Weak-strong uniqueness for
energy-reaction-diffusion systems

Katharina Hopf

submitted: January 27, 2021

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: katharina.hopf@wias-berlin.de

Weak-strong uniqueness for energy-reaction-diffusion systems

Katharina Hopf

ABSTRACT. We establish weak-strong uniqueness and stability properties of renormalised solutions to a class of energy-reaction-diffusion systems, which genuinely feature cross-diffusion effects. The systems considered are motivated by thermodynamically consistent models, and their formal entropy structure allows us to use as a key tool a suitably adjusted relative entropy method. Weak-strong uniqueness is obtained for general entropy-dissipating reactions without growth restrictions, and certain models with a non-integrable diffusive flux. The results also apply to a class of (isoenergetic) reaction-cross-diffusion systems.

1. INTRODUCTION

It is well acknowledged that the evolution of a system of diffusing and reacting chemicals is influenced by the thermal state of the system. *Energy-reaction-diffusion systems* (ERDS) take into account this thermal dependency by consistently coupling the evolution of the chemical concentrations $c = (c_1, \ldots, c_n)$ to a heat-type equation for the *internal energy density* u. Choosing the internal energy density as the thermal variable (as opposed to temperature for instance) has the advantage that the underlying physical entropy is jointly concave in the state variables $z = (u, c)$ [34, 35].

Recently, global existence of weak and renormalised solutions has been obtained for a class of thermodynamically consistent ERDS [24] taking the form (with $z := (u, c)$)

$$
\begin{cases}
\partial_t u = \nabla \cdot \big(A_{0j}(z) \nabla z_j \big), & t > 0, x \in \Omega, \\
\partial_t c_i = \nabla \cdot \big(A_{ij}(z) \nabla z_j \big) + R_i(z), & t > 0, x \in \Omega, \quad i \in \{1, \ldots, n\}, \\
0 = A_{ij}(z) \nabla z_j \cdot \nu, & t > 0, x \in \partial\Omega, \quad i \in \{0, \ldots, n\},
\end{cases}
\tag{1.1}
$$

see also the more explicit system (2.5). Eq. (1.1) is supplemented with an initial condition $(u, c)_{|t=0} = (u^{\mathrm{in}}, c^{\mathrm{in}})$ for $x \in \Omega$, where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain with outer unit normal ν. Note that we use the summation convention omitting the summation symbol in repeatedly occurring indices (here $\sum_{j=0}^n$). The diffusion matrix $A(z) = (A_{ij}(z))_{i,j=0,\ldots,n}$ and the reactions $(R_i(z))_{i=1,\ldots,n}$ are obtained from an underlying formal gradient structure based on entropy functionals $H(z) = \int_\Omega h(z)\, \mathrm{d}x$ with convex densities $h(z) = h(u, c)$ taking the form

$$
h(u, c_1, \ldots, c_n) = -\sigma(u) + \sum_{i=1}^n b(c_i, w_i(u)).
$$

Here, σ denotes the thermal part when the concentrations c_i are in their thermodynamic equilibrium $w_i = w_i(u)$, and $b(s, e) := e\lambda(s/e)$ with λ denoting the Boltzmann function, see Section 1.1 for details. The absence of reactions in the u-component of (1.1) reflects the property of conservation of the total (internal) energy $\int_\Omega u$.

ERDS genuinely feature cross-diffusion effects, such as concentration flux driven by gradients of the internal energy density and energy flux due to concentration gradients, which are one of the main sources of difficulties in their analysis. These phenomena are closely linked to the thermodynamic origin of ERDS, and are related to the *Soret effect* and the *Dufour effect*, well-known in physics, which describe concentration flux due to temperature gradients resp. heat flux driven by concentration gradients.

In the present manuscript, we aim to derive stability properties including a weak-strong uniqueness result for ERDS (1.1) based on their thermodynamic structure. This can be seen as a way of justifying the weak solution concept in [24]. The main contribution of [24] was to identify non-trivial classes of thermodynamically consistent models that allow for an existence theory of generalised solutions. Interestingly, even in a cross-diffusion dominant regime and without physically restrictive growth conditions on the reactions, existence has been obtained in [24] based on the notion of renormalised solutions [22]. Our weak-strong uniqueness principle covers models involving various cross-diffusion phenomena (Soret effect, Dufour effect, cross diffusion between species), and applies in particular to a class of isoenergetic reaction-cross-diffusion systems, thus generalising [12]. Weak-strong uniqueness is obtained from a weak stability estimate for a generalised distance involving as in [23] an adjusted relative entropy. By suitably exploiting the thermodynamic structure of the system some of the technical issues arising in the proof of [12] will be avoided. We also obtain an asymptotic stability result.

1.1. **Thermodynamic modelling.**

Let us now briefly specify the thermodynamic structure considered in the present manuscript. For more background on the modelling, we refer to [24, 39, 28, 36]. Models compatible with thermodynamics can be derived using the Onsager formalism in [36]. Here, the main ingredient is a triple $(Z, H, \mathbb{K})$ consisting of a state space Z, a driving functional H, and a so-called Onsager operator $\mathbb{K}$. Typically, $Z \subset X$ is a convex subset of a Banach space X, $H : Z \to \mathbb{R} \cup \{\infty\}$ a differentiable and convex functional on Z (below usually referred to as entropy due to its correspondence to the negative of the physical entropy), while $\mathbb{K}$ can be seen as a generalised inverse Riemannian metric tensor on Z. More specifically, for every $z \in Z$, $\mathbb{K}(z)$ defines a positive semi-definite and symmetric (unbounded) operator from $T_z^* Z$ to $T_z Z$. If $\mathbb{G} = \mathbb{K}^{-1}$ exists, the triple $(Z, H, \mathbb{G})$ forms a gradient system. Then, motivated by the classical gradient flow equation $\mathbb{G}(z)\dot{z} = -DH(z)$, with $\dot{z}$ denoting the time derivative of $z = z(t)$, one considers the evolution law

$$\dot{z} = -\mathbb{K}(z)DH(z),$$

where here DH denotes the Fréchet derivative of the functional H. An advantage of this Onsager form is that it facilitates the consistent coupling of different physical phenomena, which can be realised by an additive decomposition of $\mathbb{K}$ [36]. Observe that, formally, the above structure encodes the following core entropy dissipation property

$$\frac{\mathrm{d}}{\mathrm{d}t} H(z) = -\langle DH(z), \mathbb{K}(z)DH(z) \rangle \leq 0 \tag{1.2}$$

along any solution curve $z = z(t)$ of the above law. Conservation of the total energy $E(z)$, with E denoting the energy functional on Z, can be guaranteed by imposing the condition $\mathbb{K}DE = \mathbb{K}^* DE \stackrel{!}{=} 0$, which implies that $\frac{\mathrm{d}}{\mathrm{d}t} E(z) = -\langle DE(z), \mathbb{K}(z)DH(z) \rangle = 0$.

In the context of ERDS, we consider, as introduced above, $z = (u, c)$ with u the internal energy density and $c = (c_1, \ldots, c_n)$ the vector of concentrations. We focus on entropies of the form

$$H(z) = \int_\Omega h(z) \, \mathrm{d}x$$

with densities

$$h(z) = h(u, c) = -\sigma(u) + \sum_{i=1}^n b(c_i, w_i(u)),$$

composed of a thermal part $\sigma(u)$ and a relative Boltzmann entropy $b(s, e) = e\lambda(s/e)$, where

$$\lambda(r) := r \log(r) - r + 1, \tag{1.3}$$

and with $w_i = w_i(u)$, $i = 1, \ldots, n$, denoting the thermodynamic equilibria of the concentrations c_i. The dependence of w_i on the internal energy density u results in a strong coupling of the system and

is one of the main sources of difficulties in the analysis. It will be convenient to introduce the function $\hat{\sigma}(u) = \sigma(u) - \sum_{i=1}^{n} w_i(u) + n$ and rewrite $h(u, c)$ in the following more explicit form

$$h(u, c) = -\hat{\sigma}(u) + \sum_{i=1}^{n} \Big(\lambda(c_i) - c_i \log w_i(u) \Big). \tag{h1}$$

To simplify the exposition, we will impose the following concrete conditions on the coefficient functions (cf. [24]):

$$\left. \begin{array}{l} \hat{\sigma} \in C^2((0, \infty)) \text{ strictly concave \& non-decreasing.} \\ w_i \in C([0, \infty)) \cap C^2((0, \infty)) \text{ concave \& non-decr. with } w_i(0) > 0 \text{ for all } i. \\ \lim_{u \downarrow 0} \hat{\sigma}'(u) = +\infty, \lim_{u \uparrow \infty} \hat{\sigma}'(u) = 0; \ \ \sup_{u \in (0,1]} \hat{\sigma}''(u) < 0. \\ \exists \beta \in (0, 1) \text{ such that } w_i(u) \lesssim 1 + u^\beta \text{ for all } i \in \{1, \ldots, n\}. \end{array} \right\} \tag{h2}$$

Typical choices are $\hat{\sigma}(u) = a \log(u)$ or $\hat{\sigma}(u) = au^\nu$ for some $\nu \in (0, 1)$, $a > 0$, and $w_i(u) = (b_{i,1} u + b_{i,0})^{\beta_i}$ or $w_i(u) = b_{i,1} u^{\beta_i} + b_{i,0}$ for $\beta_i \in (0, 1)$, $b_{i,0} > 0$, $b_{i,1} \geq 0$.

Our weak-strong uniqueness principle does not rely, in an essential way, on this specific form of the entropy density.[1] In fact, in the proof of our main theorem (Thm 2.8), identity (h1) is only used to guarantee the coercivity properties in Proposition 3.2. The crucial point in the assumptions (h2) on the coefficient functions is that they ensure good convexity properties, and more specifically the locally uniform positive definiteness of the Hessian $D^2 h$, which is essential for estimate (3.5) in Prop. 3.2. The monotonicity assumptions on $\hat{\sigma}$ and w_i are relevant from the modelling point of view, since they ensure that $u \mapsto h(u, c)$ is non-increasing, so that temperature, which is given by $-\frac{1}{\partial_u h}$, is non-negative.

As in [39, 24] we are primarily interested in Onsager operators $\mathbb{K}$ of the form

$$\mathbb{K}(z)\zeta = \mathbb{K}_{\text{diff}}(z)\zeta + \mathbb{K}_{\text{react}}(z)\zeta = -\nabla \cdot (\mathbb{M}(z)\nabla\zeta) + \mathbb{L}(z)\zeta,$$

where $\mathbb{M}(z), \mathbb{L}(z) \in \mathbb{R}^{(1+n) \times (1+n)}$ are positive semi-definite symmetric matrices and where $\nabla = \nabla_x$ is the gradient with respect to $x \in \Omega$. We will complement $\mathbb{K}$ with the no-flux boundary conditions $\mathbb{M}\nabla\zeta \cdot \nu = 0$ on $\partial\Omega$, where ν denotes the outer unit normal vector to $\partial\Omega$. Observing that $E(u, c) = \int_\Omega u$ describes the total (internal) energy, the condition $\mathbb{K}DE \equiv 0$, ensuring energy conservation, means that $\ker \mathbb{L}(z) \supseteq \text{span}\{(1, 0)^T\}$. Thus, by the symmetry of $\mathbb{L}$, the zeroth component R_0 of $R(z) := -\mathbb{L}(z)Dh(z)$ vanishes. Moreover, positive semi-definiteness of $\mathbb{L}$ implies the inequality

$$D_i h(z) R_i(z) \leq 0. \tag{1.4}$$

In this paper, the specific form of $\mathbb{L}(z)$ will not be relevant. Instead, we directly work with reactions $R(z)$ of the form

$$R(z) := (0, R_1(z), \ldots, R_n(z))$$

satisfying (1.4).

With $\mathbb{K}$ as above, the equation $\dot{z} = -\mathbb{K}(z)DH(z)$ can be written in the form (1.1) by choosing

$$A(z) := \mathbb{M}(z)D^2 h(z).$$

In short,

$$\begin{aligned} \partial_t z &= \nabla \cdot \big(A(z)\nabla z \big) + R(z), & t > 0, x \in \Omega, \\ 0 &= A(z)\nabla z \cdot \nu, & t > 0, x \in \partial\Omega, \end{aligned} \tag{ERDS}$$

subject to an initial condition $z_{|t=0} = z^{\text{in}}$. In the above setting, the entropy dissipation property (1.2) takes the form

$$\frac{\text{d}}{\text{d}t} H(z) + \int_\Omega \mathscr{P}(z)\,\text{d}x = \int_\Omega D_i h(z) R_i(z)\,\text{d}x \leq 0,$$

---

[1]See Sec. 2.3.2 for an example of a different entropy density that our technique can be adapted to.

where $\mathscr{P}(z) := \nabla D_i h(z) \cdot \mathbb{M}_{il}(z) \nabla D_l h(z) \geq 0$, by the positive semi-definiteness of the mobility matrix $\mathbb{M}$, which will be assumed throughout. Supposing, for instance, that $\mathscr{P}(z) \gtrsim \sum_{i=1}^{n} |\nabla \sqrt{c_i}|^2$ (as it can be proved for many of the models considered in [24], see Section 2.3 and Lemma 6.1), and using conservation of $\int_{\Omega} u$ together with suitable bounds on $H(z)$ (cf. Lemma 6.3), the entropy dissipation property provides a priori control of $\sum_{i=1}^{n} \|\nabla \sqrt{c_i}\|_{L_{t,x}^2}^2$. Let us further note that the fact that $R_0(z)$ vanishes gives to some extent a scalar-like structure to the u-component of (ERDS), and if, for instance, $\mathbb{M}$ is chosen such that $A_{0j} = a\,\delta_{0j}$ for some function $a = a(z) \geq 0$, any L^p-type energy $\frac{1}{p}\|u(t)\|_{L^p}^p$, $p \in (1, \infty)$, is formally non-increasing in time.

1.2. Motivation and strategy.

Being able, for a given PDE, to identify concepts of solutions for which both existence and uniqueness can be established is a fundamental concern in modelling and analysis. For *scalar* equations, there are various tools to identify frameworks allowing for the existence of a *unique* solution, even in regimes of low regularity and with strong nonlinearities. One approach is based on the 'doubling variables' technique first employed by Kružkov [32] to entropy solutions of first-order equations, and extended by Carrillo [6] to hyperbolic-parabolic-elliptic equations. The concept was adapted to situations where L^∞ bounds are not available to give uniqueness in a class of renormalised solutions [3, 7]. See also [2, 40, 1] for more recent developments. Let us also mention the Young measure approach to conservation laws going back to Tartar [43] and DiPerna [19] who obtained, in the scalar case, uniqueness of solutions obeying an entropy inequality. For second order parabolic and elliptic equations the viscosity solution technique and associated comparison principles [30, 14] are powerful tools and the key to a variety of wellposedness results in geometric, highly nonlinear or degenerate settings, see e.g. [20, 13, 5, 8]. Some extensions of the viscosity solution approach to systems are available for weakly coupled problems with a monotonicity condition [29]. Further uniqueness results applying to specific systems and typically in more regular situations include [25, 31, 10, 41, 26, 4].

In general, the case of *strongly coupled* (parabolic) *systems* tends to be much more difficult. While under a parabolicity condition the existence of suitable generalised solutions (here referred to as 'weak' solutions) can often be established, positive uniqueness results in such general settings are rare. It is therefore common, to relax the quest for uniqueness to the problem of whether weak solutions are uniquely determined in situations where a sufficiently regular solution (a 'strong' solution) happens to exist. In other words, one is interested in the question of whether such strong solutions are unique in a potentially much larger class of weak solutions. The question of weak-strong uniqueness is classical in fluid dynamics problems and goes back to Leray's fundamental work [33], where it was established for the incompressible Navier–Stokes equations. We refer to the survey by Wiedemann [45] for more details and further references. For recent advances on conditional uniqueness results for dissipative measure-valued solutions to conservation laws, see [27] and references therein. Quite interesting in the thermodynamics context is moreover the relative entropy technique employed by DiPerna [18] and Dafermos [16] for hyperbolic conservation laws.

Relative entropy methods are nowadays a standard tool to study weak stability properties of nonlinear systems endowed with a (convex) entropy structure. Generally speaking (using the notation introduced in Sec. 1.1), a relative entropy of the form

$$H_{\mathrm{rel}}(z, \tilde{z}) = H(z) - \int_{\Omega} D_i h(\tilde{z})(z_i - \tilde{z}_i)\,\mathrm{d}x - H(\tilde{z}) \tag{1.5}$$

is used to measure the distance between a weak solution z and a strong solution $\tilde{z}$. Observe that for convex entropies $H(z)$, the map $z \mapsto H_{\mathrm{rel}}(z, \tilde{z})$ is a non-negative, convex functional vanishing in $z = \tilde{z}$. The thermodynamic structure ensures that regular solutions automatically satisfy an entropy dissipation balance (cf. eq. (1.2)). Physically relevant processes may, however, in general possess less regularity and here an entropy inequality is often added as an admissibility criterion for weak solutions.

The goal is then to obtain an upper bound on the time evolution of the relative entropy that implies stability of a regular flow on a finite time horizon among generalised solutions. This means that given a regular flow, any weak solution that is initially close (in relative entropy) will remain close for some time.

In the present paper, we pursue such a strategy in the context of ERDS. Weak-strong uniqueness has recently been obtained for entropy-dissipating reaction-diffusion systems with a uniformly elliptic and bounded diagonal diffusion matrix [23], where the main difficulty consists in a lack of control of the reaction rates. Extensions to a cross-diffusion system from population dynamics with weak cross diffusion can be found in [12]. Both references are based on the relative entropy method, but their arguments rely on the specific structure of the diffusion matrix of their systems. Here, we would like to present a more general strategy to deduce stability from an underlying thermodynamic structure.

Given the strong coupling and lack of a priori bounds in L^∞, there are several difficulties in our ERDS that require an adaptation of the classical relative entropy approach to weak-strong stability. First, due to the lack of growth restrictions on the reactions and in some cases even the flux term (see the models in [24]), the evolution of the classical relative entropy used to measure the distance between a renormalised solution z and a strong solution $\tilde{z}$, cannot be properly controlled. This is due to the term

$$- \int_\Omega D_i h(\tilde{z}) \frac{\mathrm{d}}{\mathrm{d}t} z_i \, \mathrm{d}x = \int_\Omega \nabla D_i h(\tilde{z}) \cdot (A_{ij}(z) \nabla z_j) \, \mathrm{d}x - \int_\Omega D_i h(\tilde{z}) R_i(z) \, \mathrm{d}x \qquad (1.6)$$

arising in the formal computation of the time derivative of $H_{\mathrm{rel}}(z, \tilde{z})$. In fact, the available a priori estimates do not ensure that $A_{ij}(z) \nabla z_j \in L^1(\Omega)$ and $R_i(z) \in L^1(\Omega)$ for a.e. time. At the same time, the corresponding integrands in (1.6) do not have a sign, and there is no hope for the uncontrolled parts to cancel with some of the remaining terms appearing in $\frac{\mathrm{d}}{\mathrm{d}t} H_{\mathrm{rel}}(z, \tilde{z})$. It is therefore necessary to adjust the relative entropy $H_{\mathrm{rel}}(z, \tilde{z})$. This issue has already been encountered in [23]; it can be resolved by introducing a suitable smooth and compactly supported truncation function $\xi^* = \xi^*(z)$ with $\xi^*(z) = 1$ if $\sum_{i=0}^n z_i \leq E$ for some $E \gg \sum_{i=0}^n \tilde{z}_i$ (see Section 3.1 for details) in the formula for the relative entropy via

$$H_{\mathrm{rel}}^*(z, \tilde{z}) := H(z) - \int_\Omega D_i h(\tilde{z})(\xi^*(z) z_i - \tilde{z}_i) \, \mathrm{d}x - H(\tilde{z}).$$

The relative entropy density adjusted in this fashion allows to remove the issue pointed out above. (Strictly speaking, in the term $D_0 h(\tilde{z})(z_0 - \tilde{z}_0)$ the truncation function is not needed in the models considered in this paper, and for other applications it may be helpful to use a different choice such as $D_i h(\tilde{z})(\xi_i^*(z) z_i - \tilde{z}_i)$ with $\xi_i^*(z) \equiv 1$ for $i = 0$, or versions thereof.)

A second difficulty arising in the case of ERDS is the inherent coupling between concentrations and energy density, which manifests itself in the circumstance that the entropy density cannot be additively decomposed into terms depending only on an individual component z_i. This in turn leads to a non-diagonal diffusion matrix $A(z)$ and renders estimating the evolution of $H_{\mathrm{rel}}^*(z, \tilde{z})$ substantially more delicate than in the diagonal case. One of the main contributions of this manuscript is to show that such estimates can be achieved, with relatively little technical effort, by carefully exploiting the entropy structure.

The energy component u plays a distinguished role in ERDS that has to be taken advantage of when interested in a general analysis. At a technical level, the physical constraint of the convex function $h(u, c)$ being non-increasing in u (to ensure a non-negative temperature) restricts the range of relevant functions $\hat{\sigma}$ to sublinearities such as $\hat{\sigma}(u) = u^\nu$ for some $\nu \in [0, 1)$ (with $\nu = 0$ corresponding to $\log$). Unless $\hat{\sigma}(u)$ has close to linear growth for large values of u, even the possibility of an existence theory solely based on the entropy estimate is questionable in general dimensions. We are interested in covering more degenerate choices of $\hat{\sigma}$, and therefore cannot purely rely on the (adjusted) relative entropy to measure the distance of a weak to the strong solution. Instead, we exploit the absence of

source terms in the u-component of the evolution system, which allows to give an extra, scalar-like structure to the evolution law for u. Here, we content ourselves with the arguably simplest choice of an L^2-structure, meaning that we consider weighted generalised distances of the form

$$\mathrm{Dist}^*_\alpha(z, \tilde{z}) = H^*_{\mathrm{rel}}(z, \tilde{z}) + \tfrac{\alpha}{2}\|u - \tilde{u}\|^2_{L^2(\Omega)}, \quad \alpha \in (0, \infty).$$

This is consistent with the approach in [24] and allows us in particular to show the weak-strong uniqueness property for the potentially pathological solutions constructed in [24, Theorem 1.8], where a cross-diffusion dominant regime was considered with gradients of the internal energy density inducing a (possibly) non-integrable concentration flux. Furthermore, by exploiting the existence of such an additional quantity that up to some error term is dissipated along the flow, we can relax the conditions on the entropy functional in [39] required for proving exponential convergence to equilibrium.

1.3. **Technique.** Here, we briefly outline, at a formal level, the main points of our argument showing a weak-strong stability estimate of the form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{Dist}^*_\alpha(z, \tilde{z}) \lesssim_{T,\alpha,\xi^*} \mathrm{Dist}^*_\alpha(z, \tilde{z}), \tag{1.7}$$

on any finite time horizon $(0, T)$, $T < T^*$, where z is assumed to be a 'weak' (renormalised) solution and $\tilde{z}$ a 'strong' solution of (ERDS) in $(0, T^*) \times \Omega$ for some $T^* \in (0, \infty]$.

First, letting $\mathrm{dist}^*_\alpha(z, \tilde{z}) = h^*_{\mathrm{rel}}(z, \tilde{z}) + \tfrac{\alpha}{2}|u - \tilde{u}|^2$, where

$$h^*_{\mathrm{rel}}(z, \tilde{z}) = h(z) - D_i h(\tilde{z})(\xi^*(z)z_i - \tilde{z}_i) - h(\tilde{z}), \tag{1.8}$$

we can write $\mathrm{Dist}^*_\alpha(z, \tilde{z}) = \int_\Omega \mathrm{dist}^*_\alpha(z, \tilde{z})\,\mathrm{d}x$. We further recall that the function $\xi^* = \xi^*(z)$ will be chosen such that $\xi^*(z) = 1$ if $\sum_{i=0}^n z_i \leq E$ for an auxiliary parameter $E \gg \sum_{i=0}^n \tilde{z}_i$. Then, if $E = E(\tilde{z}, \min\{\alpha, 1\})$ is chosen large enough, $\mathrm{dist}^*_\alpha(z, \tilde{z}) \geq 0$ for all $z \in [0, \infty)^{1+n}$ with equality if and only if $z = \tilde{z}$.

To sketch the argument leading to (1.7), let us for simplicity only consider the case where $A_{0j}(z) = a(z)\delta_{0j}$ with $a \gtrsim 1$. In this case, it will suffice to take $\alpha \geq 1$. We now assume that z and $\tilde{z}$ are sufficiently regular solutions of (ERDS) (with $A = \mathbb{M}D^2 h$, $\mathbb{M} \geq 0$, $D_i h R_i \leq 0$), where the strong solution $\tilde{z}$ be such that $\|\tilde{z}\|_{C^{0,1}([0,T]\times\bar{\Omega})} < \infty$ and $\inf_{(0,T)\times\Omega} \tilde{z}_i > 0$ for all $i \in \{0, \ldots, n\}$ and all $T < T^*$. To estimate the time evolution of $\mathrm{Dist}^*_\alpha(z, \tilde{z})$, one formally computes

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{Dist}^*_\alpha(z, \tilde{z}) = \int_\Omega \rho^{(h)}\,\mathrm{d}x + \alpha\int_\Omega \rho^{(g)}\,\mathrm{d}x,$$

where (see Lemma 3.4)

$$\begin{aligned}
\rho^{(h)} := {} & -\nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
& + \nabla(D_i(\xi^*(z)z_j)D_j h(\tilde{z})) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
& + \nabla\Big(D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)\Big) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z}) \\
& - D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)R_i(\tilde{z}) \\
& + (D_i h(z) - D_j h(\tilde{z})D_i(\xi^*(z)z_j))R_i(z),
\end{aligned} \tag{1.9}$$

and

$$\begin{aligned}
\rho^{(g)} := {} & -a(z)|\nabla u|^2 - a(\tilde{z})|\nabla\tilde{u}|^2 + a(z)\nabla u \cdot \nabla\tilde{u} + a(\tilde{z})\nabla u \cdot \nabla\tilde{u} \\
= {} & -a(z)|\nabla u - \nabla\tilde{u}|^2 - (a(z) - a(\tilde{z}))(\nabla u - \nabla\tilde{u}) \cdot \nabla\tilde{u}.
\end{aligned} \tag{1.10}$$

Thus, to show (1.7) it suffices to obtain a pointwise upper bound of the form

$$\rho_\alpha := \rho^{(h)} + \alpha\rho^{(g)} \lesssim \mathrm{dist}^*_\alpha(z, \tilde{z}). \tag{1.11}$$

This pointwise estimate will be proved by distinguishing four cases determined by the value of the weak solution $z = z(t, x) \in [0, \infty)^{1+n}$ at any given point (t, x). This case distinction is motivated by the following observations:

First, if $z \in [0, \infty)^{1+n}$ with $\sum_{i=0}^{n} z_i \leq E$ for $E = E(\tilde{z})$ large enough, we want $h_{\mathrm{rel}}^*(z, \tilde{z})$ to coincide with the classical relative entropy density $h_{\mathrm{rel}}(z, \tilde{z}) = h(z) - D_i h(\tilde{z})(z_i - \tilde{z}_i) - h(\tilde{z})$ to be able to use its distance-like properties. This will be ensured by choosing $\xi^*(z) = 1$ with $D^k \xi^*(z) = 0$ for all $k \in \mathbb{N}^+$ whenever $\sum_{i=0}^{n} z_i \leq E$ (cf. the definition of ξ^* in Sect. 3.1). Thus, if $|z - \tilde{z}|$ is close to zero, the strict convexity, non-negativity and vanishing in $z = \tilde{z}$ of $\mathrm{dist}_\alpha^*(\cdot, \tilde{z})$ imply that $\mathrm{dist}_\alpha^*(z, \tilde{z}) \sim_{\|\tilde{z}\|_{L^\infty}, E} |z - \tilde{z}|^2$ for $|z| \leq E$. In this case, to show that ρ_α is quadratically small in $|z - \tilde{z}|$, we write (using $\sum_i z_i \leq E$)

$$
\begin{aligned}
\rho^{(h)} = & -\nabla(D_i h(z) - D_i h(\tilde{z})) \cdot \mathbb{M}_{il}(z) \nabla(D_l h(z) - D_l h(\tilde{z})) \\
& - \nabla(D_i h(z) - D_i h(\tilde{z})) \cdot (\mathbb{M}_{il}(z) - \mathbb{M}_{il}(\tilde{z})) \nabla D_l h(\tilde{z}) \\
& - \nabla\big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j)\big) \cdot \mathbb{M}_{il}(\tilde{z}) \nabla D_l h(\tilde{z}) \\
& + \big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j)\big) R_i(\tilde{z}) \\
& + (D_i h(z) - D_i h(\tilde{z}))(R_i(z) - R_i(\tilde{z})),
\end{aligned}
$$

see case $\mathcal{A}_+$ in the proof of Theorem 2.8 for details. In order to deal with the terms involving a gradient of z that appear in the second and the third term on the RHS, one would like to exploit the non-positive first term on the RHS. Typically (such as in the ERDS models considered in [24]), the submatrix $(\mathbb{M}_{il}(z))_{i,l=1,\ldots,n}$ will, however, degenerate as soon as $c_i \searrow 0$ for some $i \in \{1, \ldots, n\}$. Yet if $\min\{z_1, \ldots, z_n\} \geq \iota$ for some $\iota > 0$, then it is possible to assume that $(\mathbb{M}_{il}(z))_{i,l=1,\ldots,n} \gtrsim_\iota \mathbb{I}_n$. This, combined with the second line in (1.10) and suitable smoothness assumptions on $\mathbb{M}$ and h, will allow us to infer that

$$
\rho_\alpha \lesssim_{E,\iota,\tilde{z}} |z - \tilde{z}|^2
$$

whenever $z(t, x) \in \mathcal{A}_+ := \{z' \in [0, \infty)^{1+n} : \min\{z_0', \ldots, z_n'\} \geq \iota, \ \sum_{i=0}^{n} z_i \leq E\}$ for some $\iota > 0$ and sufficiently large $E \geq 1$.

To deal with the case $z(t, x) \in \mathcal{A}_0 := \{z' : \min\{z_0', \ldots, z_n'\} < \iota, \ \sum_{i=0}^{n} z_i \leq E\}$, we fix $\iota = \iota(\tilde{z}) > 0$ small enough such that $\inf \tilde{z}_i \geq 2\iota$ for all $i = 0, \ldots, n$. This implies that $|z - \tilde{z}| \geq \iota$ whenever $z \in \mathcal{A}_0$. Thus, since $|z - \tilde{z}| \gtrsim 1$ is bounded away from zero, so is $\mathrm{dist}_\alpha^*(z, \tilde{z})$ (see Prop. 3.2). It then suffices to have suitable coercivity estimates on $\mathscr{P}(z)$ that allow to absorb those terms on the RHS of (1.9) that involve gradients of z and do not have a sign by the first term on the RHS, which equals $-\mathscr{P}(z)$. (Such coercivity estimates are typically already needed in the construction of solutions.)

It remains to consider the case $\sum_i z_i > E$, where E will be chosen large enough, in particular such that $E \geq \|\sum_i \tilde{z}_i\|_{L^\infty((0,T)\times\Omega)} + 1$ and $E \geq E_0$ with E_0 being such that $\mathrm{dist}_\alpha^*(z, \tilde{z}) \gtrsim \sum_i c_i \log_+(c_i) + u^2 + 1$ for all $\tilde{E} \geq E_0$. If $z \notin \mathrm{supp}\, \xi^*$, $\rho^{(h)}$ takes a simple form. The entropy dissipating property of diffusion and reactions, and the Lipschitz regularity of the strong solution $\tilde{z}$ are sufficient to deduce (1.11) in this case (referred to as $z \in \mathcal{C}$).

The intermediate case (below referred to as case $z \in \mathcal{B}$), where $0 < \xi^* < 1$, is somewhat more delicate as can be seen in formula (1.9), where the second term on the RHS involves products of gradients of the weak solution. In order to be able to absorb this term by the first term on the RHS of (1.9), another scale $E' \gg E$ is introduced (for convenience we choose $E' = E^N$, as in [23]), and ξ^* will be taken such that $\xi^*(z) = 0$ if and only if $\sum_{i=0}^{n} z_i \geq E'$, $\xi^*(z) = 1$ if and only if $\sum_{i=0}^{n} z_i \leq E$. On these scales, derivatives of ξ^* typically have an additional decay property enabling the desired absorption if $\frac{E'}{E}$ is large enough.

1.4. **Outline.** The rest of the article is structured as follows. In Section 2 we introduce relevant definitions and hypotheses, and formulate our main results: a weak-strong uniqueness principle for renormalised solutions to (ERDS) (see Thm 2.8), a strong entropy dissipation inequality as used in the proof of Theorem 2.8 (see Prop. 2.9), and a result on the exponential convergence to equilibrium (see Prop. 2.10). In Section 2.3 we present selected examples that our main results apply to, including the class of ERDS considered in [24] as well as a class of models with cross diffusion between species. The weak-strong uniqueness principle is proved in Section 3, starting with several auxiliary results with the actual proof of Theorem 2.8 being given in Section 3.4. The proofs of the entropy dissipation inequality and the exponential convergence to equilibrium are given in Sections 4 and 5 respectively. Some auxiliary results are gathered in Appendix 6.

1.5. **Notations.**

- *Summation convention*: any unspecified summations of the form $\sum_i$ are to be understood as $\sum_{i=0}^n$. For brevity, we use a summation convention for summing over the system's components $i = 0, \ldots, n$ in case of repeatedly occurring indices while omitting the summation symbol. In ambiguous situations the summation symbol will be used. Summations restricted to $i = 1, \ldots, n$ (excluding the u-component) will always be made explicit. In our convention, summation over repeated indices has priority over other mathematical operations such as integration or taking the absolute value. For instance, by default we let $|A_{ik}(z)\nabla z_k| = |\sum_{k=0}^n A_{ik}(z)\nabla z_k|$.
- For technical concerns regarding the notation $\mathbb{M}_{il}(z)\nabla D_l h(z)$, we refer to Remark 2.3.
- We denote by $R = (0, R_1, \ldots, R_n)^T$ the vectors of reaction rates.
- Given $T^* \in (0, \infty]$, we let $I = [0, T^*)$ denote the time horizon of interest. For $T > 0$, we abbreviate $\Omega_T := (0, T) \times \Omega$.
- For functions $f = f(z_0, \ldots, z_n)$ we let $D_i f = \frac{\partial f}{\partial z_i}$ and $D_{ij} f = \frac{\partial^2 f}{\partial z_i \partial z_j}$ for $i, j \in \{0, \ldots, n\}$.
- In estimates, $C < \infty$ typically denotes a finite (sufficiently large) constant that may change from line to line, while we often use $\epsilon > 0$ to denote a (sufficiently small) positive constant.
- For quantities $A, B \geq 0$ we write $A \lesssim B$ if there exists a fixed constant $C < \infty$ such that $A \leq CB$. The notation $A \gtrsim B$ means $B \lesssim A$, while $A \sim B$ is to be understood as both $A \lesssim B$ and $A \gtrsim B$ being satisfied. In order to indicate dependencies of the constant $C = C(p_1, \ldots, p_k)$ on certain parameters $p_1, \ldots, p_k$, we write $A \lesssim_{p_1, \ldots, p_k} B$, and analogously for $\gtrsim$ and $\sim$.
- Any dependence of constants and estimates on the regular solution $\tilde{z} \in C^{0,1}$ will usually not be explicitly indicated.
- We let $\min(z) := \min\{z_0, z_1, \ldots, z_n\}$ for $z = (z_0, \ldots, z_n) \in [0, \infty)^{1+n}$.
- By default, $|\cdot|$ denotes the Euclidean norm, e.g. $|z| = (\sum_i |z_i|^2)^{\frac{1}{2}}$
- $|z|_1 = \sum_{i=0}^n z_i$ and $|c|_1 = \sum_{i=1}^n c_i$.
- For time-dependent integral functionals $\int_\Omega f((t,x), z(t,x)) \, \mathrm{d}x$, where $z = z(t,x)$ denotes a 'weak' solution of (ERDS) taking in a suitable sense the data z^{in}, we use the convention

$$\int_\Omega f((t,x), z(t,x)) \, \mathrm{d}x \Big|_{t=0}^{t=T} := \int_\Omega f((T,x), z(T,x)) \, \mathrm{d}x - \int_\Omega f((0,x), z^{\mathrm{in}}(x)) \, \mathrm{d}x$$

provided the terms on the RHS are well-defined.
- For an open set $U \subset \mathbb{R}^N$, $C^k(U)$ denotes the space of continuous functions on U that are continuously differentiable up to order $k \in \mathbb{N}$. By $C^{k,\nu}(U) = C^{k,\nu}_{\mathrm{loc}}(U)$, we denote the space of functions in $C^k(U)$, whose k-th derivative is ν-Hölder continuous for some $\nu \in (0, 1]$ on compact subsets $K \subseteq U$. (We use the symbol $C^{k,\nu}_{\mathrm{loc}}(U)$ for clarity's sake.)

2. MAIN RESULTS

2.1. **Assumptions.** Throughout these notes, we let $d \geq 1$ and $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with $|\Omega| = 1$. We further let $T^* \in (0, \infty]$ and $I = [0, T^*)$.

To prepare for stating our main result, the weak-strong uniqueness principle (Theorem 2.8), we gather the following conditions.

(A1) Entropy: $h \in C^4((0, \infty)^{1+n})$ is of the form (h1), where $\hat{\sigma}$ and w_i are supposed to satisfy (h2).

(A2) Reactions: $R = (0, R_1, \ldots, R_n) \in C([0, \infty)^{1+n})^{1+n}$
    (i) satisfy $\sum_{i=1}^{n} D_i h(z) R_i(z) \leq 0$ in $(0, \infty)^{1+n}$
    (ii) are locally Lipschitz continuous in $(0, \infty)^{1+n}$.

(A3) Mobility matrix: $\mathbb{M} \in \left(C_{\mathrm{loc}}^{0,1}((0, \infty)^{1+n}) \cap C([0, \infty)^{1+n}) \right)^{(1+n) \times (1+n)}$ and there exist non-negative functions $m, a \in C_{\mathrm{loc}}^{0,1}((0, \infty)^{1+n})$ and $\varpi \in \{0, 1\}$ with $0 \leq m \lesssim \varpi$ and $a \gtrsim 1$ such that

$$\mathbb{M}_{0l} = \widetilde{\mathbb{M}}_{0l} + \delta_{0l} m \quad \text{for } l = 0, \ldots, n, \tag{A3.a}$$

for suitable $\widetilde{\mathbb{M}}_{0l}$ satisfying $\sum_{l=0}^{n} \widetilde{\mathbb{M}}_{0l} D_{lj} h = \delta_{0j} a$.
    Moreover, for all $z \in [0, \infty)^{1+n}$ with $\min_i z_i \geq \iota$ for some $\iota > 0$ there exists $\epsilon(\iota) > 0$ such that

$$\mathbb{M}(z) \geq \mathrm{diag}(m(z), 0, \ldots, 0) + \epsilon(\iota) \, \mathrm{diag}(0, 1, \ldots, 1). \tag{A3.b}$$

By continuity, when $\iota = 0$, (A3.b) holds true with $\epsilon(\iota) = 0$.

Using our standard notation $A(z) = \mathbb{M}(z) D^2 h(z)$, hp. (A3.a) implies that, formally,

$$\sum_{j=0}^{n} A_{0j}(z) \nabla z_j = \sum_{l=0}^{n} \mathbb{M}_{0l}(z) \nabla D_l h(z) = a(z) \nabla u + m(z) \nabla D_0 h(z). \tag{A3.c}$$

We further need certain bounds on the flux and the concentration gradients in terms of the entropy dissipation. For this purpose we define for non-negative functions $z_j \in L_{\mathrm{loc}}^1(I, L^1(\Omega))$ such that $\nabla(z_j^s) \in L_{\mathrm{loc}}^2(I; L^2(\Omega))$ for some $s \in \{\frac{1}{2}, 1\}$ for each $j \in \{0, \ldots, n\}$, the quantity

$$\begin{aligned}
\mathscr{P}(z) &:= \nabla z : (D^2 h(z) A(z) \nabla z) \\
&= \nabla D_i h(z) \cdot (\mathbb{M}_{il}(z) \nabla D_l h(z)),
\end{aligned}$$

where the second equality is to be understood in a formal sense, see Remark 2.3. By the positive semi-definiteness of $\mathbb{M}$ imposed by hp. (A3.b), we have $\mathscr{P}(z) \geq 0$ for any such z, and more specifically, $\mathscr{P}(z) \geq m(z) |\nabla D_0 h(z)|^2$.

(A4) For all $K \geq 1$

$$\chi_{\{|z| \leq K\}} |\nabla c| \lesssim_K \sqrt{\mathscr{P}(z)}, \tag{A4.a}$$

$$\chi_{\{|z| \leq K\}} \left| \sum_{j=0}^{n} A_{ij}(z) \nabla z_j \right| \lesssim_K \sqrt{\mathscr{P}(z)} \tag{A4.b}$$

    for all $i = 0, \ldots, n$.

(A5) For $a(z)$ as in (A3),

$$|a(z) \nabla u| \lesssim (1 + u) \sqrt{\mathscr{P}(z)}.$$

Additionally, we often impose the following bound:

(A6) For all $0 \leq i \leq n$

$$\chi_{\{|z|_1 \geq 1\}} |\nabla z| \left| \sum_{j=0}^{n} A_{ij}(z) \nabla z_j \right| \lesssim |z| \mathscr{P}(z) + |z| |\nabla u|^2.$$

If (A6) is not satisfied, we have to assume that $\varpi = 0$ in (A3) together with the condition:

(A6') For all $0 \leq i \leq n$ and all $\underline{u} \in (0, 1]$

$$\chi_{\{u \geq \underline{u}\}} \chi_{\{|z|_1 \geq 1\}} |\nabla z| \left| \sum_{j=0}^{n} A_{ij}(z) \nabla z_j \right| \lesssim |z| \mathscr{P}(z) + C(|z|, \underline{u}) |\nabla u|^2.$$

Let us observe that (A6) $\Longrightarrow$ (A6').

A selection of relevant examples fulfilling the above hypotheses is provided in Section 2.3.

2.2. **Definitions and Results.** Throughout this text, we write $A(z) := \mathbb{M}(z) D^2 h(z)$, where h takes the form (h1), (h2). We further recall our summation convention (see Notations 1.5).

Definition 2.1 (Renormalised solutions). *Let $I = [0, T^*)$ and suppose that the vector-valued function $z = (u, c_1, \ldots, c_n)$ has non-negative components $z_i \geq 0$ satisfying $\sqrt{z_i} \in L^2_{\mathrm{loc}}(I; H^1(\Omega))$ or $z_i \in L^2_{\mathrm{loc}}(I; H^1(\Omega))$ for all $i = 0, \ldots, n$. Further suppose that for all $E \geq 1$*

$$\chi_{\{|z| \leq E\}} A_{ik}(z) \nabla z_k \in L^2_{\mathrm{loc}}(I; L^2(\Omega)),$$

for every $i \in \{0, \ldots, n\}$.

We call such z a renormalised solution of the energy-reaction-diffusion system (ERDS) *in $\Omega_{T^*} := (0, T^*) \times \Omega$ with initial data z^{in} if for all $\xi \in C^\infty(\mathbb{R}_{\geq 0}^{1+n})$ with compactly supported derivative $D\xi$, all $\psi \in C^\infty(I \times \bar{\Omega})$ and almost all $T \in (0, T^*)$*

$$\int_\Omega \xi(z(T, \cdot)) \psi(T, \cdot) \, \mathrm{d}x - \int_\Omega \xi(z^{\mathrm{in}}) \psi(0, \cdot) \, \mathrm{d}x - \int_0^T \int_\Omega \xi(z) \partial_t \psi \, \mathrm{d}x \mathrm{d}t$$

$$= -\int_0^T \int_\Omega D_{ij} \xi(z) A_{ik}(z) \nabla z_k \cdot \nabla z_j \psi \, \mathrm{d}x \mathrm{d}t \tag{2.1}$$

$$- \int_0^T \int_\Omega D_i \xi(z) A_{ik}(z) \nabla z_k \cdot \nabla \psi \, \mathrm{d}x \mathrm{d}t + \int_0^T \int_\Omega D_i \xi(z) R_i(z) \psi \, \mathrm{d}x \mathrm{d}t.$$

Remark 2.2. *By approximation, given a renormalised solution z, the equality (2.1) can be seen to hold true for a larger set of test functions $\psi \in C(I \times \bar{\Omega})$ with $\partial_t \psi \in L^1_{\mathrm{loc}}(I; L^1(\Omega))$, $\nabla \psi \in L^2_{\mathrm{loc}}(I; L^2(\Omega))$, and for truncation functions $\xi \in C^2(\mathbb{R}_{\geq 0}^{1+n})$ with $\mathrm{supp}\, D\xi$ compact.*

Remark 2.3 (Notation). *Let z denote a renormalised solution of* (ERDS) *in the sense of Def. 2.1. To keep notation simple and better emphasise the entropy structure of the diffusive part, we will often use a 'symbolic' notation writing $\mathbb{M}_{il}(z) \nabla D_l h(z)$ instead of $A_{ik}(z) \nabla z_k$, where as before the summation convention is used. Likewise, we write $\nabla D_i h(z)$ instead of the more precise notation $D_{ij} h(z) \nabla z_j$. The point here is that while, by hypothesis, the weak derivatives ∇z_j, $j = 0, \ldots, n$, are well-defined (in the standard Sobolev/distributional sense), the function $D_i h(z)$ may not be weakly differentiable.*

Remark 2.4. *The following equation, equivalent to (2.1), can be obtained by 'reversing the product rule' for ∇*

$$\int_\Omega \xi(z(T, \cdot)) \psi(T, \cdot) \, \mathrm{d}x - \int_\Omega \xi(z^{\mathrm{in}}) \psi(0, \cdot) \, \mathrm{d}x - \int_0^T \int_\Omega \xi(z) \partial_t \psi \, \mathrm{d}x \mathrm{d}t$$

$$= -\int_0^T \int_\Omega \nabla \big(D_i \xi(z) \psi \big) \cdot \mathbb{M}_{il}(z) \nabla D_l(z) \, \mathrm{d}x \mathrm{d}t + \int_0^T \int_\Omega D_i \xi(z) R_i(z) \psi \, \mathrm{d}x \mathrm{d}t, \tag{2.2}$$

where we recall our convention $\mathbb{M}_{il}(z) \nabla D_l h(z) := A_{ik}(z) \nabla z_k$, see Remark 2.3.

Notice that thanks to the hypothesis of $D\xi$ being compactly supported, no growth restrictions have to be imposed on the reaction term $R(z)$ in (2.1), and none of the flux terms $A_{ik}(z)\nabla z_k$ is necessarily required to be integrable in order for the integrals in (2.1) to converge. At the same time, this restrictive condition on the set of admissible truncation functions ξ means that recovering a separate (e.g. weak) formulation of a single component $i_0 \in \{0, \dots, n\}$ of the system (assuming integrability of $A_{i_0 k}(z)\nabla z_k$ and $R_{i_0}(z)$) is not immediate unless all components of the flux and the reactions are known to be integrable. Thus, consistent with the existence result for (ERDS) in [24], our weak-strong uniqueness principle for renormalised solutions additionally assumes a weak formulation for the energy component.

Definition 2.5 (Weak solutions of energy equation). *Let $z = (u, c)$ be as in Def. 2.1. We say that u is a weak solution of the energy component $\partial_t u = \operatorname{div}(A_{0j}(z)\nabla z_j)$ in Ω_{T^*} with no-flux boundary conditions and initial condition u^{in} if $A_{0j}(z)\nabla z_j \in L^1_{\mathrm{loc}}(I; L^1(\Omega))$ and if for all $\varphi \in C^1(I \times \bar{\Omega})$ and almost all $T < T^*$*

$$
\int_\Omega u(T, \cdot)\varphi(T, \cdot)\,\mathrm{d}x - \int_\Omega u^{\mathrm{in}}\varphi(0, \cdot)\,\mathrm{d}x - \int_0^T \int_\Omega u\partial_t\varphi\,\mathrm{d}x\mathrm{d}t
$$
$$
= -\int_0^T \int_\Omega A_{0j}(z)\nabla z_j \cdot \nabla\varphi\,\mathrm{d}x\mathrm{d}t. \tag{2.3}
$$

By carefully using lower-semicontinuity type properties of the entropy and entropy dissipation, the existence proof of global-in-time weak and renormalised solutions to ERDS as provided in [24] typically allows to show that for almost all $T \in (0, \infty)$ the constructed solutions satisfy the entropy dissipation inequality

$$
H(z(T)) - H(z^{\mathrm{in}}) \le -\int_0^T \int_\Omega \mathscr{P}(z)\,\mathrm{d}x\mathrm{d}t + \int_0^T \int_\Omega R_i(z)D_i h(z)\,\mathrm{d}x\mathrm{d}t, \tag{ED}
$$

where we recall the notation $\mathscr{P}(z) := \nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \ge 0$. Observe that thanks to the non-negativity of $\mathscr{P}(z)$ and $-R_i(z)D_i h(z)$ (to be assumed throughout), any function $z = (u, c)$ with $u \in L^\infty_t L^1_x$ and well-defined, measurable $\mathscr{P}(z)$ that satisfies (ED) with $h(z^{\mathrm{in}}) \in L^1(\Omega)$ necessarily has the regularity $\mathscr{P}(z) \in L^1(\Omega_T)$ and $R_i(z)D_i h(z) \in L^1(\Omega_T)$. (For the estimate on $H(z(T))$ required in this argument, see (6.7) in the appendix.)

We further note that the energy equation $\partial_t u = \operatorname{div}(A_{0j}(z)\nabla z_j)$ is satisfied in the weak sense by the solutions constructed in ref. [24], that $\nabla u \in L^2(\Omega_T)$, and that the quantity $G(u) = \frac{1}{2}\|u\|^2_{L^2(\Omega)}$ satisfies (with an equality)

$$
G(u(T)) - G(u^{\mathrm{in}}) \le -\int_0^T \int_\Omega a(z)|\nabla u|^2\,\mathrm{d}x\mathrm{d}t - \int_0^T \int_\Omega m(z)\nabla D_0 h(z) \cdot \nabla u\,\mathrm{d}x\mathrm{d}t \tag{ENE}
$$

for almost all $T \in (0, \infty)$. (In [24], the case $m \equiv 0$ was considered.)

Our main theorem asserts a weak-strong uniqueness property of entropy-dissipating renormalised solutions to the energy-reaction-diffusion problem (ERDS) under a set of hypotheses that is motivated by and compatible with the models in [24] (see system (2.5) below). Before stating the theorem, we need to specify what we understand by a 'strong' solution.

Definition 2.6 (Weak solution). *Let $I = [0, T^*)$. We call a function $z \in L^1_{\mathrm{loc}}(I; W^{1,1}(\Omega))$, $z = (u, c_1, \dots, c_n)$ with $z_i \ge 0$ for all i, a weak solution of (ERDS) in Ω_{T^*} with initial data z^{in} if $A_{ij}(z)\nabla z_j, R_i(z) \in$*

$L^1_{\text{loc}}(I; L^1(\Omega))$ *for all* i, *and if for all* $\psi \in C^\infty(I \times \bar{\Omega})^{1+n}$ *and almost all* $T \in (0, T^*)$

$$\int_\Omega z_i(T, \cdot)\psi_i(T, \cdot)\,\mathrm{d}x - \int_\Omega z_i^{\text{in}}\psi_i(0, \cdot)\,\mathrm{d}x - \int_0^T \int_\Omega z_i \partial_t \psi_i \,\mathrm{d}x\mathrm{d}t$$

$$= -\int_0^T \int_\Omega A_{ij}(z)\nabla z_j \cdot \nabla \psi_i \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_\Omega R_i(z)\psi_i \,\mathrm{d}x\mathrm{d}t. \tag{2.4}$$

Definition 2.7 ('Strong' solution). *Let* $I = [0, T^*)$. *We call* $z = z(t, x)$ *a strong solution of system* (ERDS) *in* Ω_{T^*} *with data* z^{in} *if it is a weak solution in the sense of Definition 2.6, has the regularity* $z \in C^{0,1}_{\text{loc}}(I \times \bar{\Omega})$ *and satisfies* $\inf_{\Omega_T} z_i > 0$ *for* $i = 0, \ldots, n$ *and every* $T \in (0, T^*)$.

Theorem 2.8 (Weak-strong uniqueness). *Let* $T^* \in (0, \infty]$. *Assume hp.* (A1)–(A5). *Further suppose that* (A6) *holds true, or alternatively that* $\varpi = 0$ *and that* (A6') *is fulfilled. Let* $z^{\text{in}} = (u^{\text{in}}, c^{\text{in}}) \in L^1(\Omega)^{1+n}$, $z_i^{\text{in}} \geq 0$ *for all* i, *and* $h(z^{\text{in}}) \in L^1(\Omega)$, $u^{\text{in}} \in L^2(\Omega)$. *Let* $z = (u, c)$ *be a renormalised solution of system* (ERDS) *in* Ω_{T^*} *taking the initial data* z^{in}, *and let* u *be a weak solution of the energy equation in the sense of Definition 2.5. Further suppose that, for almost all* $T \in (0, T^*)$, z *satisfies the entropy inequality* (ED) *and the energy inequality* (ENE) *with* $\nabla u \in L^2(\Omega_T)$. *If* $\tilde{z}$ *is a strong solution in* Ω_{T^*} *in the sense of Definition 2.7 taking the same initial data* z^{in}, *then* $z = \tilde{z}$ *a.e. in* Ω_{T^*}.

The proof of Theorem 2.8 will be completed in Section 3.4. In Section 2.3 we provide a list of examples covered by this theorem, including a class of (isoenergetic) cross-diffusion systems.

As mentioned in the introduction, the proof of Theorem 2.8 is based on a weak-strong stability type estimate with a generalised distance involving a modified relative entropy and an L^2-part for the energy component (cf. Sec. 1.3). The entropy inequality (ED) is a fundamental ingredient in the proof. As pointed out above, the solutions constructed in ref. [24] enjoy this estimate. The hypothesis in Theorem 2.8 that admissible weak solutions satisfy some kind of entropy/energy dissipation inequality is not uncommon in systems with strong nonlinearities; similar assumptions are encountered in the context of weak-strong uniqueness in fluid dynamics problems, see e.g. [45, 42] and references therein. Yet for many of the models we are interested in, the entropy dissipation inequality (ED) as well as the energy (in)equality (ENE) can be derived for general renormalised solutions z in the sense of Definition 2.1 (with the u-component satisfying a scalar problem as in Def. 2.5) if one assumes integrability of the quantities $\mathscr{P}(z)$ and $a(z)|\nabla u|^2$. In order to illustrate the ideas, we provide a proof of inequality (ED) for one of the models in the ref. [24] considering ERDS of the form

$$\dot{u} = \mathrm{div}\big(a(u, c)\nabla u + m(z)\nabla D_0 h(z)\big), \tag{2.5a}$$

$$\dot{c}_i = \mathrm{div}\left(m_i(u, c)\nabla \log\left(\tfrac{c_i}{w_i(u)}\right) + a(u, c)c_i \tfrac{w_i'(u)}{w_i(u)}\nabla u\right) + R_i(u, c), \tag{2.5b}$$

where, as before, $h(u, c)$ satisfies (h1), (h2), and where $a(z) := \pi_1(z)\gamma(z)$, with

$$\gamma(u, c) := \big(\partial_u^2 h - \sum_{i=1}^n c_i\big(\tfrac{w_i'}{w_i}\big)^2\big) = -\hat{\sigma}''(u) - \sum_{i=1}^n c_i \tfrac{w_i''(u)}{w_i(u)} > 0. \tag{2.6}$$

System (2.5) is obtained as a special case of (ERDS) by choosing

$$\mathbb{M}(z) := \mathrm{diag}\big(m, m_1, \ldots, m_n\big) + \pi_1 \mu \otimes \mu \tag{2.7}$$

with non-negative functions $m, m_i, \pi_1 \in C([0, \infty)^{1+n})$ to be specified below, and $\mu = (1, \mu_1, \ldots, \mu_n)$ determined by

$$\mu_i(u, c) := \frac{w_i'(u)}{w_i(u)}c_i \qquad \text{for } i \in \{1, \ldots, n\}.$$

Throughout, the bound

$$0 \leq \sqrt{\pi_1(u, c)} \lesssim 1 + u \tag{2.8}$$

and the mild regularity condition $\sqrt{\pi_1(u, c)} w_i'(u) \in C([0, \infty)^{1+n})$, $i = 1, \ldots, n$, will be imposed, the latter ensuring the continuity of $\mathbb{M}$ in $[0, \infty)^{1+n}$.

The functions $m_i(u, c)$, $i \in \{1, \ldots, n\}$, are assumed to take the following form for certain $a_i \in C([0, \infty)^{1+n})$ and $\kappa_{0,i}, \kappa_{1,i} \geq 0$:

$$m_i(u, c) = c_i a_i(u, c), \qquad \text{where} \quad a_i(u, c) \sim \kappa_{0,i} + \kappa_{1,i} c_i. \tag{2.9}$$

Hypotheses of Model (M0). Model **(M0)** consists of the following conditions:

- Entropy density h is given by (h1), (h2) with coefficient functions satisfying

$$(w_i')^2 \lesssim -w_i'' w_i \tag{2.10}$$

  for all $i \in \{1, \ldots, n\}$.
- Reactions $R_i \in C([0, \infty)^{1+n})$, $i = 1, \ldots, n$, satisfy (A2.i), where $R_0 \equiv 0$.
- $\mathbb{M}$ is given by (2.7)–(2.9) with
  - rank-one part: $\pi_1(z) \sim \frac{1}{\gamma(z)}$, where γ is given by (2.6).
  - diagonal part: $0 \leq m(z) \lesssim \varpi$ for some $\varpi \in \{0, 1\}$, and $m_i(u, c)$ given by (2.9) with $\kappa_{0,i} = 1, \kappa_{1,i} = 0$ for all $i \in \{1, \ldots, n\}$.

We recall that the evolution law (ERDS) associated with Model **(M0)** takes the form (2.5a), (2.5b). (Cf. Lemma 6.1 and [24] for details.) It is further easy to see that condition (2.8) is compatible with the choice $\pi_1(z) \sim \frac{1}{\gamma(z)}$ for any power law $\hat{\sigma}(u) = u^\nu$, $\nu \in (0, 1)$ and $\hat{\sigma}(u) = \log(u)$.

Model **(M0)** generalises the special case $\mathbb{M}(u, c) = \text{const} \cdot (D^2 h(u, c))^{-1}$ considered in [28]. It allows for species-dependent diffusivities, and genuinely contains cross terms in this case. More precisely, for models with species-dependent diffusion coefficients, thermodynamical consistency always leads to cross-diffusion effects, since for a diagonal diffusion matrix $A = \text{diag}(\ldots) \in \mathbb{R}^{(1+n) \times (1+n)}$ that is not a multiple of the identity matrix the product $\mathbb{M} = A(D^2 h)^{-1}$ is not symmetric.

In the derivation of the entropy dissipation inequality (ED), we need to assume that

$$m(u, c) \sum_{l=1}^{n} c_l \lesssim (1 + u)^2, \tag{2.11}$$

and have to impose the following conditions mainly serving to avoid issues for small values of u close to zero:

$$\begin{cases} m(u, c) |\hat{\sigma}''(u)| \lesssim 1, \\ -\hat{\sigma}''(v) \lesssim -\hat{\sigma}''(u) + 1 & \text{for all } v \geq u. \end{cases} \tag{2.12}$$

Proposition 2.9 (Strong entropy dissipation inequality). *Let $T^* \in (0, \infty]$. Let the hypotheses of Model **(M0)** hold, and assume locally ϵ_0-Hölder continuous reactions $R \in C_{\text{loc}}^{\epsilon_0}([0, \infty)^{1+n})$ for some $\epsilon_0 \in (0, 1)$. In addition, assume the bounds (2.11), (2.12), and suppose that $w_i \in C^2([0, \infty))$ for $i = 1, \ldots, n$. Let $z^{\text{in}} = (u^{\text{in}}, c^{\text{in}})$, $z_i^{\text{in}} \geq 0$, and $u^{\text{in}}, \hat{\sigma}_-(u^{\text{in}}) \in L^1(\Omega)$ and $c_i^{\text{in}} \in L \log L(\Omega)$ for all i. Let $z = (u, c)$ have non-negative components and suppose that*

$$u \in L_{\text{loc}}^\infty([0, T^*); L^1(\Omega)),$$
$$\mathscr{P}(z) \in L^1(\Omega_T), \quad a(z) |\nabla u|^2 \in L^1(\Omega_T) \quad \text{for all } T < T^*.$$

If z is a renormalised solution (in the sense of Def. 2.1) of system (ERDS) in Ω_{T^*} with initial data z^{in}, then the strong entropy dissipation inequality is satisfied, i.e.

$$H(z(t)) - H(z(s)) \leq -\int_s^t \int_\Omega \mathscr{P}(z) \, \mathrm{d}x \mathrm{d}\tau + \int_s^t \int_\Omega R_i(z) D_i h(z) \, \mathrm{d}x \mathrm{d}\tau \tag{ED.s}$$

for a.e. $0 \le s < t < T^*$, and for $s = 0$[2] and a.e. $t \in (0, T^*)$. In particular, ineq. (ED) *holds true for a.e. $T \in (0, T^*)$.*

For the proof of Proposition 2.9, see Section 4. Some comments on generalisations of Proposition 2.9 to other models are provided in Section 2.3.

In our final main result, we illustrate that a version of the generalised distance can further be used to prove exponential convergence to equilibrium. Exponential convergence in relative entropy has been studied at a formal level in [39] by means of log-Sobolev type inequalities leading to entropy-entropy dissipation estimates (see also [28]). In contrast to the present approach, [39] solely relies on the relative entropy, which restricts the results to thermal parts σ close to linear; for instance, the choice $\sigma(u) \sim \log u$ related to gas dynamics is only admitted in dimensions $d \le 2$. For simplicity, in the following result we disregard the reaction term and refer to [39] for applications on mass action-type reactions. We will further assume the strong energy inequality i.e.

$$G(u(t)) - G(u(s)) \le - \int_s^t \int_\Omega a(z)|\nabla u|^2 \, \mathrm{d}x\mathrm{d}\tau - \int_s^t \int_\Omega m(z)\nabla D_0 h(z) \cdot \nabla u \, \mathrm{d}x\mathrm{d}\tau \quad \text{(ENE.s)}$$

for a.e. $0 \le s < t < T^*$, and for $s = 0$ and a.e. $t \in (0, T^*)$.

Proposition 2.10 (Exponential convergence to equilibrium). *Recall that $|\Omega| = 1$ and let the hypotheses of Model* **(M0)** *hold. Let $z^{\mathrm{in}} = (u^{\mathrm{in}}, c^{\mathrm{in}})$ have non-negative components with $\hat{\sigma}_-(u^{\mathrm{in}}) \in L^1(\Omega)$, $u^{\mathrm{in}} \in L^2(\Omega)$, $c_i^{\mathrm{in}} \in L \log L$. Let $z = (u, c)$ with $u \in L^\infty_{\mathrm{loc}}([0, \infty), L^1(\Omega))$ and $\nabla u \in L^2_{\mathrm{loc}}([0, \infty), L^2(\Omega))$ be a global-in-time renormalised solution of* (ERDS) *with $R \equiv 0$, and suppose that* (ED.s) *and* (ENE.s) *are satisfied (with $T^* = \infty$).*

Then $\int_\Omega z_i(t, x) \, \mathrm{d}x = \int_\Omega z_i^{\mathrm{in}}(x) \, \mathrm{d}x =: \bar{z}_i$ for all $i \in \{0, \dots, n\}$ and a.a. $t > 0$, and there exist constants $\alpha \in (0, 1]$ and $\lambda = \lambda(\bar{z}, \alpha, \Omega) > 0$ such that for a.e. $t > 0$

$$\mathrm{Dist}_\alpha(z(t), \bar{z}) \le \mathrm{e}^{-\lambda t} \mathrm{Dist}_\alpha(z^{\mathrm{in}}, \bar{z}), \qquad (2.13)$$

where $\mathrm{Dist}_\alpha(z, \bar{z}) = H_{\mathrm{rel}}(z, \bar{z}) + \alpha G_{\mathrm{rel}}(u, \bar{u})$ (cf. eq. (1.5)).

See Section 5 for the proof of Prop. 2.10. For non-trivial reactions obeying mass action kinetics, the steady state $\bar{z}$ associated with (ERDS) is determined by solving a constrained minimisation problem for the convex entropy functional $H(u, c)$ imposing energy conservation $E(u, c) = \int_\Omega u \overset{!}{=} E_0$ and further linear constraints taking into account possible conservation laws for the concentrations (see [28, 39]). Prop. 2.10 considers the simplest case, where all species have a conserved mass. Extension to more general reactions is usually achieved by means of suitable coercivity estimates for the dissipation term $-D_i h(z) R_i(z) \ge 0$ associated with the reactions. See [28, 39] for applications in a non-isothermal setting; for previous works in the isothermal case, we refer to [38, 17, 37] and references therein. Let us observe that since **(M0)** allows for $m \neq 0$, leading to energy flux induced by temperature gradients, a maximum principle for the internal energy density (as is crucially used in [28]) is not available here. We further note that, with a standard Csiszár–Kullback–Pinsker inequality [15, 44], the bound (2.13) can be used to deduce exponential convergence to equilibrium in $L^2(\Omega) \times (L^1(\Omega))^n$.

Remark 2.11. *Observe that the condition* (ED.s) *in Prop 2.10 is satisfied under the additional hypotheses on the coefficient functions imposed in Prop. 2.9. Under suitable regularity hypotheses, the strong energy inequality* (ENE.s) *(with an equality) can be proved similarly as in* [24, Lemma 6.1].

2.3. **Examples.** Below, we provide relevant applications of the weak-strong uniqueness result.

---

[2]With the understanding that $H(z(0)) = H(z^{\mathrm{in}})$.

2.3.1. *Energy-reaction-diffusion systems.* The hypotheses of Theorem 2.8 are compatible with the class of ERDS introduced in [24]. In that work, the existence of generalised solutions (weak or renormalised) has been established for two classes of models, both taking the form (2.5) with $m \equiv 0$.

One of the models in [24] is a special case of **(M0)** (see page 13) with $m \equiv 0$. A brief verification of the model hypotheses of Theorem 2.8 for **(M0)** is provided in the appendix, see Lemma 6.2. The existence analysis for **(M0)** focuses on reactions obeying suitable growth conditions, in which case there are global-in-time weak solutions. However, renormalised solutions can be constructed by adapting the proof of [24, Theorem 1.8], and in this case no growth restrictions on $|R(z)|$ are required. Conceptually, the construction of renormalised solutions for **(M0)** is simpler than in [24, Theorem 1.8], since the diffusive flux is integrable and the strong convergence of ∇u is not required in case of **(M0)**.

The second class of models considered in [24] will here be referred to as **(M1)**. It again takes the form (2.5) with $m \equiv 0$, and assumes the following conditions:

- Entropy density h is given by (h1), (h2).
- Reactions $R_i \in C([0, \infty)^{1+n})$, $i = 1, \ldots, n$, satisfy (A2.i), where $R_0 \equiv 0$.
- $\mathbb{M}$ is given by (2.7)–(2.9) with $m(z) \equiv 0$, and $m_i(u, c)$ given by (2.9), where $\kappa_{0,i}, \kappa_{1,i}$ satisfy one of the following:
    - (i) $\kappa_{0,i} = 1$ and $\kappa_{1,i} = 0$ for all i
    - (ii) $\kappa_{0,i} \geq 0$ and $\kappa_{1,i} = 1$ for all i
- Moreover,
    - $\pi_1 \gamma \gtrsim 1$ \qquad (γ as in (2.6))
    - $w_i'(u) \lesssim -w_i''(u)\sqrt{\pi_1}$
    - $\sqrt{\pi_1} \frac{w_i'}{w_i} \lesssim 1$.

Global weak solutions in case (ii) have been constructed in [24, Theorem 1.2] for reactions obeying suitable growth hypotheses. More interesting is case (i), in which the existence of global renormalised solutions has been established in [24, Theorem 1.8] for general continuous reactions satisfying (A2.i). A pecularity of this model lies in the circumstance that the renormalised formulation is needed not only to give a meaning to the reaction rates, but also to the diffusion flux $A(z)\nabla z$, which may be non-integrable. Model (M1.i) satisfies conditions (A3.a) (with $\varpi = 0$), (A3.b), (A4), (A5) and (A6') of Theorem 2.8. We refer to [24, Section 2.2] (notably the proof of Lemma 2.3 therein), where the necessary estimates can be found. For verifying (A6'), one should also use the fact that the coefficient function $a(z) = \pi_1(z)\gamma(z)$ satisfies the bound $\chi_{\{u \geq \underline{u}\}}|a(u, c)| \lesssim C(|z|, \underline{u})$ for any $\underline{u} > 0$. Thus, under the extra smoothness assumptions $\hat{\sigma}, w_i \in C^4((0, \infty))$ and $R_i, m, m_i, \pi_1 \in C^{0,1}_{\text{loc}}((0, \infty)^{1+n})$, Theorem 2.8 is applicable. We caveat that verifying (ED) as it was done in Prop. 2.9 for **(M0)** is much more delicate for Model (M1.i) due to the possibility of strong cross diffusion caused by a non-integrable diffusion flux. Whether or not (M1.i) admits an analogue of Prop. 2.9 is an open question.

2.3.2. *Reaction-diffusion systems.* Our results further apply to (isoenergetic) population models generalising the two-species system for pattern formation by Shigesada, Kawasaki, and Teramoto (SKT). Reduction to the isoenergetic case is achieved by choosing $\varpi = 0$ (see (A3)) and $u \equiv u^{\text{in}}$ to be spatially constant, which is consistent with the evolution law for the energy density u if $\varpi = 0$. Then, the given energy density $u = u^{\text{in}} \in \mathbb{R}^+$ can be regarded as a fixed system's parameter (in particular $\nabla u = 0$) and one can write $h = h(c), A = A(c), R = R(c)$, and $\mathscr{P} = \mathscr{P}(c)$ etc.

The generalised SKT system as considered in [9] states

$$
\begin{aligned}
\partial_t c_i &= \nabla \cdot (A_{ij}(c)\nabla c_j) + R_i(c), &\quad t > 0, x \in \Omega, \\
0 &= A_{ij}(c)\nabla c_j \cdot \nu, &\quad t > 0, x \in \partial\Omega,
\end{aligned}
\tag{SKT}
$$

where $A_{ij}(c) = \delta_{ij}p_i(c) + c_i\frac{\partial p_i}{\partial c_j}(c)$ for $i,j = 1,\dots,n$ with $p_i(c) = a_{i0} + \sum_{k=1}^n a_{ik}c_k^s$ for suitable $a_{il} \geq 0$ and some $s > 0$. Under certain hypotheses, this system has a generalised gradient structure with entropy density given by $h(c) = \sum_{i=1}^n \pi_i\lambda_s(c_i)$ for constants $\pi_i > 0$ and λ_1 given by (1.3), $\lambda_s(r) = \frac{r^s - sr}{s-1} + 1$ for $s \neq 1$. Under a weak cross-diffusion (wc) condition (see eq. (12) in [9]) or the detailed balance (db) condition $\pi_i a_{ij} = \pi_j a_{ji}$ for all $i,j \in \{1,\dots,n\}$ together with $a_{i0} > 0$, $a_{ii} > 0$ for all i, the matrix $\mathbb{M}(z) = A(z)(D^2h(z))^{-1}$ satisfies the non-degeneracy condition (A3.b), i.e. $\mathbb{M}(c) \gtrsim_\iota I_n$ whenever $\min\{c_1,\dots,c_n\} \geq \iota$, see the explicit estimate in [12, Lemma 2.1] for $s = 1$, and [9, Section 2] for the general case under certain extra hypotheses. Observe that the detailed balance condition is equivalent to the symmetry of the mobility matrix $\mathbb{M} = A(D^2h)^{-1}$, and thus ensures the symmetry of the diffusive part $\mathbb{K}_{\text{diff}}$ of the Onsager operator.

a) Linear transition rates $s = 1$. In this case, assuming (wc) or (db) with $a_{i0} > 0$, $a_{ii} > 0$, one has $\mathscr{P}(c) \gtrsim \sum_{i=1}^n \left(|\nabla c_i|^2 + |\nabla\sqrt{c_i}|^2\right)$ and $|A(c)\nabla c| \lesssim |c||\nabla c|$. Thus, assumptions (A1), (A3)–(A6) on the entropy density and the mobility matrix are satisfied. (The more general weights $\pi_i > 0$ in $h(c)$ as opposed to the unit weights in (A1) do not impact the analysis.) We leave it to reader to verify that, by adapting the proof of Prop. 2.9 (see also [23, Prop. 5]), the entropy dissipation inequality (ED) can be proved for renormalised solutions to this system, when assuming the regularity

$$c_i, \sqrt{c_i} \in L^2_{\text{loc}}(I; H^1(\Omega)) \quad \text{for all } i. \tag{2.14}$$

Observe that this regularity is essentially equivalent to the assumption in Prop. 2.9 that $\mathscr{P}(z) \in L^1(\Omega_T)$ for all $T < T^*$. Thus, for (SKT) with $s = 1$ and reactions satisfying (A2), Theorem 2.8 yields the weak-strong uniqueness of renormalised solutions of the regularity (2.14). A similar result has been obtained previously in [12, Theorem 1]. Here, we should caveat that the regularity assumption $c_i \in L^2_{\text{loc}}(I; H^1(\Omega))$ *and* $\sqrt{c_i} \in L^2_{\text{loc}}(I; H^1(\Omega))$ is also needed in the proof of [12, Theorem 1], is, however, incompletely stated in this theorem. Moreover, our result shows that hypothesis '(H2.iii)' in [12] on the reactions can be dropped.

b) Nonlinear transition rates: an adjustment of our hypotheses further allows to treat system (SKT) with superlinear transition rates $s \in (1, 2]$. In this case, condition (A6) is no longer fulfilled, but with a suitable adjustment of the truncation function $\xi^*(c)$ in the definition of the modified relative entropy density $h_{\text{rel}}^*(z, \tilde{z})$ (given by (1.8)) our technique can still be applied. See Remark 3.1 for technical details.

For sublinear transition rates, $s < 1$, the problem becomes more delicate since a direct analogue of condition (A6) is not available. When relying exclusively on entropy estimates for a priori bounds, even the construction of renormalised solutions, which to the author's knowledge is currently only available in the case of linear transition rates [11] (but likely to be extendable to the case $s \in [1, 2]$), is open for small $s \in (0, 1)$.

3. WEAK-STRONG UNIQUENESS PRINCIPLE

In this section we will establish a stability estimate implying Theorem 2.8. Throughout this section, we therefore assume the hypotheses of Theorem 2.8. Before turning to the actual proof in Subsection 3.4, we gather some technical auxiliary results.

3.1. **Truncation function** $\xi^*(z)$. Let ι, B be fixed constants, $0 < \iota < 1 < B < \infty$, that are such that for all $i \in \{0, \dots, n\}$

$$2\iota \leq \tilde{z}_i \leq B. \tag{3.1}$$

Different regimes/case distinction. We henceforth abbreviate $|z|_1 = \sum_{i=0}^{n} z_i$. For sufficiently large but fixed parameters $E, N \geq 2$ (to be chosen later) we decompose $[0, \infty)^{1+n}$ into the following three sets:

$$\mathcal{A} = \{z \in [0, \infty)^{1+n} : |z|_1 \leq E\},$$
$$\mathcal{B} = \{z \in [0, \infty)^{1+n} : E < |z|_1 < E^N\},$$
$$\mathcal{C} = \{z \in [0, \infty)^{1+n} : |z|_1 \geq E^N\}.$$

A motivation for this decomposition is given in the introduction (see Section 1.3). Let us also mention that in the proof of Theorem 2.8 the set $\mathcal{A}$ will be further decomposed into $\mathcal{A}_+$ and $\mathcal{A}_0$ (defined in (3.18)). The parameter E will always be supposed to satisfy $E \geq 2B$. A finite number of further lower bounds on E will be imposed later on.

Adjusted relative entropy. We now define

$$h_{\mathrm{rel}}^*(z, \tilde{z}) = h(z) - D_i h(\tilde{z})(\xi^*(z)z_i - \tilde{z}_i) - h(\tilde{z}),$$

where $\xi^* = \xi^{(E,N)} \in C^\infty([0, \infty)^{1+n})$, $0 \leq \xi^* \leq 1$, is a truncation function subordinate to the above case distinction enjoying the following properties:

(t1) $\xi^*(z) = 1$ and $D^k \xi^*(z) = 0$ for all $k \in \mathbb{N}_+$ if $z \in \mathcal{A}$,
(t2) $\xi^*(z) = 0$ and $D^k \xi^*(z) = 0$ for all $k \in \mathbb{N}_+$ if $z \in \mathcal{C}$,

and

(t3) $|D\xi^*(z)| \lesssim \frac{1}{N|z|_1}$, $|D^2\xi^*(z)| \lesssim \frac{1}{N(|z|_1)^2}$ for all $z \in [0, \infty)^{1+n}$.

A function ξ^* that has these properties can be obtained as follows: let $\vartheta \in C_b^\infty(\mathbb{R})$ be non-increasing with $\vartheta(r) = 1$ for $r \leq 0$ and $\vartheta(r) = 0$ for $r \geq 1$ and define (cf. [23])

$$\xi^*(z) = \xi^{(E,N)}(z) = \vartheta\left(\frac{\log(|z|_1) - \log(E)}{\log(E^N) - \log(E)}\right).$$

It is elementary to check that this choice satisfies the above properties. For instance, note that

$$|D\xi^{(E,N)}(z)| \lesssim \frac{1}{N|z|_1} \cdot \frac{1}{\log(E)}$$

if $N \geq 2$.

Remark 3.1 (Superlinear transition rates). *When dealing with (isoenergetic) reaction-cross-diffusion population systems with* superlinear *transition rates $s \in (1, 2]$ for concentrations $c_1, \ldots, c_n$ (see Example b) in Sec. 2.3.2), the decay property* (t3) *is no longer sufficient, and the above choice of ξ^* should be replaced by*

$$\xi^*(c) := \xi^{(E,N,s)}(c) := \vartheta\left(\frac{\log(\rho_s^s(c)) - \log(E^s)}{\log(E^{sN}) - \log(E^s)}\right) = \vartheta\left(\frac{\log(\rho_s(c)) - \log(E)}{\log(E^N) - \log(E)}\right), \qquad (3.2)$$

where ϑ is as before, and

$$\rho_s(c) := \left(\sum_{i=1}^{n}(c_i + \delta)^s\right)^{\frac{1}{s}}$$

with $\delta = \delta(s) = 1$ for $s \in (1, 2)$ and $\delta(s) = 0$ for $s = 2$. The parameter $\delta \in \{0, 1\}$ ensures the smoothness of ρ_s^s and thus of ξ^*.

Introducing the regimes $\mathcal{A}^{(s)} = \{\rho_s(c) \leq E\}$, $\mathcal{B}^{(s)} = \{E < \rho_s(c) < E^N\}$, $\mathcal{C}^{(s)} = \{\rho_s(c) \geq E^N\}$, properties (t1) *and* (t2) *remain valid with $\mathcal{A}, \mathcal{C}$ replaced by $\mathcal{A}^{(s)}$ and $\mathcal{C}^{(s)}$, respectively. Derivatives of*

ξ^* *given by* (3.2) *enjoy the following decay properties, specifically adapted to the problem at hand,*

$$|D_i\xi^*(c)| \lesssim \frac{(c_i+\delta)^{s-1}}{N|c|_s^s}, \qquad |D_{ik}\xi^*(c)| \lesssim \frac{(c_i+\delta)^{s-1}(c_k+\delta)^{s-1}}{N|c|_s^{2s}} + \frac{(c_i+\delta)^{s-2}}{N|c|_s^s}\delta_{ik},$$

where δ_{ik} *denotes the Kronecker delta. As will become clear in Section 3.4, if* $c \in \mathcal{A}^{(s)}$, *one can simply follow the reasoning in the proof of Theorem 2.8 using the fact that* (SKT) *satisfies (under reasonable hypotheses) a non-degeneracy condition analogous to* (A3.b).

If $c \in \mathcal{B}^{(s)}$, *we need an analogue of assumption* (A6) *to be able to absorb terms without a good sign involving products of gradients of the renormalised solution by the entropy dissipation. Typical systems* (SKT) *satisfy the coercivity bound (cf.* [9])

$$\mathscr{P}(c) \gtrsim \sum_{i=1}^{n} |\nabla c_i^s|^2 + \sum_{i=1}^{n} |\nabla c_i^{s/2}|^2. \tag{3.3}$$

Confining to systems (SKT) *enjoying this bound, a suitable generalisation of hp.* (A6) *(in the isoenergetic case) that is satisfied by such systems is*

$$\chi_{\{|c|_s \geq 1\}} \cdot \left(c_i^{s-1} + c_i^{s/2-1}\right)|A_{ij}(c)\nabla c_j| \lesssim |c|_s^s\sqrt{\mathscr{P}(c)}.$$

Observe that since $s \in (1,2]$, *the factor* $\left(c_i^{s-1} + c_i^{s/2-1}\right)$ *in this condition can be equivalently replaced by* $\left((c_i+\delta)^{s-1} + c_i^{s/2-1}\right)$. *Using the above model bounds and decay properties of derivatives of* ξ^*, *one can verify the estimate*

$$\chi_{\{|c|_s \geq 1\}}|\nabla D_i(\xi^*(c)c_l)||A_{ij}(c)\nabla c_j| \lesssim \frac{1}{N}\mathscr{P}(c),$$

which allows to deal with the case $c \in \mathcal{B}^{(s)}$ *(cf. ineq.* (3.24) *in the proof of Thm 2.8). When verifying this bound, one uses the fact that* $|\nabla c_l| \lesssim |\nabla c_l^s| + |\nabla c_l^{s/2}|$, *which holds true since* $s \in (1,2]$.

If $c \in \mathcal{C}^{(s)}$, *one can follow the reasoning in the proof of Thm 2.8.*

Thus, for models with superlinear transition rates $s \in (1,2]$ *that satisfy* (3.3), *weak-strong uniqueness is obtained by adapting the proof of Theorem 2.8 as sketched above. Here, we also use the fact that thanks to the locally uniform convexity of the entropy density associated with* (SKT) *(see Sec. 2.3.2) an analogue of Prop. 3.2 is immediate.*

3.2. **Coercivity properties of the generalised distance.** We henceforth let $g(s) = \frac{1}{2}s^2$,

$$g_{\mathrm{rel}}(u,\tilde{u}) := g(u) - g'(\tilde{u})(u - \tilde{u}) - g(\tilde{u}) = \frac{1}{2}|u - \tilde{u}|^2, \tag{3.4}$$

and define for $\alpha \in (0,\infty)$

$$\mathrm{dist}^*_\alpha(z,\tilde{z}) = h^*_{\mathrm{rel}}(z,\tilde{z}) + \alpha g_{\mathrm{rel}}(u,\tilde{u}).$$

For the following assertion we recall that $z \in \mathcal{A}$ if and only if $\sum_{i=0}^{n} z_i \leq E$.

Proposition 3.2 (Coercivity properties). *Recall that* $h(z) = h(u,c)$ *is given by* (h1), (h2), *and that* $\tilde{z}$ *satisfies* (3.1). *For any* $E \in [1,\infty)$, *we have*[3]

$$|z - \tilde{z}|^2 \lesssim_E h^*_{\mathrm{rel}}(z,\tilde{z}) \quad \text{if } z \in \mathcal{A}. \tag{3.5}$$

There exists $\underline{E} = \underline{E}(\tilde{z}) < \infty$ *such that for any* $E \geq \underline{E}$ *and any* $\alpha \geq 1$

$$h^*_{\mathrm{rel}}(z,\tilde{z}) + \alpha g_{\mathrm{rel}}(u,\tilde{u}) \geq \epsilon\left(\sum_{i=1}^{n} c_i \log_+ c_i + u^2\right) + 1 \quad \text{if } z \in \mathcal{A}^c, \tag{3.6}$$

where $\epsilon > 0$ *is a positive constant only depending on model parameters.*

---

[3]We recall that any dependence of estimates on ι and B (see (3.1)) will usually not be indicated explicitly.

For any $\alpha \in (0, 1]$ there exists $\underline{E} = \underline{E}(\alpha, \tilde{z}) < \infty$ such that for any $E \geq \underline{E}$

$$h_{\mathrm{rel}}^*(z, \tilde{z}) + \alpha g_{\mathrm{rel}}(u, \tilde{u}) \geq \epsilon \left(\sum_{i=1}^{n} c_i \log_+ c_i + \alpha \, u^2 \right) + 1 \quad \text{if } z \in \mathcal{A}^c, \tag{3.7}$$

where $\epsilon > 0$ is a positive constant only depending on model parameters.

Remark 3.3. *Note that as long as $\alpha \in [1, \infty)$, the lower bound $\underline{E}$ can be chosen independently of α. This property is essential for obtaining the stability estimate (1.7) in the case where (A6) is not satisfied (including model **(M1)** in case (i), where $\kappa_{1,i} = 0$; cf. Section 2.3).*

Proof of Prop. 3.2. For the first assertion, we note that, as can be seen from the proof of [39, Prop. 2.1], the entropy density h is locally uniformly convex on $[0, \infty)^{1+n}$ with $D_{ij}h(z) \geq \epsilon_0(E)\delta_{ij}$ if $|z|_1 \leq E$. Since, by construction, $\xi^*(z) = 1$ whenever $z \in \mathcal{A}$, we thus infer

$$h_{\mathrm{rel}}^*(z, \tilde{z}) = h(z) - D_i h(\tilde{z})(z - \tilde{z})_i - h(\tilde{z}) \gtrsim_E |z - \tilde{z}|^2.$$

Let us now turn to assertions (3.6) and (3.7). Since $\tilde{u} \leq B$, we have the bound $|u - \tilde{u}|^2 \geq \frac{1}{4} u^2 \chi_{\{u \geq 2B\}}$. By Lemma 6.3, we further have for some $\nu \in [0, 1)$ and some positive constant $\epsilon > 0$

$$h(z) \geq \epsilon \sum_{i=1}^{n} c_i \log_+ c_i - C(\hat{\sigma}_+(u) + u^\nu) - C.$$

Hence,

$$h_{\mathrm{rel}}^*(z, \tilde{z}) + \alpha g_{\mathrm{rel}}(u, \tilde{u}) \geq \epsilon \sum_{i=1}^{n} c_i \log_+ c_i + \alpha u^2 \chi_{\{u \geq 2B\}} - C \sum_{i=0}^{n} z_i - C. \tag{3.8}$$

Inequality (3.6) is now immediate since $\sum_{i=1}^{n} c_i \log(c_i) + u^2$ dominates $\sum_{0=1}^{n} z_i$ for $|z|_1 \geq \underline{E}$ whenever $\underline{E}$ is large enough. Observe that the lower bound on $\underline{E}$ can be chosen independently of $\alpha \in [1, \infty)$.

Let now $\alpha \in (0, 1]$ be given. Inequality (3.8) shows that if $\underline{E} = \underline{E}(\alpha)$ is large enough, we obtain (3.7) if $E \geq \underline{E}$. $\qquad \square$

3.3. **Evolution inequality for the generalised distance.** We will abbreviate

$$H_{\mathrm{rel}}^*(z, \tilde{z}) := \int_\Omega h_{\mathrm{rel}}^*(z, \tilde{z}) \, \mathrm{d}x, \qquad G_{\mathrm{rel}}(u, \tilde{u}) := \int_\Omega g_{\mathrm{rel}}(u, \tilde{u}) \, \mathrm{d}x.$$

In this subsection, we exploit the evolution laws satisfied by the renormalised solution z and the strong solution $\tilde{z}$ of Theorem 2.8 to derive an evolution inequality for the generalised distance

$$\mathrm{Dist}_\alpha^*(z, \tilde{z}) := \int_\Omega \mathrm{dist}_\alpha^*(z, \tilde{z}) \, \mathrm{d}x = H_{\mathrm{rel}}^*(z, \tilde{z}) + \alpha \, G_{\mathrm{rel}}(u, \tilde{u}), \tag{3.9}$$

where $\alpha \in (0, \infty)$ is a suitably chosen weight (to be specified in Section 3.4).

We recall that $I := [0, T^*)$. Furthermore, we note that since $\partial_t \tilde{z} \in L^\infty(\Omega_T)$ for any $T < T^*$, we can integrate by parts with respect to time in the weak formulation (2.4) satisfied by the strong solution $\tilde{z}$ to find

$$\int_0^T \int_\Omega \partial_t \tilde{z}_i \psi_i \, \mathrm{d}x \mathrm{d}t = - \int_0^T \int_\Omega A_{ij}(\tilde{z}) \nabla \tilde{z}_j \cdot \nabla \psi_i \, \mathrm{d}x \mathrm{d}t + \int_0^T \int_\Omega R_i(\tilde{z}) \psi_i \, \mathrm{d}x \mathrm{d}t. \tag{3.10}$$

By a density argument, one can see that eq. (3.10) holds true for all $\psi \in L^1_{\mathrm{loc}}(I; W^{1,1}(\Omega))$.

We consider separately the two quantities $H_{\mathrm{rel}}^*(z, \tilde{z})$ and $G_{\mathrm{rel}}(u, \tilde{u})$ appearing in (3.9), beginning with the former.

Lemma 3.4 (Evolution of the entropic part). *For a.e. $T < T^*$ one has*

$$H_{\text{rel}}^*(z, \tilde{z})\Big|_{t=0}^{t=T} \leq \int_0^T \int_\Omega \rho^{(h)} \, \mathrm{d}x\mathrm{d}t,$$

where

$$
\begin{aligned}
\rho^{(h)} :={}& -\nabla D_i h(z) \cdot \mathbb{M}_{il}(z) \nabla D_l h(z) \\
& + \nabla(D_i(\xi^*(z)z_j)D_j h(\tilde{z})) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
& + \nabla\Big(D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)\Big) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z}) \\
& - D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)R_i(\tilde{z}) \\
& + (D_i h(z) - D_j h(\tilde{z})D_i(\xi^*(z)z_j))R_i(z).
\end{aligned}
\tag{3.11}
$$

Proof of Lemma 3.4. The subsequent observations apply to a.e. $T < T^*$.

We write

$$H_{\text{rel}}^*(z, \tilde{z}) = H(z) - \int_\Omega D_i h(\tilde{z})\xi^*(z)z_i \, \mathrm{d}x + \int_\Omega (D_i h(\tilde{z})\tilde{z}_i - h(\tilde{z})) \, \mathrm{d}x.$$

For the first term on the RHS we use the fact that, by hypothesis, z satisfies (ED), i.e.

$$H(z)\Big|_{t=0}^{t=T} \leq -\int_0^T \int_\Omega \nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \, \mathrm{d}x\mathrm{d}t + \int_0^T \int_\Omega D_i h(z)R_i(z) \, \mathrm{d}x\mathrm{d}t.$$

For the second term, we want to use the fact that z satisfies the renormalised formulations (2.1) and (2.2) with the truncation function $\xi(z) = \xi^*(z)z_j$ and the test function $\psi = D_j h(\tilde{z}) \in W^{1,\infty}(\Omega_T)$. (For the admissibility of this choice, see Remark 2.2.) Inserting these choices in eq. (2.2), we obtain

$$
\begin{aligned}
-\int_\Omega D_j h(\tilde{z})\xi^*(z)z_j \, \mathrm{d}x\Big|_{t=0}^{t=T} & + \int_0^T \int_\Omega \xi^*(z)z_j \tfrac{\mathrm{d}}{\mathrm{d}t}D_j h(\tilde{z}) \, \mathrm{d}x\mathrm{d}t \\
& = \int_0^T \int_\Omega \nabla(D_i(\xi^*(z)z_j)D_j h(\tilde{z})) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \, \mathrm{d}x\mathrm{d}t \\
& \quad - \int_0^T \int_\Omega D_i(\xi^*(z)z_j)R_i(z)D_j h(\tilde{z}) \, \mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

We next rewrite the second term on the LHS choosing in the weak form (3.10) for $\tilde{z}$ the test function $\psi := D_{ij}h(\tilde{z})\xi^*(z)z_j$. This yields

$$
\begin{aligned}
\int_0^T \int_\Omega D_{ij}h(\tilde{z})\xi^*(z)z_j\partial_t\tilde{z}_i \, \mathrm{d}x\mathrm{d}t = & -\int_0^T \int_\Omega \nabla\big(D_{ij}h(\tilde{z})\xi^*(z)z_j\big) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z}) \, \mathrm{d}x\mathrm{d}t \\
& + \int_0^T \int_\Omega D_{ij}h(\tilde{z})\xi^*(z)z_j R_i(\tilde{z}) \, \mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

Observe that since $\chi_{\{|z|\leq E\}}\nabla z \in L^2_{\text{loc}}(I; L^2(\Omega))$ for any $E < \infty$, the function $\psi := D_{ij}h(\tilde{z})\xi^*(z)z_j \in L^2_{\text{loc}}(I; H^1(\Omega))$ is indeed admissible in the weak equation (3.10) for $\tilde{z}$.

We finally need to determine the evolution of the term

$$\int_\Omega (D_i h(\tilde{z})\tilde{z}_i - h(\tilde{z})) \, \mathrm{d}x.$$

To this end, note that thanks to the regularity of $\tilde{z}$,

$$\int_\Omega h(\tilde{z})\,\mathrm{d}x\Big|_{t=0}^{t=T} = \int_0^T\int_\Omega D_i h(\tilde{z})\partial_t\tilde{z}_i\,\mathrm{d}x\mathrm{d}t,$$

$$\int_\Omega D_i h(\tilde{z})\tilde{z}_i\,\mathrm{d}x\Big|_{t=0}^{t=T} = \int_0^T\int_\Omega D_i h(\tilde{z})\partial_t\tilde{z}_i\,\mathrm{d}x\mathrm{d}t + \int_0^T\int_\Omega \left(\tfrac{\mathrm{d}}{\mathrm{d}t}D_i h(\tilde{z})\right)\tilde{z}_i\,\mathrm{d}x\mathrm{d}t.$$

Subtracting the first from the second equality then yields

$$\begin{aligned}
\int_\Omega \big(D_i h(\tilde{z})\tilde{z}_i - h(\tilde{z})\big)\,\mathrm{d}x\Big|_{t=0}^{t=T} &= \int_0^T\int_\Omega D_{ij}h(\tilde{z})\tilde{z}_i\partial_t\tilde{z}_j\,\mathrm{d}x\mathrm{d}t \\
&= \int_0^T\int_\Omega D_{ij}h(\tilde{z})\tilde{z}_j\partial_t\tilde{z}_i\,\mathrm{d}x\mathrm{d}t \\
&= -\int_0^T\int_\Omega \nabla\big(D_{ij}h(\tilde{z})\tilde{z}_j\big)\cdot\mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z})\,\mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega D_{ij}h(\tilde{z})\tilde{z}_j R_i(\tilde{z})\,\mathrm{d}x\mathrm{d}t,
\end{aligned}$$

where in the second step we have used the symmetry of the Hessian of h.

In combination, the above estimates yield the bound

$$\begin{aligned}
H^*_{\mathrm{rel}}(z,\tilde{z})\Big|_{t=0}^{t=T} &\leq -\int_0^T\int_\Omega \nabla D_i h(z)\cdot\mathbb{M}_{il}(z)\nabla D_l h(z)\,\mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega D_i h(z)R_i(z)\,\mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega \nabla(D_i(\xi^*(z)z_j)D_j h(\tilde{z}))\cdot M_{il}(z)\nabla D_l h(z)\,\mathrm{d}x\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega D_i(\xi^*(z)z_j)R_i(z)D_j h(\tilde{z})\,\mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega \nabla\big(D_{ij}h(\tilde{z})\xi^*(z)z_j\big)\cdot\mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z})\,\mathrm{d}x\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega D_{ij}h(\tilde{z})\xi^*(z)z_j R_i(\tilde{z})\,\mathrm{d}x\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega \nabla\big(D_{ij}h(\tilde{z})\tilde{z}_j\big)\cdot\mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z})\,\mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T\int_\Omega D_{ij}h(\tilde{z})\tilde{z}_j R_i(\tilde{z})\,\mathrm{d}x\mathrm{d}t.
\end{aligned}$$

The asserted inequality is now obtained upon rearranging the integrals on the RHS. $\qquad\square$

We next turn to the energetic part. We first note that equation (3.10) and the fact that $R_0 \equiv 0$ imply that

$$\int_0^T\int_\Omega \partial_t\tilde{u}\varphi\,\mathrm{d}x\mathrm{d}t = -\int_0^T\int_\Omega A_{0j}(\tilde{z})\nabla\tilde{z}_j\cdot\nabla\varphi\,\mathrm{d}x\mathrm{d}t \tag{3.12}$$

for all $\varphi \in L^1_{\mathrm{loc}}(I; W^{1,1}(\Omega))$.

Lemma 3.5 (Evolution of the energetic part). *Recall the definition of g_{rel} in (3.4) and the notation $G_{\mathrm{rel}}(u, \tilde{u}) := \int_\Omega g_{\mathrm{rel}}(u, \tilde{u}) \, \mathrm{d}x$. For almost every $T > 0$, we have*

$$G_{\mathrm{rel}}(u, \tilde{u}) \Big|_{t=0}^{t=T} = \int_0^T \int_\Omega \rho^g \, \mathrm{d}x \mathrm{d}t, \tag{3.13}$$

where

$$\begin{aligned}
\rho^{(g)} := &-a(z)|\nabla u - \nabla \tilde{u}|^2 \\
&- (a(z) - a(\tilde{z}))(\nabla u - \nabla \tilde{u}) \cdot \nabla \tilde{u} \\
&- m(z)(\nabla u - \nabla \tilde{u}) \cdot (\nabla D_0 h(z) - \nabla D_0 h(\tilde{z})) \\
&- (m(z) - m(\tilde{z}))(\nabla u - \nabla \tilde{u}) \cdot \nabla D_0 h(\tilde{z}).
\end{aligned} \tag{3.14}$$

Proof. We expand $g_{\mathrm{rel}}(u, \tilde{u}) = \frac{1}{2}u^2 - u\tilde{u} + \frac{1}{2}\tilde{u}^2$.

To deal with the first term on the RHS, we use the energy inequality (ENE), i.e. the property that for a.e. $T < T^*$

$$G(u) \Big|_{t=0}^{t=T} \leq -\int_0^T \int_\Omega a(z)|\nabla u|^2 \, \mathrm{d}x \mathrm{d}t - \int_0^T \int_\Omega m(z)\nabla D_0 h(z) \cdot \nabla u \, \mathrm{d}x \mathrm{d}t.$$

To determine the time evolution of the term $\int_\Omega u\tilde{u} \, \mathrm{d}x$, we assert that the Lipschitz function $\tilde{u}$ is admissible in the weak formulation (2.3) of the equation for u, thus yielding

$$\begin{aligned}
&-\int_\Omega u\tilde{u} \, \mathrm{d}x \Big|_{t=0}^{t=T} + \int_0^T \int_\Omega u\partial_t\tilde{u} \, \mathrm{d}x \mathrm{d}t \\
&\qquad = \int_0^T \int_\Omega \big(a(z)\nabla u \cdot \nabla \tilde{u} + m(z)\nabla D_0 h(z) \cdot \nabla \tilde{u}\big) \, \mathrm{d}x \mathrm{d}t,
\end{aligned}$$

where we used (A3.c). The admissibility of $\tilde{u}$ can be shown as follows: first exploit the regularity properties of $\nabla u \in L^2_{\mathrm{loc}}(I; L^2(\Omega))$, $u \in L^\infty_{\mathrm{loc}}(I; L^2(\Omega))$, which hold true by hypothesis resp. follow from (ENE) and the fact that $\mathscr{P}(z) \in L^1_{\mathrm{loc}}(I; L^1(\Omega))$. Assumption (A5), Gagliardo–Nirenberg interpolation applied to u and the estimate

$$m|\nabla D_0 h(z)| \leq \sqrt{m}\sqrt{\mathscr{P}(z)} \lesssim \sqrt{\mathscr{P}(z)}$$

then imply improved integrability of the flux term, namely for some $s = s(d) > 1$

$$A_{0j}(z)\nabla z_j = a(z)\nabla u + m(z)\nabla D_0 h(z) \in L^s_{\mathrm{loc}}(I; L^s(\Omega)).$$

With these bounds one can now use an approximation argument to show that, under the current hypotheses, eq. (2.3) can be extended in particular to Lipschitz functions $\varphi \in C^{0,1}(I \times \bar{\Omega})$.

Finally, using the test function $\varphi = u - \tilde{u} \in L^1_{\mathrm{loc}}(I; W^{1,1}(\Omega))$ in the weak equation (3.12) for $\tilde{u}$ gives

$$-\int_0^T \int_\Omega \partial_t\tilde{u}(u - \tilde{u}) \, \mathrm{d}x \mathrm{d}t = \int_0^T \int_\Omega a(\tilde{z})\nabla \tilde{u} \cdot \nabla(u - \tilde{u}) \, \mathrm{d}x \mathrm{d}t.$$

The asserted identity (3.13) is now obtained by adding up the above equations and rearranging appropriately the terms on the RHS. $\qquad \square$

The evolution inequality for our generalised distance is an immediate consequence of the previous two propositions.

Corollary 3.6. *Let $\alpha \in (0, \infty)$. We have*

$$\mathrm{Dist}^*_\alpha(z, \tilde{z}) \Big|_{t=0}^{t=T} \leq \int_0^T \int_\Omega \rho_\alpha \, \mathrm{d}x \mathrm{d}t, \tag{3.15}$$

where $\rho_\alpha := \rho^{(h)} + \alpha\rho^{(g)}$ *with* $\rho^{(h)}$, $\rho^{(g)}$ *given by* (3.11) *resp.* (3.14).

3.4. **Stability estimate.**

Proof of Theorem 2.8. Since (A6) implies (A6'), it suffices to prove the assertion for the case '$\varpi = 1$ (and thus, by hp., (A6))' and the case '$\varpi = 0$ and (A6')', henceforth referred to as Case $\varpi = 1$ resp. Case $\varpi = 0$.

We will show the following.

- Case $\varpi = 1$: if $\alpha \in (0, 1]$ is sufficiently small, and if $E = E(\alpha)$ and $N = N(E)$ are large enough, then for almost all $T \in (0, T^*)$

$$\mathrm{Dist}^*_\alpha(z(t, \cdot), \tilde{z}(t, \cdot))\Big|_{t=0}^{t=T} \lesssim_{E,N,\alpha} \int_0^T \mathrm{Dist}^*_\alpha(z, \tilde{z})\,\mathrm{d}t. \tag{3.16}$$

- Case $\varpi = 0$: if E, $N = N(E)$ and $\alpha \in [1, \infty)$ are chosen large enough (α possibly depending on E, N), then for a.a. $T \in (0, T^*)$ ineq. (3.16) holds true.

Once ineq. (3.16) has been established, we can invoke Gronwall's inequality to infer that for a.e. $T \in (0, T^*)$

$$\mathrm{Dist}^*_\alpha(z(T, \cdot), \tilde{z}(T, \cdot)) \leq \mathrm{Dist}^*_\alpha(z^{\mathrm{in}}, \widetilde{z^{\mathrm{in}}}) \exp(kT),$$

where $k = k(E, N, \alpha) > 0$. The estimates will also depend on the fixed constant $\iota > 0$, i.e. on the pointwise lower bound for $\min(\tilde{z})$. This dependency will only be indicated occasionally and for the sake of clarity.

In view of inequality (3.15) it suffices to show the pointwise bound

$$\rho_\alpha \lesssim_{E,N,\alpha} \mathrm{dist}^*_\alpha(z, \tilde{z}). \tag{3.17}$$

An elementary ingredient in the proof of this bound will be the coercivity properties of dist^*_α (see Prop. 3.2). We anticipate that referring to Prop. 3.2 will be the only instance, where the present proof makes use of the more specific form of the entropy density $h(u, c)$ assumed in (A1). Loosely speaking, besides the locally strict convexity ensuring (3.5), we will rely on a lower bound on the generalised distance of the form $\mathrm{dist}^*_\alpha(z, \tilde{z}) \gtrsim 1 + u^2$ for $|z| \gg 1$.

We will distinguish the cases $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ introduced on page 17, where, owing to the degeneracies of $\mathbb{M}(z)$ occurring when one of the concentrations vanishes, we further decompose the set $\mathcal{A}$ into

$$\mathcal{A}_+ := \{z' : \min(z') \geq \iota\} \cap \mathcal{A} \qquad \text{and} \qquad \mathcal{A}_0 := \{z' : \min(z') < \iota\} \cap \mathcal{A}, \tag{3.18}$$

where $\min(z') := \min\{z'_0, \ldots, z'_n\}$ for $z' = (z'_0, \ldots, z'_n) \in [0, \infty)^{1+n}$. This decomposition further serves to avoid regularity issues of h as $z_i \searrow 0$ for some $i \in \{0, \ldots, n\}$.

If $z \in (\mathcal{A}_+)^c$, we will make use of the following equivalent formula for $\rho^{(g)}$

$$\begin{aligned}
\rho^{(g)} = {} & -a(z)|\nabla u|^2 - a(\tilde{z})|\nabla \tilde{u}|^2 + a(z)\nabla u \cdot \nabla \tilde{u} + a(\tilde{z})\nabla u \cdot \nabla \tilde{u} \\
& - m(z)(\nabla u - \nabla \tilde{u}) \cdot \nabla D_0 h(z) \\
& + m(\tilde{z})(\nabla u - \nabla \tilde{u}) \cdot \nabla D_0 h(\tilde{z}).
\end{aligned}$$

Since, by hypothesis, $0 \leq m(z) \lesssim \varpi$ and $a(z) \gtrsim 1$, this implies that

$$\rho^{(g)} \leq -\tfrac{a(z)}{2}|\nabla u|^2 + C|a(z)\nabla u| + Cm(z)\varpi|\nabla D_0 h(z)|^2 + C. \tag{3.19}$$

Using this form in the case when $z \in (\mathcal{A}_+)^c$, we can avoid for instance issues due to $a(z)$ becoming singular as $u \to 0$ by using the bound (A5) on the energy flux.

Finally, note that $z \in \mathcal{A}$ implies $\xi^*(z) = 1$ and $D^k \xi^*(z) = 0 \ \forall k \in \mathbb{N}_+$, so that, if $z \in \mathcal{A}$, one has by formula (3.11)

$$
\begin{aligned}
\rho^{(h)} = &-\nabla D_i h(z) \cdot \mathbb{M}_{il}(z) \nabla D_l h(z) \\
&+ \nabla D_i h(\tilde{z}) \cdot M_{il}(z) \nabla D_l h(z) \\
&+ \nabla \Big(D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) \Big) \cdot \mathbb{M}_{il}(\tilde{z}) \nabla D_l h(\tilde{z}) \, \mathrm{d}x \mathrm{d}t \\
&- D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) R_i(\tilde{z}) \\
&+ (D_i h(z) - D_i h(\tilde{z})) R_i(z).
\end{aligned}
\tag{3.20}
$$

We are now ready to tackle the four cases.

Case $z \in \mathcal{A}_+$: in this case we have the control $\iota \leq z_i \leq E$ for all $i \in \{0, \dots, n\}$, and we need to show that $\rho_\alpha \lesssim |z - \tilde{z}|^2$. We therefore rewrite formula (3.20) as

$$
\begin{aligned}
\rho^{(h)} = &-\nabla (D_i h(z) - D_i h(\tilde{z})) \cdot \mathbb{M}_{il}(z) \nabla D_l h(z) \\
&+ \nabla \big(D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) \big) \cdot \mathbb{M}_{il}(\tilde{z}) \nabla D_l h(\tilde{z}) \\
&+ \big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) \big) R_i(\tilde{z}) \\
&+ (D_i h(z) - D_i h(\tilde{z}))(R_i(z) - R_i(\tilde{z})) \\
= &-\nabla (D_i h(z) - D_i h(\tilde{z})) \cdot \mathbb{M}_{il}(z) \nabla (D_l h(z) - D_l h(\tilde{z})) \\
&- \nabla (D_i h(z) - D_i h(\tilde{z})) \cdot (\mathbb{M}_{il}(z) - \mathbb{M}_{il}(\tilde{z})) \nabla D_l h(\tilde{z}) \\
&- \nabla \big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) \big) \cdot \mathbb{M}_{il}(\tilde{z}) \nabla D_l h(\tilde{z}) \\
&+ \big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j) \big) R_i(\tilde{z}) \\
&+ (D_i h(z) - D_i h(\tilde{z}))(R_i(z) - R_i(\tilde{z})).
\end{aligned}
\tag{3.21}
$$

Since, by hp. (A3.b), $\mathbb{M}(z) \geq \mathrm{diag}(m(z)\varpi, 0, \dots, 0) + \epsilon(\iota) \mathrm{diag}(0, 1, \dots, 1)$ for a suitable constant $\epsilon(\iota) > 0$, we have

$$
\begin{aligned}
\nabla (D_i h(z) - D_i h(\tilde{z})) \cdot \mathbb{M}_{il}(z) \nabla (D_l h(z) - D_l h(\tilde{z})) & \\
\geq m(z)\varpi |\nabla D_0 h(z) - \nabla D_0 h(\tilde{z})|^2 &+ \epsilon(\iota)\delta |\nabla D_c h(z) - \nabla D_c h(\tilde{z})|^2
\end{aligned}
$$

for any $\delta \in (0, 1]$. The auxiliary parameter $\delta = \delta(\alpha)$ will eventually be chosen small enough to be specified below. For $i \in \{1, \dots, n\}$ we observe that since $\iota \leq z_j \leq E$ for all $j \in \{0, \dots, n\}$, the triangle inequality yields

$$
\begin{aligned}
|\nabla D_i h(z) - \nabla D_i h(\tilde{z})| &\geq |\nabla \log(c_i) - \nabla \log(\tilde{c}_i)| - |\nabla \log(w_i(u)) - \nabla \log(w_i(\tilde{u}))| \\
&\geq \tfrac{1}{c_i} |\nabla c_i - \nabla \tilde{c}_i| - \big| \tfrac{1}{c_i} - \tfrac{1}{\tilde{c}_i} \big| |\nabla \tilde{c}_i| - |\nabla u - \nabla \tilde{u}| \tfrac{w_i'(u)}{w_i(u)} \\
&\quad - \big| \tfrac{w_i'(u)}{w_i(u)} - \tfrac{w_i'(\tilde{u})}{w_i(\tilde{u})} \big| |\nabla \tilde{u}| \\
&\geq \epsilon(E, \iota) |\nabla c - \nabla \tilde{c}| - C(E, \iota) |\nabla u - \nabla \tilde{u}| - C(E, \iota) |z - \tilde{z}|,
\end{aligned}
$$

where $\epsilon(E, \iota) > 0$ is some sufficiently small constant.

We next estimate, using the local Lipschitz continuity of $\mathbb{M}$,

$$
\begin{aligned}
|\nabla (D_i h(z) - D_i h(\tilde{z})) \cdot (\mathbb{M}_{il}(z) - \mathbb{M}_{il}(\tilde{z})) \nabla D_l h(\tilde{z})| &\lesssim_{E,\iota} |\nabla D h(z) - \nabla D h(\tilde{z})| |z - \tilde{z}| \\
&\lesssim_{E,\iota} |z - \tilde{z}| |\nabla z - \nabla \tilde{z}| + |z - \tilde{z}|^2.
\end{aligned}
$$

Before estimating the remaining terms, we compute for $f \in C^3_{\text{loc}}((0,\infty)^{1+n})$

$$\nabla\big(f(z) - f(\tilde{z}) - D_j f(\tilde{z})(z_j - \tilde{z}_j)\big) = D_j f(z)\nabla z_j - D_j f(\tilde{z})\nabla \tilde{z}_j - D_j f(\tilde{z})\nabla(z_j - \tilde{z}_j)$$
$$- D_{jk} f(\tilde{z})(z_j - \tilde{z}_j)\nabla \tilde{z}_k$$
$$= (D_j f(z) - D_j f(\tilde{z}))\nabla z_j - D_{jk} f(\tilde{z})(z_j - \tilde{z}_j)\nabla \tilde{z}_k$$
$$= (D_j f(z) - D_j f(\tilde{z}))\nabla(z_j - \tilde{z}_j)$$
$$+ [D_k f(z) - D_k f(\tilde{z}) - D_{jk} f(\tilde{z})(z_j - \tilde{z}_j)]\nabla \tilde{z}_k$$

and, using Taylor's theorem, for $k = 0, \ldots, n$

$$|D_k f(z) - D_k f(\tilde{z}) - D_{jk} f(\tilde{z})(z_j - \tilde{z}_j)| \lesssim_{E,\iota,f} |z - \tilde{z}|^2.$$

Letting $f(z) = D_i h(z)$, we infer since $h \in C^4_{\text{loc}}((0,\infty)^{1+n})$ (see hp. (A1)) that

$$|\nabla\big(D_i h(z) - D_i h(\tilde{z}) - D_{ij} h(\tilde{z})(z_j - \tilde{z}_j)\big)| \lesssim_{\iota,E} |z - \tilde{z}||\nabla z - \nabla \tilde{z}| + |z - \tilde{z}|^2.$$

Using the previous bounds to estimate the RHS of (3.21), recalling also hp. (A2.ii), and applying Young's inequality and an absorption argument, we thus infer for suitable $\epsilon(\iota, E) > 0$

$$\rho^{(h)} \leq -\epsilon(\iota, E)\delta|\nabla c - \nabla \tilde{c}|^2 - m(z)\varpi|\nabla D_0 h(z) - \nabla D_0 h(\tilde{z})|^2$$
$$+ C_1(E,\iota)\delta|\nabla u - \nabla \tilde{u}|^2 + C(\delta, E, \iota)|z - \tilde{z}|^2.$$

On the other hand, using the fact that $z \mapsto a(z)$ is locally Lipschitz continuous in $(0,\infty)^{1+n}$, we deduce from eq. (3.14) for suitable $\epsilon_1 > 0$ and $C_2 < \infty$ (independent of E, ι)

$$\rho^{(g)} \leq -\epsilon_1|\nabla u - \nabla \tilde{u}|^2 + C_2 m(z)\varpi|\nabla D_0 h(z) - \nabla D_0 h(\tilde{z})|^2 + C(E,\iota)|z - \tilde{z}|^2,$$

where we used the fact that $0 \leq m(z) \lesssim \varpi$.

If $\varpi = 1$, we choose $\alpha \in (0, 1]$ small enough such that $\alpha C_2 \leq 1$ and subsequently $\delta = \delta(\alpha, E, \iota)$ sufficiently small such that $\delta C_1(E,\iota) \leq \alpha\epsilon_1$. We may then conclude that

$$\rho_\alpha = \rho^{(h)} + \alpha\rho^{(g)} \leq C(\alpha, E, \iota)|z - \tilde{z}|^2.$$

Let us emphasise that the smallness condition of α is independent of E.

If instead $\varpi = 0$, we choose for given[4] $\alpha \in [1,\infty)$ the parameter $\delta = \delta(\alpha, E, \iota)$ small enough such that $\delta C_1(E,\iota) \leq \epsilon_1\alpha$, and obtain as before

$$\rho_\alpha \leq C(\alpha, E, \iota)|z - \tilde{z}|^2.$$

Case $z \in \mathcal{A}_0$: in this case, we have no lower bound on z_i away from zero, but since $\min(z) \leq \iota$ and $\min(\tilde{z}) \geq 2\iota$, we know that $|z - \tilde{z}| \geq \iota$. By Prop. 3.2 (cf. (3.5)) it thus suffices to prove that $\rho_\alpha \lesssim_E 1$.

Recalling (3.20) and (A2.i), we estimate

$$\rho^{(h)} \leq -\nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)$$
$$+ \nabla D_i h(\tilde{z}) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)$$
$$+ \nabla\big(D_{ij} h(\tilde{z})(z_j - \tilde{z}_j)\big) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z})$$
$$+ C(E)$$
$$\leq -\mathscr{P}(z) + C(E)\sqrt{\mathscr{P}(z)} + C|\nabla u| + C(E),$$

where the last step uses hp. (A4.a) and hp. (A4.b). Hence,

$$\rho^{(h)} \leq -\tfrac{1}{2}\mathscr{P}(z) + C(E)|\nabla u| + C(E).$$

---

[4]In the case $\varpi = 0$, it suffices to restrict α to the range $1 \leq \alpha < \infty$.

Next, by ineq. (3.19) and hp. (A4.b),

$$\rho^{(g)} \leq -\frac{a(z)}{2}|\nabla u|^2 + C(\delta_1, E) + \frac{\delta_1}{4}\mathscr{P}(z) + Cm(z)\varpi|\nabla D_0 h(z)|^2 \qquad (3.22)$$

for any $\delta_1 > 0$.

If $\varpi = 1$, we let $\delta_1 = 1$ and use the estimate $m(z)\varpi|\nabla D_0 h(z)|^2 \leq \mathscr{P}(z)$, which follows from hp. (A3), to see that after possibly decreasing $\alpha \in (0, 1]$ we have $\rho_\alpha \leq -\alpha\frac{a(z)}{2}|\nabla u|^2 + C(E)|\nabla u| + C(E) \leq -\alpha\frac{a(z)}{4}|\nabla u|^2 + C(\alpha, E)$.

If $\varpi = 0$, we choose $\delta_1 = \frac{1}{\alpha} \leq 1$. Then

$$\rho_\alpha \leq -\frac{1}{2}\mathscr{P}(z) - \alpha\frac{a(z)}{2}|\nabla u|^2 + \frac{a(z)}{4}|\nabla u|^2 + C(\alpha, E) + \frac{1}{4}\mathscr{P}(z)$$
$$\lesssim_{\alpha, E} 1.$$

<u>Case $z \in \mathcal{B}$</u> : in this case derivatives of ξ^* do in general not vanish, but we know that $1 \ll E < |z|_1 < E^N$. We estimate

$$\begin{aligned}
\rho^{(h)} &= -\nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
&\quad + \nabla(D_i(\xi^*(z)z_j)D_j h(\tilde{z})) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
&\quad + \nabla\Big(D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)\Big) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z}) \\
&\quad - D_{ij}h(\tilde{z})(\xi^*(z)z_j - \tilde{z}_j)R_i(\tilde{z}) \\
&\quad + D_i h(z)R_i(z) - D_j h(\tilde{z})D_i(\xi^*(z)z_j)R_i(z) \\
&\leq -\mathscr{P}(z) \\
&\quad + C|\nabla D_i(\xi^*(z)z_j) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)| \\
&\quad + C(E, N)\sqrt{\mathscr{P}(z)} + C(E, N)|\nabla u| \qquad \text{(by hp. (A4.b), (A4.a))} \\
&\quad + C(E, N) \qquad \text{(using hp. (A2.i))} \\
&\leq -\frac{1}{2}\mathscr{P}(z) \\
&\quad + C|\nabla D_i(\xi^*(z)z_j) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)| \\
&\quad + C(E, N)|\nabla u| + C(E, N).
\end{aligned}$$

In order to estimate the term $|\nabla D_i(\xi^*(z)z_j) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)|$, we observe that

$$\begin{aligned}
|\nabla D_i(\xi^*(z)z_j)| &\lesssim |D\xi^*(z)||\nabla z| + |D^2\xi^*(z)z||\nabla z| \\
&\lesssim \frac{1}{N|z|_1}|\nabla z| \qquad \text{(by (t3))}.
\end{aligned} \qquad (3.23)$$

We first consider the case $\varpi = 1$. Then (A6) is at our disposal, which yields using (3.23)

$$|\nabla D_i(\xi^*(z)z_j)||\mathbb{M}_{il}(z)\nabla D_l h(z)| \lesssim \frac{1}{N|z|_1}|\nabla z||\sum_j A_{ij}(z)\nabla z_j| \lesssim \frac{1}{N}\mathscr{P}(z) + \frac{1}{N}|\nabla u|^2.$$

Thus, choosing $N = N(\min\{1, \alpha\})$ sufficiently large, we infer

$$\rho^{(h)} \leq -\frac{1}{4}\mathscr{P}(z) + \min\{1, \alpha\}\frac{a(z)}{8}|\nabla u|^2 + C(E, N, \alpha).$$

Next, simiarly as in (3.22), we estimate

$$\rho^{(g)} \leq -\frac{a(z)}{2}|\nabla u|^2 + C(E, N) + \frac{1}{8}\mathscr{P}(z) + C_4 m(z)\varpi|\nabla D_0 h(z)|^2.$$

Decreasing $\alpha \in (0, 1]$, if necessary, to ensure that $\alpha C_4 m(z)\varpi|\nabla D_0 h(z)|^2 \leq \frac{1}{8}\mathscr{P}(z)$, we obtain $\rho_\alpha \lesssim_{N, E} 1$.

It remains to consider the case where $\varpi = 0$ and (A6') are fulfilled. Since $\underline{u} := \inf u^{\text{in}} = \inf \tilde{u}(0, \cdot) \geq 2\iota$, Lemma 6.4 yields $\inf u \geq \underline{u} > 0$. By hp. (A6'), we infer for all $0 \leq i \leq n$

$$|\nabla z||\mathbb{M}_{il}(z)\nabla D_l h(z)| \leq C|z|_1 \mathscr{P}(z) + C(E, N, \underline{u})|\nabla u|^2.$$

Thus, recalling ineq. (3.23), we can estimate for N large enough

$$|\nabla D_i(\xi^*(z)z_j) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z)| \leq \tfrac{1}{4}\mathscr{P}(z) + C(E, N, \underline{u})|\nabla u|^2 \qquad (3.24)$$

to infer

$$\rho^{(h)} \leq -\tfrac{1}{4}\mathscr{P}(z) + C_1(E, N, \underline{u})|\nabla u|^2 + C(E, N).$$

Next, since $\varpi = 0$, ineq. (3.19) yields

$$\rho^{(g)} \leq -\tfrac{a(z)}{2}|\nabla u|^2 + C(E, N)\sqrt{\mathscr{P}(z)} + C$$
$$\leq -\tfrac{a(z)}{2}|\nabla u|^2 + \tfrac{1}{8\alpha}\mathscr{P}(z) + C(\alpha, E, N).$$

Increasing $\alpha = \alpha(E, N, \underline{u})$, if necessary, to ensure that $\alpha\tfrac{a(z)}{2} \geq C_1(E, N, \underline{u})$, we conclude

$$\rho_\alpha \leq -\tfrac{1}{8}\mathscr{P}(z) + C(\alpha, E, N, \underline{u}).$$

Case $z \in \mathcal{C}$: in this case, $\xi^*(z) = 0$ and $D^k\xi^*(z) = 0$ for all $k \in \mathbb{N}_+$. Thus

$$\begin{aligned}
\rho^{(h)} &= -\nabla D_i h(z) \cdot \mathbb{M}_{il}(z)\nabla D_l h(z) \\
&\quad - \nabla(D_{ij}h(\tilde{z})\tilde{z}_j) \cdot \mathbb{M}_{il}(\tilde{z})\nabla D_l h(\tilde{z}) \\
&\quad + D_{ij}h(\tilde{z})\tilde{z}_j R_i(\tilde{z}) \\
&\quad + D_i h(z)R_i(z) \\
&\leq -\mathscr{P}(z) + C,
\end{aligned}$$

where we used hp. (A2.i).

If $\varpi = 1$, we have thanks to (3.19) and hp. (A3), (A5)

$$\rho^{(g)} \leq -\tfrac{a(z)}{2}|\nabla u|^2 + C_4\mathscr{P}(z) + u^2 + C,$$

where C_4 is independent of E, N. Hence, after possibly decreasing $\alpha \in (0, 1]$ to ensure that $\alpha C_4 \leq 1$, we find

$$\rho_\alpha \leq C + \alpha u^2 \lesssim_\alpha \text{dist}_\alpha^*(z, \tilde{z}),$$

where the second step follows from (3.7) (after choosing $E = E(\alpha)$ large enough).

If $\varpi = 0$, we estimate using again (3.19) and hp. (A5)

$$\rho^{(g)} \leq -\tfrac{a(z)}{2}|\nabla u|^2 + C(1 + u)\sqrt{\mathscr{P}(z)} + C$$
$$\leq -\tfrac{a(z)}{2}|\nabla u|^2 + \tfrac{1}{2\alpha}\mathscr{P}(z) + C(\alpha)u^2 + C(\alpha),$$

and infer

$$\rho_\alpha \leq -\tfrac{1}{2}\mathscr{P}(z) - \alpha\tfrac{a(z)}{2}|\nabla u|^2 + C(\alpha)u^2 + C(\alpha)$$
$$\lesssim_\alpha \text{dist}_\alpha^*(z, \tilde{z}).$$

The second step follows from the coercivity property (3.6) and the fact that $\alpha \geq 1$.

This proves the bound (3.17) and thus completes the proof of Theorem 2.8. □

4. Strong Entropy dissipation property

Proof of Proposition 2.9. We first establish (ED.s) for $s = 0$ and a.e. $t = T \in (0, T^*)$, that is, we first prove (ED). In a second step (see page 34), we point out how to extend the result to a.e. $0 < s < t < T^*$.

Case 1: $s = 0, t = T \in (0, T^*)$.

We consider for a small parameter $\delta > 0$ the regularised entropy density

$$h_\delta(u, c) = \delta u + h(u, c).$$

The additive term δu serves to ensure coercivity, since the original density $h(z)$ may in general allow for cancellations at infinity reflecting the coupling between concentrations and energy component. In particular, for every $L \in \mathbb{N}$ the sublevel set $\{z \in \mathbb{R}_{\geq 0}^{1+n} : h_\delta(z) \leq L\}$ is bounded. This coercivity property easily follows from the lower bound

$$h(z) \geq -\hat{\sigma}(u) + \epsilon_* \sum_{i=1}^n c_i \log c_i - Cu^\nu - C, \tag{4.1}$$

valid for suitable $\nu \in [0, 1)$, $\epsilon_* > 0$ (see Lemma 6.3), together with the sublinearity of the increasing function $\hat{\sigma}(u)$ as $u \to \infty$ (see (h2)).

In order to define an admissible truncation function, we consider as in [23, Proof of Prop. 5] for $L \geq 2$ an auxiliary function $\theta_L \in C^\infty(\mathbb{R})$ satisfying $\theta_L(s) = s$ for $|s| \leq L$, $0 \leq \theta_L' \leq 1$,

$$|\theta_L''(s)| \lesssim \frac{C}{1 + |s| \log(|s| + e)} \tag{4.2}$$

for all $s \in \mathbb{R}$ and $\theta_L'(s) = 0$ for $|s| \geq L^C$ for some sufficiently large constant $C \geq 2$, which is kept fixed throughout the proof.

To derive the entropy dissipation inequality, we would like to choose the truncation function $\theta_L(h_\delta(\cdot))$ and the test function $\psi \equiv 1$ in the renormalised formulation satisfied by z, and then let $L \to \infty$ and subsequently $\delta \to 0$. Since derivatives of h_δ are in general not bounded as $u \searrow 0$ or $c_i \searrow 0$, further regularisation is required. We let

$$h_{\delta,\varepsilon}(u, c) = h_\delta(z) + \hat{\sigma}(u) - \hat{\sigma}(u + \varepsilon),$$

abbreviate for $z = (u, c_1, \ldots, c_n)$

$$z^{\tilde{\varepsilon}} := (u, c_1 + \tilde{\varepsilon}, \ldots, c_n + \tilde{\varepsilon}), \quad \tilde{\varepsilon} \in (0, 1],$$

and then consider the function $z \mapsto h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \in C^2([0, \infty)^{1+n})$. Thanks to (4.1), it is easy to see that for fixed $\delta \in (0, 1]$ sublevel sets of $z \mapsto h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})$ are bounded; in fact

$$h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \geq \frac{\delta}{2}u + \epsilon_* \sum_{i=1}^n c_i \log_+ c_i - C_\delta \tag{4.3}$$

for $\delta, \varepsilon, \tilde{\varepsilon} \in (0, 1]$. Hence, the C^2-function $\xi(z) := \theta_L(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))$ has compactly supported derivative $D\xi$, and is thus an admissible truncation in eq. (2.1) (cf. Remark 2.2). This, combined with the choice

$\psi \equiv 1$ in eq. (2.1), yields for a.e. $T < T^*$

$$
\begin{aligned}
LHS &:= \int_\Omega \theta_L(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) \, \mathrm{d}x \Big|_{t=0}^{t=T} \\
&= -\int_0^T \int_\Omega \theta'_L(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) D_{ij} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) \nabla z_j \cdot A_{ik}(z) \nabla z_k \, \mathrm{d}x\mathrm{d}t \\
&\quad -\int_0^T \int_\Omega \theta''_L(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) D_j h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) D_i h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) \nabla z_j \cdot A_{ik}(z) \nabla z_k \, \mathrm{d}x\mathrm{d}t \\
&\quad +\int_0^T \int_\Omega \theta'_L(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) D_i h(z^{\tilde\varepsilon}) R_i(z) \, \mathrm{d}x\mathrm{d}t \\
&=: I + II + III,
\end{aligned}
\tag{4.4}
$$

where we recall the summation convention (see Notations 1.5). In the reaction term we have used the fact that $D_i h_{\delta,\varepsilon}(z) = D_i h(z)$ whenever $i \neq 0$ together with $R_0 \equiv 0$.

We will establish the asserted inequality (ED) by taking the $\liminf$ of the LHS and the $\limsup$ of the RHS of the above equation (4.4) as $\tilde\varepsilon \to 0$, $L \to \infty$ and $\delta, \varepsilon \to 0$, in the stated order. We perform the corresponding limits separately in LHS and in each of the three terms I, II, III. Below we use, without explicit reference, the following basic properties satisfied under the hypotheses of Model **(M0)**, see Lemma 6.1:

$$
a(z)|\nabla u|^2 \sim |\nabla u|^2,
$$

$$
\mathscr{P}(z) \gtrsim \sum_{i=1}^n |\nabla \sqrt{c_i}|^2 + |\sqrt\gamma \, \nabla u|^2 + |\sqrt m \nabla D_0 h(z)|^2,
$$

where $\gamma(u,c) = -\hat\sigma''(u) - \sum_{l=1}^n \frac{w_l''(u)}{w_l(u)} c_l$,

$$
|A_{ik}(z)\nabla z_k| \lesssim \sqrt{c_i}\sqrt{\mathscr{P}(z)} \qquad \text{for } i \geq 1,
$$

$$
|A_{0k}(z)\nabla z_k| \lesssim |\nabla u| + \sqrt m \sqrt{\mathscr{P}(z)}.
$$

LHS:

The limit $\tilde\varepsilon \to 0$ of the LHS is immediate due to the boundedness of θ_L, and yields

$$
\int_\Omega \theta_L(h_{\delta,\varepsilon}(z)) \, \mathrm{d}x \Big|_{t=0}^{t=T}.
$$

We next take

$$
\liminf_{\varepsilon,\delta\to\infty} \liminf_{L\to\infty}
$$

of the last expression using a combination of the dominated convergence theorem and Fatou's lemma:

At initial time, we estimate using the lower and upper bounds on h in Lemma 6.3

$$
|\theta_L(h_{\delta,\varepsilon}(z^{\mathrm{in}}))| \leq |h_{\delta,\varepsilon}(z^{\mathrm{in}})| \lesssim u^{\mathrm{in}} + |\hat\sigma_-(u^{\mathrm{in}})| + \sum_{i=0}^n c_i^{\mathrm{in}} \log_+ c_i^{\mathrm{in}} + 1.
$$

We can hence use dominated convergence to deduce that, as $L \to \infty$ and $\delta, \varepsilon \to 0$,

$$
\int_\Omega \theta_L(h_{\delta,\varepsilon}(z^{\mathrm{in}})) \, \mathrm{d}x \to \int_\Omega h(z^{\mathrm{in}}) \, \mathrm{d}x.
$$

To deal with the integral at time $t = T$, we first observe that, thanks to the regularity $u \in L^\infty_{\mathrm{loc}}([0, T^*), L^1(\Omega))$ and (4.1), the negative part of $\theta_L(h_{\delta,\varepsilon}(z(T, \cdot)))$ is controlled pointwise in x, for a.e. $T \in (0, T^*]$, by

an integrable function, uniformly in L, δ, ε, and its integral can thus be shown to converge in the same way as the term at initial time. For the positive part of $\theta_L(h_{\delta,\varepsilon}(z(T, \cdot)))$ we use Fatou's lemma:

$$\int_\Omega \max\{h_{\delta,\varepsilon}(z(T, \cdot)), 0\}\,\mathrm{d}x \le \liminf_{L \to \infty} \int_\Omega \max\{\theta_L(h_{\delta,\varepsilon}(z(T, \cdot))), 0\}\,\mathrm{d}x,$$

$$\int_\Omega \max\{h(z(T, \cdot)), 0\}\,\mathrm{d}x \le \liminf_{\delta,\varepsilon \to 0} \int_\Omega \max\{h_{\delta,\varepsilon}(z(T, \cdot)), 0\}\,\mathrm{d}x.$$

In combination, we infer

$$\int_\Omega h(z)\,\mathrm{d}x\bigg|_{t=0}^{t=T} \le \liminf_{\delta,\varepsilon \to 0} \liminf_{L \to \infty} \lim_{\tilde\varepsilon \to 0} \int_\Omega \theta_L(h_{\delta,\varepsilon}(z^{\tilde\varepsilon}))\,\mathrm{d}x\bigg|_{t=0}^{t=T}.$$

Diffusive dissipation term I:
We assert that

$$\limsup_{\delta,\varepsilon \to 0} \limsup_{L \to \infty} \limsup_{\tilde\varepsilon \to 0} \left(-\int_0^T \int_\Omega \theta_L'(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) \nabla z_j \cdot D_{ji} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) A_{ik}(z) \nabla z_k\,\mathrm{d}x\mathrm{d}t \right)$$

can be bounded above by the non-positive term

$$-\int_0^T \int_\Omega \nabla z_j \cdot D_{ji} h(z) A_{ik}(z) \nabla z_k\,\mathrm{d}x\mathrm{d}t.$$

To show this, we will mainly rely on the dominated convergence theorem. We therefore start by listing several uniform pointwise estimates on the terms involved.

We first estimate for i, j fixed the term

$$p_{ij} := \theta_L'(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) D_{ji} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) \nabla z_j \cdot \sum_k A_{ik}(z) \nabla z_k,$$

where we recall that $0 \le \theta_L' \le 1$.

Case $i, j \ge 1$: in this case $D_{ij} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) = \frac{1}{c_i + \tilde\varepsilon} \delta_{ij}$ and thus

$$|p_{ij}| \le |\tfrac{1}{c_i + \tilde\varepsilon} \nabla c_i \cdot \sum_k A_{ik}(z) \nabla z_k| \lesssim \tfrac{1}{c_i} |\nabla c_i| \sqrt{c_i} \sqrt{\mathscr{P}(z)} \lesssim \mathscr{P}(z).$$

Case $i \ge 1, j = 0$: observing that $-D_{i0} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) = \frac{w_i'}{w_i} \lesssim \sqrt{\frac{-w_i''}{w_i}}$ (cf. (2.10)), we find

$$|p_{ij}| \le |\tfrac{w_i'}{w_i} \nabla u \cdot \sum_k A_{ik}(z) \nabla z_k| \lesssim |\sqrt{\tfrac{-w_i''}{w_i}} c_i \nabla u| \sqrt{\mathscr{P}(z)} \lesssim \mathscr{P}(z).$$

Case $i = 0, j \ge 0$: we split the sum over k in the definition of p_{0j} into two parts:

$$\sum_k A_{0k}(z) \nabla z_k = a(z) \nabla u + m(z) \nabla D_0 h(z)$$

and split p_{0j} accordingly into $p_{0j} = p_{0j}^{(0)} + p_{0j}^{(1)}$.

For $j \ge 1$ we estimate as above

$$|p_{0j}^{(0)}| \le |\tfrac{w_j'}{w_j} \nabla c_j \cdot a(z) \nabla u| \lesssim |\sqrt{\tfrac{-w_j''}{w_j}} c_j \nabla u| |\nabla \sqrt{c_j}| \lesssim \mathscr{P}(z).$$

For $j = 0$, we have $p_{00}^{(0)} = \theta_L'(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) D_{00} h_{\delta,\varepsilon}(z^{\tilde\varepsilon}) a(z) |\nabla u|^2 \ge 0$, and since in the equation it comes with a minus sign, its integral can easily be handled using Fatou's lemma.

It remains to estimate the part $p_{0j}^{(1)}$. To deal with the limit $\tilde{\varepsilon} \to 0$ (for finite L), we estimate

$$|\nabla D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \cdot m(z) \nabla D_0 h(z)| \lesssim \sqrt{m} |\nabla D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})| \sqrt{\mathscr{P}(z)}$$

and for $|z^{\tilde{\varepsilon}}| \leq C(L, \delta)$

$$\sqrt{m} |\nabla D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})| \leq |D_{00} h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \nabla u| + \sum_{j=1}^{n} |D_{0j} h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \nabla c_j|$$

$$\lesssim_{L,\delta,\varepsilon} |\nabla u| + \sqrt{\mathscr{P}(z)}.$$

Thus the limit $\tilde{\varepsilon} \to 0$ can be handled using dominated convergence.

We can now let $\tilde{\varepsilon} = 0$ and compute

$$\nabla D_0 h_{\delta,\varepsilon}(z) \cdot m(z) \nabla D_0 h(z)$$
$$= m(z) |\nabla D_0 h(z)|^2 - (\hat{\sigma}''(u+\varepsilon) - \hat{\sigma}''(u)) \nabla u \cdot m(z) \nabla D_0 h(z).$$

Note that the first term on the RHS is bounded above by $\mathscr{P}(z) \in L^1(\Omega_T)$. Concerning the second term, we estimate

$$|(\hat{\sigma}''(u+\varepsilon) - \hat{\sigma}''(u)) \nabla u \cdot m(z) \nabla D_0 h(z)|$$
$$\lesssim \sqrt{m} |(\hat{\sigma}''(u+\varepsilon) - \hat{\sigma}''(u)) \nabla u| \sqrt{\mathscr{P}(z)}$$
$$\lesssim |(1 + \sqrt{-\hat{\sigma}''(u)}) \nabla u| \sqrt{\mathscr{P}(z)}$$
$$\lesssim \mathscr{P}(z) + |\nabla u|^2,$$

where in the penultimate step we have used hypothesis (2.12).

Combining the above estimates allows to take the successive limits

$$\limsup_{\varepsilon,\delta \to 0} \limsup_{L \to \infty} \limsup_{\tilde{\varepsilon} \to 0} \dots$$

as above, thus yielding the asserted inequality for term I.

Remainder gradient term II:
We will show that

$$\limsup_{L \to \infty} \limsup_{\tilde{\varepsilon} \to 0} \left(-\int_0^T \int_\Omega \theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})) D_j h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \nabla z_j \cdot D_i h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) A_{ik}(z) \nabla z_k \, \mathrm{d}x \mathrm{d}t \right) \leq 0. \quad (4.5)$$

As in the previous paragraph, the main task is to obtain uniform pointwise estimates on the terms involved, where here we can afford a dependence of our estimates on δ and ε. We introduce for i, j fixed the term

$$q_{ij} = -\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})) D_j h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \nabla z_j \cdot D_i h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \sum_k A_{ik}(z) \nabla z_k$$

and observe that, by (4.2) and (4.3), for $L \geq L_0(\delta)$ (henceforth to be assumed)

$$|\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))| \lesssim_\delta \frac{1}{1 + (u + \sum_i c_i \log_+(c_i)) \log(|z|_1 + \mathrm{e})}. \quad (4.6)$$

Case $i, j \geq 1$: in this case,

$$|q_{ij}| \lesssim |\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))| \left| \log\left(\frac{c_i + \tilde{\varepsilon}}{w_i(u)}\right) \right| \left| \log\left(\frac{c_j + \tilde{\varepsilon}}{w_j(u)}\right) \right| \sqrt{c_i} \sqrt{c_j} |\nabla \sqrt{c_j}| \sqrt{\mathscr{P}(z)}.$$

Estimating

$$\sqrt{c_i} \left| \log\left(\frac{c_i + \tilde{\varepsilon}}{w_i(u)}\right) \right| \lesssim \sqrt{c_i} \log_+(c_i) + \sqrt{w_i(u)} \log_+(w_i(u)) + 1,$$

we find, using (h2) and (4.6),

$$|\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))| \sqrt{c_i}| \log\left(\tfrac{c_i+\tilde{\varepsilon}}{w_i(u)}\right)| \sqrt{c_j}| \log\left(\tfrac{c_j+\tilde{\varepsilon}}{w_j(u)}\right)|$$
$$\lesssim |\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))|(\sqrt{c_i}\log_+(c_i) + \sqrt{u} + 1)(\sqrt{c_j}\log_+(c_j) + \sqrt{u} + 1) \lesssim_\delta 1.$$

Thus, $|q_{ij}| \lesssim_\delta \mathscr{P}(z)$.

Case $i \geq 1, j = 0$: here,

$$|q_{ij}| \lesssim \sqrt{c_i}| \log\left(\tfrac{c_i+\tilde{\varepsilon}}{w_i(u)}\right)||D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})\nabla u|\sqrt{\mathscr{P}(z)} \, |\theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))|.$$

In view of the factor $\sqrt{c_i}$, this shows that $|q_{ij}| \leq C(L,\delta,\varepsilon)|\nabla u|\sqrt{\mathscr{P}(z)}$ uniformly in $\tilde{\varepsilon}$, allowing us to infer by dominated convergence

$$\limsup_{\tilde{\varepsilon}\to 0} \int_0^T \int_\Omega q_{ij} \, \mathrm{d}x\mathrm{d}t \leq - \int_0^T \int_\Omega \theta_L''(h_{\delta,\varepsilon}(z))D_0 h_{\delta,\varepsilon}(z)\nabla u \cdot \log\left(\tfrac{c_i}{w_i(u)}\right) A_{ik}(z)\nabla z_k \, \mathrm{d}x\mathrm{d}t.$$

Since $\varepsilon > 0$ and $\tfrac{w_l'(u)}{w_l(u)} \lesssim \sqrt{\tfrac{-w_l''(u)}{w_l(u)}}$, we have the rough bound

$$|D_0 h_{\delta,\varepsilon}(z)\nabla u| \lesssim_\varepsilon |\nabla u| + \sum_l \sqrt{c_l}\sqrt{\gamma}|\nabla u|$$
$$\lesssim |\nabla u| + \sum_l \sqrt{c_l}\sqrt{\mathscr{P}(z)}. \tag{4.7}$$

Moreover,

$$| \log\left(\tfrac{c_i}{w_i(u)}\right) A_{ik}(z)\nabla z_k| \lesssim (\log_+(c_i) + 1 + \log_+(w_i(u)))\sqrt{c_i}\sqrt{\mathscr{P}(z)},$$

and hence

$$|\theta_L''(h_{\delta,\varepsilon}(z))D_0 h_{\delta,\varepsilon}(z)\nabla u \cdot D_i h_{\delta,\varepsilon}(z)A_{ik}(z)\nabla z_k| \lesssim_\varepsilon \mathscr{P}(z) + |\nabla u|^2.$$

Case $i = 0, j \geq 0$: using the fact that for $|z|_1 \leq C(L,\delta)$

$$|D_j h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})\nabla z_j||D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})||a(z)\nabla u + m(z)\nabla D_0 h(z)| \lesssim_{L,\delta,\varepsilon} \mathscr{P}(z) + |\nabla u|^2,$$

one can take the limit $\tilde{\varepsilon} \to 0$

$$\limsup_{\tilde{\varepsilon}\to 0} \big(- \int_0^T \int_\Omega \theta_L''(h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}))D_j h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}})\nabla z_j \cdot D_0 h_{\delta,\varepsilon}(z^{\tilde{\varepsilon}}) \sum_k A_{0k}(z)\nabla z_k \, \mathrm{d}x\mathrm{d}t\big)$$
$$\leq - \int_0^T \int_\Omega \theta_L''(h_{\delta,\varepsilon}(z))D_j h_{\delta,\varepsilon}(z)\nabla z_j \cdot D_0 h_{\delta,\varepsilon}(z) \sum_k A_{0k}(z)\nabla z_k \, \mathrm{d}x\mathrm{d}t.$$

To obtain L-uniform bounds of the integrand on the RHS, we split

$$\sum_k A_{0k}(z)\nabla z_k = a(z)\nabla u + m(z)\nabla D_0 h(z).$$

Using (4.7) we have for $j \geq 1$

$$|\nabla z_j D_j h_{\delta,\varepsilon}(z)||D_0 h_{\delta,\varepsilon}(z)a(z)\nabla u|$$
$$\lesssim_\varepsilon \sqrt{c_j}(\log_+(c_j) + 1 + \log_+(w_j(u)))(|\nabla u|\sqrt{\mathscr{P}(z)} + \sum_l \sqrt{c_l}\mathscr{P}(z)),$$

and for $j = 0$

$$|\nabla u D_0 h_{\delta,\varepsilon}(z)||D_0 h_{\delta,\varepsilon}(z)a(z)\nabla u| \lesssim_\epsilon \big(|\nabla u| + \sqrt{|c|_1}\sqrt{\mathscr{P}(z)}\big)^2 \lesssim |\nabla u|^2 + |c|_1\mathscr{P}(z).$$

It remains to consider the term involving $m(z)$. We estimate for $j \geq 1$

$$|D_j h_{\delta,\varepsilon}(z) \nabla z_j \cdot D_0 h_{\delta,\varepsilon}(z)||m(z) \nabla D_0 h(z)|$$
$$\lesssim_\varepsilon \sqrt{c_j}(\log_+(c_j) + 1 + \log_+(w_j(u)))(1 + \tfrac{\sum_l c_l}{1+u})\sqrt{m}\mathscr{P}(z),$$

where we used the fact that $\frac{w_l'(u)}{w_l(u)} \lesssim \frac{1}{1+u}$ for all l allowing us to estimate

$$|D_0 h_{\delta,\varepsilon}(z)| \lesssim_\varepsilon 1 + \tfrac{\sum_{l=1}^n c_l}{1+u}.$$

Similarly, for $j = 0$ we estimate, using also the bound (4.7),

$$|D_0 h_{\delta,\varepsilon}(z) \nabla u \cdot D_0 h_{\delta,\varepsilon}(z)||m(z) \nabla D_0 h(z)|$$
$$\lesssim_\varepsilon (1 + \tfrac{\sum_{l'=1}^n c_{l'}}{1+u})\big(|\nabla u| + \sum_{l=1}^n \sqrt{c_l}\sqrt{\mathscr{P}(z)}\big)\sqrt{m}\sqrt{\mathscr{P}(z)}.$$

Thus, recalling the conditions (2.11), (2.12) on $m(z)$, we infer the L-uniform bound

$$|\theta_L''(h_{\delta,\varepsilon}(z))D_j h_{\delta,\varepsilon}(z) \nabla z_j \cdot D_0 h_{\delta,\varepsilon}(z) \sum_k A_{0k}(z) \nabla z_k| \lesssim_{\delta,\varepsilon} \mathscr{P}(z) + |\nabla u|^2.$$

Combining the above estimates, we can take the limits

$$\limsup_{L\to\infty} \limsup_{\tilde\varepsilon\to 0} \ldots$$

of term II and obtain ineq. (4.5) by the pointwise convergence $\theta_L''(s) \to 0$ as $L \to \infty$.

Reactions III:
Concerning the reaction term III, we first need to take care of the fact that $D_i h(z)$ is unbounded near $c_i = 0$. By (4.3), we have $|z|, |z^{\tilde\varepsilon}| \leq C(L, \delta)$ unless $\theta_L'(h_{\delta,\varepsilon}(z^{\tilde\varepsilon})) = 0$. Using the local ϵ_0-Hölder regularity of R_i, we then compute for $|z^{\tilde\varepsilon}| \leq C(L, \delta)$

$$D_i h(z^{\tilde\varepsilon})R_i(z) = D_i h(z^{\tilde\varepsilon})[R_i(z) - R_i(z^{\tilde\varepsilon})] + D_i h(z^{\tilde\varepsilon})R_i(z^{\tilde\varepsilon})$$
$$\leq C(L, \delta)(1 + \sum_{i=1}^n \log(c_i + \tilde\varepsilon))\tilde\varepsilon^{\epsilon_0} + D_i h(z^{\tilde\varepsilon})R_i(z^{\tilde\varepsilon}).$$

The first term in the last line converges, as $\tilde\varepsilon \to 0$, uniformly to zero on the set $\{|z^{\tilde\varepsilon}| \leq C(L, \delta)\}$. The second term is non-positive by hp. (A2.i). Thus, since $\theta_L' \geq 0$ and $\lim_{L\to\infty} \theta_L'(s) = 1$ for all $s \in \mathbb{R}$ we can use Fatou's lemma to infer

$$\limsup_{L\to\infty} \limsup_{\tilde\varepsilon\to 0} \int_0^T \int_\Omega \theta_L'(h_{\delta,\varepsilon}(z^{\tilde\varepsilon}))D_i h(z^{\tilde\varepsilon})R_i(z)\,\mathrm{d}x\mathrm{d}t$$
$$\leq \limsup_{L\to\infty} \int_0^T \int_\Omega \theta_L'(h_{\delta,\varepsilon}(z))D_i h(z)R_i(z)\,\mathrm{d}x\mathrm{d}t$$
$$\leq \int_0^T \int_\Omega D_i h(z)R_i(z)\,\mathrm{d}x\mathrm{d}t.$$

Observe that the last line is independent of δ and ε.

Put together, the above inequalities and equation (4.4) imply the entropy dissipation inequality (ED).

<u>Case 2:</u> $0 < s < t < T^*$.

We assert that the fact that z is a renormalised solution in Ω_{T^*} in the sense of Definition 2.1 implies that for a.e. $0 < s < t < T^*$, and all $\tilde{\psi} \in C^\infty([0, T^*) \times \bar{\Omega})$

$$\int_\Omega \xi(z(t,\cdot))\tilde{\psi}(t,\cdot)\,\mathrm{d}x - \int_\Omega \xi(z(s,\cdot))\tilde{\psi}(s,\cdot)\,\mathrm{d}x - \int_s^t \int_\Omega \xi(z)\partial_\tau\tilde{\psi}\,\mathrm{d}x\mathrm{d}\tau$$

$$= -\int_s^t \int_\Omega D_{ij}\xi(z)A_{ik}(z)\nabla z_k \cdot \nabla z_j \tilde{\psi}\,\mathrm{d}x\mathrm{d}\tau \tag{4.8}$$

$$- \int_s^t \int_\Omega D_i\xi(z)A_{ik}(z)\nabla z_k \cdot \nabla\tilde{\psi}\,\mathrm{d}x\mathrm{d}\tau + \int_s^t \int_\Omega D_i\xi(z)R_i(z)\tilde{\psi}\,\mathrm{d}x\mathrm{d}\tau$$

for all $\xi \in C^\infty(\mathbb{R}_{\geq 0}^{1+n})$.

This can be proved as follows: take $\eta \in C^\infty(\mathbb{R})$, $\eta' \geq 0$, $\eta = 0$ on $(-\infty, -1]$, $\eta = 1$ on $[0, \infty)$, and define $\eta_{s,\delta}(\tau) := \eta(\frac{\tau-s}{\delta})$ for $0 < \delta \ll 1$. Observe that $\eta_{s,\delta}(\tau) = 0$ for $\tau \leq s - \delta$ and $\eta_{s,\delta}(\tau) = 1$ for $\tau \geq s$. In the renormalised formulation (2.1) with $T = t$, we choose the test function $\psi(\tau, x) := \tilde{\psi}(\tau, x)\eta_{s,\delta}(\tau)$. Then, the corresponding right-hand side of (2.1) converges, as $\delta \to 0$, to the right-hand side of eq. (4.8) by the dominated convergence theorem. The corresponding left-hand side takes the form

$$\int_\Omega \xi(z(t,\cdot))\tilde{\psi}(t,\cdot)\,\mathrm{d}x - \int_{s-\delta}^s \int_\Omega \xi(z)\tilde{\psi}\,\mathrm{d}x\,\partial_\tau\eta_{s,\delta}\mathrm{d}\tau + O(\delta) - \int_s^t \int_\Omega \xi(z)\partial_\tau\tilde{\psi}\,\mathrm{d}x\mathrm{d}\tau.$$

The second term in the last line can be rewritten as $\left(F * (\frac{1}{\delta}\eta'(-\frac{\cdot}{\delta}))\right)(s)$ for the measurable, bounded function $F(\tau) := -\int_\Omega \xi(z(\tau,x))\tilde{\psi}(\tau,x)\,\mathrm{d}x$, and since $\int_\mathbb{R} \eta'(-\tau)\mathrm{d}\tau = 1$, we have the convergence $\left(F * (\frac{1}{\delta}\eta'(-\frac{\cdot}{\delta}))\right)(s) \to F(s) = -\int_\Omega \xi(z(s,x))\tilde{\psi}(s,x)\,\mathrm{d}x$ for a.e. $s \in (0, T^*)$ as $\delta \to 0$. This establishes the asserted identity (4.8).

From (4.8) we infer that, for a.e. $0 < s < T^*$, the function $(\tau, x) \mapsto z(s+\tau, x)$ is a renormalised solution in $(0, T^* - s) \times \Omega$ with initial data $z(s, \cdot)$. Now we can invoke Case 1 and deduce (ED.s). $\square$

5. Exponential convergence to equilibrium

Proof of Proposition 2.10. Let us first observe that the regularity hypotheses on $z = (u, c)$ together with the bounds (ED.s) and (ENE.s) imply that $\mathscr{P}(z) + |\nabla u|^2 \in L^1(\Omega_T)$ for any $T < \infty$ and that $u \in L^\infty_{\text{loc}}([0,\infty), L^2(\Omega))$, $c_i \log c_i \in L^\infty_{\text{loc}}([0,\infty), L^1(\Omega))$ for all $i \in \{1, \ldots, n\}$. Here, we also used the lower and upper bounds on $H(z)$ provided in Lemma 6.3.

The energy and mass conservation properties $\int_\Omega z_l(t, x)\,\mathrm{d}x = \bar{z}_l$ for a.e. $t > 0$, where $l \in \{0, 1, \ldots, n\}$, can be seen as follows. In the renormalised formulation (2.1), we choose $\psi \equiv 1$ and $\xi(z) := \varphi_l^E(z)$ for $E \geq 1$, where $\varphi_l^E(z) = E\varphi_l(E^{-1}z)$ for some $\varphi_l \in C^\infty(\mathbb{R}_{\geq 0}^{1+n})$ with $\text{supp}\,D\varphi_l$ compact and $\varphi_l(z) = z_l$ for $|z|_1 < 1$ (see [24] for an example of such φ_l). This gives

$$\int_\Omega \varphi_l^E(z(T,\cdot))\,\mathrm{d}x - \int_\Omega \varphi_l^E(z^{\text{in}})\,\mathrm{d}x$$
$$= -\int_0^T \int_\Omega D_{ij}\varphi_l^E(z)A_{ik}(z)\nabla z_k \cdot \nabla z_j\,\mathrm{d}x\mathrm{d}t. \tag{5.1}$$

By the dominated convergence theorem, the LHS converges, as $E \to \infty$, to

$$\int_\Omega z_l(T, x)\,\mathrm{d}x - \int_\Omega z_l^{\text{in}}\,\mathrm{d}x.$$

Since $D^2 \varphi_l(z) = 0$ for $|z|_1 < 1$, we have $\lim_{E \to \infty} D_{ij} \varphi_l^E(z) = 0$ for every $z \in \mathbb{R}_{\geq 0}^{1+n}$. At the same time, using the bounds in Lemma 6.1 it is easy to see that for all $i, j \in \{0, \ldots, n\}$

$$\left| D_{ij} \varphi_l^E(z) \sum_{k=0}^{n} A_{ik}(z) \nabla z_k \cdot \nabla z_j \right| \lesssim \mathscr{P}(z) + |\nabla u|^2$$

uniformly in $E \geq 1$. Hence, the integral on the RHS of (5.1) converges to zero as $E \to \infty$ thanks to dominated convergence.

Let us now sketch the proof showing the exponential convergence to equilibrium. Below, ϵ_k, $k = 0, 1, \ldots$, denote fixed positive constants. As we only consider Model **(M0)**, it suffices to take $\alpha \in (0, 1]$. Conservation of the total energy and the mass of each species, combined with $\bar{z}$ being spatially constant, yields for a.e. time

$$H_{\text{rel}}(z, \bar{z}) = H(z) - D_i h(\bar{z}) \int_\Omega (z_i - \bar{z}_i) \, \mathrm{d}x - H(\bar{z})$$
$$= H(z) - H(\bar{z}).$$

Hence, inequality (ED.s) gives for a.e. $t \geq s \geq 0$

$$H_{\text{rel}}(z(\tau), \bar{z}) \Big|_{\tau=s}^{\tau=t} \leq -\int_s^t \int_\Omega \mathscr{P}(z) \, \mathrm{d}x \mathrm{d}\tau.$$

Writing $B(c, \bar{c}) := \sum_{i=1}^{n} b(c_i, \bar{c}_i)$, where $b(s, \bar{s}) = \bar{s} \lambda(s/\bar{s})$, we have by estimate (6.2) and the logarithmic Sobolev inequality (cf. [39])

$$\int_\Omega \mathscr{P}(z) \, \mathrm{d}x \geq \epsilon_0 \sum_{i=1}^{n} \int_\Omega |\nabla \sqrt{c_i}|^2 \, \mathrm{d}x \geq \epsilon_1 \int_\Omega B(c, \bar{c}) \, \mathrm{d}x, \tag{5.2}$$

where $\epsilon_1 = \epsilon_1(\Omega) > 0$.

Recalling the definition of g_{rel} in (3.4), and using $\int u \, \mathrm{d}x = \int \bar{u} \, \mathrm{d}x$, (ENE.s), and the Poincaré–Wirtinger inequality, we can further estimate for some $\epsilon_2 = \epsilon_2(\Omega) > 0$

$$G_{\text{rel}}(u(\tau), \bar{u}) \Big|_{\tau=s}^{\tau=t} \leq -\epsilon_2 \int_s^t \int_\Omega |u - \bar{u}|^2 \, \mathrm{d}x \mathrm{d}\tau + \varpi C \int_s^t \int_\Omega \mathscr{P}(z) \, \mathrm{d}x \mathrm{d}\tau. \tag{5.3}$$

We next let $E \geq |\bar{z}|$, to be fixed later. Then, by uniform convexity, for $|z| \leq E$,

$$B(c, \bar{c}) + \alpha |u - \bar{u}|^2 \gtrsim_{E,\alpha} |z - \bar{z}|^2 \gtrsim_E h_{\text{rel}}(z, \bar{z}).$$

(The argument leading to the second inequality is as in the proof of (3.5).)

At the same time,

$$B(c, \bar{c}) + \alpha |u - \bar{u}|^2 \geq \tfrac{1}{2} \left(\sum_{i=1}^{n} c_i \log c_i + \alpha u^2 \right) - C(\bar{z}).$$

Hence, for $E = E(\bar{z}, \alpha)$ large enough, we obtain

$$B(c, \bar{c}) + \alpha |u - \bar{u}|^2 \gtrsim_E h_{\text{rel}}(z, \bar{z}).$$

Combining the above estimates and choosing $\alpha \in (0, 1]$ such that $\alpha C \varpi \leq \tfrac{1}{2}$ (with C as in (5.3)), we infer

$$\left[H_{\text{rel}}(z(\tau), \bar{z}) + \alpha G_{\text{rel}}(u(\tau), \bar{u}) \right] \Big|_{\tau=s}^{\tau=t} \leq -\frac{1}{2} \int_s^t \int_\Omega \mathscr{P}(z) \, \mathrm{d}x \mathrm{d}\tau - \epsilon_2 \alpha \int_s^t \|u - \bar{u}\|_{L^2(\Omega)}^2 \mathrm{d}\tau$$

$$\leq -\epsilon_3 \int_s^t \left[H_{\text{rel}}(z(\tau), \bar{z}) + \alpha G_{\text{rel}}(u(\tau), \bar{u}) \right] \mathrm{d}\tau,$$

where $\epsilon_3 = \epsilon_3(\bar{z}, \alpha, \Omega) > 0$. A version of Gronwall's inequality (see e.g. [21, p. 702]) yields the asserted bound (2.13) for $\lambda = \epsilon_3 > 0$. $\qquad\square$

Let us note that in the above proof we have not used the fact that $\pi_1 \gamma^2 |\nabla u|^2$ is dominated by $\mathscr{P}(z)$ (see (5.2)). If $\sigma(u)$ is sufficiently close to a linear function for $u \gg 1$, e.g. $\sigma(u) = u^{1-\epsilon(d)}$ for $\epsilon(d) > 0$ small enough, this term may be exploited as in [39, Section 3] to quantitatively improve the decay rate.

6. APPENDIX

Lemma 6.1 (Estimates for Model **(M0)**). *Let the hypotheses of Model* **(M0)** *be satisfied. Then, formally,*

$$\sum_{k=0}^{n} A_{0k}(z) \nabla z_k = a(z) \nabla z + m(z) \nabla D_0 h(z), \tag{6.1a}$$

$$\sum_{k=0}^{n} A_{ik}(z) \nabla z_k = m_i(z) \nabla D_i h(z) + a(z) c_i \frac{w_i'(u)}{w_i(u)} \nabla u \quad \text{for } i \geq 1, \tag{6.1b}$$

where $a(z) = \pi_1(z)\gamma(z)$ and γ is given by (2.6).
Moreover, for any sufficiently regular function $z = (u, c_1, \ldots, c_n)$ with positive components, we have the following estimates:

Abbreviating

$$\mathscr{P}(z) := \nabla z : (D^2 h(z) A(z) \nabla z) = \nabla D_i h(z) \cdot (\mathbb{M}_{il}(z) \nabla D_l h(z))$$

and $\gamma(u, c) := -\hat{\sigma}''(u) - \sum_{l=1}^{n} \frac{w_l''(u)}{w_l(u)} c_l$, one has

$$\mathscr{P}(z) \gtrsim \sum_{i=1}^{n} |\nabla \sqrt{c_i}|^2 + |\sqrt{\gamma}\, \nabla u|^2 + |\sqrt{m} \nabla D_0 h(z)|^2, \tag{6.2}$$

and

$$a(z) |\nabla u|^2 \sim |\nabla u|^2. \tag{6.3}$$

Furthermore,

$$|A(z) \nabla z| \lesssim \left(\max_{i=1,\ldots,n} \sqrt{c_i} + \sqrt{\pi_1(z)} + \sqrt{m(z)} \right) \sqrt{\mathscr{P}(z)}, \tag{6.4}$$

$$\left| \sum_{k=0}^{n} A_{ik}(z) \nabla z_k \right| \lesssim \sqrt{c_i} \sqrt{\mathscr{P}(z)} \quad \text{for } i \geq 1, \tag{6.5}$$

$$\left| \sum_{k=0}^{n} A_{0k}(z) \nabla z_k \right| \lesssim |\nabla u| + \sqrt{m} \sqrt{\mathscr{P}(z)}. \tag{6.6}$$

Proof. Identities (6.1a)–(6.1b) follow from a straightforward computation using the definition of μ_i (see [24], if necessary). Except for the last term in (6.2), estimate (6.2) is a consequence of [24, Lemma 2.1] (in [24]: $m \equiv 0$). The additional control of $|\sqrt{m} \nabla D_0 h(z)|^2$ for $m = m(z) \geq 0$ easily follows from the definition of $\mathscr{P}(z)$ and $\mathbb{M}$. Eq. (6.3) is immediate since $a = \pi \gamma \sim 1$ by hypothesis. Estimate (6.4) has been established in [24, Lemma 2.3, eq. (2.10)] for $m \equiv 0$, and the current version thus follows estimate (6.2), which implies the bound $|m(z) \nabla D_0 h(z)| \leq \sqrt{m} \sqrt{\mathscr{P}(z)}$. Estimate (6.5) is a consequence of the proof of [24, Lemma 2.3], while estimate (6.6) follows from the fact that $a \sim 1$ and the bound on $m(z) \nabla D_0 h(z)$ observed before. $\qquad\square$

Lemma 6.2. *Model* **(M0)** *(see page 13) fulfils conditions* (A1)–(A5) *and* (A6) *of Theorem 2.8 when assuming additionally the regularity hypotheses $\hat{\sigma}, w_i \in C^4((0, \infty))$,* (A2.ii) *and $m, m_i, \pi_1 \in C^{0,1}_{\text{loc}}((0, \infty)^{1+n})$.*

Proof. The asserted estimates can be verified using Lemma 6.1: the bounds in (A4) are immediate consequences of estimates (6.2), (6.4) combined with the bound $0 \leq \sqrt{\pi_1(z)} \lesssim (1+u)$. Condition (A5) follows from estimating $|a(z)\nabla u| \lesssim \sqrt{\pi_1}|\sqrt{\gamma}\nabla u|$ and using (6.2). Condition (A6) easily follows from (6.5) and (6.6). (We have not aimed at optimising the conditions on $m(z)$, which are far from being sharp.) The conditions in (A3) follow from the definition of $\mathbb{M}$ in (2.7). $\qquad\square$

Lemma 6.3 (Lower and upper entropy bounds). *Let $h = h(u,c)$ be given by* (h1) *with* (h2) *being satisfied. There exist positive constants $\epsilon_\beta > 0$, $\kappa_\beta \in (0,1)$ and $C_\beta, C \in (0,\infty)$ such that for all $(u,c) \in [0,\infty)^{1+n}$*

$$h(u,c) \geq -\hat{\sigma}(u) + \epsilon_\beta \sum_{i=1}^{n} c_i \log(c_i) - Cu^{\kappa_\beta} - C_\beta, \tag{6.7}$$

$$h(u,c) \leq -\hat{\sigma}(u) + C \sum_{i=1}^{n} c_i \log(c_i) + C. \tag{6.8}$$

Proof. Letting $\beta_* = \frac{\beta+1}{2} \in (\beta, 1)$, we estimate using (h2)

$$c_i \log(w_i(u)) = c_i \log(w_i(u))\chi_{\{w_i(u) \leq c_i^{\beta_*}\}} + c_i \log(w_i(u))\chi_{\{w_i(u) > c_i^{\beta_*}\}}$$

$$\leq \beta_* c_i \log(c_i) + w_i(u)^{\frac{1}{\beta_*}} \log(w_i(u)) + C$$

$$\leq \beta_* c_i \log(c_i) + Cu^{\frac{1}{2}(1+\beta/\beta_*)} + C.$$

Thus,

$$h(u,c) = -\hat{\sigma}(u) + \sum_{i=1}^{n} \big(\lambda(c_i) - c_i \log(w_i(u))\big)$$

$$\geq -\hat{\sigma}(u) + \sum_{i=1}^{n} \big((1-\beta_*)c_i \log(c_i) - c_i + 1\big) - Cu^{\frac{1}{2}(1+\beta/\beta_*)} - C.$$

This yields (6.7) with $\kappa_\beta := \frac{1}{2}(1+\beta/\beta_*) < 1$, $\epsilon_\beta = \frac{1}{2}(1-\beta_*) > 0$ and a suitable constant $C_\beta < \infty$.

Estimate (6.8) easily follows from the hypothesis that $w_i(0) > 0$ for all i (see also the proof of [24, eq. (2.9)]). $\qquad\square$

Lemma 6.4 (Minimum principle). *In addition to the hypotheses of Theorem 2.8 assume that $\varpi = 0$. Let $T \in (0, T^*)$ and $\underline{u} := \inf_{\Omega_T} u^{\mathrm{in}} > 0$. Then the energy component u of the renormalised solution $z = (u,c)$ satisfies $u \geq \underline{u}$ almost everywhere in Ω_T.*

Sketch proof. The hypotheses imply that $u \in L^\infty_{\mathrm{loc}}(I; L^2(\Omega))$ and that there exists $r > 1$ such that $a(z)\nabla u \in L^r_{\mathrm{loc}}(I; L^r(\Omega))$, $\partial_t u \in L^r_{\mathrm{loc}}(I; (W^{1,r'}(\Omega))^*)$, $\frac{1}{r} + \frac{1}{r'} = 1$. The weak formulation of the energy component (2.3) can therefore be integrated by parts with respect to time to give

$$\int_0^{T'} \langle \partial_t u, \varphi \rangle \, \mathrm{d}t = - \int_0^{T'} \int_\Omega a(z)\nabla u \cdot \nabla\varphi \, \mathrm{d}x\mathrm{d}t. \tag{6.9}$$

for a.a. $T' \in (0, T]$. Ignoring regularity issues for the moment and testing the equation with $\varphi = (u - \underline{u})_-$ leads to

$$\frac{1}{2} \int_\Omega |(u-\underline{u})_-|^2 \, \mathrm{d}x \bigg|_{t=0}^{t=T'} + \int_0^{T'} \int_\Omega a(z)|\nabla(u-\underline{u})_-|^2 \, \mathrm{d}x\mathrm{d}t = 0. \tag{6.10}$$

This implies that $(u - \underline{u})_- = 0$ and hence $u \geq \underline{u}$ a.e. in Ω_T.

To make the argument rigorous, one considers a smooth partition of unity $(\chi_k)_{k=1}^N$ on $\bar{\Omega}$ as in the proof of the L^2 identities in [24, Lemma 6.1], see also [11, Lemma 12] and [22, Lemma 4]. For simplicity, we

only outline the reasoning in the case of $\psi := \chi_k$ being compactly supported in Ω, and refer, for the general case, to the first and the third of the references provided before.

Denote by $\tilde{\rho}$ the standard mollifying kernel, let $\rho := \tilde{\rho} * \tilde{\rho}$ and $\rho_\varepsilon(x) := \frac{1}{\varepsilon^d}\rho(\frac{x}{\varepsilon})$. Then, for $\varepsilon > 0$ small enough (only depending on $\mathrm{dist}(\mathrm{supp}\,\psi, \partial\Omega)$), choose in (6.9) the test function

$$\varphi = \rho_\varepsilon * ((\rho_\varepsilon * u - \underline{u})_-\psi),$$

which lies in $W^{1,r}_{\mathrm{loc}}(I; H^s(\Omega))$ for any $s \in \mathbb{N}$. Abbreviate $u_\varepsilon = \rho_\varepsilon * u$ and compute

$$\partial_t u_\varepsilon (u_\varepsilon - \underline{u})_-\psi = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|(u_\varepsilon - \underline{u})_-|^2\psi.$$

One the other hand, the term

$$\int_0^{T'}\int_\Omega \rho_\varepsilon * (a(z)\nabla u) \cdot \nabla((u_\varepsilon - \underline{u})_-\psi)\,\mathrm{d}x\mathrm{d}t$$

can be shown to converge to

$$\int_0^{T'}\int_\Omega a(z)\nabla u \cdot \nabla((u - \underline{u})_-\psi)\,\mathrm{d}x\mathrm{d}t$$

by arguing similarly as in the proof of the L^2-energy identity in [24, Lemma 6.1], see also [11, Lemma 12], where one should use the fact that $\|(u_\varepsilon - \underline{u})_-\|_{L^\infty} \leq \underline{u}$, which follows from the non-negativity of u. Thus,

$$\frac{1}{2}\int_\Omega |(u - \underline{u})_-|^2\chi_k\,\mathrm{d}x\bigg|_{t=0}^{t=T'} + \int_0^{T'}\int_\Omega a(z)\nabla u \cdot \nabla((u - \underline{u})_-\chi_k)\,\mathrm{d}x\mathrm{d}t = 0,$$

and upon summation over k one arrives at (6.10). $\qquad\square$

References

[1] N. Alibaud, B. Andreianov, and A. Ouédraogo. Nonlocal dissipation measure and L^1 kinetic theory for fractional conservation laws. *Comm. Partial Differential Equations*, 45(9):1213–1251, 2020.

[2] B. Andreianov and N. Igbida. On uniqueness techniques for degenerate convection-diffusion problems. *Int. J. Dyn. Syst. Differ. Equ.*, 4(1-2):3–34, 2012.

[3] P. Bénilan, J. Carrillo, and P. Wittbold. Renormalized entropy solutions of scalar conservation laws. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 29(2):313–327, 2000.

[4] J. Berendsen, M. Burger, V. Ehrlacher, and J.-F. Pietschmann. Uniqueness of strong solutions and weak-strong stability in a system of cross-diffusion equations. *J. Evol. Equ.*, 20(2):459–483, 2020.

[5] L. A. Caffarelli and X. Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[6] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.*, 147(4):269–361, 1999.

[7] J. Carrillo and P. Wittbold. Uniqueness of renormalized solutions of degenerate elliptic-parabolic problems. *J. Differential Equations*, 156(1):93–121, 1999.

[8] J. A. Carrillo, K. Hopf, and J. L. Rodrigo. On the singularity formation and relaxation to equilibrium in 1D Fokker-Planck model with superlinear drift. *Adv. Math.*, 360:106883, 66, 2020.

[9] X. Chen, E. S. Daus, and A. Jüngel. Global existence analysis of cross-diffusion population systems for multiple species. *Arch. Ration. Mech. Anal.*, 227(2):715–747, 2018.

[10] X. Chen and A. Jüngel. A note on the uniqueness of weak solutions to a class of cross-diffusion systems. *J. Evol. Equ.*, 18(2):805–820, 2018.

[11] X. Chen and A. Jüngel. Global renormalized solutions to reaction-cross-diffusion systems with self-diffusion. *J. Differential Equations*, 267(10):5901–5937, 2019.

[12] X. Chen and A. Jüngel. Weak-strong uniqueness of renormalized solutions to reaction-cross-diffusion systems. *Math. Models Methods Appl. Sci.*, 29(2):237–270, 2019.

[13] Y. G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33(3):749–786, 1991.

[14] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[15] I. Csiszár. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 8:85–108, 1963.

[16] C. M. Dafermos. Hyperbolic balance laws in continuum physics. In *Nonlinear problems in theoretical physics (Proc. IX G.I.F.T. Internat. Sem. Theoret. Phys., Univ. Zaragoza, Jaca, 1978)*, volume 98 of *Lecture Notes in Phys.*, pages 107–121. Springer, Berlin-New York, 1979.

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

[18] R. J. DiPerna. Uniqueness of solutions to hyperbolic conservation laws. *Indiana Univ. Math. J.*, 28(1):137–188, 1979.

[19] R. J. DiPerna. Measure-valued solutions to conservation laws. *Arch. Rational Mech. Anal.*, 88(3):223–270, 1985.

[20] L. C. Evans and J. Spruck. Motion of level sets by mean curvature. I. *J. Differential Geom.*, 33(3):635–681, 1991.

[21] K. Fellner and E.-H. Laamri. Exponential decay towards equilibrium and global classical solutions for nonlinear reaction-diffusion systems. *J. Evol. Equ.*, 16(3):681–704, 2016.

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

[23] J. Fischer. Weak-strong uniqueness of solutions to entropy-dissipating reaction-diffusion equations. *Nonlinear Anal.*, 159:181–207, 2017.

[24] J. Fischer, K. Hopf, M. Kniely, and A. Mielke. Global existence analysis of energy-reaction-diffusion equations. *arXiv preprint 2012.03792v2*, 2020.

[25] H. Gajewski. On a variant of monotonicity and its application to differential equations. *Nonlinear Anal.*, 22(1):73–80, 1994.

[26] G. Galiano and J. Velasco. Well-posedness of a cross-diffusion population model with nonlocal diffusion. *SIAM J. Math. Anal.*, 51(4):2884–2902, 2019.

[27] P. Gwiazda, O. Kreml, and A. Świerczewska Gwiazda. Dissipative measure-valued solutions for general conservation laws. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 37(3):683–707, 2020.

[28] J. Haskovec, S. Hittmeir, P. Markowich, and A. Mielke. Decay to equilibrium for energy-reaction-diffusion systems. *SIAM J. Math. Anal.*, 50(1):1037–1075, 2018.

[29] H. Ishii and S. Koike. Viscosity solutions for monotone systems of second-order elliptic PDEs. *Comm. Partial Differential Equations*, 16(6-7):1095–1128, 1991.

[30] R. Jensen. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Arch. Rational Mech. Anal.*, 101(1):1–27, 1988.

[31] A. Jüngel. Regularity and uniqueness of solutions to a parabolic system in nonequilibrium thermodynamics. *Nonlinear Anal.*, 41(5-6, Ser. A: Theory Methods):669–688, 2000.

[32] S. N. Kružkov. First order quasilinear equations in several independent variables. *Mathematics of the USSR-Sbornik*, 10(2):217–243, feb 1970.

[33] J. Leray. Sur le mouvement d'un liquide visqueux emplissant l'espace. *Acta Math.*, 63(1):193–248, 1934.

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

[35] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.

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

[37] A. Mielke. Uniform exponential decay for reaction-diffusion systems with complex-balanced mass-action kinetics. In *Patterns of dynamics*, volume 205 of *Springer Proc. Math. Stat.*, pages 149–171. Springer, Cham, 2017.

[38] A. Mielke, J. Haskovec, and P. A. Markowich. On uniform decay of the entropy for reaction-diffusion systems. *J. Dynam. Differential Equations*, 27(3-4):897–928, 2015.

[39] A. Mielke and M. Mittnenzweig. Convergence to equilibrium in energy-reaction-diffusion systems using vector-valued functional inequalities. *J. Nonlinear Sci.*, 28(2):765–806, 2018.

[40] B. Perthame. *Kinetic formulation of conservation laws*, volume 21 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[41] D. Pham and R. Temam. A result of uniqueness of solutions of the Shigesada-Kawasaki-Teramoto equations. *Adv. Nonlinear Anal.*, 8(1):497–507, 2019.

[42] J. C. Robinson, J. L. Rodrigo, and W. Sadowski. *The three-dimensional Navier-Stokes equations*, volume 157 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2016. Classical theory.

[43] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, volume 39 of *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass.-London, 1979.

[44] A. Unterreiter, A. Arnold, P. Markowich, and G. Toscani. On generalized Csiszár-Kullback inequalities. *Monatsh. Math.*, 131(3):235–253, 2000.

[45] E. Wiedemann. Weak-strong uniqueness in fluid dynamics. In *Partial differential equations in fluid mechanics*, volume 452 of *London Math. Soc. Lecture Note Ser.*, pages 289–326. Cambridge Univ. Press, Cambridge, 2018.