A SECOND-ORDER PARTITIONED METHOD FOR BIOCONVECTIVE FLOWS WITH CONCENTRATION DEPENDENT VISCOSITY

MADELINE EDWARDS *, MARTINA BUKAC †, AND CATALIN TRENCHEA ‡

Abstract. This work is focused on the mathematical and computational modeling of bioconvection, which describes the mixing of fluid and micro-organisms exhibiting negative geotaxis movement under the force of gravity. The collective population moves towards the surface of the fluid, generating a Rayleigh–Taylor instability, where initial fingers of organisms plummet to the bottom. The inherent drive to swim vertically generates large collective flow patterns that persist in time. We model the flow using the Navier-Stokes equations for an incompressible, viscous fluid, coupled with the transport equation describing the concentration of the micro-organisms. We use a nonlinear semigroup approach to prove the existence of solutions. We propose a partitioned, second-order, time adaptive numerical method based on the Cauchy's one-legged 'θ-like' scheme. We prove that the method is energy-stable, and for small time steps, the iterative procedure in the partitioned algorithm is linearly convergent. The numerical results confirm the expected second-order of accuracy. We also present a computational study of a chaotic system describing bioconvection of motile flagellates.

Key words. Bioconvection, partitioned method, adaptive time-stepping, chaotic flow

1. Introduction. The following work presents a mathematical and computational modeling of bioconvection, which describes the mixing of fluid and micro-organisms exhibiting negative geotaxis movement under the force of gravity. The collective population moves towards the surface of the fluid, generating instability at the top layer causing initial fingers of organisms to plummet to the bottom of the container. The inherent drive to swim vertically generates large collective flow patterns that persist in time. The physical system is mathematically modeled by the incompressible, viscous, Navier-Stokes equations, coupled with a transport equation for the concentration of the micro-organisms.

1.1. The mathematical model. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded domain with smooth boundary $\partial\Omega$. Let $c(\boldsymbol{x}, t)$ denote the concentration of microorganisms at point $\boldsymbol{x} = (x_1, \cdots, x_d) \in \Omega$, at time $t > 0$, and let $\boldsymbol{u} = \{u_j(\boldsymbol{x}, t)\}_{j=1}^d$ and $p(\boldsymbol{x}, t)$ denote the velocity and the pressure of the culture fluid.

The collective dynamics of negative geotactic micro-organisms immersed in a viscous fluid can be modelled by the system of partial differential equations, coupling the Navier–Stokes equations with a convective transport equation:

$$\frac{\partial \boldsymbol{u}}{\partial t}(\boldsymbol{x}, t) - \nabla \cdot \big(\nu(c(\boldsymbol{x}, t))D(\boldsymbol{u}(\boldsymbol{x}, t))\big) + (\boldsymbol{u}(\boldsymbol{x}, t) \cdot \nabla)\boldsymbol{u}(\boldsymbol{x}, t) + \nabla p(\boldsymbol{x}, t) = -g(1 + \gamma c(\boldsymbol{x}, t))\boldsymbol{i}_d + \boldsymbol{f}(\boldsymbol{x}, t), \quad (1.1)$$

$$(\nabla \cdot \boldsymbol{u})(\boldsymbol{x}, t) = 0, \quad (1.2)$$

$$\frac{\partial c}{\partial t}(\boldsymbol{x}, t) - \Theta \Delta c(\boldsymbol{x}, t) + (\boldsymbol{u}(\boldsymbol{x}, t) \cdot \nabla)c(\boldsymbol{x}, t) + U\frac{\partial c}{\partial x_d}(\boldsymbol{x}, t) = 0, \quad (1.3)$$

for all $\boldsymbol{x} \in \Omega, t > 0$. Here $D(\boldsymbol{u}) = \frac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)$ denotes the symmetric part of the stress tensor, $\boldsymbol{f}$ is a body force, g represents the acceleration due to gravity, and Θ and U are the diffusion rate and mean velocity of upward swimming of the micro-organism, respectively. The parameter $\gamma > 0$ represents the relative difference of the density ρ_0 of the micro-organism from that of the density ρ_m of the culture fluid, that is, $\gamma = \frac{\rho_0}{\rho_m} - 1$. The vector $\boldsymbol{i}_d = (0, \cdots, 1)^T$ denotes the vertical unit vector in $\mathbb{R}^d$, so that the term $-g(1 + \gamma c)\boldsymbol{i}_d$ captures the effect of gravity on organisms at the upper surface. The equations (1.1) and (1.2) are the Navier–Stokes type equations for the incompressible viscous culture fluid, and represent the conservation of momentum and conservation of mass for Newtonian fluids. The equation (1.3) describes the transport of micro-organisms, where $U > 0$, and the $U\frac{\partial c}{\partial x_d}$ term represents the effect of mean upward swimming of micro-organisms. The equation (1.3) can be written as a conservation equation:

$$\frac{d}{dt}c + \nabla \cdot J = 0, \qquad \boldsymbol{x} \in \Omega, t > 0,$$

where $\frac{d}{dt} = \frac{\partial}{\partial t} + (\boldsymbol{u} \cdot \nabla)$ is the derivative along the fluid particle, and $J = -\Theta \nabla c + Uc\boldsymbol{i}_d$ is the flux of micro-organisms.

---

*Department of Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260 USA. mme58@pitt.edu.

†Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556 USA. Email: mbukac@nd.edu.

‡Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 USA. Email: trenchea@pitt.edu.

The boundary conditions are of Dirichlet type for the fluid, and no-flux for the concentration, such that at each point $\boldsymbol{x} \in \partial\Omega, t > 0$:

$$\boldsymbol{u} = 0, \qquad \Theta \frac{\partial c}{\partial \boldsymbol{n}} - cU\boldsymbol{n}_d = 0, \tag{1.4}$$

where $\boldsymbol{n} = (n_1, \dots, n_d)$ denotes the exterior unit normal vector on $\partial\Omega$. We also prescribe initial conditions on $\boldsymbol{u}, p$, and c, that is, $\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{u}_0(\boldsymbol{x}), p(\boldsymbol{x}, 0) = p_0(\boldsymbol{x})$, and $c(\boldsymbol{x}, 0) = c_0(\boldsymbol{x})$, for all $\boldsymbol{x} \in \Omega$.

In an ideal Newtonian fluid, the viscosity ν is assumed to be constant. However, this assumption has been shown not to hold generally. In real-life suspensions it was observed that the viscosity $\nu = \nu(c)$ depends on the concentration of the micro-organisms, see e.g., [17] for a list of references with proven explicit expressions of $\nu(c)$. The nonlinear viscosity yields better approximations, especially for the bottom heavy micro-organisms, since the viscous drag force is the major reason for their upward swimming. Hence, we consider the 3D generalized bioconvection model whose viscosity depends on the concentration $\nu = \nu(c)$.

First we point out that the concentration is positive, $c(\boldsymbol{x}, t) > 0$ for all $\boldsymbol{x} \in \Omega, t > 0$. This is an immediate consequence of the maximum principle in [45, pp. 174], [5] (for its proof in the constant diffusion case, see Lemma 4.2 in [36, pp. 141]). We also note that the no-flux boundary condition for the concentration (1.4) ensures the conservation of the total mass of micro-organisms: $\alpha = \int_\Omega c(\boldsymbol{x}, t) d\boldsymbol{x} = \int_\Omega c_0(\boldsymbol{x}) d\boldsymbol{x}$, for any $t \geq 0$. Integrating (1.3) over Ω, we obtain:

$$0 = \frac{\partial}{\partial t} \int_\Omega c(\boldsymbol{x}, t) d\boldsymbol{x} - \int_\Omega \left(\Theta \Delta c(\boldsymbol{x}, t) - U \frac{\partial c}{\partial x_d}(\boldsymbol{x}, t) \right) d\boldsymbol{x} + \int_\Omega \boldsymbol{u}(\boldsymbol{x}, t) \cdot \nabla c(\boldsymbol{x}, t) d\boldsymbol{x},$$

hence the total mass of micro-organisms is constant in time:

$$\frac{\partial}{\partial t} \int_\Omega c(\boldsymbol{x}, t) d\boldsymbol{x} = 0, \qquad \forall t > 0. \tag{1.5}$$

Indeed, the second term in the right-hand side above (1.5) vanishes, by using the divergence theorem and the no-flux condition in (1.4). The third terms also vanishes, by applying the incompressibility condition (1.2) and the homogeneous Dirichlet boundary condition for the velocity.

Let us define the kinetic energy of the fluid flow, and respectively the potential energy of the system by

$$\text{KE}(t) = \frac{1}{2} \int_\Omega |\boldsymbol{u}(\boldsymbol{x}, t)|^2 d\boldsymbol{x}, \qquad \text{PE}(t) = \alpha g \int_\Omega c(\boldsymbol{x}, t) \cdot (x_d - x_d^{\min}) d\boldsymbol{x},$$

where $x_d^{\min}$ is the 'bottom' of Ω in the vertical direction. Testing the Navier-Stokes equation (1.1) with $\boldsymbol{u}(\boldsymbol{x}, t)$, and using the incompressibility condition (1.2) and the homogeneous boundary conditions on the flow, we obtain

$$\frac{1}{2} \frac{d}{dt} \int_\Omega |\boldsymbol{u}(\boldsymbol{x}, t)|^2 d\boldsymbol{x} + \int_\Omega \nu(c)|\nabla \boldsymbol{u}|^2 d\boldsymbol{x} = \int_\Omega -g(1 + \gamma c(\boldsymbol{x}, t))\boldsymbol{i}_d \boldsymbol{u} d\boldsymbol{x} + \int_\Omega \boldsymbol{f}(\boldsymbol{x}, t)\boldsymbol{u} d\boldsymbol{x}.$$

Similarly, testing the transport equation (1.3) with $\gamma g(x_d - x_d^{\min})$ gives

$$0 = \frac{d}{dt} \alpha g \int_\Omega c(\boldsymbol{x}, t)(x_d - x_d^{\min}) d\boldsymbol{x} + \gamma g \int_\Omega \Theta \frac{\partial c}{\partial x_d} d\boldsymbol{x} - \gamma g \int_\Omega c\boldsymbol{u} \cdot \boldsymbol{i}_d d\boldsymbol{x} - \gamma g U \int_\Omega c d\boldsymbol{x},$$

and therefore we have the total energy balance equation:

$$\frac{d}{dt}(\text{KE} + \text{PE}) + \int_\Omega \nu(c)|\nabla \boldsymbol{u}|^2 d\boldsymbol{x} = \int_\Omega \boldsymbol{f}(\boldsymbol{x}, t)\boldsymbol{u} d\boldsymbol{x} + g \int_\Omega \left(\gamma(Uc - \Theta \nabla c \cdot \boldsymbol{i}_d) - \boldsymbol{u} \cdot \boldsymbol{i}_d \right) d\boldsymbol{x}. \tag{1.6}$$

1.2. Previous work. Within the works of [27, 37], the problem of bioconvection is extended to two-dimensions with a rigid bottom, lateral boundaries, and a stress-free top. In the context of wide chambers under these conditions, multiple plumes form and are periodic in the axial direction. The stability of a single plume is considered and studied. The Rayleigh number is modified to be independent of the height of the chamber. The system is discretized using a conservative finite-difference scheme on a staggered mesh. The stability of the system and single plume is analyzed across changes in aspect ratio (grid size), swimming speed of the micro-organisms, and gryotaxis number. The results

are consistent with the findings of [19, 20, 31, 42] where instability of the top layer forms as micro-organisms collect in the top layer of the fluid. These instabilities are observed as potential energy converts into kinetic energy [31].

The study of bioconvection and stability is extended to periodic conditions [28, 29, 37] and periodic solutions [21]. The existence and uniqueness of a 3D time-periodic solution is proven for weak solutions [21]. The three-dimensional chamber allows for the formation of bioconvective plumes which become unstable as depth increases. This instability causes 'blobs' to appear periodically around the plume [28, 29]. The wavelength of the periodic plumes is analyzed across changes of concentration, diffusion coefficient, and depth of chamber impacting the resulting bioconvective patterns [29]. The wavelength is found to be most sensitive to the ratio of swimming speed and diffusivity but less sensitive to changes in depth of the chamber.

In the case of viscosity depending on the concentration of particles, Einstein derived the theoretical expression for viscosity of small particles at relatively low concentrations in a suspension. This expression holds under the assumption that particles do not interact due to low levels of concentration. However, in the work by Mooney [44], this model is extended to higher concentrations of rigid spherical particles, which provide good agreement with experimental data. In particular, Mooney investigates the crowding effect of two-component rigid spheres in a finite volume, where the size and concentration of partial volume occupied by component one crowds component two. Experimental data for increases in concentration of uniform sphere sizes relative to viscosity are in good agreement with the analytical findings of this work, but the experimental comparisons of concentration due to varying particle sizes or particle interaction is limited. The author concludes that the crowding effect of particle interaction with changes in concentration can be analytically described as basic geometric crowding of spheres.

Following the motivation to extend Einstein's equation of a dilute suspension of rigid particles to higher concentrations, the work of [25] finds that Einstein's formula can only be extended to the situation when the confinement, which is defined as the ratio of particle size width with respect to the width of the square channel, is less than 0.02. By using a two-way coupling method of second-order accuracy, the authors find that the differences of the *relative viscosity* compared to *viscosity calculated* from Einstein's formula increases as confinement increases. This finding indicates the discrepancy and inability for Einstein's equation to accurately capture measurements at high concentration levels. For non-Newtonian fluids, increasing Reynold's numbers changes the fluid's properties from thixotropy to dilatancy (shear-thinning to shear thickening) resulting from the outward particle mitigation. This property change occurs when the power-law index of the Reynold's number increases from less than one to greater than one.

The authors of [10] prove the existence and uniqueness of weak solutions for the generalized stationary system of bioconvection where viscosity changes with concentration. The system is constructed in three dimensions and assumes the upward flow of organisms to be small. The challenge of the generalized system results from nonlinear terms, which the authors are able to approximate using a Helmholtz decomposition.

In [17], the authors study the existence of weak solutions for bioconvective flow in two dimensions, using the Galerkin method to construct a sequence of approximating solutions within finite-dimensional subspaces. The uniqueness is proved assuming the uniform boundedness of the stress tensor $\|D(\boldsymbol{u})\|_{L^\infty(\Omega)} \le C$. The numerical experiments are performed using a backward-Euler time integration and finite element approximations.

The existence of a $3D$ solution for the system (1.1)-(1.3) with constant $\nu(c) = \nu_*$ was proven in [36], using a Galerkin approximation. The corresponding numerical approximations were performed in [32], using an implicit-explicit type method for time integration, combining the backward-forward Euler and the leapfrog scheme. Inverse problems of optimal control type for stationary and time dependent bioconvection were treated in [1, 23], with possible applications in forecasting the ecological state of atmosphere and ocean.

1.3. The focus of this work. This work focuses on the development of a partitioned, variable time-stepping method for the bioconvection problem. We consider a bioconvection model as in [17], consisting of a fluid equation and a transport equation, where the fluid viscosity depends on the concentration. We use a nonlinear semigroup approach to prove the existence of weak solutions for the continuous problem. In contrast with previous work by Cao et al. [10], we do not have a restriction on the up-swimming speed U. For the numerical solution of the problem, a novel partitioned scheme is presented, where the fluid and the transport problems are decoupled and sub-iteratively solved using the backward-Euler method with a fractional step until convergence. After that, the solutions are extrapolated and the time step is adapted. The backward-Euler discretization combined with extrapolation corresponds to the refactorized Cauchy's 'θ-like' method, which is symplectic and second-order accurate in time when $\theta = \frac{1}{2}$, in which case the midpoint method is obtained [16]. Application of this method to bioconvection, in combination with partitioning, raises new challenges in proving stability and convergence. Here we show that, under certain assumptions,

the sub-iterative process in the algorithm is linearly convergent, and that the proposed numerical method is stable. The time adaptivity is based on Milne's device and a modified Adams-Bashforth, second-order accurate method for the estimation of the local truncation error [13–16, 41]. The expected convergence results are verified in numerical examples on a benchmark problem based on the method of manufactured solutions. We also numerically investigate bioconvection patterns in a culture of motile flagellates, where the system gives rise to chaotic solutions.

The rest of this paper is structured as follows. The mathematical model is analyzed in Chapter 2. The numerical method is presented in Chapter 3, together with the stability and convergence analysis. Numerical examples are presented in Chapter 4. Finally, conclusions are drawn in Chapter 5.

2. A nonlinear semigroup approach to the bioconvection system. Throughout the article, we assume that the kinematic viscosity $v(\cdot) : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous and that there exist positive constants v_*, v^*, L such that:

$$v_* \leq v(c_1) \leq v^*, \qquad |v(c_1) - v(c_2)| \leq L|c_1 - c_2|, \qquad \forall c_1, c_2 \in \mathbb{R}. \tag{2.1}$$

The assumption (2.1) is relatively weak [21], and is justified by applications [8, 11, 24, 38, 44], with the lower bound v_* representing the viscosity of the fluid with no micro-organisms. The upper bound v^* is guaranteed to exist since the volume fraction of the micro-organisms in the container has a maximum value, due to the geometric structure of the spherical particles [27]. We recall the Poincaré inequality for the velocity (for Ω a subset bounded at least in one direction):

$$\|\boldsymbol{u}\| \leq C_\Omega \|\nabla \boldsymbol{u}\|, \qquad \forall \boldsymbol{u} \in H_0^1(\Omega), \tag{2.2}$$

and the Poincaré–Wirtinger inequality for the concentration (for Ω a Lipschitz domain, i.e., a bounded connected open subset with a Lipschitz boundary):

$$\|c - \alpha\| \leq C_\Omega \|\nabla c\|, \qquad \forall c \in H^1(\Omega) \tag{2.3}$$

where $\alpha = \frac{1}{|\Omega|} \int_\Omega c$.

Let $W^{m,p}(\Omega)$, for an integer $m \geq 1$ and $1 \leq p \leq \infty$, denote the Sobolev space of real valued functions (and their vector-valued counterparts) which are in $L^p(\Omega)$ together with their weak derivatives of order less than or equal to m. Let $H^m(\Omega) = W^{m,2}(\Omega)$, and denote by $\|\cdot\|_m$ both the norm in $H^m(\Omega)$ and $\left(H^m(\Omega)\right)^d$, and by $\|\cdot\|$ the $L^2(\Omega)$ and $\left(L^2(\Omega)\right)^d$ norms.

By a classical device due to J. Leray, the boundary value problem (1.1)-(1.3) can be written as an infinite-dimensional Cauchy problem in an appropriate function space on Ω. To this end, we introduce the following spaces:

$$H = \{\boldsymbol{u} \in (L^2(\Omega))^d; \nabla \cdot \boldsymbol{u} = 0\} \times L^2(\Omega), \qquad V = \{\boldsymbol{u} \in (H_0^1(\Omega))^d; \nabla \cdot \boldsymbol{u} = 0\} \times H^1(\Omega).$$

We define the operator $\mathscr{A} \in \mathscr{L}(V, V')$ by setting

$$\langle \mathscr{A}(\boldsymbol{u}_1, c_1), (\boldsymbol{u}_2, c_2)\rangle = \int_\Omega \left(v_* \nabla \boldsymbol{u}_1 \cdot \nabla \boldsymbol{u}_2 + \Theta \nabla c_1 \cdot \nabla c_2\right) d\boldsymbol{x} \tag{2.4}$$

for all $(\boldsymbol{u}_i, c_i) \in V$. The operator $\mathscr{A}$ is a self-adjoint, unbounded operator on H, with the domain $D(\mathscr{A}) = \{(\boldsymbol{u}, c) \in V \cap (H^2(\Omega))^{d+1}, \Theta \frac{\partial c}{\partial n} - U c n_d = 0 \text{ on } \partial\Omega\}$. We define also a continuous trilinear form $\mathscr{B}_0$ on $V \times V \times V$ by setting

$$\mathscr{B}_0\big((\boldsymbol{u}_1, c_1), (\boldsymbol{u}_2, c_2), (\boldsymbol{u}_3, c_3)\big) \tag{2.5}$$
$$= \int_\Omega \left((v(c_2) - v_*)\nabla \boldsymbol{u}_1 \cdot \nabla \boldsymbol{u}_3 + (\boldsymbol{u}_1 \cdot \nabla)\boldsymbol{u}_2 \boldsymbol{u}_3 + (\boldsymbol{u}_1 \cdot \nabla)c_2 c_3 + g(1 + \gamma c_1) i_d \boldsymbol{u}_3 + U \frac{\partial c_1}{\partial x_d} c_3\right) d\boldsymbol{x},$$

and a continuous bilinear operator $\mathscr{B} : V \to V'$ with

$$\langle \mathscr{B}(\boldsymbol{u}_1, c_1), (\boldsymbol{u}_2, c_2)\rangle = \mathscr{B}_0\big((\boldsymbol{u}_1, c_1), (\boldsymbol{u}_1, c_1), (\boldsymbol{u}_2, c_2)\big), \tag{2.6}$$

for all $(\boldsymbol{u}_i, c_i) \in V$. We recall (see e.g. [6, 22, 26, 43, 46])

$$\int_\Omega \left((\boldsymbol{u}_1 \cdot \nabla)\boldsymbol{u}_2 \boldsymbol{u}_3 + (\boldsymbol{u}_1 \cdot \nabla)c_2 c_3\right) d\boldsymbol{x} \leq C_\Omega \|(\boldsymbol{u}_1, c_1)\|_{m_1} \|(\boldsymbol{u}_2, c_2)\|_{m_2+1} \|(\boldsymbol{u}_3, c_3)\|_{m_3} \tag{2.7}$$

4

for all $(\boldsymbol{u}_1,c_1) \in H^{m_1}(\Omega), (\boldsymbol{u}_2,c_2) \in H^{m_2+1}(\Omega), (\boldsymbol{u}_3,c_3) \in H^{m_3}(\Omega)$, where

$$m_1 + m_2 + m_3 \geq \frac{d}{2}, \qquad \text{if } m_i \neq \frac{d}{2} \text{ for all } i = 1,\dots,d$$

and

$$m_1 + m_2 + m_3 > \frac{d}{2}, \qquad \text{if } m_i = \frac{d}{2} \text{ for any of } i = 1,\dots,d.$$

In particular, it follows by (2.7) that the operator $\mathscr{B}$ is continuous from V to V'. Indeed, we have

$$\langle \mathscr{B}(\boldsymbol{u}_1,c_1) - \mathscr{B}(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_3,c_3) \rangle$$
$$= \mathscr{B}_0\big((\boldsymbol{u}_1,c_1), (\boldsymbol{u}_1,c_1) - (\boldsymbol{u}_2,c_2), (\boldsymbol{u}_3,c_3)\big) - \mathscr{B}_0\big((\boldsymbol{u}_1,c_1) - (\boldsymbol{u}_2,c_2), (\boldsymbol{u}_2,c_2), (\boldsymbol{u}_3,c_3)\big)$$
$$\leq C_\Omega \big(\|(\boldsymbol{u}_1,c_1)\|_1 \|(\boldsymbol{u}_1,c_1) - (\boldsymbol{u}_2,c_2)\|_1 \|(\boldsymbol{u}_3,c_3)\|_1 + \|(\boldsymbol{u}_1,c_1) - (\boldsymbol{u}_2,c_2)\|_1 \|(\boldsymbol{u}_2,c_2)\|_1 \|(\boldsymbol{u}_3,c_3)\|_1 \big),$$

hence

$$\|\mathscr{B}(\boldsymbol{u}_1,c_1) - \mathscr{B}(\boldsymbol{u}_2,c_2)\|_{V'} \leq C_\Omega \|(\boldsymbol{u}_1,c_1) - (\boldsymbol{u}_2,c_2)\|_1 \big(\|(\boldsymbol{u}_1,c_1)\|_1 + \|(\boldsymbol{u}_2,c_2)\|_1 \big),$$

for all $(\boldsymbol{u}_1,c_1), (\boldsymbol{u}_2,c_2) \in V$.

The trilinear form $\mathscr{B}_0\big((\boldsymbol{u}_1,c_1), (\boldsymbol{u}_2,c_2), (\boldsymbol{u}_2,c_2)\big)$ is well defined and the following property holds:

$$\mathscr{B}_0\big((\boldsymbol{u}_1,c_1), (\boldsymbol{u}_2,c_2), (\boldsymbol{u}_2,c_2)\big) = \int_\Omega \Big((\nu(c_2) - \nu_*) \nabla \boldsymbol{u}_1 \cdot \nabla \boldsymbol{u}_2 + g(1 + \gamma c_1) i_d \boldsymbol{u}_2 + U \frac{\partial c_1}{\partial x_d} c_2 \Big) d\boldsymbol{x}. \qquad (2.8)$$

DEFINITION 2.1. *Let* $(\boldsymbol{f},0) \in L^2(0,T;V')$ *and* $(\boldsymbol{u}_0,c_0) \in H$. *Then* $(\boldsymbol{u},c) : [0,T] \to H$ *is said to be a weak solution to equation* (1.1)-(1.3) *if*

$$(\boldsymbol{u},c) \in L^2(0,T;V) \cap C_w([0,T];H) \cap W^{1,1}([0,T];V'), \qquad (2.9)$$

$$\frac{d(\boldsymbol{u},c)}{dt}(t) + \mathscr{A}(\boldsymbol{u},c)(t) + \mathscr{B}(\boldsymbol{u},c)(t) = (P\boldsymbol{f}(t),0), \qquad a.e., \ t \in (0,T), \qquad (2.10)$$

and $(\boldsymbol{u},c)(0) = (\boldsymbol{u}_0,c_0) \in H$.

Here $C_w([0,T];H)$ *denotes the space of weakly continuous function* $(\boldsymbol{u},c) : [0,T] \to H$, *while* $\frac{d(\boldsymbol{u},c)}{dt}$ *is the strong derivative of* $(\boldsymbol{u},c) : [0,T] \to V'$, *and* $P : (L^2(\Omega))^d \to \{\boldsymbol{u} \in (L^2(\Omega))^d; \nabla \cdot \boldsymbol{u} = 0\}$ *is the Hodge projection.*

Also $(\boldsymbol{u},c)$ *is said to be a strong solution to equation* (1.1)-(1.3) *if* $(\boldsymbol{u},c) \in W^{1,1}([0,T];H) \cap L^2(0,T;D(\mathscr{A}))$ *and* (2.10) *holds with* $\frac{d(\boldsymbol{u},c)}{dt} \in L^1(0,T;H)$ *being the strong derivative of* $(\boldsymbol{u},c) : [0,T] \to H$.

We would like to treat (2.10) as a nonlinear Cauchy problem in the space H. The proof for the existence of weak solutions is based on the semigroup approach proposed in [4, 7, 39] for the Navier–Stokes equations. Because the operator $\mathscr{A} + \mathscr{B}$ is not quasi-m-accretive in H, we first consider a quasi-m-accretive approximation of the operator $\mathscr{B}$. For each $M > 0$, we define the following modified nonlinearity $\mathscr{B}_M : V \to V$ by setting

$$\mathscr{B}_M(\boldsymbol{u},c) = \begin{cases} \mathscr{B}(\boldsymbol{u},c), & \text{if } \|(\boldsymbol{u},c)\|_{H^1(\Omega)} \leq M, \\ \Big(\frac{M}{\|(\boldsymbol{u},c)\|_{H^1(\Omega)}} \Big)^2 \mathscr{B}(\boldsymbol{u},c), & \text{if } \|(\boldsymbol{u},c)\|_{H^1(\Omega)} > M, \end{cases} \qquad (2.11)$$

and consider the operator $\mathscr{G}_M : D(\mathscr{G}_M) \subset H \to H$:

$$\mathscr{G}_M = \mathscr{A} + \mathscr{B}_M, \qquad D(\mathscr{G}_M) = D(\mathscr{A}). \qquad (2.12)$$

The operator $\mathscr{G}_M$ is well-defined:

$$\|\mathscr{G}_M(\boldsymbol{u},c)\| \leq \|\mathscr{A}(\boldsymbol{u},c)\| + C_\Omega \|(\boldsymbol{u},c)\|_1^{3/2} \|(\boldsymbol{u},c)\|_2^{1/2} + (\nu(c) - \nu_*) \|(\boldsymbol{u},c)\|_2 + g|\Omega|^{1/2} + g\gamma\|c\| + U\|\nabla c\|, \quad (2.13)$$

for all $(\boldsymbol{u},c) \in D(\mathscr{A})$ (see Lemma A.1 in Appendix A), and moreover $\mathscr{G}_M$ is quasi-m-accretive [3] (see Lemma A.3 in Appendix A). Now, for each $M > 0$, we consider the equation

$$\frac{d(\boldsymbol{u},c)}{dt}(t) + \mathscr{A}(\boldsymbol{u},c)(t) + \mathscr{B}_M(\boldsymbol{u},c)(t) = (P\boldsymbol{f}(t),0), \qquad \text{a.e. } t \in (0,T), \tag{2.14}$$

and $(\boldsymbol{u},c)(0) = (\boldsymbol{u}_0,c_0)$.

PROPOSITION 2.2. *Let $(\boldsymbol{u}_0,c_0) \in D(\mathscr{A})$ and $(\boldsymbol{f},0) \in W^{1,1}([0,T];H)$ be given. Then there exists a unique solution $(\boldsymbol{u}_M,c_M) \in W^{1,\infty}([0,T;H]) \cap L^\infty(0,T;D(\mathscr{A})) \cap C([0,T];V)$ to (2.14). Moreover, $\frac{d^+}{dt}(\boldsymbol{u}_M,c_M)(t)$ exists for all $t \in [0,T)$ and*

$$\frac{d^+}{dt}(\boldsymbol{u}_M,c_M)(t) + \mathscr{A}(\boldsymbol{u}_M,c_M)(t) + \mathscr{B}_M(\boldsymbol{u}_M,c_M)(t) = (P\boldsymbol{f}(t),0), \qquad \forall t \in [0,T). \tag{2.15}$$

Proof. The conclusion follows from a classical result on existence and uniqueness of strong solutions for Cauchy problems in reflexive Banach, or uniformly convex spaces, with quasi-m-accretive operators see [3, Theorem 1.16]. Since $\mathscr{G}_M(\boldsymbol{u}_M,c_M) \in L^\infty(0,T;H)$, we derive that $\frac{d(\boldsymbol{u}_M,c_M)}{dt} \in L^\infty(0,T;H)$, and therefore $(\boldsymbol{u}_M,c_M) \in C([0,T];V)$. □

We state now the main existence result for the strong solutions to (1.1)-(1.3). The result follows from the fact that for M sufficiently large, the solution $(\boldsymbol{u}_M,c_M)(t)$, defined by Proposition 2.2, is independent of M on each interval $[0,T]$ if $d = 2$, or on $[0,T_{(\boldsymbol{u}_0,c_0)}]$ if $d = 3$.

THEOREM 2.3. *Assume (2.1) and $(\boldsymbol{f},0) \in W^{1,1}([0,T];H)$, $(\boldsymbol{u}_0,c_0) \in D(\mathscr{A})$. Then there exists a unique solution $(\boldsymbol{u},c) \in W^{1,\infty}([0,T^*];H) \cap L^\infty(0,T^*;D(\mathscr{A})) \cap C([0,T^*];V)$ such that*

$$\frac{d(\boldsymbol{u},c)}{dt}(t) + \mathscr{A}(\boldsymbol{u},c)(t) + \mathscr{B}(\boldsymbol{u},c)(t) = (P\boldsymbol{f}(t),0), \qquad a.e., \ t \in (0,T^*), \tag{2.16}$$

and $(\boldsymbol{u},c)(0) = (\boldsymbol{u}_0,c_0)$, for some $T^* = T^*_{(\boldsymbol{u}_0,c_0)} \leq T$. In the $d = 2$ case, $T^* = T$.
Moreover, $(\boldsymbol{u},c)(t)$ is right differentiable and

$$\frac{d^+(\boldsymbol{u},c)}{dt}(t) + \mathscr{A}(\boldsymbol{u},c)(t) + \mathscr{B}(\boldsymbol{u},c)(t) = (P\boldsymbol{f}(t),0), \qquad \forall t \in [0,T^*).$$

Proof. Let $M > 0$ be fixed, but arbitrary. We first test equation (2.15) with $(\boldsymbol{0},c_M)$, and then with $(\boldsymbol{u}_M,0)$, to obtain:

$$\|\boldsymbol{u}_M(t)\|^2 + \|c_M(t)\|^2 + \int_0^t \|\nu(c_M)^{1/2}\nabla\boldsymbol{u}_M(s)\|^2 ds + \int_0^t \Theta\|\nabla c_M(s)\|^2 ds \tag{2.17}$$

$$\leq 4g^2 C_\Omega^2 \frac{1}{\nu_*}|\Omega|t + 4g^2 C_\Omega^2 \frac{1}{\nu_*}\gamma^2\|c_0\|^2 \frac{\Theta}{U^2}\Big(\exp(\frac{U^2}{\Theta}t)-1\Big) + \|\boldsymbol{u}_0\|^2 + \|c_0\|^2\exp(\frac{U^2}{\Theta}t) + \frac{2}{\nu_*}\int_0^t \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds,$$

where we used (2.4), (2.8), the Poincaré inequality (2.2), and the Grönwall inequality.

Recall that by Proposition 2.2, $(\boldsymbol{u}_M,c_M) \in L^\infty(0,T;D(\mathscr{A}))$. We now test (2.15) with $(-\Delta\boldsymbol{u}_M,-\Delta c_M)$ to obtain:

$$\|\nabla\boldsymbol{u}_M(t)\|^2 + \|\nabla c_M(t)\|^2 + \int_0^t \Big(\|\nu(c_M)^{1/2}\Delta\boldsymbol{u}_M(s)\|^2 + \Theta\|\Delta c_M(s)\|^2\Big)ds \tag{2.18}$$

$$\leq 2\int_0^t \int_\Omega \Big((\boldsymbol{u}_M \cdot \nabla)\boldsymbol{u}_M\Delta\boldsymbol{u}_M + (\boldsymbol{u}_M \cdot \nabla)c_M\Delta c_M\Big)d\boldsymbol{x}ds + 2g^2\frac{1}{\nu_*}\int_0^t \|1+\gamma c_M(s)\|^2 ds$$

$$+ \frac{U^2}{\Theta}\int_0^t \Big\|\frac{\partial c_M(s)}{\partial x_d}\Big\|^2 ds + \|\nabla\boldsymbol{u}_0\|^2 + \|\nabla c_0\|^2 + \frac{2}{\nu_*}\int_0^t \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds.$$

In order to prove that the $L^\infty(0,T;H^1(\Omega))$ norm of $(\boldsymbol{u}_M,c_M)$ is independent of M, we treat separately the spatial dimensions $d = 2$ and $d = 3$.

(i) For $d = 3$, by (2.7) and the interpolation inequality, we have

$$\int_\Omega \Big((\boldsymbol{u}_M \cdot \nabla)\boldsymbol{u}_M \Delta \boldsymbol{u}_M + (\boldsymbol{u}_M \cdot \nabla)c_M \Delta c_M \Big) d\boldsymbol{x}$$
$$\leq \frac{\nu_*}{2}\|\Delta \boldsymbol{u}_M\|^2 + \frac{\Theta}{2}\|\Delta c_M\|^2 + C_0 \big(\|\nabla \boldsymbol{u}_M\|^6 + \|\nabla c_M\|^6 \big) + C \big(\|\boldsymbol{u}_M\|^2 + \|c_M\|^2 + \|c_M\|^3 + \|c_M\|^6 \big),$$

which by (2.18) gives

$$\|\nabla \boldsymbol{u}_M(t)\|^2 + \|\nabla c_M(t)\|^2 + \frac{1}{2}\int_0^t \Big(\nu(c_M)\|\Delta \boldsymbol{u}_M(s)\|^2 + \Theta \|\Delta c_M(s)\|^2 \Big) ds \leq \phi(t),$$

where

$$\phi(t) := C_0 \int_0^t \Big(\|\nabla \boldsymbol{u}_M(s)\|^6 + \|\nabla c_M(s)\|^6 \Big) ds$$
$$+ C\int_0^T \big(\|\boldsymbol{u}_M(s)\|^2 + \|c_M(s)\|^2 + \|c_M(s)\|^3 + \|c_M(s)\|^6 \big) ds + 2g^2 \frac{1}{\nu_*}\int_0^T \|1 + \gamma c_M(s)\|^2 ds$$
$$+ \frac{U^2}{\Theta}\int_0^T \Big\| \frac{\partial c_M(s)}{\partial x_d} \Big\|^2 ds + \|\nabla \boldsymbol{u}_0\|^2 + \|\nabla c_0\|^2 + \frac{2}{\nu_*}\int_0^T \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds.$$

We note that $\phi(t)$ satisfies the integral inequality

$$\phi'(t) \leq 2C_0 \phi^3(t), \qquad \forall t \in (0,T),$$

and

$$\phi(0) := C_0 \int_0^T \big(\|\boldsymbol{u}_M(s)\|^2 + \|c_M(s)\|^2 + \|c_M(s)\|^3 + \|c_M(s)\|^6 \big) ds + 2g^2 \frac{1}{\nu_*}\int_0^T \|1 + \gamma c_M(s)\|^2 ds$$
$$+ \frac{U^2}{\Theta}\int_0^T \Big\| \frac{\partial c_M(s)}{\partial x_d} \Big\|^2 ds + \|\nabla \boldsymbol{u}_0\|^2 + \|\nabla c_0\|^2 + \frac{2}{\nu_*}\int_0^T \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds.$$

Moreover, by (2.17), we have that $\phi(0)$ is uniformly bounded above:

$$\phi(0) \leq C_1 < \infty, \qquad \forall M > 0.$$

Therefore,

$$\|\nabla \boldsymbol{u}_M(t)\|^2 + \|\nabla c_M(t)\|^2 + \frac{1}{2}\int_0^t \Big(\nu(c_M)\|\Delta \boldsymbol{u}_M(s)\|^2 + \Theta \|\Delta c_M(s)\|^2 \Big) ds \tag{2.19}$$
$$\leq \Big(\frac{\phi^2(0)}{1 - 4C_0 t\phi^2(0)} \Big)^{1/2} \leq \Big(\frac{C_1^2}{1 - 4C_0 C_1^2 t} \Big)^{1/2}, \qquad \forall t \in (0,T^*),$$

where

$$T^* = \frac{1}{4C_0 C_1^2} \leq \frac{1}{4C_0 \phi^2(0)}.$$

(ii) For $d = 2$, we similarly have

$$\int_\Omega \Big((\boldsymbol{u}_M \cdot \nabla)\boldsymbol{u}_M \Delta \boldsymbol{u}_M + (\boldsymbol{u}_M \cdot \nabla)c_M \Delta c_M \Big) d\boldsymbol{x} \leq \frac{\nu_*}{2}\|\Delta \boldsymbol{u}_M\|^2 + \frac{\Theta}{2}\|\Delta c_M\|^2 + C\|\boldsymbol{u}_M\|^2 \|\nabla \boldsymbol{u}_M\|^4$$
$$+ C\Big(\|\boldsymbol{u}_M\|^2 \|\nabla c_M\|^4 + \|\boldsymbol{u}_M\|^2 \|\nabla \boldsymbol{u}_M\|^2 + \|c_M\|^4 + \|c_M\|^2 \|\nabla c_M\|^2 + \|\boldsymbol{u}_M\|^2 \|c_M\|^4 \Big),$$

7

which by (2.18) gives

$$\|\nabla \boldsymbol{u}_M(t)\|^2 + \|\nabla c_M(t)\|^2 + \frac{1}{2}\int_0^t \Big(\nu(c_M)\|\Delta \boldsymbol{u}_M(s)\|^2 + \Theta\|\Delta c_M(s)\|^2\Big)ds$$

$$\leq C\int_0^t \Big(\|\boldsymbol{u}_M(s)\|^2\|\nabla\boldsymbol{u}_M(s)\|^4 + \|\boldsymbol{u}_M(s)\|^2\|\nabla c_M(s)\|^4\Big)ds$$

$$+ C\int_0^T \Big(\|\boldsymbol{u}_M(s)\|^2\|\nabla\boldsymbol{u}_M(s)\|^2 + \|c_M(s)\|^4 + \|c_M(s)\|^2\|\nabla c_M(s)\|^2 + \|\boldsymbol{u}_M(s)\|^2\|c_M(s)\|^4\Big)ds$$

$$+ 2g^2\frac{1}{\nu_*}\int_0^T \|1 + \gamma c_M(s)\|^2 ds + \frac{U^2}{\Theta}\int_0^T \Big\|\frac{\partial c_M(s)}{\partial x_d}\Big\|^2 ds + \|\nabla\boldsymbol{u}_0\|^2 + \|\nabla c_0\|^2$$

$$+ \frac{2}{\nu_*}\int_0^T \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds.$$

By (2.17) and the Grönwall inequality this yields

$$\|\nabla \boldsymbol{u}_M(t)\|^2 + \|\nabla c_M(t)\|^2 + \frac{1}{2}\int_0^t \Big(\nu(c_M)\|\Delta \boldsymbol{u}_M(s)\|^2 + \Theta\|\Delta c_M(s)\|^2\Big)ds \qquad (2.20)$$

$$\leq C\Big(\|\nabla\boldsymbol{u}_0\|^2 + \|\nabla c_0\|^2 + \int_0^T \|(P\boldsymbol{f}(s),0)\|_{V'}^2 ds\Big), \qquad \forall t \in (0,T).$$

In conclusion, from (2.19) and (2.20) we infer that for M large enough, $\|(\boldsymbol{u}_M,c_M)\|_1 \leq M$ on $(0,T^*)$ if $d = 3$, or on $(0,T)$ if $d = 2$. Hence, by definition (2.11), we have that $\mathscr{B}_M(\boldsymbol{u}_M,c_M) = \mathscr{B}(\boldsymbol{u}_M,c_M)$ on $(0,T^*)$ (respectively on $(0,T)$), and consequently $(\boldsymbol{u}_M,c_M) \equiv (\boldsymbol{u},c)$ is a solution to (2.16). This completes the proof of existence. The uniqueness is immediate. $\square$

We state now the existence of a weak solution for (1.1)-(1.3).

THEOREM 2.4. *Assume* (2.1). *For any initial condition* $(\boldsymbol{u}_0,c_0) \in H$ *and* $(\boldsymbol{f},0) \in L^2(0,T;V')$ *there is at least one weak solution* $\boldsymbol{u} \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$, $c \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$ *to* (1.1)-(1.3), *given by*

$$(\boldsymbol{u},c) = w - \lim_{m\to\infty}(\boldsymbol{u}_{M_m},c_{M_m}) \quad \text{in } L^2(0,T;V), \text{ weak-star in } L^\infty(0,T;H),$$

$$\frac{d}{dt}(\boldsymbol{u},c) = w - \lim_{m\to\infty}\frac{d}{dt}(\boldsymbol{u}_{M_m},c_{M_m}) \quad \text{in } L^{\frac{4}{3}}(0,T;V'),$$

for some $m \to \infty$.

Proof. By (2.5) and (2.7) we have

$$\|\mathscr{B}_M(\boldsymbol{u}_M,c_M)\|_{V'}$$

$$\leq C\Big(\|\boldsymbol{u}_M\|^{1/2}\|\nabla\boldsymbol{u}_M\|^{3/2} + \|\boldsymbol{u}_M\|^{1/2}\|\nabla\boldsymbol{u}_M\|^{1/2}\|\nabla c_M\| + \|\nabla\boldsymbol{u}_M\| + \|\nabla c_M\| + g\|1 + \gamma c_M\|\Big),$$

and therefore by (2.17) we obtain

$$\int_0^T \Big\|\frac{d}{dt}(\boldsymbol{u}_M,c_M)\Big\|_{V'}^{\frac{4}{3}}dt \leq C.$$

This allows passing to the limit, on subsequences, in (2.15). $\square$

3. Numerical method. For the semi-discretization in time, let $\{t_n\}_{0 \leq n \leq N}$ denote the mesh points based on a non-uniform time step τ_n, such that $t_{n+1} = t_n + \tau_n$. We also denote $t_{n+\theta_n} = t_n + \theta_n\tau_n$, for any $\theta_n \in [0,1]$. The main steps of the proposed method are described as follows. First, on the interval $[t_n,t_{n+\theta}]$, we approximate the solution using the implicit backward-Euler method:

$$\frac{\boldsymbol{u}_{n+\theta_n} - \boldsymbol{u}_n}{\theta_n\tau_n} - \nabla\cdot(\nu(c_{n+\theta_n})D(\boldsymbol{u}_{n+\theta_n})) + (\boldsymbol{u}_{n+\theta_n}\cdot\nabla)\boldsymbol{u}_{n+\theta_n} + \nabla p_{n+\theta_n} = -g(1 + \gamma c_{n+\theta_n})\boldsymbol{i}_d + \boldsymbol{f}_{n+\theta_n},$$

$$\nabla\cdot\boldsymbol{u}_{n+\theta_n} = 0, \qquad (3.1)$$

$$\frac{c_{n+\theta_n} - c_n}{\theta_n\tau_n} - \Theta\Delta c_{n+\theta_n} + \boldsymbol{u}_{n+\theta_n}\cdot\nabla c_{n+\theta_n} + U\frac{\partial c_{n+\theta_n}}{\partial x_d} = 0.$$

8

Then, on $[t_{n+\theta_n}, t_{n+1}]$, we use the explicit forward Euler method:

$$\frac{\boldsymbol{u}_{n+1} - \boldsymbol{u}_{n+\theta_n}}{(1-\theta_n)\tau_n} - \nabla \cdot (\nu(c_{n+\theta_n})D(\boldsymbol{u}_{n+\theta_n})) + (\boldsymbol{u}_{n+\theta_n} \cdot \nabla)\boldsymbol{u}_{n+\theta_n} + \nabla p_{n+\theta_n} = -g(1+\gamma c_{n+\theta_n})\boldsymbol{i}_d + \boldsymbol{f}_{n+\theta_n}$$

$$\frac{c_{n+1} - c_{n+\theta_n}}{(1-\theta_n)\tau_n} - \Theta\Delta c_{n+\theta_n} + \boldsymbol{u}_{n+\theta_n} \cdot \nabla c_{n+\theta_n} + U\frac{\partial c_{n+\theta_n}}{\partial x_d} = 0,$$

(3.2)

or, equivalently, the linear extrapolations (see e.g., [16]):

$$\boldsymbol{u}_{n+1} = \frac{1}{\theta_n}\boldsymbol{u}_{n+\theta_n} - \left(\frac{1}{\theta_n}-1\right)\boldsymbol{u}_n, \quad c_{n+1} = \frac{1}{\theta_n}c_{n+\theta_n} - \left(\frac{1}{\theta_n}-1\right)c_n.$$

(3.3)

Finally, for time-adaptivity, we compute the local truncation error $\widehat{T}_{n+1}$ using Milne's device [13–16, 41], and given a tolerance $\varepsilon_{\Delta t}$, and positive parameters r_{min}, r_{max} and s, adapt the time step:

$$\tau_{new} = \tau_n \min\left\{r_{max}, \max\left\{r_{min}, s\left(\frac{\delta}{\|\widehat{T}_{n+1}\|}\right)^{\frac{1}{3}}\right\}\right\}.$$

(3.4)

Numbers r_{min} and r_{max} are added so that the ratio of τ_{new} and τ_n stays between these values. The coefficient s is a 'safety' parameter, routinely used to reduce the number of rejected time steps in the adaptive algorithm. If $\|\widehat{T}_{n+1}\| \leq \varepsilon_{\Delta t}$, we set $\tau_{n+1} = \tau_{new}$, chose θ_{n+1}, and evolve the time interval $t_{n+2} = t_{n+1} + \theta_{n+1}\tau_{n+1}$. Otherwise, we set $\tau_n = \tau_{new}$ and go back to the backward-Euler problem (3.1).

We note that the equations (3.1) and (3.2) can be equivalently written as the following one-legged θ-method ([18] originally used by Cauchy to prove the existence of a solution to ordinary differential equations):

$$\frac{\boldsymbol{u}_{n+1} - \boldsymbol{u}_n}{\tau_n} - \nabla \cdot (\nu(c_{n+\theta_n})D(\boldsymbol{u}_{n+\theta_n})) + (\boldsymbol{u}_{n+\theta_n} \cdot \nabla)\boldsymbol{u}_{n+\theta_n} + \nabla p_{n+\theta_n} = -g(1+\gamma c_{n+\theta_n})\boldsymbol{i}_d + \boldsymbol{f}_{n+\theta_n}$$

$$\nabla \cdot \boldsymbol{u}_{n+1} = 0,$$

$$\frac{c_{n+1} - c_n}{\tau_n} - \Theta\Delta c_{n+\theta_n} + \boldsymbol{u}_{n+\theta_n} \cdot \nabla c_{n+\theta_n} + U\frac{\partial c_{n+\theta_n}}{\partial x_d} = 0,$$

(3.5)

where, by (3.3),

$$\boldsymbol{u}_{n+\theta_n} = \theta_n\boldsymbol{u}_{n+1} + (1-\theta_n)\boldsymbol{u}_n, \quad c_{n+\theta_n} = \theta_n c_{n+1}(1-\theta_n)c_n.$$

When $\theta_n = \frac{1}{2}$, (3.5) gives the midpoint rule. Moreover, under the assumptions in Theorems 2.4 and 2.3, the steady state equations in (3.5) have weak and respectively strong solutions, see e.g. [10].

We also note that the one-leg method (3.5) for the finite-difference time-discretization and the finite elements methods for the spatial-discretization, gives rise to suitable weak solutions in the sense of Scheffer and Caffarelli-Kohn-Nirenberg, see e.g. [9, 16].

REMARK 1. *Assume that the initial datum and the forcing term satisfy the assumptions of Theorem 2.3, namely $(\boldsymbol{u}_0, c_0) \in D(\mathscr{A})$ and $(\boldsymbol{f}, 0) \in W^{1,1}([0,T];H)$. Then, for small time steps τ_n, depending inverse proportionally with ν_*^3 and $\nu_*^2\Theta$, the operator $\frac{1}{\tau_n}I + \mathscr{A} + \mathscr{B}$ is a contraction on $D(\mathscr{A})$. Therefore, for all $n \geq 0$, the Cauchy method (3.5) has a unique strong solution $(\boldsymbol{u}_n, c_n) \in D(\mathscr{A})$.*

The solution to the backward-Euler step (3.1) can be computed in a partitioned way, i.e., by solving the Navier-Stokes equations and transport equations separately. Let us denote by $\boldsymbol{u}_{n+\theta_n}^{(\kappa)}$ and $c_{n+\theta_n}^{(\kappa)}$ the κth-iterates of the velocity and concentration, respectively, where the initial iterates are taken to be

$$\boldsymbol{u}_{n+\theta_n}^{(0)} = (1+\theta_n)\boldsymbol{u}_n - \theta_n\boldsymbol{u}_{n-1}, \qquad c_{n+\theta_n}^{(0)} = (1+\theta_n)c_n - \theta_n c_{n-1}.$$

Then, the iterates are defined by

$$\frac{\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} - \boldsymbol{u}_n}{\theta_n\tau_n} - \nabla\cdot(\nu(c_{n+\theta_n}^{(\kappa)})D(\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)})) + (\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)}\cdot\nabla)\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} + \nabla p_{n+\theta_n}^{(\kappa+1)} = -g(1+\gamma c_{n+\theta_n}^{(\kappa)})\boldsymbol{i}_d + \boldsymbol{f}_{n+\theta_n},$$

$$\nabla \cdot \boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} = 0,$$

$$\frac{c_{n+\theta_n}^{(\kappa+1)} - c_n}{\theta_n\tau_n} - \Theta\Delta c_{n+\theta_n}^{(\kappa+1)} + \boldsymbol{u}_{n+\theta_n}^{(\kappa)}\cdot\nabla c_{n+\theta_n}^{(\kappa+1)} + U\frac{\partial c_{n+\theta_n}^{(\kappa+1)}}{\partial x_d} = 0,$$

(3.6)

9

for all $\kappa \geq 0$. In Section 3.2, we show that, as $\kappa \nearrow \infty$, the sequences of iterates converge

$$\boldsymbol{u}_{n+\theta_n}^{(\kappa)} \longrightarrow \boldsymbol{u}_{n+\theta_n}, \qquad c_{n+\theta_n}^{(\kappa)} \longrightarrow c_{n+\theta_n},$$

and in the limit, the system (3.6) solves (3.1).

3.1. Stability Bounds. First we note that, analogously to the continuous case (1.5), the total mass is conserved at the semi-discrete in time level, giving an $L^1(\Omega)$ stability bound for c_n. Indeed, integrating (3.5) over Ω, we have

$$0 = \int_\Omega \frac{c_{n+1}(\boldsymbol{x}) - c_n(\boldsymbol{x})}{\tau_n} d\boldsymbol{x} - \int_\Omega \left(\Theta \Delta c_{n+\theta_n}(\boldsymbol{x}) - U \frac{\partial c_{n+\theta_n}(\boldsymbol{x})}{\partial x_d}\right) d\boldsymbol{x} + \int_\Omega \boldsymbol{u}_{n+\theta_n}(\boldsymbol{x}) \cdot \nabla c_{n+\theta_n}(\boldsymbol{x}) d\boldsymbol{x},$$

which yields that the total mass of micro-organisms is constant in time:

$$\int_\Omega c_{n+1}(\boldsymbol{x}) d\boldsymbol{x} = \int_\Omega c_n(\boldsymbol{x}) d\boldsymbol{x}, \qquad \forall n \geq 0. \tag{3.7}$$

Secondly, similarly to (1.6), we have

$$\frac{1}{\tau_n}\left(\mathrm{KE}_{n+1} + \mathrm{PE}_{n+1} - \mathrm{KE}_n - \mathrm{PE}_n\right) + \frac{2\theta_n - 1}{2}\|u_{n+1}(\boldsymbol{x}) - u_n(\boldsymbol{x})\|^2 + \int_\Omega \nu(c_{n+\theta_n}(\boldsymbol{x}))|\nabla u_{n+\theta_n}(\boldsymbol{x})|^2 d\boldsymbol{x}$$

$$= \int_\Omega \boldsymbol{f}_{n+\theta_n}(\boldsymbol{x})\boldsymbol{u}_{n+\theta_n}(\boldsymbol{x}) d\boldsymbol{x} + g \int_\Omega \left(\gamma(U c_{n+\theta_n}(\boldsymbol{x}) - \Theta \nabla c_{n+\theta_n}(\boldsymbol{x}) \cdot \boldsymbol{i}_d) - \boldsymbol{u}_{n+\theta_n}(\boldsymbol{x}) \cdot \boldsymbol{i}_d\right) d\boldsymbol{x}$$

$$= \int_\Omega \boldsymbol{f}_{n+\theta_n}(\boldsymbol{x})\boldsymbol{u}_{n+\theta_n}(\boldsymbol{x}) d\boldsymbol{x} + g\gamma U \int_\Omega c_0(\boldsymbol{x}) d\boldsymbol{x} - g \int_\Omega \left(\gamma\Theta\nabla c_{n+\theta_n} + \boldsymbol{u}_{n+\theta_n}\right) \cdot \boldsymbol{i}_d d\boldsymbol{x}.$$

Hence, for a insulated system ($f = 0$) and the midpoint method ($\theta_n = 1/2$), the total energy balance at the semi-discrete level writes:

$$\frac{1}{\tau_n}\left(\mathrm{KE}_{n+1} + \mathrm{PE}_{n+1} - \mathrm{KE}_n - \mathrm{PE}_n\right) + \int_\Omega \nu(c_{n+1/2}(\boldsymbol{x}))|\nabla u_{n+1/2}(\boldsymbol{x})|^2 d\boldsymbol{x}$$

$$= g\gamma U \int_\Omega c_0(\boldsymbol{x}) d\boldsymbol{x} - g \int_\Omega \left(\gamma\Theta\nabla c_{n+1/2} + \boldsymbol{u}_{n+1/2}\right) \cdot \boldsymbol{i}_d d\boldsymbol{x}.$$

Next, we prove that the semi-discrete system (3.5) is well posed, namely its solution satisfies energy estimates similar to the continuous case, which are often referred to as a stability result. Let us denoted by K the following constant depending only on the parameters of the problem:

$$K = \frac{C_\Omega^2}{\nu_*}g^2\left(|\Omega| + 2\gamma^2\alpha^2\right)T + \frac{2}{\nu_*}g^2\gamma^2 C_\Omega^4 \frac{1}{\Theta}\exp(TU^2 C_\Omega/\Theta)\|c_0\|^2 + \frac{1}{2}.$$

THEOREM 3.1. *Under assumptions* (2.1) *we have that the solution to the semi-discrete problem* (3.5) *is bounded in* $L^\infty(0,T,L^2(\Omega)) \cap L^2(0,T,H^1(\Omega))$, *and the following estimate holds*

$$\frac{1}{2}\left(\|\boldsymbol{u}_N\|^2 + \|c_N\|^2\right) + \sum_{n=0}^{N-1}\frac{2\theta_n - 1}{2}\left(\|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n\|^2 + \|c_{n+1} - c_n\|^2\right) \tag{3.8}$$

$$+ \sum_{n=0}^{N-1}\tau_n \int_\Omega \left(\nu(c_{n+\theta_n}) - \nu_*\right)|D(\boldsymbol{u}_{n+\theta_n})|^2 + \frac{\Theta}{2}\sum_{n=0}^{N-1}\tau_n\|\nabla c_{n+\theta_n}\|^2$$

$$\leq \frac{1}{2}\left(\|\boldsymbol{u}_0\|^2 + \|c_0\|^2\right) + K + \frac{C_\Omega^2}{2\nu_*}\sum_{n=0}^{N-1}\tau_n\|\boldsymbol{f}_{n+\theta_n}\|^2.$$

Proof. Testing the momentum equation in (3.5) with $\tau_n \boldsymbol{u}_{n+\theta_n}$, the transport equations with $\tau_n c_{n+\theta_n}$, using the incompressibility condition and the boundary conditions (1.4), and by summing from $n = 0$ to $N-1$, we obtain:

$$\frac{1}{2}\|\boldsymbol{u}_N\|^2 + \sum_{n=0}^{N-1}\frac{2\theta_n - 1}{2}\|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n\|^2 + \sum_{n=0}^{N-1}\tau_n\int_\Omega \nu(c_{n+\theta_n})|D(\boldsymbol{u}_{n+\theta_n})|^2 \tag{3.9}$$

$$= \frac{1}{2}\|\boldsymbol{u}_0\|^2 - \sum_{n=0}^{N-1}\tau_n\int_\Omega g(1 + \gamma c_{n+\theta_n})i_2\,\boldsymbol{u}_{n+\theta_n} + \sum_{n=0}^{N-1}\tau_n\int_\Omega \boldsymbol{f}_{n+\theta_n}\boldsymbol{u}_{n+\theta_n},$$

$$\frac{1}{2}\|c_N\|^2 + \sum_{n=0}^{N-1}\frac{2\theta_n - 1}{2}\|c_{n+1} - c_n\|^2 + \Theta\sum_{n=0}^{N-1}\tau_n\|\nabla c_{n+\theta_n}\|^2 = \frac{1}{2}\|c_0\|^2 + \sum_{n=0}^{N-1}\tau_n\int_\Omega U\frac{\partial c_{n+\theta_n}}{\partial x_d}c_{n+\theta_n}d\boldsymbol{x}. \tag{3.10}$$

First, using the boundary conditions (1.4), the trace inequality [30], and the discrete Grönwall's inequality [34], we get:

$$\frac{1}{2}\|c_N\|^2 + \sum_{n=0}^{N-1}\frac{2\theta_n - 1}{2}\|c_{n+1} - c_n\|^2 + \frac{\Theta}{2}\sum_{n=0}^{N-1}\tau_n\|\nabla c_{n+\theta_n}\|^2 \leq \frac{1}{2}\exp(2TU^2 C_\Omega/\Theta)\|c_0\|^2. \tag{3.11}$$

Next, using the Cauchy-Schwarz, the Poincaré (2.2) and Poincaré-Wirtinger (2.3) inequalities, and the (2.1) lower bound on the fluid viscosity, from (3.9) we get the following energy estimate for the flow velocity:

$$\frac{1}{2}\|\boldsymbol{u}_N\|^2 + \sum_{n=0}^{N-1}\frac{2\theta_n - 1}{2}\|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n\|^2 + \sum_{n=0}^{N-1}\tau_n\int_\Omega \big(\nu(c_{n+\theta_n}) - \nu_*\big)|D(\boldsymbol{u}_{n+\theta_n})|^2$$

$$\leq \frac{1}{2}\|\boldsymbol{u}_0\|^2 + \frac{C_\Omega^2}{\nu_*}g^2\big(|\Omega| + 2\gamma^2\alpha^2\big)T + \frac{2}{\nu_*}g^2\gamma^2 C_\Omega^4\sum_{n=0}^{N-1}\tau_n\|\nabla c_{n+\theta_n}\|^2 + \frac{C_\Omega^2}{2\nu_*}\sum_{n=0}^{N-1}\tau_n\|\boldsymbol{f}_{n+\theta_n}\|^2.$$

Using the mass balance relation (3.11) we obtain (3.8), i.e., the time-discretization is nonlinearly energy stable for all $\theta_n \in [\frac{1}{2}, 1]$. From the energy equality (3.9)-(3.11) we note that the method has nonzero numerical dissipation unless $\theta_n = 1/2$, which corresponds to the midpoint rule. □

3.2. Convergence of the sub-iterations. PROPOSITION 3.2. *For small enough time steps*

$$\tau_n \leq \frac{\Theta}{U^2}, \tag{3.12}$$

under assumption (2.1)*, the iterations* $\{\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)}, c_{n+\theta_n}^{(\kappa+1)}\}$ *satisfying the partitioned method* (3.6) *converge weakly in* $H^1(\Omega)$*, and strongly in* $L^2(\Omega)$ *to the solution* $\{\boldsymbol{u}_{n+\theta_n}, c_{n+\theta_n}\}$ *of the coupled system* (3.1)*.*

Proof. We test (3.6) with $\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)}$ and $c_{n+\theta_n}^{(\kappa+1)}$, respectively, to obtain

$$\frac{1}{2\theta_n\tau_n}\Big(\|\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)}\|^2 - \|\boldsymbol{u}_n\|^2 + \|\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} - \boldsymbol{u}_n\|^2\Big) + \int_\Omega \nu(c_{n+\theta_n}^{(\kappa)})|D(\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)})|^2$$

$$= -g\int_\Omega(1 + \gamma c_{n+\theta_n}^{(\kappa)})i_2\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} + \int_\Omega \boldsymbol{f}_{n+\theta_n}\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)},$$

$$\frac{1}{2\theta_n\tau_n}\Big(\|c_{n+\theta_n}^{(\kappa+1)}\|^2 - \|c_n\|^2 + \|c_{n+\theta_n}^{(\kappa+1)} - c_n\|^2\Big) + \Theta\|\nabla c_{n+\theta_n}^{(\kappa+1)}\|^2 = \int_\Omega U\frac{\partial c_{n+\theta_n}^{(\kappa+1)}}{\partial x_d}c_{n+\theta_n}^{(\kappa+1)}d\boldsymbol{x}.$$

From the second relation, with an argument similar to the one used for (3.11), we have that hence, for small enough time steps, i.e., satisfying (3.12),

$$\Big(\frac{1}{\theta_n\tau_n} - \frac{1}{\Theta}U^2\Big)\|c_{n+\theta_n}^{(\kappa+1)}\|^2 + \frac{1}{2\theta_n\tau_n}\|c_{n+\theta_n}^{(\kappa+1)} - c_n\|^2 + \frac{\Theta}{2}\|\nabla c_{n+\theta_n}^{(\kappa+1)}\|^2 \leq \frac{1}{\theta_n\tau_n}\|c_n\|^2,$$

the sequence $\{c_{n+\theta_n}^{(\kappa)}\}_\kappa$ is bounded in $H^1(\Omega)$, and therefore

$$c_{n+\theta_n}^{(\kappa)} \rightharpoonup c_{n+\theta_n} \qquad \text{weakly in } H^1(\Omega), \text{ strongly in } L^2(\Omega).$$

11

Using the uniform (lower) bound of the viscosity $\nu(\cdot)$, from the NSE energy equality we obtain similarly that

$$\boldsymbol{u}_{n+\theta_n}^{(\kappa)} \rightharpoonup \boldsymbol{u}_{n+\theta_n} \qquad \text{weakly in } H^1(\Omega), \text{ strongly in } L^2(\Omega).$$

By the Lipschitz continuity of the viscosity function (2.1) we have that also

$$\nu(c_{n+\theta_n}^{(\kappa)}) \to \nu(c_{n+\theta_n}) \qquad \text{strongly in } L^2(\Omega),$$

and therefore, letting $\kappa \nearrow \infty$ in (3.6), we see that the limits $\boldsymbol{u}_{n+\theta_n}, c_{n+\theta_n}$ satisfy (3.1). $\square$

3.2.1. Linear convergence. Denote

$$\delta_u^\kappa = \boldsymbol{u}_{n+\theta_n} - \boldsymbol{u}_{n+\theta_n}^{(\kappa)}, \quad \delta_p^\kappa = p_{n+\theta_n} - p_{n+\theta_n}^{(\kappa)}, \quad \delta_u^\kappa = c_{n+\theta_n} - c_{n+\theta_n}^{(\kappa)}.$$

Assume now that the initial data is smooth and the time step is small enough, such that by Remark 1, we have that (3.5) admits a unique strong solution $(\boldsymbol{u}_n, c_n) \in D(\mathscr{A})$. Then we have the following stronger result, regarding the order of convergence of the iterative process (3.6).

PROPOSITION 3.3. *For small time steps, the sequence* $\{\boldsymbol{u}_{n+\theta_n}^{(\kappa)}, c_{n+\theta_n}^{(\kappa)}\}$ *converges linearly to the solution of* (3.1) $\{\boldsymbol{u}_{n+\theta_n}, c_{n+\theta_n}\}$.

Proof. Subtracting (3.6) from (3.1) we obtain

$$\frac{\delta_u^{\kappa+1}}{\theta_n \tau_n} - \nabla \cdot \left((\nu(c_{n+\theta_n}) - \nu(c_{n+\theta_n}^{(\kappa)})) D(\boldsymbol{u}_{n+\theta_n}) \right) - \nabla \cdot \left(\nu(c_{n+\theta_n}^{(\kappa)}) D(\delta_u^{\kappa+1}) \right)$$
$$+ \left(\delta_u^{\kappa+1} \cdot \nabla \right) \boldsymbol{u}_{n+\theta_n} + (\boldsymbol{u}_{n+\theta_n}^{(\kappa+1)} \cdot \nabla) \delta_u^{\kappa+1} + \nabla \delta_p^{\kappa+1} = -g\gamma \delta_c^\kappa i_2,$$
$$\nabla \cdot \delta_u^{\kappa+1} = 0,$$
$$\frac{\delta_c^{\kappa+1}}{\theta_n \tau_n} - \Theta \Delta \delta_c^{\kappa+1} + \delta_u^\kappa \cdot \nabla c_{n+\theta_n} + \boldsymbol{u}_{n+\theta_n}^{(\kappa)} \cdot \nabla \delta_c^{\kappa+1} + U \frac{\partial \delta_c^{(\kappa+1)}}{\partial x_2} = 0.$$

Proceeding similarly as in the proof of Theorem 3.1, using the Cauchy-Schwarz inequality and the bounds (2.1) on the flow viscosity, we have

$$\left(\frac{1}{\theta_n \tau_n} - \frac{1}{2} g\gamma \right) \|\delta_u^{\kappa+1}\|^2 + \nu_* \|D(\delta_u^{\kappa+1})\|^2$$
$$\leq \frac{1}{2} g\gamma \|\delta_c^\kappa\|^2 + L \int_\Omega |\delta_c^\kappa| \, |D(\boldsymbol{u}_{n+\theta_n})| \, |D(\delta_u^{\kappa+1})| + \int_\Omega \left| \left(\delta_u^{\kappa+1} \cdot \nabla \right) \boldsymbol{u}_{n+\theta_n} \delta_u^{\kappa+1} \right|,$$
$$\frac{1}{\theta_n \tau_n} \|\delta_c^{\kappa+1}\|^2 + \Theta \|\nabla \delta_c^{\kappa+1}\|^2 \leq \int_\Omega \left| \delta_u^\kappa \cdot \nabla c_{n+\theta_n} \delta_c^{\kappa+1} \right|.$$

Using the Ladyzhenskaya (Gagliardo-Nirenberg) inequalities [40], in the $d = 2$ case, for example, we obtain

$$\left(\frac{1}{\theta_n \tau_n} - \frac{1}{2} g\gamma - \frac{1}{\nu_*} \right) \|\delta_u^{\kappa+1}\|^2 + \left(\frac{1}{\theta_n \tau_n} - \frac{1}{2\Theta} \right) \|\delta_c^{\kappa+1}\|^2 + \frac{\nu_*}{2} \|D(\delta_u^{\kappa+1})\|^2 + \frac{\Theta}{2} \|\nabla \delta_c^{\kappa+1}\|^2$$
$$\leq \frac{C^4}{16\nu_*} \left(\|\nabla \boldsymbol{u}_{n+\theta_n}\|^2 + \|\nabla c_{n+\theta_n}\|^2 \right)^2 \|\delta_u^\kappa\|^2 + \left(\frac{1}{2} g\gamma + \frac{C^4 L^4}{|\nu_*|^2 \Theta} \|\nabla \boldsymbol{u}_{n+\theta_n}\|^2 \|\Delta \boldsymbol{u}_{n+\theta_n}\|^2 \right) \|\delta_c^\kappa\|^2$$
$$+ \frac{\nu_*}{4} \|\nabla \delta_u^\kappa\|^2 + \frac{\Theta}{4} \|\nabla \delta_c^\kappa\|^2.$$

So for small time steps τ_n such that

$$\frac{C^4}{16\nu_*} \left(\|\nabla \boldsymbol{u}_{n+\theta_n}\|^2 + \|\nabla c_{n+\theta_n}\|^2 \right)^2 \leq \left(\frac{1}{\theta_n \tau_n} - \frac{1}{2} g\gamma - \frac{1}{\nu_*} \right)$$
$$\frac{1}{2} g\gamma + \frac{C^4 L^4}{|\nu_*|^2 \Theta} \|\nabla \boldsymbol{u}_{n+\theta_n}\|^2 \|\Delta \boldsymbol{u}_{n+\theta_n}\|^2 \leq \left(\frac{1}{\theta_n \tau_n} - \frac{1}{2\Theta} \right),$$

i.e.,

$$\theta_n \tau_n \leq \min\left\{ \left(\frac{1}{2}g\gamma + \frac{1}{\nu_*} + \frac{C^4}{16\nu_*}\left(\|\nabla \boldsymbol{u}_{n+\theta_n}\|^2 + \|\nabla c_{n+\theta_n}\|^2\right)^2\right)^{-1}, \right.$$
$$\left. \left(\frac{1}{2}g\gamma + \frac{1}{2\Theta} + \frac{C^4 L^4}{|\nu_*|^2\Theta}\|\nabla \boldsymbol{u}_{n+\theta_n}\|^2\|\Delta \boldsymbol{u}_{n+\theta_n}\|^2\right)^{-1}\right\}$$

we have linear convergence. □

4. Numerical examples. This section is focused on the numerical simulations of bioconvection. The problem is discretized in space using the finite element method with uniform, conforming meshes, where mesh size is denoted as Δx. The numerical method is implemented in the finite element solver, FreeFem++ [33]. We first investigate the rates of convergence of the proposed partitioned numerical method using both the constant and variable viscosity in Example 4.1. In Example 4.2, we simulate the bioconvection of motile flagellates, following the work in [32]. As in [32], we use a constant viscosity, however, we extend the simulation time and study the effects of small perturbations of the initial concentration. Similar problem is simulated in Example 4.3, but now using concentration-dependent viscosity.

4.1. Example 1. In this example, we present the rates of convergence for our method. We consider two test cases for viscosity: (*i*) viscosity is constant and independent of concentration, and (*ii*) viscosity is defined as a function of concentration. In both examples, a fixed time step is used. Furthermore, we use $\mathbb{P}_2$ elements for the velocity and concentration, and $\mathbb{P}_1$ elements for pressure. The tolerance for the sub-iterative backward-Euler problem is set to $\varepsilon = 10^{-4}$ and the final simulation time is $T = 1$ s.

Rates of Convergence - Constant Viscosity. We consider a benchmark problem, similar to [17], where the domain and parameters are defined as follows:

$$\Omega = [-1,1] \times [-1,1], \quad g = 10, \quad \gamma = 0.1, \quad U = 0.1, \quad \Theta = 0.1, \quad \nu_0 = 0.1.$$

We define viscosity as $\nu = \nu_0$. The exact solution is given by:

$$\boldsymbol{u}_{ref}(\boldsymbol{x},t) = t^2 \sin(\pi x)\sin(\pi y), \quad p = c_{ref}(\boldsymbol{x},t) = t^2 x^2 y^2. \tag{4.1}$$

We compute the relative errors between the approximated and exact solutions, which are defined as:

$$e_{\boldsymbol{u}}^{ref} = \frac{\|\boldsymbol{u} - \boldsymbol{u}_{ref}\|_{L^2(\Omega)}}{\|\boldsymbol{u}_{ref}\|_{L^2(\Omega)}}, \quad e_c^{ref} = \frac{\|c - c_{ref}\|_{L^2(\Omega)}}{\|c_{ref}\|_{L^2(\Omega)}}.$$

To compute the rates of convergence, the time step and the mesh size are refined four times, and their values are given as follows:

$$\{\tau, \Delta x\} = \left\{\frac{0.05}{2^i}, \frac{0.08}{2^i}\right\}_{i=0}^4.$$

The rates of convergence in the constant viscosity case are presented in Table 4.1. We expect first order when $\theta_n = 1$ and second order when $\theta_n = 1/2$. The computed rates are in good agreement with the expected values.

Rates of Convergence - Viscosity as Function of Concentration. We extend and modify the previous example to a more complex problem where the viscosity depends on the concentration, $\nu = \nu(c)$. Changing viscosity to be a function of concentration is motivated by the practical application of the problem, where the collective flow of micro-organisms results in areas of high and low concentration with respect to the organisms. As organisms pool together and become density populated in an area, this increase in concentration increases the viscosity of the flow. For this problem, we assume that the exact solution is the same as in the previous example (see 4.1), and the following parameters are used:

$$\Omega = [0,0.2] \times [0,0.2], \quad g = 1.0, \quad \gamma = 1.0, \quad U = 1.0, \quad \Theta = 1.0, \quad \nu_0 = 1.0.$$

We define the viscosity as a function of concentration,

$$\nu(c) = \begin{cases} \nu_0 & c \leq 10^{-8} \\ \nu_0(1 + 0.01\, c(\boldsymbol{x},t)) & c > 10^{-8}. \end{cases}$$

$\theta_n = 1$	τ	Δx	$e_{\boldsymbol{u}}^{ref}$	$\text{Rate}_{\boldsymbol{u}}$	e_c^{ref}	Rate_c
	τ	0.08	$4.11679 \cdot 10^{-2}$	—	$1.09574 \cdot 10^{-2}$	—
	$\tau/2$	0.04	$2.15777 \cdot 10^{-2}$	0.931979	$5.63213 \cdot 10^{-3}$	0.960153
	$\tau/4$	0.02	$1.10509 \cdot 10^{-2}$	1.89736	$2.85981 \cdot 10^{-3}$	1.93791
	$\tau/8$	0.01	$5.59194 \cdot 10^{-3}$	0.982743	$1.43772 \cdot 10^{-3}$	0.992137
$\theta_n = 1/2$	τ	Δx	$e_{\boldsymbol{u}}^{ref}$	$\text{Rate}_{\boldsymbol{u}}$	e_c^{ref}	Rate_c
	τ	0.08	$3.88742 \cdot 10^{-3}$	—	$5.49568 \cdot 10^{-3}$	—
	$\tau/2$	0.04	$6.61822 \cdot 10^{-5}$	2.5543	$1.46636 \cdot 10^{-4}$	1.90606
	$\tau/4$	0.02	$1.49089 \cdot 10^{-5}$	2.15027	$3.79530 \cdot 10^{-5}$	1.94995
	$\tau/8$	0.01	$3.92655 \cdot 10^{-6}$	1.92484	$9.64102 \cdot 10^{-6}$	1.97696

TABLE 4.1

Rates of convergence for $\theta = 1, 1/2$ for $v = v_0$.

The time steps and mesh sizes are defined as:

$$\{\tau, \Delta x\} = \left\{ \frac{0.05}{2^i}, \frac{0.04}{2^i} \right\}_{i=0}^{4}.$$

The rates of convergence are presented in Table 4.2.

$\theta_n = 1$	τ	Δx	$e_{\boldsymbol{u}}^{ref}$	$\text{Rate}_{\boldsymbol{u}}$	e_c^{ref}	Rate_c
	τ	0.04	$5.80196 \cdot 10^{-5}$	—	$1.29159 \cdot 10^{-3}$	—
	$\tau/2$	0.02	$1.42646 \cdot 10^{-5}$	2.0241	$1.61036 \cdot 10^{-4}$	3.00369
	$\tau/4$	0.01	$6.63115 \cdot 10^{-6}$	1.10511	$2.21916 \cdot 10^{-5}$	2.8593
	$\tau/8$	0.005	$3.22078 \cdot 10^{-6}$	1.04185	$5.65963 \cdot 10^{-6}$	1.97123
$\theta_n = 1/2$	τ	Δx	$e_{\boldsymbol{u}}^{ref}$	$\text{Rate}_{\boldsymbol{u}}$	e_c^{ref}	Rate_c
	τ	0.04	$1.6856 \cdot 10^{-4}$	—	$1.36356 \cdot 10^{-3}$	—
	$\tau/2$	0.02	$4.8812 \cdot 10^{-5}$	1.78795	$2.01521 \cdot 10^{-4}$	2.75838
	$\tau/4$	0.01	$1.30425 \cdot 10^{-5}$	1.90402	$3.96018 \cdot 10^{-5}$	2.34729
	$\tau/8$	0.005	$3.49384 \cdot 10^{-6}$	1.90033	$9.2243 \cdot 10^{-6}$	2.10205

TABLE 4.2

Rates of convergence for $\theta = 1, 1/2$ for $v = v(c)$.

We observe that the backward-Euler method ($\theta_n = 1$) has a higher rate of convergence for the first refinement step, but with continued refinements of the spatial and temporal meshes, the rates of convergence approach order one for the velocity and order two for concentration. The relative errors for $\theta_n = 1$ saturate quickly with refinements, but we note that the relative error is small at all rates. The rates of convergence for the midpoint method ($\theta_n = 1/2$) are approximately second order. More precisely, a slightly smaller value is obtained for the velocity, and a slightly larger value is obtained for the concentration.

4.2. Example 2.

In this example, we consider the evolution of bioconvection patterns in a culture of motile flagellates, as described in [32, run 16]. The computation domain is defined as $\Omega = (0, L) \times (0, H)$, with $L = 16$ and $H = 2$. At the bottom and side boundaries we prescribe $\boldsymbol{u} = 0$, while at the top boundary we have zero normal flux and zero shear traction:

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0, \quad \boldsymbol{\tau} \cdot \boldsymbol{\sigma} \boldsymbol{n} = 0.$$

For the concentration, we prescribe a no-flux boundary condition on the entire boundary as defined in (1.4). The parameter values used in this example are obtained from [32] and summarized in Table 4.3.

14

Parameters	ν (cm^2/s)	g (cm/s^2)	γ (cm^2/cells)	Θ (cm^2/s)	U (cm/s)
Values	0.01	980.665	$5 \cdot 10^{-10}$	0.01	0.1

TABLE 4.3
Parameters used in Example 2.

It is worth noting that the model we are using is similar to the one in [32], with the only difference being the lack of $-g\gamma c$ term in the momentum equation in [32]. However, since $\gamma = 5 \cdot 10^{-10}$, this term should not affect much the model/computations. The parameter values used in this example yield the sublayer Rayleigh number $R = 9.80$, Schmidt number $\sigma = \nu_0/\Theta = 1$, sublayer thickness $h = \nu/U = 0.1$ cm, vertical swimming time $T_s = H/U = 20$ s, and vertical diffusion time $T_d = H^2/\Theta = 400$ s.

The simulations are performed until $T = 1200$ s is reached. We let $\theta_n = 0.5$, for all n, and $\varepsilon = 10^{-4}$. $\mathbb{P}_2 - \mathbb{P}_1$ elements are used for the fluid velocity and pressure, respectively, and $\mathbb{P}_2$ elements are used for concentration on a structured mesh containing 14,400 elements.

Initially, we set $\boldsymbol{u}_0 = 0$, and consider a set of different initial conditions for the concentration, defined as $c_0 = 10^5 + \delta$ cells/cm^2, where $\delta \in \{\pm 4 \cdot 10^{-10}, \pm 3 \cdot 10^{-10}, \pm 2 \cdot 10^{-10}, \pm 10^{-10}, 0\}$. In these simulations, we use a constant time step, $\tau = 10^{-2}$. For cases $\delta = 0, \delta = \pm 10^{-10}$, we also perform simulations using adaptive time stepping with parameters $\varepsilon_{\Delta t} = 10^{-4}, r_{min} = 0.5, r_{max} = 1.2$ and $s = 0.95$. The maximal allowed time step is set to be 0.1. The local truncation error, $\hat{T}_{n+1}$, is computed using an explicit, second order Adams-Bashforth two-step method as described in [12].

In all simulations, the concentration at the beginning forms layers, with largest values at the top of the domain (see Figure 4.1, panel (a)). Around $t = 40$ s, the patterns start to emerge. However, the system is chaotic, and different parameters give different results. The number of falling fingers ranges between four and seven until the stable solution is reached. The concentration and velocity at time $t = 100$ s are shown in Figure 4.1, panel (b). Around $t = 300$ s, the solution starts to stabilize, and a stable, steady solution is eventually obtained. Interestingly, when the patterns begin to stabilize after the chaotic regime, the same stable solution is reached in all considered cases. The concentration and velocity at the final time are shown in panel (c) of Figure 4.1.

Similarly as in [32], we calculate the total kinetic (K.E.) and potential energy (P.E.), defined as

$$\text{K.E.} = \frac{1}{2} \int_\Omega \boldsymbol{u}^2, \quad \text{P.E.} = \gamma g \int_\Omega cy.$$

Figure 4.2 shows the spaghetti plot [35] of the total kinetic energy versus the total potential energy obtained using different initial conditions, and both fixed and variable time-stepping. In all the cases considered here, the chaotic regime begins around $t = 68$ s, and lasts until roughly $t = 150$ s. At that time, all the solutions start behaving in a similar fashion, and seem to converge to the same stable solution.

We denote some points of interest relative to the energy in Figure 4.3. We denote by O the initial energy, and by A the time of onset of Rayleigh-Taylor instability, which occurs at 39 s. It is computed as the time when the kinetic energy becomes greater than 10^{-3}. Point B occurs at 43.3 and 48 s. During that time two fingers in the center of the domain join into one (as shown in the top right panel). Point C occurs at 74 and 81.25 s, which is when a similar even happens at the ends of the domain, as shown in the bottom right panel. Finally, we also denote by D the point when the stable state is reached, which occurs at 190 s.

The final number of falling fingers being $N = 3$ (the fingers attached to the side wall are counted as 0.5), as seen in Figure 4.1, implies that the aspect ratio of the convection cells is $\mu = L/(NH) \approx 2.6667$. In [32, pp. 768], the authors remark that among the quotients $L/(NH)$, 2.6667 is the closest to 2.34, which is the aspect ratio of the *optimum* convection cell, predicted by the linear theory of the thermally driven Bénard-Rayleigh convection. Nonetheless, they also argue that this may be just a coincidence for the wavenumber selection, as the 'final' state of the numerical results is in a nonlinear range ($T = 400$ s in their case, or $T = 1200$ s in our case).

In the Bénard-Rayleigh problem, the buoyancy is exerted by the difference between the boundary conditions, while for bioconvection problem, the buoyancy flux $b_f = \alpha g n_{pi} U$ is specified by the upward swimming of the microorganisms. The authors argue in [32, pp. 772] that the flux-type Rayleigh number $R_{af} = b_f H^4 \Theta^{-1} \nu^{-1} \approx 7.84 \cdot 10^5$ is the relevant nondimensional parameter which characterizes the bioconvection at the final stage. Therefore, the authors clarify the transient selection mode and claim that their numerical results qualitatively have good 'coincidence' with the patterns of *Heterosigma akashiwo*. They also explain the transient selection mode proposing and testing the
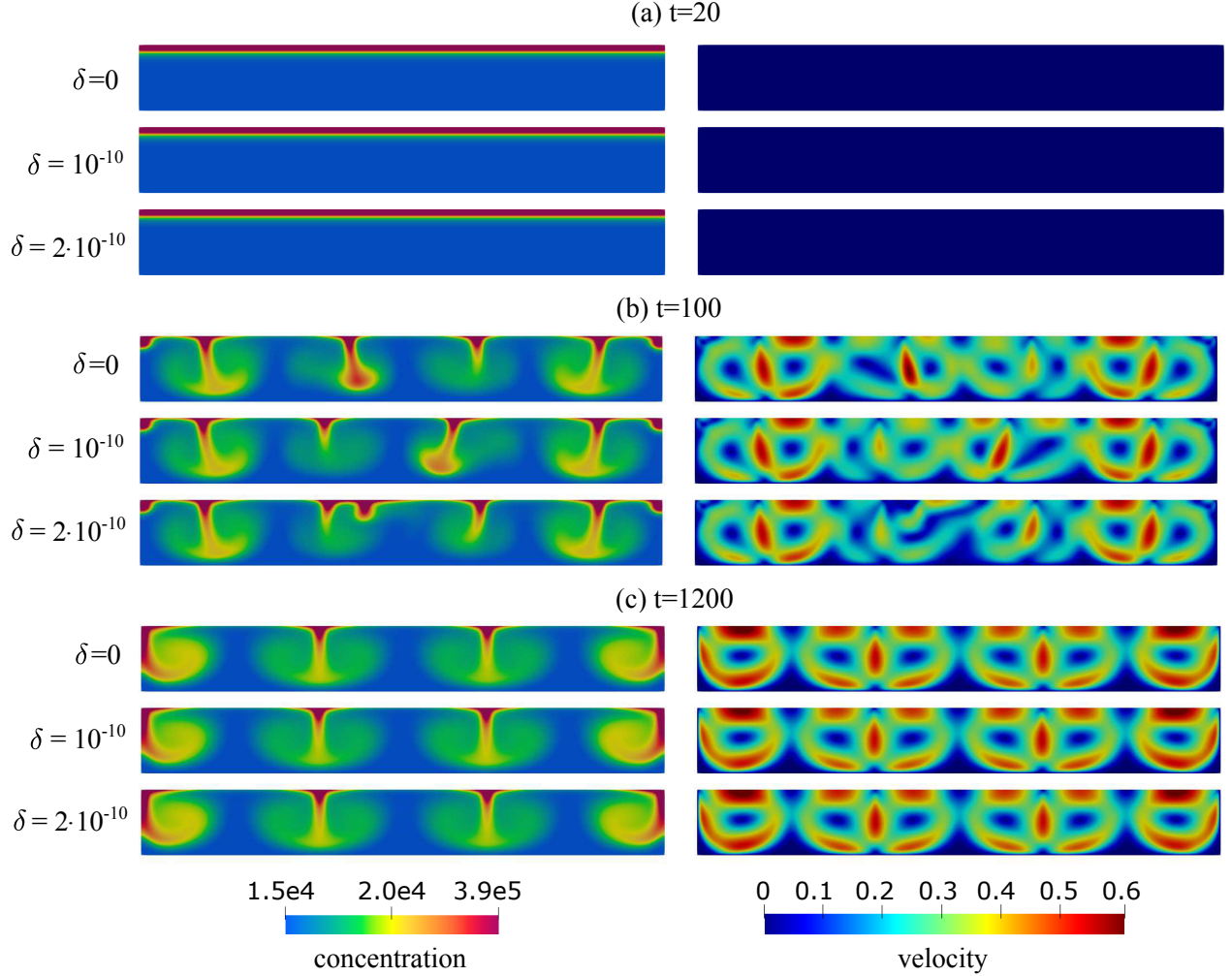
FIG. 4.1. *Concentration (left) and velocity magnitude (right) at times (a) $t = 20$ s, (b) $t = 100$ s, and (c) $t = 1200$ s obtained using $\delta = 0$, $\delta = 10^{-10}$ and $\delta = 2 \cdot 10^{-10}$.*

following hypothesis: "the convection system evolves such that the total potential energy of the system is minimized", also arguing that the system would be more stable against certain perturbations in the state of lower potential energy than in higher potential energy. This seems to not agree with our plots in Figure 4.2.

For the cases obtained using the fixed time-stepping, we plot the relative L^2-variations for each simulation with respect to the instance when $\delta = 0$. Figure 4.4 shows the variations for concentration (top panel) and velocity (bottom panel). Since the initial conditions are very close together, the variations are initially nearly zero. As the system becomes chaotic at around $t = 68$ s, the variations start to grow, leading to the maximum values of 83.5% and 100% for concentration and velocity, respectively, obtained at around $t = 150$ s. After that, the variations decrease and the concentration variation stays at 7%, while the velocity variation stays at 4.6%. Even though the concentration and velocity profiles at the end of the simulations shown in Figure 4.1, panel (c), are nearly identical, the small L^2-variations we observe are concentrated at the top of the domain, as shown in Figure 4.5 for the case when $\delta = -10^{-10}$.

Finally, we plot the time steps used in the adaptive time-stepping algorithm in Figure 4.6 for $\delta = 0$, $\delta = -10^{-10}$ and $\delta = 10^{-10}$. We observe that in all three cases, the time steps exhibit similar behavior. Initially, the time steps oscillate between small and large values, and reach the maximum allowed time step for a short while. As the patterns start to emerge, the time steps decrease, and then finally increase again to the maximum size as the simulations reach
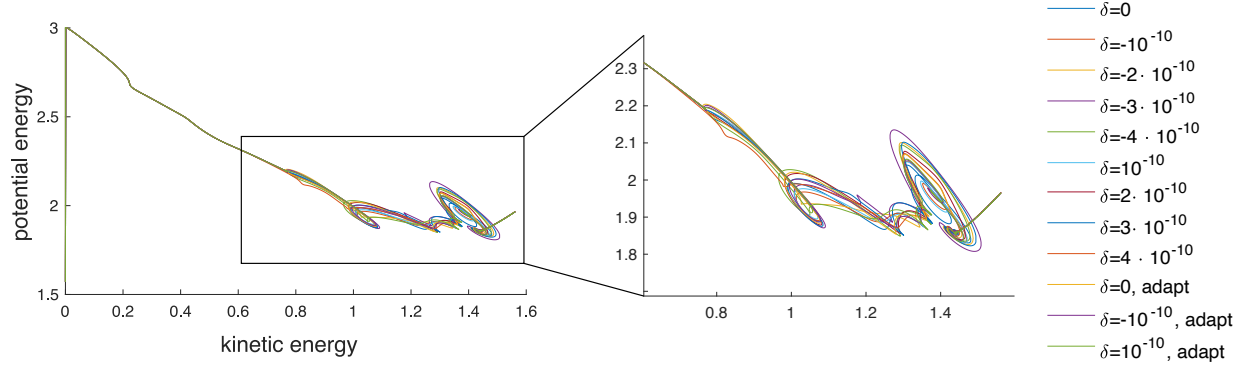
FIG. 4.2. *The total kinetic energy (horizontal axis) versus the total potential energy (vertical axis) obtained using different variation of initial conditions, and fixed and adaptive time stepping.*
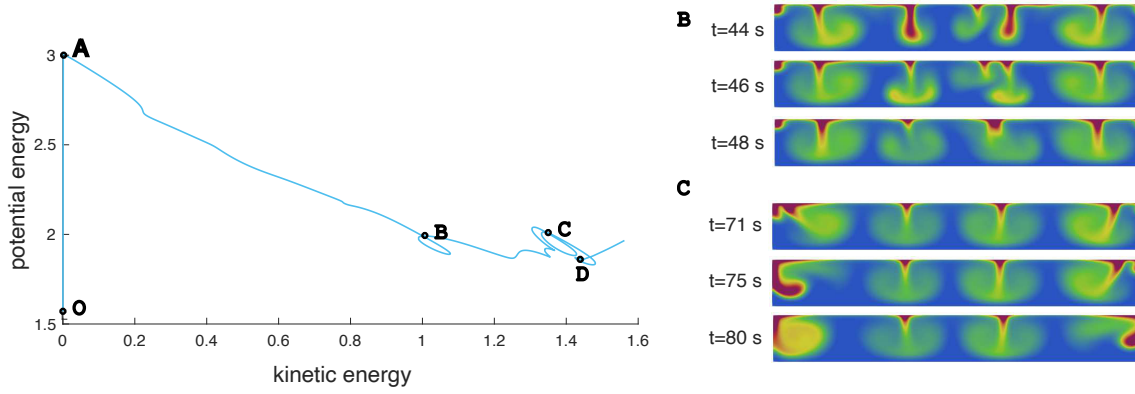

FIG. 4.3. *Times of points of interest superimposed with the plot showing the total kinetic energy (horizontal axis) versus the total potential energy (vertical axis) obtained using $\delta = 0$ and fixed time stepping. O denotes the initial energy, and A denotes the onset of Rayleigh-Taylor instability (when the kinetic energy becomes greater than 10^{-3}), occurring at 39 s. Point B occurs at 43.3 and 48 s. The concentration during that time frame is shown in the top right panel. Point C occurs at 74 and 81.25 s, and the concentration during that time is shown in the bottom right panel. Finally, point D occurs at 190 s.*
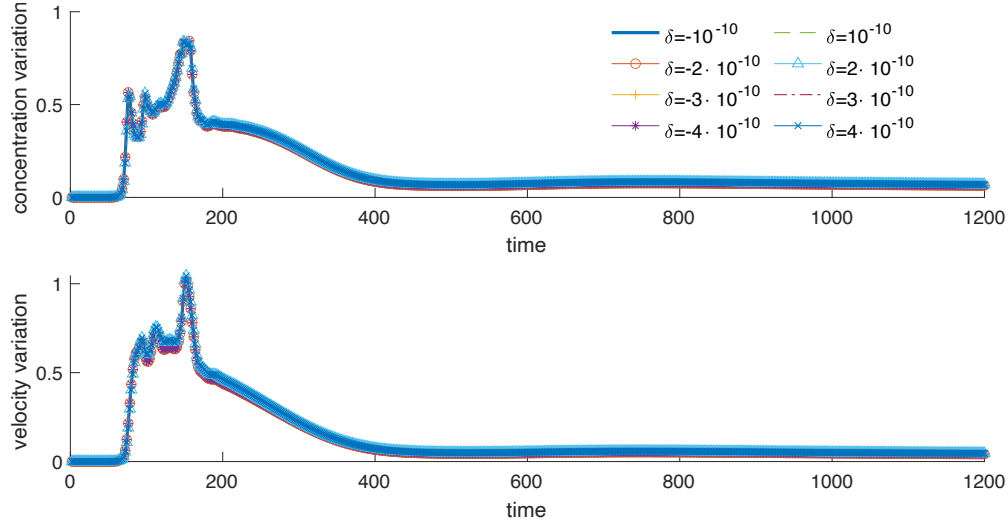
FIG. 4.4. *The L^2-norm of concentration (top) and velocity (bottom) variations with respect to the solution obtained using $\delta = 0$ over time.*

-2.5e+05 -1e+05   0   1e+05   2.5e+05
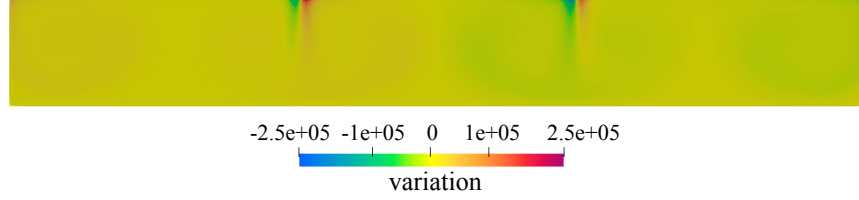
variation

FIG. 4.5. *The surface plot of the variation between the solutions obtained with $\delta = -10^{-10}$ and $\delta = 0$ at time $t = 1200$ s.*

a stable, steady solution. We had a total of 4770 rejected trials with $\delta = 0$, 4620 rejected trials when $\delta = -10^{-10}$, and 4494 rejected trials when $\delta = 10^{-10}$. Despite the large number of rejected trials, the adaptive time stepping is more efficient than the fixed time stepping, since it allows the use of large time steps during the majority of the simulation, while maintaining accuracy by using smaller time steps when needed.
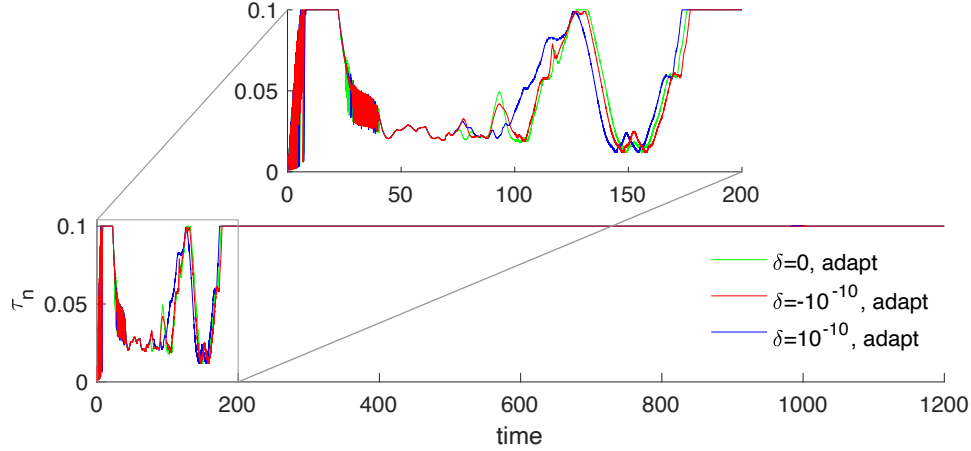
FIG. 4.6. *The time steps, τ_n, used in the adaptive time-stepping algorithm obtained with $\delta = 0$ (green line), $\delta = -10^{-10}$ (red line) and $\delta = 10^{-10}$ (blue line).*

4.3. Example 3. In this example, we consider similar settings as in Example 4.2, but with viscosity which depends on the concentration. In particular, we consider the following two models for the viscosity:

$$v_1(c) = \begin{cases} v_0, & c_r \leq 0, \\ v_0(1 + 2.5c_r + 5.3c_r^2), & c_r > 0, \end{cases} \tag{4.2}$$

and

$$v_2(c) = \begin{cases} v_0, & c_r \leq 0, \\ v_0(1 + 2.5c_r + 5.3c_r^2), & 0 < c_r \leq 10\%, \\ v_0 \exp\left(\dfrac{2.5c_r}{1 - 1.4c_r}\right), & 10 < c_r \leq 60\%, \\ v_0 \exp(9.375), & c_r > 60\%, \end{cases} \tag{4.3}$$

where $c_r = c/c_{max}$ is the relative concentration. We define c_{max} to be $c_{max} = 6.3 \cdot 10^7$ in (4.2), and $c_{max} = 7 \cdot 10^6$ in (4.3). In that way, (4.2) describes the constitutive relation for the viscosity in low concentration regime, based on work by Batchelor [8], and (4.3) combines the low concentration regime with the constitutive relation for high concentrations, proposed by Mooney [44]. This relation was also used in [17] to capture both low and high concentration regimes.

We use fixed time-stepping with $\tau_n = 10^{-2}$, for all n, and a structured, Union Jack mesh, where we consider 3 mesh sizes: mesh 1 consisting of 15×120 elements, mesh 2 consisting of 22×176 elements, and mesh 3 consisting of 30×240 elements. Other parameter settings are the same as in Table 4.3.
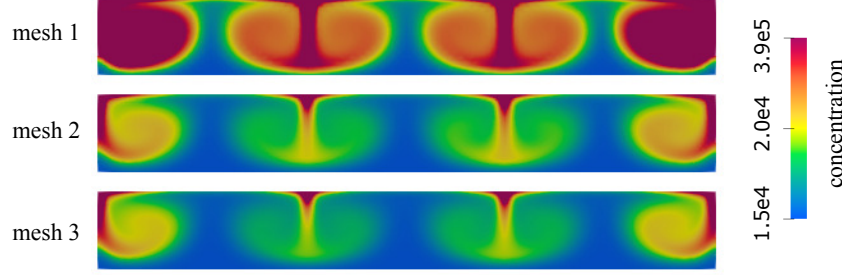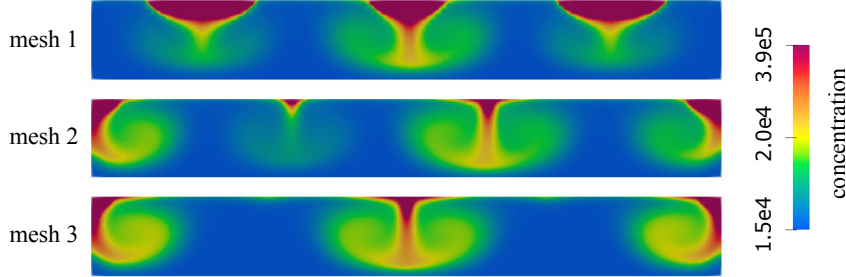
18

FIG. 4.7. *Concentration obtained using the viscosity model given by* (4.2) *on three different meshes at time* $t = 1200$ *s.*

Figure 4.7 shows the concentration obtained using the viscosity model described in (4.2) on three different meshes. We observe that as we refine the mesh, the results appear to converge, with concentration pattern showing 3 falling fingers (two in the middle and one half on each side). This agrees very well with the results obtained in Example 4.2, indicating that the constant viscosity model predicts the same solution as the low concentration viscosity model.

FIG. 4.8. *Concentration obtained using the viscosity model given by* (4.3) *on three different meshes at time* $t = 1200$ *s.*

The results obtained using both low and high concentration model (4.3) are shown in Figure 4.8. When the coarsest mesh (mesh 1) is used, the concentration forms three falling fingers. When the intermediate mesh is used, we obtain three fingers again. However, one of the inner fingers does not appear to be fully developed. That finger first starts to form around $t = 200$ s, but then it moves to the right and merges with the next finger. This is repeated around $t = 300$ s. At $t = 450$ s, the finger starts forming again, but this time it stays in its position until the end of the simulation at $T = 1200$ s. When the finest mesh (mesh 3) is used, we obtain only two fingers (one in the middle and one half on each side). In this example, the aspect ratio is $\mu = \ell/H = L/(NH) = 4$, and is closer to $2\sqrt{3} \approx 3.4641$, which is the value the authors in [32] argue that μ should grow to.

5. Conclusions. In this work, we investigate the dynamics of cultive fluid with negative geotaxis orientation and movement. To solve this system, we propose an adaptive, partitioned numerical method based on Cauchy's one-legged 'θ-like' scheme. We prove that the subiterative step in the numerical method is convergent, and that the scheme is stable. Our numerical study reveals that the first order of convergence is obtained when $\theta_n = 1$, and the second order of convergence is obtained when $\theta_n = 1/2$, as expected. We also computationally investigate a more realistic example of bioconvection in a culture of motile flagellates.

Previous works [2, 31] involving pattern formation arising from bioconvection found that patterns were sensitive to changes in domain aspect ratios, initial concentration or density changes. Motivated by the study in [31] where a constant viscosity is used, we perform a similar example to the one considered in [31], but with a longer computational time. We also perturb the initial concentration by the addition of a scalar value of order 10^{-10}. The resulting solutions of the perturbed concentration yield interesting results. All solutions are identical in all cases of initial concentration until $t = 68$ s when the energy trajectories (Fig. 4.2) evolve along similar paths but deviate slightly from each other. At approximately $t = 250$ s, the energy plots come together overlapping trajectories. This overlap demonstrates that in all cases of perturbations to initial concentration, the approximations converge to the same stable solution. Some of our observations seem to contradict the hypothesis presented in [31] that the potential energy is minimized as the

system develops in time. Our findings demonstrate that the stable solution begins at approximately $t = 250$ s and continues until our final $T = 1200$ s, but both the kinetic and potential energy continue to increase slowly in time. In the Discussion section in [32, pp 774], the authors acknowledge differences between experimental results and their theoretical results and numerical conclusions, and say that the exceptions may be attributed to intermittent generation of fingers and hysteresis.

In several of the perturbed concentration tests, we investigate the time adaptivity properties of the proposed method. The adaptive scheme requires small time steps until time $t = 200$ s. After that, in all cases considered here, the time step increases to the largest allowed value. This agrees with the observed dynamics of the system, where the flow patterns across all simulations rapidly change until $t = 200$ s. This is the most complex portion of the simulated time and therefore it is understandable that a smaller time step is needed for the flow to be appropriately resolved.

We expand our long time simulations of bioconvection to also consider the case where the fluid viscosity is a function of concentration. Two different viscosity models are used: a low concentration model and a model which includes both low and high concentrations of micro-organisms. In the low concentration regime, concentration appears to converge to the same stable solution as we refine the computational mesh. In contrast, the model that includes both low and high concentration is used, larger differences are observed. Namely, final solutions obtained on different meshes do not have the same number of falling fingers. More work is necessary to fully understand these results, especially related to the chaotic properties of the system. This is out of scope of this paper, and will be considered in our future work.

Appendix A. .
LEMMA A.1. *The operator $\mathscr{G}_M$ (2.12) is well-defined and*

$$\|\mathscr{G}_M(\boldsymbol{u},c)\| \leq \|\mathscr{A}(\boldsymbol{u},c)\| + C_\Omega \|(\boldsymbol{u},c)\|_1^{3/2} \|(\boldsymbol{u},c)\|_2^{1/2} + (\nu(c) - \nu_*)\|(\boldsymbol{u},c)\|_2 + g|\Omega|^{1/2} + g\gamma\|c\| + U\|\nabla c\|,$$

for all $(\boldsymbol{u},c) \in D(\mathscr{A})$.

Proof of Lemma A.1 By the definition (2.12) we have that

$$\|\mathscr{G}_M(\boldsymbol{u},c)\| \leq \|\mathscr{A}(\boldsymbol{u},c)\| + \|\mathscr{B}_M(\boldsymbol{u},c)\|, \qquad \forall(\boldsymbol{u},c) \in D(\mathscr{A}).$$

First we consider the case $\|(\boldsymbol{u}_1,c_1)\|_1 \leq M$. Then, using (2.7) with $m_1 = 1, m_2 = \frac{1}{2}, m_3 = 0$, and the interpolation inequality, we have

$$|\langle \mathscr{B}_M(\boldsymbol{u}_1,c_1),(\boldsymbol{u}_2,c_2)\rangle| = |\mathscr{B}_0\big((\boldsymbol{u}_1,c_1),(\boldsymbol{u}_1,c_1),(\boldsymbol{u}_2,c_2)\big)|$$
$$\leq C_\Omega \|(\boldsymbol{u}_1,c_1)\|_1^{3/2} \|(\boldsymbol{u}_1,c_1)\|_2^{1/2} \|(\boldsymbol{u}_2,c_2)\|$$
$$+ \Big((\nu(c_1) - \nu_*)\|(\boldsymbol{u}_1,c_1)\|_2 + g|\Omega|^{1/2} + g\gamma\|c_1\| + U\|\nabla c_1\|\Big)\|(\boldsymbol{u}_2,c_2)\|,$$

and therefore

$$\|\mathscr{B}_M(\boldsymbol{u}_1,c_1)\| \leq C_\Omega \|(\boldsymbol{u}_1,c_1)\|_1^{3/2} \|(\boldsymbol{u}_1,c_1)\|_2^{1/2} + (\nu(c_1) - \nu_*)\|(\boldsymbol{u}_1,c_1)\|_2 + g|\Omega|^{1/2} + g\gamma\|c_1\| + U\|\nabla c_1\|,$$

for all $(\boldsymbol{u}_1,c_1) \in D(\mathscr{A}), \|(\boldsymbol{u}_1,c_1)\|_1 \leq M$.
In the other case, when $\|(\boldsymbol{u}_1,c_1)\|_1 > M$, we have similarly that

$$\|\mathscr{B}_M(\boldsymbol{u}_1,c_1)\|$$
$$\leq \frac{M^2}{\|(\boldsymbol{u}_1,c_1)\|_1^2}\Big(C_\Omega \|(\boldsymbol{u}_1,c_1)\|_1^{3/2} \|(\boldsymbol{u}_1,c_1)\|_2^{1/2} + (\nu(c_1) - \nu_*)\|(\boldsymbol{u}_1,c_1)\|_2 + g|\Omega|^{1/2} + g\gamma\|c_1\| + U\|\nabla c_1\|\Big)$$
$$\leq C_\Omega \|(\boldsymbol{u}_1,c_1)\|_1^{3/2} \|(\boldsymbol{u}_1,c_1)\|_2^{1/2} + (\nu(c_1) - \nu_*)\|(\boldsymbol{u}_1,c_1)\|_2 + g|\Omega|^{1/2} + g\gamma\|c_1\| + U\|\nabla c_1\|,$$

and therefore (2.13) follows for all $(\boldsymbol{u},c) \in D(\mathscr{A})$. □

In order to prove that the operator $\mathscr{G}_M$ is quasi-m-accretive, we need the following preliminary result on monotonicity of the quantised operator $\mathscr{B}_M$.

LEMMA A.2. *There exists a positive constant C_M such that*

$$\langle \mathscr{B}_M(\boldsymbol{u}_1,c_1) - \mathscr{B}_M(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2) \rangle \tag{A.1}$$
$$\geq \int_\Omega (\nu(c_2) - \nu_*)|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)|^2 dx$$
$$- \frac{\nu_*}{2}\|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|^2 - \frac{\Theta}{2}\|\nabla(c_1 - c_2)\|^2 - C_M(\|c_1 - c_2\|^2 + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|^2),$$

for all $(\boldsymbol{u}_1,c_1),(\boldsymbol{u}_2,c_2) \in V$.

Proof. We split the argument in three cases. First we note that, for $d = 3$, by (2.1), (2.7) and Hölder's inequality we have

$$\int_\Omega (\nu(c_1) - \nu(c_2))\nabla\boldsymbol{u}_1 \cdot \nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)dx + \int_\Omega \Big(((\boldsymbol{u}_1 - \boldsymbol{u}_2)\cdot\nabla)\boldsymbol{u}_1(\boldsymbol{u}_1 - \boldsymbol{u}_2) + ((\boldsymbol{u}_1 - \boldsymbol{u}_2)\cdot\nabla)c_1(c_1 - c_2)\Big)dx$$
$$\geq -\frac{\nu_*}{2}\|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|^2 - \frac{\Theta}{4}\|\nabla(c_1 - c_2)\|^2 - C(\|c_1 - c_2\|^2 + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|^2)(\|\nabla\boldsymbol{u}_1\|^4 + \|\nabla c_1\|^4). \tag{A.2}$$

(a) For the case $\|(\boldsymbol{u}_1,c_1)\|_{H^1(\Omega)}, \|(\boldsymbol{u}_2,c_2)\|_{H^1(\Omega)} \leq M$ we have that

$$\langle \mathscr{B}_M(\boldsymbol{u}_1,c_1) - \mathscr{B}_M(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2) \rangle = \langle \mathscr{B}(\boldsymbol{u}_1,c_1) - \mathscr{B}(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2) \rangle$$
$$= \int_\Omega (\nu(c_2) - \nu_*)|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)|^2 dx + \int_\Omega (\nu(c_1) - \nu(c_2))\nabla\boldsymbol{u}_1 \cdot \nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)dx$$
$$+ \int_\Omega \Big(((\boldsymbol{u}_1 - \boldsymbol{u}_2)\cdot\nabla)\boldsymbol{u}_1(\boldsymbol{u}_1 - \boldsymbol{u}_2) + ((\boldsymbol{u}_1 - \boldsymbol{u}_2)\cdot\nabla)c_1(c_1 - c_2)\Big)dx$$
$$+ g\gamma \int_\Omega (c_1 - c_2)i_d(\boldsymbol{u}_1 - \boldsymbol{u}_2)dx + \int_\Omega U\frac{\partial(c_1 - c_2)}{\partial x_d}(c_1 - c_2)dx,$$

and therefore by (A.2) and the Cauchy-Schwarz inequality

$$\langle \mathscr{B}_M(\boldsymbol{u}_1,c_1) - \mathscr{B}_M(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2) \rangle$$
$$\geq \int_\Omega (\nu(c_2) - \nu_*)|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)|^2 dx - \frac{\nu_*}{2}\|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|^2 - \frac{\Theta}{2}\|\nabla(c_1 - c_2)\|^2 - C_M(\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|^2 + \|c_1 - c_2\|^2).$$

(b) For the case of $\|(\boldsymbol{u}_1,c_1)\|_1 > M$, $\|(\boldsymbol{u}_2,c_2)\|_1 \leq M$ (similar estimates are obtained when $\|(\boldsymbol{u}_1,c_1)\|_1 \leq M$, $\|(\boldsymbol{u}_2,c_2)\|_1 >$ ▮ M) we have

$$\langle \mathscr{B}_M(\boldsymbol{u}_1,c_1) - \mathscr{B}_M(\boldsymbol{u}_2,c_2), (\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2) \rangle \tag{A.3}$$
$$= \frac{M^2}{\|(\boldsymbol{u}_1,c_1)\|_1^2}\mathscr{B}_0((\boldsymbol{u}_1,c_1),(\boldsymbol{u}_1,c_1),(\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2)) - \mathscr{B}_0((\boldsymbol{u}_2,c_2),(\boldsymbol{u}_2,c_2),(\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2))$$
$$= \frac{M^2}{\|(\boldsymbol{u}_1,c_1)\|_1^2}\Big(\mathscr{B}_0((\boldsymbol{u}_1,c_1),(\boldsymbol{u}_1,c_1),(\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2)) - \mathscr{B}_0((\boldsymbol{u}_2,c_2),(\boldsymbol{u}_2,c_2),(\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2))\Big)$$
$$+ \Big(\frac{M^2}{\|(\boldsymbol{u}_1,c_1)\|_1^2} - 1\Big)\mathscr{B}_0((\boldsymbol{u}_2,c_2),(\boldsymbol{u}_2,c_2),(\boldsymbol{u}_1 - \boldsymbol{u}_2, c_1 - c_2)).$$

In a manner similar to the derivation of (A.2), the first term in (A.3) gives

$$
\frac{M^2}{\|(u_1,c_1)\|_1^2}\Big(\mathscr{B}_0\big((u_1,c_1),(u_1,c_1),(u_1-u_2,c_1-c_2)\big)-\mathscr{B}_0\big((u_2,c_2),(u_2,c_2),(u_1-u_2,c_1-c_2)\big)\Big)
$$

$$
=\frac{M^2}{\|(u_1,c_1)\|_1^2}\bigg[\int_\Omega(\nu(c_2)-\nu_*)|\nabla(u_1-u_2)|^2dx+\int_\Omega(\nu(c_1)-\nu(c_2))\nabla u_1\cdot\nabla(u_1-u_2)dx
$$

$$
+\int_\Omega\Big(((u_1-u_2)\cdot\nabla)u_1(u_1-u_2)+((u_1-u_2)\cdot\nabla)c_1(c_1-c_2)\Big)dx
$$

$$
+g\gamma\int_\Omega(c_1-c_2)i_d(u_1-u_2)dx+\int_\Omega U\frac{\partial(c_1-c_2)}{\partial x_d}(c_1-c_2)dx\bigg]
$$

$$
\geq\frac{M^2}{\|(u_1,c_1)\|_1^2}\int_\Omega(\nu(c_2)-\nu_*)|\nabla(u_1-u_2)|^2dx
$$

$$
-\frac{\nu_*}{4}\|\nabla(u_1-u_2)\|^2-\frac{\Theta}{4}\|\nabla(c_1-c_2)\|^2-C_M\big(\|c_1-c_2\|^2+\|u_1-u_2\|^2\big).
$$

Using the fact that $\|(u_2,c_2)\|_1^2\leq M^2$, the triangular inequality, the Poincaré and Holder inequalities, we derive similarly that the last term in (A.3) can be bounded below as follows

$$
\Big(\frac{M^2}{\|(u_1,c_1)\|_1^2}-1\Big)\mathscr{B}_0\big((u_2,c_2),(u_2,c_2),(u_1-u_2,c_1-c_2)\big)
$$

$$
=-\frac{\|(u_1,c_1)\|_1^2-M^2}{\|(u_1,c_1)\|_1^2}\int_\Omega\Big((\nu(c_2)-\nu_*)\nabla u_2\cdot\nabla(u_1-u_2)+(u_2\cdot\nabla)u_2(u_1-u_2)
$$

$$
+(u_2\cdot\nabla)c_2(c_1-c_2)+g(1+\gamma c_2)i_d(u_1-u_2)+\int_\Omega U\frac{\partial c_2}{\partial x_d}(c_1-c_2)dx\Big)dx
$$

$$
\geq-\frac{(\|(u_1,c_1)\|_1-\|(u_2,c_2)\|_1)(\|(u_1,c_1)\|_1+\|(u_2,c_2)\|_1)}{\|(u_1,c_1)\|_1^2}\int_\Omega\Big|(\nu(c_2)-\nu_*)\nabla u_2\cdot\nabla(u_1-u_2)
$$

$$
+(u_2\cdot\nabla)u_2(u_1-u_2)+(u_2\cdot\nabla)c_2(c_1-c_2)+g(1+\gamma c_2)i_d(u_1-u_2)+U\frac{\partial c_2}{\partial x_d}(c_1-c_2)\Big|dx
$$

$$
\geq-\frac{\nu_*}{4}\|\nabla(u_1-u_2)\|^2-\frac{\Theta}{4}\|\nabla(c_1-c_2)\|^2-C_M\big(\|u_1-u_2\|^2+\|c_1-c_2\|^2\big),
$$

therefore

$$
\langle\mathscr{B}_M(u_1,c_1)-\mathscr{B}_M(u_2,c_2),(u_1-u_2,c_1-c_2)\rangle
$$

$$
\geq\frac{M^2}{\|(u_1,c_1)\|_1^2}\int_\Omega(\nu(c_2)-\nu_*)|\nabla(u_1-u_2)|^2dx
$$

$$
-\frac{\nu_*}{2}\|\nabla(u_1-u_2)\|^2-\frac{\Theta}{2}\|\nabla(c_1-c_2)\|^2-C_M\big(\|c_1-c_2\|^2+\|u_1-u_2\|^2\big).
$$

(c) Finally, when $\|(u_1,c_1)\|_{H^1(\Omega)},\|(u_2,c_2)\|_{H^1(\Omega)}\geq M$ we have that

$$
\langle\mathscr{B}_M(u_1,c_1)-\mathscr{B}_M(u_2,c_2),(u_1-u_2,c_1-c_2)\rangle
$$

$$
=\frac{M^2}{\|(u_1,c_1)\|_1^2}\mathscr{B}_0\big((u_1-u_2,c_1-c_2),(u_1,c_1),(u_1-u_2,c_1-c_2)\big)
$$

$$
+\Big(\frac{M^2}{\|(u_1,c_1)\|_1^2}-\frac{M^2}{\|(u_2,c_2)\|_1^2}\Big)\mathscr{B}_0\big((u_2,c_2),(u_2,c_2),(u_1-u_2,c_1-c_2)\big),
$$

and the argument follows in a manner similar to the previous cases.

□

LEMMA A.3. *The operator $\mathscr{G}_M$ is quasi-m-accretive, i.e., there exists α_M such that $\mathscr{G}_M + \alpha_M I$ is m-accretive (maximal monotone) in $H \times H$.*

Proof. Using (A.1) in Lemma A.2, and (2.4), we have that

$$
\begin{aligned}
&\langle (\mathscr{G}_M + \alpha_M I)(\boldsymbol{u}_1, c_1) - (\mathscr{G}_M + \alpha_M I)(\boldsymbol{u}_2, c_2), (\boldsymbol{u}_1, c_1) - (\boldsymbol{u}_2, c_2) \rangle \\
&= \nu_* \|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|^2 + \Theta \|\nabla(c_1 - c_2)\|^2 + \langle \mathscr{B}_M(\boldsymbol{u}_1, c_1) - \mathscr{B}_M(\boldsymbol{u}_2, c_2), (\boldsymbol{u}_1, c_1) - (\boldsymbol{u}_2, c_2) \rangle \\
&\quad + \alpha_M \left(\|\boldsymbol{u}_1 - \boldsymbol{u}_2\|^2 + \|c_1 - c_2\|^2 \right) \\
&\geq \frac{\nu_*}{2} \|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|^2 + \frac{\Theta}{2} \|\nabla(c_1 - c_2)\|^2 + \frac{M^2}{\|(\boldsymbol{u}_1, c_1)\|_1^2} \int_\Omega (\nu(c_2) - \nu_*) |\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)|^2 d\boldsymbol{x}.
\end{aligned}
$$

for all $(\boldsymbol{u}_1, c_1), (\boldsymbol{u}_2, c_2) \in D(\mathscr{G}_M)$, and any $\alpha_M \geq C_M$. □

REFERENCES

[1] G. V. ALEKSEEV AND E. ADOMAVICHUS, *Theoretical analysis of inverse extremal problems of admixture diffusion in viscous fluids*, Journal of Inverse and Ill-posed Problems, 9 (2001), pp. 435–468.

[2] A. A. AVRAMENKO, Y. Y. KOVETSKA, AND I. V. SHEVCHUK, *Lorenz approach for analysis of bioconvection instability of gyrotactic motile microorganisms*, Chaos Solitons and Fractals, 166 (2023), p. 112957.

[3] V. BARBU, *Analysis and control of nonlinear infinite-dimensional systems*, vol. 190 of Mathematics in Science and Engineering, Academic Press Inc., Boston, MA, 1993.

[4] ———, *The time optimal control of Navier-Stokes equations*, Systems Control Lett., 30 (1997), pp. 93–100.

[5] ———, *Partial differential equations and boundary value problems*, vol. 441 of Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 1998. Translated and revised from the 1993 Romanian original by the author.

[6] ———, *Stabilization of Navier-Stokes flows*, Communications and Control Engineering Series, Springer, London, 2011.

[7] V. BARBU AND S. S. SRITHARAN, *Flow invariance preserving feedback controllers for the Navier-Stokes equation*, J. Math. Anal. Appl., 255 (2001), pp. 281–307.

[8] G. K. BATCHELOR AND J. T. GREEN, *The determination of the bulk stress in a suspension of spherical particles to order c^2*, Journal of Fluid Mechanics, 56 (1972), p. 401–427.

[9] L. C. BERSELLI, S. FAGIOLI, AND S. SPIRITO, *Suitable weak solutions of the Navier-Stokes equations constructed by a space-time numerical discretization*, J. Math. Pures Appl. (9), 125 (2019), pp. 189–208.

[10] J. L. BOLDRINI, M. A. ROJAS-MEDAR, AND M. D. ROJAS-MEDAR, *Existence and uniqueness of stationary solutions to bioconvective flow equations*, Electron. J. Differential Equations, (2013), pp. No. 110, 15.

[11] J. F. BRADY, *The rheological behavior of concentrated colloidal dispersions*, The Journal of Chemical Physics, 99 (1993), pp. 567–581.

[12] M. BUKAČ, G. FU, A. SEBOLDT, AND C. TRENCHEA, *Time-adaptive partitioned method for fluid-structure interaction problems with thick structures*, J. Comput. Phys., (2022).

[13] ———, *Time-adaptive partitioned method for fluid-structure interaction problems with thick structures*, J. Comput. Phys., 473 (2023), pp. Paper No. 111708, 19.

[14] M. BUKAČ, A. SEBOLDT, AND C. TRENCHEA, *Refactorization of Cauchy's Method: A Second-Order Partitioned Method for Fluid-Thick Structure Interaction Problems*, J. Math. Fluid Mech., 23 (2021), p. 64.

[15] M. BUKAČ AND C. TRENCHEA, *Adaptive, second-order, unconditionally stable partitioned method for fluid-structure interaction*, Comput. Methods Appl. Mech. Engrg., 393 (2022), pp. Paper No. 114847, 24.

[16] J. BURKARDT AND C. TRENCHEA, *Refactorization of the midpoint rule*, Appl. Math. Lett., 107 (2020), p. 106438.

[17] Y. CAO, S. CHEN, AND H.-W. VAN WYK, *Well-posedness and finite element approximation of time dependent generalized bioconvective flow*, Numer. Methods Partial Differential Equations, 36 (2020), pp. 709–733.

[18] A.-L. CAUCHY, *Équations différentielles ordinaires*, Éditions Études Vivantes, Ltée., Ville Saint-Laurent, QC; Johnson Reprint Corp., New York, 1981. Cours inédit. Fragment.

[19] S. CHILDRESS, M. LEVANDOWSKY, AND E. A. SPIEGEL, *Pattern formation in a suspension of swimming microorganisms: equations and stability theory*, Journal of Fluid Mechanics, 69 (1975), pp. 591–613.

[20] S. CHILDRESS AND J. PERCUSS, *Nonlinear aspects of chemotaxis*, Math. Biosc., 56 (1981), pp. 217–237.

[21] B. CLIMENT-EZQUERRA, L. FRIZ, AND M. A. ROJAS-MEDAR, *Time-reproductive solutions for a bioconvective flow*, Ann. Mat. Pura Appl. (4), 192 (2013), pp. 763–782.

[22] P. CONSTANTIN AND C. FOIAS, *Navier-Stokes equations*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1988.

[23] A. CĂPĂŢÎNĂ AND R. STAVRE, *A control problem in biconvective flow*, J. Math. Kyoto Univ., 37 (1997), pp. 585–595.

[24] A. EINSTEIN, *Eine neue bestimmung der moleküldimensionen*, Annalen der Physik, 324 (1906), pp. 289–306.

[25] T. FUKUI, M. KAWAGUCHI, AND K. MORINISHI, *Numerical study on the inertial effects of particles on rheology of a suspension*, Advances in Mechanical Engineering, 11 (2019), pp. 1–10.

[26] A. FURSIKOV, *Optimal control of distributed systems. Theory and applications*, vol. 187 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 2000. Translated from the 1999 Russian original by Tamara Rozhkovskaya.

[27] S. GHORAI AND N. A. HILL, *Development and stability of gyrotactic plumes in bioconvection*, J. Fluid Mech., 400 (1999), pp. 1–31.

[28] ———, *Periodic arrays of gyrotactic plumes in bioconvection*, Physics of Fluids, 12 (2000), pp. 5–22.

[29] S. GHORAI, R. SINGH, AND N. A. HILL, *Wavelength selection in gyrotactic bioconvection*, Bulletin of Mathematical Biology, 77 (2015).

[30] P. GRISVARD, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[31] A. HARASHIMA, M. WATANABE, AND I. FUJISHIRO, *Evolution of bioconvection patterns in a culture of motile flagellates*, The Physics of Fluids, 31 (1988), pp. 764–775.

[32] ———, *Evolution of bioconvection patterns in a culture of motile flagellates*, The Physics of Fluids, 31 (1988), pp. 764–775.

[33] F. HECHT, *New development in freefem++*, J. Numer. Math., 20 (2012), pp. 251–265.

[34] J. G. HEYWOOD AND R. RANNACHER, *Finite-element approximation of the nonstationary Navier-Stokes problem. IV. Error analysis for second-order time discretization*, SIAM J. Numer. Anal., 27 (1990), pp. 353–384.

[35] E. KALNAY, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2003.

[36] Y. KAN-ON, K. NARUKAWA, AND Y. TERAMOTO, *On the equations of bioconvective flow*, J. Math. Kyoto Univ., 32 (1992), pp. 135–153.

[37] A. KARIMI AND A. M. ARDEKANI, *Gyrotactic bioconvection at pycnoclines*, Journal of Fluid Mechanics, 733 (2013), pp. 245–267.

[38] I. M. KRIEGER AND T. J. DOUGHERTY, *A Mechanism for Non-Newtonian Flow in Suspensions of Rigid Spheres*, Transactions of The Society of Rheology, 3 (1959), pp. 137–152.

[39] A. LABOVSKY AND C. TRENCHEA, *Large eddy simulation for turbulent magnetohydrodynamic flows*, J. Math. Anal. Appl., 377 (2011), pp. 516–533.

[40] O. A. LADYŽENSKAJA, *Solution "in the large" to the boundary-value problem for the Navier-Stokes equations in two space variables*, Soviet Physics. Dokl., 123 (3) (1958), pp. 1128–1131 (427–429 Dokl. Akad. Nauk SSSR).

[41] W. LAYTON, W. PEI, Y. QIN, AND C. TRENCHEA, *Time step adaptivity in the method of Dahlquist, Liniger and Nevanlinna*, tech. rep., University of Pittsburgh, 2022.

[42] M. LEVANDOWSKY, W. S. CHILDRESS, E. A. SPIEGEL, AND S. H. HUTNER, *A mathematical model of pattern formation by swimming microorganisms*, The Journal of Protozoology, 22 (1975), pp. 296–306.

[43] J.-L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris; Gauthier-Villars, Paris, 1969.

[44] M. MOONEY, *The viscosity of a concentrated suspension of spherical particles*, Journal of Colloid Science, 6 (1951), pp. 162–170.

[45] M. H. PROTTER AND H. F. WEINBERGER, *Maximum principles in differential equations*, Springer-Verlag, New York, 1984. Corrected reprint of the 1967 original.

[46] R. TEMAM, *Navier-Stokes equations*, AMS Chelsea Publishing, Providence, RI, 2001. Theory and numerical analysis, Reprint of the 1984 edition.