ESAIM: M2AN 52 (2018) 1109–1135
https://doi.org/10.1051/m2an/2018036

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

STABILITY ESTIMATES FOR SYSTEMS WITH SMALL CROSS-DIFFUSION[☆,☆☆]

Luca Alasio, Maria Bruna and Yves Capdeboscq[*]

Abstract. We discuss the analysis and stability of a family of cross-diffusion boundary value problems with nonlinear diffusion and drift terms. We assume that these systems are close, in a suitable sense, to a set of decoupled and linear problems. We focus on stability estimates, that is, continuous dependence of solutions with respect to the nonlinearities in the diffusion and in the drift terms. We establish well-posedness and stability estimates in an appropriate Banach space. Under additional assumptions we show that these estimates are time independent. These results apply to several problems from mathematical biology; they allow comparisons between the solutions of different models *a priori*. For specific cell motility models from the literature, we illustrate the limit of the stability estimates we have derived numerically, and we document the behaviour of the solutions for extremal values of the parameters.

Mathematics Subject Classification. 35K55, 35B30, 35Q92, 65M15

Received January 19, 2018. Accepted May 24, 2018.

1. Introduction

1.1. Background and motivation

In this paper we analyse a class of nonlinear cross-diffusion systems of PDEs which model multi-species populations in presence of short-range interactions between individuals. We assume that these systems are close, in a suitable sense, to decoupled sets of linear parabolic evolution problems. Such problems arise in many applications in mathematical biology, such as chemotactic cell migration, ion transport through cell membranes, and spatial segregation in interacting species. The strength of the interactions (and therefore of the nonlinear terms) is quantified with a small parameter ϵ, so that when $\epsilon = 0$ the system becomes diagonal and linear. The biological justification for these models comes from weakly-interacting species, whereby interactions between populations (such as excluded-volume or chemotactic interactions) are present but are not dominant over the isolated species behaviour.

The cross-diffusion systems we are interested in have the form

$$\partial_t u - \operatorname{div}\left[\mathfrak{D}(t,x,u)\nabla u - \mathfrak{F}(t,x,u)u\right] = 0, \quad \text{in} \quad \Omega,\ t>0, \tag{1.1a}$$

[☆]L. Alasio was supported by the Engineering and Physical Sciences Research Council grant [EP/L015811/1].
[☆☆]M. Bruna was supported by the L'Oréal UK and Ireland Fellowship For Women In Science.

Keywords and phrases: Cross diffusion, continuous dependence, quasilinear parabolic systems.

Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK.

[*] Corresponding author: yves.capdeboscq@maths.ox.ac.uk

Article published by EDP Sciences

© EDP Sciences, SMAI 2018

with boundary and initial conditions

$$[\mathfrak{D}(t,x,u)\nabla u - \mathfrak{F}(t,x,u)u]\cdot\nu = 0, \quad \text{on} \quad \partial\Omega,\ t>0, \tag{1.1b}$$
$$u(0,\cdot) = u^0, \quad \text{in} \quad \Omega, \tag{1.1c}$$

where Ω is a smooth, bounded, and connected domain in $\mathbb{R}^d$ ($d = 1,2,3$), ν denotes the outward normal on $\partial\Omega$, and $u = (u_1,\ldots,u_m)$ is the vector of densities of each species. The divergence div and gradient ∇ represent derivatives with respect to the d spatial variables. Here $\mathfrak{D}(t,x,u)$ and $\mathfrak{F}(t,x,u)$ are $m\times m$ matrices of diffusion tensors and drift vectors, respectively (see (1.12) for further details). In particular, the entries of the diffusion tensor $\mathfrak{D}$ may be scalars in the case of isotropic diffusion, or $d\times d$ tensors in the case of anisotropic diffusion. The drift matrix elements $\mathfrak{F}_{ij}$ are d-dimensional vectors. In our class of cross-diffusion systems, the matrices $\mathfrak{D}$ and $\mathfrak{F}$ are close to matrices that are diagonal and independent of u, that is, they can be written in the form

$$\begin{aligned}\mathfrak{D}(t,x,u) &= \mathfrak{D}^{(0)}(t,x) + \epsilon\mathfrak{D}^{(1)}(t,x,u) + O(\epsilon^2),\\ \mathfrak{F}(t,x,u) &= \mathfrak{F}^{(0)}(t,x) + \epsilon\mathfrak{F}^{(1)}(t,x,u) + O(\epsilon^2),\end{aligned} \tag{1.2}$$

where ϵ is a small parameter.

The focus of this paper is to study the stability of the solutions to (1.1) under perturbations of order ϵ. We establish that the solutions depend continuously on the nonlinearities $\mathfrak{D}^{(1)}$ and $\mathfrak{F}^{(1)}$ for ϵ small enough. The cross-diffusion model (1.1) is a non-linear system, and this combines two types of difficulties, namely the non-linearity and the fact that fully coupled parabolic systems of equations do not enjoy, in general, the same smoothness properties as parabolic equations (see, for example, [10], chap. 9, and [11]). Our results are detailed in Proposition 1.5 and Theorem 1.7. They are quantitative, in the sense that we provide a bound on ϵ below which our perturbation result applies. The novelty of our analysis consists in the unified approach to the study of regularity and stability properties in "strong" Sobolev norms for a relatively wide class of nonlinear cross-diffusion systems.

Our stability estimate uses the underlying regularity of the system, which, as we will see, it inherits from the leading order model, consisting of decoupled linear evolution equations. We show that for small perturbations at least some of the regularity is preserved and, using a fixed point argument, we deduce a stability estimate with respect to the nonlinearities of the model.

Similar results concerning nonlinear systems where interactions between species (or components) are limited to lower order term (so-called weakly coupled systems) are available in the work of Camilli and Marchi [5]. They extend the results available for scalar equations in terms of continuous dependence estimates in the sup norm using the doubling variable method [14] and viscosity solutions. Their results do not apply to fully coupled systems with cross diffusion present such as the ones we are considering. Continuous dependence for fully coupled quasilinear systems was studied by Cannon *et al.* [6]. They established existence and uniqueness, following arguments of Ladyzhenskaya, Solonnikov, and Ural'tseva [15] in larger Sobolev spaces (weaker norms). They derive stability estimates under additional integrability properties assumptions for the gradients. We establish existence and uniqueness in stronger norms, removing the need of additional regularity assumptions.

There are several models, especially in mathematical biology, that fit into the class of systems (1.1) and (1.2). This is the case for models describing the transport of cells or ions while accounting for the finite-size of particles [3, 4, 17, 21]. These models were derived from stochastic agent-based models assuming that the concentration of cells or ions is not too large, so that the transport dominates over the finite-size interactions between cells or ions. The diffusion and drift matrices become density-dependent due to the interactions, but this correction is small since it scales with the excluded volume in the system. Below we present three of such models, and show how they fit into our framework.

Example 1.1 (Random walk on a lattice with size exclusion)**.** A cross-diffusion model for two interacting species was employed to describe the motility of biological cells by Simpson *et al.* [21] or ion transport by Burger *et al.* [4]. The models were derived assuming that particles are restricted to a regular square lattice and undergo a simple exclusion random walk, in which a particle can only jump to a site if it is presently unoccupied. In order to obtain a continuum model such as (1.1) from these so-called lattice-based models, it is generally assumed that the occupancies of adjacent sides are independent, so that the jumping probabilities take a simple form and do not require correlation functions [4, 21]. Clearly, such an approximation is poor when the overall occupancy of the lattice is high. As a result, these models are generally considered valid for low-lattice occupancies.

The models in [4, 21] consider two species of equal size, whose diameter is given by the lattice spacing ε, that undergo a random walk with isotropic diffusion D_i and external potential $V_i(x)$, for $i = 1, 2$ (the jumping rates increase with D_i and the jumps are biased in the direction of $-\nabla V_i(x)$). There are N_1 particles of the first species, and N_2 of the second species. Under these assumptions, a cross-diffusion model of the form (1.1) is obtained, where the population densities $u_1(t, x)$ and $u_2(t, x)$ represent the probability that a particle from first or second species respectively is at $x \in \Omega$ at time t. The diffusion and drift matrices are given by [4]

$$\mathfrak{D}(u) = \begin{pmatrix} D_1(1 - \epsilon \bar{N}_2 u_2) & \epsilon D_1 \bar{N}_2 u_1 \\ \epsilon D_2 \bar{N}_1 u_2 & D_2(1 - \epsilon \bar{N}_1 u_1) \end{pmatrix}, \tag{1.3a}$$

$$\mathfrak{F}(u) = \begin{pmatrix} -\nabla V_1(1 - \epsilon \bar{N}_1 u_1) & \epsilon \bar{N}_2 u_1 \nabla V_1 \\ \epsilon \bar{N}_1 u_2 \nabla V_2 & -\nabla V_2(1 - \epsilon \bar{N}_2 u_2) \end{pmatrix}, \tag{1.3b}$$

where $\epsilon = (N_1 + N_2)\varepsilon^d/|\Omega| \ll 1$ represents the total volume fraction of the lattice occupied by particles and $\bar{N}_i = N_i/(N_1 + N_2)$. We have written (1.3) in a form consistent with our notations, which differ slightly from those used in [4]. Global existence for such model was shown in [7]. In that paper, as in most works using lattice-based models, the continuum model is written in terms of the volume concentrations $\hat{u}_i$, so that the mass of $\hat{u}_i$ equals the total volume occupied by species i (that is, $\int_\Omega \hat{u}_i(t, x)\mathrm{d}x = N_i\varepsilon^d/|\Omega|$). We write (1.3) in terms of probability densities u_i, which implies that $\int_\Omega u_i \mathrm{d}x = 1$. The two quantities are related by the identity $\hat{u}_i = \bar{N}_i \epsilon u_i$. The potentials appearing in (1.3b), V_i, are not rescaled by the diffusion coefficient as it is done in [4]. The number of species can take any values provided that ϵ, is small. The matrices in (1.3) are of the form (1.2) that we consider in this paper. There are also other lattice-based models that fit well into such framework, such as that derived by Shigesada *et al.* [20] to describe spatial segregation of interacting animal populations.

Example 1.2 (Brownian motion with size exclusion)**.** A cross-diffusion model for two interacting species of diffusive particles was obtained by Bruna and Chapman for $d = 2, 3$ in [3], starting from a system with two types of Brownian hard spheres. The population densities $u_i(t, x)$, $i = 1, 2$, represent the probability that a particle of species i is at $x \in \Omega$ at time t, and so $\int_\Omega u_i(t, x)\mathrm{d}x = 1$. The model assumes there are N_i particles of species i, of diameter ε_i and isotropic diffusion constant D_i. The position X_i of each particle in species i evolves in time according to the stochastic differential equation

$$\mathrm{d}X_i(t) = \sqrt{2D_i}\mathrm{d}W(t) - \nabla V_i(X_i(t))\mathrm{d}t, \tag{1.4}$$

where $i = 1$ or 2, and W are independent, d-dimensional standard Brownian motions. Reflective boundary conditions are imposed whenever two particles are in contact ($\|X_i - X_j\| = (\varepsilon_i + \varepsilon_j)/2$, when X_i and X_j are of type i and j, respectively), as well as on the boundary of the domain $\partial\Omega$.

The cross-diffusion model is derived using the method of matched asymptotic expansions under the assumption that the volume fraction of the system is small, or equivalently, that $(N_1\varepsilon_1^d + N_2\varepsilon_2^d)/|\Omega| \sim \epsilon \ll 1$, where ϵ is defined as in Example 1.1 with $\varepsilon = (\varepsilon_1 + \varepsilon_2)/2$. When the number of particles in each species is large, the cross-diffusion model in [3] is of the form (1.1), with diffusion matrix

$$\mathfrak{D}(u) = \begin{pmatrix} D_1(1 + \epsilon a_1 u_1 - \epsilon c_1 u_2) & \epsilon D_1 b_1 u_1 \\ \epsilon D_2 b_2 u_2 & D_2(1 + \epsilon a_2 u_2 - \epsilon c_2 u_1) \end{pmatrix}, \tag{1.5a}$$

and drift matrix

$$\mathfrak{F}(u) = \begin{pmatrix} -\nabla V_1 & \epsilon c_1 \nabla (V_1 - V_2) u_1 \\ \epsilon c_2 \nabla (V_2 - V_1) u_2 & -\nabla V_2 \end{pmatrix}. \tag{1.5b}$$

The parameters a_i, b_i, c_i $(i = 1, 2)$ are all positive numbers that depend on the problem dimension, particle sizes, numbers, and relative diffusion coefficients (see specific values in Sect. 3). Model (1.5) also fits into the form (1.2), with $\epsilon = 0$ when particles are non-interactive (point particles) and evolve according to two decoupled linear drift-diffusion equations.

Example 1.3 (Asymptotic gradient-flow structures)**.** Certain cross-diffusion systems possess a formal gradient-flow structure, that is, they can be formulated as

$$\partial_t u - \nabla \cdot \left(M \nabla \frac{\delta E}{\delta u} \right) = 0, \tag{1.6}$$

where $M \in \mathbb{R}^{m\times m}$ is known as mobility matrix and $\delta E/\delta u$ is the variational derivative of the entropy (or free energy) function $E[u]$. While the underlying microscopic model (1.4) of Example 1.2 has a natural entropy, in [2] it was noted that model (1.5) does not have an obvious gradient-flow structure, but that it is close to one that does have such convenient structure. More specifically, consider the following entropy

$$E_\epsilon[u] = \int_\Omega \left[u_1 \log u_1 + u_2 \log u_2 + u_1 \frac{V_1}{D_1} + u_2 \frac{V_2}{D_2} + \frac{\epsilon}{2} \left(a_1 u_1^2 + 2 a_{12} u_1 u_2 + a_2 u_2^2 \right) \right] \mathrm{d}x, \tag{1.7a}$$

with $a_{12} = (d-1)(c_1 + c_2)$, and the mobility matrix

$$M_\epsilon(u) = \begin{pmatrix} D_1 u_1 (1 - \epsilon c_1 u_2) & D_1 c_2 \epsilon u_1 u_2 \\ D_2 c_1 \epsilon u_1 u_2 & D_2 u_2 (1 - c_2 \epsilon u_1) \end{pmatrix}. \tag{1.7b}$$

The cross-diffusion system (1.1) with diffusion and drift matrices (1.5) and $N_1 = N_2$,[1] can be rewritten as

$$\partial_t u = \nabla \cdot \left(M_\epsilon \nabla \frac{\delta E_\epsilon}{\delta u} - \epsilon^2 G \right), \tag{1.8}$$

where $G = G(u, \nabla u)$ (see more details in Sect. 3). In particular, the discrepancy between the system in Example 1.2 and the gradient-flow induced by (1.7) is of order ϵ^2, an order higher than that of the model.[2] Does this legitimise the use of (1.7) as a gradient-flow structure of the system? Having a formal gradient-flow structure can facilitate the analysis of cross-diffusion models [13]. The gradient-flow model (1.6)–(1.7) was studied in [2]; stability, uniqueness of the stationary solutions, and a global-in-time existence result was shown.

It is natural to ask whether the approximation argument in Example 1.3 can be made rigorous, and, more generally if minor changes in the models can be safely ignored. For instance, given a two-species biological system, does it matter if we choose a lattice-based model (like in Example 1.1), or an off-lattice model (like in Example 1.2) with equal particle number, size, diffusivity, etc.? If so, can we quantify the differences? Lattice-based approaches have become very common, as they offer a simple way to derive continuum PDE models. They can be unrealistic since most biological transport processes modelled by these are not constrained on a lattice [18]. Nevertheless, if one is solely interested in the population-level behaviour of the system, is it worth using a more realistic off-lattice model? When is the even simpler model (linear advection-diffusion) sufficiently

[1] In [2] the more general case when $N_1 \neq N_2$ was also considered, by writing the system in terms of number densities $N_i u_i$.

[2] Systems (1.1)–(1.5) and (1.6)–(1.7) are in fact identical when both species have the same particle sizes, $\varepsilon_1 = \varepsilon_2$, and diffusivities, $D_1 = D_2$, since G vanishes in that particular case.

accurate? The aim of this paper is to answer these questions and quantify the differences between models of the form (1.1).

1.2. Outline of the results

As we are working with systems of equations, we use different indices to refer to the ambient space variables and the component or species number. Greek indices $1 \leq \alpha, \beta \leq d$ refer to directions in the ambient space, $\mathbb{R}^d$, for $d = 1, 2, 3$. Latin indices $1 \leq i, j \leq m$ are used to refer to the species number. The domain Ω where the problem is formulated is bounded, connected and of class C^2 in $\mathbb{R}^d$. The outward normal on $\partial\Omega$ is written ν.

The parabolic models we consider are weak formulations of problems of the form

$$
\begin{aligned}
&\partial_t u_i - \partial_\alpha \left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,u)\partial_\beta u_j - \mathfrak{F}_{ij}^{\alpha}(t,x,u)u_j\right] = 0, \quad \text{in} \quad \Omega,\\
&\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,u)\partial_\beta u_j - \mathfrak{F}_{ij}^{\alpha}(t,x,u)u_j\right]\cdot \nu_\alpha = 0, \quad \text{on} \quad \partial\Omega,\\
&u(0,\cdot) = u^0, \quad \text{in} \quad \Omega,
\end{aligned} \tag{1.9}
$$

for $1 \leq i \leq m$. The Einstein summation convention is used, that is, repeated indices are implicitly summed.

Our main result is a stability estimate for cross-diffusion systems that are close to diagonal, decoupled, linear diffusion problems. Our reference problem will be the weak formulation of

$$
\begin{aligned}
&\partial_t u_i - \partial_\alpha \left[D_i^{\alpha\beta}(t,x)\partial_\beta u_i - F_i^{\alpha}(t,x)u_i\right] = 0, \quad \text{in} \quad \Omega,\\
&\left[D_i^{\alpha\beta}(t,x)\partial_\beta u_i - F_i^{\alpha}(t,x)u_i\right]\cdot \nu_\alpha = 0, \quad \text{on} \quad \partial\Omega,\\
&u(0,\cdot) = u^0, \quad \text{in} \quad \Omega.
\end{aligned} \tag{1.10}
$$

The initial datum u^0 in (1.9) and (1.10) belongs to $H^2(\Omega)$. Note that throughout the paper we write $H^2(\Omega)$ for $H^2(\Omega;\mathbb{R}^m)$, and similarly for other spaces.

Compared to the general system (1.9), in (1.10) we have specified that $\mathfrak{D}_{ij} = \mathfrak{F}_{ij} = 0$ if $i \neq j$, and $\mathfrak{D}$ and $\mathfrak{F}$ do not depend on u. In Examples 1.1, 1.2, and 1.3, the reference problem corresponds to the case $\epsilon = 0$, with $D_i^{\alpha\beta}(x,t) = \delta_{\alpha\beta}D_i$ and $F_i = -\nabla V_i$. We allow time and space variations of the diffusion coefficients as it does not affect the analysis. We could also have safely included lower-order terms, but it would have resulted in somewhat longer and relatively routine developments. Additionally such terms do not appear in the three examples of interest.

System (1.10) is strongly parabolic, that is, there exist a positive constant λ such that for every $t \in [0,\infty)$, $x \in \Omega$ and $\xi \in \mathbb{R}^d$, there holds

$$
D_i^{\alpha\beta}(t,x)\xi^\alpha\xi^\beta \geq \lambda\,|\xi|^2\,, \quad i = 1,\ldots,m. \tag{1.11}
$$

Furthermore, we shall assume that D is symmetric in the space indices α and β.

We allow perturbations of system (1.10) scaled by a small parameter ϵ. Namely we consider (1.9) with

$$
\begin{aligned}
\mathfrak{D}_{ij}^{\alpha\beta}(t,x,u) &= D_i^{\alpha\beta}(t,x) + \epsilon a_{ij}^{\alpha\beta}(t,x)\phi_{ij}^{\alpha\beta}(u),\\
\mathfrak{F}_{ij}^{\alpha}(t,x,u) &= F_i^{\alpha}(t,x) + \epsilon b_{ij}^{\alpha}(t,x)\psi_{ij}^{\alpha}(u).
\end{aligned} \tag{1.12}
$$

The variations of the coefficients a and b are of class C^2 in time and space, that is,

$$
\|(a,b)\|_{C^2([0,\infty)\times\mathbb{R}^d)} \leq M, \tag{1.13}
$$

and the dependence on u of the perturbations is also of class C^2,

$$\phi, \psi \in C^2\left(\mathbb{R}^m\right)^{m\times m}, \qquad \phi(0) = \psi(0) = 0. \tag{1.14}$$

Furthermore, we assume that D and F satisfy the bound

$$\sum_{\alpha,\beta,i} \|D_i^{\alpha\beta}\|_{C^1([0,\infty)\times\mathbb{R}^d)} + \sum_{\alpha,i} \|F_i^{\alpha}\|_{C^1([0,\infty)\times\mathbb{R}^d)} \le M. \tag{1.15}$$

In the context of biological models, one is often interested in arbitrarily long behaviour and, in turn, convergence to a steady state. Along this line, we prove sharper estimates when the coefficients D and F of the reference problem (1.10) do not depend on time and F is derived from a potential (as in Examples 1.1, 1.2, and 1.3). In particular, consider the following additional assumption:

(H) for each $i \in \{1, \dots, m\}$, D_i is independent of time and there exists V_i such that $F_i = -D_i \nabla V_i$.

Our estimates will be expressed in terms of the constants appearing in assumptions (1.11), (1.13), (1.14) and (1.15). More specifically, the following positive-valued functions will appear:

$$L_i : R \to \|(\phi,\psi)\|_{C^i\left(\overline{B_R(0)}\right)} \quad i = 0, 1, 2, \tag{1.16}$$

$$K_0 : R \to M\left(5L_0\left(C_S^\infty R\right) + 2C_S^2 L_1\left(C_S^\infty R\right) R\right), \tag{1.17}$$

$$K_1 : R \to C_S M\left(L_1(R)R + L_2(R)R^2\right), \tag{1.18}$$

$$K_2 : R \to 6RC_{T/\infty} C_S \max\left(\left(L_0(R) + L_1(R)R\right), M(1+R)\right), \tag{1.19}$$

where C_S^2, C_S, and C_S^∞ depend on Ω and d and are given by (2.7), (2.8), and (A.35) respectively. The constant $C_{T/\infty}$ determines the dependence on a final time $T > 0$ of our estimates and is given by

$$C_{T/\infty} = \begin{cases} C_T & \text{when } \textbf{(H)} \text{ does not apply,} \\ C_\infty & \text{when } \textbf{(H)} \text{ applies,} \end{cases} \tag{1.20}$$

where C_T is given by (A.30) and depends on M, Ω, L_0, L_1 and T only, and C_∞ is specified in (A.31) and it depends on M, Ω, L_0 and L_1 only – not T. The upper bound ϵ_0 on the range of values ϵ allowed will be determined by means of the following function

$$\epsilon_0 : R \to \min\left(\frac{1}{2 + 2K_0(R)}, \frac{1}{1 + K_1(R)}\right). \tag{1.21}$$

Our first result, which is instrumental to our main theorem, provides an existence result and a regularity estimate for solutions of system (1.9). Given $T > 0$, we denote the parabolic cylinder by $Q_T = (0, T) \times \Omega$.

Definition 1.4. We name $W(Q_T)$ the Banach space of functions with two weak derivatives in space in $L^2(\Omega)$ continuously in time, and one time derivative in $H^1(Q_T)$, that is,

$$W\left(Q_T\right) = \left\{u \in C\left([0,T]; H^2(\Omega)\right), \partial_t u \in H^1\left(Q_T\right)\right\}.$$

We are now ready to state our first result, concerning existence and uniqueness of solutions of (1.9).

Proposition 1.5. *Assume that hypothesis (1.11), (1.12), (1.13), (1.14) and (1.15) hold. Consider $u^0 \in W(Q_T)$ satisfying the compatibility condition*

$$\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,u^0)\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,u^0)u_j^0\right]\cdot\nu = 0 \quad \text{on } \partial\Omega, \quad i=1,\dots,m. \tag{1.22}$$

Let

$$Y_0 = C_{T/\infty}\|u^0\|_{H^2(\Omega)}. \tag{1.23}$$

If $\epsilon < \epsilon_0(Y_0)$, then system (1.9) admits a unique solution $u \in W(Q_T)$ and there holds

$$\|u\|_{W(Q_T)} \le Y_0.$$

Remark 1.6. Any compatible initial data in $H^2(\Omega)$ is allowed, provided ϵ is small enough. Note that the compatibility condition (1.22) holds for any initial data compactly supported in Ω. All ϵ within the range $[0,\epsilon_0(Y_0))$ are allowed, and the solution u is bounded linearly by its initial condition. When assumption **(H)** holds, the solution is bounded for all times.

Our result holds for space dimension $d = 1, 2$ and 3, but not above. Two embeddings are used in our proofs: $L^4(\Omega) \subset H^1(\Omega)$, which does not hold when $d \ge 5$, and $L^\infty(\Omega) \subset H^2(\Omega)$, which does not hold when $d \ge 4$.

Our purpose is to establish a stability result under perturbations. Therefore we consider a second problem with $\mathfrak{D}$ and $\mathfrak{F}$ replaced by

$$\begin{aligned}\widetilde{\mathfrak{D}}_{ij}^{\alpha\beta}(t,x,u) &= D_i^{\alpha\beta}(t,x) + \epsilon\tilde{a}_{ij}^{\alpha\beta}(t,x)\tilde{\phi}_{ij}^{\alpha\beta}(u),\\ \widetilde{\mathfrak{F}}_{ij}^{\alpha}(t,x,u) &= F_i^{\alpha}(t,x) + \epsilon\tilde{b}_{ij}^{\alpha}(t,x)\tilde{\psi}_{ij}^{\alpha}(u),\end{aligned} \tag{1.24}$$

where $\tilde{a}$, $\tilde{b}$, $\tilde{\phi}$ and $\tilde{\psi}$ satisfy hypothesis (1.13), (1.14) and, without loss of generality,

$$\left\|(\tilde{\phi},\tilde{\psi})\right\|_{C^i(\overline{B_R(0)})} \le L_i(R) \quad \text{for all} \quad 0 \le R, \quad i = 0,1,2.$$

for L_i defined in (1.16). Our main result is as follows.

Theorem 1.7. *Given $u^0, \tilde{u}^0 \in H^2(\Omega)$ compactly supported in Ω, write*

$$Y_1 = C_{T/\infty}\max\left(\|u^0\|_{H^2(\Omega)}, \|\tilde{u}^0\|_{H^2(\Omega)}\right),$$

and assume $\epsilon < \epsilon_0(Y_1)$ so that Proposition 1.5 applies for both sets of parameters. Let $u \in W(Q_T)$ be the solution of (1.9) and $\tilde{u} \in W(Q_T)$ be the solution of (1.9) with $\mathfrak{D}$, $\mathfrak{F}$ and u^0 are replaced by $\widetilde{\mathfrak{D}}$ and $\widetilde{\mathfrak{F}}$ and $\tilde{u}^0$, respectively. Then the following stability estimate holds:

$$\|\tilde{u}-u\|_{W(Q_T)} \le \Gamma_1\|\tilde{u}^0 - u^0\|_{H^2(\Omega)} + \epsilon\Gamma_2\left(\|(\tilde{a},\tilde{b})-(a,b)\|_{C^1([0,\infty)\times\mathbb{R}^d)} + \|(\tilde{\phi},\tilde{\psi})-(\phi,\psi)\|_{C^1\left(\overline{B_{Y_1}(0)}\right)}\right), \tag{1.25}$$

where $\Gamma_1 = (1+K_1(Y_1))C_{T/\infty}$, $\Gamma_2 = (1+K_1(Y_1))K_2(Y_1)$ and K_1, K_2 are non decreasing functions given by (1.18), (1.19) respectively. They depend on Ω, M, λ, L_0, L_1 and L_2 and $C_{T/\infty}$ only.

Theorem 1.7 implies, for example, that we can control the differences between the solutions of the models in Examples 1.1 and 1.2, by considering the differences in their respective diffusion and drift matrices, which appear at order ϵ. Similarly, we can also use this result to predict the error we will make by approximating model (1.5) in Example 1.2 as the gradient flow in Example 1.3. Since the differences between models appear at order ϵ^2 in this case, provided the initial data are equal, the error will be bounded and of order ϵ^2 for all times (see Sect. 3).

Remark 1.8. The compatibility condition (1.22) appearing in Proposition 1.5 is automatically satisfied by compactly supported initial data as we have assumed in Theorem 1.7. However, Theorem 1.7 also holds (with the same proof) provided that u^0 and $\tilde{u}^0$ satisfy the following four conditions:

$$
\begin{aligned}
&\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,u^0)\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,u^0)u_j^0\right]\cdot\nu = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m,\\
&\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,\tilde{u}^0)\partial_\beta \tilde{u}_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,\tilde{u}^0)\tilde{u}_j^0\right]\cdot\nu = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m,\\
&\left[\tilde{\mathfrak{D}}_{ij}^{\alpha\beta}(t,x,u^0)\partial_\beta u_j^0 - \tilde{\mathfrak{F}}_{ij}^{\alpha}(t,x,u^0)u_j^0\right]\cdot\nu = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m,\\
&\left[\tilde{\mathfrak{D}}_{ij}^{\alpha\beta}(t,x,\tilde{u}^0)\partial_\beta \tilde{u}_j^0 - \tilde{\mathfrak{F}}_{ij}^{\alpha}(t,x,\tilde{u}^0)\tilde{u}_j^0\right]\cdot\nu = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m.
\end{aligned}
$$

We choose to write the result for compactly supported initial data to improve readability.

2. Proof of Proposition 1.5 and Theorem 1.7

In Lemma 2.1, we derive an estimate for a linearisation of system (1.9).

Lemma 2.1. *Assume that $\mathfrak{D}$ and $\mathfrak{F}$ are given by (1.12), and that a, b and ϕ, ψ satisfy (1.13) and (1.14) respectively. Suppose that $h \in W(Q_T)$ satisfies*

$$\epsilon K_0\left(\|h\|_{W(Q_T)}\right) < 1, \tag{2.1}$$

where K_0 is given by (1.17).

For all $u^0 \in H^2(\Omega)$ and $f \in C([0,T]; H^1(Q_T)) \cap H^1(0,T;L^2(\Omega))$ such that

$$\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j^0 + f_i^\alpha(t=0)\right]\cdot\nu = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m, \tag{2.2}$$

there exists a unique weak solution $u \in W(Q_T)$ to the linearised system

$$
\begin{aligned}
&\partial_t u_i - \partial_\alpha\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j + f_i^\alpha\right] = 0, \quad \text{in } \mathcal{D}'(\Omega),\\
&\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j + f_i^\alpha\right]\nu_\alpha = 0, \quad \text{on } \partial\Omega, \quad i=1,\ldots,m,\\
&u(0,x) = u^0, \quad \text{in } \Omega.
\end{aligned}
\tag{2.3}
$$

Furthermore, the solution map

$$S : \left(h, u^0, f\right) \to u, \text{ where } u \text{ is the solution of (2.3)}, \tag{2.4}$$

satisfies

$$\left\|S(h,u^0,f)\right\|_{W(Q_T)}\left[1-\epsilon K_0(\|h\|_{W(Q_T)})\right] \le \frac{1}{2}C_{T/\infty}\left(\|u^0\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(Q_T))\cap H^1(0,T;L^2(\Omega))}\right),$$

*where $C_{T/\infty} > 0$ is given by (1.20) and does not depend on T if **(H)** holds.*

The proof of Lemma 2.1 is in Appendix A. This first result has an immediate corollary.

Corollary 2.2. *For any u^0 and h in $W(Q_T)$, suppose that*

$$\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j^0\right]\cdot\nu = 0 \quad \text{on } \partial\Omega,$$

and

$$\epsilon \leq \frac{1}{2+2K_0\left(C_{T/\infty}\|u^0\|_{H^2(\Omega)}\right)}, \qquad \|h\|_{W(Q_T)} \leq Y_0,$$

where K_0, $C_{T/\infty}$, and Y_0 are defined in (1.17), (1.20), and (1.23) respectively. Then

$$\left\|S(h,u^0,0)\right\|_{W(Q_T)} < Y_0.$$

Proof. Since K_0 is a non decreasing function, we obtain

$$\epsilon K_0\left(\|h\|_{W(Q_T)}\right) \leq \frac{K_0\left(Y_0\right)}{2+2K_0\left(Y_0\right)} < \frac{1}{2},$$

hence (2.1) is satisfied. Applying Lemma 2.1 with $f=0$, we obtain the announced estimate. $\square$

In a second step, we establish a contraction property.

Lemma 2.3. *Given $\epsilon>0$, $u^0\in H^2(\Omega)$, and $h,\tilde{h}\in W(Q_T)$, suppose that on $\partial\Omega$*

$$\left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j^0\right]\cdot\nu = 0, \qquad \left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,\tilde{h})\partial_\beta u_j^0 - \mathfrak{F}_{ij}^{\alpha}(t,x,\tilde{h})u_j^0\right]\cdot\nu = 0.$$

Suppose also that

$$\epsilon \leq \frac{1}{2[1+K_0(Y_0)]}, \qquad \max(\|h\|_{W(Q_T)}, \|\tilde{h}\|_{W(Q_T)}) \leq Y_0,$$

where K_0, $C_{T/\infty}$, and Y_0 are defined in (1.17), (1.20), and (1.23) respectively. Then we have

$$\left\|S(h,u^0,0) - S(\tilde{h},u^0,0)\right\|_{W(Q_T)} \leq \epsilon K_1(Y_0)\|h-\tilde{h}\|_{W(Q_T)},$$

with K_1 given by (1.18).

Proof. Write $u=S(h,u^0,0)$ and $\tilde{u}=S(\tilde{h},u^0,0)$. We have

$$u-\tilde{u} = \epsilon S\left(h,0,g\right),$$

where

$$g_i^\alpha = a_{ij}^{\alpha\beta}(t,x)\left[\phi_{ij}^{\alpha\beta}(h)-\phi_{ij}^{\alpha\beta}(\tilde{h})\right]\partial_\beta\tilde{u}_j + b_{ij}^{\alpha}(t,x)\left[\psi_{ij}^{\alpha\beta}(h)-\psi_{ij}^{\alpha\beta}(\tilde{h})\right]\tilde{u}_j. \tag{2.5}$$

Noting that

$$\left|\phi_{ij}^{\alpha\beta}(h)-\phi_{ij}^{\alpha\beta}(\tilde{h})\right| \leq L_1(Y_0)\left|h-\tilde{h}\right|,$$

we find

$$\max_{[0,T]} \|g\|_{L^2(\Omega)} \leq ML_1(Y_0)\|h-\tilde{h}\|_{W(Q_T)} \|\tilde{u}\|_{W(Q_T)} \leq ML_1(Y_0)M_0\|h-\tilde{h}\|_{W(Q_T)}.$$

Similarly, we can estimate the gradient as follows

$$\begin{aligned}|\nabla g| \leq M \Big(L_1(Y_0)\big|h-\tilde{h}\big| + L_2(Y_0)\big|h-\tilde{h}\big| \, |\nabla h| + L_1(Y_0)\big|\nabla h - \nabla\tilde{h}\big| \Big) \left(|\nabla\tilde{u}| + |\tilde{u}|\right) \\ +ML_1(Y_0)\big|h-\tilde{h}\big| \left(\big|\nabla^2\tilde{u}\big| + |\nabla\tilde{u}|\right).\end{aligned}$$

Therefore

$$\begin{aligned}\|\nabla g\|_{L^2(\Omega)} \leq ML_1(Y_0) \Big[2\|h-\tilde{h}\|_{L^\infty(Q_T)}\|\tilde{u}\|_{H^2(\Omega)} + \|h-\tilde{h}\|_{L^4(\Omega)} \left(\|\nabla\tilde{u}\|_{L^4(\Omega)} + \|\tilde{u}\|_{L^4(\Omega)}\right)\Big] \\ +ML_2(Y_0)\|h-\tilde{h}\|_{L^\infty(Q_T)} \|\nabla h\|_{L^4(\Omega)} \Big(\|\nabla\tilde{u}\|_{L^4(\Omega)} + \|\tilde{u}\|_{L^4(\Omega)}\Big).\end{aligned}$$

Thanks to the Ladyzhenskaya (or Gagliardo–Nirenberg) inequality, we obtain

$$\max_{[0,T]} \|\nabla g\|_{L^2(\Omega)} \leq C_S M \left(L_1(Y_0)Y_0 + L_2(Y_0)Y_0^2\right) \|h-\tilde{h}\|_{W(Q_T)},$$

where C_S^1 is a product of Sobolev embedding constants, depending on Ω and d, namely

$$C_S^1 = \max\Big(1, C\left(H^2(\Omega) \hookrightarrow L^\infty(\Omega)\right)^3, C\left(H^2(\Omega) \hookrightarrow W^{1,4}(\Omega)\right)^3\Big). \tag{2.6}$$

We now turn to the time derivative

$$\begin{aligned}|\partial_t g| \leq M \Big(L_1(Y_0)\big|h-\tilde{h}\big| + L_2(Y_0)\big|h-\tilde{h}\big| \, |\partial_t h| + L_1(Y_0)\big|\partial_t h - \partial_t\tilde{h}\big| \Big) \left(|\nabla\tilde{u}| + |\tilde{u}|\right) \\ +ML_1(Y_0)\big|h-\tilde{h}\big| \left(|\nabla\partial_t\tilde{u}| + |\partial_t\tilde{u}|\right).\end{aligned}$$

Thus, using that $\partial_t h, \partial_t\tilde{h} \in L^4(Q_T)$ and $\partial_t\nabla\tilde{u} \in L^2(Q_T)$, we have

$$\|\partial_t g\|_{L^2(Q_T)} \leq C_S^2 M \left(L_1(Y_0)Y_0 + L_2(Y_0)Y_0^2\right) \left\|h-\tilde{h}\right\|_{W(Q_T)},$$

where C_S^2 is also a product of Sobolev embedding constants, depending on Ω and d, namely

$$C_S^2 = \max\Big(C\left(H^1(\Omega) \hookrightarrow L^4(\Omega)\right)^2, 1\Big). \tag{2.7}$$

Finally, we apply Lemma 2.1 to obtain

$$\|u-\tilde{u}\|_{W(Q_T)} \leq \epsilon Y_0 C_S M \left[L_1(Y_0)Y_0 + L_2(Y_0)Y_0^2\right] \|h-\tilde{h}\|_{W(Q_T)},$$

with

$$C_S = C_S^1 + C_S^2. \tag{2.8}$$

□

We now turn to the proof of Proposition 1.5.

Proof of Proposition 1.5. Recall that

$$\epsilon_0 : R \to \min\left(\frac{1}{2+2K_0(R)}, \frac{1}{1+K_1(R)}\right).$$

where K_0 and K_1 are defined in (1.17) and (1.18), respectively.

Given $u^0 \in W(Q_T)$ we introduce the sequence v_n given by $v_0 = u^0$ and, for all $n \geq 0$,

$$v_{n+1} = S(v_n, u^0, 0),$$

where S is the solution map defined in (2.4). Note that the compatibility condition (1.22) is satisfied at every step. Corollary 2.2 shows that $\|v_n\|_{W(Q_T)} \leq Y_0$ for each n. Furthermore, thanks to Lemma 2.3,

$$\|v_{n+2} - v_{n+1}\|_{W(Q_T)} \leq \epsilon_0 K_1(Y_0)\|v_{n+1} - v_n\|_{W(Q_T)} \leq \frac{K_1(Y_0)}{1+K_1(Y_0)}\|v_{n+1} - v_n\|_{W(Q_T)}.$$

The sequence thus converges to a solution of (1.9), thanks to the contraction mapping theorem. □

We now turn to the proof of the perturbation result in Theorem 1.7. Consider the linearised system given by

$$\begin{aligned}
&\partial_t \tilde{u}_i - \partial_\alpha\left[\widetilde{\mathfrak{D}}_{ij}^{\alpha\beta}(t,x,\tilde{h})\partial_\beta \tilde{u}_j - \widetilde{\mathfrak{F}}_{ij}^{\alpha}(t,x,\tilde{h})\tilde{u}_j\right] = \tilde{f}_i, \quad \text{in} \quad \mathcal{D}'(\Omega),\\
&\left[\widetilde{\mathfrak{D}}_{ij}^{\alpha\beta}(t,x,\tilde{h})\partial_\beta \tilde{u}_j - \widetilde{\mathfrak{F}}_{ij}^{\alpha}(t,x,\tilde{h})\tilde{u}_j\right]\nu_\alpha = 0, \quad \text{on} \quad \partial\Omega,\\
&\tilde{u}(0,\cdot) = \tilde{u}^0, \quad \text{in} \quad \Omega,
\end{aligned} \tag{2.9}$$

Following the notation of Lemma 2.1 (see (2.4)), the solution operator associated to (2.9) is denoted by $\tilde{S}(\tilde{h}, \tilde{u}^0, \tilde{f})$.

Proposition 2.4. *Let $h, \tilde{h} \in W(Q_T)$ be compactly supported in Ω for $t = 0$ and write*

$$Y_1 = C_{T/\infty}\max\left(\|\tilde{h}\|_{W(Q_T)}, \|h\|_{W(Q_T)}\right).$$

Assume $\epsilon < \epsilon_0(Y_1)$, so that the solution operators S and $\tilde{S}$ corresponding to (2.3) and (2.9) respectively are well defined. For any $u^0, \tilde{u}^0 \in H^2(\Omega)$ with compact support in Ω there holds

$$\begin{aligned}
\left\|\tilde{S}(\tilde{h}, \tilde{u}^0, 0) - S(h, u^0, 0)\right\|_{W(Q_T)} &\leq C_{T/\infty}\|\tilde{u}^0 - u^0\|_{H^2(\Omega)} + \epsilon K_1(Y_1)\|\tilde{h} - h\|_{W(Q_T)}\\
&\quad + \epsilon K_2(Y_1)\left(\|(\tilde{a},\tilde{b}) - (a,b)\|_{C^1([0,\infty)\times\mathbb{R}^d)} + \|(\tilde{\phi},\tilde{\psi}) - (\phi,\psi)\|_{C^1\left(\overline{B_{Y_1}(0)}\right)}\right).
\end{aligned}$$

where K_2 depends on L_0, L_1, Ω, M and $C_{T/\infty}$ and is given by (1.19).

Proof. We write

$$\tilde{S}(\tilde{h}, \tilde{u}^0, 0) - S(h, u^0, 0) = \tilde{S}(\tilde{h}, \tilde{u}^0, 0) - S(\tilde{h}, \tilde{u}^0, 0) + S(\tilde{h}, \tilde{u}^0, 0) - S(h, \tilde{u}^0, 0) + S(h, \tilde{u}^0, 0) - S(h, u^0, 0).$$

Thanks to Lemma 2.1 and to the linearity of S with respect to the initial data, we have

$$\left\|S\left(h,\tilde{u}^0,0\right)-S\left(h,u^0,0\right)\right\|_{W(Q_T)}\leq C_{T/\infty}\|u^0-\tilde{u}^0\|_{H^2(\Omega)}.$$

Note that the compatibility condition is satisfied due to the compact support of u^0, $\tilde{u}^0$ and $h(t=0)$ in Ω. On the other hand, Lemma 2.3 shows that

$$\left\|S(\tilde{h},\tilde{u}^0,0)-S(h,\tilde{u}^0,0)\right\|_{W(Q_T)}\leq\epsilon K_1(Y_1)\|\tilde{h}-h\|_{W(Q_T)}.$$

We write

$$\tilde{S}(\tilde{h},\tilde{u}^0,0)-S(\tilde{h},\tilde{u}^0,0)=\epsilon\tilde{S}(\tilde{h},0,\tilde{g}),$$

where $\tilde{g}$ is given by

$$\begin{aligned}\tilde{g}_i^\alpha&=\epsilon^{-1}\left[\left(\tilde{\mathfrak{D}}_{ij}^{\alpha\beta}-\mathfrak{D}_{ij}^{\alpha\beta}\right)(t,x,\tilde{h})\partial_\beta\tilde{u}_j-\left(\tilde{\mathfrak{F}}_{ij}^{\alpha}-\mathfrak{F}_{ij}^{\alpha}\right)(t,x,\tilde{h})\tilde{u}^j\right]\\&=\left(\tilde{a}_{ij}^{\alpha\beta}\tilde{\phi}_{ij}^{\alpha\beta}-a_{ij}^{\alpha\beta}\phi_{ij}^{\alpha\beta}\right)(t,x,\tilde{h})\partial_\beta\tilde{u}_j+\left(\tilde{b}_i^{\alpha\beta}\tilde{\psi}_{ij}^{\alpha}-b_{ij}^{\alpha}\psi_{ij}^{\alpha}\right)(t,x,\tilde{h})\tilde{u}^j,\end{aligned}$$

and $\tilde{u}=S(\tilde{h},\tilde{u}^0,0)$. In other words, $\tilde{g}$ is of the form

$$\tilde{g}=\left[(\tilde{a}-a)\tilde{\phi}+a(\tilde{\phi}-\phi)\right]\nabla\tilde{u}+\left[(\tilde{b}-b)\tilde{\psi}+b(\tilde{\psi}-\psi)\right]\tilde{u},$$

and thus we are in a setting similar to that of the proof of Lemma 2.3. In particular we have

$$\begin{aligned}(|\nabla g|+|g|)\leq&\left(\left\|(\tilde{a},\tilde{b})-(a,b)\right\|_{C^1([0,\infty)\times\mathbb{R}^d)}L_0(Y_1)+M\max_{B_{Y_1}(0)}\left|(\tilde{\phi},\tilde{\psi})-(\phi,\psi)\right|\right)(|\nabla^2\tilde{u}|+2|\nabla\tilde{u}|+|\tilde{u}|)\\&+\left(\left|(\tilde{a},\tilde{b})-(a,b)\right|L_1(Y_1)+M\max_{B_{Y_1}(0)}\left|(\tilde{D\phi},\tilde{D\psi})-(D\phi,D\psi)\right|\right)|\nabla\tilde{h}|(|\nabla\tilde{u}|+|\tilde{u}|).\end{aligned}$$

As in the proof of Lemma 2.3, using Gagliardo–Nirenberg's inequality to bound the last term, we find

$$\begin{aligned}\frac{1}{C_S}\max_{[0,T]}\|\tilde{g}\|_{H^1(\Omega)}\leq&\left\|(\tilde{a},\tilde{b})-(a,b)\right\|_{C^1([0,\infty)\times\mathbb{R}^d)}\left[2L_0(Y_1)Y_1+L_1(Y_1)Y_1^2\right]\\&+M\left\|(\tilde{\phi},\tilde{\psi})-(\phi,\psi)\right\|_{C^1(\overline{B_{Y_1}(0)})}(2Y_1+Y_1^2),\end{aligned}$$

where C_S is given by (2.8). Finally, we bound $\partial_t g$ to show that $\tilde{g}\in C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))$ in the same way, namely

$$\begin{aligned}\frac{1}{C_S}\|\partial_t\tilde{g}\|_{L^2(Q_T)}\leq&\left\|(\tilde{a},\tilde{b})-(a,b)\right\|_{C^1([0,\infty)\times\mathbb{R}^d)}\left[L_0(Y_1)Y_1+L_1(Y_1)Y_1^2\right]\\&+M\left\|(\tilde{\phi},\tilde{\psi})-(\phi,\psi)\right\|_{C^1(\overline{B_{Y_1}(0)})}(Y_1+Y_1^2).\end{aligned}$$

Because of the compact support of u^0, $\tilde{u}^0$, $\tilde{h}(t=0)$ and $h(t=0)$ in Ω, we can conclude thanks to Lemma 2.1 that

$$\left\|\tilde{S}(\tilde{h},\tilde{u}^0,0)-S(\tilde{h},\tilde{u}^0,0)\right\|_{W(Q_T)}\leq\epsilon K_2(Y_1)\left(\left\|(\tilde{a},\tilde{b})-(a,b)\right\|_{C^1([0,\infty)\times\mathbb{R}^d)}+\left\|(\tilde{\phi},\tilde{\psi})-(\phi,\psi)\right\|_{C^1(\overline{B_{Y_1}(0)})}\right).$$

$\square$

Proof of Theorem 1.7. As in the proof of Proposition 1.5, the sequences $v_{n+1} = S\left(v_n, u^0, 0\right)$ and $\tilde{v}_{n+1} = \tilde{S}\left(\tilde{v}_n, \tilde{u}^0, 0\right)$ for all $n \geq 1$, with $v_0 = u^0$ and $\tilde{v}_0 = \tilde{u}^0$, converge to u and $\tilde{u}$, respectively as $n \to \infty$. Thanks to Proposition 2.4 we have

$$\begin{aligned}\|\tilde{v}_{n+1} - v_{n+1}\|_{W(Q_T)} \leq C_{T/\infty}\|\tilde{u}^0 - u^0\|_{H^2(\Omega)} + \epsilon K_1(Y_1)\|\tilde{v}_n - v_n\|_{W(Q_T)} \\ + \epsilon K_2(Y_1)\left(\|(\tilde{a}, \tilde{b}) - (a, b)\|_{C^1([0,\infty)\times\mathbb{R}^d)} + \|(\tilde{\phi}, \tilde{\psi}) - (\phi, \psi)\|_{C^1(\overline{B_{Y_1}(0)})}\right).\end{aligned}$$

Passing to the limit as $n \to \infty$, we obtain

$$\begin{aligned}\|\tilde{u} - u\|_{W(Q_T)} \leq [1 + K_1(Y_1)]C_{T/\infty}\|\tilde{u}^0 - u^0\|_{H^2(\Omega)} \\ + \epsilon[1 + K_1(Y_1)]K_2(Y_1)\left(\|(\tilde{a}, \tilde{b}) - (a, b)\|_{C^1([0,\infty)\times\mathbb{R}^d)} + \|(\tilde{\phi}, \tilde{\psi}) - (\phi, \psi)\|_{C^1(\overline{B_{Y_1}(0)})}\right),\end{aligned}$$

as required. ◻

3. Numerical simulations

In this section, we present numerical simulations for the cross-diffusion systems described as Examples 1.1,1.2 and 1.3 in the introduction. We consider these examples when the physical dimension is $d = 2$, but with initial data and potentials V_i varying in one direction such that the solutions of (1.1) can be represented as one-dimensional. We solve (1.1) in the domain $\Omega = (-1/2, 1/2)$ using a second-order accurate finite-difference scheme in space and the method of lines with the inbuilt Matlab ode solver `ode15s` in time. We use an equidistant mesh of size $|\Omega|/J$, with nodes $x_n = -1 + n\Delta x$, $0 \leq n \leq J$. The fluxes are evaluated at the nodes x_n to ensure the no-flux conditions are imposed accurately, while the solutions u_i are computed at the midpoints $x_{n+1/2}$. The unknowns are $u_{i,n}(t) \approx u_i(x_n, t)$, $i = 1, 2$. The discretisation of the spatial derivatives is done in the spirit of the positivity-preserving scheme proposed in [24]. For example, the terms of the form $u_i \nabla u_j$ are discretised as

$$\left(u_i \frac{\partial u_j}{\partial x}\right)(x_{n+1/2}) \approx \left(\frac{2u_{i,n+1}u_{i,n}}{u_{i,n+1} + u_{i,n}}\right)\left(\frac{u_{j,n+1} - u_{j,n}}{\Delta x}\right).$$

We begin with a simulation of the model in Example 1.2 for fixed ϵ to demonstrate a typical evolution of a cross-diffusion system. The value of the parameters used in the numerical implementation are given below. Recall that the model describes two species of hard sphere particles in $\mathbb{R}^d$, $d = 2, 3$, possibly with different numbers N_i, diffusions D_i, and diameters ε_i. The coefficients in (1.5) are given by

$$a_i = \frac{2\pi}{d}(d-1)\bar{N}_i\bar{\varepsilon}_i^d, \quad b_i = \frac{2\pi}{d}\frac{[(d-1)D_i + dD_j]}{D_i + D_j}\bar{N}_i, \quad c_i = \frac{2\pi}{d}\frac{D_i}{D_i + D_j}\bar{N}_j, \tag{3.1}$$

for $i, j = 1, 2$ $(j \neq i)$, where $\bar{N}_i = N_i/(N_1 + N_2)$, $\varepsilon = (\varepsilon_1 + \varepsilon_2)/2$, $\bar{\varepsilon}_i = \varepsilon_i/\varepsilon$. In particular $\bar{N}_1 + \bar{N}_2 = 1$ and $\bar{\varepsilon}_1 + \bar{\varepsilon}_2 = 2$. The small parameter ϵ is then defined as

$$\epsilon = (N_1 + N_2)\varepsilon^d/|\Omega|. \tag{3.2}$$

For our first example, we choose $N_1 = N_2 = 100$, $D_1 = D_2 = 1$, $\varepsilon_1 = \varepsilon_2 = 0.0354$, $d = 2$. This gives the value $\epsilon = 0.25$. We set initial data $u_{1,0} = C\exp(-80(x + 0.2)^2)$, where C is the normalisation constant, and $u_{2,0} = 1$, and external potentials $V_1(x) = 1 - \exp(-120x^2)$ and $V_2(x) = 0$. We run the time-dependent simulation until $T = 1$ and plot the results in Figure 1. We observe the evolution of u_1 towards a non-trivial steady state, governed by V_1, while u_2 diffuses away from the centre (despite having no external potential) due to the cross-species interaction.

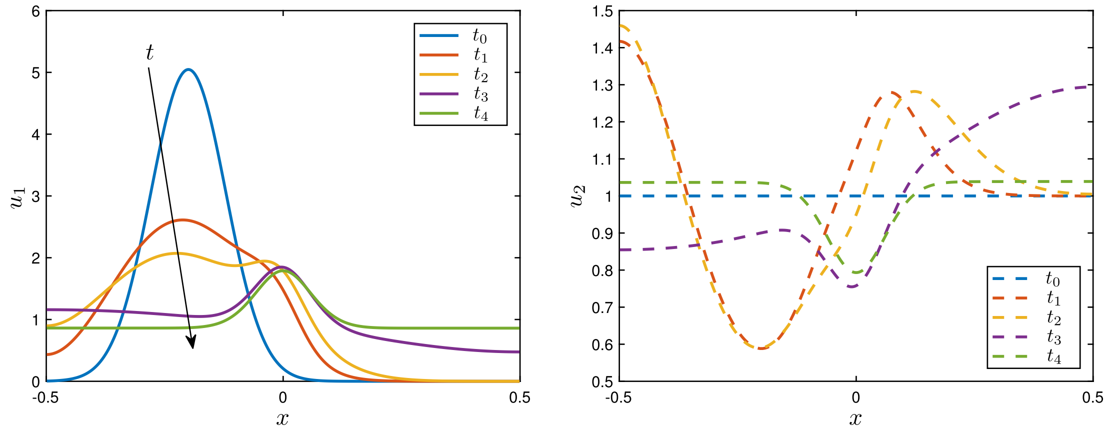

FIGURE 1. Time-dependent simulation of the model (1.5) in Example 1.2. Time-evolution of the population densities u_1 (*left*) and u_2 (*right*) with initial data $u_1^0 = C\exp(-80(x+0.2)^2)$, where C is the normalisation constant, and $u_2^0 = 1$, and final time $T = 1$ (times shown $t_0 = 0, t_1 = 0.005, t_2 = 0.01, t_3 = 0.1, t_4 = T$). The external potentials are $V_1(x) = 1 - \exp(-120x^2)$ and $V_2(x) = 0$ and volume fraction parameter $\epsilon = 0.25$. The other parameters are: $\bar{N}_i = 1/2$, $\bar{\varepsilon}_i = 1$, $d = 2$, $D_i = 1$, $J = 500$.

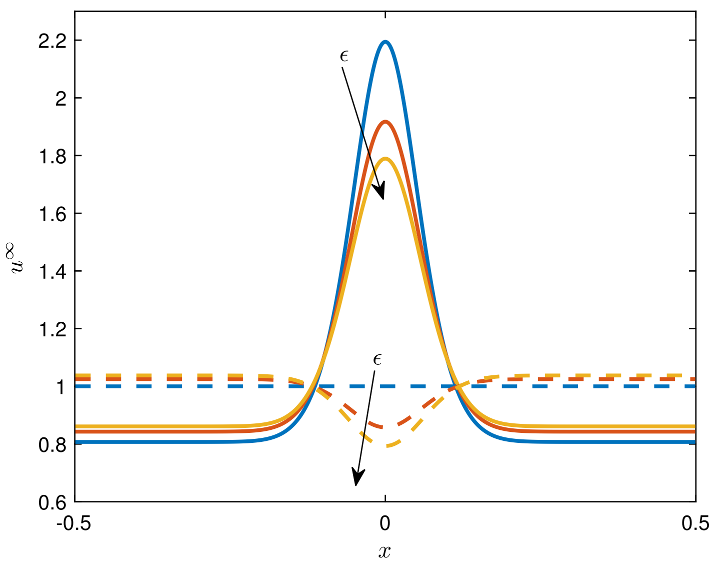

FIGURE 2. Steady state solutions u_1^∞ (*solid lines*) and u_2^∞ (*dashed lines*) of the model (1.5) in Example 1.2 for different values of ϵ, $\epsilon = 0, 0.125, 0.25$ (arrows show the direction of increasing ϵ). The other parameters are given in Figure 1.

To show the dependence of the solutions of (1.5) with the small parameter ϵ, in Figure 2 we plot the steady state solution u^∞ for three values of the occupied volume ϵ, namely $\epsilon = 0, 0.125, 0.25$. This is obtained by running the time-dependent solver for long times; we found $T = 20$ to be sufficient. Convergence to a unique steady state is guaranteed by the results in [2] and our time-independent estimates. We observe the effects of ϵ: for $\epsilon = 0$ (no interactions), $u_2 = 1$ is already the steady state solution. As we increase ϵ, the maximum of u_1^∞ decreases, as not so many particles can fit where the potential is minimised, and a minimum in u_2^∞ appears where u_1^∞ has its maximum, showing that particles from species 2 are pushed out driven by gradients in u_1.

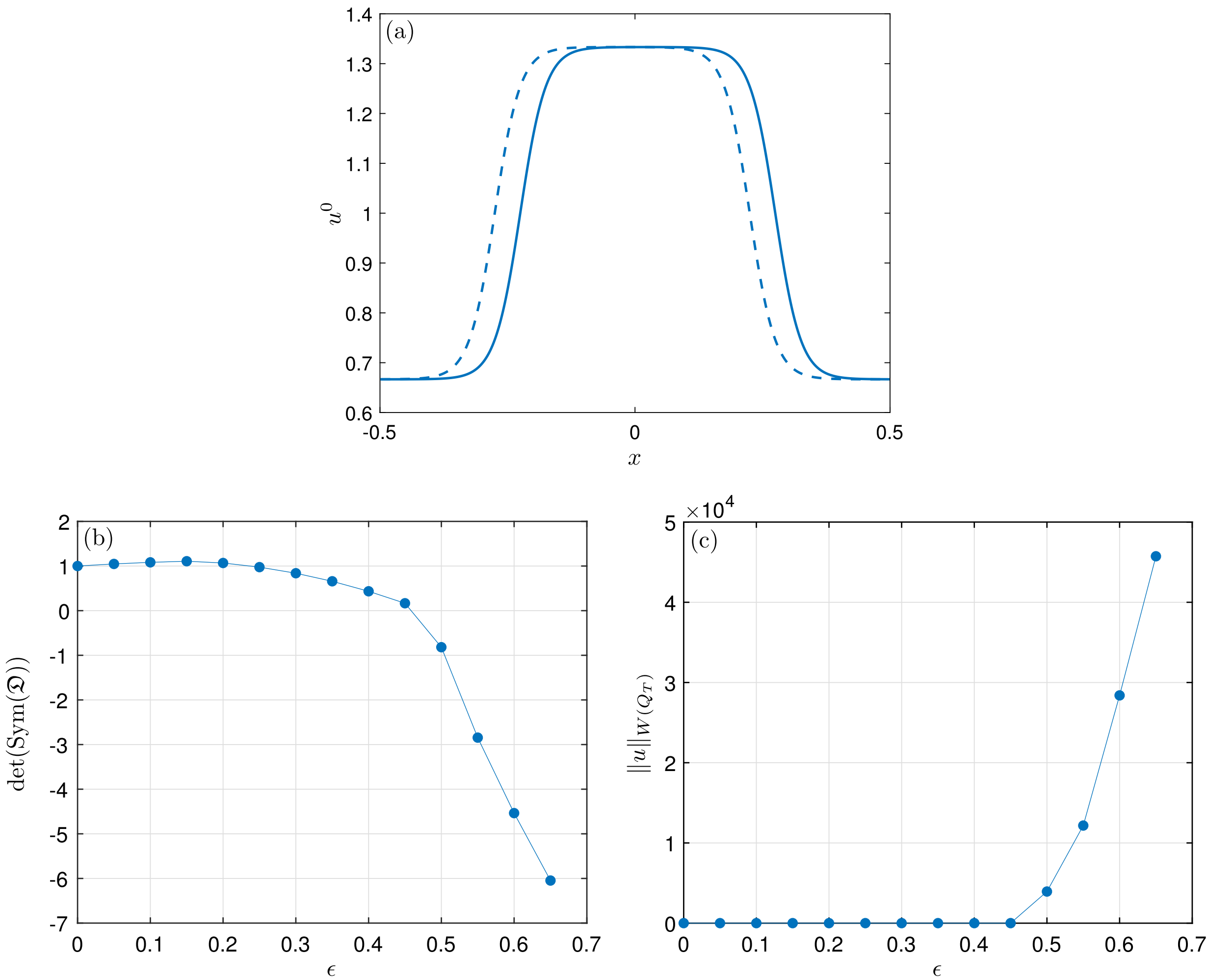

FIGURE 3. Simulation of model (1.5) for increasing values of ϵ. (a) Initial data: $u_1^0 = 1 + 0.5\tanh[10(2x + a - b)] + 0.5\tanh[-10(2x - a - b)]$ (*solid line*) and $u_2^0 = 1 + 0.5\tanh[10(2x + a + b)] + 0.5\tanh[-10(2x - a + b)]$ (*dashed line*) with $a = 0.5$ and $b = 0.05$. (b) Determinant of the symmetrised diffusion matrix (1.5a) as a function of ϵ. (c) Norm $\|u\|_{W(Q_T)}$ computed using (3.3) as a function of ϵ. Parameters used: $\bar{N}_i = 1/2$, $\bar{\varepsilon}_i = D_i = 1$, $V_i = 0$, $d = 2$, and $T = 0.1$, $J = 500$ and $M = 100$.

In the next simulation, we want to test the behaviour of the system in Example 1.2 as the perturbation in ϵ increases. To make the calculation of the bounds simpler, we assume that $\bar{\varepsilon}_i = 1$, $\bar{N}_i = 1/2$, $d = 2$, and that the two components of the solution coincide at at least one point, that is, $u_1 = u_2 = u^*$ for some $u^* > 0$. We choose the initial data shown in Figure 3(a) and $V_i = 0$, so that $u^* = \max_x u^0 \approx 1.333$. We have already introduced a bound ϵ_0 in (1.21), ensuring that the existence result in Theorem 1.7 holds. The expression of ϵ_0 is found in the proof of Lemma 2.1 for a general system, but it can be improved for the specific system at hand. However, in this section we will use another bound, which we denote by ϵ^*, that ensures ellipticity of the diffusion matrix (1.5a). This is in fact the practical bound required to obtain meaningful numerical results, and it is in general less restrictive than ϵ_0.

Lemma 3.1 (Ellipticity bound). *The following condition is necessary to ensure coercivity of the diffusive term. Suppose that the solution of (1.1) with matrices (1.5) satisfies* $\max_{Q_T}|u| = u^* > 0$, $d = 2$, *and one of the following cases apply:*

(i) Different diffusivities: $\bar{\varepsilon}_i = 1$, $\bar{N}_i = 1/2$, *and* $\theta = (D_1 - D_2)^2/4D_1D_2 \geq 0$.
(ii) Different particle sizes: $D_i = 1$, $\bar{N}_i = 1/2$, $\bar{\varepsilon}_2 = 2 - \bar{\varepsilon}_1$, *and* $\theta = 1 - 2\bar{\varepsilon}_1 + \bar{\varepsilon}_1^2 \geq 0$.
(iii) Different particle numbers: $D_i = 1$, $\bar{\varepsilon}_i = 1$, $\bar{N}_2 = 1 - \bar{N}_1$, *and* $\theta = 9(1/4 - \bar{N}_1 + \bar{N}_1^2) \geq 0$.

Then the symmetrised version of the diffusion matrix (1.5a) is non-degenerate provided that

$$\epsilon \leq \epsilon^* = \frac{1+\sqrt{9+4\theta}}{2+\theta}(\pi u^*)^{-1},$$

where θ *takes the values specified above. The bound is sharp in the case that both components* u_1 *and* u_2 *attain* u^* *at the same point.*

Proof. Recall that the diffusion matrix of Example 1.2 is

$$\mathfrak{D}(u) = \begin{pmatrix} D_1(1+\epsilon a_1 u_1 - \epsilon c_1 u_2) & \epsilon D_1 b_1 u_1 \\ \epsilon D_2 b_2 u_2 & D_2(1+\epsilon a_2 u_2 - \epsilon c_2 u_1) \end{pmatrix}.$$

From the numerical point of view, a realistic bound can be obtained imposing that the symmetrised diffusion matrix does not degenerate. We consider the case (i), that is, $\bar{\varepsilon}_i = 1$, $\bar{N}_i = 1/2$. Suppose that both components u_1, u_2 attain the same maximum at the same point, $u_1 = u_2 = u^*$. We have

$$\det(\mathrm{Sym}(\mathfrak{D})) = \det(\mathfrak{D}) - \left(\frac{\mathfrak{D}_{12} - \mathfrak{D}_{21}}{2}\right)^2 = D_1 D_2\left[1 + \frac{1}{2}\epsilon\pi u^* - \frac{1}{4}(\epsilon\pi u^*)^2(2+\theta)\right],$$

where $\theta = (D_1 - D_2)^2/(4D_1D_2) \geq 0$. Imposing that $\det(\mathrm{Sym}(D)) = 0$ leads to

$$\epsilon\pi u^* = \frac{1+\sqrt{9+4\theta}}{2+\theta},$$

as required. The other cases, as well as the non-sharp cases when, for instance, $u_1 < u_2 = u^*$, follow in a similar way. □

To test the upper bounds on ϵ, in the next example we run a simulation of model (1.5) for increasing values of ϵ. We expect the norm $\|u\|_{W(Q_T)}$ to increase suddenly for values $\epsilon > \epsilon^*$. In the example we consider, $\epsilon^* = 2/(\pi u^*) \approx 0.4776$, and $\epsilon^* = 2\tilde{\epsilon}_0 \gg \epsilon_0 \approx 2.57 \times 10^{-5}$. In the simulations, we approximate the norm in $W_2(Q_T)$ as follows. Let $u_i(n,k)$ denote the finite-difference approximation of $u_i(x_n, t_k)$, where x_n and t_k are J and M equally spaced nodes in $\Omega = [-1/2, 1/2]$ and $[0,T]$ respectively, $x_n = -1/2 + n\Delta x$, $\Delta x = 1/J$ and $t_k = 0 + k\Delta t$, $\Delta t = T/M$. Then

$$\begin{aligned} \|u\|_{W(Q_T)} \approx & \sqrt{\Delta x \Delta t \sum_{n,k}[u_{1xx}^2(n,k) + u_{2xx}^2(n,k) + u_{1t}^2(n,k) + u_{2t}^2(n,k)]} \\ & + \max_k \sqrt{\Delta x \sum_n [u^2(n,k) + u_2^2(n,k) + u_{1x}^2(n,k) + u_{2x}^2(n,k)]}, \end{aligned} \tag{3.3}$$

where $u_{ixx}(n,k) = [u_i(n+1,k) + u_i(n-1,k) - 2u_i(n,k)]/\Delta x^2$, $u_{ix}(n,k) = [u_i(n+1,k) - u_i(n-1,k)]/(2\Delta x)$ and $u_{it}(n,k) = [u_i(n,t+1) - u_i(n,t)]/\Delta t$. We choose initial data u^0 such that the two components attain the

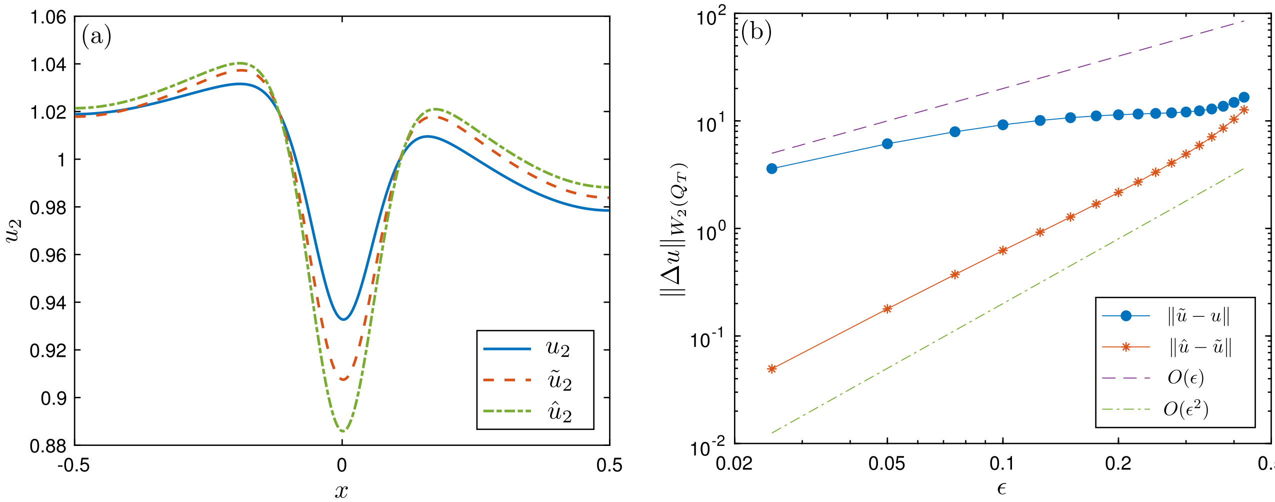

FIGURE 4. Comparison between the models in Examples 1.1, 1.2, and 1.3 for increasing values of ϵ. (a) Second component u_2 at time $t = 0.1$ for $\epsilon = 0.25$ from model (1.3) (u_2), model (1.5) ($\tilde{u}_2$), and model (1.6) ($\hat{u}_2$). (b) Norm in $W_2(Q_T)$ of the difference between solutions of models in Examples 1.1 and 1.2, $\|\tilde{u} - u\|$, and between models in Examples 1.3 and 1.1, $\|\hat{u} - \tilde{u}\|$. Norm computed using (3.3) as a function of ϵ. Dash and dot-dash lines show curves $O(\epsilon)$ and $O(\epsilon^2)$ for reference. Parameters used: $\bar{N}_i = 1/2$, $\bar{\varepsilon}_i = 1$, $D_1 = 1.5$, $D_2 = 1$, $V_1(x) = 1 - \exp(-120x^2)$ and $V_2(x) = 0$, $d = 2$, final time $T = 1$, $J = 500$ and $M = 100$. Initial data as in Figure 3(a).

same maximum u^* in regions that overlap (see Fig. 3a), and zero external potentials V_i so that we can ensure that the maximum of u^0 is also the global maximum. We consider the symmetric case when diffusivities, particle numbers and sizes are equal, $\bar{\varepsilon}_1 = \bar{\varepsilon}_2 = 1$, $\bar{N}_1 = \bar{N}_2 = 1/2$, $D_1 = D_2$, so that $\theta \equiv 0$ and $\epsilon^* = 2/(\pi u^*) =$ from Lemma 3.1. We observe that the norm $\|u\|_{W(Q_T)}$ blows up as expected for $\epsilon \geq 0.5$, when the determinant of the symmetrised diffusion matrix is negative.

Our second set of simulations relates to stability under perturbations of the matrices $\mathfrak{D}$ and $\mathfrak{F}$. We compare the solutions of Example 1.1 and Example 1.2, and the solution of Example 1.2 and the gradient-flow solution of Example 1.3. In the first case, the perturbation or differences between the models are at order ϵ, whereas in the second case the differences are at order ϵ^2. We would like to test the theoretical predictions of our analysis, namely, that we can control the difference between the solutions of the models in Examples 1.1, 1.2 and 1.3 by the difference in their diffusion and drift matrices.

We denote by $\mathfrak{D}$ and $\mathfrak{F}$ the matrices of Example 1.1, and by $\widetilde{\mathfrak{D}}$ and $\widetilde{\mathfrak{F}}$ those of Example 1.2. The difference between the models is

$$\widetilde{\mathfrak{D}} - \mathfrak{D} = \epsilon \begin{pmatrix} D_1[a_1 u_1 + u_2(\bar{N}_2 - c_1)] & D_1 u_1 (b_1 - \bar{N}_2) \\ D_2 u_2 (b_2 - \bar{N}_1) & D_2[a_2 u_2 + u_1(\bar{N}_1 - c_2)] \end{pmatrix}, \tag{3.4a}$$

and

$$\begin{aligned} \widetilde{\mathfrak{F}} - \mathfrak{F} &= \epsilon \begin{pmatrix} -u_1 \nabla V_1 \bar{N}_1 & u_1[(c_1 - \bar{N}_2)\nabla V_1 - c_1 \nabla V_2] \\ u_2[(c_2 - \bar{N}_1)\nabla V_2 - c_2 \nabla V_1] & -u_2 \nabla V_2 \bar{N}_2 \end{pmatrix} \\ &= \epsilon \begin{pmatrix} -\nabla V_1 \bar{N}_1 & [(c_1 - \bar{N}_2)\nabla V_1 - c_1 \nabla V_2] \\ [(c_2 - \bar{N}_1)\nabla V_2 - c_2 \nabla V_1] & -\nabla V_2 \bar{N}_2 \end{pmatrix} \circ \begin{pmatrix} u_1 & u_1 \\ u_2 & u_2 \end{pmatrix}, \end{aligned} \tag{3.4b}$$

In the second line, we rewrite the difference as two matrices, one dependent on x and the other on u (as required in our analysis), where $\circ$ denotes the Hadamard or entry-wise product of matrices.

The difference between the model (1.5) in Example 1.2 and the gradient-flow model (1.6) in Example 1.3 is the order ϵ^2 term G (see (1.8)), given by

$$G = (\theta_1 \nabla u_1 - \theta_2 \nabla u_2) u_1 u_2 \begin{pmatrix} -D_1 \\ D_2 \end{pmatrix}, \tag{3.5}$$

where $\theta_1 = a_1 c_1 - a_{12} c_2$, $\theta_2 = a_2 c_2 - a_{12} c_1$, and $a_{12} = (d-1)(c_1 + c_2)$. Therefore, both models have the same drift matrices and their difference is contained in their respective diffusion matrices. If we denote by $\hat{\mathfrak{D}}$ the diffusion matrix of model (1.6), then $(\hat{\mathfrak{D}} - \widetilde{\mathfrak{D}}) \nabla u = \epsilon^2 G$ (see (1.8)), that is,

$$\hat{\mathfrak{D}} - \widetilde{\mathfrak{D}} = \epsilon^2 u_1 u_2 \begin{pmatrix} -D_1 \theta_1 & D_1 \theta_2 \\ D_2 \theta_1 & -D_2 \theta_2 \end{pmatrix}. \tag{3.6}$$

To test our stability results, we next compare the solutions of the models above in a simulation with initial data as in Figure 3(a), equal particle numbers $\bar{N}_i = 1/2$, equal particle sizes $\bar{\varepsilon}_i = 1$ (since the lattice-based model in Example 1.1 only admits equal sizes), and $D_1 = 1.5$, $D_2 = 1$. We plot the results in Figure 4, using the potentials $V_1(x) = 1 - \exp(-120x^2)$ and $V_2(x) = 0$ as in Figure 2. As expected, the stability between models in Examples 1.1 and 1.2 is of order ϵ, whereas the difference between the solutions of models in Examples 1.2 and 1.3 scales with ϵ^2.

Appendix A. Proof of Lemma 2.1

Our approach is classical and the parabolic estimate mostly follows from an elliptic regularity estimate. Yet, for general cross-diffusion systems, it is well known that such elliptic results do not always hold, including for quasilinear systems with analytic dependence on u (see for example [10, 22]). Therefore this result needs to be proved in the case at hand. Some of the more technical arguments are detailed in well known references (for example [12, 23] concerning elliptic regularity and [8, 9, 15, 16] for the parabolic case), so we safely skip a certain number of intermediate steps, and we give the relevant references.

The following lemma provides the key regularity result.

Lemma A.1. *Given $\omega \in C^1\left(\overline{\Omega}; \mathbb{R}^+\right)$, for any $u^0 \in H^2\left(\Omega; \mathbb{R}^m\right)$ and*

$$g_i \in C\left([0,T]; H^1\left(\Omega; \mathbb{R}^d\right)\right) \cap H^1\left(0,T; L^2\left(\Omega; \mathbb{R}^d\right)\right), \quad i = 1, \ldots, m$$

the weak solution u of

$$\omega \partial_t u_i - \partial_\alpha \left[D_i^{\alpha\beta}(x,t) \partial_\beta u_i + F_{ij}^{\alpha}(x,t) u_j + g_i^\alpha \right] = 0, \quad \text{in} \quad \Omega, \tag{A.1}$$

$$\left[D_i^{\alpha\beta}(x,t) \partial_\beta u_i + F_{ij}^{\alpha}(x,t] u_j + g_i^\alpha \right] \cdot \nu^\alpha = 0, \quad \text{on} \quad \partial\Omega, \tag{A.2}$$

$$u_i(0) = u_i^0, \quad \text{in} \quad \Omega, \tag{A.3}$$

for $i = 1, \ldots, m$, is unique in $L^2\left(0,T; H^1(\Omega)\right) \cap C\left([0,T]; L^2(\Omega)\right)$ with $\partial_t u \in L^2(0,T;(H^1(\Omega))')$. If the compatibility condition

$$\left[D_i^{\alpha\beta}(x,t) \partial_\beta u_i^0 + F_{ij}^{\alpha} u_j^0 + g_i^\alpha \right] \cdot \nu^\alpha = 0 \quad \text{on } \partial\Omega, \quad i = 1, \ldots, m \tag{A.4}$$

holds, then u satisfies

$$\|u\|_{W(Q_T)} \le \frac{1}{2} C_T \left(\left\|u^0\right\|_{H^2(\Omega)} + \|g\|_{C([0,T];H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))} \right), \tag{A.5}$$

where the constant C_T is given by (A.30) and depends on m, λ, T, the C^1 norms of ω, D, and F, and the domain Ω only.

Furthermore, if $F_{ij}^{\alpha} = D_i^{\alpha\beta}\partial_\beta V_i$ with $V_i \in C^1\left(\overline{\Omega};\mathbb{R}\right)$, and for each i, D_i and V_i do not depend on time, then

$$\|u\|_{W(Q_T)} \leq \frac{1}{2}C_\infty\left(\left\|u^0\right\|_{H^2(\Omega)} + \|g\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))}\right), \tag{A.6}$$

where C_∞, given by (A.31), depends on m, λ, the C^1norms of ω, D, F and the domain Ω only. In particular, C_∞ is independent of T.

Proof. Note that no coupling appears in (A.3), therefore the index i can be dropped, as the result relates to equations, and not systems. For the purpose of this proof, it is convenient to modify the formulation of the problem to simplify the computations. We will write $D = A^2$, with $A \in C^1\left(\overline{Q_T};\mathbb{R}^{d\times d}\right)$ symmetric, positive definite and A satisfies

$$\left\|A^{-1}(x,t)\right\|_\infty \leq \lambda^{-1/2} \text{ in } Q_T. \tag{A.7}$$

We write $F = AB$, and $g = Af$, so that the evolution problem under consideration can be written under the form

$$\omega\partial_t u - \operatorname{div}\left(A^2\nabla u + ABu + Af\right) = 0 \quad \text{in } \mathcal{D}'\left(\Omega\right). \tag{A.8}$$

The *a priori* bounds we will use are

$$\|A\|_{L^\infty(Q_T)} + \|\nabla A\|_{L^\infty(Q_T)} \leq M_A, \tag{A.9}$$

$$\|B\|_{L^\infty(Q_T)} + \|B\|_{L^\infty(Q_T)} + \|\nabla B\|_{L^\infty(Q_T)} \leq M_B, \tag{A.10}$$

$$\left\|\omega^{-1}\right\|_{L^\infty(Q_T)} + \|\omega\|_{L^\infty(Q_T)} + \|\nabla\omega\|_{L^\infty(Q_T)} \leq M_\omega, \tag{A.11}$$

and

$$\|\partial_t A\|_{C\left(\overline{Q_T}\right)} + \left\|A^{-1}\partial_t A\right\|_{C\left(\overline{Q_T}\right)} + \left\|\partial_t A^{-1}\right\|_{C\left(\overline{Q_T}\right)} + \left\|\partial_t\left(A^{-1}B\right)\right\|_{C\left(\overline{Q_T}\right)} \leq M_T. \tag{A.12}$$

For *a.e.* $t \in [0,T]$, we define $\mathcal{A}\left(t,u,v\right) : H^1(\Omega;\mathbb{R}) \times H^1(\Omega;\mathbb{R}) \to \mathbb{R}$ by

$$\mathcal{A}(t,u,v) = \int_\Omega \left(A^2\right)^{\alpha\beta}(t,x)\,\partial_\beta u\partial_\alpha v\,\mathrm{d}x + \int_\Omega (AB)^\alpha(t,x)\,u\partial_\alpha v\,\mathrm{d}x. \tag{A.13}$$

Using the *a priori* bounds (A.9) and (A.10), we find the upper bound

$$\mathcal{A}(t,u,v) \leq M_A\left(M_A + M_B\right)\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}.$$

Furthermore, using (A.7) as well, we have the lower bound

$$\mathcal{A}(t,u,v) \geq \lambda\|u\|^2_{H^1(\Omega)} - M_A M_B\|u\|_{L^2(\Omega)}\|u\|_{H^1(\Omega)} \geq \frac{1}{2}\lambda\|u\|^2_{H^1(\Omega)} - \frac{1}{2\lambda}M_A^2M_B^2\|u\|^2_{L^2(\Omega)}. \tag{A.14}$$

We may therefore apply the parabolic version of the Lax–Milgram Theorem of Lions [1, 16] to deduce that there exists a unique solution of (A.3) $u \in L^2(0,T;H^1(\Omega;\mathbb{R})) \cap C([0,T];L^2(\Omega;\mathbb{R}))$ with $\partial_t u \in L^2(0,T;H^1(\Omega;\mathbb{R})')$.

We now derive an explicit bound. Integrating (A.8) by parts against u we find

$$\partial_t \frac{1}{2} \int_\Omega \omega u^2 \, \mathrm{d}x + \int_\Omega A^2(x,t) \nabla u \cdot \nabla u \, \mathrm{d}x + \int_\Omega u AB \cdot \nabla u \, \mathrm{d}x + \int_\Omega Af \cdot \nabla u \, \mathrm{d}x = 0.$$

Thus, using (A.14) and Cauchy–Schwarz

$$\partial_t \left(\frac{1}{2} \left\| \sqrt{\omega} u \right\|^2_{L^2(\Omega)} \right) + \frac{1}{2} \| A \nabla u \|^2_{H^1(\Omega)} \le \| f \|^2_{L^2(\Omega)} + \left\| \omega^{-\frac{1}{2}} B \right\|^2_{L^\infty(\Omega)} \left\| \sqrt{\omega} u \right\|^2_{L^2(\Omega)},$$

which leads to two bounds

$$\| u \|_{C([0,T], L^2(\Omega))} \le M_\omega^{\frac{1}{2}} \left(\exp\left(\sqrt{2} M_\omega M_B T \right) \| f \|_{L^2(Q_T)} + M_\omega^{\frac{1}{2}} \left\| u^0 \right\|_{L^2(\Omega)} \right), \tag{A.15}$$

and

$$\sqrt{\lambda/2} \| \nabla u \|_{L^2(0,T;L^2(\Omega))} \le \sqrt{1/2} \| A \nabla u \|_{L^2(0,T;L^2(\Omega))} \le \sqrt{\frac{M_\omega}{2}} \left\| u^0 \right\|_{L^2(\Omega)} + \| f \|_{L^2(Q_T)}. \tag{A.16}$$

Note that $\int_\Omega u \, \mathrm{d}x = \int_\Omega u^0 \, \mathrm{d}x$ for all times. As a result,

$$\begin{aligned} \| u \|_{L^2(0,T;H^1(\Omega))} &\le \sqrt{T} \left| \frac{1}{|\Omega|} \int_\Omega u^0 \, \mathrm{d}x \right| + \left\| u - \frac{1}{|\Omega|} \int_\Omega u \, \mathrm{d}x \right\|_{L^2(0,T;L^2(\Omega))} + \| \nabla u \|_{L^2(0,T;L^2(\Omega))} \\ &\le \sqrt{T} \left| \frac{1}{|\Omega|} \int_\Omega u^0 \, \mathrm{d}x \right| + (C_P(\Omega) + 1) \lambda^{-1/2} \| A \nabla u \|_{L^2(0,T;L^2(\Omega))} \\ &\le C_1 \left(\left\| u^0 \right\|_{L^2(\Omega)} + \| f \|_{L^2(Q_T)} \right), \end{aligned} \tag{A.17}$$

where $C_P(\Omega)$ is the Poincaré–Wirtinger constant, and

$$C_1 = \sqrt{T} \, |\Omega|^{-1/2} + (C_P(\Omega) + 1) \sqrt{\frac{M_\omega}{2\lambda}}. \tag{A.18}$$

Let us now focus on higher regularity. We are going to show that

$$u \in C\left([0,T]; H^2(\Omega) \right) \cap H^1\left(0,T; H^1(\Omega) \right).$$

We write

$$\Phi = A \nabla u + Bu + f. \tag{A.19}$$

Thanks to (A.16) and (A.17), we have

$$\| \Phi \|_{L^2(0,T;L^2(\Omega))} \le C_2 \left(\left\| u^0 \right\|_{L^2(\Omega)} + \| f \|_{L^2(Q_T)} \right),$$

with

$$C_2 = 1 + \sqrt{M_\omega} + M_B C_1. \tag{A.20}$$

Next, we are going to test (A.3) against $\eta = \partial_t u - \omega^{-1} \operatorname{div}(A\Phi)$. Notice that we have to ensure that η is a valid test function. We just sketch the procedure, namely we consider $\eta_{\tau,h} = \Delta_\tau u - \omega^{-1}\Delta_h^\alpha \left(A^{\alpha\beta}\Phi^\beta\right)$, where the difference quotient time derivative is given by $\Delta_\tau u = \left(u\left(\cdot+\tau\right) - u\left(\cdot\right)\right)\tau^{-1}$ and difference quotient space derivatives in direction i is given by $\Delta_{-h}^\alpha \psi = \left(\psi\left(\cdot + he_\alpha\right) - \psi\left(\cdot\right)\right)h^{-1}$. We have to test (A.3) against $\eta_{\tau,h}$ and subsequently pass to the limit for $\tau, h \to 0$, paying attention to the direction normal to the boundary near $\partial\Omega$. This step is somewhat technical but straightforward and it justifies the following calculations rigorously. To simplify the exposition, we use directly $\partial_t u - \omega^{-1}\operatorname{div}(A\Phi)$ as the test function in the following steps, and obtain

$$\int_\Omega \omega\left(\partial_t u\right)^2 \,\mathrm{d}x + \int_\Omega (A\Phi)\cdot\nabla\left(-\omega^{-1}\operatorname{div}(A\Phi)\right)\,\mathrm{d}x - 2\int_\Omega \partial_t u \operatorname{div}(A\Phi)\,\mathrm{d}x = 0. \tag{A.21}$$

As $A\Phi\cdot\nu = 0$, we find that

$$\int_\Omega (A\Phi)\cdot\nabla\left(-\omega^{-1}\operatorname{div}(A\Phi)\right)\,\mathrm{d}x = \int_\Omega \omega^{-1}\left(\operatorname{div}(A\Phi)\right)^2\,\mathrm{d}x. \tag{A.22}$$

Let us now turn to the mixed term. We have

$$\begin{aligned} -2\int_\Omega \partial_t u \operatorname{div}(A\Phi)\,\mathrm{d}x &= 2\int_\Omega \partial_t\left(\left(A^{-1}A\right)\nabla u\right)\cdot(A\Phi)\,\mathrm{d}x = 2\int_\Omega \left[\partial_t(A\nabla u) + A\partial_t(A^{-1})A\nabla u\right]\cdot\Phi\,\mathrm{d}x \\ &= 2\int_\Omega\left[\partial_t(\Phi) + A\partial_t(A^{-1})\Phi\right]\cdot\Phi\,\mathrm{d}x - 2\int_\Omega\left[\partial_t(Bu+f) + A\partial_t(A^{-1})(Bu+f)\right]\cdot\Phi\,\mathrm{d}x. \end{aligned} \tag{A.23}$$

Inserting (A.22) and (A.23) into (A.21) and using Cauchy–Schwarz, we obtain

$$\begin{aligned} &\left\|\sqrt{\omega}\partial_t u\right\|^2_{L^2(\Omega)} + \int_\Omega \omega^{-1}\operatorname{div}(A\Phi)^2\,\mathrm{d}x + \partial_t\left\|\Phi\right\|^2_{L^2(\Omega)} \\ &\quad\le 2M_T M_A\left(\left\|\Phi\right\|^2_{L^2(\Omega)} + \left\|f\right\|_{L^2(\Omega)}\left\|\Phi\right\|_{L^2(\Omega)}\right) + 2\left\|\partial_t f\right\|_{L^2(\Omega)}\left\|\Phi\right\|_{L^2(\Omega)} \\ &\quad\quad + 2M_T M_A M_B\left\|u\right\|_{L^2(\Omega)}\left\|\Phi\right\|_{L^2(\Omega)} + 2M_B M_\omega^{\frac12}\left\|\sqrt{\omega}\partial_t u\right\|_{L^2(\Omega)}\left\|\Phi\right\|_{L^2(\Omega)}. \end{aligned} \tag{A.24}$$

Using Young's inequality, we recombine inequality (A.24) to find

$$\frac12\left\|\sqrt{\omega}\partial_t u\right\|^2_{L^2(\Omega)} + \left\|\sqrt{\omega}\operatorname{div}(A\Phi)\right\|^2_{L^2(\Omega)} + \partial_t\left\|\Phi\right\|^2_{L^2(\Omega)} \le (2C_3+1)\left\|\Phi\right\|^2_{L^2(\Omega)} + M_B^2\left\|u\right\|^2_{L^2(\Omega)} + \left\|\partial_t f\right\|^2_{L^2(\Omega)} + \left\|f\right\|_{L^2(\Omega)},$$

with

$$C_3 = 2M_T M_A\left(1 + 2M_T M_A\right) + 2M_B^2 M_\omega.$$

Integrating in time, we find

$$\begin{aligned} &\left\|\Phi\right\|^2_{C([0,T],L^2(\Omega))} + \frac12\left\|\sqrt{\omega}\partial_t u\right\|^2_{L^2(Q_T)} + \left\|\sqrt{\omega}\operatorname{div}(A\Phi)\right\|^2_{L^2(Q_T)} \\ &\quad\le C_4\left(\left\|u^0\right\|_{L^2(\Omega)} + \left\|f\right\|_{L^2(Q_T)}\right)^2 + \left\|A\nabla u^0 + Bu^0 + f\left(t=0\right)\right\|^2_{L^2(\Omega)} + \left\|\partial_t f\right\|^2_{L^2(Q_T)} + \left\|f\right\|^2_{L^2(Q_T)}, \end{aligned}$$

with

$$C_4 = \left(2C_3+1\right)C_2^2 + M_B^2 C_1^2. \tag{A.25}$$

Let us now check that the estimate above allows us to define $\partial_t u|_{t=0}$ in an appropriate sense. Since

$$\|\nabla u\|_{C([0,T],L^2(\Omega))} \leq \lambda^{-\frac{1}{2}} \left(\|\Phi\|_{C([0,T],L^2(\Omega))} + M_B \|u\|_{C([0,T],L^2(\Omega))} + \|f\|_{C([0,T],L^2(\Omega))}\right),$$

for any $v \in H^1(\Omega)$, the map

$$t \to \int_\Omega [A(x,t)\nabla u \cdot \nabla v + Bu \cdot \nabla v + fu \cdot \nabla v] \, \mathrm{d}x,$$

is continuous on $[0,T]$. In other words, we define $\partial_t u|_{t=0} \in \left(H^1(\Omega)\right)'$ as follows

$$\begin{aligned}\int_\Omega \partial_t u|_{t=0} \, v \, \mathrm{d}x &= \lim_{t \downarrow 0} \int_\Omega [A(x,t)\nabla u \cdot \nabla v + B(x,t)u.\nabla v + f(x,0) \cdot \nabla v] \, \mathrm{d}x \\ &= \int_\Omega \left[A(x,0)\nabla u^0 \cdot \nabla v + B(x,0)u^0.\nabla v + f(x,0) \cdot \nabla v\right] \mathrm{d}x,\end{aligned}$$

provided that the compatibility condition (A.4) holds, that is,

$$\left[A(x,0)\nabla u^0 - B(x,0)u^0 - f(x,0)\right] \cdot \nu = 0.$$

An integration by parts then shows that

$$\int_\Omega \partial_t u|_{t=0} \, v \, \mathrm{d}x = \int_\Omega \operatorname{div}\left[A(x,0)\nabla u^0 + B(x,0)u^0 + f(x,0)\right] v \, \mathrm{d}x,$$

which, in turn, shows that $\partial_t u|_{t=0} \in L^2(\Omega)$ and

$$\|\partial_t u|_{t=0}\|_{L^2(\Omega)} \leq (M_A + M_B)\left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))}\right). \tag{A.26}$$

We now notice that $\partial_t u$ is a weak solution of (A.3), where f is replaced by $\partial_t f + \partial_t A \nabla u + \partial_t B u$ and u^0 is replaced by $\partial_t u|_{t=0}$. From (A.17) we obtain

$$\|\partial_t f + \partial_t A \nabla u + \partial_t B u\|_{L^2(Q_T)} \leq \max(M_T C_1, 1)\left(\left\|u^0\right\|_{L^2(\Omega)} + \|f\|_{H^1(0,T;L^2(\Omega))}\right) \tag{A.27}$$

Thus (A.16) becomes

$$\begin{aligned}\sqrt{\frac{\lambda}{2}} \|\partial_t \nabla u\|_{L^2((0,T);L^2(\Omega))} &\leq \sqrt{\frac{M_\omega}{2}} \|\partial_t u|_{t=0}\|_{L^2(\Omega)} + \|\partial_t f + \partial_t A \nabla u + \partial_t B u\|_{L^2(Q_T)} \\ &\leq \sqrt{\frac{M_\omega}{2}} (M_A + M_B)\left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))}\right) \\ &\quad + \max(M_T C_1, 1)\left(\left\|u^0\right\|_{L^2(\Omega)} + \|f\|_{H^1(0,T;L^2(\Omega))}\right),\end{aligned}$$

and (A.15) gives

$$\begin{aligned}\|\partial_t u\|_{C([0,T];L^2(\Omega))} &\leq M_\omega^{\frac{1}{2}} \left[\exp\left(\sqrt{2} M_\omega M_B T\right) \|\partial_t f + \partial_t A \nabla u + \partial_t B u\|_{L^2(Q_T)} + M_\omega^{\frac{1}{2}} \|\partial_t u|_{t=0}\|_{L^2(\Omega)}\right] \\ &\leq C_5 \left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))} + \|f\|_{H^1(0,T;L^2(\Omega))}\right),\end{aligned}$$

with

$$C_5 = M_\omega \left(M_A + M_B\right) + M_\omega^{\frac{1}{2}} \max(M_T C_1, 1) \exp\left(\sqrt{2} M_\omega M_B T\right). \tag{A.28}$$

Finally, we observe that the right-hand side of the identity

$$\operatorname{div}\left(A \nabla u\right) = \partial_t u - \operatorname{div}\left(B u + f\right),$$

belongs to $C\left([0,T]; L^2\left(\Omega\right)\right)$, and therefore the left-hand side belongs to the same space. This in turn shows that $u \in H^2(\Omega)$ for any t, in fact $u \in C\left([0,T]\,; H^2\left(\Omega\right)\right)$, see, for example, [19], with

$$\|u\|_{C([0,T];H^2(\Omega))} \leq C(\Omega, M_A, \lambda)\left(C_5 + M_B C_1\right)\left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))} + \|f\|_{H^1(0,T;L^2(\Omega))}\right).$$

Altogether, we have shown

$$\|u\|_{C([0,T];H^2(\Omega))} + \|u\|_{H^1(0,T;H^1(\Omega))} \leq \frac{1}{2} C_T \left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))} + \|f\|_{H^1(0,T;L^2(\Omega))}\right), \tag{A.29}$$

where

$$C_T = 2\left(C(\Omega, M_A, \lambda)\left(C_5 + M_B C_1\right) + \sqrt{\frac{M_\omega}{\lambda}}\left(M_A + M_B\right) + \sqrt{\frac{2}{\lambda}} \max(M_T C_1, 1)\right), \tag{A.30}$$

and C_4 and C_5 are given by (A.25) and (A.28), respectively, as announced.

Let us now turn to the particular case when $B = A\nabla V$, with $V \in C^2\left(\overline{\Omega}\right)$, and A and V are independent of time. We perform the change of unknown $w = u \exp V$ and, thanks to Lemma A.3, we can study the problem satisfied by w. We have

$$\begin{aligned}
&\exp\left(-V\right) \partial_t w - \operatorname{div}\left[A \exp\left(-V\right) \nabla w + f\right] = 0, \quad \text{in} \quad \Omega,\\
&\left[A \exp(-V) \nabla w + f\right] \cdot \nu = 0, \quad \text{on} \quad \partial\Omega,\\
&w(0) = u^0 \exp(V), \quad \text{in} \quad \Omega,
\end{aligned}$$

that is, the same system as (A.3) above, with $\omega = \exp(-V)$, $M_B = 0$ and $M_T = 0$. In this case,

$$C_2 = 1 + \sqrt{M_\omega}, \quad C_3 = 0, \quad C_4 = C_2^2, \quad C_5 = M_\omega M_A + M_\omega^{\frac{1}{2}},$$

and the constant C_T in (A.30) becomes

$$\tilde{C}' = 2\left(C(\Omega, M_A, \lambda) C_5 + \sqrt{\frac{M_\omega}{\lambda}} M_A + \sqrt{\frac{2}{\lambda}}\right),$$

and it does not depend on T. Thanks to Lemma A.3, we find that the bound (A.29) holds with the following constant

$$C_\infty = \tilde{C}\left[\left(1 + M_V'\right)^2 + M_V''\right] \exp M_V, \tag{A.31}$$

which again is independent of T. $\square$

Remark A.2 (Ellipticity bound for ϵ)**.** Suppose that an *a priori* bound for u on Q_T is known, say $u^* = \sup_{Q_T} |u|$. For any $\xi_i^\alpha \in \mathbb{R}^{d\times m}, \zeta_j \in \mathbb{R}^m$, we have the lower bound

$$\mathfrak{D}_{ij}^{\alpha\beta}(t,x,y)\xi_i^\alpha\xi_j^\beta = D_i^{\alpha\beta}(t,x)\xi_i^\alpha\xi_i^\beta + \epsilon a_{ij}^{\alpha\beta}(t,x)\phi_{ij}^{\alpha\beta}(y)\xi_i^\alpha\xi_j^\beta \geq (\lambda - \epsilon L_0(u^*)\|a\|_\infty)|\xi|^2, \tag{A.32}$$

where $\|a\|_\infty = \max_{i,j,\alpha,\beta,x} |a_{ij}^{\alpha\beta}(x)|$ and L_0 is given in (1.16). Therefore, choosing

$$\epsilon < \min\left(\frac{\lambda}{1+\|a\|_\infty L_0(u^*)}, 1\right), \tag{A.33}$$

guarantees coercivity, and this is sufficient to ensure existence and uniqueness of weak solutions of the linearised system (2.3) *via* Lax–Milgram lemma. This in turn ensures the existence and uniqueness of weak solutions of the original system (1.9). We use relation (A.33) to derive an *a priori* upper bound for ϵ in a specific case, see Lemma 3.1.

Lemma A.3. *Given $V \in C^2\left(\overline{\Omega}\right)$, the map $u \to u\exp(V)$ is a bi-continuous isomorphism in $C\left([0,T]; H^2(\Omega)\right) \cap H^1\left(0,T; H^1(\Omega)\right)$. The following inequalities hold*

$$\begin{aligned}
\|u\exp(V)\|_{W(0,T,\Omega)} &\leq \left[(1+M_V')^2 + M_V''\right]\exp M_V \|u\|_{W(0,T,\Omega)},\\
\|u\|_{W(0,T,\Omega)} &\leq \left[(1+M_V')^2 + M_V''\right]\exp M_V \|u\exp(V)\|_{W(0,T,\Omega)},
\end{aligned}$$

where $M_V = \sup_\Omega |V|$, $M_V' = \sup_\Omega |\nabla V|$ and $M_V'' = \sup_\Omega \left|\nabla^2 V\right|$.

Proof. Note that it is sufficient to prove one inequality, as replacing V by $-V$ changes the map to its inverse. Indeed, we have

$$\begin{aligned}
\|u\exp(V)\|_{L^2(\Omega)} &\leq \exp M_V \|u\|_{L^2(\Omega)},\\
\|u\exp(V)\|_{H^1(\Omega)} &\leq (1+M_V')\exp M_V \|u\|_{H^1(\Omega)},\\
\|u\exp(V)\|_{H^2(\Omega)} &\leq \left[(1+M_V')^2 + M_V''\right]\exp M_V \|u\|_{H^2(\Omega)}.
\end{aligned}$$

□

The second step in the proof of Lemma 2.1 concerns the regularity of the forcing term f, which coincides with the regularity of the cross-diffusion term, provided that h and u are in $W(Q_T)$.

Lemma A.4. *The map*

$$\begin{aligned}
&P : Q_T \times C^\infty(Q_T;\mathbb{R}^m)^2 \to C^2\left(Q_T : \mathbb{R}^{m\times d}\right)\\
&(t,x,h,u) \to a_{ij}^{\alpha\beta}(t,x)\phi_{ij}^{\alpha\beta}(h)\,\partial_\beta u_j + b_{ij}^\alpha(t,x)\psi_{ij}^\alpha(h)\,u_j,
\end{aligned} \tag{A.34}$$

has the following property

$$P(Q_T \times W(Q_T) \times W(Q_T)) \subset C\left([0,T]; H^1(\Omega;\mathbb{R}^m)\right) \cap H^1\left(0,T; L^2(\Omega;\mathbb{R}^m)\right).$$

Furthermore, there holds

$$\sup_{[0,T]}\left(\|\nabla P_i(t,x,h,u)\|_{L^2(\Omega)} + \|P_i(t,x,h,u)\|_{L^2(\Omega)}\right) + \|\partial_t P_i(t,x,h,u)\|_{L^2(Q_T)} \leq K_0\left(\|h\|_{W(Q_T)}\right)\|u\|_{W(Q_T)},$$

where K_0 is given by (1.17).

Proof. Note that $L^\infty(Q_T) \subset C\left([0,T]; H^2(\Omega;\mathbb{R}^m)\right)$. Therefore

$$\sup_{Q_T} |h| \leq C_S^\infty \|h\|_{W(Q_T)},$$

where

$$C_S^\infty = C(H^2(\Omega) \hookrightarrow L^\infty(\Omega)), \tag{A.35}$$

is the Sobolev constant associated to the embedding of $H^2(\Omega)$ into $L^\infty(\Omega)$, and depends on Ω and d. We compute the following bounds for P

$$\begin{aligned}
&|P_i(t,x,h,u)| \leq \sup_{\Omega\times[0,\infty)} (|a|,|b|) L_0\left(\sup_{Q_T}|h|\right)(|\nabla u| + |u|),\\
&\|P_i(t,x,h,u)\|_{L^2(\Omega)} \leq M\|u\|_{H^1(\Omega)} \leq M L_0\left(C_S^\infty \|h\|_{W(Q_T)}\right)\|u\|_{W(Q_T)},
\end{aligned}$$

for all $t \in [0,T]$. Similarly, for the spatial derivatives of P we have

$$\begin{aligned}
|\partial_\alpha P_i(t,x,h,u)| \leq{}& M\left(L_0\left(C_S^\infty\|h\|_{W(Q_T)}\right) + L_1\left(C_S^\infty\|h\|_{W(Q_T)}\right)|\nabla h|\right)(|\nabla u| + |u|)\\
&+ M L_0\left(C_S^\infty\|h\|_{W(Q_T)}\right)\left(|\nabla^2 u| + |\nabla u|\right).
\end{aligned}$$

Therefore, using Cauchy–Schwarz and the Sobolev embedding $H^1(\Omega) \hookrightarrow L^4(\Omega)$ we find

$$\begin{aligned}
\|\nabla P_i(t,x,h,u)\|_{L^2(\Omega)} &\leq 2ML_0\left(C_S^\infty\|h\|_{W(Q_T)}\right)\|u\|_{W(Q_T)} + ML_1\left(C_S^\infty\|h\|_{W(Q_T)}\right)\|\nabla h\|_{L^4}\left(\|\nabla u\|_{L^4} + \|u\|_{L^4}\right)\\
&\leq M\left[2L_0\left(C_S^\infty\|h\|_{W(Q_T)}\right) + C_S^2 L_1\left(C_S^\infty\|h\|_{W(Q_T)}\right)\|h\|_{W(Q_T)}\right]\|u\|_{W(Q_T)},
\end{aligned}$$

where C_S^2 is defined by (2.7). This shows that $P_i(t,x,h,u) \in C\left([0,T]; H^1(\Omega)\right)$. Finally, for the time derivative we obtain

$$\begin{aligned}
|\partial_t P_i(t,x,h,u)| \leq{}& M\left[L_0\left(C_S^\infty\|h\|_{W(Q_T)}\right) + L_1\left(C_S^\infty\|h\|_{W(Q_T)}\right)|\partial_t h|\right](|\nabla u| + |u|)\\
&+ ML_0\left(C_S^\infty\|h\|_{W(Q_T)}\right)(|\nabla\partial_t u| + |\partial_t u|),
\end{aligned}$$

and

$$\|\partial_t P_i(t,x,h,u)\|_{L^2(Q_T)} \leq M\left[2L_0\left(C_S^\infty\|h\|_{W(Q_T)}\right) + L_1\left(C_S^\infty\|h\|_{W(Q_T)}\right)\|\partial_t h\|_{L^2(Q_T)}\right]\|u\|_{W(Q_T)}.$$

Altogether we have shown that

$$\sup_{[0,T]}\left(\|\nabla P_i(t,x,h,u)\|_{L^2(\Omega)} + \|P_i(t,x,h,u)\|_{L^2(\Omega)}\right) + \|\partial_t P_i(t,x,h,u)\|_{L^2(Q_T)} \leq K_0\left(\|h\|_{W(Q_T)}\right)\|u\|_{W(Q_T)},$$

where K_0 is defined by (1.17), as announced. $\square$

Proof of Lemma 2.1. We write

$$\partial_\alpha \left[\mathfrak{D}_{ij}^{\alpha\beta}(t,x,h)\partial_\beta u_j - \mathfrak{F}_{ij}^{\alpha}(t,x,h)u_j + f_i^\alpha\right] = \partial_\alpha \left[D_i^{\alpha\beta}(t,x)\partial_\beta u_j - F_i^{\alpha}(t,x)u_j + g_i^\alpha\right],$$

with $g_i^\alpha = f_i^\alpha + \epsilon P_i^\alpha(t,x,h,u)$, and P given by (A.34). Lemma A.1 shows that

$$\|u\|_{W(Q_T)} \leq \frac{1}{2} C_T \left(\left\|u^0\right\|_{H^2(\Omega)} + \|g\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))}\right),$$

and

$$\|g\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))} \leq \|f\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))} + \epsilon \left\|P(t,x,h,u)\right\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))}.$$

Thanks to Lemma A.4, there holds

$$\sup_{[0,T]} \left(\|\nabla P_i\left(t,x,h,u\right)\|_{L^2(\Omega)}\right) + \|\partial_t P_i\left(t,x,h,u\right)\|_{L^2(Q_T)} \leq K_0\left(\|h\|_{W(Q_T)}\right)\|u\|_{W(Q_T)},$$

and therefore

$$\|u\|_{W(Q_T)}\left[1 - K_0\big(\|h\|_{W(Q_T)}\big)\right] \leq \frac{1}{2} C_T \left(\left\|u^0\right\|_{H^2(\Omega)} + \|f\|_{C([0,T];H^1(\Omega))\cap H^1(0,T;L^2(\Omega))}\right),$$

which is our thesis, as thanks to the Fredholm Alternative, boundedness implies existence and uniqueness. The proof in the time independent case is analogous and C_T is replaced by C_∞. □

Acknowledgements. The authors are very thankful for the detailed comments and suggestions of the referees which have significantly improved the quality of this manuscript. The third author was visiting Laboratoire Jacques-Louis Lions during the final stage of this paper, and he is very grateful for the hospitality and warmth of his hosts.

References

[1] H. Brezis. Analyse fonctionnelle: Théorie et applications. *Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]*. Masson, Paris (1983).
[2] M. Bruna, M. Burger, H. Ranetbauer, and M.-T. Wolfram, Cross-diffusion systems with excluded-volume effects and asymptotic gradient flow structures. *J. Nonlinear Sci.* **27** (2017) 687–719.
[3] M. Bruna and S.J. Chapman, Diffusion of multiple species with excluded-volume effects. *J. Chem. Phys.* **137** (2012) 204116.
[4] M. Burger, M. Di Francesco, J.-F. Pietschmann, and B. Schlake, Nonlinear cross-diffusion with size exclusion. *SIAM J. Math. Anal.* **42** (2010) 2842.
[5] F. Camilli and C. Marchi, Continuous dependence estimates and homogenization of quasi-monotone systems of fully nonlinear second order parabolic equations. *Nonlinear Anal. Theor.* **75** (2012) 5103–5118.
[6] J.R. Cannon, W.T. Ford, and A.V. Lair, Quasilinear parabolic systems. *J. Differ. Equ.* **20** (1976) 441–472.
[7] L. Chen and A. Jüngel, Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differ. Equ.* **224** (2006) 39–59.
[8] R. Dautray and J.-L. Lions, Mathematical Analysis and Numerical Methods for Science and Technology. Springer Verlag, Berlin (1993).
[9] L.C. Evans, Partial Differential Equations. American Mathematical Society (1998).
[10] M. Giaquinta and L. Martinazzi, An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs. Vol. 2 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa (2005).
[11] M. Giaquinta and M. Struwe, On the partial regularity of weak solutions of nonlinear parabolic systems. *Math. Z.* **179** (1982) 437–451.
[12] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Vol. 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA (1985).
[13] A. Jüngel, The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity* **28** (2015) 1963.

[14] S.N. Kružkov, First order quasilinear equations in several independent variables. *Math. USSR SB+* **10** (1970) 217.
[15] O.A. Ladyzhenskaia, V.A. Solonnikov, and N.N. Ural'tseva, Linear and Quasi-Linear Equations of Parabolic Type. Vol. 23. American Mathematical Society (1988).
[16] E. Magenes and J. Lions, Problèmes aux limites non homogènes. Dunod, Paris (1968) Vol. 1–3.
[17] B.t. Perthame, Parabolic Equations in Biology: Growth, Reaction, Movement and Diffusion. *Lecture Notes on Mathematical Modelling in the Life Sciences*. Springer, Cham (2015).
[18] M.J. Plank and M.J. Simpson, Models of collective cell behaviour with crowding effects: comparing lattice-based and lattice-free approaches. *J. R. Soc. Interface* **9** (2012) 2983–2996.
[19] M. Plum, Explicit h2-estimates and pointwise bounds for solutions of second-order elliptic boundary value problems. *J. Math. Anal. Appl.* **165** (1992) 36–61.
[20] N. Shigesada, K. Kawasaki, and E. Teramoto, Spatial segregation of interacting species. *J. Theor. Biol.* **79** (1979) 83–99.
[21] M.J. Simpson, K.A. Landman, and B.D. Hughes, Multi-species simple exclusion processes. *Phys. A: Stat. Mech. Appl.* **388** (2009) 399–406.
[22] J. Stará and O. John, Some (new) counterexamples of parabolic systems. *Commentat. Math. Univ. Carol.* **36** (1995) 503–510
[23] G.M. Troianiello, Elliptic Differential Equations and Obstacle Problems. *The University Series in Mathematics*. Plenum Press, New York (1987).
[24] L. Zhornitskaya and A.L. Bertozzi, Positivity-preserving numerical schemes for lubrication-type equations. *SIAM J. Numer. Anal.* **37** (2000) 523–555.