

## MACROSCOPIC LIMIT OF THE BECKER–DÖRING EQUATION VIA GRADIENT FLOWS

ANDRÉ SCHLICHTING\*

**Abstract.** This work considers gradient structures for the Becker–Döring equation and its macroscopic limits. The result of Niethammer [*J. Nonlinear Sci.* 13 (2003) 115–122] is extended to prove the convergence not only for solutions of the Becker–Döring equation towards the Lifshitz–Slyozov–Wagner equation of coarsening, but also the convergence of the associated gradient structures. We establish the gradient structure of the nonlocal coarsening equation rigorously and show continuous dependence on the initial data within this framework. Further, on the considered time scale the small cluster distribution of the Becker–Döring equation follows a quasistationary distribution dictated by the monomer concentration.

**Mathematics Subject Classification.** 49J40, 34A34, 35L65, 49J45, 49K15, 60J27, 82C26

Received July 29, 2016. Accepted January 25, 2018.

### 1. INTRODUCTION

#### 1.1. The Becker–Döring model

In this work, we are interested in gradient structures for the Becker–Döring equation and its macroscopic limits. The Becker–Döring equation [3] is a model for the coagulation and fragmentation of clusters consisting of identical monomers. The main modeling assumption is only monomers are able to coagulate and fragment with other clusters in a way that the total density of monomers is conserved

$$\sum_{l=1}^{\infty} l n_l(t) = \sum_{l=1}^{\infty} l n_l(0) =: \varrho_0 \quad \text{for all } t > 0. \quad (1.1)$$

Hereby,  $n_l(t)$  is the density of clusters of size  $l$  at time  $t$ . The evolution of the densities  $n_l(t)$  is given by an countable number of ordinary differential equations of the form

$$\dot{n}_l(t) = J_{l-1}(t) - J_l(t) \quad l = 2, 3, \dots \quad (1.2)$$

---

*Keywords and phrases:* Gradient flows, energy-dissipation principle, evolutionary Gamma convergence, quasistationary states, well-prepared initial conditions.

Institut für Angewandte Mathematik, Universität Bonn, Bonn, Germany.

\* Corresponding author: [Schlichting@iam.uni-bonn.de](mailto:Schlichting@iam.uni-bonn.de)

where  $J_l$  is the flux from clusters of size  $l$  to clusters of size  $l + 1$ . The system (1.2) gets closed with an equation for  $n_1$

$$\dot{n}_1(t) = - \sum_{l=1}^{\infty} J_l(t) - J_1(t) =: J_0(t) - J_1(t), \quad (1.3)$$

which is chosen, such that formally (1.1) is satisfied. The fluxes  $J_l$  are given by mass-action kinetics, that is the rate of coagulation is determined by  $a_l n_1 n_l$  and the rate of fragmentation is given by  $b_{l+1} n_{l+1}$ , where  $a_l$  and  $b_{l+1}$  are rate factors only depending on  $l$ . This leads to the constitutive relation

$$J_l(t) = a_l n_1(t) n_l(t) - b_{l+1} n_{l+1}(t), \quad l = 1, 2, \dots \quad (1.4)$$

The detailed balance condition for this system reads  $J_l(t) = 0$  for all  $l$ , satisfied by a one-parameter family of equilibrium solutions

$$\omega_l(z) := z^l Q_l, \quad \text{with } Q_1 := 1 \quad \text{and} \quad Q_l := \prod_{j=1}^{l-1} \frac{a_j}{b_{j+1}}. \quad (1.5)$$

To specify the long-time behavior, we introduce the convergence radius of the series  $z \mapsto \sum_l l z^l Q_l$  by  $z_s \in [0, \infty]$  as well as its value at the convergence radius

$$\varrho_s := \sum_{l=1}^{\infty} l z_s^l Q_l \in [0, \infty]. \quad (1.6)$$

We are interested in the regime where  $z_s \in (0, \infty)$  and  $\varrho_s \in (0, \infty)$ . We will assume that the rates are explicitly given as follows:

**Assumption 1.1** (Rates). For  $\alpha \in [0, 1)$ ,  $\gamma \in (0, 1)$  and  $z_s, q > 0$  define the coagulation and fragmentation rate of a monomer for a cluster of size  $l$  by

$$a_l := l^\alpha \quad \text{and} \quad b_l := l^\alpha (z_s + q l^{-\gamma}).$$

Hereby, the parameter  $z_s$  is consistent with its definition as radius of convergence (*cf.* Lem. 4.1) and  $\varrho_s$  as defined in (1.6) is strictly positive and finite under Assumption 1.1.

Then, as investigated by [2] solutions to the Becker–Döring equation with  $\varrho_0 \leq \varrho_s$  converge to the equilibrium state  $\omega_l(z)$ , where  $z = z(\varrho_0)$  is given such that  $\sum_{l=1}^{\infty} l z^l Q_l = \varrho_0$  and the convergence takes place in a weighted  $\ell^1$  space

$$\lim_{t \rightarrow \infty} \sum_{l=1}^{\infty} l |n_l(t) - \omega_l(z)| = 0$$

In the case  $\varrho_0 > \varrho_s$ , it holds

$$\lim_{t \rightarrow \infty} n_l(t) = \omega_l(z_s) \quad \text{for each } l \geq 1.$$

Hence, the excess mass  $\varrho_0 - \varrho_s > 0$  vanishes in the limit  $t \rightarrow \infty$ . The interpretation is, that the excess mass is contained in larger and larger clusters as times evolve. These large clusters form a new phase, *e.g.* liquid droplets

formed out of supersaturated vapor. It is the aim of this work to add some aspect to the understanding of the formation of the new phase.

The crucial ingredient for the above convergence statements is the existence of a Lyapunov functional  $\mathcal{F}$  of the form of a relative entropy  $\mathcal{F}_z(n) := \mathcal{H}(n \mid \omega(z))$ . Hereby,  $z > 0$  is a parameter selecting the stationary state and the relative entropy is defined by

$$\mathcal{H}(n \mid \omega) := \sum_{l=1}^{\infty} \omega_l \psi\left(\frac{n_l}{\omega_l}\right) \quad \text{with} \quad \psi(a) := a \log a - a + 1, \quad \text{for } a > 0. \quad (1.7)$$

A calculation shows that it is formally decreasing along solutions to the Becker–Döring equation

$$\frac{d\mathcal{F}_z(n(t))}{dt} = - \sum_{l=1}^{\infty} (a_l n_1 n_l - b_{l+1} n_{l+1}) (\log a_l n_1 n_l - \log b_{l+1} n_{l+1}) =: -\mathcal{D}(n(t)) \leq 0. \quad (1.8)$$

Hence, the Lyapunov function can be interpreted as a free energy dissipating along the flow. This indicates, that the free energy is minimized as  $t \rightarrow \infty$ . By the mass conservation (1.1), we expect the long-time limit to be the solution to the following minimization problem

$$\inf \left\{ \mathcal{F}(n) : \sum_{l=1}^{\infty} l n_l = \varrho_0 \right\} = \begin{cases} \mathcal{F}_z(\omega(z)), & \varrho_0 \leq \varrho_s; \\ \mathcal{F}_{z_s}(\omega(z_s)), & \varrho_0 > \varrho_s. \end{cases} \quad (1.9)$$

In the first case the infimum is attained and the parameter  $z = z(\varrho_0)$  is chosen such that  $\sum_{l=1}^{\infty} l \omega_l(z) = \varrho_0$ . In the second case the infimum is not attained (*cf.* [2], Thm. 4.4). From now on, we choose  $z = z(\varrho_0)$  in this particular form and omit the subscript. Hence, the functional reflects correctly the long-time behavior of the equation. Moreover, the Lyapunov function has the form of a relative entropy and the question arises, whether there exists a gradient structure for the Becker–Döring equation having this relative entropy as driving free energy.

## 1.2. Gradient flow structure

To bring the system into the framework of gradient-flows, it is helpful to interpret the Becker–Döring equation as the following system of chemical reactions



Hereby,  $X_l$  denotes a cluster of size  $l$  and the rates for coagulation  $\{a_l\}_{l \geq 1}$  and fragmentation  $\{b_l\}_{l \geq 2}$  are positive as in Assumption 1.1. In this formulation, we can use the gradient structure as observed by Mielke [15] for chemical reactions under detailed balance condition and it turns out that the Becker–Döring equation is indeed a gradient flow with respect to the Lyapunov function (1.7) under a suitable metric. The same metric was discovered by Maas [14] in the setting of reversible Markov chains.

The existence of the metric depends crucially on the detailed balance condition satisfied by the equilibrium (1.5)

$$a_l \omega_1 \omega_l = b_{l+1} \omega_{l+1} =: k^l, \quad (1.11)$$

where  $k^l$  is the stationary equilibrium flux and the implicit parameter  $z$  is chosen according to  $\varrho_0$  as described after (1.9). Equations (1.2), (1.3), (1.4) can be compactly rewritten with the help of (1.11) as

$$\dot{n} = - \sum_{l=1}^{\infty} k^l \left( \frac{n_1 n_l}{\omega_1 \omega_l} - \frac{n_{l+1}}{\omega_{l+1}} \right) (e^1 + e^l - e^{l+1}), \quad (1.12)$$

with  $e_i^l = 0$  for  $i \neq l$  and  $e_l^l = 1$  for  $l \in \mathbb{N}$ . Since the free energy  $\mathcal{F}$  is of the form of a relative entropy (1.7), we can identify its variation as

$$D\mathcal{F}(n) = \left( \log \frac{n_1}{\omega_1}, \dots, \log \frac{n_i}{\omega_i}, \dots \right).$$

Then, the gradient flow formulation of the Becker-Döring equation takes the form

$$\dot{n} = -\mathcal{K}(n) D\mathcal{F}(n), \quad (1.13)$$

where the Onsager matrix  $\mathcal{K}$  is defined by

$$\mathcal{K}(n) := \sum_{l=1}^{\infty} k^l \Lambda \left( \frac{n_1 n_l}{\omega_1 \omega_l}, \frac{n_{l+1}}{\omega_{l+1}} \right) (e^1 + e^l - e^{l+1}) \otimes (e^1 + e^l - e^{l+1}) \quad (1.14)$$

and  $\Lambda(\cdot, \cdot)$  is the logarithmic mean given for  $a, b > 0$  by

$$\Lambda(a, b) = \int_0^1 a^s b^{1-s} ds = \begin{cases} \frac{a-b}{\log a - \log b}, & a \neq b \\ a, & a = b. \end{cases} \quad (1.15)$$

The identification of (1.12) and (1.13) is based on the algebraic identity

$$\Lambda(ab, c) (\log a + \log b - \log c) = ab - c \quad \text{for } a, b, c > 0.$$

We refer to Appendix A for the more general structure behind this identities and applications to other coagulation and fragmentation models.

### 1.3. Variational characterization

The gradient flow formulation allows for a variational characterization initiated by de Giorgi and its collaborators [7] under the name of curves of maximal slope. From the interpretation of the Becker-Döring model as chemical reaction, it is clear that the total number of particles is conserved, which suggests to define the state manifold

$$\mathcal{M} := \left\{ n \in \mathbb{R}_+^{\mathbb{N}} : \sum_{l=1}^{\infty} l n_l = \varrho_0 \right\}.$$

Possible variations of the state manifold consistent with the Becker-Döring dynamic are given by the linear space  $\mathcal{TM} = \text{span}\{e^1 + e^l - e^{l+1} : l \in \mathbb{N}\}$ . By the definition of the Onsager matrix (1.14), we have that the following space is well-defined

$$\mathcal{T}_n^* \mathcal{M} := \{ \phi \in \mathbb{R}^{\mathbb{N}} : \exists s \in \mathcal{TM} \text{ such that } s = -\mathcal{K}(n)\phi \}. \quad (1.16)$$

A crucial ingredient to study the underlying metric structure is the continuity equation and curves of finite action.

**Definition 1.2** (Curves of finite action). A pair  $[0, T] \ni t \mapsto (n(t), \phi(t)) \in \mathcal{M} \times \mathcal{T}_{n(t)}^* \mathcal{M}$  is a solution to the continuity equation, denoted by  $(n, \phi) \in \mathcal{CE}_T$ , if it satisfies

- (i)  $n(\cdot) : [0, T] \rightarrow \mathcal{M}$  is absolute continuous.
- (ii) The pair  $(n, \phi)$  satisfies the continuity equation for  $t \in (0, T)$  in the weak form, that is for all  $\psi \in C_c^1((0, T), \mathbb{R})$  and all  $l \in \mathbb{N}$  holds

$$\int_0^T \left( \dot{\psi}(t) n_l(t) - \psi(t) (\mathcal{K}(n(t))\phi(t))_l \right) dt = 0. \quad (1.17)$$

The *action*  $\mathcal{A}$  of a pair  $(n, \phi) \in \mathcal{M} \times \mathcal{T}_n^* \mathcal{M}$  is defined by

$$\mathcal{A}(n, \phi) := \phi(t) \cdot \mathcal{K}(n(t))\phi(t) = \sum_{r=1}^R k^r \hat{n}_r^\omega |\nabla_r \phi(t)|^2, \quad (1.18)$$

where  $\nabla_r \phi := \phi_{r+1} - \phi_r - \phi_1$  and

$$\hat{n}_l^\omega := \Lambda \left( \frac{n_1 n_l}{\omega_1 \omega_l}, \frac{n_{l+1}}{\omega_{l+1}} \right). \quad (1.19)$$

A curve  $(n, \phi) \in \mathcal{CE}_T$  is called a *curve of finite action*, if

$$\sup_{t \in [0, T]} \mathcal{F}(n(t)) < \infty, \quad \int_0^T \mathcal{A}(n(t), \phi(t)) dt < \infty, \quad \text{and} \quad \int_0^T \mathcal{D}(n(t)) dt < \infty,$$

where  $\mathcal{D}(n) = \mathcal{A}(n, -D\mathcal{F}(n))$  is given as in (1.8).

The nonlocal gradient  $\nabla_r \phi$  in (1.18) can be avoided by interpreting the monomer concentration  $n_1 = \rho - \sum_{l=2}^\infty l n_l$  as a nonlocal boundary condition. Along this idea a Fokker-Planck equation with such type of boundary condition having similar features like the Becker–Döring model was recently introduced in [6].

Curves of finite action give a variational formulation to solutions of the Becker–Döring equation. In comparison to the direct gradient flow equation (1.13), this avoids regularity questions arising from the application of the chain rule. The concept was introduced in [7] and further investigated in [1]: For any curve  $(n, \phi) \in \mathcal{CE}_T$  of finite action holds

$$\mathcal{J}(n) := \mathcal{F}(n(T)) - \mathcal{F}(n(0)) + \frac{1}{2} \int_0^T \mathcal{D}(n(t)) dt + \frac{1}{2} \int_0^T \mathcal{A}(n(t), \phi(t)) dt \geq 0. \quad (1.20)$$

Moreover, equality is attained if and only if  $n$  is a solution of (1.12).

We provide the crucial observation of the proof, which follows formally by evaluating

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(n(t)) &= D\mathcal{F}(n(t)) \cdot \dot{n}(t) \stackrel{(1.17)}{=} D\mathcal{F}(n(t)) \cdot \mathcal{K}(n(t))\phi(t) \\ &\geq -\frac{1}{2} D\mathcal{F}(n(t)) \cdot \mathcal{K}(n(t)) D\mathcal{F}(n(t)) - \frac{1}{2} \phi(t) \cdot \mathcal{K}(n(t))\phi(t), \end{aligned}$$

where we used that  $\mathcal{K}$  is positive semidefinite and the Cauchy–Schwarz inequality. The equality case is read off from the equality case in Cauchy–Schwarz. For a rigorous treatment in a similar situation, we refer to Section 2.5 of [10]. We use this variational structure to pass to the limit after a suitable rescaling.

#### 1.4. The macroscopic limit

The connection between the Becker–Döring equation with positive excess mass  $\varrho_0 - \varrho_s = \bar{\varrho} > 0$  and a macroscopic theory of coarsening is due to Penrose [22]. He observed by formal asymptotics that the macroscopic part of the Becker–Döring dynamics converges after a suitable rescaling (*cf.* Sect. 2.1) to a classical coarsening model introduced by Lifshitz and Slyozov [13], and Wagner [25]

$$\partial_t \nu_t + \partial_\lambda (\lambda^\alpha (u(t) - q\lambda^{-\gamma}) \nu_t) = 0 \quad \text{with} \quad u(\nu_t) = \frac{q \int \lambda^{\alpha-\gamma} d\nu_t}{\int \lambda^\alpha d\nu_t}. \quad (1.21)$$

Hereby, the measure  $\nu_t(d\lambda)$  is the distribution of particles of macroscopic size  $\lambda \in \mathbb{R}_+$ . Moreover, the parameters  $\alpha$ ,  $\gamma$  and  $q$  satisfy Assumption 1.1 and we will call the nonlocal conservation law (1.21) the LSW equation in the following. Formally, the total mass is conserved and the evolution stays in the state manifold for any  $t > 0$

$$\nu_t \in M = \left\{ \nu \in C_c^0(\mathbb{R}_+)^* \mid \int \lambda \nu(d\lambda) = \bar{\varrho} \right\}.$$

The LSW equation are a gradient flow as formally observed by Niethammer ([17], Sect. 4). The driving energy of the system is given exactly by the first order expansion of the macroscopic part of a suitable rescaling of the free energy (1.7) (with  $z = z_s$ ) driving the Becker–Döring equation (*cf.* Lem. 4.2)

$$E(\nu) := \frac{q}{1-\gamma} \int \lambda^{1-\gamma} \nu(d\lambda). \quad (1.22)$$

Let us introduce a formal Riemannian structure and define a tangent space on  $M$  by  $T_\nu M := \{s : \mathbb{R}_+ \rightarrow \mathbb{R} \mid \int \lambda s d\lambda = 0\}$ . An identification of tangent and cotangent vectors is obtained *via* the operator  $K(\nu) : T_\nu^* M \rightarrow T_\nu M$  given by

$$K(\nu)w := -\partial_\lambda (\lambda^\alpha w \nu) \quad \text{for} \quad w \in T_\nu^* M := \{w : \mathbb{R}_+ \rightarrow \mathbb{R} \mid \exists s \in T_\nu M : K(\nu)v = s\}. \quad (1.23)$$

By an integration by parts of the identity  $0 = \int \lambda s \nu(d\lambda)$  holds the inclusion property  $T_\nu^* M \subseteq \{w : \mathbb{R}_+ \rightarrow \mathbb{R} \mid \int \lambda^\alpha w \nu(d\lambda) = 0\}$ . Let us formally derive the gradient structure for the LSW equation (*cf.* [17], Sect. 4), that is we assume all differentials and quantities to be smooth enough. The differential of the energy (1.22) is given for some  $s \in T_\nu M$  by using the identification  $s = -K(\nu)w$  with  $w \in T_\nu^* M$

$$\begin{aligned} DE(\nu) \cdot s &= \frac{q}{1-\gamma} \int \lambda^{1-\gamma} s d\lambda = - \int \left( u\lambda - \frac{q}{1-\gamma} \lambda^{1-\gamma} \right) s d\lambda \\ &= - \int \lambda^\alpha (u - q\lambda^{-\gamma}) w \nu(d\lambda), \end{aligned}$$

where  $u \in \mathbb{R}$  can be chosen such that  $u - DE(\nu) \in T_\nu^* M$  thanks to  $\int \lambda^\alpha w \nu(d\lambda) = 0$ . Then, the gradient flow in weak form satisfies for all  $\tilde{s} \in T_\nu M$

$$\int \lambda^\alpha w \tilde{w} \nu(d\lambda) = g_{\alpha,\nu}(\partial_t \nu, \tilde{s}) = -DE(\nu) \cdot \tilde{s} = \int \lambda^\alpha (u - q\lambda^{-\gamma}) \tilde{w} \nu(d\lambda),$$

where  $\partial_t \nu = -\partial_\lambda(\lambda^\alpha w \nu)$  and  $\tilde{s} = -\partial_\lambda(\lambda^\alpha \tilde{w} \nu)$  in distribution. Hence, we obtain the identification

$$w = u - q\lambda^{-\gamma},$$

where  $u = u(\nu)$  is a Lagrangian multiplier chosen such that  $w \in T_\nu^*M$ , that is it satisfies the constraint  $\int \lambda^\alpha w \nu(d\lambda) = 0$  and is formally given by (1.21). Hence, the gradient flow of the energy  $E$  with respect to the metric induced by  $K$  is given by

$$\partial_t \nu_t = -K(\nu)DE(\nu) = -\partial_\lambda(\lambda^\alpha(u(\nu_t) - q\lambda^{-\gamma})\nu_t),$$

where  $u(\nu_t)$  given by (1.21). To make the above observation rigorous, we use the de Giorgi formalism of curves of maximal slope. Up to technical details, which is dealt with in Section 3, we can define an action functional as follows: For a pair  $(\nu, w)$  solving the continuity equation  $\partial_t \nu_t + \partial_\lambda(\lambda^\alpha w_t \nu_t) = 0$  in distributions, denoted by  $(\nu, w) \in \text{CE}_T$ , the action is defined by

$$A(\nu_t, w_t) := \int \lambda^\alpha |w_t|^2 d\nu_t.$$

Then, by the identification of tangent and co-tangent vectors *via*  $s = -\partial_\lambda(\lambda^\alpha w \nu)$ , we obtain that the dissipation is given by

$$D(\nu_t) = A(\nu_t, u(t) - DE(\nu_t)) = \int \lambda^\alpha |u(\nu_t) - q\lambda^{-\gamma}|^2 d\nu_t, \quad (1.24)$$

where  $u(\nu_t) \in L^2((0, T))$  given by (1.21) ensures that  $u(\nu_t) - q\lambda^{-\gamma} \in T_\nu^*M$ , *i.e.* it is a valid cotangent vector satisfying  $\int \lambda^\alpha (u(\nu_t) - q\lambda^{-\gamma}) d\lambda = 0$ .

The functional  $J(\nu)$ , which completely characterizes solutions to (1.21) (*cf.* Prop. 3.6) is defined by

$$J(\nu) := E(\nu_T) - E(\nu_0) + \frac{1}{2} \int D(\nu_t) dt + \frac{1}{2} \int A(\nu_t, w_t) dt \geq 0, \quad (1.25)$$

with  $J(\nu) = 0$  if and only if  $\nu_t$  is a weak solution to the LSW equation (1.21).

The main application of this variational framework is to prove the convergence of the Becker–Döring gradient structure to the LSW gradient structure. In addition, the variational characterization of the LSW equation together with a compactness statement for curves of finite action (*cf.* Prop. 3.5) allows to proof continuous dependence on the initial data (*cf.* Cor. 3.8).

## 1.5. Passage to the limit

The macroscopic limit is rigorously derived by Niethammer [16]. There, the main technical tool was to pass to the limit in the energy-dissipation relation associated with the rescaled Becker–Döring equation to obtain the energy-dissipation relation of the LSW equation. The one for solutions to the Becker–Döring equation is obtained by integrating the identity (1.8) in time

$$\mathcal{F}(n(T)) - \mathcal{F}(n(0)) + \int_0^T \mathcal{D}(n(t)) dt = 0. \quad (1.26)$$

The functional  $\mathcal{J}$  from (1.20) contains the identity (1.26), since for solutions of the Becker–Döring equation it holds  $\mathcal{A}(n(t), -D\mathcal{F}(n(t))) = \mathcal{D}(n(t))$ .

Likewise, from (1.25) and (1.24) follows that the LSW equation satisfy the energy-dissipation identity

$$E(\nu_T) - E(\nu_0) + \int_0^T D(\nu_t) dt = 0.$$

where  $u(t)$  is given in (1.21).

The contribution of this work is to lift the convergence statement from the level of energy-dissipation relations along solutions to the functionals  $\mathcal{J}$  and  $J$  along curves of finite action. Hereby, by doing so no essential new technical difficulties arise, which underlines the fact that the gradient structure is natural for these types of equations. We prove that a suitable rescaling of the functional  $\mathcal{F}^\varepsilon$  converges to the functional  $J$  in an evolutionary  $\Gamma$ -convergence sense under the assumption of well-prepared initial data (see Thm. 2.2). In particular, the gradient structure of the Becker–Döring equation converges to the one of the LSW equation (*cf.* Thm. 2.2) and in particular it implies the convergence of solutions (*cf.* Cor. 2.3). This program follows the ideas of Sandier and Serfaty [23], and was later generalized by Serfaty [24].

The ingredients of the proof of convergence are based on: (i) the variational characterization of the Becker–Döring equations in Section 1.3, which follows the gradient structure established by [15]; (ii) the rigorous variational characterization of solutions to the LSW equation in Section 3, which extends the formal gradient structure of Section 4 of [17]; (iii) *a priori* estimates for the variational framework of the Becker–Döring gradient structure in Section 4.1, which lifts many of the results of [19] from solutions of the Becker–Döring system to curves of finite action.

Another motivation to reconsider the proof of [16] is that systems possessing a gradient structure can be well described by studying convexity properties of the free energy with respect to the implied metric. Especially, the results of [21] suggest, that the system shows dynamic metastability as described by [20] for gradient systems. Under this point of view also the additional results on quasistationarity in the next subsection are first steps towards a characterization of dynamic metastability of the Becker–Döring equations.

## 1.6. Well preparedness of initial data and quasistationarity

A crucial assumption in the approach of showing convergence *via* curves of maximal slope is the well preparedness of initial data, which assumes that the rescaled free energy of the Becker–Döring gradient structure converges to the one of the LSW gradient structure

$$\mathcal{F}^\varepsilon(n^\varepsilon(0)) \rightarrow E(\nu_0) \quad \text{as } \varepsilon \rightarrow 0.$$

The second contribution of this work is to show that on the rescaled time-scale, the Becker–Döring equation reach instantaneously a quasistationary equilibrium, which is dictated only by the monomer concentration. On the other hand, the monomer concentration follows closely a macroscopic quantity similarly defined as  $u$  in (1.21). The crucial ingredient in the proof is an energy-dissipation estimate based on a logarithmic Sobolev inequality similarly to the one used in [4] to proof convergence to equilibrium in the noncondensing case  $\varrho_0 \leq \varrho_s$ .

The quasistationary result shows, that the microscopic part of the rescaled free energy  $\mathcal{F}^\varepsilon(n^\varepsilon(t))$  vanishes for almost every  $t \geq 0$ . It does so by proving a separation of time scales. The fast scale is the relaxation time of small clusters towards a local equilibrium, which can be understood as the response to the slower coarsening time of the large clusters. On the level of convergence of gradient flows, this is a step towards showing, that only the macroscopic part of the rescaled free energy  $\mathcal{F}^\varepsilon(n^\varepsilon(0))$  has to converge towards  $E(\nu_0)$  to ensure well prepared initial data. The conjecture is, that the microscopic part is automatically well prepared on the observed rescaled time-scale. This is consistent with the continuous dependence on the initial data of the LSW equation, which is valid under the assumption of convergence of the macroscopic energy for the initial data (see Cor. 3.8).

## 1.7. Outline

The next section contains in Section 2.1 the rescaling of the Becker–Döring gradient flow structure. This enables us to state the main results in Section 2.2. In Section 3, we prove the gradient flow structure of the LSW equation and prove the continuous dependence on the initial data within this framework. Section 4 contains some *a priori* estimate for the Becker–Döring system in Section 4.1, which allow then to the limit in the gradient structure in Section 4.2 and finally we prove the quasistationary equilibrium of the small clusters in Section 4.3. We conclude the paper with Appendix A showing that also more general discrete coagulation and fragmentation models fall into this framework. Moreover, Appendix B provides an elementary estimate.

## 2. MAIN RESULTS

### 2.1. Heuristics and scaling

From now, we consider the Becker–Döring system with initial total mass  $\varrho_0 > \varrho_s$  and rates satisfying Assumption 1.1. Moreover, the reference state for the free energy is given by  $\omega = \omega(z_s)$  as defined in (1.5).

We fix a scale  $\varepsilon^{-1}$  of the large cluster for some  $\varepsilon > 0$  and consider the first order expansion of the energy in  $\varepsilon$ . For some cut-off  $l_0$ , we introduce for  $l \geq l_0$  the rescaled variable  $\lambda = \varepsilon l$  and treat  $\lambda$  as continuous variable on  $\mathbb{R}_+$ .

We rescale the cluster density  $n_l$  by  $\varepsilon^2$  and define the empirical measure by

$$\nu^\varepsilon(d\lambda) := (I_{\text{mac}}^\varepsilon n)(d\lambda) := \varepsilon \sum_{l \geq l_0} \delta_{\varepsilon l}(d\lambda) \frac{n_l}{\varepsilon^2} = \frac{1}{\varepsilon} \sum_{l \geq l_0} \delta_{\varepsilon l}(d\lambda) n_l. \quad (2.1)$$

That is for each  $\zeta \in C_c^0(\mathbb{R})$  holds

$$\int_0^\infty \zeta(\lambda) \nu^\varepsilon(d\lambda) = \frac{1}{\varepsilon} \sum_{l \geq l_0} \zeta(\varepsilon l) n_l.$$

This scaling preserves the mass in the large cluster, which follows by approximating  $\zeta(\lambda) = \lambda$  with cut-off functions.

The leading order contribution of the free energy is given by the free energy of the large clusters  $l \geq l_0$ . This part of the free energy (1.7) can be expanded (*cf.* Lem. 4.2) as follows

$$\begin{aligned} \mathcal{F}(n) &\geq \sum_{l \geq l_0} \omega_l \psi\left(\frac{n_l}{\omega_l}\right) = \left( \frac{q}{z_s(1-\gamma)} \sum_{l \geq l_0} l^{1-\gamma} n_l \right) (1 + O(l_0^{-\sigma}) + O(l_0^\gamma \omega_{l_0})) \\ &= \frac{\varepsilon^\gamma q}{z_s(1-\gamma)} \int \lambda^{1-\gamma} d\nu^\varepsilon (1 + O(l_0^{-\sigma}) + O(l_0^\gamma \omega_{l_0})), \end{aligned} \quad (2.2)$$

for some  $\sigma > 0$ . To match the macroscopic energy (1.22), we define the rescaled free energy as

$$\mathcal{F}^\varepsilon(n) = \frac{z_s}{\varepsilon^\gamma} \mathcal{F}(n).$$

The rescaled free energy is of order 1, whenever the free energy is of order  $\varepsilon^\gamma$ . The main result of [2] states that the total free energy decreases to zero as  $t \rightarrow \infty$ . Hence, one possible way to obtain initial data  $n^\varepsilon(0)$  with  $\mathcal{F}^\varepsilon(n^\varepsilon(0)) = O(1)$  is to introduce a time  $t_\varepsilon$  such that  $\mathcal{F}(n(t_\varepsilon)) = O(\varepsilon^\gamma)$  and set  $n^\varepsilon(0) = n(t_\varepsilon)$ . In particular, this implies by the results of [21], that for  $\varepsilon$  small enough, all possible existing metastable states are already broken down.

The expansion (2.2) also shows, that the cut-off  $l_0$  has to satisfy two conditions (cf. (4.2) and (4.4))

$$\lim_{\varepsilon \rightarrow 0} l_0^\gamma \omega_{l_0} = 0 \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \max\{l_0^\gamma, l_0^\alpha\} \sqrt{\mathcal{F}(n^\varepsilon(0))} = 0.$$

By taking into account the asymptotic of  $\{Q_l\}_{l \geq 1}$  (cf. Lem. 4.1) and recalling  $\omega_l = z_s^l Q_l$ , the cut-off  $l_0$  can be chosen as

$$l_0 := \lfloor \varepsilon^{-x} \rfloor \quad \text{for some} \quad x \in (0, \tfrac{1}{2}). \quad (2.3)$$

We consider only states  $n$  such that free energy is of order  $\varepsilon^\gamma$ , that is we consider the restricted state space

$$\mathcal{M}^\varepsilon := \left\{ n \in \mathbb{R}_+^{\mathbb{N}} : \sum_{l \geq 1} l n_l = \varrho_0 \quad \text{and} \quad \mathcal{F}(n) \leq \varepsilon^\gamma \right\}.$$

Likewise, the differential of the free energy for states  $n^\varepsilon \in \mathcal{M}^\varepsilon$  will be of order  $\varepsilon^\gamma$  and hence covectors will be also on scale  $z_s^{-1} \varepsilon^\gamma$ , that is we define a rescaled vector field  $w^\varepsilon$  by

$$\nabla_l \phi = (e^{l+1} - e^l - e^1) \cdot \phi = \phi_{l+1} - \phi_l - \phi_1 =: z_s^{-1} \varepsilon^\gamma w^\varepsilon(\varepsilon l). \quad (2.4)$$

The rescaling of tangent vectors is then determined by the rescaling necessary for obtaining the macroscopic Onsager operator (1.23). This follows heuristically by expanding the Onsager matrix (1.14)

$$(\mathcal{K}(n)\phi)_l = -\varepsilon^{1-\alpha} \partial_\lambda^\varepsilon (\lambda^\alpha \varepsilon^\gamma w^\varepsilon \nu^\varepsilon) (1 + o(1)) \quad \text{with} \quad \partial_\lambda^\varepsilon f(\lambda) := \frac{f(\lambda + \varepsilon) - f(\lambda)}{\varepsilon}.$$

Hence, we define the rescaled Onsager operator by

$$(\mathcal{K}^\varepsilon(n)w^\varepsilon)(\varepsilon l) := \frac{1}{\varepsilon^{1-\alpha+\gamma}} (\mathcal{K}(n)\phi)(l),$$

where  $w^\varepsilon$  and  $\phi$  are given by the relation (2.4). This rescaling translates to the action  $\mathcal{A}(n, \phi)$  (1.18) and we define the rescaled action by

$$\mathcal{A}^\varepsilon(n, w^\varepsilon) := \frac{z_s}{\varepsilon^{1-\alpha+2\gamma}} \sum_{l \geq 1} k^l \hat{n}_l^\omega |\nabla_l \phi|^2. \quad (2.5)$$

Since, the dissipation is given as  $\mathcal{D}(n) := \mathcal{A}(n, -D\mathcal{F}(n))$ , the rescaling is the same and we define  $\mathcal{D}^\varepsilon(n) = z_s \varepsilon^{-(1-\alpha+2\gamma)} \mathcal{D}(n)$ . Hence the total rescaling between cotangent and tangent vectors is  $\varepsilon^{1-\alpha+\gamma}$ , which fixes the time scale for the macroscopic process.

Now, we introduce rescaled curves of finite action in analog to Definition 1.2. By abuse of notation the new time-scale  $t/\varepsilon^{1-\alpha+\gamma}$  is still denoted by  $t$ .

**Definition 2.1** (Rescaled curves of finite action). A weak solution  $[0, T] \ni t \mapsto (n^\varepsilon(t), w^\varepsilon(t))$  to the rescaled continuity equation

$$\int_0^T \left( \dot{\psi}(t) n_i^\varepsilon(t) - \psi(t) (\mathcal{K}^\varepsilon(n^\varepsilon(t)) w^\varepsilon(t))_i \right) dt = 0, \quad \text{for all } \psi \in C_c^1((0, T); \mathbb{R})$$

denoted by  $(n^\varepsilon, w^\varepsilon) \in \mathcal{CE}_T^\varepsilon$  is called a rescaled curve of finite action if

$$\sup_{t \in [0, T]} \mathcal{F}^\varepsilon(\nu_t^\varepsilon) < \infty, \quad \int_0^T \mathcal{A}^\varepsilon(n^\varepsilon(t), w^\varepsilon(t)) dt < \infty \quad \text{and} \quad \int_0^T \mathcal{D}^\varepsilon(n^\varepsilon(t)) dt < \infty.$$

Moreover, for such a curve we define the rescaled functional characterizing curves of maximal slope by

$$\mathcal{J}^\varepsilon(n^\varepsilon) := \mathcal{F}^\varepsilon(n^\varepsilon(T)) - \mathcal{F}^\varepsilon(n^\varepsilon(0)) + \frac{1}{2} \int_0^T \mathcal{D}^\varepsilon(n^\varepsilon(t)) dt + \frac{1}{2} \int_0^T \mathcal{A}^\varepsilon(n^\varepsilon(t), w^\varepsilon(t)) dt \geq 0. \quad (2.6)$$

In particular solutions such that  $\mathcal{J}^\varepsilon(n^\varepsilon) = 0$  satisfy the time-rescaled Becker–Döring equation

$$\dot{n}^\varepsilon(t) = -\varepsilon^{1-\alpha+\gamma} \mathcal{K}(n^\varepsilon(t)) D\mathcal{F}(n^\varepsilon(t)). \quad (2.7)$$

## 2.2. Convergence of the gradient structures

The functionals  $\mathcal{J}^\varepsilon$  (2.6) and  $J$  (1.25) are used to characterize solutions of the Becker–Döring and LSW equations in a variational way, respectively. The main idea to show convergence of the Becker–Döring equation to the LSW equation, which goes back to [23] (cf. [24]), is to prove  $\liminf_{\varepsilon \rightarrow 0} \mathcal{J}^\varepsilon(n^\varepsilon) \geq J(\nu)$  for curves of finite action  $n^\varepsilon$  converging to  $\nu$ . The lower semi-continuity estimate can be established by showing individual semi-continuity estimates for the energy, action and dissipation. This is the content of Theorem 2.2.

**Theorem 2.2** (Convergence of curves of finite action). *Suppose that  $\alpha \geq 1 - 3\gamma$ . For  $T > 0$  let  $(n^\varepsilon, w^\varepsilon) \in \mathcal{CE}_T^\varepsilon$  be a rescaled curve of finite action and  $\nu_0^\varepsilon := \Pi_{\text{mac}}^\varepsilon n^\varepsilon(0)$  with  $\Pi_{\text{mac}}^\varepsilon$  as defined in (2.1) satisfy*

$$\int_R^\infty \lambda \nu_0^\varepsilon(d\lambda) \rightarrow 0 \quad \text{as } R \rightarrow \infty \quad \text{uniformly in } \varepsilon. \quad (2.8)$$

Then, there exists a limiting curve  $t \mapsto (\nu_t, w_t) \in \text{CE}_T$  such that

$$\nu_t^\varepsilon := \Pi_{\text{mac}}^\varepsilon n^\varepsilon(t) \xrightarrow{*} \nu_t \quad \text{in } C_c^0(\mathbb{R}_+)^* \quad \text{for all } t \in [0, T] \quad (2.9)$$

and

$$w_t^\varepsilon(\lambda) \nu_t^\varepsilon(d\lambda) dt \xrightarrow{*} w_t(\lambda) \nu_t(d\lambda) dt \quad \text{in } C_c^0([0, T] \times \mathbb{R}_+)^*. \quad (2.10)$$

There exists  $u \in L^2((0, T))$  such that

$$h^\varepsilon(t) := \frac{n_1(t) - z_s}{\varepsilon^\gamma} \rightharpoonup u(t), \quad \text{weakly in } L^2((0, T)), \quad (2.11)$$

and  $u(t)$  satisfies the identity

$$u(t) = \frac{q \int \lambda^{\alpha-\gamma} \nu_t(d\lambda)}{\int \lambda^\alpha \nu_t(d\lambda)}.$$

Moreover, the energy, the action and the dissipation satisfy the following lim inf estimates

$$\forall t \in [0, T] : \quad \lim_{\varepsilon \rightarrow 0} \mathcal{F}^\varepsilon(\nu_t^\varepsilon) \geq E(\nu_t), \quad (2.12)$$

$$\liminf_{\varepsilon \rightarrow 0} \int_0^T \mathcal{A}^\varepsilon(\nu_t^\varepsilon, w_t^\varepsilon) dt \geq \int_0^T A(\nu_t, w_t) dt, \quad (2.13)$$

$$\liminf_{\varepsilon \rightarrow 0} \int_0^T \mathcal{D}^\varepsilon(\nu_t^\varepsilon) dt \geq \int_0^T D(\nu_t) dt. \quad (2.14)$$

The classical conclusion from the above theorem is the convergence of curves of maximal slope under the assumption of well-prepared initial data to deal with the term  $-\mathcal{F}^\varepsilon(n^\varepsilon(0))$  inside of  $\mathcal{J}^\varepsilon(n^\varepsilon)$ . The following Corollary is an immediate consequence of Theorem 2.2 by the arguments of Theorem 2 from [24].

**Corollary 2.3** (Convergence of curves of maximal slope). *Suppose  $\alpha \geq 1 - 3\gamma$  and let  $(n^\varepsilon, w^\varepsilon) \in \mathcal{CE}_T^\varepsilon$  be a curve of finite action. Moreover assume  $\nu_0^\varepsilon := \Pi_{\text{mac}}^\varepsilon n^\varepsilon(0)$  satisfy the tightness condition (2.8) and  $n^\varepsilon(0)$  is well-prepared in the sense that*

$$\lim_{\varepsilon \rightarrow 0} \mathcal{F}^\varepsilon(n^\varepsilon(0)) = E(\nu_0).$$

Then, there exists a limiting  $(\nu, w) \in \text{CE}_T$  satisfying (2.9) and (2.10) such that

$$\liminf_{\varepsilon \rightarrow 0} \mathcal{J}^\varepsilon(n^\varepsilon) \geq J(\nu) \geq 0.$$

Especially, if  $\mathcal{J}^\varepsilon(n^\varepsilon) = 0$  then  $J(\nu) = 0$  and it holds

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \mathcal{F}^\varepsilon(n^\varepsilon(t)) &= E(\nu_t) \quad \text{for all } t \in [0, T], \\ \mathcal{A}^\varepsilon(n^\varepsilon, w^\varepsilon) &\rightarrow A(\nu, w) \quad \text{for a.e. } t \in [0, T], \\ \mathcal{D}^\varepsilon(n^\varepsilon) &\rightarrow D(\nu) \quad \text{for a.e. } t \in [0, T]. \end{aligned}$$

### 2.3. Quasistationary evolution

The statement (2.11) connects the microscopic monomer concentration with a ratio of moments of the macroscopic cluster distribution. It is possible to show this identity already on the level of rescaled Becker–Döring equation alone. That is, the monomer concentration follows closely a moment ratio of the distribution of the large clusters.

**Proposition 2.4.** *For any curve  $(n^\varepsilon, \phi^\varepsilon) \in \mathcal{CE}_T^\varepsilon$  such that  $\mathcal{J}^\varepsilon(n^\varepsilon) < \infty$  uniformly in  $\varepsilon$  and  $\nu_0^\varepsilon$  satisfying (2.8) the rescaled monomer excess concentration  $h^\varepsilon$  as defined in (2.11) satisfies*

$$\int_0^T (h^\varepsilon(t) - u^\varepsilon(t))^2 dt \leq C \int_0^T \mathcal{D}_{\text{mac}}^\varepsilon(n(t)) dt, \quad (2.15)$$

where  $\mathcal{D}_{\text{mac}}^\varepsilon$  is defined like  $\mathcal{D}^\varepsilon$  with summation restricted to  $\{l_0, \dots, \infty\}$  and

$$u^\varepsilon(t) := \frac{\sum_{l \geq l_0} (b_{l+1} n_{l+1}(t) - a_l n_l)}{\varepsilon^\gamma \sum_{l \geq l_0} a_l n_l}.$$

The above results together with a refined energy-dissipation estimate based on a logarithmic Sobolev inequality allows to establish detailed information on the distribution of the small clusters for curves of rescaled finite action and in particular for every solution of the time-rescaled Becker–Döring equation (2.7). The result makes part of the formal asymptotic contained in Section 3 of [16] rigorous.

**Theorem 2.5** (Quasistationary distribution). *For any curve  $(n^\varepsilon, \phi^\varepsilon) \in \mathcal{CE}_T^\varepsilon$  such that  $\mathcal{J}^\varepsilon(n^\varepsilon) < \infty$  uniformly in  $\varepsilon$  and  $\nu_0^\varepsilon$  satisfying (2.8) the small cluster follow a quasistationary distribution dictated by  $n_1$ : For  $l_0 = \lfloor \varepsilon^{-x} \rfloor$  with  $x$  satisfying (2.3) holds*

$$\int_0^T \mathcal{H}_{\text{mic}}(n^\varepsilon(t) \mid \omega(n_1^\varepsilon(t))) \, dt \leq C \varepsilon^{\gamma+(1-x)(1-\alpha+\gamma)} \int_0^T \mathcal{D}_{\text{mic}}^\varepsilon(n_t^\varepsilon) \, dt,$$

where  $\omega_l(z) = z^l Q_l$  as defined in (1.5),  $\mathcal{D}_{\text{mic}}^\varepsilon$  is defined like  $\mathcal{D}^\varepsilon$  with summation restricted to  $\{1, \dots, l_0 - 1\}$  and  $\mathcal{H}_{\text{mic}}$  is the microscopic relative entropy defined by

$$\mathcal{H}_{\text{mic}}(n \mid \omega(z)) := \sum_{l=1}^{l_0-1} \omega_l(z) \psi\left(\frac{n_l}{\omega_l(z)}\right) \quad \text{with} \quad \psi(x) = x \log x - x + 1.$$

In particular, for a.e.  $t \in (0, T)$  it holds

$$\lim_{\varepsilon \rightarrow 0} \mathcal{F}_{\text{mic}}^\varepsilon(n^\varepsilon(t)) = 0 \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \mathcal{F}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon) = E(\nu_t), \quad (2.16)$$

where  $\mathcal{F}_{\text{mic}}^\varepsilon$  is defined like  $\mathcal{F}^\varepsilon$  with summation restricted to  $\{1, \dots, l_0 - 1\}$ .

**Remark 2.6.** The statement (2.16) is not enough to ensure well-prepared initial data, since the statement only holds for a.e.  $t \in [0, T]$ . However, it suggests that the statement of Corollary 2.3 holds already under the assumption of *macroscopically well-prepared* initial data:

$$\lim_{\varepsilon \rightarrow 0} E(\nu_0^\varepsilon) = E(\nu_0). \quad (2.17)$$

The assumption (2.17) together with the tightness condition (2.8) are natural, since they are also a sufficient condition for establishing continuous dependency on the initial data for the limiting gradient flow (cf. Cor. 3.8).

**Remark 2.7.** It is possible to use a different rescaling of the Becker–Döring system with different assumptions on the coagulation and fragmentation rates to obtain the LSW equation in the limit (cf. [5, 12]). Recently, within this scaling regime a quasi steady approximation was used to derive a suitable boundary condition for the macroscopic limits (cf. [8]).

### 3. THE LSW EQUATION AND ITS GRADIENT STRUCTURE

To make the formal calculation from Section 1.4 rigorous, we introduce the concept of curves of finite action for the LSW equation.

**Definition 3.1** (Curves of finite action). A weakly\* continuous curve  $[0, T] \ni t \mapsto \nu_t \in M$  is called a curve of finite action, if there exists a measurable vector field  $[0, T] \ni t \mapsto w_t \in T_{\nu_t}^* M$  such that

$$A(\nu, w) := \int_0^T \int \lambda^\alpha |w_t|^2 \nu_t(d\lambda) < \infty,$$

where the pair  $(\nu, w) \in \text{CE}_T$  solves the continuity equation

$$\partial_t \nu_t + \partial_\lambda (\lambda^\alpha w_t \nu_t) = 0 \quad \text{in} \quad C_c^\infty([0, T] \times \mathbb{R}_+)^*. \quad (3.1)$$

Before formulating the compactness statement, we want to revise the definition of the dissipation (1.24) and generalize it to curves of finite action. The dissipation acts as a weak upper gradient. Hence, for a curve of finite

action  $[0, T] \ni t \mapsto \nu_t \in M$  and using the fact that  $w_t \in T_{\nu_t}^* M$  for all  $t \in [0, T]$  it formally follows

$$\begin{aligned} |E(\nu_T) - E(\nu_0)| &= \left| - \int_0^T \int q \lambda^{\alpha-\gamma} w_t \, d\nu_t \, dt \right| \leq \int_0^T \lambda^\alpha \left| \int (u(t) - q \lambda^{-\gamma}) w_t \, d\nu_t \right| dt \\ &\leq \int_0^T \left( \int \lambda^\alpha (u(\nu_t) - q \lambda^{-\gamma})^2 \, d\nu_t \right)^{\frac{1}{2}} (A(\nu_t, w_t))^{\frac{1}{2}} dt, \end{aligned} \quad (3.2)$$

where  $u(\nu_t)$  is an arbitrary function on  $M$ . The choice of  $u(\nu_t)$  is fixed by a minimization in  $L^2$ . That is, we define the dissipation as the weighted  $L^2$ -minimal upper gradient for the energy. Before doing so, we need as an auxiliary result, that a finite dissipation implies the existence of the  $\alpha$ -moment for a curve of finite action.

**Lemma 3.2** (Moment estimate). *Assume  $\alpha \geq 1 - 3\gamma$ . Let  $(\nu, w) \in \text{CE}_T$  be a curve of finite action in  $M$  such that*

$$\inf_{u \in L^2([0, T])} \int_0^T \int \lambda^\alpha (u(t) - q \lambda^{-\gamma})^2 \, d\nu_t \, dt < \infty. \quad (3.3)$$

Then, it holds the moment estimate

$$\int_0^T \int \lambda^\alpha \, d\nu_t \, dt < \infty. \quad (3.4)$$

*Proof.* Let us define  $D(\nu, u) = \int \lambda^\alpha (u - q \lambda^{-\gamma})^2 \, d\nu$ . We observe that for  $\alpha \geq 1 - \gamma$ , there is nothing to show, since the bound follows by interpolation from  $\sup_{t \in [0, T]} E(\nu_t) < \infty$  and  $\int \lambda \, d\nu_t = \bar{\varrho}$ .

Therefore, assume now  $\alpha \leq 1 - \gamma$ . Let us define  $\eta(\lambda) := \lambda \chi_{[0, 1]}(\lambda) + \chi_{(1, \infty)}(\lambda)$ . Then, we can estimate with Cauchy–Schwarz for any  $\kappa \in \mathbb{R}$

$$\int_0^T \left( \int (u(t) - q \lambda^{-\gamma}) \eta(\lambda)^\kappa \, d\nu_t \right)^2 dt \leq \int_0^T D(\nu_t, u(t)) \int \eta(\lambda)^{2\kappa} \lambda^{-\alpha} \, d\nu_t \, dt \quad (3.5)$$

$$\leq \int_0^T D(\nu_t, u(t)) \, dt \sup_{t \in [0, T]} \int \eta(\lambda)^{2\kappa} \lambda^{-\alpha} \, d\nu_t. \quad (3.6)$$

Since,  $\sup_{t \in [0, T]} E(\nu_t) < \infty$  and  $\int \lambda \, d\nu_t = \bar{\varrho}$ , we can use interpolation to bound the sup in  $t$  provided  $2\kappa - \alpha \geq 1 - \gamma$ . On, the other hand, since  $\int \lambda \, d\nu_t = \bar{\varrho}$  for all  $t \geq 0$ , there exists a constant  $\bar{\varrho}_T > 0$  for any  $T > 0$  such that  $\int \eta(\lambda) \, d\nu_t \geq \bar{\varrho}_T$  (see also Lem. 4.7 for a similar argument). We can estimate the left hand side of (3.5) from below in the case  $\kappa = 1$  by using the Young inequality for some  $0 < \tau < 1$

$$\begin{aligned} \int_0^T \left( \int (u(t) - q \lambda^{-\gamma}) \eta(\lambda) \, d\nu_t \right)^2 dt &\geq (1 - \tau) \bar{\varrho}_T^2 \int_0^T u(t)^2 \, dt \\ &\quad - \left( \frac{1}{\tau} - 1 \right) \int_0^T ((1 - \gamma) E(\nu_t))^2 \, dt. \end{aligned}$$

Since,  $E(\nu_t) \in L^\infty([0, T])$ , we obtain the first *a priori* estimate

$$\int_0^T u(t)^2 \, dt \leq C_T \int_0^T D(\nu_t, u(t)) \, dt + C_T. \quad (3.7)$$

Another choice is  $\kappa = 1 - \gamma$  thanks to  $\alpha \leq 1 - \gamma$ . Then, we estimate the left hand side of (3.5) by using again the Young inequality with  $\tau \in (0, 1)$  as follows

$$\begin{aligned} \int_0^T \left( \int (u(t) - q\lambda^{-\gamma})\eta(\lambda)^{1-\gamma} d\nu_t \right)^2 dt &\geq (1-\tau)q \int \left( \int_0^T \lambda^{1-2\gamma} d\nu_t \right)^2 dt \\ &\quad - \left( \frac{1}{\tau} - 1 \right) \int_0^T u(t)^2 ((1-\gamma)E(\nu_t))^2 dt. \end{aligned}$$

Since, we trivially have  $\int_1^\infty \lambda^{1-2\gamma} d\nu_t \leq \int \lambda d\nu_t = \bar{\rho}$ , it follows by using the first *a priori* bound (3.7) and  $E(\nu_t) \in L^\infty([0, T])$  the second *a priori* estimate

$$\int_0^T \left( \int \lambda^{1-2\gamma} d\nu_t \right)^2 dt \leq C_T \int_0^T D(\nu_t, u(t)) dt + C_T, \quad (3.8)$$

which shows (3.4) for  $\alpha \geq 1 - 2\gamma$ . Hence, we assume now  $\alpha \leq 1 - 2\gamma$ . Similarly to (3.5), we can now estimate by Cauchy–Schwarz for some  $\tilde{\kappa} \in \mathbb{R}$

$$\int_0^T \left| \int (u(t) - q\lambda^{-\gamma})\eta(\lambda)^{\tilde{\kappa}} d\nu_t \right| dt \leq \left( \int_0^T D(\nu_t, u(t)) dt \right)^{\frac{1}{2}} \left( \int_0^T \int \eta(\lambda)^{2\tilde{\kappa}} \lambda^{-\alpha} d\nu_t dt \right)^{\frac{1}{2}}. \quad (3.9)$$

The second factor is bounded for  $2\tilde{\kappa} - \alpha \geq 1 - 2\gamma$  by (3.8). Hence, a possible choice is  $\tilde{\kappa} = 1 - 2\gamma$  by the assumption  $\alpha \leq 1 - 2\gamma$ . Since  $u \in L^2((0, T))$  and  $\int \lambda^{1-2\gamma} d\nu_t \in L^2((0, T))$ , we conclude the estimate (3.4).  $\square$

The Lemma provides the crucial ingredient to conclude that the dissipation is well-defined and justifies the use of the weak formulation in the first step of (3.2).

**Proposition 3.3.** *Assume  $\alpha \geq 1 - 3\gamma$ . Let  $(\nu, w) \in \text{CE}_T$  be a curve of finite action in  $M$  such that (3.3) holds. Then the associated minimization problem has a unique solution  $u \in L^2([0, T])$  such that*

$$\lambda \mapsto u(t) - q\lambda^{-\gamma} \in T_{\nu_t}^* M \quad \text{for a.e. } t \in [0, T]. \quad (3.10)$$

Moreover, the associated functional defined for a.e.  $t \in [0, T]$  by

$$D(\nu_t) := \int \lambda^\alpha (u(t) - q\lambda^{-\gamma})^2 d\nu_t \quad \text{with} \quad u(t) := \frac{q \int \lambda^{\alpha-\gamma} d\nu_t}{\int \lambda^\alpha d\nu_t}, \quad (3.11)$$

called dissipation, is a strong upper gradient for the energy  $E$ . That is, it holds for any curve  $(\nu, w) \in \text{CE}_T$  of finite action

$$|E(\nu_t) - E(\nu_s)| \leq \int_s^t \sqrt{D(\nu_r)} \sqrt{A(\nu_r, w_r)} dr, \quad \forall 0 \leq s < t \leq T. \quad (3.12)$$

Hereby, equality in (3.12) holds if and only if  $w_t(\lambda) = \pm(u(t) - q\lambda^{-\gamma})$  for  $\nu_t$ -a.e.  $\lambda \in \mathbb{R}_+$ .

*Proof.* In the first step, we show (3.10) and (3.11). Therefore, the first variation of the minimization problem (3.3) along some  $s : \mathbb{R}_+ \rightarrow \mathbb{R}$  is given by

$$\int_0^T \int (u(t) - q\lambda^{-\gamma})\lambda^\alpha d\nu_t s(t) dt = 0.$$

We show that is is well-defined by an estimate analog to (3.9)

$$\left| \int_0^T \lambda^\alpha (u(t) - q\lambda^{-\gamma}) \, d\nu_t \, s(t) \, dt \right| \leq \left( \int_0^T D(\nu_t, u(t)) \, dt \right)^{\frac{1}{2}} \left( \int_0^T s(t)^2 \int \lambda^\alpha \, d\nu_t \, dt \right)^{\frac{1}{2}},$$

which is bounded thanks to the estimate (3.4) for  $s \in L^\infty((0, T))$ . In addition the a priori estimate (3.7) shows that minimizer is actually in  $L^2((0, T))$  and hence satisfying the Euler-Lagrange equation  $\int (u(t) - q\lambda^{-\gamma}) \lambda^\alpha \, d\nu_t = 0$  for a.e.  $t \in [0, T]$ , which is nothing else than (3.10) also showing (3.11).

It is left to show, that  $D(\nu_t)$  is a strong upper gradient for the energy. Therefore, we fix a test function  $\zeta \in C_c^\infty(\mathbb{R}_+)$  and calculate for a curve  $(\nu, w) \in \text{CE}_T$

$$\frac{d}{dt} \frac{q}{1-\gamma} \int \lambda^{1-\gamma} \zeta \, d\nu_t = q \int \lambda^{\alpha-\gamma} \zeta \, w_t \, d\nu_t + \frac{q}{1-\gamma} \int \lambda^{1+\alpha-\gamma} \zeta' \, w_t \, d\nu_t =: \text{I} + \text{II}.$$

Using the fact that  $w_t \in T_{\nu_t}^* M$ , we can smuggle in  $u(t)$  and apply Cauchy-Schwarz to the first term I, to obtain

$$\text{I} \leq A(\nu_t, w_t)^{\frac{1}{2}} \left( \int \lambda^\alpha (u - q\lambda^{-\gamma} \zeta)^2 \, d\nu_t \right)^{\frac{1}{2}},$$

Hereby, equality holds if and only if  $w_t = \pm w_t^\zeta$  with  $w_t^\zeta := u - q\lambda^{-\gamma} \zeta$ . Hence, by choosing  $\zeta_n$  converging to 1 from below the result (3.12) follows by integration in time and dominated convergence, provided the term II vanishes. By an additional approximation step, we can justify to choose the sequence  $\zeta_n(\lambda) = n\lambda\chi_{[0, 1/n)} + \chi_{[1/n, \infty)}$  and estimate II by

$$\text{II} \leq \frac{1}{1-\gamma} \left( \int_0^{\frac{1}{n}} \lambda^\alpha |w_t|^2 \, d\nu_t \right)^{\frac{1}{2}} \left( \int_0^{\frac{1}{n}} \lambda^\alpha |u - q\lambda^{-\gamma} \lambda \zeta_n'|^2 \, d\nu_t \right)^{\frac{1}{2}}.$$

Since, we can assume the r.h.s. of (3.12) to be finite, we can conclude again by dominated convergence, that  $\text{II} \rightarrow 0$  as  $n \rightarrow \infty$ , which finishes the proof.  $\square$

**Lemma 3.4** (Tightness is preserved by curves of finite action). *Let  $\{\nu_0^\varepsilon\}_{\varepsilon>0}$  be a family satisfying the tightness condition (2.8). Then for any  $T > 0$  and any family of curves  $\{(\nu^\varepsilon, w^\varepsilon) \in \text{CE}_T : \nu_{t=0}^\varepsilon = \nu_0^\varepsilon\}_{\varepsilon>0}$  of uniformly finite action the family  $\{\nu_t^\varepsilon\}_{\varepsilon>0}$  satisfies the tightness condition (2.8) uniformly in  $\varepsilon$  for any  $t \in [0, T]$ .*

*Proof.* Fix a test function  $\eta_{r,R} \in C_c^\infty(\mathbb{R}_+, [0, 1])$  such that  $\eta_{r,R}(s) = 0$  for  $s < r/2$  and  $s > 2R$ ,  $\eta_{r,R}(s) = 1$  for  $r \leq s \leq R$ ,  $|\eta'_{r,R}(s)| \leq C/r$  for  $r/2 \leq s < r$  as well as  $|\eta'_{r,R}(s)| \leq C/R$  for  $R \leq s < 2R$ . We can estimate for a fixed curve of finite action  $(\nu, w) \in \text{CE}_T$

$$\begin{aligned} \left| \frac{d}{dt} \int \lambda \eta(\lambda) \, d\nu_t \right| &\leq \int \lambda^\alpha \eta |w_t| \, d\nu_t + \int \lambda^{1+\alpha} |\eta'| |w_t| \, d\nu_t \\ &\leq \frac{C}{r^{\frac{1-\alpha}{2}}} \int_{\frac{r}{2}}^\infty \lambda^{\frac{1+\alpha}{2}} |w_t| \, d\nu_t + \frac{C}{r^{\frac{1-\alpha}{2}}} \int_{\frac{r}{2}}^r \lambda^{\frac{1+\alpha}{2}} |w_t| \, d\nu_t \\ &\quad + \frac{C}{R^{\frac{1-\alpha}{2}}} \int_R^{2R} \lambda^{\frac{1+\alpha}{2}} |w_t| \, d\nu_t \\ &\leq C \left( \frac{1}{r^{\frac{1-\alpha}{2}}} + \frac{1}{R^{\frac{1-\alpha}{2}}} \right) A(\nu_t, w_t)^{\frac{1}{2}} \left( \int \lambda \, d\nu_t \right)^{\frac{1}{2}}. \end{aligned}$$

By an integration in time, letting  $R \rightarrow \infty$  and using the assumption of finite action, we obtain for all  $t \in [0, T]$  the estimate

$$\int_r^\infty \lambda \, d\nu_t \leq \int_{\frac{r}{2}}^\infty \lambda \, d\nu_0 + \frac{C\sqrt{T}}{r^{\frac{1-\alpha}{2}}}.$$

Hereby the constant  $C$  only depends on the test function and the action of the curve. Hence, if we apply this estimate for  $\{\nu_t^\varepsilon\}_{\varepsilon>0}$ , we observe its tightness by the tightness assumption on  $\{\nu_0^\varepsilon\}_{\varepsilon>0}$  and the uniform finite action of the family.  $\square$

**Proposition 3.5** (Compactness of curves of finite action). *Assume  $\alpha \geq 1 - 3\gamma$  and let  $(\nu^n, w^n) \in \text{CE}_T$  for  $n \in \mathbb{N}$  be a family of solutions to the continuity equation with uniformly bounded action and dissipation such that  $\{\nu_0^n\}_{n \in \mathbb{N}}$  satisfies the tightness condition (2.8). Then, there exists a subsequence and a couple  $(\nu, w) \in \text{CE}_T$ , such that*

$$\begin{aligned} \nu_t^n &\xrightarrow{*} \nu_t \text{ in } C_c^0(\mathbb{R}_+)^* \quad \forall t \in [0, T], \\ w^n \nu^n &\xrightarrow{*} w \nu \text{ in } C_c^0([0, T] \times \mathbb{R}_+)^*. \end{aligned} \quad (3.13)$$

In addition, the action and dissipation satisfy the lim inf estimates

$$\int_0^T A(\nu_t, w_t) \, dt \leq \liminf_{n \rightarrow \infty} \int_0^T A(\nu_t^n, w_t^n) \, dt \quad (3.14)$$

$$\int_0^T D(\nu_t) \, dt \leq \liminf_{n \rightarrow \infty} \int_0^T D(\nu_t^n) \, dt. \quad (3.15)$$

*Proof.* For any  $\zeta \in C_c^1(\mathbb{R}_+)$  and  $0 \leq t_1 < t_2 \leq T$  holds

$$\begin{aligned} \left| \int \zeta \, d\nu_{t_2}^n - \int \zeta \, d\nu_{t_1}^n \right| &= \left| \int_{t_1}^{t_2} \int \partial_\lambda \zeta \lambda^\alpha w_t^n \nu_t(d\lambda) \, dt \right| \\ &\leq \sup_{\lambda > 0} \frac{|\partial_\lambda \zeta|}{\lambda^{\frac{\alpha}{2}}} |t_2 - t_1|^{\frac{1}{2}} \left( \int_{t_1}^{t_2} A(\nu_t^n, w_t^n) \, dt \right)^{\frac{1}{2}}, \end{aligned}$$

which shows (3.13) and the weak\* continuity of  $\nu_t$ . Moreover, it holds for  $\kappa \in \mathbb{R}$  and  $\zeta \in C_c^0([0, T] \times \mathbb{R}_+)$

$$\int_0^T \int \zeta(t, \lambda) \lambda^{\kappa+\alpha} |w_t^n| \, d\nu_t^n \, dt \leq \left( \int_0^T A(\nu_t^n, w_t^n) \, dt \right)^{\frac{1}{2}} \left( \int_0^T \int \zeta(t, \lambda)^2 \lambda^{2\kappa+\alpha} \, d\nu_t \, dt \right)^{\frac{1}{2}}. \quad (3.16)$$

By lower semi-continuity it follows  $E(\nu_t) < \infty$  and by the tightness Lemma 3.4 it follows the conservation of total mass  $\int \lambda \, d\nu_t = \int \lambda \, d\nu_0 = \int \lambda \, d\nu_0^n < \infty$ . Hence,  $\nu_t \in M$  for all  $t \in [0, T]$ . Then, by interpolation, the second term in (3.16) is finite for  $1 - \gamma \leq 2\kappa + \alpha \leq 1$ . There exists  $\mu \in C_c^0([0, T] \times \mathbb{R}_+)$  such that  $w^n \nu^n \xrightarrow{*} \mu$  and the pair  $(\nu, \mu)$  satisfies  $\partial_t \nu_t + \partial_\lambda \mu_t = 0$  in  $C_c^\infty([0, T] \times \mathbb{R}_+)^*$ . Since  $(\nu^n, w^n)$  is a curve of finite action, we find a subsequence such that

$$\lim_{n \rightarrow \infty} \int_0^T A(\nu_t^n, w_t^n) \, dt = A^* := \liminf_{n \rightarrow \infty} \int_0^T \int_0^T A(\nu_t^n, w_t^n) \, dt.$$

Hence, we get the estimate with  $\kappa = (1 - \alpha)/2$

$$\int_0^T \int \zeta(t, \lambda) \lambda^{\frac{1-\alpha}{2}} \mu_t(d\lambda) dt \leq \left( A^* \int_0^T \int \zeta(t, \lambda)^2 \lambda d\nu_t dt \right)^{\frac{1}{2}}. \quad (3.17)$$

Now, we can apply the Riesz representation theorem to find  $v \in L^2(\lambda d\nu_t dt)$  such that

$$\int_0^T \int \zeta(t, \lambda) \lambda^{\frac{1-\alpha}{2}} \mu_t(d\lambda) dt = \int_0^T \int \zeta(t, \lambda) \lambda v(t, \lambda) \nu_t(d\lambda) dt.$$

Setting  $\tilde{\zeta}(t, \lambda) = \lambda^{\frac{1-\alpha}{2}} \zeta(t, \lambda)$  and  $w_t(\lambda) = \lambda^{\frac{1-\alpha}{2}} v(t, \lambda)$ , we get that  $\mu_t(d\lambda) = v(t, \lambda) \nu_t(d\lambda)$ . Moreover, since  $w \in L^2(\lambda^\alpha d\nu_t dt)$  it is of finite action. Moreover, by approximating  $\zeta(t, \lambda) = \frac{w_t(\lambda)}{\lambda^{\frac{1-\alpha}{2}}}$  it follows from (3.17) the lower semi-continuity of the action (3.14).

Finally, (3.15) follows by noting that  $D(\nu_t^n) = A(\nu_t^n, u(\nu_t^n) - q\lambda^{-\gamma})$ , which is well-defined by (3.10).  $\square$

The formulation of the LSW gradient flow as curves of minimal action, reads now in analog to the one of the Becker-Döring equation (1.20).

**Proposition 3.6** (LSW equation as curves of maximal slope). *Let  $\alpha \geq 1 - 3\gamma$ . For  $(\nu, w) \in \text{CE}_T$  with finite action holds*

$$J(\nu) := E(\nu_T) - E(\nu_0) + \frac{1}{2} \int_0^T D(\nu_t) dt + \frac{1}{2} \int_0^T A(\nu_t, w_t) dt \geq 0. \quad (3.18)$$

Moreover, equality holds if and only if  $\nu_t$  is a solution to the LSW equation.

*Proof.* We can assume that the dissipation  $\int_0^T D(\nu_t) dt$  is bounded, because else there is nothing to show. Then, we can use the strong upper gradient property of the dissipation (3.12) after an application of the Young inequality to arrive at

$$\frac{d}{dt} E(\nu_t) \geq -\frac{1}{2} D(\nu_t) - \frac{1}{2} A(\nu_t, w_t).$$

An integration of the above estimate shows the nonnegativity of  $J$  in (3.18). The equality case follows from the equality case in (3.12) for a.e.  $t \in [0, T]$  by choosing  $w_t(\lambda) = u(\nu_t) - q\lambda^{-\gamma}$ . Then, by weak\* continuity of  $t \mapsto \nu_t$  follows the result for all  $t \in [0, T]$ .

Now, for a curve  $(\nu, w) \in \text{CE}_T$  with  $J(\nu) = 0$  follows by Proposition (3.3) and (3.12) the identity

$$-\int \lambda^\alpha q \lambda^{-\gamma} w_t d\nu_t = \sqrt{A(\nu_t, w_t) D(\nu_t)} = A(\nu_t, w_t) = D(\nu, w_t).$$

Since  $w_t \in T_{\nu_t}^* M$ , it follows that  $w_t = u(t) - q\lambda^{-\gamma}$  for  $\nu_t$  almost every  $\lambda \in \mathbb{R}_+$ . Hence, the continuity equation (3.1) takes the form

$$\partial_t \nu_t + \partial_\lambda (\lambda^\alpha (u(t) - q\lambda^{-\gamma}) \nu_t) = 0 \quad \text{in} \quad C_c^\infty([0, T] \times \mathbb{R}_+)^*,$$

which is nothing else than a weak solution to the LSW equation.  $\square$

**Remark 3.7.** The compactness statement in Proposition 3.5 is also a tool to proof existence of solution to the LSW equation by the particle method (cf. [18, 19]). Therefore, the initial distribution is approximated in the weak\* sense by a discrete sum of Dirac deltas. Solutions for such data are determined by solving the finite system

of ordinary differential equations determined by (1.21) for each particle. Then the compactness statement allows to pass to the limit in the particle number and existence for measure valued initial distributions is obtained.

In addition, the compactness statement Proposition 3.5 with the variational characterization of solutions of the LSW equation from Proposition 3.6 is the essential tool to show the continuous dependence of the solution on the initial data.

**Corollary 3.8** (Continuous dependency on the initial data). *Let  $\{\nu_0^\varepsilon\}_{\varepsilon>0}$  be a sequence of initial data satisfying the tightness condition (2.8) and*

$$\lim_{\varepsilon \rightarrow 0} E(\nu_0^\varepsilon) = E(\nu_0) < \infty. \quad (3.19)$$

*Then there exists a solution  $\nu \in C_c^\infty([0, T] \times \mathbb{R}_+)^*$  to the LSW equation such that  $\nu_t^\varepsilon \xrightarrow{*} \nu_t$  in  $C_c^0(\mathbb{R}_+)$  for all  $t \in [0, T]$ .*

*Proof.* By the compactness statement Proposition (3.5) follows that there exists a couple  $(\nu_t, w_t) \in \text{CE}_T$  being the weak\* limit of  $(\nu_t^\varepsilon, w_t^\varepsilon) \in \text{CE}_T$  and satisfying the two lim inf estimates (3.14) and (3.15). By lower semi-continuity of the energy and the assumption (3.19) follows  $0 = \liminf_{\varepsilon \rightarrow 0} J(\nu^\varepsilon) \geq J(\nu) \geq 0$  and hence  $J(\nu) = 0$ , which proves the claim.  $\square$

**Remark 3.9.** The above result is consistent with the existing literature: In Theorem 2.2 of [19], the continuous dependency on the initial data was shown under the tightness condition (2.8) with respect to weak\* convergence for continuous test functions compactly supported on  $[0, \infty)$  including 0, i.e. Borel measures on  $[0, \infty)$ . Then, it is easy to see that weak\* convergence with respect to this class implies convergence of the macroscopic energy (2.17).

## 4. PROOF OF MAIN RESULTS

### 4.1. A priori estimates for the Becker–Döring gradient structure

In this section, we consider the Becker–Döring equation and its gradient structure as introduced in Sections 1.1 and 1.2, respectively.

The reversible equilibrium distribution  $\omega$  with parameter  $z \in (0, z_s]$  (corresponding to the conserved quantity) is given by (1.5). Note, that the radius of convergence for  $z \mapsto \sum_{l=1}^{\infty} l\omega_l(z)$  is  $z_s$  and  $\sum_{l=1}^{\infty} l\omega_l =: \varrho_s < \infty$  (cf. Lem. 4.1). Hence, the equilibrium state  $\omega_l := \omega_l(z_s)$  is the one with largest total mass  $\varrho_s$ . We work in the excess mass regime and any state will have total mass larger than  $\varrho_s$  to which there doesn't exist an according equilibrium state with the same total mass. The free energy is always the relative entropy with respect to  $\omega = \omega(z_s)$ , if not stated explicitly.

**Lemma 4.1.** *Under Assumption 1.1, there exists a constant  $\mathcal{F}_0$  such that for any  $l \geq 2$*

$$Q_l = \frac{1}{l^\alpha z_s^{l-1}} \exp\left(\left(\mathcal{F}_0 - \frac{q}{z_s}(1-\gamma)l^{1-\gamma} + \frac{q^2}{2z_s^2(1-2\gamma)}(l^{1-2\gamma} - 1)\right)(1 + O(l^{-\gamma}))\right), \quad (4.1)$$

where  $\frac{l^{1-2\gamma}-1}{1-2\gamma} := \log l$  for  $\gamma = \frac{1}{2}$ .

The proof relies on elementary estimates and is included for convenience in Appendix B. The expansion of the rates allows us to easily conclude the expansion of the free energy  $\mathcal{F}$ .

**Lemma 4.2** (Expansion of free energy). *Let  $n \in \mathcal{M}$  be given such that  $\mathcal{F}(n) < \infty$  as defined in (1.7), then there exists  $\sigma > 0$  such that for any  $l_0 \geq 2$*

$$\mathcal{F}_{l_0}(n) = \mathcal{F}_{l_0}^{\text{LSW}}(n)(1 + O(l_0^{-\sigma}) + O(l_0^\gamma \omega_{l_0})), \quad (4.2)$$

where  $\mathcal{F}_{l_0}$  and  $\mathcal{F}_{l_0}^{\text{LSW}}$  are defined by

$$\mathcal{F}_{l_0}(n) := \sum_{l=l_0}^{\infty} \omega_l \psi\left(\frac{n_l}{\omega_l}\right) \quad \text{and} \quad \mathcal{F}_{l_0}^{\text{LSW}}(n) := \frac{q}{z_s(1-\gamma)} \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l.$$

*Proof.* We expand the function  $\psi$  in the definition of  $\mathcal{F}_{l_0}$

$$\mathcal{F}_{l_0}(n) = \sum_{l=l_0}^{\infty} \left( n_l \log \frac{1}{z_s^l Q_l} + n_l (\log n_l - 1) + \omega_l \right)$$

We estimate the first sum using the asymptotic expansion (4.1)

$$\begin{aligned} \sum_{l=l_0}^{\infty} \left( n_l \log \frac{1}{z_s^l Q_l} \right) &= \frac{q}{z_s(1-\gamma)} \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l + O\left( \sum_{l=l_0}^{\infty} l^{1-2\gamma} n_l \right) \\ &\quad + \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l \frac{\log l^\alpha}{l^{1-\gamma}} - \sum_{l=l_0}^{\infty} \frac{l^{1-\gamma}}{l^{1-\gamma}} n_l (\log z_s + \mathcal{F}_0(1 + O(l^{-\gamma}))) \\ &= \left( \frac{q}{z_s(1-\gamma)} \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l \right) \left( 1 + O(l_0^{-\gamma} \log l_0) + O(l_0^{-(1-\gamma)}) \right), \end{aligned}$$

Likewise, we note that for any  $\beta \in (0, 1)$  exists  $C_\beta > 0$  such that for  $x > 0$

$$|\min\{x(\log x - 1), 0\}| \leq C_\beta x^\beta$$

and with the Hölder inequality, we can estimate

$$\sum_{l=l_0}^{\infty} |\min\{n_l(\log n_l - 1), 0\}| \leq C_\beta \sum_{l=l_0}^{\infty} n_l^\beta \leq C_\beta \left( \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l \right)^\beta \left( \sum_{l=l_0}^{\infty} \frac{1}{l^{\frac{\beta}{1-\beta}(1-\gamma)}} \right)^{1-\beta}.$$

Now, we can choose  $\beta$  such that  $\frac{\beta}{1-\beta}(1-\gamma) = \kappa > 1$  and  $\beta < 1$  leading to the estimate

$$\sum_{l=l_0}^{\infty} |\min\{n_l(\log n_l - 1), 0\}| \leq C_\beta \left( \beta \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l + 1 - \beta \right) O(l_0^{-(\kappa-1)(1-\beta)}).$$

The last term evaluates with the help of (4.1) to

$$\begin{aligned} \sum_{l=l_0}^{\infty} \omega_l &\leq \sum_{l=l_0}^{\infty} \frac{z_s}{l^\alpha} \exp\left( \left( \mathcal{F}_0 - \frac{q}{z_s(1-\gamma)} l^{1-\gamma} \right) (1 + O(l^{-\gamma})) \right) \\ &\leq C \int_{l_0}^{\infty} \exp\left( -\frac{q}{z_s(1-\gamma)} l^{1-\gamma} \right) dl \leq C l_0^\gamma \exp\left( -\frac{q}{z_s(1-\gamma)} l_0^{1-\gamma} \right). \end{aligned}$$

Therefore, a combination of all the estimates leads to the result.  $\square$

Moreover, we need a Czar-Pinsker inequality for the free energy, which was already a crucial ingredient in [16]

**Proposition 4.3** (Czar-Pinsker inequality ([16], Lems. 2.1, 2.2)). *For  $n \in \mathcal{M}$  and any small  $\eta > 0$  and any  $p < \infty$  and any  $l_0 \geq 2$  holds*

$$\sum_{l=1}^{\infty} l^{1-\gamma} |n_l - \omega_l| \leq C \sqrt{\mathcal{F}(n)} \quad (4.3)$$

$$\left| \sum_{l=l_0}^{\infty} l n_l - (\varrho - \varrho_s) \right| \leq C l_0^\gamma \sqrt{\mathcal{F}(n)} + C_p l_0^{-p}. \quad (4.4)$$

For the next Lemmata, we make statements on curves of finite action to deduce certain compactness, which we later need for passing to the limit. These Lemmata are the analog of Lemmas 2.3 and 2.4 from [16], but we proof them for curves of finite action instead of solutions to the Becker–Döring equation.

**Lemma 4.4** (A priori estimates for curves of finite action). *Let  $(n, \phi) \in \mathcal{CE}_T$  be a curve of finite action as in Definition 1.2 and  $\eta \in L^2(0, T)$ , then it holds*

$$\int_0^T \eta(t) \sum_{l=l_0}^{\infty} l^\kappa k^l \hat{n}_l^\omega(t) |\nabla_l \phi(t)| dt \leq C \left( \sup_{t \in [0, T]} \mathcal{F}_{l_0}^{\text{LSW}}(n(t)) \right)^{\frac{1-\alpha-2\kappa}{2\gamma}} \int_0^T |\eta(t)| \sqrt{\mathcal{A}_{\text{mac}}(n(t), \phi(t))} dt, \quad (4.5)$$

for any  $\kappa \in \left[ \frac{1-\alpha-\gamma}{2}, \frac{1-\alpha}{2} \right]$  with  $\mathcal{A}_{\text{mac}}$  the action as defined in (1.18) restricted to  $l \geq l_0$ . Hereby,  $\nabla_l \phi := \phi_{l+1} - \phi_l - \phi_1$  and  $\hat{n}_l^\omega$  is defined in (1.19). Moreover, it also holds the estimate

$$\int_0^T \eta(t) \sum_{l=l_0}^{\infty} l^\kappa |a_l n_1(t) n_l(t) - b_{l+1} n_{l+1}(t)| \leq C \left( \sup_{t \in [0, T]} \mathcal{F}_{l_0}^{\text{LSW}}(n(t)) \right)^{\frac{1-\alpha-2\kappa}{2\gamma}} \int_0^T |\eta(t)| \sqrt{\mathcal{D}_{\text{mac}}(n(t))} dt, \quad (4.6)$$

where again  $\mathcal{D}_{\text{mac}}$  is defined as in (1.8) restricted to  $l \geq l_0$ .

*Proof.* We estimate using the Cauchy–Schwarz inequality

$$\sum_{l=l_0}^{\infty} l^\kappa k^l \hat{n}_l^\omega |\nabla_l \phi(t)| \leq \left( \sum_{l=l_0}^{\infty} k^l \hat{n}_l^\omega |\nabla_l \phi(t)|^2 \right)^{\frac{1}{2}} \left( \sum_{l=l_0}^{\infty} l^{2\kappa} k^l \hat{n}_l^\omega \right)^{\frac{1}{2}}.$$

Now, using that fact that

$$k^l \hat{n}_l^\omega = \Lambda(a_l n_1 n_l, b_{l+1} n_{l+1}) = \Lambda(\lambda^\alpha n_1 n_l, (\lambda + 1)^\alpha (z_s + q(l+1)^{-\gamma}) n_{l+1}),$$

the estimate  $\frac{(\lambda+1)^\alpha}{\lambda^\alpha} \leq 1 + \frac{\alpha}{\lambda}$  and from (4.3) the bound  $|n_1 - z_s| \leq C \sqrt{\mathcal{F}(n)}$ , it follows

$$\sum_{l=l_0}^{\infty} l^{2\kappa} k^l \hat{n}_l^\omega \leq 2 \left( z_s + \max \left\{ C \sqrt{\mathcal{F}(n)}, q l_0^{-\gamma} \right\} \right) \sum_{l=l_0}^{\infty} l^{\alpha+2\kappa} n_l.$$

Now, we use the Hölder inequality to interpolate

$$\sum_{l=l_0}^{\infty} l^{\alpha+2\kappa} n_l \leq \left( \sum_{l=l_0}^{\infty} l n_l \right)^{\frac{\alpha+2\kappa+\gamma-1}{\gamma}} \left( \sum_{l=l_0}^{\infty} l^{1-\gamma} n_l \right)^{\frac{1-\alpha-2\kappa}{\gamma}} \leq C \left( \mathcal{F}_{l_0}^{\text{LSW}}(n(t)) \right)^{\frac{1-\alpha-2\kappa}{\gamma}}$$

by assuming  $1 - \alpha - \gamma \leq 2\kappa \leq 1 - \alpha$ . The estimate (4.6) follows from (4.5) by noting that with the choice  $\phi^*(t) = D\mathcal{F}(n(t)) = \left(\log \frac{n_l(t)}{\omega_l}\right)_{l \geq 1}$  holds

$$k^l \hat{n}_l^\omega(t) |\nabla_l \phi^*(t)| = a_l n_1(t) n_l(t) - b_{l+1} n_{l+1}(t)$$

and  $\mathcal{A}_{\text{mac}}(n(t), \phi^*(t)) = \mathcal{D}_{\text{mac}}(n(t))$ .  $\square$

The last *a priori* estimate deals with tightness and how tightness is preserved for curves of finite action.

**Lemma 4.5** (Tightness is preserved for curves of finite action). *A family  $N \subset \mathcal{M}$  is called tight provided that*

$$\sup_{n \in N} \sum_{l=R}^{\infty} l n_l \rightarrow 0 \quad \text{as } R \rightarrow \infty. \quad (4.7)$$

*If the family  $N_0 \subset \mathcal{M}$  satisfy the tightness condition (4.7). Then for any  $T > 0$  and any family of curves  $\{(n, \phi) \in \mathcal{CE}_T : n(0) \in N_0\}$  of uniformly finite action the family  $\{n(t)\}$  also satisfies the tightness condition (4.7) for  $t \in [0, T]$ .*

*Proof.* The proof is similar to Lemma 3.4, where the same result is proven for the LSW gradient structure. Let  $1 \ll M_1 \ll M_2$  and let  $\eta \in C^1(\mathbb{R})$  be a cut off function such that  $\eta(l) = 0$  for  $l \leq \frac{M_1}{2}$  and  $l \geq 2M_2$ ,  $\eta(l) = 1$  for  $M_1 \leq l \leq M_2$  and such that  $\eta'(l) \leq \frac{C}{M_1}$  for  $\frac{M_1}{2} \leq l \leq M_1$  and  $|\eta'(l)| \leq \frac{C}{M_2}$  for  $M_2 \leq l \leq 2M_2$ . Moreover, we define  $\mathcal{N}_l := l\eta(l)$  and assume  $M_1 > 2$  such that  $\eta(1) = 0$ . Then, it follows for any curve of finite action  $(n, \phi) \in \Phi$

$$\begin{aligned} \frac{d}{dt} \mathcal{N} \cdot n(t) &= \mathcal{N} \cdot \partial_t n(t) = \mathcal{N} \cdot \mathcal{K}(n) \phi \\ &= \left( \sum_{l=1}^{\infty} \eta(l+1) k^l \hat{n}_l^\omega(t) \nabla_l \phi(t) + \sum_{l=1}^{\infty} k^l \hat{n}_l^\omega(t) l \nabla_l \eta \nabla_l \phi(t) \right) \\ &\leq C \left( \frac{1}{M_1^{\frac{1-\alpha}{2}}} \sum_{l=M_1/2}^{\infty} l^{\frac{1-\alpha}{2}} k^l \hat{n}_l^\omega(t) |\nabla_l \phi(t)| \right. \\ &\quad \left. + \left( \frac{1}{M_1} M_1^{1-\frac{1-\alpha}{2}} + \frac{1}{M_2} M_2^{1-\frac{1-\alpha}{2}} \right) \sum_{l=M_1/2}^{\infty} l^{\frac{1-\alpha}{2}} k^l \hat{n}_l^\omega(t) |\nabla_l \phi(t)| \right) \\ &\leq C \left( \frac{1}{M_1^{\frac{1-\alpha}{2}}} + \frac{1}{M_2^{\frac{1-\alpha}{2}}} \right) \sum_{l=M_1/2}^{\infty} l^{\frac{1-\alpha}{2}} k^l \hat{n}_l^\omega(t) |\nabla_l \phi(t)|, \end{aligned}$$

where  $C$  is the constant depending only on the cut off function  $\eta$ . Integrating over time and using (4.5) leads to

$$\sum_{l=M_1}^{M_2} l n_l(t) \leq \sum_{l=M_1/2}^{2M_2} l n_l(0) + C \left( \frac{1}{M_1^{\frac{1-\alpha}{2}}} + \frac{1}{M_2^{\frac{1-\alpha}{2}}} \right) \int_0^t \sqrt{\mathcal{A}(n(t), \phi(t))} dt.$$

Now, using the fact that  $t \mapsto n(t)$  is a curve of finite action and letting  $M_2 \rightarrow \infty$ , we obtain

$$\sum_{l=M_1}^{\infty} l n_l(t) \leq \sum_{l=M_1/2}^{\infty} l n_l(0) + \frac{Ct^{\frac{1}{2}}}{M_1^{\frac{1-\alpha}{2}}},$$

where the constant  $C$  is uniform for the family. This finishes the proof since  $N_0$  satisfies the tightness condition (4.7).  $\square$

#### 4.2. Passage to the limit: Proof of Theorem 2.2

To pass to the limit in the discrete continuity equation, we define the flux density measure for a fixed covector  $\phi$  and rescaled one  $w^\varepsilon(\varepsilon l) = z_s \varepsilon^{-\gamma} \nabla_l \phi$  (cf. (2.4)) by

$$\begin{aligned} \mu^\varepsilon(d\lambda) &:= \frac{z_s}{\varepsilon^{1-\alpha+2\gamma}} \sum_{l \geq l_0} \delta_{\varepsilon l}(d\lambda) k^l \widehat{n}_l^{\varepsilon\omega} \nabla_l \phi \\ &= \frac{1}{\varepsilon^{1-\alpha+\gamma}} \sum_{l \geq l_0} \delta_{\varepsilon l}(d\lambda) l^\alpha \Lambda(n_1^\varepsilon n_l^\varepsilon, (z_s + q(l+1)^{-\gamma}) n_{l+1}^\varepsilon) w^\varepsilon(\lambda). \end{aligned} \quad (4.8)$$

and the dissipation flux density measure

$$\hat{\mu}^\varepsilon(d\lambda) := \frac{1}{\varepsilon^{1-\alpha+\gamma}} \sum_{l \geq l_0} \delta_{\varepsilon l}(d\lambda) (a_l n_1(t) n_l(t) - b_{l+1} n_{l+1}(t)) \quad (4.9)$$

Let us note, that with the above definitions for  $l \geq l_0$  and  $\lambda = \varepsilon l$  holds

$$\dot{n}_l^\varepsilon(t) - \frac{1}{\varepsilon^{1-\alpha+\gamma}} (\mathcal{K}[n]\phi)_l = \partial_t \nu_t^\varepsilon(\lambda) + \partial_\lambda^\varepsilon \mu_t^\varepsilon(\lambda) = 0, \quad (4.10)$$

where

$$\partial_\lambda^\varepsilon \mu_t^\varepsilon(\lambda) := \frac{\mu_t^\varepsilon(\lambda + \varepsilon) - \mu_t^\varepsilon(\lambda)}{\varepsilon}.$$

Let us summarize the *a priori* estimates found in Section 4.1 and rewrite them in rescaled variables. We denote with  $\mathcal{F}_{\text{mac}}^\varepsilon(\Pi_{\text{mac}}^\varepsilon n) = \varepsilon^{-\gamma} \mathcal{F}_{\text{mac}}(n)$  and similarly for  $\mathcal{A}_{\text{mac}}^\varepsilon$  as well as  $\mathcal{D}_{\text{mac}}^\varepsilon$ .

**Proposition 4.6** (Rescaled *a priori* estimates). *With  $x$  from (2.3) holds*

(i) *The rescaled free energy satisfies*

$$C \geq \mathcal{F}^\varepsilon(n^\varepsilon) \geq \mathcal{F}_{\text{mac}}^\varepsilon(\nu^\varepsilon) = E(\nu^\varepsilon)(1 + O(\varepsilon^{x\sigma})) \quad (4.11)$$

(ii) *The total excess mass satisfies*

$$\left| \int \lambda \nu^\varepsilon(d\lambda) - (\varrho_0 - \varrho_s) \right| \leq C \varepsilon^{\gamma(\frac{1}{2}-x)} \sqrt{\mathcal{F}_{\text{mac}}^\varepsilon(\nu^\varepsilon)}. \quad (4.12)$$

(iii) *Let  $[0, T] \ni t \mapsto \nu_t^\varepsilon \in \mathcal{M}^\varepsilon$  be a rescaled curve of finite action and  $\eta \in L^2((0, T))$ , then for any  $\kappa \in [\frac{1-\alpha-\gamma}{2}, \frac{1-\alpha}{2}]$*

$$\int_0^T \eta(t) \int \lambda^\kappa |\mu_t^\varepsilon(d\lambda)| dt \leq C \left( \sup_{t \in [0, T]} \mathcal{F}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon) \right)^{\frac{1-\alpha-2\kappa}{2\gamma}} \int_0^T |\eta(t)| \sqrt{\mathcal{A}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon, w_t^\varepsilon)} dt, \quad (4.13)$$

$$\int_0^T \eta(t) \int \lambda^\kappa |\hat{\mu}_t^\varepsilon(d\lambda)| dt \leq C \left( \sup_{t \in [0, T]} \mathcal{F}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon) \right)^{\frac{1-\alpha-2\kappa}{2\gamma}} \int_0^T |\eta(t)| \sqrt{\mathcal{D}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon)} dt. \quad (4.14)$$

(iv) If  $\{\nu_0^\varepsilon\}_{\varepsilon>0}$  satisfies the tightness condition (2.8) and  $[0, T] \ni t \mapsto \nu_t^\varepsilon \in \mathcal{M}^\varepsilon$  are rescaled curves of finite action, then for all  $t \in [0, T]$  also  $\{\nu_t^\varepsilon\}_{\varepsilon>0}$  satisfies the tightness condition (2.8).

The above results enable us to conclude the lim inf estimates and proof Theorem 2.2.

*Proof of Theorem 2.2.*

**Step 1:** *Convergence of  $\nu^\varepsilon$ .* For  $\zeta \in C_c^1(\mathbb{R}_+)$  and  $0 \leq t_1 < t_2 \leq T$ , we calculate using the discrete continuity equation in the form (4.10)

$$\begin{aligned} \left| \int \zeta d\nu_{t_1}^\varepsilon - \int \zeta d\nu_{t_2}^\varepsilon \right| &= \left| \int_{t_1}^{t_2} \int \partial_\lambda^\varepsilon \zeta(\lambda) \mu_t^\varepsilon(d\lambda) dt \right| \\ &\leq \sup_{\lambda \in \mathbb{R}_+} \frac{|\partial_\lambda^\varepsilon \zeta(\lambda)|}{\lambda^{\frac{1-\alpha}{2}}} \int_{t_1}^{t_2} \int \lambda^{\frac{1-\alpha}{2}} |\mu_t^\varepsilon(d\lambda)| dt \\ &\stackrel{(4.13)}{\leq} C \sup_{\lambda \in \mathbb{R}_+} \frac{|\partial_\lambda^\varepsilon \zeta(\lambda)|}{\lambda^{\frac{1-\alpha}{2}}} \int_{t_1}^{t_2} \sqrt{\mathcal{A}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon, w_t^\varepsilon)} dt \\ &\leq C \sup_{\lambda \in \mathbb{R}_+} \frac{|\zeta'(\lambda)|}{\lambda^{\frac{1-\alpha}{2}}} \sqrt{|t_1 - t_2|}. \end{aligned}$$

This estimate together with the bound (4.11) imply *via* Arzelà-Ascoli the weak\* convergence towards a weakly\* continuous map  $t \mapsto \nu_t$ . Moreover, the *a priori* bounds (4.11), (4.12) and tightness condition (2.8) imply that  $\int \zeta d\nu^\varepsilon \rightarrow \int \zeta d\nu$  holds for  $\zeta \in C^0(\mathbb{R}_+)$  satisfying

$$\limsup_{\lambda \rightarrow \infty} \frac{|\zeta(\lambda)|}{\lambda} < \infty \quad \text{and} \quad \lim_{\lambda \rightarrow 0} \frac{|\zeta(\lambda)|}{\lambda^{1-\gamma}} = 0,$$

which implies that the excess mass is preserved

$$\int \lambda d\nu_t = \varrho_0 - \varrho_s \quad \Rightarrow \quad \nu_t \in M, \quad \text{for all } t \in [0, T].$$

Moreover, the bounds (4.11) and (4.12) also imply by weak lower semi-continuity the estimate (2.12) and especially that  $\sup_{t \in [0, T]} E(\nu_t) < \infty$ .

**Step 2:** *Convergence of  $\mu^\varepsilon$ .* The *a priori* estimate (4.13) implies the existence of a measure  $\mu \in C_c^0([0, T] \times \mathbb{R}_+)^*$  such that up to subsequences

$$\int \int \zeta(t, \lambda) d\mu_t^\varepsilon dt \rightarrow \int \int \zeta(t, \lambda) d\mu \quad \text{for all } \zeta \in C_c^0([0, T] \times \mathbb{R}_+). \quad (4.15)$$

Now, we show the limiting measure is of the form  $\mu(dt, d\lambda) = \lambda^\alpha w_t(\lambda) \nu_t(d\lambda) dt$  for some vector field  $w_t$  with finite action. Therefore, we remind at the definition of  $\mu^\varepsilon$  (4.8) and  $\mathcal{A}_{\text{mac}}^\varepsilon$  (2.5) to estimate

$$\int \int \zeta(t, \lambda) \lambda^{\frac{1-\alpha}{2}} \mu_t^\varepsilon(d\lambda) dt \leq \left( \int_0^T \mathcal{A}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon, w_t^\varepsilon) dt \right)^{\frac{1}{2}} \left( \int_0^T \frac{1}{z_s} \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l^{1-\alpha} k^l \widehat{n}_l^\varepsilon(t) dt \right)^{\frac{1}{2}}. \quad (4.16)$$

The second term on the right hand side can be bounded by using the one-homogeneity and concavity of  $(a, b) \mapsto \Lambda(a, b)$

$$\begin{aligned}
& \frac{1}{z_s} \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l^{1-\alpha} k^l \widehat{n}_l^\varepsilon(t) \\
& \leq \frac{1}{z_s} \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l \Lambda(n_1^\varepsilon(t) n_l^\varepsilon(t), (z_s + q(l+1)^{-\gamma}) n_{l+1}^\varepsilon(t)) \\
& \leq \frac{1}{z_s} \Lambda \left( n_1^\varepsilon(t) \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l n_l^\varepsilon(t), (z_s + q l_0^{-\gamma}) \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l n_{l+1}^\varepsilon(t) \right) \\
& \leq \frac{z_s + o(1)}{z_s} \sum_{l \geq l_0} \zeta^2(t, \varepsilon l) l n_l^\varepsilon(t),
\end{aligned}$$

where we used in the last estimate that  $l_0 = \varepsilon^{-x}$ ,  $|n_1^\varepsilon - z_s| \leq C \sqrt{\mathcal{F}(n)} \leq C \varepsilon^{\frac{x}{2}}$  by (4.3) and the fact that  $\zeta$  is uniformly continuous. Since  $t \mapsto n^\varepsilon(t)$  is a curve of finite action and by the convergence of the total mass, it follows that the right hand side of (4.16) is finite. Hence, we can pass to the limit  $\varepsilon \rightarrow 0$  in (4.16) by the same argument as in (4.15). It follows for a subsequence which attains

$$\int_0^T \mathcal{A}_{\text{mac}}^\varepsilon(\nu^\varepsilon, w^\varepsilon) dt \rightarrow A^* := \liminf_{\varepsilon \rightarrow 0} \int_0^T \mathcal{A}_{\text{mac}}^\varepsilon(\nu^\varepsilon, w^\varepsilon) dt.$$

the estimate

$$\int \int \zeta(t, \lambda) \lambda^{\frac{1-\alpha}{2}} \mu_t(d\lambda) dt \leq \left( A^* \int \int \zeta^2(t, \lambda) \lambda \nu_t(d\lambda) dt \right)^{\frac{1}{2}},$$

with  $\mu_t(d\lambda)$  denoting the disintegration of  $\mu$  in  $t$ . Hence, we can conclude as in the derivation of (3.17) to find  $v \in L^2(\lambda d\nu_t dt)$  by the Riesz representation theorem showing lower semi-continuity of the action (2.13).

**Step 3: Convergence of the dissipation  $\mathcal{D}^\varepsilon$ .** We observe that  $\mathcal{D}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon) = \mathcal{A}_{\text{mac}}^\varepsilon(\nu_t^\varepsilon, \tilde{w}^\varepsilon)$ , where  $\tilde{w}^\varepsilon$  is the special vector field given by  $-\nabla_l D\mathcal{F}^\varepsilon(\nu_t^\varepsilon)$ , i.e. for all  $l$

$$\varepsilon^\gamma \tilde{w}_t^\varepsilon(\varepsilon l) = \log \frac{n_1^\varepsilon n_l^\varepsilon}{\omega_1 \omega_l} - \log \frac{n_{l+1}^\varepsilon}{\omega_{l+1}}.$$

Therefore, we can apply the same arguments of step 2, but now to the dissipation flux density  $\hat{\mu}^\varepsilon$  defined in (4.9) and use the *a priori* estimate (4.14) to deduce the lim inf estimate

$$\liminf_{\varepsilon \rightarrow 0} \int_0^T \mathcal{D}^\varepsilon(\nu_t^\varepsilon) dt \geq \int_0^T A(\nu_t, \tilde{w}_t) dt = \int_0^T \int \lambda^\alpha |\tilde{w}_t|^2 \nu_t(d\lambda) dt,$$

for some  $\tilde{w} \in C_c^0([0, T] \times \mathbb{R}_+)^*$ . It is left to show that  $h^\varepsilon(t) \rightharpoonup h(t)$  in  $L^2((0, T))$  and  $\tilde{w}_t$  is of the form  $h(t) - q/\lambda^\gamma$  with  $h \in L^2([0, T])$ , however this statement follows exactly along the lines of Lemma 2.6 from [16]. The final result (2.14) follows now by the definition of  $D(\nu_t)$  as the infimum over all such  $h \in L^2([0, T])$  from Lemma 3.3.

**Step 4: Continuity equation holds.** Finally, choosing a subsequence such that both convergences (2.9) and (2.10) holds for a test function  $\zeta \in C_c^\infty([0, T] \times \mathbb{R})$ , we can pass to the limit in the weak form of the discrete continuity

equation (4.10)

$$\begin{aligned} \int_0^T \int \partial_t \zeta(t, \lambda) \nu_t^\varepsilon(d\lambda) dt + \int \int \partial_\lambda^\varepsilon \zeta(t, \lambda) \mu_t^\varepsilon(d\lambda) dt &= 0 \\ \downarrow \varepsilon \rightarrow 0 & \quad \downarrow \varepsilon \rightarrow 0 \\ \int_0^T \int \partial_t \zeta(t, \lambda) \nu_t(d\lambda) dt + \int \int \partial_\lambda \zeta(t, \lambda) \mu_t(d\lambda) dt &= 0, \end{aligned}$$

which shows  $(\nu, w) \in \text{CE}_T$ . □

### 4.3. Quasistationary expansion: Proof of Theorem 2.5

The proofs of Proposition 2.4 and Theorem 2.5 consists in several steps, which are formulated in the following Lemmata. In the proofs of this section,  $C$  is a generic constant, which is assumed to be independent of  $\varepsilon$  and only depending on the parameters inside of the rates from Assumption 1.1.

**Lemma 4.7.** *Assume that  $\mathcal{F}_{\text{mac}}^\varepsilon(\nu^\varepsilon) \leq C$ ,  $l_0$  satisfies (2.3) and  $\nu^\varepsilon$  satisfies the tightness condition (2.8). Then for any  $\kappa \in [0, 1]$ , there exists  $c > 0$  such that*

$$\int \lambda^\kappa d\nu^\varepsilon \geq c > 0 \quad \text{uniformly in } \varepsilon > 0. \quad (4.17)$$

*Proof.* The assumptions of the Lemma ensure the conservation of the excess mass (4.12). Together with the tightness assumption, we have for any  $\kappa \leq 1$

$$\begin{aligned} \int \lambda^\kappa \nu_t^\varepsilon(d\lambda) &\geq \int_0^M \lambda^\kappa \nu_t^\varepsilon(d\lambda) \\ &\geq \frac{1}{M^{1-\kappa}} \int_0^M \lambda \nu_t^\varepsilon(d\lambda) \\ &\geq \frac{1}{M^{1-\kappa}} \left( \rho_0 - \rho_s - O_\varepsilon \left( \varepsilon^{\gamma(1/2-x)} \right) - o_M(1) \right). \end{aligned}$$

Hence, we can choose  $M$  large enough but finite and  $\varepsilon$  small enough such that for some  $c > 0$  the estimate (4.17) holds. □

**Lemma 4.8.** *For any  $n \in \mathcal{M}$  holds*

$$(u - h) \log \left( \frac{1 + u}{1 + h} \right) \leq \frac{\mathcal{D}_{\text{mac}}(n)}{A(z_s)}, \quad (4.18)$$

where  $u := u(z_s)$ ,

$$h := \frac{n_1 - z_s}{z_s} \quad \text{and} \quad u(z) := \frac{B - A(z)}{A(z)},$$

and

$$A(z) := \sum_{l \geq l_0} a_l z n_l \quad \text{and} \quad B := \sum_{l \geq l_0} b_{l+1} n_{l+1}. \quad (4.19)$$

Moreover, if  $\mathcal{F}(n) \leq C\varepsilon^\gamma$ ,  $l_0$  satisfies (2.3) and  $\nu^\varepsilon$  satisfies the tightness condition (2.8), then it holds for  $\varepsilon$  small enough and some  $C > 0$  uniformly in  $\varepsilon$  the estimate (2.15) from Proposition 2.4.

*Proof.* For the proof  $n$  is fixed such that  $\mathcal{F}(n) \leq C\varepsilon^\gamma$ . Then, we introduce two measures  $\alpha$  and  $\beta$  on  $\{l_0, l_0 + 1, \dots\}$

$$\forall l \geq l_0 : \quad \alpha_l(z) := a_l z n_l \quad \text{and} \quad \beta_l := b_{l+1} n_{l+1}$$

with partition sums  $A(z)$  and  $B$  (4.19), respectively.

We introduce  $\mathcal{D}_{\text{mac},z}(n)$  the constant monomer density dissipation of the large clusters

$$\mathcal{D}_{\text{mac},z}(n) := \sum_{l \geq l_0} (a_l z n_l - b_{l+1} n_{l+1}) \log \frac{a_l z n_l}{b_{l+1} n_{l+1}} \geq 0.$$

Note, that by this definition  $\mathcal{D}_{\text{mac}}(n) = \mathcal{D}_{\text{mac},n_1}(n)$ . By the definition (4.19), it follows  $A(n_1) = (1+h)A(z_s)$  and the identity

$$u(n_1) = \frac{1}{1+h} (u(z_s) - h). \quad (4.20)$$

Now, rewrite and  $\mathcal{D}_{\text{mac},z}(n)$  and apply the Jensen inequality to the one-homogeneous convex function  $a \mapsto a \log(1+a)$

$$\begin{aligned} \mathcal{D}_{\text{mac},z}(n) &= \sum_{l \geq l_0} \alpha_l(z) \frac{\beta_l - \alpha_l(z)}{\alpha_l(z)} \log \left( 1 + \frac{\beta_l - \alpha_l(z)}{\alpha_l(z)} \right) \\ &\geq (B - A(z)) \log \frac{B}{A(z)} = A(z) (u(z) \log(1 + u(z))) \end{aligned}$$

Hence, we obtain by setting  $z = n_1$  and using (4.20), the estimate

$$A(z) (u(z) \log(1 + u(z))) = A(z_s) (u - h) \log \left( 1 + \frac{u - h}{1 + h} \right),$$

from where we conclude (4.18). By using the explicit expression of the rates (1.1) follows

$$\begin{aligned} 2A(z_s) - B &= \sum_{l \geq l_0} z_s l^\alpha \left( 1 - \frac{q}{z_s l^\gamma} \right) n_l + z_s l_0^\alpha n_{l_0} + q l_0^{\alpha-\gamma} n_{l_0} \\ &\geq \left( 1 - \frac{q}{z_s l_0^\gamma} \right) A(z_s) \geq A(z_s) (1 - O(\varepsilon^{x\gamma})), \end{aligned}$$

by the definition of  $l_0$  (2.3). Hence, we have  $B \leq A(z_s) (1 + O(\varepsilon^{x\gamma}))$  and in particular with (4.20)

$$u(n_1) \leq \frac{u(z_s)}{1 - |h|} + \frac{|h|}{1 - |h|} \leq O(\varepsilon^{x\gamma}) + O\left(\varepsilon^{\frac{\gamma}{2}}\right) = O(\varepsilon^{x\gamma}), \quad (4.21)$$

where we used that with  $\mathcal{F}(n) \leq C\varepsilon^\gamma$  also  $|h| \leq C\varepsilon^{\frac{\gamma}{2}}$  from the estimate (4.3). The estimate (4.21) allows to linearize the bound (4.18) as follows

$$(u-h) \log\left(\frac{1+h}{1+u}\right) \geq \frac{(u-h)^2}{\max\{1+h, 1+u\}} \geq \frac{(u-h)^2}{1+O(\varepsilon^{x\gamma})}.$$

Finally, to deduce the estimate (2.15), it is enough to rewrite it in rescaled variables and use the estimate (4.17) from Lemma 4.7

$$A = z_s \varepsilon^{1-\alpha} \int \lambda^\alpha \nu^\varepsilon(d\lambda) \geq c\varepsilon^{1-\alpha} > 0. \quad \square$$

The time-scale separation between the dynamic of the small clusters and the one of the large clusters is characterized by the following logarithmic Sobolev type inequality.

**Proposition 4.9** (Microscopic energy-dissipation estimate). *Let  $\omega_l(z) := z^l Q_l$ . Then for all  $n \in \mathbb{R}_+^N$  with  $\mathcal{F}_{\text{mic}}(n) \leq C\varepsilon^\gamma$  there exists  $C_{\text{EED}}$  independent of  $\varepsilon$  such that it holds*

$$\mathcal{H}_{\text{mic}}(n \mid \omega(n_1)) \leq C_{\text{EED}} \varepsilon^{-x(1-\alpha+\gamma)} \mathcal{D}_{\text{mic}}(n), \quad (4.22)$$

where  $\mathcal{H}_{\text{mic}}$  is the microscopic part of the relative entropy between  $n$  and  $\omega(n_1)$  defined by

$$\mathcal{H}_{\text{mic}}(n \mid \omega(z)) := \sum_{l=1}^{l_0-1} \omega_l(z) \psi\left(\frac{n_l}{\omega_l(z)}\right) \quad \text{with} \quad \psi(a) = a \log a - a + 1.$$

*Proof.* We note, that the function  $(a, b) \mapsto \varphi(a, b) := (a-b)(\log a - \log b)$  occurring in the definition of the dissipation is one-homogeneous. In addition, the following lower bound holds

$$(a-b)(\log a - \log b) \geq 4\left(\sqrt{a} - \sqrt{b}\right)^2.$$

Moreover, we remind that  $\omega(z)$  satisfies the detailed balance condition  $a_l \omega_1(z) \omega_l(z) = b_{l+1} \omega_{l+1}(z)$ . By choosing  $z = n_1$ , the dissipation can be rewritten and bounded from below by

$$\begin{aligned} \mathcal{D}(n) &= \sum_{l \geq 1} a_l n_1 \omega_l(n_1) \varphi\left(\frac{n_l}{\omega_l(n_1)}, \frac{n_{l+1}}{\omega_{l+1}(n_1)}\right) \\ &\geq 4 \sum_{l \geq 1} a_l n_1 \omega_l(n_1) \left(\sqrt{\frac{n_l}{\omega_l(n_1)}} - \sqrt{\frac{n_{l+1}}{\omega_{l+1}(n_1)}}\right)^2 := \overline{\mathcal{D}}(n). \end{aligned}$$

Hence, instead of showing the estimate (4.22), it is sufficient to proof

$$\mathcal{H}_{\text{mic}}(n \mid \omega(n_1)) \leq C_{\text{EED}} \overline{\mathcal{D}}(n).$$

This inequality was investigated in [4]. To apply the result there, we introduce the measures

$$l \in \{1, \dots, l_0 - 1\} : \quad \mu_l(z) := \frac{\omega_l(z)}{\sum_{l=1}^{l_0-1} \omega_l(z)} \quad \text{and} \quad \nu_l(z) := \frac{a_l \omega_l(z)}{\sum_{l=1}^{l_0-1} \omega_l(z)}.$$

Hereby, we note that  $\mu$  is a probability measure, but  $\nu$  not necessarily. Let us assume the following mixed logarithmic Sobolev inequality

$$\text{Ent}_\mu(f^2) := \sum_{l=1}^{l_0-1} \mu_l f_l^2 \log \frac{f_l^2}{\sum_{l=1}^{l_0-1} f_l^2 \mu_l} \leq C_{\text{LSI}} \sum_{l=1}^{l_0-1} \nu_l (f_l - f_{l+1})^2. \quad (4.23)$$

Then, Proposition 3.2 from [4], where by the different normalization of  $\nu$ , the constant simplifies to

$$C_{\text{EED}} \leq \frac{C_{\text{LSI}}}{n_1^2} \left( n_1^2 + 2 \left( \sum_{l=1}^{l_0-1} n_l \right) \left( \sum_{l=1}^{l_0-1} \omega_l(n_1) \right) \right). \quad (4.24)$$

To proof the mixed logarithmic Sobolev inequality (4.23), we use Corollary 2.4 and Remark 2.5 from [4], from which we obtain the bound

$$C_{\text{LSI}} \leq 480 \sup_{1 < l < l_0} W_l(n_1) \log \left( \frac{W_1(n_1)}{W_l(n_1)} \right) V_l(n_1), \quad (4.25)$$

where

$$W_l(z) = \sum_{j=l}^{l_0-1} \omega_j(z) \quad \text{and} \quad V_l(z) = \sum_{j=1}^{l-1} \frac{1}{a_j \omega_j(z)}.$$

We will establish the following estimates for  $|z - z_s| \leq C\varepsilon^{\frac{\gamma}{2}}$  and some  $C > 1$

$$\frac{1}{C} \omega_l(z) \leq W_l(z) \leq Cl^\gamma \omega_l(z) \quad (4.26)$$

$$V_l(z) \leq \frac{Cl^{\gamma-\alpha}}{\omega_l(z)} \quad (4.27)$$

We postpone the proof of the estimates and first show the final result. By a combination of (4.26) and (4.27) with (4.25), we obtain the estimate

$$C_{\text{LSI}} \leq C \sup_{1 < l < l_0} l^{2\gamma-\alpha} \log \left( \frac{C}{\omega_l(z)} \right).$$

By the expansion (4.1) follows

$$\frac{1}{\omega_l(z)} \leq Cl^\alpha \exp \left( \frac{q(1-\gamma)}{z_s} l^{1-\gamma} - l \log \frac{z}{z_s} \right) \leq C \exp(Cl^{1-\gamma}),$$

where we used that  $l_0 \left| \log \frac{z}{z_s} \right|$  is uniformly bounded, because of  $|z - z_s| \leq C\varepsilon^{\frac{\gamma}{2}}$ . We obtain the upper bound  $C_{\text{LSI}} \leq C \sup_{1 < l < l_0} l^{1-\alpha+\gamma} \leq Cl_0^{1-\alpha+\gamma}$ . A combination of this bound with (4.24) leads to the bound

$$C_{\text{EED}} \leq C\varepsilon^{-x(1-\alpha+\gamma)} \frac{n_1^2 + W_1(n_1) \sum_{l=1}^{l_0-1} n_l}{n_1^2}$$

The conclusion (4.22) follows now from (4.26),  $|n_1 - z_s| \leq C\varepsilon^{\frac{\gamma}{2}}$  and the bound  $\sum_{l=1}^{l_0-1} n_l \leq \sum_{l \geq 1} l n_l = \varrho$ . To proof the estimates (4.26) and (4.27), we first observe that by the assumption  $|z - z_s| \leq C\varepsilon^{\frac{\gamma}{2}}$ , we have the comparison

$$\frac{\omega_l(z)}{\omega_l(z_s)} = \left(\frac{z}{z_s}\right)^l \leq \left(1 + C\varepsilon^{\frac{\gamma}{2}}\right)^l \leq \exp\left(C\varepsilon^{\frac{\gamma}{2}} l_0\right) \leq C$$

by the choice of  $l_0$  (2.3). In the complete analog way, we get  $\frac{\omega_l(z)}{\omega_l(z_s)} \geq \frac{1}{C}$ . Hence, it is enough to show (4.26) and (4.27) for  $z = z_s$ .

Therefore, we use the expansion (4.1) from Lemma 4.1 in the form: For some constant  $C > 1$  and any  $l \geq 1$  holds

$$\frac{1}{C l^\alpha z_s^{l-1}} \exp\left(-\frac{q}{z_s}(1-\gamma)l^{1-\gamma}\right) \leq Q_l \leq \frac{C}{l^\alpha z_s^{l-1}} \exp\left(-\frac{q}{z_s}(1-\gamma)l^{1-\gamma}\right).$$

The estimate (4.26) with  $z = z_s$  is now proven

$$\begin{aligned} W_l(z) &\leq \frac{C}{z_s^2} \sum_{j=l}^{l_0-1} \frac{1}{j^\alpha} \exp\left(-\frac{q}{z_s}(1-\gamma)j^{1-\gamma}\right) \leq \frac{C}{l^\alpha} \sum_{j=l}^{l_0-1} \exp\left(-\frac{q}{z_s}(1-\gamma)j^{1-\gamma}\right) \\ &\leq \frac{C}{l^\alpha} \int_l^{l_0-1} \exp\left(-\frac{q}{z_s}(1-\gamma)v^{1-\gamma}\right) dv \leq C l^{\gamma-\alpha} \int_{l^{1-\gamma}}^\infty \exp\left(-\frac{q}{z_s}(1-\gamma)v\right) dv \\ &\leq C l^\gamma \omega_l. \end{aligned}$$

The estimate (4.27) with  $z = z_s$  follows similarly

$$\begin{aligned} V_l(z) &\leq \frac{C}{z_s^2} \sum_{j=1}^{l-1} \exp\left(\frac{q}{z_s}(1-\gamma)l^{1-\gamma}\right) \leq \frac{C}{z_s^2} \int_1^l \exp\left(\frac{q}{z_s}(1-\gamma)v^{1-\gamma}\right) dv \\ &\leq C \frac{(1-\gamma)}{z_s^2} l^\gamma \int_0^{l^{1-\gamma}} \exp\left(\frac{q}{z_s}(1-\gamma)v\right) dv \leq C \frac{l^{\gamma-\alpha}}{\omega_l}. \end{aligned} \quad \square$$

The estimates (2.5) and (2.16) from Theorem 2.5 are a consequence of (4.22).

*Proof of Theorem 2.5.* Finally, the estimate (2.5) follows just by rescaling  $\mathcal{D}_{\text{mic}}$  and integrating the estimate along a curve of finite action. For the statement (2.16), we first observe that  $\mathcal{F}_{\text{mic}}(n) = \mathcal{H}_{\text{mic}}(n | \omega)$  and get by writing  $z = z_s e^h$

$$\begin{aligned} \mathcal{H}_{\text{mic}}(n | \omega(z)) - \mathcal{F}_{\text{mic}}(n) &= -\sum_{l=1}^{l_0-1} l n_l \log \frac{z}{z_s} - \sum_{l=1}^{l_0-1} \omega_l \left(1 - \left(\frac{z}{z_s}\right)^l\right) \\ &= h \left(\varrho_s - \sum_{l=1}^{l_0-1} l n_l + \sum_{l=1}^{l_0-1} l \omega_l \frac{e^{lh} - 1}{lh} - \varrho_s\right). \end{aligned}$$

The first difference in the bracket can be bounded in terms of (4.4) from Lemma 4.3. The second difference can be explicitly expressed as follows

$$\begin{aligned} \left| \varrho_s - \sum_{l=1}^{l_0-1} l\omega_l \frac{e^{lh} - 1}{lh} \right| &\leq \sum_{l \geq l_0} l\omega_l + \sum_{l=1}^{l_0-1} l\omega_l \frac{e^{lh} - 1 - lh}{lh} \\ &\leq Cl_0^\gamma \omega_{l_0} + Ch \sum_{l=1}^{l_0-1} l^2 \omega_l \leq C(l_0^\gamma \omega_{l_0} + h), \end{aligned}$$

where, we used the upper in (4.26), which holds by the proof also with  $l_0 = \infty$ . Moreover,  $\omega_l$  has arbitrary high moments following from the expansion (4.1). By the choice of  $l_0$  (2.3) and again (4.1) follows that  $l_0^\gamma \omega_{l_0} \leq C\varepsilon^p$  for any  $p > 0$ . Hence, combining all these estimates and reminding that  $\mathcal{F}_{\text{mic}}(n) \leq C\varepsilon^\gamma$ , we get

$$|\mathcal{H}_{\text{mic}}(n | \omega(z)) - \mathcal{F}_{\text{mic}}(n)| \leq C|h| \left( \varepsilon^{-x\gamma} \sqrt{\mathcal{F}_{\text{mic}}(n)} + \varepsilon^p + |h| \right) \leq C|h|(\varepsilon^\sigma + |h|),$$

where  $\sigma = \gamma(\frac{1}{2} - x) > 0$  by (2.3). Hence, we can conclude for a curve of finite action

$$\begin{aligned} \int_0^T \mathcal{F}_{\text{mic}}^\varepsilon(n(t)) dt &\leq z_s \varepsilon^{-\gamma} \int_0^T \mathcal{H}_{\text{mic}}(n(t) | \omega(n_1(t))) dt \\ &\quad + C\varepsilon^{-\gamma} \int_0^T \left| \log \left( \frac{n_1(t)}{z_s} - 1 \right) \right| \left( \varepsilon^\sigma + \left| \log \left( \frac{n_1(t)}{z_s} - 1 \right) \right| \right) dt \\ &\leq C\varepsilon^{(1-x)(1-\alpha+\gamma)} \int_0^T \mathcal{D}_{\text{mic}}^\varepsilon(n(t)) dt \\ &\quad + C\varepsilon^\gamma \int_0^T \left| \frac{n_1(t) - z_s}{\varepsilon^\gamma} \right|^2 dt + C\varepsilon^\sigma \sqrt{T} \left( \int_0^T \left| \frac{n_1(t) - z_s}{\varepsilon^\gamma} \right|^2 dt \right)^{\frac{1}{2}}. \end{aligned}$$

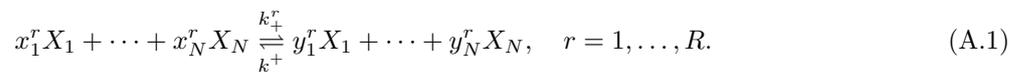
The conclusion (2.16) follows by (2.11) from Theorem 2.2.  $\square$

## APPENDIX A. GRADIENT STRUCTURES FOR COAGULATION AND FRAGMENTATION MODELS

### A.1 Reversible chemical reactions as gradient flows

This part of the appendix shows the general structure for reversible chemical reactions. Since, the Becker–Döring equation and other coagulation-fragmentation models can be interpreted as an infinite set of chemical reactions (1.10), they fall into this category. The basic observation goes back to Mielke [15], who found the entropic gradient flow structure for reversible chemical reactions.

**Definition A.1** (Reversible chemical reaction). Let  $n \in \mathbb{R}_+^N$  be the densities of  $N \in \mathbb{N} \cup \{+\infty\}$  different chemical species (or complexes)  $X_i$  reacting according to the mass action law. Each reaction  $r = 1, \dots, R$  with  $R \in \mathbb{N} \cup \{+\infty\}$  is characterized by the stoichiometric coefficients  $x^r, y^r \in \mathbb{N}_0^N$  and forward and backward reaction rates  $k_\pm^r > 0$



The chemical reaction is assumed to be reversible. That is, there exists a state  $\omega = (\omega_1, \dots, \omega_N) \in \mathbb{R}^N$  such that

$$k_+^r \omega^{x^r} = k_-^r \omega^{y^r} =: k^r. \quad (\text{A.2})$$

Here, the notation for multiindices is used:  $\omega^{x^r} = \prod_{i=1}^N \omega_i^{x_i^r}$ . The evolution equation for the density is given by

$$\dot{n} = - \sum_{r=1}^R k^r \left( \frac{n^{x^r}}{\omega^{x^r}} - \frac{n^{y^r}}{\omega^{y^r}} \right) (x^r - y^r). \quad (\text{A.3})$$

The Becker–Döring clustering equation interpreted as an infinite set of chemical reactions (1.10) fall in this framework by setting  $N = R = \infty$  and  $x_i^r := \delta_{i,1} + \delta_{i,r}$  and  $y_i^r := \delta_{i,r+1}$ . The detailed balance condition (A.2) is satisfied in terms of the one-parameter family of equilibrium distributions  $\omega(z)$  (1.5). Moreover, the more general Smoluchowski coagulation and fragmentation model fit into this framework (*cf.* Appendix A.2) under the assumption of detailed balance.

The free energy is defined as relative entropy with respect to the reversible equilibrium as in (1.7), *i.e.*  $\mathcal{F}(n) = \mathcal{H}(n | \omega)$  and hence  $D\mathcal{F}(n) = \left( \log \frac{n_1}{\omega_1}, \dots, \log \frac{n_N}{\omega_N} \right)$ . To define the manifold of states, the stoichiometric subspace and its complement are used

$$\mathcal{S} := \text{span}\{x^r - y^r : r = 1, \dots, R\} \quad \text{and} \quad \mathcal{S}^\perp := \{s \in \mathbb{R}^N : \gamma \cdot s = 0, \forall \gamma \in \mathcal{S}\}.$$

Then, the manifold is given for some fixed  $n_0 \in \mathbb{R}_+^N$  by the affine space of densities

$$\mathcal{M}_{n_0} := (n_0 + \mathcal{S}) \cap \mathbb{R}_+^N = \{n \in \mathbb{R}_+^N : n \cdot s = n_0 \cdot s, \forall s \in \mathcal{S}^\perp\}$$

The definition formalizes that  $\mathcal{S}^\perp$  contains all conservation laws of the reaction and therefore the tangent vectors on  $\mathcal{M}_{n_0}$  are given by  $\mathbb{R}^{\mathcal{S}}$ . Coagulation and fragmentation models of one species, like Becker–Döring, in this terminology are characterized by

$$\mathcal{S}^\perp = \text{span}\{\mathbf{I}\}, \quad \text{with} \quad \mathbf{I} := (1, 2, 3, 4, \dots).$$

Hence, the manifold has only one conserved quantity, which is the density  $\varrho_0 > 0$  of the total number of particles  $\mathcal{M} := \{n \in \mathbb{R}_+^N : n \cdot \mathbf{I} = \sum_{l=1}^\infty l n_l = \varrho_0\}$ .

The derivative of the energy  $D\mathcal{F}$  is a force and has to be interpreted as covector. The underlying metric can be specified by mapping covectors to (tangent-)vectors. This is done *via* the Onsager matrix to be defined as the symmetric semi-positive definite matrix

$$\mathcal{K}(n) := \sum_r k^r \Lambda \left( \frac{n^{x^r}}{\omega^{x^r}}, \frac{n^{y^r}}{\omega^{y^r}} \right) (x^r - y^r) \otimes (x^r - y^r), \quad (\text{A.4})$$

where  $\Lambda(\cdot, \cdot)$  is the logarithmic mean in (1.15). Hence, recalling that the space of vectors was given by  $\mathbb{R}^{\mathcal{S}}$ , we define the covectors with the help of the Onsager operator by (1.16), where the identification is well-defined since the image of  $\mathcal{K}$  is by definition  $\mathbb{R}^{\mathcal{S}}$ , whenever  $n$  is strictly positive in all of its components. Note, although the tangent space is state independent, this is not the case for the cotangent space.

With this preliminary definitions a reversible chemical reaction as given in Definition A.1 is formally the gradient flow of the free energy  $\mathcal{F}$  with respect to the metric structure induced by the Onsager operator (A.4)

and it holds the formal identity

$$\dot{n} = -\mathcal{K}(n)D\mathcal{F}(n). \quad (\text{A.5})$$

The property from which immediately follows that (A.5) is the same as (A.3) is

$$(x^r - y^r) \cdot D\mathcal{F}(n) = \sum_{i=1}^n x_i^r \log \frac{n_i}{\omega_i} - y_i^r \log \frac{n_i}{\omega_i} = \log \frac{n^{x^r}}{\omega^{x^r}} - \log \frac{n^{y^r}}{\omega^{y^r}},$$

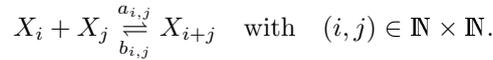
which is nothing else than the nominator of the logarithmic mean  $\Lambda\left(\frac{n^{x^r}}{\omega^{x^r}}, \frac{n^{y^r}}{\omega^{y^r}}\right)$  and resembles a discrete chain rule. The gradient flow decreases its energy along its evolution in terms of the dissipation, *i.e.*

$$\begin{aligned} \frac{d}{dt}\mathcal{F}(n) &= D\mathcal{F}(n) \cdot \dot{n} = -D\mathcal{F}(n) \cdot \mathcal{K}(n)D\mathcal{F}(n) \\ &= -\sum_r k^r \left( \frac{n^{x^r}}{\omega^{x^r}} - \frac{n^{y^r}}{\omega^{y^r}} \right) \left( \log \frac{n^{x^r}}{\omega^{x^r}} - \log \frac{n^{y^r}}{\omega^{y^r}} \right) =: -\mathcal{D}(n) \end{aligned}$$

We see that the Becker–Döring system fits into this framework. However, there is freedom in the choice of the free energy and under certain physical assumption, there are other possible choices.

## A.2 Smoluchowski coagulation and fragmentation equation

The Becker–Döring clustering equation is itself just a special case in the more general class of Smoluchowski coagulation and fragmentation equations seen as the following family of chemical reactions



Hence, the stoichiometric coefficients in (A.1) are given as  $x_k^{(i,j)} = \delta_{i,k} + \delta_{j,k}$  and  $y_k^{(i,j)} = \delta_{i+j,k}$ . A gradient flow structure can be established under the assumption of detailed balance, which in this case does not necessarily hold: There exists a state  $\omega \in \mathbb{R}^{\mathbb{N}}$  such that for  $(i, j) \in \mathbb{N} \times \mathbb{N}$  holds

$$a_{i,j}\omega_i\omega_j = b_{i,j}\omega_{i+j}.$$

Under this condition, the Smoluchowski coagulation and fragmentation equation is the gradient flow (A.5) of the free energy  $\mathcal{F}$  with respect to the Onsager operator  $\mathcal{K}$  defined in (A.4).

## A.3 Modified Becker–Döring system

The modified Becker–Döring system was introduced by Dreyer and Duderstadt [9]. The main feature is the introduction of a mixing entropy between the clusters. Hence, the free energy consists of a relative entropy part as defined in (1.7) plus a mixing entropy depending on the total number of clusters

$$\tilde{\mathcal{F}}(n) = \mathcal{H}(n \mid \omega) - N(n)(\log N(n) - 1) \quad \text{with} \quad N(n) = \sum_i n_i. \quad (\text{A.6})$$

The most compact form of the free energy is  $\tilde{\mathcal{F}}(n) = \sum_i \left( n_i \log \frac{n_i}{\omega_i N(n)} + \omega_i \right)$ . Hence, the differential of the free energy differential is given by

$$D\tilde{\mathcal{F}}(n) = \left( \log \frac{n_1}{\omega_1 N(n)}, \dots, \log \frac{n_i}{\omega_i N(n)}, \dots \right). \quad (\text{A.7})$$

The reaction is still of the same form as the classical Becker–Döring system (1.10), *i.e.*  $x_i^r = \delta_{i,1} + \delta_{i,r}$  and  $y_i^r = \delta_{i,r+1}$  in (A.1). This leads to the same detailed balance condition as for the classical Becker–Döring model  $a_r \omega_1 \omega_r = b_{r+1} \omega_{r+1} =: k^r$ . Hence, we obtain the same possible equilibrium states  $\omega_r(z) = z^r Q_r$  given in (1.5). Again  $z$  has to be determined from the formal conservation law  $\sum_{l=1}^{\infty} l \omega_l(z) = \sum_{l=1}^{\infty} l n_l$ . However, the existence as minimizer of the free energy in this case is more involved and for a detailed analysis of the equilibrium states, we refer to [11].

Now, from (A.7), we further deduce

$$(x^r - y^r) \cdot D\tilde{\mathcal{F}}(n) = \log \frac{n_1 n_r}{\omega_1 \omega_r N(n)^2} - \log \frac{n_{r+1}}{\omega_{r+1} N(n)} = \log \frac{n_1 n_r}{\omega_1 \omega_r} - \log \frac{N(n) n_{r+1}}{\omega_{r+1}}.$$

From the above identity, the modified Onsager matrix can be read off and is given by

$$\mathcal{K}(n) := \sum_r k^r \Lambda \left( \frac{n_1 n_r}{\omega_1 \omega_r}, \frac{N(n) n_{r+1}}{\omega_{r+1}} \right) (x^r - y^r) \otimes (x^r - y^r).$$

Then, we obtain the modified Becker–Döring equation as the gradient flow of the modified free energy (A.6)

$$\begin{aligned} \dot{n} &= -\tilde{\mathcal{K}}(n) D\tilde{\mathcal{F}}(n) = - \sum_r k^r \left( \frac{n_1 n_r}{\omega_1 \omega_r} - \frac{N(n) n_{r+1}}{\omega_{r+1}} \right) (x^r - y^r) \\ &= - \sum_r (a_r n_1 n_r - b_{r+1} N(n) n_{r+1}) (x^r - y^r) = - \sum_r \tilde{J}_r (x^r - y^r). \end{aligned}$$

The explicit form of the equation is given for any  $l = 1, 2, \dots$  by

$$\dot{n}_l = \tilde{J}_{l-1} - \tilde{J}_l \quad \text{with} \quad \tilde{J}_0 := - \sum_{r=1}^{\infty} \tilde{J}_r \quad \text{and} \quad \tilde{J}_r(n) := a_r n_1 n_r - b_{r+1} N(n) n_{r+1}.$$

## APPENDIX B. PROOF OF LEMMA 4.1

*Proof of Lemma 4.1.* We calculate using the definition (1.5) of  $Q_l$

$$\log(l^\alpha z_s^{l-1} Q_l) = - \sum_{j=2}^l \log \left( 1 + \frac{q}{z_s j^\gamma} \right)$$

The function  $x \mapsto \log \left( 1 + \frac{q}{z_s k^\gamma} \right)$  is positive, continuous and monotone decreasing to 0. Therefore, we can define the Euler number

$$C_1 := \lim_{l \rightarrow \infty} \left( \sum_{j=2}^l \log \left( 1 + \frac{q}{z_s j^\gamma} \right) - \int_2^l \log \left( 1 + \frac{q}{z_s x^\gamma} \right) \right).$$

Moreover, we get from the Euler-MacLaurin formula the estimate

$$\left| C_1 - \left( \sum_{j=2}^l \log \left( 1 + \frac{q}{z_s j^\gamma} \right) - \int_2^l \log \left( 1 + \frac{q}{z_s x^\gamma} \right) dx \right) \right| \leq \log \left( 1 + \frac{q}{z_s l^\gamma} \right) \leq \frac{q}{z_s l^\gamma}.$$

The following bound

$$\frac{q}{z_s x^\gamma} - \frac{1}{2} \left( \frac{q}{z_s x^\gamma} \right)^2 \leq \log \left( 1 + \frac{q}{z_s x^\gamma} \right) \leq \frac{q}{z_s x^\gamma} - \frac{1}{2} \left( \frac{q}{z_s x^\gamma} \right)^2 + \frac{1}{3} \left( \frac{q}{z_s x^\gamma} \right)^3$$

implies the estimate

$$0 \leq \frac{\int_2^l \log \left( 1 + \frac{q}{z_s x^\gamma} \right) dx}{\int_2^l \left( \frac{q}{z_s x^\gamma} - \frac{1}{2} \left( \frac{q}{z_s x^\gamma} \right)^2 \right) dx} - 1 \leq O(l^{-\gamma}),$$

hereby, we use the convention that  $\frac{l^{\kappa-1}}{\kappa} = \log l$  for  $\kappa = 0$ . Now, we can combine all the estimates to obtain

$$\begin{aligned} \log(l^\alpha z_s^{l-1} Q_l) &= - \left( \sum_{j=2}^l \log \left( 1 + \frac{q}{z_s j^\gamma} \right) - \int_2^l \log \left( 1 + \frac{q}{z_s x^\gamma} \right) dx \right) \\ &\quad - \left( 1 + \frac{\int_2^l \log \left( 1 + \frac{q}{z_s x^\gamma} \right) dx}{\int_2^l \left( \frac{q}{z_s x^\gamma} - \frac{1}{2} \left( \frac{q}{z_s x^\gamma} \right)^2 \right) dx} - 1 \right) \int_2^l \left( \frac{q}{z_s x^\gamma} - \frac{1}{2} \left( \frac{q}{z_s x^\gamma} \right)^2 \right) dx \\ &= \left( \mathcal{F}_0 - \frac{q}{z_s(1-\gamma)} l^{1-\gamma} + \frac{q^2}{2z_s^2(1-2\gamma)} l^{1-2\gamma} \right) (1 + O(l^{-\gamma})), \end{aligned}$$

which concludes the proof by setting  $\mathcal{F}_0 = \frac{q2^{1-\gamma}}{z_s(1-\gamma)} - \frac{q^2 2^{1-2\gamma}}{2z_s^2(1-2\gamma)} - C_1$ .  $\square$

*Acknowledgements.* The author wishes to thank Matthias Erbar, Stefan Luckhaus, Babara Niethammer and Juan Velázquez for many fruitful discussions on the Becker–Döring system, LSW equation, gradient flows and related topics. The author thanks the referees whose incisive and detailed comments have substantially improved the final version of the manuscript. The author gratefully acknowledges support by the German Research Foundation through the Collaborative Research Center 1060 *The Mathematics of Emergent Effects*. Part of this work was done while the author was enjoying the hospitality of the Hausdorff Research Institute for Mathematics during the Junior Trimester Program on Optimal Transport.

## REFERENCES

- [1] L. Ambrosio, N. Gigli and G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures. *Lectures in Mathematics ETH Zürich*. Birkhäuser-Verlag, Basel (2005).
- [2] J.M. Ball, J. Carr and O. Penrose, The Becker–Döring cluster equations: basic properties and asymptotic behaviour of solutions. *Commun. Math. Phys.* **104** (1986) 657–692.
- [3] R. Becker and W. Döring, Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Ann. Phys.* **24** (1935) 719–752.
- [4] J. Cañizo, A. Einav and B. Lods, Trend to equilibrium for the Becker–Döring equations: an analogue of Cercignani’s conjecture. *Anal. & PDE* **10** (2017) 1663–1708.
- [5] J.-F. Collet, T. Goudon, F. Poupaud and A. Vasseur, The Becker–Döring system and its Lifshitz-Slyozov limit. *SIAM J. Appl. Math.* **62** (2002) 1488–1500.

- [6] J.G. Conlon and A. Schlichting, A Non-local Problem for the Fokker-Planck Equation Related to the Becker-Döring Model. Preprint [arXiv:1711.00782](https://arxiv.org/abs/1711.00782) (2017).
- [7] E. De Giorgi, A. Marino and M. Tosques, Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **68** (1980) 180–187.
- [8] J. Deschamps, E. Hingant and R. Yvinec, Quasi Steady State Approximation of the Small Clusters in Becker–Döring Equations Leads to Boundary Conditions in the Lifshitz-Slyozov Limit. *Commun. Math. Sci.* **15** (2017) 1353-1384.
- [9] W. Dreyer and F. Duderstadt, On the Becker–Döring theory of nucleation of liquid droplets in solids. *J. Stat. Phys.* **123** (2006) 55–87.
- [10] M. Erbar, M. Fathi, V. Laschos and A. Schlichting, Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discret. Contin. Dyn. Syst.* **36** (2016) 6799–6833.
- [11] M. Herrmann, M. Naldzhieva and B. Niethammer, On a thermodynamically consistent modification of the Becker–Döring equations. *Physica D* **222** (2006) 116–130.
- [12] P. Laurençot and S. Mischler, From the Becker–Döring to the Lifshitz–Slyozov–Wagner equations. *J. Stat. Phys.* **106** (2002) 957–991.
- [13] I.M. Lifshitz and V.V. Slyozov, The kinetics of precipitation from supersaturated solid solutions. *J. Phys. Chem. Solids* **19** (1961) 35–50.
- [14] J. Maas, Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261** (2011) 2250–2292.
- [15] A. Mielke, A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity* **24** (2011) 1329.
- [16] B. Niethammer, On the evolution of large clusters in the Becker–Döring model. *J. Nonlinear Sci.* **13** (2003) 115–122.
- [17] B. Niethammer, Macroscopic limits of the Becker–Döring equations. *Commun. Math. Sci.* **2** (2004) 85–92.
- [18] B. Niethammer and R.L. Pego, On the initial-value problem in the Lifshitz–Slyozov–Wagner theory of Ostwald ripening. *SIAM J. Math. Anal.* **31** (2000) 467–485.
- [19] B. Niethammer and R.L. Pego, Well-posedness for measure transport in a family of nonlocal domain coarsening models. *Indiana Univ. Math. J.* **54** (2005) 499–530.
- [20] F. Otto and M. Reznikoff, Slow motion of gradient flows. *J. Differ. Equ.* **237** (2007) 372–420.
- [21] O. Penrose, Metastable states for the Becker–Döring cluster equations. *Commun. Math. Phys.* **541** (1989) 515–541.
- [22] O. Penrose, The Becker–Döring equations at large times and their connection with the LSW theory of coarsening. *J. Stat. Phys.* **89** (1997) 305–320.
- [23] E. Sandier and S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Commun. Pure Appl. Math.* **57** (2004) 1627–1672.
- [24] S. Serfaty, Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discret. Contin. Dyn. Syst.* **31** (2011) 1427–1451.
- [25] C. Wagner, Theorie der Alterung von Niederschlägen durch Umlösen (Ostwald-Reifung). *Z. Elektrochem.* **65** (1961) 581–591.