Arch. Rational Mech. Anal. (2023) 247:5
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-022-01834-9

Nonlinear and Linearized Models in Thermoviscoelasticity

RUFAT BADAL, MANUEL FRIEDRICH & MARTIN KRUŽÍK

Communicated by M. ORTIZ

Abstract

We consider a quasistatic nonlinear model in thermoviscoelasticity at a finite-strain setting in the Kelvin–Voigt rheology, where both the elastic and viscous stress tensors comply with the principle of frame indifference under rotations. The force balance is formulated in the reference configuration by resorting to the concept of nonsimple materials, whereas the heat transfer equation is governed by the Fourier law in the deformed configurations. Weak solutions are obtained by means of a staggered in-time discretization where the deformation and the temperature are updated alternatingly. Our result refines a recent work by Mielke and Roubíček (Arch Ration Mech Anal 238:1–45, 2020) since our approximation does not require any regularization of the viscosity term. Afterwards, we focus on the case of deformations near the identity and small temperatures, and we show by a rigorous linearization procedure that weak solutions of the nonlinear system converge in a suitable sense to solutions of a system in linearized thermoviscoelasticity. The same property holds for time-discrete approximations and we provide a corresponding commutativity result.

1. Introduction

Nonlinear and large strain continuum mechanics has become a thriving field of research over the last few decades; it is still subject of important advancements and, at the same time, offers many challenging open questions. For instance, rigorous studies on large strain viscoelastic materials [19,24,26,31] or nonlinear models in thermoviscoelasticity [33] have been initiated only recently. Besides analytical intricacies, the usage of large strain models in engineering practice is often impeded due to nonconvex behavior that complicates numerical implementations. On many occasions, however, linearized models are still sufficient to describe observed phenomena and are significantly easier to treat, both analytically and numerically.

Published online: 05 January 2023

Roughly speaking, heuristic calculations suggest that, if the deformation of the body is "close" to the identity, nonlinear models can be replaced by linear ones with a negligible error. Clearly, the reliability of such predictions depends on the rigorous derivation of simplified linearized models, e.g., via Γ-convergence [7,12]. This is an intensive research program that has been initiated in the context of linearized elastostatics in [13]. Subsequently, this work was extended in various directions, among others, models for incompressible materials [23,27,28], atomistic models [8,40], or problems without Dirichlet boundary conditions [29] have been considered. For multiwell energies allowing for phase transitions we refer to [1,14,39], and we mention also settings beyond elasticity such as plasticity [34] or fracture [18,20]. As to evolutionary models, we refer to [19] where viscoelasticity in the Kelvin–Voigt rheology and its linearized version are treated.

The goal of this contribution is to couple the nonlinear equations of viscoelasticity with a heat transfer equation. We first analyze a corresponding frame-indifferent and thermodynamically-consistent model of thermoviscoelasticity at large strains, and refine the results obtained recently by Mielke and Roubíček [33]. Afterwards, in the spirit of the isothermal result [19], we pass to a linearized limit in terms of rescaled displacement fields and different regimes of rescaled temperatures.

We start by introducing the large strain model. Neglecting inertia effects, a nonlinear viscoelastic material in Kelvin–Voigt rheology obeys the following system of equations

$$-\operatorname{div}\big(\partial_F W(\nabla y, \theta) + \partial_{\dot F} R(\nabla y, \nabla \dot y, \theta)\big) = f \qquad \text{in } [0,T]\times\Omega. \tag{1.1}$$

here $[0, T]$ is a process time interval with $T > 0$, $\Omega \subset \mathbb{R}^d$ is a bounded domain representing the reference configuration, $y\colon [0, T] \times \Omega \to \mathbb{R}^d$ is a deformation mapping, ∇y is the deformation gradient, θ denotes the temperature, $W\colon \mathbb{R}^{d\times d} \times [0, \infty) \to \mathbb{R} \cup \{+\infty\}$ is a stored energy density, which represents a potential of the first Piola–Kirchhoff stress tensor $\partial_F W$, and $F \in \mathbb{R}^{d\times d}$ is the placeholder of ∇y. Finally, $R\colon \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d} \times [0, \infty) \to \mathbb{R}$ denotes a (pseudo)potential of dissipative forces, where $\dot F$ is the time derivative of F, and $f\colon [0, T] \times \Omega \to \mathbb{R}^d$ is a volume density of external forces acting on Ω.

The density W respects frame indifference under rotations and positivity of the determinant of the deformation gradient, i.e., local non-self-penetration is realized. (In contrast to [24], we do not consider conditions implying global non-self-penetration.) At the same time, we focus on physically correct viscous stresses, i.e., as observed by Antman [3], R must comply with a time-continuous frame indifference principle meaning that for all F it holds that

$$R(F, \dot F, \theta) = \hat R(C, \dot C, \theta)$$

for some nonnegative function $\hat R$, where $C := F^\top F$ and $\dot C := \dot F^T F + F^T \dot F$.

In contrast to the rapidly developed static theory at large strains, already in the isothermal case existence of solutions to (1.1) remains a challenging problem and results for models respecting the physically relevant frame indifference for both W and R are scarce. We refer, e.g., to [26] for local in-time existence or to [15] for the existence of measure-valued solutions. To date, weak solutions in finite strain isothermal viscoelasticity [19,24,33] can only be guaranteed by using the concept of second-grade nonsimple materials where the stored energy density (and consequently the first Piola–Kirchhoff stress tensor) additionally depends on the second gradient of the deformation. This idea was first introduced by Toupin [41,42] and proved to be useful in mathematical continuum mechanics, see e.g. [4,5,32,35]. In this spirit, we consider a version of (1.1) for nonsimple materials where the stored energy density depends also on the second gradient of y, and (1.1) is replaced by

$$-\operatorname{div}\Big(\partial_F W(\nabla y,\theta)-\operatorname{div}(\partial_G H(\nabla^2 y))+\partial_{\dot F} R(\nabla y,\nabla\dot y,\theta)\Big)=f \qquad \text{in } [0,T]\times\Omega, \tag{1.2}$$

which corresponds to an additional convex term $\int_\Omega H(\nabla^2 y)\mathrm{d}x$ in the stored energy. Let us mention that a main justification of this model lies in the observation that, in the small strain limit and under suitable scaling, the problem leads to the standard system of linear viscoelasticity without second gradient [19].

In the present contribution, we focus on a nonlinear coupling of the system (1.2) with a heat transfer equation of the form

$$\begin{aligned} c_V(\nabla y,\theta)\,\dot\theta &= \operatorname{div}(\mathcal{K}(\nabla y,\theta)\nabla\theta)+\partial_{\dot F} R(\nabla y,\nabla\dot y,\theta):\nabla\dot y\\ &\quad+\theta\partial_{F\theta}W^{\mathrm{cpl}}(\nabla y,\theta):\nabla\dot y \text{in } [0,T]\times\Omega, \end{aligned} \tag{1.3}$$

where $\partial_{F\theta} := \partial_F\partial_\theta$, W^{cpl} denotes a thermo-mechanical coupling potential, $c_V(F,\theta) = -\theta\partial_\theta^2 W^{\mathrm{cpl}}(F,\theta)$ is the heat capacity, $\mathcal{K}$ denotes the matrix of the heat-conductivity coefficients, and the last term plays the role of an adiabatic heat source. This corresponds to heat transfer modeled by the Fourier law in the deformed configuration which is however pulled back to the reference configurations, whence $\mathcal{K}$ depends on the deformation gradient. Here, following [33], we assume a rather weak thermal coupling by using the splitting of the free energy W via the explicit ansatz

$$W(F,\theta)=W^{\mathrm{el}}(F)+W^{\mathrm{cpl}}(F,\theta), \tag{1.4}$$

implying that $\partial_\theta W = \partial_\theta W^{\mathrm{cpl}}$. The coupled system (1.2)–(1.3) is equipped with suitable initial and boundary conditions, see (2.17)–(2.18) below.

Thermoviscoelasticity is a notoriously difficult problem already at small strains, e.g., there is no obvious variational structure of thermal part due to the low regularity of data. New developments in the the L^1-theory for the nonlinear heat equation [9,10] paved the way to advancements in small strain thermoviscoelasticity (for example, see [6,11,37]) which eventually culminated in the analysis of a physically sound large-strain model by Mielke and Roubíček [33]. We refer to [33, Introduction; items (α)–(ε)] for the main properties and challenges for this model which coincides with ours up to minor points, see Remark 2.1. Their existence result is based on a time-incremental approach for a regularized system which does not comply with the above mentioned frame indifferent principles, e.g., in (1.2) a term $\lambda\nabla\dot y$ is added for $\lambda > 0$. Then, they first pass to the time-continuous limit in the regularized problem and eventually recover the original system in the limit of the vanishing regularization parameter λ.

The first result of our work (Theorem 2.3) revisits their study by proposing a slightly different semidiscretization in time which directly approximates the PDE system in the limit for vanishing time steps and comes along without any regularization. Although establishing the same existence result on weak solutions, our approach sheds new light on the issue as we propose a time-discrete approximation scheme complying with frame indifference. This, combined with a spatial discretization, see e.g. [25, Section 9.3], could be the basis for a numerical implementation. As in [33], our scheme is staggered, i.e., first the deformation is updated at fixed temperature from the previous time step and then the temperature is updated. Our scheme differs in the usage of explicit or implicit steps, i.e., whether in certain terms the 'old' or the 'new' temperature is used, see Remark 2.4. By means of delicate estimates on the coupling potential, we are hereby able to establish the necessary a priori bounds without any regularization. At this point, we derive a priori estimates for different scalings of the elastic strains and the temperature which is at the basis of our subsequent analysis on small-strain limits.

In the second part of our work, we are interested in the case of small strains and temperatures, i.e., when $\nabla u := \nabla y - \mathbf{Id}$ is of order ε for some small $\varepsilon > 0$ and θ is of order ε^α for any exponent $\alpha > 0$. Here, $u := y - \mathbf{id}$ is the displacement corresponding to y with $\mathbf{id}$ and $\mathbf{Id}$ standing for the identity map and identity matrix, respectively. Such properties are certainly reasonable if initial values and boundary values for the deformation and the temperature are close to the identity or zero, respectively. Therefore, it is convenient to introduce the rescaled displacement $u_\varepsilon = \varepsilon^{-1}(y - \mathbf{id})$ and rescaled temperature $\mu_\varepsilon = \varepsilon^{-\alpha}\theta$, and to replace f by εf. We write (1.2)–(1.3) in terms of the rescaled quantities and multiply (1.2) with ε^{-1} and (1.3) with $\varepsilon^{-\alpha}$. Then, formally, we can pass to the limit and obtain the system

$$\begin{aligned} -\operatorname{div}\left(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot u) + \mathbb{B}^{(\alpha)}\mu\right) &= f, \\ \bar c_V \dot\mu - \operatorname{div}(\mathbb{K}_0 \nabla\mu) &= \mathbb{C}_D^{(\alpha)} e(\dot u) : e(\dot u), \end{aligned} \tag{1.5}$$

where $\mathbb{C}_W := \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})$ is the tensor of elastic constants (W^{el} is defined in (1.4)), $\mathbb{C}_D := \partial_{\dot F}^2 R(\mathbf{Id}, \dot F, 0)$ is the tensor of viscosity coefficients, $\mathbb{B}^{(\alpha)}$ represents a thermal expansion matrix, $\bar c_V$ is the heat capacity at zero temperature and the stress free material state, and $\mathbb{K}_0 := \mathcal{K}(\mathbf{Id}, 0)$. Finally, $e(u) := (\nabla u + (\nabla u)^\top)/2$ denotes the linear strain tensor and $e(\dot u)$ the strain rate. By different scaling properties of the two equations, it turns out that the limit is α-dependent and, as we point out later, only meaningful in the regime $\alpha \in [1, 2]$. The matrix $\mathbb{B}^{(\alpha)}$ is only active for $\alpha = 1$ and in this case it is related to the coupling potential, namely $\mathbb{B}^{(\alpha)} = \partial_{F\theta} W^{\mathrm{cpl}}(\mathbf{Id}, 0)$. On the other hand, $\mathbb{C}_D^{(\alpha)}$ is nonzero only for $\alpha = 2$ and then it coincides with $\mathbb{C}_D$. Interestingly, although the nonlinear thermoviscoelasticity system is written for a nonsimple material, in the limit we obtain linear equations without spatial gradients of $e(u)$. This relies on the fact that we assume H to have super-quadratic growth at 0. Formal derivations of such PDE systems is not new and can be found, e.g., in [21, Section 59]. The second main contribution of our work (Theorem 2.7) is to make this limit passage rigorous, i.e., we show that solutions to the nonlinear system (1.2)–(1.3) converge in a suitable sense to the unique solution of the linear system (1.5) as $\varepsilon \to 0$. Besides this convergence result, we also get analogous convergences for time-discretized problems, and we confirm that convergences for vanishing time step and $\varepsilon \to 0$ commute, see Theorem 2.8.

To the best of on knowledge, this is the first linearization result of a mechanical model coupled with heat transfer in the material. We perform linearization near the natural (i.e., stress free) state and zero temperature. Without further details, let us however mention that by a shifting argument our techniques would allow to linearize about a fixed, positive temperature θ_c, whenever the initial and boundary data lie above θ_c and the coupling potential $W^{\mathrm{cpl}}(F, \theta)$ vanishes for $\theta \leqq \theta_c$.

We now give an outline of the paper and present some fundamental ingredients of the proof. After some basic notations, we introduce the nonlinear setting in Sect. 2.1. In Sect. 2.2, we formulate our semi-discrete approximation result in the nonlinear setting and briefly highlight the differences to the scheme in [33], see Remark 2.4. In Sect. 2.3, we introduce the linearized setting and present our results on convergence of solutions in the nonlinear-to-linear passage.

In Sects. 3.1–3.2, we address the well-definedness of the staggered time-incremental scheme. The core of our approach is an inductive bound on the total energy, see Lemma 3.11: this is achieved by suitably testing the momentum balance and the heat-transfer equation, adding the two equations, and exploiting cancellation of the dissipation. In contrast to [33], see particularly [33, Remark 6.1], this cancellation is already possible in the time-discrete setting as we use a simpler, explicit, thermo-mechanical coupling term in the scheme allowing us to proceed without the necessity of regularizing terms. This, however, comes at the expense of the fact that the argument to guarantee nonnegativity of the temperature in the thermal step is more sophisticated. For this, we need a delicate estimate for the coupling potential, see Proposition 3.8.

As a preparation for the passage to the linearized system, we need an adaption of the bound on the total energy, see Sect. 3.3. In fact, due to the different scaling ε and ε^α of the mechanical and the heat-transfer equation, the above mentioned cancellation cannot be used in general for small ε. Thus, novel techniques are required to tame the contribution of the dissipation including higher integrability of the temperature variable, see Lemma 3.15 for details. Section 3 is closed with a priori bounds derived from the energy bound, see Sect. 3.4. As in [33], the main ingredients here are Gagliardo–Nirenberg interpolation inequalities and special test functions developed by Boccardo and Gallouët [10] for parabolic equations with measure-valued right-hand side. For convenience of the reader, almost complete proofs are provided since in addition to [33] we need scaling invariant estimates in terms of the small parameter ε.

In Sect. 4 we then address the passage to vanishing time steps in the nonlinear model. At this point, having settled the a priori estimates, we can essentially follow [33]. Since we work without regularization terms, however, we need to combine and adapt the techniques from Sections 5–6 of [33], and therefore we elaborate the proofs to some extent. Eventually, Sect. 5 is devoted to the linearization. In Sect. 5.1 we first deal with the passage to the time-continuous problem. The strategy in the proof is similar to the one in the nonlinear setting in Sect. 4, with the additional challenge that in each term we need to ensure that higher order terms in Taylor expansions are asymptotically negligible. In particular, we show that contributions of the second gradient vanish in the limit. As in the nonlinear setting, strong convergence of the strains and the strain rates is necessary to pass to the limit, see Lemma 4.5 and Lemma 5.4. Due to rescaling of the equations, however, this is more demanding in the passage to the linerized setting as higher integrability of the temperature is needed to control the coupling term, cf. Remark 4.3. Eventually, Sect. 5.2 is devoted to time-discrete problems which particularly involves a Γ-convergence result for the mechanical part, see Proposition 5.7.

2. The Model and Main Results

2.1. The setting and modeling assumptions

In what follows, we use standard notation for Lebesgue and Sobolev spaces. The lower index $_+$ means nonnegative elements, i.e., $L^2_+(\Omega)$ denotes the convex cone of nonnegative functions belonging to $L^2(\Omega)$ and a similar definition is used for $H^1_+(\Omega)$. We also set $\mathbb{R}_+ := [0, +\infty)$. Let $a \wedge b := \min\{a, b\}$ for $a, b \in \mathbb{R}$. Denoting by $d \geqq 2$ the dimension, we let $\mathbf{Id} \in \mathbb{R}^{d\times d}$ be the identity matrix, and $\mathbf{id}(x) := x$ stands for the identity map on $\mathbb{R}^d$. We define the subsets $SO(d) := \{A \in \mathbb{R}^{d\times d} \colon A^T A = \mathbf{Id},\ \det A = 1\}$, $GL^+(d) := \{F \in \mathbb{R}^{d\times d} \colon \det(F) > 0\}$, and $\mathbb{R}^{d\times d}_{\mathrm{sym}} := \{A \in \mathbb{R}^{d\times d} \colon A^T = A\}$. Furthermore, $F^{-T} := (F^{-1})^T = (F^T)^{-1}$, and given a tensor (of arbitrary dimension), $|F|$ will denote its Frobenius norm. We denote the scalar product between vectors, matrices, or 3rd-order tensors by $\cdot$, $:$, and $\vdots$, respectively. As usual, in the proofs generic constants C may vary from line to line. If not stated otherwise, constants depend only on d, $p > d$, Ω, $\alpha > 0$, and the potentials introduced in the sequel. We frequently use a scaled version of Young's inequality with constant $\lambda \in (0, 1)$ by which we mean $ab \leqq \lambda a^p + C b^q/\lambda$ for $a, b \geqq 0$, exponents $p, q \geqq 1$ with $1/p + 1/q = 1$, and $C > 0$ large enough.

Consider an open bounded set $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary $\Gamma := \partial\Omega$. Let Γ_D, Γ_N be disjoint Borel subsets of Γ such that $\mathcal{H}^{d-1}(\Gamma_D) > 0$, and $\Gamma = \Gamma_D \cup \Gamma_N$, representing Dirichlet and Neumann parts of the boundary, respectively. For $p > d$, we introduce the set of *admissible deformations* by

$$\mathcal{Y}_{\mathbf{id}} := \{*\} y \in W^{2,p}(\Omega; \mathbb{R}^d) \colon y = \mathbf{id} \text{ on } \Gamma_D,\ \det(\nabla y) > 0 \text{ in } \Omega, \tag{2.1}$$

and we say that the *absolute temperature* θ is admissible if $\theta \in L^1_+(\Omega)$. We also introduce the space

$$W^{2,p}_{\Gamma_D}(\Omega; \mathbb{R}^d) := \{y \in W^{2,p}(\Omega; \mathbb{R}^d) \colon y = 0 \text{ on } \Gamma_D\}. \tag{2.2}$$

Next, we discuss our variational setting. In this regard, let c_0, C_0 with $0 < c_0 < C_0 < \infty$ be some fixed constants.

2.1.1. Mechanical energy and coupling energy The *elastic energy* $\mathcal{W}^{\mathrm{el}} \colon \mathcal{Y}_{\mathbf{id}} \to \mathbb{R}_+$ is given by

$$\mathcal{W}^{\mathrm{el}}(y) := \int_\Omega W^{\mathrm{el}}(\nabla y)\,\mathrm{d}x, \tag{2.3}$$

where $W^{\mathrm{el}}\colon GL^+(d) \to \mathbb{R}_+$ is a frame indifferent elastic energy potential with the usual assumptions in nonlinear elasticity. More precisely, we require that

(W.1) W^{el} is continuous and C^3 in a neighborhood of $SO(d)$;
(W.2) Frame indifference: $W^{\mathrm{el}}(QF) = W^{\mathrm{el}}(F)$ for all $F \in GL^+(d)$ and $Q \in SO(d)$;
(W.3) Lower bound: $W^{\mathrm{el}}(F) \geqq c_0\big(|F|^2 + \det(F)^{-q}\big) - C_0$ for all $F \in GL^+(d)$, where $q \geqq \frac{pd}{p-d}$.

Adopting the concept of 2nd-grade nonsimple materials, see [41,42], we also consider a *strain gradient energy term* $\mathcal{H}\colon \mathcal{Y}_{\mathbf{id}} \to \mathbb{R}_+$, defined as

$$\mathcal{H}(y) := \int_\Omega H(\nabla^2 y)\mathrm{d}x, \tag{2.4}$$

where its potential $H\colon \mathbb{R}^{d\times d\times d} \to \mathbb{R}_+$ satisfies

(H.1) H is convex and C^1;
(H.2) Frame indifference: $H(QG) = H(G)$ for all $G \in \mathbb{R}^{d\times d\times d}$ and $Q \in SO(d)$;
(H.3) $c_0|G|^p \leqq H(G) \leqq C_0(1 + |G|^p)$ and $|\partial_G H(G)| \leqq C_0|G|^{p-1}$ for all $G \in \mathbb{R}^{d\times d\times d}$.

The *mechanical energy* $\mathcal{M}\colon \mathcal{Y}_{\mathbf{id}} \to \mathbb{R}_+$ is then defined as the sum

$$\mathcal{M}(y) := \mathcal{W}^{\mathrm{el}}(y) + \mathcal{H}(y). \tag{2.5}$$

Besides the mechanical energy, we introduce a *coupling energy* $\mathcal{W}^{\mathrm{cpl}}\colon \mathcal{Y}_{\mathbf{id}} \times L^1_+(\Omega) \to \mathbb{R}$ given by

$$\mathcal{W}^{\mathrm{cpl}}(y,\theta) := \int_\Omega W^{\mathrm{cpl}}(\nabla y, \theta)\mathrm{d}x,$$

where $W^{\mathrm{cpl}}\colon GL^+(d)\times\mathbb{R}_+ \to \mathbb{R}$ describes mutual interactions of mechanical and thermal effects (see e.g. [21]), and satisfies

(C.1) W^{cpl} is continuous and C^2 in $GL^+(d) \times (0,\infty)$;
(C.2) $W^{\mathrm{cpl}}(QF,\theta) = W^{\mathrm{cpl}}(F,\theta)$ for all $F \in GL^+(d)$, $\theta \geqq 0$, and $Q \in SO(d)$;
(C.3) $W^{\mathrm{cpl}}(F, 0) = 0$ for all $F \in GL^+(d)$;
(C.4) $|W^{\mathrm{cpl}}(F,\theta) - W^{\mathrm{cpl}}(\tilde{F},\theta)| \leqq C_0(1 + |F| + |\tilde{F}|)|F - \tilde{F}|$ for all F, $\tilde{F} \in GL^+(d)$, and $\theta \geqq 0$;
(C.5) For all $F \in GL^+(d)$ and $\theta > 0$ it holds that

$$|\partial_F^2 W^{\mathrm{cpl}}(F,\theta)| \leqq C_0, \quad |\partial_{F\theta} W^{\mathrm{cpl}}(F,\theta)| \leqq \frac{C_0(1+|F|)}{\max\{\theta, 1\}},$$
$$c_0 \leqq -\theta\partial_\theta^2 W^{\mathrm{cpl}}(F,\theta) \leqq C_0.$$

Notice that, by (C.3) and the second bound in (C.5), $\partial_F W^{\mathrm{cpl}}$ can be continuously extended to zero temperatures with $\partial_F W^{\mathrm{cpl}}(F,0) = 0$. For $F \in GL^+(d)$ and $\theta \geqq 0$, we define the *total free energy potential*

$$W(F,\theta) := W^{\mathrm{el}}(F) + W^{\mathrm{cpl}}(F,\theta). \tag{2.6}$$

We refer to [33, Examples 2.4 and 2.5] for a class of coupling potentials satisfying all assumptions above.

2.1.2. Dissipation potential The *dissipation functional* $\mathcal{R}\colon \mathcal{Y}_{\mathbf{id}} \times W^{2,p}_{\Gamma_D}(\Omega;\mathbb{R}^d)\times L^1_+(\Omega) \to \mathbb{R}_+$ is defined as

$$\mathcal{R}(\tilde{y}, y-\tilde{y}, \theta) := \int_\Omega R(\nabla\tilde{y}, \nabla y - \nabla\tilde{y}, \theta)\mathrm{d}x, \tag{2.7}$$

where $R\colon \mathbb{R}^{d\times d}\times\mathbb{R}^{d\times d}\times\mathbb{R}_+ \to \mathbb{R}_+$ is the *potential of dissipative forces* satisfying

(D.1) $R(F, \dot{F}, \theta) := \frac{1}{2}D(C,\theta)[\dot{C}, \dot{C}] := \frac{1}{2}\dot{C} : D(C,\theta)\dot{C}$, where $C := F^T F$, $\dot{C} := \dot{F}^T F + F^T \dot{F}$, and $D \in C(\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}_+; \mathbb{R}^{d\times d\times d\times d})$ with $D_{ijkl} = D_{jikl} = D_{klij}$ for $1 \leqq i, j, k, l \leqq d$;
(D.2) $c_0|\dot{C}|^2 \leqq \dot{C} : D(C,\theta)\dot{C} \leqq C_0|\dot{C}|^2$ for all $C,\ \dot{C} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, and $\theta \geqq 0$.

Notice that the fact that R can be written as a function depending on the right Cauchy-Green tensor $C = F^T F$ and its time derivative $\dot{C}$ is equivalent to *dynamic frame indifference* (see also [2]). Condition (D.1) also implies that the viscous stress $\partial_{\dot{F}} R(F, \dot{F}, \theta)$ is linear in the time derivative $\dot{C}$ as indeed a simple calculation shows for any $i,\ j \in \{1, \dots, d\}$:

$$\begin{aligned}\partial_{\dot{F}_{ij}} R(F, \dot{F}, \theta) &= \partial_{\dot{F}_{ij}}\big(\dot{F}_{mk}F_{ml} + F_{mk}\dot{F}_{ml}\big)\big(D(C,\theta)\dot{C}\big)_{kl} \\ &= (\delta_{im}\delta_{kj}F_{ml} + \delta_{im}\delta_{jl}F_{mk})\big(D(C,\theta)\dot{C}\big)_{kl} \\ &= F_{il}\big(D(C,\theta)\dot{C}\big)_{jl} + F_{ik}\big(D(C,\theta)\dot{C}\big)_{kl},\end{aligned}$$

where we have used Einstein summation convention for l, m in $\{1, \dots, d\}$, and δ_{ij} denotes the Kronecker symbol. Hence, by the symmetry of $D(C,\theta)$ (see (D.1)) and the arbitrariness of i and j this proves

$$\partial_{\dot{F}} R(F, \dot{F}, \theta) = 2F(D(C,\theta)\dot{C}). \tag{2.8}$$

The choice of a linear material viscosity is crucial in our approach and is a relevant modeling assumption for non-activated dissipative processes with rather moderate rates. We emphasize, however, that the geometrical nonlinearity of finite elasticity is still present due to $\dot{C}$ in (2.8), and that $\partial_{\dot{F}} R$ necessarily also depends on F, even for constant functions D. We also define the associated *dissipation rate* $\xi\colon \mathbb{R}^{d\times d}\times\mathbb{R}^{d\times d}\times\mathbb{R}_+ \to \mathbb{R}_+$ as

$$\begin{aligned}\xi(F, \dot{F}, \theta) &:= \partial_{\dot{F}} R(F, \dot{F}, \theta) : \dot{F} = 2F(D(C,\theta)\dot{C}) : \\ &\qquad \dot{F} = D(C,\theta)\dot{C} : (\dot{F}^T F + F^T \dot{F}) = 2R(F, \dot{F}, \theta),\end{aligned} \tag{2.9}$$

where the second identity follows from (2.8), and the third from the symmetries in (D.1).

2.1.3. Heat conductivity and internal energy The map $\mathbb{K}\colon \mathbb{R}_+ \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ will denote the *heat conductivity tensor* of the material in the deformed configuration. We require that $\mathbb{K}$ is continuous, symmetric, uniformly positive definite, and bounded. More precisely, for all $\theta \geqq 0$ it holds that

$$c_0 \leqq \mathbb{K}(\theta) \leqq C_0, \tag{2.10}$$

where the inequalities are meant in the eigenvalue sense. We define the pull-back $\mathcal{K}\colon GL^+(d) \times \mathbb{R}_+ \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ of $\mathbb{K}$ into the reference configuration by (see [33, (2.24)])

$$\mathcal{K}(F,\theta) := \det(F)F^{-1}\mathbb{K}(\theta)F^{-T}. \tag{2.11}$$

2.1.4. Internal and total energy The *(thermal part of the) internal energy* $W^{\mathrm{in}}\colon GL^+(d) \times (0,\infty) \to \mathbb{R}$ is defined as

$$W^{\mathrm{in}}(F,\theta) := W^{\mathrm{cpl}}(F,\theta) - \theta\partial_\theta W^{\mathrm{cpl}}(F,\theta). \tag{2.12}$$

Using (C.3) and the third bound in (C.5), we can easily see that W^{in} can be continuously extended to zero temperatures by setting $W^{\mathrm{in}}(F,0) = 0$ for all $F \in GL^+(d)$. Also by the third bound in (C.5), the internal energy is controlled by the temperature in the sense that

$$\partial_\theta W^{\mathrm{in}}(F,\theta) = -\theta\partial_\theta^2 W^{\mathrm{cpl}}(F,\theta) \in [c_0, C_0] \qquad \text{for all } F \in GL^+(d) \text{ and } \theta > 0, \tag{2.13}$$

which along with (C.3), yields

$$c_0\theta \leqq W^{\mathrm{in}}(F,\theta) \leqq C_0\theta. \tag{2.14}$$

Eventually, the *total energy functional* $\mathcal{E}\colon \mathcal{Y}_{\mathbf{id}} \times L^1_+(\Omega) \to \mathbb{R}_+$ is then given by

$$\mathcal{E}(y,\theta) := \mathcal{M}(y) + \mathcal{W}^{\mathrm{in}}(y,\theta) \quad \text{with } \mathcal{W}^{\mathrm{in}}(y,\theta) := \int_\Omega W^{\mathrm{in}}(\nabla y,\theta)\mathrm{d}x. \tag{2.15}$$

Remark 2.1. (Comparison to [33]) We close this part on modeling assumptions by highlighting the differences to the assumptions in [33]: Our condition in (W.3) is slightly more general than the corresponding one in [33, (2.30a)], where the term $|F|^2$ is replaced by $|F|^s$ for $s > 2$. We do not assume that W^{cpl} is bounded from below. Condition (C.3) as well as bounds similar to (C.4)–(C.5) are also required in [33], see [33, (2.15), (2.30)]. There, the bound on $\partial_{F\theta}W^{\mathrm{cpl}}$ is slightly more general for θ near zero, and only an upper bound on the eigenvalues of $\partial_F^2 W^{\mathrm{cpl}}(F,\theta)$ is required, see [33, (2.30c)]. This similarity of the assumptions will in particular allow us to employ several intermediate steps proven in [33]. For models complying with the above assumptions we refer to [33, Examples 2.4, 2.5].

2.1.5. Equations of nonlinear thermoviscoelasticity Fixing a finite time horizon $T > 0$, let us from now on shortly write $I := [0,T]$. We fix a constant $\varepsilon \in (0,1]$ which represents the *magnitude of the elastic strain*. In the first part of the paper, we are mainly interested in the large strain setting, where $\varepsilon = 1$. However, later we perform the passage to the small strain limit $\varepsilon \to 0$. To allow for a consistent notation, we include the parameter ε throughout the entire paper. Let εf with $f \in W^{1,1}(I; L^2(\Omega;\mathbb{R}^d))$ be a time-dependent *dead force*, εg with $g \in W^{1,1}(I; L^2(\Gamma_N;\mathbb{R}^d))$ be a *boundary traction*, and let $\varepsilon^\alpha\theta_\flat$ with $\theta_\flat \in W^{1,1}(I; L^2_+(\Gamma))$ and $\alpha > 0$ be an external temperature. We study the coupled system

$$\varepsilon f = -\operatorname{div}\big(\partial_F W(\nabla y,\theta) + \partial_{\dot F} R(\nabla y,\nabla\dot y,\theta) - \operatorname{div}(\partial_G H(\nabla^2 y))\big), \tag{2.16a}$$

$$-\theta\partial^2_\theta W^{\mathrm{cpl}}(\nabla y,\theta)\,\dot\theta = \operatorname{div}(\mathcal{K}(\nabla y,\theta)\nabla\theta) + \xi(\nabla y,\nabla\dot y,\theta) + \theta\partial_{F\theta}W^{\mathrm{cpl}}(\nabla y,\theta) : \nabla\dot y, \tag{2.16b}$$

which, as in [33], is complemented with the boundary conditions

$$\big(\partial_F W(\nabla y,\theta) + \partial_{\dot F} R(\nabla y,\nabla\dot y,\theta)\big)\nu - \operatorname{div}_S\big(\partial_G H(\nabla^2 y)\nu\big) = \varepsilon g \quad \text{on } I\times\Gamma_N, \tag{2.17a}$$

$$y = \mathbf{id} \quad \text{on } I\times\Gamma_D, \tag{2.17b}$$

$$\partial_G H(\nabla^2 y) : (\nu\otimes\nu) = 0 \quad \text{on } I\times\Gamma, \tag{2.17c}$$

$$\mathcal{K}(\nabla y,\theta)\nabla\theta\cdot\nu + \kappa\theta = \kappa\varepsilon^\alpha\theta_\flat \quad \text{on } I\times\Gamma. \tag{2.17d}$$

here ν denotes the outward pointing unit normal on Γ and $\kappa \geqq 0$ is a *phenomenological heat-transfer coefficient* on Γ. Moreover, div_S represents the *surface divergence*, defined by $\operatorname{div}_S(\cdot) = \operatorname{tr}(\nabla_S(\cdot))$, where tr denotes the trace and $\nabla_S := (\mathbf{Id} - \nu\otimes\nu)\nabla$ denotes the surface gradient. We refer to [33, (2.28)–(2.29)] for an explanation and derivation of the boundary conditions. Note that by (2.9) the system (2.16) indeed coincides with (1.2)–(1.3). The mechanical evolution (2.16a) is the quasistatic version of the Kelvin–Voigt rheological model (neglecting inertia), corresponding to the sum of the conservative and the dissipative forces. The equation (2.16b) follows from the entropy equation $\theta\dot s = \xi - \operatorname{div} q$, where the *entropy* s is expressed in terms of the free energy by $s = -\partial_\theta W = -\partial_\theta W^{\mathrm{cpl}}$. Furthermore, the dissipation rate ξ is defined in (2.9) and the *heat flux* q is modeled by the *Fourier law* in the deformed configuration, pulled back to the reference configuration, i.e., $q = -\mathcal{K}(F,\theta)\nabla\theta$. The term $-\theta\partial^2_\theta W^{\mathrm{cpl}}(\nabla y,\theta)$ corresponds to the *heat capacity* at constant volume and the last term in (2.16b) is an *adiabatic heat* source. We again refer to [33] or to [25, Section 8.1] for details. Notice that the the purely mechanical stored energy W^{el}, see (2.3), does not influence the heat production and transfer in (2.16b).

We consider a corresponding initial-value problem, by imposing the initial conditions

$$y(0,\cdot) = y_{0,\varepsilon} := \mathbf{id} + \varepsilon u_0 \quad\text{and}\quad \theta(0,\cdot) = \theta_{0,\varepsilon} := \varepsilon^\alpha\mu_0 \tag{2.18}$$

for some $\mu_0 \in L^2_+(\Omega)$ and some $u_0 \in W^{2,p}_{\Gamma_D}(\Omega;\mathbb{R}^d)$. We now define weak solutions associated to the initial-boundary-value problem (2.16)–(2.18).

Definition 2.2. (*Weak solution of the nonlinear system*) A couple $(y,\theta)\colon I\times\Omega \to \mathbb{R}^d\times\mathbb{R}$ is called a *weak solution* to the initial-boundary-value problem (2.16)–(2.18) if $y \in L^\infty(I;\mathcal{Y}_{\mathbf{id}}) \cap H^1(I;H^1(\Omega;\mathbb{R}^d))$ with $y(0,\cdot) = y_{0,\varepsilon}$, $\theta \in L^1(I;W^{1,1}(\Omega))$ with $\theta \geqq 0$ a.e., and if it satisfies the identities

$$\begin{aligned}\int_0^T\int_\Omega &\partial_G H(\nabla^2 y)\,\vdots\,\nabla^2 z \\ &+ \Big(\partial_F W(\nabla y,\theta) + \partial_{\dot F} R(\nabla y,\nabla\dot y,\theta)\Big) : \nabla z \,\mathrm{d}x\,\mathrm{d}t \\ &= \varepsilon\int_0^T\int_\Omega f\cdot z\,\mathrm{d}x\,\mathrm{d}t + \varepsilon\int_0^T\int_{\Gamma_N} g\cdot z\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t\end{aligned} \tag{2.19}$$

for any test function $z \in C^\infty(I \times \overline{\Omega}; \mathbb{R}^d)$ with $z = 0$ on $I \times \Gamma_D$, as well as

$$\begin{aligned}&\int_0^T \int_\Omega \mathcal{K}(\nabla y, \theta)\nabla\theta \cdot \nabla\varphi - \big(\xi(\nabla y, \nabla\dot{y}, \theta) + \partial_F W^{\mathrm{cpl}}(\nabla y, \theta) : \nabla\dot{y}\big)\varphi - W^{\mathrm{in}}(\nabla y, \theta)\dot{\varphi}\,\mathrm{d}x\,\mathrm{d}t \\ &\quad + \kappa \int_0^T \int_\Gamma \theta\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t = \kappa\varepsilon^\alpha \int_0^T \int_\Gamma \theta_\flat\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t + \int_\Omega W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon})\,\varphi(0)\,\mathrm{d}x \quad (2.20)\end{aligned}$$

for any test function $\varphi \in C^\infty(I \times \overline{\Omega})$ with $\varphi(T) = 0$.

One can indeed show that sufficiently smooth weak solutions lead to the classical formulation (2.16) along with the boundary conditions (2.17), see [33]. We refer to [33, (2.28)–(2.29)] for details on the derivation of (2.16a), particularly how to treat the boundary terms. For the derivation of (2.16b), one uses standard integration by parts and the fact that by the definition in (2.12) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(W^{\mathrm{in}}(\nabla y, \theta)) = \partial_F W^{\mathrm{cpl}}(\nabla y, \theta) : \nabla\dot{y} - \theta\partial_{F\theta} W^{\mathrm{cpl}}(\nabla y, \theta) : \nabla\dot{y} - \theta\partial_\theta^2 W^{\mathrm{cpl}}(\nabla y, \theta)\dot{\theta}.$$

Moreover, using test functions with $\varphi(0) \neq 0$ we obtain $W^{\mathrm{in}}(\nabla y(0), \theta(0)) = W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon})$, and by the strict monotonicity in (2.13) along with $y(0) = y_{0,\varepsilon}$ we conclude $\theta(0) = \theta_{0,\varepsilon}$. We emphasize that one can only expect the regularity $\nabla\dot{y} \in L^2(I \times \Omega; \mathbb{R}^{d\times d})$ and thus $\xi(\nabla y, \nabla\dot{y}, \theta) \in L^1(I \times \Omega)$ by (2.9). Therefore, (2.16b) can be understood as a heat equation with L^1-data. For this, (2.20) is a standard weak formulation, see e.g. [37].

2.2. Approximation of solutions in the nonlinear setting

In this subsection, we study the nonlinear system and therefore we fix $\varepsilon = 1$. (In the notation, ε is still included, as before.) The existence of energy-conserving weak solutions to (2.16) in the sense of Definition 2.2 has been proven in [33, Theorem 2.2]. In contrast to this work, we show here that the solutions can be obtained directly as limits of a staggered time-incremental scheme without using any additional regularization.

We fix a discrete time step size $\tau \in (0, 1]$. For the sake of notational clarity, we assume without a further mention that any τ we encounter evenly divides the time interval $[0, T]$. Given any sequence $(a_l)_{l\geqq 0}$, it will be useful to introduce the following notation for discrete differences

$$\delta_\tau a_l := \frac{a_l - a_{l-1}}{\tau}, \quad l \in \mathbb{N}.$$

Our time-discrete staggered scheme is initialized by setting

$$y_\tau^{(0)} := y_{0,\varepsilon} \qquad \text{and} \qquad \theta_\tau^{(0)} := \theta_{0,\varepsilon}, \tag{2.21}$$

where $y_{0,\varepsilon}$ and $\theta_{0,\varepsilon}$ are as in (2.18). We then alternate between a *mechanical step*, deforming the material while keeping the temperature fixed, and a *thermal step*, adjusting the temperature distribution inside the material without changing the deformation. More precisely, suppose that we have already constructed

$y_\tau^{(0)}, \ldots, y_\tau^{(k-1)} \in \mathcal{Y}_{\mathbf{id}}$, and $\theta_\tau^{(0)}, \ldots, \theta_\tau^{(k-1)} \in L^2_+(\Omega)$ for some $k \in \{1, \ldots, T/\tau\}$. The next deformation $y_\tau^{(k)}$ is a solution of the minimization problem

$$\min_{y \in \mathcal{Y}_{\mathbf{id}}} \Big\{ \mathcal{M}(y) + \mathcal{W}^{\mathrm{cpl}}\big(y, \theta_\tau^{(k-1)}\big) + \frac{1}{\tau}\mathcal{R}(y_\tau^{(k-1)}, y - y_\tau^{(k-1)}, \theta_\tau^{(k-1)}) - \varepsilon\langle \ell_\tau^{(k)}, y\rangle \Big\}, \tag{2.22}$$

where

$$\langle \ell_\tau^{(k)}, y\rangle := \int_\Omega f_\tau^{(k)} \cdot y \mathrm{d}x + \int_{\Gamma_N} g_\tau^{(k)} \cdot y \mathrm{d}\mathcal{H}^{d-1} \tag{2.23}$$

for $f_\tau^{(k)} := \tau^{-1} \int_{(k-1)\tau}^{k\tau} f(t)\mathrm{d}t$ and $g_\tau^{(k)} := \tau^{-1} \int_{(k-1)\tau}^{k\tau} g(t)\mathrm{d}t$. We define the k-th temperature step $\theta_\tau^{(k)}$ as a solution of the minimization problem

$$\begin{aligned} \min_{\theta \in H^1_+(\Omega)} \Big\{ &\int_\Omega \int_0^\theta \frac{1}{\tau}\big(W^{\mathrm{in}}(\nabla y_\tau^{(k)}, s) - W^{\mathrm{in}}(\nabla y_\tau^{(k-1)}, \theta_\tau^{(k-1)})\big)\mathrm{d}s\mathrm{d}x \\ &+ \int_\Omega \frac{1}{2}\nabla\theta \cdot \mathcal{K}(\nabla y_\tau^{(k-1)}, \theta_\tau^{(k-1)})\nabla\theta\mathrm{d}x \\ &- \int_\Omega h_\tau(y_\tau^{(k)}, y_\tau^{(k-1)}, \theta_\tau^{(k-1)})\,\theta\mathrm{d}x + \frac{\kappa}{2}\int_\Gamma (\theta - \varepsilon^\alpha \theta_{\flat,\tau}^{(k)})^2 \mathrm{d}\mathcal{H}^{d-1} \Big\}, \end{aligned} \tag{2.24}$$

where h_τ plays the role of a heat source given by

$$\begin{aligned} &h_\tau(y_\tau^{(k)}, y_\tau^{(k-1)}, \theta_\tau^{(k-1)}) \\ &\quad := \partial_F W^{\mathrm{cpl}}(\nabla y_\tau^{(k-1)}, \theta_\tau^{(k-1)}) : \delta_\tau \nabla y_\tau^{(k)} + \xi(\nabla y_\tau^{(k-1)}, \delta_\tau \nabla y_\tau^{(k)}, \theta_\tau^{(k-1)}) \end{aligned} \tag{2.25}$$

and $\theta_{\flat,\tau}^{(k)} := \frac{1}{\tau}\int_{(k-1)\tau}^{k\tau} \theta_\flat(t)\mathrm{d}t$. The underlying idea is that the Euler–Lagrange equations associated to (2.22) and (2.24) lead to time-discretized variants of the equations (2.16), see (3.7) and (3.11) below. Supposing that the steps $y_\tau^{(0)}, \ldots, y_\tau^{(T/\tau)}$ and $\theta_\tau^{(0)}, \ldots, \theta_\tau^{(T/\tau)}$ as described above exist, we define interpolations as follows: for $k \in \{0, \ldots, T/\tau\}$, we let $\overline{y}_\tau(k\tau) = \underline{y}_\tau(k\tau) = \hat{y}_\tau(k\tau) := y_\tau^{(k)}$ and for $t \in ((k-1)\tau, k\tau)$

$$\overline{y}_\tau(t) := y_\tau^{(k)}, \quad \underline{y}_\tau(t) := y_\tau^{(k-1)}, \quad \hat{y}_\tau(t) := \frac{k\tau - t}{\tau} y_\tau^{(k-1)} + \frac{t - (k-1)\tau}{\tau} y_\tau^{(k)}. \tag{2.26}$$

A similar notation is employed for $\overline{y}_\tau$, $\underline{y}_\tau$, and $\hat{y}_\tau$. We now formulate our first main result concerning the convergence of solutions to the staggered scheme towards a weak solution of (2.16)–(2.18).

Theorem 2.3. (Staggered time-incremental scheme and convergence to solutions) *Given any $T > 0$ there exists $\tau_0 \in (0, 1]$ such that for any $\tau \in (0, \tau_0)$ the following holds:*

(i) *(Existence of the scheme) The sequences $y_\tau^{(0)}, \ldots, y_\tau^{(T/\tau)}$ and $\theta_\tau^{(0)}, \ldots, \theta_\tau^{(T/\tau)}$ satisfying (2.21), (2.22), and (2.24) exist.*

(ii) *(Convergence to solutions) There exist* $y \in L^\infty(I; \mathcal{Y}_{\mathbf{id}}) \cap H^1(I; H^1(\Omega; \mathbb{R}^d))$ *and* $\theta \in L^1(I; W^{1,1}(\Omega))$ *such that the couple* (y, θ) *is a weak solution to* (2.16)–(2.18) *in the sense of Definition* 2.2, *and up to selecting a subsequence, it holds that*

$$\hat{y}_\tau \to y \text{ in } L^\infty(I; W^{1,\infty}(\Omega; \mathbb{R}^d)) \quad \text{and} \quad \dot{\hat{y}}_\tau \to \dot{y}_\varepsilon \text{ strongly in } L^2(I; H^1(\Omega; \mathbb{R}^d)), \tag{2.27}$$

$$\hat{\theta}_\tau \to \theta \text{ in } L^s(I \times \Omega) \quad \text{and} \quad \hat{\theta}_\tau \rightharpoonup \theta \text{ weakly in } L^r(I; W^{1,r}(\Omega)) \tag{2.28}$$

as $\tau \to 0$ *for any* $r \in [1, \frac{d+2}{d+1})$ *and* $s \in [1, \frac{d+2}{d})$. *The same holds true if we replace* $\hat{y}_\tau$ *by* $\overline{y}_\tau$ *or* $\underline{y}_\tau$ *in the first part of* (2.27), *and* $\hat{\theta}_\tau$ *by* $\overline{\theta}_\tau$ *or* $\underline{\theta}_\tau$ *in* (2.28).

Let us mention that the proof shows that weak solutions satisfy a total energy balance of the form

$$\frac{d}{\mathrm{d}t}\mathcal{E}(y, \theta) = \varepsilon \int_\Omega f \cdot \dot{y} \mathrm{d}x + \varepsilon \int_{\Gamma_N} g \cdot \dot{y} \mathrm{d}\mathcal{H}^{d-1} - \kappa \int_\Gamma (\theta - \varepsilon^\alpha \theta_\flat) \mathrm{d}\mathcal{H}^{d-1},$$

i.e., the total energy is conserved up to the work of the external loadings and the heat flux through Γ.

Remark 2.4. (Difference to scheme in [33]) The scheme has several differences to the one considered in [33, (4.5)–(4.7)]. On the one hand, both steps in [33] are suitably regularized. More precisely, in (2.22) an additional dissipative term $\frac{\lambda}{2\tau} \|\nabla y - \nabla y_\tau^{(k-1)}\|^2_{L^2(\Omega)}$ is considered, where $\lambda > 0$ is a regularization parameter (called ε there), and in (2.24) the dissipation rate ξ is replaced by a smoothly truncated version $\frac{\xi}{1+\lambda\xi}$. On the other hand, the term $\partial_F W^{\mathrm{cpl}}(\nabla y_\tau^{(k-1)}, \theta_\tau^{(k-1)})\, \theta$ in (2.24)–(2.25) is replaced by the more involved term $\int_0^\theta \partial_F W^{\mathrm{cpl}}(\nabla y_\tau^{(k)}, s) \mathrm{d}s$. One of the main novelties in the present work is that the same result on existence and time-discrete approximations is achieved for the simpler, explicit, thermo-mechanical coupling term $\partial_F W^{\mathrm{cpl}}(\nabla y_\tau^{(k-1)}, \theta_\tau^{(k-1)})$ and without regularizing terms.

2.3. Passage to linearized thermoviscoelasticity

We are now interested in the passage to a small strain regime $\varepsilon \to 0$. This is induced by small external loading, boundary traction, and external temperature as $\varepsilon \to 0$, see (2.16a) and the boundary conditions in (2.17). In a similar fashion, we suppose that the initial values are small when ε is small, see (2.18). At this point, we additionally require that

(W.4) $W^{\mathrm{el}}(F) \geqq c_0 \,\mathrm{dist}^2(F, SO(d))$ for all $F \in GL^+(d)$, and $W^{\mathrm{el}}(F) = 0$ if $F \in SO(d)$;
(H.4) $H(0) = 0$;
(C.6) The *heat capacity* $c_V(F, \theta) := -\theta \partial^2_\theta W^{\mathrm{cpl}}(F, \theta)$ for $F \in GL^+(d)$ and $\theta > 0$ as well as $\partial_{F\theta} W^{\mathrm{cpl}}$ can be continuously extended to $GL^+(d) \times \mathbb{R}_+$;
(C.7) For all $F \in GL^+(d)$ and $\theta > 0$ it holds that $\partial_{FF\theta} W^{\mathrm{cpl}}(F, \theta) \leqq \frac{C_0}{\max\{\theta, 1\}}$.

In order to ensure the compatibility of (W.4) with (W.3), we assume $C_0 \geqq c_0(d+1)$ from now on. We write the equations (2.16) and the boundary conditions (2.17) equivalently in terms of the *rescaled displacement field* $u = \varepsilon^{-1}(y - \mathbf{id})$ and the *rescaled temperature* $\mu = \varepsilon^{-\alpha}\theta$. Then, for $\alpha \in [1, 2]$, rescaling the equations by ε^{-1} and $\varepsilon^{-\alpha}$, respectively, and letting $\varepsilon \to 0$ we obtain, at least formally, the system

$$\begin{aligned} - \operatorname{div}\left(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot{u}) + \mathbb{B}^{(\alpha)}\mu\right) &= f, \\ \bar{c}_V \dot{\mu} - \operatorname{div}(\mathbb{K}_0 \nabla \mu) &= \mathbb{C}_D^{(\alpha)} e(\dot{u}) : e(\dot{u}), \end{aligned} \tag{2.29}$$

along with the boundary conditions

$$\begin{aligned} &u = 0 \text{ on } I \times \Gamma_D, \\ &\quad \left(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot{u}) + \mathbb{B}^{(\alpha)}\mu\right)\nu = g \text{ on } I \times \Gamma_N, \quad \mathbb{K}_0 \nabla\mu \cdot \nu + \kappa\mu = \kappa\theta_\flat \text{ on } I \times \Gamma \end{aligned} \tag{2.30}$$

and the initial conditions

$$u(0) = u_0, \quad \mu(0) = \mu_0. \tag{2.31}$$

Here, $e(u) := \frac{1}{2}(\nabla u + (\nabla u)^T)$ denotes the linear strain tensor, and the tensors of elasticity and viscosity coefficients are defined by

$$\mathbb{C}_W := \partial_F^2 W^{\mathrm{el}}(\mathbf{Id}), \quad \mathbb{C}_D := \partial_{\dot{F}}^2 R(\mathbf{Id}, 0, 0) = 4D(\mathbf{Id}, 0). \tag{2.32}$$

Moreover, the heat conductivity tensor and the heat capacity (see also (C.6)) at zero temperature and the natural material state are given by

$$\mathbb{K}_0 := \mathbb{K}(0), \quad \bar{c}_V := c_V(\mathbf{Id}, 0). \tag{2.33}$$

Eventually, we have the α-dependent quantities

$$\mathbb{B}^{(\alpha)} = \begin{cases} \partial_{F\theta} W^{\mathrm{cpl}}(\mathbf{Id}, 0) & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha \in (1, 2] \end{cases}, \quad \mathbb{C}_D^{(\alpha)} = \begin{cases} 0 & \text{if } \alpha \in [1, 2) \\ \mathbb{C}_D & \text{if } \alpha = 2, \end{cases} \tag{2.34}$$

where $\mathbb{B}^{(\alpha)}$ plays the role of a *thermal expansion matrix*. Notice that in the formal analysis above the elasticity tensor does not depend on the coupling potential. This is due to the fact that $\partial_F^2 W^{\mathrm{cpl}}(\mathbf{Id}, 0) = 0$, see (C.3).

Although the nonlinear system is given for a nonsimple material, in the limit we obtain equations without spatial gradients of $e(u)$. This is a consequence of the growth conditions in (H.3). Moreover, there is an interesting decoupling effect due to the different scaling of coupling terms in the mechanical and the heat-transfer equation, expressed in terms of the α-dependent quantities in (2.34). This computation also shows why we restrict to the range $\alpha \in [1, 2]$. Indeed, formally, we would have $\mathbb{B}^{(\alpha)} = +\infty$ for $\alpha < 1$ while $\mathbb{C}_D^{(\alpha)} = +\infty$ for $\alpha > 2$.

The second main goal of this article is to show that the above formal linearization can be made rigorous. In the case $\alpha \in [1, 2)$, our analysis requires a regularization of the thermal evolution. More precisely, we define the k-th thermal step through

$$\text{(2.24) with } \xi \text{ replaced by } \xi_\alpha^{\mathrm{reg}}, \tag{2.24$_\varepsilon$}$$

where

$$\xi_\alpha^{\mathrm{reg}} := \begin{cases} \xi & \text{if } \xi \leqq 1, \\ \xi^{\alpha/2} & \text{else.} \end{cases} \tag{2.35}$$

Due to the different scaling of the mechanical and the heat-transfer equation, the existence of a solution to the scheme is more delicate for ε small and $\alpha \neq 2$. More specifically, we need higher integrability of W^{in} defined in (2.12) in $L^{2/\alpha}$ which can be guaranteed by the choice in (2.35). We refer to Sect. 3.3 below for details. We also emphasize that for $\alpha = 2$ no regularization is applied as $\xi_\alpha^{\mathrm{reg}} = \xi$. A similar result as Theorem 2.3 holds true in the regularized setting.

Proposition 2.5. (Vanishing time-discretization in the regularized nonlinear setting) *Given any $T > 0$ there exists $\varepsilon_0, \tau_0 \in (0, 1]$ such that for any $\tau \in (0, \tau_0)$ and $\varepsilon \in (0, \varepsilon_0)$ the following holds:*

(i) *(Existence of the scheme) The sequences $y_{\varepsilon,\tau}^{(0)}, \ldots, y_{\varepsilon,\tau}^{(T/\tau)}$ and $\theta_{\varepsilon,\tau}^{(0)}, \ldots, \theta_{\varepsilon,\tau}^{(T/\tau)}$ satisfying* (2.21), (2.22), *and (2.24$_\varepsilon$) exist.*

(ii) *(Convergence to solutions) The convergences* (2.27)–(2.28) *towards a limit $(y_\varepsilon, \theta_\varepsilon)$ hold true for the interpolations of the steps from* (i). *Here $(y_\varepsilon, \theta_\varepsilon)$ is a weak solution to the system in a sense similar to Definition* 2.2, *namely* (2.19) *is satisfied and* (2.20) *holds with ξ replaced by $\xi_\alpha^{\mathrm{reg}}$.*

We will prove that (2.29) admits a unique weak solution and that solutions of the above described regularization guaranteed by Proposition 2.5(ii) converge to the solution of (2.29) in a suitable sense. Setting

$$H^1_{\Gamma_D}(\Omega; \mathbb{R}^d) := \{u \in H^1(\Omega; \mathbb{R}^d) \colon u = 0 \text{ on } \Gamma_D\} \tag{2.36}$$

we have the following definition of weak solutions for the linearized system:

Definition 2.6. (Weak solution of the linearized system) A couple $(u, \mu) \colon I \times \Omega \to \mathbb{R}^d \times \mathbb{R}$ is called a *weak solution* to the initial-boundary-value problem (2.29)–(2.31) if $u \in H^1(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ with $u(0, \cdot) = u_0$, $\mu \in L^1(I; W^{1,1}(\Omega))$ with $\mu \geqq 0$ a.e., and if it satisfies the identities

$$\int_0^T \int_\Omega \left(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot{u}) + \mu \mathbb{B}^{(\alpha)}\right) : \nabla z \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega f \cdot z \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\Gamma_N} g \cdot z \,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \tag{2.37}$$

for any $z \in C^\infty(I \times \overline{\Omega}; \mathbb{R}^d)$ with $z = 0$ on $I \times \Gamma_D$, as well as

$$\int_0^T \int_\Omega \mathbb{K}\nabla\mu \cdot \nabla\varphi - \mathbb{C}_D^{(\alpha)} e(\dot{u}) : e(\dot{u})\varphi - \bar{c}_V \mu \dot{\varphi} \,\mathrm{d}x\,\mathrm{d}t + \kappa \int_\Gamma \mu\varphi \,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t = \kappa \int_\Gamma \theta_\flat \varphi \,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t + \bar{c}_V \int_\Omega \mu_0 \varphi(0)\,\mathrm{d}x \tag{2.38}$$

for any $\varphi \in C^\infty(I \times \overline{\Omega})$ with $\varphi(T) = 0$.

Indeed, it is a standard matter to check that sufficiently smooth weak solutions lead to the classical formulation (2.29). Next, we state the relation between time-continuous or time-discrete solutions of the nonlinear system and solutions to (2.29)–(2.31).

Theorem 2.7. (Passage to linearized thermoviscoelasticity) *Under the above assumptions we have:*

(i) *There exists a unique weak solution* (u, μ) *to* (2.29)–(2.31) *in the sense of Definition* 2.6.

(ii) *Given any sequence* $(\varepsilon_k)_k$ *converging to zero and any sequence of weak solutions* $(y_{\varepsilon_k}, \theta_{\varepsilon_k})$ *given by Proposition* 2.5 (ii), *the functions* $u_{\varepsilon_k} := \varepsilon_k^{-1}(y_{\varepsilon_k} - \mathbf{id})$ *and* $\mu_k = \varepsilon_k^{-\alpha}\theta_{\varepsilon_k}$ *satisfy*

$$u_{\varepsilon_k} \to u \text{ in } L^\infty(I; H^1(\Omega; \mathbb{R}^d)), \quad \dot{\hat{u}}_k \to \dot{u} \text{ in } L^2(I; H^1(\Omega; \mathbb{R}^d)),$$
$$\mu_{\varepsilon_k} \to \mu \text{ in } L^s(I \times \Omega), \quad \mu_{\varepsilon_k} \rightharpoonup \mu \text{ weakly in } L^r(I; W^{1,r}(\Omega))$$

for any $s \in [1, \frac{d+2}{d})$ *and* $r \in [1, \frac{d+2}{d+1})$.

(iii) *Given sequences* $(\varepsilon_k)_k$, $(\tau_k)_k$ *converging to zero and any sequence* $(\overline{y}_{\varepsilon_k,\tau_k}, \overline{\theta}_{\varepsilon_k,\tau_k})$ *of time-discrete solutions given by Proposition* 2.5 (i), $\overline{u}_k := \varepsilon_k^{-1}(\overline{y}_{\varepsilon_k,\tau_k} - \mathbf{id})$ *and* $\overline{\mu}_k = \varepsilon_k^{-\alpha}\overline{\theta}_{\varepsilon_k,\tau_k}$ *satisfy*

$$\hat{u}_k \to u \text{ in } L^\infty(I; H^1(\Omega; \mathbb{R}^d)), \quad \dot{\hat{u}}_k \to u \text{ in } L^2(I; H^1(\Omega; \mathbb{R}^d)),$$
$$\hat{\mu}_k \to \mu \text{ in } L^s(I \times \Omega), \quad \hat{\mu}_k \rightharpoonup \mu \text{ weakly in } L^r(I; W^{1,r}(\Omega))$$

for any $s \in [1, \frac{d+2}{d})$ *and* $r \in [1, \frac{d+2}{d+1})$. *Apart from the convergence of* $\dot{\hat{u}}_k$, *the same holds true if we replace* $\hat{y}_{\varepsilon_k,\tau_k}$ *by* $\overline{y}_{\varepsilon_k,\tau_k}$ *or* $\underline{y}_{\varepsilon_k,\tau_k}$ *and* $\hat{\theta}_{\varepsilon_k,\tau_k}$ *by* $\overline{\theta}_{\varepsilon_k,\tau_k}$ *or* $\underline{\theta}_{\varepsilon_k,\tau_k}$, *and consider the corresponding rescaled quantities.*

Note particularly that we obtain strong convergence of strains and strain rates. Finally, we study the relation between the time-discrete solutions in the nonlinear and the linear setting, as well as the convergence of time-discrete solutions in the linearized setting under vanishing time-discretization.

Theorem 2.8. (Passage to linearized thermoviscoelasticity, time-discrete solutions) *The following properties hold:*

(i) *Let* τ *be sufficiently small. For every* $k \in \{1, \ldots, T/\tau\}$ *we have as* $\varepsilon \to 0$

$$\frac{1}{\varepsilon}(y_{\varepsilon,\tau}^{(k)} - \mathbf{id}) \to u_\tau^{(k)} \text{ strongly in } H^1(\Omega; \mathbb{R}^d), \quad \frac{1}{\varepsilon^\alpha}\theta_{\varepsilon,\tau}^{(k)} \rightharpoonup \mu_\tau^{(k)} \text{ weakly in } W^{1,r}(\Omega) \tag{2.39}$$

for any $r \in [1, \frac{d+2}{d+1})$, *where* $u_\tau^{(k)}$ *is uniquely determined by*

$$\int_\Omega \left(\mathbb{C}_W e(u_\tau^{(k)}) + \mathbb{C}_D e(\delta_\tau u_\tau^{(k)}) + \mu_\tau^{(k-1)}\mathbb{B}^{(\alpha)}\right) : \nabla z \mathrm{d}x - \langle \ell_\tau^{(k)}, z\rangle = 0 \tag{2.40}$$

for all $z \in H^1_{\Gamma_D}(\Omega; \mathbb{R}^d)$ *and* $\mu_\tau^{(k)}$ *is uniquely determined by*

$$\int_\Omega \Big(\bar{c}_V \delta_\tau \mu_\tau^{(k)} - \mathbb{C}_D^{(\alpha)} e(\delta_\tau u_\tau^{(k)}) : e(\delta_\tau u_\tau^{(k)})\Big)\varphi \mathrm{d}x + \int_\Omega \mathbb{K}_0 \nabla \mu_\tau^{(k)} \cdot \nabla \varphi \mathrm{d}x + \kappa \int_\Gamma (\mu_\tau^{(k)} - \theta_{\flat,\tau}^{(k)})\varphi \mathrm{d}\mathcal{H}^{d-1} = 0 \tag{2.41}$$

for all $\varphi \in C^\infty(\overline{\Omega})$, *where* $\delta_\tau u_\tau^{(k)} := (u_\tau^{(k)} - u_\tau^{(k-1)})/\tau$ *and* $\delta_\tau \mu_\tau^{(k)} := (\mu_\tau^{(k)} - \mu_\tau^{(k-1)})/\tau$.

(ii) *Given* $(u_\tau^{(k)})_k$ *and* $(\mu_\tau^{(k)})_k$ *from* (i), *define* $\hat{u}_\tau$ *and* $\hat{\mu}_\tau$ *similarly to* (2.26). *Then,*

$$\hat{u}_\tau \to u \text{ in } L^\infty(I; H^1(\Omega; \mathbb{R}^d)), \quad \dot{\hat{u}}_\tau \to \dot{\hat{u}} \text{ in } L^2(I; H^1(\Omega; \mathbb{R}^d)),$$
$$\hat{\mu}_\tau \to \mu \text{ in } L^s(I \times \Omega), \quad \hat{\mu}_\tau \rightharpoonup \mu \text{ weakly in } L^r(I; W^{1,r}(\Omega)) \tag{2.42}$$

as $\tau \to 0$ *for any* $s \in [1, \frac{d+2}{d})$ *and* $r \in [1, \frac{d+2}{d+1})$, *where* (u, μ) *is the unique weak solution of* (2.29)–(2.31) *in the sense of Definition* 2.6. *Apart from the time derivative, the convergences in* (2.42) *also hold for the other interpolations.*

Remark 2.9. (Variational structure in the time-discrete linear setting) With regard to Theorem 2.8, we can in fact show that $u_\tau^{(k)}$ is the unique solution of the minimization problem

$$\operatorname{argmin}_{u \in H^1_{\Gamma_D}(\Omega;\mathbb{R}^d)} \left\{ \frac{1}{2} \int_\Omega (\mathbb{C}_W e(u) + \mu_\tau^{(k-1)} \mathbb{B}^{(\alpha)}) : e(u) \mathrm{d}x + \frac{1}{2\tau} \int_\Omega \mathbb{C}_D e(u - u_\tau^{(k-1)}) : e(u - u_\tau^{(k-1)}) \mathrm{d}x - \langle \ell_\tau^{(k)}, u \rangle \right\}$$

and for $\alpha \in [1, 2)$ that the nonnegative function $\mu_\tau^{(k)}$ is the unique solution of the minimization problem

$$\operatorname{argmin}_{\mu \in H^1(\Omega)} \left\{ \frac{\bar{c}_V}{2\tau} \int_\Omega (\mu - \mu_\tau^{(k-1)})^2 \mathrm{d}x - \int_\Omega \mathbb{C}_D^{(\alpha)} e(\delta_\tau u_\tau^{(k)}) : e(\delta_\tau u_\tau^{(k)}) \mu + \frac{1}{2} \mathbb{K}_0 \nabla \mu \cdot \nabla \mu \mathrm{d}x + \frac{\kappa}{2} \int_\Gamma (\mu - \theta_{\flat,\tau}^{(k)})^2 \mathrm{d}x \right\}. \tag{2.43}$$

From the a priori bounds in the nonlinear setting, we will be only able to prove that $\mathbb{C}_D e(\delta_\tau u_\tau^{(k)}) : e(\delta_\tau u_\tau^{(k)}) \in L^1(\Omega)$. Consequently, the functional in (2.43) might not be well-defined on $H^1(\Omega)$ for $\alpha = 2$. Nevertheless, for sufficiently smooth Γ, smooth functions f and $\theta_\flat$, and $\Gamma_D = \Gamma$, it follows by elliptic regularity theory that $\mathbb{C}_D e(\delta_\tau u_\tau^{(k)}) : e(\delta_\tau u_\tau^{(k)}) \in L^2(\Omega)$. In this case, $\mu_\tau^{(k)}$ is a minimizer of (2.43) also for $\alpha = 2$.

Section 3 is devoted to existence of the staggered time-incremental scheme leading to Theorem 2.3(i) and Proposition 2.5(i). Then, in Sect. 4 we pass to the limit $\tau \to 0$ and show Theorem 2.3(ii) and Proposition 2.5(ii). Eventually, in Sect. 5 we address the limit $\varepsilon \to 0$ and prove Theorems 2.7 and 2.8.

3. Staggered Time-Incremental Scheme

This section is devoted to the analysis of the staggered time-incremental scheme described in the previous section. Let us start with some fundamental auxiliary results.

Lemma 3.1. (A priori estimates, positivity of determinant) *Given $M > 0$ there exists a constant $C_M > 0$ such that for all $y \in \mathcal{Y}_{\mathbf{id}}$ with $\mathcal{M}(y) \leqq M$ (where $\mathcal{M}$ is defined in* (2.5)*) it holds that*

$$\|y\|_{W^{2,p}(\Omega)} \leqq C_M, \quad \|y\|_{C^{1,1-d/p}(\Omega)} \leqq C_M,$$
$$\|(\nabla y)^{-1}\|_{C^{1-d/p}(\Omega)} \leqq C_M, \quad \det(\nabla y) \geqq \frac{1}{C_M} \text{ in } \Omega. \tag{3.1}$$

If W additionally satisfies (W.4)*, there exists a universal constant C and a constant $C^*_M > 0$ with $C^*_M \to 0$ as $M \to 0$ such that*

$$\|y - \mathbf{id}\|_{H^1(\Omega)} \leqq C\|\mathrm{dist}(\nabla y, SO(d))\|_{L^2(\Omega)}, \tag{3.2}$$
$$\|y - \mathbf{id}\|_{W^{1,\infty}(\Omega)} \leqq C^*_M. \tag{3.3}$$

Proof. For a proof of the first part we refer to [33, Theorem 3.1] relying on a result in [22]. The second part can be found in [19, Lemma 4.2], where $\mathscr{S}^M_\delta$ therein simply corresponds to $\mathcal{M}(y) \leqq M\delta^2$. □

Lemma 3.2. (Generalized Korn's inequality) *Given $M > 0$ there exists a constant $c_M > 0$ such that for all $v \in H^1_{\Gamma_D}(\Omega; \mathbb{R}^d)$ and $y \in \mathcal{Y}_{\mathbf{id}}$ with $\mathcal{M}(y) \leqq M$ it holds that*

$$\int_\Omega |*|(\nabla v)^T \nabla y + (\nabla y)^T \nabla v^2 \mathrm{d}x \geqq c_M \|v\|^2_{H^1(\Omega)}.$$

Proof. The statement can be found in [33, Corollary 3.4], relying on the result in [36]. □

Lemma 3.3. (Heat conductivity) *For any $M > 0$ there exist constants c_M, $C_M > 0$ such that for $y \in \mathcal{Y}_{\mathbf{id}}$ satisfying $\mathcal{M}(y) \leqq M$ and $\theta \in L^1(\Omega)$ we have that $\mathcal{K}(\nabla y, \theta)$ is well-defined and*

$$c_M \leqq \mathcal{K}(\nabla y, \theta) \leqq C_M. \tag{3.4}$$

Proof. By Lemma 3.1 we see that $(\nabla y(x))^{-1}$ exists for every $x \in \Omega$ which shows the well-definedness of $\mathcal{K}(\nabla y, \theta)$, see (2.11). The bound in (3.4) is a direct consequence of the latter three estimates in (3.1) combined with (2.10). □

Lemma 3.4. (Estimate on coupling potential) *For all $F \in GL^+(d)$ and $\theta \geqq 0$ it holds that*

$$|\partial_F W^{\mathrm{cpl}}(F, \theta)| \leqq 2C_0(\theta \wedge 1)(1 + |F|). \tag{3.5}$$

Proof. We start by proving (3.5) for $\theta \leqq 1$. To this end, we use that $\partial_F W^{\mathrm{cpl}}(F, 0) = 0$ (see (C.3) and comments thereafter), (C.5), and apply the Fundamental Theorem of Calculus to get that

$$\begin{aligned}|\partial_F W^{\mathrm{cpl}}(F,\theta)| &= \Big|\partial_F W^{\mathrm{cpl}}(F,0) + \int_0^\theta \partial_{F\theta} W^{\mathrm{cpl}}(F,s)\mathrm{d}s\Big| \\ &\leqq \int_0^\theta |\partial_{F\theta} W^{\mathrm{cpl}}(F,s)|\mathrm{d}s \\ &\leqq C_0(1+|F|)\int_0^\theta \max\{s,1\}^{-1}\mathrm{d}s = C_0\theta(1+|F|).\end{aligned}$$

On the other hand, for $\theta \geqq 1$, we use (C.4) in the limit $\tilde{F} \to F$ to find $|\partial_F W^{\mathrm{cpl}}(F,\theta)| \leqq C_0(1+2|F|)$ for every $F \in GL^+(d)$. □

3.1. Existence of solutions to time-discretized schemes

In this subsection, we show that for $\tau \in (0, \tau_0]$ a single time step of the staggered time-discretization scheme introduced in (2.22)–(2.24) is well-defined. The parameter τ_0 in principle depends on a bound of the mechanical energy of previous deformations, but we stress that a posteriori τ_0 can be chosen independently of the step. Here, we treat the case $\alpha = 2$ and $\varepsilon \in (0, 1]$ postponing necessary adaptions for $\alpha < 2$ to Sect. 3.3 below. We assume the same set-up of Sect. 2.2. More precisely, consider initial steps $y^{(0)}_{\varepsilon,\tau} := y_{0,\varepsilon} \in \mathcal{Y}_{\mathbf{id}}$ and $\theta^{(0)}_{\varepsilon,\tau} := \theta_{0,\varepsilon} \in L^2_+(\Omega)$ with $y_{0,\varepsilon}$ and $\theta_{0,\varepsilon}$ as in (2.18). Further, let $f \in W^{1,1}(I; L^2(\Omega; \mathbb{R}^d))$, $g \in W^{1,1}(I; L^2(\Gamma_N; \mathbb{R}^d))$, $\theta_\flat \in W^{1,1}(I; L^2_+(\Gamma))$, and for each $k \in \{1, \dots, T/\tau\}$ let $\ell^{(k)}_\tau$ be as in (2.23). Suppose that we have already constructed $y^{(0)}_{\varepsilon,\tau}, \dots, y^{(k-1)}_{\varepsilon,\tau} \in \mathcal{Y}_{\mathbf{id}}$ and $\theta^{(0)}_{\varepsilon,\tau}, \dots, \theta^{(k-1)}_{\varepsilon,\tau} \in L^2_+(\Omega)$ for some $k \in \{1, \dots, T/\tau\}$. (We always add an index ε for clarification.) We first investigate the existence of the k-th mechanical step.

Proposition 3.5. (Mechanical step) *For any $M > 0$ there exists $\tau_0 \in (0, 1]$ such that if $k \in \{1, \dots, T/\tau\}$, $\tau \in (0, \tau_0)$, and $\mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) \leqq M$ the minimization problem (2.22) is well-posed, i.e.,*

$$\min_{y\in\mathcal{Y}_{\mathbf{id}}} \Big\{\mathcal{M}(y) + \mathcal{W}^{\mathrm{cpl}}\big(y, \theta^{(k-1)}_{\varepsilon,\tau}\big) + \frac{1}{\tau}\mathcal{R}(y^{(k-1)}_{\varepsilon,\tau}, y - y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) - \varepsilon\langle \ell^{(k)}_\tau, y\rangle\Big\} \tag{3.6}$$

attains a solution. Furthermore, such a minimizer $y^{(k)}_{\varepsilon,\tau}$ solves the corresponding Euler–Lagrange equation, i.e., it holds for all $z \in W^{2,p}_{\Gamma_D}(\Omega; \mathbb{R}^d)$ (see (2.2)) that

$$\begin{aligned}&\int_\Omega \big(\partial_F W(\nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) + \partial_{\dot F} R(\nabla y^{(k-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\big) : \\ &\quad \nabla z + \partial_G H(\nabla^2 y^{(k)}_{\varepsilon,\tau}) \vdots \nabla^2 z \mathrm{d}x - \varepsilon\langle \ell^{(k)}_\tau, z\rangle = 0. \end{aligned}\tag{3.7}$$

Proof. We provide the proof for the coercivity in $W^{2,p}(\Omega; \mathbb{R}^d)$. The remaining argument coincides with the one in [33, Proposition 4.1], and we only include a brief sketch for convenience of the reader. Let us shortly write $\tilde{y} := y_{\varepsilon,\tau}^{(k-1)}$ and $\tilde{\theta} := \theta_{\varepsilon,\tau}^{(k-1)}$. Let $(y_n)_n \subset \mathcal{Y}_{\mathbf{id}}$ be a minimizing sequence for the problem in (3.6). Using $\tilde{y}$ as a competitor we can, without loss of generality, assume that for all $n \in \mathbb{N}$

$$\mathcal{M}(y_n)+\mathcal{W}^{\mathrm{cpl}}(y_n,\tilde{\theta})+\frac{1}{\tau}\mathcal{R}(\tilde{y},y_n-\tilde{y},\tilde{\theta})-\varepsilon\langle \ell_\tau^{(k)},y_n\rangle \leqq \mathcal{M}(\tilde{y})+\mathcal{W}^{\mathrm{cpl}}(\tilde{y},\tilde{\theta})-\varepsilon\langle \ell_\tau^{(k)},\tilde{y}\rangle$$

and therefore

$$\mathcal{M}(y_n)+\frac{1}{\tau}\mathcal{R}(\tilde{y},y_n-\tilde{y},\tilde{\theta}) \leqq \mathcal{M}(\tilde{y})+|\mathcal{W}^{\mathrm{cpl}}(y_n,\tilde{\theta})-\mathcal{W}^{\mathrm{cpl}}(\tilde{y},\tilde{\theta})|+\varepsilon|\langle \ell_\tau^{(k)},y_n-\tilde{y}\rangle|. \tag{3.8}$$

By Lemma 3.2 and (D.2) there exists $c_M > 0$ (only depending on M) such that

$$\frac{1}{\tau}\mathcal{R}(\tilde{y},y_n-\tilde{y},\tilde{\theta}) \geqq \frac{c_M}{\tau}\int_\Omega |\nabla y_n-\nabla\tilde{y}|^2\mathrm{d}x.$$

By (3.5), the Fundamental Theorem of Calculus, Young's inequality with constant $c_M/(2\tau)$, and (W.3) we derive

$$\begin{aligned}
\left|\mathcal{W}^{\mathrm{cpl}}(y_n,\tilde{\theta})-\mathcal{W}^{\mathrm{cpl}}(\tilde{y},\tilde{\theta})\right| &\leqq 2C_0\int_\Omega(\tilde{\theta}\wedge 1)(1+|\nabla y_n|+|\nabla\tilde{y}|)|\nabla y_n-\nabla\tilde{y}|\mathrm{d}x\\
&\leqq C_M\tau\int_\Omega(\tilde{\theta}\wedge 1)^2\Big(1+2C_0c_0^{-1}+c_0^{-1}W^{\mathrm{el}}(\nabla y_n)+c_0^{-1}W^{\mathrm{el}}(\nabla\tilde{y})\Big)\mathrm{d}x\\
&\quad+\frac{c_M}{4\tau}\int_\Omega|\nabla y_n-\nabla\tilde{y}|^2\mathrm{d}x\\
&\leqq C_M\tau\big(\|\tilde{\theta}\wedge 1\|^2_{L^2(\Omega)}+\mathcal{W}^{\mathrm{el}}(y_n)+\mathcal{W}^{\mathrm{el}}(\tilde{y})\big)+\frac{c_M}{4\tau}\int_\Omega|\nabla y_n-\nabla\tilde{y}|^2\mathrm{d}x
\end{aligned}$$

for C_M sufficiently large depending on M and c_0. By using Poincaré's inequality, the trace estimate on the bulk and surface term, respectively, and Young's inequality with constant $c_M/(4C\tau\varepsilon)$ we derive that

$$\begin{aligned}
\varepsilon|\langle \ell_\tau^{(k)},y_n-\tilde{y}\rangle| &= \varepsilon\left|\int_\Omega f_\tau^k\cdot(y_n-\tilde{y})\mathrm{d}x+\int_{\Gamma_N} g_\tau^k\cdot(y_n-\tilde{y})\mathrm{d}\mathcal{H}^{d-1}\right|\\
&\leqq C\varepsilon\big(\|f_\tau^k\|_{L^2(\Omega)}+\|g_\tau^k\|_{L^2(\Gamma_N)}\big)\|\nabla y_n-\nabla\tilde{y}\|_{L^2(\Omega)}\\
&\leqq C_M\tau\varepsilon^2(\|f_\tau^k\|^2_{L^2(\Omega)}+\|g_\tau^k\|^2_{L^2(\Gamma_N)})+\frac{c_M}{4\tau}\|\nabla y_n-\nabla\tilde{y}\|^2_{L^2(\Omega)}.
\end{aligned}$$

Combining the aforementioned estimates with (3.8), and using $\mathcal{M} \geqq \mathcal{W}^{\mathrm{el}}$ we get that

$$\begin{aligned}
&(1-C_M\tau)\mathcal{M}(y_n)+\frac{c_M}{2\tau}\|\nabla y_n-\nabla\tilde{y}\|^2_{L^2(\Omega)}\\
&\quad\leqq(1+C_M\tau)\mathcal{M}(\tilde{y})+C_M\tau(\|\tilde{\theta}\wedge 1\|^2_{L^2(\Omega)}+\varepsilon^2\|f_\tau^k\|^2_{L^2(\Omega)}+\varepsilon^2\|g_\tau^k\|^2_{L^2(\Gamma_N)}).
\end{aligned} \tag{3.9}$$

For τ_0 sufficiently small such that $C_M\tau_0 \leqq 1/2$, Lemma 3.1 then shows the desired coercivity in $W^{2,p}(\Omega;\mathbb{R}^d)$. The functional is weakly lower semicontinuous on $W^{2,p}(\Omega;\mathbb{R}^d)$ by the convexity of H, see (H.1), the compact embedding

$W^{2,p}(\Omega;\mathbb{R}^d) \subset W^{1,\infty}(\Omega;\mathbb{R}^d)$, and the continuity of W^{el}, W^{cpl}, and R. This proves the existence of a minimizer.

For the derivation of the Euler–Lagrange equation, we recall the definitions in (2.4)–(2.6). The treatment of the convex term $\mathcal{H}$ is standard by (H.3) and (H.1). The Gâteaux differentiability of the other terms relies on the uniform bound on gradients and the control on the determinant, see (3.1). We refer also to [33, Proposition 3.2]. □

From the previous proof, we directly deduce

Lemma 3.6. (Bound on mechanical energy and dissipation) *For any $M > 0$ there exist constants c_M, $C_M > 0$ and $\tau_0 \in (0, 1]$ such that if $k \in \{1, \dots, T/\tau\}$, $\tau \in (0, \tau_0)$, and $\mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) \leqq M$ it holds that*

$$\begin{aligned}\mathcal{M}(y^{(k)}_{\varepsilon,\tau}) &+ c_M\tau\|\delta_\tau\nabla y^{(k)}_{\varepsilon,\tau}\|^2_{L^2(\Omega)}\\ &\leqq (1 + C_M\tau)\mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) + C_M\tau\big(\|\theta^{(k-1)}_{\varepsilon,\tau}\wedge 1\|^2_{L^2(\Omega)}\\ &+ \varepsilon^2\|f^{(k)}_\tau\|^2_{L^2(\Omega)} + \varepsilon^2\|g^{(k)}_\tau\|^2_{L^2(\Gamma_N)}\big).\end{aligned} \tag{3.10}$$

Proof. Let C_M as in (3.9). For τ_0 sufficiently small with respect to C_M we derive $\frac{1}{1-C_M\tau} \leqq 1 + 2C_M\tau$ for all $\tau \in (0, \tau_0)$. Dividing (3.9) (for $y^{(k)}_{\varepsilon,\tau}$ in place of y_n) by $1 - C_M\tau$ we get the desired estimate, up to changing the constants C_M and c_M. □

Remark 3.7. By $1\wedge s \leqq \sqrt{s}$ for $s \geqq 0$, (2.14), (2.15), by the definition below (2.23), and by a standard application of Hölder's inequality, we deduce from (3.10) that

$$\begin{aligned}&\mathcal{M}(y^{(k)}_{\varepsilon,\tau}) + c_M\tau\|\delta_\tau\nabla y^{(k)}_{\varepsilon,\tau}\|^2_{L^2(\Omega)}\\ &\quad\leqq \mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) + C_M(\tau\mathcal{E}(y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) + \varepsilon^2\|f\|^2_{L^2(I\times\Omega)} + \varepsilon^2\|g\|^2_{L^2(I\times\Gamma_N)}).\end{aligned}$$

In fact, we have $\|f^{(k)}_\tau\|^2_{L^2(\Omega)} = \tau^{-2}\int_\Omega\big|\int_{(k-1)\tau}^{k\tau} f(t,x)\mathrm{d}t\big|^2\mathrm{d}x \leqq \tau^{-1}\int_0^T\|f(t)\|^2_{L^2(\Omega)}\mathrm{d}t$ and a similar computation holds for g.

In the next lemma we discuss the well-definedness of the thermal step.

Proposition 3.8. (Thermal step) *For any $M > 0$ there exists $\tau_0 \in (0, 1]$ such that if the minimizer given by Proposition 3.5 exists, $\tau \in (0, \tau_0)$, and $\mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) \leqq M$ the minimization problem (2.24) is well-posed on $H^1_+(\Omega)$. More precisely,*

$$\begin{aligned}\mathcal{T}(\theta) := &\int_\Omega\int_0^\theta \frac{1}{\tau}\big(W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, s) - W^{\mathrm{in}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\big)\mathrm{d}s\mathrm{d}x\\ &+ \frac{1}{2}\int_\Omega \nabla\theta\cdot\mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\nabla\theta\mathrm{d}x\\ &- \int_\Omega h_\tau(y^{(k)}_{\varepsilon,\tau}, y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\theta\mathrm{d}x + \frac{\kappa}{2}\int_\Gamma(\theta - \varepsilon^2\theta^{(k)}_{\flat,\tau})^2\mathrm{d}\mathcal{H}^{d-1}\end{aligned}$$

is finite on $H^1(\Omega)$ *and attains a unique minimizer* $\theta^{(k)}_{\varepsilon,\tau}$ *on* $H^1_+(\Omega)$. *Moreover,* $\theta^{(k)}_{\varepsilon,\tau}$ *satisfies*

$$\begin{aligned}\int_\Omega \Bigg(&\frac{w^{(k)}_{\varepsilon,\tau} - w^{(k-1)}_{\varepsilon,\tau}}{\tau} - \partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \\ &\delta_\tau \nabla y^{(k)}_{\varepsilon,\tau} - \xi(\nabla y^{(k-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\Bigg)\varphi \mathrm{d}x \\ &+ \int_\Omega \mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\nabla\theta^{(k)}_{\varepsilon,\tau} \cdot \nabla\varphi \mathrm{d}x + \kappa \int_\Gamma (\theta^{(k)}_{\varepsilon,\tau} - \varepsilon^2 \theta^{(k)}_{\flat,\tau})\varphi \mathrm{d}\mathcal{H}^{d-1} = 0\end{aligned} \tag{3.11}$$

for any $\varphi \in H^1(\Omega)$, *where, for brevity,* $w^{(k-1)}_{\varepsilon,\tau} := W^{\mathrm{in}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})$ *and* $w^{(k)}_{\varepsilon,\tau} := W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k)}_{\varepsilon,\tau})$.

Remarkably, the nonnegativity constraint in the minimization problem (2.24) does *not* influence the stationarity condition (3.11). We also emphasize that in contrast to [33] we can ensure uniqueness of the minimizer. This is due to the fact that we use a simpler (explicit) thermo-mechanical coupling term in the scheme; see Remark 2.4 for details.

Proof. Step 1 (Finiteness) We start by showing that all terms of $\mathcal{T}$ are well-defined and integrable. First, by (2.14) we find that

$$\int_0^\theta W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, s)\mathrm{d}s \in [\tfrac{c_0}{2}\theta^2, \tfrac{C_0}{2}\theta^2] \tag{3.12}$$

and $\int_0^\theta w^{(k-1)}_{\varepsilon,\tau}\mathrm{d}s \leqq C_0\theta\theta^{(k-1)}_{\varepsilon,\tau}$ a.e. on Ω, which both lie in $L^1(\Omega)$ by Hölder's inequality. By Lemma 3.3, $\mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})$ is well-defined in Ω, and the corresponding term in $\mathcal{T}$ is integrable. Finally, by (3.5), (D.2), (2.9), and the second estimate in (3.1) we get that the term h_τ defined in (2.25) satisfies $h_\tau(y^{(k)}_{\varepsilon,\tau}, y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) \in L^\infty(\Omega)$, i.e., the third term is also well-defined. This completes the proof of the well-definedness of $\mathcal{T}$.

Step 2 (Existence) The functional is coercive on $H^1_+(\Omega)$ due to $\int_0^\theta W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, s)\mathrm{d}s \geqq \frac{c_0}{2}\theta^2$ by (3.12), the estimate $\nabla\theta \cdot \mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\nabla\theta \geqq c_M|\nabla\theta|^2$ by (3.4), and the fact that all other terms are either nonnegative or linear in θ. Moreover, the functional is weakly lower semicontinuous on $H^1_+(\Omega)$. To see this, we again use (3.4), the weak continuity of the trace operator in $H^1(\Omega)$, and the fact that all other bulk terms are continuous in $L^2(\Omega)$ by the reasoning in Step 1. This shows that a minimizer $\theta^{(k)}_{\varepsilon,\tau}$ exists.

Step 3 (Euler–Lagrange equation) In order to prove (3.11) for test functions $\varphi \in H^1(\Omega)$ which are not constrained to be nonnegative, we extend the minimization problem (2.24) to possibly negative functions $\theta \in H^1(\Omega)$ and we show that $\theta^{(k)}_{\varepsilon,\tau}$ minimizes $\mathcal{T}$ on $H^1(\Omega)$. To this end, recalling that $W^{\mathrm{in}}(F, 0) = 0$ for $F \in GL^+(d)$ (see below (2.12)), we continuously extend W^{in} to negative temperatures by setting

$W^{\mathrm{in}}(F,\theta) = 0$ for $\theta < 0$. It now suffices to check that there exists a constant $c_M > 0$ such that for all $\theta \in H^1(\Omega)$ it holds that

$$\mathcal{T}(\theta) \geqq \mathcal{T}(\theta^+) + \frac{c_M}{2}\|\nabla\theta^-\|^2_{L^2(\Omega)}, \tag{3.13}$$

where $\theta^- := \max\{-\theta, 0\}$ and $\theta^+ := \max\{\theta, 0\}$, i.e., $\theta = \theta^+ - \theta^-$. This guarantees that minimizers of $\mathcal{T}$ are nonnegative, and because $\mathcal{T}$ is strictly convex (to see this, use (2.13)), $\theta^{(k)}_{\varepsilon,\tau}$ is its unique minimizer on $H^1(\Omega)$. Once this is achieved, in view of (2.25) and (3.12), by taking first variations it is a standard matter to check that (3.11) holds true.

Hence, it remains to prove (3.13). First, as $\theta^{(k)}_{\flat,\tau} \geqq 0$ $\mathcal{H}^{d-1}$-a.e. on Γ, we find that

$$\int_\Gamma (\theta - \varepsilon^2\theta^{(k)}_{\flat,\tau})^2 \mathrm{d}\mathcal{H}^{d-1} \geqq \int_\Gamma (\theta^+ - \varepsilon^2\theta^{(k)}_{\flat,\tau})^2 \mathrm{d}\mathcal{H}^{d-1}. \tag{3.14}$$

Next, by using (3.4) we see that

$$\begin{aligned}\frac{1}{2}\int_\Omega \nabla\theta \cdot \mathcal{K}^{(k-1)}_{\varepsilon,\tau}\nabla\theta \mathrm{d}x &= \frac{1}{2}\int_\Omega \nabla\theta^+ \cdot \mathcal{K}^{(k-1)}_{\varepsilon,\tau}\nabla\theta^+ \mathrm{d}x + \frac{1}{2}\int_\Omega \nabla\theta^- \cdot \mathcal{K}^{(k-1)}_{\varepsilon,\tau}\nabla\theta^- \mathrm{d}x \\ &\geqq \frac{1}{2}\int_\Omega \nabla\theta^+ \cdot \mathcal{K}^{(k-1)}_{\varepsilon,\tau}\nabla\theta^+ \mathrm{d}x + \frac{c_M}{2}\int_\Omega |\nabla\theta^-|^2 \mathrm{d}x, \end{aligned} \tag{3.15}$$

where for brevity we have set $\mathcal{K}^{(k-1)}_{\varepsilon,\tau} := \mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})$. Moreover, for a.e. $x \in \Omega$ we have that

$$\int_0^{\theta(x)} W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, s)\mathrm{d}s \geqq \int_0^{\theta^+(x)} W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, s)\mathrm{d}s. \tag{3.16}$$

This follows from $W^{\mathrm{in}}(F,s) = 0$ for all $(F,s) \in GL^+(d) \times (-\infty, 0)$. Eventually, we consider the terms involving h_τ and $w^{(k-1)}_{\varepsilon,\tau}$. At this point, our argument for proving nonnegativity of the temperature is more delicate compared to [33] as we use the backward approximation $\theta^{(k-1)}_{\varepsilon,\tau}$, see Remark 2.4. By (C.2) there exists a function $\hat{W}^{\mathrm{cpl}}$ such that $W^{\mathrm{cpl}}(F,\theta) = \hat{W}^{\mathrm{cpl}}(C,\theta)$ with $C = F^TF$. Clearly, $\partial_C \hat{W}^{\mathrm{cpl}}$ is symmetric which implies with the chain rule that

$$\partial_F W^{\mathrm{cpl}}(F,\theta) = F\big(\partial_C \hat{W}^{\mathrm{cpl}}(C,\theta) + (\partial_C \hat{W}^{\mathrm{cpl}}(C,\theta))^T\big) = 2F\partial_C \hat{W}^{\mathrm{cpl}}(C,\theta). \tag{3.17}$$

By Lemma 3.1, $\nabla y^{(k-1)}_{\varepsilon,\tau}$ is invertible at every point in Ω. Hence, setting $C^{(k-1)}_{\varepsilon,\tau} := (\nabla y^{(k-1)}_{\varepsilon,\tau})^T \nabla y^{(k-1)}_{\varepsilon,\tau}$, we derive, by the second and third bound in (3.1), (3.5), (3.17), and the fact that $t \wedge 1 \leqq \sqrt{t}$ for all $t \geqq 0$, that

$$\begin{aligned}\big|\partial_C \hat{W}^{\mathrm{cpl}}(C^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\big| &= \frac{1}{2}\Big|(\nabla y^{(k-1)}_{\varepsilon,\tau})^{-1}\partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\Big| \\ &\leqq 2C_0|(\nabla y^{(k-1)}_{\varepsilon,\tau})^{-1}|(\theta^{(k-1)}_{\varepsilon,\tau} \wedge 1)(1 + |\nabla y^{(k-1)}_{\varepsilon,\tau}|) \leqq C_M\sqrt{\theta^{(k-1)}_{\varepsilon,\tau}} \end{aligned} \tag{3.18}$$

for $C_M > 0$ sufficiently large. Let us further define $\dot{C}^{(k)}_{\varepsilon,\tau} := (\delta_\tau \nabla y^{(k)}_{\varepsilon,\tau})^T \nabla y^{(k-1)}_{\varepsilon,\tau} + (\nabla y^{(k-1)}_{\varepsilon,\tau})^T \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}$. By the symmetry of $\partial_C \hat{W}^{\mathrm{cpl}}$ we have for all $F \in GL^+(d)$, $G \in \mathbb{R}^{d\times d}$, and $\theta \geqq 0$

$$F\partial_C \hat{W}^{\mathrm{cpl}}(C,\theta) : G = \partial_C \hat{W}^{\mathrm{cpl}}(C,\theta) : F^T G = \partial_C \hat{W}^{\mathrm{cpl}}(C,\theta) : G^T F,$$

where, again, $C := F^T F$. We now use this identity with $F = \nabla y^{(k-1)}_{\varepsilon,\tau}$ and $G = \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}$. By (2.14), (3.17), (3.18), and Young's inequality with constant τ it follows that

$$\begin{aligned}|\partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}| &= 2\big|\nabla y^{(k-1)}_{\varepsilon,\tau} \partial_C \hat{W}^{\mathrm{cpl}}(C^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}\big| \\ = |\partial_C \hat{W}^{\mathrm{cpl}}(C^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \dot{C}^{(k)}_{\varepsilon,\tau}| &\leqq C_M \sqrt{w^{(k-1)}_{\varepsilon,\tau}} |\dot{C}^{(k)}_{\varepsilon,\tau}| \leqq \frac{w^{(k-1)}_{\varepsilon,\tau}}{\tau} + C_M^2 \tau |\dot{C}^{(k)}_{\varepsilon,\tau}|^2.\end{aligned}$$

Choosing τ_0 sufficiently small such that $C_M^2 \tau_0 \leqq c_0$, we derive by (D.1), (D.2), and (2.9) for all $\tau \in (0, \tau_0)$ that

$$\partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau} \geqq -\frac{w^{(k-1)}_{\varepsilon,\tau}}{\tau} - \xi(\nabla y^{(k-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}). \tag{3.19}$$

This shows $\tau^{-1} w^{(k-1)}_{\varepsilon,\tau} + h_\tau(y^{(k)}_{\varepsilon,\tau}, y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) \geqq 0$ a.e. on Ω. From this we deduce that

$$\begin{aligned}&-\int_\Omega \Big(\frac{w^{(k-1)}_{\varepsilon,\tau}}{\tau} + h_\tau(y^{(k)}_{\varepsilon,\tau}, y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\Big)\theta \mathrm{d}x \\ &\quad\geqq -\int_\Omega \Big(\frac{w^{(k-1)}_{\varepsilon,\tau}}{\tau} + h_\tau(y^{(k)}_{\varepsilon,\tau}, y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})\Big)\theta^+ \mathrm{d}x.\end{aligned} \tag{3.20}$$

Combining the estimates (3.14)–(3.20) leads to (3.13), which concludes the proof. □

Remark 3.9. (Nonnegativity of temperature without dissipation rate) To derive estimate (3.19), it was essential that $\xi(F, \dot{F}, \theta) \geqq c|\dot{F}^T F + F^T \dot{F}|^q$ for some $q > 1$. The pointwise nonnegativity can still be established only under the assumption that $\xi \geqq 0$, at the expense of assuming that $\mathcal{M}(y^{(k-1)}_{\varepsilon,\tau}) \leqq \eta$ and $\mathcal{M}(y^{(k)}_{\varepsilon,\tau}) \leqq \eta$ for some η sufficiently small, and that W additionally satisfies (W.4). Indeed, in this case we can show that

$$\partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau} \geqq -\frac{w^{(k-1)}_{\varepsilon,\tau}}{\tau} \tag{3.21}$$

a.e. in Ω, which, along with $\xi \geqq 0$, implies (3.20). To see this, by (2.14), (3.3), and (3.5), we can estimate that

$$\begin{aligned}|\partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}| &\leqq 2C_0 \theta^{(k-1)}_{\varepsilon,\tau}(1 + |\nabla y^{(k-1)}_{\varepsilon,\tau}|)|\delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}| \\ &\leqq \frac{2C_0}{c_0} w^{(k-1)}_{\varepsilon,\tau}(1 + |\mathbf{Id}| + C^*_\eta)\frac{2C^*_\eta}{\tau}.\end{aligned}$$

Since $C^*_\eta \to 0$ as $\eta \to 0$, (3.21) indeed follows for η small enough. This property will be exploited in the adaptions to the case $\alpha < 2$ in Sect. 3.3 below.

For any $y^{(k)}_{\varepsilon,\tau}$ and $\theta^{(k)}_{\varepsilon,\tau}$ as given in this subsection, we define from now on

$$w^{(k)}_{\varepsilon,\tau} := W^{\mathrm{in}}(\nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k)}_{\varepsilon,\tau}).$$

3.2. Well-definedness of the scheme

For fixed time horizon $T > 0$ and time step $\tau \in (0, 1]$ small enough, we will now prove the well-definedness of the staggered time-discretization scheme described in the previous subsection. In this part, we are interested in the large-strain setting, and treat the case $\varepsilon = 1$ and $\alpha = 2$, where ξ is not regularized. For later purposes, we again include ε in the estimates. (The reader only interested in large strains, can readily set $\varepsilon = 1$.) As before, we assume for the sake of simplicity that T/τ is an integer. Although not being necessary, for convenience we suppose that (W.4) holds. At the end of the subsection, we briefly indicate the changes if (W.4) is not assumed.

We start with a bound on the total energy $\mathcal{E}$ defined in (2.15). We also need to take the work of the external forces into account. To this end, similarly to the notation in (2.23), we consider, for each $t \in I$, the functionals $\ell(t)$ on $H^1(\Omega; \mathbb{R}^d)$ defined by

$$\langle \ell(t), v\rangle := \int_\Omega f(t) \cdot v \mathrm{d}x + \int_{\Gamma_N} g(t) \cdot v \mathrm{d}\mathcal{H}^{d-1} \tag{3.22}$$

for all $v \in H^1(\Omega; \mathbb{R}^d)$. Furthermore, we define

$$C_{f,g} := \|f\|_{W^{1,1}(I;L^2(\Omega))} + \|g\|_{W^{1,1}(I;L^2(\Gamma_N))}. \tag{3.23}$$

Note that the trace estimate in $H^1(\Omega; \mathbb{R}^d)$ shows that

$$\|\ell(t)\|_{H^{-1}} \leqq C\big(\|f(t)\|_{L^2(\Omega)} + \|g(t)\|_{L^2(\Gamma_N)}\big),$$

and, hence, by the Fundamental Theorem of Calculus in $W^{1,1}(I; L^2(\Omega))$ and $W^{1,1}(I; L^2(\Gamma_N))$ we get that

$$\|\ell(t)\|_{H^{-1}} \leqq C_T C_{f,g} \tag{3.24}$$

for a constant C_T only depending on T. Given the sequences $y^{(0)}_{\varepsilon,\tau}, \ldots, y^{(k)}_{\varepsilon,\tau}$ and $\theta^{(0)}_{\varepsilon,\tau}, \ldots, \theta^{(k)}_{\varepsilon,\tau}$ for some $k \in \{1, \ldots, T/\tau\}$, as described in Sect. 3.1, we define, for $l \in \{0, \ldots, k\}$,

$$\mathcal{F}^{(l)} := \mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) - \varepsilon\langle \ell(l\tau), y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\rangle, \tag{3.25}$$

and observe the following relation between $\mathcal{F}^{(l)}$ and the total energy $\mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau})$:

Lemma 3.10. *There exists a constant $C_T > 0$ only depending on T such that for all $l \in \{0, \ldots, k\}$ with $k \in \{1, \ldots, T/\tau\}$ it holds that*

$$\varepsilon|\langle \ell(l\tau), y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\rangle| \leqq \min\{\mathcal{F}^{(l)}, \mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau})\} + \varepsilon^2 C_T C^2_{f,g},$$

with $C_{f,g}$ as defined in (3.23).

Proof. By $y^{(l)}_{\varepsilon,\tau} \in \mathcal{Y}_{\mathbf{id}}$, Poincaré's inequality, (3.2), and (W.4) we derive that

$$\|y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\|^2_{H^1(\Omega)} \leqq C\|\nabla y^{(l)}_{\varepsilon,\tau} - \mathbf{Id}\|^2_{L^2(\Omega)} \leqq C\mathcal{W}^{\mathrm{el}}(y^{(l)}_{\varepsilon,\tau}).$$

Hence, by (3.24) and Young's inequality with constant λ/ε (to be chosen below) it follows that

$$\begin{aligned}
|\langle \ell(l\tau), y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\rangle| &\leqq \|\ell(l\tau)\|_{H^{-1}}\|y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\|_{H^1(\Omega)} \\
&\leqq C_T C_{f,g}\|y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\|_{H^1(\Omega)} \\
&\leqq \frac{C_T\varepsilon}{\lambda}C^2_{f,g} + \frac{\lambda}{\varepsilon}\|y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\|^2_{H^1(\Omega)} \\
&\leqq \frac{C_T\varepsilon}{\lambda}C^2_{f,g} + C\frac{\lambda}{\varepsilon}\mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}).
\end{aligned}$$

Now, take λ small enough such that $C\lambda \leqq \frac{1}{2}$. Then, by the definition of $\mathcal{F}^{(l)}$ we discover that

$$\mathcal{F}^{(l)} = \mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) - \varepsilon\langle \ell(l\tau), y^{(l)}_{\varepsilon,\tau} - \mathbf{id}\rangle \geqq \frac{1}{2}\mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) - \varepsilon^2 C_T C^2_{f,g},$$

and the statement follows. ◻

We now proceed with the bound on the total energy. For definiteness, we set $\ell(t) = 0$ for $t \notin I$.

Lemma 3.11. (Inductive bound on the total energy) *For any $M > 0$ there exist C_M such that, if the sequences $y^{(0)}_{\varepsilon,\tau}, \dots, y^{(k)}_{\varepsilon,\tau}$ and $\theta^{(0)}_{\varepsilon,\tau}, \dots, \theta^{(k)}_{\varepsilon,\tau}$, as described in Sect. 3.1, for some $k \in \{1, \dots, T/\tau\}$ exist satisfying $\mathcal{F}^{(l)} \leqq M$ for all $l = 0, \dots, k-1$ with $\mathcal{F}^{(l)}$ defined in (3.25), it holds that*

$$\begin{aligned}
\mathcal{F}^{(k)} \leqq \mathcal{F}^{(0)} + C_M\tau V_k + \varepsilon^2 C_T(1 + C^3_{f,g}) + C\sum_{l=0}^{k}\mathcal{F}^{(l)} \\
\times \int_{(l-1)\tau}^{l\tau}\big(\|\dot\ell(t)\|_{H^{-1}} + \|\dot\ell(t+\tau)\|_{H^{-1}}\big)\mathrm{d}t \\
+ \kappa\varepsilon^2\int_0^{k\tau}\int_\Gamma \theta_\flat \,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t,
\end{aligned} \tag{3.26}$$

where C is a universal constant, C_T a constant only depending on T, and

$$V_k := \sum_{l=1}^{k}\tau\int_\Omega |\delta_\tau\nabla y^{(l)}_{\varepsilon,\tau}|^2\mathrm{d}x. \tag{3.27}$$

Proof. *Step 1* Let us fix $l \in \{1, \dots, k\}$. Using Proposition 3.5 for l in place of k, (2.9), and testing (3.7) with $z = \delta_\tau y^{(l)}_{\varepsilon,\tau}$ it follows that

$$\begin{aligned}
0 = \int_\Omega \partial_F W(\nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau}) : \delta_\tau\nabla y^{(l)}_{\varepsilon,\tau} + \partial_G H(\nabla^2 y^{(l)}_{\varepsilon,\tau}) \,\vdots\, \delta_\tau\nabla^2 y^{(l)}_{\varepsilon,\tau}\mathrm{d}x \\
+ \int_\Omega \xi(\nabla y^{(l-1)}_{\varepsilon,\tau}, \delta_\tau\nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau})\mathrm{d}x - \varepsilon\langle \ell^{(l)}_\tau, \delta_\tau y^{(l)}_{\varepsilon,\tau}\rangle.
\end{aligned} \tag{3.28}$$

Similarly, using Proposition 3.8 for l in place of k we test (3.11) with $\varphi = 1$ to obtain

$$0 = \int_\Omega \delta_\tau w_{\varepsilon,\tau}^{(l)} - \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)} \\ - \int_\Omega \xi(\nabla y_{\varepsilon,\tau}^{(l-1)}, \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)})\mathrm{d}x + \kappa \int_\Gamma (\theta_{\varepsilon,\tau}^{(l)} - \varepsilon^2 \theta_{\flat,\tau}^{(l)})\mathrm{d}\mathcal{H}^{d-1}. \quad (3.29)$$

Adding (3.28) to (3.29), multiplying by τ, and eventually summing over $l = 1, \ldots, k$ we discover that

$$\int_\Omega w_{0,\varepsilon}\mathrm{d}x = \tau \sum_{l=1}^{k} \Big(\int_\Omega \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(l)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}\mathrm{d}x + \int_\Omega \partial_G H(\nabla^2 y_{\varepsilon,\tau}^{(l)}) \vdots \delta_\tau \nabla^2 y_{\varepsilon,\tau}^{(l)}\mathrm{d}x \Big) \\ + \tau \sum_{l=1}^{k} \int_\Omega \big(\partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)}) - \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)})\big) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}\mathrm{d}x \\ - \sum_{l=1}^{k} \Big(\tau\kappa \int_\Gamma (\varepsilon^2 \theta_{\flat,\tau}^{(l)} - \theta_{\varepsilon,\tau}^{(l)})\mathrm{d}\mathcal{H}^{d-1} + \tau\varepsilon\langle \ell_\tau^{(l)}, \delta_\tau y_{\varepsilon,\tau}^{(l)}\rangle \Big) + \int_\Omega w_{\varepsilon,\tau}^{(k)}\mathrm{d}x, \quad (3.30)$$

where $w_{0,\varepsilon} := W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon})$. Here, we also used that $W = W^{\mathrm{el}} + W^{\mathrm{cpl}}$.

Step 2 We continue by bounding the first two sums on the right-hand side of (3.30) from below. By the convexity of H (see (H.1)) it follows for $l \in \{1, \ldots, k\}$ that

$$H(\nabla^2 y_{\varepsilon,\tau}^{(l-1)}) \geqq H(\nabla^2 y_{\varepsilon,\tau}^{(l)}) + \partial_G H(\nabla^2 y_{\varepsilon,\tau}^{(l)}) \vdots (\nabla^2 y_{\varepsilon,\tau}^{(l-1)} - \nabla^2 y_{\varepsilon,\tau}^{(l)}).$$

Integrating the above inequality over Ω and summing over $l = 1, \ldots, k$ leads to

$$\tau \sum_{l=1}^{k} \int_\Omega \partial_G H(\nabla^2 y_{\varepsilon,\tau}^{(l)}) \vdots \delta_\tau \nabla^2 y_{\varepsilon,\tau}^{(l)}\mathrm{d}x \geqq \mathcal{H}(y_{\varepsilon,\tau}^{(k)}) - \mathcal{H}(y_{0,\varepsilon}), \quad (3.31)$$

where we recall the notation in (2.4). By using the piecewise affine function $\hat{y}_{\varepsilon,\tau}$ introduced in (2.26), and that $\mathcal{W}^{\mathrm{el}}$ is Gateaux differentiable (see [33, Proposition 3.2]) we get that

$$\sum_{l=1}^{k} \int_{(l-1)\tau}^{l\tau} \int_\Omega \partial_F W^{\mathrm{el}}(\nabla \hat{y}_{\varepsilon,\tau}(t)) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}\mathrm{d}x\mathrm{d}t \\ = \int_0^{k\tau} \int_\Omega \partial_F W^{\mathrm{el}}(\nabla \hat{y}_{\varepsilon,\tau}(t)) : \nabla \dot{\hat{y}}_{\varepsilon,\tau}(t)\mathrm{d}x\mathrm{d}t = \mathcal{W}^{\mathrm{el}}(y_{\varepsilon,\tau}^{(k)}) - \mathcal{W}^{\mathrm{el}}(y_{0,\varepsilon}). \quad (3.32)$$

For τ_0 sufficiently small, we can apply Lemma 3.6 in the version of Remark 3.7. This along with Lemma 3.10, $\mathcal{F}^{(l)} \leqq M$ for $l \in \{0, \ldots, k-1\}$, and (3.23) implies for all $l \in \{1, \ldots, k\}$ that

$$\begin{aligned}\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) &\leqq \mathcal{M}(y_{\varepsilon,\tau}^{(l-1)}) + C_M\big(\mathcal{E}(y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)}) + \varepsilon^2\|f\|^2_{L^2(I\times\Omega)} + \varepsilon^2\|g\|^2_{L^2(I\times\Gamma_N)}\big)\\ &\leqq 2(1+C_M)\mathcal{F}^{(l-1)} + C_M\varepsilon^2 C_T C_{f,g}^2 \leqq 2(1+C_M)M + C_M\varepsilon^2 C_T C_{f,g}^2.\end{aligned}$$

Together with Lemma 3.1 we get that there exists a compact convex set K, only depending on M, T, f, and g, such that $\nabla y_{\varepsilon,\tau}^{(l)} \in K$ a.e. on Ω for all $l \in \{0, \dots, k\}$. Then, by the regularity of W^{el}, setting $C_M := \sup_{F\in K}|\partial_{FF}W^{\mathrm{el}}(F)|$, we can estimate for any $t \in [(l-1)\tau, l\tau]$ with $l \in \{1, \dots, k\}$ that

$$\begin{aligned}&|\partial_F W^{\mathrm{el}}(\nabla\hat{y}_{\varepsilon,\tau}(t)) - \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(l)})|\\ &\quad\leqq C_M|\nabla\hat{y}_{\varepsilon,\tau}(t) - \nabla y_{\varepsilon,\tau}^{(l)}| = C_M\frac{l\tau - t}{\tau}|\nabla y_{\varepsilon,\tau}^{(l)} - \nabla y_{\varepsilon,\tau}^{(l-1)}| \leqq C_M|\nabla y_{\varepsilon,\tau}^{(l)} - \nabla y_{\varepsilon,\tau}^{(l-1)}|.\end{aligned}$$

Consequently, we get

$$\begin{aligned}&\sum_{l=1}^k\left|\int_{(l-1)\tau}^{l\tau}\int_\Omega \partial_F W^{\mathrm{el}}(\nabla\hat{y}_{\varepsilon,\tau}(t)) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}\,\mathrm{d}x\mathrm{d}t - \tau\int_\Omega \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(l)}) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}\,\mathrm{d}x\right|\\ &\quad\leqq C_M\sum_{l=1}^k\tau\int_\Omega|\nabla y_{\varepsilon,\tau}^{(l)} - \nabla y_{\varepsilon,\tau}^{(l-1)}||\delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}|\,\mathrm{d}x = C_M\tau\sum_{l=1}^k\tau\int_\Omega|\delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}|^2\mathrm{d}x = C_M\tau V_k.\end{aligned}$$

Combined with (3.32), this leads to

$$\tau\sum_{l=1}^k\int_\Omega \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(l)}) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}\,\mathrm{d}x \geqq \mathcal{W}^{\mathrm{el}}(y_{\varepsilon,\tau}^{(k)}) - \mathcal{W}^{\mathrm{el}}(y_{0,\varepsilon}) - C_M\tau V_k. \tag{3.33}$$

In a similar fashion, using the first bound in (C.5), we can estimate

$$\tau\sum_{l=1}^k\int_\Omega\big(\partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)}) - \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)})\big) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}\,\mathrm{d}x \geqq -C_0\tau V_k. \tag{3.34}$$

Now, employing (3.31), (3.33), and (3.34) in (3.30), and using the definition of the total energy $\mathcal{E}$ we conclude that

$$\mathcal{E}(y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) \leqq \mathcal{E}(y_{0,\varepsilon}, \theta_{0,\varepsilon}) + C_M\tau V_k + \sum_{l=1}^k\tau\varepsilon\langle\ell_\tau^{(l)}, \delta_\tau y_{\varepsilon,\tau}^{(l)}\rangle + \sum_{l=1}^k\tau\kappa\int_\Gamma(\varepsilon^2\theta_{\flat,\tau}^{(l)} - \theta_{\varepsilon,\tau}^{(l)})\mathrm{d}\mathcal{H}^{d-1}. \tag{3.35}$$

Step 3: It remains to estimate the last two terms on the right-hand side of (3.35). By the nonnegativity of $\theta_{\varepsilon,\tau}^{(l)}$ and the definition of $\theta_{\flat,\tau}^{(l)}$ below (2.25) we can bound

$$\sum_{l=1}^k\tau\kappa\int_\Gamma(\varepsilon^2\theta_{\flat,\tau}^{(l)} - \theta_{\varepsilon,\tau}^{(l)})\mathrm{d}\mathcal{H}^{d-1} \leqq \sum_{l=1}^k\tau\kappa\varepsilon^2\int_\Gamma\theta_{\flat,\tau}^{(l)}\mathrm{d}\mathcal{H}^{d-1} = \kappa\varepsilon^2\int_0^{k\tau}\int_\Gamma\theta_\flat\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t. \tag{3.36}$$

Note that for any $l \in \{1, \ldots, k\}$ and $t \in ((l-1)\tau, l\tau)$ we have that $\delta_\tau y^{(l)}_{\varepsilon,\tau} = \dot{\hat{y}}_{\varepsilon,\tau}(t)$. Consequently, integration by parts yields

$$\sum_{l=1}^{k} \tau \langle \ell^{(l)}_\tau, \delta_\tau y^{(l)}_{\varepsilon,\tau} \rangle = \int_0^{k\tau} \langle \ell(t), \dot{\hat{y}}_{\varepsilon,\tau}(t) \rangle \mathrm{d}t \tag{3.37}$$

$$= \langle \ell(k\tau), \hat{y}_{\varepsilon,\tau}(k\tau) - \mathbf{id} \rangle - \langle \ell(0), \hat{y}_{\varepsilon,\tau}(0) - \mathbf{id} \rangle - \int_0^{k\tau} \langle \dot{\ell}(t), \hat{y}_{\varepsilon,\tau}(t) - \mathbf{id} \rangle \mathrm{d}t$$

$$\leqq \langle \ell(k\tau), \hat{y}_{\varepsilon,\tau}(k\tau) - \mathbf{id} \rangle - \langle \ell(0), \hat{y}_{\varepsilon,\tau}(0) - \mathbf{id} \rangle$$

$$+ \int_0^{k\tau} \|\dot{\ell}(t)\|_{H^{-1}} \|\hat{y}_{\varepsilon,\tau}(t) - \mathbf{id}\|_{H^1(\Omega)} \mathrm{d}t. \tag{3.38}$$

By Poincaré's inequality, (3.2), and (W.4), for $t \in [(l-1)\tau, l\tau]$, we have that

$$\|\hat{y}_{\varepsilon,\tau}(t) - \mathbf{id}\|^2_{H^1(\Omega)} \leqq 2(\|y^{(l-1)}_{\varepsilon,\tau} - \mathbf{id}\|^2_{H^1(\Omega)} + \|y^{(l)}_{\varepsilon,\tau}$$
$$-\mathbf{id}\|^2_{H^1(\Omega)}) \leqq C\big(\mathcal{W}^{\mathrm{el}}(y^{(l-1)}_{\varepsilon,\tau}) + \mathcal{W}^{\mathrm{el}}(y^{(l)}_{\varepsilon,\tau})\big).$$

Therefore, by Lemma 3.10, (3.24), and $\sqrt{s} \leqq s/\varepsilon$ for all $s \geqq \varepsilon^2$ we get that

$$\int_0^{k\tau} \|\dot{\ell}(t)\|_{H^{-1}} \|\hat{y}_{\varepsilon,\tau}(t) - \mathbf{id}\|_{H^1(\Omega)} \mathrm{d}t$$
$$\leqq C \sum_{l=1}^{k} \Big(\varepsilon + \varepsilon^{-1} \mathcal{E}(y^{(l-1)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau}) + \varepsilon^{-1} \mathcal{E}(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) \Big) \int_{(l-1)\tau}^{l\tau} \|\dot{\ell}(t)\|_{H^{-1}} \mathrm{d}t$$
$$\leqq \frac{C}{\varepsilon} \sum_{l=1}^{k} \Big(\big(\mathcal{F}^{(l-1)} + \mathcal{F}^{(l)}\big) \int_{(l-1)\tau}^{l\tau} \|\dot{\ell}(t)\|_{H^{-1}} \mathrm{d}t \Big) + \varepsilon C_T (C_{f,g} + C^3_{f,g}).$$

Then, using an index shift and $C_{f,g} \leqq \frac{2}{3} + \frac{1}{3} C^3_{f,g}$ we get that

$$\int_0^{k\tau} \|\dot{\ell}(t)\|_{H^{-1}} \|\hat{y}_\tau(t) - \mathbf{id}\|_{H^1(\Omega)} \mathrm{d}t \tag{3.39}$$

$$\leqq \frac{C}{\varepsilon} \sum_{l=0}^{k} \Big(\mathcal{F}^{(l)} \int_{(l-1)\tau}^{l\tau} \big(\|\dot{\ell}(t)\|_{H^{-1}} + \|\dot{\ell}(t+\tau)\|_{H^{-1}} \big) \mathrm{d}t \Big) + \varepsilon C_T (1 + C^3_{f,g}) \tag{3.40}$$

for a possibly larger $C_T > 0$. We plug this into (3.37) and use (3.36) to estimate the terms on the right-hand side of (3.35), which by (3.25) concludes the proof. □

We proceed with a bound on the (discrete) strain rates V_k defined in (3.27).

Lemma 3.12. (Inductive bound on the strain rates) *Given M, $T > 0$, there exist a constant C_M and $\tau_0 \in (0, 1]$ only depending on M, and a constant C_T only depending on T such that for $\tau \in (0, \tau_0)$ the following holds: Suppose that there* *exist the sequences* $y_{\varepsilon,\tau}^{(0)}, \dots, y_{\varepsilon,\tau}^{(k)}$ *and* $\theta_{\varepsilon,\tau}^{(0)}, \dots, \theta_{\varepsilon,\tau}^{(k)}$ *for some* $k \in \{1, \dots, T/\tau\}$*, as described in Sect.* 3.1*, with* $\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) \leqq M$ *for all* $l \in \{0, \dots, k-1\}$*. Then,*

$$\sum_{l=1}^{k} \tau \int_{\Omega} |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x \leqq C_M \mathcal{M}(y_{0,\varepsilon}) + \varepsilon^2 C_M C_T C_{f,g}^2 + C_M \tau \sum_{l=0}^{k-1} \big(\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) + \|\theta_{\varepsilon,\tau}^{(l)} \wedge 1\|_{L^2(\Omega)}^2\big). \tag{3.41}$$

Proof. By Lemma 3.6 there exist constants c_M, $C_M > 0$ depending on M such that we have for $l \in \{1, \dots, k\}$

$$\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) + c_M \tau \int_{\Omega} |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x \leqq (1 + C_M \tau) \mathcal{M}(y_{\varepsilon,\tau}^{(l-1)})$$
$$+ C_M \tau \big(\|\theta_{\varepsilon,\tau}^{(l)} \wedge 1\|_{L^2(\Omega)}^2 + \varepsilon^2 \|f_\tau^{(l)}\|_{L^2(\Omega)}^2 + \varepsilon^2 \|g_\tau^{(l)}\|_{L^2(\Gamma_N)}^2\big).$$

Summing the above inequality over $l = 1, \dots, k$ and recalling the definition of $f_\tau^{(l)}$, $g_\tau^{(l)}$ below (2.23), we arrive at

$$\mathcal{M}(y_{\varepsilon,\tau}^{(k)}) - \mathcal{M}(y_{0,\varepsilon}) + c_M \sum_{l=1}^{k} \tau \int_{\Omega} |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x$$
$$= C_M \varepsilon^2 \int_0^{k\tau} \big(\|f(t)\|_{L^2(\Omega)}^2 + \|g(t)\|_{L^2(\Gamma_N)}^2\big) \mathrm{d}t$$
$$+ C_M \tau \sum_{l=0}^{k-1} \big(\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) + \|\theta_{\varepsilon,\tau}^{(l)} \wedge 1\|_{L^2(\Omega)}^2\big).$$

As $\mathcal{M}(y_{\varepsilon,\tau}^{(k)}) \geqq 0$, we conclude the proof by (3.24). □

We are ready to prove the well-definedness of our time-discretization scheme, i.e., Theorem 2.3(i). At the same time, we will also derive two useful a priori bounds, namely on the total energy and on the (discrete) strain rate, respectively.

Theorem 3.13. (Well-definedness of the scheme) *For any* $T > 0$ *there exist a constant* $\bar{C}_T > 0$*, corresponding constants*

$$M' := 2e^{\bar{C}_T C_{f,g}} \Big(\varepsilon^{-2} \mathcal{F}^{(0)} + \bar{C}_T (1 + C_{f,g}^3) + \kappa \int_0^T \int_\Gamma \theta_\flat \mathrm{d}\mathcal{H}^{d-1} \mathrm{d}t\Big),$$
$$M := 2M' + \bar{C}_T C_{f,g}^2, \tag{3.42}$$

as well as constants $C_M > 0$ *and* $\tau_0 \in (0, 1]$ *depending on* M *such that the following holds true: for each* $\tau \in (0, \tau_0)$ *such that* $T/\tau \in \mathbb{N}$ *the sequences* $y_{\varepsilon,\tau}^{(0)}, \dots, y_{\varepsilon,\tau}^{(T/\tau)}$ *and* $\theta_{\varepsilon,\tau}^{(0)}, \dots, \theta_{\varepsilon,\tau}^{(T/\tau)}$ *as described in Sect.* 3.1 *exist, and for all* $k \in \{0, \dots, T/\tau\}$ *we have that*

$$\mathcal{E}(y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) \leqq \varepsilon^2 M, \tag{3.43}$$

$$\sum_{l=1}^{k} \tau \int_{\Omega} |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x \leqq \varepsilon^2 C_M M (1 + T) + \varepsilon^2 C_M \bar{C}_T C_{f,g}^2. \tag{3.44}$$

Proof. Step 1 Let C_T be the maximum of the constants C_T from Lemmas 3.10, 3.11, 3.12, and equation (3.24), and let C be the universal constant of Lemma 3.11. Define $\bar{C}_T = \max\{2CTC_T, 2C_T, 2\}$, and let M' and M be as in (3.42). Then, let $C_M > 0$ be the maximum of the constants C_M from Lemmas 3.11 and 3.12. Moreover, let $\tau_0 \in (0, 1]$ be chosen sufficiently small so that Lemma 3.5, Lemma 3.6, Proposition 3.8, Lemma 3.11, and Lemma 3.12 hold true (all applied for M from (3.42)). In place of (3.43), we focus on showing

$$\mathcal{F}^{(k)} = \mathcal{E}(y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) - \varepsilon\langle \ell(k\tau), y_{\varepsilon,\tau}^{(k)}\rangle \leqq \varepsilon^2 M', \tag{3.45}$$

as then (3.43) follows directly by Lemma 3.10.

We will prove the statement by induction over K. In the base case $K = 0$, (3.45) is satisfied by our choice of M', and the fact that $y_{\varepsilon,\tau}^{(0)} = y_{0,\varepsilon}$ and $\theta_{\varepsilon,\tau}^{(0)} = \theta_{0,\varepsilon}$. Given $K \in \{1, \ldots, T/\tau\}$, let us assume that the statement as well as (3.45) hold true for $K - 1$. We now show that the statement holds true for K. Applying first Proposition 3.5 and then Proposition 3.8 we see that $y_{\varepsilon,\tau}^{(k)}$ and $\theta_{\varepsilon,\tau}^{(k)}$ exist, where for both propositions we use the induction hypothesis (3.43) for $K - 1$.

Step 2 In this step, we prove that for τ_0 small enough we have that

$$C_M \tau V_K \leqq \varepsilon^2, \tag{3.46}$$

where V_K is defined in (3.27). By $\varepsilon^2 M \leqq M$, Remark 3.7, and (3.23) there exists a constant $\tilde{C}_M$ only depending on M such that

$$\frac{1}{\tilde{C}_M \tau} \|\nabla y_{\varepsilon,\tau}^{(K)} - \nabla y_{\varepsilon,\tau}^{(K-1)}\|^2_{L^2(\Omega)} \leqq \varepsilon^2 M + \varepsilon^2 \tilde{C}_M (M + T C_T^2 C_{f,g}^2),$$

where we again used the hypothesis (3.43) for $K - 1$. Hence, by possibly further decreasing τ_0 (depending only on M, f, g, T, and the initial values) we can ensure that

$$C_M \|\nabla y_{\varepsilon,\tau}^{(K)} - \nabla y_{\varepsilon,\tau}^{(K-1)}\|^2_{L^2(\Omega)} \leqq \frac{\varepsilon^2}{2}.$$

Furthermore, by possibly decreasing τ_0 (depending only on M, u_0, μ_0, f, g, and T) and using the hypothesis (3.44) for $K - 1$ in place of K we get $C_M \tau V_{K-1} \leqq \varepsilon^2/2$. Consequently, combining the previous estimates and using $\tau V_K = \tau V_{K-1} + \|\nabla y_{\varepsilon,\tau}^{(K)} - \nabla y_{\varepsilon,\tau}^{(K-1)}\|^2_{L^2(\Omega)}$, the desired bound (3.46) follows.

Step 3 By hypothesis the energy bound in (3.45) is satisfied for $k \in \{0, \ldots, K - 1\}$. Consequently, Lemma 3.11 applies for any $k \in \{0, \ldots, K\}$. By (3.46) we have that

$$\begin{aligned}\mathcal{F}^{(k)} &\leqq \mathcal{F}^{(0)} + \varepsilon^2 + \varepsilon^2 C_T (1 + C_{f,g}^3) + C \sum_{l=0}^{k} \mathcal{F}^{(l)} \\ &\quad \times \int_{(l-1)\tau}^{l\tau} (\|\dot{\ell}(t)\|_{H^{-1}} + \|\dot{\ell}(t+\tau)\|_{H^{-1}}) \mathrm{d}t + \kappa\varepsilon^2 \int_0^T \int_\Gamma \theta_\flat \mathrm{d}\mathcal{H}^{d-1} \mathrm{d}t.\end{aligned}$$

We now use the following discrete version of Grönwall's Lemma: if $\beta > 0$, $(a_l)_l$ is a nonnegative sequence, $(b_l)_l \subset (0, 1/2)$, and

$$a_k \leqq \beta + \sum_{l=0}^{k} b_l a_l \quad \text{for } k \geqq 0,$$

then

$$a_k \leqq 2\beta \exp\Big(\sum_{l=0}^{k-1} 2b_l\Big) \quad \text{for } k \geqq 0.$$

Indeed, as $b_l \leqq 1/2$, we get $a_k \leqq 2\beta + \sum_{l=0}^{k-1} 2b_l a_l$, and then the statement follows from the elementary discrete Grönwall inequality. We apply this result for

$$\begin{aligned} &\beta := \mathcal{F}^{(0)} + \varepsilon^2 + \varepsilon^2 C_T(1 + C_{f,g}^3) + \kappa\varepsilon^2 \int_0^T \int_\Gamma \theta_\flat \, \mathrm{d}\mathcal{H}^{d-1} \mathrm{d}t, \\ &a_l := \mathcal{F}^{(l)}, \qquad b_l := C \int_{(l-1)\tau}^{l\tau} \big(\|\dot{\ell}(t)\|_{H^{-1}} + \|\dot{\ell}(t+\tau)\|_{H^{-1}}\big) \mathrm{d}t, \end{aligned} \tag{3.47}$$

where we note that $b_l \leqq 1/2$ for all l, provided that τ_0 is chosen small enough depending on f and g. In view of (3.24), we then see that (3.45) for K is true. Finally, (3.44) directly follows from the application of Lemma 3.12 and the fact that the last term in (3.41) can be controlled by $\mathcal{E}$, see e.g. Remark 3.7. □

Eventually, if (W.4) is not assumed, we get additional additive constants in Lemma 3.10 and in the derivation of (3.40), leading to an additional constant in (3.26) which however does not scale as ε^2. This does influence the proof of the well-definedness, only the scaling of the energy in terms of ε.

3.3. Adaptions for exponents $\alpha < 2$

In this subsection, we prove Proposition 2.5(i). This part can be skipped by a reader only interested in the proof of Theorem 2.3. In the previous subsection, we have already established the well-definedness of the scheme in the large-strain setting, as well as the energy bound (3.43). The latter will be essential to obtain a priori bounds for the limit passage $\tau \to 0$ in Sect. 4. In the case $\alpha < 2$, for the passage to the linearized setting $\varepsilon \to 0$, however, the bound (3.43) and the induced a priori bounds are not expedient. This is due to the different scaling of the internal and mechanical energy, being of order ε^α and ε^2, respectively. To this end, it is necessary to establish energy bounds for rescaled versions of the energy functionals from Sect. 2.1, namely $\mathcal{M}_\varepsilon := \frac{1}{\varepsilon^2}\mathcal{M}$, $\mathcal{W}_\varepsilon^{\mathrm{cpl}} := \frac{1}{\varepsilon^2}\mathcal{W}^{\mathrm{cpl}}$, and for $\alpha \in [1, 2]$,

$$\mathcal{E}_\varepsilon(y, \theta) := \mathcal{M}_\varepsilon(y) + \frac{\alpha}{2\varepsilon^2} \int_\Omega W^{\mathrm{in}}(\nabla y, \theta)^{\frac{2}{\alpha}} \mathrm{d}x, \tag{3.48}$$

where both 'types of energy' are of the same order. Controlling this energy is more delicate compared to Proposition 3.11, as the mechanical and thermal equation

(3.28)–(3.29) scale with different powers of ε and cannot simply be added up. Therefore, novel ideas are required to control the contributions of W^{cpl} and ξ. To achieve this, higher integrability of W^{in} in $L^{2/\alpha}$ is needed which can be guaranteed by using the regularization of $\xi_\alpha^{\mathrm{reg}}$ introduced in (2.35). This in turn induces new challenges for the analysis of the time-discrete scheme since showing the nonnegativity of the temperature in the thermal step, see Proposition 3.8, is more delicate. For this, it will be essential to assume that strains are small, i.e., we suppose that the parameter $\varepsilon \in (0, 1]$ is sufficiently small.

Note that for the entire subsection we can assume that $\alpha \in [1, 2)$ since in the case $\alpha = 2$ there is no regularization of the dissipation rate, the existence of the scheme is already guaranteed by Theorem 2.3(i), and also an energy bound for $\mathcal{E}_\varepsilon$ follows already from (3.43). The mechanical step is not affected by the regularization, but Proposition 3.8 needs to be adapted.

Proposition 3.14. (Thermal step with regularization) *For any $M > 0$ there exists $\varepsilon_0 > 0$ such that if $\varepsilon \in (0, \varepsilon_0)$, if the minimizer $y_{\varepsilon,\tau}^{(k)}$ given in Proposition 3.5 exists, and if $\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(k-1)}) \leqq M$ and $\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(k)}) \leqq M$ the minimization problem (2.24$_\varepsilon$) attains a unique solution $\theta_{\varepsilon,\tau}^{(k)}$ satisfying (3.11) for all $\varphi \in H^1(\Omega)$ with ξ replaced by $\xi_\alpha^{\mathrm{reg}}$.*

Proof. As $\xi \geqq \xi_\alpha^{\mathrm{reg}}$, the existence and uniqueness of $\theta_{\varepsilon,\tau}^{(k)}$ follows by the same reasoning as in Steps 1–2 of the proof of Proposition 3.8. Since $\xi_\alpha^{\mathrm{reg}} \geqq 0$, the nonnegativity of the temperature follows by Remark 3.9 for ε_0 sufficiently small, where we use $\mathcal{M}(y_{\varepsilon,\tau}^{(k-1)}) \leqq M\varepsilon^2$ and $\mathcal{M}(y_{\varepsilon,\tau}^{(k)}) \leqq M\varepsilon^2$. □

Our next goal is to adapt Proposition 3.11 to the present setting. As a preparation, supposing that for $k \in \{0, \dots, T/\tau\}$ the steps $y_{\varepsilon,\tau}^{(k)}$ and $\theta_{\varepsilon,\tau}^{(k)}$ exist, we define

$$\mathcal{F}_\varepsilon^{(k)} := \mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) - \varepsilon^{-1}\langle \ell(k\tau), y_{\varepsilon,\tau}^{(k)} - \mathbf{id}\rangle, \tag{3.49}$$

where ℓ is defined in (3.22). By repeating the proof of Lemma 3.10 we find that

$$\varepsilon^{-1}|\langle \ell(k\tau), y_{\varepsilon,\tau}^{(k)} - \mathbf{id}\rangle| \leqq \min\{\mathcal{F}_\varepsilon^{(k)}, \mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)})\} + C_T C_{f,g}^2 \tag{3.50}$$

for $k \in \{0, \dots, T/\tau\}$, for a constant $C_T > 0$ only depending on T and $C_{f,g}$ as in (3.23).

Lemma 3.15. (Inductive bound on the rescaled total energy) *There exists $\tau_0 \in (0, 1]$ and, given $M > 0$, $\varepsilon_0 \in (0, 1]$ such that the following holds true: suppose that for $\tau \in (0, \tau_0)$, $\varepsilon \in (0, \varepsilon_0)$, and $k \in \{1, \dots, T/\tau\}$ the steps $y_{\varepsilon,\tau}^{(0)}, \dots, y_{\varepsilon,\tau}^{(k)}$ and $\theta_{\varepsilon,\tau}^{(0)}, \dots, \theta_{\varepsilon,\tau}^{(k)}$ exist such that $\mathcal{F}_\varepsilon^{(l)} \leqq M$ for all $l \in \{0, \dots, k-1\}$. Then, for a a universal constant C and a constant C_T possibly depending on T it holds that*

$$\begin{aligned}\mathcal{F}_\varepsilon^{(k)} \leqq C\Big(&\mathcal{F}_\varepsilon^{(0)} + \sum_{l=0}^{k} \mathcal{F}_\varepsilon^{(l)} \int_{(l-1)\tau}^{l\tau} \big(1 + \|\dot{\ell}(t)\|_{H^{-1}} + \|\dot{\ell}(t+\tau)\|_{H^{-1}}\big)\mathrm{d}t \\ &+ \kappa \int_0^{k\tau} \int_\Gamma \theta_\flat^2 \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t\Big) \\ &+ C_T(1 + C_{f,g}^3).\end{aligned}$$

Proof. As a preliminary step, we show that the assumption $\mathcal{F}_\varepsilon^{(l)} \leqq M$ for all $l \in \{0, \ldots, k-1\}$ implies bounds on the rescaled mechanical energy for all $l \in \{0, \ldots, k\}$. First, by (3.50) we get for $l \in \{0, \ldots, k-1\}$ that

$$\mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l)}) = \mathcal{F}_\varepsilon^{(l)} + \varepsilon^{-1}\langle \ell(k\tau), y_{\varepsilon,\tau}^{(k)} - \mathbf{id}\rangle \leqq 2\mathcal{F}_\varepsilon^{(l)} + C_T C_{f,g}^2 \leqq 2M + C_T C_{f,g}^2. \tag{3.51}$$

Consequently, we can choose ε_0 sufficiently small such that $\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) \leqq 1$ for $l \in \{0, \ldots, k-1\}$. Then, we apply (3.10) for $M = 1$ to get τ_0 such that for $\tau \in (0, \tau_0]$ it holds that

$$\begin{aligned}\mathcal{M}(y_{\varepsilon,\tau}^{(k)}) \leqq (1 + C_1\tau)\mathcal{M}(y_{\varepsilon,\tau}^{(k-1)}) + C_1\tau\big(&\|\theta_{\varepsilon,\tau}^{(k-1)} \wedge 1\|_{L^2(\Omega)}^2 \\ &+ \varepsilon^2\|f_\tau^{(k)}\|_{L^2(\Omega)}^2 + \varepsilon^2\|g_\tau^{(k)}\|_{L^2(\Gamma)}^2\big),\end{aligned}$$

where C_1 is a universal constant. By (2.14) and the fact that $1 \wedge t \leqq t^{1/\alpha}$ for $t \geqq 0$ we find that

$$\|\theta_{\varepsilon,\tau}^{(k-1)} \wedge 1\|_{L^2(\Omega)}^2 \leqq C_0\|w_{\varepsilon,\tau}^{(k-1)}\|_{L^{\frac{2}{\alpha}}(\Omega)}^{\frac{2}{\alpha}}, \tag{3.52}$$

such that, dividing the above estimate by ε^2 and recalling (3.48) as well as Remark 3.7 we get that

$$\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(k)}) \leqq (1 + C_1\tau)\mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) + C_1\big(\|f\|_{L^2(I\times\Omega)}^2 + \|g\|_{L^2(I\times\Gamma_N)}^2\big).$$

This along with (3.51) shows that, possibly decreasing ε_0, we have $\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(l)}) \leqq 1$ for all $l \in \{0, \ldots, k\}$. This induces that in the following proof the constants coming from Lemmas 3.1, 3.6, and 3.12 are universal and denoted by C_1.

As in the proof of Proposition 3.11, the strategy relies on a suitable test of the mechanical and the thermal equation, see also (3.28)–(3.29). In contrast, however, the resulting equations cannot be summed up, but have to be treated separately. This will allow us to show the estimates

$$\begin{aligned}&\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(k)}) - \varepsilon^{-1}\langle \ell(k\tau), y_{\varepsilon,\tau}^{(k)} - \mathbf{id}\rangle + \frac{\tau}{\varepsilon^2}\sum_{l=1}^{k}\int_\Omega \xi(\nabla y_{\varepsilon,\tau}^{(l-1)}, \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)})\,\mathrm{d}x \\ &\quad\leqq C\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(0)}) + C_T(1 + C_{f,g}^3) \\ &\qquad + C\sum_{l=0}^{k}\mathcal{F}_\varepsilon^{(l)}\int_{(l-1)\tau}^{l\tau}\big(1 + \|\dot{\ell}(t)\|_{H^{-1}} + \|\dot{\ell}(t+\tau)\|_{H^{-1}}\big)\,\mathrm{d}t,\end{aligned} \tag{3.53}$$

and

$$\begin{aligned}&\frac{\alpha}{2\varepsilon^2}\int_\Omega (w_{\varepsilon,\tau}^{(k)})^{\frac{2}{\alpha}}\,\mathrm{d}x - \frac{\tau}{\varepsilon^2}\sum_{l=1}^{k}\int_\Omega \xi(\nabla y_{\varepsilon,\tau}^{(l-1)}, \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)})\,\mathrm{d}x \\ &\quad\leqq \frac{\alpha}{\varepsilon^2}\int_\Omega (w_{\varepsilon,\tau}^{(0)})^{\frac{2}{\alpha}}\,\mathrm{d}x + C\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(0)}) + C_T(1 + C_{f,g}^2) \\ &\qquad + C\tau\sum_{l=0}^{k}\mathcal{F}_\varepsilon^{(l)} + \kappa\int_0^{k\tau}\int_\Gamma \theta_\flat^2\,\mathrm{d}\mathcal{H}^{d-1}\,\mathrm{d}t,\end{aligned} \tag{3.54}$$

where C is a universal constant and C_T possibly depends on T. Then, in view of (3.48), (3.49), and (3.51) for $l = 0$, the result follows by summing up the two estimates. We now treat (3.53) and (3.54) separately. *Step 1 (Inductive bound on the mechanical energy):* The first part is achieved by bounds similar to the ones obtained in the proof of Proposition 3.11, and we therefore refer to estimates therein. Testing (3.7) for l in place of k with $z = \delta_\tau y_{\varepsilon,\tau}^{(l)}$ we get (3.28). Then, multiplying both sides by $\frac{\tau}{\varepsilon^2}$, summing over $l = 1, \ldots, k$, and using $W = W^{\mathrm{el}} + W^{\mathrm{cpl}}$, (3.31), as well as (3.33), by possibly increasing C_1 we derive that

$$
\begin{aligned}
&\mathcal{M}_\varepsilon(y_{\varepsilon,\tau}^{(k)}) - \mathcal{M}_\varepsilon(y_{0,\varepsilon}) + \frac{\tau}{\varepsilon^2}\sum_{l=1}^{k} \\
&\qquad \int_\Omega \big(\partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)} + \xi(\nabla y_{\varepsilon,\tau}^{(l-1)}, \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)})\big)\mathrm{d}x \\
&\quad \leqq C_1 \frac{\tau^2}{\varepsilon^2}\sum_{l=1}^{k}\int_\Omega |\delta_\tau \nabla^2 y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x + \frac{\tau}{\varepsilon}\sum_{l=1}^{k}\langle \ell_\tau^{(l)}, \delta_\tau y_{\varepsilon,\tau}^{(l)}\rangle.
\end{aligned} \tag{3.55}
$$

Here, we also used the definition of V_k in (3.27), and the fact that the initial value is given by $y_{0,\varepsilon}$. By (3.5), (3.52), and Young's inequality it follows that

$$
\begin{aligned}
&\frac{1}{\varepsilon^2}\Big|\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)} \mathrm{d}x\Big| \\
&\quad \leqq \frac{C}{\varepsilon^2}\int_\Omega (\theta_{\varepsilon,\tau}^{(l-1)} \wedge 1)(1 + |\nabla y_{\varepsilon,\tau}^{(l)} - \mathbf{Id}|)|\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|\mathrm{d}x \\
&\quad \leqq \frac{C}{\varepsilon^2}\int_\Omega \big((w_{\varepsilon,\tau}^{(l-1)})^{\frac{2}{\alpha}} + |\nabla y_{\varepsilon,\tau}^{(l)} - \mathbf{Id}|^2\big)\mathrm{d}x + \frac{C}{\varepsilon^2}\int_\Omega |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x.
\end{aligned} \tag{3.56}
$$

By Lemma 3.12 and (3.52) we get that

$$
\begin{aligned}
\sum_{l=1}^{k} \tau \int_\Omega |\delta_\tau \nabla y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x &\leqq \varepsilon^2 C_1 \mathcal{M}_\varepsilon(y_{0,\varepsilon}) + \varepsilon^2 C_1 C_T C_{f,g}^2 \\
&\qquad + C_1 \tau \sum_{l=0}^{k-1}\big(\mathcal{M}(y_{\varepsilon,\tau}^{(l)}) + \|(w_{\varepsilon,\tau}^{(l-1)})^{\frac{1}{\alpha}}\|_{L^2(\Omega)}^2\big).
\end{aligned}
$$

Using the definition of the total energy in (3.48), the definition of $\mathcal{W}_\varepsilon^{\mathrm{el}}$, and (W.4), we insert this in (3.56) to obtain

$$
\frac{\tau}{\varepsilon^2}\sum_{l=1}^{k}\Big|\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)} \mathrm{d}x\Big| + \frac{\tau}{\varepsilon^2}\sum_{l=1}^{k}\int_\Omega |\delta_\tau \nabla^2 y_{\varepsilon,\tau}^{(l)}|^2 \mathrm{d}x \tag{3.57}
$$

$$
\leqq C\tau \sum_{l=0}^{k}\mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l)}) + C\mathcal{M}_\varepsilon(y_{0,\varepsilon}) + C C_T C_{f,g}^2. \tag{3.58}
$$

Next, by repeating the argument in (3.37)–(3.40) we find that

$$\begin{aligned}&\frac{\tau}{\varepsilon}\sum_{l=1}^{k}\langle \ell_\tau^{(l)}, \delta_\tau y_{\varepsilon,\tau}^{(l)}\rangle\\ &\leqq \varepsilon^{-1}\langle \ell(k\tau), \hat{y}_\tau(k\tau)-\mathbf{id}\rangle - \varepsilon^{-1}\langle \ell(0), \hat{y}_\tau(0)-\mathbf{id}\rangle\\ &\quad + C\sum_{l=0}^{k}\Big(\mathcal{F}_\varepsilon^{(l)}\int_{(l-1)\tau}^{l\tau}\big(\|\dot{\ell}(t)\|_{H^{-1}}+\|\dot{\ell}(t+\tau)\|_{H^{-1}}\big)\mathrm{d}t\Big) + C_T(1+C_{f,g}^3).\end{aligned} \tag{3.59}$$

Employing (3.57) and (3.59) in (3.55), and using again (3.50) we arrive at (3.53).

Step 2 (Inductive bound on the temperature): For $\alpha \in [1,2)$, let $\chi(t) := \frac{\alpha}{2}(\varepsilon^\alpha + t)^{\frac{2}{\alpha}}$ for $t \geqq 0$. The convexity of χ implies that

$$\int_\Omega (w_{\varepsilon,\tau}^{(l)} - w_{\varepsilon,\tau}^{(l-1)})\chi'(w_{\varepsilon,\tau}^{(l)})\mathrm{d}x \geqq \int_\Omega \chi(w_{\varepsilon,\tau}^{(l)})\mathrm{d}x - \int_\Omega \chi(w_{\varepsilon,\tau}^{(l-1)})\mathrm{d}x.$$

Summation of this estimate over $l = 1,\ldots,k$ leads to

$$\begin{aligned}&\frac{\alpha}{2}\int_\Omega (w_{\varepsilon,\tau}^{(k)})^{\frac{2}{\alpha}}\mathrm{d}x\\ &\leqq \int_\Omega \chi(w_{\varepsilon,\tau}^{(k)})\mathrm{d}x \leqq \int_\Omega \chi(w_{\varepsilon,\tau}^{(0)})\mathrm{d}x + \sum_{l=1}^{k}\int_\Omega (w_{\varepsilon,\tau}^{(l)} - w_{\varepsilon,\tau}^{(l-1)})\chi'(w_{\varepsilon,\tau}^{(l)})\mathrm{d}x.\end{aligned} \tag{3.60}$$

This suggests to test (3.11) (for l in place of k, $\xi_\alpha^{\mathrm{reg}}$ in place of ξ, and $\varepsilon^\alpha\theta_{\flat,\tau}^{(l)}$ in place of $\varepsilon^2\theta_{\flat,\tau}^{(l)}$) with $\varphi = \chi'(w_{\varepsilon,\tau}^{(l)})$ which yields

$$\begin{aligned}0 = &\int_\Omega \Big(\delta_\tau w_{\varepsilon,\tau}^{(l)} - \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)}) :\\ &\qquad \delta_\tau \nabla y_{\varepsilon,\tau}^{(l)} - \xi_\alpha^{\mathrm{reg}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}, \theta_{\varepsilon,\tau}^{(l-1)})\Big)\chi'(w_{\varepsilon,\tau}^{(l)})\mathrm{d}x\\ &+ \int_\Omega \mathcal{K}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)})\nabla\theta_{\varepsilon,\tau}^{(l)}\cdot\nabla(\chi'(w_{\varepsilon,\tau}^{(l)}))\mathrm{d}x + \kappa\int_\Gamma (\theta_{\varepsilon,\tau}^{(l)}\\ &- \varepsilon^\alpha\theta_{\flat,\tau}^{(l)})\chi'(w_{\varepsilon,\tau}^{(l)})\mathrm{d}\mathcal{H}^{d-1}.\end{aligned} \tag{3.61}$$

We now estimate the various terms separately. First, we employ (3.5), (2.14), (3.1), and Young's inequality with powers $2/\alpha$ and $2/(2-\alpha)$ to obtain

$$\begin{aligned}&\Big|\int_\Omega \Big[\partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(l-1)}, \theta_{\varepsilon,\tau}^{(l-1)}) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}\Big]\chi'(w_{\varepsilon,\tau}^{(l)})\mathrm{d}x\Big|\\ &\leqq 2C_0\int_\Omega (\theta_{\varepsilon,\tau}^{(l-1)}\wedge 1)(1+|\nabla y_{\varepsilon,\tau}^{(l-1)}|)|\delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}|(\varepsilon^\alpha + w_{\varepsilon,\tau}^{(l)})^{\frac{2}{\alpha}-1}\mathrm{d}x\\ &\leqq C(1+C_1)\int_\Omega \Big((w_{\varepsilon,\tau}^{(l-1)}\wedge 1)^{\frac{2}{\alpha}}|\delta_\tau\nabla y_{\varepsilon,\tau}^{(l)}|^{\frac{2}{\alpha}} + \big(\varepsilon^2 + (w_{\varepsilon,\tau}^{(l)})^{\frac{2}{\alpha}}\big)\Big)\mathrm{d}x.\end{aligned} \tag{3.62}$$

If $\alpha \in (1,2)$, we use in the last estimate another Young's inequality, now with powers $\alpha/(\alpha-1)$ and α, as well as $t \wedge 1 \leqq t^{(\alpha-1)/\alpha}$ for all $t \geqq 0$ to show that

$$\begin{aligned}
&\Big|\int_\Omega \Big[\partial_F W^{\mathrm{cpl}}(\nabla y^{(l-1)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(l)}_{\varepsilon,\tau}\Big]\chi'(w^{(l)}_{\varepsilon,\tau})\mathrm{d}x\Big| \\
&\quad\leqq C\int_\Omega \big(\varepsilon^2 + (w^{(l-1)}_{\varepsilon,\tau})^{\frac{2}{\alpha}} + (w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}} + |\delta_\tau \nabla y^{(l)}_{\varepsilon,\tau}|^2\big)\mathrm{d}x \\
&\quad\leqq C\varepsilon^2\Big(1 + \mathcal{E}_\varepsilon(y^{(l-1)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau}) + \mathcal{E}_\varepsilon(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) + \frac{1}{\varepsilon^2}\int_\Omega |\delta_\tau \nabla y^{(l)}_{\varepsilon,\tau}|^2\mathrm{d}x\Big).
\end{aligned} \tag{3.63}$$

Notice that for $\alpha = 1$ the above bound follows directly from (3.62), simply using $w^{(l-1)}_{\varepsilon,\tau} \wedge 1 \leqq 1$.

Next, we estimate the $\xi^{\mathrm{reg}}_\alpha$-term. From the definition of $\xi^{\mathrm{reg}}_\alpha$ in (2.35), we have that $\xi^{\mathrm{reg}}_\alpha \leqq \xi^{\frac{\alpha}{2}}$. Hence, by Young's inequality with power $2/\alpha$ and $2/(2-\alpha)$, and by a similar reasoning as before, it follows that

$$\begin{aligned}
&\int_\Omega \xi^{\mathrm{reg}}_\alpha(\nabla y^{(l-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau})\chi'(w^{(l)}_{\varepsilon,\tau})\mathrm{d}x \\
&\quad\leqq \int_\Omega \xi(\nabla y^{(l-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau})\mathrm{d}x + C\varepsilon^2(1 + \mathcal{E}_\varepsilon(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau})).
\end{aligned} \tag{3.64}$$

We continue by investigating the $\mathcal{K}$-term. By (2.12) and the chain rule we have that

$$\begin{aligned}
\nabla(\chi'(w^{(l)}_{\varepsilon,\tau})) = {}& \frac{2-\alpha}{\alpha}(\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}-2} \\
&\times \Big[\Big(\partial_F W^{\mathrm{cpl}}(\nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}) - \theta^{(l)}_{\varepsilon,\tau}\partial_{F\theta}W^{\mathrm{cpl}}(\nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau})\Big) : \nabla^2 y^{(l)}_{\varepsilon,\tau} \\
&- \theta^{(l)}_{\varepsilon,\tau}\partial^2_\theta W^{\mathrm{cpl}}(\nabla y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau})\nabla\theta^{(l)}_{\varepsilon,\tau}\Big].
\end{aligned}$$

This combined with (3.5), the second and third bound in (C.5), (2.14), (3.3), and (3.4) leads to

$$\begin{aligned}
&\mathcal{K}^{(l-1)}_{\varepsilon,\tau}\nabla\theta^{(l)}_{\varepsilon,\tau} \cdot \nabla(\chi'(w^{(l)}_{\varepsilon,\tau})) \\
&\quad\geqq \frac{2-\alpha}{\alpha}(\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}-2}\Big(c|\nabla\theta^{(l)}_{\varepsilon,\tau}|^2 - C(w^{(l)}_{\varepsilon,\tau} \wedge 1)|\nabla^2 y^{(l)}_{\varepsilon,\tau}||\nabla\theta^{(l)}_{\varepsilon,\tau}|\Big)
\end{aligned} \tag{3.65}$$

for some $c > 0$, where we set $\mathcal{K}^{(l-1)}_{\varepsilon,\tau} := \mathcal{K}(\nabla y^{(l-1)}_{\varepsilon,\tau}, \theta^{(l-1)}_{\varepsilon,\tau})$ for brevity. (In the definition of χ, the addend ε^α appears to ensure that $(\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}-2}$ is well-defined for $\alpha > 1$.) By $t \wedge 1 \leqq t^{1-2/(p\alpha)}$ for all $t \geqq 0$, Young's inequality twice (firstly with power 2 and constant $\lambda \in (0,1)$, secondly with powers $p/(p-2)$ and $p/2$) we derive that

$$\begin{aligned}
(w^{(l)}_{\varepsilon,\tau} \wedge 1)|\nabla^2 y^{(l)}_{\varepsilon,\tau}||\nabla\theta^{(l)}_{\varepsilon,\tau}| &\leqq \lambda|\nabla\theta^{(l)}_{\varepsilon,\tau}|^2 + \frac{1}{\lambda}(w^{(l)}_{\varepsilon,\tau})^{2\frac{p-2}{p}}(w^{(l)}_{\varepsilon,\tau})^{\frac{4(\alpha-1)}{p\alpha}}|\nabla^2 y^{(l)}_{\varepsilon,\tau}|^2 \\
&\leqq \lambda|\nabla\theta^{(l)}_{\varepsilon,\tau}|^2 + \frac{1}{\lambda}\Big((w^{(l)}_{\varepsilon,\tau})^2 + (w^{(l)}_{\varepsilon,\tau})^{2\frac{\alpha-1}{\alpha}}|\nabla^2 y^{(l)}_{\varepsilon,\tau}|^p\Big).
\end{aligned}$$

Choosing λ small enough such that $C\lambda < c/2$ (with c and C as in (3.65)), we derive, with (3.65), that

$$\int_\Omega \mathcal{K}^{(l-1)}_{\varepsilon,\tau} \nabla\theta^{(l)}_{\varepsilon,\tau} \cdot \nabla(\chi'(w^{(l)}_{\varepsilon,\tau}))\mathrm{d}x \geqq \frac{c}{2}\frac{2-\alpha}{\alpha}(\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}-2}|\nabla\theta^{(l)}_{\varepsilon,\tau}|^2 - C\varepsilon^2\mathcal{E}_\varepsilon(y^{(l)}_{\varepsilon,\tau}, \theta^{(l)}_{\varepsilon,\tau}). \tag{3.66}$$

Lastly, for the boundary term, we use (2.14) as well as Young's inequality with powers $2/\alpha$ and $2/(2-\alpha)$ and constant $\lambda \in (0, 1)$ to arrive at

$$\begin{aligned}\int_\Gamma (\theta^{(l)}_{\varepsilon,\tau} - \varepsilon^\alpha\theta^{(l)}_{\flat,\tau})\chi'(w^{(l)}_{\varepsilon,\tau})\mathrm{d}\mathcal{H}^{d-1} &\geqq \int_\Gamma (C_0^{-1}w^{(l)}_{\varepsilon,\tau} - \varepsilon^\alpha\theta^{(l)}_{\flat,\tau})(\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}-1}\mathrm{d}\mathcal{H}^{d-1}\\ &\geqq \frac{1}{C_0}\int_\Gamma (w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}}\mathrm{d}\mathcal{H}^{d-1} - \frac{\varepsilon^2}{\lambda}\int_\Gamma (\theta^{(l)}_{\flat,\tau})^{\frac{2}{\alpha}}\mathrm{d}\mathcal{H}^{d-1} - \lambda\int_\Gamma (\varepsilon^\alpha + w^{(l)}_{\varepsilon,\tau})^{\frac{2}{\alpha}}\mathrm{d}\mathcal{H}^{d-1}.\end{aligned}$$

Therefore, choosing λ sufficiently small with respect to $1/C_0$, we get that

$$\int_\Gamma (\theta^{(l)}_{\varepsilon,\tau} - \varepsilon^\alpha\theta^{(l)}_{\flat,\tau})\chi'(w^{(l)}_{\varepsilon,\tau})\mathrm{d}\mathcal{H}^{d-1} \geqq -C\varepsilon^2\Big(1 + \int_\Gamma (\theta^{(l)}_{\flat,\tau})^{\frac{2}{\alpha}}\mathrm{d}\mathcal{H}^{d-1}\Big). \tag{3.67}$$

We then divide (3.60) by ε^2, insert (3.61) multiplied by τ in this inequality, and use (3.63), (3.64), (3.66), and (3.67) to estimate the various terms. This together with the bounds from (3.50) and (3.57), and the fact that

$$\tau\sum_{l=1}^k \|(\theta^{(l)}_{\flat,\tau})^{\frac{2}{\alpha}}\|_{L^1(\Gamma)} \leqq C\tau\sum_{l=1}^k \big(1 + \|\theta^{(l)}_{\flat,\tau}\|^2_{L^2(\Gamma)}\big) \leqq C_T + C\|\theta_\flat\|^2_{L^2([0,k\tau]\times\Gamma)}$$

by Hölder's inequality yields (3.54). This concludes the proof. □

Theorem 3.16. (Well-definedness of the scheme) *For any $T > 0$ there exist a constant $\bar{C}_T$, corresponding constants*

$$\begin{aligned}M' &:= 2e^{\bar{C}_T(1+C_{f,g})}\Big(\bar{C}_T\mathcal{F}^{(0)} + \bar{C}_T(1 + C^3_{f,g}) + \kappa\int_0^T\int_\Gamma \theta_\flat^2\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t\Big),\\ M &:= 2M' + \bar{C}_T C^2_{f,g},\end{aligned}$$

as well as constants ε_0, $\tau_0 \in (0, 1]$ depending also on M such that the following holds true: for each $\varepsilon \in (0, \varepsilon_0)$ and $\tau \in (0, \tau_0)$ such that $T/\tau \in \mathbb{N}$ the sequences $y^{(0)}_{\varepsilon,\tau}, \dots, y^{(T/\tau)}_{\varepsilon,\tau}$ and $\theta^{(0)}_{\varepsilon,\tau}, \dots, \theta^{(T/\tau)}_{\varepsilon,\tau}$ exist, and for all $k \in \{0, \dots, T/\tau\}$ we have that

$$\begin{aligned}&\mathcal{E}_\varepsilon(y^{(k)}_{\varepsilon,\tau}, \theta^{(k)}_{\varepsilon,\tau}) \leqq M,\\ &\sum_{k=1}^{T/\tau}\frac{\tau}{\varepsilon^2}\int_\Omega |\delta_\tau\nabla y^{(l)}_{\varepsilon,\tau}|^2\mathrm{d}x \leqq \bar{C}_T M(1 + T) + \bar{C}_T C^2_{f,g}.\end{aligned} \tag{3.68}$$

Proof. The theorem is a consequence of Lemma 3.15 and Lemma 3.12. The argument is similar to the one of Theorem 3.13 and we therefore omit the details. Let us just mention that the energy bound follows in the same way by induction, up to using different values β, a_l, and b_l in (3.47), and by employing (3.50) in place of Lemma 3.10. Based on the uniform energy bound, Proposition 3.14 indeed shows that the scheme is well-defined, provided that ε_0 is chosen sufficiently small. Eventually, the bound on the strain rates follows from Lemma 3.12, see particularly (3.57) in the previous proof. □

Remark 3.17. Due to our regularization of the dissipation rate, in the case $\alpha \in [1, 2)$ we obtain the additional control

$$\int_0^T \int_\Omega \frac{|\nabla \overline{\mu}_{\varepsilon,\tau}|^2}{(1+\overline{\mu}_{\varepsilon,\tau})^{2(1-\frac{1}{\alpha})}} \mathrm{d}x\mathrm{d}t \leqq C_\alpha < \infty \tag{3.69}$$

for a constant C_α depending on α, but independent of ε and τ, where we shortly wrote $\overline{\mu}_{\varepsilon,\tau} := \varepsilon^{-\alpha}\overline{\theta}_{\varepsilon,\tau}$ (see also (2.26) for the definition of $\overline{\theta}_{\varepsilon,\tau}$). This follows by using the positive term on the right-hand side of (3.66).

3.4. A Priori bounds

Fix initial values $(y_{0,\varepsilon}, \theta_{0,\varepsilon})$ with $\mathcal{E}_\varepsilon(y_{0,\varepsilon}, \theta_{0,\varepsilon}) \leqq E_0$ for some $E_0 > 0$. Without further notice, we suppose in this subsection that the sequences $y^{(0)}_{\varepsilon,\tau}, \ldots, y^{(T/\tau)}_{\varepsilon,\tau}$ and $\theta^{(0)}_{\varepsilon,\tau}, \ldots, \theta^{(T/\tau)}_{\varepsilon,\tau}$ exist by Theorem 2.3(i) or Proposition 2.5(i), respectively, for $\varepsilon \in (0, \varepsilon_0)$ for some ε_0 depending only on α, E_0, f, g, $\theta_\flat$, and T. (In the case $\alpha = 2$, we can set $\varepsilon = 1$). We derive a priori bounds on the rescaled displacements $\varepsilon^{-1}(y^{(l)}_{\varepsilon,\tau} - \mathbf{id})$ and the rescaled temperatures $\varepsilon^{-\alpha}\theta^{(l)}_{\varepsilon,\tau}$ for $l \in \{1, \ldots, T/\tau\}$. To this end, for small ε, we will again assume (W.4). Recall the definition of the interpolations in (2.26). In a similar way, we write $\overline{w}_{\varepsilon,\tau} = W^{\mathrm{in}}(\overline{y}_{\varepsilon,\tau}, \overline{\theta}_{\varepsilon,\tau})$, and similarly for the other interpolations. The next lemma is a direct consequence of Theorem 3.13 and Theorem 3.16.

Lemma 3.18. (First a priori bounds) *Let $E_0 > 0$ such that $\mathcal{E}_\varepsilon(y_{0,\varepsilon}, \theta_{0,\varepsilon}) \leqq E_0$. Then, there exists a constant $C > 0$ depending on α, E_0, f, g, $\theta_\flat$, and T such that $\mathcal{E}_\varepsilon(y^{(k)}_{\varepsilon,\tau}, \theta^{(k)}_{\varepsilon,\tau}) \leqq C$ for all $k \in \{1, \ldots, T/\tau\}$, and the interpolants constructed from the discrete solutions satisfy*

$$\|\overline{y}_{\varepsilon,\tau} - \mathbf{id}\|_{L^\infty(I;W^{1,\infty}(\Omega;\mathbb{R}^d))} + \|\nabla^2 \overline{y}_{\varepsilon,\tau}\|_{L^\infty(I;L^p(\Omega;\mathbb{R}^d))} \leqq C\varepsilon^{2/p}, \tag{3.70a}$$

$$\|\overline{y}_{\varepsilon,\tau} - \mathbf{id}\|_{L^\infty(I;H^1(\Omega;\mathbb{R}^d))} \leqq C\varepsilon, \tag{3.70b}$$

$$\|\nabla \dot{\hat{y}}_{\varepsilon,\tau}\|_{L^2(I\times\Omega;\mathbb{R}^{d\times d})} \leqq C\varepsilon, \tag{3.70c}$$

$$\|\overline{\theta}_{\varepsilon,\tau}\|_{L^\infty(I;L^1(\Omega))} + \|\overline{w}_{\varepsilon,\tau}\|_{L^\infty(I;L^1(\Omega))} \leqq C\varepsilon^\alpha \tag{3.70d}$$

Estimates (3.70a)–(3.70b) *also hold for $\underline{y}_{\varepsilon,\tau}$, and* (3.70d) *holds for $\underline{\theta}_{\varepsilon,\tau}$, $\hat{\theta}_{\varepsilon,\tau}$, $\underline{w}_{\varepsilon,\tau}$, and $\hat{w}_{\varepsilon,\tau}$, as well.*

Proof. Let us first suppose that (W.4) holds. The energy bound on $\mathcal{E}_\varepsilon$ for $\alpha = 2$ and $\alpha \in [1, 2)$ follows directly from Theorems 3.13 and 3.16, respectively. The first two estimates can be shown from the uniform bound on the energy, (W.4), (3.2), (5.21), and Poincaré's inequality. In a similar way, the bound on $\overline{w}_{\varepsilon,\tau}$ in (3.70d) follows from the bound on the total rescaled energy, (3.48), and Hölder's inequality. Then, the proof of (3.70d) is concluded by (2.14). Finally, (3.70c) is a direct consequence of (3.44) and (3.68), respectively. Eventually, for $\alpha = 2$ and ε near 1, the result also holds without assuming (W.4) as (W.3) allows us to derive (3.70a)–(3.70b) with C in place of $C\varepsilon^{2/p}$ and $C\varepsilon$ on the right-hand side. □

In order to pass to the limit $\tau \to 0$ in the next section, we need additional a priori bounds for the temperature. Testing the equation (3.11) turns out to be delicate since for $\alpha = 2$ the viscous dissipation $\xi(\nabla y^{(k-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau})$ is only bounded in $L^1(I \times \Omega)$. Thus, to obtain improved estimates that work in this case, we employ special test functions developed by Boccardo and Gallouët [10] for parabolic equations with a measure-valued right-hand side, see also [16]. We follow here the approach in [33]. However, almost complete proofs are provided since compared to their setting we perform the estimates in the time discrete setting and we derive fine estimates in terms of the small parameter ε.

Lemma 3.19. (Weighted L^2-bound) *For any $\eta \in (0, 1)$ there exists a constant C independent of ε, τ, and α such that*

$$\sum_{k=1}^{T/\tau} \tau \int_\Omega \frac{\eta}{\left(1 + \varepsilon^{-\alpha} w^{(k)}_{\varepsilon,\tau}\right)^{1+\eta}} \left|\nabla w^{(k)}_{\varepsilon,\tau}\right|^2 \mathrm{d}x \leqq C\varepsilon^{2\alpha}. \tag{3.71}$$

Actually, this statement is needed only for $\alpha = 2$ since for $\alpha \in [1, 2)$ we have a better estimate by Remark 3.17. Still, we state and prove the result for any α since the following argument does not depend on α.

Proof. Step 1: In the following, C will denote a constant independent of k, ε, τ, α, and η. Given $k \in \{1, \ldots, T/\tau\}$, we have by (3.11) (for $\xi^{\mathrm{reg}}_\alpha$ in place of ξ) that for any $\varphi_k \in H^1(\Omega)$

$$\begin{aligned} \int_\Omega \delta_\tau w^{(k)}_{\varepsilon,\tau} \varphi_k \mathrm{d}x &= \int_\Omega h^k_{\varepsilon,\tau} \varphi_k \mathrm{d}x \\ &\quad - \int_\Omega \mathcal{K}^{(k-1)}_{\varepsilon,\tau} \nabla \theta^{(k)}_{\varepsilon,\tau} \cdot \nabla \varphi_k \mathrm{d}x - \kappa \int_\Gamma (\theta^{(k)}_{\varepsilon,\tau} - \varepsilon^\alpha \theta^{(k)}_{\flat,\tau}) \varphi_k \mathrm{d}\mathcal{H}^{d-1}, \end{aligned} \tag{3.72}$$

where we write

$$\begin{aligned} h^{(k)}_{\varepsilon,\tau} &:= \partial_F W^{\mathrm{cpl}}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) : \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau} + \xi^{\mathrm{reg}}_\alpha(\nabla y^{(k-1)}_{\varepsilon,\tau}, \delta_\tau \nabla y^{(k)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}), \\ \mathcal{K}^{(k-1)}_{\varepsilon,\tau} &:= \mathcal{K}(\nabla y^{(k-1)}_{\varepsilon,\tau}, \theta^{(k-1)}_{\varepsilon,\tau}) \end{aligned} \tag{3.73}$$

for brevity. Given $\eta \in (0, 1)$, let $\chi_{\eta,\varepsilon} : \mathbb{R} \to \mathbb{R}$ be the function uniquely determined by $\chi_{\eta,\varepsilon}(0) = 0$ and $\chi'_{\eta,\varepsilon}(t) = 1 - \frac{1}{(1+\varepsilon^{-\alpha} t)^\eta}$ for all $t \geqq 0$. Choosing $\varphi_k := \chi'_{\eta,\varepsilon}(w^{(k)}_{\varepsilon,\tau})$

in (3.72), multiplying both sides by τ, and summing over $k = 1, \dots, T/\tau$, we arrive at

$$\begin{aligned}\sum_{k=1}^{T/\tau} \int_\Omega (w_{\varepsilon,\tau}^{(k)} - w_{\varepsilon,\tau}^{(k-1)}) \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}x &= \sum_{k=1}^{T/\tau} \tau \int_\Omega h_{\varepsilon,\tau}^{(k)} \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}x \\ &- \sum_{k=1}^{T/\tau} \tau \int_\Omega \chi''_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathcal{K}_{\varepsilon,\tau}^{(k-1)} \nabla \theta_{\varepsilon,\tau}^{(k)} \cdot \nabla w_{\varepsilon,\tau}^{(k)} \mathrm{d}x \\ &- \kappa \sum_{k=1}^{T/\tau} \tau \int_\Gamma (\theta_{\varepsilon,\tau}^{(k)} - \varepsilon^\alpha \theta_{\flat,\tau}^{(k)}) \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}\mathcal{H}^{d-1}. \end{aligned} \tag{3.74}$$

Our goal is to show that

$$\sum_{k=1}^{T/\tau} \tau \int_\Omega \chi''_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathcal{K}_{\varepsilon,\tau}^{(k-1)} \nabla \theta_{\varepsilon,\tau}^{(k)} \cdot \nabla w_{\varepsilon,\tau}^{(k)} \mathrm{d}x \leqq C\varepsilon^\alpha. \tag{3.75}$$

To this end, we estimate the various terms in (3.74). First, notice that by the convexity of $\chi_{\eta,\varepsilon}$ we have, for any $k \in \{1, \dots, T/\tau\}$, that

$$\chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k-1)}) \geqq \chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) + \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)})(w_{\varepsilon,\tau}^{(k-1)} - w_{\varepsilon,\tau}^{(k)}),$$

and therefore,

$$\begin{aligned}\sum_{k=1}^{T/\tau} \int_\Omega \big(w_{\varepsilon,\tau}^{(k)} - w_{\varepsilon,\tau}^{(k-1)}\big) \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}x &\geqq \sum_{k=1}^{T/\tau} \int_\Omega \big(\chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) - \chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k-1)})\big) \mathrm{d}x \\ = \int_\Omega \chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(T/\tau)}) \mathrm{d}x - \int_\Omega \chi_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(0)}) \mathrm{d}x &\geqq - \int_\Omega w_{\varepsilon,\tau}^{(0)} \mathrm{d}x \geqq -C\varepsilon^\alpha,\end{aligned}$$

where we used $\chi_{\eta,\varepsilon} \geqq 0$ and $\chi_{\eta,\varepsilon}(t) \leqq t$ for all $t \geqq 0$, and in the last step also (3.70d). Using (3.5), $\xi_\alpha^{\mathrm{reg}} \leqq \xi$, (2.9), (D.2), and (3.70a) we see that

$$\sum_{k=1}^{T/\tau} \int_\Omega |h_{\varepsilon,\tau}^{(k)}| \mathrm{d}x \leqq C \sum_{k=1}^{T/\tau} \int_\Omega \Big(\sqrt{\theta_{\varepsilon,\tau}^{(k-1)}} |\delta_\tau \nabla y_{\varepsilon,\tau}^{(k)}| + |\delta_\tau \nabla y_{\varepsilon,\tau}^{(k)}|^2\Big) \mathrm{d}x,$$

where we used that $t \wedge 1 \leqq \sqrt{t}$ for $t \geqq 0$. Then, by Young's inequality, $\chi'_{\eta,\varepsilon} \leqq 1$, (3.70c), and (3.70d) we get

$$\begin{aligned}&\sum_{k=1}^{T/\tau} \tau \int_\Omega h_{\varepsilon,\tau}^{(k)} \chi'_{\eta,\varepsilon}(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}x \\ &\leqq \sum_{k=1}^{T/\tau} \tau \int_\Omega |h_{\varepsilon,\tau}^{(k)}| \mathrm{d}x \leqq C \sum_{k=1}^{T/\tau} \tau \int_\Omega \big(\theta_{\varepsilon,\tau}^{(k-1)} + |\delta_\tau \nabla y_{\varepsilon,\tau}^{(k)}|^2\big) \mathrm{d}x \leqq C\varepsilon^\alpha, \end{aligned} \tag{3.76}$$

where we have used $\alpha \leqq 2$. Lastly, by $\theta_{\varepsilon,\tau}^{(k)} \geqq 0$, $\kappa \geqq 0$, $\chi_{\eta,\varepsilon}' \in [0,1]$, and the definition of $\theta_{\flat,\tau}^{(k)}$ it follows that

$$
\begin{aligned}
& -\kappa \sum_{k=1}^{T/\tau} \tau \int_{\Omega} (\theta_{\varepsilon,\tau}^{(k)} - \varepsilon^{\alpha}\theta_{\flat,\tau}^{(k)}) \chi_{\eta,\varepsilon}'(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}\mathcal{H}^{d-1} \\
& \leqq \kappa \sum_{k=1}^{T/\tau} \tau \int_{\Omega} \varepsilon^{\alpha}\theta_{\flat,\tau}^{(k)} \chi_{\eta,\varepsilon}'(w_{\varepsilon,\tau}^{(k)}) \mathrm{d}\mathcal{H}^{d-1} \\
& \leqq \kappa\varepsilon^{\alpha} \int_0^T \int_{\Omega} \theta_{\flat} \mathrm{d}x\mathrm{d}t \leqq C\varepsilon^{\alpha},
\end{aligned}
$$

where C also depends on $\theta_\flat$. Employing all the aforementioned estimates in (3.74) we obtain (3.75).

Step 2: We are now ready to show (3.71). In this regard, first notice the following relation between $\nabla w_{\varepsilon,\tau}^{(k)}$ and $\nabla\theta_{\varepsilon,\tau}^{(k)}$: since $w_{\varepsilon,\tau}^{(k)} = W^{\mathrm{in}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)})$, (2.12) implies that

$$
\begin{aligned}
\nabla w_{\varepsilon,\tau}^{(k)} &= \Big[\partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) - \theta_{\varepsilon,\tau}^{(k)} \partial_{F\theta} W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)})\Big] : \\
&\quad \nabla^2 y_{\varepsilon,\tau}^{(k)} - \theta_{\varepsilon,\tau}^{(k)} \partial_\theta^2 W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k)}) \nabla\theta_{\varepsilon,\tau}^{(k)} \\
&=: \tilde{W}_1^{(k)} : \nabla^2 y_{\varepsilon,\tau}^{(k)} + \tilde{W}_2^{(k)} \nabla\theta_{\varepsilon,\tau}^{(k)}.
\end{aligned} \tag{3.77}
$$

By (3.5), (C.5), and (3.70a), we find that the abbreviations $\tilde{W}_1^{(k)}$ and $\tilde{W}_2^{(k)}$ satisfy $\tilde{W}_1^{(k)} \leqq C(\theta_{\varepsilon,\tau}^{(k)} \wedge 1)$ and $\tilde{W}_2^{(k)} \in [c_0, C_0]$, respectively. Then, using (3.77), Lemma 3.3, and the energy bound from Lemma 3.18 we see that there exists a constant $c > 0$ such that

$$
\begin{aligned}
\frac{c}{C_0} \chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) |\nabla w_{\varepsilon,\tau}^{(k)}|^2 &\leqq \big(\tilde{W}_2^{(k)}\big)^{-1} \chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) \mathcal{K}_{\varepsilon,\tau}^{(k-1)} \nabla w_{\varepsilon,\tau}^{(k)} \cdot \nabla w_{\varepsilon,\tau}^{(k)} \\
&\leqq \chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) \mathcal{K}_{\varepsilon,\tau}^{(k-1)} \nabla\theta_{\varepsilon,\tau}^{(k)} \cdot \nabla w_{\varepsilon,\tau}^{(k)} \\
&\quad + C\chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) (\theta_{\varepsilon,\tau}^{(k)} \wedge 1) |\nabla^2 y_{\varepsilon,\tau}^{(k)}| |\nabla w_{\varepsilon,\tau}^{(k)}|.
\end{aligned} \tag{3.78}
$$

We now control the second term above. By $t \wedge 1 \leqq t^{\frac{p-1}{p}}$ for all $t \geqq 0$ and Young's inequality with constant $\lambda \in (0,1)$ (to be chosen later), we estimate by (2.14)

$$
\begin{aligned}
& C\chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) (\theta_{\varepsilon,\tau}^{(k)} \wedge 1) |\nabla^2 y_{\varepsilon,\tau}^{(k)}| |\nabla w_{\varepsilon,\tau}^{(k)}| \\
& \leqq C\chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) \Big(\lambda |\nabla w_{\varepsilon,\tau}^{(k)}|^2 + \frac{1}{\lambda} (w_{\varepsilon,\tau}^{(k)})^{2\frac{p-1}{p}} |\nabla^2 y_{\varepsilon,\tau}^{(k)}|^2\Big).
\end{aligned} \tag{3.79}
$$

Using the elementary fact

$$
\chi_{\eta,\varepsilon}''(w_{\varepsilon,\tau}^{(k)}) = \frac{\eta}{\varepsilon^{\alpha}\Big(1 + \varepsilon^{-\alpha} w_{\varepsilon,\tau}^{(k)}\Big)^{1+\eta}} \leqq \frac{1}{\varepsilon^{\alpha} + w_{\varepsilon,\tau}^{(k)}}, \tag{3.80}
$$

we derive by Young's inequality with powers $p/(p-2)$ and $p/2$ that

$$\chi''_{\eta,\varepsilon}(w^{(k)}_{\varepsilon,\tau})(w^{(k)}_{\varepsilon,\tau})^{2\frac{p-1}{p}}|\nabla^2 y^{(k)}_{\varepsilon,\tau}|^2 \leqq (w^{(k)}_{\varepsilon,\tau})^{2\frac{p-1}{p}-1}|\nabla^2 y^{(k)}_{\varepsilon,\tau}|^2$$
$$= (w^{(k)}_{\varepsilon,\tau})^{\frac{p-2}{p}}|\nabla^2 y^{(k)}_{\varepsilon,\tau}|^2 \leqq C(w^{(k)}_{\varepsilon,\tau} + |\nabla^2 y^{(k)}_{\varepsilon,\tau}|^p).$$

Let us take λ small enough so that $C\lambda \leqq c/(2C_0)$ where c is as in (3.78) and C is as in (3.79). Then, inserting (3.79) into (3.78) we derive that

$$\frac{c}{2C_0}\chi''_{\eta,\varepsilon}(w^{(k)}_{\varepsilon,\tau})|\nabla w^{(k)}_{\varepsilon,\tau}|^2 \leqq C\Big(\chi''_{\eta,\varepsilon}(w^{(k)}_{\varepsilon,\tau})\mathcal{K}^{(k-1)}_{\varepsilon,\tau}\nabla\theta^{(k)}_{\varepsilon,\tau}\cdot\nabla w^{(k)}_{\varepsilon,\tau}+w^{(k)}_{\varepsilon,\tau}+|\nabla^2 y^{(k)}_{\varepsilon,\tau}|^p\Big).$$

Integrating the above inequality over Ω, multiplying by τ, and summing over $k = 1, \dots, T/\tau$ we derive by (3.75), (3.70a), and (3.70d) that

$$\sum_{k=1}^{T/\tau}\tau\int_\Omega \chi''_{\eta,\varepsilon}(w^{(k)}_{\varepsilon,\tau})|\nabla w^{(k)}_{\varepsilon,\tau}|^2 \mathrm{d}x \leqq C(1+T)\varepsilon^\alpha.$$

where in the final step we used $\alpha \leqq 2$. By using the first identity in (3.80), we conclude the proof of (3.71). □

Theorem 3.20. (Further a priori bounds on the temperature) *For any $q \in [1, \frac{d+2}{d})$ and $r \in [1, \frac{d+2}{d+1})$ there exist constants C_q and C_r, respectively, both independent of ε and τ such that*

$$\sum_{k=0}^{T/\tau}\tau\int_\Omega \big(|\theta^{(k)}_{\varepsilon,\tau}|^q + |w^{(k)}_{\varepsilon,\tau}|^q\big)\mathrm{d}x \leqq C_q\varepsilon^{\alpha q}, \tag{3.81}$$

$$\sum_{k=1}^{T/\tau}\tau\int_\Omega \big(|\nabla\theta^{(k)}_{\varepsilon,\tau}|^r + |\nabla w^{(k)}_{\varepsilon,\tau}|^r\big)\mathrm{d}x \leqq C_r\varepsilon^{\alpha r}. \tag{3.82}$$

Moreover, we can find a constant C independent of ε and τ such that

$$\sum_{k=1}^{T/\tau}\tau\|\delta_\tau w^{(k)}_{\varepsilon,\tau}\|_{W^{1,\infty}(\Omega)^*} \leqq C\varepsilon^\alpha. \tag{3.83}$$

Proof. Let q, r be as in the statement. As $w^{(k)}_{\varepsilon,\tau} \in H^1(\Omega)$ (see (2.24)), it follows that $\|\overline{w}_{\varepsilon,\tau}\|_{L^\infty(I;H^1(\Omega))} < \infty$. Therefore, by using the a priori estimate $\|1+\varepsilon^{-\alpha}\overline{w}_{\varepsilon,\tau}\|_{L^\infty(I;L^1(\Omega))} \leqq C + \mathcal{L}^d(\Omega)$ (see (3.70d)) as well as Lemma 3.19, we can repeat the argument from the proof of [33, Proposition 6.3, equation (6.6)] for $\varepsilon^{-\alpha}\overline{w}_{\varepsilon,\tau}$ in place of w_ε, cf. also Remark 3.21 below. This gives the existence of constants C_q, C_r such that

$$\sum_{k=0}^{T/\tau}\tau\int_\Omega|w^{(k)}_{\varepsilon,\tau}|^q\mathrm{d}x \leqq C_q\varepsilon^{\alpha q}, \quad \sum_{k=1}^{T/\tau}\tau\int_\Omega|\nabla w^{(k)}_{\varepsilon,\tau}|^r\mathrm{d}x \leqq C_r\varepsilon^{\alpha p}. \tag{3.84}$$

By (2.14) we then directly see that (for a possibly larger C_q)

$$\sum_{k=0}^{T/\tau} \tau \int_\Omega |\theta_{\varepsilon,\tau}^{(k)}|^q \mathrm{d}x \leqq C_q \varepsilon^{\alpha q}. \tag{3.85}$$

To conclude the proof of (3.81)–(3.82), it remains to control the gradient of the temperature. Employing the relation between $\nabla w_{\varepsilon,\tau}^{(k)}$ and $\nabla \theta_{\varepsilon,\tau}^{(k)}$ in (3.77), by (3.5) and (C.5) we see that

$$|\nabla \theta_{\varepsilon,\tau}^{(k)}| \leqq C\big(|\nabla w_{\varepsilon,\tau}^{(k)}| + (\theta_{\varepsilon,\tau}^{(k)} \wedge 1)|\nabla^2 y_{\varepsilon,\tau}^{(k)}|\big).$$

Consequently, using $t \wedge 1 \leqq t^{\frac{p-1}{p}}$ for all $t \geqq 0$ and Young's inequality with powers $p/(p-r)$ and p/r we derive that

$$\begin{aligned}\int_\Omega |\nabla \theta_{\varepsilon,\tau}^{(k)}|^r \mathrm{d}x &\leqq C \int_\Omega |\nabla w_{\varepsilon,\tau}^{(k)}|^r \mathrm{d}x + C\varepsilon^{\alpha r} \int_\Omega (\varepsilon^{-\alpha} \theta_{\varepsilon,\tau}^{(k)})^{r\frac{p-1}{p}} |\varepsilon^{-\frac{\alpha}{p}} \nabla^2 y_{\varepsilon,\tau}^{(k)}|^r \mathrm{d}x \\ &\leqq C \int_\Omega |\nabla w_{\varepsilon,\tau}^{(k)}|^r \mathrm{d}x + C\varepsilon^{\alpha r} \int_\Omega \Big((\varepsilon^{-\alpha}\theta_{\varepsilon,\tau}^{(k)})^{r\frac{p-1}{p-r}} + \frac{1}{\varepsilon^\alpha}|\nabla^2 y_{\varepsilon,\tau}^{(k)}|^p\Big)\mathrm{d}x.\end{aligned} \tag{3.86}$$

As r was chosen strictly smaller than $\frac{d+2}{d+1}$, we see by $p \geqq 2$ that

$$r\frac{p-1}{p-r} < \frac{d+2}{d+1}\frac{p-1}{p-\frac{d+2}{d+1}} = \frac{d+2}{d}\frac{1}{1+\frac{p-2}{d(p-1)}} \leqq \frac{d+2}{d}.$$

Consequently, multiplying (3.86) with τ, summing over $k = 1, \ldots, T/\tau$, and using (3.70a), (3.84), and (3.85) we conclude the proof of (3.82). Here, we again used $\alpha \leqq 2$. Lastly, we show (3.83). Testing (3.11) for the k-th step with arbitrary $\varphi \in W^{1,\infty}(\Omega)$, and using the shorthand notation for $h_{\varepsilon,\tau}^{(k)}$ and $\mathcal{K}_{\varepsilon,\tau}^{(k-1)}$ from (3.73), we see by (3.4) and the continuity of the trace operator in $W^{1,1}(\Omega)$ that

$$\begin{aligned}&\Big|\int_\Omega \delta_\tau w_{\varepsilon,\tau}^{(k)} \varphi \mathrm{d}x\Big| = \Big|\int_\Omega h_{\varepsilon,\tau}^{(k)}\varphi \mathrm{d}x - \int_\Omega \mathcal{K}_{\varepsilon,\tau}^{(k-1)} \nabla \theta_{\varepsilon,\tau}^{(k)} \cdot \nabla \varphi \mathrm{d}x \\ &\quad - \kappa \int_\Gamma (\theta_{\varepsilon,\tau}^{(k)} - \varepsilon^\alpha \theta_{\flat,\tau}^{(k)})\varphi \mathrm{d}\mathcal{H}^{d-1}\Big| \\ &\leqq \|h_{\varepsilon,\tau}^{(k)}\|_{L^1(\Omega)} \|\varphi\|_{L^\infty(\Omega)} + C\|\nabla \theta_{\varepsilon,\tau}^{(k)}\|_{L^1(\Omega)} \|\nabla \varphi\|_{L^\infty(\Omega)} \\ &\quad + \Big(C\kappa \|\theta_{\varepsilon,\tau}^{(k)}\|_{W^{1,1}(\Omega)} + \kappa\varepsilon^\alpha \int_\Gamma \theta_{\flat,\tau}^{(k)} \mathrm{d}\mathcal{H}^{d-1}\Big)\|\varphi\|_{L^\infty(\Omega)} \\ &\leqq \Big(\|h_{\varepsilon,\tau}^{(k)}\|_{L^1(\Omega)} + C\|\theta_{\varepsilon,\tau}^{(k)}\|_{W^{1,1}(\Omega)} + C\varepsilon^\alpha \int_\Gamma \theta_{\flat,\tau}^{(k)} \mathrm{d}\mathcal{H}^{d-1}\Big)\|\varphi\|_{W^{1,\infty}(\Omega)}.\end{aligned}$$

By the arbitrariness of φ this shows that

$$\|\delta_\tau w_{\varepsilon,\tau}^{(k)}\|_{W^{1,\infty}(\Omega)^*} \leqq \|h_\tau^{(k)}\|_{L^1(\Omega)} + C\|\theta_{\varepsilon,\tau}^{(k)}\|_{W^{1,1}(\Omega)} + C\varepsilon^\alpha \int_\Gamma \theta_{\flat,\tau}^{(k)} \mathrm{d}\mathcal{H}^{d-1}. \tag{3.87}$$

We have already seen in the proof of Lemma 3.19 (see in particular (3.76)) that

$$\sum_{k=1}^{T/\tau} \tau \|h_{\varepsilon,\tau}^{(k)}\|_{L^1(\Omega)} \leqq C\varepsilon^{\alpha}.$$

Consequently, by (3.81)–(3.82) for $q = r = 1$ and (3.87) the desired bound (3.83) follows. □

Remark 3.21. For $\alpha \in [1, 2)$, by means of Remark 3.17 we obtain a stronger bound on the temperature: given $q = \frac{2}{\alpha} + \frac{4}{\alpha d}$ and $r = 2\frac{d+2}{\alpha d+2}$, we can find a constant C independent of ε and τ such that

$$\sum_{k=1}^{T/\tau} \tau \int_\Omega |\theta_{\varepsilon,\tau}^{(k)}|^q \mathrm{d}x \leqq C\varepsilon^{\alpha q}, \quad \sum_{k=1}^{T/\tau} \tau \int_\Omega |\nabla\theta_{\varepsilon,\tau}^{(k)}|^r \mathrm{d}x \leqq C\varepsilon^{\alpha r}. \tag{3.88}$$

This can be seen as follows: We start with the second bound. In this regard, by a For $\alpha = 1$, this directly follows from (3.69), where we recall $\overline{\mu}_{\varepsilon,\tau} = \varepsilon^{-\alpha}\overline{\theta}_{\varepsilon,\tau}$. Let $\alpha \in (1, 2)$. Note that $r \in [1, 2)$ and let $m := r(1 - \frac{1}{\alpha})$. Employing a standard truncation and approximation argument we can assume, without loss of generality, that $\overline{\mu}_{\varepsilon,\tau} \in L^\infty(I \times \Omega)$. Then, by (3.69) and Hölder's inequality with powers $\frac{2}{2-r}$ and $\frac{2}{r}$ we derive that

$$\begin{aligned}
\|\nabla\overline{\mu}_{\varepsilon,\tau}\|^r_{L^r(I\times\Omega)} &= \int_0^T \int_\Omega (1+\overline{\mu}_{\varepsilon,\tau})^m \frac{|\nabla\overline{\mu}_{\varepsilon,\tau}|^r}{(1+\overline{\mu}_{\varepsilon,\tau})^m} \mathrm{d}x\mathrm{d}t \\
&\leqq \|1+\overline{\mu}_{\varepsilon,\tau}\|^m_{L^{\frac{2m}{2-r}}(I\times\Omega)} \left(\int_0^T \int_\Omega \frac{|\nabla\overline{\mu}_{\varepsilon,\tau}|^2}{(1+\overline{\mu}_{\varepsilon,\tau})^{2(1-\frac{1}{\alpha})}} \mathrm{d}x\mathrm{d}t\right)^{\frac{r}{2}} \\
&\leqq C\|1+\overline{\mu}_{\varepsilon,\tau}\|^m_{L^{\frac{2m}{2-r}}(I\times\Omega)}.
\end{aligned} \tag{3.89}$$

With $r = 2\frac{d+2}{\alpha d+2} = 2 - 2\frac{\alpha d}{\alpha d+2}(1 - \frac{1}{\alpha})$ we can use the anisotropic Gagliardo–Nirenberg interpolation inequality (see e.g. [30, Lemma 4.2]) with $\theta = \frac{\alpha d}{\alpha d+2}$, $s = p = \frac{r}{\theta}$, $s_1 = \infty$, $s_2 = p_2 = r$, and $p_1 = \frac{2}{\alpha}$ to get

$$\begin{aligned}
\|1+\overline{\mu}_{\varepsilon,\tau}\|_{L^{\frac{2m}{2-r}}(I\times\Omega)} &\leqq C\|1+\overline{\mu}_{\varepsilon,\tau}\|^{\frac{2}{\alpha d+2}}_{L^\infty(I;L^{\frac{2}{\alpha}}(\Omega))} \\
&\quad\times \left(\|1+\overline{\mu}_{\varepsilon,\tau}\|_{L^\infty(I;L^{\frac{2}{\alpha}}(\Omega))} + \|\nabla\overline{\mu}_{\varepsilon,\tau}\|_{L^r(I\times\Omega)}\right)^{\frac{\alpha d}{\alpha d+2}},
\end{aligned} \tag{3.90}$$

where we use $\frac{r}{\theta} = \frac{2m}{2-r}$. Notice that by (3.48) and the energy bound in Lemma 3.18 we have that $\|1+\overline{\mu}_{\varepsilon,\tau}\|_{L^\infty(I;L^{2/\alpha}(\Omega))}$ is uniformly bounded in ε and τ. Hence, with (3.89) and $m\frac{\alpha d}{\alpha d+2} = \frac{(\alpha-1)d}{\alpha d+2}r$ we derive that

$$\|\nabla\overline{\mu}_{\varepsilon,\tau}\|^r_{L^r(I\times\Omega)} \leqq C\left(1 + \|\nabla\overline{\mu}_{\varepsilon,\tau}\|^{\frac{(\alpha-1)d}{\alpha d+2}r}_{L^r(I\times\Omega)}\right).$$

As $\frac{(\alpha-1)d}{\alpha d+2} < 1$, this shows the second bound in (3.88) for the case $\alpha \in (1, 2)$. The first estimate in (3.88) then follows from the second one and (3.90), where we use that $\frac{2m}{2-r} = \frac{r}{\theta} = q$.

4. Existence of Solutions in the Nonlinear Setting

In this section we pass from time-discrete to time-continuous solutions by letting $\tau \to 0$ and establish Proposition 2.5(ii). Notice that for the special case $\alpha = 2$ and $\varepsilon = 1$ this will lead to Theorem 2.3(ii). For the deformation and the momentum balance we can closely follow [33, Section 5], and therefore proofs are omitted or sketched only. For the limit passage in the heat equation, however, our arguments are different as we work without regularization terms, cf. Remark 2.4. We first use the a priori estimates on the interpolants in order to extract convergent subsequences. Afterwards, we pass to the limit in the discretized weak forms of the momentum balance and the heat equation. Here, the most delicate term is the dissipation rate ξ which is quadratic in $\dot{F}$. Therefore, strong convergence in $L^2(I; H^1(\Omega))$ for the strain rates is required. As before, we assume for simplicity that $T/\tau \in \mathbb{N}$. Moreover, without further notice, we suppose from now on that $\tau \in (0, \tau_0)$ and $\varepsilon \in (0, \varepsilon_0]$, where τ_0 and $\varepsilon_0 = \varepsilon_0(\alpha)$ are chosen such that all statements from Sects. 3.1–3.4 are satisfied. In particular, $\varepsilon_0 = 1$ for $\alpha = 2$. The corresponding time-discrete solutions are denoted by $y^{(0)}_{\varepsilon,\tau}, \ldots, y^{(T/\tau)}_{\varepsilon,\tau} \in \mathcal{Y}_{\mathbf{id}}$ and $\theta^{(0)}_{\varepsilon,\tau}, \ldots, \theta^{(T/\tau)}_{\varepsilon,\tau} \in L^2_+(\Omega)$. We recall the definition of the interpolations in (2.26) and employ similar notation for $\overline{\theta}_{\varepsilon,\tau}$, $\underline{\theta}_{\varepsilon,\tau}$, and $\hat{\theta}_{\varepsilon,\tau}$, as well as $\overline{w}_{\varepsilon,\tau}$, $\underline{w}_{\varepsilon,\tau}$, and $\hat{w}_{\varepsilon,\tau}$. All generic constants $C > 0$ are always assumed to be independent of τ and ε.

We start with the convergence of the deformations under vanishing time-discretization.

Lemma 4.1. (Convergence of deformations) *For each $\varepsilon \in (0, \varepsilon_0]$, we can find $y_\varepsilon \in L^\infty(I; \mathcal{Y}_{\mathbf{id}}) \cap H^1(I; H^1(\Omega; \mathbb{R}^d))$ with $y_\varepsilon(0, \cdot) = y_{0,\varepsilon}$ such that, up to a subsequence (not relabeled), it holds that*

$$\hat{y}_{\varepsilon,\tau} \overset{*}{\rightharpoonup} y_\varepsilon \text{ weakly* in } L^\infty(I; \mathcal{Y}_{\mathbf{id}}) \text{ and } \hat{y}_{\varepsilon,\tau} \rightharpoonup y_\varepsilon \text{ weakly in } H^1(I; H^1(\Omega; \mathbb{R}^d)), \tag{4.1a}$$

$$\nabla \hat{y}_{\varepsilon,\tau} \to \nabla y_\varepsilon \text{ in } L^\infty(I; L^\infty(\Omega; \mathbb{R}^{d\times d})) \tag{4.1b}$$

as $\tau \to 0$. In the first convergence of (4.1a), and in (4.1b), the same holds true if we replace $\hat{y}_{\varepsilon,\tau}$ by $\underline{y}_{\varepsilon,\tau}$ or $\overline{y}_{\varepsilon,\tau}$.

Proof. First, (4.1a) follows from the a priori estimates (3.70a), (3.70c) and by Banach's selection principle. For (4.1b), one uses the embedding $W^{2,p}(\Omega; \mathbb{R}^d) \subset C^{1,1-\frac{d}{p}}(\Omega; \mathbb{R}^d)$ to obtain a Hölder estimate in space and (3.70c) for a Hölder estimate in time. Then, by an interpolation estimate one can show that the sequence is bounded in $C^\gamma(I; C^{1,\gamma}(\Omega; \mathbb{R}^d))$ for some $\gamma > 0$, and the uniform convergence of the gradients follows then from the Arzelà-Ascoli theorem. We refer to [33, Proof of Proposition 5.1, Step 1] for more details. To conclude that the first convergences in (4.1a) and (4.1b) also hold for $\underline{y}_{\varepsilon,\tau}$ or $\overline{y}_{\varepsilon,\tau}$ one again uses (3.70b)–(3.70c) to see $\|\nabla\hat{y}_{\varepsilon,\tau} - \nabla\overline{y}_{\varepsilon,\tau}\|_{L^\infty(I;L^2(\Omega))} \leqq C\tau^{\frac{1}{2}}$. □

We proceed with the convergence of the temperatures.

Lemma 4.2. (Convergence of temperatures) *For each $\varepsilon \in (0, \varepsilon_0]$, there exists $\theta_\varepsilon \in L^1(I; W^{1,1}(\Omega))$ with $\theta_\varepsilon \geqq 0$ a.e. such that, up to a subsequence (not relabeled), it holds that*

$$\overline{\theta}_{\varepsilon,\tau} \rightharpoonup \theta_\varepsilon \quad \textit{and} \quad \overline{w}_{\varepsilon,\tau} \rightharpoonup w_\varepsilon \quad \textit{weakly in } L^r(I; W^{1,r}(\Omega)) \textit{ for any } r \in \left[1, \tfrac{d+2}{d+1}\right), \tag{4.2a}$$

$$\hat{\theta}_{\varepsilon,\tau} \to \theta_\varepsilon \quad \textit{and} \quad \hat{w}_{\varepsilon,\tau} \to w_\varepsilon \textit{ in } L^s(I \times \Omega) \textit{ for any } s \in \left[1, \tfrac{d+2}{d}\right), \tag{4.2b}$$

as $\tau \to 0$ where $w_\varepsilon := W^{\mathrm{in}}(\nabla y_\varepsilon, \theta_\varepsilon)$ for y_ε as in Lemma 4.1. In (4.2b), the same holds true if we replace $\hat{\theta}_{\varepsilon,\tau}$ with $\underline{\theta}_{\varepsilon,\tau}$ or $\overline{\theta}_{\varepsilon,\tau}$ and $\hat{w}_{\varepsilon,\tau}$ with $\underline{w}_{\varepsilon,\tau}$ or $\overline{w}_{\varepsilon,\tau}$, respectively.

Proof. The existence of the limit and the convergences in (4.2a) follow from the a priori bounds in Theorem 3.20 together with Banach's selection principle.

Let $t_0 \in (0, T)$ and $r \in [1, \frac{d+2}{d+1})$. By Theorem 3.20, $(\hat{w}_{\varepsilon,\tau})_\tau$ is bounded in

$$L^r([t_0, T]; W^{1,r}(\Omega)) \cap W^{1,1}([t_0, T]; W^{1,\infty}(\Omega)^*).$$

Hence, for any $\tilde{r} < r^* := \frac{rd}{d-r}$, due to the compact embedding $W^{1,r}(\Omega) \subset\subset L^{\tilde{r}}(\Omega)$, the Aubin-Lions' theorem shows that there exists $\hat{w}_\varepsilon \in L^r([t_0, T]; L^{\tilde{r}}(\Omega))$ such that $(\hat{w}_{\varepsilon,\tau})_\tau \to \hat{w}_\varepsilon$ in $L^r([t_0, T]; L^{\tilde{r}}(\Omega))$, up to taking a subsequence. We observe that $\hat{w}_\varepsilon = w_\varepsilon$. Indeed, it is elementary to check that by (3.83)

$$\begin{aligned}\|\hat{w}_{\varepsilon,\tau} - \overline{w}_{\varepsilon,\tau}\|_{L^1(I;W^{1,\infty}(\Omega)^*)} &\leqq \|\overline{w}_{\varepsilon,\tau} - \underline{w}_{\varepsilon,\tau}\|_{L^1(I;W^{1,\infty}(\Omega)^*)} \\ &\leqq \tau\|\dot{\hat{w}}_{\varepsilon,\tau}\|_{L^1(I;W^{1,\infty}(\Omega)^*)} \to 0\end{aligned} \tag{4.3}$$

as $\tau \to 0$. Next, we show that the convergence $\hat{w}_{\varepsilon,\tau} \to w_\varepsilon$ in $L^r([t_0, T]; L^{\tilde{r}}(\Omega))$ as $\tau \to 0$ can be improved to convergence in $L^s([t_0, T]; L^s(\Omega))$ for any exponent $s \in [1, \frac{d+2}{d})$. To this end, we will interpolate with the bound

$$\|w_\varepsilon\|_{L^\infty(I;L^1(\Omega))} \leqq \sup_{\tau>0} \|\overline{w}_{\varepsilon,\tau}\|_{L^\infty(I;L^1(\Omega))} < \infty, \tag{4.4}$$

which follows from (3.70d). Fix $s \in (1, \frac{d+2}{d})$ and consider $r \in (1, \frac{d+2}{d+1})$, $\tilde{r} \in (1, r^*)$, both to be specified later. Now, as $\lim_{r\to\frac{d+2}{d+1}} \frac{rd}{d-r} \geqq \frac{d+2}{d} > s$, notice that for r, $\tilde{r}$ large enough it holds that $\lambda := \frac{\tilde{r}-s}{s(\tilde{r}-1)} \in (0, 1)$. Writing $v_\tau := \hat{w}_{\varepsilon,\tau} - w_\varepsilon$ for shorthand and using Hölder's inequality in the integral over Ω with powers $q_1 = \frac{\tilde{r}-1}{\tilde{r}-s}$ and $q_1' = \frac{\tilde{r}-1}{s-1}$, we derive that

$$\begin{aligned}\|v_\tau\|^s_{L^s([t_0,T];L^s(\Omega))} &= \int_{t_0}^T \int_\Omega |v_\tau|^{\lambda s}|v_\tau|^{(1-\lambda)s}\mathrm{d}x\mathrm{d}t \\ &\leqq \int_{t_0}^T \left(\int_\Omega |v_\tau|\mathrm{d}x\right)^{\frac{1}{q_1}} \left(\int_\Omega |v_\tau|^{\tilde{r}}\mathrm{d}x\right)^{\frac{1}{q_1'}} \mathrm{d}t,\end{aligned} \tag{4.5}$$

where we have used $\lambda s q_1 = 1$ and $(1-\lambda)sq_1' = \tilde{r}$. Let $q_2 := \frac{r(\tilde{r}-1)}{\tilde{r}(s-1)}$ and notice that

$$\lim_{r\to\frac{d+2}{d+1}}\lim_{\tilde{r}\to r^*} q_2 = \lim_{r\to\frac{d+2}{d+1}} \frac{r(d+1)-d}{d(s-1)} = \frac{2}{d(s-1)} > 1$$

where the last inequality is due to $s < 1+\frac{2}{d}$. Hence, by possibly increasing r and $\tilde{r}$ we can assure that $q_2 > 1$. We denote by q_2' the conjugate of q_2. Consequently, by $\hat{w}_{\varepsilon,\tau} \to \hat{w}_\varepsilon$ in $L^r([t_0,T]; L^{\tilde{r}}(\Omega))$ as $\tau\to 0$, by (4.4), and by Hölder's inequality in the integral in (4.5) over $[t_0,T]$ with powers q_2' and q_2 we get

$$\begin{aligned}
\|v_\tau\|^s_{L^s([t_0,T];L^s(\Omega))} &\leqq \left(\int_{t_0}^T \left(\int_\Omega |v_\tau|\,\mathrm{d}x\right)^{\frac{q_2'}{q_1}}\mathrm{d}t\right)^{\frac{1}{q_2'}} \left(\int_{t_0}^T \left(\int_\Omega |v_\tau|^{\tilde{r}}\,\mathrm{d}x\right)^{\frac{r}{\tilde{r}}}\mathrm{d}t\right)^{\frac{1}{q_2}} \\
&\leqq \left(2\sup\nolimits_{\tau>0}\|\overline{w}_{\varepsilon,\tau}\|_{L^\infty(I;L^1(\Omega))}\right)^{\frac{1}{q_1}} \|\hat{w}_{\varepsilon,\tau}-w_\varepsilon\|^{\frac{r}{q_2}}_{L^r([t_0,T];L^{\tilde{r}}(\Omega))} \\
&\to 0 \text{ as } \tau\to 0.
\end{aligned} \tag{4.6}$$

Sending $t_0\to 0$ and using (3.81), this shows (4.2b) for the sequence $(\hat{w}_{\varepsilon,\tau})_\tau$. To obtain the same convergence for $\overline{w}_{\varepsilon,\tau}$ and $\underline{w}_{\varepsilon,\tau}$, we use a more general version of Aubin-Lions for time-derivatives as measures, see Corollary 7.9 in [38]. To this end it suffices to see that $\overline{w}_{\varepsilon,\tau}$ and $\underline{w}_{\varepsilon,\tau}$ are bounded in $L^r([t_0,T]; W^{1,r}(\Omega)) \cap BV([t_0,T]; W^{1,\infty}(\Omega)^*)$, and then by repeating (4.5)–(4.6) we get (4.2b) for $\overline{w}_{\varepsilon,\tau}$ and $\underline{w}_{\varepsilon,\tau}$, up to taking a subsequence.

It remains to show (4.2b) for the three different interpolations of the temperatures. In view of (2.13), for any $F\in GL^+(d)$, the map $W^{\mathrm{in}}(F,\cdot)$ is invertible with $\frac{\mathrm{d}}{\mathrm{d}\theta}(W^{\mathrm{in}}(F,\cdot)^{-1}) \leqq \frac{1}{c_0}$. Thus, from the definition $\overline{w}_{\varepsilon,\tau} = W^{\mathrm{in}}(\overline{y}_{\varepsilon,\tau},\overline{\theta}_{\varepsilon,\tau})$ we get $\overline{\theta}_{\varepsilon,\tau} = W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon,\tau},\cdot)^{-1}(\overline{w}_{\varepsilon,\tau})$. Setting $\theta_\varepsilon := W^{\mathrm{in}}(\nabla y_\varepsilon,\cdot)^{-1}(w_\varepsilon)$, by (4.1b) for $\overline{y}_{\varepsilon,\tau}$ and by $\overline{w}_{\varepsilon,\tau}\to w_\varepsilon$ in $L^s(I\times\Omega)$ (see (4.2b)), we get

$$\overline{\theta}_{\varepsilon,\tau} = W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon,\tau},\cdot)^{-1}(\overline{w}_{\varepsilon,\tau}) \to W^{\mathrm{in}}(\nabla y_\varepsilon,\cdot)^{-1}(w_\varepsilon) = \theta_\varepsilon \quad \text{in } L^s(I\times\Omega).$$

The convergence for $(\underline{\theta}_{\varepsilon,\tau})_\tau$ follows in a similar fashion. Lastly, combining the convergence of $(\overline{\theta}_{\varepsilon,\tau})_\tau$ and $(\underline{\theta}_{\varepsilon,\tau})_\tau$ we obtain (4.2b) also for $\hat{\theta}_{\varepsilon,\tau}$. ◻

Remark 4.3. (i) Note that (4.2a) does not holds in general for $\hat{\theta}_{\varepsilon,\tau}$, $\underline{\theta}_{\varepsilon,\tau}$, $\hat{w}_{\varepsilon,\tau}$, and $\underline{w}_{\varepsilon,\tau}$ as we did not assume Sobolev regularity for the initial datum $\theta_{0,\varepsilon}\in L^2_+(\Omega)$. Yet, the statement could be obtained on any subinterval $I'\subset I$ with $0\notin I'$.

(ii) The result only relies on the a priori bounds in Theorem 3.20. Consequently, the same convergence result holds true for the *rescaled temperature and rescaled internal energy*, namely along (interpolations of) the sequences $(\varepsilon_k^{-\alpha}\theta^{(k)}_{\varepsilon_k,\tau_k})_k$ and $(\varepsilon_k^{-\alpha}w^{(k)}_{\varepsilon_k,\tau_k})_k$ for sequences $(\varepsilon_k,\tau_k)_k$ with $\varepsilon_k\to 0$ as $k\to\infty$. Namely, the proof of $\varepsilon_k^{-\alpha}\overline{w}_{\varepsilon_k,\tau_k}\to\tilde{w}$ in $L^s(I\times\Omega)$ for some $\tilde{w}$ is the same, taking the a priori bounds in (3.70d) and Theorem 3.20 into account. In view of (C.6), $\bar{c}_V = c_V(\mathbf{Id},0)$ exists and by the third estimate in (C.5) we have $\bar{c}_V \geqq c_0$. Hence, we can define $\tilde{\theta} := \tilde{w}/\bar{c}_V$. Furthermore, by

$W^{\mathrm{in}}(F,0) = 0$ for all $F \in GL^+(d)$, $c_V = \partial_\theta W^{\mathrm{in}}$ (see (2.13)), and the Fundamental Theorem of Calculus we find

$$\begin{aligned}\overline{\theta}_{\varepsilon_k,\tau_k} &= W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon_k,\tau_k},\cdot)^{-1}(\overline{w}_{\varepsilon_k,\tau_k})\\ &= \int_0^{\overline{w}_{\varepsilon_k,\tau_k}} c_V(\nabla\overline{y}_{\varepsilon_k,\tau_k}, W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon_k,\tau_k},\cdot)^{-1}(s))^{-1}\mathrm{d}s\\ &= \varepsilon_k^\alpha \int_0^{\varepsilon_k^{-\alpha}\overline{w}_{\varepsilon_k,\tau_k}} c_V(\nabla\overline{y}_{\varepsilon_k,\tau_k}, W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon_k,\tau_k},\cdot)^{-1}(\varepsilon_k^\alpha s))^{-1}\mathrm{d}s, \qquad (4.7)\end{aligned}$$

where we changed coordinates in the last identity. Consequently, using the third inequality in (C.5), we can derive the bound

$$\begin{aligned}|\varepsilon_k^{-\alpha}\overline{\theta}_{\varepsilon_k,\tau_k} - \bar{c}_V^{-1}\tilde{w}| &= \Big| \int_0^{\varepsilon_k^{-\alpha}\overline{w}_{\varepsilon_k,\tau_k}} c_V(\nabla\overline{y}_{\varepsilon_k,\tau_k}, W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon_k,\tau_k},\cdot)^{-1}(\varepsilon_k^\alpha s))^{-1}\mathrm{d}s\\ &\qquad - \int_0^{\tilde{w}} \bar{c}_V^{-1}\mathrm{d}s\Big|\\ &\leqq \frac{1}{c_0}|\varepsilon_k^{-\alpha}\overline{w}_{\varepsilon_k,\tau_k} - \tilde{w}| + f_k,\end{aligned}$$

where

$$f_k := \int_0^{\tilde{w}} |c_V(\nabla\overline{y}_{\varepsilon_k,\tau_k}, W^{\mathrm{in}}(\nabla\overline{y}_{\varepsilon_k,\tau_k},\cdot)^{-1}(\varepsilon_k^\alpha s))^{-1} - \bar{c}_V^{-1}|\mathrm{d}s.$$

It remains to show that $f_k \to 0$ in $L^s(I\times\Omega)$. By the third bound in (C.5) we see that $|f_k| \leqq \frac{2}{c_0}\tilde{w} \in L^s(I\times\Omega)$. Then, by (C.6) and the definition of $\bar{c}_V$, it follows that $f_k \to 0$ a.e. in $I\times\Omega$. Dominated Convergence yields the desired result. The same argument holds for the other interpolations.

(iii) In the case $\alpha = 1$, the convergence can be improved to $\varepsilon_k^{-1}\underline{\theta}_{\varepsilon_k,\tau_k} \to \tilde{\theta}$ in $L^2(I; L^2(\Omega))$. Indeed, by Remark 3.21 and $\theta_{0,\varepsilon} \in L^2_+(\Omega)$ we get that

$$\|\overline{\theta}_{\varepsilon_k,\tau_k}\|_{L^2(I\times\Omega)} + \|\underline{\theta}_{\varepsilon_k,\tau_k}\|_{L^2(I\times\Omega)} + \|\nabla\overline{\theta}_{\varepsilon_k,\tau_k}\|_{L^2(I\times\Omega)} \leqq C\varepsilon.$$

Then, the convergence in $L^2(I; L^2(\Omega))$ follows by repeating the argument above via Aubin-Lions' theorem, simply using the compact embedding $H^1(\Omega) \subset\subset L^2(\Omega)$.

We are ready to pass to the limit in the time-discrete mechanical evolution.

Proposition 4.4. (Convergence of the mechanical equation) *Let y_ε be as in Lemma 4.1 and θ_ε as in Lemma 4.2. Then, for any test-function $z \in C^\infty(I\times\overline{\Omega})$ with $z = 0$ on $I\times\Gamma_D$ we have that (2.19) holds.*

Proof. The statement is proved in [33, Proof of Proposition 5.1, Step 2] and we include a sketch for the reader's convenience. For $y \in \mathcal{Y}_{\mathbf{id}}$ we define a functional on $X := W^{2,p}(\Omega;\mathbb{R}^d)$ by

$$\langle \mathbf{H}(y), z\rangle = \int_0^T\int_\Omega \partial_G H(\nabla^2 y) \vdots \nabla^2 z.$$

Note that $\mathbf{H}$ is a hemicontinuous and monotone operator as H is convex. We further choose $b_{\varepsilon\tau},\ b_\varepsilon \in X^*$ such that (3.7) can be written as

$$\langle \mathbf{H}(\overline{y}_{\varepsilon,\tau}), z\rangle = \langle b_{\varepsilon\tau}, z\rangle \tag{4.8}$$

for all $z \in W^{2,p}_{\Gamma_D}(\Omega;\mathbb{R}^d)$, and (2.19) can be written as

$$\langle \mathbf{H}(y_\varepsilon), z\rangle = \langle b_\varepsilon, z\rangle \tag{4.9}$$

for all $z \in W^{2,p}_{\Gamma_D}(\Omega;\mathbb{R}^d)$. Note that (4.8) holds by Proposition 3.5, and that our goal is to confirm (4.9).

First, $b_{\varepsilon\tau} \overset{*}{\rightharpoonup} b_\varepsilon$ weakly* in X^* for $\tau \to 0$ as in each of the three terms of $b_{\varepsilon\tau}$ (i.e., $\partial_F W$, $\partial_{\dot F} R$, and $\ell^{(k)}_\tau$, respectively, see (3.7)) one can pass to the limit by using weak convergence of $(\nabla \dot{\hat{y}}_{\varepsilon,\tau})_\tau$ in $L^2(I;H^1(\Omega;\mathbb{R}^d))$ (see (4.1a)), uniform convergence of $(\nabla\overline{y}_{\varepsilon,\tau})_\tau$, $(\nabla \underline{y}_{\varepsilon,\tau})_\tau$ on $I\times\Omega$ (see (4.1b)), and pointwise a.e. convergence of $(\underline{\theta}_{\varepsilon,\tau})_\tau$ on $I\times\Omega$ (up to a subsequence, see (4.2b)). At this point, we use in particular that $\partial_{\dot F} R$ is linear in $\nabla\dot{\hat{y}}_{\varepsilon,\tau}$ and that $\partial_F W(\overline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})$ is bounded due to (W.1), (3.5), and (3.70a). Moreover, due to uniform convergence of the gradients we also have $\langle b_{\varepsilon\tau}, \overline{y}_{\varepsilon,\tau}\rangle \to \langle b_\varepsilon, y_\varepsilon\rangle$. We now use Minty's trick for the monotone operator $\mathbf{H}$: identity (4.8) and the convergences $\overline{y}_{\varepsilon,\tau} \rightharpoonup y_\varepsilon$ weakly in X, $b_{\varepsilon\tau} \overset{*}{\rightharpoonup} b_\varepsilon$ weakly* in X^*, and $\langle b_{\varepsilon\tau}, \overline{y}_{\varepsilon,\tau}\rangle \to \langle b_\varepsilon, y_\varepsilon\rangle$ imply $\mathbf{H}(y_\varepsilon) = b_\varepsilon$ as elements of X^*, i.e., (4.9) holds. □

For the limit passage in the time-discrete heat equation, we will need the strong convergence of the strain rates $(\nabla\dot{\hat{y}}_{\varepsilon,\tau})_\tau$ in $L^2(I;L^2(\Omega;\mathbb{R}^{d\times d}))$ since the dissipation rate $\xi(\nabla \underline{y}_{\varepsilon,\tau}, \nabla\dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})$ is quadratic in $\nabla\dot{\hat{y}}_{\varepsilon,\tau}$. Note that our a priori bounds currently only guarantee weak convergence. The next lemma improves this convergence.

Lemma 4.5. (Strong convergence of the strain rates) *For y_ε as in Lemma 4.1, we have that, up to taking a subsequence,*

$$\dot{\hat{y}}_{\varepsilon,\tau} \to \dot{y}_\varepsilon \textit{ strongly in } L^2(I;H^1(\Omega;\mathbb{R}^d)) \textit{ as } \tau \to 0. \tag{4.10}$$

Proof. The proof follows essentially by combining Steps 4 in the proof of [33, Proposition 5.1, Proposition 6.4]. We give the main steps here in our setting because we work completely without regularization. First, in the time-continuous setting, one derives the energy balance

$$\begin{aligned}\mathcal{M}(y_\varepsilon(T)) + 2\int_0^T \mathcal{R}(y_\varepsilon, \dot{y}_\varepsilon, \theta_\varepsilon)\mathrm{d}t &= \mathcal{M}(y_{0,\varepsilon}) + \varepsilon\int_0^T \langle \ell(t), \dot{y}_\varepsilon\rangle \mathrm{d}t\\ &\quad - \int_0^T\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_\varepsilon) : \nabla\dot{y}_\varepsilon \mathrm{d}x\mathrm{d}t,\end{aligned} \tag{4.11}$$

where we recall the notation in (2.5), (2.7), and (3.22). This follows by testing the momentum balance (2.19) derived in Proposition 4.4 with $\dot{y}_\varepsilon \in L^2(I;H^1(\Omega))$, employing (2.9), and using a chain rule for the Λ-convex functional $\mathcal{M}$, see [33, Proposition 3.6]. Our next goal is to show a similar balance in the time-discrete setting. To this end, we test the Euler–Lagrange equation (3.7) of the k-th mechanical step with $y_{\varepsilon,\tau}^{(k)} - y_{\varepsilon,\tau}^{(k-1)}$ to get that

$$\begin{aligned} 2\tau\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, \delta_\tau y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k-1)}) &= \tau\varepsilon\langle \ell_\tau^{(k)}, \delta_\tau y_{\varepsilon,\tau}^{(k)}\rangle - \tau\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k-1)}) : \delta_\tau \nabla y_{\varepsilon,\tau}^{(k)} \mathrm{d}x \\ &- \int_\Omega \partial_G H(\nabla^2 y_{\varepsilon,\tau}^{(k)}) \vdots (\nabla^2 y_{\varepsilon,\tau}^{(k)} - \nabla^2 y_{\varepsilon,\tau}^{(k-1)}) - \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(k)}) : (\nabla y_{\varepsilon,\tau}^{(k)} - \nabla y_{\varepsilon,\tau}^{(k-1)})\mathrm{d}x. \end{aligned} \tag{4.12}$$

By the Λ-convexity of $\mathcal{M}$ derived in [33, Proposition 3.2], we can find $\Lambda > 0$ depending on the energy bound in Lemma 3.18 and the bound in (3.1) but independent of ε, τ, and k such that

$$\begin{aligned} \mathcal{M}(y_{\varepsilon,\tau}^{(k-1)}) \geqq \mathcal{M}(y_{\varepsilon,\tau}^{(k)}) &- \Lambda\|\nabla y_{\varepsilon,\tau}^{(k-1)} - \nabla y_{\varepsilon,\tau}^{(k)}\|_{L^2(\Omega)}^2 \\ &+ \int_\Omega \partial_G H(\nabla^2 y_{\varepsilon,\tau}^{(k)}) \vdots (\nabla^2 y_{\varepsilon,\tau}^{(k-1)} - \nabla^2 y_{\varepsilon,\tau}^{(k)})\mathrm{d}x \\ &+ \int_\Omega \partial_F W^{\mathrm{el}}(\nabla y_{\varepsilon,\tau}^{(k)}) : (\nabla y_{\varepsilon,\tau}^{(k-1)} - \nabla y_{\varepsilon,\tau}^{(k)})\mathrm{d}x. \end{aligned}$$

Using this bound in (4.12) then leads to

$$\begin{aligned} &\mathcal{M}(y_{\varepsilon,\tau}^{(k)}) - \mathcal{M}(y_{\varepsilon,\tau}^{(k-1)}) + 2\tau\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, \delta_\tau y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k-1)}) - \Lambda\tau^2\|\delta_\tau\nabla y_{\varepsilon,\tau}^{(k)}\|_{L^2(\Omega)}^2 \\ &\quad\leqq \tau\varepsilon\langle \ell_\tau^{(k)}, \delta_\tau y_{\varepsilon,\tau}^{(k)}\rangle - \tau\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k)}, \theta_{\varepsilon,\tau}^{(k-1)}) : \delta_\tau\nabla y_{\varepsilon,\tau}^{(k)}\mathrm{d}x. \end{aligned}$$

Summing the above inequality over $k \in \{1, \ldots, T/\tau\}$ we arrive at a discrete analog of (4.11), namely,

$$\begin{aligned} &\mathcal{M}(\overline{y}_{\varepsilon,\tau}(T)) + 2\int_0^T \mathcal{R}(\underline{y}_{\varepsilon,\tau}, \dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\mathrm{d}t - \Lambda\tau\int_0^T\int_\Omega |\nabla\dot{\hat{y}}_{\varepsilon,\tau}|^2\mathrm{d}x\mathrm{d}t \\ &\quad\leqq \mathcal{M}(y_{0,\varepsilon}) + \varepsilon\int_0^T \langle \ell(t), \dot{\hat{y}}_{\varepsilon,\tau}\rangle\mathrm{d}t - \int_0^T\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla\overline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) : \nabla\dot{\hat{y}}_{\varepsilon,\tau}\mathrm{d}x\mathrm{d}t, \end{aligned} \tag{4.13}$$

where in the integral for the force terms we used the definition in (2.23). Up to selecting a further subsequence, we can suppose that the convergences in Lemma 4.1 and Lemma 4.2 hold true, and that $\underline{\theta}_{\varepsilon,\tau} \to \theta_\varepsilon$ pointwise a.e. in $I \times \Omega$, $\dot{\hat{y}}_{\varepsilon,\tau} \rightharpoonup \dot{y}_\varepsilon$ weakly in $L^2(I; H^1(\Omega; \mathbb{R}^d))$, and $\overline{y}_{\varepsilon,\tau}(T) \rightharpoonup y_\varepsilon(T)$ weakly in $W^{2,p}(\Omega)$ as $\tau \to 0$. This shows that

$$\begin{aligned} I_\varepsilon^{(1)} &:= \lim_{\tau\to 0}\left(\varepsilon\int_0^T \langle \ell(t), \dot{\hat{y}}_{\varepsilon,\tau}\rangle\mathrm{d}t - \int_0^T\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla\overline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) : \nabla\dot{\hat{y}}_{\varepsilon,\tau}\mathrm{d}x\mathrm{d}t\right) \\ &= \varepsilon\int_0^T \langle \ell(t), \dot{y}_\varepsilon\rangle\mathrm{d}t - \int_0^T\int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_\varepsilon) : \nabla\dot{y}_\varepsilon\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{4.14}$$

Setting

$$\dot{C}_{\varepsilon\tau} := (\nabla \dot{\hat{y}}_{\varepsilon,\tau})^T \nabla \underline{y}_{\varepsilon,\tau} + (\nabla \underline{y}_{\varepsilon,\tau})^T \nabla \dot{\hat{y}}_{\varepsilon,\tau}, \quad \dot{C}_\varepsilon := (\nabla \dot{y}_\varepsilon)^T \nabla y_\varepsilon + (\nabla y_\varepsilon)^T \nabla \dot{y}_\varepsilon$$

we see by (4.1) that $\dot{C}_{\varepsilon\tau} \rightharpoonup \dot{C}_\varepsilon$ weakly in $L^2(I \times \Omega; \mathbb{R}^{d\times d})$. Consequently, by the convexity of H and the fact that $\mathcal{R}$ is convex in $\dot{C} = \dot{F}^T F + F^T \dot{F}$, standard lower semicontinuity arguments (see also [17, Theorem 7.5]) imply that

$$\begin{aligned} I_\varepsilon^{(2)} &:= \liminf_{\tau\to 0} \mathcal{M}(\overline{y}_{\varepsilon,\tau}(T)) \geqq \mathcal{M}(y_\varepsilon(T)), \\ I_\varepsilon^{(3)} &:= \liminf_{\tau\to 0} \int_0^T \mathcal{R}(\underline{y}_{\varepsilon,\tau}, \dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\mathrm{d}t \geqq \int_0^T \mathcal{R}(y_\varepsilon, \dot{y}_\varepsilon, \theta_\varepsilon)\mathrm{d}t. \end{aligned} \tag{4.15}$$

Combining (4.11), (4.13), (4.14), and (4.15), and using that $\lim_{\tau\to 0} \tau \int_0^T \int_\Omega |\nabla \dot{\hat{y}}_{\varepsilon,\tau}|^2 \mathrm{d}x\mathrm{d}t = 0$ we get that

$$\begin{aligned} \mathcal{M}(y_\varepsilon(T)) + 2\int_0^T \mathcal{R}(y_\varepsilon, \dot{y}_\varepsilon, \theta_\varepsilon)\mathrm{d}t &= \mathcal{M}(y_{0,\varepsilon}) + I_\varepsilon^{(1)} \\ &\geqq I_\varepsilon^{(2)} + 2I_\varepsilon^{(3)} \geqq \mathcal{M}(y_\varepsilon(T)) + 2\int_0^T \mathcal{R}(y_\varepsilon, \dot{y}_\varepsilon, \theta_\varepsilon)\mathrm{d}t, \end{aligned}$$

and thus both inequalities in (4.15) are actually equalities. Consequently, we get by (2.7) and (D.1) that

$$\int_0^T \int_\Omega D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})\, \dot{C}_{\varepsilon\tau} : \dot{C}_{\varepsilon\tau} \mathrm{d}x\mathrm{d}t \to \int_0^T \int_\Omega D(C_\varepsilon, \theta_\varepsilon)\, \dot{C}_\varepsilon : \dot{C}_\varepsilon \mathrm{d}x\mathrm{d}t, \tag{4.16}$$

where we shortly write $C_{\varepsilon\tau} := (\nabla \underline{y}_{\varepsilon,\tau})^T \nabla \underline{y}_{\varepsilon,\tau}$ and $C_\varepsilon := (\nabla y_\varepsilon)^T \nabla y_\varepsilon$. Based on this, we show the strong convergence of the strain rates. By (D.2) it follows that

$$\begin{aligned} c_0 &\int_0^T \int_\Omega |\dot{C}_{\varepsilon\tau} - \dot{C}_\varepsilon|^2 \mathrm{d}x\mathrm{d}t \\ &\leqq \int_0^T \int_\Omega D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})(\dot{C}_{\varepsilon\tau} - \dot{C}_\varepsilon) : (\dot{C}_{\varepsilon\tau} - \dot{C}_\varepsilon)\mathrm{d}x\mathrm{d}t \\ &= \int_0^T \int_\Omega D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})\, \dot{C}_{\varepsilon\tau} : \dot{C}_{\varepsilon\tau}\mathrm{d}x\mathrm{d}t - 2\int_0^T \int_\Omega D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})\, \dot{C}_\varepsilon : \dot{C}_{\varepsilon\tau}\mathrm{d}x\mathrm{d}t \\ &\quad + \int_0^T \int_\Omega D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})\, \dot{C}_\varepsilon : \dot{C}_\varepsilon \mathrm{d}x\mathrm{d}t. \end{aligned}$$

By a weak-strong convergence argument and (4.1) we get that $\dot{C}_{\varepsilon\tau} \rightharpoonup \dot{C}_\varepsilon$ weakly in $L^2(I; L^2(\Omega; \mathbb{R}^{d\times d}))$. Moreover, by (D.2), $D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})$ is uniformly bounded and $D(C_{\varepsilon\tau}, \underline{\theta}_{\varepsilon,\tau})\dot{C}_\varepsilon$ converges to $D(C_\varepsilon, \theta_\varepsilon)\, \dot{C}_\varepsilon$ strongly in $L^2(\Omega; \mathbb{R}^{d\times d})$. Thus, (4.16) and Dominated Convergence imply that

$$\lim_{\tau\to 0} \|\dot{C}_{\varepsilon\tau} - \dot{C}_\varepsilon\|_{L^2(I\times\Omega)} = 0. \tag{4.17}$$

It remains to show that $\nabla \dot{\hat{y}}_{\varepsilon,\tau} \to \nabla \dot{y}_\varepsilon$ strongly in $L^2(I; L^2(\Omega; \mathbb{R}^{d\times d}))$ as then (4.10) follows from Poincaré's inequality. By the uniform bound on the energy in

Lemma 3.18, we can apply the generalized Korn's inequality stated in Lemma 3.2 for a constant c depending only on the initial data and f, g, $\theta_\flat$, and T. This shows that

$$
\begin{aligned}
&c\|\nabla\dot{\hat{y}}_{\varepsilon,\tau} - \nabla\dot{y}_\varepsilon\|_{L^2(I\times\Omega)} \\
&\quad\leqq \|(\nabla\dot{\hat{y}}_{\varepsilon,\tau} - \nabla\dot{y}_\varepsilon)^T\nabla y_\varepsilon + (\nabla y_\varepsilon)^T(\nabla\dot{\hat{y}}_{\varepsilon,\tau} - \nabla\dot{y}_\varepsilon)\|_{L^2(I\times\Omega)} \\
&\quad\leqq \|(\nabla\dot{\hat{y}}_{\varepsilon,\tau})^T\nabla\underline{y}_{\varepsilon,\tau} + (\nabla\underline{y}_{\varepsilon,\tau})^T\nabla\dot{\hat{y}}_{\varepsilon,\tau} - (\nabla\dot{y}_\varepsilon)^T\nabla y_\varepsilon - (\nabla y_\varepsilon)^T\nabla\dot{y}_\varepsilon\|_{L^2(I\times\Omega)} \\
&\qquad + 2\|\nabla\dot{\hat{y}}_{\varepsilon,\tau}\|_{L^2(I\times\Omega)}\|\nabla\underline{y}_{\varepsilon,\tau} - \nabla y_\varepsilon\|_{L^\infty(I\times\Omega)}.
\end{aligned}
$$

Now, (4.1b), (4.17), and $\sup_{\tau>0}\|\nabla\dot{\hat{y}}_{\varepsilon,\tau}\|_{L^2(I\times\Omega)} < +\infty$ by (4.1a) show $\|\nabla\dot{\hat{y}}_{\varepsilon,\tau} - \nabla\dot{y}_\varepsilon\|_{L^2(I\times\Omega)} \to 0$ as $\tau \to 0$. This concludes the proof. □

The last step in the proof of Theorem 2.3(ii) and Proposition 2.5(ii) consists in passing to the limit of the thermal evolution.

Proposition 4.6. (Convergence of the heat-transfer equation) *Let y_ε be as in Lemma 4.1 and θ_ε as in Lemma 4.2. Then, for any test-function $\varphi \in C^\infty(I \times \overline{\Omega})$ with $\varphi(T) = 0$, we have that $(y_\varepsilon, \theta_\varepsilon)$ satisfies (2.20) with $\xi_\alpha^{\mathrm{reg}}$ in place of ξ.*

Proof. Suppose that we have already selected a subsequence such that Lemmas 4.1 and 4.2 apply. By possibly taking a further subsequence we can also assume that $\underline{\theta}_{\varepsilon,\tau} \to \theta_\varepsilon$ pointwise a.e. in $I \times \Omega$. Furthermore, let φ as in the statement. Summing the Euler–Lagrange equation (3.11) (for $\xi_\alpha^{\mathrm{reg}}$ in place of ξ) for each step and integrating by parts we get that

$$
\begin{aligned}
&\int_0^T\int_\Omega \mathcal{K}(\nabla\underline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\nabla\overline{\theta}_{\varepsilon,\tau}\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t + \kappa\int_0^T\int_\Gamma \overline{\theta}_{\varepsilon,\tau}\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega \Big(\xi_\alpha^{\mathrm{reg}}(\nabla\underline{y}_{\varepsilon,\tau}, \nabla\dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) + \partial_F W^{\mathrm{cpl}}(\nabla\underline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) : \nabla\dot{\hat{y}}_{\varepsilon,\tau}\Big)\varphi\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T\int_\Omega \hat{w}_{\varepsilon,\tau}\dot{\varphi}\,\mathrm{d}x\,\mathrm{d}t \\
&= \kappa\varepsilon^\alpha\int_0^T\int_\Gamma \overline{\theta}_{\flat,\tau}\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t + \int_\Omega W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon})\varphi(0)\,\mathrm{d}x,
\end{aligned}
\tag{4.18}
$$

where $\overline{\theta}_{\flat,\tau}(t) := \theta_{\flat,\tau}^{(k)}$ for $t \in ((k-1)\tau, k\tau]$ and $k \in \{1, \dots, T/\tau\}$. As $\theta_\flat \in W^{1,1}(I; L^2(\Gamma))$ we see $\|\overline{\theta}_{\flat,\tau} - \theta_\flat\|_{L^1(I;L^1(\Gamma))} \leqq \tau\|\dot{\theta}_\flat\|_{L^1(I;L^2(\Gamma))}$. Consequently,

$$
\int_0^T\int_\Gamma \overline{\theta}_{\flat,\tau}\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \to \int_0^T\int_\Gamma \theta_\flat\varphi\,\mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \text{ as } \tau \to 0. \tag{4.19}
$$

It thus remains to show that the left-hand side of the above equality converges towards the left-hand side of (2.20) (with $\xi_\alpha^{\mathrm{reg}}$ in place of ξ) as $\tau \to 0$. By Lemma 3.3

and our choice of φ we have $|\mathcal{K}(\nabla \underline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\nabla\varphi| \leqq C|\nabla\varphi|$ a.e. in $I \times \Omega$. Consequently, by the weak convergence of $(\overline{\theta}_{\varepsilon,\tau})_\tau$ in $L^r(I; W^{1,r}(\Omega))$, see (4.2), it follows that

$$\int_0^T \int_\Omega \mathcal{K}(\nabla \underline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\nabla\overline{\theta}_{\varepsilon,\tau} \cdot \nabla\varphi \mathrm{d}x + \kappa \int_0^T \int_\Gamma \overline{\theta}_{\varepsilon,\tau}\varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t$$
$$\to \int_0^T \int_\Omega \mathcal{K}(\nabla y_\varepsilon, \theta_\varepsilon)\nabla\theta_\varepsilon \cdot \nabla\varphi \mathrm{d}x + \kappa \int_0^T \int_\Gamma \theta_\varepsilon\varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t.$$

The strong convergence of $(\hat{w}_{\varepsilon,\tau})_\tau$ in $L^s(I \times \Omega)$ for some $s \in (1, \frac{d+2}{d})$, see (4.2b), leads to

$$-\int_0^T \int_\Omega \hat{w}_{\varepsilon,\tau}\dot{\varphi}\mathrm{d}x\mathrm{d}t \to -\int_0^T \int_\Omega w_\varepsilon\dot{\varphi}\mathrm{d}x\mathrm{d}t = -\int_0^T \int_\Omega W^{\mathrm{in}}(\nabla y_\varepsilon, \theta_\varepsilon)\dot{\varphi}\mathrm{d}x\mathrm{d}t.$$

As in the proof of Lemma 4.5, (see (4.14)), we obtain

$$\int_0^T \int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla \underline{y}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) : \nabla\dot{\hat{y}}_{\varepsilon,\tau}\varphi \mathrm{d}x\mathrm{d}t \to \int_0^T \int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_\varepsilon) : \nabla\dot{y}_\varepsilon\varphi \mathrm{d}x\mathrm{d}t.$$

Note that by (D.2), (2.9), and by $\xi_\alpha^{\mathrm{reg}} \leqq \xi$, we have that

$$\xi_\alpha^{\mathrm{reg}}(\nabla \underline{y}_{\varepsilon,\tau}, \nabla\dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau}) \leqq 2C_0\left|(\nabla\dot{\hat{y}}_{\varepsilon,\tau})^T\nabla\underline{y}_{\varepsilon,\tau} + (\nabla\underline{y}_{\varepsilon,\tau})^T\nabla\dot{\hat{y}}_{\varepsilon,\tau}\right|^2.$$

By Lemma 4.5 and (4.1b) $((\nabla\dot{\hat{y}}_{\varepsilon,\tau})^T\nabla\underline{y}_{\varepsilon,\tau} + (\nabla\underline{y}_{\varepsilon,\tau})^T\nabla\dot{\hat{y}}_{\varepsilon,\tau})_\tau$ converges strongly in $L^2(I; L^2(\Omega; \mathbb{R}^{d\times d}))$. Consequently, we get that $\left(\xi_\alpha^{\mathrm{reg}}(\nabla \underline{y}_{\varepsilon,\tau}, \nabla\dot{\hat{y}}_{\varepsilon,\tau}, \underline{\theta}_{\varepsilon,\tau})\right)_\tau$ is equi-integrable. Using the pointwise convergence of $(\nabla\underline{y}_{\varepsilon,\tau})_\tau$ and $(\underline{\theta}_{\varepsilon,\tau})_\tau$ as well as the continuity of $\xi_\alpha^{\mathrm{reg}}$, we can also pass to the limit in the $\xi_\alpha^{\mathrm{reg}}$-term by an application of Vitali's convergence theorem. As we passed to the limit in each term, the proof is concluded. □

5. Passage to the Linearized System

This section is devoted to the proofs of Theorems 2.7–2.8. In the following, let $(\varepsilon_k)_k$ and $(\tau_k)_k$ be sequences with $\varepsilon_k \to 0$ and either $\tau_k = \tau$ constant or $\tau_k \to 0$. Suppose that initial data $(y_{0,\varepsilon_k}, \theta_{0,\varepsilon_k})$ as in (2.18) are given. For brevity, we denote the corresponding time-discrete interpolations by $\overline{y}_k := \overline{y}_{\varepsilon_k,\tau_k}$, $\underline{y}_k := \underline{y}_{\varepsilon_k,\tau_k}$, and $\hat{y}_k := \hat{y}_{\varepsilon_k,\tau_k}$, see (2.26). A similar shorthand notation is also used for the interpolation of the temperatures as well as the internal energies. Recall that the objects exist by Proposition 2.5(i). In a similar way, we denote the time-continuous solutions obtained in Proposition 2.5(ii) by $(y_{\varepsilon_k}, \theta_{\varepsilon_k})$. It will be useful to use a similar notation for the rescaled quantities: for time-discrete solutions we define

$$\overline{u}_k := \frac{\overline{y}_k - \mathbf{id}}{\varepsilon_k}, \quad \underline{u}_k := \frac{\underline{y}_k - \mathbf{id}}{\varepsilon_k}, \quad \hat{u}_k := \frac{\hat{y}_k - \mathbf{id}}{\varepsilon_k}, \quad \overline{\mu}_k := \frac{\overline{\theta}_k}{\varepsilon_k^\alpha}, \quad \underline{\mu}_k := \frac{\underline{\theta}_k}{\varepsilon_k^\alpha},$$

and for time-continuous solutions we let

$$u_{\varepsilon_k} := \frac{y_{\varepsilon_k} - \mathbf{id}}{\varepsilon_k}, \qquad \mu_{\varepsilon_k} := \frac{\theta_{\varepsilon_k}}{\varepsilon_k^{\alpha}}.$$

For any $v \in L^2(I; H^1(\Omega; \mathbb{R}^d))$ we denote the *symmetrized gradient* by $e(v) := \frac{1}{2}(\nabla v + \nabla v^T)$. Finally, all constants we encounter in this section are implicitly assumed to be independent of k.

We start with compactness results for the rescaled quantities which directly follow from the a priori estimates for the nonlinear system. Recall the definition of $H^1_{\Gamma_D}$ in (2.36).

Lemma 5.1. (Compactness for the rescaled displacements) *There exist $u, \tilde{u} \in H^1(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ with $u(0) = \tilde{u}(0) = u_0$ such that, up to possibly taking a subsequence, it holds that*

$$\hat{u}_k \to u \text{ in } L^\infty(I; L^2(\Omega; \mathbb{R}^d)), \quad \hat{u}_k \rightharpoonup u \text{ weakly in } H^1(I; H^1(\Omega; \mathbb{R}^d)), \tag{5.1}$$

$$u_{\varepsilon_k} \to \tilde{u} \text{ in } L^\infty(I; L^2(\Omega; \mathbb{R}^d)), \quad u_{\varepsilon_k} \rightharpoonup \tilde{u} \text{ weakly in } H^1(I; H^1(\Omega; \mathbb{R}^d)). \tag{5.2}$$

Moreover, if $\tau_k \to 0$, we also have

$$\overline{u}_k, \underline{u}_k \rightharpoonup u \text{ weakly in } H^1(I; H^1(\Omega; \mathbb{R}^d)). \tag{5.3}$$

Later, by uniqueness of the solution to the linear system, we will see that actually $u = \tilde{u}$.

Proof. By the definition of $\overline{u}_k$ and (3.70b) we derive for any $t \in I$ that $\|\overline{u}_k(t)\|_{H^1(\Omega)} = \varepsilon_k^{-1}\|\overline{y}_k - \mathbf{id}\|_{H^1(\Omega)} \leqq C$. For the other interpolations, we proceed in a similar fashion and get for all $t \in I$ that

$$\|\hat{u}_k(t)\|_{H^1(\Omega)} \leqq C. \tag{5.4}$$

Moreover, using Poincaré's inequality, (3.70c), and the definition of $\hat{u}_k$ we have that

$$\|\dot{\hat{u}}_k\|_{L^2(I;H^1(\Omega))} \leqq C\|\nabla\dot{\hat{u}}_k\|_{L^2(I;L^2(\Omega))} = \frac{1}{\varepsilon_k}\|\nabla\dot{\hat{y}}_k\|_{L^2(I\times\Omega)} \leqq C. \tag{5.5}$$

Combining (5.4)–(5.5) we discover that $(\hat{u}_k)_k$ is bounded in $L^\infty(I; H^1(\Omega; \mathbb{R}^d)) \cap H^1(I; H^1(\Omega; \mathbb{R}^d))$ and thus $(\hat{u}_k)_k$ is compact in $C(I; L^2(\Omega; \mathbb{R}^d))$ by the Aubin-Lions' theorem. This together with Banach's selection principle shows (5.1). Moreover, (5.3) follows from (5.5) and the definition of the interpolations. Finally, due to (5.1) and the fact that $\hat{u}_k \in H^1(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ with $\hat{u}_k(0) = u_0$ (see (2.1) and (2.18)), it directly follows that $u \in H^1(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ with $u(0) = u_0$.

We now show (5.2). To this end, suppose that for each $k \in \mathbb{N}$ the solution $(y_{\varepsilon_k}, \theta_{\varepsilon_k})$ is obtained as the limit of time discrete solutions $(\hat{y}_{\varepsilon_k\tau_l}, \hat{\theta}_{\varepsilon_k\tau_l})$ for a sequence $(\tau_l)_l$ converging to zero. Repeating (5.4)–(5.5) the corresponding rescaled quantities satisfy $\|\hat{u}_{\varepsilon_k\tau_l}\|_{L^\infty(I;H^1(\Omega))} \leqq C$ and $\|\dot{\hat{u}}_{\varepsilon_k\tau_l}\|_{L^2(I;H^1(\Omega))} \leqq C$ for a constant C independent of l. Then, using (2.27) we get

$$\|u_{\varepsilon_k}\|_{L^\infty(I;H^1(\Omega))} \leqq C \quad \text{and} \quad \|\dot{u}_{\varepsilon_k}\|_{L^2(I;H^1(\Omega))} \leqq C.$$

Now, (5.2) and the other properties of $\tilde{u}$ again follow by the Aubin-Lions' theorem. □

Lemma 5.2. (Compactness for the rescaled temperatures) *There exist* $\mu, \tilde{\mu} \in L^1(I; W^{1,1}(\Omega))$ *with* $\mu, \tilde{\mu} \geqq 0$ *such that, up to possibly taking a subsequence, for any* $s \in [1, \frac{d+2}{d})$ *and* $r \in [1, \frac{d+2}{d+1})$ *it holds that*

$$\overline{\mu}_k \to \mu \text{ in } L^s(I \times \Omega), \quad \overline{\mu}_k \rightharpoonup \mu \text{ weakly in } L^r(I; W^{1,r}(\Omega)), \tag{5.6}$$

$$\mu_{\varepsilon_k} \to \tilde{\mu} \text{ in } L^s(I \times \Omega), \quad \mu_{\varepsilon_k} \rightharpoonup \tilde{\mu} \text{ weakly in } L^r(I; W^{1,r}(\Omega)). \tag{5.7}$$

Moreover, if $\tau_k \to 0$*, we also have that*

$$\underline{\mu}_k,\ \hat{\mu}_k \to \mu \text{ in } L^s(I \times \Omega). \tag{5.8}$$

Later, by uniqueness of the solution to the linear system, we will see that actually $\mu = \tilde{\mu}$.

Proof. Let r and s be as in the statement. The proof of (5.6) relies on the a priori bounds on the internal energy in Theorem 3.20, i.e.,

$$\|\overline{\theta}_k\|_{L^r(I;W^{1,r}(\Omega))} + \|\overline{w}_k\|_{L^s(I\times\Omega)} + \|\overline{w}_k\|_{L^r(I;W^{1,r}(\Omega))} + \|\dot{\hat{w}}_k\|_{L^1(I;W^{1,\infty}(\Omega)^*)} \leqq C\varepsilon_k^\alpha. \tag{5.9}$$

In fact, we can follow closely the lines of the proof of Lemma 4.2, see Remark 4.3(ii). In particular, one first shows the convergence of the internal energies and then by (4.7) the convergence of the temperatures. Here, we also see that for $\tau_k \to 0$ property (4.3) implies (5.8).

To see (5.7), we suppose that for each $k \in \mathbb{N}$ the solution $(y_{\varepsilon_k}, \theta_{\varepsilon_k})$ is obtained as the limit of time discrete solutions $(\hat{y}_{\varepsilon_k\tau_l}, \hat{\theta}_{\varepsilon_k\tau_l})$ for a sequence $(\tau_l)_l$ converging to zero. By the above reasoning we obtain (5.9) for $\overline{\theta}_{\varepsilon_k\tau_l}$ in place of $\overline{\theta}_k$ and $\overline{w}_{\varepsilon_k\tau_l} := W^{\rm in}(\nabla\overline{y}_{\varepsilon_k\tau_l}, \overline{\theta}_{\varepsilon_k\tau_l})$ in place of $\overline{w}_k$. Then by (4.2) and the lower semicontinuity of the norms we get that

$$\|\theta_{\varepsilon_k}\|_{L^r(I;W^{1,r}(\Omega))} + \|w_{\varepsilon_k}\|_{L^s(I\times\Omega)} + \|w_{\varepsilon_k}\|_{L^r(I;W^{1,r}(\Omega))} \leqq C\varepsilon_k^\alpha,$$

where $w_{\varepsilon_k} := W^{\rm in}(\nabla y_{\varepsilon_k}, \theta_{\varepsilon_k})$. It now suffices to check that, also,

$$\|\dot{w}_{\varepsilon_k}\|_{L^1(I;W^{1,\infty}(\Omega)^*)} \leqq C\varepsilon_k^\alpha \tag{5.10}$$

holds, as then the statement follows by repeating the proof of Lemma 4.2, see again Remark 4.3(ii). To derive (5.10), we use (2.20) (for $\xi_\alpha^{\rm reg}$ in place of ξ) to get that $\dot{w}_{\varepsilon_k}$ coincides in the distributional sense with σ where for each $t \in I$ and each $\varphi \in C_c^\infty(\Omega)$ we set that

$$\begin{aligned}\langle\sigma(t), \varphi\rangle := \kappa \int_\Gamma \big(\varepsilon_k^\alpha\theta_\flat - \theta_{\varepsilon_k}\big)\varphi \,\mathrm{d}\mathcal{H}^{d-1}(x) \\ - \int_\Omega \Big(\mathcal{K}(\nabla y_{\varepsilon_k}, \theta_{\varepsilon_k})\nabla\theta_{\varepsilon_k} \cdot \nabla\varphi - \big(\xi_\alpha^{\rm reg}(\nabla y_{\varepsilon_k}, \nabla\dot{y}_{\varepsilon_k}, \theta_{\varepsilon_k}) \\ + \partial_F W^{\rm cpl}(\nabla y_{\varepsilon_k}, \theta_{\varepsilon_k}) : \nabla\dot{y}_{\varepsilon_k}\big)\varphi\Big)\mathrm{d}x,\end{aligned}$$

where all functions on the right-hand side are evaluated at $t \in I$. By passing to the limit $\tau \to 0$ in (3.70a)–(3.70c) and (3.81)–(3.82) we obtain the a priori bounds $\|y_{\varepsilon_k} - \mathbf{id}\|_{L^\infty(I;W^{1,\infty}(\Omega))} \leqq C\varepsilon_k^{2/p}$, $\|y_{\varepsilon_k} - \mathbf{id}\|_{H^1(I;H^1(\Omega))} \leqq C\varepsilon_k$, and $\|\theta_{\varepsilon_k}\|_{L^1(I;W^{1,1}(\Omega))} \leqq C\varepsilon_k^\alpha$. This along with (D.2), $\xi_\alpha^{\mathrm{reg}} \leqq \xi$, (3.4), (3.5), $\alpha \leqq 2$, and the trace estimate shows that $t \mapsto \|\sigma(t)\|_{W^{1,\infty}(\Omega)^*}$ lies in $L^1(I)$ with $\|\sigma\|_{L^1(I;W^{1,\infty}(\Omega)^*)} \leqq C\varepsilon_k^\alpha$. This concludes the proof of (5.10). □

We now proceed with the proofs of Theorems 2.7 and 2.8 which we split into two subsections.

5.1. Proof of Theorem 2.7

We will only prove Theorem 2.7(iii) as item (ii) of the statement can be obtained along similar lines by performing the linearization directly in the weak formulation (2.19)–(2.20) in place of the Euler–Lagrange equations (3.7) and (3.11). Note that the proof of Theorem 2.7(iii) will also imply the existence statement in Theorem 2.7(i). In this subsection, we also address the uniqueness of the solutions to the linearized system.

Proposition 5.3. (Linearization of the mechanical equation) *Let u and μ be given as in Lemmas 5.1–5.2. Then, for any $z \in C^\infty(I \times \overline{\Omega}; \mathbb{R}^d)$ with $z = 0$ on $I \times \Gamma_D$ we have that* (2.37) *holds.*

Proof. Let z be as in the statement. As $z \in W^{2,p}_{\Gamma_D}(\Omega; \mathbb{R}^d)$, we can multiply (3.7) with τ_k/ε_k and sum over all steps $1, \ldots, T/\tau$ to get that

$$\frac{1}{\varepsilon_k}\int_0^T\int_\Omega \Big(\partial_F W(\nabla\overline{y}_k, \underline{\theta}_k) + \partial_{\dot F} R(\nabla\underline{y}_k, \nabla\dot{\hat{y}}_k, \underline{\theta}_k)\Big) : \nabla z + \partial_G H(\nabla^2\overline{y}_k) \vdots \nabla^2 z \,\mathrm{d}x\,\mathrm{d}t$$
$$= \int_0^T \langle \overline{\ell}_{\tau_k}(t), z\rangle \,\mathrm{d}t, \tag{5.11}$$

where $\overline{\ell}_{\tau_k}(t) := \ell_{\tau_k}^{(l)}$ for $t \in ((l-1)\tau, l\tau]$ and $l \in \{1, \ldots, T/\tau\}$. Our goal now is to show that (2.37) arises as the limit of the above equation as $k \to \infty$. First, recalling (2.23) we can easily check that

$$\int_0^T \langle \overline{\ell}_{\tau_k}(t), z(t)\rangle \,\mathrm{d}t \to \int_0^T\int_\Omega f \cdot z + \int_0^T\int_{\Gamma_N} g \cdot z \,\mathrm{d}x\,\mathrm{d}t \tag{5.12}$$

as $k \to \infty$. By (H.3) for $\partial_G H$, (3.70a), and Hölder's inequality with powers $\frac{p}{p-1}$ and p, we derive that

$$\frac{1}{\varepsilon_k}\left|\int_0^T\int_\Omega \partial_G H(\nabla^2\overline{y}_k) \vdots \nabla^2 z \,\mathrm{d}x\,\mathrm{d}t\right| \leqq \frac{C_0}{\varepsilon_k}\int_0^T\int_\Omega |\nabla^2\overline{y}_k|^{p-1}|\nabla^2 z| \,\mathrm{d}x\,\mathrm{d}t$$
$$\leqq \frac{C_0}{\varepsilon_k}\int_0^T \|\nabla^2\overline{y}_k\|^{p-1}_{L^p(\Omega)}\|\nabla^2 z\|_{L^p(\Omega)} \,\mathrm{d}t \leqq C\varepsilon_k^{\frac{2(p-1)}{p}-1}$$
$$= C\varepsilon_k^{1-\frac{2}{p}} \to 0, \tag{5.13}$$

as $p > d \geqq 2$. We now address the coupling term. In view of $\partial_F W^{\mathrm{cpl}}(\mathbf{Id}, 0) = 0$, (3.70a), and (3.5), a Taylor expansion implies that

$$\begin{aligned}&\Big|\partial_F W^{\mathrm{cpl}}(\nabla \overline{y}_k, \underline{\theta}_k) - \Big(\partial_F^2 W^{\mathrm{cpl}}(\mathbf{Id}, 0)\varepsilon_k \nabla \overline{u}_k + \partial_{F\theta} W^{\mathrm{cpl}}(\mathbf{Id}, 0)(\varepsilon_k^\alpha \underline{\mu}_k \wedge 1)\Big)\Big|\\ &\quad \leqq C|\varepsilon_k \nabla \overline{u}_k|^2 + C\big(|\varepsilon_k^\alpha \underline{\mu}_k|^2 \wedge 1\big)\end{aligned} \tag{5.14}$$

pointwise a.e. in $I \times \Omega$. Thus, by (5.1) and (5.8), along with $t^2 \wedge 1 \leqq t^s$ for $t \geqq 0$ for some fixed $s \in (1, \frac{d+2}{d})$ it follows that

$$\begin{aligned}&\lim_{k\to\infty} \frac{1}{\varepsilon_k} \int_0^T \int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla \overline{y}_k, \underline{\theta}_k) : \nabla z \,\mathrm{d}x\,\mathrm{d}t\\ &\quad = \lim_{k\to\infty} \int_0^T \int_\Omega \big(\partial_F^2 W^{\mathrm{cpl}}(\mathbf{Id}, 0)\nabla \overline{u}_k + \varepsilon_k^{-1}\partial_{F\theta} W^{\mathrm{cpl}}(\mathbf{Id}, 0)(\varepsilon_k^\alpha \underline{\mu}_k \wedge 1)\big) : \nabla z \,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

Recalling $\partial_F^2 W^{\mathrm{cpl}}(\mathbf{Id}, 0) = 0$ (cf. (C.3)) and the definition of $\mathbb{B}^{(\alpha)}$ in (2.34) we find that

$$\lim_{k\to\infty} \frac{1}{\varepsilon_k} \int_0^T \int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla \overline{y}_k, \underline{\theta}_k) : \nabla z \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega \mathbb{B}^{(\alpha)}\mu : \nabla z \,\mathrm{d}x\,\mathrm{d}t. \tag{5.15}$$

By a Taylor expansion, (3.70a), and the fact that W^{el} is C^3 we have that

$$\Big|\varepsilon_k^{-1}\partial_F W^{\mathrm{el}}(\nabla \overline{y}_k) - \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})\nabla \overline{u}_k\Big| \leqq \frac{C}{\varepsilon_k}|\nabla \overline{y}_k - \mathbf{Id}|^2.$$

Integrating the above inequality over $I \times \Omega$ and using (3.70b) we get that

$$\begin{aligned}&\Big|\int_0^T \int_\Omega \Big(\varepsilon_k^{-1}\partial_F W^{\mathrm{el}}(\nabla \overline{y}_k) - \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})\nabla \overline{u}_k\Big) : \nabla z \,\mathrm{d}x\,\mathrm{d}t\Big|\\ &\quad \leqq C\varepsilon_k^{-1}\|\nabla \overline{y}_k - \mathbf{Id}\|^2_{L^2(I\times\Omega)} \leqq CT\varepsilon_k \to 0.\end{aligned} \tag{5.16}$$

By (2.8)

$$\begin{aligned}&\partial_{\dot F} R(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k) : \nabla z = 2\nabla \underline{y}_k (D(C_k, \underline{\theta}_k)\varepsilon_k \dot{C}_k) :\\ &\nabla z = \varepsilon_k \dot{C}_k : D(C_k, \underline{\theta}_k)(\nabla z^T \nabla \underline{y}_k + \nabla \underline{y}_k^T \nabla z),\end{aligned} \tag{5.17}$$

where

$$C_k := \nabla \underline{y}_k^T \nabla \underline{y}_k, \quad \dot{C}_k := \nabla \dot{\hat{u}}_k^T \nabla \underline{y}_k + \nabla \underline{y}_k^T \nabla \dot{\hat{u}}_k. \tag{5.18}$$

Note that the second identity is obtained by an elementary computation using the symmetries of D stated in (D.1). By (3.70a) and (5.1) we then see that

$$\dot{C}_k \rightharpoonup 2e(\dot{u}) \quad \text{weakly in } L^2(I \times \Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{5.19}$$

Using (D.2) we also have that

$$|D(C_k, \underline{\theta}_k)(\nabla z^T \nabla \underline{y}_k + \nabla \underline{y}_k^T \nabla z)| \leqq 2C_0 \|\nabla z\|_{L^\infty(\Omega)} \|\nabla \underline{y}_k\|_{L^\infty(\Omega)}.$$

Up to taking a subsequence (not relabeled), we can suppose that $\nabla \underline{y}_k \to \mathbf{Id}$ and $\underline{\theta}_k \to 0$ a.e. in $I \times \Omega$. Thus, Dominated Convergence implies

$$D(C_k, \underline{\theta}_k)(\nabla z^T \nabla \underline{y}_k + \nabla \underline{y}_k^T \nabla z) \to D(\mathbf{Id}, 0)(\nabla z + \nabla z^T) = 2D(\mathbf{Id}, 0)\nabla z$$

strongly in $L^2(I \times \Omega; \mathbb{R}^{d\times d})$. This along with (5.17) and (5.19) leads to

$$\varepsilon_k^{-1} \int_0^T \int_\Omega \partial_{\dot{F}} R(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k) : \nabla z \mathrm{d}x\mathrm{d}t \to \int_0^T \int_\Omega 4D(\mathbf{Id}, 0)e(\dot{u}) : \nabla z \mathrm{d}x\mathrm{d}t. \tag{5.20}$$

Recalling the definition of $\mathbb{C}_D$ and $\mathbb{C}_W$ in (2.32), as well as collecting (5.12), (5.13), (5.15), (5.16), and (5.20) we conclude the proof. $\square$

Similarly as in Sect. 4, for the limit passage in the heat-transfer equation, we will need the strong convergence of the strain rates $(\nabla \dot{\hat{u}}_k)_k$ in $L^2(I; L^2(\Omega; \mathbb{R}^{d\times d}))$ since the dissipation rate is quadratic in $\nabla \dot{\hat{u}}_k$. We now improve the compactness in Lemma 5.1 as follows. At this state, we need the additional assumption (H.4) which combined with the bound on $\partial_G H(G)$ from (H.3) leads to

$$|H(G)| \leqq C_0 |G|^p \quad \text{for all } G \in \mathbb{R}^{d\times d\times d}. \tag{5.21}$$

Lemma 5.4. (Strong convergence of the rescaled strains and strain rates) *With u as in Lemma 5.1, up to possibly taking a subsequence, we have that*

$$\begin{aligned} &\hat{u}_k(t) \to u(t) \text{ strongly in } H^1(\Omega; \mathbb{R}^d) \text{ for all } t \in I, \\ &\nabla \dot{\hat{u}}_k \to \nabla \dot{u} \text{ strongly in } L^2(I; L^2(\Omega; \mathbb{R}^{d\times d})). \end{aligned} \tag{5.22}$$

The first convergence also holds with $\overline{u}_k$ or $\underline{u}_k$ in place of $\hat{u}_k$.

Proof. Step 1 (Lower bounds for elastic energy and dissipation) Suppose we have already selected a subsequence so that the convergences of Lemma 5.1 as well as Lemma 5.2 hold true. Recall the definition of $\mathcal{M}_{\varepsilon_k}$ before (3.48). For convenience, for any $v \in H^1(\Omega; \mathbb{R}^d)$, we define

$$\overline{\mathcal{M}}_0(v) := \frac{1}{2}\int_\Omega \mathbb{C}_W e(v) : e(v) \mathrm{d}x,$$

where $\mathbb{C}_W = \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})$ is as in (2.32). Let us fix an arbitrary $t \in I$. By the non-negativity of H, a Taylor expansion, and (3.70a) we derive that

$$\begin{aligned} \mathcal{M}_{\varepsilon_k}(\overline{y}_k(t)) &\geqq \varepsilon_k^{-2} \int_\Omega W^{\mathrm{el}}(\nabla \overline{y}_k(t))\mathrm{d}x \\ &\geqq \frac{1}{2}\int_\Omega \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})\nabla \overline{u}_k(t) : \nabla \overline{u}_k(t) - C\int_\Omega |\overline{y}_k(t) - \mathbf{Id}||\nabla \overline{u}_k(t)|^2 \mathrm{d}x \\ &\geqq \frac{1}{2}\int_\Omega \partial_F^2 W^{\mathrm{el}}(\mathbf{Id})\nabla \overline{u}_k(t) : \nabla \overline{u}_k(t) - C\varepsilon_k^{2/p}\int_\Omega |\nabla \overline{u}_k(t)|^2 \mathrm{d}x. \end{aligned} \tag{5.23}$$

Consequently, using (5.3), it follows that

$$I_1 := \liminf_{k\to\infty} \mathcal{M}_{\varepsilon_k}(\overline{y}_k(t)) \geqq \liminf_{k\to\infty} \overline{\mathcal{M}}_0(\overline{u}_k(t)) \geqq \overline{\mathcal{M}}_0(u(t)). \tag{5.24}$$

Let C_k and $\dot{C}_k$ be as in (5.18). In (5.19) we have seen that $\dot{C}_k \rightharpoonup 2e(\dot{u})$ weakly in $L^2(I \times \Omega; \mathbb{R}^{d\times d})$. This along with the definition in (2.9), $\mathbb{C}_D = 4D(\mathbf{Id}, 0)$, the pointwise convergences of $(\nabla \underline{y}_k)_k$ and $(\underline{\theta}_k)_k$, and standard lower semicontinuity arguments (see also [17, Theorem 7.5]) show

$$\begin{aligned} I_2 &:= \liminf_{k\to\infty} \varepsilon_k^{-2} \int_0^t \int_\Omega \xi(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k)\mathrm{d}x\mathrm{d}s \\ &= \liminf_{k\to\infty} \int_0^t \int_\Omega D(C_k, \underline{\theta}_k)\dot{C}_k : \dot{C}_k \mathrm{d}x\mathrm{d}s \\ &\geqq \int_0^t \int_\Omega \mathbb{C}_D e(\dot{u}) : e(\dot{u})\mathrm{d}x\mathrm{d}s. \end{aligned} \tag{5.25}$$

Step 2 (Convergence of elastic energies and dissipations) Our next goal is to show the reverse inequalities for the lim sup. To this end, we draw ideas from the proof of Lemma 4.5 and compare an energy balance on the nonlinear time-discrete level with a time-continuous energy balance in the linearized setting. First, recall from (4.13) that for $K \in \mathbb{N}$ with $K\tau_k \in [t, t+\tau_k)$ it holds that

$$\begin{aligned} &\mathcal{M}_{\varepsilon_k}(\overline{y}_k(K\tau_k)) + \varepsilon_k^{-2} \int_0^{K\tau_k} \int_\Omega \xi(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k)\mathrm{d}x\mathrm{d}s - \tau_k \Lambda \int_0^{K\tau_k} \int_\Omega |\nabla \dot{\hat{u}}_k|^2 \mathrm{d}x\mathrm{d}s \\ &\quad \leqq \mathcal{M}_{\varepsilon_k}(y_{0,\varepsilon_k}) + \frac{1}{\varepsilon_k} \int_0^{K\tau_k} \langle \ell(s), \dot{\hat{y}}_k(s)\rangle \mathrm{d}s - \int_0^{K\tau_k} \int_\Omega \varepsilon_k^{-1} \partial_F W^{\mathrm{cpl}}(\nabla \overline{y}_k, \underline{\theta}_k) : \nabla \dot{\hat{u}}_k \mathrm{d}x\mathrm{d}s, \end{aligned} \tag{5.26}$$

where $\Lambda > 0$ does not depend on k. Here, we also used (2.9) to replace R by ξ. Now, in a similar fashion, testing (2.37) with $z = \dot{u}$ we see that

$$\overline{\mathcal{M}}_0(u(t)) - \overline{\mathcal{M}}_0(u_0) + \int_0^t \int_\Omega \big(\mathbb{C}_D e(\dot{u}) : e(\dot{u}) + \mu \mathbb{B}^{(\alpha)} : \nabla \dot{u}\big)\mathrm{d}x\mathrm{d}s = \int_0^t \langle \ell(s), \dot{u}(s)\rangle \mathrm{d}s. \tag{5.27}$$

We now address the convergence of the various terms. First of all, by (5.1) we clearly have that

$$\frac{1}{\varepsilon_k} \int_0^{K\tau_k} \langle \ell(s), \dot{\hat{y}}_k(s)\rangle \mathrm{d}s = \int_0^{K\tau_k} \langle \ell(s), \dot{\hat{u}}_k(s)\rangle \mathrm{d}s \to \int_0^t \langle \ell(s), \dot{u}(s)\rangle \mathrm{d}s. \tag{5.28}$$

For $\alpha = 1$, by arguing similarly as in (5.14)–(5.15), and using (3.70a) as well as $\underline{\mu}_k \to \mu$ strongly in $L^2(I \times \Omega)$, by Remark 4.3(iii) we find that

$$I_3 := \lim_{k\to\infty} \frac{1}{\varepsilon_k} \int_0^{K\tau_k} \int_\Omega \partial_F W^{\mathrm{cpl}}(\nabla \overline{y}_k, \underline{\theta}_k) : \nabla \dot{\hat{u}}_k \mathrm{d}x\mathrm{d}s = \int_0^t \int_\Omega \mathbb{B}^{(\alpha)} \mu : \nabla \dot{u} \mathrm{d}x\mathrm{d}t, \tag{5.29}$$

where we also used the definition of $\mathbb{B}^{(\alpha)}$ in (2.34). For $\alpha \in (1, 2]$, (5.29) also holds (with $\mathbb{B}^{(\alpha)} = 0$), since by Remark 3.21 we find that $\underline{\mu}_k$ is bounded in $L^q(I; L^q(\Omega))$

for some $q \in (2/\alpha, 2]$, and therefore using $t \wedge 1 \leqq t^{q/2}$ for $t \geqq 0$ and Young's inequality with constant $\varepsilon_k^{\alpha q/2}$ we get that

$$\begin{aligned}&\varepsilon_k^{-1} \int_0^T \int_\Omega \big|(\varepsilon_k^\alpha \underline{\mu}_k) \wedge 1\big| |\nabla \dot{\hat{u}}_k| \mathrm{d}x\mathrm{d}t \\ &\quad \leqq \varepsilon_k^{-1}\Big(\varepsilon_k^{-\alpha q/2} \|\varepsilon_k^\alpha \underline{\mu}_k\|^q_{L^q(I\times\Omega)} + \varepsilon_k^{\alpha q/2} \|\nabla \dot{\hat{u}}_k\|^2_{L^2(I\times\Omega)}\Big) \to 0. \qquad (5.30)\end{aligned}$$

Eventually, we get that

$$\lim_{k\to\infty} \mathcal{M}_{\varepsilon_k}(y_{0,\varepsilon_k}) = \lim_{k\to\infty} \mathcal{M}_{\varepsilon_k}(\mathbf{id} + \varepsilon_k u_0) = \overline{\mathcal{M}}_0(u_0). \qquad (5.31)$$

In fact, for the convergence of the elastic energy we repeat the Taylor expansion in (5.23) (with equality), and for the second-gradient term we get by (5.21), $u_0 \in W^{2,p}(\Omega; \mathbb{R}^d)$, and $p > 2$ that

$$\varepsilon_k^{-2}\Big|\int_\Omega H(\varepsilon_k \nabla^2 u_0)\mathrm{d}x\Big| \leqq C\varepsilon_k^{p-2} \int_\Omega |\nabla^2 u_0|^p \mathrm{d}x \leqq C\varepsilon_k^{p-2} \to 0.$$

Combining (5.26)–(5.27), $K\tau_k \geqq t$, the convergences (5.24), (5.25), (5.28), (5.29), and (5.31), as well as using that $\tau_k \int_0^{K\tau_k} \int_\Omega |\nabla \dot{\hat{u}}_k|^2 \mathrm{d}x\mathrm{d}s \to 0$ as $\tau_k \to 0$ we get that

$$\begin{aligned}&\overline{\mathcal{M}}_0(u(t)) + \int_0^t \int_\Omega \big(\mathbb{C}_D e(\dot{u}) : e(\dot{u}) + \mu \mathbb{B}^{(\alpha)} : \nabla \dot{u}\big)\mathrm{d}x\mathrm{d}s = \overline{\mathcal{M}}_0(u_0) + \int_0^t \langle \ell(s), \dot{u}(s)\rangle \mathrm{d}s \\ &\quad \geqq I_1 + I_2 + I_3 \geqq \overline{\mathcal{M}}_0(u(t)) + \int_0^t \int_\Omega \big(\mathbb{C}_D e(\dot{u}) : e(\dot{u}) + \mu \mathbb{B}^{(\alpha)} : \nabla \dot{u}\big)\mathrm{d}x\mathrm{d}s.\end{aligned}$$

Thus, all inequalities in (5.24) and (5.25) are equalities. In particular, we derive that

$$\lim_{k\to\infty} \frac{1}{2}\int_\Omega \mathbb{C}_W e(\overline{u}_k(t)) : e(\overline{u}_k(t))\mathrm{d}x = \frac{1}{2}\int_\Omega \mathbb{C}_W e(u(t)) : e(u(t))\mathrm{d}x, \qquad (5.32)$$

$$\lim_{k\to\infty} \frac{1}{\varepsilon_k^2}\int_0^t \int_\Omega \xi(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k)\mathrm{d}x\mathrm{d}s = \int_0^t \int_\Omega 4D(\mathbf{Id}, 0)e(\dot{u}) : e(\dot{u})\mathrm{d}x\mathrm{d}s, \qquad (5.33)$$

where we also used the definition of $\mathbb{C}_D$ in (2.32).

Step 3 (Strong convergence) Strong convergence for $\overline{u}_k$ in $H^1(\Omega; \mathbb{R}^d)$, i.e., the first part of (5.22), follows directly from (5.32), Korn's and Poincaré's inequality, and the fact that $\mathbb{C}_W$ is positive definite on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. In the same way we obtain convergence of $\underline{u}_k$ by employing $\underline{y}_k(t)$ in place of $\overline{y}_k(t)$ in (5.23). Hence, the statement also holds for $\hat{u}_k$. For the second part of (5.22), we will first show strong convergence of $(\dot{C}_k)_k$ defined in (5.18): by (D.2) we estimate

$$\begin{aligned}&c_0 \int_0^T \int_\Omega |\dot{C}_k - 2e(\dot{u})|^2 \mathrm{d}x\mathrm{d}t \leqq \int_0^T \int_\Omega D(C_k, \underline{\theta}_k)(\dot{C}_k - 2e(\dot{u})) : (\dot{C}_k - 2e(\dot{u}))\mathrm{d}x\mathrm{d}t \\ &\quad = \varepsilon_k^{-2} \int_0^T \int_\Omega \xi(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k)\mathrm{d}x\mathrm{d}t - 2\int_0^T \int_\Omega 2D(C_k, \underline{\theta}_k)e(\dot{u}) : \dot{C}_k \mathrm{d}x\mathrm{d}t \\ &\qquad + \int_0^T \int_\Omega 4D(C_k, \underline{\theta}_k)e(\dot{u}) : e(\dot{u})\mathrm{d}x\mathrm{d}t.\end{aligned}$$

By (5.33) for $t = T$, the pointwise convergence of $(\nabla \underline{y}_k)_k$ and $(\underline{\theta}_k)_k$ to $\mathbf{Id}$ and 0, respectively (see (5.1)–(5.3) and (5.6)), and the already shown weak convergence of $\dot{C}_k$ towards $2e(\dot{u})$ (cf. (5.19)) we see that the above derived upper bound converges to 0 as $k \to \infty$. Then, the desired strong convergence of $(\nabla \dot{\hat{u}}_k)_k$ is derived as follows: by using Poincaré's and Korn's inequality, (5.1), and (3.70a) we get that

$$\begin{aligned}
\int_0^T \int_\Omega |\nabla \dot{\hat{u}}_k - \nabla \dot{u}|^2 \mathrm{d}x\mathrm{d}t &\leqq C \int_0^T \int_\Omega |\mathrm{sym}(\nabla \dot{\hat{u}}_k - \nabla \dot{u})|^2 \mathrm{d}x\mathrm{d}t \\
&\leqq C \int_0^T \int_\Omega |\dot{C}_k - 2e(\dot{u})|^2 \mathrm{d}x\mathrm{d}t + C \int_0^T \int_\Omega |\nabla \underline{y}_k - \mathbf{Id}|^2 |\nabla \dot{\hat{u}}_k|^2 \mathrm{d}x\mathrm{d}t \\
&\leqq C \int_0^T \int_\Omega |\dot{C}_k - 2e(\dot{u})|^2 \mathrm{d}x\mathrm{d}t + C\varepsilon_k^{4/p} \int_0^T \int_\Omega |\nabla \dot{\hat{u}}_k|^2 \mathrm{d}x\mathrm{d}t \to 0.
\end{aligned}$$

This concludes the proof. □

Proposition 5.5. (Linearization of the heat-transfer equation) *Let u be as in Lemma 5.1 and μ as in Lemma 5.2. Then, for any $\varphi \in C^\infty(I \times \overline{\Omega})$ with $\varphi(T) = 0$ we have that (2.38) holds.*

Proof. Similarly to the proof of Proposition 4.6, see (4.18), we can show that

$$\begin{aligned}
&\int_0^T \int_\Omega \mathcal{K}(\nabla \underline{y}_k, \underline{\theta}_k) \nabla \overline{\mu}_k \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t - \int_0^T \int_\Omega \varepsilon_k^{-\alpha} \hat{w}_k \dot{\varphi} \mathrm{d}x\mathrm{d}t + \kappa \int_0^T \int_\Gamma \overline{\mu}_k \varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \\
&\quad - \int_0^T \int_\Omega \Big(\varepsilon_k^{-\alpha} \xi_\alpha^{\mathrm{reg}}(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k) + \varepsilon_k^{1-\alpha} \partial_F W^{\mathrm{cpl}}(\nabla \underline{y}_k, \underline{\theta}_k) : \nabla \dot{\hat{u}}_k \Big) \varphi \mathrm{d}x\mathrm{d}t \\
&= \kappa \int_0^T \int_\Gamma \overline{\theta}_{\flat,\tau} \varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t + \varepsilon_k^{-\alpha} \int_\Omega W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon}) \varphi(0) \mathrm{d}x,
\end{aligned} \tag{5.34}$$

where $\hat{w}_k := \hat{w}_{\varepsilon_k,\tau_k}$, $\nabla y_{0,\varepsilon} = \mathbf{Id} + \varepsilon_k \nabla u_0$, and $\theta_{0,\varepsilon} = \varepsilon_k^\alpha \mu_0$. Note that in contrast to (4.18), we rescaled both sides with $\varepsilon_k^{-\alpha}$. We will now pass to the limit in each integral above as $k \to \infty$. Recall that $c_V(F, \theta) := -\theta \partial_\theta^2 W^{\mathrm{cpl}}(F, \theta)$ for any $F \in GL^+(d)$ and $\theta \geqq 0$.

Using (C.6) and Dominated Convergence we can show in a similar fashion as in Remark 4.3 that

$$\varepsilon_k^{-\alpha} \int_\Omega W^{\mathrm{in}}(\nabla y_{0,\varepsilon}, \theta_{0,\varepsilon}) \varphi(0) \mathrm{d}x \to \int_\Omega \bar{c}_V \mu_0 \varphi(0) \mathrm{d}x.$$

By Lemma 3.3 we have that $|\mathcal{K}(\nabla \underline{y}_k, \underline{\theta}_k)|$ is uniformly bounded. Consequently, from the pointwise convergence of $\nabla \underline{y}_k$ and $\underline{\theta}_k$ to $\mathbf{Id}$ and 0, respectively, see (5.1)–(5.3) and (5.6), we derive that

$$\begin{aligned}
&\int_0^T \int_\Omega \mathcal{K}(\nabla \underline{y}_k, \underline{\theta}_k) \nabla \overline{\mu}_k \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t + \kappa \int_0^T \int_\Gamma \overline{\mu}_k \varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t \\
&\quad \to \int_0^T \int_\Omega \mathbb{K}_0 \nabla \mu \cdot \nabla \varphi \mathrm{d}x\mathrm{d}t + \kappa \int_\Gamma \mu \varphi \mathrm{d}\mathcal{H}^{d-1}\mathrm{d}t,
\end{aligned}$$

where $\mathbb{K}_0$ is defined in (2.33). By a change of variables and Dominated Convergence we find that

$$\varepsilon_k^{-\alpha} W^{\mathrm{in}}(\nabla \overline{y}_k, \overline{\theta}_k) = \int_0^{\overline{\mu}_k} c_V(\nabla \overline{y}_k, \varepsilon_k^\alpha s) \mathrm{d}s = \int_0^{\mu} c_V(\nabla \overline{y}_k, \varepsilon_k^\alpha s) \mathrm{d}s + \mathrm{O}(|\overline{\mu}_k - \mu|)$$
$$\to c_V(\mathbf{Id}, 0)\, \mu \tag{5.35}$$

pointwise a.e. in $I \times \Omega$, where we again used that by (C.5) the function c_V is bounded, the pointwise convergence of $(\nabla \overline{y}_k)_k$ to $\mathbf{Id}$, and the pointwise convergence $\overline{\mu}_k \to \mu$ (see (5.6), up to a subsequence). By Dominated Convergence this convergence also holds in $L^1(I \times \Omega)$. The same holds true for $\underline{y}_k, \underline{\theta}_k$ in place of $\overline{y}_k, \overline{\theta}_k$. Thus, recalling the definition of $\hat{w}_k$, we have shown that

$$\int_0^T \int_\Omega \varepsilon_k^{-\alpha} \hat{w}_k \dot{\varphi} \mathrm{d}x \mathrm{d}t \to \int_0^T \int_\Omega c_V(\mathbf{Id}, 0) \mu \dot{\varphi} \mathrm{d}x \mathrm{d}t = \int_0^T \int_\Omega \bar{c}_V \mu \dot{\varphi} \mathrm{d}x \mathrm{d}t. \tag{5.36}$$

We now prove that the contribution of the coupling potential vanishes in the limit. Indeed, by (3.5), (3.70a), (3.81), (5.1), and $t \wedge 1 \leqq t^{s/2}$ for some $s > \frac{2(\alpha-1)}{\alpha}$ with $s \in (1, \frac{d+2}{d})$, the Cauchy-Schwarz and Hölder's inequality we see that

$$\begin{aligned}
&\Big| \int_0^T \int_\Omega \varepsilon_k^{1-\alpha} W^{\mathrm{cpl}}(\nabla \underline{y}_k, \underline{\theta}_k) : \nabla \dot{\hat{u}}_k \varphi \mathrm{d}x \mathrm{d}t \Big| \\
&\quad \leqq \varepsilon_k^{1-\alpha} \int_0^T \int_\Omega C(\underline{\theta}_k \wedge 1)(1 + |\nabla \underline{y}_k - \mathbf{Id}|) |\nabla \dot{\hat{u}}_k| |\varphi| \mathrm{d}x \mathrm{d}t \\
&\quad \leqq C \varepsilon_k^{1-\alpha} \| \underline{\theta}_k^{\frac{s}{2}} \|_{L^2(\Omega)} \| \nabla \dot{\hat{u}}_k \|_{L^2(\Omega)} \| \varphi \|_{L^\infty(\Omega)} \\
&\quad \leqq C \varepsilon_k^{1-\alpha+\alpha s/2} \| \underline{\mu}_k \|_{L^s(\Omega)}^{\frac{s}{2}} \| \nabla \dot{\hat{u}}_k \|_{L^2(\Omega)} \| \varphi \|_{L^\infty(\Omega)} \to 0.
\end{aligned}$$

Lastly, by (5.18), by the second convergence in (5.22), (2.9), and the continuity of D one can show for $\alpha = 2$ that

$$\int_0^T \int_\Omega \varepsilon_k^{-\alpha} \xi_\alpha^{\mathrm{reg}}(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k) \varphi = \int_0^T \int_\Omega \varepsilon_k^{-\alpha} \xi(\nabla \underline{y}_k, \nabla \dot{\hat{y}}_k, \underline{\theta}_k) \varphi$$
$$\to \int_0^T \int_\Omega \mathbb{C}_D e(\dot{u}) : e(\dot{u}) \varphi \mathrm{d}x \mathrm{d}t.$$

For $\alpha < 2$ instead, it is easy to check using $\xi_\alpha^{\mathrm{reg}} \leqq \xi$ that the term vanishes as $k \to \infty$. Collecting all convergences and recalling the definition of $\mathbb{C}_D^{(\alpha)}$ in (2.34), we get that (2.38) holds true, where for the external temperature we use (4.19). □

Lemma 5.6. (Uniqueness of the linearized system) *There exists at most one solution in the sense of Definition 2.6.*

Proof. We start with $\alpha \in (1, 2]$. In this case, (2.37) is independent of the variable μ. We show uniqueness of u. To this end, we suppose that there exist two solutions u_1, u_2, and set $u := u_1 - u_2$. Then $u = 0$ on $I \times \Gamma_D$ and $u(0) = 0$. Subtracting the weak formulations (2.37) for both u_1 and u_2, we see that for any $z \in C^\infty(I \times \overline{\Omega}; \mathbb{R}^d)$ with $z = 0$ on $I \times \Gamma_D$ it holds that

$$\int_0^T \int_\Omega \Big(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot{u})\Big) : \nabla z \mathrm{d}x \mathrm{d}t = 0. \tag{5.37}$$

Let us now define

$$a(t) := \frac{1}{2} \int_\Omega \mathbb{C}_D e(u(t)) : e(u(t)) \mathrm{d}x \quad \text{for } t \in I.$$

Note that $a \in W^{1,1}(I)$ with

$$\dot{a}(t) = \int_\Omega \mathbb{C}_D e(\dot{u}(t)) : e(u(t)) \mathrm{d}x = \int_\Omega \mathbb{C}_D e(\dot{u}(t)) : \nabla u(t) \mathrm{d}x$$

for a.e. $t \in I$. Let $\tilde{\varphi} \in C^\infty(I)$. Testing (5.37) with a sequence of smooth maps $(z_h)_h$ vanishing on $I \times \Gamma_D$ and converging to $\tilde{\varphi} u$ in $L^2(I; H^1(\Omega))$ we derive that

$$\int_0^T \tilde{\varphi} \int_\Omega \Big(\mathbb{C}_W e(u) + \mathbb{C}_D e(\dot{u})\Big) : \nabla u \mathrm{d}x \mathrm{d}t = 0.$$

By the arbitrariness of $\tilde{\varphi}$ it then follows for almost all $t \in I$ that

$$\int_\Omega \big(\mathbb{C}_D e(\dot{u}(t)) + \mathbb{C}_W e(u(t))\big) : \nabla u(t) \mathrm{d}x = 0.$$

This shows that

$$\dot{a}(t) = \int_\Omega \mathbb{C}_D e(\dot{u}(t)) : \nabla u(t) \mathrm{d}x = -\int_\Omega \mathbb{C}_W e(u(t)) : e(u(t)) \mathrm{d}x \leqq 0.$$

As $a(0) = 0$, it follows that $a = 0$, and therefore $u = 0$. Now, given a unique $u \in H^1(I; H^1(\Omega))$, we see that (2.38) is an equation in the variable μ only. More precisely, it corresponds to the weak formulation of a heat equation with L^1-data. Uniqueness has been provided in [37, Proposition 1]. This finishes the proof in the case $\alpha \in (1, 2]$.

We now briefly give the argument for $\alpha = 1$. In this case, (2.38) does not depend on u and uniqueness follows again from [37, Proposition 1]. Then, the term $\int_0^T \int_\Omega \mathbb{B}^\alpha \mu : \nabla z \mathrm{d}x \mathrm{d}t$ in (2.37) is only a datum, and uniqueness of u follows by repeating the argument starting with (5.37). □

We are now ready to prove Theorem 2.7.

Proof of Theorem 2.7. We start with the proof of Theorem 2.7(iii). First, by Lemmas 5.1–5.2, we obtain limits $u \in H^1(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ and $\mu \in L^1(I; W^{1,1}(\Omega))$. In view of (5.1), (5.6), and Lemma 5.4, the convergences stated in the statement hold, up to selecting a subsequence. In particular, (5.3) and (5.8) show that the convergence holds for all three different interpolations. By Propositions 5.3 and 5.5 we see that (u, μ) is a weak solution in the sense of Definition 2.6. As the weak solution is unique by Lemma 5.6, Urysohn's subsequence principle implies that the convergence holds for the whole sequence. This concludes the proof of Theorem 2.7(i),(iii).

We briefly describe the adaptions for Theorem 2.7(ii). First, in the compactness result we replace (5.1) and (5.6) by (5.2) and (5.7), respectively. The linearization of the mechanical equation and the heat-transfer equation in Propositions 5.3 and 5.5, respectively, can be derived along similar lines, by replacing the time discrete equations (5.11) and (5.34) with their time-continuous analogs in (2.19) and (2.20), respectively. In a similar fashion, for the proof Lemma 5.4, we use the time-continuous energy balance (4.11) in place of (5.26). The rest of the argument remains unchanged. □

5.2. Proof of Theorem 2.8

We start with a Γ-convergence result. With the notation from Sects. 2.1–2.2 we define for $k \in \{1, \dots, T/\tau\}$ the functional $E_\varepsilon^{(k)} \colon H^1_{\Gamma_D}(\Omega; \mathbb{R}^d) \to \mathbb{R}$ through $E_\varepsilon^{(k)}(u) = +\infty$ if $u \notin W^{2,p}(\Omega; \mathbb{R}^d)$ and

$$\begin{aligned} E_\varepsilon^{(k)}(u) := {} & \frac{1}{\varepsilon^2}\mathcal{M}(\mathbf{id} + \varepsilon u) + \frac{1}{\varepsilon^2}\mathcal{W}^{\mathrm{cpl}}(\mathbf{id} + \varepsilon u, \theta_{\varepsilon,\tau}^{(k-1)}) \\ & + \frac{1}{\tau\varepsilon^2}\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, \mathbf{id} + \varepsilon u - y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) \\ & - \langle \ell_\tau^{(k)}, u\rangle - \frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)} \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\mathrm{d}x \end{aligned} \tag{5.38}$$

if $u \in W^{2,p}(\Omega; \mathbb{R}^d)$. Although the last term in (5.38) does not influence the minimizers of $E_\varepsilon^{(k)}$ for fixed k, it is needed to ensure the boundedness of $(|E_\varepsilon^{(k)}|)_\varepsilon$ as $\varepsilon \to 0$ along sequences of minimizers. Recall also $\mathcal{E}_\varepsilon$ from (3.48).

Proposition 5.7. *Suppose that* $\sup_{\varepsilon>0} \mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) < +\infty$ *and* $u_{\varepsilon,\tau}^{(k-1)} := \varepsilon^{-1}(y_{\varepsilon,\tau}^{(k-1)} - \mathbf{id}) \to u_\tau^{(k-1)}$ *strongly in* $H^1(\Omega; \mathbb{R}^d)$ *as* $\varepsilon \to 0$. *Suppose that* $\varepsilon^{-\alpha}\theta_{\varepsilon,\tau}^{(k-1)} \to \mu_\tau^{(k-1)}$ *in* $L^1(\Omega)$ *and that the convergence holds in* $L^2(\Omega)$ *if* $\alpha = 1$. *Then, the sequence* $(E_\varepsilon^{(k)})_\varepsilon$*, defined in* (5.38)*,* Γ*-converges in the weak* H^1*-topology to* $\bar{E}_0^{(k)} \colon H^1_{\Gamma_D}(\Omega; \mathbb{R}^d) \to \mathbb{R}$ *given by*

$$\begin{aligned} \bar{E}_0^{(k)}(u) := {} & \int_\Omega \Big(\frac{1}{2}\mathbb{C}_W e(u) : e(u)\mathrm{d}x + \frac{1}{2\tau}\mathbb{C}_D e(\tilde{u}) : e(\tilde{u}) + \bar{c}_V \mu_\tau^{(k-1)} + \mu_\tau^{(k-1)}\mathbb{B}^{(\alpha)} \colon \nabla\tilde{u}\Big)\mathrm{d}x \\ & - \langle \ell_\tau^{(k)}, u\rangle, \end{aligned}$$

where $\tilde{u} := u - u_\tau^{(k-1)}$, $\mathbb{C}_W$, $\mathbb{C}_D$ *as in* (2.32), $\bar{c}_V$ *as in* (2.33), *and* $\mathbb{B}^{(\alpha)}$ *as in* (2.34).

Proof. All constants we encounter in this proof are implicitly assumed to be independent of ε. We will work with the equivalent representation

$$\begin{aligned} E_\varepsilon^{(k)}(u) = & \frac{1}{\varepsilon^2}\mathcal{M}(\mathbf{id} + \varepsilon u) + \frac{1}{\tau\varepsilon^2}\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, \mathbf{id} + \varepsilon u - y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) \\ & + \frac{1}{\varepsilon^2}\int_\Omega W_\varepsilon^{\mathrm{in}}(\mathbf{Id} + \varepsilon\nabla u, \theta_{\varepsilon,\tau}^{(k-1)})\mathrm{d}x \\ & + \frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\big(\partial_\theta W^{\mathrm{cpl}}(\mathbf{Id} + \varepsilon\nabla u, \theta_{\varepsilon,\tau}^{(k-1)}) \\ & - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big)\mathrm{d}x - \langle \ell_\tau^{(k)}, u\rangle, \end{aligned} \tag{5.39}$$

which can be derived from (5.38) by adding and subtracting

$$\frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\partial_\theta W^{\mathrm{cpl}}(\mathbf{Id} + \varepsilon\nabla u, \theta_{\varepsilon,\tau}^{(k-1)})$$

and using the definition of W^{in} in (2.12).

Step 1 (Mechanical energy bound): Let $(u_\varepsilon)_\varepsilon \subset W^{2,p}_{\Gamma_D}(\Omega;\mathbb{R}^d)$ be a sequence such that $\sup_{\varepsilon>0} E_\varepsilon^{(k)}(u_\varepsilon) < \infty$. We will show that then also $\sup_{\varepsilon>0}\varepsilon^{-2}\mathcal{M}(y_\varepsilon) < \infty$, where we shortly wrote $y_\varepsilon := \mathbf{id} + \varepsilon u_\varepsilon$. By the nonnegativity of W^{in} and R we derive that

$$\begin{aligned} E_\varepsilon^{(k)}(u_\varepsilon) \geqq & \frac{1}{\varepsilon^2}\mathcal{M}(y_\varepsilon) + \frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\big(\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)}) \\ & - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big)\mathrm{d}x - \langle \ell_\tau^{(k)}, u_\varepsilon\rangle. \end{aligned} \tag{5.40}$$

By the second bound in (C.5), Young's inequality with constant λ, and $1 \wedge t \leqq \sqrt{t}$ for $t \geqq 0$ it follows that

$$\begin{aligned} & \theta_{\varepsilon,\tau}^{(k-1)}\big|\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)}) - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big| \\ & \leqq C(\theta_{\varepsilon,\tau}^{(k-1)} \wedge 1)(1 + |\nabla y_\varepsilon - \mathbf{Id}| + |\nabla y_{\varepsilon,\tau}^{(k-1)} - \mathbf{Id}|) \\ & \leqq \frac{C}{\lambda}\theta_{\varepsilon,\tau}^{(k-1)} + C\lambda|\nabla y_\varepsilon - \mathbf{Id}|^2 + C\lambda|\nabla y_{\varepsilon,\tau}^{(k-1)} - \mathbf{Id}|^2. \end{aligned}$$

Integrating over Ω and using (3.2) as well as (W.4) we get that

$$\begin{aligned} & \frac{1}{\varepsilon^2}\Big|\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\big(\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)}) - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big)\mathrm{d}x\Big| \\ & \leqq \frac{C}{\lambda\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\mathrm{d}x + \frac{C\lambda}{\varepsilon^2}\mathcal{W}^{\mathrm{el}}(y_{\varepsilon,\tau}^{(k-1)}) + \frac{C\lambda}{\varepsilon^2}\mathcal{W}^{\mathrm{el}}(y_\varepsilon). \end{aligned} \tag{5.41}$$

Again by (3.2), Poincaré's inequality, and Young's inequality with constant λ/ε we see that

$$|\langle \ell_\tau^{(k)}, u_\varepsilon\rangle| = \varepsilon^{-1}|\langle \ell_\tau^{(k)}, y_\varepsilon - \mathbf{id}\rangle| \leqq \frac{C}{\lambda}\|\ell_\tau^{(k)}\|^2_{H^{-1}} + C\frac{\lambda}{\varepsilon^2}\mathcal{W}^{\mathrm{el}}(y_\varepsilon).$$

Hence, combining the above estimate with (5.41) and (5.40), and using Hölder's inequality, we arrive at

$$E_\varepsilon^{(k)}(y_\varepsilon) \geqq (1 - C\lambda)\varepsilon^{-2}\mathcal{M}(y_\varepsilon) - \frac{C}{\lambda}\big(\mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) + \|\ell_\tau^{(k)}\|_{H^{-1}}^2\big).$$

Choosing λ sufficiently small such that $1 - C\lambda \geqq 1/2$ this leads the desired bound. Consequently, in the sequel, we can assume that (3.70a)–(3.70b) holds for both y_ε and $y_{\varepsilon,\tau}^{(k-1)}$.

Step 2 (Γ-lim inf*):* Let $(u_\varepsilon)_\varepsilon \subset H^1_{\Gamma_D}(\Omega;\mathbb{R}^d)$ be such that $u_\varepsilon \rightharpoonup u$ weakly in $H^1(\Omega;\mathbb{R}^d)$. Without loss of generality we can assume that $\sup_{\varepsilon>0} E_\varepsilon^{(k)}(y_\varepsilon) < \infty$ and $\liminf_{\varepsilon\to0} E_\varepsilon^{(k)}(u_\varepsilon) = \lim_{\varepsilon\to0} E_\varepsilon^{(k)}(u_\varepsilon)$. In particular, we can select a subsequence (without relabeling) such that $\theta_{\varepsilon,\tau}^{(k-1)} \to 0$ a.e. in Ω. We are now ready to compute the lim inf of the various terms of $E_\varepsilon^{(k)}(u_\varepsilon)$. By (3.70a) we see that $\nabla y_\varepsilon \to \mathbf{Id}$ uniformly. Hence, by the weak convergence of $(u_\varepsilon)_\varepsilon$ in $H^1(\Omega;\mathbb{R}^d)$ we can show similarly to the derivation of (5.24) that

$$\liminf_{\varepsilon\to0} \frac{1}{\varepsilon^2}\mathcal{M}(y_\varepsilon) \geqq \liminf_{\varepsilon\to0} \frac{1}{2}\int_\Omega \mathbb{C}_W \nabla u_\varepsilon : \nabla u_\varepsilon \mathrm{d}x \geqq \frac{1}{2}\int_\Omega \mathbb{C}_W e(u) : e(u)\mathrm{d}x. \tag{5.42}$$

As in the proof of (5.25), it follows from the pointwise convergence of $(y_{\varepsilon,\tau}^{(k-1)})_\varepsilon$ and $(\theta_{\varepsilon,\tau}^{(k-1)})_\varepsilon$ that

$$\liminf_{\varepsilon\to0} \frac{1}{\tau\varepsilon^2}\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, y_\varepsilon - y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) \geqq \frac{1}{2\tau}\int_\Omega \mathbb{C}_D e(u - u_\tau^{(k-1)}) : e(u - u_\tau^{(k-1)})\mathrm{d}x.$$

By the same argument as in (5.36), the L^1-convergence of $\varepsilon^{-\alpha}\theta_{\varepsilon\tau}^{(k-1)}$ implies that

$$\lim_{\varepsilon\to0} \frac{1}{\varepsilon^\alpha}\int_\Omega W^{\mathrm{in}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)})\mathrm{d}x = \bar{c}_V \int_\Omega \mu_\tau^{(k-1)}\mathrm{d}x. \tag{5.43}$$

For the remaining coupling term in (5.39), we Taylor expand around $(\mathbf{Id}, \theta_{\varepsilon,\tau}^{(k-1)})$ and get by the second bound in (C.5), (C.7), and (3.70a), applied for both y_ε and $y_{\varepsilon,\tau}^{(k-1)}$, that

$$\begin{aligned}
&\theta_{\varepsilon,\tau}^{(k-1)}\big|\big(\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)}) - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big) \\
&\quad - \partial_{\theta F} W^{\mathrm{cpl}}(\mathbf{Id}, \theta_{\varepsilon,\tau}^{(k-1)}) : \nabla(y_\varepsilon - y_{\varepsilon,\tau}^{(k-1)})\big| \\
&\quad \leqq C(\theta_{\varepsilon,\tau}^{(k-1)} \wedge 1)\big(|\nabla y_\varepsilon - \mathbf{Id}|^2 + |\nabla y_{\varepsilon,\tau}^{(k-1)} - \mathbf{Id}|^2\big) \\
&\quad \leqq C\varepsilon^{1+\frac{2}{p}}(\theta_{\varepsilon,\tau}^{(k-1)} \wedge 1)\big(|\nabla u_\varepsilon| + |\nabla u_{\varepsilon,\tau}^{(k-1)}|\big).
\end{aligned}$$

pointwise a.e. in Ω. Thus, by repeating the argument in (5.30) we derive that

$$\begin{aligned}
&\lim_{\varepsilon\to0} \frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\big(\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon, \theta_{\varepsilon,\tau}^{(k-1)}) - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)})\big)\mathrm{d}x \\
&\quad = \lim_{\varepsilon\to0} \int_\Omega \varepsilon^{-1}\theta_{\varepsilon,\tau}^{(k-1)}\partial_{\theta F} W^{\mathrm{cpl}}(\mathbf{Id}, \theta_{\varepsilon,\tau}^{(k-1)}) : \nabla(u_\varepsilon - u_{\varepsilon,\tau}^{(k-1)}).
\end{aligned}$$

Thus, by the definition of $\mathbb{B}^{(\alpha)}$ in (2.34) and by repeating the argument in (5.29)–(5.30) we conclude that

$$\frac{1}{\varepsilon^2}\int_\Omega \theta_{\varepsilon,\tau}^{(k-1)}\big(\partial_\theta W^{\mathrm{cpl}}(\nabla y_\varepsilon,\theta_{\varepsilon,\tau}^{(k-1)}) - \partial_\theta W^{\mathrm{cpl}}(\nabla y_{\varepsilon,\tau}^{(k-1)},\theta_{\varepsilon,\tau}^{(k-1)})\big)\mathrm{d}x$$
$$\to \int_\Omega \mu_\tau^{(k-1)}\mathbb{B}^{(\alpha)} : \nabla(u-u_\tau^{(k-1)})\mathrm{d}x \tag{5.44}$$

as $\varepsilon\to 0$. Finally, notice that the weak convergence also implies $\lim_{\varepsilon\to0}\langle \ell_\tau^{(k)}, u_\varepsilon\rangle = \langle \ell_\tau^{(k)}, u\rangle$. Combining all aforementioned estimates we conclude the proof of the Γ-lim inf.

Step 3 (Γ-lim sup) Let $u\in H^1(\Omega;\mathbb{R}^d)$ with $u=0$ on Γ_D. By a standard approximation argument in Sobolev spaces we can assume without loss of generality that $u\in C^\infty(\Omega;\mathbb{R}^d)$. Choose $u_\varepsilon = u$ for all ε. We only need to check the convergence of the energy. First, notice that by (5.21) and $p>2$

$$\frac{1}{\varepsilon^2}\int_\Omega H(\nabla^2 y_\varepsilon)\mathrm{d}x \leqq \frac{1}{\varepsilon^2}\int_\Omega C_0|\varepsilon\nabla^2 u|^p\mathrm{d}x = C_0\varepsilon^{p-2}\int_\Omega |\nabla^2 u|^p\mathrm{d}x \to 0,$$

where $y_\varepsilon := \mathbf{id}+\varepsilon u$. By a Taylor expansion we also see that

$$\frac{1}{\varepsilon^2}\int_\Omega W^{\mathrm{el}}(\nabla y_\varepsilon)\mathrm{d}x = \frac{1}{2\varepsilon^2}\int_\Omega \mathbb{C}_W\varepsilon\nabla u : \varepsilon\nabla u\mathrm{d}x + \mathrm{O}\Big(\varepsilon\int_\Omega|\nabla^3 u|\mathrm{d}x\Big)$$
$$\to \frac{1}{2}\int_\Omega \mathbb{C}_W e(u) : e(u)\mathrm{d}x.$$

Furthermore, using (D.1) we can write that

$$\frac{1}{\varepsilon^2}\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, y_\varepsilon - y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) = \frac{1}{2}\int_\Omega D(C_\varepsilon,\theta_{\varepsilon,\tau}^{(k-1)})\dot C_\varepsilon : \dot C_\varepsilon,$$

where $C_\varepsilon := (\nabla y_{\varepsilon,\tau}^{(k-1)})^T\nabla y_{\varepsilon,\tau}^{(k-1)}$ and $\dot C_\varepsilon := (\nabla u - \nabla u_{\varepsilon,\tau}^{(k-1)})^T\nabla y_{\varepsilon,\tau}^{(k-1)} + (\nabla y_{\varepsilon,\tau}^{(k-1)})^T(\nabla u - \nabla u_{\varepsilon,\tau}^{(k-1)})$. By the strong convergence of $(u_{\varepsilon,\tau}^{(k-1)})_\varepsilon$ in $H^1(\Omega;\mathbb{R}^d)$ it follows that $\dot C_\varepsilon \to 2e(u-u_\tau^{(k-1)})$ strongly in $L^2(\Omega;\mathbb{R}^{d\times d})$. Consequently,

$$\frac{1}{\tau\varepsilon^2}\mathcal{R}(y_{\varepsilon,\tau}^{(k-1)}, y_\varepsilon - y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) \to \frac{1}{2\tau}\int_\Omega \mathbb{C}_D e(u-u_\tau^{(k-1)}) : e(u-u_\tau^{(k-1)}).$$

The convergence of the terms (5.43) and (5.44) follows as in the previous step. This concludes the proof. □

We close with the proof of Theorem 2.8.

Proof of Theorem 2.8. We prove the result by induction on k. For the base case $k=0$, we only need to check the convergences and the energy convergence. In fact, setting $u_\tau^{(0)} := u_0$ and $\mu_\tau^{(0)} := \mu_0$, this directly follows from (2.18) and repeating the argument in the Γ-lim sup above.

Suppose now that the statement is true for $k-1$ where $k\in\{1,\ldots,T/\tau\}$. With (3.70b) we have $u_{\varepsilon,\tau}^{(k)} = \varepsilon^{-1}(y_{\varepsilon,\tau}^{(k)} - \mathbf{id}) \rightharpoonup u_\tau^{(k)}$ weakly in $H^1(\Omega;\mathbb{R}^d)$ (up to a subsequence). By the induction hypothesis it holds that $u_{\varepsilon,\tau}^{(k-1)} = \varepsilon^{-1}(y_{\varepsilon,\tau}^{(k-1)} - \mathbf{id}) \to$

$u_\tau^{(k-1)}$ strongly in $H^1(\Omega;\mathbb{R}^d)$ and $\varepsilon^{-\alpha}\theta_{\varepsilon,\tau}^{(k-1)} \rightharpoonup \mu_\tau^{(k-1)}$ weakly in $W^{1,r}(\Omega)$ for any $r \in [1, \frac{d+2}{d+1})$. Therefore, we also find $\varepsilon^{-\alpha}\theta_{\varepsilon,\tau}^{(k-1)} \to \mu_\tau^{(k-1)}$ strongly in $L^1(\Omega)$. If $\alpha = 1$, Remark 4.3(iii) even yields convergence in $L^2(\Omega)$. As also $\sup_{\varepsilon>0}\mathcal{E}_\varepsilon(y_{\varepsilon,\tau}^{(k-1)}, \theta_{\varepsilon,\tau}^{(k-1)}) < +\infty$ due to Lemma 3.18, we can apply Proposition 5.7. By the fundamental theorem of Γ-convergence, $u_\tau^{(k)}$ is a minimizer of $\bar{E}_0^{(k)}$ and $E_\varepsilon^{(k)}(u_{\varepsilon,\tau}^{(k)}) \to \bar{E}_0^{(k)}(u_\tau^{(k)})$. As $\bar{E}_0^{(k)}$ is strictly convex, $u_\tau^{(k)}$ is the unique minimizer of the corresponding minimization problem. In particular, the weak H^1-convergence of $(u_{\varepsilon,\tau}^{(k)})_\varepsilon$ holds true without selecting a subsequence. Moreover, energy convergence implies that in (5.42) equality holds. This along with weak convergence, as well as Korn's and Poincaré's inequality yields $u_{\varepsilon,\tau}^{(k)} \to u_\tau^{(k)}$ strongly in $H^1(\Omega;\mathbb{R}^d)$. Clearly, $u_\tau^{(k)}$ satisfies (2.40).

Let $r \in [1, \frac{d+2}{d+1})$ and $s \in [1, \frac{d+2}{d})$. As $\tau > 0$ was fixed, we see by (5.6) that, up to selecting a subsequence, $\varepsilon^{-\alpha}\theta_{\varepsilon,\tau}^{(k)} \to \mu_\tau^{(k)}$ weakly in $W^{1,r}(\Omega)$ and strongly in $L^s(\Omega)$. This limit $\mu_\tau^{(k)}$ solves (2.41). Indeed, testing (3.11) (with $\xi_\alpha^{\mathrm{reg}}$ in place of ξ) with $\varphi \in C^\infty(\bar{\Omega})$ and dividing by ε^α we can pass to the limit $\varepsilon \to 0$, and obtain (2.41) by an argument similar to the one in the proof Proposition 5.5 neglecting the time dependence. The main difference is that we do not perform integration by parts in time, but by using the argument in (5.35) we pass directly to the limit in the term

$$\frac{1}{\varepsilon^\alpha}\int_\Omega \tau^{-1}\big(w_{\varepsilon,\tau}^{(k)} - w_{\varepsilon,\tau}^{(k-1)}\big)\varphi \mathrm{d}x \to \int_\Omega \bar{c}_V \tau^{-1}\big(\mu_\tau^{(k)} - \mu_\tau^{(k-1)}\big)\varphi \mathrm{d}x.$$

To conclude the induction step, it remains to show the uniqueness of $\mu_\tau^{(k)}$, which in particular will imply that the weak $W^{1,r}$-convergence holds true without selecting a subsequence. Suppose that $\tilde{\mu}_\tau^{(k)}$ also satisfies (2.41). Then, for the difference $\mu := \mu_\tau^{(k)} - \tilde{\mu}_\tau^{(k)}$ it holds that

$$\int_\Omega \big(\bar{c}_V \frac{\mu}{\tau}\varphi + \mathbb{K}_0 \nabla\mu\cdot\nabla\varphi\big)\mathrm{d}x + \kappa\int_\Gamma \mu\varphi \mathrm{d}\mathcal{H}^{d-1} = 0.$$

Taking a smooth sequence $(\varphi_h)_h \subset C_c^\infty(\Omega)$ converging to $\chi(\mu)$ in C^1, where $\chi(t) := \arctan(t)$, this shows with (2.10), $\chi(t)t \geqq 0$ for all t, and $\chi' \geqq 0$ that $\int_\Omega \frac{\mu}{\tau}\chi(\mu)\mathrm{d}x = 0$. As $\chi(t)t \geqq 0$ for all t and $\chi(t) = 0$ if and only if $t = 0$, we have proved $\mu \equiv 0$, and thus uniqueness holds.

(ii) We only sketch the proof as it follows along the lines of the reasoning in Sect. 4. Let $\hat{u}_\tau$, $\overline{u}_\tau$, $\underline{u}_\tau$ be defined similar to (2.26), and use similar notation for μ. We first observe that $(\hat{u}_\tau)_\tau$ is bounded in $H^1(I; H^1(\Omega;\mathbb{R}^d))$ and $(\hat{\mu}_\tau)_\tau$ is bounded in $L^r(I; W^{1,r}(\Omega))$. This follows from Lemmas 5.1–5.2 and (2.39). Additional control can be recovered from the estimates stated in Theorem 3.20. Thus, we can find $u \in H^1(I; H^1_{\Gamma_D}(\Omega;\mathbb{R}^d))$ such that $\nabla\dot{\hat{u}}_\tau \rightharpoonup \nabla\dot{u}$ and $\nabla\overline{u}_\tau \rightharpoonup \nabla u$ weakly in $L^2(I\times\Omega;\mathbb{R}^{d\times d})$. Moreover, there exists $\mu \in L^1(I; W^{1,1}(\Omega))$ with $\mu \geqq 0$ a.e. such that the latter two convergences in (2.42) can be derived (up to a subsequence) using the Aubin-Lions' theorem and by following the reasoning in Lemma 4.2.

Using (2.40) for every smooth $z \in L^2(I; H^1_{\Gamma_D}(\Omega; \mathbb{R}^d))$ and summing over every $k \in \{1, \dots, T/\tau\}$ we derive that

$$\int_0^T \int_\Omega \big(\mathbb{C}_W e(\overline{u}_\tau) + \underline{\mu}_\tau \mathbb{B}^{(\alpha)} + \mathbb{C}_D e(\dot{\hat{u}}_\tau)\big) : \nabla z \mathrm{d}x\mathrm{d}t - \int_0^T \langle \overline{\ell}_\tau(t), z(t)\rangle \mathrm{d}t = 0.$$

Consequently, we can then pass to the limit $\tau \to 0$ in the above equality which results in (2.37). Using (2.41) for every $k \in \{1, \dots, T/\tau\}$ we also see that for any $\varphi \in C^\infty(I \times \bar{\Omega})$ with $\varphi(T) = 0$ it holds that

$$\int_0^T \int_\Omega \big(\mathbb{C}_D^{(\alpha)} e(\dot{\hat{u}}_\tau) : e(\dot{\hat{u}}_\tau)\, \varphi + \mathbb{K}_0 \nabla \overline{\mu}_\tau \cdot \nabla \varphi - \bar{c}_V \hat{\mu}_\tau \dot{\varphi}\big) \mathrm{d}x\mathrm{d}t + \kappa \int_0^T \int_\Gamma (\overline{\mu}_\tau - \overline{\theta}_{\flat,\tau}) \varphi \mathrm{d}\mathcal{H}^{d-1} = \bar{c}_V \int_\Omega \mu_0 \varphi(0) \mathrm{d}x,$$

where as usual we applied integration by parts. In particular, as $\tau \to 0$ by (2.42) we see that

$$\begin{aligned} &\lim_{\tau \to 0} \int_0^T \int_\Omega (-\bar{c}_V \hat{\mu}_\tau \dot{\varphi} + \mathbb{K}_0 \nabla \overline{\mu}_\tau \cdot \nabla \varphi) \mathrm{d}x\mathrm{d}t + \kappa \int_0^T \int_\Gamma (\overline{\mu}_\tau - \overline{\theta}_{\flat,\tau}) \varphi \mathrm{d}\mathcal{H}^{d-1} \\ &\quad = \int_0^T \int_\Omega (-\bar{c}_V \mu \dot{\varphi} + \mathbb{K}_0 \nabla \mu \cdot \nabla \varphi) \mathrm{d}x\mathrm{d}t + \kappa \int_0^T \int_\Gamma (\mu - \theta_\flat) \varphi \mathrm{d}\mathcal{H}^{d-1}. \end{aligned} \tag{5.45}$$

We also find that

$$\begin{aligned} \lim_{\tau \to 0} \frac{1}{2} \int_\Omega \mathbb{C}_W e(\bar{u}(t)) : e(\bar{u}(t)) \mathrm{d}x &= \frac{1}{2} \int_\Omega \mathbb{C}_W e(u_\tau(t)) : e(u_\tau(t)) \mathrm{d}x, \\ \lim_{\tau \to 0} \int_0^T \int_\Omega \mathbb{C}_D e(\dot{\hat{u}}_\tau) : e(\dot{\hat{u}}_\tau)\, \varphi \mathrm{d}x\mathrm{d}t &= \int_0^T \int_\Omega \mathbb{C}_D e(\dot{u}) : e(\dot{u})\, \varphi \mathrm{d}x\mathrm{d}t. \end{aligned} \tag{5.46}$$

Indeed, inequalities follow from weak convergence, and the equalities are recovered by resorting to energy balances in the time-discrete and time-continuous setting, see Lemma 4.5, in particular (4.15)–(4.16), for details. Let us highlight that at this point for $\alpha = 1$ we exploit $\int_0^T \int_\Omega \underline{\mu}_\tau \mathbb{B}^{(\alpha)} : \nabla \dot{\hat{u}}_\tau \mathrm{d}x \to \int_0^T \int_\Omega \mu \mathbb{B}^{(\alpha)} : \nabla \dot{u} \mathrm{d}x$ since we can assume $\underline{\mu}_\tau \to \mu$ in $L^2(I; L^2(\Omega))$ by Remark 4.3(iii).

The second part of (5.46) along with (5.45) implies that (2.38) holds. This shows that (u, μ) is a weak solution of (2.29)–(2.31) in the sense of Definition 2.6. This solution is unique (see Theorem 2.7(i)), all aforementioned convergences hold true without selecting a subsequence. Energy convergence in (5.46) along with weak convergence implies $\overline{u}_\tau(t) \to \overline{u}(t)$ strongly in $H^1(\Omega; \mathbb{R}^d)$ for every $t \in I$. For the other interpolations, one can argue in a similar fashion by replacing $\overline{u}_\tau(t)$ by $\underline{u}_\tau(t)$ in (5.46).

Acknowledgements. This work was funded by the DFG Project FR 4083/5-1 and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044 -390685587, Mathematics Münster: Dynamics–Geometry–Structure. The work was further supported by the DAAD Project 57600633, and by the Project DAAD-22-03. MARTIN KRUŽÍK acknowledges support by GAČR-FWF Project 21-06569K and by the Erwin Schrödinger International Institute for Mathematics and Physics during his stay in Vienna in 2022.

References

1. Alicandro, R., Dal Maso, G., Lazzaroni, G., Palombaro, M.: Derivation of a linearised elasticity model from singularly perturbed multiwell energy functionals. *Arch. Ration. Mech. Anal.* **230**, 1–45, 2018
2. Antman, S.S.: Physically unacceptable viscous stresses. *Z. Angew. Math. Phys.* **49**, 980–988, 1998
3. Antman, S.S.: *Nonlinear Problems of Elasticity*. Springer, New York, 2004
4. Ball, J.M., Currie, J.C., Olver, P.L.: Null Lagrangians, weak continuity, and variational problems of arbitrary order. *J. Funct. Anal.* **41**, 135–174, 1981
5. Batra, R.C.: Thermodynamics of non-simple elastic materials. *J. Elast.* **6**, 451–456, 1976
6. Blanchard, D., Guibé, O.: Existence of a solution for a nonlinear system in thermoviscoelasticity. *Adv. Differ. Equ.* **5**, 1221–1252, 2000
7. Braides, A.: Γ-*Convergence for Beginners*. Oxford University Press, Oxford, 2002
8. Braides, A., Solci, M., Vitali, E.: A derivation of linear elastic energies from pair-interaction atomistic systems. *Netw. Heterog. Media* **2**, 551–567, 2007
9. Boccardo, L., Dall'aglio, A., Gallouët, T., Orsina, L.: Nonlinear parabolic equations with measure data. *J. Funct. Anal.* **147**, 237–258, 1997
10. Boccardo, L., Gallouët, T.: Non-linear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* **87**, 149–169, 1989
11. Bonetti, E., Bonfanti, G.: Existence and uniqueness of the solution to a 3D thermoelastic system. *Electron. J. Differ. Equ.* **50**, 1–15, 2003
12. Dal Maso, G.: *An Introduction to* Γ-*Convergence*. Birkhäuser, Boston, 1993
13. Dal Maso, G., Negri, M., Percivale, D.: Linearized elasticity as Γ-limit of finite elasticity. *Set-valued Anal.* **10**, 165–183, 2002
14. Davoli, E., Friedrich, M.: Two-well linearization for solid–solid phase transitions. Preprint at arXiv:2005.03892v1, 2020
15. Demoulini, S.: Weak solutions for a class of nonlinear systems of viscoelasticity. *Arch. Ration. Mech. Anal.* **155**, 299–334, 2000
16. Feireisl, E., Málek J.: On the Navier–Stokes equations with temperature-dependent transport coefficients. *Differ. Equ. Nonlinear Mech.* 2006
17. Fonseca, I., Leoni, G.: *Modern Methods in the Calculus of Variations:* L^p *Spaces*. Springer, Berlin, 2007
18. Friedrich, M.: A derivation of linearized Griffith energies from nonlinear models. *Arch. Ration. Mech. Anal.* **225**, 425–467, 2017
19. Friedrich, M., Kružík, M.: On the passage from nonlinear to linearized viscoelasticity. *SIAM J. Math. Anal.* **50**, 4426–4456, 2018
20. Friedrich, M.: Griffith energies as small strain limit of nonlinear models for nonsimple brittle materials. *Math. Eng.* **2**, 75–100, 2020
21. Gurtin, M.E., Fried, E., Anand, L.: *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010
22. Healey, T.J., Krömer, S.: Injective weak solutions in second-gradient nonlinear elasticity. *ESAIM Control Optim. Cal. Var.* **15**, 863–871, 2009

23. JESENKO, M., SCHMIDT, B.: Geometric linearization of theories for incompressible elastic materials and applications. *Math. Mod. Methods Appl. Sci.* **31**, 829–860, 2021
24. KRÖMER, S., ROUBÍCEK, T.: Quasistatic viscoelasticity with self-contact at large strains. *J. Elast.* **142**, 433–445, 2020
25. KRUŽÍK, M., ROUBÍCEK, T.: *Mathematical Methods in Continuum Mechanics of Solids*. Springer, Cham, 2019
26. LEWICKA, M., MUCHA, P.B.: A local existence result for system of viscoelasticity with physical viscosity. *Evol. Equ. Control Theory* **2**, 337–353, 2013
27. MAININI, E., PERCIVALE, D.: Variational linearization of pure traction problems in incompressible elasticity. *Zeit. Angew. Math. Phys.* **71**, 146, 2020
28. MAININI, E., PERCIVALE, D.: Linearization of elasticity models for incompressible materials. Preprint at arXiv:2004.09286, 2020
29. MAOR, C., MORA, M.G.: Reference configurations versus optimal rotations: a derivation of linear elasticity from finite elasticity for all traction forces. *J. Nonlinear Sci.* **31**, 62, 2021
30. MIELKE, A., NAUMANN, J.: On the existence of global-in-time weak solutions and scaling laws for Kolmogorov's two-equation model of turbulence. Preprints: WIAS, Berlin, 2545 and arXiv:1801.02039, 2018
31. MIELKE, A., ORTNER, C., ŞENGÜL, Y.: An approach to nonlinear viscoelasticity via metric gradient flows. *SIAM J. Math. Anal.* **46**, 1317–1347, 2014
32. MIELKE, A., ROUBÍCEK, T.: Rate-independent elastoplasticity at finite strains and its numerical approximation. *Math. Models Methods Appl. Sci.* **26**, 2203–2236, 2016
33. MIELKE, A., ROUBÍCEK, T.: Thermoviscoelasticity in Kelvin–Voigt rheology at large strains. *Arch. Ration. Mech. Anal.* **238**, 1–45, 2020
34. MIELKE, A., STEFANELLI, U.: Linearized plasticity is the evolutionary Γ-limit of finite plasticity. *J. Eur. Math. Soc. (JEMS)* **15**, 923–948, 2013
35. PODIO-GUIDUGLI, P.: Contact interactions, stress, and material symmetry for nonsimple elastic materials. *Theor. Appl. Mech.* **28–29**, 261–276, 2002
36. POMPE, W.: Korn's first inequality with variable coefficients and its generalization. *Comment. Math. Univ. Carolinae* **44**, 57–70, 2003
37. ROUBÍCEK, T.: Nonlinear heat equation with L^1-data. *NoDEA Nonlinear Differ. Equ. Appl.* **5**(4), 517–527, 1998
38. ROUBÍCEK, T.: *Nonlinear Partial Differential Equations with Applications*, vol. 153. Springer, 2013
39. SCHMIDT, B.: Linear Γ-limits of multiwell energies in nonlinear elasticity theory. *Contin. Mech. Thermodyn.* **20**, 375–396, 2008
40. SCHMIDT, B.: On the derivation of linear elasticity from atomistic models. *Netw. Heterog. Media* **4**, 789–812, 2009
41. TOUPIN, R.A.: Elastic materials with couple stresses. *Arch. Ration. Mech. Anal.* **11**, 385–414, 1962
42. TOUPIN, R.A.: Theory of elasticity with couple stress. *Arch. Ration. Mech. Anal.* **17**, 85–112, 1964

Rufat Badal
Department of Mathematics,
Friedrich-Alexander Universität Erlangen-Nürnberg,
Cauerstr. 11,
91058 Erlangen
Germany.
e-mail: rufat.badal@fau.de

and

Manuel Friedrich
Mathematics Münster,
University of Münster,
Einsteinstr. 62,
48149 Münster
Germany.
e-mail: manuel.friedrich@fau.de

and

Martin Kružík
Institute of Information Theory and Automation,
Czech Academy of Sciences,
Pod vodárenskou věží 4,
182 00 Prague 8
Czechia.
e-mail: kruzik@utia.cas.cz

and

Martin Kružík
Faculty of Civil Engineering,
Czech Technical University,
Thákurova 7,
166 29 Prague 6
Czechia.

(*Received March 4, 2022 / Accepted December 12, 2022*)