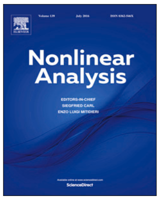
Nonlinear Analysis 223 (2022) 113064

Contents lists available at ScienceDirect

Nonlinear Analysis

www.elsevier.com/locate/na



Existence and regularity for a system of porous medium equations with small cross-diffusion and nonlocal drifts

Luca Alasio[a], Maria Bruna[b,*], Simone Fagioli[c], Simon Schulz[d]

[a] LJLL, Sorbonne Université, 4 Place Jussieu 75005 Paris, France
[b] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK
[c] DISIM, Università degli Studi dell'Aquila, Via Vetoio 1, 67100 Coppito, L'Aquila, Italy
[d] Department of Mathematics, University of Wisconsin-Madison, Van Vleck Hall, 480 Lincoln Dr, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 3 March 2022
Accepted 26 June 2022
Communicated by Jan Kristensen

MSC:
35K65
35B30
35Q92
35K51
35K55

Keywords:
Degenerate cross-diffusion
Vanishing diffusivity
Schauder fixed point
Fisher information

ABSTRACT

We prove the existence and Sobolev regularity of solutions of a nonlinear system of degenerate-parabolic PDEs with self- and cross-diffusion, transport/confinement and nonlocal interaction terms. The macroscopic system of PDEs models the evolution of an arbitrary number of species with quadratic porous-medium interactions in a bounded domain Ω in any spatial dimension and originates from a many-particle system. The cross interactions between different species are scaled by a parameter $\delta < 1$, with the $\delta = 0$ case corresponding to no interactions across species. A smallness condition on δ ensures existence of solutions up to an arbitrary time $T > 0$ in a subspace of $L^2(0, T; H^1(\Omega))$. This is shown via a Schauder fixed point argument for a regularised system followed by a vanishing diffusivity approach. The proof uses the lower semicontinuity of the Fisher information in combination with the div–curl Lemma. An ad hoc weak–strong uniqueness result ensures equivalence between weak formulations of the regularised problem; this is proved by studying a related dual problem. We provide numerical evidence showing blow-up of the Sobolev norm for $\delta \to 1$.

1. Introduction

Partial differential equations involving nonlinear diffusion in combination with local or nonlocal transport are commonly used to model the macroscopic behaviour of a large number of agents or individuals in the natural, life, and social sciences. Systems involving many species or types of agents have been used to study multiple chemotactic populations in competition for nutrient [1,2], tumour growth [3–5], pedestrian dynamics [6,7], and opinion formation [8]. Further applications can be found in population biology [9,10], in neural networks [11], in semiconductor devices [12,13] and even in modelling rival gangs in a city [14].

---

[*] Corresponding author.

E-mail addresses: luca.alasio@sorbonne-universite.fr (L. Alasio), bruna@maths.cam.ac.uk (M. Bruna), simone.fagioli@univaq.it (S. Fagioli), smschulz2@wisc.edu (S. Schulz).

https://doi.org/10.1016/j.na.2022.113064

In this paper we study a class of drift–diffusion systems taking the following form:

$$\partial_t u_i = \operatorname{div}\left[u_i(\nabla u_i + \nabla L_i(t,x,u_i) + \delta F_i(t,x,u,\nabla u))\right], \qquad i = 1,\dots,M, \tag{1}$$

where $u = (u_1,\dots,u_M)$ is a vector of non-negative functions defined on a bounded domain $Q_T = (0,T)\times\Omega$ describing the densities of M subpopulations. The transport term L_i models the presence of external forces and nonlocal self-interactions, $L_i(t,x,u_i) = V_i(t,x) + \big(W_i * u_i(t,\cdot)\big)(x)$, while F_i represents the interaction with the other species. The latter is of particular importance since it includes the contribution of *cross-diffusion terms* through the dependence on the vector ∇u. Eq. (1) can be seen as a macroscopic description of a particle-based model. As a particular example, consider an interacting particle system with M species, each composed of N_i identical particles at position $X_k^i(t) \in \Omega$ and evolving according to

$$\mathrm{d}X_k^i = -\nabla V_i(t,X_k^i)\mathrm{d}t - \sum_{\substack{j=1\\(\ell,j)\neq(k,i)}}^{M}\sum_{\ell=1}^{N_i} \nabla K_{ij}(X_k^i - X_\ell^j)\mathrm{d}t, \tag{2}$$

where K_{ij} is the pairwise interacting potential on a particle of species i due to a particle of species j. As we discuss in Section 3.1, the formal mean-field limit of (2) as $N = \sum_{i=1}^M N_i$ going to infinity is given by (23) with $L_i = V_i$ and $F_i(t,x,u,\nabla u) = \sum_{j=1,j\neq i}^M \nabla u_j$.

One of the main features of (1) is that it can be used to describe cell sorting and the resulting pattern formation. This is a reorganisation process in which cells of different species – which in principle react differently to external forces such as a chemical signal or attraction/repulsion with the other species – have the propensity to group in a delimited region [15–18]. This biological phenomenon can also be interpreted as the inhibition or activation of growth whenever two populations occupy the same habitat, which can be attributed to volume or size constraints of the individual cells forming the different populations. In the seminal papers [19–23] it was shown that segregation is induced by the presence of cross-diffusion terms. Nonlinear diffusion may also help in describing volume filling effects and in preventing blow-up in biological aggregation models, see [24–28].

The main goal of the present paper is to provide a well-posedness result for (1), which can be seen as a δ-order perturbation of a set of M decoupled drift–diffusion equations with degenerate diffusion of porous-media type and nonlocal interactions. One of the main difficulties for these systems is the lack of a suitable maximum principle, meaning that Sobolev estimates must be obtained in an alternative fashion. Under a *smallness* assumption on the perturbation parameter δ, we can estimate each density in $L^2(0,T;H^1(\Omega))$, as well as the relative time derivatives in a dual Sobolev space. Such estimates allow us to show convergence to a weak solution for a proper regularised approximating sequence. The structure of the system and the numerical simulations indicate that the expected critical value for the $L^2(0,T;H^1(\Omega))$-framework is $\delta = 1$. However, technical conditions impose a stricter bound on δ. The BV norm is the only norm expected to remain bounded as $\delta \to 1$. We highlight that the upper bound for δ depends not on the initial datum but on the final time. However, this dependence appears to be an effect of the technical steps in the proof. An illustrative example in which the smallness condition on δ is explicitly computed is provided in Section 3.2. The value of δ corresponding to the short final time can be considered an effective upper bound for δ.

The one-species counterpart of system (1) has been largely studied in the literature, see for example [29,30] and references therein. Nevertheless, a complete well-posedness theory for cross-diffusion systems in the presence of transport terms is not currently available. Indeed, the separation process between species described at the beginning of this section may lead to discontinuities of the densities and their derivatives at the interface between different species. These issues are also accentuated by the presence of degenerate diffusion terms that, on the one hand, determine the finite speed of propagation in the supports of the solutions and, on the other hand, cause a possible loss of regularity at their boundary.

The main difficulty lies in providing *a priori* estimates for the single components u_1 and u_2 and on their space derivatives. Several attempts were made in this direction, even to extend the concepts of parabolicity, see the classical Refs. [31–34]. In presence of reaction terms instead of the transport terms, an existence theory was obtained in a one-dimensional BV-setting for the case $\delta = 1$ in [19,20,35]; see also [36] for a multi-dimensional result. The BV-setting is somehow natural since the emergence of "segregated solutions" is highlighted in several contexts [37–39], see also [40, Chap. 1, Sect. 4]. A general existence theory for systems with arbitrary cross-diffusion terms and local/nonlocal transport is incomplete.

To achieve a satisfactory theory, many results in the literature have been inspired by the gradient flow structure that can be associated with systems in the form of (3). These can be split into two categories: formal gradient flow structure and a Wasserstein gradient flow theory. In the first group we mention the works [41–45], where a formal gradient flow formulation provides the estimates needed to prove global existence. The second approach concerns the many-species version of Wasserstein gradient flow theory of [29,46]. Such an approach has been already successfully used in [47] for a system of nonlocal interaction equations with two species and non-symmetric cross-interactions, and was first used in a system with cross-diffusion terms but no transport terms in [48]. Other results, only apply to diagonal diffusion and in some cases only in bounded domains [49,50], or with *dominant diagonal parts,di2018nonlinear*, see also [51–54]. In one space dimension, [55] provides an existence result for cross-diffusion systems with *ordered* external potentials.

In order to show existence and Sobolev regularity of solutions to (1) we first regularise the original system by adding an artificial diffusivity term and "freezing" the cross-diffusion terms. We prove existence of the aforementioned regularised problems by employing a Schauder fixed-point approach, where the smallness condition on δ appears explicitly. Central to this method is the requirement that the regularised solution operator be strongly compact, which we prove in two steps; first, by showing that any bounded sequence to which the operator is applied converges weakly, and then by showing convergence to this weak limit in norm. This last step is performed by using semicontinuity properties of the Fisher information in conjunction and the div–curl Lemma. Crucial to this approach is the observation that the previously mentioned weak limit itself satisfies an equation endowed with regularity properties. To exploit this fact, we prove an ad hoc weak–strong uniqueness result to show the equivalence of different notions of solution; see Lemma 4.14. We highlight that this latter result is of independent interest, as it extends the uniqueness theorems in [56] to the case of a system of PDEs with nonlocal drifts. One of the main advantages of our approach is that it does not rely on optimal transport nor on an underlying gradient flow structure of the PDE system under consideration, as done for instance in [57,58]. A detailed summary of our strategy is found in Section 2.2.

In order to better explain the difficulties that cross-diffusion terms bring to the analysis, let us consider the following special case of (1) for $M = 2$:

$$\partial_t u_i = \partial_x \left[u_i \partial_x (u_i + \delta u_j) - u_i V_i'(x) \right], \quad i, j = 1, 2,\ j \neq i. \tag{3}$$

According to the classical theory in [59], the well-posedness of (3) is related to the positive definiteness of the diffusion matrix

$$D(u_1, u_2) = \begin{pmatrix} u_1 & \delta u_1 \\ \delta u_2 & u_2 \end{pmatrix}.$$

Since the matrix above is not symmetric, we must consider its symmetric part $\frac{1}{2}(D + D^T)$, which has determinant

$$\det\left(\frac{1}{2}(D + D^T)\right) = \left(1 - \frac{\delta^2}{2}\right) u_1 u_2 - \frac{\delta^2}{4}(u_1^2 + u_2^2).$$

From the above, it is clear that cross-diffusion may induce a negative quadratic form. The lack of uniform parabolicity (namely the failure of $D(u_1, u_2) \geq c\mathbb{I}$, for some $c > 0$) is present in several applications such as [60,61], and has attracted a lot of interest in recent years.

The paper is organised as follows. In Section 2.1 we provide the general set of assumptions and the statement of the main result, Theorem 1. We sketch the main steps in the strategy of the proof of the main result in Section 2.2. In Section 3.1 we present a formal derivation of system (1) from a system of interacting particles. Section 3.2 is devoted to the numerical investigation of a particular case of system (1), which helps to highlight the influence of the parameter δ in the time evolution of solutions and its norms. The remainder of the paper focuses on the proof of Theorem 1. In Section 4 we introduce a regularised system, providing existence, uniqueness and uniform estimates for these regularised solutions. In Section 5 we present a fixed point argument and show compactness properties for the solution map of the regularised system. Convergence of approximate solutions to weak solutions of system (1) is established in Section 6 through vanishing diffusivity. Appendix A collects some technical results used in the paper.

2. Main result

2.1. Assumptions and notion of solution

Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain of class C^2 with outward normal denoted by ν, and let $T > 0$. We denote the parabolic cylinder by $Q_T := (0,T) \times \Omega$, the lateral boundary by $\Sigma_T := (0,T) \times \partial\Omega$, and the closure by $\bar{Q}_T := [0,T] \times \bar{\Omega}$. Consider also M functions

$$F_i : \bar{Q}_T \times \mathbb{R}^M \times \mathbb{R}^{d\times M} \to \mathbb{R}^d$$

such that the dependence on the last argument is affine, namely

$$F_i(t,x,z,p) = G_i^0(t,x,z) + \sum_{j=1}^{M} G_{ij}^1(t,x,z)p_j, \qquad \text{for } i \in \{1,\dots,M\}, \tag{4}$$

where $\{G_i^0\}_{i=1}^M$ are vector functions with values in $\mathbb{R}^d$ and $\{G_{ij}^1\}_{i,j=1}^M$ take values in the space of $d\times d$ matrices; all of the functions above are assumed to be C^2-regular and uniformly bounded with respect their arguments. The assumption of regularity is not optimal and can be relaxed, but it avoids unpleasant technicalities in Section 4.1 (see also A.1). We emphasise that, in (4), for each $j \in \{1,\dots,M\}$, the quantity p_j is a column vector in $\mathbb{R}^d$, while p is a $d \times M$ matrix whose jth column is the column vector p_j. We assume that there exists a positive constant $C_F = C_F(T,\Omega)$ such that

$$|F_i(t,x,z,p)| \le C_F(1+|p|), \quad \forall (t,x,z,p) \in \bar{Q}_T \times \mathbb{R}^M \times \mathbb{R}^{d\times M}. \tag{5}$$

For $i \in \{1,\dots,M\}$ and $\delta \in \mathbb{R}$ some small constant to be made precise, consider the following system of equations:

$$\begin{cases} \partial_t u_i = \operatorname{div}\left[u_i(\nabla u_i + \nabla L_i(t,x,u_i) + \delta F_i(t,x,u,\nabla u))\right] & \text{in } Q_T, \\ 0 = \nu \cdot \left[u_i(\nabla u_i + \nabla L_i(t,x,u_i) + \delta F_i(t,x,u,\nabla u))\right] & \text{on } \Sigma_T, \\ u_i(0,\cdot) = u_{i,0} & \text{on } \Omega, \end{cases} \tag{6}$$

where $u = (u_i)_{i=1}^M$ is the unknown vector-valued function, and each $u_{i,0}$ is a given non-negative function in $L^p(\Omega)$ for $p > 1$. The transport terms $\{L_i\}_{i=1}^M$ are prescribed by

$$L_i(t,x,u_i) = V_i(t,x) + \big(W_i * u_i(t,\cdot)\big)(x), \tag{7}$$

where we assume W_i to be radially symmetric for each $i \in \{1,\dots,M\}$, and the convolution to be only with respect to the space variable, i.e.,

$$\big(W_i * u_i(t,\cdot)\big)(x) = \int_\Omega W_i(x-y)u_i(t,y)\,\mathrm{d}y \qquad \text{a.e. } (t,x) \in Q_T.$$

Again, to avoid unpleasant technicalities in Section 4.1, we assume $\{V_i, W_i\}_{i=1}^M$ to be C^2-regular and uniformly bounded in all of their respective arguments. In particular, there exists a positive constant C_L such that

$$\max_{i\in\{1,\dots,M\}} \|V_i\|_{C^2(\mathbb{R}^{d+1})} + \max_{i\in\{1,\dots,M\}} \|W_i\|_{C^2(\mathbb{R}^d)} \le C_L. \tag{8}$$

Remark 2.1. We note that minor adaptations of our approach allow to treat additional cross-interaction terms of the form $\operatorname{div}\left(u_i \nabla \sum_{j,k=1}^M W_j * u_k(t,\cdot)(x)\right)$ in (6). However, in order to do this, cross-interaction terms must be included as part of the term F_i, i.e., they must be premultiplied by the small parameter δ.

The definition of the function space that we use depends on the initial data as follows.

Definition 2.2 (*Function Space*)**.** We define the following Banach space:

$$\Xi := \Big\{ u : Q_T \to \mathbb{R}^M \big| u \in (L^2(0,T;H^1(\Omega)))^M,\ \partial_t u \in (X')^M, \text{ and, for } i \in \{1,\dots,M\}, \\ u_i \ge 0 \text{ a.e. in } Q_T, \quad \int_\Omega u_i(t,x)\,\mathrm{d}x = \int_\Omega u_{i,0}(x)\,\mathrm{d}x \text{ a.e. } t\in[0,T] \Big\},$$

where

$$X := L^r(0,T;W^{1,r}(\Omega)), \qquad X' = L^{r'}(0,T;(W^{1,r}(\Omega))'),$$

with $r := 2(d+1)$ and $r' = (2d+2)/(2d+1)$.

Definition 2.3 (*Weak Solution*)**.** Fix an arbitrary $T > 0$. Given the non-negative functions $(u_{i,0})_{i=1}^M$ belonging to $L^p(\Omega)$ for $p > 1$, we say that the vector-valued function $u = (u_i)_{i=1}^M \in \Xi$, is a weak solution of (6) if:

1. for any test function $\phi \in C^1(\bar{Q}_T)$ and, for each $i \in \{1,\dots,M\}$, there holds

$$\langle \partial_t u_i, \phi\rangle_{X'\times X} + \int_{Q_T} u_i(\nabla u_i + \nabla L_i(t,x,u_i) + \delta F_i(t,x,u,\nabla u)) \cdot \nabla\phi\,\mathrm{d}x\,\mathrm{d}t = 0; \tag{9}$$

2. for each $i \in \{1,\dots,M\}$, the function u_i is non-negative a.e. in Q_T and conserves its initial mass, i.e.,

$$\int_\Omega u_i(t,x)\,\mathrm{d}x = \int_\Omega u_{i,0}(x)\,\mathrm{d}x \qquad \text{a.e. } t \in (0,T); \tag{10}$$

3. for each $i \in \{1,\dots,M\}$, we have $u_i \in C([0,T];(W^{1,r}(\Omega))')$ (*cf.* Remark 2.5) and the initial datum is satisfied in the $(W^{1,r}(\Omega))'$ sense, i.e., $\lim_{t\to 0^+} \|u_i(t,\cdot) - u_{i,0}\|_{(W^{1,r}(\Omega))'} = 0$.

Remark 2.4. Observe that the second condition in Definition 2.3 implies that any weak solution $u = (u_i)_{i=1}^M$, for $i \in \{1,\dots,M\}$, belongs to $L^\infty(0,T;L^1(\Omega))$ and $\|u_i(t,\cdot)\|_{L^1(\Omega)} = \int_\Omega u_i(t,x)\,\mathrm{d}x = \int_\Omega u_{i,0}(x)\,\mathrm{d}x$, $\forall t \in [0,T]$.

Remark 2.5. Observe that $\Xi \subset (C([0,T];(W^{1,r}(\Omega))'))^M$. Indeed, let $u \in \Xi$ and $\varphi \in X$ with $\|\varphi\|_X \le 1$ be arbitrary. Then, using the Hölder inequality, for any $i \in \{1,\dots,M\}$,

$$\left| \int_{Q_T} u_i \varphi\,\mathrm{d}x\,\mathrm{d}t \right| \le \|u_i\|_{L^2(Q_T)} (|\Omega| T)^{\frac{d}{r}} \|\varphi\|_X.$$

Thus, $\|u_i\|_{X'} \le \|u_i\|_{L^2(Q_T)}(|\Omega|T)^{\frac{d}{r}}$. Meanwhile, we also have $\partial_t u_i \in X'$ by the definition of Ξ, and it therefore follows that u_i belongs to $W^{1,r'}(0,T;(W^{1,r}(\Omega))')$. By [62, Theorem 2 of Section 5.9.2], it follows that $u_i \in C([0,T];(W^{1,r}(\Omega))')$ for every $i \in \{1,\dots,M\}$.

We identify

$$\langle \partial_t u_i, \phi \rangle_{X' \times X} = \int_0^T \langle \partial_t u_i(t,\cdot), \phi(t,\cdot) \rangle_{\Omega} \, \mathrm{d}t, \qquad \forall \phi \in C^1(\bar{Q}_T), \tag{11}$$

where $\langle \cdot, \cdot \rangle_{\Omega}$ is the duality product of $W^{1,r}(\Omega)$. Moreover, given $u = (u_i)_{i=1}^M \in \Xi$, the weak formulation (9) is equivalent to

$$\int_{Q_T} \big(-u_i \partial_t \phi + u_i (\nabla u_i + \nabla L_i(t,x,u_i) + \delta F_i(t,x,u,\nabla u)) \cdot \nabla \phi \big) \, \mathrm{d}x \, \mathrm{d}t = 0,$$

for $i \in \{1, \ldots, M\}$, for any $\phi \in C^1(\bar{Q}_T)$ with $\phi(0,\cdot) = \phi(T,\cdot) = 0$ in $\bar{\Omega}$; see Lemma 4.13.

Our main result is the following.

Theorem 1. *Fix $T > 0$ and let $(u_{i,0})_{i=1}^M$ be non-negative functions belonging to $L^p(\Omega)$ for $p > 1$, and $\delta \in \mathbb{R}$ be such that*

$$|\delta| < \frac{1}{\sqrt{\alpha C_{\Omega}} C_F}, \tag{12}$$

where C_F is specified in (5), α depends only on Ω and the smoothing operator (103), and C_{Ω} is given in (35). Then there exists a weak solution $u = (u_i)_{i=1}^M$ of (6), in the sense of Definition 2.3. Moreover, there exists a positive constant $C = C(\Omega, T, d, \delta)$, prescribed by

$$C = C_{\Omega} \big(1 - \delta^2 C_F^2 C_{\Omega}\big)^{-1}, \tag{13}$$

such that, for $i \in \{1, \ldots, M\}$,

$$\|u_i\|^2_{L^2(0,T;H^1(\Omega))} \le C \left(1 + \|u_{i,0}\|^2_{L^1(\Omega)} + \int_{\Omega} u_{i,0} \log u_{i,0} \, \mathrm{d}x \right), \tag{14}$$

and there exists another positive constant $C' = C'(\Omega, T, d, \delta)$, such that, for $i \in \{1, \ldots, M\}$,

$$\|\partial_t u_i\|_{X'} \le C' \left(1 + \|u_{i,0}\|^2_{L^1(\Omega)} + \int_{\Omega} u_{i,0} \log u_{i,0} \, \mathrm{d}x \right). \tag{15}$$

Remark 2.6 (*How Small Does δ Need to Be?*)**.** Notice that the smallness condition (12) is not particularly restrictive since it does not depend on the initial datum, but only on L, F, and on the geometry of the domain. In turn, to extend the solution beyond the final time T, it suffices to pose a new initial boundary value problem with initial condition $u(T,\cdot)$. Some explicit examples are presented in Section 3.2.

2.2. *Strategy*

We summarise the strategy for the proof of Theorem 1 as follows:

- *Weak solution for regularised frozen system* [Section 4.1]: We consider a decoupled, regularised system with unknown z instead of u to distinguish it from the solution of the original coupled system. The decoupled system, namely (26), is obtained by "freezing" the cross-diffusion terms. In particular, we replace the unknown vector-valued function z with a given function $\bar{z}$ and, eventually, we shall identify z and $\bar{z}$ via a fixed point argument. We study solutions $z \in (C^{2,1}(\bar{Q}_T))^M$ according to Definition 4.3. Existence, uniqueness, non-negativity and mass preservation of the solutions are shown in Lemma 4.8.
- *Uniform estimates and uniqueness for regularised frozen system* [Sections 4.2 and 4.3]: In Lemmas 4.9 and 4.11, we derive uniform estimates with respect to the regularisation parameters for solutions of the regularised system. We obtain H^1-type bounds for z and bounds in a dual Sobolev space for $\partial_t z$. Uniqueness in Ξ is obtained by introducing a suitable dual problem; this is a weak–strong uniqueness result.

- *Weak compactness of the solution map* [Section 5.1]: We construct a solution operator S for the regularised system (26) associating z to $\bar{z}$ as in (61). In Lemma 5.4, we show that the map S, composed with a suitable regularising operator (62), is sequentially weakly compact in a suitable Sobolev space.
- *Strong compactness of the solution map* [Section 5.2]: In Lemma 5.9 we improve the compactness result and show that the solution map is strongly compact in $\varXi$. To do this, we exploit the lower semicontinuity of the Fisher information and apply the div–curl Lemma.
- *Vanishing diffusivity* [Section 6]: Thanks to a variant of Schauder's Fixed Point Theorem, in Proposition 5.10, we obtain existence of solutions of the coupled system (60), which corresponds to original system (6) with artificial diffusivity. Finally, we let the diffusivity vanish and prove Theorem 1.

3. Formal model derivation and numerical investigation

3.1. Formal derivation

In this subsection we sketch a formal derivation of (1) starting from the interacting particle system with M species, each composed of N_i identical particles, $i = 1, \dots, M$. To simplify the presentation, the derivation is shown for the case for $L_i(t, x, u_i) \equiv V_i(x)$, that is, we drop the nonlocal interactions W_i and the time-dependence, and a simple cross-diffusion term $F_i(t, x, u, \nabla u) \equiv \sum_{j=1, j\neq i}^{M} \nabla u_j$. The addition of nonlocal interactions and time is straightforward. We denote by N the total number of particles, $N = \sum_i N_i$. We consider the following model:

$$\mathrm{d}X_k^i(t) = -\nabla V_i(X_k^i)\mathrm{d}t - \sum_{j=1}^{M} \sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^{N_i} \nabla K_{ij}(X_k^i - X_\ell^j)\mathrm{d}t, \tag{16a}$$

$$X_k^i(0) = \xi_k^i, \qquad k = 1, \dots, N_i, \tag{16b}$$

where $X_k^i(t)$ is the position of the kth particle in the ith species at time t, evolving in a bounded domain $\Omega \subset \mathbb{R}^d$ such that $|\Omega| = 1$. Particles are initialised with $\xi_1^i, \dots, \xi_{N_i}^i$ independent and identically distributed random variables with the common probability density function $u_{i,0}$. Here K_{ii} denotes the self-interaction potential in species i, and K_{ij} denotes the cross-interaction potential between species i and j. Note that K_{ij} and K_{ji} may be different to represent an asymmetric interaction between the two species. The potentials are assumed to be obtained from some fixed function K_0 by the scaling

$$K_{ij}(x) = \chi_{ij} K_0\left(\frac{|x|}{\varepsilon_{ij}}\right), \tag{17}$$

where the parameters χ_{ij} and ε_{ij} represent the strength and the range of the interactions respectively, and depend on N_j in a way that will be made specific later on. The scale-free potential $K_0 : \mathbb{R}^d \to \mathbb{R}$ is a radial, nonnegative function whose gradient is locally Lipschitz outside the origin. Moreover, it is assumed that $\|K_0\|_{L^1} < \infty$. Without loss of generality, we set $\|K_0\|_{L^1} = 1$.

Depending on χ and ε, one expects different limit equations [63]. For example, when the interactions are long-range ($\varepsilon \sim 1$) and weak ($\chi \sim N^{-1}$), then one particle interacts on average with an order N particles as $N \to \infty$ one recovers a mean-field limit for weakly interacting particles. In contrast, the case of moderately interacting particles corresponds to stronger but more localised interactions, so one particle interacts with fewer particles. As a result, one expects interactions to emerge as local terms in the limit equation.

We define the total interaction potential of the ith species as

$$K_i(\vec{x}) = \sum_{k=1}^{N_i}\left[\sum_{\ell>k}^{N_i} K_{ii}(x_k^i - x_\ell^i) + \sum_{\substack{j=1\\j\neq i}}^{M}\sum_{\ell=1}^{N_j} K_{ij}(x_k^i - x_\ell^j)\right], \tag{18}$$

where $\vec{x} = (x_k^i)_{k=1,\dots,N_i, i=1,\dots,M}$. Then the joint probability density $P_N(\vec{x}, t) = \mathrm{Prob}(X(t) = \vec{x})$ of N particles evolving according to (16) satisfies the following equation

$$\partial_t P_N = \sum_{i=1}^{M} \sum_{k=1}^{N_i} \nabla_{x_k^i} \cdot \left[\nabla V_i(x_k^i) P_N + \nabla_{x_k^i} K_i(\vec{x}) P_N\right], \qquad \vec{x} \in \Omega^N, t > 0, \tag{19a}$$

together with boundary conditions

$$\nu \cdot \left[\nabla V_i(x_k^i) P_N + \nabla_{x_k^i} K_i(\vec{x}) P_N\right] = 0, \qquad x_k^i \in \partial\Omega, t > 0, \tag{19b}$$

for $k = 1, \dots, N_i$ and $i = 1, \dots, M$, where ν is the outward normal on $\partial\Omega$ and the other coordinates are in Ω.

We consider the one-particle densities for each species as

$$u_i(x, t) = \int_{\Omega^N} P_N(\vec{x}, t) \delta(x_1^i - x) \mathrm{d}\vec{x}, \tag{20}$$

where we note that the choice of x_1^i is unimportant (since within a subpopulation, particles are indistinguishable). To obtain the equation for $u_1(x)$, we integrate (19a) over all particle positions except one particle in the first species and use the boundary conditions (19b):

$$\partial_t u_1 = \mathrm{div}\left[\nabla V_1(x) u_1 + G_1(x, t)\right], \tag{21}$$

where

$$\begin{aligned} G_1(x, t) &= \int_{\Omega^N} \nabla_{x_1^1} K_1(\vec{x}) P_N \delta(x_1^i - x) \mathrm{d}\vec{x} \\ &= \int_{\Omega^N} \left[\sum_{\ell=2}^{N_1} \nabla_x K_{11}(x - x_\ell^1) P_N + \sum_{j=2}^{M} \sum_{\ell=1}^{N_j} \nabla_x K_{1j}(x - x_\ell^j) P_N\right] \mathrm{d}\vec{x} \\ &= (N_1 - 1) \int_{\Omega} \nabla_x K_{11}(x - y) P_2^{11}(x, y, t) \mathrm{d}y + \sum_{j=2}^{M} N_j \int_{\Omega} \nabla_x K_{1j}(x - y) P_2^{1j}(x, y, t) \mathrm{d}y. \end{aligned}$$

Here $P_2^{1j}, j = 1, \dots, M$ stands for the following two-particle density

$$P_2^{1j}(x, y, t) = \int_{\Omega^N} P_N(\vec{x}, t) \delta(x_1^1 - x) \delta(x_2^j - y) \mathrm{d}\vec{x}.$$

Oelschläger [64] proved propagation of chaos (meaning that any fixed number of particles remains approximately independent in time despite the interaction) for the single-species cases similar to (16) under quite restrictive initial conditions. These conditions were relaxed by Philipowski [65] using regularising Brownian motions, that is, adding terms $\epsilon \mathrm{d}B_k^i$ in (16a) and taking $\epsilon \to 0$ at a suitable rate depending on N, ε. So including such a term would make sense when considering a rigorous derivation. An alternative approach taken by [66,67] in the multiple species case is to include the interactions between particles in the diffusion term (note that, in their case, the mean-field limit model still contains linear diffusion terms). Concerning Shigesada–Kawasaki–Teramoto cross-diffusion system, a rigorous derivation has been obtained in the papers [68,69]. For our purposes, here, we simply assume that an analogous propagation of chaos for P_N exists. In particular, this means that the two-particle marginals may be approximated by

$$P_2^{11}(x, y, t) = u_1(x, t) u_1(y, t), \qquad P_2^{1j}(x, y, t) = u_1(x, t) u_j(y, t).$$

Using these expressions in G_1, it reduces to

$$G_1(x, t) = (N_1 - 1) u_1 \nabla(K_{11} * u_1) + \sum_{j=2}^{M} N_j u_1 \nabla(K_{1j} * u_j).$$

Finally, if we consider the scaling (17) with $\varepsilon_{11}, \varepsilon_{1j} \ll 1$, we can localise the convolution terms and arrive at

$$G_1(x,t) = (N_1 - 1)\varepsilon_{11}^d \chi_{11} u_1 \nabla u_1 + \sum_{j=2}^{M} N_j \varepsilon_{1j}^d \chi_{1j} u_1 \nabla u_j, \tag{22}$$

using that $\|K_0\|_{L^1} = 1$. The analogous calculation can be done for any of the other species to obtain $G_i(x,t)$ for $i = 1, \dots, M$. Now we can determine the suitable scaling for interactions that lead to the structure in (1). Namely, we set the strengths to be

$$\chi_{ii} = \frac{1}{(N_i - 1)\varepsilon_{ii}^d}, \qquad \chi_{ij} = \frac{\delta}{N_j \varepsilon_{ij}^d},$$

with $\delta \ll 1$, and we let $N_i \to \infty, \varepsilon_{ij} \to 0$ in such a way that $N_j \gg 1/\varepsilon_{ij}$. Using this and combining (21) and (22), we arrive at

$$\partial_t u_i = \operatorname{div}\Big[u_i\Big(\nabla u_i + \nabla V_i(x) + \delta \sum_{j=1, j\neq i}^{M} \nabla u_j\Big)\Big]. \tag{23}$$

3.2. *Numerical simulations*

In this subsection we present numerical simulations of (1) with two species ($M = 2$) in one dimension ($d = 1$). We consider the case with $L_i(t,x,u_i) = V_i(x)$ and $F_i(t,x,u,\nabla u) \equiv u_j$, leading to the following system of equations:

$$\partial_t u_1 = \operatorname{div}\left[u_1\left(\nabla u_1 + \nabla V_1(x) + \delta \nabla u_2\right)\right], \tag{24a}$$

$$\partial_t u_2 = \operatorname{div}\left[u_2\left(\nabla u_2 + \nabla V_2(x) + \delta \nabla u_1\right)\right]. \tag{24b}$$

Throughout this section we consider the domain $\Omega = [-1,1]$ with no-flux boundary conditions, and initial conditions $u_1(0,x) = u_{1,0}(x)$ and $u_2(0,x) = u_{2,0}(x)$ with unit mass. We solve (24) using the positivity-preserving finite-volume scheme presented in [39], which is first order in space and time. We use $J = 64$ grid points in space and a fixed timestep $\Delta t = 10^{-6}$.

We consider what the bound (12) on δ is for our particular examples. For our choice of cross-term F_i, we have that $C_F = 1$ (see (5)) and C_Ω in (35) simplifies to

$$C_\Omega(T) = 2\max\left\{(1 + C_P), \frac{T}{2} + 4(1 + C_P)\big(e^{-1} + 2TC_L^2\big)\right\},$$

with Poincaré constant $C_P = (2/\pi)^2$ and $C_L = \max_i \|V_i\|_{L^2}$. For the purposes of the numerical simulation, it is convenient to consider $C_\Omega(\Delta t)$. In the limit of $\Delta t \to 0$, we have

$$C_\Omega(0) = 8e^{-1}(1 + C_P) \approx 4.136,$$

independent of the external potentials L_i. Therefore, the upper bound on δ is given by

$$\delta < \delta_{\max} = 1/\sqrt{C_\Omega(0)} \approx 0.492.$$

Below we numerically investigate the behaviour beyond such value, and close to the critical value $\delta = 1$.

Example 1 (*Left and Right Initial Conditions*). In the first example we consider the initial conditions

$$u_{1,0}(x) = C_1[(x + 0.5)(-0.9 - x)]_+, \qquad u_{2,0}(x) = C_2[(x - 0.5)(0.9 - x)]_+,$$

where C_1, C_2 are such that the initial densities are normalised to unit mass. We consider the external potentials $V_1 = 0$ and $V_2(x) = 2x^2$, and four different values of $\delta = 0.4, 0.6, 0.8, 0.99$. For this choice of potentials, we have $C_L = 6$ and

$$\delta_{\max}(T) = 0.0203, \quad \delta_{\max}(\varDelta t) = 0.492,$$

where the upper bound on δ is $\delta_{\max}(t) = 1/\sqrt{C_\varOmega(t)}$.

In the left column of Fig. 1 we show the time evolution at ten equally spaced times t_k between 0 and $T = 3$ of u_1 and u_2 (solid blue and red lines, respectively) as well as the corresponding steady states $u_{1,\infty}(x)$ and $u_{2,\infty}(x)$ (dashed green and purple lines respectively), which are computed as the minimisers of the energy

$$E[u_1, u_2](t) = \int_\varOmega \left[\frac{1}{2}u_1^2 + \frac{1}{2}u_2^2 + \delta u_1 u_2 + V_1 u_1 + V_2 u_2\right] \mathrm{d}x. \tag{25}$$

As we increase δ closer to one (the value at which E stops being strictly convex), we observe the formation of sharper interface between the two components. For the smallest value of δ, $\delta = 0.5$, there is no “vacuum region” for u_1 due to u_2 (that is, $\operatorname{supp} u_1 = \varOmega$) and by $t = 3$ the solution is very close to the steady state. Increasing δ changes this: for larger δ, the stationary solution u_1 has a vacuum region in the middle of the domain in which $u_1 = 0$ to numerical precision (which grows closer to $\operatorname{supp} u_2$ as δ approaches one). This vacuum region implies that it takes much longer for half of the mass of u_1 to transfer from the left to the right on the domain, implying that the equilibration to the stationary state is much slower (this can be seen in the bottom row, where the final time solution $u_1(T, x)$ is still very far from the steady-state minimiser $u_{1,\infty}$).

In the right column of Fig. 1 we plot the time evolution of the spatial L^2 norms of u_i and ∇u_i, as well as the Total Variation (TV), all computed using the partition given by the spatial grid used in the finite-volume scheme. The key point to note is that the effect of increasing δ is noticed markedly by the semi-norm $\|\nabla u_i\|_{L^2}$, whereas the other two norms, $\|u_i\|_{L^2}$ and $\|u_i\|_{TV}$, remain mostly unchanged by δ.

Example 2 (*Uniform Initial Conditions*)**.** Here, we consider the same set-up as in the previous example, except that now both components start with uniform initial conditions

$$u_{1,0}(x) = u_{2,0}(x) = 1/|\varOmega|.$$

Therefore, we expect the same stationary states (since for $\delta < 1$, E is strictly convex). We show the results of this example in Fig. 2. Because of the symmetry in the initial conditions, in this case the convergence to the steady state is much faster, as there is no mass that has to “cross” through the vacuum region as the latter is formed.

Example 3 (*Stronger External Potentials*)**.** We now consider the left and right initial conditions as in Example 1 while changing the external potentials to $V_1(x) = x^2/2$ and $V_2(x) = 50x^2$. For this choice, we have that

$$\delta_{\max}(T) = 0.00126, \quad \delta_{\max}(\varDelta t) = 0.480,$$

that is, a ten-fold reduction in $\delta_{\max}(T)$ with respect to Example 1 but a barely noticeable change in $\delta_{\max}(\varDelta t)$ (as expected given that $\delta_{\max}$ is independent of L_i in the small time limit). The stronger confinement potential in the second species leads to a vacuum region in the first species for smaller values of δ than in the previous examples, and the associated slower convergence (see Fig. 3).

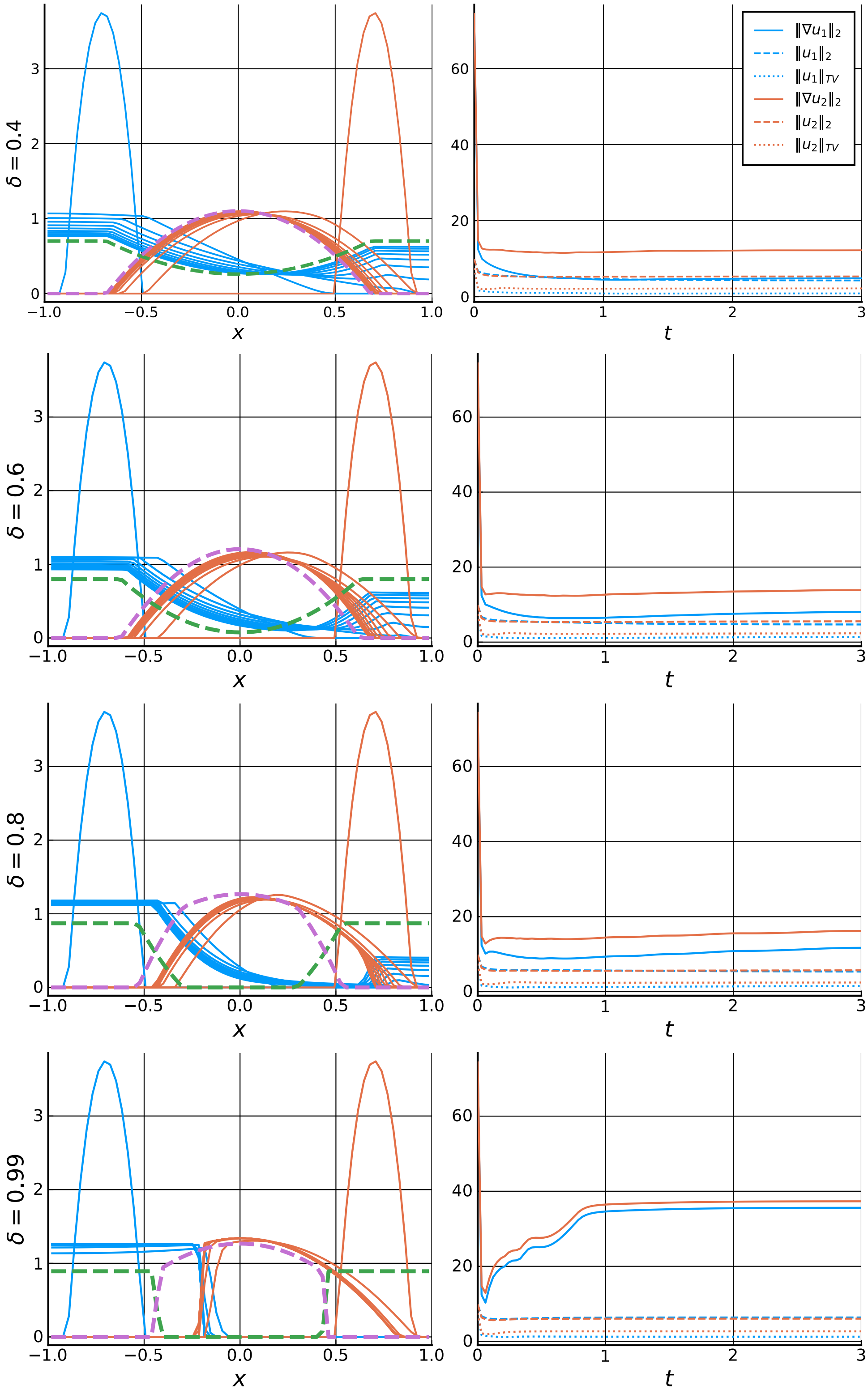


Fig. 1. Time evolution with left and right initial conditions, $V_1 = 0$ and $V_2 = 2x^2$ and final time $T = 3$, for various values of δ (Example 1). The left column shows ten equally spaced timepoints in $[0, T]$ in solid lines, and the stationary stated in dashed lines. The right column shows the temporal evolution of three norms.

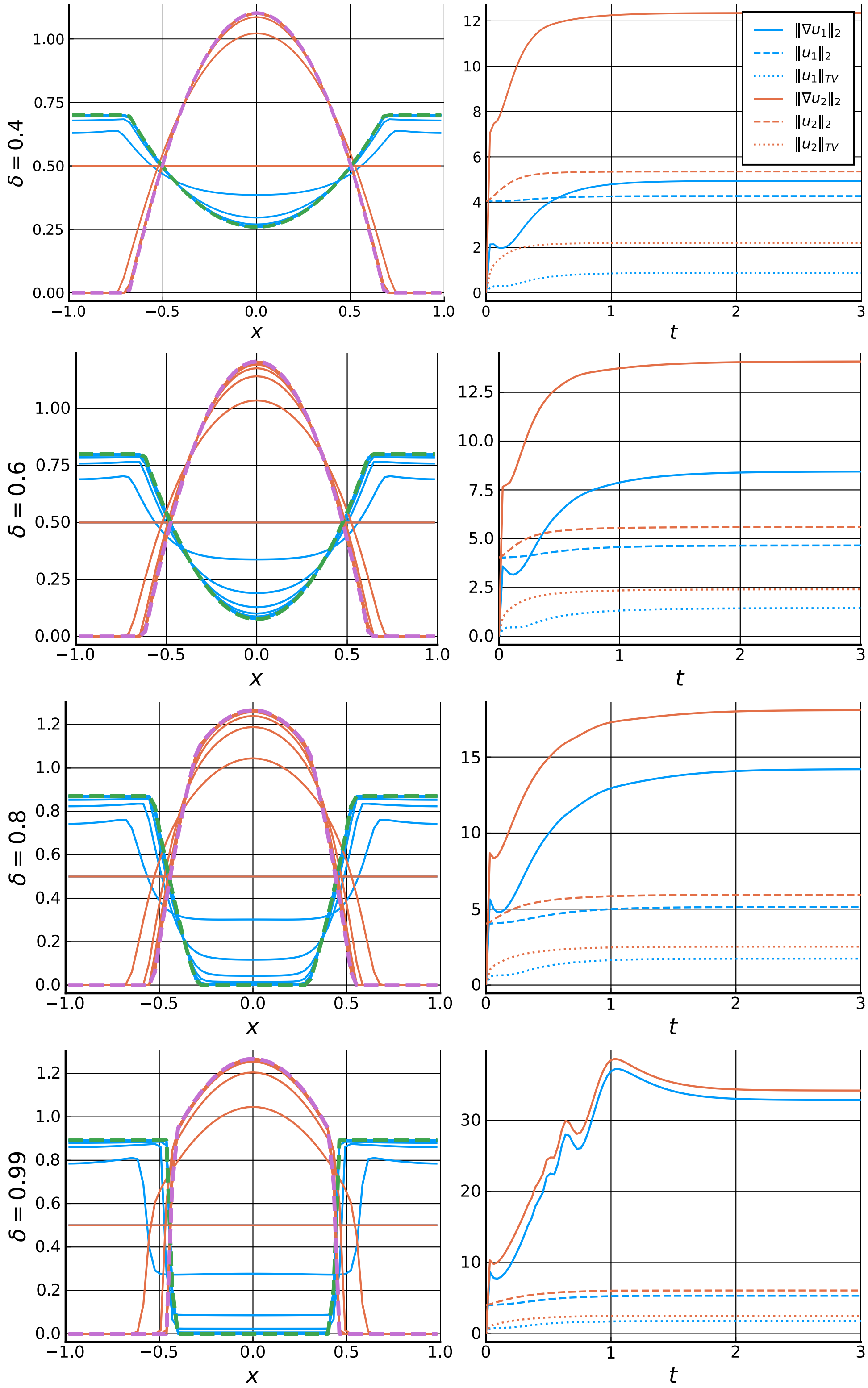


Fig. 2. Time evolution with uniform initial conditions, $V_1 = 0$ and $V_2 = 2x^2$ and final time $T = 3$, for various values of δ (Example 2). The left column shows ten equally spaced timepoints in $[0, T]$ in solid lines, and the stationary stated in dashed lines. The right column shows the temporal evolution of three norms.

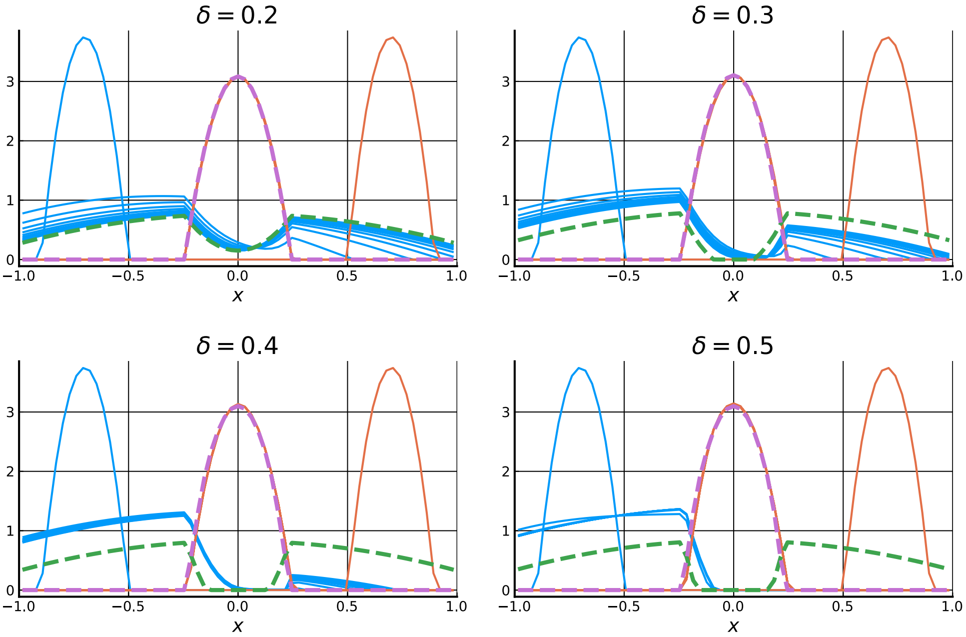


Fig. 3. Time evolution with uniform initial conditions, $V_1(x) = x^2/2$ and $V_2(x) = 50x^2$ and final time $T = 3$, for various values of δ (Example 3). There are ten equally spaced timepoints in $[0, T]$ in solid lines, and stationary stated in dashed lines.

Example 4 (*Evolution of Norms in Time and Space as a Function of* δ)**.** In this final example, we look at the evolution of the norms in δ instead of time. To this end, we consider the following integrated-in-time norms

$$\|u\|_{2,T} = \left\{ \sum_{t_k=0}^{T} \left[\|u_1(t_k,\cdot)\|_{L^2}^2 + \|u_2(t_k,\cdot)\|_{L^2}^2 \right] \right\}^{1/2},$$

$$\|\nabla u\|_{2,T} = \left\{ \sum_{t_k=0}^{T} \left[\|\nabla u_1(t_k,\cdot)\|_{L^2}^2 + \|\nabla u_2(t_k,\cdot)\|_{L^2}^2 \right] \right\}^{1/2},$$

$$\|u\|_{TV,T} = \left\{ \sum_{t_k=0}^{T} \left[\|u_1(t_k,\cdot)\|_{TV} + \|u_2(t_k,\cdot)\|_{TV} \right] \right\}^{1/2}.$$

We use uniform initial conditions, a final time $T = 5$, and values for $\delta = 0, 0.1, \ldots, 0.9, 0.95, 0.99$. We show the evolution of the three norms in Fig. 4 for two cases: first, for the potentials used in Examples 1 and 2, namely $V_1 = 0$ and $V_2 = 2x^2$; and second, for $V_1 = x^2/2$ and $V_2 = 50x^2$. In the latter, the combination of external potentials makes the interface between the two components sharper (since the second component has a very strong confining potential, but also the first component now wants to concentrate around the origin). This fact is clearly visible in the trend of $\|\nabla u\|_{2,T}$ for increasing δ. In contrast, the TV norm remains unchanged. As mentioned in the introduction, this observation indicates that a smallness assumption on δ is necessary to keep to the functional framework of $L^2(0, T; H^1(\Omega))$ for the analysis. The plot also suggests one can hope for a more general existence theory for solutions belonging to the space BV when δ is close to one.

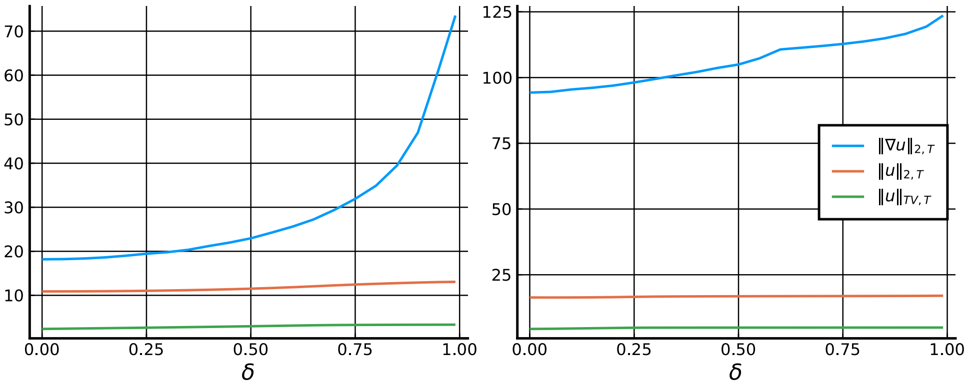


Fig. 4. Norms (rescaled by $T|\Omega|$) in time and space as a function of δ with uniform initial conditions and final time $T = 3$. (Left) $V_1 = 0$ and $V_2 = 2x^2$. (Right) $V_1 = x^2/2$ and $V_2 = 50x^2$.

4. Regularised "frozen" system

We introduce below the regularised system with frozen cross-diffusion. Let $\bar{z} = (\bar{z}_i)_{i=1}^M \in (C^\infty(\bar{Q}_T))^M$ be a given vector function. Throughout this section, we denote by $z = (z_i)_{i=1}^M$ the solution of the *regularised frozen system*

$$\begin{cases} \partial_t z_i = \operatorname{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,\bar{z},\nabla\bar{z})) + \varepsilon \nabla z_i] & \text{in } Q_T, \\ 0 = \nu \cdot [z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,\bar{z},\nabla\bar{z})) + \varepsilon \nabla z_i] & \text{on } \Sigma_T, \\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega. \end{cases} \tag{26}$$

Remark 4.1. The constant $\varepsilon > 0$ and the vector of non-negative functions $z_0 = (z_{i,0})_{i=1}^M \in (C_c^\infty(\Omega))^M$ do not change throughout the present section and Section 5. The initial functions $z_{i,0}$ are chosen such that $\int_\Omega z_{i,0}\,\mathrm{d}x = \int_\Omega u_{i,0}\,\mathrm{d}x$ for $i \in \{1,\dots,M\}$.

In the following subsection, we introduce the definition of weak solution to the above regularised frozen system, and show the mass conservation and non-negativity for such weak solutions. We then prove the existence of these solutions and deduce from their regularity that they satisfy the system of equations in the classical sense. Then we prove some Sobolev estimates independent of ε and conclude with a uniqueness result. Below we will use the shorthand $\bar{F}_i$ to refer to the function

$$\bar{F}_i(t,x) := F_i(t,x,\bar{z}(t,x),\nabla\bar{z}(t,x)) \qquad \forall (t,x) \in \bar{Q}_T. \tag{27}$$

Remark 4.2. In view of the condition that $\{G_i^0\}_{i=1}^M$ and $\{G_{ij}^1\}_{i,j=1}^M$ in (4) be C^2-regular and uniformly bounded with respect to all arguments (see Section 2.1), for each $\bar{z} = (\bar{z}_i)_{i=1}^M \in (C^\infty(\bar{Q}_T))^M$ fixed, there exists a positive constant $\Lambda_{\bar{z}}$, where

$$\Lambda_{\bar{z}} = \Lambda_{\bar{z}}\left(\max_{1\le i\le M}\|\bar{z}_i\|_{C^2(\bar{Q}_T)}, \max_{1\le i\le M}\|F_i\|_{C^2(\bar{Q}_T\times\mathbb{R}^M\times\mathbb{R}^{Md})}\right),$$

such that, for all $(t,x) \in \bar{Q}_T$ and every $i \in \{1,\dots,M\}$,

$$|\bar{F}_i(t,x)| + \sum_{j=1}^d \left|\frac{\partial \bar{F}_i}{\partial x_j}(t,x)\right| + \left|\frac{\partial \bar{F}_i}{\partial t}(t,x)\right| + \sum_{j=1}^d \left|\frac{\partial^2 \bar{F}_i}{\partial t \partial x_j}(t,x)\right| + \sum_{k=1}^d\sum_{j=1}^d \left|\frac{\partial^2 \bar{F}_i}{\partial x_k \partial x_j}(t,x)\right| \le \Lambda_{\bar{z}}. \tag{28}$$

Additionally, there exists a positive constant Λ_0 depending only on $(z_{i,0})_{i=1}^M$ such that

$$\|z_{i,0}\|_{C^2(\bar{\Omega})} \le \Lambda_0. \tag{29}$$

4.1. Definition and existence of regularised solutions

Per the definition of weak solution given in [59, Section 5.7], we provide the following notion of solution to the regularised frozen problem.

Definition 4.3 (*Weak Solution for Regularised Frozen System*)**.** We say that $z \in (C^{2,1}(\bar{Q}_T))^M$ *solves the weak form* of (26) if, for any test function $\phi \in C^1(\bar{Q}_T)$, for $i \in \{1, \dots, M\}$, for $t \in [0, T]$,

$$\begin{aligned}\int_\Omega z_i(t,x)\phi(t,x)\,\mathrm{d}x - \int_\Omega z_{i,0}(x)\phi(0,x)\,\mathrm{d}x - \int_0^t \int_\Omega z_i \partial_t \phi \,\mathrm{d}x\,\mathrm{d}t \\ + \int_0^t \int_\Omega \left[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,\bar{z},\nabla\bar{z})) + \varepsilon \nabla z_i\right] \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t = 0.\end{aligned} \tag{30}$$

Correspondingly, we define $S_\varepsilon : \bar{z} \to z$ to be the *solution operator*, whose image is the weak solution of (26).

Remark 4.4. Note that the following compatibility condition has been implicitly imposed in the previous weak formulation,

$$0 = \nu \cdot \left[z_{i,0}(\nabla z_{i,0} + \nabla L_i(0,x,z_{i,0}) + \delta F_i(0,x,\bar{z}(0,x),\nabla\bar{z}(0,x))) + \varepsilon \nabla z_{i,0}\right] \qquad \text{on } \partial\Omega, \tag{31}$$

which is manifestly satisfied for all choices of $\bar{z} \in (C^\infty(\bar{Q}_T))^M$, as the fixed initial data $z_{i,0}$ is identically zero on the boundary $\partial\Omega$ due to its compact support (see Remark 4.1).

Lemma 4.5 (*Mass Conservation for Regularised Frozen System*)**.** *Suppose there exists a weak solution z of the problem* (26) *in the sense of* Definition 4.3*. Then, for $i \in \{1, \dots, M\}$,*

$$\int_\Omega z_i(t,x)\,\mathrm{d}x = \int_\Omega z_{i,0}(x)\,\mathrm{d}x \qquad \forall t \in [0,T].$$

Proof. The assertion is immediate from using the test function $\phi = 1$ in Definition 4.3. □

Lemma 4.6 (*Sign Preservation for Regularised Frozen System*)**.** *Suppose there exists a weak solution z of the problem* (26) *in the sense of* Definition 4.3*. Then, for $i \in \{1, \dots, M\}$,*

$$z_i(t,x) \ge 0 \qquad \text{for a.e. } (t,x) \in \bar{Q}_T.$$

Proof. Define the function $\theta(t,x) := [z_i(t,x)]_-$ to be the negative part of the weak solution in question. Noting that $\theta = -z_i \mathbb{1}_{z_i \le 0}$, we observe that this function is non-negative and supported in the set $\{(t,x) \in \bar{Q}_T : z_i(t,x) \le 0\}$. Moreover, we find that

$$\nabla\theta = -\nabla z_i \mathbb{1}_{z_i \le 0}, \qquad \partial_t \theta = -\partial_t z_i \mathbb{1}_{z_i \le 0},$$

in the sense of distributions. It follows that $\theta \in L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^1(\Omega))$ and $\partial_t\theta \in X' \cap L^2(0,T;(H^1(\Omega))')$. Using standard density arguments in Sobolev spaces, we may test against θ in the weak formulation of Definition 4.3. In turn, we obtain, for a.e. $t \in (0,T)$,

$$\frac{d}{dt}\int_\Omega \frac{1}{2}\theta^2(t,x)\,\mathrm{d}x + \int_\Omega \left(\theta|\nabla\theta|^2 + \theta\nabla\theta \cdot \nabla L_i(t,x,z_i) + \delta\theta\nabla\theta \cdot F_i(t,x,\bar{z},\nabla\bar{z}) + \varepsilon|\nabla\theta|^2\right)\mathrm{d}x = 0. \tag{32}$$

Given the form of the terms $\{L_i\}_{i=1}^M$ from (7), we have

$$\left|\int_\Omega \theta\nabla\theta \cdot \nabla L_i(t,x,z_i)\,\mathrm{d}x\right| \le \left(\|\nabla V_i(t,\cdot)\|_{L^\infty(\Omega)} + \|\nabla W_i * z_i(t,\cdot)\|_{L^\infty(\Omega)}\right)\int_\Omega \theta|\nabla\theta|\,\mathrm{d}x,$$

and

$$\|\nabla W_i * z_i(t,\cdot)\|_{L^\infty(\Omega)} \le \|\nabla W_i\|_{L^\infty(\mathbb{R}^d)} \|z_i(t,\cdot)\|_{L^1(\Omega)} \qquad \text{a.e. } t \in (0,T).$$

It therefore follows, using the fact that $z_i \in L^\infty(0,T;L^1(\Omega))$ since $z_i \in C^{2,1}(\bar{Q}_T)$ as per Definition 4.3, that

$$\left| \int_\Omega \theta \nabla\theta \cdot \nabla L_i(t,x,z_i)\,\mathrm{d}x \right| \le C_L \big(1 + \|z_i\|_{L^\infty(0,T;L^1(\Omega))}\big) \int_\Omega \theta |\nabla\theta|\,\mathrm{d}x \qquad \text{a.e. } t \in (0,T).$$

Meanwhile, using the boundedness of $\bar{z}$ and that of F in C^2 to control $F_i(t,x,\bar{z},\nabla\bar{z})$ from (28) (see Remark 4.2), we obtain

$$\left| \int_\Omega \delta\theta\nabla\theta \cdot F_i(t,x,\bar{z},\nabla\bar{z})\,\mathrm{d}x \right| \le |\delta| \Lambda_{\bar{z}} \int_\Omega \theta|\nabla\theta|\,\mathrm{d}x \qquad \text{a.e. } t \in (0,T).$$

Integrating (32) with respect to the time variable, and using the previous estimates, we find

$$\int_\Omega \frac{1}{2}\theta^2(t,x)\,\mathrm{d}x + \int_0^t \int_\Omega \theta|\nabla\theta|^2\,\mathrm{d}x\,\mathrm{d}\tau + \varepsilon \int_0^t \int_\Omega |\nabla\theta|^2\,\mathrm{d}x\,\mathrm{d}\tau \le \Lambda \int_0^t \int_\Omega \theta|\nabla\theta|\,\mathrm{d}x,$$

where the positive constant

$$\Lambda := |\delta|\Lambda_{\bar{z}} + C_L\big(1 + \|z_i\|_{L^\infty(0,T;L^1(\Omega))}\big),$$

is independent of time. An application of the Cauchy–Young inequality gives

$$\int_\Omega \frac{1}{2}\theta^2(t,x)\,\mathrm{d}x + \int_0^t \int_\Omega \theta|\nabla\theta|^2\,\mathrm{d}x\,\mathrm{d}\tau + \frac{\varepsilon}{2} \int_0^t \int_\Omega |\nabla\theta|^2\,\mathrm{d}x\,\mathrm{d}\tau \le \frac{\Lambda^2}{\varepsilon} \int_0^t \int_\Omega \frac{1}{2}\theta^2\,\mathrm{d}x\,\mathrm{d}\tau.$$

Dropping the last two terms in the left-hand side of the inequality above, Grönwall's Lemma yields

$$\int_\Omega \theta^2(t,x)\,\mathrm{d}x \le \left(\int_\Omega \theta^2(0,x)\,\mathrm{d}x\right) e^{\frac{\Lambda^2}{\varepsilon}t} = 0 \qquad \text{for a.e. } t \in (0,T),$$

where the final equality follows from the non-negativity of the initial data $z_{i,0}$ (Remark 4.1). The result follows. $\square$

Remark 4.7. It is a priori not clear how to prove such a sign preservation result for the original system (6) directly from Definition 2.3, due to the presence of cross-terms of the form $\int_\Omega \nabla u_j(t,x)[u_i(t,x)]_-\,\mathrm{d}x$ with $i \neq j$. The non-negativity of the solution of the original system (6) will therefore be deduced via a limiting procedure from the non-negativity of the regularised solutions of (26).

In what follows, we apply the classical theory of Ladyzhenskaia, Solonnikov, and Ural'tseva [59, Chap. 5, Sec. 7, Thm. 7.4] to deduce the existence and uniqueness of classical solutions to the regularised system (26). The proof is given in Appendix A.1.

Lemma 4.8 (*Existence and Uniqueness of Regularised Solutions*)**.** *There exists a unique $z = (z_i)_{i=1}^M \in C^{2,1}(\bar{Q}_T)$ solving (26) as a pointwise equality between continuous functions. Moreover, for $i \in \{1,\dots,M\}$,*

$$z_i(t,x) \ge 0 \quad \text{for } (t,x) \in \bar{Q}_T, \qquad z_i(0,x) = z_{i,0}(x) \quad \text{for } x \in \bar{\Omega},$$
$$\int_\Omega z_i(t,x)\,\mathrm{d}x = \int_\Omega z_{i,0}(x)\,\mathrm{d}x \quad \text{for } t \in [0,T].$$

4.2. Uniform estimates

In this subsection, we derive the uniform estimates of the solutions of the regularised frozen system by testing the equation against the logarithm of the solution.

Lemma 4.9 (*Energy Estimates*)**.** *Let $z = (z_i)_{i=1}^M \in (C^{2,1}(\bar{Q}_T))^M$ be the solution of (26) provided by Lemma 4.8. Then, for any $t \in [0,T]$, there holds the estimate*

$$\int_\Omega z_i(t)(\log z_i(t))_+ \,\mathrm{d}x + \frac{1}{2}\int_0^t \int_\Omega |\nabla z_i|^2 \,\mathrm{d}x\,\mathrm{d}\tau + \varepsilon \int_0^t \int_\Omega \frac{|\nabla z_i|^2}{z_i} \,\mathrm{d}x\,\mathrm{d}\tau$$
$$\le (e^{-1}|\Omega| + 2|\Omega| T C_L^2) + \int_\Omega z_{i,0} \log z_{i,0} \,\mathrm{d}x + 2|\Omega| T C_L^2 \left(\int_\Omega z_{i,0} \,\mathrm{d}x\right)^2 + \int_0^T \int_\Omega \delta^2 |\bar{F}_i|^2 \,\mathrm{d}x\,\mathrm{d}\tau, \quad (33)$$

for $i \in \{1,\dots,M\}$, along with the Sobolev estimate

$$\sup_{t\in[0,T]} \int_\Omega z_i(t)(\log z_i(t))_+ \,\mathrm{d}x + \| z_i \|^2_{L^2(0,T;H^1(\Omega))}$$
$$\le C_\Omega \left(1 + \|z_{i,0}\|^2_{L^1(\Omega)} + \int_\Omega z_{i,0}\log z_{i,0} \,\mathrm{d}x + \int_0^T \int_\Omega \delta^2 |\bar{F}_i|^2 \,\mathrm{d}x\,\mathrm{d}t\right), \quad (34)$$

for $i \in \{1,\dots,M\}$. The positive constant $C_\Omega = C_\Omega(\Omega, T, d)$, which is independent of ε and $z_0 = (z_{i,0})_{i=1}^M$, is given by

$$C_\Omega = 2 \cdot \max\left\{(1 + C_P), \frac{T}{|\Omega|} + 2|\Omega|(1 + C_P)\left(e^{-1} + 2TC_L^2\right)\right\}, \quad (35)$$

where $C_P = C_P(\Omega, d)$ is the Poincaré constant.

Proof. Begin by assuming that z_i is strictly positive in Q_T and multiply (26) by $\log z_i$. By rewriting the time derivative, we get

$$\partial_t(z_i \log z_i) - \partial_t z_i = (\log z_i)\,\mathrm{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta \bar{F}_i) + \varepsilon \nabla z_i] \qquad \forall (t,x) \in Q_T.$$

Integrating in space and time then yields, for any $t \in [0,T]$,

$$\int_0^t \frac{d}{d\tau}\left(\int_\Omega z_i(\tau)\log z_i(\tau)\,\mathrm{d}x\right)\mathrm{d}\tau - \int_0^t \frac{d}{d\tau}\left(\int_\Omega z_i(\tau)\,\mathrm{d}x\right)\mathrm{d}\tau =$$
$$- \int_0^t \int_\Omega \left[\nabla z_i \cdot (\nabla z_i + \nabla L_i(t,x,z_i) + \delta \bar{F}_i) + \varepsilon \frac{|\nabla z_i|^2}{z_i}\right] \mathrm{d}x\,\mathrm{d}\tau, \quad (36)$$

where the no-flux boundary condition makes the boundary term vanish.

When z_i is merely non-negative, we multiply by $\log(\beta + z_i)$ for some $\beta > 0$ and get

$$\partial_t(z_i \log(\beta + z_i)) - (\partial_t z_i)\frac{z_i}{\beta + z_i} = (\log(\beta + z_i))\,\mathrm{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta \bar{F}_i) + \varepsilon \nabla z_i], \qquad \forall (t,x) \in Q_T.$$

Integrating in time and space, and using that $|\partial_t z_i|$ is integrable on Q_T since z_i is $C^{2,1}(\bar{Q}_T)$, the Dominated Convergence Theorem implies

$$\lim_{\beta\to 0} \int_0^t \int_\Omega (\partial_t z_i)\frac{z_i}{\beta + z_i}\,\mathrm{d}x\,\mathrm{d}\tau = \int_0^t \int_\Omega (\partial_t z_i)\,\mathrm{d}x\,\mathrm{d}\tau.$$

All other terms may be treated similarly with the exception of the Fisher information, where one obtains

$$\lim_{\beta\to 0} \int_0^t \int_\Omega \frac{|\nabla z_i|^2}{\beta + z_i}\,\mathrm{d}x\,\mathrm{d}\tau = \int_0^t \int_\Omega \frac{|\nabla z_i|^2}{z_i}\,\mathrm{d}x\,\mathrm{d}\tau,$$

by the Monotone Convergence Theorem, since the sequence of integrands $\left\{|\nabla z_i|^2/(\beta+z_i)\right\}_{\beta>0}$ is pointwise increasing as $\beta\to 0$ and non-negative by virtue of Lemma 4.8. In turn, we recover (36).

Recall from Lemma 4.8 that $\int_\Omega z_i(\tau)\,\mathrm{d}x$ is constant. Eq. (36) therefore simplifies to

$$\begin{aligned}\int_\Omega z_i(t)\log z_i(t)\,\mathrm{d}x &= \int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x - \delta\int_0^t\int_\Omega \nabla z_i\cdot\bar{F}_i\,\mathrm{d}x\,\mathrm{d}\tau - \int_0^t\int_\Omega |\nabla z_i|^2\,\mathrm{d}x\,\mathrm{d}\tau \\ &\quad - \int_0^t\int_\Omega \nabla z_i\cdot\nabla L_i(t,x,z_i)\,\mathrm{d}x\,\mathrm{d}\tau - \varepsilon\int_0^t\int_\Omega \frac{|\nabla z_i|^2}{z_i}\,\mathrm{d}x\,\mathrm{d}\tau,\end{aligned} \tag{37}$$

for any $t\in[0,T]$. Observe also that $z\log z = z\log z\mathbb{1}_{\{z\geq 1\}} + z\log z\mathbb{1}_{\{0\leq z<1\}}$, $\forall z\in[0,\infty)$, and, since $x\mapsto x\log x$ is non-positive over the interval $[0,1]$ and achieves its minimum (with value $-e^{-1}$) at the point $x=e^{-1}$, it follows that

$$z\log z \geq z\log z\mathbb{1}_{\{z\geq 1\}} - e^{-1} \qquad \forall z\in[0,\infty). \tag{38}$$

Also, using (7), (8), and an application of the triangle inequality followed by Hölder's inequality yields

$$\begin{aligned}\left|\int_0^t\int_\Omega \nabla z_i\cdot\nabla L_i(t,x,z_i)\,\mathrm{d}x\,\mathrm{d}\tau\right| &\leq \|\nabla V_i\|_{L^2(Q_T)}\|\nabla z_i\|_{L^2(Q_T)} \\ &\quad + \|\nabla z_i\|_{L^2(Q_T)}\left(\int_0^t\int_\Omega |\nabla W_i * z_i(\tau,\cdot)(x)|^2\,\mathrm{d}x\,\mathrm{d}\tau\right)^{\frac{1}{2}} \\ &\leq (|\Omega|T)^{\frac{1}{2}}C_L\left(1+\int_\Omega z_{i,0}(y)\,\mathrm{d}y\right)\|\nabla z_i\|_{L^2(Q_T)},\end{aligned} \tag{39}$$

where we bounded the convolution term as follows

$$\begin{aligned}\int_0^t\int_\Omega |\nabla W_i * z_i(\tau,\cdot)(x)|^2\,\mathrm{d}x\,\mathrm{d}\tau &= \int_0^t\int_\Omega\left|\int_\Omega \nabla W_i(x-y)z_i(\tau,y)\,\mathrm{d}y\right|^2\mathrm{d}x\,\mathrm{d}\tau \\ &\leq C_L^2\int_0^t\int_\Omega\left(\int_\Omega z_i(\tau,y)\,\mathrm{d}y\right)^2\mathrm{d}x\,\mathrm{d}\tau = C_L^2|\Omega|T\left(\int_\Omega z_{i,0}(y)\,\mathrm{d}y\right)^2,\end{aligned} \tag{40}$$

where we used the boundedness of $\{\nabla W_i\}_{i=1}^M$ inherited from (8), the non-negativity of z_i due to Lemma 4.6, and the fact that $x\mapsto x^2$ is increasing on $[0,\infty)$ to obtain the inequality, and the mass conservation from Lemma 4.5 to obtain the final equality. Using estimates (38) and (39), along with an application of the weighted Cauchy–Young inequality to the terms on the right-hand side of (39) and to $\int_0^t\int_\Omega \nabla z_i\cdot\bar{F}_i\,\mathrm{d}x\,\mathrm{d}t$, we deduce the estimate (33) from (37).

The Poincaré–Wirtinger inequality

$$\int_\Omega\left(z_i - \frac{1}{|\Omega|}\int_\Omega z_i\,\mathrm{d}x\right)^2\mathrm{d}x \leq C_P\int_\Omega |\nabla z_i|^2\,\mathrm{d}x,$$

where $C_P = C_P(\Omega,d)$ is the Poincaré constant, implies

$$\begin{aligned}C_P\int_\Omega |\nabla z_i|^2\,\mathrm{d}x &\geq \int_\Omega\left(z_i - \frac{1}{|\Omega|}\int_\Omega z_i\,\mathrm{d}x'\right)^2\mathrm{d}x \\ &= \int_\Omega\left[z_i^2 - \frac{2}{|\Omega|}z_i\int_\Omega z_i\,\mathrm{d}x' + \frac{1}{|\Omega|^2}\left(\int_\Omega z_i\,\mathrm{d}x'\right)^2\right]\mathrm{d}x \\ &= \int_\Omega z_i^2\,\mathrm{d}x - \frac{1}{|\Omega|}\left(\int_\Omega z_{i,0}\,\mathrm{d}x\right)^2,\end{aligned}$$

where we used the conservation of mass in the final equality. Substituting back into (33), we get

$$\begin{aligned}\int_0^T\int_\Omega \frac{1}{2}z_i^2\,\mathrm{d}x\,\mathrm{d}t &\leq \frac{T}{2|\Omega|}\left(\int_\Omega z_{i,0}\,\mathrm{d}x\right)^2 + C_P(e^{-1}|\Omega| + 2|\Omega|TC_L^2) + C_P\int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x \\ &\quad + 2C_P|\Omega|TC_L^2\left(\int_\Omega z_{i,0}\,\mathrm{d}x\right)^2 + C_P\int_0^T\int_\Omega \delta^2|\bar{F}_i|^2\,\mathrm{d}x\,\mathrm{d}\tau.\end{aligned}$$

Combining the above with (33), we obtain

$$\int_0^T\int_\Omega \frac{1}{2}(z_i^2+|\nabla z_i|^2)\,\mathrm{d}x\,\mathrm{d}t \le \left(\frac{T}{2|\Omega|}+2(1+C_P)|\Omega|TC_L^2\right)\left(\int_\Omega z_{i,0}\,\mathrm{d}x\right)^2$$
$$+\ (1+C_P)(e^{-1}|\Omega|+2|\Omega|TC_L^2)+(1+C_P)\int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x$$
$$+\ (1+C_P)\int_0^T\int_\Omega \delta^2|\bar{F}_i|^2\,\mathrm{d}x\,\mathrm{d}t,$$

from which we recover (34) with the appropriate constant C_Ω given by (35). □

Before proceeding to the next lemma, which covers the uniform estimate for the time derivative of the regularised solutions, we recall the interpolation result of Di Benedetto.

Proposition 4.10 (*Proposition 3.2 of [70]*)**.** *Let $m,p\ge 1$ and assume that $\partial\Omega$ is piecewise smooth. There exists a constant γ depending only on d,m,p and the structure of $\partial\Omega$ such that, for any $v\in V^{m.p}$, where*

$$V^{m.p} := L^\infty(0,T;L^m(\Omega))\cap L^p(0,T;W^{1,p}(\Omega)), \tag{41}$$

we have

$$\|v\|_{L^q(Q_T)} \le \gamma\left(1+\frac{T}{|\Omega|^{\frac{d(p-m)+pm}{dm}}}\right)\|v\|_{V^{m.p}}, \qquad \text{where } q=p\frac{d+m}{d}. \tag{42}$$

Lemma 4.11 (*Time Derivative Estimate of the Regularised Solutions*)**.** *Recall the space X introduced in Definition 2.2. There holds the uniform estimate in the dual space*

$$\|\partial_t z_i\|_{X'} \le C_{X'}\left(1+\|z_{i,0}\|_{L^1(\Omega)}^2+\int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x+\delta^2\left\|\bar{F}_i\right\|_{L^2(Q_T)}^2\right), \tag{43}$$

where the positive constant $C_{X'}=C_{X'}(\Omega,T,d)$, which is independent of ε and $z_0=(z_{i,0})_{i=1}^M$, is given by

$$C_{X'} = 2\gamma\left[\left(1+\frac{T}{|\Omega|^{\frac{d+2}{d}}}\right)+(|\Omega|T)^{\frac{d}{r}}\right]\left(2+3C_\Omega+2|\Omega|TC_L^2\right). \tag{44}$$

Proof. Applying Proposition 4.10 with $v=z_i$, $m=1$, and $p=2$ yields

$$\|z_i\|_{L^q(Q_T)} \le \gamma\left(1+\frac{T}{|\Omega|^{\frac{d+2}{d}}}\right)\|z_i\|_{V^{1.2}}, \quad q=2\frac{d+1}{d}, \tag{45}$$

where the space $V^{1,2}$ is defined in (41). Notice that $q>2$ for all choices of dimension d.

Fix $\eta\in C^\infty(\bar{Q}_T)$. Going back to (26), writing $\langle\partial_t z_i,\eta\rangle=\int_0^T\int_\Omega \partial_t z_i\eta\,\mathrm{d}x\,\mathrm{d}t$, and using the divergence theorem in conjunction with the no-flux boundary condition, we find

$$|\langle\partial_t z_i,\eta\rangle| = \left|\int_0^T\int_\Omega \nabla\eta\cdot[z_i(\nabla z_i+\nabla L_i(t,x,z_i)+\delta\bar{F}_i)+\varepsilon\nabla z_i]\,\mathrm{d}x\,\mathrm{d}t\right|, \tag{46}$$

from which we obtain, using the triangle inequality,

$$|\langle\partial_t z_i,\eta\rangle| \le \int_0^T\int_\Omega |\nabla\eta| z_i|\nabla z_i+\nabla L_i(t,x,z_i)+\delta\bar{F}_i|\,\mathrm{d}x\,\mathrm{d}t+\varepsilon\int_0^T\int_\Omega|\nabla\eta||\nabla z_i|\,\mathrm{d}x\,\mathrm{d}t.$$

Then, using Hölder's inequality, we get

$$|\langle \partial_t z_i, \eta\rangle| \le \|z_i\|_{L^q(Q_T)} \left\|\nabla z_i + \nabla L_i(\cdot,\cdot,z_i) + \delta \bar{F}_i\right\|_{L^2(Q_T)} \|\nabla \eta\|_{L^r(Q_T)} + \varepsilon \|\nabla\eta\|_{L^2(Q_T)} \|\nabla z_i\|_{L^2(Q_T)}, \tag{47}$$

where r satisfies $\frac{1}{q}+\frac{1}{r}=\frac{1}{2}$, and q is as given in (45). Hence $r = 2(d+1) > 2$ and, since $Q_T = (0,T)\times\Omega$ is a bounded domain, an application of the Hölder inequality shows that

$$\begin{aligned}\|\nabla\eta\|_{L^2(Q_T)} &= \left(\int_0^T\int_\Omega |\nabla\eta|^2 \,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2}} \\ &\le \left(\int_0^T\int_\Omega |\nabla\eta|^{2(d+1)} \,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2(d+1)}} \left(\int_0^T\int_\Omega 1 \,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{d}{2(d+1)}} = (|\Omega|T)^{\frac{d}{r}} \|\nabla\eta\|_{L^r(Q_T)}.\end{aligned}$$

Combining the above with (45) and (47) shows that, with $C_\gamma = C_\gamma(\Omega,T,d)$ a positive constant given by

$$C_\gamma = \gamma\left(1 + \frac{T}{|\Omega|^{\frac{d+2}{d}}}\right) + (|\Omega|T)^{\frac{d}{r}},$$

which is independent of $\varepsilon\in(0,1)$ and $z_0 = (z_{i,0})_{i=1}^M$, there holds

$$\begin{aligned}|\langle \partial_t z_i, \eta\rangle| \le C_\gamma \|\nabla\eta\|_{L^r(Q_T)} \Bigg(&\|z_i\|_{L^2(0,T;H^1(\Omega))} \\ &+ \|z_i\|_{V^{1,2}} \left(\|z_i\|_{L^2(0,T;H^1(\Omega))} + \|\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)} + \delta\left\|\bar{F}_i\right\|_{L^2(Q_T)}\right)\Bigg),\end{aligned}$$

for any $\eta \in C^\infty(\bar{Q}_T)$, where we also used Minkowski's inequality. Using the Cauchy–Young inequality on both of the terms inside the large brackets gives

$$\begin{aligned}|\langle \partial_t z_i, \eta\rangle| \le 2C_\gamma \Bigg(&1 + \|z_i\|_{V^{1,2}}^2 \\ &+ \|z_i\|^2_{L^2(0,T;H^1(\Omega))} + \|\nabla L_i(\cdot,\cdot,z_i)\|^2_{L^2(Q_T)} + \delta^2\left\|\bar{F}_i\right\|^2_{L^2(Q_T)}\Bigg)\|\nabla\eta\|_{L^r(Q_T)},\end{aligned} \tag{48}$$

for any $\eta \in C^\infty(\bar{Q}_T)$. Observe then that

$$\|\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)} \le \|\nabla V_i\|_{L^2(Q_T)} + \left(\int_0^T\int_\Omega |\nabla W_i * z_i(t,\cdot)(x)|^2 \,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2}},$$

where we estimate the second term on the right-hand side in the same way as in (40). It follows that

$$\|\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)} \le (|\Omega|T)^{\frac{1}{2}} C_L \left(1 + \int_\Omega z_{i,0}\,\mathrm{d}x\right). \tag{49}$$

Note also that the mass conservation of Lemma 4.5 yields

$$\|z_i\|_{V^{1,2}} \le \|z_{i,0}\|_{L^1(\Omega)} + \|z_i\|_{L^2(0,T;H^1(\Omega))}.$$

By combining the above with (48) along with (49), and with the estimate (34) of Lemma 4.9, we obtain

$$|\langle \partial_t z_i, \eta\rangle| \le C_{X'}\left(1 + \|z_{i,0}\|^2_{L^1(\Omega)} + \int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x + \delta^2\left\|\bar{F}_i\right\|^2_{L^2(Q_T)}\right)\|\nabla\eta\|_{L^r(Q_T)},$$

for any $\eta\in C^\infty(\bar{Q}_T)$, with $C_{X'} = 2C_\gamma\left(2 + 3C_\Omega + 2|\Omega|TC_L^2\right)$, i.e., as given in (44), which is independent of ε and $z_0 = (z_{i,0})_{i=1}^M$. Using the density of the smooth functions in the space $L^r(0,T;W^{1,r}(\Omega))$, we take the supremum over all test functions $\eta\in L^r(0,T;W^{1,r}(\Omega))$ in the previous estimate, and deduce the uniform estimate (43). □

We also note the following estimate on the second derivatives of the regularised solutions. For clarity of exposition, we delay the proof to Appendix A.2, which relies on first proving a H^2 type estimate for a nonlinear transformation of the regularised solution.

Lemma 4.12 (*Second Derivative Estimate*)**.** *For the regularised frozen system (26), there holds, for $i \in \{1, \dots, M\}$, the estimate*

$$\|\Delta z_i\|_{L^2(Q_T)} \leq C(\varepsilon, \delta, T, \Omega, C_L, \|z_{i,0}\|_{C^1(\bar{\Omega})}, \|\bar{F}_i\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}, \|\partial_t \bar{F}_i\|_{L^1(0,T;L^2(\Omega))}), \tag{50}$$

where the right-hand side is a positive quantity depending only on the parameters in its parentheses.

4.3. Uniqueness of weak solutions to the regularised frozen problem

The following result provides a correspondence between equivalent weak formulations of the regularised frozen problem (26). We omit the proof, as it is standard. We note in passing that, by the same method of proof, the analogous equivalences also hold for the weak formulation of the original problem, as mentioned in Remark 2.5.

Lemma 4.13 (*Equivalence of Weak Formulations of* (26))**.** *Let $z = (z_i)_{i=1}^M \in \Xi$ and denote the flux by $\mathcal{F}_i := z_i(\nabla z_i + \nabla L_i(t, x, z_i) + \delta F_i(t, x, \bar{z}, \nabla \bar{z})) + \varepsilon \nabla z_i$. The following formulations are equivalent:*

1 for each $i \in \{1, \dots, M\}$, for every $\eta \in C^1(\bar{\Omega})$, for a.e. $t \in [0, T]$,

$$\langle \partial_t z_i(t, \cdot), \eta \rangle_\Omega + \int_\Omega \mathcal{F}_i \cdot \nabla \eta \, \mathrm{d}x = 0, \tag{51}$$

where $\langle \cdot, \cdot \rangle_\Omega$ is the duality product of $W^{1,r}(\Omega)$;

2 for each $i \in \{1, \dots, M\}$, for every $\phi \in C^1(\bar{Q}_T)$,

$$\langle \partial_t z_i, \phi \rangle_{X' \times X} + \int_{Q_T} \mathcal{F}_i \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t = 0; \tag{52}$$

3 for each $i \in \{1, \dots, M\}$, for every $\phi \in C^1(\bar{Q}_T)$ with $\phi(0, \cdot) = \phi(T, \cdot) = 0$ on $\bar{\Omega}$,

$$\int_{Q_T} \left(-z_i \partial_t \phi + \mathcal{F}_i \cdot \nabla \phi \right) \mathrm{d}x \, \mathrm{d}t = 0. \tag{53}$$

The following result is the focal point of this subsection and is concerned with the uniqueness of solutions to the regularised frozen problem in the larger space Ξ.

Lemma 4.14 (*Uniqueness for Regularised Frozen Problem in Ξ*)**.** *There exists a unique $z = (z_i)_{i=1}^M \in \Xi$ such that, for every $i \in \{1, \dots, M\}$, $\int_\Omega z_i(t, x) \, \mathrm{d}x = \int_\Omega z_{i,0}(x) \, \mathrm{d}x$, $z_i(0, \cdot) = z_{i,0}$ in $(W^{1,r}(\Omega))'$ and, for every $\eta \in C^1(\bar{Q}_T)$,*

$$\langle \partial_t z_i, \eta \rangle_{X' \times X} + \int_{Q_T} [z_i(\nabla z_i + \nabla L_i(t, x, z_i) + \delta F_i(t, x, \bar{z}, \nabla \bar{z})) + \varepsilon \nabla z_i] \cdot \nabla \eta \, \mathrm{d}x \, \mathrm{d}t = 0. \tag{54}$$

Proof. *Step 1 (related dual problem):* Let z and z^* satisfy the hypotheses and let $w := z_i - z_i^*$. Problem (54) is nonlinear. However, we can formally derive a dual equation through integration by parts and properties of the convolution. Specifically, supposing that the convolution kernel H is radially symmetric and that the functions below are sufficiently integrable, we note the following property:

$$\int_\Omega f(x)(H * g)(x) \, \mathrm{d}x = \int_{\Omega \times \Omega} f(x) H(x - y) g(y) \, \mathrm{d}y \, \mathrm{d}x = \int_\Omega (H * f)(y) g(y) \, \mathrm{d}y,$$

which follows directly from the Tonelli–Fubini Theorem. In the sequel we use the notation $\bar{F}_i$ introduced in (27), and we drop the arguments of $\bar{F}_i$, V_i and W_i. Define $\phi \in L^2(0,T;H^2(\Omega)) \cap H^1(0,T;L^2(\Omega))$ as a strong solution of the following linear and strongly parabolic *dual equation*:

$$\begin{aligned} \partial_t\phi - (\varepsilon + a_\kappa)\Delta\phi + (\nabla V_i + \delta\bar{F}_i + \nabla W_i * z_i)\cdot\nabla\phi + \sum_{k=1}^{d}(\partial_{x_k}W_i) * (z_i^*\partial_{x_k}\phi) &= -\xi \quad \text{in } Q_T,\\ \nabla\phi\cdot\nu &= 0 \quad \text{on } \Sigma_T,\\ \phi(0,\cdot) &= 0 \quad \text{on } \Omega, \end{aligned} \tag{55}$$

where $\xi \in C_c^\infty(Q_T)$ is arbitrary and $\{a_\kappa\}_{\kappa\in\mathbb{N}}$ is a monotone sequence of bounded functions that approximate $\frac{z_i+z_i^*}{2}$; in particular, we choose

$$a_\kappa := \min\left\{\left(\frac{z_i+z_i^*}{2}\right), \kappa\right\}.$$

Note that $a_\kappa \geq 0$ belongs to $L^\infty(Q_T)$ for every $\kappa \in \mathbb{N}$.

Step 2 (estimates on time-shifted quantity): Consider now the equation for the time-shifted function $\psi(t,x) = \phi(T-t,x)$:

$$\begin{aligned} \partial_t\psi + (\varepsilon + a_\kappa)\Delta\psi - (\nabla V_i + \delta\bar{F}_i + \nabla W_i * z_i)\cdot\nabla\psi - \sum_{k=1}^{d}(\partial_{x_k}W_i) * (z_i^*\partial_{x_k}\psi) &= \xi \quad \text{in } Q_T,\\ \nabla\psi\cdot\nu &= 0 \quad \text{on } \Sigma_T,\\ \psi(T,\cdot) &= 0 \quad \text{on } \Omega, \end{aligned} \tag{56}$$

and test against a bounded C^1 function $\theta : [0,T] \to [1,\infty)$ such that $\partial_t\theta \geq 1$. We have, for every $t \in [0,T]$,

$$\begin{aligned} \int_t^T\int_\Omega \partial_t\psi\,\theta\Delta\psi\,\mathrm{d}x\,\mathrm{d}\tau &= -\int_t^T\int_\Omega \frac{\theta}{2}\partial_t\left(|\nabla\psi|^2\right)\mathrm{d}x\,\mathrm{d}\tau\\ &= -\int_\Omega \frac{1}{2}\left(\theta(T)|\nabla\psi(T,x)|^2 - \theta(t)|\nabla\psi(t,x)|^2\right)\mathrm{d}x + \int_t^T\int_\Omega |\nabla\psi|^2\partial_t\left(\frac{\theta}{2}\right)\mathrm{d}x\,\mathrm{d}\tau\\ &\geq \int_\Omega \frac{1}{2}|\nabla\psi(t,x)|^2\,\mathrm{d}x + \int_t^T\int_\Omega \frac{1}{2}|\nabla\psi|^2\,\mathrm{d}x\,\mathrm{d}\tau, \end{aligned}$$

where we used $\phi(0,\cdot) = \psi(T,\cdot) = 0$ and the lower bound on θ. On the other hand, we also have

$$\begin{aligned} \int_t^T\int_\Omega \partial_t\psi\,\theta\Delta\psi\,\mathrm{d}x\,\mathrm{d}\tau = \int_t^T\int_\Omega \theta\Big[&-(\varepsilon + a_\kappa)|\Delta\psi|^2 + (\nabla V_i + \delta\bar{F}_i + \nabla W_i * z_i)\cdot\nabla\psi\Delta\psi\\ &+ \sum_{k=1}^{d}(\partial_{x_k}W_i) * (z_i^*\partial_{x_k}\psi)\Delta\psi + \xi\Delta\psi\Big]\,\mathrm{d}x\,\mathrm{d}\tau. \end{aligned} \tag{57}$$

Firstly, we estimate the drift terms: letting $\mathcal{G} := \nabla V_i + \delta\bar{F}_i + \nabla W_i * z_i$,

$$\begin{aligned} \int_t^T\int_\Omega \theta\mathcal{G}\cdot\nabla\psi\,\Delta\psi\,\mathrm{d}x\,\mathrm{d}\tau &= \int_t^T\int_\Omega \frac{\theta}{2}|\nabla\psi|^2\operatorname{div}\mathcal{G}\,\mathrm{d}x\,\mathrm{d}\tau - \int_t^T\int_\Omega \theta\sum_{k,l}\partial_{x_k}\psi\,\partial_{x_l}\mathcal{G}_k\,\partial_{x_l}\psi\,\mathrm{d}x\,\mathrm{d}\tau\\ &\leq \left(\frac{1}{2} + d^2\right)\|\theta\|_{L^\infty([0,T])}\|\nabla\mathcal{G}\|_{L^\infty(Q_T)}\int_{Q_T}|\nabla\psi|^2\,\mathrm{d}x\,\mathrm{d}t, \end{aligned}$$

where we used that $\mathcal{G}$ and its space derivative are bounded, since the nonlocal drift is bounded by $\|\nabla W_i * z_i\|_{L^\infty(Q_T)} \leq \|W_i\|_{C^1(\mathbb{R}^d)}\|z_{i,0}\|_{L^1(\Omega)}$. Secondly, we estimate the nonlocal term:

$$\begin{aligned}
&\left| \int_t^T \int_\Omega \theta \sum_{k=1}^d (\partial_{x_k} W_i) * (z_i^* \partial_{x_k}\psi) \Delta\psi \,\mathrm{d}x\,\mathrm{d}\tau \right| \\
&\qquad = \left| \int_t^T \int_\Omega \theta \sum_{k=1}^d (\nabla\partial_{x_k} W_i) * (z_i^* \partial_{x_k}\psi) \cdot \nabla\psi \,\mathrm{d}x\,\mathrm{d}\tau \right| \\
&\qquad \leq d\|\theta\|_{L^\infty([0,T])}\|W_i\|_{C^2(\mathbb{R}^d)} \int_t^T \int_\Omega |\nabla\psi(\tau,x)| \left(\int_\Omega |z_i^*(\tau,y)\nabla\psi(\tau,y)| \,\mathrm{d}y \right) \mathrm{d}x\,\mathrm{d}\tau \\
&\qquad \leq d|\Omega|^{\frac{1}{2}} \|\theta\|_{L^\infty([0,T])}\|W_i\|_{C^2(\mathbb{R}^d)} \int_t^T \|z_i^*(\tau,\cdot)\|_{L^2(\Omega)} \|\nabla\psi(\tau,\cdot)\|^2_{L^2(\Omega)} \,\mathrm{d}\tau,
\end{aligned}$$

where we integrated by parts to obtain the first equality and used the Hölder inequality to obtain both the second and the final lines. By integrating the term in (57) involving $\xi\Delta\psi$ by parts, and then using the Cauchy–Schwarz integral inequality followed by the Young inequality, we have obtained

$$\begin{aligned}
&\frac{1}{2} \|\nabla\psi(t,\cdot)\|^2_{L^2(\Omega)} + \int_t^T \int_\Omega \left(\frac{1}{2}|\nabla\psi|^2 + (\varepsilon + a_\kappa)|\Delta\psi|^2 \right) \mathrm{d}x\,\mathrm{d}\tau \\
&\quad \leq \frac{1}{2}\|\nabla\xi\|^2_{L^2(Q_T)} \qquad\qquad (58) \\
&\qquad + C(d, |\Omega|, \|\theta\|_{L^\infty}, \|\nabla\mathcal{G}\|_{L^\infty}, \|W_i\|_{C^2}) \int_t^T (1 + \|z_i^*(\tau,\cdot)\|_{L^2(\Omega)}) \frac{1}{2} \|\nabla\psi(\tau,\cdot)\|^2_{L^2(\Omega)} \,\mathrm{d}x\,\mathrm{d}\tau.
\end{aligned}$$

From the above we deduce, by first estimating the integrand in the right-hand side by its supremum over $[\tau, T]$ and then bounding the left-hand side similarly, that for every $t \in [0,T]$ there holds

$$\frac{1}{2} \sup_{\tau\in[t,T]} \|\nabla\psi(\tau,\cdot)\|^2_{L^2(\Omega)} \leq \frac{1}{2}\|\nabla\xi\|^2_{L^2(Q_T)} + C \int_t^T (1 + \|z_i^*(\tau,\cdot)\|_{L^2(\Omega)}) \frac{1}{2} \sup_{y\in[\tau,T]} \|\nabla\psi(y,\cdot)\|^2_{L^2(\Omega)} \,\mathrm{d}x\,\mathrm{d}\tau.$$

where we used the shorthand C to denote the quantity with the same name appearing on the right-hand side of (58). Applying the Grönwall Lemma (starting from the initial point $t = T$, where $\psi(T,\cdot) = \phi(0,\cdot) = 0$) to this latter inequality yields

$$\sup_{t\in[0,T]} \|\nabla\psi(t,\cdot)\|^2_{L^2(\Omega)} \leq \|\nabla\xi\|^2_{L^2(Q_T)} \exp\left(C \int_0^T (1 + \|z_i^*(\tau,\cdot)\|_{L^2(\Omega)}) \,\mathrm{d}\tau \right),$$

Using the Hölder inequality, the integral inside the exponential is bounded by $T + T^{\frac{1}{2}}\|z_i^*\|^2_{L^2(Q_T)}$. Hence, returning to (58) and using the previous estimate, we obtain

$$\int_{Q_T} (\varepsilon + a_\kappa)|\Delta\psi|^2 \,\mathrm{d}x\,\mathrm{d}t \leq C_\xi, \tag{59}$$

where C_ξ depends on d, $|\Omega|$, T, $\|\nabla\xi\|_{L^2(Q_T)}$, $\|\theta\|_{L^\infty}$, $\|\nabla\mathcal{G}\|_{L^\infty}$, $\|W_i\|_{C^2}$, $\|z_i^*\|_{L^2(Q_T)}$, and not on κ.

Step 3 (estimates on the difference of solutions w): The regularity of ϕ implies that $\psi \in X$. Recalling that $w(0,\cdot) = \psi(T,\cdot) = 0$, by proceeding as in the proof of Lemma 4.13, we obtain

$$\begin{aligned}
\langle \partial_t w, \psi \rangle_{X'\times X} &= -\int_{Q_T} w\partial_t\psi \,\mathrm{d}x\,\mathrm{d}t = -\int_{Q_T} [\varepsilon\nabla w + z_i\nabla z_i - z_i^*\nabla z_i^*] \cdot \nabla\psi \,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_{Q_T} \left[w(\nabla V_i + \delta\bar{F}_i) \cdot \nabla\psi + (z_i\nabla W_i * z_i - z_i^*\nabla W_i * z_i^*) \cdot \nabla\psi \right] \mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Testing against the solution of the dual problem (56), we have

$$\int_{Q_T} w\left[\partial_t\psi + \left(\varepsilon + \frac{z_i+z_i^*}{2}\right)\Delta\psi - (\nabla V_i + \delta\bar{F}_i + \nabla W_i * z_i)\cdot\nabla\psi \right.$$
$$\left. + \sum_{k=1}^{d}(\partial_{x_k}W_i) * (z_i^*\partial_{x_k}\psi)\right]\mathrm{d}x\,\mathrm{d}t = \int_{Q_T} w\,\xi\,\mathrm{d}x\,\mathrm{d}t.$$

It follows that

$$\int_{Q_T} w\,\xi\,\mathrm{d}x\,\mathrm{d}t = \int_{Q_T} w\left(\frac{z_i+z_i^*}{2} - a_\kappa\right)\Delta\psi\,\mathrm{d}x\,\mathrm{d}t$$
$$\leq \left(\int_{Q_T}(\varepsilon + a_\kappa)|\Delta\psi|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2}}\left(\int_{Q_T}\frac{w^2}{\varepsilon + a_\kappa}\left(\frac{z_i+z_i^*}{2} - a_\kappa\right)^2\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2}}.$$

Thus, using the non-negativity of a_κ and the estimate (59), we have obtained

$$\int_{Q_T} w\,\xi\,\mathrm{d}x\,\mathrm{d}t \leq \varepsilon^{-1}C_\xi^{\frac{1}{2}}\left(\int_{Q_T} w^2\left(\frac{z_i+z_i^*}{2} - a_\kappa\right)^2\mathrm{d}x\,\mathrm{d}t\right)^{\frac{1}{2}},$$

and we recall that C_ξ is independent of κ. Even if the integral on the right-hand side above may be infinite, since $a_\kappa \nearrow \frac{z_i+z_i^*}{2}$ a.e. monotonically as $\kappa \to \infty$, a direct application of the Monotone Convergence Theorem leads to

$$\int_{Q_T} w\,\xi\,\mathrm{d}x\,\mathrm{d}t \leq 0,$$

for any $\xi \in C_c^\infty(Q_T)$. In particular, we can choose ξ to be strictly positive or strictly negative, so we must have $w = z_i - z_i^* = 0$ a.e. in Q_T. □

5. Fixed point with diffusivity

In this section, we use a fixed point argument in order to go from the regularised frozen system (26) to the *regularised coupled system*

$$\begin{cases} \partial_t z_i = \operatorname{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon\nabla z_i] & \text{in } Q_T, \\ 0 = \nu\cdot[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon\nabla z_i] & \text{on } \Sigma_T, \\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega. \end{cases} \tag{60}$$

We begin by recalling the Leray–Schauder–Schaefer Fixed Point Theorem and its simple corollary.

Theorem (*Leray–Schauder–Schaefer*)**.** Let S be a compact map from a Banach space B into itself. Suppose that the set $\{\xi \in B : \xi = \lambda S(\xi) \text{ for some } \lambda \in [0,1]\}$ is bounded. Then S has a fixed point.

Corollary 5.1. *Let S be a compact map from a Banach space B into itself. Suppose that there exist two constants $a \in [0,1)$ and $b > 0$ such that $\|S(\xi)\|_B \leq a\|\xi\|_B + b$ for all $\xi \in B$. Then S has a fixed point.*

We emphasise that, throughout this entire section, $\varepsilon > 0$ and the initial data $z_0 = (z_{i,0})_{i=1}^M \in (C_c^\infty(\Omega))^M$ prescribed in Section 4 (*cf.* Remark 4.1) are fixed.

5.1. Weak compactness of the solution map

Recall the regularised frozen system (26) of Section 4, i.e.,

$$\begin{cases} \partial_t z_i = \operatorname{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,\bar z,\nabla \bar z)) + \varepsilon \nabla z_i] & \text{in } Q_T, \\ 0 = \nu \cdot [z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,\bar z,\nabla \bar z)) + \varepsilon \nabla z_i] & \text{on } \Sigma_T, \\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega. \end{cases}$$

From Section 4 we know that the above admits a unique classical solution for each smooth vector function $\bar z$. We also recall the space Ξ introduced in Definition 2.2, and note that, since it is a closed subspace of a Banach space, it is itself Banach.

Consider the solution operator S_ε of Definition 4.3:

$$\begin{aligned} S_\varepsilon : (C^\infty(\bar Q_T))^M &\to (C^{2,1}(\bar Q_T))^M \subset \Xi \\ \bar z &\mapsto z, \end{aligned} \tag{61}$$

where z solves (26) and $(C^\infty(\bar Q_T))^M$ is equipped with the subspace topology of Ξ. Let us introduce, for every $\mu > 0$, the following linear smoothing operator R_μ (which is defined by extension and mollification in Appendix A.4):

$$\begin{aligned} R_\mu : \Xi &\to (C^\infty(\bar Q_T))^M \\ w &\mapsto R_\mu(w), \end{aligned} \tag{62}$$

with the property that $R_\mu(w)$ converges to w strongly in $(L^2(0,T;H^1(\Omega)))^M$ as $\mu \to 0$, along with

$$\|R_\mu(w)\|_{(L^2(0,T;H^1(\Omega)))^M} \le C_{reg} \|w\|_{(L^2(0,T;H^1(\Omega)))^M}, \tag{63}$$

for some fixed constant C_{reg} independent of μ, and

$$\|R_\mu(w)\|_{(L^\infty(0,T;W^{2,\infty}(\Omega)))^M} + \|\partial_t R_\mu(w)\|_{(L^\infty(0,T;W^{1,\infty}(\Omega)))^M} \le C_\mu \|w\|_{(L^2(0,T;H^1(\Omega)))^M}, \tag{64}$$

for some positive constant C_μ depending on μ (*cf.* Lemma A.6 and Remark A.7). This constant explodes in the limit as $\mu \to 0$.

We will also repeatedly use the following two technical lemmas in later arguments. Their proofs are elementary and are therefore omitted for concision.

Lemma 5.2. *Let $\{\zeta^n\}_{n\in\mathbb{N}}$ be a sequence in $L^2(Q_T)$ such that:*

1. *$\{\zeta^n\}_{n\in\mathbb{N}}$ converges weakly to ζ in $L^2(Q_T)$;*
2. *ζ^n is non-negative a.e. in Q_T for every $n \in \mathbb{N}$;*
3. *$\int_{Q_T} \zeta^n(t,x)\,\mathrm{d}x = \Lambda$ a.e. $t \in (0,T)$ for some non-negative constant Λ.*

Then $\zeta \in L^\infty(0,T;L^1(\Omega))$, $\zeta \ge 0$ a.e. in Q_T, and $\displaystyle\int_\Omega \zeta(t,x)\,\mathrm{d}x = \Lambda$.

Lemma 5.3. *Let $\{\zeta^n\}_{n\in\mathbb{N}}$ be a sequence in $L^2(Q_T)$ such that:*

1. *$\{\zeta^n\}_{n\in\mathbb{N}}$ converges weakly to ζ in $L^2(Q_T)$;*
2. *For every $n \in \mathbb{N}$, $\zeta^n(0,\cdot) = \zeta_0$ in $(W^{1,r}(\Omega))'$ for some fixed $\zeta_0 \in L^p(\Omega)$;*
3. *$\{\partial_t \zeta^n\}_{n\in\mathbb{N}}$ converges weakly-* to $\partial_t \zeta$ in X'.*

Then, $\zeta \in C([0,T];(W^{1,r}(\Omega))')$, and there exists a positive constant Λ, independent of n such that, given any $\phi \in W^{1,r}(\Omega)$, there holds

$$\left| \int_\Omega \zeta(t,x)\phi(x)\,\mathrm{d}x - \int_\Omega \zeta(s,x)\phi(x)\,\mathrm{d}x \right| \le (t-s)^{\frac{1}{r}} \Lambda \|\phi\|_{W^{1,r}(\Omega)} \quad \forall 0 < s \le t \le T. \tag{65}$$

Moreover, $\|\zeta(t,\cdot) - \zeta_0\|_{(W^{1,r}(\Omega))'} \le t^{\frac{1}{r}} \Lambda$ for all $0 \le t \le T$, so that $\zeta(0,\cdot) = \zeta_0$ in $(W^{1,r}(\Omega))'$.

Lemma 5.4 (*Weak Compactness*)**.** *The map $S_\varepsilon \circ R_\mu$ from Ξ into itself is weakly sequentially compact with respect to the subspace topology of $(L^2(0,T;H^1(\Omega))^M$.*

Proof. *Step 1 (the bounded sequence):* For $i \in \{1,\dots,M\}$, let $\{\bar{z}_i^n\}_{n\in\mathbb{N}}$ be a uniformly bounded sequence in Ξ, i.e., there exists $C_i > 0$, independent of n, μ, ε, such that

$$\|\bar{z}_i^n\|_{L^2(0,T;H^1(\Omega))} \le C_i, \qquad \|\partial_t \bar{z}_i^n\|_{X'} \le C_i \qquad \forall n \in \mathbb{N}. \tag{66}$$

Let us introduce

$$\hat{z}^n := R_\mu(\bar{z}^n), \qquad z^n := S_\varepsilon(\hat{z}^n),$$

and notice that by the estimate (64), up to the multiplicative constant C_{reg}, we have that $\bar{z} = (\bar{z}_i)_{i=1}^M$ and $\hat{z} = (\hat{z}_i)_{i=1}^M$ satisfy the same uniform bound (66). We also define

$$\hat{F}_i^n := F_i(t, x, \hat{z}^n, \nabla \hat{z}^n).$$

and notice that, due to (5) and (66), we have

$$\|\hat{F}_i^n\|_{L^2(Q_T)} \le C_F(1 + C_i) \qquad \forall n \in \mathbb{N},$$

where we omitted the C_{reg} factor, for clarity of presentation. By Lemmas 4.8, 4.9, and 4.11, there exist a positive constants C_i' independent of n, μ, ε such that

$$\|z_i^n\|_{L^2(0,T;H^1(\Omega))} \le C_i', \qquad \|\partial_t z_i^n\|_{X'} \le C_i', \qquad \forall n \in \mathbb{N}.$$

Step 2 (the weakly convergent subsequence): It follows that all of $\{\bar{z}_i^n\}_{n\in\mathbb{N}}$, $\{\hat{z}_i^n\}_{n\in\mathbb{N}}$, and $\{z_i^n\}_{n\in\mathbb{N}}$ are bounded sequences in Ξ. An application of the theorems of Banach–Alaoglu and Aubin–Lions [71, Theorem II.5.16] implies that there exists a common subsequence (still indexed by n) such that

$$\begin{aligned}
\bar{z}_i^n \rightharpoonup \bar{z}_i, \quad \hat{z}_i^n \rightharpoonup \hat{z}_i, \quad z_i^n &\rightharpoonup z_i && \text{weakly in } L^2(0,T;H^1(\Omega)),\\
\hat{F}_i^n &\rightharpoonup \hat{F}_i^* && \text{weakly in } L^2(Q_T),\\
\bar{z}_i^n \to \bar{z}_i, \quad z_i^n &\to z_i && \text{strongly in } L^2(Q_T),\\
\partial_t z_i^n &\overset{*}{\rightharpoonup} v_i && \text{weakly-* in } X',
\end{aligned}$$

where $\bar{z}_i, \hat{z}_i, z_i \in L^2(0,T;H^1(\Omega))$, $\hat{F}_i^* \in L^2(Q_T)$, and $v_i \in X'$. From the linearity and the continuity property of the regularisation operator (*cf.* Corollary A.8),

$$\|\hat{z}_i^n - R_\mu \bar{z}_i\|_{L^2(Q_T)} = \|R_\mu \bar{z}_i^n - R_\mu \bar{z}_i\|_{L^2(Q_T)} \le C\|\bar{z}_i^n - \bar{z}_i\|_{L^2(Q_T)} \to 0 \qquad \text{as } n \to \infty,$$

for some positive constant C, which, incidentally, does not depend on μ. Hence it follows that $\hat{z}_i = R_\mu \bar{z}_i$ as elements of $L^2(Q_T)$, and that, additionally, $\hat{z}_i^n \to \hat{z}_i$ strongly in $L^2(Q_T)$. Moreover, given any $\Theta \in (C_c^1(Q_T))^d$, there holds, from the definition of weak derivative,

$$\left| \int_{Q_T} \Theta \cdot \nabla(R_\mu \bar{z}_i - \hat{z}_i)\, \mathrm{d}x\, \mathrm{d}t \right| = \left| \int_{Q_T} \operatorname{div} \Theta (R_\mu \bar{z}_i - \hat{z}_i)\, \mathrm{d}x\, \mathrm{d}t \right| \le \|\operatorname{div} \Theta\|_{L^2(Q_T)} \|R_\mu \bar{z}_i - \hat{z}_i\|_{L^2(Q_T)},$$

from which we deduce that $\|\nabla(R_\mu \bar{z}_i - \hat{z}_i)\|_{L^2(Q_T)} = 0$, due to the density of $C_c^1(Q_T)$ in $L^2(Q_T)$. Thus, $\hat{z}_i = R_\mu \bar{z}_i$ as elements of $L^2(0,T;H^1(\Omega))$.

Additionally, we note that, in view of Lemma 5.2, we have $z_i \ge 0$ a.e. in Q_T and

$$\int_\Omega |z_i(t,x)|\, \mathrm{d}x = \int_\Omega z_i(t,x)\, \mathrm{d}x = \int_\Omega z_{i,0}\, \mathrm{d}x \qquad \text{a.e. } t \in (0,T),$$

which fulfils the requirement for $z_i \in L^\infty(0,T;L^1(\Omega))$.

Step 3 (convergence of cross-diffusion terms): Using the structure of $\hat{F}_i^n$ provided by (4), we know that

$$\hat{F}_i^* = F_i(t, x, \hat{z}, \nabla \hat{z}) = G_i^0(t, x, \hat{z}) + \sum_{j=1}^{M} G_{ij}^1(t, x, \hat{z}) \nabla \hat{z}_j.$$

Indeed, letting $\Theta \in (C_c^1(Q_T))^d$ be arbitrary, the structure (4) implies that

$$\int_{Q_T} \Theta \cdot \hat{F}_i^n \, \mathrm{d}x \, \mathrm{d}t = \int_{Q_T} \Theta \cdot G_i^0(t, x, \hat{z}^n) \, \mathrm{d}x \, \mathrm{d}t + \sum_{j=1}^{M} \int_{Q_T} \Theta \cdot \nabla G_{ij}^1(t, x, \hat{z}^n) \nabla \hat{z}_j \, \mathrm{d}x \, \mathrm{d}t.$$

Then, the fundamental theorem of calculus and the strong convergence in $L^2(Q_T)$ yield

$$\|G_i^0(\cdot, \cdot, \hat{z}^n) - G_i^0(\cdot, \cdot, \hat{z})\|_{L^2(Q_T)} \le \|\nabla_z G_i^0\|_{L^\infty(Q_T \times \mathbb{R}^M)} \max_{i \in \{1, \dots, M\}} \|\hat{z}_i^n - \hat{z}_i\|_{L^2(Q_T)} \to 0, \tag{67}$$

and similarly

$$\|G_{ij}^1(\cdot, \cdot, \hat{z}^n) - G_{ij}^1(\cdot, \cdot, \hat{z})\|_{L^2(Q_T)} \le \|\nabla_z G_{ij}^1\|_{L^\infty(Q_T \times \mathbb{R}^M)} \max_{i \in \{1, \dots, M\}} \|\hat{z}_i^n - \hat{z}_i\|_{L^2(Q_T)} \to 0, \tag{68}$$

so that $G_i^0(\cdot, \cdot, \hat{z}^n) \to G_i^0(\cdot, \cdot, \hat{z})$ and $G_{ij}^1(\cdot, \cdot, \hat{z}^n) \to G_{ij}^1(\cdot, \cdot, \hat{z})$ strongly in $L^2(Q_T)$. The weak convergence also implies $\nabla \hat{z}_j^n \rightharpoonup \nabla \hat{z}_j$ weakly in $L^2(Q_T)$, so that, using the fact that the product of a strongly converging sequence with a weakly converging sequence converges itself in the weak sense,

$$\lim_{n \to \infty} \int_{Q_T} \Theta \cdot \hat{F}_i^n \, \mathrm{d}x \, \mathrm{d}t = \int_{Q_T} \Theta \cdot \hat{F}_i \, \mathrm{d}x \, \mathrm{d}t.$$

Step 4 (convergence of drift terms): Similarly, with the term L_i, we have

$$\begin{aligned} \|L_i(t, x, z_i^n) - L_i(t, x, z_i)\|_{L^2(Q_T)}^2 &= \int_0^T \int_\Omega \left|(W_i * (z_i^n(t, \cdot) - z_i(t, \cdot)))(x)\right|^2 \mathrm{d}x \, \mathrm{d}t \\ &= \int_0^T \int_\Omega \left| \int_\Omega W_i(x - y)(z_i^n(t, y) - z_i(t, y)) \, \mathrm{d}y \right|^2 \mathrm{d}x \, \mathrm{d}t \\ &\le |\Omega| C_L^2 \int_0^T \int_\Omega \left(\int_\Omega |z_i^n(t, y) - z_i(t, y)|^2 \, \mathrm{d}y \right) \mathrm{d}x \, \mathrm{d}t \\ &= |\Omega|^2 C_L^2 \|z_i^n - z_i\|_{L^2(Q_T)}^2 \to 0, \end{aligned}$$

where we applied Jensen's inequality and used the boundedness (8) to obtain the third line, and the Tonelli–Fubini theorem to obtain the final line. An identical strategy yields

$$\begin{aligned} \|\nabla L_i(t, x, z_i^n) - \nabla L_i(t, x, z_i)\|_{L^2(Q_T)}^2 &= \int_0^T \int_\Omega \left|(\nabla W_i * (z_i^n(t, \cdot) - z_i(t, \cdot)))(x)\right|^2 \mathrm{d}x \, \mathrm{d}t \\ &\le |\Omega|^2 C_L^2 \|z_i^n - z_i\|_{L^2(Q_T)}^2 \to 0, \end{aligned}$$

so that we obtain the convergence $L_i(\cdot, \cdot, z_i^n) \to L(\cdot, \cdot, z_i)$ strongly in $L^2(0, T; H^1(\Omega))$.

Step 5 (sequence of time derivatives and conclusion): Furthermore, given any $\theta \in C_c^1(Q_T)$, an integration by parts with respect to the time variable yields

$$\int_{Q_T} \theta \partial_t z_i^n \, \mathrm{d}x \, \mathrm{d}t = - \int_{Q_T} z_i^n \partial_t \theta \, \mathrm{d}x \, \mathrm{d}t.$$

By taking the weak limits on both sides, we get $\int_{Q_T} \theta v_i \, \mathrm{d}x \, \mathrm{d}t = - \int_{Q_T} z_i \partial_t \theta \, \mathrm{d}x \, \mathrm{d}t$ and deduce $v_i = \partial_t z_i$. Hence, since any weakly-* convergent sequence is bounded, we have that $\|\partial_t z_i\|_{X'} < +\infty$, and the final requirement for z_i belonging to Ξ is fulfilled. We therefore have the following convergences for the subsequence:

$$
\begin{array}{lll}
\bar{z}_i^n \rightharpoonup \bar{z}_i, \quad \hat{z}_i^n \rightharpoonup R_\mu \bar{z}_i, & z_i^n \rightharpoonup z_i & \text{weakly in } L^2(0,T;H^1(\Omega)), \\
\bar{z}_i^n \to \bar{z}_i, \quad \hat{z}_i^n \to R_\mu \bar{z}_i, & z_i^n \to z_i & \text{strongly in } L^2(Q_T), \\
 & L_i(\cdot,\cdot,z_i^n) \to L_i(\cdot,\cdot,z_i) & \text{strongly in } L^2(0,T;H^1(\Omega)), \\
 & \hat{F}_i^n \rightharpoonup F_i(\cdot,\cdot,R_\mu \bar{z}, \nabla R_\mu \bar{z}) & \text{weakly in } L^2(Q_T), \\
 & \partial_t z_i^n \overset{*}{\rightharpoonup} \partial_t z_i & \text{weakly-* in } X',
\end{array}
\tag{69}
$$

with $z_i \in \Xi$, and the lemma is proved, since we have shown the weak convergence in Ξ of a subsequence of $\{S_\varepsilon \circ R_\mu(\bar{z}^n)\}_{n\in\mathbb{N}}$ towards $z = (z_i)_{i=1}^M$ for any bounded sequence $\{\bar{z}_n\}_{n\in\mathbb{N}}$ in Ξ. □

Remark 5.5. We emphasise that, in order to obtain (69), the Aubin–Lions Lemma (*cf.* [71, Theorem II.5.16]) was used to ensure the strong convergences $\bar{z}_i^n \to \bar{z}_i$ and $z_i^n \to z_i$ in $L^2(Q_T)$. The application of [71, Theorem II.5.16] is justified due to the uniform bound in $L^2(0,T;H^1(\Omega))$ for the sequence of functions $\{\bar{z}_i^n, z_i^n\}_{n\in\mathbb{N}}$ and the uniform bound in $X' = L^{r'}(0,T;(W^{1,r}(\Omega))')$ for the corresponding sequence of derivatives $\{\partial_t \bar{z}_i^n, \partial_t z_i^n\}_{n\in\mathbb{N}}$. The strong convergence $\hat{z}_i^n \to R_\mu \bar{z}_i$ in $L^2(Q_T)$ was not deduced directly from the Aubin–Lions Lemma (though, alternatively, this can also be done) and was later obtained from properties of the regularisation operator R_μ and the strong convergence $\bar{z}_i^n \to \bar{z}_i$ in $L^2(Q_T)$.

Remark 5.6. As a consequence of (69), for any $\phi \in C^1(\bar{Q}_T)$,

$$
\begin{aligned}
&\int_{Q_T} z_i^n \nabla z_i^n \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \to \int_{Q_T} z_i \nabla z_i \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t, \\
&\int_{Q_T} z_i^n \nabla L_i(t,x,z_i^n) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \to \int_{Q_T} z_i \nabla L_i(t,x,z_i) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t, \\
&\int_{Q_T} z_i^n F_i(t,x,R_\mu \bar{z}^n, \nabla R_\mu \bar{z}^n) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \to \int_{Q_T} z_i F_i(t,x,R_\mu \bar{z}, \nabla R_\mu \bar{z}) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

It is then clear that the limit function z_i satisfies the following weak formulation:

$$
\langle \partial_t z_i, \phi \rangle_{X' \times X} + \int_{Q_T} \nabla\phi \cdot \big([z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,R_\mu \bar{z}, \nabla R_\mu \bar{z})] + \varepsilon \nabla z_i \big) \,\mathrm{d}x\,\mathrm{d}t = 0,
$$

for every $\phi \in C^1(\bar{Q}_T)$. Note the similarity between the no-flux weak formulation above and the formulation (52) in Lemma 4.13.

5.2. Strong compactness of the solution map

In this subsection, we improve the compactness result in Lemma 5.4 and show that the solution map is actually strongly compact from Ξ to itself. To begin with, we recall the following result concerning semicontinuity properties of the Fisher information.

Lemma 5.7 (*Properties of Fisher Information, [72] Lemma 4.10*)**.** *Let K be a closed subset of $\mathbb{R}^d$. The Fisher information, i.e., the functional*

$$
\mathcal{F}[w] := \int_{\{x\in K \mid w(x)>0\}} \frac{|\nabla w|^2}{w} \,\mathrm{d}x,
$$

is convex and sequentially lower semicontinuous with respect to the weak topology of $L^1(K)$.

Consequently, we have the following lemma, the proof of which is contained in Appendix A.3.

Lemma 5.8. *Let $\{\zeta^n\}_{n\in\mathbb{N}}$ be a sequence of non-negative functions in $L^1(Q_T)$, and suppose that $\zeta^n \to \zeta$ strongly in $L^1(Q_T)$. Then there exists a subsequence of $\{\zeta^n\}_{n\in\mathbb{N}}$, still indexed by n, for which there holds*

$$\int_{Q_T} \frac{|\nabla\zeta|^2}{\zeta}\,\mathrm{d}x\,\mathrm{d}t \le \liminf_{n\to\infty} \int_{Q_T} \frac{|\nabla\zeta^n|^2}{\zeta^n}\,\mathrm{d}x\,\mathrm{d}t.$$

Lemma 5.9 (*Strong Compactness*)**.** *The map $S_\varepsilon \circ R_\mu$ from $\varXi$ into itself is strongly sequentially compact.*

Proof. *Step 1 (difference of two "weak formulations"):* We emphasise that ε and μ are fixed throughout this proof. For $i \in \{1,\dots,M\}$, let $\{\bar{z}_i^n\}_{n\in\mathbb{N}}$ be a uniformly bounded sequence in $\varXi$. Arguing as in Lemma 5.4, we consider $\hat{z}^n = R_\mu \bar{z}^n$ and $z^n = S_\varepsilon \hat{z}^n = S_\varepsilon \circ R_\mu(\bar{z}^n)$. Recalling (69), we have that, for a suitable subsequence (still indexed by n),

$$\begin{aligned}
&\bar{z}_i^n \rightharpoonup \bar{z}_i, \quad \hat{z}_i^n \rightharpoonup R_\mu\bar{z}_i, \quad z_i^n \rightharpoonup z_i && \text{weakly in } L^2(0,T;H^1(\varOmega)),\\
&\bar{z}_i^n \to \bar{z}_i, \quad \hat{z}_i^n \to R_\mu\bar{z}_i, \quad z_i^n \to z_i && \text{strongly in } L^2(Q_T),\\
&L_i(\cdot,\cdot,z_i^n) \to L_i(\cdot,\cdot,z_i) && \text{strongly in } L^2(0,T;H^1(\varOmega)),\\
&\hat{F}_i^n \rightharpoonup F_i(\cdot,\cdot,R_\mu\bar{z},\nabla R_\mu\bar{z}) && \text{weakly in } L^2(Q_T),\\
&\partial_t z_i^n \overset{*}{\rightharpoonup} \partial_t z_i && \text{weakly-* in } X',
\end{aligned}$$

for some non-negative $z_i \in \varXi$, and we define $\hat{z} := R_\mu\bar{z}$. By integrating the equality (36) from Lemma 4.9 over the time interval $[t_0,t] \subsetneq [0,T]$, we obtain

$$\begin{aligned}
&\int_\varOmega \left[z_i^n(t)\log z_i^n(t) - z_i^n(t_0)\log z_i^n(t_0)\right]\mathrm{d}x\\
&\quad = -\int_{t_0}^t\int_\varOmega \left[\delta\nabla z_i^n\cdot F_i(t,x,\hat{z}^n,\nabla\hat{z}^n) + \nabla z_i^n\cdot\nabla L_i(t,x,z_i^n) + |\nabla z_i^n|^2 + \varepsilon\frac{|\nabla z_i^n|^2}{z_i^n}\right]\mathrm{d}x\,\mathrm{d}\tau,
\end{aligned}$$

for a.e. $t > t_0 > 0$. By Remark 5.6, the uniqueness in the space $\varXi$ due to Lemma 4.14, and the added regularity inherited from Lemma 4.8, we deduce that $z_i \in C^{2,1}(\bar{Q}_T)$ satisfies (26) classically (with $\hat{z} = R_\mu\bar{z}$ featuring in the arguments of F_i). Then, similarly to what we had in Section 4.2, we also obtain

$$\begin{aligned}
&\int_\varOmega \left[z_i(t)\log z_i(t) - z_i(t_0)\log z_i(t_0)\right]\mathrm{d}x\\
&\quad = -\int_{t_0}^t\int_\varOmega \left[\nabla z_i\cdot\delta F_i(t,x,\hat{z},\nabla\hat{z}) + \nabla z_i\cdot\nabla L_i(t,x,z_i) + |\nabla z_i|^2 + \varepsilon\frac{|\nabla z_i|^2}{z_i}\right]\mathrm{d}x\,\mathrm{d}\tau,
\end{aligned}$$

for a.e. $t > t_0 > 0$. Taking the difference of the two relations above we obtain

$$\begin{aligned}
&\int_{t_0}^t\int_\varOmega \left[\left(1+\frac{\varepsilon}{z_i^n}\right)|\nabla z_i^n|^2 - \left(1+\frac{\varepsilon}{z_i}\right)|\nabla z_i|^2\right]\mathrm{d}x\,\mathrm{d}\tau\\
&\quad = \int_\varOmega \big(z_i(t_0)\log z_i(t_0) - z_i^n(t)\log z_i^n(t) - z_i(t_0)\log z_i(t_0) + z_i(t)\log z_i(t)\big)\,\mathrm{d}x\\
&\qquad - \int_{t_0}^t\int_\varOmega \big(\nabla z_i^n\cdot\nabla L_i(t,x,z_i^n) - \nabla z_i\cdot\nabla L_i(t,x,z_i)\big)\,\mathrm{d}x\,\mathrm{d}\tau\\
&\qquad - \delta\int_{t_0}^t\int_\varOmega \big(\nabla z_i^n\cdot F_i(t,x,\hat{z}^n,\nabla\hat{z}^n) - \nabla z_i\cdot F_i(t,x,\hat{z},\nabla\hat{z})\big)\,\mathrm{d}x\,\mathrm{d}\tau.
\end{aligned} \tag{70}$$

One can show, by following Step 1 of the proof of Lemma 5.8 (*cf.* Appendix A.3), that the strong convergence $z_i^n \to z_i$ in $L^2(Q_T)$ implies that, for a subsequence, for a.e. $t \in (0,T)$, we have $\|z_i^n(t,\cdot) - z_i(t,\cdot)\|_{L^2(\varOmega)} \to 0$.

Then, defining $f : x \mapsto x(\log x)\mathbb{1}_{[0,\infty)}(x)$, we note that $|f(x)| \leq C(1+x^2)$ globally for some universal constant C. Using the Generalised Dominated Convergence Theorem, we deduce that f maps $L^2(\Omega)$ continuously into $L^1(\Omega)$, whence the entire first term on the right-hand side of (70) vanishes for this subsequence. The second term on the right-hand side of (70) also vanishes, due to the strong convergence in $L^2(0,T;H^1(\Omega))$ of the terms involving ∇L_i, *cf.* (69). On the other hand, the final term on the right-hand side of (70) requires additional work.

Step 2 (div–curl Lemma to make right-hand side of (70) *vanish):* Recall that, due to the structure provided by (4),

$$F_i(t,x,\hat{z},\nabla\hat{z}) = G_i^0(t,x,\hat{z}) + \sum_{j=1}^{M} G_{ij}^1(t,x,\hat{z})\nabla\hat{z}_j,$$

and we consider, in particular, the term

$$\int_{t_0}^{t}\int_{\Omega}\big(G_{ij}^1(\tau,x,\hat{z}^n)\nabla\hat{z}_j^n\cdot\nabla z_i^n - G_{ij}^1(\tau,x,\hat{z})\nabla\hat{z}_j\cdot\nabla z_i\big)\,\mathrm{d}x\,\mathrm{d}\tau. \tag{71}$$

For what follows we define the $(d+1)$-dimensional vector fields

$$\hat{v}_j^n(\tau,x) := (0, G_{ij}^1(\tau,x,\hat{z}^n(\tau,x))\nabla\hat{z}_j^n(\tau,x)), \qquad v_i^n(\tau,x) := (0,\nabla z_i^n(\tau,x)).$$

Note that the strong convergence in $L^2(Q_T)$ of the terms involving G_{ij}^1, *cf.* (68), and the weak convergences $\nabla\hat{z}_j^n \rightharpoonup \nabla\hat{z}_j, \nabla z_i^n \rightharpoonup \nabla z_i$ in $L^2(Q_T)$ implies that both sequences $\{\hat{v}_j^n\}_{n\in\mathbb{N}}, \{v_i^n\}_{n\in\mathbb{N}}$ are weakly convergent in $(L^2(Q_T))^{d+1}$. In the next paragraph, we pass to the limit in (71) using the div–curl Lemma.

By Lemma 4.12, $\|\Delta z^n\|_{L^2(Q_T)}$ is bounded by $\|R_\mu\bar{z}^n\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} + \|\partial_t R_\mu\bar{z}^n\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}$, which, from the estimate (64), is bounded by $\|\bar{z}^n\|_{L^2(0,T;H^1(\Omega))}$, and this is bounded independently of n. Therefore, for every $i \in \{1,\dots,M\}$,

$$\operatorname{div}_{t,x} v_i^n = \Delta z_i^n \qquad \text{is bounded in } L^2(Q_T) \text{ independently of } n,$$

and thus, by the Rellich Theorem, the sequence $\{\operatorname{div}_{t,x} v_i^n\}_{n\in\mathbb{N}}$ is confined to a compact subset of $H^{-1}(Q_T)$. Additionally, an explicit computation using the chain rule shows that, for every $j \in \{1,\dots,M\}$,

$$|\operatorname{curl}_{t,x}\hat{v}_j^n| \leq \max_{i,j\in\{1,\dots,M\}}\|G_{ij}^1\|_{C^1(\bar{Q}_T\times\mathbb{R}^M)}\big(|\partial_t\nabla\hat{z}^n| + |\nabla^2\hat{z}^n| + |\nabla\hat{z}^n| + |\partial_t\hat{z}^n||\nabla\hat{z}^n| + |\nabla\hat{z}^n|^2\big).$$

Recall that $\mu > 0$ is fixed. Hence, estimate (64) shows that $\|\hat{z}_i^n\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} = \|R_\mu\bar{z}_i^n\|_{L^\infty(0,T;W^{2,\infty}(\Omega))}$ is bounded independently of n and likewise for $\|\partial_t\hat{z}_i^n\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} = \|\partial_t R_\mu\bar{z}_i^n\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}$. In turn, $\{\operatorname{curl}_{t,x}\hat{v}_j^n\}_{n\in\mathbb{N}}$ is bounded in $L^\infty(Q_T)$ independently of n, and therefore also uniformly bounded in $L^2(Q_T)$, whence the Rellich Theorem implies that this sequence is confined to a compact subset of $H^{-1}(Q_T)$. A direct application of the div –curl Lemma (*cf.* [73, Theorem 1]) to the product $\{\hat{v}_j^n\cdot v_i^n\}_{n\in\mathbb{N}}$ yields that

$$G_{ij}^1(t,x,\hat{z}^n)\nabla\hat{z}_j^n\cdot\nabla z_i^n \to G_{ij}^1(t,x,\hat{z})\nabla\hat{z}_j\cdot\nabla z_i \qquad \text{in } \mathcal{D}'(Q_T).$$

We conclude that the term in (71) vanishes in the as $n\to\infty$. Note that the term

$$\int_{t_0}^{t}\int_{\Omega}\big(G_i^0(t,x,\hat{z}^n)\cdot\nabla z_i^n - G_i^0(t,x,\hat{z})\cdot\nabla z_i\big)\,\mathrm{d}x\,\mathrm{d}\tau,$$

also vanishes in the limit as $n\to\infty$, since $\nabla z_i^n \rightharpoonup \nabla z_i$ weakly in $L^2(Q_T)$ and $G_i^0(\cdot,\cdot,\hat{z}^n) \to G_i^0(\cdot,\cdot,\hat{z})$ strongly in $L^2(Q_T)$, as per the estimate (67).

Step 3 (lower semicontinuity of Fisher information to deduce convergence of norms): Returning to (70) and using the fact that $0 < t_0 < t < T$ were arbitrary, it follows (by possibly taking a further subsequence to let $t_0 \to 0$ and $t \to T$) that

$$\int_0^T\int_\Omega\left[\left(1+\frac{\varepsilon}{z_i^n}\right)|\nabla z_i^n|^2 - \left(1+\frac{\varepsilon}{z_i}\right)|\nabla z_i|^2\right]\mathrm{d}x\,\mathrm{d}\tau \to 0 \qquad \text{as } n\to\infty. \tag{72}$$

Observe that the strong convergence $z_i^n \to z_i$ in $L^2(Q_T)$ and the lower semicontinuity result Lemma 5.8 imply, after passing to a further subsequence if necessary,

$$\liminf_{n\to\infty} \int_0^T \int_\Omega \left[\frac{1}{z_i^n}|\nabla z_i^n|^2 - \frac{1}{z_i}|\nabla z_i|^2\right] \mathrm{d}x \,\mathrm{d}\tau \geq 0,$$

which, combining with (72), implies

$$\limsup_{n\to\infty}\big(\|\nabla z_i^n\|^2_{L^2(Q_T)} - \|\nabla z_i\|^2_{L^2(Q_T)}\big) \leq 0.$$

Combining the above with $\|\nabla z_i\|_{L^2(Q_T)} \leq \liminf_{n\to\infty}\|\nabla z_i^n\|_{L^2(Q_T)}$, which holds true because of the weak lower semicontinuity of the norm and $\nabla z_i^n \rightharpoonup \nabla z_i$ weakly in $L^2(Q_T)$, we deduce

$$\|\nabla z_i^n\|_{L^2(Q_T)} \to \|\nabla z_i\|_{L^2(Q_T)} \qquad \text{as } n \to \infty. \tag{73}$$

The combination of weak convergence of $\{z_i^n\}_{n\in\mathbb{N}}$ in $L^2(0,T;H^1(\Omega))$ with the convergence of the norm establishes strong convergence in $L^2(0,T;H^1(\Omega))$.

Step 4 (strong convergence of time derivatives in dual space): Finally, we verify the strong convergence in the dual space X' for the sequence of time derivative $\{\partial_t z_i^n\}_{n\in\mathbb{N}}$. Taking the difference of the weak formulations we obtain, for any $\theta \in C_c^1(Q_T)$,

$$\begin{aligned}\langle \partial_t(z_i^n - z_i), \theta\rangle = &-\int_{Q_T}\big(\varepsilon\nabla(z_i^n - z_i) + z_i^n\nabla z_i^n - z_i\nabla z_i\big)\cdot\nabla\theta \,\mathrm{d}x\,\mathrm{d}t\\ &-\int_{Q_T}\big(z_i^n\nabla L_i(t,x,z_i^n) - z_i\nabla L_i(t,x,z_i)\big)\cdot\nabla\theta\,\mathrm{d}x\,\mathrm{d}t\\ &-\int_{Q_T}\delta\big(z_i^n F_i(t,x,\hat{z}^n,\nabla\hat{z}^n) - z_i F_i(t,x,\hat{z},\nabla\hat{z})\big)\cdot\nabla\theta\,\mathrm{d}x\,\mathrm{d}t,\end{aligned}$$

and, as per the convergences identified in Remark 5.6, the right-hand side vanishes in the limit as $n \to \infty$. However, we must study this limit quantitatively. To this end, observe from this previous equation that

$$\begin{aligned}|\langle\partial_t(z_i^n - z_i),\theta\rangle| \leq \big(&\varepsilon\|\nabla z_i^n - \nabla z_i\|_{L^2(Q_T)} + \|z_i^n\nabla z_i^n - z_i\nabla z_i\|_{L^2(Q_T)}\\ &+ \|z_i^n\nabla L_i(\cdot,\cdot,z_i^n) - z_i\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)}\\ &+ |\delta|\|z_i^n F_i(\cdot,\cdot,\hat{z}^n,\nabla\hat{z}^n) - z_i F_i(\cdot,\cdot,\hat{z},\nabla\hat{z})\|_{L^2(Q_T)}\big)\|\nabla\theta\|_{L^2(Q_T)},\end{aligned}$$

after which an application of the Hölder inequality yields

$$\begin{aligned}|\langle\partial_t(z_i^n - z_i),\theta\rangle| \leq \big(&\varepsilon\|\nabla z_i^n - \nabla z_i\|_{L^2(Q_T)} + \|z_i^n\nabla z_i^n - z_i\nabla z_i\|_{L^2(Q_T)}\\ &+ \|z_i^n\nabla L_i(\cdot,\cdot,z_i^n) - z_i\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)}\\ &+ |\delta|\|z_i^n F_i(\cdot,\cdot,\hat{z}^n,\nabla\hat{z}^n) - z_i F_i(\cdot,\cdot,\hat{z},\nabla\hat{z})\|_{L^2(Q_T)}\big)(|\Omega|T)^{\frac{d}{2(d+1)}}\|\nabla\theta\|_{L^r(Q_T)}.\end{aligned}$$

Taking the supremum over all $\theta \in C_c^1(Q_T)$ with $\|\theta\|_X \leq 1$ and using the density of $C_c^1(Q_T)$ in X, we get

$$\begin{aligned}\|\partial_t(z_i^n - z_i)\|_{X'} \leq (|\Omega|T)^{\frac{d}{2(d+1)}}\big(&\varepsilon\|\nabla z_i^n - \nabla z_i\|_{L^2(Q_T)} + \|z_i^n\nabla z_i^n - z_i\nabla z_i\|_{L^2(Q_T)}\\ &+ \|z_i^n\nabla L_i(\cdot,\cdot,z_i^n) - z_i\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)}\\ &+ |\delta|\|z_i^n F_i(\cdot,\cdot,\hat{z}^n,\nabla\hat{z}^n) - z_i F_i(\cdot,\cdot,\hat{z},\nabla\hat{z})\|_{L^2(Q_T)}\big).\end{aligned}$$

Using the strong convergence in $L^2(Q_T)$ of $\nabla z_i^n \to \nabla z_i$, $z_i^n \to z_i$, and $\nabla L_i(\cdot,\cdot,z_i^n) \to \nabla L_i(\cdot,\cdot,z_i)$, we immediately deduce that the first three terms on the right-hand side of the previous equation vanish in the limit as $n \to \infty$, since the product of two strongly convergent sequences in $L^2(Q_T)$ is itself strongly convergent in $L^2(Q_T)$. Finally, recall that $\{\hat{z}^n\}_{n\in\mathbb{N}}$ is precisely the sequence $\{R_\mu\bar{z}^n\}_{n\in\mathbb{N}}$, and that the regularisation parameter $\mu > 0$ is fixed throughout this procedure. Thus, since $\|\bar{z}_i^n\|_{L^2(0,T;H^1(\Omega))}$ is bounded independently of n, the estimate (64) gives the uniform boundedness in $L^2(Q_T)$ of the sequences of higher derivatives $\{\partial_t \nabla \hat{z}_i^n\}_{n\in\mathbb{N}}$ and $\{\nabla^2 \hat{z}_i^n\}_{n\in\mathbb{N}}$. It then follows from the Aubin–Lions Lemma (*cf.* [71, Theorem II.5.16]) that $\nabla \hat{z}_i^n \to \nabla \hat{z}_i$ strongly in $L^2(Q_T)$, since we already knew from (69) that the sequence converged weakly in $L^2(0,T;H^1(\Omega))$ to this same limit. Combining with the strong convergence $\hat{z}_i^n \to \hat{z}_i$, we get

$$\hat{z}_i^n \to \hat{z}_i \qquad \text{strongly in } L^2(0,T;H^1(\Omega)).$$

It is then immediate from the structure of F_i given in (4) that we have the strong convergence

$$\|F_i(\cdot,\cdot,\hat{z}^n,\nabla\hat{z}^n) - F_i(\cdot,\cdot,\hat{z},\nabla\hat{z})\|_{L^2(Q_T)} \to 0 \qquad \text{as } n\to\infty,$$

whence $\|\partial_t(z_i^n - z_i)\|_{X'}$ vanishes in the limit as $n\to\infty$, again using the fact that the product of two sequences converging strongly in $L^2(Q_T)$ is itself strongly convergent in $L^2(Q_T)$. The proof is complete. $\square$

We now arrive at the existence of solutions to the regularised coupled system (60).

Proposition 5.10 (*Existence for regularised coupled system*)**.** *Fix $\varepsilon > 0$ and $z_{i,0} \in C_c^\infty(\Omega)$ to be non-negative functions such that $\int_\Omega z_{i,0}\,\mathrm{d}x = \int_\Omega u_{i,0}\,\mathrm{d}x$ for $i \in \{1,\dots,M\}$. There exists $z = (z_i)_{i=1}^M$, belonging to the space Ξ, which solves the regularised coupled system (60), i.e.,*

$$\begin{cases} \partial_t z_i = \operatorname{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon \nabla z_i] & \text{in } Q_T, \\ 0 = \nu\cdot[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon\nabla z_i] & \text{on } \Sigma_T, \\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega, \end{cases}$$

in the weak sense: for any test function $\phi \in C^1(\bar{Q}_T)$, for $i \in \{1,\dots,M\}$,

$$\langle \partial_t z_i, \phi\rangle_{X'\times X} + \int_{Q_T} \big[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon\nabla z_i\big]\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t = 0, \tag{74}$$

with $z_i(0,\cdot) = z_{i,0}$ in $(W^{1,r}(\Omega))'$. Moreover, each z_i is non-negative and conserves its initial mass, and there exists a positive constant $C = C(\Omega,T,d,\delta)$, which is independent of ε and $z_0 = (z_{i,0})_{i=1}^M$, such that, for $i \in \{1,\dots,M\}$,

$$\|z_i\|^2_{L^2(0,T;H^1(\Omega))} + \|\partial_t z_i\|_{X'} \le C\bigg(1 + \|z_{i,0}\|^2_{L^1(\Omega)} + \int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x\bigg). \tag{75}$$

Proof. *Step 1 (solution with smoothing operator R_μ in the right-hand side):* Begin by showing that for each $\mu > 0$, there exists $z = (z_i)_{i=1}^M$, belonging to the space $C^{2,1}(\bar{Q}_T)$, which solves the system

$$\begin{cases} \partial_t z_i = \operatorname{div}[z_i(\nabla z_i + L_i(t,x,z_i) + \delta F_i(t,x,R_\mu z,\nabla R_\mu z)) + \varepsilon\nabla z_i] & \text{in } Q_T, \\ 0 = \nu\cdot[z_i(\nabla z_i + L_i(t,x,z_i) + \delta F_i(t,x,R_\mu z,\nabla R_\mu z)) + \varepsilon\nabla z_i] & \text{on } \Sigma_T, \\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega, \end{cases} \tag{76}$$

in the classical sense, and that there exists a positive constant $C = C(\Omega,T,d,C_{reg},\delta)$, independent of μ,ε, and $z_{i,0}$, such that, for $i\in\{1,\dots,M\}$,

$$\|z_i\|^2_{L^2(0,T;H^1(\Omega))} + \|\partial_t z_i\|_{X'} \le C\bigg(1 + \|z_{i,0}\|^2_{L^1(\Omega)} + \int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x\bigg). \tag{77}$$

To this end, recall that $S_\varepsilon \circ R_\mu$ maps the Banach space Ξ into itself compactly (*cf.* Lemmas 4.9 and 5.9, and (12) the smallness condition on δ). Subsequently, an application of Corollary 5.1 shows that there exists $z \in \Xi$ which is a fixed point of $S_\varepsilon \circ R_\mu$. Since $S_\varepsilon \circ R_\mu$ maps Ξ into $(C^{2,1}(\bar{Q}_T))^M$, the equality in Ξ $z = S_\varepsilon(R_\mu(z))$ implies that there exists a representative of z belonging to $(C^{2,1}(\bar{Q}_T))^M$ and, additionally, such representative solves (76) in the classical sense. In view of the definition of Ξ, we automatically obtain that, for $i \in \{1, \dots, M\}$, z_i is non-negative and conserves its initial mass.

By integrating (76) against any test function $\phi \in C^1(\bar{Q}_T)$, we get, for $i \in \{1, \dots, M\}$,

$$\langle \partial_t z_i, \phi \rangle_{X' \times X} + \int_{Q_T} \big(z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,R_\mu z, \nabla R_\mu z)) + \varepsilon \nabla z_i \big) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t = 0, \tag{78}$$

and we also have $z_i(0,\cdot) = z_{i,0}$ in $(W^{1,r}(\Omega))'$, since we already know that this latter equality holds in the pointwise sense.

Meanwhile, the estimate on $\|z_i\|^2_{L^2(0,T;H^1(\Omega))}$ in (77) follows from Lemma 4.9, using the smallness of δ along with the bound

$$|F_i(t,x,R_\mu z(t,x), \nabla R_\mu z(t,x))| \leq C_F(T,\Omega)(1 + |\nabla R_\mu z(t,x)|),$$

due to (5), and the estimate (63) (*cf.* Lemma A.6). Using this latter bound and the one for $\|z_i\|^2_{L^2(0,T;H^1(\Omega))}$, we then obtain the estimate on $\|\partial_t z_i\|_{X'}$ in (77) from Lemma 4.11.

Step 2 (taking the limit $\mu \to 0$. Part I: drift terms): For each $\mu > 0$, define $z^\mu = (z_i^\mu)_{i=1}^M$ to be the solution of (76) provided by Step 1, and recall that (78) holds. Observe from the estimates (77) that $\{z^\mu\}_{\mu>0}$ is a bounded sequence in $(L^2(0,T;H^1(\Omega)))^M$, and this bound is independent of μ. Hence, as in the proof of Lemma 5.4, an application of the theorem of Banach–Alaoglu for reflexive spaces and the Aubin–Lions Lemma (*cf.* [71, Theorem II.5.16]) implies the existence of a subsequence, which we still label as $\{z^\mu\}_{\mu>0}$, converging weakly in $(L^2(0,T;H^1(\Omega)))^M$ and strongly in $(L^2(Q_T))^M$ to some $z = (z_i)_{i=1}^M \in (L^2(0,T;H^1(\Omega)))^M$, and such that $\partial_t z^\mu \overset{*}{\rightharpoonup} \partial_t z$ weakly-* in $(X')^M$. This latter weak-* convergence in $(X')^M$ is manifestly enough to pass to the limit in the first term of the weak formulation (78), i.e., the duality product, and for the final term, we have

$$\begin{aligned} I_\mu := \bigg| \int_{Q_T} & \big[z_i^\mu(\nabla z_i^\mu + \nabla L_i(t,x,z_i^\mu) + \delta F_i(t,x,R_\mu z^\mu, \nabla R_\mu z^\mu)) + \varepsilon \nabla z_i^\mu \big] \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \\ & - \int_{Q_T} \big[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z)) + \varepsilon \nabla z_i \big] \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg|, \end{aligned}$$

which can be expanded as $I_\mu \leq J_\mu + K_{1,\mu} + K_{2,\mu}$, where

$$J_\mu := \bigg| \int_{Q_T} z_i^\mu(\nabla z_i^\mu + \delta F_i(t,x,R_\mu z^\mu, \nabla R_\mu z^\mu)) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t - \int_{Q_T} z_i(\nabla z_i + \delta F_i(t,x,z,\nabla z)) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg|,$$

while

$$K_{1,\mu} := \bigg| \int_{Q_T} \varepsilon \nabla z_i^\mu \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t - \int_{Q_T} \varepsilon \nabla z_i \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg|,$$

and

$$K_{2,\mu} := \bigg| \int_{Q_T} z_i^\mu \nabla L_i(t,x,z_i^\mu) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t - \int_{Q_T} z_i \nabla L_i(t,x,z_i) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg|.$$

Observe that

$$K_{1,\mu} = \varepsilon \bigg| \int_{Q_T} (\nabla z_i^\mu - \nabla z_i) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg| \to 0,$$

in view of the weak convergence $z_i^\mu \rightharpoonup z_i$ in $L^2(0,T;H^1(\Omega))$, which naturally implies $\nabla z_i^\mu \rightharpoonup \nabla z_i$ weakly in $(L^2(Q_T))^M$. Additionally, observe that

$$K_{2,\mu} = \bigg| \int_{Q_T} (z_i^\mu \nabla L_i(t,x,z_i^\mu) - z_i \nabla L_i(t,x,z_i)) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \bigg|,$$

and

$$\begin{aligned}\|\nabla L_i(\cdot,\cdot,z_i^\mu) - \nabla L_i(\cdot,\cdot,z_i)\|^2_{L^2(Q_T)} &= \int_0^T \int_\Omega |\nabla W_i * (z_i^\mu(t,\cdot) - z_i(t,\cdot))(x)|^2 \,\mathrm{d}x\,\mathrm{d}t \\ &= \int_0^T \int_\Omega \left| \int_\Omega \nabla W_i(x-y)(z_i^\mu(t,y) - z_i(t,y))\,\mathrm{d}y \right|^2 \mathrm{d}x\,\mathrm{d}t \\ &\le |\Omega| C_L^2 \int_0^T \int_\Omega \left(\int_\Omega |z_i^\mu(t,y) - z_i(t,y)|^2 \,\mathrm{d}y \right) \mathrm{d}x\,\mathrm{d}t \\ &= |\Omega|^2 C_L^2 \|z_i^\mu - z_i\|^2_{L^2(Q_T)} \to 0 \qquad \text{as } \mu \to 0,\end{aligned}$$

where we used the boundedness from (8) and Jensen's inequality to obtain the third line, and the Fubini–Tonelli theorem to obtain the final equality, and the strong convergence $z_i^\mu \to z_i$ in $L^2(Q_T)$ to show that the limit vanishes. Thus,

$$\|\nabla L_i(\cdot,\cdot,z_i^\mu) - \nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)} \to 0 \qquad \text{as } \mu \to 0,$$

and hence, being the product of two strongly convergent sequences, we get $z_i^\mu \nabla L_i(\cdot,\cdot,z_i^\mu) \to z_i \nabla L_i(\cdot,\cdot,z_i)$ strongly in $L^2(Q_T)$, whence the Cauchy–Schwarz integral inequality yields that $K_{2,\mu} \to 0$ in the limit as $\mu \to 0$.

Step 3 (taking the limit $\mu \to 0$. Part II: cross-diffusion terms): We return to J_μ, which we write as $J_\mu \le J_{1,\mu} + |\delta| J_{2,\mu}$, where

$$J_{1,\mu} := \left| \int_{Q_T} (z_i^\mu \nabla z_i^\mu - z_i \nabla z_i) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \right|,$$

and

$$J_{2,\mu} := \left| \int_{Q_T} \big(z_i^\mu F_i(t,x,R_\mu z^\mu, \nabla R_\mu z^\mu) - z_i F_i(t,x,z,\nabla z)\big) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \right|.$$

Recall that since $z_i^\mu \to z_i$ strongly in $L^2(Q_T)$ and $\nabla z_i^\mu \rightharpoonup \nabla z_i$ weakly in $(L^2(Q_T))^M$, the product $z_i^\mu \nabla z_i^\mu$ converges weakly to $z_i \nabla z_i$ in $(L^2(Q_T))^M$, whence $J_{1,\mu} \to 0$ in the limit as μ vanishes. It remains to control $J_{2,\mu}$. Observe that an application of the triangle inequality yields

$$\begin{aligned} J_{2,\mu} \le & \left| \int_{Q_T} (z_i^\mu - z_i) F_i(t,x,R_\mu z^\mu, \nabla R_\mu z^\mu) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \right| \\ & + \left| \int_{Q_T} z_i \big(F_i(t,x,R_\mu z^\mu,\nabla R_\mu z^\mu) - F_i(t,x,z,\nabla z)\big) \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \right| =: L_{1,\mu} + L_{2,\mu}. \end{aligned}$$

Note that

$$\begin{aligned} L_{1,\mu} &\le \|\nabla\phi\|_{L^\infty(Q_T)} \max_{i\in\{1,\dots,M\}} \|F_i(\cdot,\cdot,R_\mu z^\mu, \nabla R_\mu z^\mu)\|_{L^2(Q_T)} \|z_i^\mu - z_i\|_{L^2(Q_T)} \\ &\le \|\nabla\phi\|_{L^\infty(Q_T)} C_F \big(1 + \max_{i\in\{1,\dots,M\}} \|\nabla R_\mu z^\mu\|_{L^2(Q_T)}\big) \|z_i^\mu - z_i\|_{L^2(Q_T)}, \end{aligned}$$

where we used the usual bound (5) for F_i, and from which it follows, using the estimate (63), that

$$L_{1,\mu} \le \|\nabla\phi\|_{L^\infty(Q_T)} C_F \big(1 + C_{reg} \max_{i\in\{1,\dots,M\}} \|z_i^\mu\|_{L^2(0,T;H^1(\Omega))}\big) \|z_i^\mu - z_i\|_{L^2(Q_T)}.$$

Furthermore, due to the weak convergence $z_i^\mu \rightharpoonup z_i$ in $L^2(0,T;H^1(\Omega))$ for each $i \in \{1,\dots,M\}$, we know that for each $i \in \{1,\dots,M\}$ this sequence is bounded, i.e., there exists a positive constant C_{seq} (independent of μ) such that

$$\max_{i\in\{1,\dots,M\}} \|z_i^\mu\|_{L^2(0,T;H^1(\Omega))} \le C_{seq} \qquad \forall \mu > 0.$$

Thus, we obtain

$$L_{1,\mu} \le \|\nabla\phi\|_{L^\infty(Q_T)} C_F (1 + C_{reg} C_{seq}) \|z_i^\mu - z_i\|_{L^2(Q_T)} \to 0 \qquad \text{as } \mu \to 0.$$

The term $L_{2,\mu}$ is more delicate. To study it, we expand the terms in F_i. Indeed, an application of the triangle inequality yields

$$L_{2,\mu} \le \left| \int_{Q_T} z_i \big(G_i^0(t,x,R_\mu z^\mu) - G_i^0(t,x,z) \big) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \right| \\ + \sum_{j=1}^{M} \left| \int_{Q_T} z_i \big(G_{ij}^1(t,x,R_\mu z^\mu) \nabla R_\mu z_j^\mu - G_{ij}^1(t,x,z) \nabla z_j \big) \cdot \nabla \phi \, \mathrm{d}x \, \mathrm{d}t \right| =: \Lambda_{1,\mu} + \Lambda_{2,\mu}.$$

Observe then that, using Hölder's inequality along with the mean value inequality and the uniform boundedness of the derivatives of $(G_i^0)_{i=1}^M$, we control the first term by

$$\Lambda_{1,\mu} \le \|\nabla \phi\|_{L^\infty(Q_T)} \|z_i\|_{L^2(Q_T)} \|\nabla_z G_i^0\|_{L^\infty(Q_T \times \mathbb{R}^M)} \max_{i \in \{1,\dots,M\}} \|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)}. \tag{79}$$

Then, for each fixed $i \in \{1,\dots,M\}$, we have

$$\begin{aligned} \|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)} &\le \|R_\mu z_i^\mu - R_\mu z_i\|_{L^2(Q_T)} + \|R_\mu z_i - z_i\|_{L^2(Q_T)} \\ &\le C_{reg} \|z_i^\mu - z_i\|_{L^2(Q_T)} + \|R_\mu z_i - z_i\|_{L^2(Q_T)} \to 0 \qquad \text{as } \mu \to 0, \end{aligned} \tag{80}$$

by virtue of the strong convergence $z_i^\mu \to z_i$ in $L^2(Q_T)$ and Corollary A.8, where we also used the linearity of the operator R_μ to obtain the second inequality. Thus, by returning to (79), we conclude that $\Lambda_{1,\mu}$ vanishes in the limit as $\mu \to 0$. For the term $\Lambda_{2,\mu}$, observe that, for each fixed $i, j \in \{1,\dots,M\}$, we have firstly that

$$\|G_{ij}^1(\cdot,\cdot,R_\mu z^\mu) - G_{ij}^1(\cdot,\cdot,z)\|_{L^2(Q_T)} \le \|\nabla_z G_{ij}^1\|_{L^\infty(Q_T \times \mathbb{R}^M)} \max_{i \in \{1,\dots,M\}} \|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)},$$

where, as before, we used the mean value inequality and the uniform boundedness of the derivatives of $(G_{ij}^1)_{i,j=1}^M$, so that we deduce from (80) that

$$G_{ij}^1(\cdot,\cdot,R_\mu z^\mu) \to G_{ij}^1(\cdot,\cdot,z) \qquad \text{strongly in } L^2(Q_T). \tag{81}$$

Meanwhile, for the first term, we have the following claim, to be proved later.

Claim 5.11.

$$\nabla R_\mu z_j^\mu \rightharpoonup \nabla z_j \qquad \textit{weakly in } (L^2(Q_T))^M.$$

Using the above, we then deduce from (81) that we have the weak convergence of the product, i.e.,

$$G_{ij}^1(\cdot,\cdot,R_\mu z^\mu) \nabla R_\mu z_j^\mu \rightharpoonup G_{ij}^1(\cdot,\cdot,z) \nabla z_j \qquad \text{weakly in } (L^2(Q_T))^M. \tag{82}$$

Thus, using the Hölder inequality to verify that $z_i \nabla \phi \in (L^2(Q_T))^M$, the weak convergence of (82) implies that $\Lambda_{2,\mu} \to 0$ in the limit as $\mu \to 0$. The limit function $z \in L^2(0,T;H^1(\Omega))$ thereby satisfies the weak formulation (74), and the estimates (75) follow from the weak (and weak-* for the bound in X') lower semi-continuity of the norms considered in the estimate (77). The non-negativity and mass conservation follow again directly from Lemma 5.2. The attainment of the initial data in the dual Sobolev space follows from a direct application of Lemma 5.3. The proof is complete. □

Proof of Claim 5.11. Recall from (80) that we have

$$\|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)} \to 0 \qquad \text{as } \mu \to 0.$$

It therefore follows that, given any test function $\psi \in (C_c^\infty(Q_T))^d$, by definition of weak derivative, there holds the integration by parts relation

$$\int_{Q_T} \nabla (R_\mu z_i^\mu - z_i) \cdot \psi \, \mathrm{d}x \, \mathrm{d}t = - \int_{Q_T} (R_\mu z_i^\mu - z_i) \operatorname{div} \psi \, \mathrm{d}x \, \mathrm{d}t,$$

where we note that there are no boundary terms due to the compact support of ψ. Thus, we get from the above, and the Hölder inequality, that

$$\left|\int_{Q_T} \nabla(R_\mu z_i^\mu - z_i)\cdot\psi \,\mathrm{d}x\,\mathrm{d}t\right| \le \|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)}\|\nabla\psi\|_{L^2(Q_T)} \to 0 \qquad \text{as } \mu\to 0.$$

Finally, due to the density of $C_c^\infty(Q_T)$ in $L^2(Q_T)$, it follows that given any $\psi\in(L^2(Q_T))^d$, there exists a sequence $\{\psi_m\}_{m\in\mathbb{N}}$ of elements of $C_c^\infty(Q_T)$ such that $\|\psi_m-\psi\|_{L^2(Q_T)}$ vanishes as $m\to\infty$. Then, we get

$$\begin{aligned}\left|\int_{Q_T} \nabla(R_\mu z_i^\mu - z_i)\cdot\psi \,\mathrm{d}x\,\mathrm{d}t\right| &\le \left|\int_{Q_T} \nabla(R_\mu z_i^\mu - z_i)\cdot(\psi-\psi_m) \,\mathrm{d}x\,\mathrm{d}t\right| + \left|\int_{Q_T} \nabla(R_\mu z_i^\mu - z_i)\cdot\psi_m \,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \|\nabla R_\mu z_i^\mu - \nabla z_i\|_{L^2(Q_T)}\|\psi-\psi_m\|_{L^2(Q_T)} \\ &\quad + \|R_\mu z_i^\mu - z_i\|_{L^2(Q_T)}\|\nabla\psi_m\|_{L^2(Q_T)},\end{aligned} \tag{83}$$

where, for the second term in the final line, we integrated by parts and then used the Hölder inequality. Now observe that, for the first term, the triangle inequality yields

$$\begin{aligned}\|\nabla R_\mu z_i^\mu - \nabla z_i\|_{L^2(Q_T)} &\le \|\nabla R_\mu z_i^\mu\|_{L^2(Q_T)} + \|\nabla z_i\|_{L^2(Q_T)} \\ &\le \|R_\mu z_i^\mu\|_{L^2(0,T;H^1(\Omega))} + \|\nabla z_i\|_{L^2(0,T;H^1(\Omega))}.\end{aligned}$$

However, since $z_i^\mu \in L^2(0,T;H^1(\Omega))$ for each fixed $\mu>0$, it follows from estimate (63) (*cf.* (104) of Lemma A.6) that, for each fixed $\mu>0$,

$$\|R_\mu z_i^\mu\|_{L^2(0,T;H^1(\Omega))} \le C_{reg}\|z_i^\mu\|_{L^2(0,T;H^1(\Omega))},$$

whence, since $\{z_i^\mu\}_{\mu>0}$ is a bounded sequence in $L^2(0,T;H^1(\Omega))$ on account of being a weakly convergent sequence therein, there exists a positive constant C_{seq} independent of m and μ, but depending on $\|\nabla z_i\|_{L^2(0,T;H^1(\Omega))}$, such that

$$\|\nabla R_\mu z_i^\mu - \nabla z_i\|_{L^2(Q_T)} \le C_{seq} \qquad \forall\mu>0.$$

In turn, by first taking the limit as $m\to\infty$ and then making μ vanish, it follows that the right-hand side of (83) tends to zero as $\mu\to 0$. To summarise, given any $\psi\in(L^2(Q_T))^d$,

$$\int_{Q_T} \nabla(R_\mu z_i^\mu - z_i)\cdot\psi \,\mathrm{d}x\,\mathrm{d}t \to 0 \qquad \text{as } \mu\to 0,$$

i.e., $\nabla R_\mu z_i^\mu \rightharpoonup \nabla z_i$ weakly in $L^2(Q_T)$, as required. □

6. Vanishing diffusivity and proof of the main result

In this section, we take the vanishing diffusivity limit as $\varepsilon\to 0$ in the regularised coupled system (60).

Lemma 6.1 (*Existence for the Degenerate System*)**.** *Fix $z_{i,0}\in C_c^\infty(\Omega)$, for $i\in\{1,\dots,M\}$, to be non-negative functions such that $\int_\Omega z_{i,0}\,\mathrm{d}x = \int_\Omega u_{i,0}\,\mathrm{d}x$ for $i\in\{1,\dots,M\}$. There exists $z=(z_i)_{i=1}^M$, belonging to the space Ξ, which solves the system*

$$\begin{cases}\partial_t z_i = \mathrm{div}[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z))] & \text{in } Q_T,\\ 0 = \nu\cdot[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z))] & \text{on } \Sigma_T,\\ z_i(0,\cdot) = z_{i,0} & \text{on } \Omega,\end{cases} \tag{84}$$

in the weak sense prescribed by Definition 2.3, with $z_i(0,\cdot) = z_{i,0}$ in $(W^{1,r}(\Omega))'$. Moreover, each z_i is non-negative and conserves its initial mass, and there exists a positive constant $C = C(\Omega, T, d, \delta)$ such that, for $i \in \{1,\dots,M\}$,

$$\|z_i\|^2_{L^2(0,T;H^1(\Omega))} + \|\partial_t z_i\|_{X'} \le C\Big(1 + \|z_{i,0}\|^2_{L^1(\Omega)} + \int_\Omega z_{i,0}\log z_{i,0}\,\mathrm{d}x\Big). \tag{85}$$

Proof. The proof is similar to that of Proposition 5.10. Recall the weak formulation prescribed by Definition 2.3, i.e., for any test function $\phi \in C^1(\bar{Q}_T)$, for $i \in \{1,\dots,M\}$,

$$\langle \partial_t z_i, \phi\rangle_{X'\times X} + \int_{Q_T} \big[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z))\big]\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t = 0. \tag{86}$$

For this proof, we define $z^\varepsilon = (z_i^\varepsilon)_{i=1}^M$ to be the weak solution of (74) for each $\varepsilon > 0$, provided by Proposition 5.10.

Observe from the estimates (75) that $\{z^\varepsilon\}_{\varepsilon>0}$ is a bounded sequence in $(L^2(0,T;H^1(\Omega)))^M$, and this bound is independent of ε. Hence, as in the proofs of Lemma 5.4 and Proposition 5.10, an application of the theorem of Banach–Alaoglu for reflexive spaces and the Aubin–Lions Lemma (*cf.* [71, Theorem II.5.16]) implies the existence of a subsequence, which we still label as $\{z^\varepsilon\}_{\varepsilon>0}$, converging weakly in $(L^2(0,T;H^1(\Omega)))^M$ and strongly in $(L^2(Q_T))^M$ to some $z = (z_i)_{i=1}^M \in (L^2(0,T;H^1(\Omega)))^M$, and such that $\partial_t z^\varepsilon \overset{*}{\rightharpoonup} \partial_t z$ weakly-* in $(X')^M$; see Lemma 4.11 for the estimate independent of ε. This latter weak-* convergence in $(X')^M$ is manifestly enough to pass to the limit in the first term of the weak formulation (74), i.e., the duality product. For the remaining term we have

$$\begin{aligned} I_\varepsilon := \Big| &\int_{Q_T} \big[z_i^\varepsilon(\nabla z_i^\varepsilon + \nabla L_i(t,x,z_i^\varepsilon) + \delta F_i(t,x,z^\varepsilon,\nabla z^\varepsilon)) + \varepsilon\nabla z_i^\varepsilon\big]\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t \\ &\qquad - \int_{Q_T}\big[z_i(\nabla z_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,z,\nabla z))\big]\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\Big|, \end{aligned}$$

which can be expanded as $I_\varepsilon \le J_\varepsilon + K_{1,\varepsilon} + K_{2,\varepsilon}$, where

$$J_\varepsilon := \Big|\int_{Q_T} z_i^\varepsilon(\nabla z_i^\varepsilon + \delta F_i(t,x,z^\varepsilon,\nabla z^\varepsilon))\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t - \int_{Q_T} z_i(\nabla z_i + \delta F_i(t,x,z,\nabla z))\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\Big|,$$

while

$$K_{1,\varepsilon} := \varepsilon\Big|\int_{Q_T}\nabla z_i^\varepsilon\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\Big|,$$

and

$$K_{2,\varepsilon} := \Big|\int_{Q_T} z_i^\varepsilon\nabla L_i(t,x,z_i^\varepsilon)\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t - \int_{Q_T} z_i\nabla L_i(t,x,z_i)\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\Big|.$$

Observe that, using Hölder's inequality,

$$K_{1,\varepsilon} \le \varepsilon\|\nabla z_i^\varepsilon\|_{L^2(Q_T)}\|\nabla\phi\|_{L^2(Q_T)},$$

and note that $\|\nabla z_i^\varepsilon\|_{L^2(Q_T)} \le \|z_i^\varepsilon\|_{L^2(0,T;H^1(\Omega))}$, which is uniformly bounded independently of ε on account of $\{z_i^\varepsilon\}_{\varepsilon>0}$ being weakly convergent in the space $L^2(0,T;H^1(\Omega))$. It follows that there exists a positive constant C independent of ε such that

$$K_{1,\varepsilon} \le \varepsilon C\|\nabla\phi\|_{L^2(Q_T)} \to 0 \qquad \text{as } \varepsilon \to 0.$$

As for $K_{2,\varepsilon}$, we estimate

$$\begin{aligned}\|\nabla L_i(\cdot,\cdot,z_i^\varepsilon)-\nabla L_i(\cdot,\cdot,z_i)\|^2_{L^2(Q_T)} &= \int_0^T\int_\Omega |\nabla W_i*(z_i^\varepsilon(t,\cdot)-z_i(t,\cdot))(x)|^2\,\mathrm{d}x\,\mathrm{d}t\\ &= \int_0^T\int_\Omega \left|\int_\Omega \nabla W_i(x-y)(z_i^\varepsilon(t,y)-z_i(t,y))\,\mathrm{d}y\right|^2\mathrm{d}x\,\mathrm{d}t\\ &\le |\Omega|C_L^2\int_0^T\int_\Omega\left(\int_\Omega |z_i^\varepsilon(t,y)-z_i(t,y)|^2\,\mathrm{d}y\right)\mathrm{d}x\,\mathrm{d}t\\ &= |\Omega|^2C_L^2\|z_i^\varepsilon-z_i\|^2_{L^2(Q_T)}\to 0\qquad \text{as } \varepsilon\to 0,\end{aligned}$$

where we used the boundedness from (8) and Jensen's inequality to obtain the third line, and the Fubini–Tonelli theorem to obtain the final equality, and the strong convergence $z_i^\varepsilon\to z_i$ in $L^2(Q_T)$ to show that the limit vanishes. Thus,

$$\|\nabla L_i(\cdot,\cdot,z_i^\varepsilon)-\nabla L_i(\cdot,\cdot,z_i)\|_{L^2(Q_T)}\to 0\qquad \text{as } \varepsilon\to 0,$$

and hence, being the product of two strongly convergent sequences, we get $z_i^\varepsilon\nabla L_i(\cdot,\cdot,z_i^\varepsilon)\to z_i\nabla L_i(\cdot,\cdot,z_i)$ strongly in $L^2(Q_T)$, whence the Cauchy–Schwarz integral inequality yields that $K_{2,\varepsilon}\to 0$ in the limit as $\varepsilon\to 0$.

We return to J_ε, which we write as $J_\varepsilon\le J_{1,\varepsilon}+|\delta|J_{2,\varepsilon}$, where

$$J_{1,\varepsilon}:=\left|\int_{Q_T}(z_i^\varepsilon\nabla z_i^\varepsilon-z_i\nabla z_i)\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\right|,$$

and

$$J_{2,\varepsilon}:=\left|\int_{Q_T}\big(z_i^\varepsilon F_i(t,x,z^\varepsilon,\nabla z^\varepsilon)-z_iF_i(t,x,z,\nabla z)\big)\cdot\nabla\phi\,\mathrm{d}x\,\mathrm{d}t\right|.$$

Recall that since $z_i^\varepsilon\to z_i$ strongly in $L^2(Q_T)$ and $\nabla z_i^\varepsilon\rightharpoonup\nabla z_i$ weakly in $(L^2(Q_T))^M$, the product $z_i^\varepsilon\nabla z_i^\varepsilon$ converges weakly to $z_i\nabla z_i$ in $(L^2(Q_T))^M$, whence $J_{1,\varepsilon}\to 0$ in the limit as ε vanishes. It remains to control $J_{2,\varepsilon}$. We already remarked that $z_i^\varepsilon\to z_i$ strongly in $L^2(Q_T)$. It therefore suffices to show that

$$F_i(\cdot,\cdot,z^\varepsilon,\nabla z^\varepsilon)\rightharpoonup F_i(\cdot,\cdot,z,\nabla z)\qquad \text{weakly in } L^2(Q_T). \tag{87}$$

To verify this, we expand the above terms in F_i. Observe that

$$F_i(t,x,z^\varepsilon,\nabla z^\varepsilon)=G_i^0(t,x,z^\varepsilon)+\sum_{j=1}^M G_{ij}^1(t,x,z^\varepsilon)\nabla z_j^\varepsilon,$$

and so, using the mean value theorem and the uniform boundedness of the derivatives of $(G_i)_{i=1}^M$, we see that

$$\|G_i^0(\cdot,\cdot,z^\varepsilon)-G_i^0(\cdot,\cdot,z)\|_{L^2(Q_T)}\le\|\nabla_zG_i^0\|_{L^\infty(Q_T\times\mathbb{R}^M)}\max_{i\in\{1,\dots,M\}}\|z_i^\varepsilon-z_i\|_{L^2(Q_T)}, \tag{88}$$

and the right-hand side vanishes due to the strong convergence $z_i^\varepsilon\to z_i$ in $L^2(Q_T)$. For the other term, we have the following claim, to be proved later.

Claim 6.2.

$$\sum_{j=1}^M G_{ij}^1(\cdot,\cdot,z^\varepsilon)\nabla z_j^\varepsilon\rightharpoonup\sum_{j=1}^M G_{ij}^1(\cdot,\cdot,z)\nabla z_j\qquad \textit{weakly in } L^2(Q_T).$$

It then follows immediately from the previous claim and (88) that we obtain the desired weak convergence (87). Now, due to the strong convergence $z_i^\varepsilon\to z_i$ strongly in $L^2(Q_T)$, it follows that we have weak convergence of the product, i.e.,

$$z_i^\varepsilon F_i(\cdot,\cdot,z^\varepsilon,\nabla z^\varepsilon)\rightharpoonup z_iF_i(\cdot,\cdot,z,\nabla z)\qquad \text{weakly in } L^2(Q_T),$$

whence it follows that $J_{2,\varepsilon} \to 0$ as $\varepsilon \to 0$. The limit function $z \in L^2(0,T;H^1(\Omega))$ thereby satisfies the weak formulation (86), and the estimates (85) follow from the weak (and weak-* for the bound in X') lower semicontinuity of the norms considered in the estimate (75). The non-negativity and mass conservation follow again directly from Lemma 5.2, while the attainment of the initial data in the dual Sobolev space follows from a direct application of Lemma 5.3. $\square$

Proof of Claim 6.2. Fix $j \in \{1,\dots,M\}$ in the sum, and by testing against any $\psi \in (C_c^\infty(Q_T))^d$, we obtain

$$\left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi \,\mathrm{d}x\,\mathrm{d}t\right| \leq \left|\int_{Q_T} (G^1_{ij}(t,x,z^\varepsilon) - G^1_{ij}(t,x,z))\nabla z^\varepsilon_j\cdot\psi\,\mathrm{d}x\,\mathrm{d}t\right| + \left|\int_{Q_T} G^1_{ij}(t,x,z)(\nabla z^\varepsilon_j - \nabla z_j)\cdot\psi\,\mathrm{d}x\,\mathrm{d}t\right|,$$

and the right-hand side of the above is bounded by

$$\|G^0_i(\cdot,\cdot,z^\varepsilon) - G^0_i(\cdot,\cdot,z)\|_{L^2(Q_T)}\|\nabla z^\varepsilon_j\|_{L^2(Q_T)}\|\psi\|_{L^\infty(Q_T)} + \left|\int_{Q_T}(\nabla z^\varepsilon_j - \nabla z_j)\cdot(G^1_{ij})^T(t,x,z)\psi\,\mathrm{d}x\,\mathrm{d}t\right|,$$

where $(G^1_{ij})^T$ is the transpose of the matrix G^1_{ij}. The first term on the right-hand side of the above vanishes as $\varepsilon \to 0$ using the strong convergence $z^\varepsilon_i \to z_i$ strongly in $L^2(Q_T)$ and an estimate analogous to the one in (88) and the uniform boundedness of $\|\nabla z^\varepsilon_j\|_{L^2(Q_T)}$ independently of ε, on account of the weak convergence $z^\varepsilon_i \rightharpoonup z_i$ in $L^2(0,T;H^1(\Omega))$. The second term on the right-hand side of the above also vanishes, on account of $(G^1_{ij})^T\psi \in L^\infty(Q_T) \subset L^2(Q_T)$ and the weak convergence $z^\varepsilon_j \rightharpoonup z_j$ in $L^2(0,T;H^1(\Omega))$.

We have therefore shown that, given any $\psi \in (C_c^\infty(Q_T))^d$, we have

$$\left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi\,\mathrm{d}x\,\mathrm{d}t\right| \to 0 \qquad \text{as } \varepsilon \to 0. \tag{89}$$

Now fix any $\psi \in (L^2(Q_T))^d$. By density of $C_c^\infty(Q_T)$ in $L^2(Q_T)$, there exists a sequence $(\psi_m)_{m\in\mathbb{N}}$ of elements of $(C_c^\infty(Q_T))^d$ converging strongly to ψ in the sense of $(L^2(Q_T))^d$. Then, by splitting the term $|\int_{Q_T}\big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi\,\mathrm{d}x\,\mathrm{d}t|$ as

$$\left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot(\psi-\psi_m)\,\mathrm{d}x\,\mathrm{d}t\right| + \left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi_m\,\mathrm{d}x\,\mathrm{d}t\right|,$$

which is itself bounded by

$$\|G^1_{ij}\|_{L^\infty(Q_T\times\mathbb{R}^M)}(\|\nabla z^\varepsilon_j\|_{L^2(Q_T)} + \|\nabla z_j\|_{L^2(Q_T)})\|\psi-\psi_m\|_{L^2(Q_T)} + \left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi_m\,\mathrm{d}x\,\mathrm{d}t\right|.$$

Using the weak convergence in $L^2(0,T;H^1(\Omega))$ to deduce the boundedness of $\|\nabla z^\varepsilon_j\|_{L^2(Q_T)}$, independently of ε, we now use an argument identical to the one in the proof of Claim 5.11 to deduce that

$$\left|\int_{Q_T} \big(G^1_{ij}(t,x,z^\varepsilon)\nabla z^\varepsilon_j - G^1_{ij}(t,x,z)\nabla z_j\big)\cdot\psi\,\mathrm{d}x\,\mathrm{d}t\right| \to 0 \qquad \text{as } \varepsilon \to 0,$$

for any $\psi \in (L^2(Q_T))^d$, as required. $\square$

This completes the vanishing diffusivity procedure. It remains to relax the assumption on the initial data, which we also do by a limiting strategy. This is contained below, which is the proof of the main result.

Proof of Theorem 1. Begin by assuming $p \in (1,\infty)$. In view of the density of $C_c^\infty(\Omega)$ in $L^p(\Omega)$, given $u_{i,0} \in L^p(\Omega)$ as in the statement of the theorem, there exists a sequence $\{z_{i,0}^m\}_{m\in\mathbb{N}}$ of elements of $C_c^\infty(\Omega)$, all of which are non-negative and satisfy the initial mass assumption $\int_\Omega z_{i,0}^m \,\mathrm{d}x = \int_\Omega u_{i,0}\,\mathrm{d}x$ as per Remark 4.1, such that

$$\|u_{i,0} - z_{i,0}^m\|_{L^p(\Omega)} \to 0 \qquad \text{as } m \to \infty,$$

for each $i \in \{1,\dots,M\}$. An explicit construction for such an approximating sequence is to set

$$z_{i,0}^m(x) := \big((u_{i,0}\mathbb{1}_\Omega) * \rho_m(x)\big)\eta_{\sigma(m)}(x)\frac{\|u_{i,0}\|_{L^1(\Omega)}}{\|\big((u_{i,0}\mathbb{1}_\Omega) * \rho_m\big)\eta_{\sigma(m)}\|_{L^1(\Omega)}} \qquad \forall x \in \Omega,$$

where ρ_m is the usual Friedrichs mollifier, η_m is a smooth non-negative cutoff function chosen such that $\eta \equiv 1$ on $\{x \in \Omega : d(x,\partial\Omega) \geq 1/m\}$ and $\eta \equiv 0$ outside $\{x \in \Omega : d(x,\partial\Omega) \geq 1/2\, m\}$, and σ is an appropriate scaling function depending on p, d (i.e. $\sigma(m) = m^q$ for some exponent $q = q(p,d)$ suitably chosen).

For the rest of this proof, for each $m \in \mathbb{N}$, we define $z^m = (z_i^m)_{i=1}^M$ to be the weak solution of (86) with initial data $z_{i,0}^m$, for $i \in \{1,\dots,M\}$, provided by Lemma 6.1. Observe from the estimates (85) that $\{z^m\}_{m\in\mathbb{N}}$ is a bounded sequence in $(L^2(0,T;H^1(\Omega)))^M$, and this bound is independent of m. Hence, as in the proofs of Lemma 5.4, Proposition 5.10, and Lemma 6.1, an application of the theorem of Banach–Alaoglu for reflexive spaces and the Aubin–Lions Lemma (*cf.* [71, Theorem II.5.16]) implies the existence of a subsequence, which we still label as $\{z^m\}_{m\in\mathbb{N}}$, converging weakly in $(L^2(0,T;H^1(\Omega)))^M$ and strongly in $(L^2(Q_T))^M$ to some $u = (u_i)_{i=1}^M \in (L^2(0,T;H^1(\Omega)))^M$, and such that $\partial_t z^m \overset{*}{\rightharpoonup} \partial_t u$ weakly-* in $(X')^M$. This latter weak-* convergence in $(X')^M$ is manifestly enough to pass to the limit in the first term of the weak formulation (86), i.e., the duality product, and for the final term, we have

$$\begin{aligned} I_m := \Bigg| \int_{Q_T} &\left[z_i^m(\nabla z_i^m + \nabla L_i(t,x,z_i^m) + \delta F_i(t,x,z^m,\nabla z^m))\right] \cdot \nabla\phi \,\mathrm{d}x\,\mathrm{d}t \\ &- \int_{Q_T} \left[u_i(\nabla u_i + \nabla L_i(t,x,z_i) + \delta F_i(t,x,u,\nabla u))\right] \cdot \nabla\phi\,\mathrm{d}x\,\mathrm{d}t\Bigg|, \end{aligned}$$

Following a procedure identical to that of the proof of Lemma 6.1, we deduce that $I_m \to 0$ as $m \to \infty$. It follows that the limit function $z \in (L^2(0,T;H^1(\Omega)))^M$ thereby satisfies the weak formulation (86).

The estimates (14)–(15) follow from the weak (and weak-* for the bound in X') lower semicontinuity of the norms considered in the estimate (85). Indeed, note that we chose $z_{i,0}^m$ at the start of the proof such that $\|z_{i,0}^m\|_{L^1(\Omega)} = \|u_{i,0}\|_{L^1(\Omega)}$ for all $m \in \mathbb{N}$. Meanwhile, the function $f : x \mapsto (x\log x)\mathbb{1}_{[0,\infty)}(x)$ is continuous and satisfies the global bound $|f(x)| \leq C(1+|x|^p)$ for some positive constant C depending only on p. As a result, by a consequence of the Generalised Dominated Convergence Theorem, f maps $L^p(\Omega)$ continuously into $L^1(\Omega)$. It therefore follows that, since $z_{i,0}^m \to u_{i,0}$ strongly in $L^p(\Omega)$,

$$z_{i,0}^m \log z_{i,0}^m \to u_{i,0}\log u_{i,0} \qquad \text{in } L^1(\Omega),$$

and hence $\lim_{m\to\infty}\int_\Omega z_{i,0}^m \log z_{i,0}^m\,\mathrm{d}x = \int_\Omega u_{i,0}\log u_{i,0}\,\mathrm{d}x$, as required.

The non-negativity and mass conservation follow again from Lemma 5.2. The convergence to the initial data in the sense of the third point of Definition 2.3 follows from a direct application of Lemma 5.3.

In the case $p = \infty$, we have $u_{i,0} \in L^q(\Omega)$ for any finite $q \in (1,\infty)$, and we can follow the same argument as before, approximating the initial data in $L^q(\Omega)$—as opposed to in $L^\infty(\Omega)$, which (in general) cannot be done using smooth compactly supported functions. The proof is complete. ☐

Acknowledgements

L. Alasio was partially supported by the Gran Sasso Science Institute and by the Paris Region Fellowship Programme (945298-ParisRegionFP). M. Bruna was supported by a Royal Society University Research

Fellowship (URF/R1/180040). S. Schulz was supported by a Royal Society Research Fellows Enhancement Award (RGF/EA/181043). This work is the result of a collaboration that started at the workshop "Recent Advances in Degenerate Parabolic Systems with Applications to Mathematical Biology" (jointly funded by the BOUM programme of SMAI and by the support of the ERC advanced grant Adora), which took place at the Laboratoire Jacques-Louis Lions (UPMC Paris VI) in February 2020.

Appendix A

A.1. Proof of Lemma 4.8 : Existence of solutions to the regularised frozen system

Proof of Lemma 4.8. We recast the problem as one with homogeneous initial condition defining $v_i(t,x) := z_i(t,x) - z_{i,0}(x)$ for each $i \in \{1,\dots,M\}$. Problem (26) may be rewritten in diagonal non-divergence form:

$$\begin{cases} \partial_t v_i + L v_i = 0 & \text{in } Q_T, \\ 0 = \nu \cdot \left[(v_i + z_{i,0} + \varepsilon)\nabla v_i + (\nabla z_{i,0} + \nabla L_i(t,x,v_i+z_{i,0}) + \delta \bar{F}_i) v_i \right] & \text{on } \Sigma_T, \\ v_i(0,\cdot) = 0 & \text{on } \Omega, \end{cases} \tag{90}$$

with

$$\begin{cases} \quad Lv = -\sum_{j,k=1}^{d} a^{jk}(v)\partial^2_{jk} v + b(t,x,v,\nabla v), \\ \quad a^{jk}(v) = (v + z_{i,0} + \varepsilon)\delta^{jk}, \\ b(t,x,v,p) = -|p|^2 - (2\nabla z_{i,0} + \nabla L_i(t,x,v+z_{i,0}) + \delta \bar{F}_i)\cdot p \\ \qquad - \operatorname{div}[\nabla z_{i,0} + \nabla L_i(t,x,v+z_{i,0}) + \delta \bar{F}_i] v - \operatorname{div}[z_{i,0}\nabla L_i(t,x,v+z_{i,0})] \\ \qquad - \operatorname{div}[z_{i,0}(\nabla z_{i,0} + \delta \bar{F}_i) + \varepsilon \nabla z_{i,0}]. \end{cases} \tag{91}$$

Thanks to (28)–(29) and (8), conditions (a), (b), and (c) of [59, Theorem 7.4 in Section 7 of Chapter 5] (which rely on estimates (7.4)–(7.6), (7.15), (7.34), and (7.36) therein) are satisfied by the scalar equation for each i in question (since the system is diagonal) in the form with homogeneous initial condition.

Subsequently there exists a unique weak solution (in the sense of Definition 4.3) of the no-flux problem with homogeneous initial condition (90) in the space of functions with continuous derivatives of the type $\partial_t^{n_1}\partial_{x_j}^{n_2}$ for $2n_1 + n_2 < 3$; denoted by $H^{3,\frac{3}{2}}(\bar{Q}_T)$ in [59]. Note in particular that this class of functions is contained in $C^{2,1}(\bar{Q}_T)$ and that uniqueness in such space can be proved by standard methods.

Notice that the requirement that F_i be C^2 in each argument is clear from the previous estimates on the derivatives of b, since, for instance, one needs to bound $\partial_t \operatorname{div}\big(F_i(t,x,\bar{z}(t,x),\nabla \bar{z}(t,x))\big)$. □

A.2. Proof of Lemma 4.12 Quantitative second derivative estimate for the regularised frozen system

Throughout this section, it will be used that the regularised frozen system (26) is diagonal. We therefore consider the single equation

$$\begin{cases} \partial_t w = \operatorname{div}\left[w(\nabla w + \nabla L + \delta \bar{F}) + \varepsilon \nabla w \right] & \text{in } Q_T, \\ 0 = \nu \cdot \left[w(\nabla w + \nabla L + \delta \bar{F}) + \varepsilon \nabla w \right] & \text{on } \Sigma_T, \\ w(0,\cdot) = w_0 & \text{on } \Omega, \end{cases} \tag{92}$$

where we omitted the i subscripts for clarity of presentation, and use the notation established in (27). Recall that (as per Remark 4.2) $\bar{F}$ is bounded in C^2-norm and (as per Remark 4.1) $w_0 \in C_c^\infty(\Omega)$. We already know from Lemmas 4.8 that there exists a non-negative $w \in C^{2,1}(\bar{Q}_T)$ solving the above equation in the classical sense, which satisfies the estimates of Lemma 4.9.

To begin with, we prove a quantitative L^∞-bound on the solution of the regularised frozen system (26).

Lemma A.1 (L^∞-*bound*)**.** *Suppose that* $z = (z_i)_{i=1}^M$ *is a* $C^{2,1}(\bar{Q}_T)$ *solution of the regularised frozen system* (26)*. Then, there holds, for each* $i \in \{1, \dots, M\}$,

$$\|z_i\|_{L^\infty(Q_T)} \le C(\varepsilon, \delta, T, \Omega, C_L, \|z_{i,0}\|_{L^\infty(\Omega)}, \|\bar{F}_i\|_{L^\infty(Q_T)}), \tag{93}$$

where the right-hand side is a positive quantity depending only on the parameters in its parentheses.

Proof. We only write the proof for $d \ge 2$. By multiplying (92) by the continuously differentiable function qw^{q-1} (for any finite $q \ge 2$) and by closely following the argument in [56, Proof of Theorem 3.1], we obtain, for every $t \in [0, T]$,

$$\begin{aligned} \frac{d}{dt} \int_\Omega w^q \,\mathrm{d}x + \varepsilon \int_\Omega |\nabla(w^{q/2})|^2 \,\mathrm{d}x &\le \frac{q(q-1)}{\varepsilon} \int_\Omega w^q (|\nabla L|^2 + \delta^2 |\bar{F}|^2) \,\mathrm{d}x \\ &\le C(\varepsilon, \delta, C_L, \|w_0\|_{L^1(\Omega)}, \|\bar{F}\|_{L^\infty(Q_T)}) \left(q^2 \int_\Omega w^q \,\mathrm{d}x \right), \end{aligned} \tag{94}$$

where we used $|\nabla L(\cdot, \cdot, w)| \le \|V\|_{C^1(\mathbb{R}^{d+1})} + |(\nabla W * w(t, \cdot))| \le \|V\|_{C^1(\mathbb{R}^{d+1})} + \|W\|_{C^1(\mathbb{R}^d)} \|w_0\|_{L^1(\Omega)}$, using Hölder's inequality for the convolution, to get the bound $\|\nabla L\|_{L^\infty(Q_T)} \le C_L(1 + \|w_0\|_{L^1(\Omega)})$. We now use the Gagliardo–Nirenberg inequality to write

$$\int_\Omega w^q \,\mathrm{d}x = \|w^{q/2}\|^2_{L^2(\Omega)} \le C(\Omega) \big(\|\nabla(w^{q/2})\|^{2\gamma}_{L^2(\Omega)} \|w^{q/2}\|^{2(1-\gamma)}_{L^1(\Omega)} + \|w^{q/2}\|^2_{L^1(\Omega)} \big),$$

for $d \ge 3$, where $\gamma = d/(d+2)$. The case $d = 2$ can be dealt with analogously using Ladyzhenskaia's inequality. Then, using the above and the weighted Young inequality, the right-hand side of (94) is bounded by

$$\frac{\varepsilon}{2} \|\nabla(w^{q/2})\|^2_{L^2(\Omega)} + C(\varepsilon, \delta, \Omega, d, C_L, \|w_0\|_{L^1(\Omega)}, \|\bar{F}\|_{L^\infty(Q_T)}) q^{\frac{2\gamma}{1-\gamma}} \|w^{q/2}\|^2_{L^1(\Omega)}.$$

Using the Gagliardo–Nirenberg and weighted Young inequalities again to bound $\|\nabla(w^{q/2})\|_{L^2(\Omega)}$ from below by $C_1(\Omega, d)\|w^{q/2}\|^2_{L^2(\Omega)} - C_2(\Omega, d)\|w^{q/2}\|^2_{L^1(\Omega)}$ and returning to (94), we obtain, for every $t \in [0, T]$,

$$\frac{d}{dt} \int_\Omega w^q \,\mathrm{d}x + \int_\Omega w^q \,\mathrm{d}x \le C(\varepsilon, \delta, \Omega, d, C_L, \|w_0\|_{L^1(\Omega)}, \|\bar{F}\|_{L^\infty(Q_T)}) \left(q^{2+\frac{2\gamma}{1-\gamma}} \left(\int_\Omega w^{q/2} \,\mathrm{d}x \right)^2 + q^2 \right),$$

and the above holds for every $q \in [2, \infty)$. We now conclude using the iterative result [56, Lemma 3.2], as per [56, Proof of Theorem 3.1]. □

With the previous estimate in hand, we proceed to the bound on the second derivative.

Proof of Lemma 4.12. We neglect the drift terms for the time being, for clarity of presentation, though we show how to treat them at the end of the proof. Hence, we restrict our focus to the higher regularity of the single equation

$$\begin{cases} \partial_t w = \operatorname{div} \big[w(\nabla w + \delta \bar{F}) + \varepsilon \nabla w \big] & \text{in } Q_T, \\ 0 = \nu \cdot \big[w(\nabla w + \delta \bar{F}) + \varepsilon \nabla w \big] & \text{on } \Sigma_T, \\ w(0, \cdot) = w_0 & \text{on } \Omega, \end{cases}$$

Recall that, by Lemma 4.9, there holds

$$\|w\|^2_{L^2(0,T;H^1(\Omega))} \le C(\varepsilon, \delta, T, |\Omega|, \|w_0\|_{L^\infty(\Omega)}, \|\bar{F}\|_{L^\infty(Q_T)}), \tag{95}$$

where we bound the right-hand side of (34) using $\|\bar{F}\|^2_{L^2(Q_T)} \le T|\Omega|\|\bar{F}\|^2_{L^\infty(Q_T)}$, $\|w_0\|_{L^1(\Omega)} \le |\Omega|\|w_0\|_{L^\infty(\Omega)}$, and $\int_\Omega w_0 \log w_0 \,\mathrm{d}x \le \int_\Omega w_0(\log w_0)_+ \,\mathrm{d}x \le C|\Omega|(1+\|w_0\|^2_{L^\infty(\Omega)})$ for some universal constant C.

Step 1 (H^2 estimate via a nonlinear transformation): Define the new (non-negative) function

$$\psi(t,x) := \frac{1}{2}w(t,x)^2 + \varepsilon w(t,x) \qquad \forall (t,x) \in \bar{Q}_T, \tag{96}$$

from which it follows that we may write, since w is non-negative itself,

$$w(t,x) = -\varepsilon + \sqrt{\varepsilon^2 + 2\psi(t,x)} \qquad \forall (t,x) \in \bar{Q}_T,$$

and we note the formula

$$\partial_t w = \frac{\partial_t \psi}{\sqrt{\varepsilon^2 + 2\psi}} \qquad \text{in } \bar{Q}_T. \tag{97}$$

The evolution equation for w may be rewritten as

$$\begin{cases} \partial_t \psi = (\varepsilon^2 + 2\psi)^{1/2} \operatorname{div}\left[\nabla\psi + \delta w \bar{F}\right] & \text{in } Q_T, \\ 0 = \nu \cdot \left[\nabla\psi + \delta w\bar{F}\right] & \text{on } \Sigma_T, \\ \psi(0,\cdot) = \frac{1}{2}w_0^2 + \varepsilon w_0 & \text{on } \Omega. \end{cases} \tag{98}$$

We now test the above against $\operatorname{div}[\nabla\psi + \delta w\bar{F}]$. An integration by parts on the left-hand side (with no boundary terms because of the no-flux condition) yields

$$-\int_\Omega \partial_t \nabla\psi \cdot [\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x = \int_\Omega (\varepsilon^2 + 2\psi)^{1/2}(\Delta\psi + \delta \operatorname{div}(w\bar{F}))^2\,\mathrm{d}x \ge \varepsilon \int_\Omega (\Delta\psi + \delta\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x,$$

where the final inequality follows from the non-negativity of ψ. Thus,

$$\frac{1}{2}\frac{d}{dt}\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2 \mathrm{d}x + \varepsilon\int_\Omega (\Delta\psi + \delta\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x \le -\delta\int_\Omega \partial_t(w\bar{F})\cdot[\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x. \tag{99}$$

The right-hand side may be expanded as

$$-\delta\int_\Omega (\partial_t w)\bar{F}\cdot[\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x - \delta\int_\Omega w(\partial_t\bar{F})\cdot[\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x, \tag{100}$$

and, by the Cauchy–Schwarz integral inequality, the second term of the above is bounded above by

$$\delta^2|\Omega|^{1/2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^2(\Omega)}\|\bar{F}\|_{L^\infty(Q_T)} + \delta\|w\|_{L^\infty(\Omega)}\|\partial_t\bar{F}\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)}.$$

The first term in (100) may be rewritten, using the formula (97) and Eq. (98), as

$$-\delta\int_\Omega \operatorname{div}[\nabla\psi + \delta w\bar{F}]\bar{F}\cdot[\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x = -\delta\int_\Omega (\Delta\psi + \delta\operatorname{div}(w\bar{F}))\bar{F}\cdot[\nabla\psi + \delta w\bar{F}]\,\mathrm{d}x.$$

Using the Cauchy–Schwarz integral inequality and the Young inequality, the right-hand side of the above is bounded above by

$$\frac{\varepsilon}{2}\int_\Omega (\Delta\psi + \delta\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x + \frac{\delta^2}{2\varepsilon}\|\bar{F}\|^2_{L^\infty(Q_T)}\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2\mathrm{d}x.$$

Returning to (99), we therefore have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2\mathrm{d}x + \frac{\varepsilon}{2}\int_\Omega (\Delta\psi + \delta\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x &\le \delta^2|\Omega|^{1/2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^2(\Omega)}\|\bar{F}\|_{L^\infty(Q_T)} \\ &+ \delta\|w\|_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} + \frac{\delta^2}{2\varepsilon}\|\bar{F}\|^2_{L^\infty(Q_T)}\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2\mathrm{d}x. \end{aligned}$$

Integrating in time and using the Hölder inequality in the second term on the right-hand side, we get, for every $t \in [0,T]$,

$$\begin{aligned}
&\frac{1}{2}\int_\Omega |\nabla\psi + \delta w\bar{F}|^2\,\mathrm{d}x + \frac{\varepsilon}{2}\int_0^t\int_\Omega (\Delta\psi + \delta\,\mathrm{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\le \frac{1}{2}\int_\Omega \left|\nabla\psi(0,x) + \delta w_0(x)\bar{F}(0,x)\right|^2\mathrm{d}x + \delta^2|\Omega|^{1/2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\|\bar{F}\|_{L^\infty(Q_T)}\\
&\qquad + \delta\|w\|_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\sup_{\tau\in[0,t]}\|\nabla\psi(\tau,\cdot)\|_{L^2(\Omega)}\\
&\qquad + \frac{\delta^2}{2\varepsilon}\|\bar{F}\|^2_{L^\infty(Q_T)}\int_0^t\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2\mathrm{d}x.
\end{aligned}$$

By the triangle inequality, $\|\nabla\psi(\tau,\cdot)\|_{L^2(\Omega)} \le \|\nabla\psi(\tau,\cdot) + \delta w(\tau,\cdot)\bar{F}(\tau,\cdot)\|_{L^2(\Omega)} + \|\delta w\bar{F}(\tau,\cdot)\|_{L^2(\Omega)}$, and using the Young inequality as well, we get, for every $t \in [0,T]$,

$$\begin{aligned}
&\frac{1}{2}\int_\Omega |\nabla\psi + \delta w\bar{F}|^2\,\mathrm{d}x + \frac{\varepsilon}{2}\int_0^t\int_\Omega (\Delta\psi + \delta\,\mathrm{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t \le \frac{1}{2}\int_\Omega \left|\nabla\psi(0,x) + \delta w_0(x)\bar{F}(0,x)\right|^2\mathrm{d}x\\
&\qquad + \delta^2|\Omega|^{1/2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\|\bar{F}\|_{L^\infty(Q_T)}\\
&\qquad + \frac{1}{4}\sup_{\tau\in[0,t]}\|\nabla\psi(\tau,\cdot) + \delta w(\tau,\cdot)\bar{F}(\tau,\cdot)\|^2_{L^2(\Omega)}\\
&\qquad + \delta^2\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|^2_{L^1(0,T;L^2(\Omega))} + \delta^2|\Omega|T\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\|\bar{F}\|_{L^\infty(Q_T)}\\
&\qquad + \frac{\delta^2}{2\varepsilon}\|\bar{F}\|^2_{L^\infty(Q_T)}\int_0^t\int_\Omega \left|\nabla\psi + \delta w\bar{F}\right|^2\mathrm{d}x.
\end{aligned}$$

As a result, we write

$$\begin{aligned}
&\frac{1}{4}\sup_{\tau\in[0,t]}\|\nabla\psi(\tau,\cdot) + \delta w(\tau,\cdot)\bar{F}(\tau,\cdot)\,\|^2_{L^2(\Omega)} + \frac{\varepsilon}{2}\int_0^t\int_\Omega (\Delta\psi + \delta\,\mathrm{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t \le\\
&\qquad \frac{1}{2}\int_\Omega \left|\nabla\psi(0,x) + \delta w_0(x)\bar{F}(0,x)\right|^2\mathrm{d}x + \delta^2|\Omega|^{1/2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\|\bar{F}\|_{L^\infty(Q_T)}\\
&\qquad + \delta^2|\Omega|T\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}\|\bar{F}\|_{L^\infty(Q_T)} + \frac{\delta^2}{2}\|w\|^2_{L^\infty(Q_T)}\|\partial_t\bar{F}\|^2_{L^1(0,T;L^2(\Omega))}\\
&\qquad + \frac{\delta^2}{2\varepsilon}\|\bar{F}\|^2_{L^\infty(Q_T)}\int_0^t \sup_{y\in[0,\tau]}\|\nabla\psi(y,\cdot) + \delta w(y,\cdot)\bar{F}(y,\cdot)\|^2_{L^2(\Omega)}\,\mathrm{d}\tau.
\end{aligned}$$

Using a combination of (93), the formula (96), and (95), we deduce from the previous inequality that, for every $t \in [0,T]$,

$$\begin{aligned}
&\sup_{\tau\in[0,t]}\|\,\nabla\psi(\tau,\cdot) + \delta w(\tau,\cdot)\bar{F}(\tau,\cdot)\,\|^2_{L^2(\Omega)} + \varepsilon\int_0^t\int_\Omega (\Delta\psi + \delta\,\mathrm{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t\\
&\qquad\qquad \le C(\varepsilon,\delta,T,\Omega,\|w_0\|_{L^\infty(\Omega)},\|\nabla w_0\|_{L^\infty(\Omega)},\|\bar{F}\|_{L^\infty(Q_T)},\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))})\cdot\\
&\qquad\qquad\quad \left(1 + \int_0^t \sup_{y\in[0,\tau]}\|\nabla\psi(y,\cdot) + \delta w(y,\cdot)\bar{F}(y,\cdot)\|^2_{L^2(\Omega)}\,\mathrm{d}\tau\right),
\end{aligned} \tag{101}$$

where C on the right-hand side denotes a quantity depending only the parameters inside its brackets. Then, an application of Grönwall's Lemma yields

$$\sup_{t\in[0,T]}\|\,\nabla\psi(t,\cdot) + \delta w(t,\cdot)\bar{F}(t,\cdot)\,\|^2_{L^2(\Omega)} \le C(\varepsilon,\delta,T,\Omega,\|w_0\|_{C^1(\bar{\Omega})},\|\bar{F}\|_{L^\infty(Q_T)},\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}).$$

Hence, returning to (101) and using (93), we have

$$\varepsilon\int_{Q_T}\left(\Delta\psi + \delta\mathrm{div}(w\bar{F})\right)^2\mathrm{d}x\,\mathrm{d}t \le C(\varepsilon,\delta,T,\Omega,\|w_0\|_{C^1(\bar{\Omega})},\|\bar{F}\|_{L^\infty(Q_T)},\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))}). \tag{102}$$

Using $\int_{Q_T}(\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t \le C(\|\bar{F}\|_{L^\infty(Q_T)}, \|\nabla\bar{F}\|_{L^\infty(Q_T)})\|w\|^2_{L^2(0,T;H^1(\Omega))}$ along with the estimate (95), and using the triangle inequality, we obtain

$$\int_{Q_T}(\Delta\psi)^2\,\mathrm{d}x\,\mathrm{d}t \le C(\varepsilon,\delta,T,\Omega,\|w_0\|_{C^1(\bar{\Omega})},\|\bar{F}\|_{L^\infty(Q_T)},\|\partial_t\bar{F}\|_{L^1(0,T;L^2(\Omega))},\|\nabla\bar{F}\|_{L^\infty(Q_T)}).$$

Note that $\Delta\psi = (\varepsilon + w)\Delta w + |\nabla w|^2$ and that w is non-negative, whence the previous estimate and (95) yield the desired estimate (50).

Step 2 (control of drift terms): We emphasise that the addition of drift terms changes nothing to this argument – and one would still define ψ as per (96) – since, by (7), they can be bounded in L^∞ as follows

$$\begin{aligned}\|\nabla L(\cdot,\cdot,w)\|_{L^\infty(Q_T)} &\le \|V\|_{C^1(\mathbb{R}^{d+1})} + \|\nabla W * w(t,\cdot)\|_{L^\infty(Q_T)}\\ &\le \|V\|_{C^1(\mathbb{R}^{d+1})} + |\Omega|\|W\|_{C^1(\mathbb{R}^d)}\|w\|_{L^\infty(Q_T)},\end{aligned}$$

in conjunction with (93), where we omitted the i subscripts. Similarly, any additional space derivatives fall directly on V and W, not on w, and so

$$\|\nabla^2 L(\cdot,\cdot,w)\|_{L^\infty(Q_T)} \le \|V\|_{C^2(\mathbb{R}^d)} + |\Omega|\|W\|_{C^2(\mathbb{R}^d)}\|w\|_{L^\infty(Q_T)}.$$

For an additional time derivative, we have

$$\|\partial_t\nabla L(\cdot,\cdot,w)\|_{L^1(0,T;L^2(\Omega))} \le |\Omega|^{1/2}T\|V\|_{C^2(\mathbb{R}^{d+1})} + \|\nabla W * \partial_t w(t,\cdot)\|_{L^1(0,T;L^2(\Omega))},$$

and the second term on the right-hand side of the above is controlled – using (97), (98), the Cauchy–Schwarz integral inequality, and the Young inequality – as

$$\begin{aligned}\int_0^T \|\nabla W * \partial_t w(t,\cdot)\|_{L^2(\Omega)}\,\mathrm{d}t &\le \|W\|_{C^1(\mathbb{R}^d)}T^{1/2}\|\partial_t w\|_{L^2(Q_T)}\\ &\le \frac{\varepsilon}{4}\int_{Q_T}(\Delta\psi + \delta\operatorname{div}(w\bar{F}))^2\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{\varepsilon}\|W\|^2_{C^1(\mathbb{R}^d)}T,\end{aligned}$$

and the first term on the right-hand side can be absorbed into the left-hand side of (102). Thus, we have shown that the term of the form $w\nabla L$ can be handled in exactly the same way as $\delta w\bar{F}$ — notice that we never make use of the δ smallness assumption. ☐

A.3. Proof of Lemma 5.8 : Lower semicontinuity of Fisher information

Proof of Lemma 5.8.

Step 1 (convergence in x for fixed t): To begin with, observe that since $\zeta^n \to \zeta$ strongly in $L^1(Q_T)$, there exists a subsequence (still indexed by n) such that $\sum_{n\in\mathbb{N}}\|\zeta^n - \zeta\|_{L^1(Q_T)} < +\infty$. From here until the end of this proof, we only consider this particular subsequence, and we never pass to further subsequences. In what follows we use the identification $L^1(Q_T) = L^1(0,T;L^1(\Omega))$ and write $\boldsymbol{\zeta}^n(t) = \zeta^n(t,\cdot)$ and $\boldsymbol{\zeta}(t) = \zeta(t,\cdot)$. We consider, for a.e. $t \in (0,T)$, the convergence of $\boldsymbol{\zeta}^n(t)$ towards $\boldsymbol{\zeta}(t)$ in the norm of $L^1(\Omega)$. In particular, the Minkowski inequality for infinite sums in $L^1((0,T))$ yields

$$\left\|\sum_{n\in\mathbb{N}}\|\boldsymbol{\zeta}^n(\cdot) - \boldsymbol{\zeta}(\cdot)\|_{L^1(\Omega)}\right\|_{L^1((0,T))} \le \sum_{n\in\mathbb{N}}\|\zeta^n - \zeta\|_{L^1(Q_T)} < +\infty,$$

and hence the infinite series $\sum_{n\in\mathbb{N}}\|\boldsymbol{\zeta}^n(\cdot) - \boldsymbol{\zeta}(\cdot)\|_{L^1(\Omega)}$ is well-defined as an element of $L^1((0,T))$. Since any integrable function is finite almost everywhere, it follows that $\sum_{n\in\mathbb{N}}\|\boldsymbol{\zeta}^n(t) - \boldsymbol{\zeta}(t)\|_{L^1(\Omega)} < +\infty$ a.e. $t \in (0,T)$. It immediately follows from the summability of this latter series that, for a.e. $t \in (0,T)$, we have $\|\boldsymbol{\zeta}^n(t) - \boldsymbol{\zeta}(t)\|_{L^1(\Omega)} \to 0$ as $n \to \infty$. Rewriting in terms of the original functions, we have shown that, for a.e. $t \in (0,T)$, $\lim_{n\to\infty} \|\zeta^n(t,\cdot) - \zeta(t,\cdot)\|_{L^1(\Omega)} = 0$.

Step 2 (conclusion via Fatou's Lemma): Note that ζ is also non-negative from the assumption that $\zeta^n \to \zeta$ in $L^1(Q_T)$. The convergence obtained at the end of Step 1 is sufficient to satisfy the hypothesis of Lemma 5.7. An application of this latter result gives, for a.e. $t \in (0,T)$, $\int_\Omega \frac{|\nabla\zeta(t,x)|^2}{\zeta(t,x)}\,\mathrm{d}x \le \liminf_{n\to\infty} \int_\Omega \frac{|\nabla\zeta^n(t,x)|^2}{\zeta^n(t,x)}\,\mathrm{d}x$. By integrating the previous inequality with respect to time and then applying the Fatou Lemma, we obtain

$$\int_{Q_T} \frac{|\nabla\zeta|^2}{\zeta}\,\mathrm{d}x\,\mathrm{d}t \le \int_0^T \liminf_{n\to\infty} \left(\int_\Omega \frac{|\nabla\zeta^n(t,x)|^2}{\zeta^n(t,x)}\,\mathrm{d}x \right) \mathrm{d}t \le \liminf_{n\to\infty} \int_{Q_T} \frac{|\nabla\zeta^n|^2}{\zeta^n}\,\mathrm{d}x\,\mathrm{d}t,$$

as required. □

A.4. Smoothing operator

The purpose of this appendix is to verify the properties of the smoothing operator R_μ in (62) stated in Section 5.1. To this end, we make the following definitions, and we note that, throughout this section only, ϕ refers to the Friedrichs bump function (see below) and not to a generic test function.

Definition A.2. Define $\phi \in C_c^\infty(\mathbb{R})$ to be the standard Friedrichs bump function $\phi(y) = e^{-\frac{1}{1-|y|^2}} \mathbb{1}_{\{|y|<1\}}$, and define, for $\mu > 0$, $\phi_\mu(t,x) := c_\phi \mu^{-(d+1)} \phi(\mu^{-1}\sqrt{t^2 + |x|^2})$ for all $(t,x) \in \mathbb{R}^{d+1}$, where the positive constant c_ϕ is chosen such that $\int_{\mathbb{R}^{d+1}} \phi_1(y)\,\mathrm{d}y = 1$.

Note that $\|\phi\|_{L^\infty(\mathbb{R})} = 1$ and

$$\operatorname{supp} \phi_\mu = \bar{B}(0,\mu),$$

where the latter is the closed ball of radius μ, and $\int_{\mathbb{R}^{d+1}} \phi_\mu(y)\,\mathrm{d}y = 1$ for every $\mu > 0$.

Definition A.3. We define the operator $A_\mu : L^2(\mathbb{R}; H^1(\mathbb{R}^d)) \to C^\infty(\bar{Q}_T)$ to be the restriction of the mollification to the parabolic cylinder, i.e., $A_\mu u := \int_{\mathbb{R}} \int_{\mathbb{R}^d} \phi_\mu(\cdot - \tau, \cdot - y) u(\tau, y)\,\mathrm{d}y\,\mathrm{d}\tau\big|_{\bar{Q}_T}$.

Lemma A.4 (*Preliminary Spatial Extension*)**.** *There exists a bounded linear operator $E' : L^2(0,T; H^1(\Omega)) \to L^2(0,T; H^1(\mathbb{R}^d))$ such that $E'f(t,x) = f(t,x)$ a.e. $(t,x) \in Q_T$, and, with $\Omega_1 := \{x \in \mathbb{R}^d : d(x,\Omega) \le 1\}$, in which Ω is compactly contained, $\operatorname{supp} E'f(t,\cdot) \subset \Omega_1$ a.e. $t \in (0,T)$. Moreover, there exists a constant $C > 0$, depending only on Ω, such that $\|E'f\|_{L^2((0,T)\times\mathbb{R}^d)} \le C\|f\|_{L^2(Q_T)}$ for all $f \in L^2(0,T; H^1(\Omega))$.*

Proof. By definition of $L^2(0,T; H^1(\Omega)) \ni f$, we have that for a.e. $t \in (0,T)$, the element $f(t,\cdot)$ belongs to $H^1(\Omega)$. The result follows at once from repeating the argument of the proof Theorem 1 of [62, Section 5.4] on the function $f(t,\cdot)$, with the time coordinate kept fixed. □

The next result follows easily by considering $Ef(t,x) := E'f(t,x)\mathbb{1}(t)_{[0,T]}$ for a.e. $(t,x) \in \mathbb{R}^{d+1}$, where E' is the bounded linear operator of Lemma A.4.

Lemma A.5 (*Sobolev Extension for Spaces Involving Time*)**.** *There exists a bounded linear operator $E : L^2(0,T; H^1(\Omega)) \to L^2(\mathbb{R}; H^1(\mathbb{R}^d))$ such that $Ef(t,x) = f(t,x)$ a.e. $(t,x) \in Q_T$, and, with $\Omega_1 := \{x \in \mathbb{R}^d : d(x,\Omega) \le 1\}$, in which Ω is compactly contained, $\operatorname{supp} Ef \subset [0,T] \times \Omega_1$. Moreover, there exists a constant $C > 0$, depending only on Ω, such that $\|Eu\|_{L^2(\mathbb{R}^{d+1})} \le C\|u\|_{L^2(Q_T)}$ for all $u \in L^2(0,T; H^1(\Omega))$.*

We finally arrive at the smoothing operator in question. The proof is a standard exercise in real analysis, which we omit for clarity of presentation.

Lemma A.6 (*Smoothing Operator*)**.** *Fix $\mu > 0$. The smoothing operator R_μ, defined explicitly by*

$$R_\mu := A_\mu \circ E : L^2(0,T;H^1(\Omega)) \to C^\infty(\bar{Q}_T), \tag{103}$$

admits, for some positive constant C_{reg} independent of μ, the estimate

$$\|R_\mu u\|_{L^2(0,T;H^1(\Omega))} \le C_{reg}\|u\|_{L^2(0,T;H^1(\Omega))} \qquad \forall u \in L^2(0,T;H^1(\Omega)). \tag{104}$$

As such, it is a bounded linear operator from $L^2(0,T;H^1(\Omega))$ to $C^\infty(\bar{Q}_T)$ equipped with the subspace norm-topology of $L^2(0,T;H^1(\Omega))$. Moreover, given any $u \in L^2(0,T;H^1(\Omega))$, we have the strong convergence

$$\|R_\mu u - u\|_{L^2(0,T;H^1(\Omega))} \to 0 \qquad \text{as } \mu \to 0. \tag{105}$$

Also, for some positive constant C_μ depending on μ, ϕ, Ω, T,

$$\|R_\mu u\|_{L^\infty(0,T;W^{2,\infty}(\Omega))} \le C_\mu\|u\|_{L^2(0,T;H^1(\Omega))} \qquad \forall u \in L^2(0,T;H^1(\Omega)). \tag{106}$$

Remark A.7. Note $|\partial_t R_\mu u(t,x)| = |\int_{\mathbb{R}}\int_{\mathbb{R}^d} \partial_t\phi_\mu(t-\tau,x-y)Eu(\tau,y)\,\mathrm{d}y\,\mathrm{d}\tau| \le \|\partial_t\phi_\mu\|_{L^2(\mathbb{R}^{d+1})}\|Eu\|_{L^2(\mathbb{R}^{d+1})}$ by the Hölder inequality, and from which it follows from Lemma A.5 that $\|\partial_t R_\mu u\|_{L^\infty(Q_T)} \le C_\mu\|u\|_{L^2(Q_T)}$, where the finiteness of the constant C_μ follows from the smoothness and compact support of ϕ_μ. Similarly, $\|\partial_t\nabla R_\mu u\|_{L^\infty(Q_T)} \le \|\partial_t\nabla\phi_\mu\|_{L^2(\mathbb{R}^{d+1})}\|Eu\|_{L^2(Q_T)}$. In view of this, relabelling C_μ as the relevant positive constant depending on μ, Ω, T, (which, incidentally, will blow up in the limit as $\mu \to 0$), we have shown that there holds

$$\|\partial_t R_\mu u\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} \le C_\mu\|u\|^2_{L^2(Q_T)} \qquad \forall u \in L^2(0,T;H^1(\Omega)). \tag{107}$$

The following result is a direct corollary of the proof of Lemma A.6.

Corollary A.8. *Fix $\mu > 0$. The smoothing operator R_μ of Lemma A.6, defined explicitly by*

$$R_\mu := A_\mu \circ E : L^2(0,T;H^1(\Omega)) \to C^\infty(\bar{Q}_T), \tag{108}$$

admits, for some positive constant C independent of μ, the estimate

$$\|R_\mu u\|_{L^2(Q_T)} \le C\|u\|_{L^2(Q_T)} \qquad \forall u \in L^2(0,T;H^1(\Omega)). \tag{109}$$

As such, it is a bounded linear operator from $L^2(Q_T)$ to $C^\infty(\bar{Q}_T)$ equipped with the subspace norm-topology of $L^2(Q_T)$. Moreover, given any $u \in L^2(0,T;H^1(\Omega))$, we have the strong convergence $\|R_\mu u - u\|_{L^2(Q_T)} \to 0$.

References

[1] C. Conca, E. Espejo, K. Vilches, Remarks on the blowup and global existence for a two species chemotactic Keller–Segel system in $\mathbb{R}^2$, Eur. J. Appl. Math. 22 (6) (2011) 553–580.

[2] E.E. Espejo, A. Stevens, J.J.L. Velázquez, A note on non-simultaneous blow-up for a drift-diffusion model, Differential Integral Equations 23 (2010) 451–462.

[3] V. Cristini, J. Lowengrub, Multiscale Modeling of Cancer: An Integrated Experimental and Mathematical Modeling Approach, Cambridge University Press, 2010.

[4] B. Perthame, Some Mathematical Models of Tumor Growth, Université Pierre et Marie Curie, Paris 2015, 2014, pp. 23–32.

[5] L. Preziosi, L. Graziano, Multiphase models of tumor growth: General framework and particular cases, in: V. Capasso (Ed.), Mathematical Modelling and Computing in Biology and Medicine, Societa' Editrice Esculapio, 2003, pp. 622–628.

[6] C. Appert-Rolland, P. Degond, S. Motsch, Two-way multi-lane traffic model for pedestrians in corridors, Netw. Heterog. Media 6 (3) (2011) 351–381.

[7] R.M. Colombo, M. Lécureux-Mercier, Nonlocal crowd dynamics models for several populations, Acta Math. Sci. 32 (1) (2012) 177–196.

[8] B. Düring, P. Markowich, J.-F. Pietschmann, M.-T. Wolfram, Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders, P. R. Soc. A 465 (2112) (2009) 3687–3708.

[9] M. Bruna, S.J. Chapman, Diffusion of multiple species with excluded-volume effects, J. Chem. Phys. 137 (20) (2012) 204116.

[10] Y. Chen, T. Kolokolnikov, A minimal model of predator-swarm interactions, J. R. Soc. Interface 11 (94) (2014) 20131208.

[11] A. Jüngel, Stefan S. Portisch, A. Zurek, Nonlocal cross-diffusion systems for multi-species populations and networks, 2021, arXiv preprint arXiv:2104.06292.

[12] L. Chen, A. Jüngel, Analysis of a parabolic cross-diffusion semiconductor model with electron–hole scattering, Comm. Partial Differential Equations 32 (1) (2007) 127–148.

[13] P.A. Markowich, P. Szmolyan, A system of convection–diffusion equations with small diffusion coefficient arising in semiconductor physics, J. Differential Equations 81 (2) (1989) 234–254.

[14] A.B.T. Barbaro, N. Rodriguez, H. Yoldas, N. Zamponi, Analysis of a cross-diffusion model for rival gangs interaction in a city, 2020, arXiv preprint arXiv:2009.04189.

[15] K. Kang, I. Primi, J. Velazquez, A 2d-model of cell sorting induced by propagation of chemical signals along spiral waves, Comm. Partial Differential Equations 38 (2013) 1069–1122.

[16] G. Lemon, J.R. King, A functional differential equation model for biological cell sorting due to differential adhesion, Math. Models Methods Appl. Sci. 23 (01) (2013) 93–126.

[17] N. Shigesada, K. Kawasaki, E. Teramoto, Spatial segregation of interacting species, J. Theoret. Biol. 79 (1) (1979) 83–99.

[18] B. Vasiev, C. Weijer, Modeling chemotactic cell sorting during dictyostelium discoideum mound formation, Biophys. J. 76 (1999) 595–605.

[19] M. Bertsch, M. Gurtin, D. Hilhorst, On a degenerate diffusion equation of the form $c(z)t = \varphi(zx)x$ with application to population dynamics, J. Differential Equations 67 (1987) 56–89.

[20] M. Bertsch, M. Gurtin, D. Hilhorst, L. Peletier, On interacting populations that disperse to avoid crowding: Preservation of segregation, J. Math. Biol. 23 (1–13) (1985).

[21] M. Bertsch, D. Hilhorst, H. Izuhara, M. Mimura, A nonlinear parabolic-hyperbolic system for contact inhibition of cell-growth, Differ. Equ. Appl. 4 (2012) 137–157.

[22] M. Bertsch, R. Dal Passo, M. Mimura, A free boundary problem arising in a simplified tumour growth model of contact inhibition, Interface Free Bound. 12 (2010) 235–250.

[23] M.E. Gurtin, A.C. Pipkin, A note on interacting populations that disperse to avoid crowding, Quart. Appl. Math. 42 (1984) 87–94.

[24] M. Bruna, M. Burger, H. Ranetbauer, M.-T. Wolfram, Cross-diffusion systems with excluded-volume effects and asymptotic gradient flow structures, J. Nonlinear Sci. 27 (2) (2017) 687–719.

[25] V. Calvez, J.A. Carrillo, Volume effects in the Keller-Segel model: energy estimates preventing blow-up, J. Math. Pure. Appl. 86 (2) (2006) 155–175.

[26] S. Hittmeir, A. Jüngel, Cross diffusion preventing blow-up in the two-dimensional Keller–Segel model, SIAM J. Math. Anal. 43 (2) (2011) 997–1022.

[27] K.J. Painter, T. Hillen, Volume-filling and quorum-sensing in models for chemosensitive movement, Can. Appl. Math. Q. 10 (2002) 501–543.

[28] D. Wilson, H. Byrne, M. Bruna. Reactions, Diffusion and volume exclusion in a heterogeneous system of interacting particles, Phys. Rev. E 97 (2018) 062137.

[29] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows: In Metric Spaces and in the Space of Probability Measures, Springer Science & Business Media, 2008.

[30] A.L. Bertozzi, D. Slepcev, Existence and uniqueness of solutions to an aggregation equation with degenerate diffusion, Commun. Pure Appl. Anal. 9 (6) (2009) 1617.

[31] H. Amann, Dynamic theory of quasilinear parabolic systems, Math. Z. 202 (2) (1989) 219–250.

[32] S.N. Busenberg, C.C. Travis, Epidemic models with spatial spread due to population migration, J. Math. Biol. 16 (1983) 181–198.

[33] D. Le, T.T. Nguyen, Everywhere regularity of solutions to a class of strongly coupled degenerate parabolic systems, Comm. Partial Differential Equations 31 (2) (2006) 307–324.

[34] M. Pierre, D. Schmitt, Blowup in reaction–diffusion systems with dissipation of mass, SIAM J. Math. Anal. 28 (1997) 259–269.

[35] J.A. Carrillo, S. Fagioli, F. Santambrogio, M. Schmidtchen, Splitting schemes and segregation in reaction cross-diffusion systems, SIAM J. Math. Anal. 50 (5) (2018) 5695–5718.

[36] P. Gwiazda, B. Perthame, A. Swierczewska-Gwiazda, A two-species hyperbolic-parabolic model of tissue growth, Comm. Partial Differential Equations 44 (12) (2019) 1605–1618.

[37] M. Burger, J.A. Carrillo, J.-F. Pietschmann, M. Schmidtchen, Segregation effects and gap formation in cross-diffusion models, Interface Free Bound. 22 (2) (2020) 175–203.

[38] M. Burger, M. Di Francesco, S. Fagioli, A. Stevens, Sorting phenomena in a mathematical model for two mutual attracting/repelling species, SIAM J. Math. Anal. 50 (3) (2018) 3210–3250.

[39] J.A. Carrillo, Y. Huang, M. Schmidtchen, Zoology of a nonlocal cross-diffusion model for two species, SIAM J. Appl. Math. 78 (2) (2018) 1078–1104.

[40] L. Ambrosio, Transport equation and Cauchy problem for non-smooth vector fields, in: Lecture Notes in Mathematics Calculus of Variations and Non-Linear Partial Differential Equations (CIME Series, Cetraro, 2005), 2008.

[41] M. Burger, M. Di Francesco, J.-F. Pietschmann, B. Schlake, Nonlinear cross-diffusion with size exclusion, SIAM J. Math. Anal. 42 (6) (2010) 2842.

[42] L. Chen, A. Jüngel, Analysis of a parabolic cross-diffusion population model without self-diffusion, J. Differential Equations 224 (2006) 39–59.
[43] L. Desvillettes, T. Lepoutre, A. Moussa, A. Trescases, On the entropic structure of reaction-cross diffusion systems, Commun. Part. Differentail Equations 40 (9) (2015) 1705–1747.
[44] A. Jüngel, The boundedness-by-entropy method for cross-diffusion systems, Nonlinearity 28 (6) (2015) 1963.
[45] A. Jüngel, I.V. Stelzer, Entropy structure of a cross-diffusion tumor-growth model, Math. Models Methods Appl. Sci. 22 (07) (2012).
[46] F. Santambrogio, Optimal Transport for Applied Mathematicians, Springer, Birkhäuser, 2015.
[47] M. Di Francesco, S. Fagioli, Measure solutions for nonlocal interaction PDEs with two species, Nonlinearity 26 (2013) 2777–2808.
[48] P. Laurencot, B. Matioc, A gradient flow approach to a thin film approximation of the muskat problem, Calc. Var. Partial Differential Equations 47 (1) (2013) 319–341.
[49] G. Carlier, M. Laborde, Remarks on continuity equations with nonlinear diffusion and nonlocal drifts, J. Math. Anal. Appl. 444 (2) (2017) 1690–1702.
[50] G. Carlier, M. Laborde, A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts, Nonlinear Anal. 150 (2017) 1–18.
[51] L. Alasio, M. Bruna, Y. Capdeboscq, Stability estimates for systems with small cross-diffusion, ESAIM-Math. Mod. Num. 52 (3) (2018) 21109–21135.
[52] L. Chen, A. Jüngel, Analysis of a multidimensional parabolic population model with strong cross-diffusion, SIAM J. Math. Anal. 36 (1) (2004) 301–322.
[53] L. Desvillettes, A. Trescases, New results for triangular reaction cross diffusion system, J. Math. Anal. Appl. 430 (1) (2015) 32–59.
[54] G. Galiano, V. Selgas, On a cross-diffusion segregation problem arising from a model of interacting particles, Nonlinear Anal. RWA 18 (2014) 34–49.
[55] I. Kim, A. Mészáros, On nonlinear cross-diffusion systems: an optimal transport approach, Calc. Var. Partial Differential Equations 57 (3) (2018) 79.
[56] I. Kim, Y.P. Zhang, Regularity properties of degenerate diffusion equations with drift, SIAM J. Math. Anal. 50 (2018) 4371–4406.
[57] L. Beck, D. Matthes, M. Zizza, Exponential convergence to equilibrium for coupled systems of nonlinear degenerate drift diffusion equations, 2021, arXiv preprint arXiv:2112.05810.
[58] M. Di Francesco, A. Esposito, S. Fagioli, Nonlinear degenerate cross-diffusion systems with nonlocal interaction, Nonlinear Anal. 169 (2018) 94–117.
[59] O.A. Ladyzhenskaia, V.A. Solonnikov, N.N. Ural'tseva, Linear and Quasi-Linear Equations of Parabolic Type, Vol. 23, American Mathematical Soc, 1988.
[60] A. Blanchet, J. Dolbeault, B. Perthame, Two-dimensional Keller–Segel model: Optimal critical mass and qualitative properties of the solutions, Electron. J. Differ. Equations (2006).
[61] J. Stará, O. John, Some (new) counterexamples of parabolic systems, Comment. Math. Univ. Carolin. 36 (3) (1995) 503–510.
[62] L.C. Evans, Partial Differential Equations. Graduate Studies in Mathematics, American Mathematical Society, 1998.
[63] M. Bodnar, J.J.L. Velázquez, Derivation of macroscopic equations for individual cell-based models: A formal approach, Math. Methods Appl. Sci. 28 (15) (2005) 1757–1779.
[64] K. Oelschläger, Large systems of interacting particles and the porous medium equation, J. Differential Equations 88 (2) (1990) 294–346.
[65] R. Philipowski, Interacting diffusions approximating the porous medium equation and propagation of chaos, Stochastic Process. Appl. 117 (4) (2007) 526–538.
[66] L. Chen, E.S. Daus, A. Holzinger, A. Jüngel, Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems, J. Nonlinear Sci. 31 (6) (2021) 1–38.
[67] L. Chen, E.S. Daus, A. Jüngel, Rigorous mean-field limit and cross-diffusion, Z. Angew. Math. Phys. 70 (2019) 122.
[68] H. Dietert, A. Moussa, Persisting entropy structure for nonlocal cross-diffusion systems, 2021, Preprint arXiv:2101.02893.
[69] J. Fontbona, S. Méléard, Non local lotka-volterra system with cross-diffusion in an heterogeneous medium, J. Math. Biol. 70 (4) (2015) 829–854.
[70] E. Di Benedetto, Degenerate Parabolic Equations, Springer Science & Business Media, 2012.
[71] F. Boyer, P. Fabrie, Mathematical Tools for the Study of the Incompressible Navier–Stokes Equations and Related Models, Vol. 183, Springer Science & Business Media, 2012.
[72] L. Ambrosio, N. Gigli, G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, Invent. Math. 195 (2) (2014) 289–391.
[73] F. Murat, Compacité par compensation, Ann. Sc. Norm. Super. Pisa, Cl. Sci. 4 (5) (1978) 489–507.