J Nonlinear Sci (2018) 28:765–806
https://doi.org/10.1007/s00332-017-9427-9

Journal of Nonlinear Science

Convergence to Equilibrium in Energy-Reaction–Diffusion Systems Using Vector-Valued Functional Inequalities

Alexander Mielke[1,2] · **Markus Mittnenzweig**[1]

Received: 11 March 2017 / Accepted: 11 November 2017 / Published online: 27 November 2017

Abstract We discuss how the recently developed energy dissipation methods for reaction diffusion systems can be generalized to the non-isothermal case. For this, we use concave entropies in terms of the densities of the species and the internal energy, where the importance is that the equilibrium densities may depend on the internal energy. Using the log-Sobolev estimate and variants for lower-order entropies as well as estimates for the entropy production of the nonlinear reactions, we give two methods to estimate the relative entropy by the total entropy production, namely a somewhat restrictive convexity method, which provides explicit decay rates, and a very general, but weaker compactness method.

Keywords Energy-reaction-diffusion systems · Vector-valued inequalities · Cross diffusion · log-Sobolev inequality · Entropy functional · Exponential decay of relative entropy · Convexity method

Mathematics Subject Classification 35K57 · 35B40 · 35Q79 · 92E20

Dedicated to Peter Markowich on the occasion of his sixtieth birthday.

Communicated by Felix Otto.

This research has been partially funded by European Research Council (ERC) via AdG 267802 *AnaMultiScale* and by Deutsche Forschungsgemeinschaft (DFG) through SFB 1114, Subproject C05.

✉ Alexander Mielke
alexander.mielke@wias-berlin.de

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

2 Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 12489 Berlin-Adlershof, Germany

1 Introduction

The aim of this paper is to generalize the recently developed exponential decay estimates for reaction–diffusion systems based on the entropy method from the isothermal case to the non-isothermal case. Entropy methods for reaction–diffusion systems (RDS) satisfying a detailed balance condition were first introduced in Gröger (1983, 1986), and a first result on uniform exponential decay was derived in Gröger (1992), Thm. 2 via a compactness argument. Applications to semiconductor models were obtained in Glitzky et al. (1994, 1996) and Glitzky and Hünlich (1997), where the self-consistent coupling through the Poisson equation led to the restriction $d \leq 2$ for the space dimension. Starting from Desvillettes and Fellner (2006), an independent and more constructive approach for isothermal RDS was developed, see Fellner and Tang (2017) and Desvillettes et al. (2017) as well as the references therein. In Mielke et al. (2015), a convexity argument was introduced that is less general in applications but leads to sharper decay estimates if applicable, see also the survey (Mielke 2017).

In this work, we discuss how some of the above methods can be transferred to the non-isothermal case, where the RDS is complemented with an energy balance either expressed as a heat equation or as a balance for the internal energy u. The coupled system is called energy-reaction–diffusion systems (ERDS). A first step toward the non-isothermal case was done in Haskovec et al. (2017), where a more explicit approach close to Desvillettes and Fellner (2006) and Fellner and Tang (2017) was developed. Here, we want to show that also the convexity method from Mielke et al. (2015) and the compactness method from Gröger (1992) and Glitzky and Hünlich (1997) can be adapted to ERDS.

Following Albinus et al. (2002) and Mielke (2011b, 2013), we use the internal energy density $u(t, \cdot) : \Omega \to \mathbb{R}$ as the main thermodynamic variable, while the absolute temperature is a dependent variable obtained from the constitutive entropy function $s = S(x, \boldsymbol{c}, u)$ via

$$\theta(t, x) = \frac{1}{\partial_u S(x, \boldsymbol{c}(t, x), u(t, x))},$$

where $\boldsymbol{c} = (c_1, \ldots, c_I) \in [0, \infty[^I$ is the vector of concentrations. The major advantage is that thermodynamics imposes that $S(x, \cdot, \cdot) : [0, \infty[^{I+1} \to \mathbb{R}$ is a concave function that allows us to apply entropy entropy-production (EEP) estimates. Moreover, $S(x, \boldsymbol{c}, u)$ is a strictly increasing function in u to make the absolute temperature nonnegative.

Our typical model for relevant entropy relations is given by

$$S(x, \boldsymbol{c}, u) = s(u) - \mathfrak{B}\big(\boldsymbol{c} \big| \boldsymbol{w}(x, u)\big) = \widehat{s}(x, u) + \sum_{i=1}^{I} \Big(c_i \log\big(w_i(x, u)\big) - \lambda_{\mathfrak{B}}(c_i)\Big), \tag{1.1}$$

where the Boltzmann function $\lambda_{\mathfrak{B}}$ and the relative Boltzmann entropy are defined via

$$\lambda_{\mathfrak{B}}(z) := z \log z - z + 1 \quad \text{and } \mathfrak{B}\big(\boldsymbol{c} \big| \boldsymbol{b}\big) := \sum_{i=1}^{I} b_i \lambda_{\mathfrak{B}}(c_i / b_i). \tag{1.2}$$

Here each $w_i(x,\cdot) : [0,\infty[\;\to [0,\infty[$ is a non-decreasing and concave function as well as $\widehat{s}(x,\cdot) : u \mapsto s(x,u) + I - \sum_{i=1}^{I} w_i(x,u)$, see the thermodynamic discussions in Sect. 2 where Proposition 2.1 provides the necessary concavity result. From $\partial_{\boldsymbol{c}} S(x,\boldsymbol{c},u) = \big(\log(c_i/w_i(x,u))\big)_{i=1,\dots,I}$, we see that the vector $\boldsymbol{w}(x,u) = \big(w_i(x,u)\big)_{i=1,\dots,I}$ gives the thermodynamic equilibrium concentrations for a given internal energy u. Further on, we assume that the constitutive function $S(x,\boldsymbol{c},u)$ and hence w_i do not depend on x, i.e., the material is homogeneous. However, following the ideas in Glitzky and Hünlich (1997) and Mielke et al. (2015), Sect. 5.1, it should be possible to generalize the methods developed here also to the inhomogeneous case.

One way of generating thermodynamically consistent energy-reaction–diffusion systems (ERDS) on a physical domain $\Omega \subset \mathbb{R}^d$ is that of using the theory of gradient systems. For this, we use the entropy functional

$$\mathcal{S}(\boldsymbol{c},u) := \int_\Omega S(\boldsymbol{c}(x),u(x))\,\mathrm{d}x$$

and an Onsager operator $\mathbb{K}(\boldsymbol{c},u) = \mathbb{K}_{\text{diff}}(\boldsymbol{c},u) + \mathbb{K}_{\text{R}}(\boldsymbol{c},u)$ modeling the entropy production by diffusion or heat transfer and by reaction, respectively. We always assume that they are of the form

$$\mathbb{K}_{\text{diff}}(\boldsymbol{c},u)\begin{pmatrix}\boldsymbol{\zeta}\\ \mu\end{pmatrix} = -\operatorname{div}\left(\mathbb{M}(\boldsymbol{c},u)\nabla\begin{pmatrix}\boldsymbol{\zeta}\\ \mu\end{pmatrix}\right) \text{ and } \mathbb{K}_{\text{R}}(\boldsymbol{c},u) = \begin{pmatrix}\mathbb{L}(\boldsymbol{c},u) & 0\\ 0 & 0\end{pmatrix},$$

where $\mathbb{K}_{\text{diff}}$ is complemented by the no-flux conditions $\mathbb{M}(\boldsymbol{c},u)\nabla\binom{\zeta}{\mu}\cdot\nu = 0$ on $\partial\Omega$. True Onsager operators $\mathbb{K}$ satisfy the symmetry (reciprocal) relation (cf. Onsager 1931) $\mathbb{K} = \mathbb{K}^*$. Moreover, the second law of thermodynamics enforces $\mathbb{K}+\mathbb{K}^* \geq 0$. Obviously, this is satisfied if the matrices $\mathbb{M}$ and $\mathbb{L}$ are symmetric and positive semi-definite. For our EEP estimates, the symmetry will not be important, since we will only rely on the positive semi-definiteness

$$\boldsymbol{A} : \mathbb{M}(\boldsymbol{c},u)\boldsymbol{A} \geq 0 \quad\text{and}\quad \boldsymbol{\zeta}\cdot\mathbb{L}(\boldsymbol{c},u)\boldsymbol{\zeta} \geq 0.$$

Of course, we will need quantitative lower bounds for $\mathbb{M}$ and $\mathbb{L}$. Indeed, starting from the entropy functional $\mathcal{S}$ we consider the ERDS for $\boldsymbol{y} = (\boldsymbol{c},u)$, which is derived formally from $\dot{\boldsymbol{y}} = \mathbb{K}(\boldsymbol{y})\mathrm{D}\mathcal{S}(\boldsymbol{y})$ giving

$$\begin{pmatrix}\dot{\boldsymbol{c}}\\ \dot{u}\end{pmatrix} = -\operatorname{div}\left(\mathbb{M}(\boldsymbol{c},u)\begin{pmatrix}\nabla\mathrm{D}_{\boldsymbol{c}}S(\boldsymbol{c},u)\\ \nabla\partial_u S(\boldsymbol{c},u)\end{pmatrix}\right) + \begin{pmatrix}\mathbb{L}(\boldsymbol{c},u)\mathrm{D}_{\boldsymbol{c}}S(\boldsymbol{c},u)\\ 0\end{pmatrix} \quad \text{in } \Omega,$$
$$\nu\cdot\nabla\mathrm{D}_{\boldsymbol{c}}S(\boldsymbol{c},u) = \boldsymbol{0} \text{ and } \nu\cdot\nabla\partial_u S(\boldsymbol{c},u) = 0 \quad \text{on } \partial\Omega.$$

Obviously the divergence form of the equation for u together with the no-flux boundary condition imply that the average internal energy

$$\overline{u(t)} := \frac{1}{|\Omega|}\int_\Omega u(t,x)\,\mathrm{d}x$$

is constant along solutions of the ERDS. Throughout this work, we assume that $\Omega \subset \mathbb{R}^d$ is a bounded domain with Lipschitz boundary and, without loss of generality, with unit mass, i.e., $|\Omega| = 1$. Thus, averages on Ω are simply equal to the integral over Ω.

Typically in RDS, there are further conserved quantities (like the total atomic masses or the charges in semiconductor models), which we assume to be given by a matrix $\mathsf{Q} : \mathbb{R}^I \to \mathbb{R}^m$ such that $\mathsf{Q}\mathbb{L}(\boldsymbol{c}, u)\boldsymbol{\zeta} = 0$. Then, again using the no-flux boundary conditions, we see that

$$\mathsf{Q}\overline{\boldsymbol{c}(t)} = \int_\Omega \mathsf{Q}\boldsymbol{c}(t, x)\,\mathrm{d}x \in \mathbb{R}^m$$

is constant along solutions. In particular, we define the associated flow-invariant subsets

$$\mathfrak{S}(\mathsf{q}, \mathsf{U}) := \left\{ (\boldsymbol{c}, u) \in \mathrm{L}^1(\Omega)^{I+1} \;\middle|\; c_i, u \geq 0,\ \overline{u} = \mathsf{U},\ \mathsf{Q}\overline{c} = \mathsf{q} \right\}.$$

The strictly concave functional $\mathcal{S}$ has a unique maximizer, which is spatially constant and which we denote by $\mathsf{w}(\mathsf{q}, \mathsf{U}) \in\,]0, \infty[^I$. With this we define the relative entropy

$$\begin{aligned}
&\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) := \int_\Omega H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}(x), u(x))\,\mathrm{d}x \text{ with} \\
&H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) = S(\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U}) + \mathrm{D}S(\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U}) \cdot \begin{pmatrix} \boldsymbol{c} - \mathsf{w}(\mathsf{q}, \mathsf{U}) \\ u - \mathsf{U} \end{pmatrix} - S(\boldsymbol{c}, u).
\end{aligned}$$

Hence, we have $\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) \geq 0$ with equality if and only if $(\boldsymbol{c}, u) = (\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U})$.

The entropy production $\mathcal{P}$ is now defined as the increase of $\mathcal{S}$ along solutions $\boldsymbol{y}(t) = (\boldsymbol{c}(t), u(t))$, namely

$$\mathcal{P}(\boldsymbol{y}) := \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\boldsymbol{y}) = \langle \mathrm{D}\mathcal{S}(\boldsymbol{y}), \mathbb{K}(\boldsymbol{y})\mathrm{D}\mathcal{S}(\boldsymbol{y}) \rangle,$$

which has a diffusive contribution $\mathcal{P}_\mathrm{D}$ and a reactive contribution $\mathcal{P}_\mathrm{R}$ given by

$$\begin{aligned}
&\mathcal{P}_\mathrm{D}(\boldsymbol{y}) := \int_\Omega \nabla \boldsymbol{y}(x) : \mathbb{W}(\boldsymbol{y}(x)) \nabla \boldsymbol{y}(x)\,\mathrm{d}x \;\text{ and } \mathcal{P}_\mathrm{R}(\boldsymbol{y}) := \int_\Omega P_\mathrm{R}(\boldsymbol{y}(x))\,\mathrm{d}x \\
&\text{with } \mathbb{W}(\boldsymbol{y}) = \mathrm{D}^2 S(\boldsymbol{y}) \mathbb{M}(\boldsymbol{y}) \mathrm{D}^2 S(\boldsymbol{y}) \text{ and } P_\mathrm{R}(\boldsymbol{y}) = \mathrm{D}S(\boldsymbol{y}) \cdot \mathbb{L}(\boldsymbol{y}) \mathrm{D}S(\boldsymbol{y}).
\end{aligned} \tag{1.3}$$

The aim of this work is to provide methods to derive entropy entropy-production (EEP) estimates in the form (see (2.16) for the definition of $\mathfrak{Q}$)

$$\forall\, (\mathsf{q}, \mathsf{U}) \in \mathfrak{Q}\ \exists\, K(\mathsf{q}, \mathsf{U}) > 0\ \forall\, (\boldsymbol{c}, u) \in \mathfrak{S}(\mathsf{q}, \mathsf{U}) :\ \mathcal{P}(\boldsymbol{c}, u) \geq K(\mathsf{q}, \mathsf{U}) \mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u), \tag{1.4}$$

which can be seen as a vector-valued generalization of the usually scalar functional inequalities of log-Sobolev type.

Indeed, we emphasize that we do neither address the question of existence of solutions for equations of the type (1.3) (see Fischer 2015; Jüngel 2015; Dreyer et al. 2016 and the references therein for recent advances) nor the question in which norms the exponential decay of the solutions can be established. We are content with the construction of $K(\mathsf{q}, \mathsf{U}) > 0$, which implies that along sufficiently good notions of weak solutions we have the uniform exponential decay of the relative entropy, namely

$$(\boldsymbol{c}(0), u(0)) \in \mathfrak{S}(\mathsf{q}, \mathsf{U}) \quad \Longrightarrow \quad \mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}(t), u(t)) \leq \mathrm{e}^{-K(\mathsf{q},\mathsf{U})\,t}\, \mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}(0), u(0)) \text{ for all } t \geq 0.$$

Clearly, this implies convergence $(\boldsymbol{c}(t), u(t)) \to (\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U})$ in $\mathrm{L}^1(\Omega)$ when exploiting a suitable variant of the Csiszár–Kullback–Pinsker estimate.

We discuss two possible strategies to establish the fundamental EEP estimate (1.4). Both rely on estimating $\mathbb{W}$ and P_R by suitable, simplified functions, namely

$$\mathbb{W}(\boldsymbol{c}, u) \geq \delta_* \operatorname{diag}\left(\frac{1}{c_1}, \ldots, \frac{1}{c_I}, \frac{1}{u^{2-\gamma}}\right) \text{ and} \tag{1.5a}$$

$$P_\mathrm{R}(\boldsymbol{c}, u) \geq \kappa_* \widehat{P}_\mathrm{R}(\boldsymbol{c}, u) \text{ with } \widehat{P}_\mathrm{R}(\boldsymbol{c}, u) := \sum_{r=1}^R \widehat{\kappa}_r(u) G\left(\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\alpha}^r}}, \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\beta}^r}}\right) \tag{1.5b}$$

with $G(a, b) = (a-b)\big(\log a - \log b\big) \geq 0$. We show that these estimates can be realized for ERDS of the form

$$\begin{pmatrix} \dot{\boldsymbol{c}} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} \big(\delta_i \Delta c_i\big)_i \\ \delta_\mathrm{u} \Delta u \end{pmatrix} + \begin{pmatrix} \boldsymbol{R}(\boldsymbol{c}, u) \\ 0 \end{pmatrix} \text{ with } \boldsymbol{R}(\boldsymbol{c}, u) = \sum_{r=1}^R \kappa_r \big(\tfrac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\alpha}^r}} - \tfrac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\beta}^r}}\big)\big(\boldsymbol{\beta}^r - \boldsymbol{\alpha}^r\big), \tag{1.6}$$

if the reaction rate coefficients satisfy $\kappa_r \geq \kappa_*$ and if for all $i = 1, \ldots, I$ there exist $\rho_i > 0$ such that the diffusion coefficients δ_i and the equilibrium functions w_i satisfy the relations

$$\rho_i \leq \frac{4\delta_\mathrm{u}(\delta_i - \delta_*)}{(\delta_i + \delta_\mathrm{u})^2} \quad \text{and} \quad \frac{1}{\rho_i}\big(w_i'(u)\big)^2 \leq \big(w_i'(u)\big)^2 - w_i(u) w_i''(u) \text{ for } u > 0,$$

see Proposition 2.3, which establishes (1.5a). We interpret these relations as a bound on cross-diffusion, since the heat equation without cross-diffusion should have the form

$$\dot{u} = -\operatorname{div}\left(M_{\mathrm{u},\mathrm{u}}(\boldsymbol{c}, u)\nabla\frac{1}{\theta}\right) \text{ with } \frac{1}{\theta} = \partial_u S(\boldsymbol{c}, u) = s'(u) + \sum_1^I c_i w_i'(u)/w_i(u),$$

see the discussion in Sect. 2.2. The assumption (1.5a) provides the lower bound

$$\mathcal{P}_\mathrm{D}(\boldsymbol{c}, u) \geq \delta_* p_\Omega \int_\Omega \left(\overline{u}\lambda_{\mathfrak{B}}\left(\frac{u}{\overline{u}}\right) + \sum_{i=1}^I \overline{c}_i \lambda_{\mathfrak{B}}\left(\frac{c_i}{\overline{c}_i}\right)\right) \mathrm{d}x$$

with $p_\Omega = \min\{\rho(\Omega, \gamma, 1), \rho(\Omega, 1, 1)\}$, where the optimal generalized log-Sobolev constants $\rho(\Omega, \gamma, \alpha)$ are defined in (3.6).

For the reactive part $\widehat{P}_\mathrm{R}$ of the entropy production (see (1.5b) for the definition), we impose the finite-dimensional conditions

$$\forall\,(\mathsf{q},\mathsf{U})\in\mathfrak{Q}\;\exists\,K_\mathrm{R}(\mathsf{q},\mathsf{U})>0\;\forall\,\boldsymbol{c}\text{ with }\mathsf{Q}\boldsymbol{c}=\mathsf{q}:\;\widehat{P}_\mathrm{R}(\boldsymbol{c},\mathsf{U})\geq K_\mathrm{R}(\mathsf{q},\mathsf{U})\mathfrak{B}\big(\boldsymbol{c}\,\big|\,\mathsf{w}(\mathsf{q},\mathsf{U})\big), \tag{1.7}$$

which, for our mass action kinetics, is equivalent to the unique-equilibrium condition that $\boldsymbol{c}=\mathsf{w}(\mathsf{q},\mathsf{U})$ is the only solution of $\boldsymbol{R}(\boldsymbol{c},\mathsf{U})=0$ satisfying $\mathsf{Q}\boldsymbol{c}=\mathsf{q}$.

The counterpart of the convexity method for isothermal RDS developed in Mielke et al. (2015) is formulated in Theorem 3.2. Assuming that there exists a $\mu(\mathsf{U})\geq 0$ such that the function

$$(\boldsymbol{c},u)\;\mapsto\;\mu(\mathsf{U})\Big(\mathsf{U}^\gamma\lambda_\mathfrak{B}(u/\mathsf{U})+\sum_i\lambda_\mathfrak{B}(c_i)\Big)+\widehat{P}_\mathrm{R}(\boldsymbol{c},u)\;\text{ is convex}$$

we obtain the constructive lower bound

$$K(\mathsf{q},\mathsf{U})\geq\frac{1}{C_H(\mathsf{q},\mathsf{U})}\;\min\Big\{\kappa_*K_\mathrm{R}(\mathsf{q},\mathsf{U}),\,\delta_*p_\Omega\tfrac{K_\mathrm{R}(\mathsf{q},\mathsf{U})}{\mu(\mathsf{U})+K_\mathrm{R}(\mathsf{q},\mathsf{U})}\Big\},$$

where the C_H is characterized in Lemma 3.1.

In Sect. 4, we discuss the generalization of the compactness argument developed in Glitzky and Hünlich (1997) for isothermal semiconductor models. There only the case $d\leq 2$ is treated because of the coupling via the electric potential, but without this there is no dimension restriction, see also Mielke (2017). In contrast to the convexity method, the theory is much more general, but the result is also significantly weaker, since the constant $K(\mathsf{q},\mathsf{U})$ in the EEP estimate is not constructive, and moreover, it also depends on an upper bound M for $\mathcal{H}_{\mathsf{q},\mathsf{U}}$, namely

$$\begin{aligned}&\forall\,(\mathsf{q},\mathsf{U})\in\mathfrak{Q}\;\;M>0\;\exists\,K_M(\mathsf{q},\mathsf{U})>0\;\forall\,(\boldsymbol{c},u)\in\mathfrak{S}(\mathsf{q},\mathsf{U}):\\&\qquad\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u)\leq M\;\Longrightarrow\;\mathcal{P}(\boldsymbol{c},u)\geq K_M(\mathsf{q},\mathsf{U})\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u),\end{aligned} \tag{1.8}$$

2 Thermodynamical Modeling

Here, we describe the general approach to obtain thermodynamically consistent models for ERDS. We start from a concave entropy and derive the coupled system of PDEs as a formal gradient system as introduced in Mielke (2011a, 2013). Using the Boltzmann entropy for the concentrations, it is possible to reproduce the exact mass action kinetics, if it satisfies a detailed balance condition, which we assume throughout this work. The detailed balance seems to be necessary for the gradient structure; however, the uniform exponential decay estimates can be obtained under the weaker notion of complex balancing, see Desvillettes et al. (2017) and Mielke (2017). Indeed, it would be straightforward to generalize the present work on ERDS to the case of complex-balanced kinetics, but we refrain from doing so in the interest of a simpler notation.

2.1 The Entropy Functional

We study general systems that are thermodynamically consistent, in the sense that they can be described as a gradient system in terms of the total entropy. We allow for I different species $X_1, \dots, X_I$, whose concentrations denoted by $c_i(t,x) \geq 0$ form the density vector $\boldsymbol{c} = (c_1, \dots, c_I) : \Omega \to [0, \infty[^I$. In addition, we have the scalar internal energy density $u(t,x) \in \mathbb{R}$. The total entropy $\mathcal{S}$ is the integral over the entropy density $s = S(x, \boldsymbol{c}, u)$, where $S(x, \cdot)$ is a strictly concave function of $(\boldsymbol{c}, u) \in [0, \infty[^I \times \mathbb{R}$, see e.g., Lieb and Yngvason (1999). Thus, we set

$$\mathcal{S}(\boldsymbol{c}, u) = \int_\Omega S(x, \boldsymbol{c}(x), u(x))\, \mathrm{d}x.$$

The absolute temperature is now a dependent variable given by

$$\theta(t,x) = \Theta(x, \boldsymbol{c}, u) := \frac{1}{\partial_u S(x, \boldsymbol{c}(t,x), u(t,x))}. \tag{2.1}$$

At a given point $x \in \Omega$ and for a fixed internal energy, the equilibrium densities are obtained by maximizing the $\boldsymbol{c} \mapsto S(x, \boldsymbol{c}, u)$ providing the vector

$$\boldsymbol{w}(x,u) = \big(w_1(x,u), \dots, w_I(x,u)\big) \in \,]0, \infty[^I,$$

i.e., $\mathrm{D}_{\boldsymbol{c}} S(x, \boldsymbol{w}(x,u), u) = 0$. One of the major difficulties in this work is that the vector $\boldsymbol{w}$ depends on u, i.e., the local equilibrium concentrations depend on u, and hence implicitly on θ.

So far, we also allowed for an explicit dependence on the material point $x \in \Omega$. Such a dependence is crucial for ERDS posed inside of solids, e.g., in semiconductor physics where c_i describes the concentrations of charge carriers. The dependence on x then models the domains of the different materials including doping structures. However, for our work this dependence is not essential, so we will drop it for notational simplicity. Following Glitzky and Hünlich (1997), it should be possible to adapt the subsequent analysis to the general case by using suitable relative densities.

For later usage, we give a suitable class of entropy functions based on the Boltzmann entropy using the function $\lambda_{\mathfrak{B}}(z) := z \log z - z + 1$:

$$S(\boldsymbol{c}, u) = s(u) - \sum_{i=1}^I w_i(u) \lambda_{\mathfrak{B}}\Big(\frac{c_i}{w_i(u)}\Big) = \widehat{s}(u) + \sum_{i=1}^I \Big(c_i \log\big(w_i(u)\big) - \lambda_{\mathfrak{B}}(c_i)\Big),$$
$$\text{where } \widehat{s}(u) = s(u) + I - \sum_{i=1}^I w_i(u). \tag{2.2}$$

The following result gives necessary and sufficient conditions on s and w_i for the concavity of this function S.

Proposition 2.1 (Concavity and monotonicity of S) *Consider smooth functions* $s : [0,\infty[\to \mathbb{R}$ *and* $w_i : [0,\infty[\to [0,\infty[$ *and define* $S : [0,\infty[^{I+1} \to \mathbb{R}$ *as in* (2.2). *Then,*

a. *S is concave if and only if*

$$w_1, w_2, \ldots, w_I, \text{ and } \widehat{s} \text{ are concave}; \tag{2.3a}$$

b. *S is non-decreasing in u if and only if*

$$w_1, w_2, \ldots, w_I, \text{ and } \widehat{s} \text{ are increasing}. \tag{2.3b}$$

Proof We simply calculate the second derivative, which yields

$$\mathrm{D}^2 S(\boldsymbol{c},u) = \begin{pmatrix} -\frac{1}{c_1} & 0 & \cdots & 0 & \frac{w_1'(u)}{w_1(u)} \\ 0 & -\frac{1}{c_2} & 0 & & \frac{w_2'(u)}{w_2(u)} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -\frac{1}{c_I} & \frac{w_I'(u)}{w_I(u)} \\ \frac{w_1'(u)}{w_1(u)} & \frac{w_2'(u)}{w_2(u)} & \cdots & \frac{w_I'(u)}{w_I(u)} & \partial_u^2 S(\boldsymbol{c},u) \end{pmatrix} \tag{2.4}$$

with $\partial_u^2 S(\boldsymbol{c},u) = \widehat{s}''(u) + \sum_{i=1}^I c_i \frac{w_i(u)w_i''(u)-(w_i'(u))^2}{(w_i(u))^2} < 0$. In particular, we find

$$\begin{pmatrix}\boldsymbol{\zeta}\\ \tau\end{pmatrix} \cdot \mathrm{D}^2 S(\boldsymbol{c},u) \begin{pmatrix}\boldsymbol{\zeta}\\ \tau\end{pmatrix} = \left(\widehat{s}''(u) + \sum_{i=1}^I c_i \frac{w_i''(u)}{w_i(u)}\right)\tau^2 - \sum_{i=1}^I c_i \left(\frac{\zeta_i}{c_i} + \frac{w_i'(u)}{w_i(u)}\tau\right)^2,$$

and (2.3a) follows immediately since c_i and u may range freely in $[0,\infty[$.

For (2.3b) part (b), we simply use $\partial_u S(\boldsymbol{c},u) = \widehat{s}'(u) + \sum_{i=1}^I c_i w_i'(u)/w_i(u)$ and the fact that c_i and u may range freely in $[0,\infty[$. □

The simplest example of a suitable entropy functional $S(\boldsymbol{c},u)$ (cf. Mielke 2013, (3.9) or Haskovec et al. 2017, Sect. 4.3) is given by the choices

$$\widehat{s}(u) = s_0 \log u \quad \text{and } w_i(u) = r_i u^{\rho_i} \text{ with } s_0, r_i > 0 \text{ and } \rho_i \in [0,1].$$

With $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_I)$ we then have the explicit relations

$$S(\boldsymbol{c},u) = (s_0 + \boldsymbol{\rho}\cdot\boldsymbol{c}) \log u + \widetilde{\sigma}(\boldsymbol{c}) \text{ with } \widetilde{\sigma}(\boldsymbol{c}) := \sum_{i=1}^I \Big(c_i \log r_i - \lambda_{\mathfrak{B}}(c_i)\Big) \text{ and}$$

$$\theta = \Theta(\boldsymbol{c},u) = u/(s_0 + \boldsymbol{\rho}\cdot\boldsymbol{c}), \quad u = U(\boldsymbol{c},\theta) = (s_0 + \boldsymbol{\rho}\cdot\boldsymbol{c})\theta.$$

Moreover, we may express the entropy and the free energy $\psi = u - \theta s$ in $\boldsymbol{c}$ and θ, which gives, using the shorthand $c^{\boldsymbol{\rho}} := s_0 + \boldsymbol{\rho}\cdot\boldsymbol{c} > 0$,

$$\overline{S}(\boldsymbol{c},\theta) = S(\boldsymbol{c}, U(\boldsymbol{c},\theta)) = c^{\boldsymbol{\rho}}\big(\log\theta + \log c^{\boldsymbol{\rho}}\big) + \widetilde{\sigma}(\boldsymbol{c}),$$
$$\overline{\psi}(\boldsymbol{c},\theta) = U(\boldsymbol{c},\theta) - \theta\overline{S}(\boldsymbol{c},\theta) = c^{\boldsymbol{\rho}}\big(1 - \lambda_{\mathfrak{B}}(\theta)\big) - \theta\big(\widetilde{\sigma}(\boldsymbol{c}) + c^{\boldsymbol{\rho}} \log c^{\boldsymbol{\rho}}\big).$$

It can be shown that $\overline{S}$ is concave if and only if $\rho_i \in [0, 1/2]$ for all i. Similarly, $\overline{\psi}(\cdot, \theta)$ is convex if and only if $\rho_i \in [0, 1/2]$ for all i. This shows that θ is not a good variable to describe ERDS.

This family can be generalized substantially by choosing a continuously differentiable function $V : [0, \infty[\to]0, \infty[$ such that

$$V'(u) \geq \max\{ \rho_i \mid i = 1, \ldots, I \} \text{ for all } u > 0.$$

Then we define $\phi : [0, \infty[\to]0, \infty[$ via $\phi'(u) = \phi(u)/V(u)$, $\phi(0) = \phi_0 \geq 0$ and set

$$w_i(u) = r_i \phi(u)^{\rho_i} \text{ for } i = 1, \ldots, I \quad \text{and } \widehat{s}(u) = s_0 \log\big(\phi(u)\big),$$

with r_i and ρ_i as above. The monotonicity of w_i and $\widehat{s}$ is obvious, and the concavities follow from

$$w_i''(u) \leq 0 \iff (1-\rho_i)(\phi')^2 \geq \phi\phi'' \iff V' \geq \rho_i.$$

The result for $\widehat{s}$ is obtained in the same way but with $\rho = 0$. The special form with the same ϕ for each w_i leads to the explicit relation $V(u) = \big(s_0 + \boldsymbol{\rho}\cdot\boldsymbol{c}\big)\theta$, which now may be nonlinear, but is still monotone.

Remark 2.2 (Other entropies) Our choice of entropies is different from the choice of entropies induced from free energies $F(\boldsymbol{c}, \theta)$ given in the separated from

$$\widehat{F}(\boldsymbol{c}, \theta) = \psi(\theta) + \sum\nolimits_{i=1}^{I} \big(E_i(\theta)c_i + \theta\lambda_{\mathfrak{B}}(c_i)\big),$$

cf. e.g., Albinus et al. (2002). With $\overline{S}(\boldsymbol{c}, \theta) = -\partial_\theta \overline{F}(\boldsymbol{c}, \theta)$ and $\overline{U} = F + \theta\overline{S}$ we easily find

$$\overline{S}(\boldsymbol{c}, \theta) = -\psi'(\theta) - \sum_{i=1}^{I} \big(E_i'(\theta)c_i + \lambda_{\mathfrak{B}}(c_i)\big) \text{ and } \overline{U}(\boldsymbol{c}, \theta) = \psi(\theta) - \theta\psi'(\theta) + \sum_{i=1}^{I} (E_i(\theta) - \theta E_i'(\theta))c_i.$$

Upon solving $u = \overline{U}(\boldsymbol{c}, \theta)$ for $\theta = \Theta(\boldsymbol{c}, u)$, we obtain $S(\boldsymbol{c}, u) = \overline{S}(\boldsymbol{c}, \Theta(\boldsymbol{c}, u))$, but in general we will lose the property $\partial_{c_i}\partial_{c_j} S(\boldsymbol{c}, u) = 0$ for $i \neq j$. However, the diagonal structure of $\mathrm{D}_{\boldsymbol{c}}^2 S(\boldsymbol{c}, u)$ was useful above.

To see the non-diagonal structure, we consider $\nu > 1$ and let

$$\psi(\theta) = -\frac{s_0}{\nu-1}\theta^\nu \quad \text{and} \quad E_i(\theta) = a_i - \frac{b_i}{\nu-1}\theta^\nu,$$

where $s_0 > 0$ and $a_i, b_i \geq 0$. From this, we obtain the explicit relations

$$\Theta(\boldsymbol{c}, u) = \Big(\frac{u - \boldsymbol{a}\cdot\boldsymbol{c}}{s_0 + \boldsymbol{b}\cdot\boldsymbol{c}}\Big)^{1/\nu} \quad \text{and} \quad S(\boldsymbol{c}, u) = \frac{\nu}{\nu-1}(u - \boldsymbol{a}\cdot\boldsymbol{c})^{1-1/\nu}(s_0 + \boldsymbol{b}\cdot\boldsymbol{c})^{1/\nu} - \sum_{i=1}^{I} \lambda_{\mathfrak{B}}(c_i).$$

If at least two of the coefficients a_i and b_i are positive, then $\mathrm{D}^2_{\boldsymbol{c}} S(\boldsymbol{c}, u)$ is non-diagonal.

2.2 Diffusion and Heat Transfer

Our ERDS will be of the general form

$$\begin{pmatrix} \dot{\boldsymbol{c}} \\ \dot{u} \end{pmatrix} = \operatorname{div}\Big(M(\boldsymbol{c}, u)\nabla\begin{pmatrix} \boldsymbol{c} \\ u \end{pmatrix}\Big) + \begin{pmatrix} \boldsymbol{R}(\boldsymbol{c}, u) \\ 0 \end{pmatrix} \text{ in } \Omega, \qquad \nabla\begin{pmatrix} \boldsymbol{c} \\ u \end{pmatrix}\cdot\nu = 0 \text{ on } \partial\Omega. \quad (2.5)$$

Here $M(\boldsymbol{c}, u)$ could be a tensor of fourth order to model anisotropy, however, only consider $M(\boldsymbol{c}, u) \in \mathbb{R}^{(I+1)\times(I+1)}$ such that M contains the mobilities of the species X_i with densities c_i, the heat conductivity $M_{I+1,I+1}$ as well as the cross-terms $M_{i,I+1}$ and $M_{I+1,i}$. Of course, we also allow for cross-diffusion between the species, i.e., M_{ij} may be nonzero for $i \neq j$. One way of generating suitable matrices M is by using a diffusion Onsager operator

$$\mathbb{K}_{\text{diff}}(\boldsymbol{c}, u)\begin{pmatrix} \boldsymbol{\zeta} \\ \mu \end{pmatrix} := -\operatorname{div}\Big(\mathbb{M}(\boldsymbol{c}, u)\nabla\begin{pmatrix} \boldsymbol{\zeta} \\ \mu \end{pmatrix}\Big), \text{ where } \mathbb{M}(\boldsymbol{c}, u) = \mathbb{M}(\boldsymbol{c}, u)^\top \geq 0.$$

Then, we generate the matrix M by applying $\mathbb{K}_{\text{diff}}$ to $\mathcal{S}$, namely

$$\begin{pmatrix} \dot{\boldsymbol{c}} \\ \dot{u} \end{pmatrix} = \operatorname{div}\Big(M(\boldsymbol{c}, u)\nabla\begin{pmatrix} \boldsymbol{c} \\ u \end{pmatrix}\Big) = \mathbb{K}(\boldsymbol{c}, u)\mathrm{D}\mathcal{S}(\boldsymbol{c}, u) \quad \text{in } \Omega. \quad (2.6)$$

By using that $\binom{\boldsymbol{\zeta}}{\mu} = \mathrm{D}\mathcal{S}(\boldsymbol{c}, u) = \binom{\mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c})}{\partial_u S(\boldsymbol{c}, u)}$, we find the relation

$$M(\boldsymbol{c}, u) = \mathbb{M}(\boldsymbol{c}, u)\big(-\mathrm{D}^2 S(\boldsymbol{c}, u)\big).$$

Hence M may be non-symmetric, but it is the product of two symmetric matrices, which implies that it is still diagonalizable with real eigenvalues.

Of course, the simplest choice for $\mathbb{M}(\boldsymbol{c}, u)$, which is often used nevertheless, is to take $\mathbb{M}$ as a scalar multiple δ of $\big(-\mathrm{D}^2 S(\boldsymbol{c}, u)\big)^{-1}$ like in the case of Wasserstein diffusion, see Jordan et al. (1998), Otto (2001) and Liero and Mielke (2013). It is still an open problem to construct sufficiently large classes of mobility tensors $\mathbb{M}$, such that $M = \mathbb{M}(-\mathrm{D}^2\mathcal{S})$ is sufficiently simple to allow for existence results for solutions.

In the present work, we do not discuss the structure of $\mathbb{M}$, since we do not need the pure diffusion equation (2.5) with $\boldsymbol{R} \equiv 0$ to be a gradient system with respect to $\mathcal{S}$. The Onsager form in (2.6) and the choice $M = -\mathbb{M}\mathrm{D}^2 S$ is just an easy way to generate a large class of admissible mobilities. Our focus concerns the entropy production generated by M along solutions of $\dot{\boldsymbol{y}} = \operatorname{div}\big(M(\boldsymbol{y})\nabla \boldsymbol{y}\big)$, viz.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\boldsymbol{y}) = \int_\Omega \mathrm{D}S(\boldsymbol{y})\cdot\dot{\boldsymbol{y}}\,\mathrm{d}x = \int_\Omega \nabla\boldsymbol{y}\cdot\big(-M(\boldsymbol{y})\mathrm{D}^2S(\boldsymbol{y})\big)\nabla\boldsymbol{y}\,\mathrm{d}x =: \mathcal{P}_{\mathrm{diff}}(\boldsymbol{y}). \quad (2.7)$$

Thus, we have to choose M such that $\mathcal{P}_{\mathrm{diff}}(\boldsymbol{y}) \geq 0$ for all $\boldsymbol{y}$. This is clearly implied by the condition that the symmetric part of $-M(\boldsymbol{y})\mathrm{D}^2S(\boldsymbol{y})$ is positive semi-definite. However, using suitable integration by parts (which is equivalent to exploit null-Lagrangians) there may be more cases, see e.g., the theory developed in Jüngel and Matthes (2006).

To obtain our desired uniform exponential decay estimates, we need a suitable quantitative lower bound $\mathcal{P}_{\mathrm{diff}}(\boldsymbol{y}) \geq \delta_*\widehat{\mathcal{P}}_\gamma(\boldsymbol{y})$, for which we choose

$$\widehat{\mathcal{P}}_\gamma(\boldsymbol{c},u) := \int_\Omega\left(\frac{|\nabla u|^2}{u^{2-\gamma}} + \sum_{i=1}^I\frac{|\nabla c_i|^2}{c_i}\right)\mathrm{d}x, \quad \text{where } \gamma\in[0,1[. \quad (2.8)$$

The parameter γ for the internal energy relates to the fact that typical entropy functions s are of the form $s(u) = s_1u^\gamma$ for $\gamma\in]0,1[$ or $s(u) = s_0\log u$ for $\gamma = 0$, while the Boltzmann entropy $\lambda_{\mathfrak{B}}(z) = z\log z - z + 1$ for the densities c_i corresponds to the case $\gamma = 1$.

The following result provides conditions on the entropy S in (2.2) and a constant diagonal matrix M such that $\mathcal{P}_{\mathrm{diff}}$ can be bounded from below by $\widehat{\mathcal{P}}_\gamma$.

Proposition 2.3 (Entropy production by linear diffusion) *Assume that S is given by (2.2) with w_i and $\widehat{s}$ strictly increasing and concave and that there exists $\gamma, \rho_i\in]0,1[$ for $i = 1,\ldots,I$ and $\sigma > 0$ such that*

$$\forall\, i,\ u>0:\quad -w_i(u)w_i''(u) \geq \frac{1-\rho_i}{\rho_i}\big(w_i'(u)\big)^2 \quad \textit{and} \quad -\widehat{s}''(u) \geq \frac{\widehat{s}_0}{u^{2-\gamma}}. \quad (2.9)$$

Then, for the constant diagonal matrix $M(\boldsymbol{y}) = \mathrm{diag}(\delta_1,\ldots,\delta_I,\delta_\mathrm{u})$ with $\delta_i,\delta_\mathrm{u} > 0$ the entropy production $\mathcal{P}_{\mathrm{diff}}$ defined in (2.7) satisfies the lower bound

$$\mathcal{P}_{\mathrm{diff}}(\boldsymbol{c},u) \geq \delta_*\widehat{\mathcal{P}}_\gamma(\boldsymbol{c},u) \quad \textit{with } \delta_* = \min\left\{\delta_\mathrm{u}\widehat{s}_0,\ \left(\delta_i - \frac{\rho_i(\delta_i+\delta_\mathrm{u})^2}{4\delta_\mathrm{u}}\right)\ \middle|\ i=1,\ldots,I\right\} \quad (2.10)$$

for all $(\boldsymbol{c},u)$ with $\sqrt{c_i}, u^{\gamma/2}\in\mathrm{H}^1(\Omega)$.

Proof An explicit calculation using D^2S from (2.4) gives $\mathcal{P}_{\mathrm{diff}}(\boldsymbol{y}) = \int_\Omega Q(\boldsymbol{y};\nabla\boldsymbol{y})\,\mathrm{d}x$ with

$$\begin{aligned} Q(\boldsymbol{c},u;\nabla c,\nabla u) &= -\delta_\mathrm{u}\widehat{s}''|\nabla u|^2 \\ &\quad + \sum_{i=1}^I\left(\delta_i\frac{|\nabla c_i|^2}{c_i} - (\delta_i+\delta_\mathrm{u})\frac{w_i'}{w_i}\nabla u\cdot\nabla c_i + \delta_\mathrm{u}c_i\frac{w_i'^2 - w_iw_i''}{w_i^2}|\nabla u|^2\right), \end{aligned}$$

where we dropped the argument u from $\widehat{s}$ and w_i for notational simplicity.

Abbreviating $a_i = w_i'/w_i$ and using the assumption (2.9), we obtain the lower bound

$$\begin{aligned} Q(\boldsymbol{c}, u; \nabla c, \nabla u) &\geq \frac{\delta_\mathrm{u}\widehat{s}_0}{u^{2-\gamma}}|\nabla u|^2 + \sum_{i=1}^{I}\left(\delta_i \frac{|\nabla c_i|^2}{c_i} - (\delta_i + \delta_\mathrm{u}) a_i \nabla u \cdot \nabla c_i + \frac{\delta_\mathrm{u} c_i a_i^2}{\rho_i}|\nabla u|^2\right) \\ &\geq \frac{\delta_\mathrm{u}\widehat{s}_0}{u^{2-\gamma}}|\nabla u|^2 + \sum_{i=1}^{I}\left(\delta_i - \frac{\rho_i(\delta_i+\delta_\mathrm{u})^2}{4\delta_\mathrm{u}}\right)\frac{|\nabla c_i|^2}{c_i}, \end{aligned} \tag{2.11}$$

where we estimated by Young's inequality each term in the sum individually. This proves the assertion. □

Remark 2.4 (Dropped mixed terms) In estimate (2.11) we dropped a positive term to eliminate the mixed terms $\nabla u \cdot \nabla c_i$. Dropping the index i for simplicity and letting

$$\widetilde{\kappa} = \frac{2\delta_\mathrm{u}}{\rho_i(\delta_i+\delta_\mathrm{u})}, \quad \widetilde{\rho} = \frac{\rho_i(\delta_i+\delta_\mathrm{u})^2}{4\delta_\mathrm{u}}, \quad \text{and} \quad A(u) = w_i(u)^{\widetilde{\kappa}}$$

this term takes the form

$$\begin{aligned} &\widetilde{\rho}\left(\frac{|\nabla c|^2}{c} - 2A'(u)\nabla u \cdot \nabla c + cA'(u)^2|\nabla u|^2\right) = c\widetilde{\rho}\big|\nabla(\log c) - \nabla(\log A(u))\big|^2 \\ &= c\widetilde{\rho}\,|\nabla \log(c/A(u))|^2 = \frac{\widetilde{\rho}}{\alpha^2}\, c^{1-2\alpha} A(u)^{2\alpha}\left|\nabla\left(\frac{c^\alpha}{A(u)^\alpha}\right)\right|^2, \end{aligned}$$

where $\alpha > 0$ is arbitrary. For a way that this term is exploited, we refer to Haskovec et al. (2017), Prop. 6.2, where $w(u) = u^{1/2}$, $\delta_\mathrm{u} = \delta$ and $\rho = \alpha = 1/2$ giving $A(u) = u$.

The estimate (2.10) is of course only useful for $\delta_* > 0$. For the case $M = \delta I_{I+1}$, this is the case with $\delta_* = \delta \min\{\, 1-\rho_i \mid i = 1, \ldots, I\,\}$. If the diffusion constants are different, we obtain restrictions on the ρ_i, namely we need

$$\rho_i < \frac{4\delta_i\delta_\mathrm{u}}{(\delta_i+\delta_\mathrm{u})^2} \leq 1. \tag{2.12}$$

This provides an interesting connection between the entropy structure and the diffusion mechanisms. Indeed, this connection may be attributed to cross-diffusion, if we take into account that the temperature is defined via $\theta = 1/\partial_u S(\boldsymbol{c}, u)$. In thermodynamical modeling, it is usually assumed that heat flow is not driven by ∇u but rather by $\nabla\theta$ (or more precisely by $\nabla(1/\theta) = \nabla(\partial_u S(\boldsymbol{c}, u))$, see e.g., Albinus et al. (2002) and Mielke (2013). Thus, our system written in terms of $(\boldsymbol{c}, u)$ must have cross-diffusion when written in terms of $(\boldsymbol{c}, \theta)$. More precisely, for S in (2.2) we have

$$\nabla \mathrm{D}S(\boldsymbol{c}, u) = \begin{pmatrix} -\frac{1}{c_i}\nabla c_i + a_i(u)\nabla u \\ \big(\widehat{s}''(u) + \sum_1^I c_i a_i'(u)\big)\nabla u + \sum_1^I a_i(u)\nabla c_i \end{pmatrix} \quad \text{with } a_i(u) = \frac{w_i'(u)}{w_i(u)}.$$

To generate a term that is always proportional to ∇u, we need to take a very specific combination of the components of $\nabla \mathrm{D}S(\boldsymbol{c}, u)$, namely

$$\nabla\big(\partial_u S(\boldsymbol{c}, u)\big) + \sum_{i=1}^{I} a_i c_i \nabla\big(\partial_{c_i} S(\boldsymbol{c}, u)\big) = \left(\widehat{s}''(u) + \sum_{i=1}^{I} c_i \frac{w_i''(u)}{w_i(u)}\right) \nabla u.$$

In turn, this implies that the linear decoupled diffusion system $\dot{\boldsymbol{y}} = M \Delta \boldsymbol{y}$ with the constant matrix $M = \mathrm{diag}(\delta_1, \ldots, \delta_I, \delta_\mathrm{u})$ can be written in the Onsager form (2.6), i.e.,

$$\begin{aligned} \begin{pmatrix} \dot{\boldsymbol{c}} \\ \dot{u} \end{pmatrix} = M \begin{pmatrix} \Delta \boldsymbol{c} \\ \Delta u \end{pmatrix} = \begin{pmatrix} (\delta_i \Delta c_i)_{1,\ldots,I} \\ \delta_\mathrm{u} \Delta u \end{pmatrix} &= \mathbb{K}(\boldsymbol{c}, u) \mathrm{D}\mathcal{S}(\boldsymbol{c}, u) \\ &= -\,\mathrm{div}\Big(\mathbb{M}(\boldsymbol{c}, u) \nabla\big(\mathrm{D}S(\boldsymbol{c}, u)\big)\Big), \end{aligned}$$

if the mobility matrix $\mathbb{M}(\boldsymbol{c}, u) \in \mathbb{R}^{(I+1)\times(I+1)}$ satisfies the relation

$$\mathbb{M}(\boldsymbol{c}, u)\big(0, \ldots, 0, 1\big)^\top = \widetilde{m}(\boldsymbol{c}, u)\big(c_1 a_1(u), \ldots, c_I a_I(u), 1\big)^\top$$

for a suitable scalar function $\widetilde{m}$. Recall that the case $w_i(u) = r_i u^{\rho_i}$ leads to $a_i(u) = \rho_i / u$, i.e., we have

$$\mathbb{M}_{i,I+1}(\boldsymbol{c}, u) = \widetilde{m}(\boldsymbol{c}, u) \frac{\rho_i c_i}{u} \quad \text{for } i = 1, \ldots, I.$$

Thus, we may interpret the condition (2.12) on ρ_i as a condition on the strength of cross-diffusion.

Remark 2.5 We may check whether $M = \mathrm{diag}(\delta_1, \ldots, \delta_\mathrm{u})$ can actually be generated by an Onsager operator with $\mathbb{M}$. For this, we use the block inversion formula

$$\left(\begin{array}{c|c} \boldsymbol{C}^{-1} & \boldsymbol{C}^{-1}\boldsymbol{a} \\ \hline (\boldsymbol{C}^{-1}\boldsymbol{a})^\top & \sigma_0 \end{array}\right)^{-1} = \left(\begin{array}{c|c} \boldsymbol{C}+\mu \boldsymbol{a}\otimes\boldsymbol{a} & -\mu \boldsymbol{a} \\ \hline -\mu \boldsymbol{a}^\top & \mu \end{array}\right) \text{ with } \frac{1}{\mu} = \sigma_0 - \boldsymbol{a}\cdot\boldsymbol{C}^{-1}\boldsymbol{a}.$$

Thus, using (2.4), we can calculate $\mathbb{M}(\boldsymbol{c}, u) = M\big(-\mathrm{D}^2 S(\boldsymbol{c}, u)\big)^{-1}$ explicitly, which leads to

$$\mathbb{M}(\boldsymbol{c}, u) = \begin{pmatrix} \delta_1 c_1 + \mu\big(\frac{c_1 w_1'}{w_1}\big)^2 & \cdots & \mu \frac{c_1 c_I w_1' w_I'}{w_1 w_I} & \delta_1 \mu c_1 \frac{w_1'}{w_1} \\ \vdots & \ddots & \vdots & \vdots \\ \mu \frac{c_I c_1 w_I' w_1'}{w_I w_1} & \cdots & \delta_I c_I + \mu\big(\frac{c_I w_I'}{w_I}\big)^2 & \delta_I \mu c_I \frac{w_I'}{w_I} \\ \delta_\mathrm{u} \mu c_1 \frac{w_1'}{w_1} & \cdots & \delta_\mathrm{u} \mu c_I \frac{w_I'}{w_I} & \delta_\mathrm{u} \mu \end{pmatrix}$$

with $\frac{1}{\mu} = -\widehat{s}''(u) - \sum_{i=1}^{I} c_i \frac{w_i''(u)}{w_i(u)} > 0$. Obviously, this matrix is only symmetric if $\delta_i = \delta_\mathrm{u}$ for all i; however, one can check that the symmetric part of $\mathbb{M}$ is positive definite under the conditions for which δ_* in (2.10) is positive.

It remains a challenging open question to find suitable symmetric and positive semi-definite mobilities $\mathbb{M}(\boldsymbol{c}, u)$ and suitable entropies $S(\boldsymbol{c}, u)$ such that the induced diffusion matrix $M(\boldsymbol{c}, u) = -\mathbb{M}(\boldsymbol{c}, u)\mathrm{D}^2 S(\boldsymbol{c}, u)$ is nontrivial but still simple enough to allow for an existence theory for solutions. Some results are available for the isothermal case (e.g., Burger et al. 2010; Liero and Mielke 2013), but nothing exists in the non-isothermal case.

2.3 Mass Action Reaction Kinetics

Next, we discuss the reaction part $\dot{\boldsymbol{c}} = \boldsymbol{R}(\boldsymbol{c}, u)$ in the ERDS (2.5), which is a simple ODE, where the internal energy $u > 0$ acts as a parameter. Thus, the theory for the kinetics is very close to the isothermal one, if we take care properly of the parametric dependence on u.

First, we observe that for a fixed u our entropy $\boldsymbol{c} \mapsto \mathcal{S}(\boldsymbol{c}, u)$ in (2.2) is, up to a constant, a classical Boltzmann entropy with the thermal equilibrium $\boldsymbol{c} = \boldsymbol{w}(u)$. Thus, our basic assumption is that $\mathcal{S}(\cdot, u)$ is a Lyapunov function for $\dot{\boldsymbol{c}} = \boldsymbol{R}(\boldsymbol{c}, u)$, i.e.,

$$P_\mathrm{react}(\boldsymbol{c}, u) := \boldsymbol{R}(\boldsymbol{c}, u) \cdot \mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c}, u) \geq 0.$$

Indeed, as for the diffusion part, we want to quantify this relation to obtain exponential decay.

Second, we specify the reaction terms further and assume mass action kinetics with R reaction pairs for the species $X_1, \ldots, X_I$ given by

$$\alpha_1^r X_1 + \cdots \alpha_I^r X_I \underset{k_\mathrm{bw}^r}{\overset{k_\mathrm{fw}^r}{\rightleftharpoons}} \beta_1^r X_1 + \cdots \beta_I^r X_I.$$

Here $\alpha_i^r, \beta_i^r \in \mathbb{N}_0$ are the stoichiometric coefficients and k_fw^r and k_bw^r are the forward and backward reaction rate coefficients of the rth reaction, which may depend on u or more generally on $\boldsymbol{c}, u)$, see Remark 2.6. Defining the stoichiometric vectors $\boldsymbol{\alpha}^r := (\alpha_i^r)_i$ and $\boldsymbol{\beta}^r := (\beta_i^r)_i$ in $\mathbb{N}_0^I$ the associated reaction rate equation reads

$$\dot{\boldsymbol{c}} = \boldsymbol{R}(c, u) = \sum_{r=1}^{R} \big(k_\mathrm{fw}^r(u)\boldsymbol{c}^{\boldsymbol{\alpha}^r} - k_\mathrm{bw}^r(u)\boldsymbol{c}^{\boldsymbol{\beta}^r}\big)\big(\boldsymbol{\beta}^r - \boldsymbol{\alpha}^r\big).$$

Third, we assume the *detailed balance condition* for the R reaction pairs

$$\text{(DBC)} \quad \forall\, r = 1, \ldots, R \;\forall\, u > 0 : \;\; k_\mathrm{fw}^r(u)\big(\boldsymbol{w}(u)\big)^{\boldsymbol{\alpha}^r} = k_\mathrm{bw}^r(u)\big(\boldsymbol{w}(u)\big)^{\boldsymbol{\beta}^r} =: \kappa_r(u), \tag{2.13}$$

which means that all the reactions pairs are simultaneously in balance at the thermodynamic equilibrium $\boldsymbol{w}(u)$. We emphasize that this condition is stronger than needed: following Desvillettes et al. (2017), Mielke (2017) it should be possible to generalize our arguments to the case of complex-balanced reaction rate equation, where still a quantitative estimate between the entropy production and the entropy is possible. For simplicity, we stay with the DBC (2.13) and follow the approach developed in Mielke (2011a, 2013) and Haskovec et al. (2017). Using $\mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c}, u) = -\big(\log(c_i/w_i(u))\big)_{i=1,\ldots,I}$ we obtain the special form

$$\dot{\boldsymbol{c}} = \boldsymbol{R}(\boldsymbol{c}, u) = \sum_{r=1}^{R} \kappa_r(u)\Big(\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\alpha}^r}} - \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\beta}^r}}\Big)\big(\boldsymbol{\beta}^r - \boldsymbol{\alpha}^r\big) \;=\; \mathbb{L}(\boldsymbol{c}, u)\mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c}, u), \tag{2.14}$$

where the monomials $\boldsymbol{c}^{\boldsymbol{\gamma}}$ are defined via $\prod_{i=1}^{I} c_i^{\gamma_i}$ and where the Onsager matrix $\mathbb{L} \in \mathbb{R}^{I\times I}$ is a sum of rank-one contributions for each reaction, viz.

$$\mathbb{L}(\boldsymbol{c}, u) = \sum_{r=1}^{R} \kappa_r(u)\Lambda\left(\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\alpha}^r}}, \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\beta}^r}}\right)\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \otimes \big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \text{ with } \Lambda(a, b) = \frac{a-b}{\log(a/b)}.$$

We now define the entropy production through reaction via

$$P_{\text{react}}(\boldsymbol{c}, u) := -\boldsymbol{R}(\boldsymbol{c}, u)\cdot\big(\log(c_i/w_i(u))\big)_i = \mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c}, u)\cdot\mathbb{L}(\boldsymbol{c}, u)\mathrm{D}_{\boldsymbol{c}} S(\boldsymbol{c}, u) \geq 0. \tag{2.15}$$

Remark 2.6 (Arrhenius-type reaction rate coefficients) It is also possible to consider reaction rate coefficients $\kappa_r > 0$ depending on $\boldsymbol{c}$ and u, which leads to an implicit dependence on the temperature, e.g., $\kappa_r(\boldsymbol{c}, u) = A_r \exp\big(-E_r\partial_u S(\boldsymbol{c}, u)/k_{\mathfrak{B}}\big) = A_r \exp\big(\frac{-E_r}{\mathsf{R}\theta}\big)$ with the activation energy E_r and Boltzmann constant $k_{\mathfrak{B}}$, see Bothe and Dreyer (2015) and Dreyer et al. (2016).

Fourth, we remark that $\mathbb{L}$ usually has lower rank, because the stoichiometric subspace

$$\mathbb{S} := \mathrm{span}\{\,\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r \mid r = 1, \ldots, R\,\} \subset \mathbb{R}^I$$

may have a dimension $m_S := \dim\mathbb{S}$ less than I. In that case (2.14) has natural conserved quantities (often called conservation of atomic masses) which can be described by a matrix $\mathsf{Q} : \mathbb{R}^I \to \mathbb{R}^m$ where $m = I - m_S$ such that

$$\mathrm{kernel}(\mathsf{Q}) = \mathbb{S} \quad\text{and}\quad \mathrm{range}\big(\mathsf{Q}^\top\big) = \mathbb{S}^\perp := \{\,\boldsymbol{\xi} \in \mathbb{R}^I \mid \boldsymbol{\xi}\cdot\boldsymbol{v} = 0 \text{ for } \boldsymbol{v} \in \mathbb{S}\,\}.$$

Obviously, solutions of (2.14) satisfy $\mathsf{Q}\boldsymbol{c}(t) = \mathsf{Q}\boldsymbol{c}(0)$ and thus stay inside the convex sets

$$\boldsymbol{C}_{\mathsf{q}} := \boldsymbol{c}(0) + \mathbb{S} \cap [0, \infty[^I, \quad \text{where } \mathsf{q} := \mathsf{Q}\boldsymbol{c}(0).$$

For later usage, we also define $\mathfrak{Q} \subset \mathbb{R}^m \times]0, \infty[$ as the set of all relevant conservation values, namely

$$\mathfrak{Q} := \{ (\mathsf{q}, \mathsf{U}) = (\mathsf{Q}\boldsymbol{c}, u) \in \mathbb{R}^{m+1} \mid (\boldsymbol{c}, u) \in [0, \infty[^I \times]0, \infty[\}. \tag{2.16}$$

Note that $\mathsf{U} = 0$ is the case of zero energy, which implies $u(t, x) = 0$ for all t and x, which reduces to the case treated in Desvillettes and Fellner (2006), Desvillettes et al. (2017), Fellner and Tang (2017) and Mielke (2017). Hence, it is no restriction to assume $\mathsf{U} > 0$ subsequently.

Finally, we remark that for each fixed u and $\mathsf{q} = \mathsf{Q}\boldsymbol{c}(0)$ there exists a unique maximizer $\boldsymbol{w} = \mathsf{w}(\mathsf{q}, u)$ of $\mathcal{S}(\cdot, u)$ on $\boldsymbol{C_q}$ such that

$$\mathsf{w}(\mathsf{q}, u) : = \text{argmax}\big\{ S(\boldsymbol{c}, u) \big| \boldsymbol{c} \in \boldsymbol{C}_\mathsf{q} \big\} \Longrightarrow \exists\, \eta \in \mathbb{R}^m\colon \left(\log\left(\frac{\mathsf{w}_i(\mathsf{q}, u)}{w_i(u)}\right)\right)_{1,\dots,I} = \mathsf{Q}^\top \eta. \tag{2.17}$$

Here $\eta \in \mathbb{R}^m$ is the Lagrange multiplier associated with the constraint $\mathsf{Q}\boldsymbol{c} = \boldsymbol{q}$. Since $\boldsymbol{w}(u)$ is a global maximizer of $S(\cdot, u)$ we obviously have $\mathsf{w}(\mathsf{q}_u, u) = \boldsymbol{w}(u)$ if $\mathsf{q}_u = \mathsf{Q}\boldsymbol{w}(u)$.

Since $\mathsf{Q}(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r) = 0$ for all r, it is easy to see that each $\mathsf{w}(\mathsf{q}, u)$ is a steady state of (2.14) and, moreover, it still satisfies the DBC (2.13), see also Mielke et al. (2015), Sect. 2.1. Thus, if necessary, we may assume that for a given initial condition we already have $\mathsf{Q}\boldsymbol{c}(0) = \mathsf{Q}\boldsymbol{w}(u)$.

We now return to the question of quantitative estimates for the exponential decay. Assuming that all reaction rate coefficients $\kappa_r(u)$ are bounded from below by κ_*, the reactive entropy production $P_{\text{react}}(\boldsymbol{c}, u)$ as defined in (2.15) is bounded from below by κ_* times the simplified entropy production potential

$$\widehat{P}_{\text{R}}(\boldsymbol{c}, u) := \sum_{r=1}^{I} \widehat{\kappa}_r(u)\, G\left(\frac{\boldsymbol{c}^{\boldsymbol{\alpha}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\alpha}^r}}, \frac{\boldsymbol{c}^{\boldsymbol{\beta}^r}}{\boldsymbol{w}(u)^{\boldsymbol{\beta}^r}}\right) \text{ with } G(a, b) := (a - b)\log(a/b) \geq 0. \tag{2.18}$$

Our major assumption for both methods is that $\widehat{P}_{\text{R}}(\cdot, u)$ vanishes in each stoichiometric subspace $\boldsymbol{C}_\mathsf{q}$ exactly in one point, namely $\mathsf{w}(\mathsf{q}, u)$. Using the special structure of $\boldsymbol{R}(\cdot, u)$ induced by the mass action kinetic and the detailed balance condition, this is equivalent to the condition that $\boldsymbol{R}(\boldsymbol{c}, u) = 0$ has exactly one solution in each $\boldsymbol{C}_\mathsf{q}$, , namely $\mathsf{w}(\mathsf{q}, u)$. We formulate this *unique-equilibrium condition* (UEC) for later reference:

$$\text{(UEC)} \qquad \{ \boldsymbol{c} \in \boldsymbol{C}_\mathsf{q} \mid \widehat{P}_{\text{R}}(\boldsymbol{c}, u) = 0 \} = \{\mathsf{w}(\mathsf{q}, u)\} = \{ \boldsymbol{c} \in \boldsymbol{C}_\mathsf{q} \mid \boldsymbol{R}(\boldsymbol{c}, u) = 0 \}. \tag{2.19}$$

Since it is well known that any additional solutions of $\boldsymbol{R}(\boldsymbol{c}, u) = 0$ have to lie on the boundary of $[0, \infty[^I$, one easy sufficient condition for the UEC is that no reaction contains an autocatalytic species, which means that α_i^r and β_i^r are both positive. Hence, we have the implication

$$\Big(\forall\, i = 1, \dots, I \text{ and } r = 1, \dots, R : \alpha_i^r \beta_i^r = 0\Big) \Longrightarrow \text{UEC holds.}$$

We refer to Feinberg (1972/73) and Glitzky and Hünlich (1997) for more details. Using the relative Boltzmann entropy $\mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}\big)$ defined in (1.2), the UEC allows us to show that

$$\forall\,(\mathsf{q},\mathsf{U})\in\mathfrak{Q}\;\exists\,K_{\mathrm{R}}(\mathsf{q},\mathsf{U})>0\;\forall\,\boldsymbol{c}\in\boldsymbol{C}_{\mathsf{q}}:\widehat{P}_{\mathrm{R}}(\boldsymbol{c},\mathsf{U})\geq K_{\mathrm{R}}(\mathsf{q},\mathsf{U})\mathfrak{B}\big(\boldsymbol{c}\big|\mathsf{w}(\mathsf{q},\mathsf{U})\big). \tag{2.20}$$

Indeed, this follows from a compactness argument for $\boldsymbol{c}\in\boldsymbol{C}_{\mathsf{Q}\boldsymbol{w}(u)}$, which is a special case of the result in Sect. 4.

2.4 Combining Entropy Production from Diffusion and Reaction

We return to the full ERDS

$$\begin{pmatrix}\dot{\boldsymbol{c}}\\ \dot{u}\end{pmatrix}=\operatorname{div}\Big(M(\boldsymbol{c},u)\nabla\begin{pmatrix}\boldsymbol{c}\\ u\end{pmatrix}\Big)+\begin{pmatrix}\boldsymbol{R}(\boldsymbol{c},u)\\ 0\end{pmatrix}\text{ in }\Omega,\quad \nu\cdot\nabla\begin{pmatrix}\boldsymbol{c}\\ u\end{pmatrix}=0\text{ on }\partial\Omega, \tag{2.21}$$

and consider sufficiently smooth solutions $(\boldsymbol{c}(t,\cdot),u(t,\cdot)):\Omega\to[0,\infty[^{I+1}$. From the no-flux boundary conditions on $\partial\Omega$ and the fact $\boldsymbol{R}(\boldsymbol{c},u)\in\mathbb{S}$, we deduce the conservation relations

$$\int_\Omega u(t,x)\,\mathrm{d}x=\int_\Omega u(0,x)\,\mathrm{d}x=:\mathsf{U}\quad\text{and}\quad\int_\Omega \mathsf{Q}\boldsymbol{c}(t,x)\,\mathrm{d}x=\int_\Omega \mathsf{Q}\boldsymbol{c}(0,x)\,\mathrm{d}x=:\mathsf{q}.$$

Thus, the solutions $(\boldsymbol{c},u)$ will stay inside the sets

$$\mathfrak{S}(\mathsf{q},\mathsf{U}):=\Big\{\,(\boldsymbol{c},u)\in\mathrm{L}^1(\Omega)^{I+1}\;\Big|\;u,c_i\geq 0,\ \int_\Omega u\,\mathrm{d}x=\mathsf{U},\ \int_\Omega \mathsf{Q}\boldsymbol{c}\,\mathrm{d}x=\mathsf{q}\,\Big\}. \tag{2.22}$$

Clearly, $\mathfrak{S}(\mathsf{q},\mathsf{U})$ contains exactly one constant state, namely $(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})$.

For a fixed (q,U) and using the concave entropy density S, we define the nonnegative and convex relative entropy density

$$H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u):=S(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})+\mathrm{D}S(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})\cdot\big(\boldsymbol{c}-\mathsf{w}(\mathsf{q},\mathsf{U}),u-\mathsf{U}\big)-S(\boldsymbol{c},u).$$

Similarly, for entropies $s(\cdot)$ only dependent on the internal energy u, we define

$$h_{\mathsf{U}}(u)=s(\mathsf{U})+s'(\mathsf{U})\cdot(u-\mathsf{U})-s(u)\quad\text{and}\quad\widehat{h}_{\mathsf{U}}(u)=h_{\mathsf{U}}(u)+I-\sum_{i=1}^{I}w_i(u). \tag{2.23}$$

Clearly, $h_{\mathsf{U}}(u)\geq 0$ and $H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u)\geq 0$ with equality if and only if $(\boldsymbol{c},u)=(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})$.

We define a relative entropy functional on $\mathrm{L}^1(\Omega)^{I+1}$ via

$$\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u):=\int_\Omega H_{\mathsf{q},\mathsf{U}}\big(\boldsymbol{c}(x),u(x)\big)\,\mathrm{d}x$$

which is convex and nonnegative. The unique minimizer is the constant function $(\boldsymbol{c},u)\equiv(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})$ which lies in the convex set $\mathfrak{S}(\mathsf{q},\mathsf{U})$. Moreover, by construction we know that

$$\forall\,(\boldsymbol{c},u)\in\mathfrak{S}(\mathsf{q},\mathsf{U}):\quad \mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u)=\mathcal{S}(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U}))-\mathcal{S}(\boldsymbol{c},u).$$

Thus, for solutions $(\boldsymbol{c}(t),u(t))$ in $\mathfrak{S}(\mathsf{q},\mathsf{U})$ the entropy production can be calculated via

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\boldsymbol{c},u)=-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_{\boldsymbol{q}}(\boldsymbol{c},u)=\mathcal{P}(\boldsymbol{c},u)=\mathcal{P}_{\mathrm{diff}}(\boldsymbol{c},u)+\mathcal{P}_{\mathrm{react}}(\boldsymbol{c},u) \tag{2.24}$$

$$\text{where } \mathcal{P}_{\mathrm{react}}(\boldsymbol{c},u):=\int_\Omega P_{\mathrm{react}}(\boldsymbol{c}(x),u(x))\,\mathrm{d}x. \tag{2.25}$$

Under the assumptions of Sects. 2.2 and 2.3, we know that $\mathcal{P}_{\mathrm{diff}}(\boldsymbol{c},u)=0$ if and only if $(\boldsymbol{c},u)=(\overline{\boldsymbol{c}},\overline{u})$, i.e., all functions are constant. In contrast, the condition $\mathcal{P}_{\mathrm{react}}(\boldsymbol{c},u)=0$ implies that $(\boldsymbol{c},u)$ are in local equilibrium only, i.e., $(\boldsymbol{c}(x),u(x))=(\mathsf{w}(\widetilde{q}(x),\widetilde{u}(x)),\widetilde{u}(x))$ for arbitrary functions $(\widetilde{q},\widetilde{u}):\Omega\to\mathbb{R}^m\times[0,\infty[$. However, combining these two conditions and the property $(\boldsymbol{c},u)\in\mathfrak{S}(\mathsf{q},\mathsf{U})$, we see that the entropy production $\mathcal{P}(\boldsymbol{c},u)$ restricted to $\mathfrak{S}(\mathsf{q},\mathsf{U})$ is 0 if and only if $(\boldsymbol{c},u)=(\mathsf{w}(\mathsf{q},\mathsf{U}),\mathsf{U})$.

Thus, it is natural to ask whether there is a EEP estimate of the form

$$\forall\,(\mathsf{q},\mathsf{U})\in\mathfrak{Q}\;\exists\,K(\mathsf{q},\mathsf{U})>0\;\forall\,(\boldsymbol{c},u)\in\mathfrak{S}(\mathsf{q},\mathsf{U}):\quad \mathcal{P}(\boldsymbol{c},u)\geq K(\mathsf{q},\mathsf{U})\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u). \tag{2.26}$$

To establish the above estimate, it is convenient to estimate the two terms in $\mathcal{P}$ by its model contributions, namely

$$\begin{aligned}&\mathcal{P}_{\mathrm{diff}}(\boldsymbol{c},u)\geq\delta_*\widehat{\mathcal{P}}_\gamma(\boldsymbol{c},u)\ \text{ and }\ \mathcal{P}_{\mathrm{react}}(\boldsymbol{c},u)\geq\kappa_*\widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c},u),\\&\text{where } \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c},u):=\int_\Omega\widehat{P}_{\mathrm{R}}(\boldsymbol{c}(x),u(x))\,\mathrm{d}x\end{aligned} \tag{2.27}$$

with $\widehat{\mathcal{P}}_\gamma$ and $\widehat{P}_{\mathrm{R}}$ from (2.8) and (2.18), respectively. In the sequel, we will hence simplify the analysis and the notation by only estimating $\widehat{\mathcal{P}}_\gamma(\boldsymbol{c},u)+\widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c},u)$ from below by $\mathcal{H}_{\mathsf{q},\mathsf{U}}$, and moreover by always assuming that $\mathsf{q}=\mathsf{Q}\boldsymbol{w}(\mathsf{U})$. As was explained in Sect. 2.3, this is always possible by adjusting $\boldsymbol{w}$ suitably. However, we need to be sure that this does not change the necessary structural properties of the vector-valued function $u\mapsto\boldsymbol{w}(u)$, such as monotonicity and concavity, see Sect. 2.1. However, when fixing on U and q we can change $\boldsymbol{w}(\cdot)$ by pre-multiplying it with the diagonal matrix $A=\mathrm{diag}(\mathrm{e}^{a_1},\ldots,\mathrm{e}^{a_I})$, where $\boldsymbol{a}=\mathsf{Q}^\top\eta$ is the vector given in (2.17). Then, obviously the new vector-valued function $\widetilde{\boldsymbol{w}}(u)=A\boldsymbol{w}(u)$ satisfies $\widetilde{w}(\mathsf{U})=\mathsf{w}(\mathsf{q},\mathsf{U})$. However, such a pre-multiplication does not change any of the properties we assumed so far or we will use later. Of course, this will change all the constants, in particular $K_R(\mathsf{U},q)$.

Thus, with a slight abuse of notation we will use the notation

$$\widehat{\mathfrak{S}}(\mathsf{U}):=\left\{\,(\boldsymbol{c},u)\in\mathrm{L}^1(\Omega)^{I+1}\;\middle|\;u,c_i\geq0,\ \int_\Omega u\,\mathrm{d}x=\mathsf{U},\ \int_\Omega\mathsf{Q}\boldsymbol{c}\,\mathrm{d}x=\mathsf{Q}\boldsymbol{w}(\mathsf{U})\,\right\},$$

which leads us to our remaining task, which is to show the following estimate:

$$\forall\, \mathsf{U} > 0\; \exists\, \widehat{K}(\mathsf{U}) > 0\; \forall\, (\boldsymbol{c}, u) \in \widehat{\mathfrak{S}}(\mathsf{U}) : \quad \big(\widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) + \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c}, u)\big) \ge \widehat{K}(\mathsf{U})\widehat{\mathcal{H}}_{\mathsf{U}}(\boldsymbol{c}, u), \tag{2.28}$$

where $\widehat{H}_{\mathsf{U}} = H_{\mathsf{Q}\boldsymbol{w}(\mathsf{U}),\mathsf{U}}$ and $\widehat{\mathcal{H}}_{\mathsf{U}}(\boldsymbol{c}, u) = \int_\Omega \widehat{H}_{\mathsf{U}}(\boldsymbol{c}, u)\,\mathrm{d}x = \mathcal{H}_{\mathsf{Q}\boldsymbol{w}(\mathsf{U}),\mathsf{U}}(\boldsymbol{c}, u)$. Clearly, if (2.28) holds, then we obtain the bound $K(\mathsf{Q}\boldsymbol{w}(\mathsf{U}), \mathsf{U}) \ge \min\{\delta_*, \kappa_*\}\widehat{K}(\mathsf{U})$ in (2.26).

3 EEP Estimates via Convexity

In this section, we describe the general method on how to obtain estimates of the entropy in terms of the entropy production by using a convexity argument in the way as derived in Mielke et al. (2015) for the isothermal case. It is based on the fact that the diffusion part $\widehat{\mathcal{P}}_{\mathrm{diff}}$ can be estimated by the log-Sobolev inequality and variants of it, see Sect. 3.1. If this method works, it provides stronger results that are not dependent on the entropy of the initial condition as in the case of the compactness method from Glitzky and Hünlich (1997) discussed in Sect. 4.

3.1 Generalized Log-Sobolev Inequalities

For $\alpha \ge 0$, we consider the entropy functions

$$F_\alpha(z) = \begin{cases} \frac{1}{\alpha(\alpha-1)}\big(z^\alpha - \alpha z + \alpha - 1\big) & \text{for } \alpha \in \mathbb{R} \setminus \{0, 1\}, \\ \lambda_{\mathfrak{B}}(z) = z \log z - z + 1 & \text{for } \alpha = 1, \\ z - \log z - 1 & \text{for } \alpha = 0. \end{cases} \tag{3.1}$$

We see that all F_α are convex and satisfy $F_\alpha(1) = F'_\alpha(1) = 0 \le F_\alpha(z)$. In particular, we have $F''_\alpha(z) = z^{\alpha-2} > 0$. Moreover, for $\alpha > 0$, we have the monotonicities

$$\partial_\alpha\big(\alpha F_\alpha(z)\big) = \frac{z}{(1-\alpha)^2} F_1\big(z^{\alpha-1}\big) \ge 0 \;\text{ and }\; \partial_\alpha\big((1-\alpha)F_\alpha(z)\big) = -\frac{1}{\alpha^2} F_1\big(z^\alpha\big) \le 0. \tag{3.2}$$

Thus, for all $\alpha \in\,]0, 1[$ all these entropies are equivalent:

$$0 < \alpha_1 < \alpha_2 < 1 \quad \Longrightarrow \quad \frac{1-\alpha_2}{1-\alpha_1} F_{\alpha_2}(z) \;\le\; F_{\alpha_1}(z) \;\le\; \frac{\alpha_2}{\alpha_1} F_{\alpha_2}(z).$$

Moreover, we have $F_1 = \lambda_{\mathfrak{B}}$ and the following identities

$$F_1(uv) = vF_1(u) + uF_1(v) + (u-1)(v-1), \tag{3.3}$$

$$a^\alpha F_\alpha(u/a) = F_\alpha(u) + \frac{a^\alpha - 1}{\alpha} - u\frac{a^{\alpha-1} - 1}{\alpha - 1}. \tag{3.4}$$

Using $\overline{u} = \int_\Omega u(x)\,\mathrm{d}x$ an integration of (3.4) gives

$$\overline{u}^\alpha \int_\Omega F_\alpha\big(u/\overline{u}\big)\,\mathrm{d}x = \int_\Omega F_\alpha(u)\,\mathrm{d}x - F_\alpha(\overline{u}) \quad (\text{recall } |\Omega| = 1). \tag{3.5}$$

Crucial for the forthcoming analysis will be the following estimate, which can be seen as a generalization of Poincaré and the log-Sobolev estimates, namely

$$\forall\, u > 0 \text{ with } u^{\gamma/2} \in \mathrm{H}^1(\Omega): \quad \int_\Omega \frac{|\nabla u|^2}{u^{2-\gamma}}\,\mathrm{d}x \geq \rho(\Omega,\gamma,\alpha)\,\overline{u}^\gamma \int_\Omega F_\alpha(u/\overline{u})\,\mathrm{d}x, \tag{3.6}$$

where $\rho(\Omega,\gamma,\alpha)$ is supposed to be the optimal constant. The monotonicity (3.2) implies

$$0 < \alpha_1 < \alpha_2 \quad \Longrightarrow \quad \rho(\Omega,\gamma,\alpha_1) \leq \frac{\alpha_1}{\alpha_2}\rho(\Omega,\gamma,\alpha_2).$$

If 2_d is the Lebesgue power for the embedding $\mathrm{H}^1(\Omega) \subset \mathrm{L}^{2_d}(\Omega)$ for a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ (namely $2_d = 2d/(d-2)$ for $d \geq 3$), then $\rho(\Omega,\gamma,\alpha)$ is positive if $\gamma\, 2_d \geq 2$ and $\alpha \leq \gamma + 2/d$, see Appendix A. The case $\alpha = \gamma = 2$ is the *Poincaré estimate*, while $\alpha = \gamma = 1$ is the *log-Sobolev estimate*, which will be crucial for the concentrations c_i.

However, for the internal energy, we want to use an entropy function F_α with $\alpha = \gamma \in [0, 1[$. Since the entropy functions are equivalent for $\alpha \in]0, 1[$ and have a linear growth, a positive constant $\rho(\Omega,\gamma,\alpha)$ can only exist if $2_d\gamma \geq 2$. Subsequently, we will often use $\rho(\Omega,\gamma,1) > 0$ which certainly holds for $\gamma 2_d > 2$. Indeed, this is a special and simple subcase of the compactness arguments developed in Sect. 4.

Considering the case of gas dynamics, where the entropy $s(u) = s_0 \log u$ is relevant, it is interesting to note that the case $\alpha = \gamma = 0$ is still possible, but only in dimensions $d \leq 2$, see Appendix B.

3.2 The General Setup of the Convexity Method

We now want to develop ideas to establish estimate (2.28) for fixed $\mathsf{U} > 0$ and $(\boldsymbol{c}, u) \in \widehat{\mathfrak{S}}(\mathsf{U})$. The main point is that the diffusion part $\widehat{\mathcal{P}}_\gamma$ of the entropy production has a lower bound in terms of functionals that only depend on the point values $(\boldsymbol{c}(x), u(x))$ and no derivatives. These functionals then need to be combined with $\widehat{\mathcal{P}}_\mathrm{R}$ to find an upper bound for $\widehat{\mathcal{H}}_\mathsf{U}$. Thus, all the latter functionals only involve point values $(\boldsymbol{c}(x), u(x))$ and no derivatives, so it is natural to seek for pointwise estimates of the integrands. However, in general this cannot work, since we need to involve the constraints arising from $(\boldsymbol{c}, u) \in \widehat{\mathfrak{S}}(\mathsf{U})$, namely $\mathsf{Q}\overline{\boldsymbol{c}} = \mathsf{Q}\boldsymbol{w}(\mathsf{U})$ and $\overline{u} = \mathsf{U}$, and to exploit the ODE decay coefficient $\widehat{K}_\mathrm{R}(\mathsf{U}) := K_\mathrm{R}(\mathsf{Q}\boldsymbol{w}(\mathsf{U}), \mathsf{U})$ from (2.20).

For the densities c_i, we use the classical log-Sobolev estimate with $\rho_\Omega^{\mathrm{lS}} = \rho(\Omega, 1, 1)$:

$$\int_\Omega \frac{|\nabla c_i|^2}{c_i}\,\mathrm{d}x \geq \rho_\Omega^{\mathrm{lS}}\mathfrak{B}\big(c_i\big|\overline{c}_i\big)=\rho_\Omega^{\mathrm{lS}}\int_\Omega \overline{c}_i\lambda_{\mathfrak{B}}(c_i/\overline{c}_i)\,\mathrm{d}x=\rho_\Omega^{\mathrm{lS}}\Big(\int_\Omega \lambda_{\mathfrak{B}}(c_i)\,\mathrm{d}x-\lambda_{\mathfrak{B}}(\overline{c}_i)\Big).$$

For the internal energy, the necessary estimate in the general case is $\int_\Omega \widehat{h}''_{\mathsf{U}}(u)|\nabla u|^2\,\mathrm{d}x \geq C_*\int_\Omega h_{\mathsf{U}}(u)\,\mathrm{d}x$ for all u with $\overline{u} = \mathsf{U}$. Recall the definition of h_{U} and $\widehat{h}_{\mathsf{U}}$ in (2.23). The assumption in Sect. 2.2 already imposed $\widehat{h}''_{\mathsf{U}}(u) \geq \delta/u^{2-\gamma}$ for some $\gamma \in]0, 1[$. Hence, setting $\rho_\Omega^{\gamma,\alpha} := \rho(\Omega, \gamma, \alpha) > 0$ (cf. (3.6)) for a suitable $\alpha \geq \gamma$ we can use the following EEP estimate for the internal energy:

$$\forall\, u \text{ with } \overline{u} = \mathsf{U} : \quad \int_\Omega \frac{|\nabla u|^2}{u^{2-\gamma}}\,\mathrm{d}x \geq \rho_\Omega^{\gamma,\alpha}\mathsf{U}^\gamma\int_\Omega F_\alpha\big(u/\mathsf{U}\big)\,\mathrm{d}x. \tag{3.7}$$

For controlling the deviation of u from $\overline{u}$, it would be sufficient to use $\alpha = \gamma$; however, we will use F_α with $\alpha \geq 1$, because a higher value for α is advantageous for the subsequent convexity condition. We further assume

$$\forall\, \mathsf{U}\,\exists\, C_h(\mathsf{U}) > 0\ \forall\, u > 0 : \quad \mathsf{U}^\gamma F_\alpha\big(u/\mathsf{U}\big) \geq C_h(\mathsf{U}) h_{\mathsf{U}}(u). \tag{3.8}$$

For the case $s(u) = s_0 u^\gamma$, we have $h_{\mathsf{U}}(u) = s_0\gamma(1-\gamma)\mathsf{U}^\gamma F_\gamma\big(u/\mathsf{U}\big)$, i.e., $C_h(\mathsf{U}) = 1/\big(s_0\gamma(1-\gamma)\big)$ can be chosen independently of U.

Combining this with the estimates for the densities, we obtain the desired lower bound for $\widehat{\mathcal{P}}_\gamma$ in terms of derivative-free functionals, namely

$$\begin{aligned} &\widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) \geq p_\Omega \int_\Omega \Big(\mathsf{U}^\gamma F_\alpha(u(x)/\mathsf{U}) + \mathfrak{B}\big(\boldsymbol{c}(x)\big|\mathbf{1}\big) - \mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\mathbf{1}\big)\Big)\,\mathrm{d}x, \\ &\text{where } \overline{u} = \mathsf{U}, \quad p_\Omega = \min\{\rho_\Omega^{\mathrm{lS}}, \rho_\Omega^{\gamma,\alpha}\}, \text{ and } \mathbf{1} = (1, 1, \ldots, 1). \end{aligned} \tag{3.9}$$

Since the desired relative entropy reads $\widehat{H}_{\mathsf{U}}(\boldsymbol{c}, u) = h_{\mathsf{U}}(u) + \mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(u)\big)$ (where $\widehat{H}_{\mathsf{U}}$ is defined after (2.28)), but the ODE decay rate $\widehat{K}_{\mathrm{R}}(\mathsf{U}) = K_{\mathrm{R}}(\mathsf{Q}\boldsymbol{w}(\mathsf{U}), \mathsf{U})$ involves the relative Boltzmann entropy $\mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(\mathsf{U})\big)$ with respect to the fixed value $\boldsymbol{w}(\mathsf{U})$, we need the following useful lemma.

Lemma 3.1 (Comparing relative entropies) *Assume that the functions* $w_i : [0, \infty[\to [0, \infty[$ *are increasing, concave, and satisfy* $w_i(0) > 0$ *and that* (3.8) *holds, then there exists a constant* $C_H(\mathsf{U})$ *such that*

$$\begin{aligned} \forall\, (\boldsymbol{c}, u) \in [0, \infty[^{I+1} : \quad \widehat{H}_{\mathsf{U}}(\boldsymbol{c}, u) &= h_{\mathsf{U}}(u)+\mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(u)\big) \\ &\leq C_H(\mathsf{U})\big(\mathsf{U}^\gamma F_\alpha(u/\mathsf{U})+\mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(\mathsf{U})\big)\big). \end{aligned}$$

Proof Using (3.8) it suffices to show the assertion with $\mathsf{U}^\gamma F_\alpha(u/\mathsf{U})$ replaced by $h_{\mathsf{U}}(u)$. Hence, we define $\widetilde{H}(\boldsymbol{c}, u) = h_{\mathsf{U}}(u)+\mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(\mathsf{U})\big)$ and see that $\widetilde{H}$ and $\widehat{H}_{\mathsf{U}}$ are nonnegative and vanish only at $(\boldsymbol{c}, u) = (\boldsymbol{w}(\mathsf{U}), \mathsf{U})$. Since both functions behave quadratically near this point, the upper estimates follows by compactness for all bounded domains.

For large $(\boldsymbol{c}, u)$ we have $\widetilde{H}(\boldsymbol{c}, u) \geq \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big) + \frac{1}{C}(u + |\boldsymbol{c}|_1) - C$ and $\widehat{H}_{\mathsf{U}}(\boldsymbol{c}, u) \leq \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big) + C(1+u+\mu|\boldsymbol{c}|_1+|\boldsymbol{w}(u)|_1)$, where $\mu = -\min_i \log w_i(0)$. Since $|\boldsymbol{w}(u)|_1 \leq C(1+u)$ the desired result follows also for large $(\boldsymbol{c}, u)$. □

Now we are able to state our major result for the convexity method. The resulting decay rate $\widetilde{K}(\mathsf{U})$ is dominated by $\kappa_* \widehat{K}_{\mathrm{R}}(\mathsf{U})$, which is the pure reaction decay and by $\delta_* p_\Omega$, which is the decay rate for pure diffusion. However, as in the isothermal case, see Mielke et al. (2015), Thm. 3.1, the latter coefficient is multiplied by a factor $\eta = \widehat{K}_{\mathrm{R}}(\mathsf{U})/(\mu(\mathsf{U})+\widehat{K}_{\mathrm{R}}(\mathsf{U})) \le 1$ that stands for the interplay of reaction and diffusion, since $\mu(\mathsf{U})$ measures how much convexity induced by diffusion is needed to convexify the typically non-convex reactive entropy production $\widehat{P}_{\mathrm{R}}$.

Theorem 3.2 (Convexity method) *Consider $S(\boldsymbol{c}, u) = s(u) - \mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(u)\big)$ and let h_{U} be defined as above, satisfying $p_\Omega > 0$ in (3.9), $C_H(\mathsf{U}) > 0$ in Lemma 3.1, and the unique equilibrium condition (2.19), i.e., $\widehat{K}_{\mathrm{R}}(\mathsf{U}) > 0$ in (2.20). We define the function*

$$G_{\mu,\mathsf{U}}(\boldsymbol{c}, u) := \mu\big(\mathsf{U}^\gamma F_\alpha(u/\mathsf{U}) + \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big)\big) + \widehat{P}_{\mathrm{R}}(\boldsymbol{c}, u)$$

and assume the following convexity condition:

$$\forall\, \mathsf{U}\ \exists\, \mu(\mathsf{U}) \ge 0: \quad G_{\mu(\mathsf{U}),\mathsf{U}} : [0,\infty[^{I+1} \to [0,\infty[\ \text{ is convex.} \tag{3.10}$$

Then, we have the following EEP estimate

$$\begin{aligned} &\forall\, (\boldsymbol{c}, u) \in \widehat{\mathfrak{S}}(\mathsf{U}): \quad \delta_* \widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) + \kappa_* \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c}, u) \ge \widetilde{K}(\mathsf{U}) \widehat{\mathcal{H}}_{\mathsf{U}}(\boldsymbol{c}, u) \\ &\text{with } \widetilde{K}(\mathsf{U}) = \frac{1}{C_H(\mathsf{U})} \min\Big\{\kappa_* \widehat{K}_{\mathrm{R}}(\mathsf{U}),\, \delta_* p_\Omega \frac{\widehat{K}_{\mathrm{R}}(\mathsf{U})}{\mu(\mathsf{U}) + \widehat{K}_{\mathrm{R}}(\mathsf{U})}\Big\}. \end{aligned} \tag{3.11}$$

Proof Defining $H_{\mathrm{B}}(\boldsymbol{c}, u) = U^\gamma F_\alpha(u/\mathsf{U}) + \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big) - \mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\mathbf{1}\big)$, choosing $\theta \in]0, 1[$, and setting $M(\theta, \mathsf{U}) = \min\{\kappa_*, (1-\theta)\delta_* p_\Omega/\mu(\mathsf{U})\}$ the assumptions yield the estimate

$$\begin{aligned} \delta_* \widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) + \kappa_* \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c}, u) &\ge \int_\Omega \Big(\delta_* p_\Omega H_{\mathrm{B}}(\boldsymbol{c}, u) + \kappa_* \widehat{P}_{\mathrm{R}}(\boldsymbol{c}, u)\Big)\,\mathrm{d}x \\ \ge\ \theta \delta_* p_\Omega \int_\Omega H_{\mathrm{B}}(\boldsymbol{c}, u)\,\mathrm{d}x &+ M(\theta, \mathsf{U}) \int_\Omega \big(G_{\mu(\mathsf{U}),\mathsf{U}}(\boldsymbol{c}, u) - \mu(\mathsf{U})\mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\mathbf{1}\big)\big)\,\mathrm{d}x \\ \overset{\text{Jensen}}{\underset{(3.10)}{\ge}}\ \theta \delta_* p_\Omega \int_\Omega H_{\mathrm{B}}(\boldsymbol{c}, u)\,\mathrm{d}x &+ M(\theta, \mathsf{U})\big(G_{\mu(\mathsf{U}),\mathsf{U}}(\overline{\boldsymbol{c}}, \overline{u}) - \mu(\mathsf{U})\mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\mathbf{1}\big)\big) \\ =\ \theta \delta_* p_\Omega \int_\Omega H_{\mathrm{B}}(\boldsymbol{c}, u)\,\mathrm{d}x &+ M(\theta, \mathsf{U})\big(\mu(\mathsf{U})\, 0 + \widehat{P}_{\mathrm{R}}(\overline{\boldsymbol{c}}, \mathsf{U})\big) \end{aligned}$$

where we used $\overline{u} = \mathsf{U}$ and $F_\alpha(\overline{u}/\mathsf{U}) = 0$ in the last identity. Employing the ODE estimate (2.20), where we use $\mathsf{Q}\overline{\boldsymbol{c}} = \mathsf{Q}\boldsymbol{w}(\mathsf{U})$ following from $(\boldsymbol{c}, u) \in \widehat{\mathfrak{S}}(\mathsf{U})$, we continue

$$\begin{aligned} \delta_* \widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) + \kappa_* \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c}, u) &\ge \theta \delta_* p_\Omega \int_\Omega H_{\mathrm{B}}(\boldsymbol{c}, u)\,\mathrm{d}x + M(\theta, \mathsf{U}) \widehat{K}_{\mathrm{R}}(\mathsf{U}) \mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\boldsymbol{w}(\mathsf{U})\big) \\ \ge N(\theta, \mathsf{U}) \int_\Omega \big(\mathsf{U}^\gamma F_\alpha(u/\mathsf{U}) &+ \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big) - \mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\mathbf{1}\big) + \mathfrak{B}\big(\overline{\boldsymbol{c}}\big|\boldsymbol{w}(\mathsf{U})\big)\big)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&=_{(1)} N(\theta, \mathsf{U}) \int_\Omega \big(\mathsf{U}^\gamma F_\alpha(u/\mathsf{U}) + \mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(\mathsf{U})\big)\big)\,\mathrm{d}x \\
&\geq_{(2)} \frac{N(\theta, \mathsf{U})}{C_H(\mathsf{U})} \int_\Omega \big(h_{\mathsf{U}}(u) + \mathfrak{B}\big(\boldsymbol{c}\big|\boldsymbol{w}(u)\big)\big)\,\mathrm{d}x.
\end{aligned}
$$

where $N(\theta, \mathsf{U}) = \min\{\theta\delta_* p_\Omega, M(\theta, \mathsf{U})\widehat{K}_\mathrm{R}(\mathsf{U})\}$. For "$=_{(*)}$" we used the general identity

$$
\mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big) - \mathfrak{B}\big(\boldsymbol{b}\big|\mathbf{1}\big) + \mathfrak{B}\big(\boldsymbol{b}\big|\widetilde{\boldsymbol{w}}\big) = \mathfrak{B}\big(\boldsymbol{c}\big|\widetilde{\boldsymbol{w}}\big) + \sum\nolimits_{i=1}^{I} (c_i - b_i) \log \widetilde{w}_i
$$

for $\boldsymbol{b} = \overline{\boldsymbol{c}}$. Since $\widetilde{\boldsymbol{w}} = \boldsymbol{w}(\mathsf{U})$ is spatially constant, the sum on the right-hand side cancels upon integration over Ω. For the estimate "$\geq_{(2)}$" we used Lemma 3.1, and the result is established if we choose the optimal $\theta = \widehat{K}_\mathrm{R}/(\mu + \widehat{K}_\mathrm{R})$. □

In the above theorem, we have given only the simplest form of the convexity method. As is shown in Mielke et al. (2015), there are several ways to generalize the method (i) by replacing the convexity assumption by an estimate on the convexification (cf. (Mielke et al. 2015, Sec. 3.3)) or (ii) by using higher entropies F_α with $\alpha > 1$ instead of $F_1 = \lambda_\mathfrak{B}$ for the densities c_i (cf. (Mielke et al. 2015, Sec. 3.4)).

Note that we already used the higher-order entropy F_α for the internal energy, since F_γ and h_U only have linear growth, which would give weaker strict convexity to be exploited for compensating the missing convexity of $\widehat{P}_\mathrm{R}$. Indeed, we will see in Sect. 3.3.3 that $\alpha = 1 > \gamma = 1/2$ is useful. Thus, it is crucial that we have $\rho(\Omega, \gamma, \alpha) > 0$ for some $\alpha \geq 1$, which is a well-known fact in finite space dimensions. The optimal constant $\rho(\Omega, \gamma, \alpha)$ is positive for $\alpha \in\,]0, \gamma{+}2/d]$ if $\gamma \in\,]0, 1[$ and $d \leq 2$ and for $\alpha \in\,]0, \gamma{+}2/d]$ if $\gamma \in\,]1{-}2/d, 1[$ and $d \geq 3$, see e.g., Mittnenzweig (2014) and Appendix A. In particular, $\alpha = 1$ is always admissible.

In the following, we will apply the method to a simple example and leave a general study for future work.

3.3 A Simple ERDS with Scalar *c*

To give a first impression of the convexity method, we consider one chemical species with density $c > 0$ that can react with the background by absorption or generation. This reaction is steered by the local value of the internal energy $u(t, x)$.

3.3.1 General Description of the Simple ERDS

More precisely, we consider the semilinear system

$$
\dot{c} = \delta_\mathrm{u}\Delta c + \kappa(c, u)\big(f(w(u)) - f(c)\big), \qquad \dot{u} = \delta_\mathrm{u}\Delta u, \tag{3.12}
$$

on a bounded Lipschitz domain with no-flux boundary conditions. Here $w(u)$ is a given function determining the equilibrium concentration $c = w(u)$ for a given internal energy. The function $f : [0, \infty[\, \to [0, \infty[$ is assumed to be smooth, strictly increasing and satisfying $f(0) = 0$. The positive reaction rate coefficient $\kappa = \kappa(c, u)$ gives the strength of the reaction.

We now redo the above estimate on this concrete example and try to make the constants as explicit as possible. For a fixed U and an entropy density $S(c, u) = s(u) - w(u)\lambda_{\mathfrak{B}}(c/w(u))$, we construct the relative entropy via

$$\mathcal{H}_{\mathsf{U}}(c, u) = \int_\Omega H_{\mathsf{U}}(c(x), u(x))\,\mathrm{d}x, \quad \text{where}$$
$$H_{\mathsf{U}}(c, u) = h_{\mathsf{U}}(u) + w(u)\lambda_{\mathrm{B}}\big(\tfrac{c}{w(u)}\big) \text{ with } h_{\mathsf{U}}(u) = s(\mathsf{U}) - s'(\mathsf{U})(u-\mathsf{U}) - s(u). \tag{3.13}$$

The entropy production terms read

$$\mathcal{P}(c, u) = \delta_{\mathrm{c}}\mathcal{P}_{\mathrm{c}}(c, u) + \delta_{\mathrm{u}}\mathcal{P}_{\mathrm{u}}(c, u) + \mathcal{P}_{\mathrm{R}}(c, u) \quad \text{with} \tag{3.14a}$$
$$\mathcal{P}_{\mathrm{c}}(c, u) = \int_\Omega \frac{1}{c}|\nabla c|^2 - \frac{w'(u)}{w(u)}\nabla c \cdot \nabla u\,\mathrm{d}x, \tag{3.14b}$$
$$\mathcal{P}_{\mathrm{u}}(c, u) = \int_\Omega -\frac{w'(u)}{w(u)}\nabla c \cdot \nabla u + \Big(c\frac{w'^2 - ww''}{w^2} + \widehat{h}''\Big)|\nabla u|^2\,\mathrm{d}x, \tag{3.14c}$$
$$\mathcal{P}_{\mathrm{R}}(c, u) = \int_\Omega P_{\mathrm{R}}(c(x), u(x))\,\mathrm{d}x$$
$$\text{with } P_{\mathrm{R}}(c, u) = \kappa(c, u)\big(f(c) - f(w(u))\big)\log\big(\tfrac{c}{w(u)}\big). \tag{3.14d}$$

Obviously, we have $P_{\mathrm{R}}(c, u) \geq 0$, but $\mathcal{P}_{\mathrm{c}}$ and $\mathcal{P}_{\mathrm{u}}$ are only nonnegative individually if $w'(u) \equiv 0$, which is a manifestation of cross-diffusion.

Proceeding as in Proposition 2.1, we assume that there exists $\rho \in]0, 1[$ such that

$$\rho < \frac{4\delta_{\mathrm{c}}\delta_{\mathrm{u}}}{(\delta_{\mathrm{c}}+\delta_{\mathrm{u}})^2} \leq 1 \quad \text{and} \quad \forall u > 0: \quad -w(u)w''(u) \geq \frac{1-\rho}{\rho}\,w'(u)^2. \tag{3.15}$$

Then, the sum is nonnegative and has the lower bound

$$\delta_{\mathrm{c}}\mathcal{P}_{\mathrm{c}}(c, u) + \delta_{\mathrm{u}}\mathcal{P}_{\mathrm{u}}(c, u) \geq \int_\Omega \Big(\theta\delta_{\mathrm{c}}^{\mathrm{red}}\frac{|\nabla c|^2}{c} + \delta_{\mathrm{u}}\widehat{h}''_{\mathsf{U}}(u)|\nabla u|^2\Big)\,\mathrm{d}x,$$

where the reduced diffusion coefficient reads $\delta_{\mathrm{c}}^{\mathrm{red}} = \delta_{\mathrm{c}}\big(1 - \rho(\delta_{\mathrm{c}}+\delta_{\mathrm{u}})^2/(4\delta_{\mathrm{c}}\delta_{\mathrm{u}})\big) > 0$, and $\widehat{h}_{\mathsf{U}}(u) = h_{\mathsf{U}}(u) + w(u)$ is as above, e.g., convex.

3.3.2 A First Simplistic EEP Estimate

A very simplistic first result is the following, which assumes that the reactive entropy production $P_{\mathrm{R}}(c, u)$ dominates the relative entropy $w(u)\lambda_{\mathfrak{B}}(c/w(u))$ for all (c, u), which essentially means that the ODE $\dot{c} = R(c, \mathsf{U})$ has a uniform exponential decay toward the equilibrium $w(\mathsf{U})$, independently of U.

Proposition 3.3 (Simple decay result) *If for all* $\mathsf{U} > 0$ *there exist positive constants* $k_\mathrm{D}(\mathsf{U})$ *and* k_R *such that*

$$\forall\, u \in \mathrm{L}^1_{\geq 0}(\Omega),\ \overline{u} = \mathsf{U} : \quad \int_\Omega \widehat{h}''(u)|\nabla u|^2 \,\mathrm{d}x \geq k_\mathrm{D}(\mathsf{U}) \int_\Omega h_\mathsf{U}(u)\,\mathrm{d}x,$$
$$\forall\,(c,u) \in \left]0,\infty\right[^2 : \quad P_\mathrm{R}(c,u) \geq k_\mathrm{R}\, w(u) \lambda_{\mathfrak{B}}(c/w(u)),$$

then we have the EEP estimate

$$\forall\,(c,u) \in \widehat{\mathfrak{S}}(\mathsf{U}) : \ \mathcal{P}(c,u) \geq \min\{\delta_u k_\mathrm{D}(\mathsf{U}), k_\mathrm{R}\}\, \mathcal{H}_\mathsf{U}(c,u).$$

Obviously, this result follows simply by estimating $\int_\Omega w(u)\lambda_{\mathfrak{B}}(c/w(u))\,\mathrm{d}x$ by $\mathcal{P}_\mathrm{R}$ and $\int_\Omega h_\mathsf{U}(u)\,\mathrm{d}x$ by $\mathcal{P}_\mathrm{u}$ separately, which only works in very special cases.

The result can be applied to the ERDS (3.12) for the choices

$$w(u) = u^\gamma, \quad s(u) = s_0 u^\gamma \text{ with } s_0 > 1 > \gamma > 0,$$
$$f(c) = c^\beta \text{ with } \beta \geq 1, \quad \text{and } \kappa(c,u) = \kappa_* u^{\gamma(1-\beta)}. \tag{3.16}$$

Then, setting $c = z w(u)$, we have

$$\frac{P_\mathrm{R}(c,u)}{w(u)\lambda_{\mathfrak{B}}(c/w(u)} = \kappa_* u^{\gamma(1-\beta)} \frac{(c^\beta - w(u)^\beta)\log(c/w(u))}{w(u)\lambda_{\mathfrak{B}}(c/w(u)} = \kappa_* \frac{(z^\beta - 1)\log z}{\lambda_{\mathfrak{B}}(z)} \geq \kappa_*\,.$$

Furthermore, we have $h_\mathsf{U}(u) = s_0\gamma(1-\gamma)\mathsf{U}^\gamma F_\gamma(u/\mathsf{U})$ and $\widehat{h}_\mathsf{U}(u) = (1-1/s_0)h_\mathsf{U}(u)$, which makes $k_\mathrm{D}(\mathsf{U})$ independent of U.

We summarize the result as follows:

Corollary 3.4 (Simplistic EEP) *We consider the ERDS* (3.12) *with the choices* (3.16) *and set* $S(c,u) = s_0 u^\gamma - u^\gamma \lambda_{\mathfrak{B}}(c/u^\gamma)$. *Assume further*

$$1 - 2/d \leq \gamma \leq \frac{4\delta_c\delta_u}{(\delta_c + \delta_u)^2}, \tag{3.17}$$

then we have the following EEP estimate

$$\forall\,(c,u) \in \widehat{\mathfrak{S}}(\mathsf{U}) : \quad \mathcal{P}(c,u) \geq K\mathcal{H}_\mathsf{U}(c,u) \text{ with } K = \min\{\kappa_*, \delta_\mathrm{u}\tfrac{s_0-1}{s_0}\rho(\Omega,\gamma,\gamma)\} > 0.$$

Of course, in the given case with $\kappa(c,u) = \kappa_* u^{\gamma(1-\beta)}$ we have uniform decay of c in the ODE $\dot{c} = \kappa(c,u)(u^{\gamma\beta} - c^\beta)$. Hence, it is very easy to show that every solution of (c,u) of the ERDS (3.12) converges to the equilibrium $(w(\mathsf{U}),\mathsf{U})$. Nevertheless, the above result is interesting, since it is not so clear that the functional $\mathcal{H}_\mathsf{U}$ satisfies the differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_\mathsf{U}(c(t),u(t)) \leq -K\mathcal{H}_\mathsf{U}(c(t),u(t))$$

uniformly of all solutions in $\widehat{\mathfrak{S}}(\mathsf{U})$.

3.3.3 A Nontrivial Case of the Simple ERDS

The strength of the convexity method lies in the fact that it also can be used in some cases where the reactive entropy production $P_{\rm R}(\boldsymbol{c}, u)$ does not provide an upper bound for the relative entropy $\mathfrak{B}\big(\boldsymbol{c}\,\big|\,\boldsymbol{w}(u)\big)$. We only give one nontrivial example, namely the special case (3.16), but now specify γ and β as well, namely

$$w(u) = (u{+}b)^{1/2}, \quad f(c) = c^2, \quad \kappa(c, u) = \kappa_* = \text{const.}, \quad s(u) = 2u^{1/2}, \quad (3.18)$$

for some $b > 0$. Thus, the relevant relative entropy reads

$$\begin{aligned} H_{\mathsf{U}}(c, u) &= h_{\mathsf{U}}(u) + (u{+}b)^{1/2}\lambda_{\mathfrak{B}}\big(c/(u{+}b)^{1/2}\big) \text{ with} \\ h_{\mathsf{U}}(u) &= \mathsf{U}^{-1/2}\big(u^{1/2} - \mathsf{U}^{1/2}\big)^2 = \tfrac{\mathsf{U}^{1/2}}{2} F_{1/2}(u/\mathsf{U}) \end{aligned} \quad (3.19)$$

However, we now use a higher-order entropy for u, namely $\alpha = 1 > \gamma = 1/2$, see also (3.7). Again relying on an enforced version of the restriction (3.17), we now assume

$$d \le 3 \quad \text{and} \quad \delta_{\rm c}^{\rm red} := \delta_{\rm c}\left(1 - \frac{(\delta_{\rm c}{+}\delta_{\rm u})^2}{8\delta_{\rm c}\delta_{\rm u}}\right) > 0, \quad (3.20)$$

where the last condition restricts $\delta_{\rm c}/\delta_u$ to $]3{-}2\sqrt{2}, 3{+}2\sqrt{2}[$. Using $\widehat{h}''(u) = w''(u) - s''(u) \ge \frac{1}{4}u^{-3/2}$, Proposition 2.3 provides an estimate for the diffusive entropy production, viz.

$$\delta_{\rm c}\mathcal{P}_{\rm c}(c, u) + \delta_{\rm u}\mathcal{P}_{\rm u}(c, u) \ge \int_\Omega \Big(p_\Omega(\overline{u})\big(\lambda_{\mathfrak{B}}(u) {-} \lambda_{\mathfrak{B}}(\overline{u})\big) + q_\Omega\big(\lambda_{\mathfrak{B}}(c) {-} \lambda_{\mathfrak{B}}(\overline{c})\big)\Big)\,\mathrm{d}x,$$

where $p_\Omega(\mathsf{U}) = \delta_{\rm u}\rho(\Omega, 1/2, 1)/(4\mathsf{U}^{1/2})$ and $q_\Omega = \delta_{\rm c}^{\rm red}\rho(\Omega, 1, 1)$. Theorem A.1 shows that both constants are strictly positive for $d \le 3$. On the right-hand side, we have now chosen a higher entropy for u, which leads to an extra factor $\mathsf{U}^{-1/2}$ in p_Ω, see (3.5). Note also that we have dropped the term $\theta\delta_{\rm c}\overline{c}F_1(c/\overline{c})$, since we will not need it.

For the total entropy production, we now have the lower bound

$$\mathcal{P}(c, u) \ge \int_\Omega \Big(p_\Omega(\overline{u})\big(\lambda_{\mathfrak{B}}(u) {-} \lambda_{\mathfrak{B}}(\overline{u})\big) + q_\Omega\big(\lambda_{\mathfrak{B}}(c) {-} \lambda_{\mathfrak{B}}(\overline{c})\big) + \frac{\kappa_*}{2}G(c^2, u{+}b)\Big)\,\mathrm{d}x,$$

where $G(a, b) = (a{-}b)\log(a/b) \ge 0$. Now we can exploit the convexity result from (Mielke et al. 2015, Lem. 4.3) which states that the function $(c, u) \mapsto L_{a,b}(c, u) = \lambda_{\mathfrak{B}}(u) + aG(c^2, u{+}b)$ is convex for $b = 0$ if $0 \le a \le a_{\rm crit} \approx 0.8565$. Calculating the second derivative gives

$$\mathrm{D}^2 L_{a,b}(c, u) = \mathrm{D}^2 L_{a,0}(c, u{+}b) + \left(\frac{1}{u} - \frac{1}{u{+}b}\right) e_2{\otimes}e_2 \ge \mathrm{D}^2 L_{a,0}(c, u{+}b) \ge 0,$$

which shows that $L_{a,b}$ is convex for all $b \geq 0$ and $a \in [0, a_{\mathrm{crit}}]$. Hence, we choose $\theta \in]0, 1[$ and set

$$\mathcal{F}_{\theta,\mathsf{U}}(c,u) := \int_\Omega \Big(\theta p_\Omega(\mathsf{U})\big(\lambda_{\mathfrak{B}}(u)-\lambda_{\mathfrak{B}}(\mathsf{U})\big) + q_\Omega\big(\lambda_{\mathfrak{B}}(c)-\lambda_{\mathfrak{B}}(\overline{c})\big)\Big)\,\mathrm{d}x.$$

Using $G(\overline{c}^2, r) \geq r\lambda_{\mathfrak{B}}(\overline{c}/r^{1/2})/6$ we obtain, for all $(c,u) \in \widehat{\mathfrak{S}}(\mathsf{U})$, the estimate

$$\begin{aligned}
\mathcal{P}(c,u) \geq\ & \mathcal{F}_{\theta,\mathsf{U}}(c,u) + \int_\Omega \Big((1-\theta)p_\Omega(\mathsf{U})\big(\lambda_{\mathfrak{B}}(u)-\lambda_{\mathfrak{B}}(\mathsf{U})\big) + \frac{\kappa_*}{2}G(c^2, u+b)\Big)\,\mathrm{d}x \\
\geq\ & \mathcal{F}_{\theta,\mathsf{U}}(c,u) + M_{\theta,\mathsf{U}}\int_\Omega \Big(\lambda_{\mathfrak{B}}(u) - \lambda_{\mathfrak{B}}(\mathsf{U}) + a_{\mathrm{crit}}G(c^2, u+b)\Big)\,\mathrm{d}x \\
\overset{\text{Jensen}}{\geq}\ & \mathcal{F}_{\theta,\mathsf{U}}(c,u) + M_{\theta,\mathsf{U}}\Big(\lambda_{\mathfrak{B}}(\mathsf{U}) - \lambda_{\mathfrak{B}}(\mathsf{U}) + G(\overline{c}^2, \mathsf{U}+b)\Big) \\
\geq\ & \mathcal{F}_{\theta,\mathsf{U}}(c,u) + \tfrac{1}{6}M_{\theta,\mathsf{U}}(\mathsf{U}+b)\lambda_{\mathfrak{B}}\big(\overline{c}/(\mathsf{U}+b)^{1/2}\big) \\
& \text{with } M_{\theta,\mathsf{U}} = \min\Big\{\tfrac{\kappa_*}{2a_{\mathrm{crit}}}, (1-\theta)p_\Omega(\mathsf{U})\Big\}.
\end{aligned}$$

Using the relation $\int_\Omega \big(\lambda_{\mathfrak{B}}(c) - \lambda_{\mathfrak{B}}(\overline{c}) + a\lambda_{\mathfrak{B}}(\overline{c}/a)\big)\,\mathrm{d}x = \int_\Omega a\lambda_{\mathfrak{B}}(c/a)\,\mathrm{d}x$ for $a = (\mathsf{U}+b)^{1/2}$ and setting $N_{\theta,\mathsf{U}} = \min\big\{q_\Omega, (\mathsf{U}+b)^{1/2}M_{\theta,\mathsf{U}}/6\big\}$, we obtain the lower estimate

$$\begin{aligned}
\mathcal{P}(c,u) &\geq \int_\Omega \Big(\theta p_\Omega(\mathsf{U})\mathsf{U}\,\lambda_{\mathfrak{B}}\big(u/\mathsf{U}\big) + N_{\theta,\mathsf{U}}(\mathsf{U}+b)^{1/2}\lambda_{\mathfrak{B}}\big(c/(\mathsf{U}^{1/2}+b)\big)\Big)\,\mathrm{d}x \\
&\geq L_{\theta,\mathsf{U}}\int_\Omega \widetilde{H}_{\mathsf{U}}(c,u)\,\mathrm{d}x \\
&\quad \text{where } \widetilde{H}_{\mathsf{U}}(c,u) = h_{\mathsf{U}}(u) + (\mathsf{U}+b)^{1/2}\,\lambda_{\mathfrak{B}}\big(c/(\mathsf{U}+b)^{1/2}\big)
\end{aligned}$$

and $L_{\theta,\mathsf{U}} = \min\big\{\theta p_\Omega(\mathsf{U})\mathsf{U}^{1/2}, N_{\theta,\mathsf{U}}\big\}$. We finally use a quantitative version of Lemma 3.1 for our specific application.

Lemma 3.5 *The functions* H_{U} (cf. (3.19)) *and* $\widetilde{H}_{\mathsf{U}}$ *satisfy the estimate*

$$\forall\, c, u, \mathsf{U}, b > 0 : \quad \widetilde{H}_{\mathsf{U}}(c,u) \geq C_H(b,\mathsf{U})H_{\mathsf{U}}(c,u)$$
$$\textit{with } C_H(b,\mathsf{U}) = \frac{8}{9 + 2\log(1+\mathsf{U}/b))}.$$

Proof We let $\beta = b/\mathsf{U}$ and define the three auxiliary functions

$$h_\beta(v) = \Big(\big(v^2 - \tfrac{\beta}{1+\beta}\big)^{1/2} - \big(\tfrac{1}{1+\beta}\big)^{1/2}\Big)^2,$$
$$B_\beta(a,v) = h_\beta(v) + v\lambda_{\mathfrak{B}}(a/v),\ \ \widetilde{B}_\beta(a,v) = h_\beta(v) + \lambda_{\mathfrak{B}}(a).$$

Inserting $c = (\mathsf{U}{+}b)^{1/2}a$ and $u = (\mathsf{U}{+}b)v^2 - b$, we obtain

$$\frac{1}{C(\mathsf{U},b)} := \sup_{c,u\geq 0} \frac{H_{\mathsf{U}}(c,u)}{\widetilde{H}_{\mathsf{U}}(c,u)} = \sup\Big\{ \frac{B_\beta(a,v)}{\widetilde{B}_\beta(a,v)} \,\Big|\, a\geq 0,\ v \geq \big(\tfrac{\beta}{1+\beta}\big)^{1/2} \Big\} =: g(\beta),\ \text{ where } \beta = b/\mathsf{U}.$$

We decompose the quotient $B_\beta/\widetilde{B}_\beta$ into three parts that can be estimated separately.

$$\begin{aligned}\frac{B_\beta(a,v)}{\widetilde{B}_\beta(a,v)} &= 1 + \frac{v{-}1{-}\log v}{\widetilde{B}_\beta(a,v)} + \frac{(1{-}a)\log v}{\widetilde{B}_\beta(a,v)} =: 1 + T_1(a,v) + T_2(v)\\ &\leq 1 + \frac{v{-}1{-}\log v}{(v-1)^2} + \frac{|1{-}a|\,|v{-}1|}{(v{-}1)^2 + \frac{1}{2}(a{-}1)^2} \leq 1 + \frac{1}{2} + 2 = 7/2,\end{aligned}$$

Using $h_\beta(v) \geq (v{-}1)^2$ we bound T_1 for $a > 0$ and $v \geq v_\beta := \big(\beta/(1{+}\beta)\big)^{1/2} \in\,]0,1[$ via

$$T_1(a,v) \leq \frac{v{-}1{-}\log v}{(v{-}1)^2} \leq \frac{1}{2} + \log\big(1/v_\beta\big).$$

For T_2 we observe that it is negative for $(1{-}a)(v{-}1) < 0$. For $(a,v) \in [0,1]\times[1,\infty[$ we easily find $T_2(a,v) \leq 2^{-1/2}$, where the supremum is attained for $(a,v) \to (1,1)$. For $(a,v) \in [1,\infty[\,\times\,]0,1]\times$ one can show numerically that

$$T_2(a,v) \leq \log(1/v) + 2^{-1/2} \leq \log(1/v_\beta) + 2^{-1/2}.$$

Analytically one chooses $a_1 > 1$, then for $a \geq a_1$ we have

$$T_2(a,v) \leq \frac{(a{-}1)\log(1/v)}{\lambda_{\mathfrak{B}}(a)} \leq \frac{a_1{-}1}{\lambda_{\mathfrak{B}}(a_1)}\log(1/v_\beta). \tag{3.21}$$

For $a \in [1,a_1]$ we have $\lambda_{\mathfrak{B}}(a) \geq \rho_1(a{-}1)^2$ with $\rho_1 := \lambda_{\mathfrak{B}}(a_1)/(a_1{-}1)^2 < 1/2$. For $v \in [v_\beta,1]$ we further have $\log(1/v) \leq \big(1{+}\log(1/v_\beta)\big)(1{-}v)$ and obtain

$$T_2(a,v) \leq \frac{(a{-}1)\big(1{+}\log(1/v_\beta)\big)(1{-}v)}{(v{-}1)^2 + \rho_1(a{-}1)^2} \leq \frac{1{+}\log(1/v_\beta)}{2\sqrt{\rho_1}}$$

Optimizing the position of a_1 one gets an estimate that is qualitatively the same as (3.21).

Together we found $g(\beta) \leq 2\log(1/v_\beta) + 9/4$, which gives the desired result using $\log(1/v_\beta) = \frac{1}{2}\log(1{+}\mathsf{U}/b)$. □

We can now summarize the EEP estimate for the simple ERDS (3.12) under the specification (3.18). Inspecting all the above definitions of constants depending on θ, we see that the optimal value is $\theta = 1/7$. With this choice we obtain the following result.

Theorem 3.6 (Constructive EEP estimate for (3.12)) *For $\mathcal{H}_{\mathsf{U}}$ and $\mathcal{P}$ as defined above via the choices* (3.18), (3.19)*, and* (3.20)*, we have the following estimate*

$$\forall\,(c,u)\in\widehat{\mathfrak{S}}(\mathsf{U}):\quad \mathcal{P}(c,u)\geq K_b(\mathsf{U})\mathcal{H}_{\mathsf{U}}(c,u),\quad \textit{where}$$
$$K_b(\mathsf{U})=\frac{8}{9+\log(1+\mathsf{U}/b)}\ \min\left\{\frac{\delta_{\mathrm{u}}}{28}\rho(\Omega,1/2,1)\,,\ \delta_{\mathrm{c}}^{\mathrm{red}}\rho(\Omega,1,1)\,,\ \frac{\kappa_*}{11}(\mathsf{U}+b)^{1/2}\right\}. \tag{3.22}$$

This result is almost quasi-optimal in the sense, that it behaves linearly in the three limiting factors, namely the diffusion in c, the diffusion in u, and the convergence in the ODE $\dot{c}=\kappa_*(\mathsf{U}+b-c^2)$. It remains unclear, whether the slowly degenerating prefactor $\big(9+2\log(1+\mathsf{U}/b)\big)^{-1}$ is really needed or simply an artifact of our method.

However, we see that the condition $b>0$ is really needed to obtain a nontrivial decay rate $K_b(\mathsf{U})$. The origin of the dependence on b is the estimate in Lemma 3.5, and it is open whether the result still holds in the case $b=0$. The following remark indicates a possible way to treat cases with $w_i(0)=0$.

Remark 3.7 (Case $w_i(0)=0$) Throughout our methods, we experience difficulties in handling the case $w_i(0)=0$. The point is the $H_{\mathsf{U}}(\boldsymbol{c},u)=h_{\mathsf{U}}(u)+\mathfrak{B}\big(\boldsymbol{c}\,\big|\,\boldsymbol{w}(u)\big)$ does not have a good upper bound if $w_i(u)$ is very small.

A possible way to handle such cases might be to use an entropy s that is unbounded at $u=0$ like $s(u)=s_0\log u$ leading to $h_{\mathsf{U}}(u)=F_0(u/\mathsf{U})$. We conjecture that for all $\mathsf{U}>0$ there exists $C_H(\mathsf{U})>0$ such that the estimate

$$F_0(u/\mathsf{U})+\mathsf{U}\lambda_{\mathfrak{B}}(u/\mathsf{U})+\mathsf{U}^{1/2}\lambda_{\mathfrak{B}}\big(c/\mathsf{U}^{1/2}\big)\geq C_H(\mathsf{U})\Big(F_0(u/\mathsf{U})+\mathsf{U}\lambda_{\mathfrak{B}}(u/\mathsf{U})+u^{1/2}\lambda_{\mathfrak{B}}\big(c/u\big)\Big)$$

holds true for all $c,u>0$. However, in that case we are forced to use the generalized log-Sobolev estimate (3.6) for the case $\gamma=0$, which is only possible for space dimensions $d\leq 2$, see Appendix B.

Remark 3.8 It was shown in Mielke (2017), Sec. 3.3 that for $\beta\in[1,22]$ the function $(c,u)\mapsto\mu(\beta)\lambda_{\mathfrak{B}}(u)+(c^\beta-u)\log(c^\beta/u)$ is convex, if $\mu(\beta)$ is chosen sufficiently large. Hence, the above convexity method can be adapted to cover the simple ERDS

$$\dot{c}=\delta_{\mathrm{c}}\Delta c+\kappa\big(w(u)^\beta-c^\beta\big),\qquad \dot{u}=\delta_{\mathrm{u}}\Delta u,$$

if we choose $w(u)=r(b+u)^{1/\beta}$ with $r,b>0$ and a suitable $s(\cdot)$, e.g., $s(u)=s_0u^\gamma+w(u)$ with $\gamma>1-2/d$.

3.4 Additivity of the Convexity Method

The crucial condition in the convexity method is the convexity of the function

$$(\boldsymbol{c},u)\ \mapsto\ G_{\mu,\mathsf{U}}(\boldsymbol{c},u):=\mu\left(U^\gamma F_\alpha(u/\mathsf{U})+\mathfrak{B}\big(\boldsymbol{c}\,\big|\,\mathbf{1}\big)\right)+\widehat{P}_{\mathrm{R}}(\boldsymbol{c},u). \tag{3.23}$$

Note that $\widehat{P}_\mathrm{R}(\boldsymbol{c}, u) \geq 0$ but the set $\{ (\boldsymbol{c}, u) \in [0, \infty[^{I+1} \mid \widehat{P}_\mathrm{R}(\boldsymbol{c}, u) = 0 \}$ is not convex in general. Hence, $\widehat{\widehat{P}}_\mathrm{R}$ will not be convex, but the terms multiplied by μ may restore the convexity if $\mu > 0$ is sufficiently large.

From the additive structure of $\widehat{P}_\mathrm{R}$ given in (1.5b), it is sufficient to consider the R reactions individually. More precisely, if for $r = 1, \ldots, R$ we find $\mu_r \geq 0$ such that

$$(\boldsymbol{c}, u) \;\mapsto\; \mu_r \left(U^\gamma F_\alpha(u/\mathsf{U}) + \mathfrak{B}\big(\boldsymbol{c}\big|\mathbf{1}\big)\right) + \widehat{\kappa}_r(u) G\Big(\tfrac{\boldsymbol{c}^{\alpha^r}}{\boldsymbol{w}(u)^{\alpha^r}}, \tfrac{\boldsymbol{c}^{\beta^r}}{\boldsymbol{w}(u)^{\beta^r}}\Big)$$

is convex, then $G_{\mu,\mathsf{U}}$ in (3.23) is convex for $\mu \geq \sum_{r=1}^R \mu_r$.

In Mielke (2017), Sec. 3.3 it is shown that there exist $\mu_1, \mu_2 \geq 0$ such that the function

$$(v_1, v_2) \;\mapsto\; \mu_1 \lambda_\mathfrak{B}(v_1) + \mu_2 \lambda_\mathfrak{B}(v_2) + G(v_1^{\beta_1}, v_2^{\beta_2}) \;\text{ is convex,}$$

if either $\beta_1 = \beta_2 \in [1, 2]$, or $\beta_1 = 1$ and $\beta_2 \in [1, m^*[$, or $\beta_1 \in [1, m^*[$ and $\beta_2 = 1$, where $m^* \approx 22.062$. Thus, one can treat exchange reactions $\beta_1 X_1 \rightleftharpoons \beta_2 X_2$ but not the important reactions $\alpha_1 X_1 + \alpha_2 X_2 \rightleftharpoons \beta_3 X_3$, for which the much more complicated "generalized convexity method" described in Mielke et al. (2015), Thm. 3.1 needs still to be developed.

Based on the above exchange reactions, we are able to consider the following ERDS with three species and four reactions, namely

$$\emptyset \rightleftharpoons \beta_1 X_1, \quad \emptyset \rightleftharpoons \beta_2 X_2, \quad \alpha_1 X_1 \rightleftharpoons X_3, \quad \alpha_3 X_3 \rightleftharpoons X_2.$$

We choose the equilibrium densities in the form $w_i(u) = a_i (u{+}b)^{1/\beta_i}$ with $a_i, b > 0$ and $\beta_i > 1$ and assume that condition (2.12) with $\rho_i = 1/\beta_i$ holds. The ERDS reads

$$\begin{aligned}
\dot{c}_1 &= \delta_1 \Delta c_1 + \kappa_1 \beta_1 \big(w_1(u)^{\beta_1} {-} c_1^{\beta_1}\big) + \kappa_3 \alpha_1 \big(\mu_3 c_3 {-} c_1^{\alpha_1}\big),\\
\dot{c}_2 &= \delta_2 \Delta c_2 + \kappa_2 \beta_2 \big(w_2(u)^{\beta_2} {-} c_2^{\beta_2}\big) + \kappa_4 \big(c_3^{\alpha_3} {-} \mu_2 c_2\big),\\
\dot{c}_3 &= \delta_3 \Delta c_3 + \kappa_3 \big(c_1^{\alpha_1} {-} \mu_3 c_3\big) + \kappa_4 \alpha_3 \big(\mu_2 c_2 {-} c_3^{\alpha_3}\big),\\
\dot{u} &= \delta_\mathrm{u} \Delta u.
\end{aligned}$$

The detailed balance condition is satisfied for all $u > 0$, if we additionally have

$$\mu_3 a_3 = a_1^{\alpha_1}, \quad \mu_2 a_2 = a_3^{\alpha_3}, \quad \beta_1 = \alpha_1 \beta_3, \quad \beta_3 = \beta_2 \alpha_3.$$

To apply the convexity result of Mielke (2017), Sec. 3.3 the exponents need to satisfy $\beta_1, \beta_2, \alpha_1, \alpha_3 \in [1, m^*[$. Thus, we certainly have a variety of nontrivial choices, e.g.,

$$(\beta_1, \beta_2, \beta_3, \alpha_1, \alpha_3) \in \big\{(2, 2, 2, 1, 1), (16, 2, 8, 2, 2), (9, 3, 9, 1, 3)\big\}.$$

Finally, the reaction part $\widehat{P}_\mathrm{R}$ of the dissipation reads

$$\begin{aligned}
\widehat{P}_\mathrm{R}(\boldsymbol{c}, u) = \kappa_1 G\big(c_1^{\beta_1}, a_1^{\beta_1}(u{+}b)\big) + \kappa_2 G\big(c_2^{\beta_2}, a_2^{\beta_2}(u{+}b)\big) + \kappa_3 G(c_1^{\alpha_1}, \mu_3 c_3)\\
+ \kappa_4 G(\mu_2 c_2, c_3^{\alpha_3}),
\end{aligned}$$

such that the convexity method is applicable as it works for all four terms individually.

4 The Glitzky–Gröger–Hünlich Approach for ERDS

In this section, we discuss the quite general result for uniform exponential decay developed in Gröger (1992), Thm. 2 and generalized to semiconductor models in Glitzky et al. (1994, 1996) and Glitzky and Hünlich (1997). In the latter works, the theory is restricted to space dimensions 1 or 2, because charged particles are considered which are coupled via the Poisson equation. In Mielke (2017), it is shown that the result transfers to general space dimensions, if electronic charges are not present or can be ignored. Here we generalize the method to the case of non-isothermal systems, where the detailed balance equilibria $\boldsymbol{w}(u)$ depend on the internal energy u. This is nontrivial and leads to dimension restrictions for entropies of the form $s(u) = cu^\gamma$, namely we will need $2_d\gamma > 2$, where $2_d > 2$ is the Lebesgue exponent such that $\mathrm{H}^1(\Omega)$ embeds continuously into $\mathrm{L}^{2_d}(\Omega)$.

4.1 General Setup and Result

We recall the definition of the equilibria $\mathsf{w}(\mathsf{q}, \mathsf{U})$ which are assumed to be the unique equilibria of the ODEs $\dot{\boldsymbol{c}} = \boldsymbol{R}(\boldsymbol{c}, \mathsf{U})$ satisfying condition $\boldsymbol{c} \in \boldsymbol{C}_\mathsf{q}$, i.e., $\mathsf{Q}\boldsymbol{c} = \mathsf{q}$. As before we define the relative entropy

$$\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) = \int_\Omega H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}(x), u(x))\,\mathrm{d}x \text{ with } H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) = h_\mathsf{U}(u) + \mathfrak{B}\big(\boldsymbol{c}\,\big|\,\mathsf{w}(\mathsf{q}, \mathsf{U})\big).$$

As before we let $\widehat{h}_\mathsf{U}(u) = h_\mathsf{U}(u) - I + \sum_{i=1}^I w_i(u)$ and assume that

$$\begin{aligned} &0 = h_\mathsf{U}(\mathsf{U}) < h_\mathsf{U}(u) \text{ for } u \neq \mathsf{U}, \quad \widehat{h}_\mathsf{U} \text{ is convex,}\\ &w_i \text{ is concave for } i = 1, \ldots, I. \end{aligned}$$

As a consequence $(\boldsymbol{c}, u) \mapsto H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u)$ is convex, nonnegative, and $H_{\mathsf{q},\mathsf{U}}(\boldsymbol{c}, u) = 0$ if and only if $(\boldsymbol{c}, u) = (\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U})$.

Since we are not interested in quantitative results, we can use the simplified form $\mathcal{P} = \widehat{\mathcal{P}}_\gamma + \widehat{\mathcal{P}}_\mathrm{R}$ with $\gamma \in\,]0, 1[$ for the entropy production, where we recall

$$\begin{aligned} &\widehat{\mathcal{P}}_\gamma(\boldsymbol{c}, u) = \int_\Omega \Big(\frac{|\nabla u|^2}{u^{2-\gamma}} + \sum_{i=1}^I \frac{|\nabla c_i|^2}{c_i}\Big)\,\mathrm{d}x \;\text{ and }\; \widehat{\mathcal{P}}_\mathrm{R}(\boldsymbol{c}, u) := \int_\Omega \widehat{P}_\mathrm{R}(\boldsymbol{c}, u)\,\mathrm{d}x,\\ &\text{where } \widehat{P}_\mathrm{R}(\boldsymbol{c}, u) = \sum_{r=1}^R G\Big(\frac{\boldsymbol{c}^{\alpha^r}}{\boldsymbol{w}(u)^{\alpha^r}}, \frac{\boldsymbol{c}^{\beta^r}}{\boldsymbol{w}(u)^{\beta^r}}\Big)\,\mathrm{d}x \text{ with } G(a, b) = (a-b)\log(a/b). \end{aligned} \tag{4.1}$$

We will use the following basic assumptions on the entropy function h_U and on the equilibria relations $c_i = w_i(u)$. Throughout, we assume that (q, U) and $\gamma \in]0, 1[$ are fixed, and all subsequent constants may depend on these values with special indication.

$$\exists\, C>0\ \forall\, u\geq 0:\ \tfrac{1}{C}\min\{|u-\mathsf{U}|,(u-\mathsf{U})^2\}\leq h_{\mathsf{U}}(u)\leq C\min\{|u-\mathsf{U}|,(u-\mathsf{U})^2\}; \tag{4.2a}$$

$$\exists\, C>0\ \forall\, u\geq 0,\ i=1,\ldots,I:\quad 0<w_i(0)\leq w_i(u)\leq C(1+u)^{\gamma}; \tag{4.2b}$$

$$\forall\, i=1,\ldots,I:\quad \text{the function } w_i \text{ is concave;} \tag{4.2c}$$

$$\text{the function } \widehat{h}_{\mathsf{U}}=h_{\mathsf{U}}-I+\sum\nolimits_{i=1}^{I} w_i \text{ is convex.} \tag{4.2d}$$

Obviously, the functions w_i must be increasing, since they are concave and bounded from below by the positive value $w_i(0)$.

The main result of this section is the follow EEP estimate that will be obtained by a compactness method that is much more general than the convexity method discussed in the previous section. However, in this result the constant $K_M(\mathsf{q},\mathsf{U})$ also depends on the upper bound M for $\mathcal{H}_{\mathsf{q},\mathsf{U}}$.

Theorem 4.1 (EEP estimate via compactness) *Consider a bounded Lipschitz domain $\Omega\subset\mathbb{R}^d$ and let the functions h and w_i satisfy the assumptions (4.2) with*

$$2_d\gamma>2,\quad \textit{where } 2_d=\frac{2d}{d-2} \textit{ for } d\geq 3 \textit{ and } \infty \textit{ else.} \tag{4.3}$$

Moreover, assume the unique equilibrium condition UEC (2.19). Then, for each $M>0$ and each $(\mathsf{q},\mathsf{U})\in\mathfrak{Q}$, there exists a constant $K_M(\mathsf{q},\mathsf{U})>0$ such that

$$\forall\,(\boldsymbol{c},u)\in\mathfrak{S}(\mathsf{q},\mathsf{U}) \textit{ with } \mathcal{H}(\boldsymbol{c},u)\leq M:\ \widehat{\mathcal{P}}_{\gamma}(\boldsymbol{c},u)+\widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{c},u)\geq K_M(\mathsf{q},\mathsf{U})\mathcal{H}_{\mathsf{q},\mathsf{U}}(\boldsymbol{c},u). \tag{4.4}$$

The proof of this result is the content of the remaining subsection of Sect. 4. It relies on two basic facts about functionals $\mathcal{F}(\lambda,z)=\int_\Omega F(\lambda,z(x))\,\mathrm{d}x$, where Ω is an open and bounded subset of $\mathbb{R}^d$.

a. If $F:[0,1]\times\mathbb{R}^m\to\mathbb{R}\cup\{\infty\}$ is nonnegative and lower semicontinuous, then

$$(\lambda_n,\boldsymbol{y}_n)\to(\lambda^*,\boldsymbol{y}^*) \text{ in } \mathbb{R}\times\mathrm{L}^1(\Omega)^m\quad\Longrightarrow\quad \mathcal{F}(\lambda^*,z^*)\leq\liminf_{n\to\infty}\mathcal{F}(\lambda_n,z_n). \tag{4.5a}$$

b. If $F:[0,1]\times\mathbb{R}^m\to\mathbb{R}$ is continuous and satisfies $0\leq F(\lambda,z)\leq C(1+z|)^p$ for some $p\geq 1$ and $C>0$, then

$$(\lambda_n,\boldsymbol{y}_n)\to(\lambda^*,\boldsymbol{y}^*) \text{ in } \mathbb{R}\times\mathrm{L}^p(\Omega)^m\quad\Longrightarrow\quad \mathcal{F}(\lambda_n,z_n)\to\mathcal{F}(\lambda^*,z^*). \tag{4.5b}$$

Clearly, (4.5a) follows from Fatou's lemma and (4.5b) from Lebesgue's dominated convergence theorem, see (Fonseca and Leoni 2007, Lem. 1.83) and (Fonseca and Leoni 2007, Thm. 1.85). The major point is that we will apply (4.5a) for the functional $\widehat{\mathcal{P}}_{\mathrm{R}}$, where the density $\widehat{P}_{\mathrm{R}}$ may have arbitrarily high growth because $\boldsymbol{\alpha}^r,\boldsymbol{\beta}^r\in\mathbb{N}_0^I$ are not restricted. In contrast, the continuity result (4.5b) will be used for $\mathcal{H}$ or a rescaled version of it, where the growth is fixed and independent of the stoichiometric vectors $\boldsymbol{\alpha}^r$ and $\boldsymbol{\beta}^r$.

In the rest of this section, the value of $(\mathsf{q},\mathsf{U})\in\mathfrak{Q}$ is fixed. Hence, we drop the subscripts at $\mathcal{H}$ and H in the sequel.

4.2 Some Preliminary Estimates

Here, we provide upper and lower bounds that are needed to obtain coercivity as well as the upper bound for applying (4.5b).

We first provide a lower bound on the relative entropy density H. We use that the function $\lambda_{\mathfrak{B}} : z \mapsto z \log z - z + 1$ satisfies the elementary estimates

$$\forall \delta > 0 \, \exists C_\delta > 0 \, \forall z \geq 0 : \quad \big(1-\sqrt{z}\big)^2 \leq \lambda_{\mathfrak{B}}(z) \leq C_\delta \big(1-\sqrt{z}\big)^2 (1+z)^\delta. \tag{4.6}$$

This implies $w\lambda_{\mathfrak{B}}(c/w) \geq \big(\sqrt{c}-\sqrt{w}\big)^2 \geq \frac{c}{2} - w$ and with (4.2a), (4.2b), and $\gamma < 1$, we obtain

$$\exists C > 0 \, \forall \boldsymbol{c}, u : \quad H(\boldsymbol{c}, u) \geq \frac{1}{C} u - C + \frac{1}{2} \sum_{i=1}^{I} c_i. \tag{4.7}$$

Since $\widehat{\mathcal{P}}_{\text{diff}}$ gives a bound on $\nabla\big(u^{\gamma/2}\big)$ it will be important to control h and w_i in terms of $v = u^{\gamma/2}$. We will use the following results.

Lemma 4.2 *Assume that h and w_i satisfy (4.2) with $\mathsf{U} > 0$ and $\gamma \in \,]0, 1[$. Then, there exists $\widehat{c}_h, C, C_w > 0$ such that*

$$\forall u \geq 0 : \quad \widehat{c}_h \big|u^{\gamma/2}-\mathsf{U}^{\gamma/2}\big|^2 \leq h(u) \leq C\Big(\big|u^{\gamma/2}-\mathsf{U}^{\gamma/2}\big|^2 + \big|u^{\gamma/2}-\mathsf{U}^{\gamma/2}\big|^{2/\gamma}\Big), \tag{4.8a}$$

$$\forall v \geq -\mathsf{U}^{\gamma/2} : \quad \Big|\Big(w_i\big[(\mathsf{U}^{\gamma/2}+v)^{2/\gamma}\big]\Big)^{1/2} - \big(w_i(\mathsf{U})\big)^{1/2}\Big| \leq C_w \, |v|. \tag{4.8b}$$

Proof The lower bound in (4.8a) follows simply by observing that the lower bound h in (4.2a) and the left-hand side of (4.8a) both behave as $(u-\mathsf{U})^2$ for u near U. Otherwise, h is strictly positive with linear growth for $u \to \infty$, while the left-hand side in (4.8a) has only growth like u^γ with $\gamma < 1$.

For the upper bound, we argue similarly that h and the right-hand side behave like $(u-\mathsf{U})^2$ for u near U. Otherwise, h and the right-hand side are bounded linearly, so the result follows.

To show (4.8b) we let $m(v) = (\mathsf{U}^{\gamma/2}+v)^{2/\gamma}$ and $\widetilde{w}(v) = \big[w(m(v))\big]^{1/2}$. Hence, we have to find C such that $|\widetilde{w}(v) - \widetilde{w}(0)| \leq C|v|$.

Case $v \in [-\mathsf{U}^{\gamma/2}, 0]$: Concavity and monotonicity of w_i implies concavity and monotonicity of $w_i^{1/2}$. Using $m(v) \leq m(0) = \mathsf{U}$ we immediately find $\widetilde{w}(0) \geq \widetilde{w}(v)$. Thus, it suffices to estimate $\widetilde{w}(v) - \widetilde{w}(0)$ by $C_- v$ from below. Using $m(0) = \mathsf{U}$ we have

$$\begin{aligned} \widetilde{w}(v) - \widetilde{w}(0) &= w(m(v))^{1/2} - w(\mathsf{U})^{1/2} \\ &\geq \big(1 - \tfrac{m(v)}{\mathsf{U}}\big) w(0)^{1/2} + \tfrac{m(v)}{\mathsf{U}} w(\mathsf{U})^{1/2} - w(\mathsf{U})^{1/2} \\ &= \big(m(v) - m(0)\big) \frac{1}{\mathsf{U}} \big(w(\mathsf{U})^{1/2} - w(0)^{1/2}\big). \end{aligned}$$

Since $w(\mathsf{U}) \geq w(0)$ it suffices to estimate $m(v) - m(0)$ from below. Because m is convex we find $m(v) - m(0) \geq m'(0)v$, where $m'(0) > 0$. Thus, for $v \in [-\overline{u}^{\gamma/2}, 0]$ we have

$$\widetilde{w}(v) - \widetilde{w}(0) \geq C_- v \quad \text{with } C_- = 2\big(w(\mathsf{U})^{1/2} - w(0)^{1/2}\big)/\big(\gamma\, \mathsf{U}^{\gamma/2}\big).$$

Case $v \geq 0$: We have $\widetilde{w}'(v) = \frac{w'(m(v))}{2\widetilde{w}(v)} m'(v)$. Moreover, concavity of w implies $0 < w(0) \leq w(u) + w'(u)(0-u)$, which implies $w'(u) \leq w(u)/u$. With $m'(v) = \frac{2}{\gamma} m(v)^{1-\gamma/2}$ and $u = m(v)$ we have

$$\widetilde{w}'(v) = \frac{w'(u)}{\gamma\, w(u)^{1/2}} u^{1-\gamma/2} \leq \frac{w(u)^{1/2}}{\gamma u^{\gamma/2}} \leq \frac{1}{\gamma}\Big(\frac{C(1+u^\gamma)}{u^\gamma}\Big)^{1/2} \leq C_+ \text{ for } u \geq \mathsf{U}.$$

Thus, we conclude $\widetilde{w}(0) \leq \widetilde{w}(v) \leq \widetilde{w}(0) + C_+ v$ for $v \geq 0$, and (4.8b) is established with $C_w = \max\{C_-, C_+\}$. □

4.3 Extraction of a Converging Subsequence

We assume that the result is not true and will generate a contradiction at the end of Sect. 4.6. Since the result is not true, there is a sequence $\boldsymbol{y}_n = (\boldsymbol{c}_n, u_n)$ such that

$$0 < \mathcal{H}(\boldsymbol{y}_n) \leq R, \quad \mathsf{Q}\boldsymbol{y}_n = \mathsf{q}, \quad \widehat{\mathcal{P}}(\boldsymbol{y}_n) := \widehat{\mathcal{P}}_\gamma(\boldsymbol{y}_n) + \widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{y}_n) \leq \frac{1}{n}\mathcal{H}(\boldsymbol{y}_n) \leq \frac{M}{n}. \tag{4.9}$$

We define the auxiliary functions

$$a_{n,i} := c_{n,i}^{1/2} \text{ for } i = 1, \ldots, I \quad \text{and} \quad a_{n,I+1} := u_n^{\gamma/2},$$

and define the sequence of vector-valued function $\boldsymbol{a}_n : \Omega \to [0, \infty[^{I+1}$. The bound $\mathcal{H}(\boldsymbol{y}_n) \leq M$ together with the coercivity (4.7) yields the bound

$$\|\boldsymbol{a}_n\|_{\mathrm{L}^2(\Omega)^{I+1}} \leq C.$$

Moreover, $\widehat{\mathcal{P}}_\gamma(\boldsymbol{y}_n) \leq \widehat{\mathcal{P}}(\boldsymbol{y}_n) \leq M/n$ gives the bound

$$\|\nabla \boldsymbol{a}_n\|_{\mathrm{L}^2(\Omega)^{I+1}} \leq \frac{C}{\sqrt{n}}.$$

Thus, we conclude that, up to the extraction of a subsequence, the sequence $\boldsymbol{a}_n$ converges strongly in $\mathrm{H}^1(\Omega)$ to a limit $\boldsymbol{a}^*$ with $\nabla \boldsymbol{a}^* \equiv 0$, i.e., each component of the limit is a constant function.

Choosing $q > 1$ with $2q \in\,]2/\gamma, 2_d]$, which exists by assumption (4.3), we obtain $\boldsymbol{a}_n \to \boldsymbol{a}^*$ in $\mathrm{L}^{2q}(\Omega)^{I+1}$. Recalling the definition of $\boldsymbol{a}_n$ yields

$$\begin{aligned} c_{n,i} &= (a_{n,i})^2 \to (a_i^*)^2 =: c_i^* \text{ in } \mathrm{L}^q(\Omega), \\ u_n &= (a_{n,I+1})^{2/\gamma} \to (a_{I+1}^*)^{2/\gamma} =: u^* \text{ in } \mathrm{L}^{\gamma q}(\Omega). \end{aligned} \tag{4.10}$$

This defines a limit vector $\boldsymbol{y}^* = (\boldsymbol{c}^*, u^*)$.

4.4 Identification of the Limit a^*

Since the functional Q is linear, it is continuous in $\mathrm{L}^1(\Omega)$ and we conclude

$$\mathsf{Q}\boldsymbol{y}^* = \lim_{n\to\infty} \mathsf{Q}\boldsymbol{y}_n = \lim_{n\to\infty} \mathsf{q} = \mathsf{q}.$$

By energy conservation, this means in particular that $u^* = \mathsf{U}$.

To identify the other components of $\boldsymbol{a}^*$ and thus the limits of $c_{n,i}$, we use the relation $\mathcal{P}_\mathrm{R}(\boldsymbol{y}_n) \le M/n$. We note that the integrand $\widehat{P}_\mathrm{R}$ of $\mathcal{P}_\mathrm{R}$ is nonnegative and continuous, since $0 < w_i(0) \le w_i(u)$, see (4.2a). Hence, $\mathcal{P}_\mathrm{R}$ is lower semicontinuous with respect to the strong topology in $\mathrm{L}^1(\Omega)^{I+1}$, see (4.5a). This implies

$$0 \le \mathcal{P}_\mathrm{R}(\boldsymbol{c}^*, \mathsf{U}) \le \liminf_{n\to\infty} \mathcal{P}_\mathrm{R}(\boldsymbol{c}_n, u_n) \le \lim_{n\to\infty} M/n = 0.$$

Together with $\mathsf{Q}\boldsymbol{y}^* = \mathsf{q}$ and the uniqueness assumption (2.19), we conclude that $\boldsymbol{y}^*$ is the unique steady state in the stoichiometric subspace, i.e., $\boldsymbol{y}^* = (\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U})$.

4.5 Blowup Near the Equilibrium

To generate a contradiction, we set

$$\lambda_n^2 := \mathcal{H}(\boldsymbol{y}_n) > 0$$

and first show $\lambda_n \to 0$. This follows by using the fact that the density $H(\boldsymbol{c}, u) = h(u) + \sum w_i(u)\lambda_{\mathfrak{B}}\big(c_i/w_i(u)\big)$ is a continuous function and satisfies, for all $\delta > 0$, the bound

$$0 \le H(\boldsymbol{c}, u) \le C_\delta\Big(1 + u + \big(1{+}|\boldsymbol{c}|{+}u^\gamma\big)\big(1{+}|\boldsymbol{c}|\big)^\delta\Big)$$

for some $C_\delta > 0$. For this we use (4.6), giving $w_i\lambda_{\mathfrak{B}}(c_i/w_i) \le C_\delta(c_i^{1/2}{-}w_i^{1/2})^2\big(1{+}c_i/w_i\big)^\delta \le \widetilde{C}_\delta(c_i{+}w_i)(1{+}c_i)^\delta$, together with (4.2a) and (4.2b).

Choosing $\delta \in \,]0, \gamma q{-}1]$, where we use (4.3) once again, we see that $\mathcal{H}$ is strongly continuous on $\mathrm{L}^q(\Omega)^I \times \mathrm{L}^{\gamma q}(\Omega)$ such that

$$\lim_{n\to\infty} \lambda_n^2 = \lim_{n\to\infty} \mathcal{H}(\boldsymbol{y}_n) = \mathcal{H}(\boldsymbol{y}^*) = \mathcal{H}(\mathsf{w}(\mathsf{q}, \mathsf{U}), \mathsf{U}) = 0,$$

as desired.

We now define the blowup functions

$$b_{n,i} = \frac{1}{\lambda_n}\big(a_{n,i} - a_i^*\big) \quad \text{for } i = 1, \ldots, I{+}1.$$

Since $\widehat{\mathcal{P}}_\gamma$ expressed in $\boldsymbol{a}$ is exactly quadratic in ∇a_i, we obtain

$$1 = \frac{1}{\lambda_n^2}\mathcal{H}(\boldsymbol{y}_n) \geq \frac{n}{\lambda_n^2}\widehat{\mathcal{P}}_\gamma(\boldsymbol{y}_n) = \frac{n}{4\lambda_n^2}\Big(\gamma^2\|\nabla a_{n,I+1}\|^2_{\mathrm{L}^2} + \sum_1^I \|\nabla a_{n,i}\|^2_{\mathrm{L}^2}\Big)$$
$$= \frac{n}{4}\Big(\gamma^2\|\nabla b_{n,I+1}\|^2_{\mathrm{L}^2} + \sum_1^I \|\nabla b_{n,i}\|^2_{\mathrm{L}^2}\Big). \qquad (4.11)$$

To control $\boldsymbol{b}_n$ without derivatives, we use a rescaled coercivity of H, namely

$$H(\boldsymbol{c}, u) \geq \widehat{c}_h\big|u^{\gamma/2} - \mathsf{U}^{\gamma/2}\big|^2 + 4\sum_1^I \big(c_i^{1/2} - w_i(u)^{1/2}\big)^2,$$

which follows from the lower estimate in (4.8a) and $w\lambda_{\mathfrak{B}}(c/w) \geq 4(c^{1/2}-w^{1/2})^2$. Inserting $\boldsymbol{y}_n = \Phi(\lambda_n\boldsymbol{b}_n)$ given by

$$\Phi(\lambda\boldsymbol{b})_i = \big(w_i(\mathsf{U})^{1/2} + \lambda b_i\big)^2 \;\text{ and }\; \Phi(\lambda, \boldsymbol{b})_{I+1} = \big(\mathsf{U}^{\gamma/2} + \lambda b_{I+1}\big)^{2/\gamma} \qquad (4.12)$$

and choosing any $\vartheta \in\,]0, 4[$ we arrive at the estimate

$$\frac{1}{\lambda_n^2}H(\boldsymbol{c}_n, u_n) \geq \widehat{c}_h|b_{n,I+1}|^2 + \frac{\vartheta}{\lambda_n^2}\sum_{i=1}^I \Big(\lambda_n b_{n,i} + w_i(\mathsf{U})^{1/2}$$
$$- \big[w_i\big((\mathsf{U}^{\gamma/2}{+}\lambda_n b_{n,I+1})^{2/\gamma}\big)\big]^{1/2}\Big)^2$$
$$\overset{(1)}{\geq} \widehat{c}_h|b_{n,I+1}|^2 + \vartheta\sum_{i=1}^I\Big(\frac{1}{2}b_{n,i}^2 - C_w^2 b_{n,I+1}^2\Big)$$
$$\overset{(2)}{\geq} \frac{1}{C}\Big(|b_{n,I+1}|^2 + \sum_{i=1}^I \frac{1}{2}b_{n,i}^2\Big).$$

In "$\overset{(1)}{\geq}$" we first used $(a{+}b)^2 \geq \frac{1}{2}a^2 - b^2$ and then (4.8b), while "$\overset{(2)}{\geq}$" follows by choosing a suitable ϑ, e.g., $\vartheta = \min\{4, \widehat{c}_h/(2IC_w^2)\}$. As a result we conclude that

$$1 = \frac{1}{\lambda_n^2}\mathcal{H}(\boldsymbol{y}_n) \geq \frac{1}{C}\|\boldsymbol{b}_n\|^2_{\mathrm{L}^2}.$$

Together with (4.11) we see that, after choosing a subsequence (not relabeled),

$$\boldsymbol{b}_n \to \boldsymbol{b}^* = \text{const.} \quad \text{in } \mathrm{H}^1(\Omega)^{I+1}.$$

4.6 The Final Contradiction

Here and in the sequel, we shortly write $\mathsf{w} = (\mathsf{w}_1, \ldots, \mathsf{w}_I) = \mathsf{w}(\mathsf{q}, \mathsf{U})$. The mapping $\boldsymbol{y} = \Phi(\lambda \boldsymbol{b})$ introduced in (4.12) satisfies

$$\Phi(0) = (\mathsf{w}, \mathsf{U}) \ \text{ and } \ \mathrm{D}\Phi(0) = \mathrm{diag}\big(2\mathsf{w}_1^{1/2}, \ldots, 2\mathsf{w}_I^{1/2}, \frac{2}{\gamma}\mathsf{U}^{1-\gamma/2}\big),$$

which shows that the weights $w_i(\mathsf{U})^{1/2}$ are relevant. We will show that $\boldsymbol{b}^*$ satisfies the following three relations

$$\boldsymbol{s} := \big(\mathsf{w}_i^{1/2} b_i^*\big)_{i=1,\ldots,I} \in \mathbb{S} \quad \text{and } \boldsymbol{b}_{I+1}^* = 0; \tag{4.13a}$$

$$\sum_{r=1}^{R} \Big(\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \cdot \boldsymbol{d}\Big)^2 = 0 \text{ with } \boldsymbol{d} := \big(b_i^*/\mathsf{w}_i^{1/2}\big)_{i=1,\ldots,I}, \text{ giving } \boldsymbol{d} \in \mathbb{S}^{\perp}; \tag{4.13b}$$

$$\boldsymbol{b}^* \neq 0. \tag{4.13c}$$

Obviously, these three relations cannot hold simultaneously, since $\boldsymbol{s} \in \mathbb{S}$ and $\boldsymbol{d} \in \mathbb{S}^{\perp}$ implies $0 = \boldsymbol{s} \cdot \boldsymbol{d}$, but obviously $\boldsymbol{s} \cdot \boldsymbol{d} = |\boldsymbol{b}_*|^2$, which contradicts (4.13c).

To derive (4.13a), we observe $\mathsf{Q}\boldsymbol{y}_n = \mathsf{q} = \mathsf{Q}\boldsymbol{y}^*$, hence $\boldsymbol{z}_n = \frac{1}{\lambda_n}\big(\boldsymbol{y}_n - \boldsymbol{y}^*\big)$ satisfies $\mathsf{Q}\boldsymbol{z}_n = 0$. Moreover, the convergence of $\boldsymbol{b}_n$ and the relations (4.12) easily provide the convergence

$$\boldsymbol{z}_n \to \boldsymbol{z}^* := \Big(2\mathsf{w}_1^{1/2} b_1^*, \ldots, 2\mathsf{w}_I^{1/2} b_I^*, \frac{2}{\gamma}\mathsf{U}^{1-\gamma/2} b_{I+1}^*\Big) = \Big(2\boldsymbol{s}, \frac{2}{\gamma}\mathsf{U}^{1-\gamma/2} b_{I+1}^*\Big).$$

Thus, (4.13a) is established, since it is equivalent to $\mathsf{Q}\boldsymbol{z}^* = 0$.

To obtain (4.13b), we use the relation

$$1 = \frac{1}{\lambda_n^2}\mathcal{H}(\boldsymbol{y}_n) \geq \frac{n}{\lambda_n^2}\widehat{\mathcal{P}}_{\mathrm{R}}(\boldsymbol{y}_n), \text{ which implies } \frac{1}{\lambda_n^2}\widehat{\mathcal{P}}_{\mathrm{R}}\big(\Phi(\lambda_n, \boldsymbol{b}_n)\big) \to 0. \tag{4.14}$$

To exploit this, we define the auxiliary functional

$$\begin{aligned}
&\mathcal{K}_{\mathrm{R}}(\lambda, \boldsymbol{b}) := \frac{1}{\lambda^2}\widehat{\mathcal{P}}_{\mathrm{R}}\big(\Phi(\lambda, \boldsymbol{b})\big) = \int_{\Omega} K_{\mathrm{R}}(\lambda, \boldsymbol{b}(x))\,\mathrm{d}x, \text{ where} \\
&K_{\mathrm{R}}(\lambda, \boldsymbol{b}) = \frac{1}{\lambda^2}\widehat{P}_{\mathrm{R}}\big(\Phi(\lambda, \boldsymbol{b})\big) \text{ for } \lambda > 0 \text{ and} \\
&2K_{\mathrm{R}}(0, \boldsymbol{b}) = \sum_{r=1}^{R}\Big(\big(\boldsymbol{\alpha}^r - \boldsymbol{\beta}^r\big) \cdot \big(\frac{2b_i}{\mathsf{w}_i^{1/2}} - \mu\frac{\mathsf{w}_i'}{\mathsf{w}_i} b_{I+1}\big)\Big)^2 \text{ with } \mathsf{w}_i' = \partial_{\mathsf{U}}\mathsf{w}_i(\mathsf{q}, \mathsf{U}),
\end{aligned}$$

where $\mu = 2/(\gamma\mathsf{U}^{\gamma/2-1})$ and $\widehat{P}_{\mathrm{R}}$ may be extended by ∞ if $\boldsymbol{y} = \Phi(\lambda, \boldsymbol{b})$ does not lie in $[0, \infty[^{I+1}$. By this definition, the function K_{R} is nonnegative and lower semicontinuous in $(\lambda, \boldsymbol{b}) \in [0, \infty[\times \mathbb{R}^{I+1}$; hence, the functional $\mathcal{K}_\mathrm{R}$ is strongly lower semicontinuous in L^2, see (4.5a). Thus,

$$\boldsymbol{b}_n \to \boldsymbol{b}^* \text{ implies } \mathcal{K}_\mathrm{R}(\lambda_n, \boldsymbol{b}_n) \to 0 = \mathcal{K}_\mathrm{R}(0, \boldsymbol{b}^*) = K_\mathrm{R}(0, \boldsymbol{b}^*).$$

Since we already know $b^*_{I+1} = 0$, we conclude (4.13b).

For the final relation, we introduce the integral density function K_H via

$$\text{for } \lambda \in]0, 1]: \ K_\mathrm{H}(\lambda, \boldsymbol{b}) = \frac{1}{\lambda^2} H\big(\Phi(\lambda, \boldsymbol{b})\big),$$
$$\text{for } \lambda = 0: \quad K_\mathrm{H}(0, \boldsymbol{b}) = \frac{4h''(\mathsf{U})}{\gamma^2 \mathsf{U}^{\gamma-2}}\, b^2_{I+1} + \sum_{i=1}^{I} \left(2b_i - \frac{\mathsf{w}_i'}{\mathsf{w}_i^{1/2}}\, \mu\, b_{I+1}\right)^2 .$$

We define K_H only on the closed subset

$$\mathrm{dom}(K_\mathrm{H}) = \{0\} \times \mathbb{R}^{I+1} \cup \{\, (\lambda, \boldsymbol{b}) \in]0, 1] \times \mathbb{R}^{I+1} \mid b_{I+1} \geq -\tfrac{\mathsf{U}^{\gamma/2}}{\lambda},\ b_i \geq -\tfrac{\mathsf{w}_i^{1/2}}{\lambda} \text{ for } i = 1, \ldots, I \,\}.$$

By construction, the function K_H is nonnegative and continuous on its domain. We now establish an upper bound for K_H using the upper bound for h in (4.8a) and the upper bound of $w\lambda_\mathfrak{B}(c/w) \leq C_\delta (c^{1/2} - w^{1/2})^2 (1 + c/w)^\delta$, see (4.6). Using the shorthand $\Phi_{I+1}(\boldsymbol{b}) = (\mathsf{U}^{\gamma/2} + \lambda b_{I+1})^{2/\gamma}$, we have

$$\begin{aligned}
\frac{h\big(\Phi_{I+1}(\boldsymbol{b})\big)}{\lambda^2} &\leq \frac{C}{\lambda^2}\Big(\big(\lambda b_{I+1}\big)^2 + \big(\lambda b_{I+1}\big)^{2/\gamma}\Big) \leq C\big(b^2_{I+1} + b^{2/\gamma}_{I+1}\big),\\
\frac{w_i(u)\lambda_\mathfrak{B}\big(\frac{c_i}{w_i(u)}\big)}{\lambda^2} &\leq \frac{C_\delta}{\lambda^2}\Big(\mathsf{w}_i^{1/2} + \lambda b_i - w_i\big(\Phi_{I+1}(\boldsymbol{b})\big)^{1/2}\Big)^2 \left(1 + \frac{\big(\mathsf{w}_i^{1/2} + \lambda b_i\big)^2}{w_i\big(\Phi_{I+1}(\boldsymbol{b})\big)}\right)^\delta\\
&\overset{*}{\leq} C_\delta \big(|b_i| + C_w |b_{I+1}|\big)^2 \Big(1 + \big(\mathsf{w}_i^{1/2} + \lambda b_i\big)^2 / w_i(0)\Big)^\delta\\
&\leq C_{K_\mathrm{H}}\big(1 + |\boldsymbol{b}|\big)^{2+2\delta} \quad \text{for all } (\lambda, \boldsymbol{b}) \in \mathrm{dom}(K_\mathrm{H}),
\end{aligned}$$

where we used $\lambda \in]0, 1]$ and $\gamma \in]0, 1[$ and employed (4.8b) and $w_i(u) \geq w_i(0)$ for "$\overset{*}{\leq}$."

Of course, the construction of K_H was such that, for $\lambda > 0$, we have

$$\frac{1}{\lambda^2}\mathcal{H}\big(\Phi(\lambda, \boldsymbol{b})\big) = \mathcal{K}_\mathrm{H}(\lambda, \boldsymbol{b}) := \int_\Omega K_\mathrm{H}(\lambda, \boldsymbol{b}(x))\, \mathrm{d}x.$$

The upper bound on K_H and its continuity imply continuity of $\mathcal{K}_H$, see (4.5b). More precisely, from $\boldsymbol{b}_n \to \boldsymbol{b}_*$ in $\mathrm{H}^1(\Omega)^{I+1}$ we obtain strong convergence in $\mathrm{L}^{2q}(\Omega)$ for some $q > 1$. Choosing $\delta \in]0, q-1]$ we use that for all $\boldsymbol{y}_n$, we have $(\lambda, \boldsymbol{b}_n(x)) \in \mathrm{dom}(K_\mathrm{H})$ for almost all $x \in \Omega$. Then, using (4.5b), we obtain

$$1 = \frac{1}{\lambda_n^2}\mathcal{H}(\boldsymbol{y}_n) = \mathcal{K}_{\rm H}(\lambda_n, \boldsymbol{b}_n) \to \mathcal{K}_{\rm H}(0, \boldsymbol{b}_*)$$
$$= \left(\frac{4h''(u)}{\gamma^2 \mathsf{U}^{\gamma-2}} (b_{I+1}^*)^2 + \sum_{i=1}^{I} \left(2b_i^* - \frac{\mathsf{w}_i'}{\mathsf{w}_i^{1/2}} \mu\, b_{I+1}^* \right)^2 \right).$$

Thus, (4.13c) is established and the contradiction is complete.

This finishes the proof of Theorem 4.1.

Acknowledgements A.M. was partially supported by Deutsche Forschungsgemeinschaft (DFG) through SFB 1114 (Scaling Cascades in Complex Systems), subproject C05 *Effective Models for Materials and Interfaces with Multiple Scales*. M.M. was supported by the European Research Council (ERC) through the Advanced Grant No. 267802 *Analysis of MultiScale Systems Driven by Functionals*.

A Generalized EEP Estimates

This appendix gives a proof of the EEP estimate (3.6) for general γ and α. For a more detailed exposition, we refer to Mittnenzweig (2014). We will use the Poincaré-Sobolev inequality

$$\|\nabla u\|_p \geq \eta_{\rm PS}(\Omega, p, q)\, \|u - \overline{u}\|_q \tag{A.1}$$

corresponding to the Sobolev embedding $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$, where $\eta_{\rm PS}(\Omega, p, q)$ is the optimal Poincaré-Sobolev constant. This inequality holds true for $p \geq d$ or for $p < d$ and $q \leq dp/(d-p)$.

Theorem A.1 *Let* $\gamma > \max\{0, 1-2/d\}$ *and* $\alpha \leq \gamma + 2/d$ *for* $d \geq 3$ *or* $\alpha < 1+\gamma$ *for* $d \leq 2$. *Then* $\rho(\Omega, \gamma, \alpha)$ *is strictly positive and can be bounded by*

$$\rho(\Omega, \gamma, \alpha) \geq \frac{4\alpha\big(\gamma - 2/q\big)}{\gamma^2 (q-1)}\, \eta_{PS}(\Omega, 2, q)^2 \tag{A.2}$$

with $q \geq \max\{2/\gamma, 2/(1+\gamma-\alpha)\}$.

Proof We substitute $v = u^{\gamma/2}$. Then

$$\overline{u}^\gamma \int_\Omega F_{\alpha_1}\left(\frac{u}{\overline{u}}\right) {\rm d}x \leq \frac{\alpha_2}{\alpha_1} \overline{u}^\gamma \int_\Omega F_{\alpha_2}\left(\frac{u}{\overline{u}}\right) {\rm d}x = C \left(\frac{\|v\|_x^x}{\|v\|_y^{x-2}} - \|v\|_y^2 \right)$$
$$\leq C(\|v\|_z^2 - \|v\|_y^2) \leq C\,(z-1)\, \|v - \overline{v}\|_z^2 \leq \frac{C\,(z-1)}{\eta_{\rm PS}(\Omega, 2, z)^2} \|\nabla v\|_2^2$$

with $\alpha_2 > 1$, $C = 1/(\alpha_1(\alpha_2 - 1))$, $x = 2\alpha_2/\gamma$, $y = 2/\gamma$, and $z = 2/(1+\gamma-\alpha_2)$. In the first line we used monotonicity of the map $\alpha \mapsto \alpha F_\alpha(z)$, see (3.2). In the last line, we applied the Hölder inequality $\|v\|_x^x \leq \|v\|_y^{x-2} \|v\|_z^2$, the Poincaré–Sobolev inequality and the following estimate which holds true for $x \geq 2$ (see Mittnenzweig (2014) for a proof):

$$\|v\|_x^2 - \|v\|_y^2 \leq \|v\|_x^2 - \|v\|_1^2 \leq (x-1)\, \|v - \overline{v}\|_x^2$$

This proves the theorem. □

Remark A.2 The above theorem still remains true for the cases $\gamma = 1 - 2/d$ as well as $\alpha = \gamma + 1$ for $d \leq 2$, see Mittnenzweig (2014). In the case $\alpha = 2$ and $\gamma = 1$, which will be used below, we have the simple bound

$$\rho(\Omega, 1, 2) \geq 2\, \eta_{\mathrm{PS}}(\Omega, 1, 2)^2, \tag{A.3}$$

which follows from the following chain of estimates:

$$\overline{u} \int_\Omega \frac{|\nabla u|^2}{u}\,\mathrm{d}x \geq \|\nabla u\|_1^2 \geq \eta_{\mathrm{PS}}(\Omega, 1, 2)^2\, \|u - \overline{u}\|^2 = 2\, \eta_{\mathrm{PS}}(\Omega, 1, 2)^2 \cdot \overline{u}^2 \int_\Omega F_2(u/\overline{u})\,\mathrm{d}x.$$

B The EEP Estimate for $\alpha = \gamma = 0$ and $d \leq 2$

This appendix provides a proof that the EEP inequality also holds in the case $\alpha = \gamma = 0$. Here we only give a densified version of the full proof and refer to Mittnenzweig (2014) for a detailed exposition, which is based on the approach developed in Carlen and Loss (1992) and Carlen et al. (2010). We note that $\rho(\Omega, 1, 2)$ is positive only for $\Omega \subset \mathbb{R}^d$ with $d \leq 2$.

Theorem B.1 (Case $\alpha = \gamma = 0$ in dimension $d \leq 2$) *We have the estimate* $\rho(\Omega, 0, 0) \geq \varrho := \rho(\Omega, 1, 2) > 0$, *i.e.,*

$$\forall\, u > 0: \quad \int_\Omega \frac{|\nabla u(x)|^2}{u(x)^2}\,\mathrm{d}x \geq \rho(\Omega, 1, 2) \int_\Omega F_0(u(x)/\mathsf{U})\,\mathrm{d}x. \tag{B.1}$$

Proof We set $u = \mathrm{e}^v$ and observe that (B.1) is equivalent to

$$H(v) := \int_\Omega |\nabla v|^2\,\mathrm{d}x \geq \varrho G(v) \text{ with } G(v) := \Big(\log\Big(\int_\Omega \mathrm{e}^v\,\mathrm{d}x\Big) - \int_\Omega v\,\mathrm{d}x\Big). \tag{B.2}$$

We consider H and G as nonnegative, proper, and convex functionals on $\mathrm{L}^2(\Omega)$. Because of $\Omega \subset \mathbb{R}^d$ with $d \leq 2$, the Moser–Trudinger inequality G is bounded whenever H is bounded.

We establish (B.2) by Legendre transform for which we note

$$H^*(\psi) = \frac{1}{4} \int_\Omega (\psi - \overline{\psi})(-\Delta)^{-1}(\psi - \overline{\psi})\,\mathrm{d}x \text{ and } G^*(\psi) = \begin{cases} \int_\Omega \Gamma(\psi(x))\,\mathrm{d}x & \text{if } \overline{\psi} = 0, \\ \infty & \text{else,} \end{cases}$$

where $\Gamma(\psi) = \lambda_{\mathfrak{B}}(\psi + 1)$ for $\psi \geq -1$ and ∞ otherwise.

We now define $J(\psi) = G^*(\psi) - \varrho H^*(\psi)$ and show $J(\psi_0) \geq 0$ for all ψ_0. For this, we note $J(0) = 0$ and then consider J along the solutions $\psi_t := \mathrm{e}^{t\Delta}\psi_0$ of the diffusion equation $\dot{\psi} = \Delta\psi$. Setting $\psi_t = c_t - 1$ with $\overline{c}_t = 1$ and $c_t \geq 0$. By differentiating along solutions and using $\dot{c}_t = \Delta c_t$ (with Neumann boundary conditions), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} J(c_t - 1) = -\int_\Omega \frac{|\nabla c_t|^2}{c_t}\,\mathrm{d}x + \frac{\varrho}{2}\int_\Omega (c_t - 1)^2\,\mathrm{d}x \geq 0,$$

since $\varrho = \rho(\Omega, 1, 2)$ which is the optimal constant for the last estimate. Hence, $J(\psi_0) \geq J(\psi_t) \geq J(0) = 0$ since $\psi_t \to 0$.

Now, we reverse the Legendre transform using $H^*(\varrho\psi) = \varrho^2 H^*(\psi)$, hence

$$\begin{aligned} H(v) = \ & \sup\big(\langle v, \psi\rangle - H^*(\psi)\big) \overset{\psi = \varrho\widetilde{\psi}}{=} \sup\big(\langle v, \varrho\widetilde{\psi}\rangle - \varrho^2 H^*(\widetilde{\psi})\big) \\ & \overset{J \geq 0}{=} \varrho \sup\big(\langle v, \widetilde{\psi}\rangle - G^*(\widetilde{\psi})\big) = \varrho G(v). \end{aligned}$$

This proves the assertion. □

References

Albinus, G., Gajewski, H., Hünlich, R.: Thermodynamic design of energy models of semiconductor devices. Nonlinearity **15**(2), 367–383 (2002)

Burger, M., Di Franscesco, M., Pietschmann, J.-F., Schlake, B.: Nonlinear cross-diffusion with size exclusion. SIAM J. Math. Anal. **42**(6), 2842–2871 (2010)

Bothe, D., Dreyer, W.: Continuum thermodynamics of chemically reacting fluid mixtures. Acta Mech. **226**, 1757–1805 (2015)

Carlen, E.A., Loss, M.: Competing symmetries, the logarithmic HLS inequality and Onofri's inequality on $\mathbf{S}^n$. Geom. Funct. Anal. **2**(1), 90–104 (1992)

Carlen, E.A., Carrillo, J.A., Loss, M.: Hardy–Littlewood–Sobolev inequalities via fast diffusion flows. Proc. Natl. Acad. Sci. USA **107**(46), 19696–701 (2010)

Desvillettes, L., Fellner, K.: Exponential decay toward equilibrium via entropy methods for reaction–diffusion equations. J. Math. Anal. Appl. **319**(1), 157–176 (2006)

Desvillettes, L., Fellner, K., Tang, B.Q.: Trend to equilibrium for reaction–diffusion systems arising from complex balanced chemical reaction networks. SIAM J. Math. Anal. **49**(4), 2666–2709 (2017)

Dreyer, W., Druet, P.-É., Gajewski, P., Guhlke, C.: Existence of weak solutions for improved Nernst–Planck–Poisson models of compressible reacting electrolytes. WIAS preprint 2291 (2016)

Feinberg, M.: Complex balancing in general kinetic systems. Arch. Ration. Mech. Anal. **49**, 187–194 (1972/73)

Fellner, K., Tang, B.Q.: Explicit exponential convergence to equilibrium for nonlinear reaction–diffusion systems with detailed balance condition. Nonlinear Anal. **159**(C), 145–180 (2017)

Fischer, J.: Global existence of renormalized solutions to entropy-dissipating reaction–diffusion systems. Arch. Ration. Mech. Anal. **218**(1), 553–587 (2015)

Fonseca, I., Leoni, G.: Modern Methods in the Calculus of Variations: L^p Spaces. Springer, Berlin (2007)

Glitzky, A., Gröger, K., Hünlich, R.: Existence, uniqueness and asymptotic behaviour of solutions to equations modelling transport of dopants in semiconductors. In: Frehse, J., Gajewski, H. (eds), Special Topics in Semiconductor Analysis, pp. 49–78. Bonner Mathematische Schriften no. 258 (1994)

Glitzky, A., Gröger, K., Hünlich, R.: Free energy and dissipation rate for reaction diffusion processes of electrically charged species. Appl. Anal. **60**(3–4), 201–217 (1996)

Glitzky, A., Hünlich, R.: Energetic estimates and asymptotic for electro-reaction–diffusion systems. Z. Angew. Math. Mech. **77**(11), 823–832 (1997)

Gröger, K.: Asymptotic behavior of solutions to a class of diffusion–reaction equations. Math. Nachr. **112**, 19–33 (1983)

Gröger, K.: On the existence of steady states of certain reaction–diffusion systems. Arch. Ration. Mech. Anal. **92**(4), 297–306 (1986)

Gröger, K.: Free energy estimates and asymptotic behaviour of reaction–diffusion processes. WIAS preprint 20 (1992)

Haskovec, J., Hittmeir, S., Markowich, P.A., Mielke, A.: Decay to equilibrium for energy-reaction–diffusion systems. SIAM J. Math. Anal. (2017) (Accepted) WIAS preprint 2233

Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)

Jüngel, A., Matthes, D.: An algorithmic construction of entropies in higher-order nonlinear PDEs. Nonlinearity **19**(3), 633–659 (2006)

Jüngel, A.: The boundedness-by-entropy method for cross-diffusion systems. Nonlinearity **28**, 1963–2001 (2015)

Liero, M., Mielke, A.: Gradient structures and geodesic convexity for reaction–diffusion systems. Philos. Trans. Royal Soc. A **371**(2005), 20120346, 28 (2013)

Lieb, E.H., Yngvason, J.: The physics and mathematics of the second law of thermodynamics. Phys. Rep. **310**(1), 96 (1999)

Mielke, A.: Formulation of thermoelastic dissipative material behavior using GENERIC. Contin. Mech. Thermodyn. **23**(3), 233–256 (2011a)

Mielke, A.: A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems. Nonlinearity **24**, 1329–1346 (2011b)

Mielke, A.: Thermomechanical modeling of energy-reaction–diffusion systems, including bulk-interface interactions. Discr. Cont. Dyn. Syst. Ser. S **6**(2), 479–499 (2013)

Mielke, A.: Uniform exponential decay for reaction–diffusion systems with complex-balanced mass-action kinetics. In: Gurevich, P., (ed), Pattern of Dynamics, Springer Proceedings in Mathematics and Statistics, Springer (2017). WIAS preprint 2326 (in print)

Mielke, A., Haskovec, J., Markowich, P.A.: On uniform decay of the entropy for reaction–diffusion systems. J. Dyn. Differ. Equ. **27**(3–4), 897–928 (2015)

Mittnenzweig, M.: Entropy production inequalities for heat equations. Master's thesis, Freie Universität and WIAS Berlin (2014)

Onsager, L.: Reciprocal relations in irreversible processes, I+II. Phys. Rev. **37**, 405–426 (1931). (part II, 38:2265–2279)

Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. Commun. Partial Differ. Equ. **26**, 101–174 (2001)