GLOBAL WEAK SOLUTION OF THE DIFFUSIVE JOHNSON-SEGALMAN MODEL OF A VISCOELASTIC HEAT-CONDUCTING FLUID

MICHAL BATHORY, MIROSLAV BULÍČEK, AND JOSEF MÁLEK

Abstract. We prove that there exists a large-data and global-in-time weak solution to a system of partial differential equations describing an unsteady flow of an incompressible heat-conducting rate-type viscoelastic stress-diffusive fluid filling up a mechanically and thermally isolated container of any dimension. To get around the notorious ill-posedness of the diffusive Oldroyd-B model in 3D, we assume that the fluid admits a strengthened dissipation mechanism, at least for excessive elastic deformations. All the relevant material coefficients are allowed to depend continuously on the temperature, whose evolution is captured by a thermodynamically consistent equation. In fact, the studied model is derived from scratch using only the balance equations and thermodynamical laws. The only real simplification of the model, apart from the incompressibility, homogeneity and isotropicity of the fluid, is that we assume a linear relation between the temperature and the internal energy. The concept of our weak solution is considerably general as the thermal evolution of the system is governed only by the entropy inequality and the global conservation of energy. Still, this is sufficient for the weak-strong compatibility of our solution and we also specify additional conditions on the material coefficients under which the balances of the total and internal energy hold locally.

1. Introduction

Mathematical formulation of the problem. Let $\Omega \subset \mathbb{R}^d$ be an open bounded domain with a Lipschitz boundary $\partial\Omega$ and its outward unit normal vector denoted by $\boldsymbol{n}$. We also denote by $\boldsymbol{x}_\tau := \boldsymbol{x} - (\boldsymbol{x}\cdot\boldsymbol{n})\boldsymbol{x}$ the tangential part of the vector $\boldsymbol{x} \in \mathbb{R}^d$. Let $(0,T)$, $T>0$, be an arbitrarily large time interval and set $Q := (0,T)\times\Omega$. In this work we develop an existence theory for the following initial-boundary value problem in Q, with arbitrarily large data $\boldsymbol{f} : Q \to \mathbb{R}^d$, $\boldsymbol{v}_0 : \Omega \to \mathbb{R}^d$, $\mathbb{B}_0 : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ positive definite and $\theta_0 : \Omega \to (0,\infty)$. Let $a \in \mathbb{R}$, $\alpha \geq 0$, $\mu, c_v > 0$ and suppose that $\nu, \lambda, \kappa : (0,\infty) \to (0,\infty)$ are continuous functions. Furthermore, let $P : (0,\infty)\times\mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be continuous. The symmetric and antisymmetric part of a gradient $(\nabla\boldsymbol{u})_{ij} = \partial_j \boldsymbol{u}_i$ of a function $\boldsymbol{u} : \mathbb{R}^d \to \mathbb{R}^d$ are denoted by $\mathbb{D}\boldsymbol{u}$ and $\mathbb{W}\boldsymbol{u}$, respectively, so that $\nabla\boldsymbol{u} = \mathbb{D}\boldsymbol{u} + \mathbb{W}\boldsymbol{u}$ with $(\mathbb{D}\boldsymbol{u})^T = \mathbb{D}\boldsymbol{u}$ and $(\mathbb{W}\boldsymbol{u})^T = -\mathbb{W}\boldsymbol{u}$. Then,

2020 *Mathematics Subject Classification.* 35A23, 76A10, 76D03.

Key words and phrases. viscoelastic heat-conducting fluids, Johnson-Segalman, weak solution.

Michal Bathory acknowledges support from the Austrian Science Fund (FWF) project F65 and also from the project No. 1652119 financed by the Charles University Grant Agency (GAUK). Michal Bathory and Miroslav Bulíček thank to the project No. 20-11027X financed by the Czech Science foundation (GAČR). Josef Málek acknowledges the support of the project No. 18-12719S financed by the Czech Science foundation (GAČR). Miroslav Bulíček and Josef Málek are members of the Nečas Center for Mathematical Modelling.

the goal is to find functions $\boldsymbol{v} : Q \to \mathbb{R}^d$, $\mathrm{p}, \theta, e, E, \eta : Q \to \mathbb{R}$, $\mathbb{B} : Q \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ fulfilling the (physical) restrictions

$$\theta > 0, \tag{1.1}$$

$$\mathbb{B}\boldsymbol{x} \cdot \boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d \setminus \{0\}, \tag{1.2}$$

$$e = c_v \theta, \tag{1.3}$$

$$E = \tfrac{1}{2}|\boldsymbol{v}|^2 + e, \tag{1.4}$$

$$\eta = c_v \ln \theta - \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}), \tag{1.5}$$

$$\xi = \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla \ln\theta|^2 + \mu P(\theta,\mathbb{B})\cdot(\mathbb{I}-\mathbb{B}^{-1}) + \mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2, \tag{1.6}$$

and solving (in some suitable sense) the system

$$\operatorname{div}\boldsymbol{v} = 0, \tag{1.7}$$

$$\partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v} - \operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v}) + \nabla\mathrm{p} = 2a\mu\operatorname{div}(\theta\mathbb{B}) + \boldsymbol{f}, \tag{1.8}$$

$$\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + P(\theta,\mathbb{B}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \tag{1.9}$$

$$\partial_t e + \boldsymbol{v}\cdot\nabla e - \operatorname{div}(\kappa(\theta)\nabla\theta) = 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}, \tag{1.10}$$

$$\begin{aligned}&\partial_t E + \boldsymbol{v}\cdot\nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta)\\ &\qquad = \operatorname{div}(-(\mathrm{p} + 2a\mu\theta)\boldsymbol{v} + 2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v},\end{aligned} \tag{1.11}$$

$$\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa(\theta)\nabla\ln\theta) + \operatorname{div}(\mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})) = \xi \tag{1.12}$$

in Q with the boundary conditions

$$\boldsymbol{v}\cdot\boldsymbol{n} = 0, \qquad (2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{n} + 2a\mu\theta\mathbb{B}\boldsymbol{n} + \alpha\boldsymbol{v})_\tau = 0, \tag{1.13}$$

$$\boldsymbol{n}\cdot\nabla\mathbb{B} = 0, \tag{1.14}$$

$$\boldsymbol{n}\cdot\nabla\theta = 0 \tag{1.15}$$

on $(0,T)\times\partial\Omega$ and with the initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0, \qquad \mathbb{B}(0) = \mathbb{B}_0, \qquad \theta(0) = \theta_0 \tag{1.16}$$

in Ω.

The first two equations (1.7) and (1.8) form the incompressible Navier-Stokes system for the unknowns $\boldsymbol{v}$ and p, however, with an additional forcing term $2a\mu\operatorname{div}(\theta\mathbb{B})$ introducing two other unknowns θ and $\mathbb{B}$. Since the dependence of the material parameters (namely the fluid's viscosity) on the pressure p is neglected, we simplify the analysis by eliminating the pressure from the system completely, taking the Leray projection of (1.8) and searching for $\boldsymbol{v}$ in divergence-free function spaces. If needed (for example if we want to preserve the equation (1.11)), the pressure can be reconstructed at the last step. Then, it is known (cf. [10]) that the Navier-slip boundary condition (1.13) allows us to prove that p is an integrable function if the boundary of Ω is smooth enough.

The evolution of $\mathbb{B}$ is governed by (1.9), which is a generalized Johnson-Segalman model (see [23]) with a stress diffusion (cf. [38] and references therein). If $a \neq 0$, the analysis of the problem becomes difficult as it is unclear whether the terms of the type $\mathbb{B}\nabla\boldsymbol{v}$ are summable, using the known apriori estimates. This is essentially the reason, why we formulate (1.9) with a general function P: the strategy is that if $P(\cdot,\mathbb{B})$ grows sufficiently fast as $|\mathbb{B}| \to \infty$, then $\mathbb{B}$ admits enough integrability to define the right hand side of (1.9). Moreover, as the form of P can be attributed to the dissipation mechanism of the fluid, restrictions its asymptotic growth should not be seen as a significant physical drawback of our model. One of the main features of our analysis is that this way, we are able to treat (1.9) with any $a \in \mathbb{R}$.

Due to (1.3), the balance of internal energy (1.10) is also the temperature equation. Note the appearance of the term $2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$ on the right hand side of (1.10). In a sense, this is the most difficult term to control in the whole system (1.7)–(1.12) and it is also the term which is sometimes omitted in some "naive" approaches to thermoviscoelasticity that are pointed out in [22, Section 3]. The equations (1.11) and (1.12) govern the evolution of two other unknowns E and η, respectively. Since these quantities together with θ are mutually connected by simple algebraic relations (1.3), (1.4) and (1.5), the equations (1.10)–(1.12) are interchangeable and each of them alone can be used to model temperature evolution. To see this, note that (1.5) and (1.6) imply

$$\partial_t E = \boldsymbol{v} \cdot \partial_t \boldsymbol{v} + \partial_t e \tag{1.17}$$

$$\partial_t \eta = c_v \theta^{-1} \partial_t \theta - \mu(\mathbb{I} - \mathbb{B}^{-1}) \cdot \partial_t \mathbb{B}, \tag{1.18}$$

where we used the classical identities for Fréchet derivatives

$$\partial_{\mathbb{B}} \operatorname{tr} \mathbb{B} = \mathbb{I} \qquad \text{and} \qquad \partial_{\mathbb{B}} \ln \det \mathbb{B} = \mathbb{B}^{-1}. \tag{1.19}$$

Using these, one can verify that (1.10)–(1.12) are equivalent.

On the level of generalized solutions, however, we may not be allowed a priori to make the operations depicted by (1.17) and (1.18) due to insufficient regularity of some terms in (1.10) or (1.11), and thus the relation between (1.10), (1.11) and (1.12) becomes less clear. Based on the available[1] a priori estimates, it turns out that in a certain sense, we have

$$\text{"}(1.10) \Rightarrow (1.11) \Rightarrow (1.12)\text{"},$$

but not the other way around. For this reason, we choose to model the evolution of temperature by (1.12) as it leads to the least restrictive assumptions on the material coefficients, while still providing a physically sound notion of generalized solution that is consistent with the classical one, as we show below. Let us remark that the idea of replacing the temperature equation/balance of internal energy/balance of total energy by the balance of entropy is not new and it was applied (although rather implicitly), e.g., in [16], [18] or [17] for different fluid models. See also [8] for similar ideas in context of certain mixtures. Next, we must explain two more nuances leading to the appropriate notion of generalized solution to (1.7)–(1.12).

First, we require that the balance of entropy (1.12) is satisfied only in the form of the inequality

$$\partial_t \eta + \boldsymbol{v} \cdot \nabla \eta - \operatorname{div}(\kappa(\theta)\nabla \ln \theta) + \operatorname{div}(\mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})) \geq \xi \tag{1.20}$$

simply because we are unable to prove equality. This relaxation can be interpreted in a way that the fluid might produce additional entropy through mechanisms that are unseen in our model. Similarly, in those cases where it makes sense to consider the balance of internal energy (1.10), we shall replace it by the inequality

$$\partial_t e + \boldsymbol{v} \cdot \nabla e - \operatorname{div}(\kappa(\theta)\nabla\theta) \geq 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}. \tag{1.21}$$

[1]Those which we are able to prove.

Second, any weak solution for which (1.20) holds, but (1.11) does not make sense, might violate the conservation of total energy and thus become unphysical. Fortunately, we are always able to prove at least that the global conservation of total energy (in a rather strong form, see (1.25) below) and hence, this condition is enforced in our notion of generalized solution.

Weak solution. Summarizing the above, we will call the tuple $(\boldsymbol{v}, \mathbb{B}, \theta, e, E, \eta)$ a weak solution of the system (1.1)–(1.9), (1.12)–(1.16) if the constraints (1.1)–(1.7) are fulfilled almost everywhere in Q, the relations

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} - \operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v}) + \nabla \mathrm{p} = 2a\mu \operatorname{div}(\theta\mathbb{B}) + \boldsymbol{f}, \tag{1.22}$$

$$\partial_t \mathbb{B} + \boldsymbol{v} \cdot \nabla \mathbb{B} + P(\theta, \mathbb{B}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \tag{1.23}$$

$$\partial_t \eta + \boldsymbol{v} \cdot \nabla \eta - \operatorname{div}(\kappa(\theta)\nabla \ln\theta) + \operatorname{div}(\mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})) \geq \xi, \tag{1.24}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \tag{1.25}$$

hold in a distributional sense with the boundary conditions (1.13)–(1.15) and with initial conditions (1.16) satisfied in a limit sense. In Definition 1 below, we specify precisely in which distributional sense we understand (1.22)–(1.25) and also we state what regularity we expect from a solution.

All in all, even though we notably drifted away from the classical formulation of the problem (1.7)–(1.12), we can still guarantee that our weak solution will fulfil the basic physical principles such as the total energy conservation and the second law of thermodynamics. Moreover, we have compatibility in the sense that if a weak solution is smooth, then it is in fact a classical solution. Indeed, we can revert the operation (1.18) to derive (1.21) from (1.24) and (1.23). On the other hand, we can multiply (1.22) by $\boldsymbol{v}$, integrate over Ω and subtract the result from (1.25), which yields, after application of boundary conditions (1.13), (1.15) and (1.3) that

$$\int_\Omega (\partial_t e + \boldsymbol{v} \cdot \nabla e - \operatorname{div}(\kappa(\theta)\nabla\theta)) = \int_\Omega (2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}).$$

Comparing this with (1.21), we deduce (1.10). Consequently, we also get (1.12), (1.11) and thus all of the equations (1.7)–(1.12) are verified.

Our main result can be informally stated as follows.

Main result. *Under certain restrictions on the asymptotic growth of material coefficients, there exists a global-in-time weak solution of the system* (1.1)–(1.9), (1.12)–(1.16) *for any initial datum with finite total energy and entropy.*

For the precise formulation we refer to Theorem 1 below. There we also specify additional restrictions that are sufficient for validity of (1.11) and (1.21).

State of the art. Regarding the existence analysis of a viscoelastic fluid model including the full temperature evolution, there is an upcoming study [5]. There the authors show global and large-data existence of a weak solution to a rate-type incompressible viscoelastic fluid model with stress diffusion under the simplifying assumption that $\mathbb{B} = b\mathbb{I}$. This assumption leads to annihilation of irregular terms coming from the objective derivative and it also simplifies the momentum equation, where the coupling to the rest of the system is realized only via temperature and elastic stress dependent viscosity. Other than that, to the author's best knowledge, there is no existence theory in a setting that would be of similar generality as considered here. Thus, for the first time, we provide an existence analysis for a viscoelastic fluid model with a full thermal evolution and taking into account all components of the extra stress tensor. Moreover, the equation for the temperature we consider is derived from fundamental thermodynamical laws (similarly as in [5], [22], [32]) and consequently, the heating originates from both the viscous and elastic forces. Also, we would like to point out that we allow the most of the material coefficients of the model to depend on temperature. Although we place some restrictions on the growth of these coefficients, these are only asymptotic and therefore unimportant from the point of view of physical applications. Furthermore, the model considered here has the property that the evolution of the temperature can not be decoupled from the rest of the model even in the case of constant material coefficients.

Even if we confine to a much simpler class of isothermal viscoelastic models, the existence theory there is far from being complete. Although there are several relevant global-in-time existence results for large data, in most cases, they are restricted in an essential way. For example, in [26] the authors provide an existence theory for a model with the corrotational Jaumann-Zaremba derivative (the case $a = 0$). This case is much easier than for the other choices of a since the corrotational part drops out upon multiplication by any matrix that commutes with $\mathbb{B}$. Moreover, it seems that the physically preferred case is $a = 1$, which corresponds to the upper convected (Oldroyd) derivative (see [31], [34], [35], [40] or [41]). Then, the follow-up of this work is [36], where the author claims to prove existence of a weak solution to FENE-P, Giesekus and PTT viscoelastic models. However, in these works it is only shown that certain defect measures of the non-linear terms are compact. Furthermore, in the scalar case, that is if $\mathbb{B} = b\mathbb{I}$, we refer to [9] (and [6], [28] in the compressible case) for an analysis of such models. In the two-dimensional case, existence and regularity results can be found in [13]. An existence theory for related viscoelastic models (Peterlin class) was developed, e.g., in [29]. However, for these models, the energy storage mechanism depends only on the spherical part of the extra stress, which is a major simplification compared to our case. A notable exception is the thesis [24], where the author obtains a global weak solution to an Oldroyd-like diffusive model under certain growth assumptions on the material coefficients. Furthermore, there are existence results for viscoelastic models involving various approximations that improve properties of the system, see e.g. [2] or [25]. Next, the article [4] contains the existence theory for viscoelastic diffusive Oldroyd-B or Giesekus models. This result, however, relies on a certain physical correction of the energy storage mechanism away from the stress-free state and thus improving the a priori estimates of the system. Moreover, various modifications of the classical Oldroyd-B model are discussed in [11]. There are also existence results that are of local nature or for small (initial) data. Local-in-time existence of regular solutions to a viscoelastic Oldroyd-B model without diffusion was shown in [21]. It is also proved there that for small data there exists a global in time solution. For the steady case of a generalized Oldroyd-B model with small and regular data, see e.g. [1].

2. Derivation of the model from physical principles

In this section, we take a closer look at the physical interpretation of the system (1.1)–(1.16). In fact, we show that this system can be obtained quite naturally just by specifying two scalar quantities representing the energy storage & dissipation mechanisms of the fluid. For the origins of this approach, we refer to [40]. Based on these ideas, there is plenty of literature dealing with the thermodynamical derivation of various viscoelastic models, see e.g. [14], [22], [32], [34], [35]. However, since none of these studies can be directly applied to derive the model which we have in mind, we briefly recreate the procedure here.

For the convenience of the reader, we review the meaning of all physical quantities that appeared so far by making the following list:

$\boldsymbol{v}$	flow velocity,
p	pressure,
$\mathbb{B}$	elastic stress tensor,
θ	temperature,
e	internal energy,
E	total energy,
η	entropy,
ξ	rate of entropy production,
μ	coefficient of shear modulus,
c_v	heat capacity,
a	objective derivative coeficient,
α	boundary friction coeficient,
ν	kinematic viscosity,
κ	thermal conductivity,
λ	stress diffusion coeficient,
$\boldsymbol{f}$	external body forces.

Next, we introduce some more quantities, which we then use to rewrite the system (1.7)–(1.12) into a more compact and clearer form.

We remark that the constraint (1.7) is a consequence of the incompressibility of the fluid and of the balance of mass. Moreover, since we assume that the fluid is also homogeneous, the density of the fluid ρ is constant. We take advantage of this by renormalizing all the other quantities so that $\rho = 1$ is not visible in the system (1.8)–(1.12).

The first two terms in each equation (1.8)–(1.12) represent the material derivative, defined by

$$\dot{\boldsymbol{u}} := \partial_t \boldsymbol{u} + \boldsymbol{v} \cdot \nabla \boldsymbol{u}. \tag{2.1}$$

Moreover, looking at the right hand side of (1.9), we recognize the terms that are characteristic for the objective derivative, defined as

$$\overset{\circ}{\mathbb{B}} := \dot{\mathbb{B}} - (\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) - a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \quad a \in \mathbb{R}. \tag{2.2}$$

Unlike the material derivative $\dot{\mathbb{B}}$, the objective derivative $\overset{\circ}{\mathbb{B}}$ (for any a) transforms correctly (as a tensor) under a time dependent rotation of the observer. When $a \in [-1, 1]$, then $\overset{\circ}{\mathbb{B}}$ is precisely the Gordon-Schowalter derivative (cf. [20]). It is known (see e.g. [39]) that by modifying the value of a, it is possible to capture a shear-thinning behaviour of the fluid. The case $a = 0$ leads to the class of models with the corrotational objective derivative (cf. [45]), which has very special properties that simplify the analysis. The case $a = 1$ in (2.2) coincides with the upper-convected objective derivative, which is probably the most popular choice in the literature. In any case, using definition (2.2), equation (1.9) becomes simply

$$\overset{\circ}{\mathbb{B}} + P(\theta, \mathbb{B}) = \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}), \tag{2.3}$$

which is the generalized (due to an implicit form of P) Johnson-Segalman viscoelastic model with stress diffusion and temperature dependent material parameters (P and λ). Depending on the choice of a and P, we can easily recover simpler models, such as the diffusive Giesekus or Oldroyd-B models.

Let us turn our attention to equation (1.8), which represents the balance of linear momentum. Indeed, defining the Cauchy stress tensor by

$$\mathbb{T} := -p\mathbb{I} + 2\nu(\theta)\mathbb{D}\boldsymbol{v} + 2a\mu\theta\mathbb{B} \tag{2.4}$$

and using (2.1), equation (1.8) takes the standard form

$$\dot{\boldsymbol{v}} = \operatorname{div}\mathbb{T} + \boldsymbol{f}.$$

Definition (2.4) is a constitutive relation which is quite typical for viscoelastic fluids: the first part $-p\mathbb{I} + 2\nu(\theta)\mathbb{D}\boldsymbol{v}$ is a standard stress response of a viscous Newtonian fluid, while the second part $2a\mu\theta\mathbb{B}$ models the forces arising due to the elastic (i.e. reversible) part of the total deformation. Assuming that a decomposition of the total deformation of the fluid into a viscous and elastic part is indeed possible, the tensor $\mathbb{B}$ itself can be interpreted as the left Cauchy-Green tensor corresponding to the elastic deformation. In this setting, the tensor $\mathbb{B}$ is sometimes denoted by $\mathbb{B}_{\kappa_p(t)}$, referring to a stress-free configuration at a time t. For more details, we refer to [40]. To avoid certain singular elastic deformations and to ensure a viable physical meaning, it is necessary that $\mathbb{B}$ is a positive definite matrix, which is the restriction (1.2).

Next, defining the fluxes of internal energy and of entropy by

$$\boldsymbol{j}_e := -\kappa(\theta)\nabla\theta \qquad \text{and} \qquad \boldsymbol{j}_\eta := -\kappa(\theta)\nabla\ln\theta + \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}) \tag{2.5}$$

and using (2.1), (2.4), (1.7), we can rewrite equations (1.10)–(1.12) as

$$\dot{e} + \operatorname{div}\boldsymbol{j}_e = \mathbb{T}\cdot\mathbb{D}\boldsymbol{v}, \tag{2.6}$$

$$\dot{E} + \operatorname{div}\boldsymbol{j}_e = \operatorname{div}(\mathbb{T}\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v}, \tag{2.7}$$

$$\dot{\eta} + \operatorname{div}\boldsymbol{j}_\eta = \xi, \tag{2.8}$$

representing the balances of internal energy, total energy and entropy, respectively.

We thus see that the equations of system (1.8)–(1.12) are merely balance equations combined with the constitutive assumptions (2.4), (2.5) and (1.6). In fact, also the equation (2.3) itself can be seen as a constitutive relation. Below we shall demonstrate that the choice we made in (2.3), (2.4), (2.5) and also in (1.3) is not completely ad-hoc, but can be seen as a consequence of (1.6) and of the single assumption that the underlying Helmholtz free energy of the fluid is given by

$$\psi := -c_v\theta\ln\tfrac{\theta}{\theta_0} + \theta\psi_e(\mathbb{B}), \tag{2.9}$$

where

$$\psi_e := \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}). \tag{2.10}$$

Therefore, we are now specifying just two scalar functions: The Helmholtz free energy ψ characterizes how the fluid stores the energy, while the rate of entropy production ξ describes the dissipation mechanisms. Let us now explain our motivation to choose (1.6) and (2.9) in particular.

The first two terms of ξ correspond to the standard heat conducting Newtonian fluid. The next two terms of ξ describe the the energy dissipation due to the elastic deformation. In particular, the last term introduces the stress diffusion term into equation (2.3). It is useful to note that the form of these terms is related to our choice of ψ or, more precisely, to the function ψ_e. Indeed, using basic identities of matrix calculus, we get

$$\psi_e'(\mathbb{B}) = \mu(\mathbb{I} - \mathbb{B}^{-1}), \qquad \psi_e''(\mathbb{B}) = \mu\mathbb{B}^{-1} \otimes \mathbb{B}^{-1} \tag{2.11}$$

and, consequently, also

$$\mu|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2 = \mu\mathbb{B}^{-1}\nabla\mathbb{B}\mathbb{B}^{-1} \cdot \nabla\mathbb{B} = \psi_e''(\mathbb{B})\nabla\mathbb{B} \cdot \nabla\mathbb{B}.$$

Therefore, (1.6) becomes

$$\xi = \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla \ln\theta|^2 + P(\theta,\mathbb{B}) \cdot \psi_e'(\mathbb{B}) + \lambda(\theta)\psi_e''(\mathbb{B})\nabla\mathbb{B} \cdot \nabla\mathbb{B}. \tag{2.12}$$

This can be further rewritten, using the product rule, as

$$\begin{aligned}\xi = \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + (P(\theta,\mathbb{B}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B})) \cdot \psi_e'(\mathbb{B}) + \kappa(\theta)\nabla\theta \cdot \frac{\nabla\theta}{\theta^2}\\ + \operatorname{div}(\lambda(\theta)\nabla\psi_e'(\mathbb{B})).\end{aligned} \tag{2.13}$$

We remark that there seems to be no agreement on what is the correct expression for ξ in viscoelastic fluids with stress diffusion and different choices can lead to physically sound models as well. On the other hand, the assumption (2.13) leads to the simplest possible (linear in $\mathbb{B}$) form of the stress diffusion term, which is an advantage not only in analysis, but apparently also in physics, see [15] and [27].

Next, let us comment on (2.9). Together with the fundamental thermodynamical identities

$$e = \psi + \theta\eta \qquad \text{and} \qquad \eta = -\partial_\theta\psi, \tag{2.14}$$

it implies (1.3) since

$$e = -c_v\theta\ln\frac{\theta}{\theta_0} + \theta\psi_e(\mathbb{B}) + c_v\theta\ln\frac{\theta}{\theta_0} + c_v\theta - \theta\psi_e(\mathbb{B}) = c_v\theta.$$

In fact, the assumption (2.9) is the most general one, for which there is a linear relation between e and θ. Although we could also work with the function $\psi = \psi(\theta,\mathbb{B})$ implicitly (assuming concavity in θ, convexity in $\mathbb{B}$, certain growth etc.), we refrain from doing so here as it would complicate the analysis in Section 3 significantly. The issue is that if $e = c_v\theta$ does not hold, then one has to put forth an additional effort to invert the relation $e = e(\theta,\mathbb{B})$, cf. discussion in [22, Section 2]. In [5] this was solved by linearizing $e = e(\theta,\mathbb{B})$ near $\theta = 0$ and eventually removing this approximation in a limit. From this point of view, our existence result can be seen as basic step towards the existence theory for the general case with $\psi = \psi(\theta,\mathbb{B})$.

Let us now compute ξ from the balance equations and (2.9). From (2.14), we get

$$\dot{e} = \dot{\psi} + \dot{\theta}\eta + \theta\dot{\eta} = \partial_\theta\psi\dot{\theta} + \partial_\mathbb{B}\psi \cdot \dot{\mathbb{B}} - \partial_\theta\psi\dot{\theta} + \theta\dot{\eta} = \theta\dot{\mathbb{B}} \cdot \psi_e'(\mathbb{B}) + \theta\dot{\eta}$$

and using this together with (2.2), (2.3), (2.6) and (2.8), we arrive at

$$\begin{aligned}\xi &= \frac{1}{\theta}\mathbb{T}\cdot\mathbb{D}\boldsymbol{v} - \dot{\mathbb{B}}\cdot\psi_e'(\mathbb{B}) - \frac{1}{\theta}\operatorname{div}\boldsymbol{j}_e + \operatorname{div}\boldsymbol{j}_\eta \\ &= \frac{1}{\theta}\mathbb{T}\cdot\mathbb{D}\boldsymbol{v} - \overset{\circ}{\mathbb{B}}\cdot\psi_e'(\mathbb{B}) - 2a\mathbb{B}\mathbb{D}\boldsymbol{v}\cdot\psi_e'(\mathbb{B}) - \frac{\nabla\theta}{\theta^2}\cdot\boldsymbol{j}_e + \operatorname{div}\Big(\boldsymbol{j}_e - \frac{1}{\theta}\boldsymbol{j}_e\Big) \\ &= \frac{1}{\theta}(\mathbb{T} - 2a\theta\mathbb{B}\psi_e'(\mathbb{B}))\cdot\mathbb{D}\boldsymbol{v} - \overset{\circ}{\mathbb{B}}\cdot\psi_e'(\mathbb{B}) - \frac{\nabla\theta}{\theta^2}\cdot\boldsymbol{j}_e + \operatorname{div}\Big(\boldsymbol{j}_e - \frac{1}{\theta}\boldsymbol{j}_e\Big).\end{aligned}$$

Comparing this with (2.13), it is easy to see that the simplest way to make the both expressions for ξ equivalent is by prescribing the constitutive relations (2.3), (2.4) and (2.5). This kind of identification can be interpreted physically using the principle of maximum rate of entropy production, see [41].

3. The mathematical analysis of the model

In this principal section of the paper we study the mathematical properties of the system (1.1)–(1.16). First, we need to fix some notation. The sets of symmetric, positive definite and positive semi-definite matrices, respectively, are defined by

$$\begin{aligned}\mathbb{R}^{d\times d}_{\mathrm{sym}} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d} : \mathbb{A} = \mathbb{A}^T\}, \\ \mathbb{R}^{d\times d}_{>0} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d}_{\mathrm{sym}} : \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} > 0 \text{ for all } 0\neq\boldsymbol{x}\in\mathbb{R}^d\}, \\ \mathbb{R}^{d\times d}_{\geq 0} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d}_{\mathrm{sym}} : \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} \geq 0 \text{ for all } \boldsymbol{x}\in\mathbb{R}^d\}\end{aligned}$$

respectively. In the special case $d = 1$, we set $\mathbb{R}_{>0} := \mathbb{R}^{1\times 1}_{>0} = (0,\infty)$ and $\mathbb{R}_{\geq 0} := \mathbb{R}^{1\times 1}_{\geq 0} = [0,\infty)$. We use the symbol $\cdot$ to denote the standard inner product in any space of finite dimension greater than one. The outer product is denoted by $\otimes$ and it is used both on $\mathbb{R}^d$ or $\mathbb{R}^{d\times d}$. Further, the symbol $|\cdot|$ denotes the Euclidian norm of any finite-dimensional object. We extend the definition of any[2] function $f : \mathbb{R}\to\mathbb{R}$ to $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ in a standard way as follows: As a consequence of the Schur's theorem, any $\mathbb{A}\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$ admits a spectral decomposition $\mathbb{A} = QDQ^T$, where Q is orthogonal and D is diagonal with real entries. Then we set $f(\mathbb{A}) := Qf(D)Q^T$, where $f(D)$ is a diagonal matrix with entries $f(D_{ii})$, $i = 1,\dots,d$ on the diagonal. If the natural domain of f is $\mathbb{R}_{>0}$, we make the same construction with $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ replaced by $\mathbb{R}^{d\times d}_{>0}$.

If not stated otherwise, the set $\Omega\subset\mathbb{R}^d$ is an open bounded set with a Lipschitz boundary (i.e. of the class $\mathcal{C}^{0,1}$) in the sense of [37, Sect. 2.1.1]). Let V be a subset of an Euclidian space. By $(L^p(\Omega;V), \|\cdot\|_p)$ and $(W^{k,p}(\Omega;V), \|\cdot\|_{k,p})$, $1\leq p\leq\infty$, $k\in\mathbb{N}$, we denote the Lebesgue and Sobolev spaces of functions $u : \Omega\to V$, with their usual norms. In certain situations, we use the notation $\|\cdot\|_{p;\Omega}$ instead to clarify which domain is considered in the norm. The standard inner product in $L^2(\Omega;V)$ and $L^2(\partial\Omega;V)$ is often abbreviated to $(\cdot,\cdot)$ and $(\cdot,\cdot)_{\partial\Omega}$, respectively. The meaning of the duality pairing $\langle\cdot,\cdot\rangle$, is always understandable from the context. Further, if $p > 1$, we set $W^{-k,p}(\Omega;V) := (W^{k,p'}(\Omega;V))^*$, where $p' := p/(p-1)$, $k\in\mathbb{N}$, and the star symbol "*" denotes the topological (continuous) dual space. Since the velocity $\boldsymbol{v}$ in our problem is constrained by $\operatorname{div}\boldsymbol{v} = 0$ in bulk and by $\boldsymbol{v}\cdot\boldsymbol{n} = 0$ on

[2] Except for $|\cdot|$.

the boundary, we introduce the following subspaces:

$$\begin{aligned}
W^{k,p}_{\boldsymbol{n}} &\coloneqq \{\boldsymbol{u} \in W^{k,p}(\Omega;\mathbb{R}^d) : \boldsymbol{u}\cdot\boldsymbol{n} = 0\}, \quad k\in\mathbb{N}, \quad p<\infty,\\
W^{k,p}_{\boldsymbol{n},\mathrm{div}} &\coloneqq \{\boldsymbol{u} \in W^{k,p}_{\boldsymbol{n}} : \operatorname{div}\boldsymbol{u} = 0\}, \quad k\in\mathbb{N}, \quad p<\infty,\\
W^{-k,2}_{\boldsymbol{n},\mathrm{div}} &\coloneqq (W^{k,2}_{\boldsymbol{n},\mathrm{div}})^*, \quad k\in\mathbb{N},\\
L^2_{\boldsymbol{n},\mathrm{div}} &\coloneqq \overline{W^{1,2}_{\boldsymbol{n},\mathrm{div}}}^{\|\cdot\|_2}
\end{aligned}$$

The expression $\boldsymbol{u}\cdot\boldsymbol{n}$ is, of course, understood in the trace sense and we shall not use any additional notation for traces of Sobolev functions, assuming it is clear from the context.

Let X be a Banach space. The Bochner spaces $L^p(0,T;X)$ with $1\le p\le\infty$ consist of strongly measurable mappings $u:[0,T]\to X$ for which the norm

$$\|u\|_{L^p(0,T;X)} \coloneqq \begin{cases} \left(\displaystyle\int_0^T \|u\|_X^p\right)^{\frac{1}{p}} & \text{if } 1\le p<\infty;\\ \operatorname*{ess\,sup}_{(0,T)}\|u\|_X & \text{if } p=\infty, \end{cases}$$

is finite. If $X = L^q(\Omega;V)$ or $X = W^{k,q}(\Omega;V)$, with $1\le q\le\infty$, $n\in\mathbb{N}$, we shorten the notation and use the symbols $\|\cdot\|_{L^pL^q}$ or $\|\cdot\|_{L^pW^{k,q}}$, respectively, for the corresponding norms. The space $\mathcal{C}([0,T];X)$, containing continuous X-valued functions on $[0,T]$, is equipped with the norm

$$\|u\|_{\mathcal{C}([0,T];X)} \coloneqq \sup_{t\in[0,T]}\|u(t)\|_X.$$

Furthermore, the space of weakly continuous functions is defined by

$$\mathcal{C}_w([0,T];X) \coloneqq \big\{u\in L^\infty(0,T;X) : \text{the function } \langle g,u\rangle \text{ is continuous in } [0,T] \text{ for every } g\in X^*\big\}.$$

When manipulating products of matrix-valued functions, we use a simplified index-free notation, which is always easily understandable in given context. To be on the safe side, we warn the reader that we write

$\nabla\mathbb{A}\cdot\nabla\mathbb{B}$	instead of	$\sum_{i,j,k}\partial_i\mathbb{A}_{jk}\partial_i\mathbb{B}_{jk}$,
$(\boldsymbol{v}\otimes\mathbb{A})\cdot\nabla\mathbb{B}$	instead of	$\sum_{i,j,k}\boldsymbol{v}_i\mathbb{A}_{jk}\partial_i\mathbb{B}_{jk}$,
$\lvert\mathbb{A}\nabla\mathbb{B}\mathbb{C}\rvert^2$	instead of	$\sum_{i,j,k}(\mathbb{A}_{il}\partial_k\mathbb{B}_{lm}\mathbb{C}_{mj})^2$,
$(\mathbb{A}\otimes\mathbb{B})\nabla\mathbb{C}$	instead of	$\sum_{m,n}\mathbb{A}_{ij}\mathbb{B}_{mn}\partial_k\mathbb{C}_{mn}$,

for example, and so on.

Assumptions on material coefficients. The mathematical properties of the system (1.1)–(1.16) depend crucially on the behaviour of the material coefficients, which were not yet specified. We will require that the coefficients

$$\nu,\kappa,\lambda,P \quad \text{are continous functions in } \mathbb{R},\mathbb{R},\mathbb{R} \text{ and } \mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}, \text{ respectively,} \tag{3.1}$$

and admit the following growth conditions for some numbers $q, r > 0$ and constants $C, C_\alpha > 0$:

$$
\begin{aligned}
C^{-1} &\le \nu(s) \le C && \text{for all } s > 0, && (3.2)\\
C^{-1}(1+s^r) &\le \kappa(s) \le C(1+s^r) && \text{for all } s > 0, && (3.3)\\
C^{-1} &\le \lambda(s) \le C && \text{for all } s > 0, && (3.4)\\
P(s,\mathbb{A}) &= P(s,\mathbb{A})^T && \text{for all } s > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, && (3.5)\\
|P(s,\mathbb{A})| &\le C(1+|\mathbb{A}|^{q+1}) && \text{for all } s > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, && (3.6)\\
P(s,\mathbb{A})\cdot\mathbb{A}^\alpha &\ge C_\alpha|\mathbb{A}|^{q+1+\alpha} - C && \text{for all } s,\alpha > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}, && (3.7)\\
P(s,\mathbb{A})\cdot\mathbb{I} &\ge -C && \text{for all } s > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}, && (3.8)\\
P(s,\mathbb{A})\cdot(\mathbb{I}-\mathbb{A}^{-1}) &\ge 0 && \text{for all } s > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}. && (3.9)
\end{aligned}
$$

In addition to that, we need one more technical assumption concerning the behaviour of P for "nearly" singular matrices[3]: There exists $\omega_P > 0$ such that

$$
P(\cdot,\mathbb{A}+\omega_P\mathbb{I})\boldsymbol{x}\cdot\boldsymbol{x} \le 0 \quad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ and } \boldsymbol{x} \in \mathbb{R}^d \text{ such that } \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} \le 0. \quad (3.10)
$$

Assumption (3.2) is quite standard for fluids. Restriction (3.3) means that κ is a bounded function near zero and has an r-growth near infinity. In literature, one can find instances where the number r is chosen relatively large, which then helps the analysis. Assumption (3.4) is chosen just for simplicity. Condition (3.5) is an obvious necessary restriction should (1.9) hold. Assumptions (3.6) and (3.7) mean that $P(\cdot,\mathbb{A})$ behaves asymptotically as $\mathbb{A}^{q+1}$, which is a crucial information to get sufficient a priori estimates. Condition (3.8) simplifies the analysis at one step and means basically that the leading order term of $P(\cdot,\mathbb{A})$ appears with the positive sign, compare e.g. with the Giesekus model, where $P(\cdot,\mathbb{A}) = \mathbb{A}^2 - \mathbb{A}$. Property (3.9) is important for the validity of the second law of thermodynamics in our model. Roughly it means that $P(\cdot,\mathbb{A})$ should be derived from (or at least share the direction with) the quantity $\mathbb{A}-\mathbb{I}$ (and not $\mathbb{I}-\mathbb{A}$). An explicit example of function P satisfying (3.5)–(3.9) would be

$$
P(s,\mathbb{A}) = \delta(s)(1+|\mathbb{A}-\mathbb{I}|^{q-\beta})\mathbb{A}^\beta(\mathbb{A}-\mathbb{I}),
$$

where δ is a continuous positive real function and $\beta \in [0,q]$. Indeed, note that, for any $\mathbb{A} \in \mathbb{R}^{d\times d}_{>0}$, we can write

$$
\begin{aligned}
\mathbb{A}^\beta(\mathbb{A}-\mathbb{I})\cdot(\mathbb{I}-\mathbb{A}^{-1}) &= \mathbb{A}^{\frac{\beta}{2}}\mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})\mathbb{A}^{\frac12}\cdot(\mathbb{I}-\mathbb{A}^{-1})\\
&= \mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})\cdot\mathbb{A}^{\frac{\beta}{2}}(\mathbb{I}-\mathbb{A}^{-1})\mathbb{A}^{\frac12} = |\mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})|^2 \ge 0,
\end{aligned}
$$

implying (3.9). Next, properties (3.6), (3.7) and (3.8) follow easily from (4.20) below. Finally, we claim that (3.10) holds with $\omega_P = 1$. Indeed, let $0 \ne \boldsymbol{x} \in \mathbb{R}^d$ be an eigenvector of $\mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, for which $\lambda := \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}/|\boldsymbol{x}|^2 \le 0$. If $\mathbb{A}+\mathbb{I} \notin \mathbb{R}^{d\times d}_{>0}$ then we can redefine $P(\cdot,\mathbb{A}+\mathbb{I})$ as needed. Otherwise, we have $\mathbb{A}+\mathbb{I} \in \mathbb{R}^{d\times d}_{>0}$, and thus

[3]Although $P(\cdot,\mathbb{A})$ should be initially defined in some way also for $\mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \setminus \mathbb{R}^{d\times d}_{>0}$, such a definition becomes irrelevant once we prove $\mathbb{B} \in \mathbb{R}^{d\times d}_{>0}$. A similar remark applies for the temperature dependent coefficients.

$\lambda > -1$ and we can write

$$P(s, \mathbb{A}+\mathbb{I})\boldsymbol{x}\cdot\boldsymbol{x} = \delta(s)(1+|\mathbb{A}|^{q-\beta})(\mathbb{A}+\mathbb{I})^{\beta}\mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}$$
$$= \delta(s)(1+|\mathbb{A}|^{q-\beta})(\lambda+1)^{\beta}\lambda|\boldsymbol{x}|^2 \le 0.$$

In addition to the parameters q and r, we also introduce the parameter ρ as the integrability exponent of the initial datum for $\mathbb{B}$, i.e., the largest number ρ, for which

$$\int_\Omega |\mathbb{B}_0|^{\rho} < \infty.$$

It is to be expected that increasing the values of q, r and ϱ improves the mathematical properties of the system (1.1)–(1.16). This seems to be true, at least up to a certain threshold: For example, it is unclear whether the case $\rho > q$ provides some improvement compared to the case $\rho = q$. Also, it seems very unlikely in our setting that one would be able to prove any better information than $\mathbb{D}\boldsymbol{v} \in L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$. To be able to quantify this precisely, let us now introduce certain conditions which play a key role in our main result below.

Conditions on q, r and ϱ. To make sure that the quantities below are well defined, we need to restrict the parameters q, r and ϱ by the condition

$$r > 1 - \frac{2}{d}, \qquad q > 1, \qquad \varrho > 1 \tag{C$_0$}$$

(we recall that $d \ge 2$ is the dimension of the domain Ω). Let us define

$$\sigma := \min\{q, \varrho\} \qquad \text{and} \qquad r_d := r + \frac{2}{d}. \tag{3.11}$$

The conditions

$$(r_d - 1)(q-1) > 2, \tag{C$_1$}$$

$$(r_d - 1)\left(q + \sigma - \frac{2d}{d+2}\right) > \frac{4d}{d+2} \tag{C$_2$}$$

are sufficient to define every term appearing in (1.22)–(1.25) in a weak sense. As such, they are actually sufficient for the existence of a weak solution, which is the content of our main result.

Further, the conditions

$$(r_d - 1)(q + \sigma - 2) > 4 - \frac{2}{d}, \tag{C$_1^E$}$$

$$(r_d - 1)\left(q + \sigma - \frac{3d}{d+2}\right) > \frac{6d}{d+2} \tag{C$_2^E$}$$

(together with some regularity assumptions on Ω) are sufficient for the local balance of the total energy (1.11) to hold if $d = 2$ or $d = 3$. In the case $d \ge 4$, it is unclear if $|\boldsymbol{v}|^2\boldsymbol{v} \in L^1(Q;\mathbb{R}^d)$, for any choice of r, q and ϱ.

Finally, the condition

$$(r_d - 1)(q + \sigma - 2) > 4 \tag{C$^\theta$}$$

is sufficient for the validity of (1.21).

Since $\sigma \le q$, it is obvious that (C$^\theta$) implies (C$_1$) and (C$_2$), written symbolically as

$$(\mathrm{C}^\theta) \Rightarrow (\mathrm{C}_1) \wedge (\mathrm{C}_2).$$

Moreover, if $d \le 3$, it is easy to see that the relation

$$(\mathrm{C}^\theta) \Rightarrow (\mathrm{C}_1^E) \wedge (\mathrm{C}_2^E) \Rightarrow (\mathrm{C}_1) \wedge (\mathrm{C}_2)$$

is valid. Instead of (C_1), we often use one of its equivalent versions:

$$r_d > \frac{q+1}{q-1}, \quad r_d > 2q' - 1, \quad \text{or} \quad r_d' < \frac{q+1}{2}.$$

Furthermore, defining

$$r_0 := \frac{q+1}{q-1} \quad \text{and} \quad r_1 := \frac{q+\sigma+2}{q+\sigma-2},$$

we observe that

$$(\mathrm{C}_1) \quad \text{is equivalent to} \quad r_d > r_0 \quad \text{and} \quad (\mathrm{C}^\theta) \quad \text{is equivalent to} \quad r_d > r_1.$$

Let us make one important remark on the assumptions above. By imposing any of the conditions (C_0)–(C^θ) and (3.2)–(3.6), we place some restrictions on the coefficients of the model which may not agree with experimental measurements. Thus, one could argue that this renders our analysis useless in the actual applications. However, if we only care about existence of a weak solution, without any quantitative criteria, then these assumptions are not that important, from the physical point of view. Indeed, note that (3.2)–(3.6) restrict only the asymptotic behaviour of the coefficients. For example, any continuous function κ defined on some interval (θ_0, θ_1), $0 < \theta_0 < \theta_1 < \infty$, can be modified in a neighbourhood of 0 and ∞ so that (3.3) holds. The interval (θ_0, θ_1) may represent the temperature range for which the model we are considering makes sense. When the fluid starts to freeze or boil, then we are clearly outside this range and it makes no sense to prescribe the coefficients ν, κ, δ and λ there. On the other hand, it is unclear whether one can deduce some absolute bounds for the temperature, besides $\theta > 0$, using only the information that is encoded in the system. Thus, purely for mathematical reasons, we have to assume that these material coefficients are defined in some way also outside (θ_0, θ_1). A similar remark applies also for the coefficient $P(\cdot, \mathbb{A})$. If $|\mathbb{A}|$ is too large, any realistic material eventually breaks down. Thus, we may set $P(\cdot, \mathbb{A}) = \mathbb{A} - \mathbb{I}$, $|\mathbb{A}| \in [0, M)$, where M is large (to mimic the Oldroyd-B model, for example) and then extend this function continuously so that (3.7) holds with some large q.

A priori estimates. Let us now explain the motivation behind the conditions (C_1)–(C^θ) by an informal derivation of the a priori estimates corresponding to the system. These also indicate what function spaces should be used to define a weak solution. As the precise computation is carried out in the proof of our main result below, this section serves only as a reader's guide through the main ideas used in the formal proof below.

The obvious starting points are the energy equality (1.25) and the entropy inequality (1.24). The first one gives, using $\boldsymbol{f} \in L^2(0,T; L^2(\Omega; \mathbb{R}^d))$ and Gronwall's inequality, that

$$\boldsymbol{v} \in L^\infty(0,T; L^2(\Omega; \mathbb{R}^d)) \quad \text{and} \quad \theta \in L^\infty(0,T; L^1(\Omega; \mathbb{R})). \tag{3.12}$$

Next, since ξ is non-negative and the left hand side of (1.24) is in a divergence form, we deduce that

$$\xi \in L^1(Q; \mathbb{R}) \tag{3.13}$$

simply by integrating (1.24) and using $\eta_0 \in L^1(\Omega;\mathbb{R})$ and boundary conditions. Information (3.13) is important to show that a weak solution meets the physical requirements (1.1) and (1.2). Unfortunately, estimates (3.12) and (3.13) do not provide enough information to even define some terms appearing in (1.22) and (1.23). Nevertheless, with the assumptions made above, it is possible to improve this information.

Thanks to the results of [3], it turns out that to a large extent, we can treat (1.23) as a scalar equation and test it by $\mathbb{B}^{\sigma-1}$. Then, using (3.7), Young's inequality and Lemma 3 below, we can eventually get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \operatorname{tr}\mathbb{B}^\sigma + \int_\Omega |\mathbb{B}|^{q+\sigma} + \int_\Omega |\nabla \mathbb{B}^{\frac{\sigma}{2}}|^2 \le C\int_\Omega |\mathbb{B}|^\sigma |\mathbb{D}\boldsymbol{v}| + C \le C\int_\Omega |\mathbb{D}\boldsymbol{v}|^{\frac{q+\sigma}{q}} + C,$$

hence

$$\|\mathbb{B}\|_{q+\sigma;Q} \le C\|\mathbb{D}\boldsymbol{v}\|^{\frac{1}{q}}_{\frac{q+\sigma}{q};Q} + C. \tag{3.14}$$

To make the last estimate usable, we need to get sufficient control over $\mathbb{D}\boldsymbol{v}$. This is the point where the first crucial difference, compared to the existence theory for the Navier-Stokes-Fourier systems, appears. There, one is able to deduce an apriori estimate for $\nabla\boldsymbol{v}$ simply by multiplying the momentum equation (1.22) by $\boldsymbol{v}$ and absorbing the right hand side. This is generally not possible in our case as we get also the term $\operatorname{div}(\theta\mathbb{B})\cdot\boldsymbol{v}$ on the right hand side, which is difficult to control. To eliminate it, one has to add the balance of internal energy (1.21), leading only to (1.25), however. Therefore, to obtain an estimate for $\mathbb{D}\boldsymbol{v}$, we have to look elsewhere. We note that in (1.21), the viscous dissipation term appears in the form $2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2$ on the right hand side. Thus, it is a good idea to test (1.21) (on an approximate level, where (1.21) makes sense) by the function $-\theta^{-\beta}$ with $\beta > 0$ as small as possible. Eventually, this leads to the estimate

$$(r+1-R)\int_Q |\nabla\theta^{\frac{R}{2}}|^2 + \int_Q \frac{|\mathbb{D}\boldsymbol{v}|^2}{\theta^{r+1-R}} \le C, \tag{3.15}$$

where

$$R = \frac{(r+\frac{2}{d}-1)(q+\sigma)}{2} - \frac{2}{d}.$$

From this, using Hölder's inequality and Sobolev embeddings, we deduce

$$\|\mathbb{D}\boldsymbol{v}_\ell\|_{p;Q} \le \|\theta_\ell^{\frac{R-r-1}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell\|^{\frac{r+1-R}{2}}_{R+\frac{2}{d};Q} \le C,$$

where

$$p = 2 - 2\frac{r+1-R}{r_d+1} = (q+\sigma)\frac{r_d-1}{r_d+1} = \frac{q+\sigma}{2r_d'-1}.$$

Looking back at (3.14), we need to ensure that

$$\frac{q+\sigma}{q} < (q+\sigma)\frac{r_d-1}{r_d+1}$$

(the strict inequality gains compactness), which is equivalent to (C_1). If this condition is satisfied, the above ideas together with some interpolation arguments and the use of (1.1)–(1.6) lead to the following a priori information:

$$\boldsymbol{v} \in L^{\infty}(0,T;L^2(\Omega;\mathbb{R}^d)), \tag{3.16}$$

$$\nabla\boldsymbol{v} \in L^{p}(0,T;L^p(\Omega;\mathbb{R}^{d\times d})), \tag{3.17}$$

$$\boldsymbol{v} \in L^{p\frac{d+2}{d}}(0,T;L^{p\frac{d+2}{d}}(\Omega;\mathbb{R}^d)), \tag{3.18}$$

$$\mathbb{B} \in L^{\infty}(0,T;L^{\sigma}(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.19}$$

$$\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}} \in L^{2}(0,T;L^2(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{3.20}$$

$$\nabla\mathbb{B}^{\frac{\sigma}{2}} \in L^{2}(0,T;L^2(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{3.21}$$

$$\mathbb{B} \in L^{q+\sigma}(0,T;L^{q+\sigma}(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.22}$$

$$\theta \in L^{\infty}(0,T;L^1(\Omega;\mathbb{R}_{>0})), \tag{3.23}$$

$$\nabla\theta^{\frac{R}{2}} \in L^{2}(0,T;L^2(\Omega;\mathbb{R}^d)), \tag{3.24}$$

$$\theta \in L^{R+\frac{2}{d}}(0,T;L^{R+\frac{2}{d}}(\Omega;\mathbb{R}_{>0})). \tag{3.25}$$

Let us now check that (3.16)–(3.25) are sufficient to define every term of the system (1.22)–(1.25) in a weak sense. Obviously, it is enough to focus on the non-linear terms. Since $\boldsymbol{v}\cdot\nabla\boldsymbol{v} = \operatorname{div}(\boldsymbol{v}\otimes\boldsymbol{v})$, we need to ensure that $\boldsymbol{v} \in L^{2+\varepsilon}(Q;\mathbb{R}^d)$ for some $\varepsilon > 0$ which, looking at (3.18) translates as

$$p\frac{d+2}{d} > 2.$$

This condition is actually equivalent to (C_2). It turns out that conditions (C_0), (C_1) and (C_2) are really the only restrictions on the parameters q, r, ϱ needed for the existence of a weak solution. To see this, we proceed with estimation of the non-linear terms. On the right hand side of (1.23) we have the product $\theta\mathbb{B}$ (under the divergence), which (by Hölder's inequality) is integrable if

$$\frac{1}{R+\frac{2}{d}} + \frac{1}{q+\sigma} = \frac{1}{p} < 1,$$

cf. (3.25) and (3.22), which is true (due to (C_1) or (C_2)). Next, in equation (1.23), the most irregular term is obviously $\mathbb{B}\nabla\boldsymbol{v}$. Since, by ($C_1$), we have

$$\frac{1}{p} + \frac{1}{q+\sigma} = \frac{1}{q+\sigma}\left(\frac{r_d+1}{r_d-1}+1\right) < \frac{q+1}{q+\sigma} < 1,$$

these terms are integrable as well. Regarding the entropy inequality (1.24), the convective term is integrable as it turns out that η is even better than square integrable. Thus, we only need to check that $\kappa(\theta)\nabla\ln\theta$ is integrable, which is true since, using (3.3), we have

$$\|\kappa(\theta)\nabla\ln\theta\|_{1;Q} \le \|\sqrt{\kappa(\theta)}\|_{2;Q}\|\sqrt{\kappa(\theta)}\nabla\ln\theta\|_{2;Q} \le C\|\theta\|^{\frac{r}{2}}_{r;Q} + C$$

and $r < R_d$ due to (C_1). Hence we see that all the terms of (1.22)–(1.25) are well defined.

All this is true if (C_1), (C_2) hold and r is not too large. Note that the key estimate (3.15) degenerates if $r+1 > R$, which happens precisely if r is so large that the condition (C^{θ}) becomes true. The condition (C^{θ}) defines a certain sub-critical case, where the existence proof becomes easier, but the derivation of estimate (3.15) (and subsequent conclusions) must be modified accordingly. Let us also remark that in our setting, the condition (C^{θ}) holds if and only if $p = 2$. In other words, the condition (C^θ) determines whether the viscous dissipation term $2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2$ is integrable or not. Actually, also the second non-linear term appearing in (1.21), that is $\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$, is integrable if ($\mathrm{C}^\theta$) is true. We thus conclude that (C^θ) determines whether (1.21) is meaningful or not.

Definition of weak solution. In addition to σ, r_d, r_0 and r_1, let us define further numbers which are useful in the definition of weak solution below. We set

$$p := \begin{cases} \dfrac{q+\sigma}{2r_d'-1} & \text{if} \quad r_d < r_1, \\ 2) & \text{if} \quad r_d = r_1, \\ 2 & \text{if} \quad r_d > r_1, \end{cases} \tag{3.26}$$

where the symbol $x_0)$, $x_0 \in \mathbb{R}$, is used throughout as an abbreviation for any number from a (sufficiently small) left neighbourhood of x_0, excluding x_0. Further, we set

$$R_d := R + \frac{2}{d}, \quad \text{where} \quad R := \begin{cases} \dfrac{(r_d-1)(q+\sigma)}{2} - \dfrac{2}{d} & \text{if} \quad r_d < r_1, \\ r+1) & \text{if} \quad r_d \geq r_1. \end{cases} \tag{3.27}$$

Finally, we define

$$\sigma_0 := \begin{cases} \sigma & \text{if} \quad \sigma < q, \\ \sigma) & \text{if} \quad \sigma = q \end{cases} \quad \text{and} \quad z := \frac{2(q+\sigma)}{q+\sigma+2}. \tag{3.28}$$

Definition 1. *Let $T > 0$ and let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a Lipschitz domain. Assume that the constants $a \in \mathbb{R}$, $\alpha \geq 0$, $c_v, \mu > 0$ and the functions ν, κ, λ, P fulfil the assumptions* (3.1)–(3.9) *with the numbers q, r, ϱ satisfying* (C_0), (C_1), (C_2)*. Let the numbers σ, p, R, R_d, σ_0 and z be defined by* (3.11), (3.26), (3.27) *and* (3.28), *respectively. Suppose that the initial data satisfy*

$$\boldsymbol{v}_0 \in L^2_{\boldsymbol{n},\mathrm{div}}(\Omega;\mathbb{R}^d), \quad \mathbb{B}_0 \in L^\varrho(\Omega;\mathbb{R}^{d\times d}_{>0}), \quad \theta_0 \in L^1(\Omega;\mathbb{R}_{>0}), \tag{3.29}$$

$$\eta_0 := c_v \ln\theta_0 - \mu(\operatorname{tr}\mathbb{B}_0 - d - \ln\det\mathbb{B}_0) \in L^1(\Omega;\mathbb{R}) \tag{3.30}$$

and that

$$\boldsymbol{f} \in L^2(0,T;L^2(\Omega;\mathbb{R}^d)). \tag{3.31}$$

Then, we say that the function $(\boldsymbol{v}, \mathbb{B}, \theta, e, E, \eta) : Q \to \mathbb{R}^d \times \mathbb{R}^{d\times d}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}$ is a weak solution of the initial-boundary value problem (1.1)–(1.9), (1.12)–(1.16) *if all of the following conditions* (I)–(IV) *are satisfied:*

(I) *The functions* $\boldsymbol{v}$, $\mathbb{B}$, θ *and* η *fulfil the properties*

$$\boldsymbol{v} \in L^p(0,T;W^{1,p}_{\boldsymbol{n},\mathrm{div}}) \cap \mathcal{C}_w([0,T];L^2(\Omega;\mathbb{R}^d)), \tag{3.32}$$

$$\partial_t \boldsymbol{v} \in L^{p\frac{d+2}{2d}}(0,T;W^{-1,p\frac{d+2}{2d}}_{\boldsymbol{n},\mathrm{div}}), \tag{3.33}$$

$$\mathbb{B} \in L^z(0,T;W^{1,z}(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap \mathcal{C}_w([0,T];L^\sigma(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.34}$$

$$\mathbb{B} \in L^{q+\sigma}(Q;\mathbb{R}^{d\times d}_{>0}), \tag{3.35}$$

$$\mathbb{B}^{\frac{\sigma}{2}} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.36}$$

$$\partial_t \mathbb{B} \in \left(L^{z'}(0,T;W^{1,z'}(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap L^{\frac{q+\sigma}{\sigma-1}}(Q;\mathbb{R}^{d\times d}_{>0})\right)^*, \tag{3.37}$$

$$\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}} \in L^2(Q;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{3.38}$$

$$\ln\det\mathbb{B} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})), \tag{3.39}$$

$$\theta \in L^\infty(0,T;L^1(\Omega;\mathbb{R}_{>0})) \cap L^{R_d}(Q;\mathbb{R}_{>0}), \tag{3.40}$$

$$\theta^{\frac{R}{2}} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}_{>0})), \tag{3.41}$$

$$\ln\theta \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})), \tag{3.42}$$

$$\eta \in L^z(0,T;W^{1,z}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})). \tag{3.43}$$

(II) *The identities* (1.3)–(1.6) *hold almost everywhere in* Q.
(III) *Equations* (1.22)–(1.25) *are satisfied in the following sense:*

$$\left.\begin{aligned}\int_0^T (\langle\partial_t\boldsymbol{v},\boldsymbol{\varphi}\rangle - (\boldsymbol{v}\otimes\boldsymbol{v},\nabla\boldsymbol{\varphi}) + (2\nu(\theta)\mathbb{D}\boldsymbol{v},\nabla\boldsymbol{\varphi})) + \alpha\int_0^T\int_{\partial\Omega}\boldsymbol{v}_\tau\cdot\boldsymbol{\varphi}_\tau \\ = -\int_0^T(2a\mu\theta\mathbb{B},\nabla\boldsymbol{\varphi}) + \int_0^T(\boldsymbol{f},\boldsymbol{\varphi}) \\ \textit{for all}\quad \boldsymbol{\varphi}\in L^{(p\frac{d+2}{2d})'}(0,T;W^{1,(p\frac{d+2}{2d})'}_{\boldsymbol{n},\mathrm{div}}),\end{aligned}\right\} \tag{3.44}$$

$$\left.\begin{aligned}\int_0^T\langle\partial_t\mathbb{B},\mathbb{A}\rangle - \int_0^T(\mathbb{B}\otimes\boldsymbol{v},\nabla\mathbb{A}) + \int_0^T(P(\theta,\mathbb{B}),\mathbb{A}) + \int_0^T(\lambda(\theta)\nabla\mathbb{B},\nabla\mathbb{A}) \\ = \int_0^T((a\mathbb{D}\boldsymbol{v}+\mathbb{W}\boldsymbol{v})\mathbb{B},\mathbb{A}+\mathbb{A}^T) \\ \textit{for all}\quad \mathbb{A}\in L^{z'}(0,T;W^{1,z'}(\Omega;\mathbb{R}^{d\times d}))\cap L^{\frac{q+\sigma}{\sigma-1}}(Q;\mathbb{R}^{d\times d}),\end{aligned}\right\} \tag{3.45}$$

$$\left.\begin{aligned}&(\eta_0,\phi)\varphi(0) - \int_0^T(\eta,\phi)\partial_t\varphi \\ &+\int_0^T\big(\kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B}-d-\ln\det\mathbb{B}) - \eta\boldsymbol{v},\nabla\phi\big)\varphi \ge \int_0^T(\xi,\phi)\varphi \\ &\textit{for all }\ \varphi\in W^{1,\infty}((0,T);\mathbb{R}_{\ge0}),\ \varphi(T)=0,\ \textit{and all }\ \phi\in W^{1,\infty}(\Omega;\mathbb{R}_{\ge0}),\end{aligned}\right\} \tag{3.46}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\quad \textit{a.e. in } [0,T]. \tag{3.47}$$

(IV) *The initial data are attained in the following way:*

$$\operatorname*{ess\,lim}_{t\to 0_+}\|\boldsymbol{v}(t)-\boldsymbol{v}_0\|_2 = 0, \tag{3.48}$$

$$\lim_{t\to 0_+}\|\mathbb{B}(t)-\mathbb{B}_0\|_{\sigma_0} = 0, \tag{3.49}$$

$$\operatorname*{ess\,lim}_{t\to 0_+}\|\theta(t)-\theta_0\|_1 = 0, \tag{3.50}$$

$$\operatorname*{ess\,lim\,inf}_{t\to 0_+}\int_\Omega \eta(t)\phi \geq \int_\Omega \eta_0\phi \quad \text{for all} \quad 0\leq\phi\in W^{1,\infty}(\Omega). \tag{3.51}$$

It remains to show that a weak solution exists.

The main result. In this section we state and prove our main result.

Theorem 1. *Suppose that all the assumptions of Definition 1 are fulfilled.*

(I) *Then, there exists a weak solution of the system* (1.1)–(1.9), (1.12)–(1.16).

(II) *If, in addition, $d\leq 3$, $\Omega\in\mathcal{C}^{1,1}$ and* (C_1^E), (C_2^E) *hold, then there exists a weak solution of the system* (1.1)–(1.9), (1.12)–(1.16) *and also a function* $\mathrm{p}\in L^{p\frac{d+2}{2d}}(Q;\mathbb{R})$ *such that the local balance of total energy* (1.11) *holds in the sense:*

$$\begin{aligned} -\left(\tfrac{1}{2}|\boldsymbol{v}_0|^2+c_v\theta_0,\phi\right)\varphi(0) - \int_0^T (E,\phi)\partial_t\varphi - \int_0^T (E\boldsymbol{v},\nabla\phi)\varphi \\ + \alpha\int_0^T\int_{\partial\Omega}|\boldsymbol{v}_\tau|^2\phi\varphi + \int_0^T(\kappa(\theta)\nabla\theta,\nabla\phi)\varphi \\ = \int_0^T(\mathrm{p}\boldsymbol{v}-2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v}-2a\mu\theta\mathbb{B}\boldsymbol{v},\nabla\phi)\varphi \end{aligned} \tag{3.52}$$

for all $\varphi\in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T)=0$, *and every* $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$.

(III) *If, in addition, condition* (C^θ) *holds, then there exists a weak solution of the system* (1.1)–(1.9), (1.12)–(1.16) *satisfying the local balance of internal energy* (1.10) (*i.e. the temperature inequality*) *in the sense:*

$$\begin{aligned} -(c_v\theta_0,\phi)\varphi(0) - \int_0^T(\theta,\phi)\partial_t\varphi - \int_0^T(c_v\theta\boldsymbol{v},\nabla\phi)\varphi + \int_0^T(\kappa(\theta)\nabla\theta,\nabla\phi)\varphi \\ \geq \int_0^T(2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2+2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v},\phi)\varphi \end{aligned} \tag{3.53}$$

for all $\varphi\in W^{1,\infty}((0,T);\mathbb{R}_{\geq 0})$, $\varphi(T)=0$, *and all* $\phi\in W^{1,\infty}(\Omega;\mathbb{R}_{\geq 0})$.

***Proof*.** Let us do the proof only for $d\geq 3$ (the case $d=2$ is easier, of course). Also, it is clearly enough to focus on the case $\alpha>0$. In the simpler case $\alpha=0$ (corresponding to the free-slip boundary condition), the boundary term in (3.44) and (3.52) is merely not present (and $\boldsymbol{v}\in L^2(0,T;L^2(\partial\Omega;\mathbb{R}^d))$ may not be true, depending on p).

(I)

General strategy. We approximate the system (1.8), (1.9), (1.10) using several parameters to obtain a proper Galerkin approximation and we show that the resulting (ODE) system has a solution. After that, our aim is to derive the entropy equation. At this point, possibly irregular terms containing θ and $\mathbb{B}$ are cut-off and

$\boldsymbol{v}$ is smooth, hence we easily obtain uniform estimates for the Galerkin approximations of $\mathbb{B}$ and θ, which might not be positive definite or positive, respectively. However, after taking the limit with these approximations and then proving certain maximum principles, we prove invertibility of θ and $\mathbb{B}$, which, in turn, enables us to derive the entropy equation. From this we read that the positivity of $\det \mathbb{B}$ and θ is preserved uniformly, which then enables us to remove the cut-off from the system. We also improve the uniform estimates by considering appropriate test functions in the equations for θ and $\mathbb{B}$. Using these, we pass to the final limit, identify the non-linear terms and initial conditions, hereby obtaining a solution of the original problem.

Approximation. First we introduce a truncation, which is essential for the proof. We also prepare some simple estimates corresponding to this truncation, that are used later in the proof. For any $\omega \in (0, \omega_P)$, let us define the "cut-off" function

$$g_\omega(\mathbb{A}, \tau) \coloneqq \frac{\max\{0, \Lambda(\mathbb{A}) - \omega\} \max\{0, \tau - \omega\}}{(|\Lambda(\mathbb{A})| + \omega)(1 + \omega|\mathbb{A}|^2)(|\tau| + \omega)(1 + \omega\tau^2)}, \quad \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad \tau \in \mathbb{R},$$

where

$$\Lambda(\mathbb{A}) \coloneqq \min\{\lambda : \det(\mathbb{A} - \lambda \mathbb{I}) = 0\},$$

i.e., the smallest eigenvalue of $\mathbb{A}$. Note that g_ω is a continuous function in $\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}$ and satisfies $0 \le g_\omega(\mathbb{A}, \tau) < 1$ for every $(\mathbb{A}, \tau) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}$. Moreover, if $\Lambda(\mathbb{A}) \le \omega$ or $\tau \le \omega$, then $g_\omega(\mathbb{A}, \tau) = 0$, whereas if $\Lambda(\mathbb{A}) > 0$ and $\tau > 0$, then $g_\omega(\mathbb{A}, \tau) \to 1$ as $\omega \to 0_+$. Furthermore, we remark that

$$g_\omega(\mathbb{A}, \tau)(1 + |\mathbb{A}| + |\mathbb{A}|^2)(1 + \tau + \tau^2) \le C(\omega). \tag{3.54}$$

The function g_ω is used below in system (3.64)–(3.66) to handle its irregular terms. Let us truncate also the initial conditions $\mathbb{B}_0$, θ_0 by defining

$$\mathbb{B}_0^\omega(x) \coloneqq \begin{cases} \mathbb{B}_0(x) & \text{if } \Lambda(\mathbb{B}_0(x)) > \omega \text{ and } |\mathbb{B}_0(x)| < \sqrt{d}\,\omega^{-1}, \\ \mathbb{I} & \text{elsewhere;} \end{cases} \tag{3.55}$$

$$\theta_0^\omega(x) \coloneqq \begin{cases} \theta_0(x) & \text{if } \omega < \theta_0(x) < \omega^{-1}, \\ 1 & \text{elsewhere.} \end{cases} \tag{3.56}$$

With such a definition, these functions clearly satisfy

$$\Lambda(\mathbb{B}_0^\omega) > \omega, \qquad \theta_0^\omega > \omega \tag{3.57}$$

and

$$|\mathbb{B}_0^\omega| < \sqrt{d}\,\omega^{-1}, \qquad |\theta_0^\omega| < \omega^{-1} \tag{3.58}$$

in Ω. Moreover, it is evident that

$$|\mathbb{B}_0^\omega| \le \sqrt{d} + |\mathbb{B}_0|, \qquad \theta_0^\omega \le 1 + \theta_0, \tag{3.59}$$

and, since $\ln 1 = 0$, also that

$$|\ln \det \mathbb{B}_0^\omega| \le |\ln \det \mathbb{B}_0|, \qquad |\ln \theta_0^\omega| \le |\ln \theta_0| \tag{3.60}$$

a.e. in Ω. Let us further remark that, since $\mathbb{B}_0 \in L^\sigma(\Omega; \mathbb{R}^{d\times d}_{>0})$, the Lebesgue measure of the sets $\{\Lambda(\mathbb{B}_0) \le \omega\}$ and $\{|\mathbb{B}_0| \ge \omega^{-1}\}$ tends to zero as $\omega \to 0_+$, and thus

$$\|\mathbb{B}_0^\omega - \mathbb{B}_0\|_\sigma^\sigma = \int_{\Lambda(\mathbb{B}_0)\le\omega} |\mathbb{I} - \mathbb{B}_0|^\sigma + \int_{|\mathbb{B}_0|\ge\omega^{-1}} |\mathbb{I} - \mathbb{B}_0|^\sigma \to 0. \tag{3.61}$$

Analogously, relying on $\theta_0 \in L^1(\Omega;\mathbb{R}_{>0})$, we also obtain

$$\|\theta_0^\omega - \theta_0\|_1 \to 0, \quad \omega \to 0_+. \tag{3.62}$$

Next, we discretize the ω-truncated system in space by the Galerkin method.[4] Let $\{\boldsymbol{w}_i\}_{i=1}^\infty$, $\{\mathbb{W}_j\}_{j=1}^\infty$ and $\{w_k\}_{k=1}^\infty$ be bases of $W^{N,2}(\Omega;\mathbb{R}^d)\cap W^{1,2}_{\boldsymbol{n},\mathrm{div}}$, $W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $W^{N,2}(\Omega;\mathbb{R})$, respectively, with the following properties:

- The bases are L^2-orthonormal and $W^{N,2}$-orthogonal.
- The number $N \in \mathbb{N}$ is chosen so large that the elements of the bases are Lipschitz (due to embeddings of Sobolev spaces).
- $w_1 = |\Omega|^{-\frac{1}{2}}$.
- For any $\ell, n \in \mathbb{N}$, there exist L^2-orthogonal projections

$$\begin{aligned} P_\ell &: L^2(\Omega;\mathbb{R}^d) \to \operatorname{span}\{\boldsymbol{w}_i\}_{i=1}^\ell, \\ Q_n &: L^2(\Omega;\mathbb{R}^{d\times d}) \to \operatorname{span}\{\mathbb{W}_j\}_{j=1}^n, \\ R_n &: L^2(\Omega;\mathbb{R}) \to \operatorname{span}\{w_k\}_{k=1}^n \end{aligned}$$

- P_ℓ, Q_n, R_n are L^2- and $W^{N,2}$-bounded, uniformly w.r.t. ℓ, n.

Existence of these bases and corresponding projections follows from standard results (see Appendix 4 in [30]) using the eigenvectors of the generalized Laplace or Stokes operators.

We fix $\ell, n \in \mathbb{N}$ and consider the problem of finding the functions $\alpha^i_{\ell n}$, $\beta^j_{\ell n}$, $\gamma^k_{\ell n}$ of time, where $i = 1,\dots,\ell$ and $j,k = 1,\dots,n$, such that the functions $\boldsymbol{v}_{\ell n}$, $\mathbb{B}_{\ell n}$, $\theta_{\ell n}$ defined as

$$\begin{aligned} \boldsymbol{v}_{\ell n}(t,x) &= \sum_{i=1}^{\ell} \alpha^i_{\ell n}(t)\boldsymbol{w}_i(x), \\ \mathbb{B}_{\ell n}(t,x) = \sum_{j=1}^{n} \beta^j_{\ell n}(t)\mathbb{W}_j(x) &\quad \text{and} \quad \theta_{\ell n} = \sum_{k=1}^{n} \gamma^k_{\ell n}(t) w_k(x) \end{aligned} \tag{3.63}$$

satisfy the following equations[5] a.e. in $(0,T_0)$, $T_0 > 0$:

$$\begin{aligned} (\partial_t \boldsymbol{v}_{\ell n}, \boldsymbol{w}_i) - (\boldsymbol{v}_{\ell n} \otimes \boldsymbol{v}_{\ell n}, \nabla \boldsymbol{w}_i) + (2\nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}, \nabla \boldsymbol{w}_i) + \alpha(\boldsymbol{v}_{\ell \boldsymbol{n}}, \boldsymbol{\varphi})_{\partial\Omega} \\ = -(2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}, \nabla \boldsymbol{w}_i) + (\boldsymbol{f}, \boldsymbol{w}_i), \end{aligned} \tag{3.64}$$

$$\begin{aligned} (\partial_t \mathbb{B}_{\ell n}, \mathbb{W}_j) + (\boldsymbol{v}_{\ell n} \cdot \nabla \mathbb{B}_{\ell n}, \mathbb{W}_j) + (P(\theta_{\ell n}, \mathbb{B}_{\ell n}), \mathbb{W}_j) + (\lambda(\theta_{\ell n})\nabla \mathbb{B}_{\ell n}, \nabla \mathbb{W}_j) \\ = (2g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, \mathbb{W}_j), \end{aligned} \tag{3.65}$$

$$\begin{aligned} (c_v\partial_t \theta_{\ell n}, w_k) + (c_v \boldsymbol{v}_{\ell n} \cdot \nabla \theta_{\ell n}, w_k) + ((\kappa(\theta_{\ell n}) + \omega|\nabla\theta_{\ell n}|^r)\nabla\theta_{\ell n}, \nabla w_k) \\ = (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n} \cdot \mathbb{D}\boldsymbol{v}_{\ell n}, w_k), \end{aligned} \tag{3.66}$$

for all $1 \le i \le \ell$, $1 \le j,k \le n$ and with the initial conditions

$$\boldsymbol{v}_{\ell n}(0) = P_\ell \boldsymbol{v}_0, \quad \mathbb{B}_{\ell n}(0) = Q_n \mathbb{B}_0^\omega, \quad \theta_{\ell n}(0) = R_n \theta_0^\omega \quad \text{in } \Omega. \tag{3.67}$$

[4]The great advantage of our approach is that with the cut-off g_ω, we need not care about positive definiteness of the basis functions for $\mathbb{B}$.

[5]The term $\omega(|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}, \nabla w_k)$ vanishes in the limit $\omega \to 0_+$ and allows us to avoid the use of a weighted Sobolev space, where the density of smooth functions is not available in general.

By the L^2-orthonormality of the bases, we have

$$(\partial_t \boldsymbol{v}_{\ell n}, \boldsymbol{w}_i) = \sum_{m=1}^{\ell} \partial_t \alpha_{\ell n}^m (\boldsymbol{w}_m, \boldsymbol{w}_i) = (\alpha_{\ell n}^i)'$$

and similarly

$$(\partial_t \mathbb{B}_{\ell n}, \mathbb{W}_j) = (\beta_{\ell n}^j)', \quad (\partial_t \theta_{\ell n}, w_k) = (\gamma_{\ell n}^k)',$$

hence (3.64)–(3.66) is, in fact, a system of $\ell + 2n$ ordinary differential equations

$$\left.\begin{aligned} (\alpha_{\ell n}^i)' &= F_1(t, \alpha_{\ell n}^1, \dots, \alpha_{\ell n}^\ell), & i &= 1, \dots, \ell, \\ (\beta_{\ell n}^j)' &= F_2(\beta_{\ell n}^1, \dots, \beta_{\ell n}^n), & j &= 1, \dots, n, \\ (\gamma_{\ell n}^k)' &= F_3(\gamma_{\ell n}^1, \dots, \gamma_{\ell n}^n), & k &= 1, \dots, n. \end{aligned}\right\} \tag{3.68}$$

It is easy to see, using (3.1), that F_1, F_2 and F_3 are continuous with respect to the variables $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$ and $\gamma_{\ell n}^k$, respectively. Moreover, the explicit dependence of F_1 on time is controlled by

$$|(\boldsymbol{f}, \boldsymbol{w}_i)| \le \|\boldsymbol{f}\|_2 \|\boldsymbol{w}_i\|_2 \in L^2(0, T; \mathbb{R}).$$

Thus, we can apply the Caratheodory existence theorem (see [12, Chapter 2, Theorem 1] or [46, Chapter 30]) and hereby obtain absolutely continuous functions $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$, $\gamma_{\ell n}^k$, $1 \le i \le \ell$, $1 \le j, k \le n$, solving (3.68) on $(0, T_0)$, where $T_0 < T$ is the time of the first blow-up. In view of the a priori estimates derived below (see e.g. (3.71)), we are able to prove that

$$\sup_{t \in (0, T_0)} \left(\sum_{i=1}^{\ell} (\alpha_{\ell n}^i(t))^2 + \sum_{j=1}^{n} (\beta_{\ell n}^j(t))^2 + \sum_{k=1}^{n} (\gamma_{\ell n}^k(t))^2 \right) < \infty,$$

hence, there can be no blow-up and the functions $\boldsymbol{v}_{kl}, \mathbb{B}_{kl}, \theta_{kl}$ are defined on an arbitrary time interval, in particular on $[0, T]$.

n-uniform estimates. By multiplying the i-th equation in (3.64) by $\alpha_{\ell n}^i$, summing the result over all $i = 1, \dots, \ell$, integrating by parts and using $(1.13)_1$, (1.7) (so that the convective term vanishes), we obtain

$$\begin{aligned} \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{v}_{\ell n}\|_2^2 + \|\sqrt{2\nu(\theta_{\ell n})} \mathbb{D} \boldsymbol{v}_{\ell n}\|_2^2 + \alpha \|\boldsymbol{v}_{\ell n}\|_{2;\partial\Omega}^2 \\ = -(2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n}) \theta_{\ell n} \mathbb{B}_{\ell n}, \mathbb{D} \boldsymbol{v}_{\ell n}) + (\boldsymbol{f}, \boldsymbol{v}_{\ell n}) \end{aligned} \tag{3.69}$$

a.e. in $(0, T)$. Then we use (3.2), (3.54), (3.67), Korn's and Young's inequality, and deduce

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{v}_{\ell n}\|_2^2 + \|\nabla \boldsymbol{v}_{\ell n}\|_2^2 + \alpha \|\boldsymbol{v}_{\ell n}\|_{2;\partial\Omega}^2 &\le C(\omega) \int_\Omega |\mathbb{D} \boldsymbol{v}_{\ell n}| + C \|\boldsymbol{f}\|_2 \|\nabla \boldsymbol{v}_{\ell n}\|_2 \\ &\le C(\omega) + C \|\boldsymbol{f}\|_2^2 + \frac{1}{2} \|\nabla \boldsymbol{v}_{\ell n}\|_2^2 \end{aligned}$$

a.e. in $(0, T)$. Integration with respect to time and the use of (3) and (3.29) directly leads to

$$\sup_{t \in (0,T)} \|\boldsymbol{v}_{\ell n}(t)\|_2^2 + \int_0^T \|\nabla \boldsymbol{v}_{\ell n}\|_2^2 + \alpha \int_0^T \|\boldsymbol{v}_{\ell n}\|_{2;\partial\Omega}^2 \le C(\omega). \tag{3.70}$$

(the dependence of the constant C on the data is omitted as $\boldsymbol{f}$, $\boldsymbol{v}_0$, θ_0, or $\mathbb{B}_0$ are fixed functions in our setting). Recalling the construction of $\boldsymbol{v}_{\ell n}$ in (3.63) and L^2-orthonormality of the basis vectors $\{\boldsymbol{w}_i\}_{i=1}^{\ell}$, we note that

$$\|\boldsymbol{v}_{\ell n}(t)\|_2^2 = \sum_{i=1}^{\ell}(\alpha_{\ell n}^i(t))^2.$$

Hence, the estimate (3.70) yields

$$\sup_{t\in(0,T)}\sum_{i=1}^{\ell}(\alpha_{\ell n}^i(t))^2 \le C(\omega), \tag{3.71}$$

which, together with $\boldsymbol{w}_i \in W^{1,\infty}(\Omega;\mathbb{R}^d)$, $i=1,\dots,\ell$, implies

$$\|\boldsymbol{v}_{\ell n}\|_{L^\infty W^{1,\infty}} \le C(\omega,\ell). \tag{3.72}$$

Using (3.71), (3.54) and (3.2) in (3.64), we see that

$$\begin{aligned}\|(\alpha_{\ell n}^i)'\|_{2;(0,T)} &= \|(\partial_t \boldsymbol{v}_{\ell n}, \boldsymbol{w}_i)\|_{2;(0,T)}\\ &= \|(\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n} - 2\nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n} - 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}, \nabla\boldsymbol{w}_i) \qquad (3.73)\\ &\qquad - (\boldsymbol{v}_{\ell n},\boldsymbol{w}_i)_{\partial\Omega} + (\boldsymbol{f},\boldsymbol{w}_i)\|_{2;(0,T)}\\ &\le C(\ell)\|\sum_{i=1}^{\ell}\big((\alpha_{\ell n}^i)^2 + |\alpha_{\ell n}^i| + 1\big)\|_{2;(0,T)} + C(\ell)\|\boldsymbol{f}\|_{L^2L^2} \le C(\omega,\ell). \qquad (3.74)\end{aligned}$$

Thus, we get

$$\|\partial_t\boldsymbol{v}_{\ell n}\|_{L^2W^{1,\infty}} = \|\sum_{i=1}^{\ell}(\alpha_{\ell n}^i)'\boldsymbol{w}_i\|_{L^2W^{1,\infty}} \le C(\omega,\ell) \tag{3.75}$$

and, using the fundamental theorem of calculus and Hölder's inequality, also that

$$|\alpha_{\ell n}^i(t) - \alpha_{\ell n}^i(s)| \le \int_s^t |(\alpha_{\ell n}^i)'| \le C(\omega,\ell)|t-s|^{\frac12} \quad \text{for every } t,s\in[0,T] \tag{3.76}$$

and any $i=1,\dots,\ell$.

Next, we multiply the j-th equation in (3.65) by $\beta_{\ell n}^j$ and sum the result over $j=1,\dots,n$. Note that the convective term vanishes after integration by parts and use of $(1.13)_1$ and (1.7). Also the term including $\mathbb{W}\boldsymbol{v}_{\ell n}$ vanishes due to symmetry of $\mathbb{B}_{\ell n}^2$. Thus, we obtain

$$\begin{aligned}\frac12\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{B}_{\ell n}\|_2^2 + (P(\theta_{\ell n},\mathbb{B}_{\ell n}),\mathbb{B}_{\ell n}) + \|\sqrt{\lambda(\theta_{\ell n})}\nabla\mathbb{B}_{\ell n}\|_2^2&\\ = (2ag_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}\mathbb{B}_{\ell n},\mathbb{B}_{\ell n})&\end{aligned} \tag{3.77}$$

a.e. in $(0,T)$. Then using (3.67), (3.7), (3.4) and (3.54) we obtain, after integration over $(0,t)$, $t\in(0,T)$, that

$$\|\mathbb{B}_{\ell n}(t)\|_2^2 + \int_0^t \|\mathbb{B}_{\ell n}\|_{2+q}^{2+q} + \int_0^t \|\nabla\mathbb{B}_{\ell n}\|_2^2 \le \|Q_n\mathbb{B}_0^\omega\|_2^2 + C(\omega,\ell). \tag{3.78}$$

From this, using properties of Q_n and (3.58), we easily read that

$$\|\mathbb{B}_{\ell n}\|_{L^\infty L^2} + \|\mathbb{B}_{\ell n}\|_{L^{2+q}L^{2+q}} + \|\nabla\mathbb{B}_{\ell n}\|_{L^2L^2} \le C(\omega,\ell). \tag{3.79}$$

To estimate the time derivative of $\mathbb{B}_{\ell n}$, we take $\mathbb{A} \in L^{q+2}(0,T;W^{N,2}(\Omega))$ with $\|\mathbb{A}\|_{L^{q+2}W^{N,2}} \leq 1$ and use (3.65), Hölder's inequality, (3.79), (3.72), (3.3), (3.6), (3.54), properties of Q_n and $(\min\{2,\frac{q+2}{q+1}\})' = q+2$ to get

$$
\begin{aligned}
&\int_0^T \langle \partial_t \mathbb{B}_{\ell n}, \mathbb{A}\rangle = \int_0^T (\partial_t \mathbb{B}_{\ell n}, Q_n\mathbb{A}) \\
&\quad = -\int_0^T (\boldsymbol{v}_{\ell n}\cdot\nabla\mathbb{B}_{\ell n}, Q_n\mathbb{A}) - \int_0^T (P(\theta_{\ell n}, \mathbb{B}_{\ell n}, Q_n\mathbb{A}) \\
&\qquad - \int_0^T (\lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n}, \nabla Q_n\mathbb{A}) + \int_0^T (2g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, Q_n\mathbb{A}) \\
&\quad \leq C(\omega,\ell)\int_0^T\int_\Omega \big(|\nabla\mathbb{B}_{\ell n}||Q_n\mathbb{A}| + |\mathbb{B}_{\ell n}|^{q+1}|Q_n\mathbb{A}| + |\nabla\mathbb{B}_{\ell n}||\nabla Q_n\mathbb{A}| + |Q_n\mathbb{A}|\big) \\
&\quad \leq C(\omega,\ell)\int_0^T (\|\nabla\mathbb{B}_{\ell n}\|_1 + \|\mathbb{B}_{\ell n}\|_{q+1}^{q+1} + 1)\|Q_n\mathbb{A}\|_{1,\infty} \\
&\quad \leq C(\omega,\ell)\int_0^T (\|\nabla\mathbb{B}_{\ell n}\|_2 + \|\mathbb{B}_{\ell n}\|_{q+2}^{q+1} + 1)\|Q_n\mathbb{A}\|_{N,2} \\
&\quad \leq C(\omega,\ell)\|\mathbb{A}\|_{L^{q+2}W^{N,2}} \leq C(\omega,\ell),
\end{aligned}
$$

hence

$$
\|\partial_t\mathbb{B}_{\ell n}\|_{L^{\frac{q+2}{q+1}}W^{-N,2}} \leq C(\omega,\ell). \tag{3.80}
$$

Next, we multiply the k-th equation in (3.66) by $\gamma_{\ell n}^k$, sum the result over $k = 1,\dots,n$, use $(1.13)_1$, (1.7) and integration by parts in the convective term to get

$$
\begin{aligned}
\frac{c_v}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta_{\ell n}\|_2^2 + \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_2^2 + \omega\|\nabla\theta_{\ell n}\|_{r+2}^{r+2} \\
= (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n}, \theta_{\ell n})
\end{aligned} \tag{3.81}
$$

a.e. in $(0,T)$. Therefore, integrating this inequality over $(0,t)$, $t\in(0,T)$, using (3.2), (3.3), (3.54), (3.72), (3.79), Young's inequality, properties of R_n and (3.58), we deduce

$$
\|\theta_{\ell n}(t)\|_2^2 + \int_0^t \|\nabla\theta_{\ell n}^{\frac{r}{2}+1}\|_2^2 + \int_0^t \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_2^2 + \int_0^t \|\nabla\theta_{\ell n}\|_{r+2}^{r+2} \leq C(\omega,\ell). \tag{3.82}
$$

This, with the help of the interpolation inequality[6]

$$
\|\theta_{\ell n}\|_{L^{r+2+\frac{4}{d}}L^{r+2+\frac{4}{d}}} \leq \|\theta_{\ell n}\|_{L^\infty L^2}^{\frac{4}{(r+2)d+4}} \|\theta_{\ell n}^{\frac{r}{2}+1}\|_{L^2L^{\frac{2d}{d-2}}}^{\frac{2d}{(r+2)d+4}},
$$

Sobolev's inequality and also Poincaré's inequality yields

$$
\begin{aligned}
\|\theta_{\ell n}\|_{L^\infty L^2} + \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_{L^2L^2} + \|\theta_{\ell n}\|_{L^{r+2+\frac{4}{d}}L^{r+2+\frac{4}{d}}} \\
+ \|\nabla\theta_{\ell n}\|_{L^{r+2}L^{r+2}} \leq C(\omega,\ell).
\end{aligned} \tag{3.83}
$$

Furthermore, taking $\tau \in L^{r+2}(0,T;W^{N,2}(\Omega))$ with $\|\tau\|_{L^{r+2}W^{N,2}} \leq 1$ and using (3.66), Young's inequality, Hölder's inequality, (3.2), (3.3) (3.72), (3.83), (3.54) and

[6] A better estimate could be derived using $\nabla\theta_\ell \in L^{r+2}L^{r+2}$ instead. However, at this moment we do not need it, and later we shall need ω-uniform estimates only.

properties of R_n, we obtain

$$
\begin{aligned}
\int_0^T \langle \partial_t \theta_{\ell n}, \tau \rangle &= \int_0^T (\partial_t \theta_{\ell n}, R_n \tau) \\
&= -\int_0^T (c_v \boldsymbol{v}_{\ell n} \cdot \nabla \theta_{\ell n}, R_n \tau) - \int_0^T (\kappa(\theta_{\ell n}) \nabla \theta_{\ell n} + \omega |\nabla \theta_{\ell n}|^r \nabla \theta_{\ell n}, \nabla R_n \tau) \\
&\quad + \int_0^T (2\nu(\theta_{\ell n}) |\mathbb{D} \boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n}) \theta_{\ell n} \mathbb{B}_{\ell n} \cdot \mathbb{D} \boldsymbol{v}_{\ell n}, R_n \tau) \\
&\le C(\omega, \ell) \int_0^T \int_\Omega \Big(|\nabla \theta_{\ell n}| |R_n \tau| + |\theta_{\ell n}|^{\frac{r}{2}} \left| \sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n} \right| |\nabla R_n \tau| \\
&\qquad\qquad + |\nabla \theta_{\ell n}|^{r+1} |\nabla R_n \tau| + |R_n \tau| \Big) \\
&\le C(\omega, \ell) \int_0^T \int_\Omega \Big(|\nabla \theta_{\ell n}| + |\theta_{\ell n}|^{r+1} + \left| \sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n} \right|^{\frac{2r+2}{r+2}} \\
&\qquad\qquad + |\nabla \theta_{\ell n}|^{r+1} + 1 \Big) \|R_n \tau\|_{1,\infty} \\
&\le C(\omega, \ell) \int_0^T \Big(\|\nabla \theta_{\ell n}\|_{\frac{r+2}{r+1}} + \|\theta_{\ell n}\|_{r+2}^{r+1} + \|\sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n}\|_2^{\frac{2r+2}{r+2}} \\
&\qquad\qquad + \|\nabla \theta_{\ell n}\|_{r+1}^{r+1} + 1 \Big) \|R_n \tau\|_{N,2} \\
&\le C(\omega, \ell) \|\tau\|_{L^{r+2} W^{N,2}} \le C(\omega, \ell),
\end{aligned}
$$

hence

$$
\|\partial_t \theta_{\ell n}\|_{L^{\frac{r+2}{r+1}} W^{-N,2}} \le C(\omega, \ell). \tag{3.84}
$$

The limit $n \to \infty$. For every $i = 1, \ldots, \ell$, the sequence $\{\alpha_{\ell n}^i\}_{n=1}^\infty \subset \mathcal{C}([0,T]; \mathbb{R})$ is bounded due to (3.71) and uniformly equicontinuous by (3.76). Hence, using the Arzelà-Ascoli theorem, for every $i = 1, \ldots, \ell$, we obtain $\alpha_\ell^i \in \mathcal{C}([0,T]; \mathbb{R})$ and a subsequence (not relabelled) such that

$$
\alpha_{\ell n}^i \to \alpha_\ell^i \quad \text{strongly in } \mathcal{C}([0,T]; \mathbb{R}) \tag{3.85}
$$

as $n \to \infty$. Then, we define

$$
\boldsymbol{v}_\ell \coloneqq \sum_{i=1}^{\ell} \alpha_\ell^i \boldsymbol{w}_i \in \mathcal{C}([0,T]; W^{1,\infty}(\Omega; \mathbb{R}^d) \cap W^{1,2}_{\boldsymbol{n},\mathrm{div}})
$$

and note that

$$
\boldsymbol{v}_{\ell n} \to \boldsymbol{v}_\ell \quad \text{strongly in } \mathcal{C}([0,T]; W^{1,\infty}(\Omega; \mathbb{R}^d)). \tag{3.86}
$$

According to estimates (3.75), (3.79), (3.80), (3.83), (3.84) and using reflexivity of the underlying spaces and the Aubin-Lions lemma, there exist subsequences

$\{\boldsymbol{v}_{\ell n}\}_{n=1}^{\infty}$, $\{\mathbb{B}_{\ell n}\}_{n=1}^{\infty}$, $\{\theta_{\ell n}\}_{n=1}^{\infty}$ and their limits $\boldsymbol{v}_{\ell}$, $\mathbb{B}_{\ell}$, θ_{ℓ}, such that

$$\partial_t \boldsymbol{v}_{\ell n} \overset{*}{\rightharpoonup} \partial_t \boldsymbol{v}_{\ell} \qquad \text{weakly* in } L^2(0,T;W^{1,\infty}(\Omega;\mathbb{R}^d)), \tag{3.87}$$

$$\boldsymbol{v}_{\ell n} \rightharpoonup \boldsymbol{v}_{\ell} \qquad \text{weakly in } L^2(0,T;L^2(\partial\Omega;\mathbb{R}^d)), \tag{3.88}$$

$$\mathbb{B}_{\ell n} \rightharpoonup \mathbb{B}_{\ell} \qquad \text{weakly in } L^2(0,T,W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}})), \tag{3.89}$$

$$\mathbb{B}_{\ell n} \to \mathbb{B}_{\ell} \qquad \text{strongly in } L^{2+q)}(Q;\mathbb{R}^{d\times d}_{\text{sym}}) \text{ and a.e. in } Q, \tag{3.90}$$

$$\partial_t\mathbb{B}_{\ell n} \rightharpoonup \partial_t\mathbb{B}_{\ell} \qquad \text{weakly in } L^{\frac{q+2}{q+1}}(0,T;W^{-N,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}})), \tag{3.91}$$

$$\theta_{\ell n} \rightharpoonup \theta_{\ell} \qquad \text{weakly in } L^{r+2}(0,T,W^{1,r+2}(\Omega;\mathbb{R})), \tag{3.92}$$

$$\theta_{\ell n} \to \theta_{\ell} \qquad \text{strongly in } L^{r+2+\frac{4}{d})}(Q;\mathbb{R}^{d\times d}_{\text{sym}}) \text{ and a.e. in } Q, \tag{3.93}$$

$$\partial_t\theta_{\ell n} \rightharpoonup \partial_t\theta_{\ell} \qquad \text{weakly in } L^{\frac{r+2}{r+1}}(0,T;W^{-N,2}(\Omega;\mathbb{R})). \tag{3.94}$$

Now we explain how to take the limit in the non-linear terms appearing in (3.64), (3.65) and (3.66). To handle most of the terms, namely

$$\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n},\quad \nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n},\quad g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n},\quad P(\theta_{\ell n},\mathbb{B}_{\ell n}),\quad \lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n},$$
$$g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n}+\mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n},\quad \boldsymbol{v}_{\ell n}\cdot\nabla\theta_{\ell n},\quad g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n},$$

we use the standard argument: These terms can be seen as a product of a weakly converging sequence with a strongly converging sequence, obtained via Vitali's theorem, (3.1) and pointwise convergence of $\boldsymbol{v}_{\ell n}$, $\mathbb{B}_{\ell n}$ and $\theta_{\ell n}$ (due to the Aubin-Lions lemma). This argument is sufficient to take the limit $n\to\infty$ in the equations (3.64) and (3.65). In (3.65), we first multiply the equation by a function $\varphi\in\mathcal{C}^1([0,T];\mathbb{R})$, integrate over $(0,T)$, then take the limit and finally use the density of functions of the form $\varphi\mathbb{A}$, $\mathbb{A}\in\operatorname{span}\{\mathbb{W}_j\}_{j=1}^{\infty}$, in the space $L^{(q+2)'}(0,T;W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}}))$. This way, we obtain

$$\begin{aligned}(\partial_t\boldsymbol{v}_\ell,\boldsymbol{w}_i)-(\boldsymbol{v}_\ell\otimes\boldsymbol{v}_\ell,\nabla\boldsymbol{w}_i)+(2\nu(\theta_\ell)\mathbb{D}\boldsymbol{v}_\ell,\nabla\boldsymbol{w}_i)+\alpha(\boldsymbol{v}_\ell,\boldsymbol{w}_i)_{\partial\Omega}\\ =-(2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell,\nabla\boldsymbol{w}_i)+(\boldsymbol{f},\boldsymbol{w}_i)\quad\text{for every } i=1,\dots,\ell\end{aligned}\tag{3.95}$$

a.e. in $(0,T)$ and

$$\begin{aligned}&\int_0^T\langle\partial_t\mathbb{B}_\ell,\mathbb{A}\rangle+\int_0^T(\boldsymbol{v}_\ell\cdot\nabla\mathbb{B}_\ell,\mathbb{A})+\int_0^T(P(\theta_\ell,\mathbb{B}_\ell),\mathbb{A})+\int_0^T(\lambda(\theta_\ell)\nabla\mathbb{B}_\ell,\nabla\mathbb{A})\\ &\quad=\int_0^T(2g_\omega(\mathbb{B}_\ell,\theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell+\mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell,\mathbb{A})\\ &\quad\text{for all } \mathbb{A}\in L^{q+2}(0,T;W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}})).\end{aligned}\tag{3.96}$$

However, the space of test functions in (3.96) can be enlarged using a standard density argument. Indeed, using Hölder's inequality, it is easy to see that every term of (3.96) (taking aside the time derivative) is well defined provided that

$$\mathbb{A}\in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}}))\cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\text{sym}})$$

and thus, we can read from (3.96) that

$$\partial_t\mathbb{B}_\ell\in\left(L^2(0,T;W^{1,2}(\Omega,\mathbb{R}^{d\times d}_{\text{sym}}))\cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\text{sym}})\right)^*.$$

Since we also have that

$$\mathbb{B}_\ell\in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}}))\cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\text{sym}}),\tag{3.97}$$

it follows from Lemma 1 below that

$$\mathbb{B}_\ell \in \mathcal{C}([0,T]; L^2(\Omega; \mathbb{R}^{d\times d}_{\rm sym})). \tag{3.98}$$

Now let us identify $\mathbb{B}_\ell(0)$. Clearly, we can use $\mathbb{A}(t,x) = \psi(t)\mathbb{P}(x)$ in (3.96), where $\psi \in \mathcal{C}^1([0,T];\mathbb{R})$, $\psi(0) = 1$, $\psi(T) = 0$, and $\mathbb{P} \in W^{N,2}(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$, to get, after integration by parts, that

$$\begin{aligned}(\mathbb{B}_\ell(0), \mathbb{P}) = -\int_0^T \big((\mathbb{B}_\ell, \mathbb{P})\partial_t\psi + (\boldsymbol{v}_\ell \cdot \nabla\mathbb{B}_\ell, \mathbb{P})\psi + (P(\theta_\ell, \mathbb{B}_\ell), \mathbb{P})\psi \\ - (\lambda(\theta_\ell)\nabla\mathbb{B}_\ell, \nabla\mathbb{P})\psi - (2g_\omega(\mathbb{B}_\ell, \theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell, \mathbb{P})\psi\big).\end{aligned} \tag{3.99}$$

On the other hand, if we multiply (3.65) by ψ, integrate over $(0,T)$ and by parts in the time derivative using (3.67), we obtain

$$\begin{aligned}(Q_n\mathbb{B}_0^\omega, \mathbb{W}_j) = -\int_0^T \big((\mathbb{B}_{\ell n}, \mathbb{W}_j)\partial_t\psi + (\boldsymbol{v}_{\ell n}\cdot\nabla\mathbb{B}_{\ell n}, \mathbb{W}_j)\psi + (P(\theta_\ell, \mathbb{B}_\ell), \mathbb{W}_j)\psi \\ - (\lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n}, \nabla\mathbb{W}_j)\psi - (2g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, \mathbb{W}_j)\psi\big).\end{aligned} \tag{3.100}$$

for every $j = 1, \ldots, n$. Then, we use completeness of $\{\mathbb{W}_j\}_{j=1}^\infty$ in $L^2(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$ and the same arguments as before to take the limit $n \to \infty$ in (3.100). This way, using also density of $\operatorname{span}\{\mathbb{W}_j\}_{j=1}^\infty$ in $W^{N,2}(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$, we get, for all $\mathbb{P} \in W^{N,2}(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$, that

$$\begin{aligned}(\mathbb{B}_0^\omega, \mathbb{P}) = -\int_0^T \big((\mathbb{B}_\ell, \mathbb{P})\partial_t\psi + (\boldsymbol{v}_\ell \cdot \nabla\mathbb{B}_\ell, \mathbb{P})\psi + (P(\theta_\ell, \mathbb{B}_\ell), \mathbb{P})\psi \\ - (\lambda(\theta_\ell)\nabla\mathbb{B}_\ell, \nabla\mathbb{P})\psi - (2g_\omega(\mathbb{B}_\ell, \theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell, \mathbb{P})\psi\big).\end{aligned}$$

If we compare this with (3.99) and use density of $W^{N,2}(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$ in $L^2(\Omega; \mathbb{R}^{d\times d}_{\rm sym})$, we deduce

$$\mathbb{B}_\ell(0) = \mathbb{B}_0^\omega \quad \text{a.e. in } \Omega. \tag{3.101}$$

We can use an analogous procedure to identify $\boldsymbol{v}_\ell(0)$. Indeed, here the situation is even simpler since (3.86) directly implies $\boldsymbol{v}_\ell \in \mathcal{C}([0,T]; W^{1,\infty}(\Omega; \mathbb{R}^d))$ and we obtain

$$\boldsymbol{v}_\ell(0) = P_\ell \boldsymbol{v}_0. \tag{3.102}$$

Our aim is now to take the limit in equation (3.66), where we need to justify the limit in the terms $\kappa(\theta_{\ell n})\nabla\theta_{\ell n}$, $|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}$ and $2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2$. For the first one, we use (3.3), (3.93) and Vitali's theorem to get

$$\sqrt{\kappa(\theta_{\ell n})} \to \sqrt{\kappa(\theta_\ell)} \quad \text{strongly in } L^{2+\frac{4}{r}}(Q;\mathbb{R}) \tag{3.103}$$

and then we combine this with (3.92), to obtain

$$\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \sqrt{\kappa(\theta_\ell)}\nabla\theta_\ell \quad \text{weakly in } L^1(Q;\mathbb{R}^d). \tag{3.104}$$

However, by (3.83) we know that (3.104) is valid also in $L^2(Q;\mathbb{R}^d)$ up to a subsequence, and hence, using again (3.103), we obtain

$$\kappa(\theta_{\ell n})\nabla\theta_{\ell n} = \sqrt{\kappa(\theta_{\ell n})}\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \sqrt{\kappa(\theta_\ell)}\sqrt{\kappa(\theta_\ell)}\nabla\theta_\ell = \kappa(\theta_\ell)\nabla\theta_\ell \tag{3.105}$$

weakly in $L^{\frac{r+2}{r+1}}(Q;\mathbb{R}^d)$.

Next, to take the limit of the term $2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2$, we first remark, using (3.1), (3.2), (3.93) and Vitali's theorem that

$$\nu(\theta_{\ell n}) \to \nu(\theta_\ell) \quad \text{strongly in } L^{\infty)}(Q;\mathbb{R}).$$

This and

$$\mathbb{D}\boldsymbol{v}_{\ell n} \to \mathbb{D}\boldsymbol{v}_{\ell} \quad \text{strongly in } \mathcal{C}([0,T]; L^{\infty}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}))$$

(cf. (3.86)) clearly proves that

$$2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 \to 2\nu(\theta_{\ell})|\mathbb{D}\boldsymbol{v}_{\ell}|^2 \quad \text{strongly in } L^{\infty)}(Q;\mathbb{R}). \tag{3.106}$$

Finally, due to (3.83), there exists $K \in L^{(r+2)'}(Q;\mathbb{R}^d)$ such that

$$|\nabla\theta_{\ell n}|^r \nabla\theta_{\ell n} \rightharpoonup K \quad \text{weakly in } L^{(r+2)'}(Q;\mathbb{R}^d). \tag{3.107}$$

Then, using also (3.105), (3.106) and previous convergence results, we can take the limit in (3.66) and obtain, for all $\tau \in L^{r+2}(0,T;W^{N,2}(\Omega;\mathbb{R}))$, that

$$\begin{aligned} \int_0^T \langle c_v \partial_t \theta_\ell, \tau\rangle + \int_0^T (c_v \boldsymbol{v}_\ell \cdot \nabla\theta_\ell, \tau) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla\tau) + \omega \int_0^T (K, \nabla\tau) \\ = \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell \mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \tau). \end{aligned} \tag{3.108}$$

Recalling (3.94), (3.105) and (3.107), we easily conclude, using a density argument, that (3.108) is valid for all $\tau \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$ and that the time derivative extends to the functional $\partial_t\theta_\ell \in L^{(r+2)'}(0,T;W^{-1,(r+2)'}(\Omega;\mathbb{R}))$. Thus, using Lemma 1 below, we also see that

$$\theta_\ell \in \mathcal{C}([0,T];L^2(\Omega;\mathbb{R})). \tag{3.109}$$

Furthermore, choosing $\tau = \theta_\ell$ in (3.108), rewriting the time derivative term and integrating by parts in the convective term leads to

$$\begin{aligned} \omega \int_Q K \cdot \nabla\theta_\ell = -\frac{c_v}{2}\|\theta_\ell(T)\|_2^2 + \frac{c_v}{2}\|\theta_\ell(0)\|_2^2 - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \\ + \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, \theta_\ell). \end{aligned} \tag{3.110}$$

We use this information to identify K as follows. We note that weak lower semicontinuity and (3.104) (which is valid in $L^2(Q;\mathbb{R}^d)$) imply

$$\int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \le \liminf_{n\to\infty} \int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2. \tag{3.111}$$

Thus, if we integrate (3.81) over $(0,T)$ and use (3.111), (3.106), weak lower semicontinuity of $\|\cdot\|_2$ and the convergence results above to take the limes superior

$n \to \infty$ and then apply (3.110), we get

$$
\begin{aligned}
&\omega \limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \\
&\quad = -\liminf_{n\to\infty} \frac{c_v}{2}\|\theta_{\ell n}(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \liminf_{n\to\infty}\int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2 \\
&\qquad + \lim_{n\to\infty}\int_0^T (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n},\theta_{\ell n}) \\
&\quad \le -\frac{c_v}{2}\|\theta_\ell(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \\
&\qquad + \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\theta_\ell) \\
&\quad = \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \frac{c_v}{2}\|\theta_\ell(0)\|_2^2 + \omega\int_Q K\cdot\nabla\theta_\ell.
\end{aligned}
\tag{3.112}
$$

To identify the initial condition for $\theta_\ell(0)$, it is enough to show that

$$\theta_\ell(t) \rightharpoonup \theta_0^\omega \quad \text{weakly in } L^2(\Omega;\mathbb{R}) \tag{3.113}$$

as $t \to 0_+$ since then we can use (3.109) to conclude

$$\theta_\ell(0) = \theta_0^\omega \quad \text{a.e. in } \Omega \tag{3.114}$$

by the uniqueness of a (weak) limit. To prove (3.113), we return to (3.66), which we multiply by $\varphi \in W^{1,\infty}(0,T;\mathbb{R})$ fulfilling $\varphi(0) = 1$, $\varphi(T) = 0$ and integrate the result over $(0,T)$ to get

$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_{\ell n}, w_k)\partial_t\varphi = \int_0^T f_n\varphi. \tag{3.115}$$

for all $k = 1,\dots,n$, where we integrated by parts and abbreviated

$$
\begin{aligned}
f_n = &-(c_v\boldsymbol{v}_{\ell n}\cdot\nabla\theta_{\ell n}, w_k) - (\kappa(\theta_{\ell n})\nabla\theta_{\ell n} + \omega|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}, \nabla w_k) \\
&+ (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n}, w_k).
\end{aligned}
$$

It follows from the results above (cf. the derivation of (3.108)) that

$$f_n \rightharpoonup f \quad \text{weakly in } L^{(r+2)'}(0,T;\mathbb{R}),$$

where

$$
\begin{aligned}
f = &-(c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell, w_k) - (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla w_k) - \omega(K,\nabla w_k) \\
&+ (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, w_k).
\end{aligned}
$$

Thus, by taking the limit $n \to \infty$ in (3.115), we arrive at

$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_\ell, w_k)\partial_t\varphi = \int_0^T f\varphi.$$

Making now a special choice

$$\varphi_\varepsilon(s) = \begin{cases} 1 & s \le t, \\ 1 - \frac{s-t}{\varepsilon} & s \in (t, t+\varepsilon), \\ 0 & s \ge t+\varepsilon, \end{cases}$$

where $t \in (0,T)$ and $0 < \varepsilon < T - t$, leads to

$$-(c_v\theta_0^\omega, w_k) + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} (c_v\theta_\ell, w_k) = \int_0^{t+\varepsilon} f\varphi_\varepsilon.$$

Furthermore, we can take the limit $\varepsilon \to 0_+$ in this equation using (3.109) on the left hand side and absolute continuity of integral on the right hand side to get

$$-(c_v\theta_0^\omega, w_k) + (c_v\theta_\ell(t), w_k) = \int_0^t f.$$

Finally, taking the limit $t \to 0_+$ yields

$$\lim_{t\to 0_+} (\theta_\ell(t), w_k) = (\theta_0^\omega, w_k),$$

for all $k = 1, \dots, n$, from which (3.113) follows by exploiting the density of the set $\operatorname{span}\{w_k\}_{k=1}^\infty$ in $L^2(\Omega;\mathbb{R})$. Hence, the identity (3.114) is proved and (3.112) hereby simplifies to

$$\limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \le \int_Q K\cdot\nabla\theta_\ell. \tag{3.116}$$

Since the operator $\boldsymbol{u} \mapsto |\boldsymbol{u}|^r\boldsymbol{u}$ is monotone and continuous, it is standard to show, using (3.116) and the Minty method, that

$$K = |\nabla\theta_\ell|^r\nabla\theta_\ell \quad \text{a.e. in } Q.$$

Hence, we proved that

$$\begin{aligned}\int_0^T \langle c_v\partial_t\theta_\ell, \tau\rangle + \int_0^T (c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell, \tau) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell + \omega|\nabla\theta_\ell|^r\nabla\theta_\ell, \nabla\tau)\\ = \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, \tau)\end{aligned} \tag{3.117}$$

for all $\tau \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$.

Positive definiteness of $\mathbb{B}_\ell$ and positivity of θ_ℓ. Here we follow the method developed in [4], i.e., we use

$$\mathbb{A}_{\boldsymbol{x}} = \chi_{(0,t)}(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-\, \boldsymbol{x}\otimes\boldsymbol{x},$$

in (3.96), where $\boldsymbol{x}\in\mathbb{R}^d$, $t\in(0,T)$ and

$$f_+ = \max\{0,f\}, \qquad f_- = \min\{0,f\}.$$

Note that since $\boldsymbol{x}$ is a constant vector, the function $\mathbb{A}_{\boldsymbol{x}}$ belongs to the same space as $\mathbb{B}_\ell \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\text{sym}})$ and is thus a valid test function in (3.96). The key property of $\mathbb{A}_{\boldsymbol{x}}$ is that it vanishes whenever the smallest eigenvalue of $\mathbb{B}_\ell$ is greater than ω (since $\mathbb{B}_\ell\boldsymbol{y}\cdot\boldsymbol{y} \ge \omega|\boldsymbol{y}|^2$ for all $\boldsymbol{y}\in\mathbb{R}^d$ in such a case). Thus, we have

$$(\Lambda(\mathbb{B}_\ell) - \omega)_+(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- = 0,$$

which implies

$$g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{A}_x = 0 \quad \text{a.e. in } Q. \tag{3.118}$$

Let us now evaluate separately the terms arising from the choice $\mathbb{A} = \mathbb{A}_{\boldsymbol{x}}$ in (3.96). For the time derivative, we write

$$\begin{aligned}\int_0^T \langle\partial_t\mathbb{B}_\ell, \mathbb{A}_{\boldsymbol{x}}\rangle &= \int_0^t \langle\partial_t(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2), (\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-\rangle\\ &= \tfrac12\|(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t)\|_2^2 - \tfrac12\|(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(0)\|_2^2\\ &= \tfrac12\|(\mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t)\|_2^2,\end{aligned} \tag{3.119}$$

where we applied Lemma 2 below for the Lipschitz function $s \mapsto s_-$ and also (3.101) and (3.57). Furthermore, using integration by parts and $(1.13)_1$ and (1.7), we get

$$\int_0^T (\boldsymbol{v}_\ell \cdot \nabla \mathbb{B}_\ell, \mathbb{A}_{\boldsymbol{x}}) = \int_0^t (\boldsymbol{v}_\ell \cdot \nabla(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2), (\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-)$$
$$= \frac{1}{2}\int_0^t \int_{\partial\Omega} ((\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-)^2 \boldsymbol{v}_\ell \cdot \boldsymbol{n} = 0$$

and also

$$\int_0^T (\lambda(\theta_\ell)\nabla \mathbb{B}_\ell, \nabla \mathbb{A}_{\boldsymbol{x}}) = \int_0^t \|\sqrt{\lambda(\theta_\ell)}\nabla(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-\|_2^2 \geq 0.$$

Moreover, since $\omega < \omega_P$, we have $(\mathbb{B}_\ell - \omega_P \mathbb{I})\boldsymbol{x} \cdot \boldsymbol{x} < \mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2$ and thus, the assumption (3.10) yields

$$\int_0^T (P(\theta_\ell, \mathbb{B}_\ell), \mathbb{A}_{\boldsymbol{x}}) = \int_0^t \int_\Omega (\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- \, P(\theta_\ell, \mathbb{B}_\ell)\boldsymbol{x} \cdot \boldsymbol{x}$$
$$= \int_0^t \int_{\{\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} < \omega|\boldsymbol{x}|^2\}} (\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2) P(\theta_\ell, (\mathbb{B}_\ell - \omega_P \mathbb{I}) + \omega_P \mathbb{I})\boldsymbol{x} \cdot \boldsymbol{x} \geq 0.$$

In addition, the right hand side of (3.96) vanishes due to (3.118). Thus, using the above computation in (3.96), we obtain

$$\|(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t)\|_2^2 \leq 0$$

for all $t \in (0,T)$ (recall (3.98)), whence

$$\mathbb{B}_\ell(t)\boldsymbol{x} \cdot \boldsymbol{x} \geq \omega|\boldsymbol{x}|^2 \quad \text{a.e. in } \Omega, \text{ for all } t \in (0,T) \text{ and for every } \boldsymbol{x} \in \mathbb{R}^d. \tag{3.120}$$

Note that this immediately yields $\mathbb{B}_\ell \in \mathbb{R}^{d\times d}_{>0}$, $\mathbb{B}_\ell^{-1} \in \mathbb{R}^{d\times d}_{>0}$ a.e. in Q, and thus

$$|\mathbb{B}_\ell^{-1}| = |\mathbb{B}_\ell^{-\frac{1}{2}}\mathbb{B}_\ell^{-\frac{1}{2}}| \leq |\mathbb{B}_\ell^{-\frac{1}{2}}|^2 = \operatorname{tr}\mathbb{B}_\ell^{-1} \leq \frac{d}{\omega}.$$

Also, using the identity

$$\nabla \mathbb{B}_\ell^{-1} = -\mathbb{B}_\ell^{-1}\nabla \mathbb{B}_\ell \mathbb{B}_\ell^{-1},$$

(which is standard for continuously differentiable functions and in general we can approximate $\mathbb{B}_\ell$ by smooth mappings and pass to the limit) and (3.79) we conclude that $\mathbb{B}_\ell^{-1}$ exists a.e. in Q and satisfies

$$\mathbb{B}_\ell^{-1} \in L^\infty(0,T; L^\infty(\Omega; \mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T; W^{1,2}(\Omega; \mathbb{R}^{d\times d}_{>0})). \tag{3.121}$$

Moreover, recalling ψ_e from (2.10) and using the simple inequalities

$$\det \mathbb{B}_\ell \geq \omega^d \quad \text{and} \quad |\ln x| \leq x + \frac{1}{x},\ x > 0,$$

it is easy to see that

$$0 \leq \psi_e(\mathbb{B}_\ell) \leq C|\mathbb{B}_\ell| + \frac{C}{\omega}$$

and

$$|\nabla \psi_e(\mathbb{B}_\ell)| = |(\mathbb{I} - \mathbb{B}_\ell^{-1}) \cdot \nabla \mathbb{B}_\ell| \leq C\left(1 + \frac{1}{\omega}\right)|\nabla \mathbb{B}_\ell|,$$

hence

$$\psi_e(\mathbb{B}_\ell) \in L^2(0,T; W^{1,2}(\Omega; \mathbb{R}_{\geq 0})) \cap L^{q+2}(Q; \mathbb{R}_{\geq 0}).$$

Next, we prove positivity of θ_ℓ. Since $\theta_\ell \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$, we can use the analogous method as before. Indeed, we start by choosing

$$\tau = \chi_{(0,t)}(\theta_\ell - \omega)_- \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$$

as a test function in (3.117) to get

$$\begin{aligned}\frac{c_v}{2}\|(\theta_\ell-\omega)_-(t)\|_2^2 - \frac{c_v}{2}\|(\theta_\ell-\omega)_-(0)\|_2^2 \\ + \int_0^t \|\sqrt{\kappa(\theta_\ell)}\nabla(\theta_\ell-\omega)_-\|_2^2 + \int_0^t \|\nabla(\theta_\ell-\omega)_-\|_{r+2}^{r+2} \\ = \int_0^t \big(2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, (\theta_\ell-\omega)_-\big) \le 0.\end{aligned} \tag{3.122}$$

Hence, using $\theta_\ell(0) = \theta_0^\omega \ge \omega$ in Ω and (3.109), we obtain that $\|(\theta_\ell(t)-\omega)_-\|_2 = 0$ for all $t \in (0,T)$, which means

$$\theta_\ell(t) \ge \omega \quad \text{a.e. in } \Omega \text{ and for all } t \in (0,T). \tag{3.123}$$

Consequently, since $\nabla\theta_\ell^{-1} = \theta_\ell^{-2}\nabla\theta_\ell$, we also obtain

$$\theta_\ell^{-1} \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}_{>0})) \cap L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}_{>0})). \tag{3.124}$$

From these findings we also easily read that

$$|\ln\theta_\ell| \le \theta_\ell + \frac{1}{\theta_\ell} \le \theta_\ell + \frac{1}{\omega} \quad \text{and} \quad |\nabla\ln\theta_\ell| = \frac{|\nabla\theta_\ell|}{\theta_\ell} \le \frac{1}{\omega}|\nabla\theta_\ell|,$$

hence

$$\ln\theta_\ell \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R})).$$

Entropy equation. In order to take the remaining limits $\ell \to \infty$ and $\omega \to 0_+$, we need to replace (3.117) by entropy inequality, whose terms are easier to handle. From this equation, we then deduce that $\det\mathbb{B}_\ell$ and θ_ℓ remain *strictly* positive a.e. in Q.

First, we rewrite (3.117) in the form

$$\begin{aligned}\langle c_v\partial_t\theta_\ell, \tau\rangle + (c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell, \tau) + (\kappa(\theta_\ell)\nabla\theta_\ell + \omega|\nabla\theta_\ell|^r\nabla\theta_\ell, \nabla\tau) \\ = (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, \tau)\end{aligned} \tag{3.125}$$

for all $\tau \in W^{1,r+2}(\Omega;\mathbb{R})$ and a.e. in $(0,T)$. Then, we take $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$ and note that $\tau = \theta_\ell^{-1}\phi$ can be used as a test function in (3.125) thanks to (3.124). This way, we get

$$\begin{aligned}\Big\langle c_v\partial_t\theta_\ell, \frac{\phi}{\theta_\ell}\Big\rangle + (c_v\boldsymbol{v}_\ell\cdot\nabla\ln\theta_\ell, \phi) + (\kappa(\theta_\ell)\nabla\ln\theta_\ell, \nabla\phi) - (\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2, \phi) \\ + \omega(|\nabla\theta_\ell|^r\nabla\ln\theta_\ell, \nabla\phi) - \omega(|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2, \phi) \\ = \Big(\frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, \phi\Big)\end{aligned} \tag{3.126}$$

a.e. in $(0,T)$. Similarly, we observe that $\mu(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi$ is a valid test function in (the localized version of) (3.96) due to (3.121). Thus, we obtain (recall the computation between (2.11) and (2.13))

$$\begin{aligned}&\langle \partial_t \mathbb{B}_\ell, \mu(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\rangle + (\boldsymbol{v}_\ell\cdot\nabla\psi_e(\mathbb{B}_\ell),\phi)\\ &\quad + (\mu P(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1}),\phi) + (\mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2,\phi)\\ &\qquad = -(\lambda(\theta_\ell)\nabla\psi_e(\mathbb{B}_\ell),\nabla\phi) + (2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\phi)\end{aligned} \tag{3.127}$$

a.e. in $(0,T)$. If we define

$$\eta_\ell := c_v\ln\theta_\ell - \psi_e(\mathbb{B}_\ell) \tag{3.128}$$

and

$$\begin{aligned}\xi_\ell := \frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 + \kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2 + \omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2\\ + \mu P(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1}) + \mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2\end{aligned} \tag{3.129}$$

and subtract (3.127) from (3.126), we get

$$\begin{aligned}&\Big\langle c_v\partial_t\theta_\ell,\frac{\phi}{\theta_\ell}\Big\rangle - \langle\partial_t\mathbb{B}_\ell,\mu(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\rangle + (\boldsymbol{v}_\ell\cdot\nabla\eta_\ell,\phi)\\ &\quad + \big((\kappa(\theta_\ell)+\omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla\psi_e(\mathbb{B}_\ell),\nabla\phi\big) = (\xi_\ell,\phi)\end{aligned} \tag{3.130}$$

a.e. in $(0,T)$ and for all $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$.

Obviously, we need to rewrite the time derivative accordingly. Concerning the term containing $\partial_t\theta_\ell$, note that $\psi(s)=\max\{|s|,\omega\}^{-1}$, $s\in\mathbb{R}$, is a bounded Lipschitz function. Since $\theta_\ell\ge\omega$ a.e. in Q by (3.123) and $\omega<\omega_P$, we get

$$\int_{\omega_P}^{\theta_\ell}\psi(s)\,\mathrm{d}s = \int_{\omega_P}^{\theta_\ell}\frac1s\,\mathrm{d}s = \ln\theta_\ell.$$

Thus, Lemma 2 below yields

$$\Big\langle c_v\partial_t\theta_\ell,\frac{\phi}{\theta_\ell}\Big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}(c_v\ln\theta_\ell,\phi).$$

Hence, if we multiply this by $\varphi\in W^{1,\infty}((0,T);\mathbb{R})$ with $\varphi(T)=0$, integrate over $(0,T)$ and by parts, we are led to

$$\int_0^T\Big\langle c_v\partial_t\theta_\ell,\frac{\phi}{\theta_\ell}\Big\rangle\varphi = -\int_0^T\int_\Omega c_v\ln\theta_\ell\phi\partial_t\varphi - \int_\Omega c_v\ln\theta_0^\omega\phi\varphi(0), \tag{3.131}$$

where we also used (3.114).

Analogous ideas can be used to rewrite the second term of (3.130). However, since the duality $\langle\partial_t\mathbb{B}_\ell,(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\rangle$ can not be interpreted entry-wise, let us proceed more carefully. We apply Lemma 1 below to obtain functions $\mathbb{B}_\ell^\varepsilon\in\mathcal{C}^1([0,T];W^{1,2}(\Omega)\cap L^{q+2}(\Omega))$ such that

$$\|\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell\|_{L^2W^{1,2}\cap L^{q+2}L^{q+2}} + \|\partial_t\mathbb{B}_\ell^\varepsilon-\partial_t\mathbb{B}_\ell\|_{L^2W^{-1,2}+L^{\frac{q+2}{q+1}}L^{\frac{q+2}{q+1}}}\to 0 \tag{3.132}$$

as $\varepsilon\to0_+$ and also $\Lambda(\mathbb{B}_\ell^\varepsilon)\ge\omega$ a.e. in Q. Since $\mathbb{B}_\ell\in\mathcal{C}([0,T];L^2(\Omega))$ (cf. (3.98)), we know that

$$\|\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell\|_2\rightrightarrows 0\quad\text{uniformly in }[0,T]. \tag{3.133}$$

Furthermore, using (3.121), we can write, for any $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$, that

$$\begin{aligned}\big|\nabla((\mathbb{I}-(\mathbb{B}_\ell^\varepsilon)^{-1})\phi)\big| &= \big|(\mathbb{B}_\ell^\varepsilon)^{-1}\nabla\mathbb{B}_\ell^\varepsilon(\mathbb{B}_\ell^\varepsilon)^{-1}\phi + (\mathbb{I}-(\mathbb{B}_\ell^\varepsilon)^{-1})\nabla\phi\big|\\ &\le \frac{C}{\omega^2}|\nabla\mathbb{B}_\ell^\varepsilon||\phi| + \left(1+\frac{C}{\omega}\right)|\nabla\phi|\end{aligned}$$

and thus, we eventually obtain that

$$(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi \rightharpoonup (\mathbb{I} - \mathbb{B}_\ell^{-1})\phi \text{ weakly in } L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\rm sym})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\rm sym}).$$

By applying this with (3.132), we get, for all $\varphi \in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T) = 0$, that

$$\begin{aligned}
&\left|\int_0^T \langle \partial_t \mathbb{B}_\ell^\varepsilon, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\rangle\varphi - \int_0^T \langle \partial_t \mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\rangle\varphi\right| \\
&\qquad \le \int_0^T \left|\langle \partial_t \mathbb{B}_\ell^\varepsilon - \partial_t \mathbb{B}_\ell, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\rangle\right| |\varphi| \\
&\qquad\quad + \left|\int_0^T \langle \partial_t \mathbb{B}_\ell \varphi, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi - \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\rangle\right| \\
&\qquad \to 0 \quad \text{as } \varepsilon \to 0_+.
\end{aligned} \tag{3.134}$$

On the other hand, using $\partial_t \mathbb{B}_\ell^\varepsilon \in \mathcal{C}([0,T];W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\rm sym}) \cap L^{q+2}(\Omega;\mathbb{R}^{d\times d}_{\rm sym}))$, we find

$$\begin{aligned}
\int_0^T \langle \partial_t \mathbb{B}_\ell^\varepsilon, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\rangle\varphi &= \int_0^T (\partial_t \psi_e(\mathbb{B}_\ell^\varepsilon), \phi)\varphi \\
&= -\int_\Omega \psi_e(\mathbb{B}_\ell^\varepsilon(0))\phi\varphi(0) - \int_0^T \int_\Omega \psi_e(\mathbb{B}_\ell^\varepsilon)\phi\partial_t\varphi
\end{aligned} \tag{3.135}$$

To take the limit in the last two terms, let us first remark that the set

$$\{\mathbb{A} \in \mathbb{R}^{d\times d}_{\rm sym} : \ \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} \ge \omega \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d\}$$

is convex in $\mathbb{R}^{d\times d}_{\rm sym}$. This and (3.120) imply, for any $s \in (0,1)$, that

$$\left|(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1}\right| \le \operatorname{tr}(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1} \le \frac{1}{d\Lambda(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))} \le \frac{1}{d\omega}$$

a.e. in Ω. Thus, by the mean value theorem and (3.133), we get

$$\begin{aligned}
\int_\Omega |\psi_e(\mathbb{B}_\ell^\varepsilon) - \psi_e(\mathbb{B}_\ell)|^2 &= \int_\Omega \left|\int_0^1 (\mathbb{I} - (\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1}) \cdot (\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell)\,\mathrm{d}s\right|^2 \\
&\le C\left(1 + \frac{1}{\omega}\right) \|\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell\|_2^2 \rightrightarrows 0 \quad \text{uniformly in } [0,T]
\end{aligned} \tag{3.136}$$

as $\varepsilon \to 0_+$. Similarly, using (3.98) and (3.101), we can also show that

$$\psi_e(\mathbb{B}_\ell) \in \mathcal{C}(0,T;L^2(\Omega;\mathbb{R})), \quad \psi_e(\mathbb{B}_\ell(0)) = \psi_e(\mathbb{B}_0^\omega). \tag{3.137}$$

Using this and (3.136), we take the limit $\varepsilon \to 0_+$ in (3.135) and compare the result with (3.134) to obtain

$$\int_0^T \langle \partial_t \mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\rangle\varphi = -\int_\Omega \psi_e(\mathbb{B}_0^\omega)\phi\varphi(0) - \int_0^T \int_\Omega \psi_e(\mathbb{B}_\ell)\phi\partial_t\varphi \tag{3.138}$$

for all $\varphi \in W^{1,\infty}(\Omega;\mathbb{R})$, $\varphi(T) = 0$, and every $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$.

Finally, if we subtract (3.138) from (3.131) and use (3.128), we can rewrite (3.130) as

$$\begin{aligned}
&-\int_0^T (\eta_\ell, \phi)\partial_t\varphi - (\eta_0^\omega, \phi)\varphi(0) - \int_0^T (\boldsymbol{v}_\ell\eta_\ell, \nabla\phi)\varphi \\
&\quad + \int_0^T \big((\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla\psi_e(\mathbb{B}_\ell), \nabla\phi\big)\varphi = \int_0^T (\xi_\ell, \phi)\varphi
\end{aligned} \tag{3.139}$$

for all $\varphi \in W^{1,\infty}(0,T;\mathbb{R})$, $\varphi(T)=0$, and $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$, where

$$\eta_0^\omega := c_v \ln \theta_0^\omega - \psi_e(\mathbb{B}_0^\omega).$$

Moreover, since $\ln \theta_\ell \in \mathcal{C}([0,T];L^2(\Omega;\mathbb{R}))$ and (3.137) hold, we easily read

$$\eta_\ell \in \mathcal{C}([0,T];L^2(\Omega;\mathbb{R})), \quad \eta_\ell(0) = \eta_0^\omega. \tag{3.140}$$

Total energy equality. The integrated version of the total energy equality is important in the derivation of the a priori estimates below. We multiply the i-th equation in (3.95) by $(\boldsymbol{v}_\ell, \boldsymbol{w}_i)$, sum up the result over $i=1,\dots,\ell$ and then we add (3.117) with $\tau = 1$. This way, after several cancellations using also $(1.13)_1$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E_\ell = (\boldsymbol{f}, \boldsymbol{v}_\ell) \quad \text{a.e. in } (0,T), \tag{3.141}$$

where $E_\ell := \frac{1}{2}|\boldsymbol{v}_\ell|^2 + c_v\theta_\ell$.

ℓ,ω-uniform estimates. Here we shall derive uniform estimates needed for the final limit passage. The fact that these estimates are uniform with respect to both ℓ and ω saves us some work (in exchange for a slight non-optimality with respect to ω).

Let us first show that the total energy of the fluid remains bounded. In (3.141), we apply Young's inequality, (3.29) and $\theta_\ell > 0$, to estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E_\ell \le \frac{1}{2}\int_\Omega |\boldsymbol{v}_\ell|^2 + \frac{1}{2}\int_\Omega |\boldsymbol{f}|^2 \le \int_\Omega E_\ell + \frac{1}{2}\int_\Omega |\boldsymbol{f}|^2$$

a.e. in $(0,T)$. Hence, by the Gronwall inequality, we get

$$\int_\Omega E_\ell(t) \le e^t\left(\int_\Omega E_\ell(0) + \frac{1}{2}\int_0^t \|\boldsymbol{f}\|_2^2\right) \quad \text{for all } t \in [0,T].$$

Then, we apply (3.102), (3.114) to identify that

$$E_\ell(0) = \frac{1}{2}|P_\ell \boldsymbol{v}_0|^2 + c_v\theta_0^\omega$$

and if we use properties of P_ℓ, (3.59), (3.29), we arrive at

$$\|\theta_\ell\|_{L^\infty L^1} + \|\boldsymbol{v}_\ell\|_{L^\infty L^2} \le C\|E_\ell\|_{L^\infty L^1} \le C. \tag{3.142}$$

Now we turn our attention to (3.139), which we localize in time by choosing[7] $\varphi = \chi_{(0,t)}$, leading to

$$\int_\Omega \eta_\ell(t)\phi + \int_0^t\int_\Omega \boldsymbol{j}_\ell\cdot\nabla\phi = \int_\Omega \eta_0^\omega\phi + \int_0^t\int_\Omega \xi_\ell\phi \quad \text{for all } \phi \in W^{1,\infty}(\Omega;\mathbb{R}) \tag{3.143}$$

and all $t \in (0,T)$ (in fact, for all $t \in [0,T]$ due to continuity of both sides of (3.143)), where

$$\boldsymbol{j}_\ell := -\boldsymbol{v}_\ell\eta_\ell + (\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla\psi_e(\mathbb{B}_\ell) \in L^1(Q;\mathbb{R}^d).$$

In particular, taking $\phi = 1$, we deduce, using $\xi_\ell \ge 0$, that the function $t \mapsto \int_\Omega \eta_\ell(t)$ is non-decreasing, and thus

$$\int_Q \xi_\ell = \max_{t\in[0,T]}\int_0^t\int_\Omega \xi_\ell = \max_{t\in[0,T]}\int_\Omega \eta_\ell(t) - \int_\Omega \eta_0^\omega = \int_\Omega \eta_\ell(T) - \int_\Omega \eta_0^\omega. \tag{3.144}$$

---

[7] Strictly speaking, as $\chi_{(0,t)}$ is not Lipschitz, we can not use it directly in (3.139). However, a standard argument using a piecewise linear approximation of $\chi_{(0,t)}$ with the Lebesgue differentiation theorem and absolute continuity of integral shows that $\chi_{(0,t)}$ is a reasonable test function.

Then, using (3.128), the inequalities

$$\ln x \leq x - 1 \quad \text{for all } x > 0 \quad \text{and} \quad \psi_e(\mathbb{B}_\ell) \geq 0, \tag{3.145}$$

assumption (3.30) and (3.142) (recall also (3.109)), we obtain

$$\int_Q \xi_\ell \leq \int_\Omega (c_v \ln \theta_\ell(T) - \psi_e(\mathbb{B}_\ell(T))) + C \leq C \int_\Omega (\theta_\ell(T) - 1) + C \leq C, \tag{3.146}$$

hence

$$\|\xi_\ell\|_{L^1L^1} \leq C. \tag{3.147}$$

Also, it is easy to see using (3.144), (3.60), (3.59), (3.29) and (3.30) that

$$\|\eta_\ell\|_{L^\infty L^1} \leq C. \tag{3.148}$$

Estimate (3.147) implies, using (3.2) and (3.9), that

$$\begin{aligned} \|\theta_\ell^{-\frac{1}{2}} \mathbb{D}\boldsymbol{v}_\ell\|_{L^2L^2} + \|\sqrt{\kappa(\theta_\ell)} \nabla \ln \theta_\ell\|_{L^2L^2} + \omega \||\nabla\theta_\ell|^{\frac{r}{2}} \nabla \ln \theta_\ell\|_{L^2L^2} \\ + \|\mathbb{B}_\ell^{-\frac{1}{2}} \nabla \mathbb{B}_\ell \mathbb{B}_\ell^{-\frac{1}{2}}\|_{L^2L^2} \leq C. \end{aligned} \tag{3.149}$$

In what follows, we improve the uniform estimate (3.149) considerably by choosing appropriate test functions in (3.96) and (3.117) and then using (C_1) and the definitions of p, R, σ to estimate the right hand sides.

Our aim is to test (3.96) by the $\mathbb{B}_\ell^{\sigma-1}$. To verify that this is a valid test function, we show first that $\mathbb{B}_\ell$ is actually essentially bounded. Indeed, taking $\mathbb{A} = \chi_{(0,t)}\phi\mathbb{I}$, $t \in (0,T)$, $\phi \in L^{q+2}(0,T; L^{q+2}(\Omega;\mathbb{R})) \cap L^2(0,T; W^{1,2}(\Omega;\mathbb{R}))$, in (3.96) yields

$$\begin{aligned} \int_0^t \langle \partial_t \operatorname{tr} \mathbb{B}_\ell, \phi\rangle + \int_0^t (\boldsymbol{v} \cdot \nabla \operatorname{tr} \mathbb{B}_\ell, \phi) + \int_0^t (P(\theta_\ell, \mathbb{B}_\ell) \cdot \mathbb{I}, \phi) + \int_0^t (\lambda(\theta_\ell) \nabla \operatorname{tr} \mathbb{B}_\ell, \nabla\phi) \\ = \int_0^t (2a g_\omega(\mathbb{B}_\ell, \theta_\ell) \mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \phi), \end{aligned}$$

hence, recalling (3.8) there exists a constant $C_0 > 0$, such that

$$\int_0^t \langle \partial_t \operatorname{tr} \mathbb{B}_\ell, \phi\rangle + \int_0^t (\boldsymbol{v} \cdot \nabla \operatorname{tr} \mathbb{B}_\ell, \phi) + \int_0^t (\lambda(\theta_\ell) \nabla \operatorname{tr} \mathbb{B}_\ell, \nabla\phi) \leq C_0 \int_0^t |\phi|.$$

Substituting $u(\boldsymbol{x}, t) := \operatorname{tr} \mathbb{B}_\ell(\boldsymbol{x}, t) - C_0 t$ leads to

$$\int_0^t \langle \partial_t u, \phi\rangle + \int_0^t (\boldsymbol{v} \cdot \nabla u, \phi) + \int_0^t (\lambda(\theta_\ell) \nabla u, \nabla\phi) \leq C_0 \int_0^t (|\phi| - \phi).$$

If we choose $\phi = (u - K)_+$ and use (1.7), $(1.13)_1$ to eliminate the convective term, we obtain

$$\frac{1}{2}\|(u(t) - K)_+\|_2^2 + \int_0^t \|\sqrt{\lambda(\theta_\ell)} \nabla (u - K)_+\|_2^2 \leq \frac{1}{2}\|(u(0) - K)_+\|_2^2.$$

If we let $K := \frac{d}{\omega}$, then (3.101) and (3.58) imply

$$(u(0) - K)_+ = (\operatorname{tr} \mathbb{B}_0^\omega - \tfrac{d}{\omega})_+ \leq (\sqrt{d}|\mathbb{B}_0^\omega| - \tfrac{d}{\omega})_+ = 0$$

in Ω. Thus, we get $\|(u(t) - \frac{d}{\omega})_+\|_2^2 = 0$, hence

$$|\mathbb{B}_\ell| \leq \operatorname{tr} \mathbb{B}_\ell \leq \tfrac{d}{\omega} + C_0 t \leq \tfrac{d}{\omega} + C_0 T$$

and we see that indeed

$$\mathbb{B}_\ell \in L^\infty(0,T; L^\infty(\Omega; \mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T; W^{1,2}(\Omega; \mathbb{R}^{d\times d}_{>0})).$$

We claim that the same property holds for $\mathbb{B}_\ell^{\sigma-1}$. Indeed, boundedness is easy to see using a spectral decomposition, while the gradient can be computed from the identity

$$\nabla \mathbb{B}_\ell^{\sigma-1} \tag{3.150}$$
$$= (\sigma-1)\int_0^1\int_0^1 \mathbb{B}_\ell^{(1-s)(\sigma-1)}((1-t)\mathbb{I}+t\mathbb{B}_\ell)^{-1}\nabla\mathbb{B}_\ell((1-t)\mathbb{I}+t\mathbb{B}_\ell)^{-1}\mathbb{B}_\ell^{s(\sigma-1)}\,\mathrm{d}s\,\mathrm{d}t,$$

which is a consequence of the well known identities for $\nabla \exp \mathbb{A}$ and $\nabla \log \mathbb{A}$, see [43], [44] or [3] and references therein for details. Since

$$|((1-t)\mathbb{I}+t\mathbb{B}_\ell)^{-1}| \le \sqrt{d}\Lambda((1-t)\mathbb{I}+t\mathbb{B}_\ell)^{-1} \le \sqrt{d}K$$

we read from (3.150) that $\nabla\mathbb{B}_\ell^{\sigma-1} \in L^2(0,T;L^2(\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}))$, hence

$$\mathbb{B}_\ell^{\sigma-1} \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0}))$$

can be used as a test function in (3.96). This way, using (3.7), the identities[8]

$$\langle \partial_t \mathbb{B}_\ell, \mathbb{B}_\ell^{\sigma-1}\rangle = \frac{1}{\sigma}\int_\Omega \partial_t \operatorname{tr}\mathbb{B}_\ell^\sigma,$$

$$(\boldsymbol{v}\cdot\nabla\mathbb{B}_\ell, \mathbb{B}_\ell^{\sigma-1}) = \frac{1}{\sigma}\int_\Omega \boldsymbol{v}\cdot\nabla\operatorname{tr}\mathbb{B}_\ell^\sigma = 0$$

and the estimate

$$\nabla\mathbb{B}_\ell\cdot\nabla\mathbb{B}_\ell^{\sigma-1} \ge \frac{4(\sigma-1)}{\sigma^2}|\nabla\mathbb{B}_\ell^{\frac{\sigma}{2}}|^2$$

from Lemma 3 (iv), (v) below, we get

$$\frac{1}{\sigma}\int_\Omega(\operatorname{tr}\mathbb{B}_\ell^\sigma(t)-\operatorname{tr}\mathbb{B}_\ell^\sigma(0)) + C\int_0^t\int_\Omega |\mathbb{B}|^{q+\sigma} + \frac{4(\sigma-1)}{\sigma^2}\int_0^t\int_\Omega \lambda(\theta_\ell)|\nabla\mathbb{B}_\ell^{\frac{\sigma}{2}}|^2$$
$$= 2a\int_0^t\int_\Omega g(\mathbb{B}_\ell,\theta_\ell)\mathbb{D}\boldsymbol{v}_\ell\cdot\mathbb{B}_\ell^\sigma.$$

If we apply (3.101), (3.4), $g_\omega \le 1$ and $|\mathbb{B}_\ell^\sigma| \le \max\{1, d^{\frac{1-\sigma}{2}}\}|\mathbb{B}_\ell|^\sigma$ (see [3]), we deduce

$$\int_\Omega(\operatorname{tr}\mathbb{B}_\ell^\sigma(t)-\operatorname{tr}\mathbb{B}_0^\omega) + \int_0^t\int_\Omega|\mathbb{B}_\ell|^{q+\sigma} + \int_0^t\int_\Omega|\nabla\mathbb{B}_\ell^{\frac{\sigma}{2}}|^2 \le C\int_0^t\int_\Omega|\mathbb{D}\boldsymbol{v}_\ell||\mathbb{B}_\ell|^\sigma. \tag{3.151}$$

Then, by (3.59), (3.29) and Young's inequality, we arrive at

$$\|\mathbb{B}_\ell\|^\sigma_{L^\infty L^\sigma} + \|\mathbb{B}_\ell\|^{q+\sigma}_{L^{q+\sigma}L^{q+\sigma}} + \|\nabla\mathbb{B}_\ell^{\frac{\sigma}{2}}\|^2_{L^2L^2} \le C + C\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{q+\sigma}{q}}_{L^{\frac{q+\sigma}{q}}L^{\frac{q+\sigma}{q}}}, \tag{3.152}$$

where the right hand side is finite due to (3.86).

Next, we use (3.152) and (3.117) to improve the information about θ_ℓ and $\mathbb{D}\boldsymbol{v}_\ell$. For any $\beta \in [0,1)$, we can show that $\theta_\ell^{-\beta} \in L^{r+2}(0,T;W^{1,r+2}(\Omega))\cap L^\infty(0,T;L^\infty(\Omega))$ similarly as in (3.124), and thus $\tau_\beta = -\theta_\ell^{-\beta}$ is an admissible test function in (3.117).

[8] To interpret the duality pairing in the first identity, one has to approximate $\mathbb{B}_\ell$ similarly as before when dealing with $\langle\partial_t\mathbb{B}_\ell, (\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\rangle$.

Using Lemma 2 with $\psi(s) = -\max(s,\omega)^{-\beta}$ to rewrite the time derivative, (3.142) with Young's inequality, integration by parts, (1.7), $(1.13)_1$ and (3.3), we obtain

$$\begin{aligned}\int_0^T \langle c_v \partial_t \theta_\ell, \tau_\beta\rangle + \int_0^T (c_v \boldsymbol{v}_\ell \cdot \nabla\theta_\ell, \tau_\beta) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla\tau_\beta) + \omega\int_0^T (|\nabla\theta_\ell|^r \nabla\theta_\ell, \nabla\tau_\beta)&\\ \geq \frac{c_v}{1-\beta}\int_\Omega ((\theta_0^{\frac{1}{\omega}})^{1-\beta} - \theta_\ell^{1-\beta}(T)) + \beta\int_Q \theta_\ell^{-1-\beta}\kappa(\theta_\ell)|\nabla\theta_\ell|^2&\\ \geq C\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 - C.& \qquad (3.153)\end{aligned}$$

We use this estimate in (3.117) with $\tau = \tau_\beta$ to deduce, using also $g_\omega \leq 1$, Hölder's inequality and (3.152) that, in the case $\sigma < q$, we have

$$\begin{aligned}\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\leq C\int_Q \theta_\ell^{1-\beta}|\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| + C\\ &\leq C\|\theta_\ell^{1-\frac{\beta}{2}}\|_{\frac{2(q+\sigma)}{q+\sigma-2};Q}\|\mathbb{B}_\ell\|_{q+\sigma;Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q} + C\\ &\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{\frac{q+\sigma}{q};Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\\ &\quad + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q} + C\\ &\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell^{\beta}\|^{\frac{1}{2q}}_{\frac{q+\sigma}{q-\sigma};Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|^{1+\frac{1}{q}}_{2;Q}\\ &\quad + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q} + C\end{aligned} \qquad (3.154)$$

while if $\sigma = q$, we omit the final step to get

$$\begin{aligned}&\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 \qquad (3.155)\\ &\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{(2-\beta)q';Q}\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{2;Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q} + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{(2-\beta)q';Q}\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q} + C.\end{aligned}$$

Thus, using $q > 1$, $(\frac{2q}{q+1})' = 2q'$ and the Young inequality, we arrive at

$$\begin{aligned}&\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 \qquad (3.156)\\ &\leq C\|\theta_\ell\|^{(2-\beta)q'}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell^{\beta}\|^{\frac{q'}{q}}_{\frac{q+\sigma}{q-\sigma};Q} + C\|\theta_\ell\|^{2-\beta}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q} + C\end{aligned}$$

if $\sigma < q$ and

$$\begin{aligned}&\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 \qquad (3.157)\\ &\leq C\|\theta_\ell\|^{2-\beta}_{(2-\beta)q';Q}\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{2}{q}}_{2;Q} + C\|\theta_\ell\|^{2-\beta}_{(2-\beta)q';Q} + C\end{aligned}$$

if $\sigma = q$, respectively. Next we focus on the case $\sigma < q$. Let us define

$$\beta_0 := \max\Big\{0, r_d + 1 - \frac{(r_d-1)(q+\sigma)}{2}\Big\}, \qquad \beta_1 := \min\Big\{1, \frac{(r_d+1)(q-\sigma)}{2q}\Big\}$$

and note using $\sigma > 1$ and $r_d > r_0$ (i.e. (C$_1$)) that

$$0 \le \beta_0 < \beta_1 \le 1.$$

Then, by a routine computation one can verify that the inequality

$$\max\left\{(2-\beta)\frac{q+\sigma}{q+\sigma-2}, \beta\frac{q+\sigma}{q-\sigma}\right\} \le r_d + 1 - \beta \tag{3.158}$$

holds if

$$r_0 < r_d < r_1 \quad \text{and} \quad \beta_0 \le \beta \le \beta_1, \tag{3.159}$$

or

$$r_d \ge r_1 \quad \text{and} \quad 0 \le \beta \le \beta_1. \tag{3.160}$$

Thus, by application of (3.158), the Hölder inequality, an interpolation inequality, the Sobolev inequality, the Poincaré inequality and (3.142), we get

$$\begin{aligned}\|\theta_\ell\|_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q} + \|\theta_\ell\|_{\frac{\beta(q+\sigma)}{q-\sigma};Q} &\le C\|\theta_\ell\|_{r_d+1-\beta;Q} \\ &\le C\|\theta_\ell\|_{L^\infty L^1}^{\frac{2}{d(r+1-\beta)+2}} \|\theta_\ell\|_{L^{r+1-\beta}L^{\frac{d}{d-2}(r+1-\beta)}}^{\frac{d(r+1-\beta)}{d(r+1-\beta)+2}} \\ &\le C\|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{L^2L^{\frac{2d}{d-2}}}^{\frac{2d}{d(r+1-\beta)+2}} \\ &\le C\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\|_{L^2L^2}^{\frac{2}{r_d+1-\beta}} + C.\end{aligned} \tag{3.161}$$

Using this in (3.156), applying Young's inequality and $q' > 1$, we obtain

$$\begin{aligned}\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\le C\|\theta_\ell\|_{r_d+1-\beta;Q}^{2q'-\beta} + C\|\theta_\ell\|_{r_d+1-\beta;Q}^{2-\beta} + C \\ &\le C\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\|_{2;Q}^{\frac{2(2q'-\beta)}{r_d+1-\beta}} + C,\end{aligned} \tag{3.162}$$

where the last exponent is strictly less than two due to $2q' < r_d + 1$ (which is equivalent to (C$_1$)). Hence, we can apply the Young inequality in (3.162) to finally get

$$\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 \le C(\beta) \tag{3.163}$$

for any r, β fulfilling (3.159) or (3.160). In case that (3.159) holds, we make the optimal choice $\beta = \beta_0$ and note that

$$\frac{r+1-\beta_0}{2} = -\frac{1}{d} + \frac{(r_d-1)(q+\sigma)}{4} = \frac{R}{2},$$

cf. (3.27). Then, from Hölder's inequality and (3.161), we deduce

$$\|\mathbb{D}\boldsymbol{v}_\ell\|_{p;Q} \le \|\theta_\ell^{-\frac{\beta_0}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell^{\frac{\beta_0}{2}}\|_{2\frac{r_d+1-\beta_0}{\beta_0};Q} \le C,$$

where (recall (3.26))

$$p = 2\frac{r_d+1-\beta_0}{r_d+1} = \frac{r_d-1}{r_d+1}(q+\sigma).$$

In the special case $r_d = r_1$, we can repeat the above estimates without $\beta = \beta_0$, choosing instead $\beta > 0$ arbitrarily small. Finally, if $r_d > r_1$, we can improve the information on $\mathbb{D}\boldsymbol{v}_\ell$ simply by taking $\tau = -1$ in (3.117). Then, using similar computation as above with β chosen as to satisfy

$$0 < \beta < \min\left\{r_d + 1 - 2\frac{q+\sigma}{q+\sigma-2}, r_d - 1\right\} = r_d + 1 - 2\frac{q+\sigma}{q+\sigma-2} \tag{3.164}$$

and using (3.152), (3.161), $\sigma \le q$, we obtain

$$\begin{aligned}\int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 &\le C\int_Q \theta_\ell|\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| \le C\|\mathbb{B}_\ell\|_{q+\sigma;Q}\|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell\|_{\frac{2(q+\sigma)}{q+\sigma-2};Q} + C\\ &\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{\frac{q+\sigma}{q};Q} + 1\right)\|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell\|_{r_d+1-\beta;Q} + C\\ &\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{1+\frac{1}{q}}_{2;Q} + \|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\right)\left(\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\|^{\frac{2}{r_d+1-\beta}}_{2;Q} + 1\right) + C.\end{aligned} \tag{3.165}$$

Hence, using $1 + \frac{1}{q} < 2$, (3.164), Young's inequality and (3.163), we get

$$\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 \le C(\beta) \tag{3.166}$$

for any β satisfying (3.164). Finally, it remains to consider the excluded case $\sigma = q$. However, in this situation, we have $r_1 = r_0 < r_d$, and thus we can take $\tau = -1$ in (3.117) as before. This way, adding also (3.157) and using analogous estimation as in (3.165), we obtain

$$\begin{aligned}&\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2\\ &\quad\le C\int_Q \theta_\ell|\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| + C\|\theta_\ell\|^{2-\beta}_{(2-\beta)q';Q}\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{2}{q}}_{2;Q} + 1\right) + C\\ &\quad\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{1+\frac{1}{q}}_{2;Q} + \|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\right)\left(\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\|^{\frac{2}{r_d+1-\beta}}_{2;Q} + 1\right)\\ &\qquad + C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{2}{q}}_{2;Q} + 1\right)\left(\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\|^{\frac{2(2-\beta)}{r_d+1-\beta}}_{2;Q} + 1\right) + C.\end{aligned} \tag{3.167}$$

If we choose β as in (3.164) and use Young's inequality, noticing that

$$\frac{2(2-\beta)}{r_d+1-\beta}q' < 2\frac{2q'-\beta q'}{2q'-\beta} = 2\left(1 - \frac{(q'-1)\beta}{2q'-\beta}\right) < 2,$$

we again conclude that (3.166) holds.

To summarize the estimates up to this point, we proved (3.147), (3.148), (3.149),

$$\|\nabla\theta_\ell^{\frac{R}{2}}\|_{L^2L^2} + \|\theta_\ell\|_{L^{R_d}L^{R_d}} \le C \tag{3.168}$$

and

$$\|\mathbb{D}\boldsymbol{v}_\ell\|_{L^pL^p} \le C. \tag{3.169}$$

Moreover, we deduce from (3.152), (3.169) and

$$\frac{q+\sigma}{q} = (q+\sigma)\left(1 - \frac{2}{2q'}\right) < (q+\sigma)\left(1 - \frac{2}{r_d+1}\right) = p \tag{3.170}$$

that

$$\|\mathbb{B}_\ell\|_{L^\infty L^\sigma} + \|\mathbb{B}_\ell\|_{L^{q+\sigma}L^{q+\sigma}} + \|\nabla\mathbb{B}^{\frac{\sigma}{2}}\|_{L^2L^2} \le C. \tag{3.171}$$

Next, the combination of the Cauchy-Schwarz inequality and (4.20) yields

$$|\nabla\mathbb{B}_\ell| \le |\mathbb{B}_\ell^{\frac12}||\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}||\mathbb{B}_\ell^{\frac12}| = \sqrt{d}|\mathbb{B}_\ell||\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|.$$

Then, since $q+\sigma>2$, we deduce by appealing to the Hölder inequality and (3.79), (3.171) that

$$\|\nabla\mathbb{B}_\ell\|_{L^zL^z} \le C. \tag{3.172}$$

Below, we derive the uniform estimates for the time derivatives. To this end, we need to determine integrability of the non-linear terms in (3.95), (3.96) and (3.139). It follows from an interpolation inequality, Korn's inequality, (3.142) and (3.169) that

$$\|\boldsymbol{v}_\ell\|_{L^{p\frac{d+2}{d}}L^{p\frac{d+2}{d}}} \le C\|\boldsymbol{v}_\ell\|_{L^\infty L^2}^{\frac{2}{d+2}}\|\mathbb{D}\boldsymbol{v}_\ell\|_{L^pL^p}^{\frac{d}{d+2}} \le C. \tag{3.173}$$

We remark that (3.26) and (C_2) imply

$$p\left(1+\frac{2}{d}\right) > 2. \tag{3.174}$$

Furthermore, the Hölder inequality, (3.168) and (3.152) yield

$$\|\theta_\ell\mathbb{B}_\ell\|_{L^pL^p} \le C, \tag{3.175}$$

Hence, as $d\ge 2$, we read from (3.95) that

$$\|\partial_t\boldsymbol{v}_\ell\|_{L^{p\frac{d+2}{2d}}W^{-1,p\frac{d+2}{2d}}_{\boldsymbol{n},\mathrm{div}}} \le C. \tag{3.176}$$

Next, we focus on the non-linear terms in (3.96). There, we integrate by parts in the convective term with the help of $(1.13)_1$. Then, using Hölder's inequality and (3.152), (3.173), we observe that

$$\|\mathbb{B}_\ell\otimes\boldsymbol{v}_\ell\|_{L^{s_1}L^{s_1}} \le C, \tag{3.177}$$

with

$$s_1 = \left(\frac{1}{q+\sigma}+\frac{1}{p\frac{d+2}{d}}\right)^{-1} > \left(\frac{1}{q+\sigma}+\frac12\right)^{-1} = z, \tag{3.178}$$

where we used (3.174). Moreover, making use of (3.171) and (3.6), we obtain

$$\|P(\theta_\ell,\mathbb{B}_\ell)\|_{L^{s_2}L^{s_2}} \le C, \quad \text{where} \quad s_2 = \frac{q+\sigma}{q+1} \le z. \tag{3.179}$$

Furthermore, using (3.171), (3.169) and Hölder's inequality, we get

$$\|(a\mathbb{D}\boldsymbol{v}_\ell+\mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell\|_{L^{s_3}L^{s_3}} \le C, \tag{3.180}$$

where

$$s_3 = \frac{1}{\frac{1}{q+\sigma}+\frac1p} = \frac{q+\sigma}{1+\frac{r_d+1}{r_d-1}} > \frac{q+\sigma}{q+1} = s_2 \tag{3.181}$$

(using $r_d > r_0$). Thus, we read from (3.96) using (3.172), (3.177), (3.178) and (3.179), (3.180), (3.181) that

$$C \ge \|\partial_t\mathbb{B}_\ell\|_{(L^{z'}W^{1,z'}\cap L^{s_2'}L^{s_2'})^*} \ge C\|\partial_t\mathbb{B}_\ell\|_{L^{s_2}W^{-1,s_2}}, \tag{3.182}$$

using some trivial embeddings.

Finally, we examine the non-linearities related to (3.139). There, since ξ_ℓ is controlled by (3.147), the problematic terms could be only on the left hand side. To get an appropriate uniform control over the convective term, we estimate

$$\eta_\ell \le \eta_\ell + \psi_e(\mathbb{B}_\ell) = c_v\ln\theta_\ell \le c_v(\theta_\ell-1).$$

This, together with (3.142) and (3.148), yields

$$\|\ln\theta_\ell\|_{L^\infty L^1} \le C. \tag{3.183}$$

Then, since (3.149) and (3.3) give

$$\|\nabla\ln\theta_\ell\|_{L^2L^2} \le C, \tag{3.184}$$

we can use Sobolev's inequality, Poincaré's inequality and an interpolation to obtain

$$\|\ln\theta_\ell\|_{L^{2+\frac{2}{d}}L^{2+\frac{2}{d}}} \le C\|\ln\theta_\ell\|_{L^\infty L^1}^{\frac{1}{d+1}}\|\ln\theta_\ell\|_{L^2W^{1,2}}^{\frac{d}{d+1}} \le C. \tag{3.185}$$

Now we observe that a similar reasoning applies also for the quantity $\ln\det\mathbb{B}_\ell$. Indeed, using (3.148), (3.183), (3.152) and (3.128) in the form

$$\ln\det\mathbb{B}_\ell = \frac{1}{\mu}(\eta_\ell - c_v\ln\theta_\ell) + \operatorname{tr}\mathbb{B}_\ell - d,$$

it is clear that

$$\|\ln\det\mathbb{B}_\ell\|_{L^\infty L^1} \le C. \tag{3.186}$$

Further, the estimate of its derivative follows from a version of Jacobi's formula (see Lemma 3 below) and (3.149) as

$$\|\nabla\ln\det\mathbb{B}_\ell\|_{L^2L^2} = \|\operatorname{tr}(\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}})\|_{L^2L^2} \le C. \tag{3.187}$$

Hence, using again the Sobolev, the Poincaré and interpolation inequalities, we get

$$\|\ln\det\mathbb{B}_\ell\|_{L^{2+\frac{2}{d}}L^{2+\frac{2}{d}}} \le C. \tag{3.188}$$

From (3.185), (3.188), (3.152) and (3.128), we deduce

$$\|\eta_\ell\|_{L^{s_4}L^{s_4}}, \quad \text{where} \quad s_4 = \min\{2+\tfrac{2}{d}, q+\sigma\} > 2, \tag{3.189}$$

and thus

$$\|\boldsymbol{v}_\ell\eta_\ell\|_{L^{s_5}L^{s_5}} \le C, \quad \text{where} \quad s_5 = \Big(\frac{d}{p(d+2)} + \frac{1}{s_4}\Big)^{-1} > 1. \tag{3.190}$$

due to (3.174) and (3.189). We remark that, since

$$\nabla\eta_\ell = c_v\nabla\ln\theta_\ell - \mu(\operatorname{tr}\nabla\mathbb{B}_\ell - \operatorname{tr}(\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}})),$$

we also have, using (3.184), (3.172), (3.149) and Poincaré's inequality that

$$\|\eta_\ell\|_{L^zW^{1,z}} \le C. \tag{3.191}$$

Looking at (3.139), we still need to verify that the flux terms are controlled. For the term $\kappa(\theta_\ell)\nabla\ln\theta_\ell$, we use Hölder's inequality, (3.3), (3.168) and (3.149) to get

$$\|\kappa(\theta_\ell)\nabla\ln\theta_\ell\|_{\frac{2R_d}{r+R_d};Q} \le \|\sqrt{\kappa(\theta_\ell)}\|_{\frac{2R_d}{r};Q}\|\sqrt{\kappa(\theta_\ell)}\nabla\ln\theta_\ell\|_{2;Q} \le C, \tag{3.192}$$

where note that $R_d > r_d$ due to (C_1). Let us derive an estimate on $\omega|\nabla\theta_\ell|^r\nabla\ln\theta_\ell$, from which it follows that this term vanishes as $\omega\to 0_+$. The number

$$y \coloneqq \frac{R_d(r+2)}{R_d(r+1)+r} = 1 + \frac{R_d - r}{R_d(r+1)+r}$$

is greater that one due to $R_d > r_d > r$ and satisfies

$$\frac{ry}{r+2-y(r+1)} = R_d.$$

We also remark that (3.149) yields

$$\omega \int_Q \frac{|\nabla \theta_\ell|^{r+2}}{\theta_\ell^2} \le C. \tag{3.193}$$

Using this information together with (3.168) and Hölder's inequality leads to

$$\begin{aligned}
\|\omega |\nabla \theta_\ell|^r \nabla \ln \theta_\ell\|_{y;Q} &= \omega \left(\int_Q \frac{|\nabla \theta_\ell|^{(r+1)y}}{\theta_\ell^{2\frac{r+1}{r+2}y}} \theta_\ell^{\frac{r}{r+2}y} \right)^{\frac{1}{y}} \\
&\le \omega \left\| \frac{|\nabla \theta_\ell|^{(r+1)y}}{\theta_\ell^{2\frac{r+1}{r+2}y}} \right\|^{\frac{1}{y}}_{\frac{r+2}{r+1}\frac{1}{y};Q} \left\| \theta_\ell^{\frac{r}{r+2}y} \right\|^{\frac{1}{y}}_{\frac{r+2}{r+2-y(r+1)};Q} \\
&\le \omega^{\frac{1}{r+2}} \left(\omega \int_Q \frac{|\nabla \theta_\ell|^{r+2}}{\theta_\ell^2} \right)^{\frac{r+1}{r+2}} \|\theta_\ell\|^{\frac{r}{r+2}}_{R_d;Q} \le C\omega^{\frac{1}{r+2}}.
\end{aligned} \tag{3.194}$$

From this and from (3.192), (3.187), (3.149), (3.139), we see, using the definition of a weak time derivative, that

$$\|\partial_t \eta_\ell\|_{L^1 W^{-M,2}} \le C, \tag{3.195}$$

where M is so large that $W^{M,2}(\Omega;\mathbb{R}) \hookrightarrow W^{1,\infty}(\Omega;\mathbb{R})$.

The final limits $\omega \to 0$, $\ell \to \infty$. Let us note that the estimates above are independent not only of ℓ, but also of ω (except for (3.193), which is used only to infer (3.194)). Hence, to spare us some work, we set $\omega = \ell^{-1}$ and hereby, it remains to take the limit $\ell \to \infty$ only.

By collecting the estimates (3.142), (3.169), (3.176), (3.171), (3.172), (3.182), (3.189), (3.191), (3.195) and using the Aubin-Lions lemma and Vitali's convergence theorem, we get the following results:

$$\boldsymbol{v}_\ell \rightharpoonup \boldsymbol{v} \qquad \text{weakly in } L^p(0,T;W^{1,p}_{\boldsymbol{n},\mathrm{div}}), \tag{3.196}$$

$$\boldsymbol{v}_\ell \to \boldsymbol{v} \qquad \text{strongly in } L^{p\frac{d+2}{d})}(Q;\mathbb{R}^d) \text{ and a.e. in } Q, \tag{3.197}$$

$$\partial_t \boldsymbol{v}_\ell \rightharpoonup \partial_t \boldsymbol{v} \qquad \text{weakly in } L^{p\frac{d+2}{2d}}(0,T;W^{-1,p\frac{d+2}{2d}}_{\boldsymbol{n},\mathrm{div}}), \tag{3.198}$$

$$\mathbb{B}_\ell \rightharpoonup \mathbb{B} \qquad \text{weakly in } L^z(0,T;W^{1,z}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{3.199}$$

$$\mathbb{B}_\ell \to \mathbb{B} \qquad \text{strongly in } L^{q+\sigma)}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \text{ and a.e. in } Q, \tag{3.200}$$

$$\partial_t \mathbb{B}_\ell \rightharpoonup \partial_t \mathbb{B} \qquad \text{weakly in } L^{s_2}(0,T;W^{-1,s_2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{3.201}$$

$$\eta_\ell \rightharpoonup \eta \qquad \text{weakly in } L^z(0,T;W^{1,z}(\Omega;\mathbb{R})), \tag{3.202}$$

$$\eta_\ell \to \eta \qquad \text{strongly in } L^{s_4)}(Q;\mathbb{R}) \text{ and a.e. in } Q, \tag{3.203}$$

$$\theta_\ell \rightharpoonup \theta \qquad \text{weakly in } L^{R_d}(Q;\mathbb{R}). \tag{3.204}$$

Now we explain how to take the limit in equations (3.95), (3.96), (3.139), (3.141) and then, we also identify the corresponding initial conditions. First, we focus on taking the limit in the function $g_{\frac{1}{\ell}}$. From (3.120), (3.123) and (3.196), (3.199) (or (3.197), (3.200)), we obtain

$$\mathbb{B}\boldsymbol{x} \cdot \boldsymbol{x} \ge 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d \quad \text{and} \quad \theta \ge 0 \quad \text{a.e. in } Q, \tag{3.205}$$

however, we need these properties with strict inequalities. To this end, we use the Fatou lemma, (3.200) and (3.186) to get

$$\int_\Omega |\ln\det\mathbb{B}| \le \liminf_{\ell\to\infty}\int_\Omega |\ln\det\mathbb{B}_\ell| \le C \quad \text{a.e. in } (0,T).$$

Thus, by taking the essential supremum over $(0,T)$, we obtain

$$\|\ln\det\mathbb{B}\|_{L^\infty L^1} < \infty, \tag{3.206}$$

which, together with (3.205) implies

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x}\in\mathbb{R}^d \quad \text{a.e. in } Q. \tag{3.207}$$

Then, note that, by (3.128), we have

$$c_v\ln\theta_\ell = \eta_\ell + \psi_e(\mathbb{B}_\ell) \tag{3.208}$$

and, by (3.203), (3.200), the right hand side of (3.208) converges a.e. in Q. Therefore, we also have

$$c_v\ln\theta_\ell \to \eta + \psi_e(\mathbb{B}) \quad \text{a.e. in } Q \tag{3.209}$$

with the limit being finite a.e. in Q thanks to (3.207). This we can rewrite as

$$\theta_\ell \to \exp\left(\frac{1}{c_v}\big(\eta+\psi_e(\mathbb{B})\big)\right) \quad \text{a.e. in } Q,$$

where the limit is positive and finite a.e. in Q. But looking at (3.204), this yields

$$\theta = \exp\left(\frac{1}{c_v}\big(\eta+\psi_e(\mathbb{B})\big)\right) > 0 \quad \text{a.e. in } Q, \tag{3.210}$$

which is (1.5), and

$$\theta_\ell \to \theta \quad \text{a.e. in } Q. \tag{3.211}$$

From (3.207), (3.210) and the point-wise convergence (3.199), (3.211) we deduce that, at almost every point $(t,x)\in Q$, we can find $M_{t,x}\in\mathbb{N}$ such that for all $\ell > M_{t,x}$ we have

$$\Lambda(\mathbb{B}_\ell(t,x)) > \frac{1}{2}\Lambda(\mathbb{B}(t,x)) > \frac{1}{\ell} \quad \text{and} \quad \theta_\ell(t,x) > \frac{1}{2}\theta(t,x) > \frac{1}{\ell}.$$

Then, looking at the definition of g_λ, we see that at almost every point $(t,x)\in Q$ and for $\ell > M_{t,x}$, the positive parts $\max\{0,\cdot\}$ can be removed and thus, it is clear that $g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)$ converges point-wise a.e. in Q to 1. Hence, the Vitali theorem and $0\le g_{\frac{1}{\ell}} < 1$, imply

$$g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell) \to 1 \quad \text{strongly in } L^{\infty)}(Q;\mathbb{R}). \tag{3.212}$$

Therefore, regarding the first two equations (3.95) and (3.96), we can take the limit in the same way as we did in the limit $n\to\infty$. Indeed, the integrability of the resulting non-linear limits was already verified when estimating $\partial_t\boldsymbol{v}_\ell$ and $\partial_t\mathbb{B}_\ell$ ((3.173)–(3.180)). This way, taking (3.212) into account, using the density of $\operatorname{span}\{\boldsymbol{w}_i\}_{i=1}^\infty$ in $W^{1,(p\frac{d+2}{2d})'}_{\boldsymbol{n},\mathrm{div}}$, integrating by parts in the convective term of (3.96) and extending the functional $\partial_t\mathbb{B}$ to the space stated in (3.37) using (3.182) and

$$z' = \frac{2(q+\sigma)}{q+\sigma-2} \quad \text{and} \quad s_2' = \frac{q+\sigma}{\sigma-1},$$

we obtain precisely (3.44) and (3.45).

Next, we show how to take the limit in (3.139). We need to prove that

$$\begin{aligned} \eta_0^{\frac{1}{\ell}} &= c_v \ln \theta_0^{\frac{1}{\ell}} - \mu(\operatorname{tr} \mathbb{B}_0^{\frac{1}{\ell}} - d - \ln \det \mathbb{B}_0^{\frac{1}{\ell}}) \\ &\to c_v \ln \theta_0 - \mu(\operatorname{tr} \mathbb{B}_0 - d - \ln \det \mathbb{B}_0) = \eta_0, \quad \ell \to \infty, \end{aligned} \tag{3.213}$$

in $L^1(\Omega)$, at least weakly. Using (3.60) and (3.59), we estimate

$$|\eta_0^\omega| \le c_v |\ln \theta_0^\omega| + \mu(|\operatorname{tr} \mathbb{B}_0^\omega| + d + |\ln \det \mathbb{B}_0^\omega|) \le C(|\ln \theta_0| + |\mathbb{B}_0| + |\ln \det \mathbb{B}_0| + 1),$$

where the right hand side is integrable by assumptions (3.29) and (3.30). Moreover, the function $\eta_0^{1/\ell}$ converges point-wise a.e. in Ω due to (3.61) and (3.62). Thus, the limit (3.213) indeed holds (even strongly in $L^1(\Omega)$) by the dominated convergence theorem. In order to take the limit in the convective term, we use (3.197), (3.203) and (3.190). Next, in order to identify the objects $\nabla \ln \theta$ and $\nabla \ln \det \mathbb{B}$, note first that (3.185), (3.188) with (3.211), (3.200) yield

$$\begin{aligned} \ln \theta_\ell &\rightharpoonup \ln \theta && \text{weakly in } L^{2+\frac{2}{d}}(Q;\mathbb{R}), \\ \ln \det \mathbb{B}_\ell &\rightharpoonup \ln \det \mathbb{B} && \text{weakly in } L^{2+\frac{2}{d}}(Q;\mathbb{R}). \end{aligned}$$

This, together with (3.187) and (3.184), implies

$$\nabla \ln \theta_\ell \rightharpoonup \nabla \ln \theta \qquad \text{weakly in } L^2(Q;\mathbb{R}^d), \tag{3.214}$$

$$\nabla \ln \det \mathbb{B}_\ell \rightharpoonup \nabla \ln \det \mathbb{B} \qquad \text{weakly in } L^2(Q;\mathbb{R}^d). \tag{3.215}$$

Then, for the term $\kappa(\theta_\ell)\nabla \ln \theta_\ell$, we use (3.1), (3.3), (3.168) and Vitali's theorem to find that

$$\sqrt{\kappa(\theta_\ell)} \rightharpoonup \sqrt{\kappa(\theta)} \quad \text{strongly in } L^{\frac{2R_d}{r})}(Q;\mathbb{R}), \tag{3.216}$$

where we recall that $R_d > r_d$. As an immediate consequence of this and (3.214), we get

$$\sqrt{\kappa(\theta_\ell)}\nabla \ln \theta_\ell \rightharpoonup \sqrt{\kappa(\theta)}\nabla \ln \theta \quad \text{weakly in } L^1(Q;\mathbb{R}^d). \tag{3.217}$$

However, this weak convergence is true (up to a subsequence) also in $L^2(Q;\mathbb{R}^d)$ due to (3.149). Therefore, using again (3.216), we obtain

$$\kappa(\theta_\ell)\nabla \ln \theta_\ell \rightharpoonup \kappa(\theta)\nabla \ln \theta \quad \text{weakly in } L^1(Q;\mathbb{R}^d).$$

Next, the term containing $\omega|\nabla\theta_\ell|^r\nabla \ln \theta_\ell$ tends to zero by (3.194). Furthermore, in the term $\mu\lambda(\theta_\ell)\nabla \operatorname{tr} \mathbb{B}_\ell$, we use (3.1), (3.4), (3.211), Vitali's theorem and (3.199). Analogously, we take the limit in the term $\mu\lambda(\theta_\ell)\nabla \ln \det \mathbb{B}_\ell$, only we use (3.215) instead of (3.199).

Now we take the limit in the terms on the right hand side of (3.139). From (3.149), we deduce that there exists $K \in L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ such that

$$\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \rightharpoonup K \quad \text{weakly in } L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{3.218}$$

For $\varepsilon \in (0,1)$, let $h_\varepsilon : (0,\infty) \to [0,1]$ be a smooth function satisfying

$$h_\varepsilon(s) = \begin{cases} 1, & s > \varepsilon; \\ 0, & s < \frac{\varepsilon}{2}, \end{cases}$$

and define

$$f_{\varepsilon,\ell} \coloneqq h_\varepsilon(\theta_\ell)\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell, \qquad f_\varepsilon \coloneqq h_\varepsilon(\theta)\sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v}$$

For fixed $\varepsilon > 0$, the function

$$h_\varepsilon(\theta_\ell)\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}$$

is bounded independently of ℓ and converges point-wise due to (3.211) and (3.1). Thus, using the Vitali theorem and (3.196), we find

$$f_{\varepsilon,\ell} \underset{\ell\to\infty}{\rightharpoonup} f_\varepsilon \quad \text{weakly in } L^1(Q;\mathbb{R}^{d\times d}_{\text{sym}}). \tag{3.219}$$

Next note that, by Hölder's, Chebyshev's inequalities and (3.183) we have

$$\begin{aligned}\int_Q \left| f_{\varepsilon,\ell} - \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \right| &\le \int_{\{\theta_\ell<\varepsilon\}} \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}|\mathbb{D}\boldsymbol{v}_\ell| \le C|\{\theta_\ell<\varepsilon\}|^{\frac12} \\ &\le C\,|\{-\ln\theta_\ell > -\ln\varepsilon\}|^{\frac12} \le \frac{C}{\sqrt{-\ln\varepsilon}}\left(\int_Q |\ln\theta_\ell|\right)^{\frac12} \le \frac{C}{\sqrt{-\ln\varepsilon}}\end{aligned}$$

for all $\ell \in \mathbb{N}$. Hence, using (3.218), (3.219) and weak lower semi-continuity, we get

$$\int_Q |f_\varepsilon - K| \le \liminf_{\ell\to\infty}\int_Q \left| f_{\varepsilon,\ell} - \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \right| \le \frac{C}{\sqrt{-\ln\varepsilon}},$$

and thus, for $\varepsilon \to 0_+$, we obtain

$$f_\varepsilon \to K \quad \text{strongly in } L^1(Q;\mathbb{R}^{d\times d}_{\text{sym}}). \tag{3.220}$$

On the other hand, since $\theta > 0$ a.e. in Q by (3.210), it is clear that

$$f_\varepsilon \to \sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v} \quad \text{a.e. in } Q. \tag{3.221}$$

Therefore, from (3.220) and (3.221), we conclude

$$K = \sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v},$$

which, using (3.218) and weak lower semi-continuity of $\|\cdot\|_{L^2L^2}$, finally gives

$$\liminf_{\ell\to\infty}\int_Q \frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 \ge \int_Q \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2.$$

In the next term $\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2$, we can use the weak lower semi-continuity directly since we already proved that (3.217) is valid in $L^2(Q;\mathbb{R}^d)$. Moreover, the auxiliary term $\omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2$ is simply estimated from below by zero.

To take the limit in the term $P(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1})$, we use (3.211), (3.200) and apply Fatou's lemma (using (3.9)).

To handle the limit in the last term of (3.139), we use again the function h_ε, but this time, we define

$$F_{\varepsilon,\ell} = h_\varepsilon(\det\mathbb{B}_\ell)\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}, \qquad F_\varepsilon = h_\varepsilon(\det\mathbb{B})\sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}.$$

Let $G \in L^2(Q;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\text{sym}})$ such that

$$\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12} \rightharpoonup G \quad \text{weakly in } L^2(Q;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\text{sym}}). \tag{3.222}$$

For $\varepsilon > 0$ fixed, the function

$$\mathbb{H}_{\varepsilon,\ell} = \sqrt{h_\varepsilon(\det \mathbb{B}_\ell)\sqrt{\lambda(\theta_\ell)}}\mathbb{B}_\ell^{-\frac{1}{2}}$$

is bounded and converges point-wise a.e. in Q due to (3.200) and (3.211). Thus, using Vitali's theorem and (3.199), we get

$$F_{\varepsilon,\ell} = \mathbb{H}_{\varepsilon,\ell}\nabla\mathbb{B}_\ell\mathbb{H}_{\varepsilon,\ell} \underset{\ell\to\infty}{\rightharpoonup} F_\varepsilon \quad \text{weakly in } L^1(Q;\mathbb{R}^d \times \mathbb{R}^{d\times d}_{\text{sym}}). \tag{3.223}$$

Moreover, using Hölder's and Chebyshev's inequalities and (3.186), we find

$$\begin{aligned}\int_Q \left|F_{\varepsilon,\ell} - \sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\right| &\le \int\limits_{\{\det\mathbb{B}_\ell<\varepsilon\}} \sqrt{\lambda(\theta_\ell)}\left|\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\right| \\ &\le C|\{\det\mathbb{B}_\ell < \varepsilon\}|^{\frac{1}{2}} \le C\,|\{-\ln\det\mathbb{B}_\ell > -\ln\varepsilon\}|^{\frac{1}{2}} \\ &\le \frac{C}{\sqrt{-\ln\varepsilon}}\left(\int_Q |\ln\det\mathbb{B}_\ell|\right)^{\frac{1}{2}} \le \frac{C}{\sqrt{-\ln\varepsilon}}\end{aligned}$$

for every $\ell \in \mathbb{N}$. Therefore, from (3.222), (3.223) and weak lower semi-continuity, we deduce

$$\int_Q |F_\varepsilon - G| \le \liminf_{\ell\to\infty}\int_Q \left|F_{\varepsilon,\ell} - \sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\right| \le \frac{C}{\sqrt{-\ln\varepsilon}},$$

hence

$$F_\varepsilon \to G \quad \text{strongly in } L^1(Q;\mathbb{R}^d \times \mathbb{R}^{d\times d}_{\text{sym}}).$$

Since $\det\mathbb{B} > 0$ a.e. by (3.207), we also have that

$$F_\varepsilon \to \sqrt{\lambda(\theta)}\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}} \quad \text{a.e. in } Q.$$

Thus, we identified that

$$G = \sqrt{\lambda(\theta)}\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}$$

and, by (3.222) and weak lower semi-continuity, there holds

$$\liminf_{\ell\to\infty}\int_Q \lambda(\theta_\ell)\left|\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\right|^2 \ge \int_Q \lambda(\theta)\left|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}\right|^2.$$

Using the argumentation above to take the limit $\ell \to \infty$ in (3.139), we obtain (3.46).

Finally, to take the limit in (3.141), we first note, using (3.102) and (3.114), that it implies

$$-\int_0^T\int_\Omega E_\ell\partial_t\phi = \int_\Omega (\tfrac{1}{2}|P_\ell\boldsymbol{v}_0|^2 + c_v\theta_0^{\frac{1}{\ell}})\phi(0) + \int_0^T\int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}_\ell\phi \tag{3.224}$$

for all $\phi \in \mathcal{C}^1([0,T];\mathbb{R})$ with $\phi(T) = 0$. Then, recalling (3.197) with $p(1+\frac{2}{d}) > 2$ and (3.211), we see that $E_\ell = \frac{1}{2}|\boldsymbol{v}_\ell|^2 + c_v\theta_\ell$ converges strongly to E and thus, using also properties of P_ℓ and (3.62), we can take the limit in (3.224) to conclude

$$-\int_0^T\int_\Omega E\partial_t\phi = \int_\Omega E_0\phi(0) + \int_0^T\int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\phi, \tag{3.225}$$

where we set $E_0 := \frac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0$. In particular, by choosing an appropriate sequence of test functions ϕ, we obtain (3.47).

Attainment of initial conditions. To finish the proof of (I), it remains to identify the initial conditions and show that they are attained strongly. Note in particular, that we search for the initial condition for the temperature while now we only have the entropy inequality at our disposal. Let us start by an observation that $\boldsymbol{v}$ and $\mathbb{B}$ are weakly continuous in time. Indeed, first of all, we recall that

$$\begin{aligned}&\boldsymbol{v} \in L^\infty(0,T;L^2(\Omega;\mathbb{R}^d)), \quad \partial_t \boldsymbol{v} \in L^{p\frac{d+2}{2d}}(0,T;W^{-1,p\frac{d+2}{2d}}_{\boldsymbol{n},\mathrm{div}}(\Omega;\mathbb{R}^d)),\\ &\mathbb{B} \in L^\infty(0,T;L^\sigma(\Omega;\mathbb{R}^{d\times d}_{>0})), \quad \partial_t \mathbb{B} \in L^{\frac{q+\sigma}{q+1}}(0,T;W^{-1,\frac{q+\sigma}{q+1}}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})),\end{aligned} \tag{3.226}$$

cf. (3.171) and (3.182). From this and (3.226) we obtain, by a standard argument known from the theory of Navier-Stokes equations (see e.g. [33, Sect. 3.8.]), that

$$\boldsymbol{v} \in \mathcal{C}_w([0,T];L^2(\Omega;\mathbb{R}^d)) \quad \text{and} \quad \mathbb{B} \in \mathcal{C}_w([0,T];L^\sigma(\Omega;\mathbb{R}^d)). \tag{3.227}$$

Then, to identify the corresponding weak limits, we can use an analogous idea as in the part where the limit $n \to \infty$ was taken together with (3.61). This way, we obtain

$$\lim_{t\to 0_+} \int_\Omega \boldsymbol{v}(t)\cdot \boldsymbol{w} = \int_\Omega \boldsymbol{v}_0 \cdot \boldsymbol{w} \quad \text{for all } \boldsymbol{w} \in L^2(\Omega;\mathbb{R}^d) \tag{3.228}$$

and

$$\lim_{t\to 0_+} \int_\Omega \mathbb{B}(t)\cdot \mathbb{W} = \lim_{\ell\to\infty} \int_\Omega \mathbb{B}_0^{\frac{1}{\ell}} \cdot \mathbb{W} = \int_\Omega \mathbb{B}_0 \cdot \mathbb{W} \quad \text{for all } \mathbb{W} \in L^{\sigma'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{3.229}$$

Unlike in the theory of Navier-Stokes(-Fourier) systems, we can not draw information about $\limsup_{t\to 0_+} \|\boldsymbol{v}(t)\|_2^2$ from the (kinetic) energy estimate directly because of the presence of $\theta\mathbb{B}$ in (3.44).[9] Instead, we need first to combine the total energy and entropy balances to obtain the initial condition for θ. In (3.225) we choose a sequence of test functions ϕ approximating the function $\chi_{[0,t)}$, $t \in (0,T)$. This way, after taking the appropriate limit, we arrive at

$$\int_\Omega E(t) = \int_\Omega E_0 + \int_0^t \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v} \quad \text{for a.a. } t \in (0,T). \tag{3.230}$$

From this (3.29), (3.228) and weak lower semi-continuity, we get

$$\begin{aligned}\int_\Omega (\tfrac12 |\boldsymbol{v}_0|^2 + c_v\theta_0) &= \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega (\tfrac12|\boldsymbol{v}(t)|^2 + c_v\theta(t)) - \lim_{t\to 0_+}\int_0^t \|\boldsymbol{f}\|_2\|\boldsymbol{v}\|_2\\ &\ge \liminf_{t\to 0_+} \int_\Omega \tfrac12 |\boldsymbol{v}(t)|^2 + \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega c_v\theta(t)\\ &\ge \int_\Omega \tfrac12|\boldsymbol{v}_0|^2 + \operatorname*{ess\,lim\,sup}_{t\to 0_+}\int_\Omega c_v\theta(t),\end{aligned}$$

hence

$$\operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega \theta(t) \le \int_\Omega \theta_0. \tag{3.231}$$

To obtain also the corresponding lower estimate, we need to extract the available information from the entropy inequality (3.46). To this end, we localize (3.46) in

[9]For that, we would need $p > \frac{2d+2}{d+2}$. This condition is stronger than (C_2), but weaker than (C_2^E).

time, using a sequence of non-negative functions approximating $\chi_{[0,t)}$. This way, we eventually obtain

$$\int_\Omega \eta(t)\phi + \int_0^t \int_\Omega \boldsymbol{j}\cdot\nabla\phi \geq \int_\Omega \eta_0\phi + \int_0^t \int_\Omega \xi\phi \tag{3.232}$$

a.e. in $(0,T)$ and for all $\phi \in W^{1,\infty}(\Omega;\mathbb{R}_{\geq 0})$, where

$$\boldsymbol{j} := -\boldsymbol{v}\eta + \kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}) \in L^1(Q;\mathbb{R}^d).$$

Hence, by taking $\operatorname{ess\,lim\,inf}_{t\to 0_+}$ of (3.232), we deduce

$$\operatorname*{ess\,lim\,inf}_{t\to 0_+} \int_\Omega \eta(t)\phi \geq \int_\Omega \eta_0\phi,$$

which is (3.51).[10] Let us now fix $\varphi \in \mathcal{C}^1(\Omega;\mathbb{R}_{\geq 0})$ such that $\int_\Omega \varphi = 1$. Since ψ_e is convex, we get from (3.51) and (3.229) (or (3.49)) that

$$\begin{aligned}\int_\Omega c_v \ln\theta_0\varphi = \int_\Omega \eta_0\varphi + \int_\Omega \psi_e(\mathbb{B})\varphi &\leq \operatorname*{ess\,lim\,inf}_{t\to 0_+}\int_\Omega \eta(t)\varphi + \liminf_{t\to 0_+}\int_\Omega \psi_e(\mathbb{B}(t))\varphi \\ &\leq \operatorname*{ess\,lim\,inf}_{t\to 0_+}\int_\Omega c_v\ln\theta(t)\varphi.\end{aligned}$$

If we use this information together with Jensen's inequality and the fact that the function $s \mapsto \exp(\frac{s}{2})$, is increasing and convex in $\mathbb{R}$, we are led to

$$\begin{aligned}\exp\left(\frac{1}{2}\int_\Omega \ln\theta_0\varphi\right) &\leq \exp\left(\frac{1}{2}\operatorname*{ess\,lim\,inf}_{t\to 0_+}\int_\Omega \ln\theta(t)\varphi\right) \\ &= \operatorname*{ess\,lim\,inf}_{t\to 0_+}\exp\left(\int_\Omega \ln\sqrt{\theta(t)}\varphi\right) \leq \operatorname*{ess\,lim\,inf}_{t\to 0_+}\int_\Omega \sqrt{\theta(t)}\varphi.\end{aligned} \tag{3.233}$$

Since $\operatorname{ess\,sup}_{(0,T)}\|\sqrt{\theta}\|_2 < \infty$, there exists a function $h \in L^2(\Omega;\mathbb{R}_{\geq 0})$, such that

$$\operatorname*{ess\,lim}_{t\to 0_+}\int_\Omega \sqrt{\theta(t)}\psi = \int_\Omega h\psi \quad \text{for all } \psi \in L^2(\Omega;\mathbb{R}). \tag{3.234}$$

Using this in (3.233) gives

$$\exp\left(\frac{1}{2}\int_\Omega \ln\theta_0\varphi\right) \leq \int_\Omega h\varphi. \tag{3.235}$$

In every Lebesgue point $x_0 \in \Omega$ of both $\ln\theta_0$ and h, we can localize inequality (3.235) in Ω by choosing a sequence of functions φ that approximates the Dirac delta distribution at $x_0 \in \Omega$. Indeed, appealing to the Lebesgue differentiation theorem, we get this way that

$$\exp(\tfrac{1}{2}\ln\theta(x_0)) = \sqrt{\theta_0(x_0)} \leq h(x_0),$$

[10] Using a similar technique, we could prove that η is essentially lower semi-continuous on $[0,T]$ in the weak topology of measures.

and, consequently, also $\sqrt{\theta_0} \le h$ a.e. in Ω. From this, (3.234) and (3.231), we deduce that

$$\begin{aligned}
&\operatorname*{ess\,lim\,sup}_{t\to 0_+} \|\sqrt{\theta(t)} - \sqrt{\theta_0}\|_2^2 \\
&\quad \le \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega \theta(t) - 2\operatorname*{ess\,lim\,inf}_{t\to 0_+} \int_\Omega \sqrt{\theta(t)}\sqrt{\theta_0} + \int_\Omega \theta_0 \\
&\quad \le 2\int_\Omega \theta_0 - 2\int_\Omega h\sqrt{\theta_0} \le 0.
\end{aligned}$$

Hence, by Hölder's inequality, we get

$$\begin{aligned}
\operatorname*{ess\,lim\,sup}_{t\to 0_+} \|\theta(t) - \theta_0\|_1 &= \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega \left| \left(\sqrt{\theta(t)} + \sqrt{\theta_0}\right)\left(\sqrt{\theta(t)} - \sqrt{\theta_0}\right)\right| \\
&\le C \operatorname*{ess\,lim\,sup}_{t\to 0_+} \|\sqrt{\theta(t)} - \sqrt{\theta_0}\|_2 = 0,
\end{aligned}$$

which implies (3.50).

Using information above, we can now improve the initial condition for $\boldsymbol{v}$ as well. Indeed, from (3.230), (3.50) and (3.29), we obtain

$$\begin{aligned}
\operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega \tfrac12 |\boldsymbol{v}(t)|^2 &\le \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega E(t) - \operatorname*{ess\,lim\,inf}_{t\to 0_+} \int_\Omega c_v\theta(t) \\
&\le \int_\Omega E_0 + \lim_{t\to 0_+} \int_0^t (\boldsymbol{f},\boldsymbol{v}) - \int_\Omega c_v\theta_0 = \int_\Omega \tfrac12 |\boldsymbol{v}_0|^2.
\end{aligned}$$

Thus, using also (3.228), we conclude that

$$\operatorname*{ess\,lim\,sup}_{t\to 0_+} \|\boldsymbol{v}(t) - \boldsymbol{v}_0\|_2^2 = \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega |\boldsymbol{v}(t)|^2 + \int_\Omega |\boldsymbol{v}_0|^2 - 2\lim_{t\to 0_+} \int_\Omega \boldsymbol{v}(t)\cdot\boldsymbol{v}_0 \le 0,$$

which implies (3.48).

It remains to identify the initial condition for $\mathbb{B}$. To this end, we can test (3.96) by $(\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell$, where $\delta > 0$. That this is a valid test function can be verified using similar ideas as above (when testing (3.96) with $\mathbb{B}^{\sigma_0 - 1}$) and noting that

$$\nabla((\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell) = (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\big(\mathbb{I} + (\sigma_0 - 2)(\delta + |\mathbb{B}_\ell|^2)^{-1}\mathbb{B}_\ell \otimes \mathbb{B}_\ell\big)\nabla\mathbb{B}_\ell$$

Moreover, we have the following identities and estimates:

$$\begin{aligned}
\langle \partial_t \mathbb{B}_\ell, (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell\rangle &= \frac{1}{\sigma_0}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}}, \\
(\boldsymbol{v}_\ell \cdot \nabla\mathbb{B}_\ell, (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell) &= \frac{1}{\sigma_0}\int_\Omega \boldsymbol{v}_\ell \cdot \nabla(\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}} = 0, \\
P(\theta_\ell, \mathbb{B}_\ell) \cdot ((\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell) &\ge (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}(C_1|\mathbb{B}_\ell|^{q+2} - C_2),
\end{aligned}$$

and

$$\begin{aligned}
&\nabla\mathbb{B}_\ell \cdot \nabla((\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell) \\
&\quad = (\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\big(|\nabla\mathbb{B}_\ell|^2 + (\sigma_0 - 2)(\delta + |\mathbb{B}_\ell|^2)^{-1}|\mathbb{B}_\ell \cdot \nabla\mathbb{B}_\ell|^2\big) \\
&\quad \ge \min\{\sigma_0 - 1, 1\}(\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}|\nabla\mathbb{B}_\ell|^2 \ge 0.
\end{aligned}$$

Using these in (3.96) tested by $(\delta+|\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}\mathbb{B}_\ell$ and applying the Young's inequality several times together with the previously derived uniform estimate

$$\int_Q (|\mathbb{B}_\ell|^{q+\sigma} + |\mathbb{D}\boldsymbol{v}_\ell|^p) \leq C$$

and the inequality $\frac{q+\sigma_0}{q} < p$, we obtain

$$\begin{aligned}\frac{1}{\sigma_0}\int_\Omega ((\delta + |\mathbb{B}_\ell(t)|^2)^{\frac{\sigma_0}{2}} - (\delta + |\mathbb{B}_\ell(0)|^2)^{\frac{\sigma_0}{2}})&\\ \leq \int_0^t \int_\Omega \Big(C_2(\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1} + 2a(\delta + |\mathbb{B}_\ell|^2)^{\frac{\sigma_0}{2}-1}|\mathbb{B}_\ell|^2|\mathbb{D}\boldsymbol{v}_\ell|\Big)&\\ \leq C\int_0^t \int_\Omega (\delta^{\frac{\sigma_0}{2}-1} + |\mathbb{D}\boldsymbol{v}_\ell|^{\frac{q+\sigma_0}{q}} + 1)&\\ \leq C(1+\delta^{\frac{\sigma_0}{2}-1})t + Ct^\alpha.&\end{aligned}$$

for a.a. $t \in (0,T)$ and for a certain $\alpha > 0$. Here we use (3.200), (3.101) and (3.61) to take the limit $\ell \to \infty$ and arrive at

$$\int_\Omega ((\delta + |\mathbb{B}(t)|^2)^{\frac{\sigma_0}{2}} - (\delta + |\mathbb{B}_0|^2)^{\frac{\sigma_0}{2}}) \leq C(1+\delta^{\frac{\sigma_0}{2}-1})t + Ct^\alpha$$

for all $t \in [0,T]$ after appropriate redefinition of $\mathbb{B}(t)$ on a null set (which does not affect any of the properties of $\mathbb{B}$ proved so far). Taking the limit $t \to 0_+$ then yields

$$\limsup_{t\to 0_+}\int_\Omega |\mathbb{B}(t)|^{\sigma_0} \leq \limsup_{t\to 0_+}\int_\Omega (\delta + |\mathbb{B}(t)|^2)^{\frac{\sigma_0}{2}} \leq \int_\Omega (\delta + |\mathbb{B}_0|^2)^{\frac{\sigma_0}{2}}$$

and letting $\delta \to 0_+$ gives

$$\limsup_{t\to 0_+}\int_\Omega |\mathbb{B}(t)|^{\sigma_0} \leq \int_\Omega |\mathbb{B}_0|^{\sigma_0}.$$

On the other hand, due to (3.229) and weak lower semi-continuity, we have

$$\int_\Omega |\mathbb{B}_0|^{\sigma_0} \leq \liminf_{t\to 0_+}\int_\Omega |\mathbb{B}(t)|^{\sigma_0},$$

and thus $\|\mathbb{B}(t)\|_{\sigma_0} \to \|\mathbb{B}_0\|_{\sigma_0}$ as $t \to 0_+$. Since $\sigma_0 > 1$, the space $L^{\sigma_0}(\Omega;\mathbb{R}^{d\times d}_{\text{sym}})$ is uniformly convex and, consequently, property (3.49) follows.

(II)

To derive (3.52) (which is a weak version of (1.11)), we need to construct the pressure p and ensure that every term appearing (3.52) is integrable. To this end, we apply the conditions (C_1^E) and (C_2^E). Moreover, we need to be able to test the momentum equation with $\boldsymbol{v}\phi$, where ϕ is some smooth function on Q. Unfortunately, we can not do this operation in (3.44) nor at any stage of our approximation scheme. The remedy is to truncate the convection term in the balance of momentum. However, then we are just mimicking the existence proof that is done in [7] for a different non-linear fluid. Thus, let us only verify the weak compactness of weak solutions $(\boldsymbol{v}_\delta, \mathrm{p}_\delta, \mathbb{B}_\delta, \theta_\delta, \eta_\delta)$ to the system $\operatorname{div} \boldsymbol{v}_\delta = 0$, (3.45), (3.46),

$$\left.\begin{aligned} \int_0^T (\langle \partial_t \boldsymbol{v}_\delta, \boldsymbol{\varphi}\rangle - (T_\delta \boldsymbol{v}_\delta \otimes \boldsymbol{v}_\delta, \nabla \boldsymbol{\varphi}) + (2\nu(\theta_\delta)\mathbb{D}\boldsymbol{v}_\delta, \nabla\boldsymbol{\varphi})) + \alpha \int_0^T\int_{\partial\Omega} \boldsymbol{v}_\delta \cdot \boldsymbol{\varphi} \\ = \int_0^T (\mathrm{p}_\delta, \operatorname{div} \boldsymbol{\varphi}) - \int_0^T (2a\mu\theta_\delta \mathbb{B}_\delta, \nabla\boldsymbol{\varphi}) + \int_0^T (\boldsymbol{f}, \boldsymbol{\varphi}) \end{aligned}\right\} \quad (3.236)$$

for all $\boldsymbol{\varphi} \in L^\infty(0,T; W^{1,\infty}_{\boldsymbol{n}})$, with $T_\delta \boldsymbol{v}_\delta = ((\boldsymbol{v}_\delta s_\delta) * r_\delta)_{\mathrm{div}}$, where s_δ is a truncation near $\partial\Omega$, r_δ is a standard mollifier and $(\cdot)_{\mathrm{div}}$ is a Helmholtz projection onto divergence-free functions, and

$$\begin{aligned} &-(\tfrac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0, \phi)\varphi(0) - \int_0^T (\tfrac{1}{2}|\boldsymbol{v}_\delta|^2 + c_v\theta_\delta, \phi)\partial_t\varphi \\ &\quad - \int_0^T ((\tfrac{1}{2}|\boldsymbol{v}_\delta|^2 + c_v\theta_\delta)\boldsymbol{v}_\delta, \nabla\phi)\varphi + \alpha\int_0^T\int_{\partial\Omega} |\boldsymbol{v}_\delta|^2\phi\varphi + \int_0^T (\kappa(\theta_\delta)\nabla\theta_\delta, \nabla\phi)\varphi \\ &\quad = \int_0^T (\mathrm{p}_\delta \boldsymbol{v}_\delta - 2\nu(\theta_\delta)(\mathbb{D}\boldsymbol{v}_\delta)\boldsymbol{v}_\delta - 2a\mu\theta_\delta\mathbb{B}_\delta\boldsymbol{v}_\delta, \nabla\phi)\varphi \end{aligned} \quad (3.237)$$

for all $\varphi \in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T) = 0$, and every $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$. The existence of such solutions follows by combining the approximation scheme from part (I) together with the one in [7]. In view of the uniform estimates derived in part (I), we may suppose that the sequence $\{(\boldsymbol{v}_\delta, \mathrm{p}_\delta, \mathbb{B}_\delta, \theta_\delta, \eta_\delta)\}_{\delta>0}$ is uniformly bounded in the spaces depicted in (3.32)–(3.43) and that we have the same convergence results as in (3.196)–(3.204) and so forth (with ℓ replaced by δ). We may also suppose that, say $\mathrm{p}_\delta \in L^2(Q;\mathbb{R})$ with $\int_\Omega \mathrm{p}_\delta = 0$. Then, since we have $\nu(\theta_\delta)\mathbb{D}\boldsymbol{v}_\delta, \theta_\delta\mathbb{B}_\delta \in L^p(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and the convection term is truncated, equation (3.236) is valid for all $\boldsymbol{\varphi} \in L^{p'}(0,T;W^{1,p'}_{\boldsymbol{n}})$, in fact. What is missing is the uniform estimate of the pressure. By localizing (3.236) in time, choosing $\boldsymbol{\varphi} = \nabla u$ and using $\operatorname{div} \boldsymbol{v}_\delta = 0$, we obtain

$$-(\mathrm{p}_\delta, \Delta u) = (T_\delta\boldsymbol{v}_\delta \otimes \boldsymbol{v}_\delta - 2\nu(\theta_\delta)\mathbb{D}\boldsymbol{v}_\delta - 2a\mu\theta_\delta\mathbb{B}_\delta, \nabla\nabla u) - \alpha(\boldsymbol{v}_\delta, \nabla u)_{\partial\Omega} + (\boldsymbol{f}, \nabla u)$$

a.e. in $(0,T)$. There the convective term, if not truncated, is the most irregular one (recall that $\|\boldsymbol{v}_\delta\otimes\boldsymbol{v}_\delta\|_{p\frac{d+2}{2d};Q} \le C$). Thus, expecting p_δ to have the same integrability, we may choose $u \in W^{2,(p\frac{d+2}{2d})'}(\Omega;\mathbb{R})$ to be the solution to the Neumann problem

$$\begin{aligned} -\Delta u &= |\mathrm{p}_0|^{p\frac{d+2}{2d}-2}\mathrm{p}_0 - \frac{1}{|\Omega|}\int_\Omega |\mathrm{p}_0|^{p\frac{d+2}{2d}-2}\mathrm{p}_0 \quad \text{in } \Omega, \\ \nabla u \cdot \boldsymbol{n} &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

a.e. in $(0,T)$, where $\mathrm{p}_0 = \mathrm{p}_\delta - \frac{1}{|\Omega|}\int_\Omega \mathrm{p}_\delta$. Since $\|u\|_{2,(p\frac{d+2}{2d})'} \le C\|\mathrm{p}_0\|_{p\frac{d+2}{2d}}$ by the corresponding L^q-theory (here we used $\Omega \in \mathcal{C}^{1,1}$), the test function u eventually leads to

$$\|\mathrm{p}_\delta\|_{p\frac{d+2}{2d}} \le C,$$

see [7] for details.

Taking the limit $\delta \to 0_+$ in (3.236), (3.45) and (3.46) can be done analogously as in the part (I), where the limit $\ell \to \infty$ was taken. Indeed, in the additional term $\int_0^T (\mathrm{p}_\delta, \operatorname{div}\boldsymbol{\varphi})$, we simply use the fact that $\mathrm{p}_\delta \rightharpoonup \mathrm{p}$ weakly in $L^{p\frac{d+2}{2d}}(Q;\mathbb{R})$. It remains to take the limit $\delta \to 0_+$ in (3.237). Since $\boldsymbol{v}_\delta$ converges strongly in

$L^{p\frac{d+2}{d}}(Q;\mathbb{R}^d)$ and $p > \frac{3d}{d+2}$ due to (C_2^E), we deduce that the terms $p_\delta \boldsymbol{v}_\delta$ and $|\boldsymbol{v}_\delta|^2\boldsymbol{v}_\delta$ converge weakly to their limits. Clearly, as $p\frac{d+2}{2d} \le p$, the same is true for the terms $\nu(\theta_\delta)(\mathbb{D}\boldsymbol{v}_\delta)\boldsymbol{v}_\delta$, $\theta_\delta\mathbb{B}_\delta\boldsymbol{v}_\delta$ and $\theta_\delta\boldsymbol{v}_\delta$. To handle the term $\kappa(\theta_\delta)\nabla\theta_\delta$, we apply (3.3) to estimate it as

$$|\kappa(\theta_\delta)\nabla\theta_\delta| \le C|\nabla\theta_\delta| + C\theta_\delta^r|\nabla\theta_\delta| \tag{3.238}$$

and treat both terms separately. If $R < 2$, we can write

$$|\nabla\theta_\delta| = \tfrac{2}{R}\theta_\delta^{1-\frac{R}{2}}|\nabla\theta_\delta^{\frac{R}{2}}|$$

and observe that both factors are uniformly bounded in $L^2(Q;\mathbb{R})$ (as $1-\frac{R}{2} < \frac{R_d}{2}$ due to (C_1^E)). On the other hand, if $R \ge 2$, then also $R_d > 2$, and thus simply writing $|\nabla\theta_\delta| = \theta_\delta|\nabla\ln\theta_\delta|$ and using $\|\nabla\ln\theta_\delta\|_{2;Q} \le C$ (recall the entropy inequality) shows that $\|\nabla\theta_\delta\|_{1+\varepsilon;Q} \le C$ for some $\varepsilon > 0$. Regarding the other term, we rewrite it as

$$\theta_\delta^r|\nabla\theta_\delta| = \tfrac{2}{R}\theta_\delta^{r+1-\frac{R}{2}}|\nabla\theta_\delta^{\frac{R}{2}}|$$

and observe that the first factor is square integrable if

$$r + 1 - \tfrac{R}{2} > \tfrac{R_d}{2},$$

which turns out to be equivalent with (C_1^E). In total, we thus see that also the term $\kappa(\theta_\delta)\nabla\theta_\delta$ converges weakly to the corresponding limit. It remains to pass in the boundary term in (3.237) if $\alpha > 0$, for which we need to show that $\boldsymbol{v}_\delta \to \boldsymbol{v}$ strongly in $L^2(0,T;L^2(\partial\Omega;\mathbb{R}^d))$. This convergence result is indeed true provided that $p > \frac{2d+2}{d+2}$ and it can be proved using all the available estimates for $\boldsymbol{v}_\delta$ together with the Aubin-Lions lemma for Sobolev-Slobodeckij spaces, interpolation and Sobolev inequalities, see [10, Corollary 1.13.]. Since (C_2^E) yields $p > \frac{3d}{d+2} \ge \frac{2d+2}{d+2}$, the condition on p is met in our situation.

Using the above argumentation, we can pass to the limit in the δ-approximated system and obtain a weak solution satisfying (3.47).

(III)

In the proof of (I) it was shown that a weak solution $(\boldsymbol{v},\mathbb{B},\theta,\eta)$ can be constructed as a weak limit of the sequence $(\boldsymbol{v}_\ell,\mathbb{B}_\ell,\theta_\ell,\eta_\ell)$ satisfying, among other things, the equation

$$\begin{aligned} -(c_v\theta_0^{\frac{1}{\ell}},\phi)\varphi(0) &- \int_0^T(\theta_\ell,\phi)\partial_t\varphi - \int_0^T(c_v\theta_\ell\boldsymbol{v}_\ell,\nabla\phi)\varphi \\ &+ \int_0^T(\kappa(\theta_\ell)\nabla\theta_\ell + \frac{1}{\ell}|\nabla\theta_\ell|^r\nabla\theta_\ell,\nabla\phi)\varphi \\ &= \int_0^T(2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\phi)\varphi \end{aligned} \tag{3.239}$$

for all $\varphi \in W^{1,\infty}(0,T;\mathbb{R})$ and every $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$ (recall (3.117), take $\tau = \varphi\phi$ and integrate by parts in the first two terms). Moreover, as $p = 2$ and $R_d = r_d + 1$), we have the following information about the convergence of $(\boldsymbol{v}_\ell, \mathbb{B}_\ell, \theta_\ell)$ to $(\boldsymbol{v}, \mathbb{B}, \theta)$:

$$\boldsymbol{v}_\ell \rightharpoonup \boldsymbol{v} \qquad \text{weakly in } L^2(0,T; W^{1,2}_{\boldsymbol{n},\mathrm{div}}), \tag{3.240}$$

$$\boldsymbol{v}_\ell \to \boldsymbol{v} \qquad \text{strongly in } L^{2+\frac{4}{d})}(Q) \text{ and a.e. in } Q, \tag{3.241}$$

$$\mathbb{B}_\ell \to \mathbb{B} \qquad \text{strongly in } L^{q+\sigma)}(Q) \text{ and a.e. in } Q, \tag{3.242}$$

$$\theta_\ell \to \theta \qquad \text{strongly in } L^{R_d}(Q) \text{ and a.e. in } Q, \tag{3.243}$$

$$\nabla \theta_\ell^{\frac{R}{2}} \rightharpoonup \nabla \theta^{\frac{R}{2}} \qquad \text{weakly in } L^2(Q), \tag{3.244}$$

$$g_{\frac{1}{\ell}}(\mathbb{B}_\ell, \theta_\ell) \to 1 \qquad \text{strongly in } L^{\infty)}(Q). \tag{3.245}$$

(recall (3.196)–(3.204), (3.211), (3.168) and (3.212)). It remains to take the limit $\ell \to \infty$ in (3.239).

Taking the limit in the first three terms of (3.239) is easy. Next, we observe that

$$0 < r + 1 - \frac{R}{2} = r + 1 - \frac{r+1)}{2} < r + 1 - \frac{r + 1 - \frac{1}{d}}{2} = \frac{r_d + 1 - \frac{1}{d}}{2} < \frac{R_d}{2}$$

and, consequently, appealing to (3.243), (3.1), (3.3) and Vitali's theorem, we obtain

$$\kappa(\theta_\ell)\theta_\ell^{1-\frac{R}{2}} \to \kappa(\theta)\theta^{1-\frac{R}{2}} \quad \text{strongly in } L^2(Q).$$

Hence, using also (3.244) (and $\theta > 0$ a.e. in Q), we arrive at

$$\kappa(\theta_\ell)\nabla\theta_\ell = \frac{2}{R}\theta_\ell^{1-\frac{R}{2}}\kappa(\theta_\ell)\nabla\theta_\ell^{\frac{R}{2}} \rightharpoonup \frac{2}{R}\theta^{1-\frac{R}{2}}\kappa(\theta)\nabla\theta^{\frac{R}{2}} = \kappa(\theta)\nabla\theta$$

weakly in $L^1(Q)$.

To see that the term $\frac{1}{\ell}|\nabla\theta_\ell|^r\nabla\theta_\ell$ vanishes in the limit $\ell \to \infty$, we can use roughly the same argumentation as for the analogous term in the entropy inequality (recall (3.194)). First we need to realize that in the estimates that follow after (3.152), we can keep the term

$$\omega \int_Q |\nabla\theta_\ell|^r \nabla\theta_\ell \cdot \nabla\tau_\beta = \beta\omega \int_Q \frac{|\nabla\theta|^{r+2}}{\theta^{\beta+1}}$$

on the left hand side (this term was previously omitted on the first occasion since we were interested only in ω-uniform estimates). This way, we observe that the estimate (3.166) can be replaced by

$$\beta \int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right| + \beta\omega \int_Q \frac{|\nabla\theta_\ell|^{r+2}}{\theta_\ell^{\beta+1}} + \int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 \le C(\beta)$$

(recall that we are now in the case $r_d > r_1$). Then, using an analogous estimation as in (3.194), we arrive at

$$\begin{aligned} \|\omega|\nabla\theta_\ell|^r\nabla\theta_\ell\|_{\alpha;Q} &= \omega \left(\int_Q \frac{|\nabla\theta_\ell|^{(r+1)\alpha}}{\theta_\ell^{\frac{\alpha(1+\beta)(r+1)}{r+2}}} \theta_\ell^{\frac{\alpha(1+\beta)(r+1)}{r+2}} \right)^{\frac{1}{\alpha}} \\ &\le \omega^{\frac{1}{r+2}} \left(\omega \int_Q \frac{|\nabla\theta_\ell|^{r+2}}{\theta_\ell^{\beta+1}} \right)^{\frac{r+1}{r+2}} \|\theta_\ell\|^{\frac{(1+\beta)(r+1)}{r+2}}_{(\frac{r+2}{\alpha(r+1)})' \frac{\alpha(1+\beta)(r+1)}{r+2}} \\ &\le C(\beta)\omega^{\frac{1}{r+2}} \|\theta_\ell\|^{\frac{(1+\beta)(r+1)}{r+2}}_{R_d;Q} \end{aligned}$$

where α solves

$$\left(\frac{r+2}{\alpha(r+1)}\right)' \frac{\alpha(1+\beta)(r+1)}{r+2} = R_d.$$

Since this is equivalent to

$$\alpha = \frac{\frac{r+2}{r+1}}{1+\frac{1+\beta}{R_d}}$$

and we have $R_d = r_d + 1) > r+1$ we see that $\beta > 0$ can be chosen so small that $\alpha > 1$, and hereby we get

$$\|\omega|\nabla\theta_\ell|^r \nabla\theta_\ell\|_{\alpha;Q} \leq C\ell^{-\frac{1}{r+2}} \to 0 \quad \text{as} \quad \ell \to \infty.$$

Next, using (3.2), (3.243) and (3.240), it is easy to see that

$$\sqrt{2\nu(\theta_\ell)}\mathbb{D}\boldsymbol{v}_\ell \rightharpoonup \sqrt{2\nu(\theta)}\mathbb{D}\boldsymbol{v} \quad \text{weakly in } L^2(Q)$$

and then, by the weak lower semi-continuity, we get

$$\liminf_{\ell\to\infty} \int_Q 2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 \geq \int_Q 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2.$$

Finally, since

$$\frac{1}{R_d} + \frac{1}{q+\sigma} = \frac{1}{r_d+1)} + \frac{1}{q+\sigma} < \frac{1}{\frac{q+\sigma+2}{q+\sigma-2}+1} + \frac{1}{q+\sigma} = \frac{1}{2}$$

due to (C^θ), we deduce from (3.243) and (3.242) that

$$\theta_\ell\mathbb{B}_\ell \to \theta\mathbb{B} \quad \text{strongly in } L^{2+\varepsilon}(Q)$$

for some sufficiently small $\varepsilon > 0$. Hence, by (3.245), it is also true that

$$g(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \to \theta\mathbb{B} \quad \text{strongly in } L^2(Q).$$

But this together with (3.240) implies

$$g(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell \rightharpoonup \theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v} \quad \text{weakly in } L^1(Q).$$

Therefore, we can indeed take the limit $\ell \to \infty$ in every term of (3.239), leading to (3.53), and the proof of (III) and of Theorem 1 is finished. □

4. Auxiliary results

In this additional section, we prove those auxiliary results which were used above but are not completely standard in the existing literature. On the other hand, they are not new (except for parts of Lemma 3 that are taken from an upcoming work [3]) and serve only to clarify some arguments used in the proof.

For the purposes of this section, we replace the interval $(0, T)$ (or $[0, T]$) by an arbitrary bounded interval $I \subset \mathbb{R}$ and set $Q = I \times \Omega$. The set Ω is always assumed to be a bounded Lipschitz domain in $\mathbb{R}^d$, $d \in \mathbb{N}$.

Intersections of Sobolev-Bochner spaces. If $X \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} X^*$ is a Gelfand triple, it is well known that

$$\mathcal{C}^1(I;X) \overset{\text{dense}}{\hookrightarrow} \mathcal{W}^p_X \hookrightarrow \mathcal{C}(I;H), \tag{4.1}$$

where

$$\mathcal{W}^p_X := \big(\{u \in L^p(I;X);\ \partial_t u \in (L^p(I;X))^*\}, \|\cdot\|_{L^pX} + \|\partial_t\cdot\|_{L^{p'}X^*}\big), \quad 1 < p < \infty.$$

The first embedding in (4.1) is useful to manipulate certain duality pairings involving time derivatives, while the second embedding is important for the identification of boundary values (i.e. initial conditions) and the corresponding integration by parts formulas. We would like to generalize (4.1) for the space

$$\mathcal{W}^{p,q}_{X,Y} := \big(\{u \in L^p(I;X) \cap L^q(I;Y);\ \partial_t u \in (L^p(I;X) \cap L^q(I;Y))^*\},$$
$$\|\cdot\|_{L^pX\cap L^qY} + \|\partial_t\cdot\|_{(L^pX\cap L^qY)^*}\big), \quad 1 < p,q < \infty,$$

The primary application which we have in mind is the case where $X = W^{1,2}(\Omega)$, $Y = L^\omega(\Omega)$ and $\omega > \frac{2d}{d-2}$ (i.e., we know better integrability than what follows from the Sobolev embedding, recall the function $\mathbb{B}_\ell$). Thus, we may assume that both X and Y admit the Gelfand triplet structure with a common Hilbert space H (even though this could be relaxed if needed).

Lemma 1. *Let $1 < p,q < \infty$ and suppose that X, Y are separable reflexive Banach spaces and H is separable Hilbert space forming Gelfand triples in the sense that*

$$X \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} X^* \quad \textit{and} \quad Y \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} Y^*. \tag{4.2}$$

Then, we have the embeddings

$$\mathcal{C}^1(I;X\cap Y) \overset{\text{dense}}{\hookrightarrow} \mathcal{W}^{p,q}_{X,Y} \hookrightarrow \mathcal{C}(I;H). \tag{4.3}$$

Moreover, the integration by parts formula

$$(u(t_2),v(t_2))_H - (u(t_1),v(t_1))_H = \int_{t_1}^{t_2} \langle \partial_t u, v\rangle + \int_{t_1}^{t_2} \langle \partial_t v, u\rangle \tag{4.4}$$

holds for any $u,v \in \mathcal{W}^{p,q}_{X,Y}$ and any $t_1,t_2 \in I$.

Proof. The proof of the first embedding in (4.3) can be done in a standard way by extending u outside I evenly, taking the convolution with a smooth kernel and then estimating the difference from u and $\partial_t u$ in the respective norms. See [19] or [46] for details.

If $u,v \in \mathcal{C}^1(I;X\cap Y) \hookrightarrow \mathcal{C}(I;H)$, then $\partial_t u, \partial_t v \in \mathcal{C}(I;X\cap Y) \hookrightarrow \mathcal{C}(I;H)$ and, using density of the embeddings in (4.2), the duality in (4.4) can be represented as

$$\langle \partial_t u, v\rangle + \langle \partial_t v, u\rangle = (\partial_t u, v)_H + (\partial_t v, u)_H = \partial_t (u,v)_H \quad \text{a.e. in } I,$$

hence (4.4) is obvious in that case. Next, we can proceed as in [42, Lemma 7.3.] to prove that

$$\|u(t)\|_H \le C(\|u\|_{L^1H} + \|u\|_{\mathcal{W}^{p,q}_{X,Y}}) \tag{4.5}$$

for all $t \in I$ and every $u \in \mathcal{C}^1(I;X\cap Y)$. Moreover, by (4.2), we have

$$\mathcal{W}^{p,q}_{X,Y} \hookrightarrow L^p(I;X)\cap L^q(I;Y) \hookrightarrow L^1(I;X)\cap L^1(I;Y) \hookrightarrow L^1(I;X+Y) \hookrightarrow L^1(I;H),$$

and thus (4.5) yields

$$\|u\|_{\mathcal{C}(I;H)} \le C\|u\|_{\mathcal{W}^{p,q}_{X,Y}}. \tag{4.6}$$

Since $\mathcal{C}^1(I; X\cap Y)$ is dense in $\mathcal{W}^{p,q}_{X,Y}$, the estimate (4.6) and identity (4.4) remain valid for all $u\in\mathcal{W}^{p,q}_{X,Y}$. Moreover, if $u\in\mathcal{W}^{p,q}_{X,Y}$, then we can take $v=u$ and $t_2\to t_1$ in (4.4) to deduce that $u\in\mathcal{C}(I;H)$. Thus, the embedding $\mathcal{W}^{p,q}_{X,Y}\hookrightarrow\mathcal{C}(I;H)$ holds and the proof is finished. □

Since $\mathcal{W}^{p,p}_{X,X}=\mathcal{W}^{p}_{X}$, we obtain the classical result (4.1) as an obvious corollary.

Fundamental theorem of calculus in the Sobolev-Bochner setting. Let $H=L^2(\Omega)$. The formula (4.4) can be used to identify that

$$\langle \partial_t u, u\rangle = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u^2 \tag{4.7}$$

a.e. in I. However, in certain situations we would like to generalize (4.7) to

$$\langle \partial_t u, \psi(u)\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega\int_w^u \psi(s)\,\mathrm{d}s.$$

Whether this is possible depends on what kind of function ψ is and also on the choice of X. The next lemma characterizes one such situation.

Lemma 2. *Let $1<p,q<\infty$. Suppose that $\psi:\mathbb{R}\to\mathbb{R}$ is a Lipschitz function. For $w\in\mathbb{R}$, we define*

$$\Psi(x)=\int_w^x\psi(s)\,\mathrm{d}s,\quad x\in\mathbb{R}.$$

Then, for any $u\in\mathcal{W}^p_{W^{1,q}(\Omega)}$, there holds

$$\Psi(u)\in\mathcal{C}(I;L^1(\Omega)) \tag{4.8}$$

and

$$\int_{t_1}^{t_2}\langle\partial_t u,\psi(u)\rangle = \int_\Omega\Psi(u(t_2)) - \int_\Omega\Psi(u(t_1)) \quad \textit{for all } t_1,t_2\in I. \tag{4.9}$$

Moreover, if ψ is bounded, then

$$\Psi(u)\in\mathcal{C}(I;L^2(\Omega)).$$

Proof. First of all, we remark that $\psi(u)\in W^{1,q}(\Omega)$ a.e. in I, by a classical result (see e.g. [47, Theorem 2.1.11.]), and thus the duality in (4.9) is well defined. Next, we apply Theorem 1 to find $u_\varepsilon\in\mathcal{C}^1(I;W^{1,q}(\Omega))$ satisfying

$$\|u_\varepsilon-u\|_{L^pW^{1,q}}+\|\partial_t u_\varepsilon-\partial_t u\|_{L^{p'}W^{-1,q'}}\to 0\quad\text{as } \varepsilon\to 0_+. \tag{4.10}$$

Then, using the standard calculus, it is easy to see that the identity

$$\begin{aligned}\int_{t_1}^{t_2}\langle\partial_t u_\varepsilon,\psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2}\int_\Omega\psi(u_\varepsilon)\partial_t u_\varepsilon\\ &= \int_{t_1}^{t_2}\int_\Omega\partial_t\Psi(u_\varepsilon)=\int_\Omega\Psi(u_\varepsilon(t_2))-\int_\Omega\Psi(u_\varepsilon(t_1))\end{aligned} \tag{4.11}$$

holds true for any $t_1,t_2\in I$. Since ψ is Lipschitz with some Lipschitz constant $L\ge 0$, we can estimate

$$|\psi(u_\varepsilon)|\le|\psi(u_\varepsilon)-\psi(0)|+|\psi(0)|\le L|u_\varepsilon|+|\psi(0)|$$

and

$$|\nabla\psi(u_\varepsilon)|\le|\psi'(u_\varepsilon)||\nabla u_\varepsilon|\le L|\nabla u_\varepsilon|.$$

Hence, the sequence $\psi(u_\varepsilon)$ is bounded in $L^p(I;W^{1,q}(\Omega))$. As $1 < p,q < \infty$, this is a reflexive space, and thus, there exist a subsequence and its limit $\overline{\psi(u)} \in L^p(I;W^{1,q}(\Omega))$ such that

$$\psi(u_\varepsilon) \rightharpoonup \overline{\psi(u)} \quad \text{weakly in } L^p(I;W^{1,q}(\Omega)). \tag{4.12}$$

Since $p > 1$, a subsequence of u_ε converges point-wise a.e. in Q to u, and thus $\overline{\psi(u)} = \psi(u)$ using the continuity of ψ. Hence, by (4.10) and (4.12), we obtain

$$\begin{aligned}\int_{t_1}^{t_2} \langle \partial_t u_\varepsilon, \psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2} \langle \partial_t u_\varepsilon - \partial_t u, \psi(u_\varepsilon)\rangle + \int_{t_1}^{t_2} \langle \partial_t u, \psi(u_\varepsilon)\rangle \\ &\to \int_{t_1}^{t_2} \langle \partial_t u, \psi(u)\rangle\end{aligned} \tag{4.13}$$

as $\varepsilon \to 0_+$. Next, using the embedding $\mathcal{W}^p_{W^{1,q}(\Omega)} \hookrightarrow \mathcal{C}(I;L^2(\Omega))$ and (4.10), we get, for any $t_0 \in I$, that

$$\|u(t) - u(t_0)\|_2 \to 0 \quad \text{as} \quad t \to t_0 \tag{4.14}$$

and

$$\|u_\varepsilon(t_0) - u(t_0)\|_2 \to 0 \quad \text{as} \quad \varepsilon \to 0_+. \tag{4.15}$$

Then, the Lipschitz continuity of ψ, Hölder's inequality and (4.14) yield

$$\begin{aligned}\int_\Omega |\Psi(u(t)) - \Psi(u(t_0))| &= \int_\Omega \left| \int_{u(t_0)}^{u(t)} \psi(s)\,\mathrm{d}s \right| \le \int_\Omega \int_{u(t_0)}^{u(t)} (|\psi(0)| + L|s|) \\ &\le \int_\Omega \int_{u(t_0)}^{u(t)} C(1 + |u(t_0)| + |u(t)|) \le C \int_\Omega (1 + |u(t_0)| + |u(t)|)|u(t) - u(t_0)| \\ &\le C\|1 + |u(t_0)| + |u(t)|\|_2 \|u(t) - u(t_0)\|_2 \le C\|u(t) - u(t_0)\|_2 \to 0\end{aligned} \tag{4.16}$$

as $\varepsilon \to 0_+$, which proves (4.8) (and thus, the values $\Phi(u(t))$, $t \in I$, are well defined). By an analogous estimate, using (4.15) instead of (4.14), we can prove that

$$\int_\Omega |\Phi(u_\varepsilon(t_0)) - \Phi(u(t_0))| \to 0 \quad \text{as } \varepsilon \to 0_+$$

for any $t \in I$. This and (4.13) used in (4.11) to take the limit $\varepsilon \to 0_+$ proves (4.9).

If ψ is bounded, we replace (4.16) by

$$\int_\Omega |\Psi(u(t)) - \Psi(u(t_0))|^2 = \int_\Omega \left| \int_{u(t_0)}^{u(t)} \psi(s)\,\mathrm{d}s \right|^2 \le C \int_\Omega |u(t) - u(t_0)|^2$$

and the rest of the proof remains the same. □

Clearly, we can also replace ψ by $\psi\phi$, where $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$, leading to

$$\int_0^t \langle \partial_t u, \psi(u)\phi\rangle = \int_\Omega \int_w^{u(t)} \psi(s)\,\mathrm{d}s\,\phi - \int_\Omega \int_w^{u(0)} \psi(s)\,\mathrm{d}s\,\phi \quad \text{for all } t \in I. \tag{4.17}$$

Then, since ϕ is a Lipschitz (time independent) function, the proof is basically the same as the one presented above.

Calculus for positive definite matrices. We recall that the operations "$\cdot$" and $|\cdot|$ on matrices are defined by

$$\mathbb{A}_1 \cdot \mathbb{A}_2 = \sum_{i=1}^{d} \sum_{j=1}^{d} (\mathbb{A}_1)_{ij} (\mathbb{A}_2)_{ij} \quad \text{and} \quad |\mathbb{A}| = \sqrt{\mathbb{A} \cdot \mathbb{A}},$$

respectively. Then, the object $|\mathbb{A}|$ coincides, in fact, with the Frobenius matrix norm of $\mathbb{A}$.

The next lemma is formulated for a function $\mathbb{A} : Q \to \mathbb{R}^{d\times d}_{>0}$ and for simplicity, we shall assume that $\mathbb{A}$ is continuously differentiable with respect to all variables, i.e., $\mathbb{A} \in \mathcal{C}^1(Q; \mathbb{R}^{d\times d}_{>0})$. In particular situations, this assumption can be of course removed by an appropriate approximation (convolution smoothing) and the assertions of the following lemma hereby extend to the setting of weakly differentiable functions. Let us also denote any of the space-time derivatives by a generic symbol ∂.

Lemma 3. *Let* $\mathbb{A} \in \mathcal{C}^1(Q; \mathbb{R}^{d\times d}_{>0})$. *Then*

(i) $$0 \le \operatorname{tr} \mathbb{A} - d - \ln \det \mathbb{A}, \tag{4.18}$$

(ii) $$|\mathbb{A}| \le \operatorname{tr} \mathbb{A} \le \sqrt{d} |\mathbb{A}|, \tag{4.19}$$

(iii) $$\min\{1, d^{\frac{1-\alpha}{2}}\} |\mathbb{A}|^\alpha \le |\mathbb{A}^\alpha| \le \max\{1, d^{\frac{1-\alpha}{2}}\} |\mathbb{A}|^\alpha \textit{ for any } \alpha \ge 0, \tag{4.20}$$

(iv) $$\partial \mathbb{A} \cdot \mathbb{A}^\alpha = \begin{cases} \frac{1}{\alpha+1} \partial \operatorname{tr} \mathbb{A}^{\alpha+1} & \textit{if } \alpha \neq -1; \\ \partial \ln \det \mathbb{A} = \partial \operatorname{tr} \log \mathbb{A} & \textit{if } \alpha = -1, \end{cases} \tag{4.21}$$

(v) $$(\operatorname{sign} \alpha) \partial \mathbb{A} \cdot \partial \mathbb{A}^\alpha \ge \begin{cases} \frac{4|\alpha|}{(\alpha+1)^2} |\partial \mathbb{A}^{\frac{\alpha+1}{2}}|^2 & \textit{if } \alpha \neq -1; \\ |\partial \log \mathbb{A}|^2 & \textit{if } \alpha = -1. \end{cases} \tag{4.22}$$

Proof. Property (i) follows by passing to the spectral decomposition of $\mathbb{A}$ and from the fact that $x \mapsto x - 1 - \ln x$ attains its minimum at $x = 1$. Estimate (ii) is a consequence of the Cauchy-Schwarz inequality since

$$|\mathbb{A}| = |(\mathbb{A}^{\frac{1}{2}})^T \mathbb{A}^{\frac{1}{2}}| \le |\mathbb{A}^{\frac{1}{2}}|^2 = \operatorname{tr} \mathbb{A} = \mathbb{I} \cdot \mathbb{A} \le |\mathbb{I}||\mathbb{A}| = \sqrt{d}|\mathbb{A}|.$$

For (iii), we refer to [3, Theorem 4] and for (iv), (v) to [3, Theorem 1]. The relation (iv) with $\alpha = -1$ is also known as the Jacobi identity. □

References

[1] N. Arada and A. Sequeira, *Strong Steady Solutions for a Generalized Oldroyd-B Model with Shear-Dependent Viscosity in a Bounded Domain*, Mathematical Models and Methods in Applied Sciences, 13 (2003), pp. 1303–1323.

[2] J. W. Barrett and S. Boyaval, *Existence and approximation of a (regularized) Oldroyd-B model*, Mathematical Models and Methods in Applied Sciences, 21 (2011), pp. 1783–1837.

[3] M. Bathory, *Optimal inequalities in multiplication of derivatives of positive definite matrices and their powers*. (to appear) preprint arXiv:2007.15052.

[4] M. Bathory, M. Bulíček, and J. Málek, *Large data existence theory for three-dimensional unsteady flows of rate-type viscoelastic fluids with stress diffusion*, Advances in Nonlinear Analysis, 10 (2020), pp. 501–521.

[5] M. Bulíček, J. Málek, V. Průša, and E. Süli, *On incompressible heat-conducting viscoelastic rate-type fluids with stress-diffusion and purely spherical elastic response*. (to appear) preprint arXiv:2007.05749.

[6] M. Bulíček, E. Feireisl, and J. Málek, *On a class of compressible viscoelastic rate-type fluids with stress-diffusion*, Nonlinearity, 32 (2019), pp. 4665–4681.

[7] M. Bulíček, J. Málek, and K. R. Rajagopal, *Mathematical analysis of unsteady flows of fluids with pressure, shear-rate, and temperature dependent material moduli that slip at solid boundaries*, SIAM Journal on Mathematical Analysis, 41 (2009), pp. 665–707.

[8] M. Bulíček and J. Havrda, *On existence of weak solution to a model describing incompressible mixtures with thermal diffusion cross effects*, ZAMM Z. Angew. Math. Mech., 95 (2015), pp. 589–619.

[9] M. Bulíček, J. Málek, V. Průša, and E. Süli, *PDE analysis of a class of thermodynamically compatible viscoelastic rate-type fluids with stress-diffusion*, in Mathematical analysis in fluid mechanics—selected recent results, vol. 710 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2018, pp. 25–51.

[10] M. Bulíček, J. Málek, and K. R. Rajagopal, *Navier's slip and evolutionary Navier-Stokes-like systems with pressure and shear-rate dependent viscosity*, Indiana University Mathematics Journal, 56 (2007), pp. 51–85.

[11] L. Chupin, *Global Strong Solutions for Some Differential Viscoelastic Models*, SIAM Journal on Applied Mathematics, 78 (2018), pp. 2919–2949.

[12] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[13] P. Constantin and M. Kliegl, *Note on Global Regularity for Two-Dimensional Oldroyd-B Fluids with Diffusive Stress*, Archive for Rational Mechanics and Analysis, 206 (2012), pp. 725–740.

[14] M. Dostalík, V. Průša, and T. Skřivan, *On diffusive variants of some classical viscoelastic rate-type models*, AIP Conference Proceedings, 2107 (2019).

[15] A. W. El-Kareh and L. G. Leal, *Existence of solutions for all Deborah numbers for a non-Newtonian model modified to include diffusion*, Journal of Non-Newtonian Fluid Mechanics, 33 (1989), pp. 257–287.

[16] E. Feireisl, *Relative entropies in thermodynamics of complete fluid systems*, Discrete & Continuous Dynamical Systems - A, 32 (2012), pp. 3059–3080.

[17] E. Feireisl, M. Frémond, E. Rocca, and G. Schimperna, *A new approach to non-isothermal models for nematic liquid crystals*, Archive for Rational Mechanics and Analysis, 205 (2012), pp. 651–672.

[18] E. Feireisl, A. Novotný, and Y. Sun, *On the motion of viscous, compressible, and heat-conducting liquids*, Journal of Mathematical Physics, 57 (2016), p. 083101.

[19] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974. Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.

[20] R. J. Gordon and W. R. Schowalter, *Anisotropic fluid theory: A different approach to the dumbbell theory of dilute polymer solutions*, Transactions of the Society of Rheology, 16 (1972), pp. 79–97.

[21] C. Guillopé and J.-C. Saut, *Existence results for the flow of viscoelastic fluids with a differential constitutive law*, Nonlinear Analysis. Theory, Methods & Applications. An International Multidisciplinary Journal, 15 (1990), pp. 849–869.

[22] J. Hron, V. Miloš, V. Průša, O. Souček, and K. Tůma, *On thermodynamics of incompressible viscoelastic rate type fluids with temperature dependent material coefficients*, International Journal of Non-Linear Mechanics, 95 (2017), pp. 193–208.

[23] M. Johnson and D. Segalman, *A model for viscoelastic fluid behavior which allows non-affine deformation*, Journal of Non-Newtonian Fluid Mechanics, 2 (1977), pp. 255–270.

[24] O. Kreml, *Mathematical Analysis of Models for Viscoelastic Fluids*, ph.D. thesis, Charles University, 2010.

[25] O. Kreml, M. Pokorný, and P. Šalom, *On the global existence for a regularized model of viscoelastic non-Newtonian fluid*, Colloquium Mathematicum, 139 (2015), pp. 149–163.

[26] P. L. Lions and N. Masmoudi, *Global solutions for some oldroyd models of non-newtonian flows*, Chinese Annals of Mathematics, 21 (2000), pp. 131–146.

[27] C.-Y. D. Lu, P. D. Olmsted, and R. C. Ball, *Effects of nonlocal stress on the determination of shear banding flow*, Physical Review Letters, 84 (2000), pp. 642–645.

[28] Y. Lu and M. Pokorný, *Global existence of large data weak solutions for a simplified compressible Oldroyd-B model without stress diffusion*, Analysis in Theory and Applications, 36 (2020), pp. 348–372.

[29] M. Lukáčová-Medviďová, H. Mizerová, Š. Nečasová, and M. Renardy, *Global existence result for the generalized Peterlin viscoelastic model*, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 2950–2964.

[30] J. Málek, J. Nečas, M. Rokyta, and M. Růžička, *Weak and Measure-valued Solutions to Evolutionary PDEs*, Chapman & Hall, 1996.
[31] J. Málek and V. Průša, *Derivation of equations for continuum mechanics and thermodynamics of fluids*, in Handbook of mathematical analysis in mechanics of viscous fluids, Springer, Cham, 2018, pp. 3–72.
[32] J. Málek, V. Průša, T. Skřivan, and E. Süli, *Thermodynamics of viscoelastic rate-type fluids with stress diffusion*, Physics of Fluids, 30 (2018).
[33] J. Málek and K. R. Rajagopal, *Mathematical issues concerning the Navier-Stokes equations and some of its generalizations*, in Evolutionary equations. Vol. II, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2005, pp. 371–459.
[34] J. Málek, K. R. Rajagopal, and K. Tůma, *On a variant of the Maxwell and Oldroyd-B models within the context of a thermodynamic basis*, International Journal of Non-Linear Mechanics, 76 (2015), pp. 42 – 47.
[35] ———, *Derivation of the variants of the Burgers model using a thermodynamic approach and appealing to the concept of evolving natural configurations*, Fluids, 3 (2018).
[36] N. Masmoudi, *Global existence of weak solutions to macroscopic models of polymeric flows*, Journal de Mathématiques Pures et Appliquées, 96 (2011), pp. 502–520.
[37] J. Nečas, *Direct methods in the theory of elliptic equations*, Springer Monographs in Mathematics, Springer, Heidelberg, 2012. Translated from the 1967 French original by Gerard Tronel and Alois Kufner, Editorial coordination and preface by Šárka Nečasová and a contribution by Christian G. Simader.
[38] P. D. Olmsted, O. Radulescu, and C.-Y. D. Lu, *Johnson–segalman model with a diffusion term in cylindrical couette flow*, Journal of Rheology, 44 (2000), pp. 257–275.
[39] R. Pivokonský, P. Filip, and J. Zelenková, *The role of the gordon–schowalter derivative term in the constitutive models—improved flexibility of the modified XPP model*, Colloid and Polymer Science, 293 (2015), pp. 1227–1236.
[40] K. R. Rajagopal and A. R. Srinivasa, *A thermodynamic frame work for rate type fluid models*, J. Non-Newton. Fluid Mech., 88 (2000), pp. 207–227.
[41] ———, *On thermomechanical restrictions of continua*, Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci., 460 (2004), pp. 631–651.
[42] T. Roubíček, *Nonlinear partial differential equations with applications*, vol. 153 of International Series of Numerical Mathematics, Birkhäuser/Springer Basel AG, Basel, second ed., 2013.
[43] R. M. Wilcox, *Exponential operators and parameter differentiation in quantum physics*, Journal of Mathematical Physics, 8 (1967), pp. 962–982.
[44] A. Wouk, *Integral representation of the logarithm of matrices and operators*, Journal of Mathematical Analysis and Applications, 11 (1965), pp. 131–138.
[45] S. Zaremba, *Sur une forme perfectionee de la theorie de la relaxation*, Bull. Int. Acad. Sci. Cracovie, (1903), pp. 594–614.
[46] E. Zeidler, *Nonlinear functional analysis and its applications. II/A*, Springer-Verlag, New York, 1990. Linear monotone operators, Translated from the German by the author and Leo F. Boron.
[47] W. P. Ziemer, *Weakly differentiable functions*, vol. 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

Michal Bathory, University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria
Email address: `michal.bathory@univie.ac.at`

Miroslav Bulíček, Charles University, Faculty of Mathematics and Physics, Mathematical Institute, Sokolovská 83, 186 75 Praha 8, Czech Republic
Email address: `mbul8060@karlin.mff.cuni.cz`

Josef Málek, Charles University, Faculty of Mathematics and Physics, Mathematical Institute, Sokolovská 83, 186 75 Praha 8, Czech Republic
Email address: `malek@karlin.mff.cuni.cz`