Calculus of Variations

EDP-convergence for a linear reaction-diffusion system with fast reversible reaction

Artur Stephan[1]

Abstract

We perform a fast-reaction limit for a linear reaction-diffusion system consisting of two diffusion equations coupled by a linear reaction. We understand the linear reaction-diffusion system as a gradient flow of the free energy in the space of probability measures equipped with a geometric structure, which contains the Wasserstein metric for the diffusion part and cosh-type functions for the reaction part. The fast-reaction limit is done on the level of the gradient structure by proving EDP-convergence with tilting. The limit gradient system induces a diffusion system with Lagrange multipliers on the linear slow-manifold. Moreover, the limit gradient system can be equivalently described by a coarse-grained gradient system, which induces a diffusion equation with a mixed diffusion constant for the coarse-grained slow variable.

1 Introduction

Considering two species X_1 and X_2 that diffuse in a bounded medium $\Omega \subset \mathbb{R}^d$ and react linearly $X_1 \rightleftharpoons X_2$, the evolution of their concentrations $c = (c_1, c_2)$ can be described by the linear reaction-diffusion system

$$\dot{c}_1 = \delta_1 \Delta c_1 - (\tilde{\alpha} c_1 - \tilde{\beta} c_2)$$
$$\dot{c}_2 = \delta_2 \Delta c_2 + (\tilde{\alpha} c_1 - \tilde{\beta} c_2) \tag{1.1}$$

complemented with no-flux boundary conditions and initial conditions, where $\delta_1, \delta_2 > 0$ are diffusion coefficients for species X_1 and X_2, respectively, and $\tilde{\alpha}, \tilde{\beta} > 0$ are reaction rates describing the reaction speed of the linear reaction $X_1 \rightleftharpoons X_2$. The aim of the paper is to investigate system (1.1) if the reaction is much faster than the diffusion. To do this, we introduce a small parameter $\varepsilon > 0$ and assume that the reaction rates are given by $\tilde{\alpha} = \frac{1}{\varepsilon}\sqrt{\frac{\alpha}{\beta}}$,

✉ Artur Stephan
    artur.stephan@wias-berlin.de

[1] Weierstraß-Institut für Angewandte Analysis und Stochastik, Mohrenstraße 39, 10117 Berlin, Germany

$\tilde{\beta} = \frac{1}{\varepsilon}\sqrt{\frac{\beta}{\alpha}}$. Then, the system (1.1) can be rewritten in an ε-dependent reaction-diffusion system

$$\dot{c}_1^\varepsilon = \delta_1 \Delta c_1^\varepsilon - \frac{1}{\varepsilon}\left(\sqrt{\frac{\alpha}{\beta}}c_1^\varepsilon - \sqrt{\frac{\beta}{\alpha}}c_2^\varepsilon\right)$$
$$\dot{c}_2^\varepsilon = \delta_2 \Delta c_2^\varepsilon + \frac{1}{\varepsilon}\left(\sqrt{\frac{\alpha}{\beta}}c_1^\varepsilon - \sqrt{\frac{\beta}{\alpha}}c_2^\varepsilon\right). \tag{1.2}$$

Reaction systems and reaction-diffusion systems with slow and fast time scales have attracted a lot of attention in the last decades [7–10,13,17,22,30,34,36,38]. Bothe and Hilhorst proved a fast-reaction limit $\varepsilon \to 0$ for (1.2) in the following form.

Theorem 1.1 ([6]) *Let $\Omega \subset \mathbb{R}^d$ be a domain with Lipschitz boundary. Let c_1^ε and c_2^ε be weak solutions of (1.2) with no-flux boundary conditions $\nabla c_i^\varepsilon \cdot \nu = 0$ on $\partial\Omega$. Then $c_1^\varepsilon \to c_1$ and $c_2^\varepsilon \to c_2$ in $\mathrm{L}^2([0, T] \times \Omega)$ as $\varepsilon \to 0$, and we have $\frac{c_1}{\beta} = \frac{c_2}{\alpha}$. Moreover, defining the coarse-grained concentration $\hat{c} = c_1 + c_2$, then $\hat{c}$ solves the diffusion equation $\dot{\hat{c}} = \hat{\delta}\Delta\hat{c}$ with a new mixed diffusion coefficient $\hat{\delta} = \frac{\beta\delta_1 + \alpha\delta_2}{\alpha + \beta}$.*

Essentially, the proof uses the free energy as a Lyapunov function to derive ε-uniform bounds on the concentrations c_i^ε and their gradients ∇c_i^ε, which is then used to prove convergence towards the slow manifold $\{c \in [0, \infty[^2 \mid \alpha c_1 = \beta c_2\}$. This proof also works for nonlinear reactions once ε-uniform L^∞-estimates are established (see [6]). On the linear slow manifold, one easily verifies that the coarse grained concentration $\hat{c} := \frac{\alpha+\beta}{\beta}c_1 = \frac{\alpha+\beta}{\alpha}c_2 = c_1 + c_2$ solves $\dot{\hat{c}} = \hat{\delta}\Delta\hat{c}$ where $\hat{\delta} = \frac{\beta\delta_1 + \alpha\delta_2}{\alpha + \beta}$ is the effective mixed diffusion coefficient.

In this work, we are not primarily interested in convergence of solutions of system (1.2). Instead, we perform the fast-reaction limit on the level of the underlying variational structure, which then implies convergence of solutions as a byproduct. Our starting point is that reaction-diffusion systems such as (1.2) can be written as a gradient flow equation induced by a gradient system $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$, where the state space Q is the space of probability measures $Q = \mathrm{Prob}(\Omega \times \{1, 2\})$ and the driving functional is the free energy $\mathcal{E}(\mu) = \int_\Omega \sum_{j=1}^2 E_B\left(\frac{c_j}{w_j}\right) w_j \mathrm{d}x$ for measures $\mu = c\,\mathrm{d}x$, with the Boltzmann function $E_B(r) = r\log r - r + 1$ and the (in general space dependent) stationary measure $w = (w_1, w_2)^\mathrm{T}$. The dissipation potential $\mathcal{R}_\varepsilon^*$ that determines the geometry of the underlying space is given by two parts $\mathcal{R}_\varepsilon^* = \mathcal{R}_{\mathrm{diff}}^* + \mathcal{R}_{\mathrm{react},\varepsilon}^*$ describing the diffusion and reaction separately. Starting with the pioneering work of Otto and coauthors [23,37] it is known that many diffusion-type problems can be understood as gradient systems driven by the free energy in the space of probability measures equipped with the Wasserstein distance. The corresponding dissipation potential $\mathcal{R}_{\mathrm{diff}}^*$ is quadratic and given by

$$\mathcal{R}_{\mathrm{diff}}^*(\mu, \xi) = \frac{1}{2}\int_\Omega \sum_{i=1}^2 \delta_i |\nabla\xi_i|^2 \,\mathrm{d}\mu_i.$$

Later Mielke [28] proposed a quadratic gradient structure with the same driving functional also for reaction-diffusion systems with reversible reactions satisfying detailed balance. Geometric properties of that gradient structure were investigated in [20,24]. Here, we are not interested in that gradient structure, but use a different, the so-called *cosh-type gradient structure*, where the reaction part is given by

$$\mathcal{R}_{\mathrm{react},\varepsilon}^*(\mu, \xi) = \frac{1}{\varepsilon}\int_\Omega \mathsf{C}^*(\xi_1(x) - \xi_2(x))\,\sqrt{\mathrm{d}\mu_1 \mathrm{d}\mu_2},$$

with $C^*(r) = 4(\cosh(r/2) - 1)$. Setting $\mathcal{R}_\varepsilon^* = \mathcal{R}_{\text{diff}}^* + \mathcal{R}_{\text{react}, \varepsilon}^*$, the reaction-diffusion system (1.2) can now be formally written as a gradient flow equation

$$\dot{\mu} = \partial_\xi \mathcal{R}_\varepsilon^*(\mu, -D\mathcal{E}(\mu)).$$

Although there are many gradient structures for (1.2) (see e.g. [30, Sect. 4]) and the cosh-type gradient structure entails several technical difficulties as defining a nonlinear kinetic relation and not inducing metric on Q, it nevertheless has several significant features. Historically, it has its origin in [27] where, following thermodynamic considerations, chemical reactions are written in exponential terms. In recent years, the cosh-gradient structure has been derived via a large-deviation principle [32,33], and it was shown that it is stable under limit processes [25] that are similar to our approach. Moreover, it does not explicitly depend on the stationary measure w, which allows for an rigorous distinction between the energetic and dissipative part [30]. This is physically reasonable because a change of the energy by an external field should not influence the geometric structure of the underlying space. We refer also to [40], where our choice of gradient structure has been formally derived for (1.2).

The goal of the paper is to construct an effective gradient system $(Q, \mathcal{E}, \mathcal{R}_{\text{eff}}^*)$ and perform the limit $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*) \to (Q, \mathcal{E}, \mathcal{R}_{\text{eff}}^*)$ as $\varepsilon \to 0$. For this, we use the notion of *convergence of gradient systems in the sense of the energy-dissipation principle*, shortly called *EDP-convergence*. EDP-convergence was introduced in [14] and further developed in [30,31] and is based on the dissipation functional

$$\mathfrak{D}_\varepsilon^\eta(\mu) = \int_0^T \mathcal{R}_\varepsilon(\mu, \dot{\mu}) + \mathcal{R}_\varepsilon^*(\mu, \eta - D\mathcal{E}(\mu)) \, dt,$$

which, for solutions μ of the gradient flow equation describes the total dissipation between initial time $\mathcal{E}(\mu(0))$ and final time $\mathcal{E}(\mu(T))$, and can now be defined for general trajectories $\mu \in L^1([0, T], Q)$. Here, the primal dissipation potential $\mathcal{R}_\varepsilon$ is defined as the Legendre transform of $\mathcal{R}_\varepsilon^*$ with respect to the second variable. The notion of EDP-convergence with tilting requires Γ-convergences of the energies $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ and of the dissipation functionals $\mathfrak{D}_\varepsilon^\eta \xrightarrow{\Gamma} \mathfrak{D}_0^\eta$ in suitable topologies, such that for all tilts η the limit $\mathfrak{D}_0^\eta$ has the form

$$\mathfrak{D}_0^\eta(\mu) = \int_0^T \mathcal{R}_{\text{eff}}(\mu, \dot{\mu}) + \mathcal{R}_{\text{eff}}^*(\mu, \eta - D\mathcal{E}_0(\mu)) \, dt,$$

see Sect. 2.2 for a precise definition. Importantly, the effective dissipation potential $\mathcal{R}_{\text{eff}}$ in the Γ-limit is independent of the tilts, hence allowing for more general energy functionals and stationary measures. In our situation, the tilts η correspond to an external potential $V = (V_1, V_2)$ added to the energy $\mathcal{E}$. On the level of the PDE, the original reaction-diffusion system (1.2) is extended to a reaction-drift-diffusion system of the form

$$\frac{d}{dt} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \text{div} \left(\begin{pmatrix} \delta_1 \nabla c_1 \\ \delta_2 \nabla c_2 \end{pmatrix} + \begin{pmatrix} \delta_1 c_1 \nabla V_1 \\ \delta_2 c_2 \nabla V_2 \end{pmatrix} \right) + \frac{1}{\varepsilon} \begin{pmatrix} -\sqrt{\frac{\alpha}{\beta}} e^{\frac{V_1 - V_2}{2}} & \sqrt{\frac{\beta}{\alpha}} e^{\frac{V_2 - V_1}{2}} \\ \sqrt{\frac{\alpha}{\beta}} e^{\frac{V_1 - V_2}{2}} & -\sqrt{\frac{\beta}{\alpha}} e^{\frac{V_2 - V_1}{2}} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The main result of the paper is Theorem 4.2 which asserts tilt EDP-convergence of $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$ to $(Q, \mathcal{E}, \mathcal{R}_{\text{eff}}^*)$ as $\varepsilon \to 0$ where the effective dissipation potential is given by

$$\mathcal{R}_{\text{eff}}^* = \mathcal{R}_{\text{diff}}^* + \chi_{\{\xi_1 = \xi_2\}},$$

where χ_A is the characteristic function of convex analysis taking values zero in A and infinity otherwise. The effective dissipation potential describes diffusion but restricts the chemical

potential $\xi = (\xi_1, \xi_2)$ to a linear submanifold. The induced gradient flow equation of the gradient system $(Q, \mathcal{E}, \mathcal{R}^*_{\text{eff}})$ is then given by a system of drift-diffusion equations on a linear submanifold with a space and time dependent Lagrange multiplier λ

$$\begin{aligned}
\dot{c}_1 &= \text{div}\,(\delta_1 \nabla c_1 + \delta_1 c_1 \nabla V_1) - \lambda \\
\dot{c}_2 &= \text{div}\,(\delta_2 \nabla c_2 + \delta_2 c_2 \nabla V_2) + \lambda
\end{aligned} \quad, \quad \frac{c_1}{\beta e^{-V_1}} = \frac{c_2}{\alpha e^{-V_2}}.$$

Moreover, as an immediate consequence of Theorem 4.2, we obtain that the effective gradient flow equation can be equivalently described as a drift-diffusion equation of the coarse-grained concentration $\hat{c}$, see Proposition 4.4. Introducing the mixed diffusion coefficient $\hat{\delta}^V = \frac{\delta_1 \beta e^{-V_1} + \delta_2 \alpha e^{-V_2}}{\beta e^{-V_1} + \alpha e^{-V_2}}$ and the mixed potential $\hat{V} = -\log(\frac{\beta}{\alpha+\beta} e^{-V_1} + \frac{\alpha}{\alpha+\beta} e^{-V_2})$, we obtain

$$\dot{\hat{c}} = \text{div}\left(\hat{\delta}^V \nabla \hat{c} + \hat{\delta}^V \hat{c} \nabla \hat{V}\right),$$

which is in accordance with [6] in the potential-free case $V = \text{const}$. Moreover, we obtain a natural coarse-grained gradient structure $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$, where $\hat{Q} = \text{Prob}(\Omega)$ is the coarse-grained state space and $\hat{\mathcal{E}}, \hat{\mathcal{R}}^*$ are the coarse-grained energy functional and dissipation potential, respectively. This coarse-grained gradient structure $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$ contains the same information as the effective gradient structure $(Q, \mathcal{E}, \mathcal{R}^*_{\text{eff}})$, although defined on a smaller state space, see Proposition 4.4.

The result on tilt EDP-convergence is an immediate consequence of the Γ-convergence result of the dissipation functional $\mathfrak{D}^\eta_\varepsilon$ (Theorem 5.12). The primal dissipation potential $\mathcal{R}_\varepsilon$ is given by an infimal sum consisting of diffusion fluxes and reaction fluxes coupled via a generalized continuity equation, see Sect. 3.3. Theorem 5.12 follows from the following observations: $\mathcal{R}^*_\varepsilon$ converges monotonically to a singular limit $\mathcal{R}^*_{\text{eff}}$, the primal dissipation potentials $\mathcal{R}_\varepsilon$ degenerate. It is not possible to control the rates of $\dot{\mu}_1$ and $\dot{\mu}_2$ separately by $\mathcal{R}_\varepsilon$, since the reaction flux between both species may become unbounded. Instead, it is possible to prove compactness for the sum (or slow variable) $\mu_1 + \mu_2$ by $\mathcal{R}_\varepsilon$, and proving convergence towards the slow-manifold where an equilibration takes place, i.e. $\alpha c_1^\varepsilon - \beta c_2^\varepsilon \to 0$. The two pieces of complementary information provide strong convergence of the densities c^ε in $L^1([0, T] \times \Omega)$. This procedure has been already successfully applied for linear and nonlinear reaction systems [30,34] and is here applied to a space-dependent evolution system. A posteriori we conclude that the limit measure μ^0 has indeed an absolutely continuous representative using results from [1]. The construction of the recovery sequence relies on the fact that the limit dissipation functional can be equivalently written as a functional of coarse-grained variables. Only the reaction flux, which is present for positive $\varepsilon > 0$ and hidden for $\varepsilon = 0$, has to be reconstructed by using the diffusion fluxes. Since the dissipation functional considers also fluctuations which may be neither strictly positive nor smooth in contrast to the solution of the linear reaction diffusion system (1.2), the construction of a recovery sequence is completed by a suitable approximation argument.

Let us finally mention, that the same results can also be established for reaction-diffusion systems, where more than two species are involved. Applying the coarse-graining and reconstruction machinery as developed in [30], a similar Γ-convergence result for the dissipation functional can be proved. For notational convenience we restrict to the two-species situation and briefly discuss the multi-species case in Sect. 6. We refer also to [42], where coarse-graining and reconstruction for concentrations as well as the fluxes is developed.

2 Gradient structures

2.1 Gradient systems and the energy-dissipation principle

Let us briefly recall what we mean with a gradient system. Following [29], we call a triple $(Q, \mathcal{E}, \mathcal{R})$ a *gradient system* if

(1) Q is a closed convex subset of a Banach space X
(2) $\mathcal{E} : Q \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ is a functional (such as the free energy)
(3) $\mathcal{R} : Q \times X \to \mathbb{R}_\infty$ is a dissipation potential, which means that for any $u \in Q$ the functional $\mathcal{R}(u, \cdot) : X \to \mathbb{R}_\infty$ is lower semicontinuous (lsc), nonnegative and convex, and it satisfies $\mathcal{R}(u, 0) = 0$.

We define the dual dissipation potential $\mathcal{R}^* : Q \times X^* \to [0, \infty]$ using the Legendre transform via

$$\mathcal{R}^*(u, \xi) = (\mathcal{R}(u, \cdot))^*(\xi) = \sup_{v \in X} \{\langle v, \xi \rangle - \mathcal{R}(u, v)\} .$$

The gradient system is uniquely described by $(Q, \mathcal{E}, \mathcal{R})$ or, equivalently by $(Q, \mathcal{E}, \mathcal{R}^*)$ and, in particular, in this paper we use the second representation.

For a sufficiently smooth energy function $\mathcal{E}$, its Gateaux-differential $\mathrm{D}\mathcal{E}(u) \in X^*$ describes the (negative) potential force of the system. The dynamics of a gradient system can be formulated in different ways as an equation in X, in $\mathbb{R}$ or in X^* (the dual Banach space of X), respectively:

(1) **Force balance in X^*:** $0 \in \partial_{\dot{u}} \mathcal{R}(u, \dot{u}) + \mathrm{D}\mathcal{E}(u) \in X^*$,
(2) **Power balance in $\mathbb{R}$:** $\mathcal{R}(u, \dot{u}) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) = -\langle \mathrm{D}\mathcal{E}(u), \dot{u} \rangle$,
(3) **Rate equation in X:** $\dot{u} \in \partial_\xi \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) \in X$.

(Here, ∂ denotes the subdifferential of convex analysis.) Equations (1) and (3) are called *gradient flow equation* associated with $(Q, \mathcal{E}, \mathcal{R}^*)$. The equivalent formulations rely on the following fact: Let X be a reflexive Banach space and $\Psi : X \to \mathbb{R}_\infty$ be proper, convex and lsc. Then for every $\xi \in X^*$ and $v \in X$ the following three statements, the so-called *Legendre-Fenchel-equivalences*, are equivalent:

$$v \in \partial \Psi^*(\xi) \quad \Leftrightarrow \quad \Psi(v) + \Psi^*(\xi) = \langle \xi, v \rangle \quad \Leftrightarrow \quad \xi \in \partial \Psi(v).$$

Especially the second dynamic formulation, the power balance (2), is interesting for us. Integrating the power balance (2) in time from 0 to T and using the chain rule for the time-derivative of $t \mapsto \mathcal{E}(u(t))$, we get another equivalent formulation of the dynamics of the gradient system, which is called *Energy-Dissipation-Balance*:

$$\text{(EDB)} \qquad \mathcal{E}(u(T)) + \int_0^T \left[\mathcal{R}(u, \dot{u}) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) \right] \mathrm{d}t = \mathcal{E}(u(0)). \qquad (2.1)$$

Equation (EDB) compares the energy of the system at time $t = 0$ and at time $t = T$, the difference is described by the total dissipation from $t = 0$ to $t = T$. This gives rise to another definition: We define the *De Giorgi dissipation functional* as

$$\mathfrak{D}(u) = \int_0^T \left[\mathcal{R}(u, \dot{u}) + \mathcal{R}^*(u, -\mathrm{D}\mathcal{E}(u)) \right] \mathrm{d}t,$$

for $u \in \mathrm{W}^{1,1}([0, T], Q)$ and extend it to infinity otherwise. The following *energy-dissipation principle* provides the definition for solutions of the gradient flow equation, see e.g. [1, Prop.

1.4.1], [2, Def. 1.1], [31, Thm 2.5]. In particular, it is the starting point for our further analysis.

Definition 2.1 We say $u \in W^{1,1}([0, T], Q)$ is a solution of the gradient flow equation (1) or (3) induced by the gradient system $(Q, \mathcal{E}, \mathcal{R}^*)$, if $\mathcal{E}(u(0)) < \infty$ and the energy-dissipation balance holds.

2.2 Definition of EDP-convergence

The definition of EDP-convergence for gradient systems relies on the notion of Γ-convergence for functionals (cf. [12]). If Y is a Banach space and $I_\varepsilon : Y \to \mathbb{R}_\infty$ we write $I_\varepsilon \xrightarrow{\Gamma} I_0$ and $I_\varepsilon \overset{\Gamma}{\rightharpoonup} I_0$ for Γ-convergence in the strong and weak topology, respectively. If both holds this is called Mosco-convergence and written as $I_\varepsilon \xrightarrow{M} I_0$.

For families of gradient systems $(X, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$, three different levels of EDP-convergence are introduced and discussed in [14,31]: simple EDP-convergence, EDP-convergence with tilting and contact EDP-convergence with tilting. EDP-convergence with tilting is the strongest notion implying the other two. Here we will only use the first two notions. For all three notions the choice of weak or strong topology is still to be decided according to the specific problem.

Definition 2.2 *(Simple EDP-convergence)* A family of gradient structures $(Q, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ is said to *EDP-converge* to the gradient system $(Q, \mathcal{E}_0, \mathcal{R}_{\mathrm{eff}})$ if the following conditions hold:

1. $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ on $Q \subset X$;
2. $\mathfrak{D}_\varepsilon$ strongly Γ-converges to $\mathfrak{D}_0$ on $L^1([0, T], Q)$ conditioned to bounded energies (we write $\mathfrak{D}_\varepsilon \xrightarrow{\Gamma_E} \mathfrak{D}_0$), i.e. we have

   (a) (Liminf-estimate) For all strongly converging families $u_\varepsilon \to u$ in $L^1([0, T], Q)$ which satisfy $\sup_{\varepsilon>0} \operatorname{ess\,sup}_{t\in[0,T]} \mathcal{E}_\varepsilon(u_\varepsilon(t)) < \infty$, we have $\liminf_{\varepsilon\to 0^+} \mathfrak{D}_\varepsilon(u_\varepsilon) \geq \mathfrak{D}_0(u)$,
   (b) (Limsup-estimate) For all $\widetilde{u} \in L^1([0, T], Q)$ there exists a strongly converging family $u_\varepsilon \to \widetilde{u}$ in $L^1([0, T], Q)$ with $\sup_{\varepsilon>0} \operatorname{ess\,sup}_{t\in[0,T]} \mathcal{E}_\varepsilon(\widetilde{u}_\varepsilon(t)) < \infty$ such that we have $\limsup_{\varepsilon\to 0^+} \mathfrak{D}_\varepsilon(u_\varepsilon) \leq \mathfrak{D}_0(\widetilde{u})$.

3. There is an effective dissipation potential $\mathcal{R}_{\mathrm{eff}} : Q \times X \to \mathbb{R}_\infty$ such that $\mathfrak{D}_0$ takes the form of a dual sum, namely $\mathfrak{D}_0(u) = \int_0^T \{\mathcal{R}_{\mathrm{eff}}(u, \dot{u}) + \mathcal{R}_{\mathrm{eff}}^*(u, -D\mathcal{E}_{\mathrm{eff}}(u))\} dt$.

   Similarly, one can also use weak Γ- or Mosco-convergence conditioned to bounded energy, which we will then write as $\mathfrak{D}_\varepsilon \overset{\Gamma_E}{\rightharpoonup} \mathfrak{D}_0$ and $\mathfrak{D}_\varepsilon \xrightarrow{M_E} \mathfrak{D}_0$. In fact, for our fast-slow reaction systems we are going to prove $\mathfrak{D}_\varepsilon \xrightarrow{M_E} \mathfrak{D}_0$.

   A general feature of EDP-convergence is that, under suitable conditions, sufficiently smooth solutions u of the gradient flow equation $\dot{u} = \partial_\xi \mathcal{R}_{\mathrm{eff}}^*(u, -D\mathcal{E}_0(u))$ of the effective gradient system $(X, \mathcal{E}_0, \mathcal{R}_{\mathrm{eff}})$ are indeed limits of solutions u^ε of the gradient flow equation $\dot{u}^\varepsilon = \partial_\xi \mathcal{R}_\varepsilon^*(u^\varepsilon, -D\mathcal{E}_\varepsilon(u^\varepsilon))$, see e.g. [11, Thm. 11.3], [30, Lem. 3.4] and [31, Lem. 2.8].

   A strengthening of simple EDP-convergence is the so-called *EDP-convergence with tilting*. This notion involves the tilted energy functionals $\mathcal{E}_\varepsilon^\eta : Q \ni u \mapsto \mathcal{E}_\varepsilon(u) - \langle \eta, u \rangle$, where the tilt η (also called *external loading*) varies through the whole dual space X^*.

Definition 2.3 (EDP-convergence with tilting (cf. [31, Def. 2.14)) A family of gradient structures $(Q, \mathcal{E}_\varepsilon, \mathcal{R}_\varepsilon)$ is said to *EDP-converge with tilting* to the gradient system $(Q, \mathcal{E}_0, \mathcal{R}_{\mathrm{eff}})$, if for all tilts $\eta \in X^*$ we have that $(Q, \mathcal{E}_\varepsilon^\eta, \mathcal{R}_\varepsilon)$ EDP-converges to $(Q, \mathcal{E}_0^\eta, \mathcal{R}_{\mathrm{eff}})$.

Clearly, we have that $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_0$ implies $\mathcal{E}_\varepsilon^\eta \xrightarrow{\Gamma} \mathcal{E}_0^\eta$ for all $\eta \in X^*$ (and similarly for weak Γ-convergence), since the linear tilt $u \mapsto -\langle \eta, u \rangle$ is weakly continuous. The main and nontrivial assumption is that additionally

$$\mathfrak{D}_\varepsilon^\eta : u \mapsto \int_0^T \left\{ \mathcal{R}_\varepsilon(u, \dot{u}) + \mathcal{R}_\varepsilon^*(u, \eta - D\mathcal{E}_\varepsilon(u)) \right\} dt$$

Γ-converges in $L^1([0, T], Q)$ to $\mathfrak{D}_0^\eta$ for all $\eta \in X^*$ and that this limit $\mathfrak{D}_0^\eta$ is given in $\mathcal{R} \oplus \mathcal{R}^*$-form with $\mathcal{R}_{\text{eff}}$ via

$$\mathfrak{D}_0^\eta(u) = \int_0^T \left\{ \mathcal{R}_{\text{eff}}(u, \dot{u}) + \mathcal{R}_{\text{eff}}^*(u, \eta - D\mathcal{E}_{\text{eff}}(u)) \right\} dt.$$

The main point is that $\mathcal{R}_{\text{eff}}$ remains independent of $\eta \in X^*$. We refer to [31] for a discussion of this and the other two notions of EDP-convergence.

3 Gradient system of reaction-diffusion systems

Now, we present the gradient system $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$, which induces the reaction-diffusion system (1.2) and discuss its properties. In Sect. 3.2 we formally derive the gradient flow equation of the gradient system including general tilts of the energy. In Sect. 3.3 we compute the primal dissipation potential $\mathcal{R}_\varepsilon$, which is only implicitly given via a infimal-convolution, and the total dissipation functional $\mathfrak{D}_\varepsilon^\eta$, which will be the main object of interest in Sect. 4. In Sect. 3, the computations are basically formal; the precise functional analytic setting is presented in Sect. 4 which also includes the Γ-convergence and EDP-convergence result.

3.1 Gradient structure for the linear reaction system

The gradient system $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$ for the linear fast-slow reaction-diffusion system (1.2) is defined as follows: The state space is the space of probability measure on $\Omega \times \{1, 2\}$

$$Q := \text{Prob}(\Omega \times \{1, 2\}) = \{\mu = (\mu_1, \mu_2) : \mu_i \in \mathcal{M}(\Omega), \ \mu_i \geq 0, \ \sum_{i=1}^2 \mu_i(\Omega) = 1\},$$

where we assume that $\Omega \subset \mathbb{R}^d$ is a compact domain with normalized mass $|\Omega| = 1$. The driving functional $\mathcal{E} : Q \to \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ is the free energy of the reaction-diffusion system. It is finite for measures $\mu = (\mu_1, \mu_2)$ with Lebesgue density $c = (c_1, c_2)$ only and has the form

$$\mathcal{E}(\mu) := \begin{cases} \int_\Omega \sum_{j=1}^2 E_B\left(\frac{c_j}{w_j}\right) w_j dx, & \text{if } \mu = c \cdot dx \\ \infty, & \text{otherwise.} \end{cases} \tag{3.1}$$

where the Boltzmann function is defined as $E_B(r) = r \log r - r + 1$ and the positive stationary measure is given by $w = \frac{1}{\alpha+\beta}(\beta, \alpha)^{\text{T}}$. Note that the stationary measure w as well as the energy $\mathcal{E}$ is ε-independent. As usual in convex analysis the differential $D\mathcal{E}$ of the energy $\mathcal{E}$ is only defined in the domain of $\mathcal{E}$, i.e. for measures with Lebesgue density. Here, all calculations are on a formal level assuming that the densities are smooth, the functionals can be extended to a linear space and being Gateaux differentiable. For a mathematical rigorous

theory of subdifferentials in the space of probability measures we refer to Section 10 of [1]. For measures μ with Lebesgue density c we have

$$\mathrm{D}\mathcal{E}(\mu) = \sum_{j=1}^{2}\left(\log c_j - \log w_j\right) = \sum_{j=1}^{2}\left(\log\frac{c_j}{w_j}\right).$$

As the equation splits into a diffusion and reaction part, so does the dual dissipation functional. We define

$$\mathcal{R}_\varepsilon^*(\mu,\xi) := \mathcal{R}_{\mathrm{diff}}^*(\mu,\xi) + \mathcal{R}_{\mathrm{react},\varepsilon}^*(\mu,\xi),$$

with

$$\mathcal{R}_{\mathrm{diff}}^*(\mu,\xi) := \frac{1}{2}\int_\Omega\sum_{j=1}^{2}\delta_j|\nabla\xi_j(x)|^2\mathrm{d}\mu_j,$$

$$\mathcal{R}_{\mathrm{react},\varepsilon}^*(\mu,\xi) := \frac{1}{\varepsilon}\int_\Omega \mathsf{C}^*(\xi_1(x) - \xi_2(x))\,\mathrm{d}\sqrt{\mu_1\mu_2},$$

where we use the cosh-function $\mathsf{C}^*(x) = 4\left(\cosh(x/2) - 1\right)$ and for measures μ with Lebesgue density c we have $\mathrm{d}\sqrt{\mu_1\mu_2} := \sqrt{c_1 c_2}\mathrm{d}x$.

The diffusion part $\mathcal{R}_{\mathrm{diff}}^*$ induces the Wasserstein distance on Q. The ε-dependent reaction part $\mathcal{R}_{\mathrm{react},\varepsilon}^*$ forces the evolution close to a linear submanifold given by

$$\mathcal{R}_{\mathrm{react},\varepsilon}^*(\mu, -\mathrm{D}\mathcal{E}(\mu)) = 0 \quad\Leftrightarrow\quad \alpha c_1 - \beta c_2 = 0.$$

Note, that since $\mathcal{R}_{\mathrm{react},\varepsilon}^*$ is not 2-homogeneous, it does not define a metric on Q. We refer to [39] which treats similar and general dissipation potentials and understands them as generalized transport costs on discrete spaces. Note that $\mathcal{R}_\varepsilon^*$ does not depend on the stationary measure w explicitly, as highlighted in [30].

3.2 The tilted gradient flow equation

To exploit the full information of the dissipation potential, we consider general tilted energies. Considering two free energies $\mathcal{E}, \widetilde{\mathcal{E}}$ of the form (3.1) with different positive stationary measures $w, \widetilde{w}$ (which may be space dependent in general), one easily sees that

$$\widetilde{\mathcal{E}}(\mu) = \mathcal{E}(\mu) + \sum_{i=1}^{2}\int_\Omega \log\left(\frac{w_i}{\widetilde{w}_i}\right)\mathrm{d}\mu_i.$$

Hence, changing the underlying stationary measure corresponds to a linear tilt of the energy by a two component potential $V = (V_1, V_2)$ where $V_i = \log\left(\frac{w_i}{\widetilde{w}_i}\right)$. On the other hand, introducing a tilted energy

$$\mathcal{E}^V(\mu) := \mathcal{E}(\mu) + \sum_{i=1}^{2}\int_\Omega V_i\mathrm{d}\mu_i,$$

where $V \in \mathrm{C}^1(\Omega, \mathbb{R}^2)$ is a two component smooth potential, its stationary measure w^V on the space $Q = \mathrm{Prob}(\Omega \times \{1, 2\})$ has the form

$$w_i^V = \frac{1}{Z}w_i\mathrm{e}^{-V_i}, \text{ where } Z := \sum_{i=1}^{2}\int_\Omega w_i\mathrm{e}^{-V_i}\mathrm{d}x. \tag{3.2}$$

We introduce $\eta_i := e^{-V_i}$ and clearly, we have $\eta_i > 0$ on $\Omega \subset \mathbb{R}^d$.

Next, we formally compute the tilted gradient flow equation $\dot{\mu} = \partial_\xi \mathcal{R}^*_\varepsilon(\mu, -D\mathcal{E}^V(\mu))$, which is induced by the gradient system $(Q, \mathcal{E}^V, \mathcal{R}^*_\varepsilon = \mathcal{R}^*_{\text{diff}} + \mathcal{R}^*_{\text{react},\varepsilon})$. First, we observe that $\mathcal{E}(\mu) < \infty$ if and only if $\mathcal{E}^V(\mu) < \infty$. Inserting $\xi_i = \left(-D\mathcal{E}^V(\mu)\right)_i$ into $\partial_\xi \mathcal{R}^*_{\text{diff},\varepsilon}(\mu, \xi)$, we see that

$$\partial_\xi \mathcal{R}^*_{\text{diff}}(\mu, \cdot)|_{\xi=-D\mathcal{E}^V(\mu)} = -(\text{div}(\delta_i c_i \nabla(-\log(c_i/w_i) - V_i))_{i=1,2} = \text{div}\,(\delta_i \nabla c_i + \delta_i c_i \nabla V_i)_{i=1,2},$$

which is the right-hand side of a system of two uncoupled drift-diffusion equations or Fokker-Planck equations for the concentrations c_i, where the fluxes are given by a diffusion part $-\delta_i \nabla c_i$ and a drift part $-\delta_i c_i \nabla V_i$.

For the reaction part of the dual dissipation potential, we insert $\xi_i = \left(-D\mathcal{E}^V(\mu)\right)_i$ into $\partial_\xi \mathcal{R}^*_{\text{react},\varepsilon}(\mu, \xi)$. On readily verifies the identity $(\mathsf{C}^*)'(\log p - \log q) = \frac{p-q}{\sqrt{pq}}$ for the cosh-function and conclude the linear term

$$\sqrt{c_1 c_2}\,(\mathsf{C}^*)'(\xi_1(x) - \xi_2(x))|_{\xi=-D\mathcal{E}^V(\mu)} = \sqrt{c_1 c_1} \frac{\frac{c_2}{w_2 \eta_2} - \frac{c_1}{w_1 \eta_1}}{\sqrt{\frac{c_1}{w_1 \eta_1} \frac{c_2}{w_2 \eta_2}}} = \sqrt{w_1 \eta_1 w_2 \eta_2}\left(\frac{c_2}{w_2 \eta_2} - \frac{c_1}{w_1 \eta_1}\right).$$

Hence, we get a tilted Markov generator of the form

$$\partial_\xi \mathcal{R}^*_{\text{react},\varepsilon}(\mu, -D\mathcal{E}^V(\mu)) = \frac{1}{\varepsilon}\begin{pmatrix} -\sqrt{\frac{\alpha}{\beta}\frac{\eta_2}{\eta_1}} & \sqrt{\frac{\beta}{\alpha}\frac{\eta_1}{\eta_2}} \\ \sqrt{\frac{\alpha}{\beta}\frac{\eta_2}{\eta_1}} & -\sqrt{\frac{\beta}{\alpha}\frac{\eta_1}{\eta_2}} \end{pmatrix} c = \frac{1}{\varepsilon}\begin{pmatrix} -\sqrt{\frac{\alpha}{\beta}}e^{\frac{V_1-V_2}{2}} & \sqrt{\frac{\beta}{\alpha}}e^{\frac{V_2-V_1}{2}} \\ \sqrt{\frac{\alpha}{\beta}}e^{\frac{V_1-V_2}{2}} & -\sqrt{\frac{\beta}{\alpha}}e^{\frac{V_2-V_1}{2}} \end{pmatrix} c,$$

which has the space dependent stationary measure (3.2).

Summarizing, the tilted evolution equation for the density c has the form

$$\frac{d}{dt}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \text{div}\left(\begin{pmatrix} \delta_1 \nabla c_1 \\ \delta_2 \nabla c_2 \end{pmatrix} + \begin{pmatrix} \delta_1 c_1 \nabla V_1 \\ \delta_2 c_2 \nabla V_2 \end{pmatrix}\right) + \frac{1}{\varepsilon}\begin{pmatrix} -\sqrt{\frac{\alpha}{\beta}}e^{\frac{V_1-V_2}{2}} & \sqrt{\frac{\beta}{\alpha}}e^{\frac{V_2-V_1}{2}} \\ \sqrt{\frac{\alpha}{\beta}}e^{\frac{V_1-V_2}{2}} & -\sqrt{\frac{\beta}{\alpha}}e^{\frac{V_2-V_1}{2}} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

(3.3)

which is a linear reaction-drift-diffusion system with space dependent reaction rates. In the special case without external forcing $V = \text{const}$, we get the linear reaction diffusion system (1.2). Note that the reaction part has the property that the product of the off-diagonal elements is constantly one. In particular, not all general linear reaction-drift-diffusion system with space dependent reaction rates for two species can be expressed in the form (3.3) and are induced by the gradient system $(Q, \mathcal{E}^V, \mathcal{R}^*_\varepsilon = \mathcal{R}^*_{\text{diff}} + \mathcal{R}^*_{\text{react},\varepsilon})$.

3.3 The dissipation functional

The dissipation functional $\mathfrak{D}_\varepsilon$ consists of two parts: the velocity part given by the primal dissipation potential $\mathcal{R}_\varepsilon$ and the slope-part (sometimes also called *Fisher information*) $\mathcal{R}^*_\varepsilon(\mu, -D\mathcal{E}(\mu))$. Again all computations are formal and we always assume that the measure μ has a Lebesgue density c. The precise functional analytic setting is presented in the Sect. 4.

The primal dissipation potential $\mathcal{R}_\varepsilon$, given by the Legendre transform of the dual dissipation potential $\mathcal{R}^*_\varepsilon = \mathcal{R}^*_{\text{diff}} + \mathcal{R}^*_{\varepsilon,\text{react}}$ with respect to the second variable, can be computed via inf-convolution of $\mathcal{R}_{\text{diff}}$ and $\mathcal{R}_{\text{react},\varepsilon}$. First, we compute both primal dissipation potentials separately.

The primal dissipation potential of the diffusion part can be computed by the Legendre transform of $\mathcal{R}_{\text{diff}}^*$

$$\mathcal{R}_{\text{diff}}(\mu, v) = \sup_{\xi=(\xi_1,\xi_2)} \left\{ \sum_{j=1}^{2} \int_{\Omega} \xi_j v_j - \frac{1}{2}\delta_j c_j |\nabla\xi_j|^2 \, dx \right\}.$$

The force $\xi^* = (\xi_1^*, \xi_2^*)$ where the supremum is attained can be calculated, is uniquely determined up to a constant, and satisfies the following elliptic equation (in weak sense)

$$-\text{div}(\delta_j c_j \nabla\xi_j^*) = v_j \quad \text{in } \Omega, \quad (\delta_j c_j \nabla\xi_j^*) \cdot v = 0 \quad \text{on } \partial\Omega.$$

Introducing the (uniquely determined) diffusion flux $J_j = \delta_j c_j \nabla\xi_j^*$, we have the explicit form of $\mathcal{R}_{\text{diff}}$, namely,

$$\mathcal{R}_{\text{diff}}(\mu, v) = \sum_{j=1}^{2} \int_{\Omega} \xi_j^* v_j - \frac{1}{2}\delta_j c_j |\nabla\xi_j^*|^2 \, dx = \frac{1}{2}\sum_{j=1}^{2} \int_{\Omega} \xi_j^* v_j \, dx = \frac{1}{2}\sum_{j=1}^{2} \int_{\Omega} \frac{|J_j|^2}{\delta_j c_j} \, dx,$$

for positive concentrations $c_j > 0$. To capture also general non-negative concentrations $c_j \geq 0$, we introduce the following notation: For a convex, lsc. function $\mathsf{F} : X \to [0, \infty]$ on a reflexive and separable Banach space X with Legendre dual F^*, we define the function $\widetilde{\mathsf{F}} : [0, \infty[\times X \to [0, \infty]$ by

$$\widetilde{\mathsf{F}}(a, x) := \left(a \, \mathsf{F}^*(\cdot)\right)^*(x) = \begin{cases} a \, \mathsf{F}\left(\frac{1}{a}x\right) & \text{for } a > 0, \\ \chi_0(x) & \text{for } a = 0. \end{cases} \tag{3.4}$$

In particular, introducing the quadratic function $\mathsf{Q}(x) = \frac{1}{2}|x|^2$ on $\mathbb{R}^d$, the primal dissipation potential of the diffusion part $\mathcal{R}_{\text{diff}}^*$ is given by

$$\mathcal{R}_{\text{diff}}(\mu, v) = \sum_{j=1}^{2} \int_{\Omega} \widetilde{\mathsf{Q}}\left(\delta_j c_j, J_j\right) dx,$$

where J_j is, by definition, the unique solution of the equation $v_j + \text{div} J_j = 0$ with $J_j \cdot v = 0$ on $\partial\Omega$.

A direct computation shows that the primal dissipation potential of the reaction part is

$$\mathcal{R}_{\text{react},\varepsilon}^*(\mu, b) = \begin{cases} \int_{\Omega} \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right) dx, & \text{for } b_1 + b_2 = 0 \\ \infty & \text{for } b_1 + b_2 \neq 0 \end{cases},$$

where $\mathsf{C} = \left(\mathsf{C}^*\right)^*$ is the Legendre transform of the cosh-function $\mathsf{C}^*(x) = 4\left(\cosh(x/2) - 1\right)$. In the following, we use the inequality

$$\frac{1}{2}|r| \cdot \log(|r| + 1) \leq \mathsf{C}(r) \leq 2|r| \cdot \log(|r| + 1), \tag{3.5}$$

In particular, this inequality implies that the Orlicz class for $A \subset \mathbb{R}^d$ given by all functions $u \in L^1(A)$ such that $\int_A \mathsf{C}(u) dx < \infty$ is, in fact, a Banach space with the norm $\|u\|_{\mathsf{C}} = \sup_{\int_A \mathsf{C}^*(v) \leq 1} \left|\int_A uv \, dx\right|$. In the following the Orlicz space is denoted by $L^{\mathsf{C}}(A)$.

Importantly, functions $\widetilde{\mathsf{Q}}, \widetilde{\mathsf{C}}$ as well as the functionals $\mathcal{R}_{\text{diff}}, \mathcal{R}_{\text{react},\varepsilon}$ are convex on their domain of definition.

The primal dissipation potential $\mathcal{R}_\varepsilon$ is the inf-convolution of $\mathcal{R}_{\mathrm{diff}}$ and $\mathcal{R}_{\mathrm{react},\varepsilon}$, and is given by

$$\mathcal{R}_\varepsilon(\mu, v) = \inf_{v = u_1 + u_2} \{\mathcal{R}_{\mathrm{diff}}(\mu, u_1) + \mathcal{R}_{\mathrm{react},\varepsilon}(\mu, u_2)\}$$

$$= \inf_{J,b} \left\{ \sum_{j=1}^2 \int_\Omega \widetilde{\mathsf{Q}}\left(\delta_j c_j, J_j\right) \mathrm{d}x + \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2(x)\right) \mathrm{d}x : \left\{ \begin{array}{l} v_1 = -\mathrm{div}\, J_1 + b_1 \\ v_2 = -\mathrm{div}\, J_2 + b_2 \\ b_1 + b_2 = 0 \end{array} \right\} \right\}.$$

In time-integrated form we get for $v = \dot\mu$ that

$$\int_0^T \mathcal{R}_\varepsilon(\mu, \dot\mu)\, \mathrm{d}t = \inf_{v = v_1 + v_2} \int_0^T \{\mathcal{R}_{\mathrm{diff}}(\mu, v_1) + \mathcal{R}_{\mathrm{react},\varepsilon}(\mu, v_2)\}\, \mathrm{d}t$$

$$= \inf_{J,b} \left\{ \int_0^T \left\{ \sum_{j=1}^2 \int_\Omega \widetilde{\mathsf{Q}}\left(\delta_j c_j, J_j\right) \mathrm{d}x + \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2(x)\right) \right\} \mathrm{d}x\, \mathrm{d}t : (c, J, b) \in \mathrm{(gCE)} \right\},$$

where we introduce the notation of a (linear) generalized continuity equation

$$(c, J, b) \in \mathrm{(gCE)} \quad \Leftrightarrow \quad \left\{ b_1 + b_2 = 0 \text{ and } \left\{ \begin{array}{l} \dot c_1 = -\mathrm{div}\, J_1 + b_1 \\ \dot c_2 = -\mathrm{div}\, J_2 + b_2 \end{array} \right\} \text{ and } J_j \cdot v = 0 \text{ on } \partial\Omega \right\}.$$

Without the reaction part, $\int_0^T \mathcal{R}_\varepsilon \mathrm{d}t$ is the dynamic formulation à la Benamou–Brenier of the Wasserstein distance in Q [4], which can be equivalently written in the form

$$\mathcal{W}_2(\mu_0, \mu_1)^2 = \inf \left\{ \int_0^1 \int_\Omega \sum_{j=1}^2 \delta_j |v_j|^2 \mathrm{d}\mu_j : \dot\mu_j + \mathrm{div}(\mu_j v_j) = 0,\ \mu_{j,0} = \mu_0,\ \mu_{j,1} = \mu_1 \right\}$$

expressed in terms of transport velocities $v_j = J_j/c_j$. The Wasserstein distance can be interpreted as a cost in transporting mass from one measure μ_0 to μ_1. In our situation $\int_0^T \mathcal{R}_\varepsilon \mathrm{d}t$ is jointly convex in c, J and b and corresponds to a modified cost function which also takes the reaction fluxes into account. The optimal diffusion fluxes J_j and reaction fluxes b_j have to satisfy the generalized continuity equation. Note that $\int_0^T \mathcal{R}_\varepsilon \mathrm{d}t$ does not induce a metric on Q since the reaction part is not quadratic.

Next, we compute the tilted slope part $\mathcal{R}_\varepsilon^*(\mu, -\mathrm{D}\mathcal{E}^V(\mu))$. To do this, we introduce the relative densities ρ^V of μ w.r.t. the stationary measure $w^V \mathrm{d}x$, i.e. $\rho_j^V = \frac{\mathrm{d}\mu_j}{w_j^V \mathrm{d}x} = \frac{c_j}{w_j^V}$, where by (3.2) the stationary measure is $w_j^V = \frac{1}{Z} w_i \mathrm{e}^{-V_j}$. Since $V \in \mathrm{C}^1(\Omega, \mathbb{R}^2)$ and $\Omega \subset \mathbb{R}^d$ is compact, μ is absolutely continuous w.r.t. the Lebesgue measure $\mathrm{d}x$ if and only if it is w.r.t. the stationary measure $w^V \mathrm{d}x$. Inserting $\xi = -\mathrm{D}\mathcal{E}^V(\mu) = -(\log(c_i/w_i) + V_i)_{i=1,2}$ in the dual dissipation potential $\mathcal{R}_\varepsilon^*$, we get for the diffusive part

$$\mathcal{R}_{\mathrm{diff}}^*(\mu, -\mathrm{D}\mathcal{E}^V(\mu)) = \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j c_j |\nabla(\log c_j/w_j + V_j)|^2 \mathrm{d}x.$$

Using $w_j^V = \frac{1}{Z} w_j \mathrm{e}^{-V_j}$, a short calculation shows $\delta_j c_j |\nabla(\log c_j/w_j + V_j)|^2 = \delta_j w_j^V \frac{|\nabla \rho_j^V|^2}{\rho_j^V}$. Hence, we have

$$\mathcal{R}_{\mathrm{diff}}^*(\mu, -\mathrm{D}\mathcal{E}^V(\mu)) = \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{|\nabla \rho_j^V|^2}{\rho_j^V} \mathrm{d}x.$$

For the reaction part, we use the identity $C^*(\log p - \log q) = 2\frac{(\sqrt{p}-\sqrt{q})^2}{\sqrt{pq}}$ and get

$$
\begin{aligned}
\mathcal{R}^*_{\text{react},\varepsilon}(\mu, -D\mathcal{E}^V(\mu)) &= 2\int_\Omega \frac{1}{\varepsilon}\sqrt{c_1 c_2}\frac{\left(\sqrt{c_1/\eta_1 w_1} - \sqrt{c_2/\eta_2 w_2}\right)^2}{\sqrt{c_1 c_2/\eta_1 w_1 \eta_2 w_2}}\,dx \\
&= \frac{2}{\varepsilon}\int_\Omega \sqrt{w_1^V w_2^V}\left(\sqrt{\rho_1^V} - \sqrt{\rho_2^V}\right)^2 x.
\end{aligned}
$$

4 EDP-convergence result

In this section we state the EDP-convergence result for the gradient systems $(Q, \mathcal{E}, \mathcal{R}^*_\varepsilon)$ to $(Q, \mathcal{E}, \mathcal{R}^*_{\text{eff}})$ and discuss the properties of the effective gradient system $(Q, \mathcal{E}, \mathcal{R}^*_{\text{eff}})$. Since the energy $\mathcal{E}$ is ε-independent the major challenge is to prove Γ-convergence of the dissipation functional $\mathfrak{D}^V_\varepsilon$, which is a functional defined on the space of trajectories in the state space Q. To be mathematical precise, we first fix the functional analytic setting.

The state space $Q = \text{Prob}(\Omega \times \{1, 2\})$ is equipped with the p-Wasserstein distance $d_{\mathcal{W}_p}$, where in our situation either $p = 1$ or $p = 2$. Recall that for any compact Euclidean subspace $E \subset \mathbb{R}^k$ the p-Wasserstein distance is defined on the space of probability measures $\text{Prob}(E)$ by

$$
\mathcal{W}_p(\mu^1, \mu^2)^p = \min_{\gamma \in \Gamma(\mu^1, \mu^2)} \int_E |x - y|^p d\gamma(x, y),
$$

where $\Gamma(\mu^1, \mu^2)$ is the set of all transport plans with marginals μ^1 and μ^2 (see e.g. [1]). The p-Wasserstein distance $\mathcal{W}_p$ metrizises the weak*-topology of measures, i.e. convergence tested against continuous functions on E. In the following we will consider either $E = \Omega$ or $E = \Omega \times \{1, 2\}$.

To define the topology in the space of trajectories on Q, we start very coarse, where we understand the trajectories on Q as measures in space and time. We denote the space of trajectories by $L^\infty_w([0, T], Q)$ equipped with the weak*-measurability. The weak*-convergence is defined as usual by

$$
\mu^\varepsilon(\cdot) \to \mu^0(\cdot) \iff \forall i \in \{1, 2\}, \forall \phi \in C^\infty(\Omega \times [0, T]) : \int_0^T \int_\Omega \phi\,d\mu_i^\varepsilon\,dt \to \int_0^T \int_\Omega \phi\,d\mu_i^0\,dt.
\tag{4.1}
$$

A finer topology, which enables to prove compactness and evaluate the effective dissipation functional is then given by the a priori bounds

$$
\sup_{\varepsilon \in]0,1]} \mathfrak{D}^V_\varepsilon(\mu^\varepsilon) \leq C, \quad \sup_{\varepsilon \in]0,1]} \operatorname{ess\,sup}_{t \in [0,T]} \mathcal{E}(\mu^\varepsilon(t)) \leq C.
\tag{4.2}
$$

In fact, as presented in Sect. 5.1, these bounds provide that the measures μ^ε have Lebesgue densities c^ε which converge strongly in $L^1([0, T] \times \Omega, \mathbb{R}^2_{\geq 0})$. Moreover, the limiting coarse-grained measure $\hat{\mu}^0 = \mu_1^0 + \mu_2^0$ has an representative which is absolutely continuous in time with values in $(\text{Prob}(\Omega), \mathcal{W}_2)$, i.e. there is a function $m \in L^1([0, T])$ such that for all $t, s \in [0, T]$ with $s \leq t$ we have

$$
\mathcal{W}_2\left(\hat{\mu}^0(s), \hat{\mu}^0(t)\right) \leq \int_s^t m(r)\,dr.
$$

Each component μ_i^0, $i = 1, 2$ is not a trajectory in the space of probability measure, but in the space of non-negative Radon measures. Proposition 5.11 shows that μ_i^0 is absolutely continuous in time with values in $(\mathcal{M}_+(\Omega), \mathcal{W}_1)$ exploiting the dual formulation of the 1-Wasserstein distance (see e.g. [15]). This compactness result is comparable to the result of Bothe and Hilhorst [6], where also strong convergence of solutions $c = (c_1, c_2)$ is proved. In particular, similar to the space independent situation in [30,34,41] one cannot guarantee that $\mu^\varepsilon(t) \to \mu^0(t)$ in Q for all times $t \in [0, T]$ as jumps in time cannot be excluded. Instead the limit $\mu^0 = c^0 \, dx$ has an absolutely continuous representative.

4.1 Main theorem

Let us state our main EDP-convergence result. For doing this, we define for $V \in C^1(\Omega, \mathbb{R}^2)$ the total dissipation functional on $L_w^\infty([0, T], Q)$ as

$$
\mathfrak{D}_\varepsilon^V(\mu) = \begin{cases} \int_0^T \left\{ \mathcal{R}_\varepsilon(\mu, \dot\mu) + \mathcal{R}_\varepsilon^*(\mu, -D\mathcal{E}^V(\mu)) \right\} dt, & \mu \in AC([0, T], Q), \mu = c \, dx \text{ a.e. in } [0, T] \\ \infty & \text{otherwise.} \end{cases}
$$

If $\mu = c \, dx$ a.e. in $[0, T]$, then the dissipation functional is given by

$$
\int_0^T \mathcal{R}_\varepsilon(\mu, \dot\mu) + \mathcal{R}_\varepsilon^*(\mu, -V - D\mathcal{E}(\mu)) dt
$$

$$
= \inf_{(c, J, b) \in (\text{gCE})} \left\{ \int_0^T \left\{ \int_\Omega \sum_{j=1}^2 \widetilde{Q}(\delta_j c_j, J_j) dx + \int_\Omega \widetilde{C}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2(x)\right) dx \right\} dt \right\} +
$$

$$
+ \int_0^T \left\{ \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{\left|\nabla \rho_j^V\right|^2}{\rho_j^V} dx + \frac{2}{\varepsilon} \int_\Omega \sqrt{w_1^V w_2^V} \left(\sqrt{\rho_1^V} - \sqrt{\rho_2^V}\right)^2 dx \right\} dt,
$$

(4.3)

where the infimum is taken over all Borel fluxes $J_j \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$, $b_j \in \mathcal{M}([0, T] \times \Omega, \mathbb{R})$ which satisfy the generalized continuity equation (gCE) in the sense of distributions, i.e.

$$
\forall j = 1, 2 \ \forall \phi \in C_c^\infty(]0, T[\times \Omega) : \int_0^T \int_\Omega \dot\phi c_j - \nabla \phi \cdot J_j \, dx \, dt = -\int_0^T \int_\Omega b_j \phi \, dx \, dt,
$$

$$
J_j \cdot \nu = 0 \quad \text{on } \partial \Omega.
$$

Strictly speaking, the functions $\widetilde{Q}$ and $\widetilde{C}$ are not defined for measures J_j, b_j and hence, the formula for the dissipation functional (4.3) is a priori not correct. In fact, introducing the related functional as in Lemma 5.2, the dissipation functional can be expressed via the Lebesgue densities of J_j, b_j. By the a priori bounds (4.2), these densities are in $L^1([0, T] \times \Omega)$ as Lemma 5.1 shows. For notational convenience, we identify the measures with their Lebesgue densities and stick to the above expression (4.3). Only the Lebesgue density of μ is denoted by a different letter (namely c). Only in Lemma 5.7, we, in fact, show compactness for the sequence of measures $J_i^\varepsilon \in \mathcal{M}(\Omega)$, which are the diffusion fluxes that provide the minimum of $\mathfrak{D}_\varepsilon^V(\mu^\varepsilon)$ in (4.3) (see Sect. 5.1).

The main result is the Γ-convergence of $\mathfrak{D}_\varepsilon^V$ to the effective dissipation functional $\mathfrak{D}_0^V$ which is defined by

$$
\mathfrak{D}_0^V(\mu) = \begin{cases} \int_0^T \mathcal{R}_{\text{eff}}(\mu, \dot\mu) + \mathcal{R}_{\text{eff}}^*(\mu, -D\mathcal{E}^V(\mu)) dt, & \text{if } \mu \in AC([0, T], Q), \mu = c \, dx \text{ a.e. in } [0, T] \\ \infty & \text{otherwise.} \end{cases}
$$

where

$$\mathcal{R}_{\text{eff}}^*(\mu, \xi) = \mathcal{R}_{\text{diff}}^*(\mu, \xi) + \chi_{\{\xi_1 = \xi_2\}}(\xi), \quad \mathcal{R}_{\text{eff}}(\mu, v) = (\mathcal{R}_{\text{eff}}(\mu, \cdot))^*(v). \tag{4.4}$$

Theorem 4.1 *Let* $V \in \mathrm{C}^1(\Omega, \mathbb{R}^2)$. *On* $\mathrm{L}_w^\infty([0, T], Q)$, *we have* Γ-*convergence conditioned to bounded energies of* $\mathfrak{D}_\varepsilon^V$, *i.e.* $\mathfrak{D}_\varepsilon^V \xrightarrow{\mathrm{M}_E} \mathfrak{D}_0^V$ *where*

$$\mathfrak{D}_0^V(\mu) = \begin{cases} \int_0^T \mathcal{R}_{\text{eff}}(\mu, \dot\mu) + \mathcal{R}_{\text{eff}}^*(\mu, -V - \mathrm{D}\mathcal{E}(\mu))\mathrm{d}t, & \mu \in \mathrm{AC}([0, T], Q), \mu = c\,\mathrm{d}x \text{ a.e. } [0, T] \\ \infty & \text{otherwise} \end{cases}$$

$$\tag{4.5}$$

with

$$\mathcal{R}_{\text{eff}}^*(\mu, \xi) = \mathcal{R}_{\text{diff}}^*(\mu, \xi) + \chi_{\{\xi_1 = \xi_2\}}(\xi),$$

$$\mathcal{R}_{\text{eff}}(\mu, v) = \inf_{u + \tilde u = v} \{\mathcal{R}_{\text{diff}}(\mu, \tilde u) + \chi_0(u_1 + u_2)\} =$$

$$= \inf \left\{ \sum_{j=1}^2 \int_\Omega \widetilde{\mathrm{Q}}(\delta_j c_j, J_j)\mathrm{d}x : u_1 + u_2 = 0, \begin{cases} v_1 = -\operatorname{div} J_1 + u_1 \\ v_2 = -\operatorname{div} J_2 + u_2 \end{cases} \right\}.$$

The theorem states that the limit dissipation functional is again of $\mathcal{R} \oplus \mathcal{R}^*$-form with an effective dissipation potential $\mathcal{R}_{\text{eff}}^*$. The effective dissipation potential $\mathcal{R}_{\text{eff}}^*$ consists again of two terms describing the diffusion and a coupling, which forces the chemical potential $-\mathrm{D}\mathcal{E}^V$ to equilibration providing the microscopic equilibria of the densities ρ^V and defining the slow manifold of the evolution

$$\rho_1^V = \rho_2^V \quad \Leftrightarrow \quad \frac{c_1}{\beta \mathrm{e}^{-V_1}} = \frac{c_2}{\alpha \mathrm{e}^{-V_2}}.$$

Since the effective dissipation functional $\mathcal{R}_{\text{eff}}$ is independent of the tilts $\eta = V$, Theorem 4.1 immediately implies that the family of gradient systems $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$ EDP-converges with tilting to the effective gradient system $(Q, \mathcal{E}, \mathcal{R}_{\text{eff}}^*)$.

Theorem 4.2 *Let* $\mathcal{R}_{\text{eff}}^*$ *be defined by* (4.4). *Then the gradient system* $(Q, \mathcal{E}, \mathcal{R}_\varepsilon^*)$ *EDP-converges with tilting to* $(Q, \mathcal{E}, \mathcal{R}_{\text{eff}}^*)$.

In the Sect. 5, we will present the detailed proof of the Γ-convergence result. In this section, we discuss the effective gradient system and its induced gradient flow equation.

4.2 Effective gradient flow equation

Similar to the space-independent situation in [30,34] the limit gradient system can be understood in two different but mathematically equivalent ways: as a gradient system on Q, and also in the coarse-grained space of slow variables $\hat Q$. On Q, the effective gradient flow equation is given with Lagrange multipliers ensuring the projection on the slow manifold. On the coarse-grained space $\hat Q$, we obtain a scalar diffusion equation with a new mixed diffusion coefficient. Throughout the section the potential $V \in \mathrm{C}^1(\Omega, \mathbb{R}^2)$ is fixed.

4.2.1 Gradient flow equation with Lagrange multipliers

For being brief, the calculations in this section are rather formal. The effective dissipation potential $\mathcal{R}_{\text{eff}}^* = \mathcal{R}_{\text{diff}}^* + \chi_{\{\xi_1 = \xi_2\}}$ consists of two parts: the first describes the dissipation of

the evolution and the second provides the linear constraint of being on the slow manifold and also the corresponding Lagrange multiplier. The evolution equation is given by

$$\dot{\mu} \in \partial_\xi \mathcal{R}^*_{\mathrm{eff}}(\mu, -V - \mathrm{D}\mathcal{E}(\mu)) = \partial_\xi \left\{ \mathcal{R}^*_{\mathrm{diff}}(\mu, -V - \mathrm{D}\mathcal{E}(\mu)) + \chi_{\{\xi_1=\xi_2\}}(-V - \mathrm{D}\mathcal{E}(\mu)) \right\}.$$

Following [16], the subdifferential of a sum is given by the sum of the subdifferentials, if one term is continuous. For the second term, the subdifferential of the characteristic function is only defined in its domain, i.e. if

$$-V_1 - \mathrm{D}\mathcal{E}(\mu)_1 = -V_2 - \mathrm{D}\mathcal{E}(\mu)_2 \quad \Leftrightarrow \quad \frac{c_1}{\beta \mathrm{e}^{-V_1}} = \frac{c_2}{\alpha \mathrm{e}^{-V_2}},$$

hence, defining the linear slow manifold. Moreover, on its domain we have for the sub-differential that $\partial \chi_{\{\xi_1=\xi_2\}} = \mathcal{M}(\Omega) \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Hence, we conclude that

$$\dot{\mu} \in \partial_\xi \left\{ \mathcal{R}^*_{\mathrm{diff}}(\mu, -V - \mathrm{D}\mathcal{E}(\mu)) \right\} + \mathcal{M}(\Omega) \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \frac{c_1}{\beta \mathrm{e}^{-V_1}} = \frac{c_2}{\alpha \mathrm{e}^{-V_2}},$$

which implies the gradient flow equation on the slow manifold with a Lagrange multiplier $\lambda = (\lambda_1, \lambda_2)$ for the densities of the form

$$\begin{cases} \dot{c}_1 = \mathrm{div}\{\delta_1 \nabla c_1 + \delta_1 c_1 \nabla V_1\} + \lambda_1 \\ \dot{c}_2 = \mathrm{div}\{\delta_2 \nabla c_2 + \delta_2 c_2 \nabla V_2\} + \lambda_2 \end{cases}, \quad \lambda_1 + \lambda_2 = 0, \quad \frac{c_1}{\beta \mathrm{e}^{-V_1}} = \frac{c_2}{\alpha \mathrm{e}^{-V_2}}. \qquad (4.6)$$

4.2.2 Coarse-grained gradient structure and its gradient flow equation

We introduce the coarse grained probability measure $\hat{\mu} = \mu_1 + \mu_2$ on Ω and the corresponding concentrations $\hat{c} := c_1 + c_2$. Moreover, we define the equilibrated densities $\hat{\rho}^V = \rho_1^V = \rho_2^V$ and the coarse-grained stationary measure $\hat{w}^V = w_1^V + w_2^V$, for which we get $\hat{c} = \rho_1^V w_1^V + \rho_2^V w_2^V = \hat{\rho}^V (w_1^V + w_2^V) = \hat{\rho}^V \hat{w}^V$. We also introduce the coarse-grained diffusion coefficient $\hat{\delta}^V = \frac{\delta_1 w_1^V + \delta_2 w_2^V}{w_1^V + w_2^V}$.

With this notation, we may define the coarse-grained gradient structure $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$. On the state space $\hat{Q} = \mathrm{Prob}(\Omega)$, we define

$$\hat{\mathcal{R}}^*(\hat{\mu}, \hat{\xi}) := \mathcal{R}^*_{\mathrm{eff}}\left(\left(\frac{w_1^V}{w_1^V + w_2^V} \hat{\mu}, \frac{w_2^V}{w_1^V + w_2^V} \hat{\mu} \right), (\hat{\xi}, \hat{\xi}) \right) = \frac{1}{2} \int_\Omega \hat{\delta}^V |\nabla \hat{\xi}|^2 \mathrm{d}\hat{\mu},$$

$$\hat{\mathcal{E}}(\hat{\mu}) := \mathcal{E}^V\left(\frac{w_1^V}{w_1^V + w_2^V} \hat{\mu}, \frac{w_2^V}{w_1^V + w_2^V} \hat{\mu} \right). \qquad (4.7)$$

Introducing the coarse-grained potential $\hat{V} = -\log(w_1^V + w_2^V) - \log Z = -\log(\hat{w}^V) - \log Z = -\log\left(w_1 \mathrm{e}^{-V_1} + w_2 \mathrm{e}^{-V_2} \right)$, for which the exponential is given by the weighted arithmetic mean of the exponentials e^{-V_1} and e^{-V_2}, i.e. $\mathrm{e}^{-\hat{V}} = w_1 \mathrm{e}^{-V_1} + w_2 \mathrm{e}^{-V_2}$ (we used $w_1 + w_2 = 1$). Easy calculations show that the energy has the explicit form

$$\hat{\mathcal{E}}(\hat{\mu}) = \int_\Omega \left(\log \hat{\mu} + \hat{V} \right) \mathrm{d}\hat{\mu}.$$

The coarse-grained dissipation functional is defined by

$$\hat{\mathfrak{D}}(\hat{\mu}) = \int_0^T \left\{ \hat{\mathcal{R}}(\hat{\mu}, \dot{\hat{\mu}}) + \hat{\mathcal{R}}^*(\hat{\mu}, -\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})) \right\} \mathrm{d}t,$$

which incorporates the tilt via the coarse-grained variables. Note, that the coarse-grained dissipation potential $\hat{\mathcal{R}}^*$ depends explicitly on the tilt V via the diffusion coefficient $\hat{\delta}^V$. This is not a contradiction to tilt-EDP convergence (Theorem 4.2), because in original variables the effective dissipation potential (4.5) is indeed independent of the tilts. The tilts dependence of $\hat{\mathcal{R}}^*$ originates from the energy and tilt dependent slow manifold.

To relate the dissipation functional $\mathfrak{D}_0^V$ with the coarse grained dissipation functional $\hat{\mathfrak{D}}$, we first show that also an equilibration of the fluxes occurs. To do this the following convexity property is important.

Lemma 4.3 *Let X a separable and reflexive Banach space and $F : X \to \mathbb{R}_\infty$ be convex and lsc. Let the function $\tilde{F} : [0, \infty[\times X \to \mathbb{R}_\infty$ be defined as in (3.4). Then, we have*

$$\tilde{F}\left(\sum_{i=1}^I a_i, \sum_{i=1}^I x_i\right) \le \sum_{i=1}^I \tilde{F}(a_i, x_i).$$

If F is strictly convex then equality holds if and only if $(a_i, x_i) = (0, 0)$ whenever $a_i = 0$ or $x_i/a_i = x_j/a_j$ whenever $a_i, a_j > 0$. Moreover, if $F(0) = 0$, we have the following monotonicity property

$$\tilde{F}(a_1, x) \le \tilde{F}(a_2, x), \quad \text{if } a_1 \ge a_2.$$

Proof Let pairs (a_i, x_i) for $i = 1, \ldots, I$ be given. If $a_i = 0$, then either $x_i = 0$ and the claim has to be shown for $I - 1$ -number of pairs, or $x_i \ne 0$ and the right-hand side is infinite meaning that the claim is trivial. So let us assume that $a_i > 0$ for all $i = 1 \ldots, I$. Then $\tilde{F}(a_i, x_i) = a_i F(x_i/a_i)$ and the claim is equivalent to

$$\sum_{i=1}^I \frac{a_i}{\sum_{i=1}^I a_i} F(x_i/a_i) \ge F\left(\sum_{i=1}^I \frac{a_i}{\sum_{i=1}^I a_i} \frac{x_i}{a_i}\right) = F\left(\frac{\sum_{i=1}^I x_i}{\sum_{i=1}^I a_i}\right),$$

which holds since F is convex. If F is strictly convex then we immediately observe that whenever $a_i, a_j > 0$ we have $\frac{x_i}{a_i} = \frac{x_j}{a_j}$, and whenever $a_i = 0$ that also $x_i = 0$.

To see the monotonicity property, we observe that $\tilde{F}(a, 0) = 0$ for all $a \ge 0$. Hence, we have

$$\tilde{F}(a_1 + a_2, x) \le \tilde{F}(a_1, x) + \tilde{F}(a_2, 0) = \tilde{F}(a_1, x),$$

which proves the claim. $\square$

Recalling formula (4.5) of the effective dissipation functional $\mathfrak{D}_0^V$ and using the above lemma, we observe that the velocity part of the dissipation functional $\mathfrak{D}_0^V$ can now be estimated. In particular, we will see that the limit dissipation functional $\mathfrak{D}_0^V$ can be equivalently expressed in coarse-grained variables $(\hat{\mu}, \hat{J})$ by using that an equilibration of concentrations also provides an equilibration of the corresponding fluxes. In the reconstruction strategy in Sect. 5.3 this equilibration is explicitly used (see (4.11) and (5.3)).

Proposition 4.4 *Let $\mu \in \mathrm{AC}([0, T], Q)$ with $\mathfrak{D}_0^V(\mu) < \infty$ and ess $\sup_{t \in [0,T]} \mathcal{E}(\mu(t)) < \infty$. Then the following holds:*

1. We have $\mathfrak{D}_0^V(\mu) = \hat{\mathfrak{D}}(\hat{\mu})$ where $\hat{\mu} = \mu_1 + \mu_2$ and

$$\hat{\mathfrak{D}}(\hat{\mu}) = \int_0^T \hat{\mathcal{R}}(\hat{\mu}, \dot{\hat{\mu}}) + \hat{\mathcal{R}}^*(\hat{\mu}, -\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})) \, \mathrm{d}t$$

$$= \inf_{\hat{J}:\dot{\hat{\mu}}+\mathrm{div}\hat{J}=0} \left\{ \int_0^T \int_\Omega \widetilde{Q}(\hat{\delta}^V \hat{c}, \hat{J}) + \frac{\hat{\delta}^V \hat{w}^V}{2} \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V} \, \mathrm{d}x \mathrm{d}t \right\},$$

where the infimum is taken over all Borel fluxes $\hat{J} \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$ which satisfy the coarse-grained continuity equation in the sense of distributions, i.e.

$$\forall \phi \in \mathrm{C}_c^\infty(]0, T[\times\Omega) \; : \; \int_0^T \int_\Omega \dot{\phi}\hat{c} \, \mathrm{d}x \, \mathrm{d}t = \int_0^T \int_\Omega \nabla\phi \cdot \hat{J} \, \mathrm{d}x \, \mathrm{d}t, \quad \hat{J} \cdot \nu = 0 \text{ on } \partial\Omega.$$

2. The chain-rule holds for $[0, T] \ni t \mapsto \hat{\mathcal{E}}(\hat{\mu}(t)) \in \mathbb{R}$.
3. The gradient flow equation of the gradient system $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$ is given by

$$\dot{\hat{c}} = -\mathrm{div}\left(\hat{\delta}^V \hat{c} \, \nabla\left(-\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})\right)\right) = \mathrm{div}\left(\hat{\delta}^V \nabla\hat{c} + \hat{\delta}^V \hat{c}\nabla\hat{V}\right), \tag{4.8}$$

with the potential $\hat{V} = -\log\hat{w}^V$ and stationary measure $\hat{w}^V$.

Equation (4.8) shows that the coarse-grained gradient flow equation induced by $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$ is a drift-diffusion equation of the coarse-grained concentration $\hat{c}$ with mixed diffusion constant $\hat{\delta}^V$. In particular, in the tilt free case we have $\hat{\delta}^{V=\mathrm{const}} = \frac{\beta\delta_1+\alpha\delta_2}{\alpha+\beta}$, and we recover the result of [6].

Proof To prove Part 1, we first observe that the bounded energy and dissipation for the trajectory μ implies that we have $\frac{c_1}{w_1^V} = \frac{c_2}{w_2^V}$ a.e. in $[0, T] \times \Omega$, see Lemma 5.10. Using $\hat{c} = c_1 + c_2$ for the densities, we get

$$\hat{c} = \frac{w_1^V + w_2^V}{w_1^V}c_1 = \frac{w_1^V + w_2^V}{w_2^V}c_2. \tag{4.9}$$

The Fisher information $\mathcal{S}_0^V(\mu) := \mathcal{R}_{\mathrm{eff}}^*(\mu, -V - \mathrm{D}\mathcal{E}(\mu))$ has the form

$$\mathcal{S}_0^V(\mu) = \frac{1}{2}\int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{\left|\nabla\rho_j^V\right|^2}{\rho_j^V}\mathrm{d}x = \frac{1}{2}\int_\Omega \left(\frac{\delta_1 w_1^V + \delta_2 w_2^V}{w_1^V + w_2^V}\right)\left(w_1^V + w_2^V\right)\frac{|\nabla\hat{\rho}^V|^2}{\hat{\rho}^V}\mathrm{d}x$$

$$= \frac{1}{2}\int_\Omega \hat{\delta}^V \hat{w}^V \frac{|\nabla\hat{\rho}^V|^2}{\hat{\rho}^V}\mathrm{d}x = \hat{\mathcal{R}}^*(\hat{\mu}, -\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})).$$

Now, take diffusion fluxes $J_1, J_2 \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$ which provide the infimum in the velocity part of the dissipation functional $\mathfrak{D}_0^V(\mu)$. Lemma 5.1 implies that J_j have Lebesgue densities (which are also denoted by J_j) that are in $\mathrm{L}^1([0, T] \times \Omega, \mathbb{R}^d)$.

Lemma 4.3 provides that also an equilibration of the fluxes occurs. Indeed, defining the coarse-grained flux $\hat{J} = J_1 + J_2$, we conclude the pointwise estimate

$$\frac{|J_1|^2}{\delta_1 c_1} + \frac{|J_2|^2}{\delta_2 c_2} \geq \frac{|J_1 + J_2|^2}{\delta_1 c_1 + \delta_2 c_2} = \frac{|\hat{J}|^2}{\frac{\delta_1 w_1^V + \delta_2 w_2^V}{w_1^V + w_2^V}\hat{c}} = \frac{|\hat{J}|^2}{\hat{\delta}^V \hat{c}}, \tag{4.10}$$

where $\hat{\delta}^V := \frac{\delta_1 w_1^V + \delta_2 w_2^V}{w_1^V + w_2^V}$. Equality holds if and only if $(J_1, c_1) = 0$ or $(J_2, c_2) = 0$ or $J_1/\delta_1 c_1 = J_2/\delta_2 c_2 = \hat{J}/\hat{\delta}^V \hat{c}$. The last condition is equivalent to

$$\hat{J} = \frac{\delta_1 w_1^V + \delta_2 w_2^V}{\delta_1 w_1^V} J_1 = \frac{\delta_1 w_1^V + \delta_2 w_2^V}{\delta_2 w_2^V} J_2 , \tag{4.11}$$

which provides an explicit formula for the coarse-grained diffusion flux.

Since $(\hat{c}, \hat{J})$ solves the coarse-grained continuity equation, we obtain for the dissipation functional that

$$\mathfrak{D}_0^V(\mu) = \inf_{(c,J,b)\in(\text{gCE})} \int_0^T \left\{ \int_\Omega \sum_{j=1}^2 \widetilde{\mathsf{Q}}(\delta_j c_j, J_j)\mathrm{d}x + \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{|\nabla \rho_j|^2}{\rho_j} \mathrm{d}x \right\} \mathrm{d}t$$

$$\geq \inf_{\hat{c}+\mathrm{div}\hat{J}=0} \int_0^T \left\{ \int_\Omega \widetilde{\mathsf{Q}}\left(\hat{\delta}^V \hat{c}, \hat{J}\right)\mathrm{d}x + \frac{1}{2} \int_\Omega \hat{\delta}^V \hat{w}^V \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V}\mathrm{d}x \right\} \mathrm{d}t. \tag{4.12}$$

To prove equality, let $\hat{\mu}$ such that $\hat{\mathfrak{D}}(\hat{\mu}) < \infty$ be given. We define the reconstructed concentrations c_1, c_2 by (4.9). Moreover, let $\hat{J}$ be a diffusion flux satisfying the continuity equation, which can be again assumed to be in $\mathrm{L}^1([0, T] \times \Omega, \mathbb{R}^d)$. We define $J_1. J_2 \in \mathrm{L}^1([0, T] \times \Omega, \mathbb{R}^d)$ by formula (4.11) and they satisfy the same boundary conditions. Using the explicitly derived reaction flux b_1, b_2 from (5.3) defined in the sense of distributions, we observe that (c, J, b) satisfies the generalized continuity equation. Hence, taking the infimum, we conclude

$$\hat{\mathfrak{D}}(\hat{\mu}) = \int_0^T \left\{ \int_\Omega \sum_{j=1}^2 \widetilde{\mathsf{Q}}(\delta_j c_j, J_j)\mathrm{d}x + \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{|\nabla \rho_j|^2}{\rho_j}\mathrm{d}x \right\} \mathrm{d}t \geq \mathfrak{D}_0^V(\mu),$$

which proves the first part.

For the chain-rule in Part 2, we refer to the proof [18, Lem 4.8] since we consider the pure diffusive situation. The proof uses a time-regularization argument and convexity of the Fisher-information following the ideas of [35, Prop. 2.4].

For Part 3, we compute the evolution equation that is induced by the gradient system is $(\hat{Q}, \hat{\mathcal{E}}, \hat{\mathcal{R}}^*)$. We have

$$\partial_\xi \hat{\mathcal{R}}^*(\hat{\mu}, \hat{\xi}) = -\mathrm{div}\left(\hat{\delta}^V \hat{c} \nabla \hat{\xi}\right),$$

$$\mathrm{D}\hat{\mathcal{E}}(\hat{\mu}) = \log \hat{\mu} + 1 - \log \hat{w}^V - \log Z,$$

$$\nabla\left(-\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})\right) = -\frac{\nabla \hat{\mu}}{\hat{\mu}} + \frac{\nabla \hat{w}^V}{\hat{w}^V} = -\frac{\nabla \hat{\mu}}{\hat{\mu}} - \nabla \hat{V},$$

which results in

$$\dot{\hat{c}} = -\mathrm{div}\left(\hat{\delta}^V \hat{c} \nabla\left(-\mathrm{D}\hat{\mathcal{E}}(\hat{\mu})\right)\right) = \mathrm{div}\left(\hat{\delta}^V \nabla \hat{c} + \hat{\delta}^V \hat{c} \nabla \hat{V}\right).$$

$\square$

We finally remark that the coarse-grained gradient flow Eq. (4.8) is equivalent to the gradient flow equation with Lagrange multipliers (4.6). Indeed, adding both equations in (4.6) and using that the original concentrations can be expressed by the coarse-grained concentrations via (4.9), the coarse-grained gradient flow Eq. (4.8) with the drift term $\hat{\delta}^V \hat{c} \nabla \hat{V}$ can be readily derived. Conversely, using (4.9), we see that $c = (c_1, c_2)$ are on the slow manifold

and satisfy (4.6). The corresponding Lagrange multipliers $\lambda = (\lambda_1, \lambda_2)$ can be explicitly calculated. Introducing the difference of the diffusion constants $\bar{\delta} = \delta_1 - \delta_2$ and the potentials $\overline{V} = V_1 - V_2$, we have

$$\lambda_1 = \frac{w_2 e^{-V_2}}{w_1 e^{-V_1} + w_2 e^{-V_2}} \left(-\bar{\delta}\Delta c_1 + \left(\delta_2 \nabla \overline{V} - \bar{\delta}\nabla V_1 \right) \cdot \nabla c_1 + c_1 \left\{ \delta_2 \nabla \overline{V} \, \nabla V_1 - \bar{\delta}\Delta V_1 \right\} \right)$$

$$\lambda_2 = \frac{w_1 e^{-V_1}}{w_1 e^{-V_1} + w_2 e^{-V_2}} \left(\bar{\delta}\Delta c_2 + \left(-\delta_1 \nabla \overline{V} + \bar{\delta}\nabla V_2 \right) \cdot \nabla c_2 + c_2 \left\{ -\delta_1 \nabla \overline{V} \, \nabla V_2 + \bar{\delta}\Delta V_2 \right\} \right).$$

We observe that the Lagrange multiplier λ_i has the same regularity as the right-hand side of the evolution equation of c_i. Moreover, both evolution equations are completely uncoupled but contain a linear annihilation/creation term, which depends on the potential $V = (V_1, V_2)$ and the diffusion coefficient $\delta = (\delta_1, \delta_2)$. A lengthy calculation shows that indeed we have $\lambda_1 + \lambda_2 = 0$.

5 Proof of Γ-convergence

In this section we prove the Γ-convergence result of Theorem 4.1. As usual, we prove Γ-convergence in three steps: First deriving compactness, secondly establishing the liminf-estimate by exploiting the compactness, thirdly constructing the recovery sequence for the limsup-estimate.

In the following the next lemma will be useful.

Lemma 5.1 *Let* $F : \mathbb{R}^k \rightarrow [0, \infty[$ *be a convex, lsc. function of superlinear growth, i.e.* $F(r)/|r| \rightarrow \infty$ *as* $|r| \rightarrow \infty$. *Then there is a constant* $k_F > 0$ *such that for any measurable functions* $W : \Omega \rightarrow \mathbb{R}^k$ *and* $\rho : \Omega \rightarrow \mathbb{R}_{\geq 0}$ *it holds*

$$\int_\Omega |W| \, dx \leq \int_\Omega \widetilde{F}(\rho, W) \, dx + k_F \int_\Omega \rho \, dx \,.$$

Proof Let W and ρ be given. We define three measurable subsets of Ω:

$$\Omega_0 = \{x : \rho(x) = 0\}, \ \Omega_1 = \{x : \rho \neq 0, \tfrac{1}{\rho}|W| \leq F(\tfrac{1}{\rho} W)\}, \ \Omega_2 = \{x : \rho \neq 0, \tfrac{1}{\rho}|W| > F(\tfrac{1}{\rho} W)\}.$$

Since F is superlinear, there is a constant $k_F > 0$ such that on Ω_2 it holds $W/\rho \leq k_F$. Hence we can estimate

$$\int_\Omega |W| dx \leq \int_{\Omega_0} |W| dx + \int_{\Omega_1} \frac{|W|}{\rho} \rho dx + \int_{\Omega_2} \frac{|W|}{\rho} \rho dx$$

$$\leq \int_{\Omega_0} \widetilde{F}(\rho, W) dx + \int_{\Omega_1} F(\tfrac{1}{\rho} W) \rho dx + k_F \int_{\Omega_2} \rho dx \leq \int_\Omega \widetilde{F}(\rho, W) dx + k_F \int_\Omega \rho dx \,.$$

$\square$

Moreover, we need the following classical lemma. It guarantees the necessary regularity for the limits, and moreover, it provides the desired liminf-estimate.

Lemma 5.2 (Lemma 9.4.3, [1]) *Let* $F : [0, \infty[\rightarrow [0, \infty]$ *be a proper, lsc, convex function with superlinear growth. We define the related functional*

$$\mathcal{F}(\mu, \gamma) = \begin{cases} \int_A F(\frac{d\mu}{d\gamma}) d\gamma, & \text{if } \mu \ll \gamma, \\ \infty, & \text{otherwise.} \end{cases}$$

Let $\mu^\varepsilon, \gamma^\varepsilon \in \mathrm{Prob}(A)$ be two sequences with $\mu^\varepsilon \overset{*}{\rightharpoonup} \mu^0$ and $\gamma^\varepsilon \overset{*}{\rightharpoonup} \gamma^0$. Then

$$\liminf_{\varepsilon \to 0} \mathcal{F}(\mu^\varepsilon, \gamma^\varepsilon) \geq \mathcal{F}(\mu^0, \gamma^0).$$

In particular, if the left-hand side is finite, then for the limits it holds $\mu^0 \ll \gamma^0$.

5.1 Compactness

Recall that for a given potential $V \in \mathrm{C}^1(\Omega, \mathbb{R}^2)$ the dissipation functional $\mathfrak{D}_\varepsilon^V$ is defined on the space of trajectories equipped with the weak topology, see (4.1). In the following we want to derive compactness for a sequence $(\mu^\varepsilon)_{\varepsilon > 0}$ of trajectories, satisfying the a priori bounds (4.2). Using the bound of the dissipation functional $\mathfrak{D}_\varepsilon^V(\mu^\varepsilon) \leq C$, we conclude that there exist diffusive fluxes $J^\varepsilon = (J_1^\varepsilon, J_2^\varepsilon)$ and reaction fluxes $b^\varepsilon = (b_1^\varepsilon, b_2^\varepsilon)$ such that $J_i^\varepsilon \in \mathcal{M}(\Omega, \mathbb{R}^d)$, $b_i^\varepsilon \in \mathcal{M}(\Omega, \mathbb{R})$ and $(\mu^\varepsilon, J^\varepsilon, b^\varepsilon)$ satisfies the continuity equation

$$(c, J, b) \in \text{(gCE)} \quad \Leftrightarrow \quad \left\{ b_1 + b_2 = 0 \text{ and } \begin{cases} \dot{c}_1 = -\mathrm{div}\, J_1 + b_1 \\ \dot{c}_2 = -\mathrm{div}\, J_2 + b_2 \end{cases} \right\}.$$

Moreover, we get bounds:

$$i = 1, 2: \quad \int_0^T \int_\Omega \frac{|\nabla \rho_i^{V,\varepsilon}|^2}{\rho_i^{V,\varepsilon}} \mathrm{d}x \mathrm{d}t \leq C, \quad \int_0^T \int_\Omega \widetilde{\mathsf{Q}}\left(c_i^\varepsilon, J_i^\varepsilon\right) \mathrm{d}x \mathrm{d}t \leq C,$$

$$\int_0^T \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1^\varepsilon c_2^\varepsilon}}{\varepsilon}, b_2^\varepsilon(x)\right) \mathrm{d}x \mathrm{d}t \leq C, \quad \frac{1}{\varepsilon} \int_0^T \int_\Omega \left(\sqrt{\rho_1^{V,\varepsilon}} - \sqrt{\rho_2^{V,\varepsilon}}\right)^2 \mathrm{d}x \mathrm{d}t \leq C.$$

Remark 5.3 Following [26] a distributional solution (μ, J, B) of the generalized continuity equation $\dot{\mu} = -\mathrm{div}\, J + B$ satisfying $\int_0^T \int_\Omega |B| + |J| \mathrm{d}x \mathrm{d}t < \infty$ can be assumed to be absolutely continuous in time. The bounds can be obtained easily using Lemma 5.1 for fixed $\varepsilon > 0$.

Because the functional is convex in the concentration c and in the fluxes J and b, weak convergence would be sufficient to prove a liminf-estimate using a Ioffe-type argument. But, as in the PDE-result [6], we aim in proving even strong convergence for the densities $c^\varepsilon \to c^0$ in $\mathrm{L}^1([0, T] \times \Omega, \mathbb{R}_{\geq 0}^2)$. This is done in two steps: First, compactness of coarse-grained variables, and secondly, convergence towards the slow manifold is shown, which together implies strong compactness. This strategy has successfully been applied already in the space-independent case in [30,34]. Moreover, we show that the limit trajectory $\mu^0 = c^0 \mathrm{d}x$ has a representative which is in $\mathrm{AC}([0, T], Q)$. Note that it is not possible to prove pointwise convergence $\mu^\varepsilon(t) \overset{*}{\rightharpoonup} \mu^0(t)$ for all $t \in [0, T]$. Instead, pointwise convergence is only shown for the coarse-grained variables $\hat{\mu}^\varepsilon := \mu_1^\varepsilon + \mu_2^\varepsilon$.

First, we derive weak compactness in space-time, which immediately follows from the uniform bound in ε and time on the energy.

Lemma 5.4 ("Very" weak compactness in space-time) Let $(\mu^\varepsilon)_{\varepsilon > 0}$, $\mu^\varepsilon \in \mathrm{L}_w^\infty([0, T], Q)$ satisfy $\sup_{\varepsilon > 0} \operatorname*{ess\,sup}_{t \in [0,T]} \mathcal{E}(\mu^\varepsilon(t)) \leq C$. Then for a.e. $t \in [0, T]$ the measure $\mu^\varepsilon(t)$ has a Lebesgue density $c^\varepsilon(t, \cdot)$. Moreover, there is a subsequence (not relabeled), such that their densities c^ε are uniformly integrable in $\Omega \times [0, T] \times \{1, 2\}$ and hence, c_i^ε converges weakly in $\mathrm{L}^1([0, T] \times \Omega)$ to c_i^0 for $i = 1, 2$.

Proof The bound on the energy implies that a.e. $t \in [0, T]$ the measure $\mu^\varepsilon(t, \cdot)$ has a Lebesgue density $c^\varepsilon(t, \cdot)$. Moreover, the functional $\mu \mapsto \int_0^T \mathcal{E}(\mu)\mathrm{d}t$ is superlinear and convex. Hence, it follows by the Theorem of de Vallée-Poussin that the densities c^ε are uniformly integrable and hence, c_i^ε converges weakly in $\mathrm{L}^1([0, T] \times \Omega)$ to c_i^0. The corresponding limit trajectory of measures is denoted by $\mu^0(t) = (\mu_1^0(t), \mu_2^0(t)) \in Q$. □

In the following, we are going to derive compactness for the concentrations c_i^ε and the diffusive fluxes J_i^ε. It is not possible to get compactness for the fast reaction flux b_2^ε by bounding the dissipation functional, which can be seen in the next remark.

Remark 5.5 To see that compactness for the fast reaction flux b_2^ε is not possible to obtain by the a priori bounds 4.2, we set $\rho^\varepsilon = 1$ constant in $[0, T] \times \Omega \times \{1, 2\}$ and $b_2^\varepsilon = b^\varepsilon$ constant in $[0, T] \times \Omega$. Then, a bound on the dissipation functional implies a bound

$$\infty > \int_0^T \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1^\varepsilon c_2^\varepsilon}}{\varepsilon}, b_2^\varepsilon(x)\right) \mathrm{d}x\mathrm{d}t \approx \widetilde{\mathsf{C}}(\frac{1}{\varepsilon}, b^\varepsilon) = \frac{1}{\varepsilon}\mathsf{C}(\varepsilon b^\varepsilon) \approx |b^\varepsilon| \log(\varepsilon|b^\varepsilon| + 1).$$

Setting $b^\varepsilon = -\log \varepsilon$, we easily see that $|b^\varepsilon| \log(\varepsilon|b^\varepsilon|+1) \to 0$ as $\varepsilon \to 0$, however, $b^\varepsilon \to \infty$. Hence, it is not possible to obtain compactness for the fast reaction flux b_i^ε. Later in Lemma 5.15 the "converse" statement is proved: If $\iint \mathsf{C}(b)\mathrm{d}x\mathrm{d}t < \infty$ then $\iint \widetilde{\mathsf{C}}(\frac{1}{\varepsilon}, b)\mathrm{d}x\mathrm{d}t \to 0$.

Next, we are going to derive time-regularity for the sequence $(\mu^\varepsilon)_{\varepsilon>0}$ in proving compactness for the coarse-grained trajectories $\hat{\mu}^\varepsilon = \mu_1^\varepsilon + \mu_2^\varepsilon$. In particular, we are able to prove pointwise convergence in time.

Lemma 5.6 (Time Regularity of μ^ε) Let $(\mu^\varepsilon)_{\varepsilon>0}$, $\mu^\varepsilon \in \mathrm{L}_w^\infty([0, T], Q)$ satisfying the a priori bounds (4.2). Then the curves $t \mapsto \hat{\mu}^\varepsilon(t) := \mu_1^\varepsilon(t) + \mu_2^\varepsilon(t)$ have ε-uniformly bounded total variation in the space $\mathrm{Prob}(\Omega)$ equipped with the 1-Wasserstein distance, i.e.

$$\|\hat{\mu}^\varepsilon\|_{TV} := \sup \left\{ \sum_{k=1}^K \mathcal{W}_1(\hat{\mu}^\varepsilon(t_k), \hat{\mu}^\varepsilon(t_{k-1})) \ : \ 0 = t_0 < \cdots < t_k < \cdots < t_K = T \right\}.$$

In particular, by Helly's selection principle, we conclude pointwise convergence $\hat{\mu}^\varepsilon(t) := \mu_1^\varepsilon(t) + \mu_2^\varepsilon(t) \xrightarrow{*} \hat{\mu}^0(t) := \mu_1^0(t) + \mu_2^0(t)$ for all $t \in [0, T]$ in $\mathrm{Prob}(\Omega)$ along a suitable subsequence, where $\mu^0 = (\mu_1^0, \mu_2^0)$ is from Lemma 5.4.

Proof Let a partition $\{t_k\}_{k=0,\dots,K}$ of the interval $[0, T]$ be given. We exploit the dual formulation of the 1-Wasserstein distance, which is given by integrating against Lipschitz functions (see e.g. [15]),

$$\mathcal{W}_1(\hat{\mu}, \hat{v}) = \sup_{\phi: \|\phi\|_{\mathrm{W}^{1,\infty}(\Omega)} \le 1} \int_\Omega \phi(\mathrm{d}\hat{\mu} - \mathrm{d}\hat{v}).$$

Taking $\phi \in \mathrm{C}^1(\Omega)$ with $\|\phi\|_{\mathrm{W}^{1,\infty}(\Omega)} \le 1$ and using the continuity equations $b_1 + b_2 = 0$ and $\begin{Bmatrix} \dot{c}_1 = -\mathrm{div}\, J_1 + b_1 \\ \dot{c}_2 = -\mathrm{div}\, J_2 + b_2 \end{Bmatrix}$, we conclude for all $[t_{k-1}, t_k] \subset [0, T]$ that

$$\mathcal{W}_1(\hat{\mu}^\varepsilon(t_k), \hat{\mu}^\varepsilon(t_{k-1})) = \int_\Omega \phi \cdot (\mathrm{d}\hat{\mu}^\varepsilon(t_k) - \mathrm{d}\hat{\mu}^\varepsilon(t_{k-1})) = \int_{t_{k-1}}^{t_k} \langle \phi, \dot{\hat{\mu}}^\varepsilon \rangle \mathrm{d}t =$$

$$= \int_{t_{k-1}}^{t_k} \int_\Omega \nabla\phi \cdot (J_1^\varepsilon + J_2^\varepsilon) \,\mathrm{d}x\mathrm{d}t \le \int_{t_{k-1}}^{t_k} \int_\Omega \sum_{i=1}^2 |J_i^\varepsilon| \mathrm{d}x\mathrm{d}t.$$

Hence, we get that

$$\sum_{k=1}^{K} \mathcal{W}_1(\hat{\mu}^{\varepsilon}(t_k), \hat{\mu}^{\varepsilon}(t_{k-1})) \leq \int_0^T \int_{\Omega} \sum_{i=1}^{2} |J_i^{\varepsilon}| \mathrm{d}x \mathrm{d}t.$$

By the bound on the dissipation functional, we obtain ε-uniform bounds on the term $\int_0^T \int_{\Omega} \mathsf{Q}(\delta_j c_j^{\varepsilon}, J_j^{\varepsilon}) \mathrm{d}x \mathrm{d}t$. Moreover, we have $\int_{\Omega} c_j^{\varepsilon} \mathrm{d}x \leq 1$ for all $t \in [0, T]$. Hence by Lemma 5.1 we conclude that $\|\hat{\mu}^{\varepsilon}\|_{TV}$ is ε-uniformly bounded. □

Next, the compactness result from Lemma 5.2 is used in order to prove compactness for the fluxes and spatial regularity.

Lemma 5.7 (Regularity for the fluxes and spatial regularity) *Let* $(\mu^{\varepsilon})_{\varepsilon>0}$ *with* $\mu^{\varepsilon} \in \mathrm{L}_w^{\infty}([0, T], Q)$ *satisfying the a priori bounds* (4.2). *Then the corresponding diffusive fluxes* $J^{\varepsilon} : [0, T] \times \Omega \to \mathbb{R}^2$ *converge weakly-star* $J^{\varepsilon} \overset{*}{\rightharpoonup} J^0$ *in* $\mathcal{M}([0, T] \times \Omega \times \{1, 2\})$ *and* $J_j^0 \ll \mu_j^0$. *Moreover* $\nabla \rho^{V,\varepsilon} \overset{*}{\rightharpoonup} \nabla \rho^{V,0}$ *in* $\mathcal{M}([0, T] \times \Omega \times \{1, 2\})$ *and* $\nabla \rho_j^0 \ll \mu_j^0$. *In particular, we conclude that* ρ_j^{ε} *is uniformly bounded in* $\mathrm{L}^1([0, T], \mathrm{W}^{1,1}(\Omega))$, *which also implies that* $\hat{c}^{\varepsilon}$ *is uniformly in* $\mathrm{L}^1([0, T], \mathrm{W}^{1,1}(\Omega))$.

Proof By the bound on the dissipation functional, we get (after extracting a suitable subsequence of $\varepsilon \to 0$) that $J^{\varepsilon} \overset{*}{\rightharpoonup} J^0$. Moreover, we have $\int_0^T \int_{\Omega} \frac{|J_j^{\varepsilon}|^2}{\rho_j^{\varepsilon}} \mathrm{d}x \leq C$ and $\rho_j^{\varepsilon} \overset{*}{\rightharpoonup} \rho_j^0$. Hence applying the Lemma 5.2, we conclude that $J_j^0 \ll \mu_j^0$. Similarly, we conclude compactness for the gradients $\nabla \rho^{V,\varepsilon}$. The only thing that remains is to identify the limit. But this is clear by definition of the weak derivatives, i.e. integrating against smooth test functions, because this is captured in the weak*-convergence. Lemma 5.1 implies that ρ_j^{ε} is uniformly bounded in $\mathrm{L}^1([0, T], \mathrm{W}^{1,1}(\Omega))$. □

The spatial regularity and the temporal regularity provides a compactness result by a BV-generalization of the Aubin-Lions-Simon Lemma.

Theorem 5.8 ([3,21]) *Let* X, Y, Z *be Banach spaces such that* X *is compactly embedded in* Y, *and* Y *is continuously in* Z^*. *Let* u^{ε} *be a bounded sequence in* $\mathrm{L}^1([0, T], X)$ *and in* $BV([0, T], Z^*)$. *Then (up to a subsequence)* u^{ε} *strongly converges in* $\mathrm{L}^1([0, T], Y)$.

In our situation we immediately conclude that $\hat{c}^{\varepsilon}$ converges strongly.

Corollary 5.9 (Strong convergence of coarse-grained variables) *Let* $(\mu^{\varepsilon})_{\varepsilon>0}$ *with* $\mu^{\varepsilon} \in \mathrm{L}_w^{\infty}([0, T], Q)$ *satisfying the a priori bounds* (4.2). *Then the coarse-grained densities* $\hat{c}^{\varepsilon}$ *converge strongly in* $\mathrm{L}^1([0, T] \times \Omega)$.

Proof Lemma 5.6 provides that $\hat{c}^{\varepsilon}$ is bounded in $BV([0, T], \mathrm{W}^{1,\infty}(\Omega)^*)$ and Lemma 5.7 provides that $\hat{c}^{\varepsilon}$ is bounded in $\mathrm{L}^1([0, T], \mathrm{W}^{1,1}(\Omega))$. Since the embedding $\mathrm{W}^{1,1}(\Omega) \subset \mathrm{L}^1(\Omega)$ is compact and the embedding $\mathrm{L}^1(\Omega) \subset \mathrm{W}^{1,\infty}(\Omega)^*$ is continuous, Theorem 5.8 yields that the sequence $\hat{c}^{\varepsilon}$ is compact in $\mathrm{L}^1([0, T] \times \Omega)$. □

It is also clear that we get convergence towards the fast manifold, which results from the fast-reaction part of the slope term.

Lemma 5.10 (Convergence towards microscopic equilibrium and strong compactness) *Let* $(\mu^{\varepsilon})_{\varepsilon>0}$, $\mu^{\varepsilon} \in \mathrm{L}_w^{\infty}([0, T], Q)$ *satisfying the a priori bounds* (4.2). *Then there is a subsequence such that* $c^{\varepsilon} \to c^0$ *strongly in* $\mathrm{L}^1([0, T] \times \Omega, \mathbb{R}^2)$ *and, moreover, it holds* $\rho_1^{V,0} = \rho_2^{V,0}$ *a.e. in* $[0, T] \times \Omega$.

Proof The bound on the dissipation functional provides $\int_0^T \int_\Omega \left(\sqrt{\rho_1^{V,\varepsilon}} - \sqrt{\rho_2^{V,\varepsilon}} \right)^2 \mathrm{d}x \mathrm{d}t \leq$

$C\varepsilon$. Hence, we conclude $\| \sqrt{\rho_1^{V,\varepsilon}} - \sqrt{\rho_2^{V,\varepsilon}} \|_{L^2([0,T]\times\Omega)} \to 0$ as $\varepsilon \to 0$. In particular, we conclude that $\rho_1^{V,0} = \rho_2^{V,0}$. The strong convergence towards the slow manifold provides strong convergence for the whole sequence. Indeed, using Cauchy-Schwartz inequality and $x - y = \left(\sqrt{x} - \sqrt{y} \right) \left(\sqrt{x} + \sqrt{y} \right)$, we have

$$\| \rho_1^{V,\varepsilon} - \rho_2^{V,\varepsilon} \|_{L^1([0,T]\times\Omega)} \leq \| \sqrt{\rho_1^{V,\varepsilon}} - \sqrt{\rho_2^{V,\varepsilon}} \|_{L^2([0,T]\times\Omega)} \| \sqrt{\rho_1^{V,\varepsilon}} + \sqrt{\rho_2^{V,\varepsilon}} \|_{L^2([0,T]\times\Omega)}.$$

The last term can be estimated by the inequality of arithmetic and geometric mean

$$\| \sqrt{\rho_1^{V,\varepsilon}} + \sqrt{\rho_2^{V,\varepsilon}} \|_{L^2([0,T]\times\Omega)}^2 = \int_0^T \int_\Omega \rho_1^{V,\varepsilon} + \rho_2^{V,\varepsilon} + 2\sqrt{\rho_1^{V,\varepsilon} \rho_2^{V,\varepsilon}} \, \mathrm{d}x \mathrm{d}t$$

$$\leq 2 \int_0^T \int_\Omega \left(\rho_1^{V,\varepsilon} + \rho_2^{V,\varepsilon} \right) \mathrm{d}x \mathrm{d}t,$$

and the right-hand side is bounded since $\mu(t) \in Q$ for $t \in [0,T]$. Hence, we conclude that $\| \rho_1^{V,\varepsilon} - \rho_2^{V,\varepsilon} \|_{L^1([0,T]\times\Omega)} \to 0$ as $\varepsilon \to 0$.

Using this convergence, we have that also $c_i^\varepsilon \to c_i^0$ strongly in $L^1([0,T] \times \Omega)$. Indeed, a direct computation shows that

$$c_i^\varepsilon - w_i^V \frac{c_1^0 + c_2^0}{w_1^V + w_2^V} = (-1)^i \left(\frac{\frac{c_2^\varepsilon}{w_2^V} - \frac{c_1^\varepsilon}{w_1^V}}{w_1^V w_2^V \left(w_1^V + w_2^V \right)} \right) + \frac{w_i^V}{w_1^V + w_2^V} \left(c_1^\varepsilon + c_2^\varepsilon - (c_1^0 + c_2^0) \right),$$

and both terms converge strongly to zero as $\varepsilon \to 0$ by convergence of $\rho_1^{V,\varepsilon} - \rho_2^{V,\varepsilon} \to 0$ and $\hat{c}^\varepsilon \to \hat{c}^0$. $\qquad\square$

Finally, we show that the limit $\mu^0 = c^0 \mathrm{d}x$ has an absolutely continuous representative in the space of probability measures. To do this, we exploit the characterization of absolutely continuous curves as solutions of the continuity equation following [1].

Proposition 5.11 *Let $(\mu^\varepsilon)_{\varepsilon>0}$, $\mu^\varepsilon \in L^\infty([0,T], Q)$ satisfying the a priori bounds* (4.2) *and let c^0 be the limit of the densities c^ε. Then the coarse-grained slow variable $\hat{\mu} = \mu_1^0 + \mu_2^0 = \left(c_1^0 + c_2^0 \right) \mathrm{d}x \in L_w^\infty([0,T], \mathrm{Prob}(\Omega))$ has a representative (in time), which is absolutely continuous in the space of probability measures equipped with the 2-Wasserstein metric. Moreover, each component μ_i^0 has an absolutely continuous representative (in time), which is absolutely continuous in the space of non-negative Radon measures equipped with the 1-Wasserstein metric.*

Proof The coarse-grained measures $\hat{\mu}^\varepsilon$ satisfy the continuity equation $\dot{\hat{\mu}}^\varepsilon + \mathrm{div}(\hat{J}^\varepsilon) = 0$ in the sense of distributions, where $\hat{J}^\varepsilon = J_1^\varepsilon + J_2^\varepsilon$ is the coarse-grained diffusion flux. Since the linear continuity equation is stable under weak convergence, we conclude that also the limits satisfy the same continuity equation $\dot{\hat{\mu}}^0 + \mathrm{div}(\hat{J}^0) = 0$, where $\hat{J}^0$ is the weak*-limit of $\hat{J}^\varepsilon$ (see Lemma 5.7). Using (4.10), the bound on the dissipation functional (4.2) implies a bound $\int_0^T \int_\Omega Q(\hat{c}^0, \hat{J}^0) \mathrm{d}x \mathrm{d}t < \infty$. Let us define the transport velocity $\hat{v} \in \mathcal{M}([0,T] \times \Omega, \mathbb{R})$

$$\hat{v} = \begin{cases} \frac{\hat{J}}{\hat{c}} & \text{for } \hat{c} > 0 \\ 0 & \text{for } \hat{c} = 0 \end{cases}.$$

Then $\int_0^T \int_\Omega Q(\hat{c}, \hat{J}) dx dt = \frac{1}{2} \int_0^T \int_\Omega |\hat{v}|^2 \hat{c} dx dt = \frac{1}{2} \int_0^T \int_\Omega |\hat{v}|^2 d\hat{\mu} dt$ and the bound on the dissipation functional implies the bound on the Borel velocity field $\|\hat{v}\|_{L^2(\hat{\mu})} < \infty$. Hence, by Theorem 8.3.1 from [1] it follows that $t \mapsto \hat{\mu}(t) \in (\text{Prob}(\Omega), \mathcal{W}_2)$ has an absolutely continuous representative.

To prove time-regularity for μ_i^0 for $i = 1, 2$, we first observe that $\mu_i^0 = \frac{w_i^V}{w_1^V + w_2^V} \hat{\mu}$ is a non-negative Radon measure. To show that it has an absolutely continuous representative, we proceed exactly as in Lemma 5.6 and exploit again the dual formulation of the 1-Wasserstein distance on the space of non-negative Radon measures. Here, we use that $\frac{w_i^V}{w_1^V + w_2^V}$ is in $C^1(\Omega)$.

5.2 Liminf-estimate

Once the compactness is established the proof of the liminf-estimate is comparatively easy.

Theorem 5.12 *Let* $\mu^\varepsilon \to \mu^0$ *in* $L_w^\infty([0, T], Q)$ *such that* $\sup_{\varepsilon \in]0,1]} \sup_{t \in [0,T]} \mathcal{E}(\mu^\varepsilon) < \infty$. *Then, we have the liminf-estimate*

$$\liminf_{\varepsilon \to 0} \mathfrak{D}_\varepsilon^V(\mu^\varepsilon) \geq \mathfrak{D}_0^V(\mu^0),$$

where the limit dissipation functional is defined as in Theorem 4.1.

Proof We may assume that $\mathfrak{D}_\varepsilon^V(\mu^\varepsilon) \leq C < \infty$ (otherwise the claim is trivial). For the given curves $t \mapsto \mu^\varepsilon(t) \in Q$ take diffusive fluxes J^ε and reactive fluxes b^ε, which satisfy the generalized continuity equation

$$(c, J, b) \in (\text{gCE}) \quad \Leftrightarrow \quad \left\{ b_1 + b_2 = 0 \text{ and } \left\{ \begin{matrix} \dot{c}_1 = -\operatorname{div} J_1 + b_1 \\ \dot{c}_2 = -\operatorname{div} J_2 + b_2 \end{matrix} \right\} \right\},$$

and approximate the infimum in $\mathfrak{D}_\varepsilon^V(\mu^\varepsilon)$ arbitrarily close, i.e.

$$\mathfrak{D}_\varepsilon^V(\mu_\varepsilon) + \varepsilon \geq \int_0^T \mathcal{D}_\varepsilon^V(\mu_\varepsilon, J^\varepsilon, b^\varepsilon) dt.$$

The integrand $\mathcal{D}_\varepsilon^V$ consists of a velocity and a slope part and both of them split into a reaction and a diffusion part:

$$\mathcal{D}_\varepsilon^V(\mu, J, b) = \int_\Omega \sum_{j=1}^2 \tilde{Q}(\delta_j c_j, J_j) dx + \int_\Omega \tilde{C}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right) dx +$$

$$+ \frac{1}{2} \int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{|\nabla \rho_j^V|^2}{\rho_j^V} dx + \frac{2}{\varepsilon} \int_\Omega \sqrt{w_1^V w_2^V} \left(\sqrt{\rho_1^V} - \sqrt{\rho_2^V}\right)^2 dx$$

(5.1)

$$=: \mathcal{V}_{\text{diff}}(\mu, J) + \mathcal{V}_{\text{react},\varepsilon}(\mu, b) + \mathcal{S}_{\text{diff}}(\mu) + \mathcal{S}_{\text{react},\varepsilon}(\mu).$$

Clearly, we also have $\int_0^T \mathcal{D}_\varepsilon^V(\mu_\varepsilon, J^\varepsilon, b^\varepsilon) dt \leq C + \varepsilon < \infty$. By Lemma 5.10, we conclude compactness for the densities $c^\varepsilon \to c^0$ in $L^1([0, T] \times \Omega \times \{1, 2\})$ and by Lemma 5.7 that $J^\varepsilon \overset{*}{\rightharpoonup} J^0$ in $\mathcal{M}([0, T] \times \Omega \times \{1, 2\})$. Using the lower-semicontinuity result from Lemma 5.2 (which implies the liminf-estimates for $\mathcal{V}_{\text{diff}}$ and $\mathcal{S}_{\text{diff}}$) and that $\mathcal{V}_{\text{react},\varepsilon}, \mathcal{S}_{\text{react},\varepsilon} \geq 0$, we

obtain the estimate $\liminf_{\varepsilon\to 0}\int_0^T \mathcal{D}_\varepsilon^V(\mu^\varepsilon, J^\varepsilon, b^\varepsilon)\mathrm{d}t \geq \int_0^T \mathcal{D}_0^V(\mu^0, J^0, b^0)\,\mathrm{d}t$, where we analogously define

$$\mathcal{D}_0^V(\mu, J, b) := \int_\Omega \sum_{j=1}^2 \widetilde{\mathsf{Q}}(\delta_j c_j, J_j)\mathrm{d}x + \frac{1}{2}\int_\Omega \sum_{j=1}^2 \delta_j w_j^V \frac{|\nabla \rho_j|^2}{\rho_j}\mathrm{d}x. \qquad (5.2)$$

Let us define $u_i^\varepsilon = \dot{c}_i^\varepsilon + \mathrm{div}J_i^\varepsilon$. We conclude convergence for $u_i^\varepsilon \to u_i^0$ in the sense of distributions, and, moreover, we have $u_1^\varepsilon + u_2^\varepsilon \to \dot{c}_1^0 + \dot{c}_2^0 + \mathrm{div}J_1^0 + \mathrm{div}J_2^0 = 0 = u_1^0 + u_2^0$. In particular, we conclude the pointwise estimate

$$\int_\Omega \sum_{j=1}^2 \widetilde{\mathsf{Q}}(\delta_j c_j^0, J_j^0)\mathrm{d}x \geq \inf_{(J,u)}\left\{\int_\Omega \sum_{j=1}^2 \widetilde{\mathsf{Q}}(\delta_j c_j^0, J_j^0)\mathrm{d}x : \begin{cases} \dot{c}_1 + \mathrm{div}J_1 = u_1 \\ \dot{c}_2 + \mathrm{div}J_2 = u_2 \\ u_1 + u_2 = 0 \end{cases}\right\},$$

which finally establishes the liminf-estimate $\liminf_{\varepsilon\to 0}\int_0^T \mathcal{D}_\varepsilon^V(\mu^\varepsilon, J^\varepsilon, b^\varepsilon)\mathrm{d}t \geq \mathfrak{D}_0^V(\mu)$.

5.3 Construction of the recovery sequence

In this section, we construct the recovery sequence for the functional $\mathfrak{D}_0^V$ to finish the Γ-convergence result in Theorem 4.1. To be precise, we will show the following:

Theorem 5.13 *Let $\mu^0 \in \mathrm{L}_w^\infty([0, T] \times \Omega, \mathbb{R}_{\geq 0}^2)$ such that the a priori bounds $\mathfrak{D}_0^V(\mu^0) < \infty$ and $\mathrm{ess\,sup}_{t\in[0,T]}\mathcal{E}(\mu^0(t)) < \infty$ hold. Then there is a sequence $(\mu^\varepsilon)_{\varepsilon>0}$, $\mu^\varepsilon \in \mathrm{AC}([0, T], Q)$, $\sup_{\varepsilon>0}\mathrm{ess\,sup}_{t\in[0,T]}\mathcal{E}(\mu^\varepsilon(t)) < \infty$, such that the densities converge $c^\varepsilon \to c^0$ strongly in $\mathrm{L}^1([0, T] \times \Omega \times \{1, 2\})$ and we have $\mathfrak{D}_\varepsilon^V(\mu^\varepsilon) \to \mathfrak{D}_0^V(\mu^0)$.* □

Using Proposition 4.4, we see that the limit functionals do not contain more information than the functionals in coarse-grained variables and it holds $\mathfrak{D}_0^V(\mu^0) = \hat{\mathfrak{D}}(\hat{\mu}^0)$. Hence, we may reconstruct the dissipation functional with the corresponding diffusion and reaction flux (J, b) from the coarse-grained variables $(\hat{c}, \hat{J})$.

Technical difficulties arise because the dissipation functional $\mathfrak{D}_0^V$ is defined not only for solution of the evolution Eq. (1.2) but for general trajectories or fluctuations. These fluctuations are in general neither strictly positive nor smooth. The proof of the limsup-estimate is done in several steps, which are elaborated in the next lemmas. The bound $\mathfrak{D}_0^V(\mu^0) < \infty$ can be assumed without loss of generality because the other case is already treated in the liminf-estimate.

Proof of Theorem 5.13 We do the reconstruction in three steps using different approximation methods. We will do the following:

1. Proposition 5.21 shows that for $\mu^0 = c^0\mathrm{d}x$ with sufficiently smooth and positive density c^0 the constant sequence $\mu^{\theta,\gamma} = \mu^0$ satisfies $\left|\mathfrak{D}_\varepsilon^V(\mu^{\theta,\gamma}) - \mathfrak{D}_0^V(\mu^{\theta,\gamma})\right| \to 0$.
2. Lemma 5.16 overcomes the positivity assumption, i.e. it shows that for all $\mu^0 = c^0\mathrm{d}x$ there is a positive c^γ such that $c^\gamma \to c^0$ and $\mathfrak{D}_0^V(\mu^\gamma) \to \mathfrak{D}_0^V(\mu^0)$ as $\gamma \to 0$.
3. A mollification argument as in [1, Lemma 8.1.10] and stated in Lemma 5.20 allows us to overcome regularity by smoothing, which shows $\mathfrak{D}_0^V(\mu^{\theta,\gamma}) \to \mathfrak{D}_0^V(\mu^\gamma)$ as $\theta \to 0$.

Hence, defining the recovery sequence $\mu^\varepsilon := \mu^{\theta,\gamma}$, we have

$$\left|\mathfrak{D}_\varepsilon^V(\mu^\varepsilon) - \mathfrak{D}_0^V(\mu^0)\right| \leq \left|\mathfrak{D}_\varepsilon^V(\mu^{\theta,\gamma}) - \mathfrak{D}_0^V(\mu^{\theta,\gamma})\right| + \left|\mathfrak{D}_0^V(\mu^{\theta,\gamma}) - \mathfrak{D}_0^V(\mu^\gamma)\right| + \left|\mathfrak{D}_0^V(\mu^\gamma) - \mathfrak{D}_0^V(\mu^0)\right|,$$

where the first term tends to zero by the first reconstruction step, the second term tends to zero by the third reconstruction step and the third term tends to zero by the second reconstruction step, which in total proves the desired convergence.   □

Before performing the three recovery steps in Sect. 5.3.2, we first illustrate the general idea of constructing the recovery sequence neglecting positivity and regularity issues for the moment.

5.3.1 Construction of recovery sequence for smooth and positive measures

To show the general idea, let us firstly assume that the density of $\hat{\mu}$ is sufficiently smooth and positive, i.e. we assume that its Lebesgue density satisfies $\hat{c} \geq \frac{1}{C} > 0$ on $\Omega \times [0, T]$ and has enough regularity that will be specified below. Let $\hat{J}$ be the diffusion flux which provides the minimum in $\hat{\mathfrak{D}}(\hat{\mu}^0) = \mathfrak{D}_0^V(\mu)$ and satisfies $\hat{c} + \mathrm{div}(\hat{J}) = 0$. We define the reconstructed variables (pointwise in $\Omega \times [0, T]$) by

$$c_1 = \frac{w_1^V}{w_1^V + w_2^V}\hat{c}, \quad c_2 = \frac{w_2^V}{w_1^V + w_2^V}\hat{c}, \quad J_1 = \frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\hat{J}, \quad J_2 = \frac{\delta_2 w_2^V}{\delta_1 w_1^V + \delta_2 w_2^V}\hat{J}.$$

$$b_1 = \left(\frac{\delta_1 - \delta_2}{\delta_1 w_1^V + \delta_2 w_2^V}\frac{w_1^V w_2^V}{w_1^V + w_2^V}\right)\mathrm{div}\,\hat{J} + \hat{J} \cdot \nabla\left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\right), \quad b_2 = -b_1. \quad (5.3)$$

The reconstructed concentrations c and diffusion fluxes $J = (J_1, J_2)$ are proportional to the coarse grained concentration $\hat{c}$ and diffusion flux $\hat{J}$, respectively. On the coarse-grained level, which considers only one species there is no reaction flux anymore. (This changes, when considering large reaction-diffusion systems as explained in Sect. 6.) The reactive flux $b = (b_1, b_2)$ is given as a function of the coarse-grained diffusion flux $\hat{J}$, which means that in the limit the diffusion determines the hidden reaction.

Concerning regularity issues, we immediately observe the following. Since w^V is smooth and positive, c_1, c_2 have the same regularity as $\hat{c}$ and also J_1, J_2 have both the same regularity as $\hat{J}$. In contrast, the reaction fluxes b_i are a priori not well-defined by the pointwise expression (5.3) for general diffusion fluxes $\hat{J} \in \mathcal{M}(\Omega \times [0, T], \mathbb{R}^d)$. In the following we are going to regularize $\hat{J}$ to make the above formula exact.

Remark 5.14 Note, that regularity assumptions for $\mathrm{div}\,\hat{J}$ are not needed if $\delta_1 = \delta_2$, i.e. if both species diffuse with the same diffusion constant. In particular, in this situation no regularization argument as in Lemma 5.20 is necessary. Moreover, no additional regularity for $\hat{J}$ is needed if $\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V} = \vartheta \in {]}0, 1{[}$ is constant. This is equivalent to

$$V_1(x) - V_2(x) = \log\left(\frac{1 - \vartheta}{\vartheta}\frac{\delta_1}{\delta_2}\frac{\beta}{\alpha}\right) = \mathrm{const},$$

which means that the potentials V_1, V_2 differ in a constant on Ω implying that $\nabla\hat{V} = \nabla V_1 = \nabla V_2$. As we will see, enough regularity for $\hat{J}$ is already obtained from bounds on the dissipation functional.

So let us assume for the moment that b_i is well-defined. Then we conclude (c, J, b) solves the generalized continuity equation (gCE), because we have

$$
\dot{c}_1 + \operatorname{div} J_1 = \frac{w_1^V}{w_1^V + w_2^V}\dot{\hat{c}} + \operatorname{div}\left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\hat{J}\right) =
$$

$$
= \frac{w_1^V}{w_1^V + w_2^V}\dot{\hat{c}} + \hat{J}\cdot\nabla\left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\right) + \left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\right)\operatorname{div}\hat{J}
$$

$$
= -\frac{w_1^V}{w_1^V + w_2^V}\operatorname{div}\hat{J} + \hat{J}\cdot\nabla\left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\right) + \left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V}\right)\operatorname{div}\hat{J} = b_1,
$$

where we used that $(\hat{c}, \hat{J})$ solves $\dot{\hat{c}} + \operatorname{div}\hat{J} = 0$. Similarly, we see that $\dot{c}_2 + \operatorname{div}J_2 = b_2$ and, by definition, we have $b_1 + b_2 = 0$. Moreover, boundary properties of $\hat{J}$ remain for $J = (J_1, J_2)$.

For the dissipation functionals, we obtain that

$$
\mathfrak{D}_\varepsilon^V(\mu) \leq \int_0^T \mathcal{D}_0^V(\mu, J, b) = \mathfrak{D}_0^V(\mu) + \int_0^T\int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right)\mathrm{d}x\mathrm{d}t,
$$

where we used formula (5.1) and that the reconstructed concentrations satisfy $\frac{c_1}{w_1^V} = \frac{c_2}{w_2^V}$. That means, that for proving that the constant sequence $\mu^\varepsilon = \mu$ is a recovery sequence, it suffices to show that $\int_0^T\int_\Omega\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right)\mathrm{d}x\mathrm{d}t \to 0$ as $\varepsilon \to 0$. This is, in fact, shown in the next lemma under the assumption that $\operatorname{div}\hat{J}, \hat{J} \in \mathsf{L}^\mathsf{C}([0, T]\times\Omega)$. The proof basically uses the monotonicity property of the Legendre dual function $\widetilde{\mathsf{C}}(a_1, b) \leq \widetilde{\mathsf{C}}(a_2, b)$ as $a_1 \geq a_2$ (see Lemma 4.3), its superlinear growth and the dominated convergence theorem.

Lemma 5.15 *Let $\hat{c} \in \mathsf{L}^1(\Omega\times[0, T])$ with $\hat{c} \geq \frac{1}{C}$ a.e. in $[0, T]\times\Omega$ for a constant $C > 0$, and let $\hat{J} : [0, T]\times\Omega \to \mathbb{R}^d$ satisfy $\operatorname{div}\hat{J}, |\hat{J}| \in \mathsf{L}^\mathsf{C}([0, T]\times\Omega)$. Then for b_2 defined by (5.3) we have $\int_0^T\int_\Omega\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right)\mathrm{d}x\mathrm{d}t \to 0$ as $\varepsilon \to 0$.*

Proof Since $\hat{c} \geq \frac{1}{C}$ a.e., Lemma 4.3 yields the pointwise estimate $\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right) \leq \widetilde{\mathsf{C}}\left(\frac{1}{C\varepsilon}, b_2\right) = \frac{1}{C\varepsilon}\mathsf{C}(C\varepsilon b_2)$. Moreover, using the inequality (3.5) we have

$$
\int_0^T\int_\Omega\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right)\mathrm{d}x\mathrm{d}t \leq \frac{1}{C\varepsilon}\int_0^T\int_\Omega\mathsf{C}(C\varepsilon b_2)\,\mathrm{d}x\mathrm{d}t \leq \int_0^T\int_\Omega 2|b_2|\log(C\varepsilon|b_2| + 1)\mathrm{d}x\mathrm{d}t.
$$

By assumption, we have that $\operatorname{div}\hat{J}, |\hat{J}| \in \mathsf{L}^\mathsf{C}([0, T]\times\Omega)$. Since $V \in \mathsf{C}^1(\Omega)$ and the Orlicz space $\mathsf{L}^\mathsf{C}([0, T]\times\Omega)$ is a Banach space, we conclude that $b_2 \in \mathsf{L}^\mathsf{C}([0, T]\times\Omega)$. By the inequality (3.5) for C, this implies that for $\varepsilon < \frac{1}{C}$, the right-hand side is bounded. We show that (for a subsequence) the integrand converges to zero pointwise a.e. in $[0, T]\times\Omega$. By the dominated convergence theorem, this would imply that $\int_0^T\int_\Omega\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1 c_2}}{\varepsilon}, b_2\right)\mathrm{d}x\mathrm{d}t \to 0$ as $\varepsilon \to 0$.

To see that the integrand converges to zero pointwise, we firstly observe that $b_2 \in \mathsf{L}^\mathsf{C}([0, T]\times\Omega) \subset \mathsf{L}^1([0, T]\times\Omega)$, which means that

$$
\int_0^T\int_\Omega \log(\varepsilon|b_2| + 1)\mathrm{d}x\mathrm{d}t \leq \varepsilon\int_0^T\int_\Omega|b_2|\mathrm{d}x\mathrm{d}t \to 0.
$$

Hence, (for a subsequence) $\log(\varepsilon|b_2| + 1)$ converges pointwise to zero and, thus, also $|b_2|\log(C\varepsilon|b_2| + 1)$. $\qquad\square$

In fact, the above proof is quite robust and already suggests that the same convergence holds even if $\hat{c}^\varepsilon$ might slowly converge to zero in $\Omega \times [0, T]$ and $\|b_\varepsilon^2\|_{L^p} \approx \varepsilon^{-\alpha}$ for some $\alpha > 0$ and $p > 1$, (see Proposition 5.21).

5.3.2 Auxiliary results for constructing the recovery sequence for general measures

First, we show how to overcome the positivity assumption by a controlled positive shift.

Lemma 5.16 *For all $\mu^0 = c^0 \, dx$ satisfying $\mathfrak{D}_0^V(\mu^0) = \hat{\mathfrak{D}}(\hat{\mu}^0) < \infty$ there is a sequence $(\hat{c}^\gamma)_{\gamma > 0}$ of densities satisfying $\hat{c}^\gamma \geq \gamma$, $\hat{c}^\gamma \to \hat{c}$ in $L^1([0, T] \times \Omega)$ as $\gamma \to 0$ such that for their corresponding measures we have $\hat{\mathfrak{D}}(\hat{\mu}^\gamma) = \mathfrak{D}_0(\mu^\gamma) \to \mathfrak{D}_0(\mu^0) = \hat{\mathfrak{D}}(\mu^0)$ as $\gamma \to 0$ and $\sup_{\gamma \in]0,1]} \mathrm{ess} \sup_{t \in [0,T]} \mathcal{E}(\mu^\gamma(t)) < \infty$.*

Proof For small $\gamma > 0$, we define $\hat{c}^\gamma := \frac{1}{Z_\gamma}(\hat{c} + 2\gamma)$, where $Z_\gamma = 1 + 2\gamma |\Omega| = 1 + 2\gamma > 0$ is the normalization factor such that $\int_\Omega \hat{c}^\gamma \, dx = 1$. Hence, $Z_\gamma \searrow 1$, $\hat{c}^\gamma \to \hat{c}$, $\sup_{\gamma \in]0,1]} \mathrm{ess} \sup_{t \in [0,T]} \mathcal{E}(\mu^\gamma(t)) < \infty$ and w.l.o.g. we assume that $\hat{c}^\gamma \geq \gamma$. Moreover, we define $\hat{J}^\gamma := \frac{1}{Z_\gamma} \hat{J}$. Clearly, $\hat{J}^\gamma \cdot \nu = 0$ on $\partial\Omega$ and $(\hat{c}^\gamma, \hat{J}^\gamma)$ solves the continuity equation $\dot{\hat{c}}^\gamma + \mathrm{div} \hat{J}^\gamma = 0$. We compute the terms in the dissipation functional $\mathfrak{D}_0^V(\mu^\gamma)$. We have $\hat{\rho}^{V,\gamma} = \frac{\hat{c}^\gamma}{w^V} = \frac{1}{Z_\gamma} \frac{\hat{c}+2\gamma}{w^V}$. Using $\max\left\{ \|\nabla(1/w^V)\|_{L^\infty}, \|w^V\|_{L^\infty} \right\} \leq C$, we get

$$\nabla \hat{\rho}^{V,\gamma} = \frac{1}{Z_\gamma}\left\{ \nabla\left(\frac{\hat{c}}{w^V}\right) + 2\gamma \nabla\left(\frac{1}{w^V}\right) \right\} \quad \Rightarrow \quad \left|\nabla \hat{\rho}^{V,\gamma}\right| \leq \frac{1}{Z_\gamma}\left\{ \left|\nabla \hat{\rho}^V\right| + 2\gamma C \right\}.$$

Using the estimate $\frac{1}{a+\delta} \leq \frac{1}{a}$ for $a, \delta > 0$, $\frac{1}{Z_\gamma} \leq 1$ and the inequality $2xy \leq \sqrt{\gamma}x^2 + \frac{1}{\sqrt{\gamma}}y^2$, we get the pointwise estimate

$$\frac{|\nabla \hat{\rho}^{V,\gamma}|^2}{\hat{\rho}^{V,\gamma}} = \frac{1}{Z_\gamma}\frac{\left(|\nabla \hat{\rho}^V| + 2\gamma C\right)^2}{\hat{\rho}^V + \frac{2\gamma}{w^V}} = \frac{1}{Z_\gamma}\frac{|\nabla \hat{\rho}^V|^2 + 4\gamma C |\nabla \hat{\rho}^V| + 4\gamma^2 C^2}{\hat{\rho}^V + \frac{2\gamma}{w^V}}$$

$$\leq \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V} + 2\frac{2\gamma C |\nabla \hat{\rho}^V|}{\hat{\rho}^V + \frac{2\gamma}{w^V}} + \frac{4\gamma^2 C^2}{\hat{\rho}^V + \frac{2\gamma}{w^V}}$$

$$\leq \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V} + \frac{1}{\hat{\rho}^V + \frac{2\gamma}{w^V}}\left(\sqrt{\gamma}\left|\nabla \hat{\rho}^V\right|^2 + \frac{1}{\sqrt{\gamma}}\{2\gamma C\}^2\right) + 2\gamma C^2 w^V$$

$$\leq \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V}(1 + \sqrt{\gamma}) + 2\sqrt{\gamma}w^V C^2(1 + \sqrt{\gamma}).$$

Hence, $\int_0^T \int_\Omega \hat{\delta}^V w^V \frac{|\nabla \hat{\rho}^{V,\gamma}|^2}{\hat{\rho}^{V,\gamma}} dx dt \to \int_0^T \int_\Omega \hat{\delta}^V w^V \frac{|\nabla \hat{\rho}^V|^2}{\hat{\rho}^V} dx dt$ as $\gamma \to 0$.

Similarly, we get

$$\int_0^T \int_\Omega \frac{|\hat{J}^\gamma|^2}{\hat{\delta}^V \hat{c}^\gamma} dx dt = \frac{1}{Z_\gamma}\int_0^T \int_\Omega \frac{|\hat{J}|^2}{\hat{\delta}^V(\hat{c}+2\gamma)} dx dt \leq \int_0^T \int_\Omega \frac{|\hat{J}|^2}{\hat{\delta}^V \hat{c}} dx dt,$$

which implies that $\int_0^T \int_\Omega \tilde{Q}(\hat{\delta}^V \hat{c}^\gamma, \hat{J}^\gamma) dx dt \leq \int_0^T \int_\Omega \tilde{Q}(\hat{\delta}^V \hat{c}, \hat{J}) dx dt$. Hence, we conclude that $\hat{c}^\gamma \geq \gamma$, $\hat{c}^\gamma \to \hat{c}$ and $\hat{\mathfrak{D}}(\hat{c}^\gamma) \to \hat{\mathfrak{D}}(\hat{c})$ as $\gamma \to 0$. □

Next, we are going to show that the reaction flux $b_1 = -b_2$ given by (5.3) can be made sufficiently smooth, i.e. at least in L^C which would allow us to proceed similar as in Lemma 5.15.

Recalling the formula for the reconstructed flux, we have

$$b_2 = \left(\frac{\delta_1 - \delta_2}{\delta_1 w_1^V + \delta_2 w_2^V} \frac{w_1^V w_2^V}{w_1^V + w_2^V} \right) \operatorname{div} \hat{J} + \hat{J} \cdot \nabla \left(\frac{\delta_1 w_1^V}{\delta_1 w_1^V + \delta_2 w_2^V} \right) =: a_1 \operatorname{div} \hat{J} + \hat{J} \cdot a_2,$$
(5.4)

and $b_2 = -b_1$, where $a_1 \in C^1(\Omega, \mathbb{R})$, $a_2 \in C^0(\Omega, \mathbb{R}^d)$. In particular, the regularity of $b_1 = -b_2$ does not depend on a_1, a_2. We are going to prove that $\operatorname{div} \hat{J}$ and $\hat{J}$ have enough regularity. The regularity of $\hat{J}$ follows from the bound on the dissipation functional. The regularity of $\operatorname{div} \hat{J}$ is achieved by mollification.

First, we show that $\hat{J} \in L^{\tilde{p}}([0, T] \times \Omega)$, for some $\tilde{p} > 1$. Clearly, we have $\hat{J} \in L^1([0, T] \times \Omega)$ by Lemma 5.1 and the bound on the dissipation functional. To improve the regularity of $\hat{J}$, we show that $\hat{c} \in L^p([0, T] \times \Omega)$, which follows from the bound on the dissipation functional $\mathfrak{D}_0^V$ and the energy functional $\mathcal{E}$.

Lemma 5.17 *Let $\mu^0 \in L_w^\infty([0, T], Q)$ such that $\mathfrak{D}_0^V(\mu^0) < \infty$ and ess $\sup_{t \in [0, T]} \mathcal{E}(\mu^0(t)) < \infty$. Then the density $\hat{c}^0$ is in $L^p([0, T] \times \Omega)$ with $p = \frac{d+1}{d} > 1$.*

Proof The bound on the energy yields $\hat{c} \in L^\infty([0, T], L^1(\Omega))$. The bound on the slope part of the dissipation functional together with Lemma 5.1 provides that $\hat{c} \in L^1([0, T], W^{1,1}(\Omega))$ (see also Lemma 5.7). By the Sobolev embedding theorem, we have the compact embedding $W^{1,1}(\Omega) \subset L^q(\Omega)$, where $1 - \frac{1}{d} > \frac{1}{q} \Leftrightarrow q < \frac{d}{d-1}$. Thus $\hat{c} \in L^\infty([0, T], L^1(\Omega)) \cap L^1([0, T], L^{\frac{d}{d-1}}(\Omega))$.

Next, we apply a classical interpolation result for Lebesgue spaces, see e.g. Theorem 5.1.2 of [5]. We have for all $\theta \in]0, 1[$ that

$$[L^{p_1}([0, T], X_1), L^{p_2}([0, T], X_2)]_\theta \simeq L^{p_\theta}([0, T], [X, Y]_\theta),$$

where $\frac{1}{p_\theta} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$. In our situation we have $p_1 = 1$, $p_2 = \infty$, $X_1 = L^{\frac{d}{d-1}}(\Omega)$, $X_2 = L^1(\Omega)$. Hence, $p_\theta = \frac{1}{1-\theta} > 1$. Moreover, $[X_1, X_2]_\theta = [L^q(\Omega), L^1(\Omega)]_\theta \simeq L^{q_\theta}(\Omega)$, where $\frac{1}{q_\theta} = \frac{1-\theta}{q} + \frac{\theta}{1}$. Setting $p_\theta = q_\theta$, we conclude $\frac{1-\theta}{1-2\theta} = q = \frac{d}{d-1}$. Solving $p_\theta = q_\theta$ for θ, we obtain $\theta = \frac{1}{1+d}$, and hence $p_\theta = q_\theta = \frac{d+1}{d} > 1$. Summarizing, we conclude $\hat{c} \in L^{\frac{d+1}{d}}([0, T] \times \Omega)$. □

Remark 5.18 Iterating the above procedure, it is possible to show $\hat{c} \in L^{(d+2)/d}([0, T] \times \Omega)$ for $d \geq 2$.

Knowing integrability of $\hat{c}$, we get also better integrability of the fluxes $\hat{J} \in L^{\tilde{p}}([0, T] \times \Omega)$ for some $\tilde{p} > 1$, which follows from the next lemma.

Lemma 5.19 *Let $\mu^0 \in L_w^\infty([0, T] \times \Omega, \mathbb{R}_{\geq 0}^2)$ such that $\mathfrak{D}_0^V(\mu^0) < \infty$ and ess $\sup_{t \in [0, T]} \mathcal{E}(\mu^0(t)) < \infty$ and let $\hat{J} \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$ be the corresponding diffusion flux satisfying the continuity equation. Then $\hat{J} \in L^{\tilde{p}}([0, T] \times \Omega, \mathbb{R}^d)$ for $\tilde{p} = \frac{2d+2}{2d+1} > 1$.*

Proof First, we show that for all $J \in \mathbb{R}^d$ and $c > 0$ we have

$$\frac{|J|^2}{c} + \frac{1}{p} c^p \geq \left(1 + \frac{1}{p} \right) |J|^{\frac{2p}{p+1}}.$$

To see this, let us define for fixed $J \in \mathbb{R}^d$ the function $F :]0, \infty[\to \mathbb{R}$, $F(c) := \frac{|J|^2}{c} + \frac{1}{p} c^p$. Clearly, $F \geq 0$ and $F(c) \to \infty$ as $c \to 0$ or $c \to \infty$. We compute the minimum. We have $F'(c) = -|J|^2 c^{-2} + c^{p-1}$ and hence the critical point is at $c_0 = |J|^{2/(p+1)}$. Inserting c_0 into F we get $F(c) \geq F(c_0) = |J|^2 |J|^{-2/(p+1)} + \frac{1}{p} |J|^{2p/(p+1)} = (1 + \frac{1}{p}) |J|^{2p/(p+1)}$.

Having this estimate, we see that by Lemma 5.17 we have $\hat{c} \in L^p([0, T] \times \Omega)$ for $p = \frac{d+1}{d}$. This implies $\hat{J} \in L^{\tilde{p}}([0, T] \times \Omega, \mathbb{R}^d)$ for $\tilde{p} = \frac{2\frac{d+1}{d}}{\frac{d+1}{d}+1} = \frac{2d+2}{2d+1}$. □

To obtain regularity for the whole reaction flux $b_1 = -b_2$, we have to obtain regularity also for $\text{div} \hat{J}$. This is done by mollifying the solution $(\hat{c}, \hat{J})$ of the continuity equation $\dot{\hat{c}} + \text{div} \hat{J} = 0$ in time. We already know that $\hat{c} : [0, T] \to \text{Prob}(\Omega)$ is continuous by Lemma 5.11. With a slight abuse of notation, we denote by $\hat{c}$ also a continuous extension on $\mathbb{R}$ such that $\hat{c} \in L^p(\mathbb{R}, L^p(\Omega))$. Now, we mollify in time and define $\hat{c}^\theta(t) = \int_{\mathbb{R}} \hat{c}(s) \psi_\theta(t - s) \text{d}s$ where ψ_θ is a positive and symmetric mollifier that approximates a dirac-distribution as $\theta \to 0$. Analogously, we define $\hat{J}^\theta$ by convolution, i.e. $\hat{J}^\theta(t) = \int_{\mathbb{R}} \hat{J}(s) \psi_\theta(t - s) \text{d}s$. Since the continuity equation is linear, the smoothed functions $(\hat{c}^\theta, \hat{J}^\theta)$ satisfy again the continuity equation with the same no-flux boundary conditions. The next lemma shows how the dissipation functional $\mathfrak{D}_0^V(\mu)$ can be approximated using mollification.

Lemma 5.20 *Let* $\mu^0 \in L_w^\infty([0, T] \times \Omega, \mathbb{R}_{\geq 0}^2)$ *such that the a priori bounds* $\mathfrak{D}_0^V(\mu^0) < \infty$ *and* $\text{ess sup}_{t \in [0,T]} \mathcal{E}(\mu^0(t)) < \infty$ *hold. Let* $\hat{J} \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$ *be the corresponding diffusion flux satisfying the continuity equation. Let* $\psi^\theta : \mathbb{R} \to \mathbb{R}$ *be a positive and symmetric mollifier. Define* $\hat{c}^\theta(t) = \int_{\mathbb{R}} \hat{c}(s) \psi_\theta(t - s) \text{d}s$ *and* $\hat{J}^\theta(t) = \int_{\mathbb{R}} \hat{J}(s) \psi_\theta(t - s) \text{d}s$. *Then, we have* $\sup_{\theta \in]0,1]} \text{ess sup}_{t \in [0,T]} \mathcal{E}(\mu^\theta(t)) < \infty$ *and* $\int_0^T \mathfrak{D}_0^V(\mu^\gamma, J^\gamma, b^\gamma) \text{d}t \to \int_0^T \mathfrak{D}_0^V(\mu^0, J^0, b^0) \text{d}t$.

Proof The energy bound on μ^θ is trivially satisfied. The convergence for the dissipation functional $\int_0^T \mathfrak{D}_0^V \text{d}t$ follows directly as the proof of Lemma 8.1.10 in [1] since the integrand is convex in $(\hat{c}, \hat{J})$. □

With these preparations, we are able to show the remaining step in the proof of Theorem 5.13.

Proposition 5.21 *Let* $\mu^0 \in L_w^\infty([0, T] \times \Omega, \mathbb{R}_{\geq 0}^2)$ *such that* $\mathfrak{D}_0^V(\mu^0) = \hat{\mathfrak{D}}_0(\hat{\mu}) < \infty$ *and* $\text{ess sup}_{t \in [0,T]} \mathcal{E}(\mu^0(t)) < \infty$ *and let* $\hat{J} \in \mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$ *be the corresponding diffusion flux satisfying the continuity equation. Let* $\psi^\theta : \mathbb{R} \to \mathbb{R}$ *be a positive and symmetric mollifier, which is specified below. Let* $\hat{c}^\theta, \hat{J}^\theta$ *the mollified functions as in Lemma 5.20 and* $\hat{c}^{\theta,\gamma}, \hat{J}^{\theta,\gamma}$ *the mollified and shifted functions as in Lemma 5.16.*

Let ψ^θ *be such that* $\|\dot{\hat{c}}^\theta\|_{L^{\tilde{p}}([0,T] \times \Omega)} \leq C \varepsilon^{-\lambda_1}$ *and* $\gamma \geq C \varepsilon^{1-\lambda_2}$ *for* $\varepsilon, \theta, \gamma \to 0$, *where* $\tilde{p} = \frac{2d+2}{2d+1}$ *is the integrability exponent of the fluxes as in Lemma 5.19, $C > 0$ is a positive constant and* $\lambda_1, \lambda_2 \in [0, 1[$, *satisfy the inequality* $2\lambda_1(d + 1) \leq \lambda_2$. *Then, we have* $|\mathfrak{D}_\varepsilon^V(\mu^{\theta,\gamma}) - \mathfrak{D}_0^V(\mu^{\theta,\gamma})| \to 0$ *as* $\varepsilon, \theta, \gamma \to 0$.

Proof First, we observe that for given $\hat{c}$ such a mollifier and these constants λ_1, λ_2 satisfying all the conditions can be easily constructed.

To prove the convergence, we follow the same strategy as in the proof of Lemma 5.15. Defining the reconstructed concentrations and fluxes as in (5.3) and using the formulas (5.1)

and (5.2), we observe that

$$\mathfrak{D}_\varepsilon^V(\mu^{\theta,\gamma}) \le \int_0^T \mathcal{D}_\varepsilon^V(\mu^{\theta,\gamma}, J^{\theta,\gamma}, b^{\theta,\gamma}) \, dt$$

$$= \int_0^T \mathcal{D}_0^V(\mu^{\theta,\gamma}, J^{\theta,\gamma}, b^{\theta,\gamma}) \, dt + \int_0^T \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1^{\theta,\gamma} c_2^{\theta,\gamma}}}{\varepsilon}, b_2^{\theta,\gamma}\right) \, dx \, dt.$$

Lemma 5.16 and Lemma 5.20 show that the first term converges to $\int_0^T \mathcal{D}_0^V(\mu^0, J^0, b^0) dt$ as $\varepsilon, \theta, \gamma \to 0$. Hence, in order to prove the desired convergence it suffices to show that the second term tends to zero.

Using the bound from below on $\hat{c}^\epsilon$, and the inequality $\log(x+1) \le C_{\tilde{p}} x^{\tilde{p}-1}$, we get the pointwise estimate

$$\widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1^{\theta,\gamma} c_2^{\theta,\gamma}}}{\varepsilon}, b_2^{\theta,\gamma}\right) \le \widetilde{\mathsf{C}}\left(C\varepsilon^{-\lambda_2}, b_2^{\theta,\gamma}\right) \le C\varepsilon^{-\lambda_2}\mathsf{C}\left(C^{-1}\varepsilon^{\lambda_2} b_2^{\theta,\gamma}\right)$$

$$\le 2C\varepsilon^{-\lambda_2}|b_2^{\theta,\gamma}|C^{-1}\varepsilon^{\lambda_2} \log\left(C^{-1}\varepsilon^{\lambda_2}|b_2^{\theta,\gamma}| + 1\right) \le \widetilde{C}|b_2^{\theta,\gamma}|^{\tilde{p}}\varepsilon^{\lambda_2(\tilde{p}-1)},$$

where $\widetilde{C} = \widetilde{C}(C_{\tilde{p}}, C)$. Using the continuity equation for $(\hat{c}^\theta, \hat{J}^\theta)$, we get from $\|\hat{c}^\theta\|_{L^{\tilde{p}}([0,T]\times\Omega)} \le C\varepsilon^{-\lambda_1}$, that $\|\text{div}\hat{J}^\theta\|_{L^{\tilde{p}}([0,T]\times\Omega)} \le C\varepsilon^{-\lambda_1}$. By Lemma 5.19, we also have $\hat{J}^\theta \in L^{\tilde{p}}([0,T]\times\Omega)$. Using the explicit formula (5.4) for $b_2^{\theta,\gamma}$, we have $\|b_2^{\theta,\gamma}\|_{L^{\tilde{p}}([0,T]\times\Omega)} \le \bar{C}\varepsilon^{-\lambda_1}$ with a new constant $\bar{C} > 0$. Hence, we get

$$\int_0^T \int_\Omega \widetilde{\mathsf{C}}\left(\frac{\sqrt{c_1^{\theta,\gamma} c_2^{\theta,\gamma}}}{\varepsilon}, b_2^{\theta,\gamma}\right) \, dx \, dt \lesssim \varepsilon^{-\tilde{p}\lambda_1}\varepsilon^{\lambda_2(\tilde{p}-1)} = \varepsilon^{\frac{1}{2d+1}\{\lambda_2 - (2d+2)\lambda_1\}}.$$

Since λ_1, λ_2 satisfy $(2d+2)\lambda_1 \le \lambda_2$, we conclude that the right-hand side converges to zero, which proves the claim. $\qquad\square$

6 Remarks for reaction-diffusion systems involving more species

In the final section, we describe the situation not for two but for general $I \in \mathbb{N}$ species. Considering $I \in \mathbb{N}$ species X_i which diffuse and react linearly $X_i \rightleftharpoons X_j$ the evolution of their concentrations $c \in \mathbb{R}_{\ge 0}^I$ is given by

$$\dot{c} = \text{diag}(\delta_1, \ldots, \delta_I)\Delta c + A^\varepsilon c,$$

where $\delta_i >$ are diffusion coefficients, $A^\varepsilon = A^S + \frac{1}{\varepsilon}A^F$ is a Markov generator (preserving positivity and total mass), which consists of a slow part and a fast part. The main assumption is that A^ε satisfies detailed balance with respect to its stationary measure w^ε. Similar to [30,34], we are going to assume that the stationary vector w^ε satisfies $w^\varepsilon \to w^0$ as $\varepsilon \to 0$ and that $w^0 > 0$. The positivity of the limit stationary measure w^0 means that in the limit the evolution is not degenerate and the concentration for all species is present.

The gradient structure is defined on the state space

$$Q = \text{Prob}(\Omega \times \{1, \cdots, I\}) := \left\{\mu = (\mu_1, \ldots, \mu_I) \; : \; \mu_i \in \mathcal{M}(\Omega), \; \mu_i \ge 0, \sum_{i=1}^I \mu_i(\Omega) = 1\right\}.$$

The driving energy functional $\mathcal{E}_\varepsilon : X \to \mathbb{R}_\infty$ has the form

$$\mathcal{E}_\varepsilon(\mu) = \begin{cases} \int_\Omega \sum_{j=1}^I E_B\left(\frac{c_j}{w_j^\varepsilon}\right) w_j^\varepsilon \mathrm{d}x, & \text{if } \mu = c\,\mathrm{d}x \\ \infty, & \text{otherwise.} \end{cases},$$

and the dual dissipation potential splits into two parts

$$\mathcal{R}^*(\mu, \xi) = \mathcal{R}^*_{\mathrm{diff}}(\mu, \xi) + \mathcal{R}^*_{\mathrm{react}}(\mu, \xi)$$

$$\mathcal{R}^*_{\mathrm{diff}}(\mu, \xi) = \frac{1}{2}\int_\Omega \sum_{j=1}^I \delta_j |\nabla\xi_j(x)|^2 \mathrm{d}\mu_j, \quad \mathcal{R}^*_{\mathrm{react},\varepsilon}(\mu, \xi) = \int_\Omega \sum_{i<j} \kappa_{ij}^\varepsilon C^*(\xi_i(x) - \xi_j(x))\, \mathrm{d}\sqrt{\mu_i\mu_j},$$

where $\kappa_{ij}^\varepsilon := A_{ij}^\varepsilon \left(\frac{w_j^\varepsilon}{w_i^\varepsilon}\right)^{1/2}$. In particular, the reaction part of the dissipation potential splits into a fast part and a slow part

$$\mathcal{R}^*_{\mathrm{react},\varepsilon}(\mu, \xi) = \mathcal{R}^*_{\mathrm{slow},\varepsilon}(\mu, \xi) + \frac{1}{\varepsilon}\mathcal{R}^*_{\mathrm{fast},\varepsilon}(\mu, \xi)$$

$$\mathcal{R}^*_{\mathrm{xy},\varepsilon}(\mu, \xi) = \int_\Omega \sum_{i<j} \tilde{\kappa}_{ij}^\varepsilon\, C^*(\xi_i(x) - \xi_j(x))\, \mathrm{d}\sqrt{\mu_i\mu_j}, \quad \mathrm{xy} \in \{\mathrm{slow}, \mathrm{fast}\},$$

where $\tilde{\kappa}_{ij}^\varepsilon$ are bounded and positive uniformly in $\varepsilon > 0$. In particular, we call a reaction and its flux b_{ij} *slow* if $A_{ij}^\varepsilon = O(1)$ and *fast* if $A_{ij}^\varepsilon = O(\varepsilon^{-1})$. Due to the detailed balance assumption and by $w^0 > 0$, the distinction between fast and slow reactions is indeed well-defined.

In the remainder of the section, we briefly explain how to generalize the proof of the EDP-convergence result also for this situation. Major differences occur at two stages, namely in 1) deriving compactness for slow reaction fluxes, 2) proving the limsup-estimate. The reaction fluxes of the fast reactions are not seen in the limit and have to be reconstructed in an analogous way as in the two-species situation.

6.1 Compactness for slow reaction fluxes and liminf-estimate

Similarly to the two species situation, compactness for the concentrations, by using strong convergence for coarse-grained variables and convergence towards the slow manifold, can be derived (cf. Lemmas 5.6, 5.10). Moreover, compactness of diffusion fluxes and spatial regularity follows, too (cf. Lemma 5.19).

In contrast to the situation of two species connected with one fast reaction, where no slow reaction fluxes exists, compactness for slow reaction fluxes b_{ij}^ε has to be derived in the multi species case. This follows immediately from Lemma 5.2, once compactness of $\sqrt{c_i^\varepsilon c_j^\varepsilon}$ is obtained. At this point it is clear that weak convergence of $c^\varepsilon \rightharpoonup c^0$ is not sufficient. Instead the previously derived strong convergence of $c^\varepsilon \to c^0$ implies by dominated convergence also strong convergence of $\sqrt{c_i^\varepsilon c_j^\varepsilon} \to \sqrt{c_i^0 c_j^0}$ in $L^1([0, T] \times \Omega)$, and hence, compactness for the slow fluxes b_{ij}^ε. Compactness for fast reaction fluxes can not be obtained as already mentioned in Remark 5.5. Having derived compactness, the proof of the liminf-estimate is exactly the same as for Theorem 5.12, since the functional $\mathcal{D}_\varepsilon$ is jointly convex in all variables (c, J, b).

6.2 Equilibration and reconstruction of reaction fluxes und recovery sequence

A crucial observation throughout the proof of the Γ-convergence was Lemma 4.3, which provides an equilibration of fluxes assuming microscopic equilibria for the concentrations. In Lemma 4.4, we derived equilibration for the diffusion fluxes. Similarly, also an equilibration of the slow reaction fluxes can be derived. In [30] a general operator-theoretic coarse-graining and reconstruction procedure has been developed. This method can also be applied to derive coarse-grained fluxes and a coarse-grained continuity equation, see [42]. Importantly for us, as in (5.3) the reconstructed slow reaction fluxes depend linearly on the coarse-grained reaction fluxes. The fast reaction fluxes are then of the form

$$b_{ij} = a_1 \operatorname{div} \hat{J}_i + a_2 \hat{J}_i + \sum_{i,j} a_{ij} \hat{b}_{ij},$$

where $a_j, a_{ij} \in C^0(\Omega, \mathbb{R}^k)$, $k \in \mathbb{N}$.

In order to prove that the constant sequence for smooth and positive concentrations is indeed a recovery sequence, we follow the same reasoning as in the Lemmas 5.15 and 5.21. The only difference comes from the explicit dependence on the coarse-grained reaction flux $\hat{b}_{ij}$. Using the bound on the limit dissipation functional (which provides bounds on $\int_0^T \int_\Omega \widetilde{C}(\sqrt{\hat{c}_i \hat{c}_j}, \hat{b}_{ij}) \mathrm{d}x \mathrm{d}t$) and the next Lemma 6.1, we obtain that $\hat{b}_{ij} \in L^C([0, T] \times \Omega)$ (we refer to [19] for a proof). Since L^C is an Orlicz-space, we conclude that the reconstructed fluxes b_i are in L^C. This allows to proceed as in Lemma 5.21 and proves the existence of a recovery sequence.

Lemma 6.1 ([19]) *Let $p > 1$. Then, for all $a \geq 0$ and $B \in \mathbb{R}$ we have*

$$\widetilde{C}(a, B) \geq \left(1 - \frac{1}{p}\right) C(B) - \frac{2}{p} a^p.$$

In particular, setting $a = \sqrt{c_i c_j}$ and $B = b_{ij}$ we have

$$\int_0^T \int_\Omega \widetilde{C}(\sqrt{c_i c_j}, b_{ij}) \mathrm{d}x \mathrm{d}t + \|\sqrt{c_i c_j}\|_{L^p([0,T] \times \Omega)}^p \gtrsim \int_0^T \int_\Omega C(b_{ij}) \mathrm{d}x \mathrm{d}t,$$

which proves that $b_{ij} \in L^C([0, T] \times \Omega)$ if $c_i, c_j \in L^p([0, T] \times \Omega)$ for some $p > 1$.

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel (2005)
2. Arnrich, S., Mielke, A., Peletier, M.A., Savaré, G., Veneroni, M.: Passing to the limit in a Wasserstein gradient flow: from diffusion to reaction. Calc. Var. Part. Diff. Eqns. **44**, 419–454 (2012)
3. Barbu, V., Precupanu, T.: Convexity and Optimization in Banach Spaces, 3rd edn. D. Reidel Publishing Co., Dordrecht (1986)
4. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000)
5. Bergh, J., Löfström, J.: Interpolation Spaces - An Introduction. Springer-Verlag, Berlin (1976)
6. Bothe, D., Hilhorst, D.: A reaction-diffusion system with fast reversible reaction. J. Math. Anal. Appl. **286**, 125–135 (2003)
7. Bothe, D., Pierre, M.: Quasi-steady-state approximation for a reaction-diffusion system with fast intermediate. J. Math. Anal. Appl. **368**(1), 120–132 (2010)
8. Bothe, D., Pierre, M.: The instantaneous limit for reaction-diffusion systems with a fast irreversible reaction. Discr. Cont. Dynam. Systems Ser. S **8**(1), 49–59 (2011)
9. Bothe, D.: Instantaneous limits of reversible chemical reactions in presence of macroscopic convection. J. Differ. Equ. **193**(1), 27–48 (2003)
10. Bothe, D., Pierre, M., Rolland, G.: Cross-diffusion limit for a reaction-diffusion system with fast reversible reaction. Comm. Partial Differ. Equ. **37**(11), 1940–1966 (2012)
11. Braides, A.: Local minimization, Variational Evolution and Gamma-convergence. Lecture Notes Mathematics, vol. 2094. Springer, Berlin (2014)
12. Dal Maso, G.: An Introduction to Γ-Convergence. Birkhäuser Boston Inc., Boston, MA (1993)
13. Daus, E.S., Desvillettes, L., Jüngel, A.: Cross-diffusion systems and fast-reaction limits. Bull. Sci. math. **159**, 102824 (2020)
14. Dondl, P., Frenzel, T., Mielke, A.: A gradient system with a wiggly energy and relaxed EDP-convergence. ESAIM Control Optim. Calc. Var. **25**(68), 1–45 (2019)
15. Edwards, D.: On the Kantorovich-Rubinstein theorem. Exposition. Math. **29**, 387–398 (2011)
16. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. Philadelphia, Pa.: Society for Industrial and Applied Mathematics (1999). https://epubs.siam.org/doi/pdf/10.1137/1.9781611971088
17. Evans, L.C.: A convergence theorem for a chemical diffusion-reaction system. Houst. J. Math. **6**(2), 259–267 (1980)
18. Frenzel, T., Liero, M.: Effective diffusion in thin structures via generalized gradient systems and EDP-convergence. Discr. Cont. Dynam. Systs. Ser. S **14**(1), 395–425 (2021)
19. Frenzel, T., Mielke, A.: Deriving the kinetic relation for the flux through a membrane via EDP-convergence. *in preparation* (2021)
20. Gladbach, P., Kopfer, E., Maas, J., Portinale, L.: Homogenisation of one-dimensional discrete optimal transport. J. de Mathematiques Pures et Appliquees **139**(7), 204–234 (2020)
21. Heida, M., Patterson, R.I.A., Renger, D.R.M.: Topologies and measures on the space of functions of bounded variation taking values in a Banach or metric space. J. Evol. Eqns. **19**, 111–152 (2018)
22. Hilhorst, D., van der Hout, R., Peletier, L.: Nonlinear diffusion in the presence of fast reaction. Nonlinear Anal. **41**, 803–823 (2000)
23. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)
24. Liero, M., Mielke, A.: Gradient structures and geodesic convexity for reaction-diffusion systems. Phil. Trans. Royal Soc. A **371(2005), 20120346**, 28 (2013)
25. Liero, M., Mielke, A., Peletier, M.A., Renger, D.R.M.: On microscopic origins of generalized gradient structures. Discr. Cont. Dynam. Systs. Ser. S **10**(1), 1–35 (2017)
26. Maniglia, S.: Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. J. Math. Pures Appl. (9) **87**(6), 601–626 (2007)
27. Marcelin, R.: Contribution a l'étude de la cinétique physico-chimique. Ann. de Physique **III**, 120–231 (1915)
28. Mielke, A.: A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. Nonlinearity **24**, 1329–1346 (2011)
29. Mielke, A.: On evolutionary Γ-convergence for gradient systems (Ch. 3). In Muntean, A., Rademacher, J., Zagaris, A., (eds.) *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, Lecture Notes in Applied Math. Mechanics Vol. 3, pp. 187–249. Springer, (2016). Proc. of Summer School in Twente University, (June 2012)

30. Mielke, A., Stephan, A.: Coarse-graining via EDP-convergence for linear fast-slow reaction systems. M3AS: Math. Models Meth. Appl. Sci. **30**(09), 1765–1807 (2020)
31. Mielke, A., Montefusco, A., Peletier, M.A.: Exploring families of energy-dissipation landscapes via tilting: three types of EDP convergence. Contin. Mech. Thermodyn. **33**, 611–637 (2021)
32. Mielke, A., Patterson, R.I.A., Peletier, M.A., Renger, D.R.M.: Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. SIAM J. Appl. Math. **77**(4), 1562–1585 (2017)
33. Mielke, A., Peletier, M.A., Renger, D.R.M.: On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. Potential Anal. **41**(4), 1293–1327 (2014)
34. Mielke, A., Peletier, M.A., Stephan, A.: EDP-convergence for nonlinear fast-slow reaction systems with detailed balance. *WIAS preprint 2781*, submitted to Nonlinearity, (2020)
35. Mielke, A., Rossi, R., Savaré, G.: Nonsmooth analysis of doubly nonlinear evolution equations. Calc. Var. Part. Diff. Eqns. **46**(1–2), 253–310 (2013)
36. Murakawa, H., Ninomiya, H.: Fast reaction limit of a three-component reaction-diffusion system. J. Math. Anal. Appl. **379**(1), 150–170 (2011)
37. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. Comm. Partial Differ. Equ. **26**, 101–174 (2001)
38. Peletier, M.A., Renger, D.R.M. Fast reaction limits via Γ-convergence of the flux rate functional. J. Dyn. Diff. Equat. (2021). https://doi.org/10.1007/s10884-021-10024-2
39. Peletier, M.A., Rossi, R., Savaré, G., Tse, O.: Jump processes as generalized gradient flows. arXiv:2006.10624 (2020)
40. Renger, D.R.M.: Gradient and GENERIC systems in the space of fluxes, applied to reacting particle systems. Entropy **20**(8), 586 (2018)
41. Stephan, A.: On EDP-convergence for gradient systems with different time scales. Proc. Appl. Math. Mech. **19**(1), e201900366 (2019)
42. Stephan, A.: Coarse-graining and reconstruction for Markov matrices. *in preparation* (2021)