\sim}{q}|^2)\underset{\sim}{q}$, where $\underset{\sim}{q} \in D$ is the *configuration* vector (or end-to-end vector) of a dumbbell. In this paper, we consider the FENE force law [36], which in nondimensional form is:

$$U(\tfrac{1}{2}|\underset{\sim}{q}|^2) := -\frac{b}{2}\ln\left(1 - \frac{|\underset{\sim}{q}|^2}{b}\right), \qquad \underset{\sim}{F}(\underset{\sim}{q}) = \frac{\underset{\sim}{q}}{1 - |\underset{\sim}{q}|^2/b}, \tag{1.1}$$

where $D = B(0, \sqrt{b}) \subset \mathbb{R}^d$, $d = 2$ or 3. The dimensionless parameter b is typically in the range $[10, 100]$. In [19], Jourdain *et al.* showed that for the stochastic differential equation modelling a suspension of FENE dumbbells (which corresponds to the deterministic Fokker–Planck-based model considered here), the solution exists and has trajectorial uniqueness if, and only if, $b > 2$ (*cf.* also Ex. 1.2 in [4]). Hence, throughout the rest of this paper, we assume that $b \in (2, \infty)$ for the FENE potential[2].

Suppose the fluid is confined to a macroscopic physical domain Ω, assumed to be a bounded open set in $\mathbb{R}^d$. Let $\underset{\sim}{u} : (\underset{\sim}{x}, t) \in \Omega \times [0, T] \mapsto \underset{\sim}{u}(\underset{\sim}{x}, t) \in \mathbb{R}^d$ denote the macroscopic velocity field, and let $p : (\underset{\sim}{x}, t) \in \Omega \times [0, T] \mapsto p(\underset{\sim}{x}, t) \in \mathbb{R}$ denote the pressure. It is typical in this problem to let $\underset{\approx}{\kappa}$ denote the macroscopic velocity gradient, *i.e.* $\underset{\approx}{\kappa} := \underset{\sim}{\nabla}_x \underset{\sim}{u}$. Also, suppose the function $(\underset{\sim}{x}, \underset{\sim}{q}, t) \mapsto \psi(\underset{\sim}{x}, \underset{\sim}{q}, t)$ represents the probability, at time t, of finding a dumbbell with center of mass at $\underset{\sim}{x}$ (for a.e. $\underset{\sim}{x} \in \Omega$) and configuration vector in the element $\underset{\sim}{q} + \mathrm{d}\underset{\sim}{q}$. Then, the micro-macro model for a suspension of FENE dumbbells, in nondimensional form, is as follows (this nondimensionalisation was also considered in [24]):

$$\frac{\partial \underset{\sim}{u}}{\partial t} + \underset{\sim}{u} \cdot \underset{\sim}{\nabla}_x \underset{\sim}{u} + \underset{\sim}{\nabla}_x p = \frac{\gamma}{\mathrm{Re}} \Delta_x \underset{\sim}{u} + \frac{b + d + 2}{b} \frac{1 - \gamma}{\mathrm{Re\,Wi}} \underset{\sim}{\nabla}_x \cdot \underset{\approx}{\tau}, \qquad (\underset{\sim}{x}, t) \in \Omega \times (0, T], \tag{1.2}$$

$$\underset{\sim}{\nabla}_x \cdot \underset{\sim}{u} = 0, \qquad (\underset{\sim}{x}, t) \in \Omega \times (0, T], \tag{1.3}$$

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_x \cdot (\underset{\sim}{u}\psi) + \underset{\sim}{\nabla}_q \cdot \left(\underset{\approx}{\kappa}\, \underset{\sim}{q}\, \psi - \frac{1}{2\mathrm{Wi}} \underset{\sim}{F}(\underset{\sim}{q})\psi\right) = \frac{1}{2\mathrm{Wi}} \Delta_q \psi, \qquad (\underset{\sim}{x}, \underset{\sim}{q}, t) \in \Omega \times D \times (0, T], \tag{1.4}$$

$$\underset{\approx}{\tau}(\underset{\sim}{x}, t) = \int_D \underset{\sim}{F} \otimes \underset{\sim}{q}\, \psi(\underset{\sim}{x}, \underset{\sim}{q}, t)\, \mathrm{d}\underset{\sim}{q}, \qquad (\underset{\sim}{x}, t) \in \Omega \times (0, T], \tag{1.5}$$

$$\underset{\sim}{u}(\underset{\sim}{x}, 0) = \underset{\sim}{u}_0(\underset{\sim}{x}), \quad \underset{\sim}{x} \in \Omega, \qquad \psi(\underset{\sim}{x}, \underset{\sim}{q}, 0) = \psi_0(\underset{\sim}{x}, \underset{\sim}{q}), \quad (\underset{\sim}{x}, \underset{\sim}{q}) \in \Omega \times D. \tag{1.6}$$

Here, Re denotes the Reynolds number, Wi is the Weissenberg number (the ratio of microscopic to macroscopic time-scales) and $\gamma \in (0, 1)$ is the ratio of solvent viscosity to total viscosity. We will also consider the Stokes–Fokker–Planck model, which is relevant in the limit $\mathrm{Re} \to 0_+$, in which (1.2) is replaced by:

$$\underset{\sim}{\nabla}_x p = \gamma \Delta_x \underset{\sim}{u} + \frac{b + d + 2}{b} \frac{1 - \gamma}{\mathrm{Wi}} \underset{\sim}{\nabla}_x \cdot \underset{\approx}{\tau}, \qquad (\underset{\sim}{x}, t) \in \Omega \times (0, T]. \tag{1.7}$$

The tensor $\underset{\approx}{\tau}$ is referred to as the *polymeric extra-stress*, and is the contribution to the fluid stress due to the microscopic dumbbells. Equation (1.5) is known as the *Kramers expression*, and it enables computation of $\underset{\approx}{\tau}$ based on the probability density function (pdf) ψ. Note that $\underset{\approx}{\tau}$ rather than ψ enters into (1.2) (or (1.7)), hence in computations with the micro-macro model it is the accuracy of $\underset{\approx}{\tau}$ that is important, rather than the accuracy of ψ. In Section 3.7 we examine a superconvergence property for $\underset{\approx}{\tau}$, which is extremely beneficial in practice.

[2] The analysis in this paper can be generalised to a broader class of FENE-like potentials that satisfy Hypotheses A and B from [22]. For simplicity of exposition, we restrict our attention to the FENE potential here.

Since ψ is a pdf for each $\underset{\sim}{x} \in \Omega$, the initial condition should be non-negative:

$$\psi(\underset{\sim}{x}, \underset{\sim}{q}, 0) = \psi_0(\underset{\sim}{x}, \underset{\sim}{q}) \geq 0, \qquad \text{for a.e. } (\underset{\sim}{x}, \underset{\sim}{q}) \in \Omega \times D, \tag{1.8}$$

and should also satisfy the following normalisation property:

$$\int_D \psi_0(\underset{\sim}{x}, \underset{\sim}{q}) \, \mathrm{d}\underset{\sim}{q} = 1, \qquad \text{for a.e. } \underset{\sim}{x} \in \Omega. \tag{1.9}$$

It is crucial to note that (1.4) is posed in $2d$ spatial dimensions, plus time. Since the computational complexity of classical numerical methods grows exponentially with the dimension of the spatial domain, the high-dimensionality of (1.4) poses a significant computational challenge. Hence, in a coupled algorithm for (1.2)–(1.6), solving the Fokker–Planck equation is generally the bottleneck step. Therefore, the focus of this paper is on the analysis and implementation of efficient numerical methods for (1.4).

In the discussion above, we have assumed that both Ω and D are domains in $\mathbb{R}^d$ so that the Fokker–Planck equation is posed on $\Omega \times D \subset \mathbb{R}^{2d}$. However, it is not essential that this is the case and, for example, in [10] the authors considered a micro-macro model in which $\Omega \subset \mathbb{R}^2$ and $D \subset \mathbb{R}^3$. No significant complications are introduced from the theoretical or implementational point of view by allowing the dimensionality of D and Ω to be different, but for the rest of this paper we will restrict our attention to the case when these domains have the same dimensionality.

Notice that the Fokker–Planck equation (1.4) contains an unbounded advection coefficient $\underset{\sim}{F}$. This is inconvenient from the point of view of analysis. Therefore we will focus on the following Kolmogorov symmetrisation [23] of the Fokker–Planck equation, in which the spring force, $\underset{\sim}{F}$, has been absorbed into a weighted diffusion term,

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_x \cdot (\underset{\sim}{u} \psi) + \underset{\sim}{\nabla}_q \cdot (\underset{\approx}{\kappa} \, \underset{\sim}{q} \, \psi) = \frac{1}{2\mathrm{Wi}} \underset{\sim}{\nabla}_q \cdot \left(M \underset{\sim}{\nabla}_q \left(\frac{\psi}{M} \right) \right), \tag{1.10}$$

where M is the (normalised) *Maxwellian* defined by

$$\underset{\sim}{q} \mapsto M(\underset{\sim}{q}) := \frac{1}{Z} \exp\left(-U(\tfrac{1}{2}|\underset{\sim}{q}|^2) \right) \in \mathrm{L}^1(D), \qquad Z := \int_D \exp\left(-U(\tfrac{1}{2}|\underset{\sim}{q}|^2) \right) \mathrm{d}\underset{\sim}{q}, \tag{1.11}$$

which in the FENE case is:

$$M(\underset{\sim}{q}) := \frac{1}{Z}(1 - |\underset{\sim}{q}|^2/b)^{b/2}. \tag{1.12}$$

The Maxwellian transformation used in (1.10) allows us to circumvent the analytical difficulties introduced by the unbounded convection coefficient in (1.4), and also the symmetry of the principal term in (1.10) is convenient from the analytical point of view – hence, for the rest of this paper we focus on the Maxwellian-transformed Fokker–Planck equation. It should be noted, however, that an alternative transformation of (1.4) was proposed by Chauvière and Lozinski [11], in which the substitution $\hat{\psi} := \psi/M^{2s/b}$ was used[3]. It was shown in Section 3.2 of [22] that with $b \geq 4s^2/(2s-1)$ and $s > 1/2$, this also leads to a well-posed problem and a stable semidiscretisation in any number of space dimensions. Therefore, the Chauvière–Lozinski transformation would also allow us to deal with the unbounded convection term in (1.4), and indeed, the Chauvière–Lozinski method has some advantages from a practical point of view; in particular, the substitution $\hat{\psi} := \psi/M^{2s/b}$ is independent of b, whereas we use a b-dependent substitution to solve (1.10), which (as discussed in [22]) is less effective for large values of b. Nevertheless, the analysis of (1.10) is simpler, and therefore for the remainder of this paper we focus on the Maxwellian-transformed form of the Fokker–Planck equation, and when considering computations, we restrict our attention to moderate values of b, *e.g.* $b \lesssim 20$, for which, as shown in Section 7 of [22], numerical methods based on the Maxwellian transformed equation perform comparably well

[3] Based on computational experience, Chauvière and Lozinski recommended $s = 2$ and $s = 2.5$ for $d = 2$ and $d = 3$, respectively.

to the Chauvière–Lozinski formulation. If larger values of b are of interest, then it would be straightforward to extend the analysis and numerical methods introduced in this paper to the Chauvière–Lozinski formulation of the Fokker–Planck equation.

In the papers of Chauvière and Lozinski [10,11,26,27] and Helzel and Otto [17], the authors decomposed the Fokker–Planck equation (1.4) (*i.e.* in the nonsymmetrised form) according to $\frac{\partial \psi}{\partial t} + (L_x + L_q)\,\psi = 0$, where

$$L_q\psi = \underset{\sim}{\nabla}_q \cdot (\underset{\approx}{\kappa}\,\underset{\sim}{q}\,\psi) - \frac{1}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_q \cdot \underset{\sim}{F}(\underset{\sim}{q})\psi + \Delta_q\psi\right), \tag{1.13}$$

$$L_x\psi = \underset{\sim}{\nabla}_x \cdot (\underset{\sim}{u}\psi), \tag{1.14}$$

and then they used an alternating-direction approach based on the operators L_q and L_x to compute numerical solutions.

That is, suppose that $0 = t^0 < t^1 < \ldots < t^n < \ldots \leq T$ is a uniform partition of spacing Δt of the interval $[0, T]$. A (two-stage) alternating-direction scheme involves approximating the solution, ψ, by ψ_2 in the following manner: given $\psi_2(t^n)$, $n \geq 0$, with $\psi_2(t^0) = \psi_0$, find ψ_1 and ψ_2 such that,

$$\frac{\partial\psi_1}{\partial t} + L_q\psi_1 = 0, \qquad t \in (t^n, t^{n+1}], \qquad \psi_1(t^n) = \psi_2(t^n), \tag{1.15}$$

$$\frac{\partial\psi_2}{\partial t} + L_x\psi_2 = 0, \qquad t \in (t^n, t^{n+1}], \qquad \psi_2(t^n) = \psi_1(t^{n+1}), \tag{1.16}$$

subject to suitable boundary conditions. A practical alternating-direction numerical method is based on spatial and temporal discretisation of (1.15) and (1.16).

In the case of the Fokker–Planck equation, (1.15) is a convection-diffusion equation posed on D and (1.16) is a first-order hyperbolic equation on Ω. After discretising in space and time, the two-stage scheme described above can be implemented by alternating between applying L_x to Ω cross-sections of $\Omega \times D$ and L_q to D cross-sections of $\Omega \times D$. This type of scheme is also referred to as a dimension-splitting or operator-splitting approach. We use these three terms (*i.e.* alternating direction/dimension-splitting/operator-splitting) interchangeably.

Using this operator-splitting, the "curse of dimensionality" associated with the numerical solution of the Fokker–Planck equation in $2d$ dimensions is ameliorated, as the splitting leads to a sequence of d-dimensional solves at each time step rather than a single $2d$-dimensional solve. Also, this splitting of L allows different numerical methods to be used in Ω and D (resulting in, what we call, a *heterogeneous* alternating-direction scheme). In Section 3 we introduce heterogeneous alternating-direction numerical methods for the FENE Fokker–Planck equation that use a finite element method in Ω and a single-domain Galerkin spectral in D. These are appropriate choices because a finite element method is flexible enough to deal with the general domain Ω, whereas D is always a ball in $\mathbb{R}^d$ and therefore the L_q operator is well suited to a spectral discretisation *via* a polar or spherical coordinate transformation to a cartesian product domain.

Also, note that alternating-direction algorithms are particularly well suited to implementation on parallel computers since they involve solving a large number of independent equations in each time-step. As shall be seen later, our alternating-direction algorithm can be efficiently implemented in parallel, and this enables us to solve large-scale deterministic multiscale problems that may otherwise have been computationally intractable (*e.g.* an important large-scale case is when $\Omega \times D \subset \mathbb{R}^6$).

The structure of this paper is as follows. In Section 2 we briefly discuss the Galerkin spectral method in the $\underset{\sim}{q}$-direction that was examined in detail in [22]. Then we introduce alternating-direction schemes for the full Fokker–Planck equation on $\Omega \times D$ in Section 3. Finally, in Section 5 we present a range of numerical results for (i) the isolated Fokker–Planck equation for an enclosed flow model problem and (ii) the coupled Navier–Stokes–Fokker–Planck system for some channel flow problems of physical interest. We make concluding remarks in Section 6.

2. The Fokker–Planck equation in configuration space

First, we briefly summarise some results from [22] on the discretisation of the d-dimensional Fokker–Planck equation posed in configuration space:

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{\nabla}_q \cdot (\underset{\approx}{\kappa}\, \underset{\sim}{q}\, \psi) = \frac{1}{2\mathrm{Wi}}\, \underset{\sim}{\nabla}_q \cdot \left(M \underset{\sim}{\nabla}_q \frac{\psi}{M} \right), \qquad (\underset{\sim}{q}, t) \in D \times (0, T], \tag{2.1}$$

supplemented with the following initial conditions:

$$\psi(\underset{\sim}{q}, 0) = \psi_0(\underset{\sim}{q}), \qquad \text{for all } \underset{\sim}{q} \in D. \tag{2.2}$$

No boundary condition on ψ will be (explicitly) imposed along ∂D. However, the function $\psi/\sqrt{M}$ will be sought in a weighted Sobolev space $\mathrm{H}^1(D; M)$ defined below; thereby, indirectly, $\psi/\sqrt{M}$ will be forced to satisfy a homogeneous Dirichlet boundary condition on ∂D. This is consistent with the recent results of Zhang and Zhang [37] and Liu and Liu [25]; see in particular Theorem 1.1 in [25]. The implied homogeneous Dirichlet boundary condition on $\psi/\sqrt{M}$ can be seen as an asymptotic decay condition for ψ as $\underset{\sim}{q}$ approaches ∂D; *viz.*,

$$\psi(\underset{\sim}{q}, t) = o\left(\sqrt{M(\underset{\sim}{q})}\, \right), \qquad \text{as } \operatorname{dist}(\underset{\sim}{q}, \partial D) \to 0_+, \ \text{ for all } t \in (0, T]. \tag{2.3}$$

Following [22], we let $\hat{\varphi} := \varphi/\sqrt{M}$ and $\underset{\sim}{\nabla}_M \hat{\varphi} := \sqrt{M}\, \underset{\sim}{\nabla}_q \left(\hat{\varphi}/\sqrt{M} \right)$, and define the function space $\mathrm{H}^1_0(D; M)$ to be the closure of $\mathrm{C}^\infty_0(D)$ in the norm of $\mathrm{H}^1(D; M)$, and

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

Then (2.1) has the following weak formulation: 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

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_D \hat{\psi}\, \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} - \int_D \underset{\approx}{\kappa}\, \underset{\sim}{q} \hat{\psi} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} + \frac{1}{2\mathrm{Wi}} \int_D \underset{\sim}{\nabla}_M \hat{\psi} \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\, \mathrm{d}\underset{\sim}{q} = 0 \qquad \forall \hat{\varphi} \in \mathrm{H}^1_0(D; M), \tag{2.4}$$

in the sense of distributions on $(0, T)$, and $\hat{\psi}(\cdot, 0) = \hat{\psi}_0(\cdot)$. Notice that we solve for $\hat{\psi}$; ψ is recovered by setting $\psi := \sqrt{M}\hat{\psi}$.

It is shown in Section 2 of [22] that $\mathrm{H}^1(D; M) = \mathrm{H}^1_0(D; M)$ and $\mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D; M)$[4]. The connection between $\mathrm{H}^1_0(D; M)$ and $\mathrm{H}^1_0(D)$ is helpful in the development of Galerkin methods for (2.4), since the construction of finite-dimensional subspaces of $\mathrm{H}^1_0(D)$ and the analysis of their approximation properties are well understood.

In [22], a backward-Euler semidiscretisation of (2.4) was studied in detail, and was proved to be unconditionally stable. Also, the existence and uniqueness of a weak solution of (2.4) was established in Theorem 3.2 of [22]. The proof makes use of the stability result alluded to above in order to use compactness results for the bounded sequence of solutions to the semidiscrete problem as $\Delta t \to 0_+$.

We studied the following fully-discrete Galerkin spectral method for the $\underset{\sim}{q}$-direction problem in detail.

[4] In fact, these results hold for all FENE-like potentials, *cf.* footnote 2.

Let $\mathcal{P}_N(D)$ be a finite-dimensional subspace of $\mathrm{H}^1_0(D;M)$ and let $\hat{\psi}^n_N \in \mathcal{P}_N(D)$ be the solution at time level n of the following fully-discrete Galerkin method[5]:

$$\int_D \frac{\hat{\psi}^{n+1}_N - \hat{\psi}^n_N}{\Delta t}\,\hat{\varphi}\;\mathrm{d}\underset{\sim}{q} - \int_D (\underset{\approx}{\kappa}^{n+1}\,\underset{\sim}{q}\,\hat{\psi}^{n+1}_N)\cdot \underset{\sim}{\nabla}_M \hat{\varphi}\;\mathrm{d}\underset{\sim}{q} + \frac{1}{2\mathrm{Wi}}\int_D \underset{\sim}{\nabla}_M \hat{\psi}^{n+1}_N \cdot \underset{\sim}{\nabla}_M \hat{\varphi}\;\mathrm{d}\underset{\sim}{q} = 0$$

$$\forall \hat{\varphi} \in \mathcal{P}_N(D), \quad n = 0, \ldots, N_T - 1, \tag{2.5}$$

$$\hat{\psi}^0_N(\cdot) := \text{the } \mathrm{L}^2(D) \text{ orthogonal projection of } \hat{\psi}_0(\cdot) = \hat{\psi}(\cdot, 0) \text{ onto } \mathcal{P}_N(D). \tag{2.6}$$

The convergence analysis of (2.5), (2.6) in the case $D \subset \mathbb{R}^2$ was considered in [22], where we proved an optimal order convergence estimate. The convergence argument in [22] made use of a number of approximation results; these approximation results will be discussed in more detail in Section 3.5.

3. An alternating-direction scheme for the full Fokker–Planck equation

In this section, we develop numerical methods for the Maxwellian-transformed Fokker–Planck equation posed on $\Omega \times D \times (0, T]$:

$$\frac{\partial \psi}{\partial t} + \underset{\sim}{u}\cdot \underset{\sim}{\nabla}_x \psi + \underset{\sim}{\nabla}_q \cdot (\underset{\approx}{\kappa}\,\underset{\sim}{q}\,\psi) = \frac{1}{2\mathrm{Wi}}\underset{\sim}{\nabla}_q \cdot \left(M \underset{\sim}{\nabla}_q \frac{\psi}{M}\right), \qquad (\underset{\sim}{x}, \underset{\sim}{q}, t) \in \Omega \times D \times (0, T], \tag{3.1}$$

$$\psi(\underset{\sim}{x}, \underset{\sim}{q}, 0) = \psi_0(\underset{\sim}{x}, \underset{\sim}{q}), \qquad (\underset{\sim}{x}, \underset{\sim}{q}) \in \Omega \times D. \tag{3.2}$$

We assume here that $\underset{\sim}{u} : (\underset{\sim}{x}, t) \in \Omega \times [0, T] \mapsto \underset{\sim}{u}(\underset{\sim}{x}, t) \in \mathbb{R}^d$ is an *a priori* defined vector field (hence $\underset{\approx}{\kappa} = \underset{\sim}{\nabla}_x \underset{\sim}{u}$ is known *a priori* also). The precise hypotheses on $\underset{\sim}{u}$ and $\underset{\approx}{\kappa}$ shall be specified below.

The above equation will be referred to as the *full* Fokker–Planck equation, to distinguish it from the equation posed on $D \times (0, T]$ only, that was studied in [22], and discussed in Section 2 above. We focus on the Maxwellian-transformed form of the Fokker–Planck equation given above; however, it should be noted that the numerical methods developed and analysed in the forthcoming sections could just as well be based on the Chauvière–Lozinski-transformed equation that was used to solve the full FENE Fokker–Planck equation in [10,11,27], and was analysed in Section 3.2 of [22].

As discussed in the Introduction, due to the cartesian product structure of the domain $\Omega \times D$, a natural approach to solving (3.1), (3.2) is to use an operator-splitting/alternating-direction method, *cf.* (1.15), (1.16). The Galerkin spectral method on D discussed in Section 2 will be used to solve (1.15), and a finite element method for (1.16) will also be introduced. A finite element method is convenient for the $\underset{\sim}{x}$-direction solver because the physical space domain, Ω, need not have simple geometry.

We propose a fully-practical alternating-direction Galerkin method for (3.1). The approach is similar in spirit to the alternating-direction method used by Chauvière and Lozinski in [10,11,27]. However, there are some important theoretical questions related to applying alternating-direction methods in this context, which have not previously been addressed in the literature, and we focus on these questions here. In particular, we consider the stability and convergence analysis of our alternating-direction scheme for (3.1) in Sections 3.4, 3.5 and 3.6. It is not obvious *a priori* what effect applying a splitting of the form (1.15), (1.16) will have on a discretisation of (3.1), and therefore it is important to rigorously establish the stability and convergence properties of the alternating-direction numerical methods that we propose.

The reader will note that the alternating-direction method under consideration is nonstandard in the sense we consider d-dimensional cross-sections (rather than one-dimensional cross-sections) of $\Omega \times D$. This poses a formidable computational challenge because, as shall be seen below, we typically need to solve a large number problems posed in d spatial dimensions in each time-step. However, the method is extremely well suited to implementation on a parallel architecture since the $\underset{\sim}{q}$-direction solves are completely independent from one another, and similarly the $\underset{\sim}{x}$-direction solves are decoupled also. We discuss the parallel implementation

[5] Here we introduce a backward-Euler temporal discretisation. We will also consider a semi-implicit discretisation in Section 3.

of our alternating-direction scheme in Section 4.2. The computational results in Section 5 were obtained using this parallel implementation.

3.1. **Weak formulation and spatial discretisation**

The full Fokker–Planck equation considered in this section depends on $\underset{\sim}{x} \in \Omega$ as well as $\underset{\sim}{q} \in D$, and therefore we will require the use of slightly different function spaces than in Section 2. Let $\mathrm{L}^2(\Omega \times D)$ be defined in the obvious way, and let $(\cdot, \cdot)$ and $\|\cdot\|$ denote the L^2 inner-product and norm over $\Omega \times D$:

$$(f, g) := \int_{\Omega \times D} f(\underset{\sim}{x}, \underset{\sim}{q}) g(\underset{\sim}{x}, \underset{\sim}{q}) \, \mathrm{d}\underset{\sim}{x} \, \mathrm{d}\underset{\sim}{q} \qquad \text{and} \qquad \|f\|^2 := (f, f).$$

We assume throughout Section 3 that $\underset{\sim}{u}$ is a divergence-free d-component vector function, *i.e.*

$$\underset{\sim}{\nabla}_x \cdot \underset{\sim}{u}(\underset{\sim}{x}, t) = 0 \qquad \text{for a.e. } (\underset{\sim}{x}, t) \in \Omega \times [0, T]. \tag{3.3}$$

It would be straightforward to adapt the arguments in this section to the case where $\underset{\sim}{u}$ is not divergence-free, but this would make the analysis more messy and would shed no further light on the properties of the numerical methods under consideration. Therefore in the interests of clarity and brevity, we restrict our attention to the case when (3.3) is satisfied.

Also, we suppose that

$$\underset{\sim}{u} \in \mathrm{L}^\infty(0, T; \mathrm{L}^\infty(\Omega)) \quad \text{and} \quad \underset{\sim}{\nabla}_x \underset{\sim}{u} =: \underset{\approx}{\kappa} \in \mathrm{W}^{1,\infty}(0, T; \mathrm{L}^\infty(\Omega)), \tag{3.4}$$

where, to simplify notation, we do not explicitly label the d or $d \times d$ dimensionality of the function spaces for $\underset{\sim}{u}(\underset{\sim}{x}, t) \in \mathbb{R}^d$ and $\underset{\approx}{\kappa}(\underset{\sim}{x}, t) \in \mathbb{R}^{d \times d}$. The assumption in (3.4) for $\underset{\approx}{\kappa}$ is slightly stronger than the assumptions in [22], in which $\underset{\approx}{\kappa} \in \mathrm{C}[0, T]$ was required for Lemma 3.1 and $\underset{\approx}{\kappa} \in \mathrm{H}^1(0, T)$ was required for Lemma 3.6.

We shall also use the space $\mathcal{X} := \left\{ \varphi \in \mathrm{L}^2(\Omega \times D) : \varphi \in \mathrm{L}^2(\Omega; \mathrm{H}^1_0(D; M)) \cap \mathrm{H}^1(\Omega; \mathrm{L}^2(D)) \right\}$ equipped with the norm $\|\varphi\|_{\mathcal{X}} := \left\{ \int_{\Omega \times D} \left(|\varphi|^2 + |\underset{\sim}{\nabla}_M \varphi|^2 \right) \mathrm{d}\underset{\sim}{x} \, \mathrm{d}\underset{\sim}{q} \right\}^{\frac{1}{2}}$.

Employing the substitution $\hat{\psi} = \psi / \sqrt{M}$ as in Section 2, the weak formulation of (3.1), (3.2) is as follows: given $\hat{\psi}_0 \in \mathrm{L}^2(\Omega \times D)$, find $\hat{\psi} \in \mathrm{L}^\infty(0, T; \mathrm{L}^2(\Omega \times D)) \cap \mathrm{L}^2(0, T; \mathcal{X})$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\hat{\psi}, \zeta) + \left(\underset{\sim}{u} \cdot \underset{\sim}{\nabla}_x \hat{\psi} \, , \, \zeta \right) - \left(\underset{\approx}{\kappa} \, \underset{\sim}{q} \hat{\psi} \, , \, \underset{\sim}{\nabla}_M \zeta \right) + \frac{1}{2\mathrm{Wi}} \left(\underset{\sim}{\nabla}_M \hat{\psi} \, , \, \underset{\sim}{\nabla}_M \zeta \right) = 0 \qquad \forall \zeta \in \mathcal{X}, \tag{3.5}$$

$$\hat{\psi}(\underset{\sim}{x}, \underset{\sim}{q}, 0) = \hat{\psi}_0(\underset{\sim}{x}, \underset{\sim}{q}), \qquad (\underset{\sim}{x}, \underset{\sim}{q}) \in \Omega \times D, \tag{3.6}$$

in the sense of distributions on $(0, T)$. For simplicity, we avoid boundary conditions on $\partial\Omega \times D$ by assuming that the macroscopic velocity field is an *enclosed flow*, *i.e.* that

$$\underset{\sim}{u} \cdot \underset{\sim}{n} = 0 \text{ on } \partial\Omega, \tag{3.7}$$

where $\underset{\sim}{n} \in \mathbb{R}^d$ is the unit outward normal for Ω. Also, the initial condition (3.6) is understood to be imposed in a weak sense and ψ is recovered by multiplying $\hat{\psi}$ by $\sqrt{M}$.

The term containing $\underset{\approx}{\kappa}$ in (3.5) will be of particular interest since, as we shall see, it is the most difficult term to treat using an alternating-direction method. We introduce the following bilinear form notation for this term, which will be convenient later on:

$$C(\underset{\approx}{\kappa}; f, g) := \left(\underset{\approx}{\kappa} \, \underset{\sim}{q} \, f \, , \, \underset{\sim}{\nabla}_M g \right). \tag{3.8}$$

We now introduce the spatial discretisation of (3.5), (3.6). Let V_h be an N_Ω-dimensional $\mathrm{H}^1(\Omega)$-conforming finite element space corresponding to a mesh $\mathcal{T}_h$ on $\overline{\Omega}$. Also, let $\mathcal{P}_N(D) \subset \mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D;M)$ be an N_D-dimensional space spanned by a set of spectral basis functions on D. Noting that $V_h \otimes \mathcal{P}_N(D) \subset \mathcal{X}$, we obtain a spatially discrete formulation of the full Fokker–Planck equation as follows:

Let $\hat{\psi}_{h,N}(\cdot,\cdot,0) \in V_h \otimes \mathcal{P}_N(D)$ be the $\mathrm{L}^2(\Omega \times D)$ projection of $\hat{\psi}_0$ onto $V_h \otimes \mathcal{P}_N(D)$. Find $\hat{\psi}_{h,N}(\cdot,\cdot,t) \in V_h \otimes \mathcal{P}_N(D)$, $t \in (0,T]$ satisfying (3.5) for all $\zeta \in V_h \otimes \mathcal{P}_N(D)$ in the sense of distributions on $(0,T)$.

3.2. Discretisation in alternating-direction form

We now discuss Galerkin alternating-direction methods for the fully discrete formulation of (3.5), (3.6) defined above.

First of all, define the bases

$$\{Y_k \in \mathcal{P}_N(D) : 1 \leq k \leq N_D\} \quad \text{and} \quad \{X_i \in V_h : 1 \leq i \leq N_\Omega\}, \tag{3.9}$$

such that $\mathrm{span}(\{Y_k\}_{1\leq k\leq N_D}) = \mathcal{P}_N(D)$ and $\mathrm{span}(\{X_i\}_{1\leq i\leq N_\Omega}) = V_h$. Also, we define $M_q, S_q \in \mathbb{R}^{N_D\times N_D}$ as

$$(M_q)_{lk} := \int_D Y_k(\underset{\sim}{q}) Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}, \qquad (S_q)_{lk} := \int_D \underset{\sim}{\nabla}_M Y_k(\underset{\sim}{q}) \cdot \underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}. \tag{3.10}$$

Similarly, $M_x, T_x \in \mathbb{R}^{N_\Omega\times N_\Omega}$ are defined as follows:

$$(M_x)_{ij} := \int_\Omega X_i(\underset{\sim}{x}) X_j(\underset{\sim}{x})\,\mathrm{d}\underset{\sim}{x}, \qquad (T_x)_{ij} := \int_\Omega (\underset{\sim}{u} \cdot \underset{\sim}{\nabla}_x X_j(\underset{\sim}{x}))\, X_i(\underset{\sim}{x})\,\mathrm{d}\underset{\sim}{x}. \tag{3.11}$$

Following [14], a fully discrete form of (3.5) using a backward-Euler time discretisation can be written as follows: Given $\hat{\psi}_N^n = \sum_{jl} \gamma_{jl}^n X_j Y_l \in V_h \otimes \mathcal{P}_N(D)$, find the vector $\underset{\sim}{\gamma}^{n+1} \in \mathbb{R}^{N_\Omega N_D}$, defining a function $\hat{\psi}_N^{n+1} = \sum_{jl} \gamma_{jl}^{n+1} X_j Y_l \in V_h \otimes \mathcal{P}_N(D)$, such that

$$(M_x \otimes M_q)\left(\frac{\underset{\sim}{\gamma}^{n+1} - \underset{\sim}{\gamma}^n}{\Delta t}\right) + (T_x \otimes M_q)\,\underset{\sim}{\gamma}^{n+1} + \frac{1}{2\mathrm{Wi}}(M_x \otimes S_q)\,\underset{\sim}{\gamma}^{n+1} - C(\underset{\approx}{\kappa}^{n+1}; \hat{\psi}_N^{n+1}, \zeta_{ik}) = 0, \tag{3.12}$$

where $\zeta_{ik} = X_i \cdot Y_k \in V_h \otimes \mathcal{P}_N(D)$ and where the matrix tensor product is defined as follows for matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$:

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp\times nq}.$$

It is also possible to obtain a tensor product form discretisation matrix of $C(\underset{\approx}{\kappa};\cdot,\cdot)$; *i.e.* consider $C(\underset{\approx}{\kappa};\zeta_{jl},\zeta_{ik})$ as follows:

$$\begin{aligned} C(\underset{\approx}{\kappa};\zeta_{jl},\zeta_{ik}) &= \int_{\Omega\times D} \left(\underset{\approx}{\kappa}^{n+1}(x)\underset{\sim}{q} X_j(\underset{\sim}{x}) Y_l(\underset{\sim}{q})\right) \cdot \underset{\sim}{\nabla}_M (X_i(\underset{\sim}{x}) Y_k(\underset{\sim}{q}))\,\mathrm{d}\underset{\sim}{x}\,\mathrm{d}\underset{\sim}{q} \\ &= \sum_{s,t=1}^{d} \left(\int_\Omega \kappa_{st}^{n+1}(\underset{\sim}{x}) X_i(\underset{\sim}{x}) X_j(x)\,\mathrm{d}\underset{\sim}{x}\right)\left(\int_D q_t\, Y_l(\underset{\sim}{q})\,\sqrt{M}\frac{\partial}{\partial q_s}\left(\frac{Y_k(\underset{\sim}{q})}{\sqrt{M}}\right)\mathrm{d}\underset{\sim}{q}\right). \end{aligned}$$

Therefore, on defining the matrices $C_x^{st} \in \mathbb{R}^{N_\Omega \times N_\Omega}$ and $C_q^{st} \in \mathbb{R}^{N_D \times N_D}$ for $1 \le s, t \le d$ such that

$$(C_x^{st})_{ij} := \int_\Omega \kappa_{st}^{n+1}(\underset{\sim}{x}) X_i(\underset{\sim}{x}) X_j(x)\, \mathrm{d}\underset{\sim}{x}, \qquad (C_q^{st})_{kl} := \int_D q_t\, Y_l(\underset{\sim}{q})\, \sqrt{M}\, \frac{\partial}{\partial q_s}\left(\frac{Y_k(\underset{\sim}{q})}{\sqrt{M}}\right) \mathrm{d}\underset{\sim}{q}, \tag{3.13}$$

we can rewrite the final term of (3.12) as $\sum_{s,t=1}^{d} (C_x^{st} \otimes C_q^{st})\, \underset{\sim}{\gamma}^{n+1}$.

However, since this matrix expression for $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ contains neither M_x nor M_q, we can not factorise the resulting equation in the same way as in [14] to obtain a suitable alternating-direction scheme. That is, the term $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ causes difficulties because its 'coefficient', $\underset{\approx}{\kappa}(\underset{\sim}{x})\underset{\sim}{q}$, depends on both the $\underset{\sim}{x}$- and $\underset{\sim}{q}$-directions.

This issue has been considered a number of times in the literature. For example, in the context of collocation-based alternating-direction schemes, Celia and Pinder [8,9] and Bialecki and Fernandes [5], developed methods that could handle equations with general variable coefficients. However, our focus is on developing a Galerkin-based framework, and therefore, again, the work of Douglas and Dupont is the most relevant here. In [14], Douglas and Dupont developed a "Laplace modification" scheme for the heat equation with general coefficients. However, it is not obvious how to apply this kind of approach to (3.12), because our problematic term is a convection term rather than a diffusion term. The most natural idea in the spirit of Douglas and Dupont would be to move the $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ term to the right-hand side of (3.12) and treat it explicitly in time. This idea is feasible, but for the purposes of practical computations, we would like to have the option of using a fully-implicit temporal discretisation. Indeed, numerical results in Section 2.6.2 of [21] demonstrated that the semi-implicit temporal discretisation of the Fokker–Planck equation in which the term $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ was treated explicitly in time was less stable for practical computations than the backward Euler discretisation, especially for problems in which the product $\mathrm{Wi}\, \|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}$ is significantly larger than 1.

In order to circumvent this limitation, we develop a Galerkin alternating-direction approach that is an amalgamation of the Douglas and Dupont framework and a new quadrature-based method. Using this approach, we can define either a fully-implicit in time or a semi-implicit in time alternating-direction method for the Fokker–Planck equation. We shall consider both options in detail below.

3.3. The alternating-direction schemes

The first ingredient that we need is a quadrature rule on Ω. Let $\{(\underset{\sim}{x}_m, w_m), w_m > 0,\ \underset{\sim}{x}_m \in \overline{\Omega}, m = 1, \dots, Q_\Omega\}$ define an element-based quadrature rule on the mesh $\mathcal{T}_h$, where the $\underset{\sim}{x}_m$ are the quadrature points and the w_m are the corresponding weights. Therefore, for functions $f, g \in \mathrm{C}^0(\overline{\Omega})$, the quadrature sum is evaluated elementwise as follows,

$$\sum_{m=1}^{Q_\Omega} w_m f(\underset{\sim}{x}_m) g(\underset{\sim}{x}_m) = \sum_{K \in \mathcal{T}_h} \sum_{l=1}^{Q_K} w_l^K f(\underset{\sim}{x}_l^K) g(\underset{\sim}{x}_l^K), \tag{3.14}$$

where Q_K is the number of quadrature points in element K; clearly $Q_\Omega \le \sum_{K \in \mathcal{T}_h} Q_K$, with equality if there are no shared quadrature points between neighbouring elements. From now on, we will use the left-hand side of (3.14) as a shorthand for the right-hand side.

We now introduce two alternative hypotheses on the accuracy of the quadrature rule, Quadrature Hypothesis 1 (QH1) and Quadrature Hypothesis 2 (QH2).

Quadrature Hypothesis 1 (QH1). The quadrature rule satisfies

$$\sum_{m=1}^{Q_\Omega} w_m \kappa_{ij}(\underset{\sim}{x}_m) f(\underset{\sim}{x}_m) g(\underset{\sim}{x}_m) = \int_\Omega \kappa_{ij}(\underset{\sim}{x}) f(\underset{\sim}{x}) g(\underset{\sim}{x})\, \mathrm{d}\underset{\sim}{x}, \tag{3.15}$$

for all $f,\ g \in V_h$ and for each component κ_{ij} of $\underset{\approx}{\kappa}$.

In Section 5, we perform computations for the Navier–Stokes–Fokker–Planck system in which the macroscopic velocity field $\underset{\sim}{u}$ is obtained by solving the Navier–Stokes equations using a finite element method on $\mathcal{T}_h$, *i.e.* the same triangulation that is used for the alternating-direction method for the Fokker–Planck equation. As a result, it is reasonable to assume that the components of $\underset{\approx}{\kappa} = \nabla_x \underset{\sim}{u}$ are represented by piecewise polynomial functions on $\mathcal{T}_h$ and in this case it is certainly possible to satisfy QH1 by choosing an appropriate element-based quadrature rule.

Quadrature Hypothesis 2 (QH2). The quadrature rule satisfies

$$\sum_{m=1}^{Q_\Omega} w_m f(\underset{\sim}{x}_m) g(\underset{\sim}{x}_m) = \int_\Omega f(\underset{\sim}{x}) g(\underset{\sim}{x}) \,\mathrm{d}\underset{\sim}{x}, \tag{3.16}$$

for all $f, g \in V_h$.

QH1 is a stronger hypothesis than QH2, and therefore in general we will require a larger value of Q_Ω in order to satisfy QH1. Some results in the following analysis will require QH1, whereas for others QH2 will suffice.

Next, let $\hat{\psi}_{h,N} \in V_h \otimes \mathcal{P}_N(D)$ denote the numerical solution of the full Fokker–Planck equation. Recalling the bases from (3.9), $\hat{\psi}_{h,N}$ can be written in terms of coefficients $\{\hat{\psi}_{ik}\}$ as follows:

$$\hat{\psi}_{h,N} := \sum_{i=1}^{N_\Omega} \sum_{k=1}^{N_D} \hat{\psi}_{ik} X_i Y_k \in V_h \otimes \mathcal{P}_N(D). \tag{3.17}$$

Define the *line functions*, $\hat{\psi}_k$, for $k = 1, \ldots, N_D$, as follows:

$$\hat{\psi}_k := \sum_{i=1}^{N_\Omega} \hat{\psi}_{ik} X_i \in V_h, \tag{3.18}$$

and note that (3.17) can be rewritten using (3.18) as follows:

$$\hat{\psi}_{h,N}(\underset{\sim}{x}, \underset{\sim}{q}) = \sum_{k=1}^{N_D} \hat{\psi}_k(\underset{\sim}{x}) Y_k(\underset{\sim}{q}). \tag{3.19}$$

The formula (3.19) shall be useful in the discussion of the alternating-direction methods below.

As discussed above, the term $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ is the most problematic in terms of applying an alternating-direction method to the Fokker–Planck equation. Therefore we begin by considering how to use a quadrature-based scheme to derive an alternating-direction type of formulation of this term.

Suppose that QH1 is satisfied and that we have the line function decomposition (3.19) for $\hat{\psi}_{h,N}$, in which $\hat{\psi}_k \in V_h$ for $k = 1, \ldots, N_D$. Also, let $\zeta = X \cdot Y \in V_h \otimes \mathcal{P}_N(D)$. Then,

$$\begin{aligned} C(\underset{\approx}{\kappa}; \hat{\psi}_{h,N}, \zeta) &= \int_{\Omega \times D} (\underset{\approx}{\kappa}\, \underset{\sim}{q}\, \hat{\psi}_{h,N}(\underset{\sim}{x}, \underset{\sim}{q})) \cdot \underset{\sim}{\nabla}_M \zeta(\underset{\sim}{x}, \underset{\sim}{q}) \,\mathrm{d}\underset{\sim}{q} \,\mathrm{d}\underset{\sim}{x} \\ &= \int_D \sum_{k=1}^{N_D} \int_\Omega \left[\underset{\approx}{\kappa}\, \underset{\sim}{q}\, \hat{\psi}_k(\underset{\sim}{x}) Y_k(\underset{\sim}{q})\right] \cdot \underset{\sim}{\nabla}_M \Big(X(\underset{\sim}{x}) Y(\underset{\sim}{q})\Big) \mathrm{d}\underset{\sim}{x} \,\mathrm{d}\underset{\sim}{q} \\ &= \int_D \sum_{k=1}^{N_D} \sum_{m=1}^{Q_\Omega} w_m \left[\underset{\approx}{\kappa}(\underset{\sim}{x}_m)\, \underset{\sim}{q}\, \hat{\psi}_k(\underset{\sim}{x}_m) Y_k(\underset{\sim}{q})\right] \cdot \underset{\sim}{\nabla}_M \Big(X(\underset{\sim}{x}_m) Y(\underset{\sim}{q})\Big) \mathrm{d}\underset{\sim}{q} \\ &= \sum_{m=1}^{Q_\Omega} w_m\, X(\underset{\sim}{x}_m) \left\{ \sum_{k=1}^{N_D} \hat{\psi}_k(\underset{\sim}{x}_m) \left(\int_D (\underset{\approx}{\kappa}(\underset{\sim}{x}_m)\, \underset{\sim}{q}\, Y_k(\underset{\sim}{q})) \cdot \underset{\sim}{\nabla}_M Y(\underset{\sim}{q}) \,\mathrm{d}\underset{\sim}{q} \right) \right\} \cdot \end{aligned} \tag{3.20}$$

This shows the equivalence between the Galerkin formulation of $C(\underset{\approx}{\kappa};\cdot,\cdot)$ on $\Omega \times D$ and the quadrature sum over $m = 1,\dots,Q_\Omega$ of the term

$$\sum_{k=1}^{N_D} \hat{\psi}_k(\underset{\sim}{x}_m)\left(\int_D (\underset{\approx}{\kappa}(\underset{\sim}{x}_m)\,\underset{\sim}{q}\,Y_k(\underset{\sim}{q}))\cdot \underset{\sim}{\nabla}_M Y(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q}\right), \tag{3.21}$$

which is the $\underset{\sim}{q}$-direction discretisation of $C(\underset{\approx}{\kappa};\cdot,\cdot)$ where the coefficient vector (corresponding to the quadrature point $\underset{\sim}{x}_m$) is the set of sampled line functions $\hat{\psi}_k(\underset{\sim}{x}_m)$, $k = 1,\dots,N_D$.

It is clear from (3.20) that sampling functions at the quadrature points $\{\underset{\sim}{x}_m \in \overline{\Omega},\ m = 1,\dots,Q_\Omega\}$ will play an important role in the alternating-direction methods we define below. We will also require a reconstruction operator, which maps from a set of values at the quadrature points to a function in V_h. We now introduce this operator. To simplify the notation, we first define the following discrete inner product and norm over Ω for $\{f_m\},\ \{g_m\} \in \mathbb{R}^{Q_\Omega}$:

$$(\{f_m\},\{g_m\})_{\ell^2(\Omega)} := \sum_{m=1}^{Q_\Omega} w_m f_m g_m, \quad \text{and} \quad \|\{f_m\}\|_{\ell^2(\Omega)} := (\{f_m\},\{f_m\})^{\frac{1}{2}}_{\ell^2(\Omega)}. \tag{3.22}$$

Note that, by (3.15) or (3.16), for $f,\ g \in V_h$, $(\{f(\underset{\sim}{x}_m)\},\{g(\underset{\sim}{x}_m)\})_{\ell^2(\Omega)} = (f,g)_{\mathrm{L}^2(\Omega)}$, where $(\cdot,\cdot)_{\mathrm{L}^2(\Omega)}$ is the standard L^2 inner product on Ω. Next we define the reconstruction operator $\mathcal{R} : \{f_m\} \in \mathbb{R}^{Q_\Omega} \mapsto \mathcal{R}\{f_m\} \in V_h$ such that

$$(\mathcal{R}\{f_m\},X)_{\mathrm{L}^2(\Omega)} = (\{f_m\},\{X(\underset{\sim}{x}_m)\})_{\ell^2(\Omega)} \qquad \forall\, X \in V_h. \tag{3.23}$$

Remark 3.1. For any $\mathcal{R}\{f_m\} \in V_h$, there exist real numbers $\gamma_1,\dots,\gamma_{N_\Omega}$ such that $\mathcal{R}\{f_m\} = \sum_{j=1}^{N_\Omega}\gamma_j X_j$. Letting $X = X_i$, $i = 1,\dots,N_\Omega$, above it is clear that (3.23) is equivalent to the linear system $M_x\underset{\sim}{\gamma} = \underset{\sim}{F}$ where $M_x \in \mathbb{R}^{N_\Omega\times N_\Omega}$ is the V_h mass matrix, $\underset{\sim}{\gamma} = (\gamma_1,\dots,\gamma_{N_\Omega})^T$, and $\underset{\sim}{F} \in \mathbb{R}^{N_\Omega}$ is such that $F_i = (\{f_m\},\{X_i(\underset{\sim}{x}_m)\})_{\ell^2(\Omega)}$. The matrix M_x is nonsingular, and therefore the reconstruction operator $\mathcal{R}$ is correctly defined by (3.23).

We now define two alternating-direction methods for (3.5), (3.6), referred to as method I and method II. The distinction between these schemes is that method I uses a semi-implicit spectral method in the $\underset{\sim}{q}$-direction (*i.e.* the term containing $\underset{\approx}{\kappa}$ is treated explicitly in time) whereas method II uses a fully-implicit temporal discretisation.

Method I: Semi-implicit scheme. Method I is initialised by computing the $\mathrm{L}^2(\Omega\times D)$ projection, $\hat{\psi}^0_{h,N}$, of the initial datum $\hat{\psi}_0 \in \mathrm{L}^2(\Omega\times D)$ onto $V_h \otimes \mathcal{P}_N(D)$, so that $\hat{\psi}^0_{h,N} \in V_h \otimes \mathcal{P}_N(D)$ satisfies

$$\left(\hat{\psi}_0,\,\zeta\right) = \left(\hat{\psi}^0_{h,N},\zeta\right) \qquad \text{for all } \zeta \in V_h \otimes \mathcal{P}_N(D). \tag{3.24}$$

Then, as in (1.15), (1.16), this alternating-direction method consists of two stages at each time-step: the $\underset{\sim}{q}$-direction stage and the $\underset{\sim}{x}$-direction stage. We begin with the $\underset{\sim}{q}$-direction stage, in which we essentially use the Galerkin spectral method in D from Section 2.

Suppose $\hat{\psi}^n_{h,N} \in V_h \otimes \mathcal{P}_N(D)$. Then in the $\underset{\sim}{q}$-direction stage we compute $\hat{\psi}^{n*}_{h,N}(\underset{\sim}{x}_m,\cdot) \in \mathcal{P}_N(D)$ for each $m = 1,\dots,Q_\Omega$ satisfying

$$\int_D \frac{\hat{\psi}^{n*}_{h,N}(\underset{\sim}{x}_m,\underset{\sim}{q}) - \hat{\psi}^{n}_{h,N}(\underset{\sim}{x}_m,\underset{\sim}{q})}{\Delta t}\,Y_l(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q} + \frac{1}{2\mathrm{Wi}}\int_D \underset{\sim}{\nabla}_M \hat{\psi}^{n*}_{h,N}(\underset{\sim}{x}_m,\underset{\sim}{q})\cdot \underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q}$$
$$= \int_D (\underset{\approx}{\kappa}^n(\underset{\sim}{x}_m)\,\underset{\sim}{q}\,\hat{\psi}^n_{h,N}(\underset{\sim}{x}_m,\underset{\sim}{q}))\cdot\underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q}, \tag{3.25}$$

for $l = 1, \ldots, N_D$. Thus, (3.25) defines an $N_D \times N_D$ linear system at each quadrature point $\underset{\sim}{x}_m$. In order to separate out the $\underset{\sim}{x}$- and $\underset{\sim}{q}$-direction dependencies more clearly, we rewrite this equation in terms of line functions using (3.19), *i.e.*:

$$\begin{aligned}\sum_{k=1}^{N_D} \hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\left(\int_D Y_k(\underset{\sim}{q})\,Y_l(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q} + \frac{\Delta t}{2\mathrm{Wi}}\int_D \underset{\sim}{\nabla}_M Y_k(\underset{\sim}{q})\cdot\underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}\right)&\\ = \sum_{k=1}^{N_D} \hat{\psi}_k^{n}(\underset{\sim}{x}_m)\left(\int_D Y_k(\underset{\sim}{q})\,Y_l(\underset{\sim}{q})\;\mathrm{d}\underset{\sim}{q} + \Delta t\int_D (\underset{\approx}{\kappa}^n(\underset{\sim}{x}_m)\,\underset{\sim}{q}\,Y_k(\underset{\sim}{q}))\cdot\underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}\right)&,\end{aligned} \tag{3.26}$$

for $l = 1, \ldots, N_D$. This system is solved at each quadrature point $\underset{\sim}{x}_m$, $m = 1, \ldots, Q_\Omega$.

Equation (3.26) shows that in the $\underset{\sim}{q}$-direction stage, the sampled values of the line functions, *i.e.* $\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)$, $k = 1, \ldots, N_D$, $m = 1, \ldots, Q_\Omega$, are the coefficients to be computed. We determine these values by solving a different linear system at each quadrature point. Note that these linear systems are completely independent from one another. This independence enables parallel computation to be used very effectively in this context; this will be discussed in more detail later.

The $\underset{\sim}{q}$-direction stage is complete once the values $\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)$, $k = 1, \ldots, N_D$, $m = 1, \ldots, Q_\Omega$ have been computed, and then we can begin solving in the $\underset{\sim}{x}$-direction. In the $\underset{\sim}{x}$-direction stage, we use a finite element discretisation of the transport equation (1.16) to update the output data from the $\underset{\sim}{q}$-direction stage. That is, for a given k, we find $\hat{\psi}_k^{n+1} \in V_h$, satisfying:

$$\int_\Omega \hat{\psi}_k^{n+1} X_i\,\mathrm{d}\underset{\sim}{x} + \Delta t\int_\Omega \left(\underset{\sim}{u}^{n+1}\cdot\underset{\sim}{\nabla}_x\hat{\psi}_k^{n+1}\right) X_i\,\mathrm{d}\underset{\sim}{x} = \int_\Omega \mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\} X_i\,\mathrm{d}\underset{\sim}{x}, \tag{3.27}$$

for $i = 1, \ldots, N_\Omega$.

Note, however, that based on (3.23) for the right-hand side in (3.27) we have:

$$\int_\Omega \mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\} X_i\,\mathrm{d}\underset{\sim}{x} = \sum_{m=1}^{Q_\Omega} w_m\,\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\,X_i(\underset{\sim}{x}_m) =: F_i. \tag{3.28}$$

Hence we do not actually have to explicitly compute $\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\} \in V_h$ in order to solve (3.27), since the following system is equivalent to (3.27):

$$\int_\Omega \hat{\psi}_k^{n+1} X_i\,\mathrm{d}\underset{\sim}{x} + \Delta t\int_\Omega \left(\underset{\sim}{u}^{n+1}\cdot\underset{\sim}{\nabla}_x\hat{\psi}_k^{n+1}\right) X_i\,\mathrm{d}\underset{\sim}{x} = F_i, \tag{3.29}$$

for $i = 1, \ldots, N_\Omega$. We solve (3.29) for each $k = 1, \ldots, N_D$, and, just as in the $\underset{\sim}{q}$-direction, these computations are decoupled from one another.

Once the $\underset{\sim}{x}$-direction computations are complete, we have the numerical solution at time level $n+1$: $\hat{\psi}_{h,N}^{n+1} = \sum_{k=1}^{N_D}\hat{\psi}_k^{n+1} Y_k \in V_h \otimes \mathcal{P}_N(D)$. Hence method I is defined by the initialisation (3.24), the $\underset{\sim}{q}$-direction spectral method (3.26) and the $\underset{\sim}{x}$-direction finite element method (3.29).

It is straightforward to show that this numerical method is well-defined (see Lem. 3.4 in [21]). In the next lemma we derive a Galerkin formulation posed on $\Omega \times D$ for method I. The availability of an equivalent one-step method is extremely useful for the numerical analysis of Galerkin alternating-direction methods.

Lemma 3.2. *Suppose the $\utilde{x}$-direction quadrature rule satisfies* QH1*. Method* I *is equivalent to the following fully-discrete formulation:*

Given $\hat\psi^0_{h,N} \in V_h \otimes \mathcal{P}_N(D)$ defined as in (3.24), for each $n = 0, \dots, N_T - 1$, $\hat\psi^{n+1}_{h,N} \in V_h \otimes \mathcal{P}_N(D)$ satisfies

$$\left(\frac{\hat\psi^{n+1}_{h,N} - \hat\psi^{n}_{h,N}}{\Delta t}, \zeta\right) + \left(\utilde{u} \cdot \utilde{\nabla}_x \hat\psi^{n+1}_{h,N}\, , \zeta\right) + \frac{1}{2\mathrm{Wi}}\left(\utilde{\nabla}_M \hat\psi^{n+1}_{h,N}\, , \utilde{\nabla}_M \zeta\right) + \frac{\Delta t}{2\mathrm{Wi}}\left(\utilde{\nabla}_M\left(\utilde{u} \cdot \utilde{\nabla}_x \hat\psi^{n+1}_{h,N}\right), \utilde{\nabla}_M \zeta\right) - \left(\underset{\approx}{\kappa}^n\, \utilde{q}\, \hat\psi^{n}_{h,N}\, , \utilde{\nabla}_M \zeta\right) = 0, \tag{3.30}$$

for all $\zeta \in V_h \otimes \mathcal{P}_N(D)$.

Proof. Multiplying (3.26) through by $X_i(\utilde{x}_m)$, where $X_i \in V_h$, and performing the weighted sum according to (3.14) gives,

$$\begin{aligned}
\sum_{k=1}^{N_D} (\{\hat\psi^{n*}_k(\utilde{x}_m)\}, \{X_i(\utilde{x}_m)\})_{\ell^2(\Omega)} &\left(\int_D Y_k(\utilde{q})\, Y_l(\utilde{q})\, \mathrm{d}\utilde{q} + \frac{\Delta t}{2\mathrm{Wi}} \int_D \utilde{\nabla}_M Y_k(\utilde{q}) \cdot \utilde{\nabla}_M Y_l(\utilde{q})\, \mathrm{d}\utilde{q}\right) \\
&= \sum_{k=1}^{N_D} (\{\hat\psi^{n}_k(\utilde{x}_m)\}, \{X_i(\utilde{x}_m)\})_{\ell^2(\Omega)} \left(\int_D Y_k(\utilde{q})\, Y_l(\utilde{q})\, \mathrm{d}\utilde{q}\right) \\
&+ \Delta t \sum_{m=1}^{Q_\Omega} w_m X_i(\utilde{x}_m) \left\{\sum_{k=1}^{N_D} \hat\psi^{n}_k(\utilde{x}_m) \left(\int_D (\underset{\approx}{\kappa}^n(\utilde{x}_m)\, \utilde{q}\, Y_k(\utilde{q})) \cdot \utilde{\nabla}_M Y_l(\utilde{q})\, \mathrm{d}\utilde{q}\right)\right\}.
\end{aligned} \tag{3.31}$$

Using the reconstruction operator, (3.23), with the ℓ^2 inner products and the argument of (3.20) on the term on the third line[6], we obtain the following formulation for $\mathcal{R}\hat\psi^{n*}_{h,N} \in V_h \otimes \mathcal{P}_N(D)$:

$$\begin{aligned}
\int_{\Omega \times D} \frac{\mathcal{R}\hat\psi^{n*}_{h,N}(\utilde{x}, \utilde{q}) - \hat\psi^{n}_{h,N}(\utilde{x}, \utilde{q})}{\Delta t}\, \zeta(\utilde{x}, \utilde{q})\, \mathrm{d}\utilde{q}\, \mathrm{d}\utilde{x} &+ \frac{1}{2\mathrm{Wi}} \int_{\Omega \times D} \utilde{\nabla}_M \mathcal{R}\hat\psi^{n*}_{h,N}(\utilde{x}, \utilde{q}) \cdot \utilde{\nabla}_M \zeta(\utilde{x}, \utilde{q})\, \mathrm{d}\utilde{q}\, \mathrm{d}\utilde{x} \\
&= \int_{\Omega \times D} (\underset{\approx}{\kappa}^n(\utilde{x})\, \utilde{q}\, \hat\psi^{n}_{h,N}(\utilde{x}, \utilde{q})) \cdot \utilde{\nabla}_M \zeta(\utilde{x}, \utilde{q})\, \mathrm{d}\utilde{q}\, \mathrm{d}\utilde{x},
\end{aligned} \tag{3.32}$$

where $\zeta = X_i \cdot Y_l$ is an element of $V_h \otimes \mathcal{P}_N(D)$ and the numerical solution at the intermediate "time level" $n*$ is defined as:

$$\mathcal{R}\hat\psi^{n*}_{h,N} := \sum_{k=1}^{N_D} \mathcal{R}\{\hat\psi^{n*}_k(\utilde{x}_m)\} Y_k \in V_h \otimes \mathcal{P}_N(D). \tag{3.33}$$

The $\utilde{x}$-direction stage is more straightforward to deal with; we use the classical Douglas–Dupont Galerkin alternating-direction approach (*cf*. [14]) for (3.27), since it does not contain any $\utilde{q}$-dependent coefficients.

Let $\mathcal{R}\{\hat\psi^{n*}_k(\utilde{x}_m)\} = \sum_{i=1}^{N_\Omega} \gamma^{n*}_{ik} X_i$ so that according to (3.33), the vector

$$\utilde{\gamma}^{n*} = (\gamma^{n*}_{11}, \dots, \gamma^{n*}_{N_\Omega 1}, \gamma^{n*}_{12}, \dots, \gamma^{n*}_{N_\Omega N_D}) \in \mathbb{R}^{N_\Omega N_D}$$

determines $\mathcal{R}\hat\psi^{n*}_{h,N}$. Similarly, denote the coefficient vector for $\hat\psi^{n+1}_{h,N}$ as $\utilde{\gamma}^{n+1} \in \mathbb{R}^{N_\Omega N_D}$, and since the vector entries are ordered in blocks according to the $\utilde{q}$-direction degrees-of-freedom, it follows that (3.27) can be written

[6] Note that $\hat\psi_k$ in the term on the last line of (3.31) must be at time level n for the argument of (3.20) to apply since it relies on the values $\{\hat\psi^n_k(\utilde{x}_m)\}$ interpolating a function in V_h.

as a linear system where the matrices are in tensor product form, *i.e.*:

$$(I_q \otimes M_x + \Delta t I_q \otimes T_x)\,\underset{\sim}{\gamma}^{n+1} = (I_q \otimes M_x)\underset{\sim}{\gamma}^{n*}, \tag{3.34}$$

where the discretisation matrices are as in (3.11) and I_q is the $N_D \times N_D$ identity matrix.

Equation (3.32) can be written in tensor product matrix form also:

$$\left(M_q \otimes M_x + \frac{\Delta t}{2\mathrm{Wi}} S_q \otimes M_x\right)\underset{\sim}{\gamma}^{n*} = (M_q \otimes M_x)\underset{\sim}{\gamma}^{n} + \Delta t C(\underset{\approx}{\kappa}^{n}; \hat{\psi}^{n}_{h,N}, \zeta_{il}), \tag{3.35}$$

where $\zeta_{il} = X_i \cdot Y_l \in V_h \otimes \mathcal{P}_N(D)$, for $1 \leq i \leq N_\Omega$ and $1 \leq l \leq N_D$. Also, M_q and S_q are defined in (3.10).

Multiplying (3.34) by $(M_q \otimes I_x + \Delta t/(2\mathrm{Wi}) S_q \otimes I_x)$, where I_x is the $N_\Omega \times N_\Omega$ identity matrix, yields,

$$\left(M_q \otimes M_x + \Delta t M_q \otimes T_x + \frac{\Delta t}{2\mathrm{Wi}} S_q \otimes M_x + \frac{(\Delta t)^2}{2\mathrm{Wi}} S_q \otimes T_x\right)\underset{\sim}{\gamma}^{n+1} = \left(M_q \otimes M_x + \frac{\Delta t}{2\mathrm{Wi}} S_q \otimes M_x\right)\underset{\sim}{\gamma}^{n*}, \tag{3.36}$$

and equating the left-hand side of (3.35) with the right-hand side of (3.36), gives:

$$\left(M_x \otimes M_q + \Delta t M_q \otimes T_x + \frac{\Delta t}{2\mathrm{Wi}} S_q \otimes M_x + \frac{(\Delta t)^2}{2\mathrm{Wi}} S_q \otimes T_x\right)\underset{\sim}{\gamma}^{n+1} = (M_q \otimes M_x)\underset{\sim}{\gamma}^{n} + \Delta t C(\underset{\approx}{\kappa}^{n}; \hat{\psi}^{n}_{h,N}, \zeta_{il}). \tag{3.37}$$

Equation (3.37) is equivalent to the inner product form in (3.30) and hence the proof is complete. □

Method II: Fully-implicit scheme. Method II is very similar to method I, the sole difference being that the term $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ is now treated implicitly in time, and therefore we refer to method II as a fully-implicit scheme. We do not discuss the initialisation step or the $\underset{\sim}{x}$-direction scheme here because they are the same as in method I. Instead, we move immediately to discussing the $\underset{\sim}{q}$-direction stage of method II.

Using the line function notation of (3.26), the $\underset{\sim}{q}$-direction numerical method is defined as follows: Given the line functions $\hat{\psi}^{n}_{k} \in V_h$, $k = 1, \ldots, N_D$, determine the values $\hat{\psi}^{n*}_{k}(\underset{\sim}{x}_m)$ satisfying

$$\begin{aligned}\sum_{k=1}^{N_D} \hat{\psi}^{n*}_{k}(\underset{\sim}{x}_m)\Bigg(\int_D Y_k(\underset{\sim}{q})\,Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q} + \frac{\Delta t}{2\mathrm{Wi}}\int_D \underset{\sim}{\nabla}_M Y_k(\underset{\sim}{q}) \cdot \underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}\\ -\Delta t \int_D (\underset{\approx}{\kappa}^{n+1}(\underset{\sim}{x}_m)\,\underset{\sim}{q}\,Y_k(\underset{\sim}{q})) \cdot \underset{\sim}{\nabla}_M Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q}\Bigg) = \sum_{k=1}^{N_D} \hat{\psi}^{n}_{k}(\underset{\sim}{x}_m)\int_D Y_k(\underset{\sim}{q})\,Y_l(\underset{\sim}{q})\,\mathrm{d}\underset{\sim}{q},\end{aligned} \tag{3.38}$$

for all $l = 1, \ldots, N_D$, and for each quadrature point $\underset{\sim}{x}_m, m = 1, \ldots, Q_\Omega$. Equation (3.38) is exactly the backward Euler Galerkin spectral method that was studied in [22].

Unfortunately we cannot derive an equivalent one-step Galerkin formulation for method II using the same reasoning as in Lemma 3.2 because the proof of that lemma relied on the term $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ being explicit-in-time (*cf.* footnote 6). In order to derive a one-step formulation for method II, we would need to recover an integral of $\mathcal{R}\{\psi^{n*}_{k}(\underset{\sim}{x}_m)\}$ over $\Omega \times D$ from the $C(\underset{\approx}{\kappa}; \cdot, \cdot)$ term in (3.38). However, this is not possible because it would require a $\underset{\approx}{\kappa}$-weighted reconstruction operator, as distinct from the unweighted reconstruction operator defined in (3.23).

Remark 3.3. In both methods I and II we use a standard Galerkin discretisation of the convection equation in Ω. With piecewise polynomials of degree p the rate of convergence of a standard Galerkin method in the L^2 norm for this equation is $\mathcal{O}(h^p)$, which is one less than the optimal order $\mathcal{O}(h^{p+1})$. With streamline-diffusion (SUPG) stabilization, the order of convergence on a general triangulation is $\mathcal{O}(h^{p+1/2})$. It would be straightforward to replace (3.27) with an SUPG scheme such as the one described in [6], but we prefer the standard Galerkin formulation here for the sake of simplicity. In the numerical results in Section 5 we do not observe any oscillations in Ω, which indicates that the standard Galerkin discretisation performs well in practice in the present context.

3.4. Stability of methods I and II

First of all, we consider the stability of method I, which is straightforward due to the availability of an equivalent one-step method. We introduce right-hand side terms into (3.30) as follows:

$$\begin{aligned}\left(\frac{\hat\psi^{n+1}_{h,N}-\hat\psi^{n}_{h,N}}{\Delta t},\zeta\right)+\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat\psi^{n+1}_{h,N}\,,\zeta\right)+\frac{1}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\,,\underset{\sim}{\nabla}_M\zeta\right)&\\ +\frac{\Delta t}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat\psi^{n+1}_{h,N}\right),\underset{\sim}{\nabla}_M\zeta\right)-\left(\underset{\approx}{\kappa}^n\,\underset{\sim}{q}\,\hat\psi^{n}_{h,N}\,,\underset{\sim}{\nabla}_M\zeta\right)&=\left(\mu^{n+1},\zeta\right)+\left(\underset{\sim}{\nu}^{n+1}\,,\underset{\sim}{\nabla}_M\zeta\right),\end{aligned}\tag{3.39}$$

for all $\zeta\in V_h\otimes\mathcal{P}_N(D)$, where $\mu^{n+1}\in\mathrm{L}^2(\Omega\times D)$ and $\underset{\sim}{\nu}^{n+1}\in\mathrm{L}^2(\Omega\times D)^d$, $n=0,\ldots,N_T-1$. Right-hand side terms of this form will be useful when we derive convergence estimates in Section 3.6.

Lemma 3.4. *If* QH1 *holds, so that we have the equivalent one-step formulation for method* I *given in Lemma* 3.2*, then letting* $\Delta t=T/N_T$*,* $N_T\geq 1$*, for* $\hat\psi^{s}_{h,N}\in V_h\otimes\mathcal{P}_N(D)$ *we have the following stability estimate for* (3.39)*:*

$$\begin{aligned}\|\hat\psi^{s}_{h,N}\|^2+\sum_{n=0}^{s-1}\Delta t\left\|\frac{\hat\psi^{n+1}_{h,N}-\hat\psi^{n}_{h,N}}{\sqrt{\Delta t}}\right\|^2+\sum_{n=0}^{s-1}\frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\|^2&\\ \leq \mathrm{e}^{Ks\Delta t}\left\{\|\hat\psi^{0}_{h,N}\|^2+\sum_{n=0}^{s-1}2\Delta t\left(\|\mu^{n+1}\|^2+4\|\underset{\sim}{\nu}^{n+1}\|^2\right)\right\}&,\end{aligned}\tag{3.40}$$

for $1\leq s\leq N_T$*, where* $K=2(1+4\mathrm{Wi}\,b\,|\underset{\approx}{\kappa}|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))})$.

Proof. Set $\zeta=\hat\psi^{n+1}_{h,N}$ in (3.39) to get

$$\begin{aligned}\left(\frac{\hat\psi^{n+1}_{h,N}-\hat\psi^{n}_{h,N}}{\Delta t},\hat\psi^{n+1}_{h,N}\right)+\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat\psi^{n+1}_{h,N},\,\hat\psi^{n+1}_{h,N}\right)+\frac{1}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\|^2&\\ +\frac{\Delta t}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat\psi^{n+1}_{h,N}\right),\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\right)-\left(\underset{\approx}{\kappa}^n\,\underset{\sim}{q}\,\hat\psi^{n}_{h,N},\,\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\right)&\\ =\left(\mu^{n+1},\hat\psi^{n+1}_{h,N}\right)+\left(\underset{\sim}{\nu}^{n+1},\,\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\right)&.\end{aligned}\tag{3.41}$$

The $\underset{\sim}{x}$-transport term vanishes because of (3.3) and (3.7). Similarly, the first term on the second line vanishes since

$$\begin{aligned}\left(\underset{\sim}{\nabla}_M\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat\psi^{n+1}_{h,N}\right),\underset{\sim}{\nabla}_M\hat\psi^{n+1}_{h,N}\right)&=\int_{\Omega\times D}M\sum_{j=1}^{d}\left(\sum_{i=1}^{d}u_i\left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial q_j}\frac{\hat\psi^{n+1}_{h,N}}{\sqrt{M}}\right)\left(\frac{\partial}{\partial q_j}\frac{\hat\psi^{n+1}_{h,N}}{\sqrt{M}}\right)\right)\mathrm{d}\underset{\sim}{x}\,\mathrm{d}\underset{\sim}{q}\\ &=\frac{1}{2}\int_{\Omega\times D}M\sum_{j=1}^{d}\left(\sum_{i=1}^{d}u_i\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial q_j}\frac{\hat\psi^{n+1}_{h,N}}{\sqrt{M}}\right)^2\right)\mathrm{d}\underset{\sim}{x}\,\mathrm{d}\underset{\sim}{q}\\ &=-\frac{1}{2}\int_{\Omega\times D}M\sum_{j=1}^{d}\left(\left(\underset{\sim}{\nabla}_x\cdot\underset{\sim}{u}\right)\left(\frac{\partial}{\partial q_j}\frac{\hat\psi^{n+1}_{h,N}}{\sqrt{M}}\right)^2\right)\mathrm{d}\underset{\sim}{x}\,\mathrm{d}\underset{\sim}{q}=0.\end{aligned}$$

Applying the identity $2(a-b)a = a^2 - b^2 + (a-b)^2$ to the first term in (3.41), yields

$$\begin{aligned}\|\hat\psi^{n+1}_{h,N}\|^2 + \left\|\hat\psi^{n+1}_{h,N} - \hat\psi^{n}_{h,N}\right\|^2 + \frac{\Delta t}{\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\|^2 &= \|\hat\psi^{n}_{h,N}\|^2 + 2\Delta t\left(\underset{\approx}{\kappa}^n\, \underset{\sim}{q}\, \hat\psi^{n}_{h,N}, \underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\right)\\ &\quad + 2\Delta t\left(\mu^{n+1}, \hat\psi^{n+1}_{h,N}\right) + 2\Delta t\left(\underset{\sim}{\nu}^{n+1}, \underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\right)\\ &=: \|\hat\psi^{n}_{h,N}\|^2 + T_1 + T_2 + T_3. \end{aligned}\tag{3.42}$$

Applying the Cauchy–Schwarz inequality to T_1, T_2 and T_3, we obtain:

$$\begin{aligned}(1-\Delta t)\|\hat\psi^{n+1}_{h,N}\|^2 + \Delta t\left\|\frac{\hat\psi^{n+1}_{h,N} - \hat\psi^{n}_{h,N}}{\sqrt{\Delta t}}\right\|^2 + \frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\|^2&\\ \le (1 + C_0\Delta t)\|\hat\psi^{n}_{h,N}\|^2 + \Delta t\left(\|\mu^{n+1}\|^2 + 4\|\underset{\sim}{\nu}^{n+1}\|^2\right),&\end{aligned}\tag{3.43}$$

where $C_0 := 4\mathrm{Wi}\, b\, |\underset{\approx}{\kappa}|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}$. Suppose that $\Delta t \le 0.5$; then

$$\begin{aligned}\|\hat\psi^{n+1}_{h,N}\|^2 + \Delta t\left\|\frac{\hat\psi^{n+1}_{h,N} - \hat\psi^{n}_{h,N}}{\sqrt{\Delta t}}\right\|^2 + \frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\|^2 &\le \frac{1+C_0\Delta t}{1-\Delta t}\|\hat\psi^{n}_{h,N}\|^2 + 2\Delta t\left(\|\mu^{n+1}\|^2 + 4\|\underset{\sim}{\nu}^{n+1}\|^2\right)\\ &\le (1 + K\Delta t)\|\hat\psi^{n}_{h,N}\|^2 + 2\Delta t\left(\|\mu^{n+1}\|^2 + 4\|\underset{\sim}{\nu}^{n+1}\|^2\right),\end{aligned}$$

where $K = 2(1 + C_0) = 2(1 + 4\mathrm{Wi}\, b\, |\underset{\approx}{\kappa}|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))})$.

Summing over $n = 0, \ldots, s-1$ gives,

$$\begin{aligned}\|\hat\psi^{s}_{h,N}\|^2 + \sum_{n=0}^{s-1}\Delta t\left\|\frac{\hat\psi^{n+1}_{h,N} - \hat\psi^{n}_{h,N}}{\sqrt{\Delta t}}\right\|^2 + \sum_{n=0}^{s-1}\frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \hat\psi^{n+1}_{h,N}\|^2&\\ \le \left\{\|\hat\psi^{0}_{h,N}\|^2 + \sum_{n=0}^{s-1} 2\Delta t\left(\|\mu^{n+1}\|^2 + 4\|\underset{\sim}{\nu}^{n+1}\|^2\right)\right\} + K\sum_{n=0}^{s-1}\Delta t\|\hat\psi^{n}_{h,N}\|^2,&\end{aligned}$$

and applying a discrete Gronwall lemma yields (3.40). □

We cannot apply an analogous argument for method II due to the absence of an equivalent one-step method. However, by combining stability results for the $\underset{\sim}{q}$-direction and $\underset{\sim}{x}$-direction methods we can establish the stability of method II, as shown in Lemma 3.5.

Lemma 3.5. *Suppose* QH2 *is satisfied and let* $\Delta t = T/N_T$, $N_T \ge 1$. *Then for* $\hat\psi^n_{h,N} \in V_h \otimes \mathcal{P}_N(D)$ *computed using alternating-direction method* II *we have*

$$\|\hat\psi^{n}_{h,N}\| \le \mathrm{e}^{c_0 n\Delta t}\|\hat\psi^{0}_{h,N}\|\tag{3.44}$$

for $1 \le n \le N_T$, *where* $c_0 := 1 + 4\mathrm{Wi}\, b\, |\underset{\approx}{\kappa}|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}$.

Proof. From the proof of Lemma 3.1 in [22], we have the following bound for (3.38) at a given quadrature point $\underset{\sim}{x}_m \in \overline{\Omega}$,

$$\|\hat\psi^{n*}(\underset{\sim}{x}_m, \cdot)\|^2_{\mathrm{L}^2(D)} \le (1 + 2c_0\Delta t)\|\hat\psi^{n}(\underset{\sim}{x}_m, \cdot)\|^2_{\mathrm{L}^2(D)}.\tag{3.45}$$

Rewriting (3.45) in terms of a basis $\{Y_1, \dots, Y_{N_D}\}$ of $\mathcal{P}_N(D)$, which, without loss of generality may be assumed to be orthogonal in the $\mathrm{L}^2(D)$ inner product, we obtain:

$$\sum_{k=1}^{N_D} \hat{\psi}_k^{n*}(\underset{\sim}{x}_m)^2 \|Y_k\|^2_{\mathrm{L}^2(D)} \le (1 + 2c_0\Delta t) \sum_{k=1}^{N_D} \hat{\psi}_k^{n}(\underset{\sim}{x}_m)^2 \|Y_k\|^2_{\mathrm{L}^2(D)}. \tag{3.46}$$

Using (3.14) to sum (3.46) for $m = 1, \dots, Q_\Omega$, and then employing (3.22), we have

$$\sum_{k=1}^{N_D} \|\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|^2_{\ell^2(\Omega)} \, \|Y_k\|^2_{\mathrm{L}^2(D)} \le (1 + 2c_0\Delta t) \sum_{k=1}^{N_D} \|\{\hat{\psi}_k^{n}(\underset{\sim}{x}_m)\}\|^2_{\ell^2(\Omega)} \, \|Y_k\|^2_{\mathrm{L}^2(D)}. \tag{3.47}$$

Since $\hat{\psi}^n_{h,N} \in V_h \otimes \mathcal{P}_N(D)$, it follows that $\hat{\psi}^n_k \in V_h$, and therefore (as observed below (3.22)) the discrete $\ell^2(\Omega)$ norm on the right-hand side above is equal to the $\mathrm{L}^2(\Omega)$ norm, so that

$$\begin{aligned} \sum_{k=1}^{N_D} \|\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|^2_{\ell^2(\Omega)} \, \|Y_k\|^2_{\mathrm{L}^2(D)} &\le (1 + 2c_0\Delta t) \sum_{k=1}^{N_D} \|\hat{\psi}_k^{n}\|^2_{\mathrm{L}^2(\Omega)} \, \|Y_k\|^2_{\mathrm{L}^2(D)} \\ &= (1 + 2c_0\Delta t) \|\hat{\psi}^n_{h,N}\|^2. \end{aligned} \tag{3.48}$$

Also, by (3.3) and (3.7), it follows easily from (3.27) that:

$$\|\hat{\psi}_k^{n+1}\|^2_{\mathrm{L}^2(\Omega)} \le \|\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|^2_{\mathrm{L}^2(\Omega)}, \tag{3.49}$$

for each k. Multiplying through by $\|Y_k\|^2_{\mathrm{L}^2(D)}$ in (3.49) and summing over $k = 1, \dots, N_D$ gives

$$\|\hat{\psi}^{n+1}_{h,N}\|^2 = \sum_{k=1}^{N_D} \|\hat{\psi}_k^{n+1}\|^2_{\mathrm{L}^2(\Omega)} \|Y_k\|^2_{\mathrm{L}^2(D)} \le \sum_{k=1}^{N_D} \|\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|^2_{\mathrm{L}^2(\Omega)} \|Y_k\|^2_{\mathrm{L}^2(D)}. \tag{3.50}$$

By taking $\{f_m\} = \{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}$ and $X = \mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\} \in V_h$ in (3.23) and applying the Cauchy–Schwarz inequality in the ℓ^2 inner product, we have

$$\begin{aligned} \|\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|^2_{\mathrm{L}^2(\Omega)} &= \Big(\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}, \{\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}(\underset{\sim}{x}_m)\}\Big)_{\ell^2(\Omega)} \\ &\le \|\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|_{\ell^2(\Omega)} \, \|\{\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}(\underset{\sim}{x}_m)\}\|_{\ell^2(\Omega)} \\ &= \|\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|_{\ell^2(\Omega)} \, \|\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|_{\mathrm{L}^2(\Omega)}, \end{aligned}$$

and therefore,

$$\|\mathcal{R}\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|_{\mathrm{L}^2(\Omega)} \le \|\{\hat{\psi}_k^{n*}(\underset{\sim}{x}_m)\}\|_{\ell^2(\Omega)}. \tag{3.51}$$

Combining (3.48), (3.50) and (3.51), gives,

$$\|\hat{\psi}^{n+1}_{h,N}\|^2 \le (1 + 2c_0\Delta t)\|\hat{\psi}^n_{h,N}\|^2, \tag{3.52}$$

from which (3.44) follows easily on noting that $1 + 2c_0\Delta t \le \mathrm{e}^{2c_0\Delta t}$. □

Remark 3.6. The argument in Lemma 3.5 can also be applied to method I and hence it follows that method I is stable when only QH2 is satisfied.

3.5. Discrete spaces and approximation results

In Section 3.6, we will derive a convergence estimate for method I. The argument relies on approximation results on $\Omega \times D$, hence the purpose of this section is to obtain the required results.

Our approach is to deduce the appropriate results on Ω and D separately, and then it is straightforward to combine them. The approximation results for the finite element method on Ω are standard and hence we defer their discussion until later.

Spectral bases on D. In [22], we considered a Galerkin spectral method on D in the case that $d = 2$ in detail. Our approach was to perform a polar-coordinate transformation of D into the rectangle $(r, \theta) \in R := (0,1) \times (0, 2\pi)$ using $\underset{\sim}{q} = (q_1, q_2) = (\sqrt{b}\, r \cos\theta, \sqrt{b}\, r \sin\theta)$. Since $\mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D; M)$, we sought a solution $\hat{\psi} \in \mathrm{H}^1_0(D)$; also we used the notation $\tilde{\psi}(r, \theta) := \hat{\psi}(q_1, q_2)$. It was proved in Lemma 5.2 of [22] that $\tilde{\psi}$ can be decomposed in polar coordinates as follows:

$$\tilde{\psi}(r,\theta) = \tilde{\psi}_1(r) + r\tilde{\psi}_2(r,\theta), \qquad (r,\theta) \in R = (0,1) \times (0, 2\pi). \tag{3.53}$$

This is in fact a first-order version of the pole condition [15] for spectral methods in polar coordinates, see Section 7 of [22] for a detailed discussion.

We performed a detailed analysis of the approximation properties of the space spanned by the basis $\mathcal{A}$, where $\mathcal{A}$ is defined as $\mathcal{A} := \mathcal{A}_1 \cup \mathcal{A}_2$ and:

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

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 only use even trigonometric modes in $\mathcal{A}$; this is because $\hat{\psi}(\underset{\sim}{q}) = \hat{\psi}(-\underset{\sim}{q})$ (*cf.* Rem. 7.1 in [22]). Note that the structure of $\mathcal{A}$ mimics the decomposition (3.53).

We also considered a second basis, basis $\mathcal{B}$, which is essentially the basis proposed by Matsushima and Marcus [29] and Verkley [34], and satisfies the full pole condition; see [22] for more details.

In the present work, we are also interested in obtaining numerical results when $D \subset \mathbb{R}^3$, and therefore we shall now introduce a basis, referred to as basis $\mathcal{C}$, that is appropriate when $d = 3$. In this case we use the following spherical coordinate change of variables:

$$\underset{\sim}{q} = (\sqrt{b}r\cos\theta\sin\phi\,,\ \sqrt{b}r\sin\theta\sin\phi\,,\ \sqrt{b}r\cos\phi), \qquad (r,\theta,\phi) \in R := (0,1) \times (0,2\pi) \times (0,\pi).$$

Following Chauvière and Lozinski [10], we choose each of our basis functions to be a product of a spherical harmonic in (θ, ϕ) and a polynomial in r (see the definition of basis $\mathcal{C}$ below). Discretisations of this type have also been considered in the recent paper by Huang and Guo [18].

Letting $\tilde{g}(r,\theta,\phi) := g(q_1, q_2, q_3)$, it follows that

$$\|g\|^2_{\mathrm{H}^1(D)} = \int_R r^2 \sin\phi \left(b^{3/2}|\tilde{g}|^2 + \sqrt{b}\left|\frac{\partial \tilde{g}}{\partial r}\right|^2 + \sqrt{b}\,\frac{1}{r^2}\left|\frac{\partial \tilde{g}}{\partial \phi}\right|^2 + \sqrt{b}\,\frac{1}{r^2\sin^2\phi}\left|\frac{\partial \tilde{g}}{\partial \theta}\right|^2 \right) \mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi =: \|\tilde{g}\|^2_{\tilde{\mathrm{H}}^1(R)},$$

and following the approach in Section 5 of [22] for the case of $d = 2$, we define $\tilde{\mathrm{H}}^1(R)$ as:

$$\tilde{\mathrm{H}}^1(R) := \{\tilde{f} \in \mathrm{L}^2_{\mathrm{loc}}(R) : \tilde{f}(r,\cdot,\phi) \in \mathrm{H}^1_p(0,2\pi) \text{ for a.e. } (r,\phi) \in (0,1) \times (0,\pi) \text{ and } \|\tilde{f}\|_{\tilde{\mathrm{H}}^1(R)} < \infty\},$$

where the periodic Sobolev space $\mathrm{H}^t_p(0,2\pi)$ is given by

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

Also, let $\tilde{\mathrm{H}}^1_0(R)$ denote the subspace of all $\tilde{g} \in \tilde{\mathrm{H}}^1_0(R)$ such that $\tilde{g}(1,\cdot,\cdot) = 0$.

We denote the spherical harmonics by $S_{l,m} : (\theta,\phi) \mapsto S_{l,m}(\theta,\phi) \in \mathbb{R}$. They are the solutions of the equation

$$\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\,\frac{\partial}{\partial\phi}S_{l,m}(\theta,\phi)\right) + \frac{1}{\sin^2\phi}\frac{\partial^2}{\partial\theta^2}S_{l,m}(\theta,\phi) + l(l+1)S_{l,m}(\theta,\phi) = 0, \tag{3.54}$$

for $(\theta,\phi) \in (0,2\pi)\times(0,\pi)$, where (3.54) is the angular part of the Laplace equation in spherical coordinates; it is well known that the solutions of (3.54) are of the form, $S_{l,m}(\theta,\phi) = C(l,m)\,P_l^m(\cos\phi)\mathrm{e}^{im\theta}$, for $l \in \mathbb{Z}_{\geq 0}$, $|m| \leq l$, where P_l^m denotes an associated Legendre function and $C(l,m)$ is a normalisation constant. Also, the (appropriately normalised) spherical harmonics satisfy the following orthogonality property:

$$\int_0^{2\pi}\int_0^{\pi} S_{l_1,m_1}(\theta,\phi)\,\overline{S}_{l_2,m_2}(\theta,\phi)\sin\phi\,\mathrm{d}\theta\,\mathrm{d}\phi = \delta_{m_1,m_2}\delta_{l_1,l_2}, \tag{3.55}$$

where the δ_{ij} are Kronecker deltas and the overline notation denotes complex conjugation.

The next lemma shall motivate our definition of a spectral basis in the case of $d = 3$.

Lemma 3.7. *Let* $\tilde{g}(r,\theta) = \sum_{l=0}^{N_{\mathrm{sph}}}\sum_{|m|\leq l}\tilde{\gamma}_l^m(r)S_{l,m}(\theta,\phi)$, $N_{\mathrm{sph}} \in \mathbb{Z}_{\geq 0}$, $\tilde{\gamma}_0^0 \in \mathrm{H}^1_{r^2}(0,1)$, *where* $\mathrm{H}^1_{r^2}(0,1)$ *is the* r^2*-weighted* H^1*-space and*

$$\tilde{\gamma}_l^m \in \mathrm{H}^1(0,1;1,r^2) := \left\{\tilde{f} \in \mathrm{H}^1_{\mathrm{loc}}(0,1)\ :\ \int_0^1 \left(|\tilde{f}(r)|^2 + r^2|\tilde{f}'(r)|^2\right)\mathrm{d}r < \infty\right\},$$

for $l > 0$ *and* $|m| \leq l$*; then* $\tilde{g} \in \tilde{\mathrm{H}}^1(R)$.

Proof. Periodicity of $\tilde{g}$ in θ follows directly from the definition of the spherical harmonics; hence it only remains to verify that $\|\tilde{g}\|_{\tilde{\mathrm{H}}^1(R)} < \infty$.

Integrating by parts in θ and ϕ (which is valid for spherical harmonics), we obtain:

$$\begin{aligned}\|\tilde{g}\|^2_{\tilde{\mathrm{H}}^1(R)} \quad = \quad & \int_R r^2\sin\phi\left(b^{3/2}|\tilde{g}|^2 + \sqrt{b}\left|\frac{\partial\tilde{g}}{\partial r}\right|^2\right)\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi \\ & -\sqrt{b}\int_R \sin\phi\,\tilde{g}\left(\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial\tilde{g}}{\partial\phi}\right) + \frac{1}{\sin^2\phi}\frac{\partial^2\tilde{g}}{\partial\theta^2}\right)\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi,\end{aligned} \tag{3.56}$$

where the boundary conditions vanish due to periodicity. Substituting the series expression of $\tilde{g}$ into (3.56) and using (3.54) and (3.55), we get:

$$\begin{aligned}\|\tilde{g}\|^2_{\tilde{\mathrm{H}}^1(R)} &= \sum_{l=0}^{N_{\mathrm{sph}}}\sum_{|m|\leq l}\int_0^1 r^2\left\{b^{3/2}|\tilde{\gamma}_l^m(r)|^2\,\mathrm{d}r + \sqrt{b}\left|\frac{\mathrm{d}\tilde{\gamma}_l^m}{\mathrm{d}r}\right|^2\right\}\mathrm{d}r \\ &+\sqrt{b}\int_R\left\{\sum_{l_1=0}^{N_{\mathrm{sph}}}\sum_{|m_1|\leq l_1}\tilde{\gamma}_{l_1}^{m_1}(r)S_{l_1,m_1}(\theta,\phi)\right\}\left\{\sum_{l_2=0}^{N_{\mathrm{sph}}}\sum_{|m_2|\leq l_2} l_2(l_2+1)\,\tilde{\gamma}_{l_2}^{m_2}(r)\overline{S}_{l_2,m_2}(\theta,\phi)\right\}\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}r \\ &= \sum_{l=0}^{N_{\mathrm{sph}}}\sum_{|m|\leq l}\int_0^1\left\{b^{3/2}r^2|\tilde{\gamma}_l^m(r)|^2 + \sqrt{b}\,r^2\left|\frac{\mathrm{d}}{\mathrm{d}r}\tilde{\gamma}_l^m(r)\right|^2 + \sqrt{b}\,l(l+1)|\tilde{\gamma}_l^m(r)|^2\right\}\mathrm{d}r.\end{aligned} \tag{3.57}$$

By the hypotheses on the $\tilde{\gamma}_l^m$, it follows that $\|\tilde{g}\|_{\tilde{\mathrm{H}}^1(R)}$ is finite. □

Note that the $\tilde{\gamma}_l^m$ in Lemma 3.7 need not be bounded on $(0,1)$: for example, $r^{-1/4} \in \mathrm{H}^1_{r^2}(0,1) \cap \mathrm{H}^1(0,1;1,r^2)$. It will be convenient from now on to use the real and imaginary parts of the spherical harmonics:

$$S^i_{l,m}(\theta,\phi) := C(l,m)\, P_l^m(\cos\phi)((1-i)\cos(m\theta) + i\sin(m\theta)), \tag{3.58}$$

where now $0 \le l \le N_{\mathrm{sph}},\ i \in \{0,1\}$, and $i \le m \le l$. In this section, we consider basis functions of the form:

$$Y^{ik}_{lm}(r,\theta,\phi) := (1-r)Q_k(r)S^i_{l,m}(\theta,\phi). \tag{3.59}$$

Here Q_k is a polynomial of degree k, $0 \le k \le N_r - 1$, so that $(1-r)Q_k \subset \mathbb{P}_{N_r,0}(0,1)$, where $\mathbb{P}_{N_r,0}(0,1)$ denotes the set of polynomials on $(0,1)$ that vanish at $r=1$. Since $\mathbb{P}_{N_r,0}(0,1) \subset \mathrm{H}^1_{r^2}(0,1) \cap \mathrm{H}^1(0,1;1,r^2)$, it follows from Lemma 3.7 that any finite linear combination of functions of the form (3.59) is contained in $\tilde{\mathrm{H}}^1_0(R)$. This is a simpler situation than in two dimensions, since now we do not need to impose a specialised decomposition (*cf.* (3.53)) in order to guarantee inclusion in $\tilde{\mathrm{H}}^1(R)$.

Basis $\mathcal{C}$ will be introduced below, but first we consider the implications of the symmetry property $\hat{\psi}(\underset{\sim}{q}) = \hat{\psi}(-\underset{\sim}{q})$ mentioned above. In the present context, the symmetry condition we wish to impose is that for any Y^{ik}_{lm}, we have $Y^{ik}_{lm}(r,\theta,\phi) = Y^{ik}_{lm}(r,\theta+\pi,\pi-\phi)$. This, in turn, requires that $S^i_{l,m}(\theta,\phi) = S^i_{l,m}(\theta+\pi,\pi-\phi)$. Noting that $S^i_{l,m}(\theta,\phi) = P^m_l(\cos\phi)((1-i)\cos(m\theta)+i\sin(m\theta))$, and $S^i_{l,m}(\theta+\pi,\pi-\phi) = (-1)^m P^m_l(-\cos\phi)((1-i)\cos(m\theta)+ i\sin(m\theta))$, it follows that we can only use associated Legendre functions for which $P^m_l(x) = (-1)^m P^m_l(-x)$, for all $x \in [-1,1]$. Since the associated Legendre functions are defined as:

$$P^m_l(x) = (-1)^m(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}(P_l(x)),$$

where $P_l(x)$ is a Legendre polynomial of degree l (for which $P_l(x) = (-1)^l P_l(-x)$), it follows that the required symmetry condition is satisfied if, and only if, l is an even number (for any $m = 0,\dots,l$).

With these considerations in mind, we define basis $\mathcal{C}$ as follows:

$$\mathcal{C} := \{Y^{ik}_{lm}\ :\ 0 \le k \le N_r - 1,\ i \in \{0,1\},\ l \in \{0,2,4,\dots,N_{\mathrm{sph}}\} \text{ and } i \le m \le l\}\cdot$$

Remark 3.8. In [10], Chauvière and Lozinski restricted their attention to two-dimensional macroscopic velocity fields, in which case a more restrictive symmetry condition was appropriate, *i.e.* that $\psi(r,\theta,\phi) = \psi(r,\theta+\pi,\phi)$, and hence they only considered spherical harmonics for which both l and m were even numbers. Compared to the more general symmetry condition considered above, the condition of Chauvière and Lozinski leads to a reduction in computational effort because, for a given N_{sph}, fewer basis functions are used since the spherical harmonics with odd m are discarded, and also it is only necessary to consider $\theta \in (0,\pi)$. In this paper, however, we are interested in treating the case in which the macroscopic velocity field can be three-dimensional, and therefore we require the symmetry condition identified above.

Approximation results on D. In Section 5 of [22], we derived a number of approximation results for span($\mathcal{A}$) in the case of $d = 2$. We will use these results in the next section, so we summarise them here.

We must first recall several preliminary definitions from [22]. Let $R := (0,1) \times (0,2\pi)$ and let $\tilde{\mathrm{H}}^1(R)$ and $\tilde{\mathrm{H}}^1_0(R)$ be defined analogously to the corresponding spaces defined above in the case of $d = 3$. We shall need weighted Sobolev spaces[7] of the form $\mathrm{H}^{s,t}_r(R) := \mathrm{H}^s_r(0,1;\mathrm{H}^t_p(0,2\pi))$ where $\mathrm{H}^{s,t}_r(R)$ is equipped (for non-negative integers s and t) with the norm $\|\cdot\|_{\mathrm{H}^{s,t}_r(R)}$ defined by

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

[7] The weight r arises from the change of measure due to the polar coordinate transformation.

Similarly, for integers $s \geq 1$ and $t \geq 0$, we define $\mathrm{H}^{s,t}_{r,0}(R) := \mathrm{H}^{s}_{r,0}(0,1;\mathrm{H}^{t}_{p}(0,2\pi))$, where $\mathrm{H}^{s}_{r,0}(0,1) := \mathrm{H}^{s}_{r}(0,1) \cap \mathrm{H}^{1}_{r,0}(0,1)$.

Finally, we define the space $\mathcal{H}^{k+1,l+1}(D)$, with $k, l \geq 1$, where

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

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,

$$\|g\|_{\mathcal{H}^k_r(D)} := \left(\|\tilde{g}_1\|^2_{\mathrm{H}^k_r(0,1)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k,0}_r(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{k-1,1}_r(R)}\right)^{\frac{1}{2}},$$

and

$$\|g\|_{\mathcal{H}^l_\theta(D)} := \left(\|\tilde{g}_2\|^2_{\mathrm{H}^{0,l}_r(R)} + \|\tilde{g}_2\|^2_{\mathrm{H}^{1,l-1}_r(R)}\right)^{\frac{1}{2}}.$$

Also, note that $\mathcal{H}^{k,l}(D) \subset \mathrm{H}^1_0(D) \subset \mathrm{H}^1_0(D;M)$ for $k, l \geq 1$.

In Section 5 of [22], we defined a projection operator $\Pi_q : \mathcal{H}^{1,1}(D) \to \mathcal{P}_N(D)$ and established the following two optimal-order approximation results for all $\hat{\psi} \in \mathcal{H}^{k,l}(D)$, with $k,\ l \geq 1$:

$$\|\hat{\psi} - \Pi_q\hat{\psi}\|_{\mathrm{H}^1_0(D;M)} \leq C_1 N_r^{-k}\|\hat{\psi}\|_{\mathcal{H}^{k+1}_r(D)} \; + \; C_2 N_\theta^{-l}\|\hat{\psi}\|_{\mathcal{H}^{l+1}_\theta(D)}, \tag{3.60}$$

and

$$\|\hat{\psi} - \Pi_q\hat{\psi}\|_{\mathrm{L}^2(D)} \leq C_1 N_r^{-k}\|\hat{\psi}\|_{\mathcal{H}^{k}_r(D)} \; + \; C_2 N_\theta^{-l}\|\hat{\psi}\|_{\mathcal{H}^{l}_\theta(D)}. \tag{3.61}$$

The results (3.60) and (3.61) are based on the assumption that $D \subset \mathbb{R}^2$ and that we use basis $\mathcal{A}$ for the $\underset{\sim}{q}$-direction spectral method. It would be possible to derive analogous results for basis $\mathcal{B}$ when $d = 2$, or for basis $\mathcal{C}$ when $d = 3$ (indeed Huang and Guo [18] recently derived approximation results for a spectral method on the unit ball in $\mathbb{R}^3$, which could be used to obtain results analogous to (3.60) and (3.61) for basis $\mathcal{C}$) but for the sake of brevity we do not consider these topics here.

Approximation results on Ω. In the $\underset{\sim}{x}$-direction we consider a quasi-interpolation operator, $\mathcal{I}_x : \mathrm{L}^1(\Omega) \to V_h$; we refer to Section 4.8 of [7] for the details of the definition of this operator (alternatively, see [12] or [33]).

The following result holds for $\mathcal{I}_x$ (*cf*. Thm. 4.8.12 in [7]).

Theorem 3.9. *Suppose that $\mathcal{T}_h$ is nondegenerate in the sense that there exists $\rho > 0$ such that for all $K \in \mathcal{T}_h$, $\mathrm{diam}(B_K) \geq \rho\,\mathrm{diam}(K)$, where B_K is the largest ball contained in K. Suppose also that the set of shape functions for each element $K \in \mathcal{T}_h$ contains all polynomials of degree less than m. Then, there exists a positive constant C such that*

$$\left(\sum_{K\in\mathcal{T}_h} h_K^{p(s-k)}\|v - \mathcal{I}_x\, v\|^p_{\mathrm{W}^{s,p}(K)}\right)^{1/p} \leq C|v|_{\mathrm{W}^{k,p}(\Omega)},$$

for all $v \in \mathrm{W}^{k,p}(\Omega)$, $0 \leq k \leq m$, $1 \leq p \leq \infty$, $0 \leq s \leq k$, where $h_K := \mathrm{diam}(K)$.

Corollary 3.10 (*cf*. Cor. 4.8.15 in [7])**.** *Setting $s = k$ in Theorem 3.9, it follows that there exists a positive constant C such that*

$$\|\mathcal{I}_x\, v\|_{\mathrm{W}^{k,p}(\Omega)} \leq C|v|_{\mathrm{W}^{k,p}(\Omega)} \qquad \forall v \in \mathrm{W}^{k,p}(\Omega), \tag{3.62}$$

for $0 \leq k \leq m$, where m is as in Theorem 3.9, and $1 \leq p \leq \infty$. Also, letting $h = \max_{K\in\mathcal{T}_h}\mathrm{diam}(K)$ in Theorem 3.9, we obtain

$$\|v - \mathcal{I}_x v\|_{\mathrm{W}^{s,p}(\Omega)} \leq Ch^{k-s}|v|_{\mathrm{W}^{k,p}(\Omega)}, \tag{3.63}$$

for $0 \leq s \leq k$, $0 \leq k \leq m$, and m, p as in (3.62).

Approximation results on $\Omega\times D$. Suppose that $d=2$ and that we are using basis $\mathcal{A}$ for the $\underset{\sim}{q}$-direction spectral method so that $\mathcal{P}_N(D)$ is span($\mathcal{A}$) mapped from R to D. Let the projection operator $\Pi:\mathrm{L}^1(\Omega;\mathcal{H}^{1,1}(D))\rightarrow V_h\otimes\mathcal{P}_N(D)$ be defined as $\Pi:=\mathcal{I}_x\,\Pi_q=\Pi_q\,\mathcal{I}_x$, so that $\eta:=\hat{\psi}-\Pi\hat{\psi}$. With the approximation properties listed above for Π_q and $\mathcal{I}_x$, the following optimal order bounds can be derived for $0\le s\le m$, with m as above, and any $k,l\ge 1$ (see Sect. 3.6 of [21]):

$$\|\eta^0\|\le Ch^s\|\hat{\psi}_0\|_{\mathrm{H}^s(\Omega;\mathrm{L}^2(D))}+C_1N_r^{-k}\|\hat{\psi}_0\|_{\mathrm{L}^2(\Omega;\mathcal{H}^k_r(D))}\;+\;C_2N_\theta^{-l}\|\hat{\psi}_0\|_{\mathrm{L}^2(\Omega;\mathcal{H}^l_\theta(D))},$$

$$\begin{aligned}\|\eta\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\quad&\le\quad Ch^s\left\|\hat{\psi}\right\|_{\ell^2(0,t^n;\mathrm{H}^s(\Omega;\mathrm{L}^2(D)))}+C_1N_r^{-k}\left\|\hat{\psi}\right\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^k_r(D)))}\\&\quad+C_2N_\theta^{-l}\left\|\hat{\psi}\right\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^l_\theta(D)))},\end{aligned}$$

$$\begin{aligned}\left\|\frac{\partial\eta}{\partial t}\right\|_{\mathrm{L}^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\quad&\le\quad Ch^s\left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,t^n;\mathrm{H}^s(\Omega;\mathrm{L}^2(D)))}+C_1N_r^{-k}\left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^k_r(D)))}\\&\quad+C_2N_\theta^{-l}\left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^l_\theta(D)))},\end{aligned}$$

$$\begin{aligned}\|\underset{\sim}{\nabla}_x\eta\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\quad&\le\quad Ch^s\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^{s+1}(\Omega;\mathrm{L}^2(D)))}\\&\quad+C_1N_r^{-k}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^1(\Omega;\mathcal{H}^k_r(D)))}+C_2N_\theta^{-l}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^1(\Omega;\mathcal{H}^l_\theta(D)))},\end{aligned}$$

$$\begin{aligned}\|\underset{\sim}{\nabla}_M\eta\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\quad&\le\quad Ch^s\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^s(\Omega;\mathrm{H}^1_0(D;M)))}\\&\quad+C_1N_r^{-k}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^{k+1}_r(D)))}+C_2N_\theta^{-l}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega;\mathcal{H}^{l+1}_\theta(D)))},\end{aligned}$$

$$\begin{aligned}\|\underset{\sim}{\nabla}_x\underset{\sim}{\nabla}_M\eta\|_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\quad&\le\quad Ch^s\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^{s+1}(\Omega;\mathrm{H}^1_0(D;M)))}\\&\quad+C_1N_r^{-k}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^1(\Omega;\mathcal{H}^{k+1}_r(D)))}+C_2N_\theta^{-l}\|\hat{\psi}\|_{\ell^2(0,t^n;\mathrm{H}^1(\Omega;\mathcal{H}^{l+1}_\theta(D)))}.\end{aligned}$$

These results will be used in the convergence argument in the next section.

3.6. Convergence analysis for method I

In this section we outline an *a priori* convergence argument for method I using the equivalent one-step scheme (3.30) and Lemma 3.4. The argument presented here is analogous to the approach in Section 4 of [22] and therefore for the sake of brevity we omit many details; see Section 3.5 of [21] for more details. Note that we need access to an equivalent one-step formulation for this argument and therefore we only consider the convergence analysis of method I.

Let $\hat{\psi}(\cdot,\cdot,t)$ be the weak solution of (3.5), (3.6) at time $t\in(0,T)$. To simplify the notation, we write $\hat{\psi}(t):=\hat{\psi}(\cdot,\cdot,t)$ throughout the rest of this section. Define

$$e^n_{h,N}:=\hat{\psi}(t^n)-\hat{\psi}^n_{h,N}=(\hat{\psi}(t^n)-\Pi\hat{\psi}(t^n))+(\Pi\hat{\psi}(t^n)-\hat{\psi}^n_{h,N})=:\eta^n+\xi^n,$$

where $\Pi:\mathrm{L}^1(\Omega;\mathcal{H}^{1,1}(D))\rightarrow V_h\otimes\mathcal{P}_N(D)$ as in Section 3.5 (although the specific definition of Π and $\mathcal{P}_N(D)$ are irrelevant at this point in the convergence argument).

Noting that $\xi^n \in V_h \otimes \mathcal{P}_N(D)$, we apply the equivalent one-step formulation for method I, (3.30), to $\xi^n = \hat{\psi}(t^n) - \hat{\psi}^n_{h,N} - \eta^n$ and set $\zeta = \xi^{n+1}$, to obtain:

$$\begin{aligned}
&\left(\frac{\xi^{n+1}-\xi^n}{\Delta t}, \xi^{n+1}\right) + \left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \xi^{n+1}, \xi^{n+1}\right) + \frac{1}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \xi^{n+1}\|^2 \\
&\qquad + \frac{\Delta t}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \xi^{n+1}), \underset{\sim}{\nabla}_M \xi^{n+1}\right) - \left(\underset{\approx}{\kappa}^n \underset{\sim}{q}\xi^n, \underset{\sim}{\nabla}_M \xi^{n+1}\right) \\
&= \left(\frac{\hat{\psi}(t^{n+1})-\hat{\psi}(t^n)}{\Delta t}, \xi^{n+1}\right) + \left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \hat{\psi}(t^{n+1}), \xi^{n+1}\right) + \frac{1}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M \hat{\psi}(t^{n+1}), \underset{\sim}{\nabla}_M \xi^{n+1}\right) \\
&\qquad + \frac{\Delta t}{2}\left(\underset{\sim}{\nabla}_M(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \hat{\psi}(t^{n+1})), \underset{\sim}{\nabla}_M \xi^{n+1}\right) - \left(\underset{\approx}{\kappa}^n \underset{\sim}{q}\hat{\psi}(t^n), \underset{\sim}{\nabla}_M \xi^{n+1}\right) \\
&\quad - \left(\frac{\eta^{n+1}-\eta^n}{\Delta t}, \xi^{n+1}\right) - \left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \eta^{n+1}, \xi^{n+1}\right) - \frac{1}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M \eta^{n+1}, \underset{\sim}{\nabla}_M \xi^{n+1}\right) \\
&\qquad - \frac{\Delta t}{2\mathrm{Wi}}\left(\underset{\sim}{\nabla}_M(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \eta^{n+1}), \underset{\sim}{\nabla}_M \xi^{n+1}\right) + \left(\underset{\approx}{\kappa}^n \underset{\sim}{q}\eta^n, \underset{\sim}{\nabla}_M \xi^{n+1}\right),
\end{aligned} \tag{3.64}$$

where the terms containing $\hat{\psi}^n_{h,N}$ and $\hat{\psi}^{n+1}_{h,N}$ vanish since $\hat{\psi}_{h,N}$ satisfies (3.30). We apply Lemma 3.4 to this equation to obtain:

$$\|\xi^n\|^2 + \sum_{m=0}^{n-1} \frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \xi^{m+1}\|^2 \le \mathrm{e}^{Kn\Delta t}\left\{\|\xi^0\|^2 + \sum_{m=0}^{n-1} 2\Delta t\left(\|\mu^{m+1}\|^2 + 4\|\underset{\sim}{\nu}^{m+1}\|^2\right)\right\}, \tag{3.65}$$

where

$$\mu^{n+1} := \frac{\hat{\psi}(t^{n+1})-\hat{\psi}(t^n)}{\Delta t} - \frac{\partial\hat{\psi}}{\partial t}(t^{n+1}) - \frac{\eta^{n+1}-\eta^n}{\Delta t} - \underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \eta^{n+1}, \tag{3.66}$$

$$\begin{aligned}
\underset{\sim}{\nu}^{n+1} := {} & \frac{\Delta t}{2\mathrm{Wi}}\underset{\sim}{\nabla}_M(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \hat{\psi}(t^{n+1})) + \underset{\sim}{K}_1 + \underset{\sim}{K}_2 - \underset{\sim}{K}_3 - \frac{1}{2\mathrm{Wi}}\underset{\sim}{\nabla}_M \eta^{n+1} \\
& - \frac{\Delta t}{2\mathrm{Wi}}\underset{\sim}{\nabla}_M(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x \eta^{n+1}) + \underset{\approx}{\kappa}^n \underset{\sim}{q}\eta^n,
\end{aligned} \tag{3.67}$$

and $K_1 := \left(\int_{t^n}^{t^{n+1}} \frac{\partial\underset{\approx}{\kappa}}{\partial t}\,\mathrm{d}t\right)\underset{\sim}{q}\hat{\psi}(t^n)$, $K_2 := \underset{\approx}{\kappa}^{n+1}\underset{\sim}{q}\left(\int_{t^n}^{t^{n+1}} \frac{\partial\hat{\psi}}{\partial t}\,\mathrm{d}t\right)$, $K_3 := \left(\int_{t^n}^{t^{n+1}} \frac{\partial\underset{\approx}{\kappa}}{\partial t}\,\mathrm{d}t\right)\underset{\sim}{q}\left(\int_{t^n}^{t^{n+1}} \frac{\partial\hat{\psi}}{\partial t}\,\mathrm{d}t\right)$.

Bounding the terms on the right-hand side of (3.65) leads to the following result:

$$\begin{aligned}
&\|\xi^n\|^2 + \sum_{m=0}^{n-1} \frac{\Delta t}{2\mathrm{Wi}}\|\underset{\sim}{\nabla}_M \xi^{m+1}\|^2 \le \mathrm{e}^{Kn\Delta t}\Bigg\{\|\eta^0\|^2 + 6\Delta t^2\left\|\frac{\partial^2\hat{\psi}}{\partial t^2}\right\|^2_{\mathrm{L}^2(0,t^n;\mathrm{L}^2(\Omega\times D))} + 6\left\|\frac{\partial\eta}{\partial t}\right\|^2_{\mathrm{L}^2(0,t^n;\mathrm{L}^2(\Omega\times D))} \\
&\quad + 6\|\underset{\sim}{u}\|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}\|\underset{\sim}{\nabla}_x \eta\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))} \\
&\quad + \frac{14}{\mathrm{Wi}^2}\,\Delta t^2\|u\|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}\left(\|\underset{\sim}{\nabla}_x\underset{\sim}{\nabla}_M \hat{\psi}\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))} + \|\underset{\sim}{\nabla}_x\underset{\sim}{\nabla}_M \eta\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))}\right) \\
&\quad + \frac{14}{\mathrm{Wi}^2}\|\underset{\sim}{\nabla}_M \eta\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))} + 56\,b\|\underset{\approx}{\kappa}\|^2_{\mathrm{L}^\infty(0,t^n;\mathrm{L}^\infty(\Omega))}\,\|\eta\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))} \\
&\quad + 56\,b\Delta t^2\left\|\frac{\partial\underset{\approx}{\kappa}}{\partial t}\right\|^2_{\mathrm{L}^\infty(0,T;\mathrm{L}^\infty(\Omega))}\,\left\|\hat{\psi}\right\|^2_{\ell^2(0,t^n;\mathrm{L}^2(\Omega\times D))} \\
&\quad + 56\,\Delta t^2\,b\|\underset{\approx}{\kappa}\|^2_{\mathrm{W}^{1,\infty}(0,T;\mathrm{L}^\infty(\Omega))}\left\|\frac{\partial\hat{\psi}}{\partial t}\right\|^2_{\mathrm{L}^2(0,t^n;L^2(\Omega\times D))}\Bigg\}.
\end{aligned} \tag{3.68}$$

Note that up until this point the convergence argument is independent of the dimension, the choice of Π and the choice of $\mathcal{P}_N(D)$. However, we now require the approximation results on $\Omega \times D$ from Section 3.5 in order to bound the terms containing η. Hence, we let $d = 2$ and define $\mathcal{P}_N(D)$ be span$(\mathcal{A})$ mapped from R to D. Substituting the appropriate bounds into (3.68) and applying the triangle inequality, we obtain:

$$\begin{aligned}
&\|\hat{\psi} - \hat{\psi}_{h,N}\|_{\ell^\infty(0,T;\mathrm{L}^2(\Omega\times D))} + \|\underset{\sim}{\nabla}_M(\hat{\psi} - \hat{\psi}_{h,N})\|_{\ell^2(0,T;\mathrm{L}^2(\Omega\times D))} \\
&\quad \le C_1 h^s\Big(\|\hat{\psi}\|_{\ell^\infty(0,T;\mathrm{H}^s(\Omega;\mathrm{L}^2(D)))} + \left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathrm{H}^s(\Omega;\mathrm{L}^2(D)))} + \left\|\hat{\psi}\right\|_{\ell^2(0,T;\mathrm{H}^s(\Omega;\mathrm{H}^1_0(D;M)))} \\
&\qquad\quad + \left\|\hat{\psi}\right\|_{\ell^2(0,T;\mathrm{H}^{s+1}(\Omega;\mathrm{L}^2(D)))}\Big) \\
&\quad + C_2 N_r^{-k}\Big(\|\hat{\psi}\|_{\ell^\infty(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^k_r(D)))} + \left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^k_r(D)))} + \|\hat{\psi}\|_{\ell^2(0,T;\mathrm{H}^1(\Omega;\mathcal{H}^k_r(D)))} \\
&\qquad\quad + \|\hat{\psi}\|_{\ell^2(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^{k+1}_r(D)))}\Big) \\
&\quad + C_3 N_\theta^{-l}\Big(\|\hat{\psi}\|_{\ell^\infty(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^l_\theta(D)))} + \left\|\frac{\partial\hat{\psi}}{\partial t}\right\|_{\mathrm{L}^2(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^l_\theta(D)))} + \|\hat{\psi}\|_{\ell^2(0,T;\mathrm{H}^1(\Omega;\mathcal{H}^l_\theta(D)))} \\
&\qquad\quad + \left\|\hat{\psi}\right\|_{\ell^2(0,T;\mathrm{L}^2(\Omega;\mathcal{H}^{l+1}_\theta(D)))}\Big) \\
&\quad + C_4\Delta t\Big(\left\|\hat{\psi}\right\|_{\ell^2(0,T;\mathrm{L}^2(\Omega\times D))} + \left\|\hat{\psi}\right\|_{\mathrm{H}^2(0,T;\mathrm{L}^2(\Omega\times D))} + \|\underset{\sim}{\nabla}_x\underset{\sim}{\nabla}_M\hat{\psi}\|_{\ell^2(0,T;\mathrm{L}^2(\Omega\times D))} \\
&\qquad\quad + N_r^{-k}\|\hat{\psi}\|_{\ell^2(0,T;\mathrm{H}^1(\Omega;\mathcal{H}^{k+1}_r(D)))} + N_\theta^{-l}\|\hat{\psi}\|_{\ell^2(0,T;\mathrm{H}^1(\Omega;\mathcal{H}^{l+1}_\theta(D)))}\Big),
\end{aligned} \tag{3.69}$$

for $0 \le s \le m$, m as in Theorem 3.9, and $k, l \ge 1$. Note that an obvious difference between (3.69) and the corresponding error estimate for the $\underset{\sim}{q}$-direction method in [22] is that in (3.69) we require $\left\|\underset{\sim}{\nabla}_x\underset{\sim}{\nabla}_M\hat{\psi}\right\|_{\ell^2(0,T;\mathrm{L}^2(\Omega\times D))} < \infty$. This regularity condition is necessitated by the presence of the cross term $\left(\underset{\sim}{\nabla}_M\left(\underset{\sim}{u}\cdot\underset{\sim}{\nabla}_x\hat{\psi}^{n+1}_{h,N}\right), \underset{\sim}{\nabla}_M\zeta\right)$ in (3.30), which in turn arises from the structure of the alternating-direction method.

Remark 3.11. Looking at (3.69), it could be argued that there is a mismatch between the convergence rates of the finite element method in Ω and the spectral method in D, in the sense that the spectral method will generally be far more accurate. This is a reasonable point, but we believe that in practice the numerical method analysed here is appropriate. First of all, while in general an h-version finite element scheme will have a low-order convergence rate, its flexibility is invaluable when it comes to meshing physical space domains that may be complicated. Moreover, we do not have a diffusion operator in the $\underset{\sim}{x}$-direction, so it is not obvious that $\hat{\psi}$ will be highly smooth on Ω.

Nevertheless, it is certainly also reasonable to use a higher-order method for solving the transport equation in physical space, for example, Chauvière and Lozinski used a spectral element method for this purpose in [10,11]. Note that the analysis in this section would carry over essentially unchanged if we replaced the finite element discretisation of (3.27) by a spectral element method.

On the other hand, the $\underset{\sim}{q}$-direction is much better suited to the use of a high-order method since D is always a ball in $\mathbb{R}^d$, and, as seen in [22], the solution profiles in D are generally very smooth. Note that in practice the spectral convergence of the $\underset{\sim}{q}$-direction numerical method means that the discrete space $\mathcal{P}_N(D)$ need only have a rather low dimensionality (in fact, the superconvergence of $\underset{\approx}{\tau}$ discussed in the next section accentuates this effect further). This is highly advantageous because (a) each $\underset{\sim}{q}$-direction solve requires relatively modest computational resources and (b) a reduction in the dimensionality of $\mathcal{P}_N(D)$ reduces the number of $\underset{\sim}{x}$-direction solves that need to be performed each time-step (*cf.* (3.27)).

Remark 3.12. We assumed $d = 2$ and that basis $\mathcal{A}$ was used for the $\underset{\sim}{q}$-direction spectral method; it is worth emphasising that the same argument could be used for basis $\mathcal{B}$ or basis $\mathcal{C}$ as well. One would need to derive approximation results analogous to (3.60) and (3.61) for these alternative bases, which could be done using the approach in Section 5 of [22].

Remark 3.13. In the preceding argument, we made use of the (pointwise) divergence-free assumption, (3.3). This assumption was made to simplify the argument, but it is not essential, *i.e.* it follows from (3.4) that $\underset{\sim}{\nabla}_x \cdot \underset{\sim}{u} \in \mathrm{L}^\infty(\Omega)$, hence if we allowed $\underset{\sim}{\nabla}_x \cdot \underset{\sim}{u}$ to be nonzero the preceding convergence argument could be modified to use the norm $\|\underset{\sim}{\nabla}_x \cdot \underset{\sim}{u}\|_{\mathrm{L}^\infty(\Omega)}$ instead of (3.3).

3.7. Superconvergence of the polymeric extra-stress

We now consider the convergence of $\underset{\approx}{\tau}$. For consistency with the convergence analysis in the preceding sections, we again restrict our attention to the $d = 2$ case with basis $\mathcal{A}$. For simplicity, we only consider the component τ_{11} of $\underset{\approx}{\tau}$, although the other components can be treated in exactly the same way.

Using the Fourier series representation of $\tilde{\psi} \in \tilde{\mathrm{H}}^1(R)$ from Lemma 5.2 in [22] (recall that $\hat{\psi}(\underset{\sim}{x}, \underset{\sim}{q}, t) = \tilde{\psi}(\underset{\sim}{x}, r, \theta, t)$), we have (*cf.* (3.53)):

$$\tilde{\psi}(\underset{\sim}{x}, r, \theta, t) = \tilde{\psi}_1(\underset{\sim}{x}, r, t) + r \sum_{l=1}^{\infty} \left(\tilde{A}_l(\underset{\sim}{x}, r, t) \cos(2l\theta) + \tilde{B}_l(\underset{\sim}{x}, r, t) \sin(2l\theta) \right). \tag{3.70}$$

We consider τ_{11} to be a functional defined on $\hat{\psi} \in \mathrm{L}^2(D)$ as $\tau_{11}(\hat{\psi}) = \int_D F_1(\underset{\sim}{q})\, q_1 \sqrt{M(\underset{\sim}{q})}\, \hat{\psi}(\underset{\sim}{q}, t)\, \mathrm{d}\underset{\sim}{q}$, where F_1 is defined by (1.1). Hence,

$$\begin{aligned} |\tau_{11}(\hat{\psi})| &= \left| \int_D q_1^2\, U'(\tfrac{1}{2}|\underset{\sim}{q}|^2) \sqrt{M(\underset{\sim}{q})}\, \hat{\psi}\, \mathrm{d}\underset{\sim}{q} \right| \leq b \left(\int_D U'(\tfrac{1}{2}|\underset{\sim}{q}|^2)^2 M(\underset{\sim}{q})\, \mathrm{d}\underset{\sim}{q} \right)^{\frac{1}{2}} \|\hat{\psi}\|_{\mathrm{L}^2(D)} \\ &= \frac{b}{\sqrt{Z}} \left(\int_D \left(1 - |\underset{\sim}{q}|^2/b \right)^{\frac{b}{2}-2} \mathrm{d}\underset{\sim}{q} \right)^{\frac{1}{2}} \|\hat{\psi}\|_{\mathrm{L}^2(D)} = \frac{b}{\sqrt{Z}} \left(2\pi b \int_0^1 (1 - r^2)^{\frac{b}{2}-2}\, r\, \mathrm{d}r \right)^{\frac{1}{2}} \|\hat{\psi}\|_{\mathrm{L}^2(D)} \\ &\leq \frac{b}{\sqrt{Z}} \left(2^{\frac{b}{2}-1}\, \pi b \int_0^1 (1 - r)^{\frac{b}{2}-2}\, \mathrm{d}r \right)^{\frac{1}{2}} \|\hat{\psi}\|_{\mathrm{L}^2(D)}, \end{aligned} \tag{3.71}$$

where Z is the normalisation constant from (1.12). Hence, we require $b > 2$ so that $\tau_{11} \in \mathrm{L}^2(D)' = \mathrm{L}^2(D)$, which is the same condition on b that we have assumed throughout.

Applying τ_{11} to (3.70) gives:

$$\begin{aligned} \tau_{11}(\hat{\psi}) &= \frac{b^2}{\sqrt{Z}} \int_0^1 \int_0^{2\pi} (1 - r^2)^{\frac{b}{4}-1} r^3 \cos^2(\theta)\, \tilde{\psi}(\underset{\sim}{x}, r, \theta, t)\, \mathrm{d}r\, \mathrm{d}\theta \\ &= \frac{\pi\, b^2}{\sqrt{Z}} \int_0^1 r^3 (1 - r^2)^{\frac{b}{4}-1} \left(\tilde{\psi}_1(\underset{\sim}{x}, r, t) + \frac{r}{2} \left(\tilde{A}_1(\underset{\sim}{x}, r, t) \right) \right) \mathrm{d}r. \end{aligned} \tag{3.72}$$

This shows that, quite remarkably, due to orthogonality with $\cos^2(\theta) = \frac{1}{2} + \frac{1}{2}\cos(2\theta)$ over $\theta \in (0, 2\pi)$, the functional τ_{11} filters out all but two terms of the infinite series in (3.70). The same filtering occurs for Galerkin spectral methods that use trigonometric polynomials in θ in the $d = 2$ case, or methods that use spherical harmonics in the $d = 3$ case. For concreteness, we consider the numerical method that uses basis $\mathcal{A}$ in the $\underset{\sim}{q}$-direction. Then, the numerical solution is given by:

$$\tilde{\psi}_{h,N}(\underset{\sim}{x}, r, \theta) = (1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}_{0,k}(\underset{\sim}{x})P_k(r) + r(1-r)\sum_{i=0}^{1}\sum_{l=1}^{N_\theta}\sum_{k=0}^{N_r-1} \tilde{\Psi}^i_{l,k}(\underset{\sim}{x})P_k(r)\Phi_{il}(\theta),$$

where $\tilde{\Psi}_{0,k}, \tilde{\Psi}^i_{l,k} \in V_h$ are line functions as in (3.18). Applying the Cauchy–Schwarz inequality, it follows that

$$\begin{aligned} \|\tau_{11}(\hat{\psi}(t^n)) - \tau_{11}(\hat{\psi}^n_{h,N})\|^2_{\mathrm{L}^2(\Omega)} \quad \leq \quad & C_* \int_\Omega \left\| \tilde{\psi}_1(\underset{\sim}{x}, r, t^n) - (1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}^n_{0,k}(\underset{\sim}{x})P_k(r)\right\|^2_{\mathrm{L}^2_r(0,1)} \mathrm{d}\underset{\sim}{x} \\ & + \frac{C_*}{4}\int_\Omega \left\| r\tilde{A}_1(\underset{\sim}{x}, r, t^n) - r(1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}^{0,n}_{1,k}(\underset{\sim}{x})P_k(r)\right\|^2_{\mathrm{L}^2_r(0,1)} \mathrm{d}\underset{\sim}{x}, \end{aligned} \tag{3.73}$$

where,

$$C_* = \begin{cases} \frac{2\pi^2 b^4}{(b/2-1)\,C}, & 2 < b < 4, \\ \frac{\pi^2\, b^4}{3\,C}, & b \geq 4, \end{cases} \tag{3.74}$$

and $\mathrm{L}^2_r(0,1)$ is the r-weighted L^2 space on $(0,1)$.

On the other hand, we have:

$$\begin{aligned} \|\hat{\psi}(\cdot,\cdot,t^n) - \hat{\psi}^n_N(\cdot,\cdot)\|^2_{\mathrm{L}^2(\Omega\times D)} = 2\pi b \int_\Omega \left\| \tilde{\psi}_1(\underset{\sim}{x}, r, t^n) - (1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}^n_{0,k}(\underset{\sim}{x})P_k(r)\right\|^2_{\mathrm{L}^2_r(0,1)} \mathrm{d}\underset{\sim}{x} \\ + \pi b \sum_{l=1}^{N_\theta}\int_\Omega \left\| r\tilde{A}_l(\underset{\sim}{x}, r, t^n) - r(1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}^{0,n}_{l,k}(\underset{\sim}{x})P_k(r)\right\|^2_{\mathrm{L}^2_r(0,1)} \mathrm{d}\underset{\sim}{x} \\ + \pi b \sum_{l=1}^{N_\theta}\int_\Omega \left\| r\tilde{B}_l(\underset{\sim}{x}, r, t^n) - r(1-r)\sum_{k=0}^{N_r-1} \tilde{\Psi}^{1,n}_{l,k}(\underset{\sim}{x})P_k(r)\right\|^2_{\mathrm{L}^2_r(0,1)} \mathrm{d}\underset{\sim}{x} \\ + \pi b \sum_{l=N_\theta+1}^{\infty}\int_\Omega \left(\left\| r\tilde{A}_l(\underset{\sim}{x}, r, t^n)\right\|^2_{\mathrm{L}^2_r(0,1)} + \left\| r\tilde{B}_l(\underset{\sim}{x}, r, t^n)\right\|^2_{\mathrm{L}^2_r(0,1)}\right) \mathrm{d}\underset{\sim}{x}, \end{aligned} \tag{3.75}$$

and hence, the τ_{11} error only contains two terms from the infinite series in (3.75), and we have

$$\|\tau_{11}(\hat{\psi}) - \tau_{11}(\hat{\psi}_{h,N})\|_{\ell^\infty(0,T;\mathrm{L}^2(\Omega))} \leq \sqrt{\frac{C_*}{2\pi b}}\, \|\hat{\psi} - \hat{\psi}_{h,N}\|_{\ell^\infty(0,T;\mathrm{L}^2(\Omega\times D))}. \tag{3.76}$$

Note that since the line functions $\tilde{\Psi}^n_{0,k}$ and $\tilde{\Psi}^{0,n}_{1,k}$ in (3.73) are computed by solving (3.27) using the $\underset{\sim}{x}$-direction finite element method, we expect an $\mathcal{O}(h^s)$ error to dominate the spatial convergence rate of $\underset{\approx}{\tau}$, just as in (3.69). However, more importantly, by comparing (3.73) and (3.75), we can see that only relatively few terms in the $\underset{\sim}{q}$-direction spectral expansion of $\hat{\psi}_{h,N}$ contribute to the τ_{11} error. Hence, this suggests that the accuracy of $\underset{\approx}{\tau}$ will be less sensitive to the resolution of the $\underset{\sim}{q}$-direction spectral method than the accuracy of $\hat{\psi}_{h,N}$. We demonstrate in Section 5.1 that in practice this results in superconvergence of $\underset{\approx}{\tau}$ with respect to the $\underset{\sim}{q}$-direction spectral method.

4. Implementation of the alternating-direction methods

In this section we furnish further details about the implementation of the alternating-direction methods.

4.1. The $\underset{\sim}{x}$- and $\underset{\sim}{q}$-direction schemes

Recall, first of all, from Section 3.3 that from an implementational point of view method I and method II are almost identical; the only difference between the two methods is that method I uses a semi-implicit temporal discretisation whereas method II uses the backward Euler scheme. We begin by considering implementational issues for the $\underset{\sim}{q}$-direction solvers. Note that this topic was considered in Section 7 of [22], and therefore we refer the reader there for further details.

For both method I and method II, we must solve an $N_D \times N_D$ linear system Q_Ω times per time-step in the $\underset{\sim}{q}$-direction. Q_Ω can be very large in practice. For example, in Section 5 we consider some computations for which Q_Ω is on the order of 10^4. The use of parallel computation can be very helpful in this context because the $\underset{\sim}{q}$-direction linear solves are independent and therefore it is straightforward to perform them in parallel (we discuss this in detail below).

Note that in the $\underset{\sim}{q}$-direction method I requires significantly less computational effort in each time-step than method II because the matrix on the left-hand side in (3.26) is constant for all m and therefore with method I we only need to perform the LU-factorisation of this matrix once, whereas the linear system in (3.38) must be reassembled and solved afresh at each quadrature point $\underset{\sim}{x}_m$ since in general $\underset{\approx}{\kappa}(\underset{\sim}{x}_m)$ varies from one quadrature point to the next. On the other hand, numerical experiments in Section 2.6.2 of [21] indicate that the backward Euler temporal discretisation of the $\underset{\sim}{q}$-direction equation used in method II is more stable in practice than the semi-implicit scheme used in method I, especially for larger values of Wi or $\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(\Omega)}$ (note that Lems. 3.4 and 3.5 show that both method I and method II are unconditionally stable on any finite time interval, but that the stability constant grows exponentially with T, and therefore for long-time computations we can still observe rapid growth of the solution; in practice, method I tends to require Δt to be chosen to be smaller than for method II to avoid this pathological behaviour). Hence, there is a familiar trade-off in efficiency: each time-step is faster with method I, but we can take larger time-steps with method II. Therefore the optimal choice of numerical method depends on the problem at hand.

Remark 4.1. The alternating direction method used by Chauvière and Lozinski in [11] is similar to method II in that it treats the $\underset{\approx}{\kappa}$ convection term implicitly in time. In the follow-up papers [10,27] the same authors developed a fast solver in which the computational work required for each $\underset{\sim}{q}$-direction solve was significantly reduced. However, their fast solver was based on the assumption that $\underset{\approx}{\kappa}$ arises from a two-dimensional velocity field (*i.e.* that $\Omega \subset \mathbb{R}^2$) whereas in the present work we are interested in developing numerical methods that are suitable for both $\Omega \subset \mathbb{R}^2$ and $\Omega \subset \mathbb{R}^3$.

The $\underset{\sim}{q}$-direction solvers for methods I and II were implemented in the `C++` programming language and `PETSc` [3] was used to perform the linear algebra operations. `PETSc` was a natural choice in this context because it is designed for use on parallel architectures, which is a feature we made extensive use of.

In the $\underset{\sim}{x}$-direction, methods I and II are identical: For each line function, $\hat{\psi}_k^{n*}$, $k = 1, \ldots, N_D$, we solve the transport equation (3.29). This involves solving an $N_\Omega \times N_\Omega$ linear system N_D times, although the system matrix $M_x + \Delta t\, T_x$ only needs to be assembled once per time-step.

In our implementation, we used an $\mathrm{H}^1(\Omega)$-conforming finite element method with quadratic shape functions to perform the $\underset{\sim}{x}$-direction computations, and we used GMRES to solve the resulting linear systems. Hence, assuming sufficient regularity for $\hat{\psi}$, we can set $s = 2$ in (3.69), which yields $\mathcal{O}(h^2)$ terms in the error estimate.

The $\underset{\sim}{x}$-direction finite element method was implemented using the free, open source `C++` finite element library `libMesh` [20].

4.2. Parallel implementation of the alternating-direction methods

It is clear that the computational effort required to solve the high-dimensional Fokker–Planck equation can be very large, particularly in the case $d = 3$, and hence parallel computation is a key ingredient in the alternating-direction framework developed here. As indicated above, methods I and II are very well suited to implementation on a parallel architecture; indeed these algorithms are "embarrassingly parallel" in the sense that they involve performing a large number of independent solves in each time-step.

More specifically, suppose we use N_{proc} processors ($N_{\text{proc}} \geq 1$) to solve a problem (using either method I or II) with parameters N_D, N_Ω denoting the number of basis functions in the $\underset{\sim}{q}$-direction and $\underset{\sim}{x}$-direction, respectively, and Q_Ω defining the number of quadrature points in Ω, as in (3.14). At time-level n, we store a dense matrix $D^n \in \mathbb{R}^{Q_\Omega \times N_D}$, where $(D^n)_{ij} = \hat{\psi}^n_j(\underset{\sim}{x}_i)$, and $\hat{\psi}^n_j \in V_h$ is a line function as in (3.18). The entries of D^n uniquely determine $\hat{\psi}^n_{h,N} \in V_h \otimes \mathcal{P}_N(D)$. In practice D^n can be a very large matrix, so we partition it among the processors so that each processor stores a subset of the rows (for $\underset{\sim}{q}$-direction solves) or columns (for $\underset{\sim}{x}$-direction solves) of D^n. We would like these submatrices to be equally sized to obtain ideal load balancing between processors, but depending on Q_Ω, N_D and N_{proc}, this is often not possible. However, to simplify the discussion here, we will assume for the remainder of this section that N_{proc} is a common divisor of Q_Ω and N_D and hence that the submatrices are equally sized.

Now, let us consider the $\underset{\sim}{q}$-direction computations at time-level n (we do not distinguish between methods I and II here because, from the point of view of the current discussion, they are identical). We distribute D^n so that each processor stores Q_Ω/N_{proc} rows of the matrix. Then, simultaneously, each processor solves the Q_Ω/N_{proc} $\underset{\sim}{q}$-direction problems corresponding to its rows in D^n and updates the data in the matrix. In this manner, D^n is updated to D^{n*} where $(D^{n*})_{ij} = \hat{\psi}^{n*}_j(\underset{\sim}{x}_i)$.

Next, we perform the $\underset{\sim}{x}$-direction computations. First of all, however, we need to redistribute D^{n*} so that each processor stores N_D/N_{proc} columns of the matrix[8]. This involves a global communication operation between all of the processors, which can be time-consuming. The time required to perform this parallel communication step depends on the problem size and the number of processors being used. We discuss this issue with regard to some practical computations in Section 5.1, where we show that by selecting N_{proc} appropriately it is generally possible to ensure that the matrix redistribution steps take only a small proportion of the overall computation time.

So, once this matrix redistribution is complete, the $\underset{\sim}{x}$-direction computations on each processor proceed in the same way as in the $\underset{\sim}{q}$-direction. That is, each processor works sequentially through its N_D/N_{proc} columns, first solving (3.29), and then sampling the resulting line function $\hat{\psi}^{n+1}_k$ at $\underset{\sim}{x}_m$ for $m = 1, \ldots, Q_\Omega$ and writing these values back into the matrix. This yields the updated matrix D^{n+1} on completion of all of the $\underset{\sim}{x}$-direction solves.

4.3. Coupled algorithm for the micro-macro model

We now define the algorithm for solving the micro-macro model (*i.e.* the Navier–Stokes–Fokker–Planck or Stokes–Fokker–Planck system). The algorithm that we use is essentially the same as those used by Chauvière and Lozinski [10,11,27] and Helzel and Otto [17] for this purpose.

Note, first of all, that we use the well known Taylor–Hood mixed finite element method for solving the Navier–Stokes or Stokes equations, *i.e.* let V_h denote the $\mathrm{H}^1(\Omega)$-conforming finite element space based on $\mathcal{T}_h$ that uses continuous piecewise quadratic shape functions (*cf.* Sect. 4.1) and let P_h be the corresponding space based on continuous piecewise linear shape functions. Then we use V_h and P_h as, respectively, the Taylor–Hood velocity and pressure finite element spaces (*cf.* Chap. 5 of [16]) such that $\underset{\sim}{u}_h \in V_h$ and $p_h \in P_h$; these spaces are known to satisfy the inf-sup stability condition (*cf.* Sect. 12.6 of [7]).

[8] In our implementation, we performed this redistribution using `PETSc`'s transpose operation for parallel dense matrices.

We implemented this Taylor–Hood scheme in `libMesh` in order to obtain the computational results in Section 5.2, and in the Navier–Stokes case we used a Newton scheme to solve the resulting nonlinear system of equations. The linear systems were solved using GMRES with incomplete LU factorisation as a preconditioner. In order to obtain faster convergence rates for the iterative solver one could apply more advanced preconditioning techniques, such as the techniques discussed in [16] that take advantage of the structure of the linear systems arising from the discretisation of Stokes or Navier–Stokes problems. However, there is little incentive for us to accelerate the convergence of our Navier–Stokes or Stokes solvers in this way because the overall computation time for computations with the Navier–Stokes–Fokker–Planck system is dominated by solving the Fokker–Planck equation on $\Omega \times D$.

Note that in Section 3 we restricted our attention to enclosed flows to simplify the analysis, but in Section 5.2 we perform computational simulations of channel flows with the micro-macro model, and therefore we must now consider the necessary modifications to the alternating-direction framework of Section 3 for problems that have inflow and outflow boundaries. In particular, we need to define the boundary conditions for the Fokker–Planck equation on $\partial\Omega_{\text{in}}$ and $\partial\Omega_{\text{out}}$.

In fact, since the Fokker–Planck equation on Ω is a pure advection problem, we do not need to do anything different on $\partial\Omega_{\text{out}}$ since by definition we have $\underset{\sim}{u}_h \cdot \underset{\sim}{n} > 0$ there. However, we do need to treat the inflow boundary differently. Suppose we set $\underset{\sim}{u}_h^n|_{\partial\Omega_{\text{in}}} = \underset{\sim}{u}_{\text{in}}^n$ for the Stokes/Navier–Stokes system for $n = 1, \ldots, N_T$. Then that boundary data also defines $\underset{\approx}{\kappa}_{\text{in}}^n = \underset{\sim}{\nabla}_x \underset{\sim}{u}_{\text{in}}^n$ on $\partial\Omega_{\text{in}}$[9], and $\underset{\approx}{\kappa}_{\text{in}}$ in turn determines the inflow boundary data, $\hat{\psi}_{\text{in}}$, on $\partial\Omega_{\text{in}} \times D$ for the Fokker–Planck equation. That is, for $s \in \partial\Omega_{\text{in}}$, $\hat{\psi}_{\text{in}}^n(s,\cdot) : \underset{\sim}{q} \in D \mapsto \hat{\psi}_{\text{in}}^n(s,\underset{\sim}{q}) \in \mathbb{R}$ for $n = 1, \ldots, N_T$ is determined by solving the $\underset{\sim}{q}$-direction Fokker–Planck equation corresponding to $\underset{\approx}{\kappa}_{\text{in}}^n(s)$, so that $\hat{\psi}_{\text{in}}^n(s,\cdot) \in \mathcal{P}_N(D)$ for each n. Writing $\hat{\psi}_{\text{in}}(s,\underset{\sim}{q}) = \sum_{k=1}^{N_D} \hat{\psi}_{\text{in},k}(s) Y_k(q)$ for $(s,\underset{\sim}{q}) \in \partial\Omega_{\text{in}} \times D$, it then follows from (3.19) that $\hat{\psi}_{\text{in},k}$ defines the inflow boundary data on $\partial\Omega_{\text{in}}$ for $\hat{\psi}_k$ in (3.29). In practice we only solve for $\hat{\psi}_{\text{in}}$ at the nodes of $\mathcal{T}_h$ on $\partial\Omega_{\text{in}}$ so that we can impose the inflow boundary condition on the line function $\hat{\psi}_k$ in an interpolatory sense. Notice also that we can compute the inflow boundary data for $\hat{\psi}_{h,N}$ before we begin solving the Navier–Stokes–Fokker–Planck system, since $\underset{\sim}{u}_{\text{in}}$ and $\underset{\approx}{\kappa}_{\text{in}}$ are specified *a priori*.

We now define the coupled micro-macro algorithm. We first initialise the algorithm to the equilibrium state by setting $\underset{\sim}{u}_h^0 = \underset{\sim}{0}$ on Ω, and therefore $\underset{\approx}{\kappa}^0 = \underset{\sim}{\nabla}_x \underset{\sim}{u}_h^0 = \underset{\approx}{0}$ on Ω also. The corresponding equilibrium steady-state solution of the Fokker–Planck equation is $\psi = M$, and hence we set $\hat{\psi}_{h,N}^0 = \sqrt{M} \in V_h \otimes \mathcal{P}_N(D)$ on $\Omega \times D$[10]. Also, for consistency with $\hat{\psi}_{h,N}^0$, we set $\underset{\approx}{\tau}_{h,N}^0$ to be the identity tensor on Ω. Then, for $n = 0, \ldots, N_T - 1$, we perform the following steps:

1. Compute $\underset{\sim}{u}_h^{n+1} \in V_h$ and $p_h^{n+1} \in P_h$ using the mixed finite element method discussed above for either the Navier–Stokes or Stokes system. We use the tensor $\underset{\approx}{\tau}_{h,N}^n$ as the source term in the momentum equation ((1.2) or (1.7)).
2. Use method I or method II to compute $\hat{\psi}_{h,N}^{n+1} \in V_h \otimes \mathcal{P}_N(D)$ with $\underset{\approx}{\kappa}^n$ in (3.26) for method I or with $\underset{\approx}{\kappa}^{n+1}$ in (3.38) for method II, and $\underset{\sim}{u}_h^{n+1}$ in (3.29) for either method.
3. Using (1.5), compute $\underset{\approx}{\tau}_{h,N}^{n+1}$ on Ω based on $\hat{\psi}_{h,N}^{n+1} \in V_h \otimes \mathcal{P}_N(D)$.
4. Return to 1. and continue marching in time.

Note that in this algorithm the $\underset{\approx}{\tau}_{h,N}$ terms in the momentum equations are explicit in time. This allows the Stokes/Navier–Stokes equations to be coupled to the Fokker–Planck equation in a simple manner, but the drawback is that the algorithm defined in steps 1. to 4. above is only conditionally stable. In Section 5.2 we use $\Delta t = 0.01$ and this time-step size is sufficiently small to yield a reliable numerical method for the micro-macro problems that we consider.

[9] We assume that $\underset{\sim}{u}_{\text{in}}$ is a fully-developed flow, and therefore that the velocity field upstream of $\partial\Omega_{\text{in}}$ has the same profile as $\underset{\sim}{u}_{\text{in}}$; this ensures that $\underset{\sim}{\nabla}_x \underset{\sim}{u}_{\text{in}}$ is well-defined on the inflow boundary.

[10] We assume here that $\sqrt{M} \in \mathcal{P}_N(D)$, which is reasonable according to Remark 6.1 in [22].

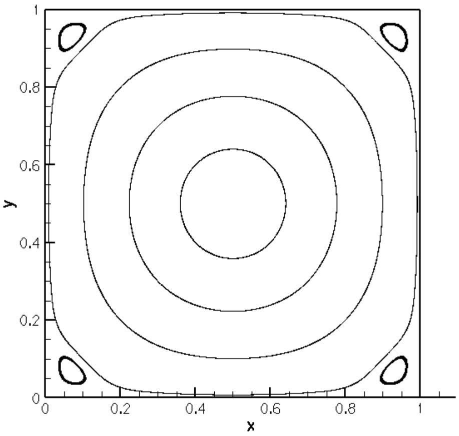


Figure 1. Streamlines of the macroscopic velocity field $\underset{\sim}{u}$ driving the enclosed flow model problem. The velocity field is the solution of the steady Navier–Stokes equation with Re $= 1$ on $\Omega = (0,1)^2$ with forcing $f(x,y) = (5\sin(2\pi y), -5\sin(2\pi x))$.

5. Numerical results

In this section we present some numerical results. First of all, in Section 5.1 we consider a model problem for the isolated FENE Fokker–Planck equation in which $\underset{\sim}{u}$ is taken to be a fixed enclosed flow. This model problem conforms to the assumptions of Section 3, and hence it provides experimental verification of the theoretical results derived therein. Then in Section 5.2 we present computational results for two channel flow problems for the micro-macro model.

5.1. **Enclosed flow model problem**

Here we take $\underset{\sim}{u}$ to be the enclosed flow velocity field in Figure 1, and we take $\underset{\sim}{u}$ to be constant in time throughout $t \in (0,T]$. We obtained $\underset{\sim}{u}$ by solving the steady incompressible Navier–Stokes equations (using the Taylor–Hood finite element scheme) with Re $= 1$, and with forcing term $f(x,y) = (5\sin(2\pi y), -5\sin(2\pi x))$, in the domain $\Omega = (0,1)^2$ and with zero Dirichlet condition for $\underset{\sim}{u}$ on $\partial\Omega$. In this case, $\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(\Omega)} \approx 2$. Note that in general the Taylor–Hood scheme for the Navier–Stokes equations does not yield a (pointwise) divergence-free velocity field, and hence the assumption (3.3) is not satisfied for the computational results in this section. However, as noted in Remark 3.13, the analysis developed in Sections 3.6 and 3.7 can be extended essentially unchanged to the case in which $\underset{\sim}{u}$ is not divergence-free.

We now consider computations using methods I and II with basis $\mathcal{A}$ in the $\underset{\sim}{q}$-direction for the model problem described above, with the parameters Wi $= 1$ and $b = 12$. Also, in each of the computations discussed below, we used the initial condition $\hat{\psi}^0_{h,N}(\underset{\sim}{x},\underset{\sim}{q}) = \sqrt{M(\underset{\sim}{q})}$ and we ensured that $N_r \geq 6$, since according to Remark 6.1 in [22], that guarantees that $\sqrt{M} \in \mathcal{P}_N(D)$ in this case. Our goal is to compare the performance of methods I and II, and to study the convergence of these methods under mesh refinement. All of the computations in this section were performed on the Lonestar parallel computer at the Texas Advanced Computing Center (TACC), `http://www.tacc.utexas.edu`, and we used the parallel implementation of the alternating direction method described in Section 4.2.

We do not know the exact solution of the Fokker–Planck equation with the velocity field in Figure 1 and therefore in order to obtain quantitative convergence results we first computed a “reference solution”, $\hat{\psi}_{\mathrm{ref}}$, and corresponding polymeric extra-stress tensor, $\underset{\approx}{\tau}_{\mathrm{ref}}$, using method I with basis $\mathcal{A}$ in the $\underset{\sim}{q}$-direction, and with a quadrature rule on Ω that satisfied QH1. We obtained this reference solution using a discrete space, $(V_h \otimes \mathcal{P}_N(D))_{\mathrm{ref}}$, for which $\mathcal{T}_h$ was a 40×40 uniform mesh of square finite elements and $(N_r, N_\theta) = (14, 14)$.

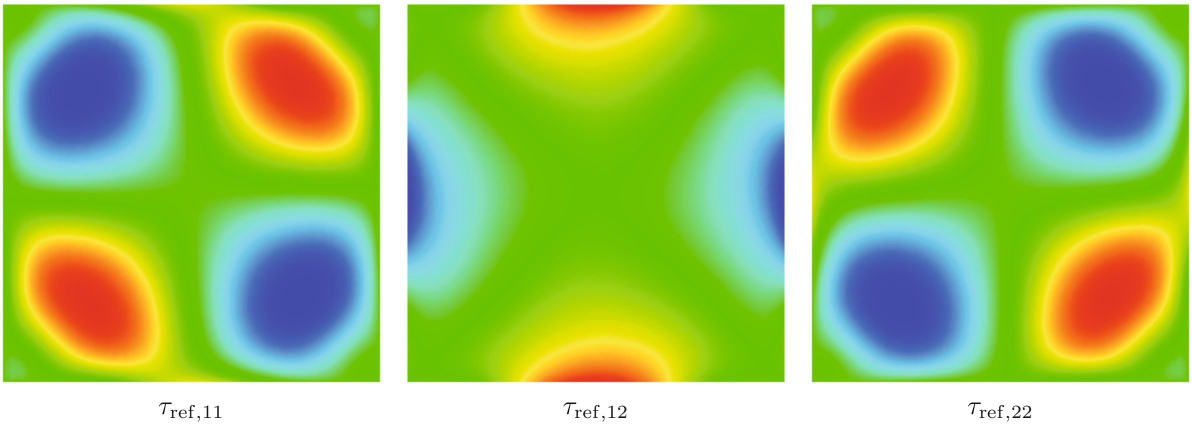


FIGURE 2. The components of $\underset{\approx}{\tau}_{\mathrm{ref}}$ at $T = 0.2$. Note that we do not show $\tau_{\mathrm{ref},21}$ since it is identical to $\tau_{\mathrm{ref},12}$. In the $\tau_{\mathrm{ref},11}$ and $\tau_{\mathrm{ref},22}$ plots, the values range from 0.882 (blue) to 1.15 (red), and in the $\tau_{\mathrm{ref},12}$ plot we have -0.229 (blue) to 0.229 (red) (figure in colour available online at http://www.esaim-m2an.org/).

In order to satisfy QH1 in this case we used a tensor product Gauss-Legendre quadrature rule with 16 quadrature points per element, and hence $Q_\Omega = 25\,600$. We took 200 time-steps with $\Delta t = 10^{-3}$ so that $T = 0.2$; this value of Δt is sufficiently small so that temporal discretisation error does not contaminate the spatial convergence results presented below. The components of $\underset{\approx}{\tau}_{\mathrm{ref}}$ at $T = 0.2$ are shown in Figure 2.

In order to obtain convergence data, we then computed $\hat{\psi}_{h,N}$ and the corresponding stress tensor $\underset{\approx}{\tau}$ for several discrete spaces coarser than $(V_h \otimes \mathcal{P}_N(D))_{\mathrm{ref}}$. First of all we carried out this process using the same numerical method with which we obtained the reference solution, *i.e.* method I with basis $\mathcal{A}$ and a quadrature rule that satisfied QH1. The solution data obtained from these computations are denoted $\hat{\psi}_{\mathrm{I}}$ and $\underset{\approx}{\tau}_{\mathrm{I}}$ below. Then, we also computed a corresponding set of numerical solutions on the same discrete spaces, but using method II with basis $\mathcal{A}$ and a quadrature rule that only satisfied QH2[11]. We denote the solution data in this second case by $\hat{\psi}_{\mathrm{II}}$ and $\underset{\approx}{\tau}_{\mathrm{II}}$.

The numerical results for $\hat{\psi}_{\mathrm{I}}$ and $\underset{\approx}{\tau}_{\mathrm{I}}$ were obtained using a numerical method that satisfies all of the hypotheses required by the convergence estimates for $\hat{\psi}$ and $\underset{\approx}{\tau}$ in Sections 3.6 and 3.7 (except the divergence-free assumption on $\underset{\sim}{u}$, but, as mentioned above, this assumption is not essential; we only introduced it to simplify the analysis). Hence, the $\hat{\psi}_{\mathrm{I}}$ and $\underset{\approx}{\tau}_{\mathrm{I}}$ convergence data in the table allow us to compare the theoretical estimates with practical convergence results. Also, the numerical results enable us to compare the convergence behaviour of method I with QH1 to method II with QH2. These two methods are very similar to one another hence we expect to observe the same convergence behaviour in the two cases, but it is important to provide experimental evidence that these two methods converge to the same solution, and at the same rate in practice because strictly speaking our convergence analysis is only valid for method I with QH1.

The convergence estimates (3.69) and (3.76) indicate that if the error due to the $\underset{\sim}{q}$-direction spectral method is negligible compared to the error from the $\underset{\sim}{x}$-direction finite element method, we should obtain $\mathcal{O}(h^2)$ convergence rates for both $\hat{\psi}$ and $\underset{\approx}{\tau}$ as $\mathcal{T}_h$ is refined. Table 1 gives the relative errors $\|\hat{\psi}_{\mathrm{I}} - \hat{\psi}_{\mathrm{ref}}\|_{\mathrm{L}^2(\Omega\times D)}/\|\hat{\psi}_{\mathrm{ref}}\|_{\mathrm{L}^2(\Omega\times D)}$ and $\|\hat{\psi}_{\mathrm{II}} - \hat{\psi}_{\mathrm{ref}}\|_{\mathrm{L}^2(\Omega\times D)}/\|\hat{\psi}_{\mathrm{ref}}\|_{\mathrm{L}^2(\Omega\times D)}$ as well as the relative τ_{11} errors $\|\tau_{\mathrm{I},11} - \tau_{\mathrm{ref},11}\|_{\mathrm{L}^2(\Omega)}/\|\tau_{\mathrm{ref},11}\|_{\mathrm{L}^2(\Omega)}$ and $\|\tau_{\mathrm{II},11} - \tau_{\mathrm{ref},11}\|_{\mathrm{L}^2(\Omega)}/\|\tau_{\mathrm{ref},11}\|_{\mathrm{L}^2(\Omega)}$, at $T = 0.2$, for the discrete spaces that we considered.

[11] We only require nine tensor product Gauss-Legendre quadrature points per element to satisfy QH2 on square finite elements.

TABLE 1. Convergence of $\hat{\psi}$ and τ_{11} with respect to the reference solution $\hat{\psi}_{\mathrm{ref}}$ and reference polymeric stress tensor $\tau_{\mathrm{ref},11}$ for a series of increasingly refined discrete spaces. The errors are calculated in the L^2 norm at $T = 0.2$, and are normalised by dividing by $\|\hat{\psi}_{\mathrm{ref}}(\cdot,\cdot,T)\|_{\mathrm{L}^2(\Omega\times D)} = 0.31$ and $\|\tau_{\mathrm{ref},11}(\cdot,T)\|_{\mathrm{L}^2(\Omega)} = 1.04$.

$\mathcal{T}_h$	(N_r, N_θ)	$\hat{\psi}_{\mathrm{I}}$ error	$\tau_{\mathrm{I},11}$ error	$\hat{\psi}_{\mathrm{II}}$ error	$\tau_{\mathrm{II},11}$ error
5×5	$(6,6)$	2.07×10^{-2}	1.63×10^{-2}	2.08×10^{-2}	1.63×10^{-2}
5×5	$(8,8)$	2.05×10^{-2}	1.63×10^{-2}	2.06×10^{-2}	1.63×10^{-2}
5×5	$(10,10)$	2.05×10^{-2}	1.63×10^{-2}	2.06×10^{-2}	1.63×10^{-2}
10×10	$(6,6)$	6.25×10^{-3}	4.22×10^{-3}	6.30×10^{-3}	4.24×10^{-3}
10×10	$(8,8)$	5.62×10^{-3}	4.22×10^{-3}	5.65×10^{-3}	4.23×10^{-3}
10×10	$(10,10)$	5.54×10^{-3}	4.22×10^{-3}	5.58×10^{-3}	4.23×10^{-3}
20×20	$(6,6)$	3.29×10^{-3}	9.95×10^{-4}	3.40×10^{-3}	1.07×10^{-3}
20×20	$(8,8)$	1.80×10^{-3}	9.90×10^{-4}	1.89×10^{-3}	1.04×10^{-3}
20×20	$(10,10)$	1.52×10^{-3}	9.90×10^{-4}	1.67×10^{-3}	1.04×10^{-3}

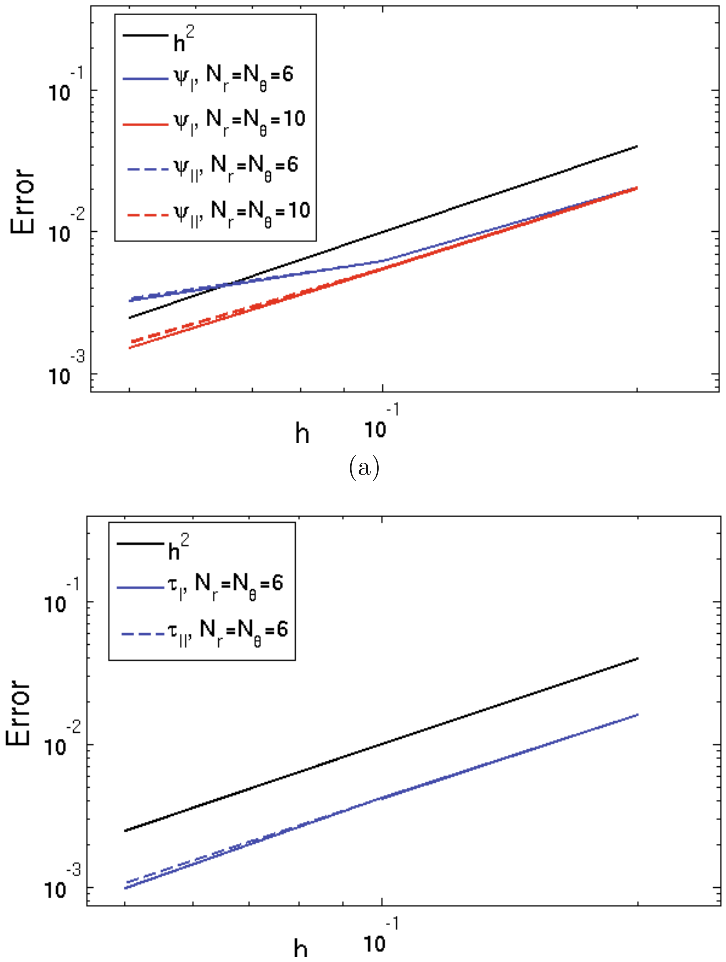
In order to gain further insight into the convergence behaviour of the numerical methods, we plotted the data in Table 1 in Figure 3.

In Figure 3a, the convergence results for $\hat{\psi}_{\mathrm{I}}$ and $\hat{\psi}_{\mathrm{II}}$ with $(N_r, N_\theta) = (6,6)$ and $(N_r, N_\theta) = (10,10)$ are plotted on a log-log scale. We have also included a plot of h^2 to show how the decay of the computed errors compare to the expected asymptotic rate. First of all, it is clear from the figure that the two numerical methods behave very similarly; the lines from $\hat{\psi}_{\mathrm{I}}$ and $\hat{\psi}_{\mathrm{II}}$ are almost indistinguishable. Also, Figure 3a shows that we obtain $\mathcal{O}(h^2)$ convergence when $(N_r, N_\theta) = (10,10)$. However, when $(N_r, N_\theta) = (6,6)$, the plots plateau, which indicates that the error due to the spectral method dominates the $\mathcal{O}(h^2)$ finite element error when $\mathcal{T}_h$ is a 20×20 mesh.

The $\tau_{\mathrm{I},11}$ and $\tau_{\mathrm{II},11}$ convergence data is plotted in Figure 3b (the data in Tab. 1 is almost identical for $(N_r, N_\theta) = (6,6), (8,8)$ and $(10,10)$, and therefore we only show the $(N_r, N_\theta) = (6,6)$ data). The plot shows that we obtained $\mathcal{O}(h^2)$ convergence for both $\tau_{\mathrm{I},11}$ and $\tau_{\mathrm{II},11}$ as $\mathcal{T}_h$ is refined from a 5×5 mesh to 20×20 mesh, when $(N_r, N_\theta) = (6,6)$. This is markedly different from the convergence behaviour of $\hat{\psi}_{h,N}$, in which the $\underset{\sim}{q}$-direction spectral error for $(N_r, N_\theta) = (6,6)$ dominated the finite element error on the 20×20 $\underset{\sim}{x}$-direction mesh. Therefore, this indicates that, as expected from Section 3.7, the D domain spectral method exhibits superconvergence for $\underset{\approx}{\tau}$ compared to $\hat{\psi}$. The superconvergence of $\underset{\approx}{\tau}$ is extremely beneficial in the context of micro-macro computations for simulating dilute polymeric fluids because in that setting the error in $\hat{\psi}$ is irrelevant; we are solely interested in the $\underset{\approx}{\tau}$ error since $\underset{\approx}{\tau}$ rather than $\hat{\psi}$ enters into the macroscopic momentum equation.

Recall from the discussion in Section 4.1 that we expect method I to require significantly less computational work per time-step in the $\underset{\sim}{q}$-direction than method II. To demonstrate this in practice, we solved the same enclosed flow model problem using both method I and method II. We used a 20×20 uniform mesh $\mathcal{T}_h$ of square finite elements with $Q_\Omega = 3600$ and basis $\mathcal{B}$ with $(N_r, N_\theta) = (15,15)$ so that $N_D = 465$. With $N_{\mathrm{proc}} = 4$, the total computation time per time-step for method I was 1.75 seconds, whereas for method II it was 3.42 seconds. This difference is due to the fact that method II took 2.37 seconds per time-step to perform the $\underset{\sim}{q}$-direction computations, whereas method I only took 0.70 seconds per time-step in the $\underset{\sim}{q}$-direction.

Nevertheless, for problems of physical interest, method II is often the preferred alternating-direction method. This is because the fully implicit temporal discretisation used by method II is more stable than the semi-implicit scheme in method I, especially for larger macroscopic velocity gradients and Weissenberg numbers (*cf.* Sect. 6.2 of [21]). Hence method I can require much smaller time-step sizes than method II, and this can often outweigh the reduced computational complexity per time-step of method I. Also, for large-scale problems we generally prefer to satisfy only QH2 rather than QH1 since with QH2 we can obtain a smaller value of Q_Ω, which in turn reduces the computational work required in each time-step of the alternating-direction method.



(a)

(b)

FIGURE 3. Plots of the $\hat{\psi}$ and $\underset{\approx}{\tau}$ convergence data in Table 1 (the lines interpolate the convergence data). (a) The black line shows the expected asymptotic decay rate, h^2, and the blue and red liness show the convergence of the two numerical methods when (N_r, N_θ) is fixed at $(6, 6)$ and $(10, 10)$, respectively. (b) The black line shows the expected asymptotic decay rate, h^2, and the solid and dashed blue lines show, respectively, the $\tau_{\mathrm{I},11}$ and $\tau_{\mathrm{II},11}$ data for $(N_r, N_\theta) = (6, 6)$. The data for the other values of (N_r, N_θ) are not plotted since the τ_{11} convergence data in Table 1 is virtually unaffected by increasing the number of spectral basis functions (figure in colour available online at http://www.esaim-m2an.org/).

We now move on to consider the scaling of the computation time as we increase the number of processors in the parallel implementation of the alternating-direction method. The enclosed-flow problem considered above provides a convenient test case with which we can quantify the parallel speedup for the alternating-direction method. We studied this speedup by, first of all, solving the enclosed flow problem on one node of the Lonestar parallel computer (each node contains four processors) to get the base computation time per time-step, which we denote $T(1)$. We then repeated the same computation, but using more computational nodes

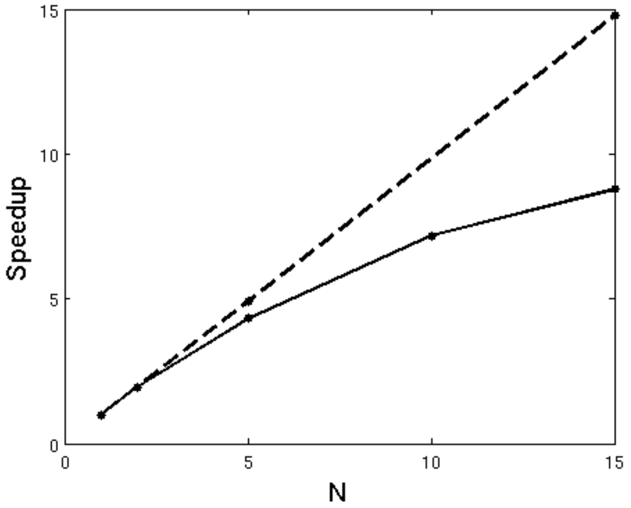


FIGURE 4. Plot of speedup, *i.e.* $T(1)/T(N)$, as the number of computational nodes is increased from 1 to 15 (the plot shows linear interpolation of data points). The speedup data for $(N_D, Q_\Omega) = (120, 3600)$ is plotted as a solid line and the dashed line shows the data for $(N_D, Q_\Omega) = (1800, 8100)$. For each computation we chose the number of nodes so that $N_{\text{proc}}(= 4N)$ was a common divisor of N_D and Q_Ω in order to ensure optimal load balancing in each case so that the comparisons of computation time are fair.

of the parallel computer and we recorded the computation time, $T(N)$, in each case, where N denotes the number of computational nodes that were used. We refer to the ratio $T(1)/T(N)$ as the *parallel speedup*.

The parameters that have the most significant effect on the computation time of the parallel alternating-direction scheme are N_D and Q_Ω, since these determine the number of $\underset{\sim}{x}$- and $\underset{\sim}{q}$-direction solves that need to be performed each time-step. Note that there are only two steps in the alternating-direction algorithm for which the computation time does not scale down proportionally to the number of processors being used: the matrix assembly for (3.29), which must be performed exactly once per time-step irrespective of N_{proc}, and also the dense matrix redistribution that precedes direction changes in the alternating-direction method. However, if the $\underset{\sim}{x}$- and $\underset{\sim}{q}$-direction solves dominate the overall computation time, then we can expect that the parallel speedup will scale linearly with the number of processors being used.

In order to examine the scaling of the parallel speedup in practice, we performed computations for two different discrete spaces, such that (i) $N_D = 120$ and $Q_\Omega = 3600$, and (ii) $N_D = 1800$ and $Q_\Omega = 8100$. We solved the enclosed flow problem for these spaces using a number of different choices of N_{proc}. We used method II with basis $\mathcal{B}$ to obtain the data below, but the parallel speedup behaviour is essentially the same whether we use method I or II or basis $\mathcal{A}$ or $\mathcal{B}$. The base computation times were $T(1) = 0.53$ seconds for the $(N_D, Q_\Omega) = (120, 3600)$ computation, and $T(1) = 157.0$ seconds for the $(N_D, Q_\Omega) = (1800, 8100)$ case.

The parallel speedup of the alternating-direction method for the two discrete spaces discussed above is plotted in Figure 4. In the case that $(N_D, Q_\Omega) = (1800, 8100)$, we obtained a parallel speedup of 14.8 when $N = 15$ (*i.e.* $N_{\text{proc}} = 60$), whereas the speedup tailed off to less than 10 when $N = 15$ for the computation with $(N_D, Q_\Omega) = (120, 3600)$. This difference in the scaling of the parallel speedup is primarily due to the fact that the overhead from the redistribution of D^n is much larger, as a proportion of the overall computation time, for the smaller problem. For example, for the $(N_D, Q_\Omega) = (120, 3600)$ problem, matrix redistribution took 8.66% of the overall computation time when $N = 1$, but when $N = 15$, it increased to 30.4%. By contrast, in the larger problem with $(N_D, Q_\Omega) = (1800, 8100)$, more time is spent on the $\underset{\sim}{q}$- and $\underset{\sim}{x}$-direction solves in each time-step, so that only 0.89% of the computation time was taken for the matrix redistribution when $N = 1$, which increased to 2.25% when $N = 15$. Since 2.25% is still only a small proportion of the overall computation time, the matrix redistribution overhead does not significantly detract from the near optimal scaling of the parallel speedup shown in Figure 4 for the $(N_D, Q_\Omega) = (1800, 8100)$ case. This indicates that as long as the values of N_D and Q_Ω are large enough, the alternating-direction method can scale efficiently to a very large number of processors.

5.2. Computational results for the micro-macro model

We now consider two micro-macro model channel flow problems. In the first, we consider the Navier–Stokes–Fokker–Planck system for a 4–to–1 contraction problem in two dimensions, and in the second, we consider the flow around a spherical obstacle in a channel with square cross-section in the case of $d = 3$. We used the coupled algorithm discussed in Section 4.3 in both cases. For each of these two problems we present numerical results for one particular discrete space $V_h \otimes \mathcal{P}_N(D)$, but in each case we performed mesh refinement studies (*i.e.* we solved using a sequence of increasingly refined spaces) to ensure that the numerical results shown below are accurate.

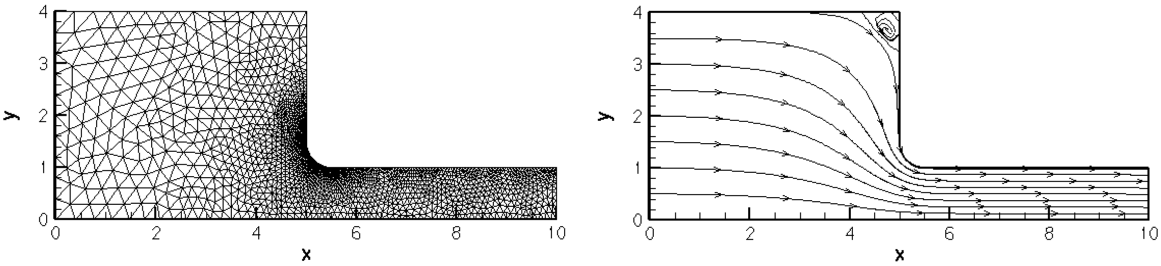
4-to-1 contraction. Contraction flows are standard benchmark problems in computational rheology because they are challenging from the numerical point of view and they also have practical relevance in industrial applications (for a detailed discussion of contraction flows see Chap. 8 of [31]). In this section we consider the coupled Navier–Stokes–Fokker–Planck model with $\mathrm{Re} = 1$ in a contracting domain, which is 10 units long, 4 units wide in the wider section and 1 unit wide in the narrow section. We imposed a parabolic inflow profile on $\underset{\sim}{u}_h$ on the left boundary $(\partial\Omega_{\text{in}})$ with maximum value of $U_{\max} = 1$, a zero normal stress outflow condition on the right boundary $(\partial\Omega_{\text{out}})$, zero Dirichlet condition on the top boundary and a symmetry condition on the bottom boundary (by setting the y-component of $\underset{\sim}{u}_h$ to zero there). Also, instead of an L-shaped domain, we use the physical space domain with a rounded corner shown in Figure 5 to avoid a singularity in the velocity field and hence ensure that $\underset{\approx}{\kappa}_h = \underset{\sim}{\nabla}_x \underset{\sim}{u}_h$ is uniformly bounded as $h \to 0_+$. Also, in order to resolve the solution satisfactorily, the finite element mesh, $\mathcal{T}_h$, has been graded so that it is finer near the corner.

We set $b = 12$, $\mathrm{Wi} = 0.8$, $\gamma = 0.59$ and took 500 time-steps with $\Delta t = 0.01$ so that $T = 5$. We used alternating-direction method II with basis $\mathcal{A}$. The mesh $\mathcal{T}_h$ contained 905 triangular finite elements and $Q_\Omega = 5430$[12]. Also, we used $(N_r, N_\theta) = (20, 20)$ for the $\underset{\sim}{q}$-direction spectral method, so that $N_D = 820$. In this computation $\mathrm{Wi}\,\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(\Omega)} \approx 10$, and the high velocity gradients are localised to the vicinity of the rounded corner. The macroscopic velocity field at $T = 5$ is plotted in Figure 5b and the corresponding components of $\underset{\approx}{\tau}_{h,N}$ are shown in Figure 6. The computation was performed using 40 processors of the Lonestar supercomputer at the Texas Advanced Computing Center using a parallel implementation of the alternating-direction method, and each time-step took 1.16 seconds to compute.

Flow around a spherical obstacle. The planar flow of a polymeric fluid around a cylindrical obstacle in a channel has also been a popular benchmark problem in the computational rheology literature (see Chap. 9 of [31]). In this section we consider a three-dimensional analogue in which we solve the micro-macro model for a suspension of FENE dumbbells for the flow around a sphere with radius 1 in a three-dimensional channel with 4×4 square cross-section. In this case $\Omega \subset \mathbb{R}^3$ and $\Omega \times D \subset \mathbb{R}^6$. We set $b = 12$, $\mathrm{Wi} = 1$, $\gamma = 0.59$ and we used the Stokes equations for the macroscopic velocity field.

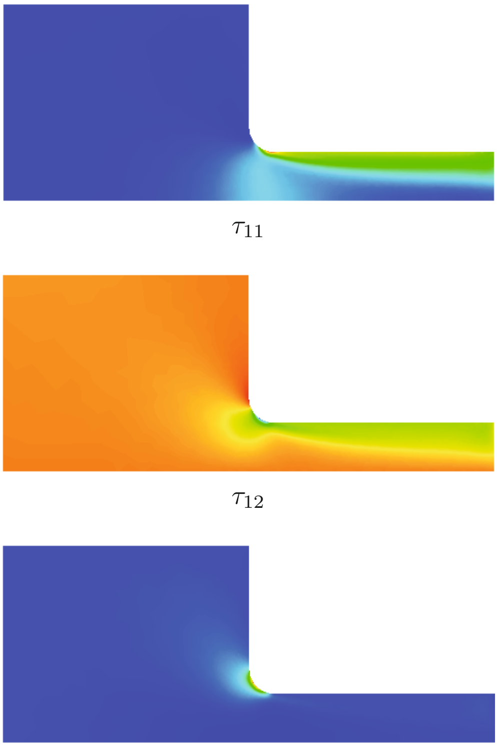
The mesh $\mathcal{T}_h$ is shown in Figure 7. We set $\underset{\sim}{u}_{\text{in}}$ to be the velocity profile corresponding to steady Stokes flow in a channel with square cross-section, with $U_{\max} = 1$. We also imposed a zero Neumann boundary condition on $\partial\Omega_{\text{out}}$, a no-slip boundary condition on the channel walls and on the spherical obstacle, and we set two symmetry boundary conditions so that we only needed to simulate the flow in one quarter of the domain.

[12] We used six quadrature points per element (with no points on element boundaries) in order to satisfy QH2, *cf.* [28].



(a) (b)

Figure 5. (a) The finite element mesh $\mathcal{T}_h$ used for the contraction flow computations. $\mathcal{T}_h$ contains 905 triangular elements. (b) Streamlines for the macroscopic velocity field; this corresponds closely to Figure 8.9 in [31], which shows computational results for planar contraction flows obtained using the fully macroscopic Oldroyd B model.



τ_{11}

τ_{12}

τ_{22}

Figure 6. The components of $\underset{\approx}{\tau}_{h,N}$ at $T = 5$. In the τ_{11} plot, values range from 0.45 (blue) to 15.7 (red), in the τ_{12} $(= \tau_{21})$ plot we have -9.75 (blue) to 1.41 (red) and in the τ_{22} plot, 0.46 (blue) to 11.5 (red). The polymeric extra-stress is largest in the region near the rounded corner (figure in colour available online at http://www.esaim-m2an.org/).

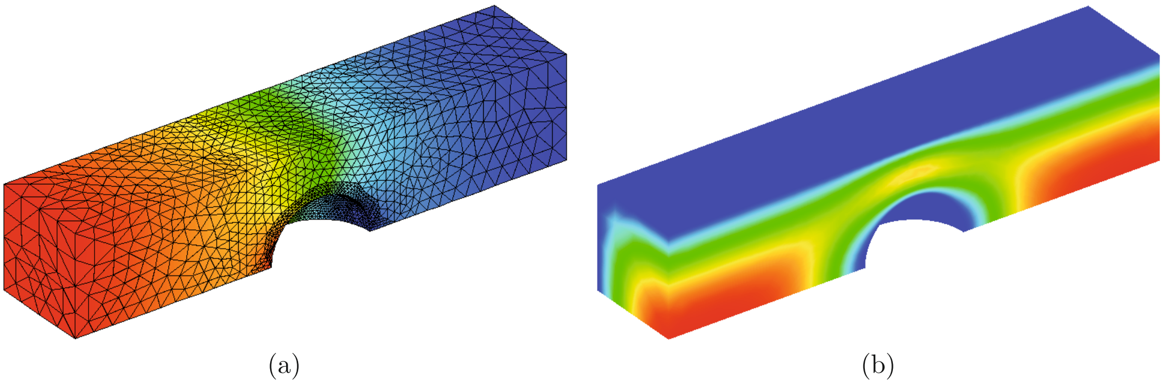


Figure 7. (a) Plot of the pressure, $p_h \in P_h$, at $T = 1$, with values ranging from 0.5 (blue) to 14.4 (red). Also, this plot shows the mesh $\mathcal{T}_h$. Note that the mesh is very fine in the vicinity of the spherical obstacle in order to resolve the solution structure in that region. (b) The x-component of the macroscopic velocity field at $T = 1$; values range from 0 (blue) to 1 (red) (figure in colour available online at `http://www.esaim-m2an.org/`).

The mesh $\mathcal{T}_h$ contains 5150 tetrahedral elements and we used $Q_\Omega = 72\,100$[13]. For the $\underset{\sim}{q}$-direction spectral method we used basis $\mathcal{C}$ with $(N_r, N_{\text{sph}}) = (12, 12)$, so that $N_D = 1092$. We took 100 time-steps with $\Delta t = 0.01$ using method II to reach $T = 1$ and in this case $\text{Wi}\,\|\underset{\approx}{\kappa}\|_{\mathrm{L}^\infty(\Omega)} \approx 5$. Plots of the x-component of $\underset{\sim}{u}_h$ and of p_h at $T = 1$ are shown in Figure 7. Also, the components of the polymeric extra-stress tensor at $T = 1$ are shown in Figure 8. This computation was performed with $N_{\text{proc}} = 128$ and it took 38.7 seconds to evaluate each time-step of the coupled Stokes–Fokker–Planck system.

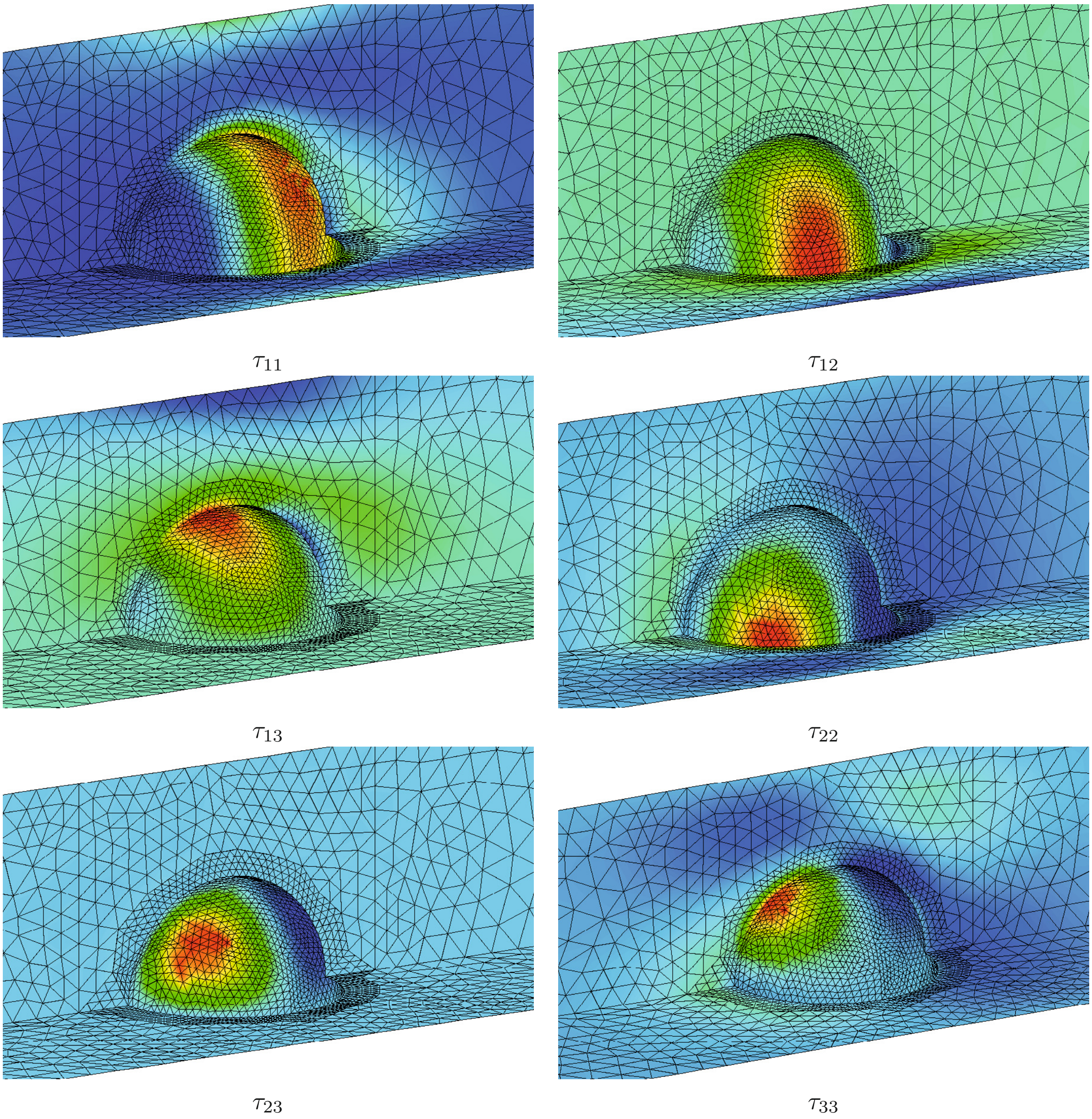
6. Conclusions

We have presented a range of theoretical results for alternating-direction methods for the Fokker–Planck equation from polymeric fluid dynamics. This analysis built upon the results in [22] for the Fokker–Planck equation in configuration space. We proved rigorous stability results and, in the case of method I, an *a priori* convergence estimate. We also established an important property related to the superconvergence of $\underset{\approx}{\tau}$ with respect to the $\underset{\sim}{q}$-direction spectral method; this property is extremely beneficial for practical computations with the micro-macro model since in that context the accuracy of $\underset{\approx}{\tau}$ is of primary importance.

We developed a computational framework for the Fokker–Planck equation that is efficient in practice and is also underpinned by rigorous theoretical analysis. We coupled the alternating-direction method to a mixed finite element method for the Navier–Stokes or Stokes equations to obtain an algorithm for the full micro-macro model for dilute polymeric fluids. We used this algorithm to obtain computational results for some channel flow problems of physical interest. Parallel computation was used in order to significantly reduce the computation time that was required for these problems and this made large-scale problems (such as the flow around a sphere problem considered in Sect. 5.2) computationally feasible. To the best of our knowledge our computation for the flow around a sphere problem is the first ever numerical implementation of the fully coupled micro-macro model in a case where $\Omega \times D \subset \mathbb{R}^6$.

An important challenge for future work is to consider the extension of the deterministic multiscale framework developed here for FENE dumbbells to the case of dumbbell chains, for which the configuration space is higher-dimensional. This topic has been considered using reduced basis methods in [1,2] and sparse grid methods in [13]

[13] We used 14 quadrature points per element (with no points on element boundaries) in order to satisfy QH2, *cf.* [35].



τ_{11}

τ_{12}

τ_{13}

τ_{22}

τ_{23}

τ_{33}

FIGURE 8. Plots of the components of the polymeric extra-stress tensor, $\underset{\approx}{\tau}_{h,N}$, at $T = 1$ for the channel flow around a spherical obstacle. The minimum (blue) and maximum (red) values in each plot are as follows; τ_{11}: 0.53 to 6.25, τ_{12}: -1.25 to 2.41, τ_{13}: -1.23 to 2.54, τ_{22}: 0.48 to 3.35, τ_{23}: -0.33 to 1.15 and τ_{33}: 0.47 to 3.46 (figure in colour available online at http://www.esaim-m2an.org/).

(see also [32]) for the Fokker–Planck equation in configuration space, and it is plausible that the alternating-direction framework developed here would be an effective approach for extending these reduced basis or sparse grid approaches to the full Fokker–Planck equation for dumbbell chains.

Acknowledgements. We would like to thank Professor Graham Carey of the University of Texas at Austin for his assistance with parallel computation. We would also like to thank Dr. Alexei Lozinski of the Institut de Mathématiques de Toulouse, and Dr. David Kay of the University of Oxford for their helpful comments.

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] S. Balay, K. Buschelman, V. Eijkhout, W.D. Gropp, D. Kaushik, M.G. Knepley, L.C. McInnes, B.F. Smith and H. Zhang, *PETSc users manual.* Tech. Rep. ANL-95/11 – Revision 2.1.5, Argonne National Laboratory (2004).

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

[5] B. Bialecki and R. Fernandes, An orthogonal spline collocation alternating direction implicit Crank-Nicolson method for linear parabolic problems on rectangles. *SIAM J. Numer. Anal.* **36** (1999) 1414–1434.

[6] P.B. Bochev, M.D. Gunzburger and J.N. Shadid, Stability of the SUPG finite element method for transient advection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.* **193** (2004) 2301–2323.

[7] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods.* Second Edn., Springer (2002).

[8] M. Celia and G. Pinder, An analysis of alternating-direction methods for parabolic equations. *Numer. Methods Part. Differ. Equ.* **1** (1985) 57–70.

[9] M. Celia and G. Pinder, Generalized alternating-direction collocation methods for parabolic equations. I. Spatially varying coefficients. *Numer. Methods Partial Differ. Equ.* **3** (1990) 193–214.

[10] 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.

[11] C. Chauvière and A. Lozinski, Simulation of dilute polymer solutions using a Fokker–Planck equation. *Comput. Fluids* **33** (2004) 687–696.

[12] P. Clément, Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Anal. Numér.* **9** (1975) 77–84.

[13] P. Delaunay, A. Lozinski and R.G. Owens, Sparse tensor-product Fokker–Planck-based methods for nonlinear bead-spring chain models of dilute polymer solutions. *CRM Proc. Lect. Notes* **41** (2007) 73–89.

[14] J. Douglas and T. Dupont, Alternating-direction Galerkin methods on rectangles. *Numer. Solution Partial Differ. Equ. II (SYNSPADE 1970)* (1971) 133–214.

[15] H. Eisen, W. Heinrichs and K. Witsch, Spectral collocation methods and polar coordinate singularities. *J. Comput. Phys.* **96** (1991) 241–257.

[16] H. Elman, D. Silvester and A. Wathen, *Finite elements and fast iterative solvers*. Oxford Science Publications, UK (2005).

[17] C. Helzel and F. Otto, Multiscale simulations of suspensions of rod-like molecules. *J. Comp. Phys.* **216** (2006) 52–75.

[18] W. Huang and B. Guo, Fully discrete Jacobi-spherical harmonic spectral method for Navier-Stokes equations. *Appl. Math. Mech.* **29** (2008) 453–476 (English Ed.).

[19] B. Jourdain, T. Lelièvre and C. Le Bris, Existence of solution for a micro-macro model of polymeric fluid: the FENE model. *J. Funct. Anal.* **209** (2004) 162–193.

[20] B.S. Kirk, J.W. Peterson, R.M. Stogner and G.F. Carey, libMesh: A C++ library for parallel adaptive mesh refinement/coarsening simulations. *Eng. Comput.* **23** (2006) 237–254.

[21] D.J. Knezevic, *Analysis and implementation of numerical methods for simulating dilute polymeric fluids.* Ph.D. Thesis, University of Oxford, UK (2008), http://www.comlab.ox.ac.uk/people/David.Knezevic.

[22] D.J. Knezevic and E. Süli, Spectral Galerkin approximation of Fokker–Planck equations with unbounded drift. *ESAIM: M2AN* **43** (2009) 445–485.

[23] A.N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104** (1931).

[24] T. Li and P. Zhang, Mathematical analysis of multi-scale models of complex fluids. *Commun. Math. Sci.* **5** (2007) 1–51.

[25] C. Liu and H. Liu, Boundary conditions for the microscopic FENE models. *SIAM J. Appl. Math.* **68** (2008) 1304–1315.

[26] A. Lozinski, *Spectral methods for kinetic theory models of viscoelastic fluids.* Ph.D. Thesis, École Polytechnique Fédérale de Lausanne, Suisse (2003).

[27] A. Lozinski and C. Chauvière, A fast solver for Fokker–Planck equation applied to viscoelastic flows calculation: 2D FENE model. *J. Computat. Phys.* **189** (2003) 607–625.

[28] J.N. Lyness and D. Jespersen, Moderate degree symmetric quadrature rules for the triangle. *J. Inst. Math. Appl.* **15** (1975) 19–32.

[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] R.G. Owens and T.N. Phillips, *Computational Rheology*. Imperial College Press (2002).

[32] C. Schwab, E. Süli and R.A. Todor, Sparse finite element approximation of high-dimensional transport-dominated diffusion problems. *ESAIM: M2AN* **42** (2008) 777–820.

[33] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* **54** (1990) 483–493.

[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.

[35] N.J. Walkington, Quadrature on simplices of arbitrary dimension. http://www.math.cmu.edu/~nw0z/publications/00-CNA-023/023abs/.

[36] H.R. Warner, Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. *Ind. Eng. Chem. Fundamentals* **11** (1972) 379–387.

[37] H. Zhang and P. Zhang, Local existence for the FENE-dumbbell model of polymeric fluids. *Arch. Ration. Mech. Anal.* **181** (2006) 373–400.