A REVISION OF RESULTS FOR STANDARD MODELS IN ELASTO-PERFECT-PLASTICITY THEORY

MIROSLAV BULÍČEK AND JENS FREHSE

ABSTRACT. We consider two most studied standard models in the theory of elasto-plasticity in arbitrary dimension $d \geq 2$, namely, the Hencky model and the Prandtl–Reus model subjected to the von Mises condition. There are many available results for these models - from the existence and the regularity theory up to the relatively sharp identification of the plastic strain in the natural function/measure space setting. In this paper we shall proceed further and improve some of known estimates in order to identify sharply the plastic strain. More specifically, we rigorously improve the integrability of the displacement and the velocity (which was known only under a nonnatural assumption that the Cauchy stress is bounded), show the BMO estimates for the stress and finally also the Morrey-like estimates for the plastic strain. In addition, we shall provide the whole theory up to the boundary. As an immediate consequence of such improved estimates, we provide a sharper identification of the plastic strain than that known up to date. In particular, in two dimensional setting, we show that the plastic strain can be point-wisely characterized in terms of the stresses *everywhere* although the stress is possibly discontinuous and thus the natural duality pairing in the space of measures could be violated.

1. INTRODUCTION

This paper focuses on the qualitative estimates for solutions to several models of linearized (possible nonlinear) elasto–plasticity. To describe the problem in more details, we shall assume that a body occupies a Lipschitz set $\mathcal{O} \subset \mathbb{R}^d$ and we a priori assume that considered deformations are small. Therefore, the initial, current and preferred (natural) configurations coincide and we can approximate the strain tensor by the linearized strain tensor $\boldsymbol{\varepsilon}(\boldsymbol{u})$, which is defined as

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) := \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)\,, \tag{1.1}$$

where $\boldsymbol{u} : (0,T) \times \mathcal{O} \to \mathbb{R}^d$ is the displacement and $T > 0$ is the length of the time interest (in this model one should prefer the notion "the loading parameter" to "time"). We also assume that the density is constant and that the inertial effects can be neglected. Then the balance of linear momentum for the quasi-static deformation takes the form

$$-\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{f} \qquad \text{in } [0,T] \times \mathcal{O}, \tag{1.2}$$

2000 *Mathematics Subject Classification.* 74G40, 35Q72, 74C05, 74G10.

Key words and phrases. elasto-perfect-plasticity, Hencky model, Prandtl–Reus model, regularity, BMO estimates, Gehring lemma, identification os the plastic strain.

M. Bulíček's work is supported by the ERC-CZ project LL1202 financed by the Ministry of Education, Youth and Sports, Czech Republic. M. Bulíček is a member of the Nečas Center for Mathematical Modeling.

where $\boldsymbol{\sigma} : (0,T) \times \mathcal{O} \to \mathbb{R}^{d\times d}_{sym}$ is the Cauchy stress and $\boldsymbol{f} : (0,T) \times \mathcal{O} \to \mathbb{R}^d$ denotes the density of given external body forces. To complete the problem (1.1)–(1.2) it remains to prescribe the boundary and initial conditions, which we shall do later, and also to characterize the relationship between $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}(\boldsymbol{u})$. In the case of linearized elasto–plasticity, we assume that the linearized strain $\boldsymbol{\varepsilon}(\boldsymbol{u})$ can be decomposed into the elastic part $\boldsymbol{e}_{\mathrm{el}}$ and the plastic part $\boldsymbol{e}_{\mathrm{p}}$, i.e.,

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{e}_{\mathrm{el}} + \boldsymbol{e}_{\mathrm{p}} \tag{1.3}$$

and that the elastic response of the material is given by the Helmholtz potential $\psi^* : \mathbb{R}^{d\times d}_{sym} \to \mathbb{R}$, which is supposed to be a convex function vanishing at zero and exploding at infinity and the elastic strain is related to the stress through

$$\boldsymbol{\sigma} = \frac{\partial \psi^*(\boldsymbol{e}_{\mathrm{el}})}{\partial \boldsymbol{e}_{\mathrm{el}}} \quad \Leftrightarrow \quad \boldsymbol{e}_{\mathrm{el}} = \frac{\partial \psi(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}\,, \tag{1.4}$$

where ψ is the conjugate function to ψ^* defined as

$$\psi(\boldsymbol{\sigma}) := \sup_{\boldsymbol{e}_{\mathrm{el}}} \left(\boldsymbol{\sigma}\cdot \boldsymbol{e}_{\mathrm{el}} - \psi^*(\boldsymbol{e}_{\mathrm{el}})\right)\,.$$

Concerning the plastic strain, we consider that it is relevant to incompressible behaviour and therefore it is reasonable to assume that

$$\operatorname{tr} \boldsymbol{e}_{\mathrm{p}} = 0\,. \tag{1.5}$$

Further, we need to specify under which conditions it may appear. Indeed, there are many possible settings (yield conditions) used in praxis, but we choose the so–called von Mieses conditions saying that

$$|\boldsymbol{\sigma}_D| < \kappa \implies \dot{\boldsymbol{e}}_{\mathrm{p}} = 0,$$

which in other words means that the response of the material is purely elastic as far as $|\boldsymbol{\sigma}_D| < \kappa$. Here the symbol $\dot{u}$ denotes the derivative of the quantity u with respect to the time variable, or more precisely with to the loading parameter t. On the other hand, if the plastic behavior takes place, then we require that

$$\kappa \frac{\dot{\boldsymbol{e}}_{\mathrm{p}}}{|\dot{\boldsymbol{e}}_{\mathrm{p}}|} = \boldsymbol{\sigma}_D\,.$$

These two conditions, can be summarized into the more compact compact Kuhn–Tucker form

$$\dot{\boldsymbol{e}}_{\mathrm{p}} = \lambda \boldsymbol{\sigma}_D \quad \text{with} \quad \lambda \ge 0\,, \quad |\boldsymbol{\sigma}_D| - \kappa \le 0 \quad \text{and} \quad \lambda\left(|\boldsymbol{\sigma}_D| - \kappa\right) = 0\,, \tag{1.6}$$

with λ given as $\lambda = |\dot{\boldsymbol{e}}_{\mathrm{p}}|/\kappa = |\dot{\boldsymbol{e}}_{\mathrm{p}}|/|\boldsymbol{\sigma}_D|$. This model with (1.4) given as $\boldsymbol{\sigma} = \mathbf{A}\boldsymbol{e}_{\mathrm{el}}$ is usually referred as the the Prandtl-Reuss model of elasto-plasticity (see [14, 15]), and for theoretical justification of the model for general form of ψ and also other possible yield conditions[1], we refer the interested reader to [6]. The second model of linearized elasto–plasticity is the Hencky model (see [12]) that is formally obtained from (1.6) by replacing $\dot{\boldsymbol{e}}_{\mathrm{p}}$ by $\boldsymbol{e}_{\mathrm{p}}$, i.e.,

$$\boldsymbol{e}_{\mathrm{p}} = \lambda \boldsymbol{\sigma}_D \quad \text{with} \quad \lambda \ge 0\,, \quad |\boldsymbol{\sigma}_D| - \kappa \le 0 \quad \text{and} \quad \lambda\left(|\boldsymbol{\sigma}_D| - \kappa\right) = 0\,. \tag{1.7}$$

At this point we can clearly specify the main problem in both above models. While the existence of the stress and the displacement fulfilling the weak formulation (see below) is known under the certain reasonable hypothesis on the data, see

[1]There are other activation criteria that may be obtained by considering anisotropic elastic response and that are connected with names such as Rankine, Saint-Venant, Tresca, etc.

[17, 18], many fundamental questions about the qualitative character of the solution remain open for several decades. The first delicate question is how to interpret (1.6), (1.7) respectively, since $\boldsymbol{\sigma}_D$ is only (Lebesgue) measurable and $\boldsymbol{e}_{\mathrm{p}}$ or $\dot{\boldsymbol{e}}_{\mathrm{p}}$ are the vector-valued Radon measures. Due to the celebrated works [1, 2, 3, 4], we know that λ appearing in (1.6), (1.7) respectively, is a nonnegative Radon measure and that instead of (1.6), (1.7), we have that

$$\dot{\boldsymbol{e}}_{\mathrm{p}} = \lambda \boldsymbol{\sigma}_p \text{ in } (0,T)\times\mathcal{O}, \quad \text{or} \quad \boldsymbol{e}_{\mathrm{p}} = \lambda \boldsymbol{\sigma}_p \text{ in } \mathcal{O}\,, \tag{1.8}$$

where $\boldsymbol{\sigma}_p$ is λ measurable and for any compact $K \subset \mathcal{O}$

$$\lim_{r\to 0}\int_K |\boldsymbol{\sigma}_D^r - \boldsymbol{\sigma}_p|\,\mathrm{d}\lambda = 0, \qquad \text{with} \qquad \boldsymbol{\sigma}_D^r(x) := \frac{1}{|B_r(x)|}\int_{B_r(x)} \boldsymbol{\sigma}_D^r(y)\,\mathrm{d}y, \tag{1.9}$$

where in case of the Prandtl–Reuss model, (1.9) holds for almost all time $t \in (0,T)$. The first result of the paper is about the characterization of the set, where $\boldsymbol{\sigma}_D \neq \boldsymbol{\sigma}_p$. In particular, we shall show that in dimension $d = 2$, we can simply set $\boldsymbol{\sigma}_p := \boldsymbol{\sigma}_D$ *everywhere in* $\mathcal{O}$, since $\boldsymbol{\sigma}_D$ is λ-measurable. In addition, we shall show that the same property holds also in higher dimension depending on a set, where $\boldsymbol{u}$ or $\dot{\boldsymbol{u}}$ do not explode too quickly (see the precise statement in Theorem 2.1).

Second result, we shall show in the paper, is that the singular measure λ and consequently also $\boldsymbol{e}_{\mathrm{p}}$ and $\dot{\boldsymbol{e}}_{\mathrm{p}}$ satisfy certain Morrey condition and consequently, these measures are absolutely continues with respect to ε-Hausdorff measure for sufficiently small $\varepsilon > 0$. In particular, they cannot concentrate at a point.

Next, it is know due to [10] that $\boldsymbol{u} \in L^{d'+\varepsilon}_{loc}(\mathcal{O})$ or $\dot{\boldsymbol{u}} \in L^\infty(0,T; L^{d'+\varepsilon}_{loc}(\mathcal{O}))$, but the proofs work only under the additional assumption that $\boldsymbol{\sigma} \in L^\infty_{loc}$, which is not the case here. Nevertheless, and it is the next key result of the paper, this claim remains true even without this additional (and probably incorrect) assumption. Moreover, we shall get these estimates up to the boundary of $\mathcal{O}$. Furthermore, as the key tool for this observation, we shall show that $\boldsymbol{\sigma} \in BMO(\mathcal{O})$, or $\boldsymbol{\sigma} \in L^\infty(0,T; BMO(\mathcal{O}))$, which seems to be also a new result. Furthermore, we also provide the uniform up to the initial time $t = 0$ fractional time derivative estimates for the solution.

The final novelty consists in the fact the we prove all estimate for an approximative problem, the Perčina-Mises model, and show that all the regularity estimates (with respect to time and also to space) remains independent of the order of the approximation and the dimension, which was also not known in many cases.

2. Weak formulation of the problem and the main result

We introduce the classical formulation of the Prandtl–Reuss and the Hencky model completed by the boundary, and if needed, also initial data. We focus here only on the most physically relevant boundary conditions, i.e., the prescribed displacement, or the traction, or the normal displacement and the tangential traction, but all results can be adapted to a more general setting. For given open bounded Lipschitz set $\mathcal{O} \subset \mathbb{R}^d$ we consider that its boundary can be decomposed onto three smooth relatively open disjoint parts of the boundary: the Dirichlet part $\partial\mathcal{O}_D$, the Neumann part $\partial\mathcal{O}_N$ and the mixed part $\partial\mathcal{O}_M$ such that $\overline{\partial\mathcal{O}_D \cup \partial\mathcal{O}_N \cup \partial\mathcal{O}_M} = \partial\mathcal{O}$. The Prandtl–Reuss model of elasto-plasticity consists in finding a quadruple

$(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{e}_{\rm el}, \boldsymbol{e}_{\rm p}) : [0,T] \times \mathcal{O} \to \mathbb{R}^{d\times d}_{sym} \times \mathbb{R}^d \times \mathbb{R}^{d\times d}_{sym} \times \mathbb{R}^{d\times d}_{sym}$ such that

$$
\begin{aligned}
&-\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f}, \quad \boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{e}_{\rm el} + \boldsymbol{e}_{\rm p}, \quad \boldsymbol{e}_{\rm el} = \frac{\partial \psi(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} && \text{in } [0,T]\times\mathcal{O}\,,\\
&\dot{\boldsymbol{e}}_{\rm p} = |\dot{\boldsymbol{e}}_{\rm p}|\frac{\boldsymbol{\sigma}_D}{\kappa}, \quad |\boldsymbol{\sigma}_D| \le \kappa \text{ and } |\dot{\boldsymbol{e}}_{\rm p}|(|\boldsymbol{\sigma}_D| - \kappa) = 0 && \text{in } [0,T]\times\mathcal{O}\,,\\
&\boldsymbol{u} = \boldsymbol{u}_0 && \text{on } [0,T]\times\partial\mathcal{O}_D\,,\\
&\boldsymbol{u}\cdot\boldsymbol{n} = \boldsymbol{u}_0\cdot\boldsymbol{n}, \quad (\boldsymbol{\sigma}\boldsymbol{n})_{\boldsymbol{\tau}} = (\boldsymbol{f}_{\boldsymbol{n}})_{\boldsymbol{\tau}} && \text{on } [0,T]\times\partial\mathcal{O}_M\,,\\
&\boldsymbol{\sigma}\boldsymbol{n} = \boldsymbol{f}_{\boldsymbol{n}} && \text{on } [0,T]\times\partial\mathcal{O}_N\,,\\
&\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0, \quad \boldsymbol{u}(0) = \boldsymbol{u}_0(0) && \text{in } \mathcal{O}\,.
\end{aligned} \tag{2.1}
$$

where $T > 0$ is the given length of the time[2] interval, the given threshold $\kappa > 0$ is a von Mieses condition, $\boldsymbol{f} : [0,T]\times\mathcal{O} \to \mathbb{R}^d$ are the given volume forces, $\boldsymbol{f}_{\boldsymbol{n}} : [0,T]\times\partial\mathcal{O} \to \mathbb{R}^d$ are the given traction forces, the initial stress is $\boldsymbol{\sigma}_0 : \mathcal{O} \to \mathbb{R}^{d\times d}_{sym}$ and the prescribed displacement on the boundary $[0,T]\times\partial\mathcal{O}_D \cup \partial\mathcal{O}_M$ and the initial displacement is represented by $\boldsymbol{u}_0 : [0,T]\times\mathcal{O} \to \mathbb{R}^d$. Here the symbol $\boldsymbol{n}$ denotes the outer normal vector on $\partial\mathcal{O}$ and for any vector $\boldsymbol{u} \in \mathbb{R}^d$ and any given $x \in \partial\mathcal{O}$ we denote $(\boldsymbol{u}(x))_{\boldsymbol{\tau}} := \boldsymbol{u}(x) - (\boldsymbol{u}(x)\cdot\boldsymbol{n}(x))\boldsymbol{n}(x)$, i.e., the projection of $\boldsymbol{u}$ to the tangent plane at the point x.

The second model, we have in mind, is the Hencky model, which can be formally formulated as: to find a quadruple $(\boldsymbol{\sigma}, \boldsymbol{u}, \boldsymbol{e}_{\rm el}, \boldsymbol{e}_{\rm p}) : \mathcal{O} \to \mathbb{R}^{d\times d}_{sym} \times \mathbb{R}^d \times \mathbb{R}^{d\times d}_{sym} \times \mathbb{R}^{d\times d}_{sym}$ such that

$$
\begin{aligned}
&-\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f}, \quad \boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{e}_{\rm el} + \boldsymbol{e}_{\rm p}, \quad \boldsymbol{e}_{\rm el} = \frac{\partial \psi(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} && \text{in } \mathcal{O}\,,\\
&\boldsymbol{e}_{\rm p} = |\boldsymbol{e}_{\rm p}|\frac{\boldsymbol{\sigma}_D}{\kappa}, \quad |\boldsymbol{\sigma}_D| \le \kappa \text{ and } |\boldsymbol{e}_{\rm p}|(|\boldsymbol{\sigma}_D| - \kappa) = 0 && \text{in } \mathcal{O}\,,\\
&\boldsymbol{u} = \boldsymbol{u}_0 && \text{on } \partial\mathcal{O}_D\,,\\
&\boldsymbol{u}\cdot\boldsymbol{n} = \boldsymbol{u}_0\cdot\boldsymbol{n}, \quad (\boldsymbol{\sigma}\boldsymbol{n})_{\boldsymbol{\tau}} = (\boldsymbol{f}_{\boldsymbol{n}})_{\boldsymbol{\tau}} && \text{on } \partial\mathcal{O}_M\,,\\
&\boldsymbol{\sigma}\boldsymbol{n} = \boldsymbol{f}_{\boldsymbol{n}} && \text{on } \partial\mathcal{O}_N\,.
\end{aligned} \tag{2.2}
$$

In general, we are not able to solve the above models in the classical sense, so we introduce a notion of a weak (variational solution). For this purpose, we define the set of admissible stresses. Before doing so let us define the subspace of the Sobolev space $W^{1,2}(\mathcal{O};\mathbb{R}^d)$, which will be used in what follows

$$\mathcal{V} := \{\boldsymbol{v} \in W^{1,2}(\mathcal{O};\mathbb{R}^d);\ \boldsymbol{v} = \boldsymbol{0} \text{ on } \partial\mathcal{O}_D,\ \boldsymbol{v}\cdot\boldsymbol{n} = 0 \text{ on } \partial\mathcal{O}_M\}.$$

Then we define the set of admissible stresses as

$$
\begin{aligned}
\mathcal{F}(t) := \Big\{ &\boldsymbol{\sigma} \in L^2(\Omega;\mathbb{R}^{d\times d}_{sym});\ |\boldsymbol{\sigma}_D| \le \kappa, \text{ and for all } \boldsymbol{v} \in \mathcal{V} \\
&\int_{\mathcal{O}} \boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}(\boldsymbol{v})\,{\rm d}x = \int_{\mathcal{O}} \boldsymbol{f}(t)\cdot\boldsymbol{v}\,{\rm d}x + \int_{\partial\mathcal{O}_{N,M}} \boldsymbol{f}_{\boldsymbol{n}}(t)\cdot\boldsymbol{v}\,{\rm d}S \Big\},
\end{aligned}
$$

where we denoted $\partial\mathcal{O}_{N,M} := \partial\mathcal{O}_N \cup \partial\mathcal{O}_N$.

Naturally, we also have to restrict on the reasonable class of possible Helmholtz potentials. We shall assume that $\psi : \mathbb{R}^{d\times d}_{sym} \to \mathbb{R}$ is a smooth nonnegative function fulfilling in addition $\psi(\boldsymbol{0}) = 0$, $\frac{\partial\psi(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}}|_{\boldsymbol{\sigma}=\boldsymbol{0}} = \boldsymbol{0}$. Moreover, there exist $C_1, C_2 > 0$

---

[2] In fact, we should not called it time interval, since t corresponds to the loading parameter.

such that for all $\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}} \in \mathbb{R}^{d\times d}_{sym}$

$$C_1|\tilde{\boldsymbol{\sigma}}|^2 \leq \frac{\partial^2\psi(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}} \cdot (\tilde{\boldsymbol{\sigma}} \otimes \tilde{\boldsymbol{\sigma}}) \leq C_2|\tilde{\boldsymbol{\sigma}}|^2. \tag{2.3}$$

In case of the Prandtl–Reuss model, we will need the further restriction and we will assume that there exists a constant fourth order tensor $\mathsf{A} \in \mathbb{R}^{d\times d}_{sym} \times \mathbb{R}^{d\times d}_{sym}$ such that for all $\boldsymbol{\sigma} \in \mathbb{R}^{d\times d}_{sym}$

$$\mathsf{A} \equiv \frac{\partial^2\psi(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}}. \tag{2.4}$$

Then we shall define the weak solution to the Prandtl–Reuss model (2.1) as follows.

Definition 2.1 (Prandtl–Reuss)**.** *Let $\mathcal{O} \subset \mathbb{R}^d$ be a Lipschitz domain. Assume that $\boldsymbol{f} \in L^2(0,T;L^2(\mathcal{O};\mathbb{R}^d))$, $\boldsymbol{f}_n \in L^2(0,T;L^2(\partial\mathcal{O};\mathbb{R}^d))$, $\boldsymbol{\sigma}_0 \in L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym})$ and $\boldsymbol{u}_0 \in W^{1,2}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}^d))$. Further, let the potential ψ satisfy (2.3)–(2.4). We say that $\boldsymbol{\sigma} \in W^{1,2}(0,T;L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym}))$ is a weak solution to (2.1) if $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0$ and for almost all $t \in (0,T)$ there holds $\boldsymbol{\sigma}(t) \in \mathcal{F}(t)$ and, in addition, we require that for almost all $t \in (0,T)$ and all $\tilde{\boldsymbol{\sigma}} \in \mathcal{F}(t)$ there holds*

$$\int_{\mathcal{O}} \mathsf{A} \cdot (\dot{\boldsymbol{\sigma}}(t) \otimes (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}}))\,\mathrm{d}x \leq \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}})\,\mathrm{d}x. \tag{2.5}$$

In a very similar way, we can also introduce the notion of a weak solution to the Hencky model (2.2), where we shall replace $\mathcal{F}(t)$ by $\mathcal{F}$ in a natural way.

Definition 2.2 (Hencky)**.** *Let $\mathcal{O} \subset \mathbb{R}^d$ be a Lipschitz domain. Assume that $\boldsymbol{f} \in L^2(\mathcal{O};\mathbb{R}^d)$, $\boldsymbol{f}_n \in L^2(\partial\mathcal{O};\mathbb{R}^d)$ and $\boldsymbol{u}_0 \in W^{1,2}(\mathcal{O};\mathbb{R}^d)$. Further, let the potential ψ satisfy (2.3). We say that $\boldsymbol{\sigma} \in L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym})$ is a weak solution to (2.2) if $\boldsymbol{\sigma} \in \mathcal{F}$ and for all $\tilde{\boldsymbol{\sigma}} \in \mathcal{F}$ there holds*

$$\int_{\mathcal{O}} \frac{\partial\psi(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}} \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\,\mathrm{d}x \leq \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\boldsymbol{u}_0) \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\,\mathrm{d}x. \tag{2.6}$$

Since our approach is constructive, we frequently use the penalization of the von Mises condition (1.6) or (1.7) (see also [13]), Hohenemser-Prager model, then one arrives at the Perčina-Mises model. Thus, introducing a new class of admissible stresses as

$$\mathcal{F}_{el}(t) := \Big\{\boldsymbol{\sigma} \in L^2(\Omega;\mathbb{R}^{d\times d}_{sym});\ \text{ and for all } \boldsymbol{v} \in \mathcal{V} \\ \int_{\mathcal{O}} \boldsymbol{\sigma} \cdot \boldsymbol{\varepsilon}(\boldsymbol{v})\,\mathrm{d}x = \int_{\mathcal{O}} \boldsymbol{f} \cdot \boldsymbol{v}\,\mathrm{d}x + \int_{\partial\mathcal{O}_{N,M}} \boldsymbol{f_n} \cdot \boldsymbol{v}\,\mathrm{d}S\Bigg\},$$

we define the approximative problems as follows.

Definition 2.3 (Prandtl–Reuss–Perčina)**.** *Let $\mathcal{O} \subset \mathbb{R}^d$ be a Lipschitz domain and $\mu > 0$. Assume that $\boldsymbol{f} \in L^2(0,T;L^2(\mathcal{O};\mathbb{R}^d))$, $\boldsymbol{f}_n \in L^2(0,T;L^2(\partial\mathcal{O};\mathbb{R}^d))$, $\boldsymbol{\sigma}_0 \in L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym})$ and $\boldsymbol{u}_0 \in W^{1,2}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}^d))$. Further, let the potential ψ satisfy (2.3)–(2.4). We say that $\boldsymbol{\sigma} \in W^{1,2}(0,T;L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym}))$ is a weak solution to Prandtl–Reuss-Perčina model if $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0$ and for almost all $t \in (0,T)$ there holds $\boldsymbol{\sigma}(t) \in \mathcal{F}_{el}(t)$ and, in addition, we require that for almost all $t \in (0,T)$ and all $\tilde{\boldsymbol{\sigma}} \in \mathcal{F}_{el}(t)$*

there holds

$$\begin{aligned}(2.7)\qquad &\int_{\mathcal{O}} \mathsf{A}\cdot(\dot{\boldsymbol{\sigma}}(t)\otimes(\boldsymbol{\sigma}(t)-\tilde{\boldsymbol{\sigma}}))\,\mathrm{d}x+\mu^{-1}\int_{\mathcal{O}}\frac{(|\boldsymbol{\sigma}_D(t)|-\kappa)_+\boldsymbol{\sigma}_D(t)}{|\boldsymbol{\sigma}_D(t)|}\cdot(\boldsymbol{\sigma}(t)-\tilde{\boldsymbol{\sigma}})\,\mathrm{d}x\\ &\qquad=\int_{\mathcal{O}}\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}(t)-\tilde{\boldsymbol{\sigma}})\,\mathrm{d}x.\end{aligned}$$

Similarly for the approximation of the Hencky model, we have the following approximation.

Definition 2.4 (Hencky–Perčina)**.** *Let $\mathcal{O}\subset\mathbb{R}^d$ be a Lipschitz domain and $\mu>0$. Assume that $\boldsymbol{f}\in L^2(\mathcal{O};\mathbb{R}^d)$, $\boldsymbol{f}_n\in L^2(\partial\mathcal{O};\mathbb{R}^d)$ and $\boldsymbol{u}_0\in W^{1,2}(\mathcal{O};\mathbb{R}^d)$. Further, let the potential ψ satisfy* (2.3)*. We say that $\boldsymbol{\sigma}\in L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym})$ is a weak solution to Hencky–Perčina model if $\boldsymbol{\sigma}\in\mathcal{F}_{el}$ and for all $\tilde{\boldsymbol{\sigma}}\in\mathcal{F}_{el}$ there holds*

$$\begin{aligned}(2.8)\qquad &\int_{\mathcal{O}}\frac{\partial\psi(\boldsymbol{\sigma})}{\partial\boldsymbol{\sigma}}\cdot(\boldsymbol{\sigma}-\tilde{\boldsymbol{\sigma}})\,\mathrm{d}x+\mu^{-1}\int_{\mathcal{O}}\frac{(|\boldsymbol{\sigma}_D|-\kappa)_+\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot(\boldsymbol{\sigma}-\tilde{\boldsymbol{\sigma}})\,\mathrm{d}x\\ &\qquad=\int_{\mathcal{O}}\boldsymbol{\varepsilon}(\boldsymbol{u}_0)\cdot(\boldsymbol{\sigma}-\tilde{\boldsymbol{\sigma}})\,\mathrm{d}x.\end{aligned}$$

The existence of weak solution to Prandtl–Reuss or Hencky model in the sense of Definitions 2.1–2.2 is very standard. However, to be able to talk also about the displacement, one needs to assume certain compatibility condition on data. Therefore, we shall require the existence of the so-called safety load condition, i.e., the existence of $\boldsymbol{\sigma}^s\in W^{1,2}(0,T;L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym}))$ fulfilling for some $\delta>0$ and all $t\in[0,T]$

$$(2.9)\qquad \boldsymbol{\sigma}^s(t)\in\mathcal{F}(t),\quad \|\boldsymbol{\sigma}^s_D(t)\|_\infty\le\kappa-\delta.$$

Similarly, in case of the Hencky model, we assume that there is $\boldsymbol{\sigma}^s\in L^2(\mathcal{O};\mathbb{R}^{d\times d}_{sym})$ fulfilling

$$(2.10)\qquad \boldsymbol{\sigma}^s\in\mathcal{F},\quad \|\boldsymbol{\sigma}^s_D\|_\infty\le\kappa-\delta.$$

Finally, we state the main results of the paper. The first one is for the Prandtl–Reuss model.

Theorem 2.1 (Prandtl–Reuss)**.** *Let all assumptions of Definition 2.3 be satisfied. Then for all $\mu>0$ there exists a unique weak solution $\boldsymbol{\sigma}^\mu$ to* (2.7)*. Moreover, there exists $\boldsymbol{u}^\mu$ such that $\boldsymbol{u}^\mu-\boldsymbol{u}_0\in W^{1,2}(0,T;W^{1,2}_0(\mathcal{O};\mathbb{R}^d))$ and*

$$(2.11)\qquad \mathsf{A}\dot{\boldsymbol{\sigma}}^\mu+\mu^{-1}(|\boldsymbol{\sigma}^\mu_D|-\kappa)_+\frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|}=\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)\qquad a.e.\ in\ (0,T)\times\mathcal{O}.$$

In addition if there exists $\boldsymbol{\sigma}^s\in\mathcal{C}^1([0,T];\mathcal{C}^2(\overline{\mathcal{O}}))$ satisfying (2.9)*, then there is a constant $\varepsilon>0$ independent of μ such that for any compact $\tilde{\mathcal{O}}\subset\mathcal{O}$*

(2.12)

$$\begin{aligned}&\sup_{t\in(0,T)}\left(\|\dot{\boldsymbol{\sigma}}^\mu(t)\|_2^2+\mu^{-1}\|(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\|_1+\|\operatorname{div}\dot{\boldsymbol{u}}^\mu(t)\|_2^2+\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu(t))\|_1+\|\dot{\boldsymbol{u}}^\mu(t)\|_{d'}\right)\\ &+\sup_{0<h<T}\frac{1}{h}\int_0^{T-h}\|\dot{\boldsymbol{\sigma}}^\mu(t+h)-\dot{\boldsymbol{\sigma}}^\mu(t)\|_2^2\,\mathrm{d}t+\sup_{t\in(0,T)}\int_{\tilde{\mathcal{O}}}|\nabla\boldsymbol{\sigma}^\mu(t,x)|^2\,\mathrm{d}x\,\mathrm{d}t\\ &+\int_0^T\int_{\tilde{\mathcal{O}}}|\dot{\boldsymbol{u}}^\mu(t,x)|^{d'+\varepsilon}\,\mathrm{d}x\,\mathrm{d}t\le C(\tilde{\mathcal{O}}),\end{aligned}$$

where the constant $C(\tilde{\mathcal{O}})$ depends only on $\boldsymbol{\sigma}^s$, A, $\mathcal{O}$ and $\tilde{\mathcal{O}}$. Moreover, there exists a subsequence that we do not relabel such that

$$\begin{aligned}
\boldsymbol{\sigma}^\mu &\rightharpoonup^* \boldsymbol{\sigma} && \text{in } W^{1,\infty}(0,T;L^2(\mathcal{O},\mathbb{R}^{d\times d}_{sym})),\\
\boldsymbol{\sigma}^\mu &\rightharpoonup \boldsymbol{\sigma} && \text{in } N^{\frac{3}{2},2}(0,T;L^2(\mathcal{O},\mathbb{R}^{d\times d}_{sym})),\\
\mu^{-1}(|\boldsymbol{\sigma}^\mu_D|-1)_+\frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|} &\rightharpoonup^* \dot{\boldsymbol{e}}_p && \text{in } L^\infty(0,T;\mathcal{M}(\overline{\mathcal{O}};\mathbb{R}^{d\times d}_{sym})),\\
\mu^{-1}(|\boldsymbol{\sigma}^\mu_D|-1)_+ &\rightharpoonup^* \lambda && \text{in } L^\infty(0,T;\mathcal{M}(\overline{\mathcal{O}})),\\
\boldsymbol{\varepsilon}(\boldsymbol{u}^\mu) &\rightharpoonup^* \boldsymbol{\varepsilon}(\boldsymbol{u}) && \text{in } W^{1,\infty}(0,T;\mathcal{M}(\overline{\mathcal{O}};\mathbb{R}^{d\times d}_{sym})),\\
\boldsymbol{u}^\mu &\rightharpoonup^* \boldsymbol{u} && \text{in } W^{1,\infty}(0,T;L^{d'}(\mathcal{O};\mathbb{R}^d)),
\end{aligned} \tag{2.13}$$

where $\boldsymbol{\sigma}$ is a weak solution in sense of Definition 2.1, (1.8) *and* (1.9) *hold and*

$$\mathsf{A}\dot{\boldsymbol{\sigma}} + \dot{\boldsymbol{e}}_p = \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) \qquad \text{in } (0,T)\times\overline{\mathcal{O}}. \tag{2.14}$$

In addition, there exists $\varepsilon > 0$ such that

$$\boldsymbol{\sigma} \in L^\infty(0,T;BMO(\mathcal{O};\mathbb{R}^{d\times d}_{sym})), \qquad \dot{\boldsymbol{u}} \in L^\infty(0,T;L^{d'+\varepsilon}(\mathcal{O};\mathbb{R}^d)). \tag{2.15}$$

Moreover, there exists positive constants C and δ such that for almost all $t\in(0,T)$, there holds

$$\begin{aligned}
&\operatorname{supp}\lambda(t) \subset \{x\in\mathcal{O}\cup\partial\mathcal{O}_D;\ M|\boldsymbol{\sigma}_D(t)|(x)=1\},\\
&\quad\dot{\boldsymbol{e}}_p(t) = \frac{1}{2}\left((\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\otimes\boldsymbol{n}+\boldsymbol{n}\otimes(\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\right)\,\mathrm{d}S \text{ on } \partial\mathcal{O}_D,\\
&\quad\dot{\boldsymbol{e}}_p(t) = \lambda(t)\boldsymbol{\sigma}_D(t) \text{ in } \mathcal{O}\setminus K(t),\\
&\lambda(t)(B_R) \le CR^\delta \text{ for all balls } B_R\subset\mathcal{O},
\end{aligned} \tag{2.16}$$

where

$$K(t) := \left\{ y\in\mathcal{O}:\ \liminf_{\varepsilon\to 0}\ \sup_{R\in(0,\varepsilon)}\ \varepsilon\int_{B_R(y)} R^{1-d-\varepsilon/2}|\dot{\boldsymbol{u}}(t,x)|^{1+\varepsilon}\,\mathrm{dx} \ge 1 \right\}.$$

In particular, if $\dot{\boldsymbol{u}}(t)\in L^q(\mathcal{O};\mathbb{R}^d)$ for some $q>d$ then

$$\dot{\boldsymbol{e}}_p(t) = \lambda(t)\boldsymbol{\sigma}_D(t) \text{ in } \mathcal{O}. \tag{2.17}$$

Consequently, due to (2.15)*, the identity* (2.17) *always holds for $d=2$.*

Please notice here that we used the notations $N^{\frac{3}{2},2}$ for the Nikoloskii space, $\mathcal{M}$ for the space of Radon measures, M is the Hardy–Littlewood maximal function and BMO for functions with bounded oscillation. For the sake of completeness we state also the theorem for the Hencky model.

Theorem 2.2 (Hencky)**.** *Let all assumptions of Definition 2.4 be satisfied. Then for all $\mu>0$ the re exists a unique weak solution $\boldsymbol{\sigma}^\mu$ to* (2.8)*. Moreover, there exists $\boldsymbol{u}^\mu$ such that $\boldsymbol{u}^\mu-\boldsymbol{u}_0\in W^{1,2}_0(\mathcal{O};\mathbb{R}^d))$ and*

$$\frac{\partial\psi(\boldsymbol{\sigma}^\mu)}{\partial\boldsymbol{\sigma}^\mu} + \mu^{-1}(|\boldsymbol{\sigma}^\mu_D|-\kappa)_+\frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|} = \boldsymbol{\varepsilon}(\boldsymbol{u}^\mu). \tag{2.18}$$

In addition if there exists $\boldsymbol{\sigma}^s \in \mathcal{C}^2(\overline{\mathcal{O}}))$ satisfying (2.10)*, then there is a constant $\varepsilon > 0$ independent of μ such that for any compact $\tilde{\mathcal{O}} \subset \mathcal{O}$*

$$\begin{aligned}
&\|\boldsymbol{\sigma}^\mu\|_2^2 + \mu^{-1}\|(|\boldsymbol{\sigma}^\mu_D| - 1)_+\|_1 + \|\operatorname{div} \boldsymbol{u}^\mu\|_2^2 + \|\boldsymbol{\varepsilon}(\boldsymbol{u}^\mu)\|_1 + \|\boldsymbol{u}^\mu(t)\|_{d'} \\
&+ \int_{\tilde{\mathcal{O}}} |\nabla \boldsymbol{\sigma}^\mu(x)|^2 \,\mathrm{d}x + \int_{\tilde{\mathcal{O}}} |\boldsymbol{u}^\mu(x)|^{d'+\varepsilon} \,\mathrm{d}x \le C(\tilde{\mathcal{O}}),
\end{aligned} \tag{2.19}$$

where the constant $C(\tilde{\mathcal{O}})$ depends only on $\boldsymbol{\sigma}^s$, A, $\mathcal{O}$ and $\tilde{\mathcal{O}}$. Moreover, there exists a subsequence that we do not relabel such that

$$\begin{aligned}
\boldsymbol{\sigma}^\mu &\rightharpoonup \boldsymbol{\sigma} && \text{in } L^2(\mathcal{O}, \mathbb{R}^{d\times d}_{sym}),\\
\mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ \frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|} &\rightharpoonup^* \boldsymbol{e}_p && \text{in } \mathcal{M}(\overline{\mathcal{O}}; \mathbb{R}^{d\times d}_{sym}),\\
\mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ &\rightharpoonup^* \lambda && \text{in } \mathcal{M}(\overline{\mathcal{O}}),\\
\boldsymbol{\varepsilon}(\boldsymbol{u}^\mu) &\rightharpoonup^* \boldsymbol{\varepsilon}(\boldsymbol{u}) && \text{in } \mathcal{M}(\overline{\mathcal{O}}; \mathbb{R}^{d\times d}_{sym}),\\
\boldsymbol{u}^\mu &\rightharpoonup \boldsymbol{u} && \text{in } L^{d'}(\mathcal{O}; \mathbb{R}^d).
\end{aligned} \tag{2.20}$$

where $\boldsymbol{\sigma}$ is a weak solution in sense of Definition 2.2, (1.8) *and* (1.9) *hold and*

$$\frac{\partial \psi(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} + \boldsymbol{e}_p = \boldsymbol{\varepsilon}(\boldsymbol{u}) \qquad \text{in } \overline{\mathcal{O}}. \tag{2.21}$$

In addition, there exists $\varepsilon > 0$ such that

$$\boldsymbol{\sigma} \in BMO(\mathcal{O}; \mathbb{R}^{d\times d}_{sym}), \qquad \boldsymbol{u} \in L^{d'+\varepsilon}(\mathcal{O}; \mathbb{R}^d). \tag{2.22}$$

Moreover, there exists positive constants C and δ such that

$$\begin{aligned}
&\operatorname{supp} \lambda \subset \{x \in \overline{\mathcal{O}} : \ M|\boldsymbol{\sigma}_D|(x) = 1\},\\
&\qquad \boldsymbol{e}_p = \frac{1}{2}\left((\boldsymbol{u} - \boldsymbol{u}_0) \otimes \boldsymbol{n} + \boldsymbol{n} \otimes (\boldsymbol{u} - \boldsymbol{u}_0)\right) \mathrm{d}S \ \text{ on } \partial\mathcal{O}_D,\\
&\qquad \boldsymbol{e}_p = \lambda \boldsymbol{\sigma}_D \ \text{ in } \mathcal{O} \setminus K,\\
&\lambda(B_R) \le C R^\delta \ \text{ for all balls } B_R \subset \mathcal{O},
\end{aligned} \tag{2.23}$$

where

$$K := \left\{ y \in \mathcal{O} : \ \liminf_{\varepsilon \to 0} \sup_{R \in (0,\varepsilon)} \varepsilon \int_{B_R(y)} R^{1-d-\varepsilon/2} |\boldsymbol{u}(x)|^{1+\varepsilon} \,\mathrm{d}x \ge 1 \right\}.$$

In particular, if $\boldsymbol{u} \in L^q(\mathcal{O}; \mathbb{R}^d)$ for some $q > d$ then

$$\boldsymbol{e}_p = \lambda \boldsymbol{\sigma}_D \ \text{ in } \mathcal{O}. \tag{2.24}$$

Consequently, due to (2.22)*, the identity* (2.24) *always holds for $d = 2$.*

To end this part of the paper, we want to emphasize the essential novelties stated in Theorems 2.1–2.2. While the existence part and the limiting part is rather standard, see [17, 18], there are several quite new results. The first one are the BMO estimates for the stress (2.15) and (2.22). Since their proof is somehow independent of the model, we shall summarize these estimates in Lemma 3.1 in the following section. Further, based on the BMO estimates we can improve the integrability of the displacement/velocity stated in (2.15) and (2.22). Notice that the same improvement was already done in [11, 10] but under the hypothesis that the Cauchy stress is bounded, which is not necessarily true for von Mieses condition. Having such improved estimates, we can then show the Morrey condition for the plastic strain, see $(2.16)_3$ and $(2.23)_3$, which is another essentially new information. Consequently, based on the Morrey condition and $W^{1,2}$ estimates for the Cauchy stress, one can identify the plastic strain point-wisely in terms of the Cauchy stress up to the set K, which in case $d=2$ or in case $\boldsymbol{u}$ or $\dot{\boldsymbol{u}} \in L^q$ with $q>d$ leads to the point-wise identification of the plastic strain by the values of the Cauchy stress everywhere in $\mathcal{O}$. The last novelty of the paper consists in the uniformity (μ independence) of the estimates (2.12) and (2.19). In particular, the μ-independent $W^{1,2}_{loc}$ estimates were not known for $d>4$ for the Perčina type approximation, and the fractional time regularity was known only for the limit problem but the estimates for the μ-approximation were still depending on μ and in fact exploding as $\mu \to 0_+$, see [9]. Furthermore, the method for the time regularity presented here, is on one hand based on the method developed in [8], but on the other hand is improved such that it leads to the global estimates over $(0,T)$ and not only to local.

In the rest of the paper, we will prove only Theorem 2.1 since the result for the Hencky model can be proven in a very similar way provided the uniform convexity of the potential ψ holds. We start the prove with the standard energy estimates, based on the safety load condition, in Subsection 3.1. The improved regularity estimates independent of μ are provided in Subsection 3.2–3.4. Next, we let $\mu \to 0$ in Subsection 3.5 to obtain the existence of solution to the original problem. Then in Subsection 3.6 we shall prove (2.15), which will be the information for proving the Morrey condition in Subsection 3.7 and finally also in the sharp identification of the limit in Subsection 3.9. Furthermore, in Subsection 3.8, we first mimic the method developed in [1, 2] to identify the plastic strain via regularization but also show that the plastic strain does not take place on the Neumann and the mixed part of the boundary, i.e., we have that the tensor-valued measures $\boldsymbol{e}_{\mathrm{p}}$ or $\dot{\boldsymbol{e}}_{\mathrm{p}}$ respectively, are not supported on $\partial\mathcal{O}_M \cup \partial\mathcal{O}_N$.

3. Proof of Theorem 2.1

As mentioned already in the introduction, we focus only on the Prandtl–Reuss model here and we shall use the Perčina approximation. We would like to notice here that a very similar procedure was developed in [18] with a slightly different approximation - the Norton-Hoff approximation. Also to simplify the presentation, we shall consider in what follows that $\kappa \equiv 1$.

3.1. **Approximation and standard a priori uniform estimates.** Thus, we shall assume that for any $\mu>0$ there exists a weak solution to Prandtl–Reuss-Perčina model according to Definition 2.3. The existence of such a σ can be shown e.g. by the Rothe approximation and we refer the interested reader to [9] or [16] for details. Moreover, one can easily find $\boldsymbol{u} \in W^{1,2}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}^d))$ such that for all $t \in (0,t)$ $\boldsymbol{u}-\boldsymbol{u}_0 \in \mathcal{V}$ and $\boldsymbol{u}(0)=\boldsymbol{u}_0(0)$. The relation (2.7) then can be point-wisely rewritten as

$$\mathsf{A}\dot{\boldsymbol{\sigma}} + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} = \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) \qquad \text{in } (0,T)\times\mathcal{O}. \tag{3.1}$$

The next step is to derive the uniform (μ independent estimates) for $(\boldsymbol{u},\boldsymbol{\sigma})$. We proceed here formally, since the estimates are known, see e.g. [18, 16, 9]. Taking the scalar product of (3.1) with $\boldsymbol{\sigma}-\boldsymbol{\sigma}^s$, recall here that $\boldsymbol{\sigma}^s$ satisfies the safety load condition (2.9), we deduce after integration over $\mathcal{O}$ that

$$
\begin{aligned}
&\int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol\sigma}-\dot{\boldsymbol\sigma}^s)\cdot(\boldsymbol\sigma-\boldsymbol\sigma^s)+\mu^{-1}(|\boldsymbol\sigma_D|-1)_+\frac{\boldsymbol\sigma_D\cdot(\boldsymbol\sigma_D-\boldsymbol\sigma^s_D)}{|\boldsymbol\sigma_D|}\,\mathrm{d}x\\
&\quad=\int_{\mathcal{O}}\boldsymbol\varepsilon(\dot{\boldsymbol u}-\dot{\boldsymbol u}_0)\cdot(\boldsymbol\sigma-\boldsymbol\sigma^s)\,\mathrm{d}x+\int_{\mathcal{O}}(\boldsymbol\varepsilon(\dot{\boldsymbol u}_0)-\mathsf{A}\cdot\boldsymbol\sigma^s)\cdot(\boldsymbol\sigma-\boldsymbol\sigma^s)\,\mathrm{d}x.
\end{aligned}
\tag{3.2}
$$

Since $\boldsymbol\sigma(t)$ and $\boldsymbol\sigma^s(t)$ belong to $\mathcal{F}_{el}(t)$ and $\boldsymbol u(t)-\boldsymbol u_0(t)$ belong to $\mathcal{V}$, we see that the first integral on the right hand side vanishes. Second, using the fact that $\boldsymbol\sigma^s$ satisfies the safety load condition, we observe

$$
\boldsymbol\sigma_D\cdot(\boldsymbol\sigma_D-\boldsymbol\sigma^s_D)=|\boldsymbol\sigma_D|^2-\boldsymbol\sigma_D\cdot\boldsymbol\sigma^s_D\ge|\boldsymbol\sigma_D|(|\boldsymbol\sigma_D|-|\boldsymbol\sigma^s_D|)\ge|\boldsymbol\sigma_D|(|\boldsymbol\sigma_D|-1+\delta).
$$

Finally, using the fact that A is symmetric, which follows from the definition of A (see (2.4)), and the fact that it is elliptic, see (2.3), we see that (3.2) leads to

$$
\begin{aligned}
&\frac{d}{dt}\int_{\mathcal{O}}\mathsf{A}(\boldsymbol\sigma-\boldsymbol\sigma^s)\cdot(\boldsymbol\sigma-\boldsymbol\sigma^s)\,\mathrm{d}x+2\int_{\mathcal{O}}\mu^{-1}\left((|\boldsymbol\sigma_D|-1)_+^2+\delta(|\boldsymbol\sigma_D|-1)_+\right)\,\mathrm{d}x\\
&\quad\le C\left(\|\boldsymbol\varepsilon(\dot{\boldsymbol u}_0)\|_2^2+\|\boldsymbol\sigma^s\|_2^2+\int_{\mathcal{O}}\mathsf{A}(\boldsymbol\sigma-\boldsymbol\sigma^s)\cdot(\boldsymbol\sigma-\boldsymbol\sigma^s)\,\mathrm{d}x\right)
\end{aligned}
$$

and consequently, by the Gronwall lemma, we deduce

$$
\begin{aligned}
&\sup_{t\in(0,T)}\|\boldsymbol\sigma(t)\|_2^2+\int_0^T\int_{\mathcal{O}}\mu^{-1}\left((|\boldsymbol\sigma_D|-1)_+^2+\delta(|\boldsymbol\sigma_D|-1)_+\right)\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\le C\int_0^T\|\boldsymbol\varepsilon(\dot{\boldsymbol u}_0)\|_2^2\,\mathrm{d}t+C\sup_{t\in(0,T)}\|\boldsymbol\sigma^s(t)\|_2^2+C\|\boldsymbol\sigma(0)\|_2^2\le C,
\end{aligned}
\tag{3.3}
$$

where the last inequality follows from the assumptions on data (namely on $\boldsymbol\sigma^s$).

The next step is to test (3.1) by $\dot{\boldsymbol\sigma}-\dot{\boldsymbol\sigma}^s$. Doing so, we get

$$
\begin{aligned}
&\int_{\mathcal{O}}\mathsf{A}\dot{\boldsymbol\sigma}\cdot\dot{\boldsymbol\sigma}+\mu^{-1}(|\boldsymbol\sigma_D|-1)_+\frac{\boldsymbol\sigma_D\cdot\dot{\boldsymbol\sigma}_D}{|\boldsymbol\sigma_D|}\,\mathrm{d}x=\int_{\mathcal{O}}\boldsymbol\varepsilon(\dot{\boldsymbol u}-\dot{\boldsymbol u}_0)\cdot(\dot{\boldsymbol\sigma}-\dot{\boldsymbol\sigma}^s)\,\mathrm{d}x\\
&\quad+\int_{\mathcal{O}}(\boldsymbol\varepsilon(\dot{\boldsymbol u}_0)\cdot(\dot{\boldsymbol\sigma}-\dot{\boldsymbol\sigma}^s)+\mathsf{A}\dot{\boldsymbol\sigma}\cdot\dot{\boldsymbol\sigma}^s+\mu^{-1}(|\boldsymbol\sigma_D|-1)_+\frac{\boldsymbol\sigma_D\cdot\dot{\boldsymbol\sigma}^s_D}{|\boldsymbol\sigma_D|}\,\mathrm{d}x.
\end{aligned}
\tag{3.4}
$$

Once again, the first term on the right hand side vanishes. For the first term on the left hand side, we use the assumption (2.3), while the second term, we shall rewrite as

$$
\mu^{-1}(|\boldsymbol\sigma_D|-1)_+\frac{\boldsymbol\sigma_D\cdot\dot{\boldsymbol\sigma}_D}{|\boldsymbol\sigma_D|}=\frac12\frac{\partial}{\partial t}\mu^{-1}(|\boldsymbol\sigma_D|-1)_+^2.
$$

Finally, using the Young and the Hölder inequality, the identity (3.4) then leads to

$$
\begin{aligned}
\frac{d}{dt}\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol\sigma_D|-1)_+^2\,\mathrm{d}x+2C_1\|\dot{\boldsymbol\sigma}\|_2^2&\le C_1\|\dot{\boldsymbol\sigma}\|_2^2+C(\|\boldsymbol\varepsilon(\dot{\boldsymbol u}_0)\|_2^2+\|\dot{\boldsymbol\sigma}^s\|_2^2)\\
&\quad+2\|\dot{\boldsymbol\sigma}^s_D\|_\infty\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol\sigma_D|-1)_+\,\mathrm{d}x.
\end{aligned}
$$

Thus, absorbing the first term on the right hand side by the corresponding term on the left hand side and integrating with respect to time, we deduce

$$
\begin{aligned}
\sup_{t\in(0,T)} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+^2 \,\mathrm{d}x + \int_0^T \|\dot{\boldsymbol{\sigma}}\|_2^2 \,\mathrm{d}t &\le C\int_0^T \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2 + \|\dot{\boldsymbol{\sigma}}^s\|_2^2 \,\mathrm{d}t \\
&\quad + C\|\dot{\boldsymbol{\sigma}}^s_D\|_{L^\infty((0,T)\times\mathcal{O})} \int_0^T \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \,\mathrm{d}x \,\mathrm{d}t \\
&\quad + \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(0)|-1)_+^2 \,\mathrm{d}x \\
&\le C,
\end{aligned} \tag{3.5}
$$

where the last inequality (with C being independent of μ) follows from the assumptions on $\boldsymbol{u}_0$, $\boldsymbol{\sigma}_0$, $\boldsymbol{\sigma}^s$, from the uniform estimate (3.3) and from the fact that $\|\boldsymbol{\sigma}_D(0)\|_\infty \le 1$ according to the assumptions. Moreover, going back to (3.1) and using (3.3) and (3.5), we see that

$$
\int_0^T \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\|_1 \,\mathrm{d}t \le C\int_0^T \int_{\mathcal{O}} |\dot{\boldsymbol{\sigma}}| + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \,\mathrm{d}x \,\mathrm{d}t \le C. \tag{3.6}
$$

Next, we shall improve the time regularity. To do so, we apply the time derivative to (3.1) and take the scalar product with $\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s$ to obtain[3]

$$
\begin{aligned}
&\int_{\mathcal{O}} \mathsf{A}(\ddot{\boldsymbol{\sigma}} - \ddot{\boldsymbol{\sigma}}^s)\cdot(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s) + \mu^{-1}\frac{\partial (|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}}{\partial t}\cdot(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\sigma}}^s)\,\mathrm{d}x \\
&\qquad = \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}} - \ddot{\boldsymbol{u}}_0)\cdot(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)\,\mathrm{d}x + \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0) - \mathsf{A}\ddot{\boldsymbol{\sigma}}^s)\cdot(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)\,\mathrm{d}x.
\end{aligned} \tag{3.7}
$$

The first term on the right hand side vanishes and for the part of the second term on the left hand side we have the estimate

$$
\mu^{-1}\frac{\partial (|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}}{\partial t}\cdot\dot{\boldsymbol{\sigma}}_D = \frac{\mu^{-1}\chi_{|\boldsymbol{\sigma}_D|>1}}{|\boldsymbol{\sigma}_D|}\left(|\dot{\boldsymbol{\sigma}}_D|^2(|\boldsymbol{\sigma}_D|-1) + |\partial_t|\boldsymbol{\sigma}_D||^2\right) \ge 0. \tag{3.8}
$$

Consequently, using the Hölder inequality and the above estimate, we see that (3.7) implies

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)\cdot(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s) - 2\mu^{-1}\frac{(|\boldsymbol{\sigma}_D|-1)_+\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot\dot{\boldsymbol{\sigma}}^s_D\,\mathrm{d}x \\
&\quad \le \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0) - \mathsf{A}\ddot{\boldsymbol{\sigma}}^s)\cdot(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s) - \mu^{-1}\frac{(|\boldsymbol{\sigma}_D|-1)_+\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot\ddot{\boldsymbol{\sigma}}^s_D\,\mathrm{d}x \\
&\quad \le C(\|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0)\|_2 + \|\ddot{\boldsymbol{\sigma}}^s\|_2)(1 + \|\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s\|_2^2) + \|\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\|_1\|\ddot{\boldsymbol{\sigma}}^s_D\|_\infty.
\end{aligned} \tag{3.9}
$$

Next, multiplying the identity (3.2) by a constant $K_1 \ge 1$, that we shall specify later, applying the time derivative and adding the result to (3.9), and then adding the term $K_2\frac{d}{dt}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}^s\|_2^2$ with $K_2 \ge 1$ to both sides of the resulting inequality, we find

$$
\begin{aligned}
\frac{d}{dt}Q_K(t) \le W_K(t)\big(&\|\dot{\boldsymbol{\sigma}}(t) - \dot{\boldsymbol{\sigma}}^s(t)\|_2^2 + \|\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t)\|_2^2 \\
&+ \|\mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+\|_1\big) + W_K(t),
\end{aligned} \tag{3.10}
$$

[3]Although $\ddot{\boldsymbol{\sigma}}$ need not to exist in general, one can make the proof rigorous by performing similar uniform estimates for the Galerkin approximation, use the ODE uniqueness and finally pass to the limiting problem, which admits the unique solution.

where

$$
\begin{aligned}
Q_K(t) &:= \frac{1}{2}\int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol{\sigma}}(t)-\dot{\boldsymbol{\sigma}}^s(t))\cdot\big(\dot{\boldsymbol{\sigma}}(t)-\dot{\boldsymbol{\sigma}}^s(t)+2K_1(\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t))\big)\,\mathrm{d}x\\
&\quad+\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+\frac{\boldsymbol{\sigma}_D(t)}{|\boldsymbol{\sigma}_D(t)|}\cdot(K_1(\boldsymbol{\sigma}_D(t)-\boldsymbol{\sigma}^s_D(t))-\dot{\boldsymbol{\sigma}}^s_D(t))\,\mathrm{d}x\\
&\quad+K_2\int_{\mathcal{O}}|\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t)|^2\,\mathrm{d}x,\\
W_K(t) &:= CK_1K_2\,(1+\|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0(t))\|_2+\|\ddot{\boldsymbol{\sigma}}^s(t)\|_2+\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))\|_2+\|\dot{\boldsymbol{\sigma}}^s(t)\|_2\\
&\qquad+\|\ddot{\boldsymbol{\sigma}}^s_D(t)\|_\infty)\,,
\end{aligned}
$$

where $K_1\ge 0$ is arbitrary constant. Finally, using (2.3) and (2.9), we can observe that

$$
\begin{aligned}
Q_K(t) &\ge \frac{C_1}{2}\|\dot{\boldsymbol{\sigma}}(t)-\dot{\boldsymbol{\sigma}}^s(t)\|_2^2-K_1C_2\|\dot{\boldsymbol{\sigma}}(t)-\dot{\boldsymbol{\sigma}}^s(t)\|_2\|\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t)\|_2\\
&\quad+\|\mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+\|_1(K_1\delta-\|\dot{\boldsymbol{\sigma}}^s_D(t)\|_\infty)+K_2\|\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t)\|_2^2\\
&\quad+K_1\|\mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+^2\|_1.
\end{aligned}
$$

Therefore, setting

$$
\begin{aligned}
K_1 &:= \frac{1+\sup_{t\in(0,T)}\|\dot{\boldsymbol{\sigma}}^s_D(t)\|_\infty}{\delta},\\
K_2 &:= 1+\frac{K_1^2C_2^2}{C_1},
\end{aligned}
$$

we obtain

$$
\begin{aligned}
Q_K(t) &\ge \frac{C_1}{4}\|\dot{\boldsymbol{\sigma}}(t)-\dot{\boldsymbol{\sigma}}^s(t)\|_2^2+\|\mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+)\|_1+\|\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t)\|_2^2\\
&\quad K_1\|\mu^{-1}(|\boldsymbol{\sigma}_D(t)|-1)_+^2\|_1
\end{aligned}
$$

and inserting this estimate into (3.10), we are led to the following inequality

$$
\frac{d}{dt}Q_K(t)\le W_K(t)Q_K(t)+W_K(t), \tag{3.11}
$$

which with the help of the Gronwall lemma and the fact that $W_K\in L^1(0,T)$ (see the assumptions on the data) imply that

$$
\sup_{t\in(0,T)} Q_K(t)\le C(1+Q_K(0)).
$$

Since $\boldsymbol{\sigma}_D(0)\equiv 0$, we also see that

$$
Q_K(0)\le C(1+\|\dot{\boldsymbol{\sigma}}(0)\|_2^2).
$$

However, since $\dot{\boldsymbol{\sigma}}(0)$ satisfies (3.1), where the second term vanishes, and also $\boldsymbol{\sigma}\in\mathcal{F}$, we see that $\dot{\boldsymbol{\sigma}}(0)$ solves the purely elastic problem, and it is rather standard to deduce the estimate

$$
Q_K(t)\le C(1+\|\dot{\boldsymbol{\sigma}}(0)\|_2^2)\le CC(1+\|\dot{\boldsymbol{\sigma}}^s(0)\|_2^2+\|\boldsymbol{\sigma}^s(0)\|_2^2).
$$

Therefore, using the assumptions on data we finally find that

$$
\sup_{t\in(0,T)}\big(\|\dot{\boldsymbol{\sigma}}(t)\|_2^2+\mu^{-1}\|(|\boldsymbol{\sigma}_D(t)|-1)_+\|_1+\mu^{-1}\|(|\boldsymbol{\sigma}_D(t)|-1)_+\|_2^2\big)\le C \tag{3.12}
$$

and using the identity (3.1) and also the embedding theorem we also get[4]
(3.13)
$$\sup_{t\in(0,T)}\Big(\|\operatorname{div}\dot{\boldsymbol{u}}\|_2^2+\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}(t))\|_1+\|\dot{\boldsymbol{u}}(t)\|_{d'}+\mu(\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}(t))\|_2^2+\|\dot{\boldsymbol{u}}(t)\|^2_{\frac{2d}{d-2}})\Big)\le C,$$
which is the first part of the uniform estimate (2.12).

3.2. **Uniform interior $W^{1,2}$ estimates.** In this subsection, we shall derive the uniform interior estimates on $\nabla\boldsymbol{\sigma}^\mu$. Already here, we present a certain novelty, since for the Perčina approximation such uniform estimates are known only for $d\le 4$. To shorten the notation we will not use the superscript μ to denote the solution of the μ-th approximation. Nevertheless, we will trace the dependence of all estimates on μ. In particular, the constant C, which may change line to line will be always independent of μ and in case there is some dependence, it will be clearly denoted. Thus, let $\xi\in\mathcal{D}(\mathcal{O})$ be arbitrary nonnegative function satisfying $|\xi|\le 1$. Next, we apply the operator ∇ to equation (3.1), take the scalar product with $\nabla\boldsymbol{\sigma}\xi^{2m}$ for some $m\in\mathbb{N}$ and integrate the result over $\mathcal{O}$. Note here that such a procedure is only formal, however can be easily justified by using the difference quotient method. Hence, we get the identity
(3.14)
$$\int_{\mathcal{O}}\mathsf{A}\nabla\dot{\boldsymbol{\sigma}}\cdot\nabla\boldsymbol{\sigma}\xi^{2m}+\mu^{-1}\nabla\left(\frac{(|\boldsymbol{\sigma}_D|-1)_+\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\right)\cdot\nabla\boldsymbol{\sigma}\xi^{2m}\,\mathrm{d}x=\int_{\mathcal{O}}\nabla\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\nabla\boldsymbol{\sigma}\xi^{2m}\,\mathrm{d}x.$$
Using the symmetry of A and the computation very similar to (3.8) we see that (3.14) leads to
$$\begin{aligned}&\frac12\frac{d}{dt}\int_{\mathcal{O}}\mathsf{A}\nabla\boldsymbol{\sigma}\cdot\nabla\boldsymbol{\sigma}\xi^{2m}\,\mathrm{d}x+\mu^{-1}\int_{|\boldsymbol{\sigma}|>1}\frac{|\nabla|\boldsymbol{\sigma}_D||^2+|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)\xi^{2m}}{|\boldsymbol{\sigma}_D|}\,\mathrm{d}x\\&\qquad=\int_{\mathcal{O}}\nabla\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\nabla\boldsymbol{\sigma}\xi^{2m}\,\mathrm{d}x.\end{aligned}\tag{3.15}$$
Next, we focus on the estimate of the term appearing on the right hand side of (3.15). Abbreviating $D_k:=\frac{\partial}{\partial x_k}$ and using the integration by parts and the fact that $\boldsymbol{\sigma}(t)\in\mathcal{F}$, we find
$$\begin{aligned}\int_{\mathcal{O}}\nabla\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\nabla\boldsymbol{\sigma}\xi^{2m}\,\mathrm{d}x&=\int_{\mathcal{O}}\sum_{i,j,k=1}^d D_kD_j\dot{\boldsymbol{u}}_iD_k\boldsymbol{\sigma}_{ij}\xi^{2m}\,\mathrm{d}x\\&=\int_{\mathcal{O}}\sum_{i,j,k=1}^d 2D_j\boldsymbol{\varepsilon}_{ik}(\dot{\boldsymbol{u}})D_k\boldsymbol{\sigma}_{ij}\xi^{2m}-D_{ij}\dot{\boldsymbol{u}}_kD_k\boldsymbol{\sigma}_{ij}\xi^{2m}\,\mathrm{d}x\\&=-2\int_{\mathcal{O}}\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\nabla\boldsymbol{f}\xi^{2m}+\sum_{i,j,k=1}^d\boldsymbol{\varepsilon}_{ik}(\dot{\boldsymbol{u}})D_k\boldsymbol{\sigma}_{ij}D_j\xi^{2m}\,\mathrm{d}x\\&\quad+\int_{\mathcal{O}}\operatorname{div}\dot{\boldsymbol{u}}\operatorname{div}\boldsymbol{f}\xi^{2m}+\operatorname{div}\boldsymbol{f}\,\dot{\boldsymbol{u}}\cdot\nabla\xi^{2m}-\nabla\boldsymbol{f}\cdot(\dot{\boldsymbol{u}}\otimes\nabla\xi^{2m})\,\mathrm{d}x\\&\quad-\int_{\mathcal{O}}\sum_{i,j,k=1}^d\dot{\boldsymbol{u}}_kD_k\boldsymbol{\sigma}_{ij}D_{ij}\xi^{2m}\,\mathrm{d}x\\&=:I_1+I_2+I_3.\end{aligned}$$

[4]In case that $d=2$, we replace the fraction $2d/(d-2)$ by arbitrary $q\in(1,\infty)$.

First, with the help of (3.13), the second term can be easily estimate by the Hölder inequality as

$$I_2 \le C(\xi,m)\|\boldsymbol{f}(t)\|_{1,d}.$$

For the first term, i.e. for the term I_1, we use the Hölder inequality and the estimate (3.13), to handle the first integral in I_1, and the identity (3.2) and the estimate (3.12) to handle the second integral in I_1 as follows

$$\begin{aligned}
I_1 &\le C(\xi,m)\|\nabla \boldsymbol{f}(t)\|_\infty - 2\int_{\mathcal{O}} \sum_{i,j,k=1}^d \boldsymbol{\varepsilon}_{ik}(\dot{\boldsymbol{u}}) D_k \boldsymbol{\sigma}_{ij} D_j \xi^{2m}\,\mathrm{d}x \\
&= C(\xi,m)\|\nabla \boldsymbol{f}(t)\|_\infty \\
&\quad - 2\int_{\mathcal{O}} \sum_{i,j,k=1}^d \left((\mathsf{A}\dot{\boldsymbol{\sigma}})_{ik} + \mu^{-1}\frac{(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}(\boldsymbol{\sigma}_D)_{ik}\right) D_k\boldsymbol{\sigma}_{ij} D_j \xi^{2m}\,\mathrm{d}x \\
&\le C(\xi,m)(\|\nabla \boldsymbol{f}(t)\|_\infty + \|\nabla \boldsymbol{\sigma}(t)\xi^m\|_2^2) \\
&\quad + C(m,\xi)\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+)|\nabla\boldsymbol{\sigma}|\xi^m\,\mathrm{d}x.
\end{aligned}$$

Finally, using the identity

$$D_k\boldsymbol{\sigma}_{ij} = D_k(\boldsymbol{\sigma}_D)_{ij} + \delta_{ij}(\boldsymbol{f}_k - \sum_{\ell=1}^d D_\ell(\boldsymbol{\sigma}_D)_{k\ell}), \tag{3.16}$$

we see that

$$\begin{aligned}
I_1 &\le C(\xi,m)(\|\nabla \boldsymbol{f}(t)\|_\infty + \|\nabla \boldsymbol{\sigma}(t)\xi^m\|_2^2) \\
&\quad + C(m,\xi)\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+)(|\boldsymbol{f}| + |\nabla\boldsymbol{\sigma}_D|)\xi^m\,\mathrm{d}x \\
&\le C(\xi,m)(\|\boldsymbol{f}(t)\|_{1,\infty} + \|\nabla \boldsymbol{\sigma}(t)\xi^m\|_2^2) \\
&\quad + \int_{\mathcal{O}} C(m,\xi)\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+)|\boldsymbol{\sigma}_D| + \frac{\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+|\nabla\boldsymbol{\sigma}_D|^2}{2|\boldsymbol{\sigma}_D|}\xi^m\,\mathrm{d}x \\
&\le C(\xi,m)(1 + \|\boldsymbol{f}(t)\|_{1,\infty} + \|\nabla \boldsymbol{\sigma}(t)\xi^m\|_2^2) + \frac{1}{2}\int_{\mathcal{O}} \frac{\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+|\nabla\boldsymbol{\sigma}_D|^2}{|\boldsymbol{\sigma}_D|}\xi^m\,\mathrm{d}x,
\end{aligned}$$

where for the last estimate we used the a priori bound (3.12). Note here that the last term on the right hand side will be absorbed by the second term on the left hand side of (3.15). Finally, for the remaining term I_3 we use the identity (3.16) and integration by parts to find (using also the estimates (3.12)–(3.13))

$$\begin{aligned}
I_3 &= -\int_{\mathcal{O}} \sum_{i,j,k=1}^d \dot{\boldsymbol{u}}_k D_k(\boldsymbol{\sigma}_D)_{ij} D_{ij}\xi^{2m} + \dot{\boldsymbol{u}}\cdot \boldsymbol{f}\triangle\xi^{2m} - \sum_{k,\ell=1}^d \dot{\boldsymbol{u}}_k D_\ell(\boldsymbol{\sigma}_D)_{k\ell})\triangle\xi^{2m}\,\mathrm{d}x \\
&\le C(\xi,m)\|\boldsymbol{f}(t)\|_d + \int_{\mathcal{O}} \operatorname{div}\dot{\boldsymbol{u}}\,\boldsymbol{\sigma}_D\cdot\nabla^2\xi^{2m} - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\boldsymbol{\sigma}_D\triangle\xi^{2m}\,\mathrm{d}x \\
&\quad + \int_{\mathcal{O}} \sum_{i,j,k=1}^d \dot{\boldsymbol{u}}_k(\boldsymbol{\sigma}_D)_{ij}D_{ij}D_k\xi^{2m} - \sum_{k,\ell=1}^d \dot{\boldsymbol{u}}_k(\boldsymbol{\sigma}_D)_{k\ell}D_\ell\triangle\xi^{2m}\,\mathrm{d}x \\
&\le C(\xi,m)\|\boldsymbol{f}(t)\|_d + C(\xi,m)\int_{\mathcal{O}} (|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})| + |\dot{\boldsymbol{u}}|)|\boldsymbol{\sigma}_D|\,\mathrm{d}x.
\end{aligned}$$

Consequently, using (3.1), we see that

$$|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})||\boldsymbol{\sigma}_D| \le C|\dot{\boldsymbol{\sigma}}||\boldsymbol{\sigma}| + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+^2 + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+$$

and by using (3.12)–(3.13) and the Hölder inequality and the embedding theorem, we may continue

$$\begin{aligned} I_3 &\le C(\xi,m)(\|\boldsymbol{f}(t)\|_d + 1) + C(\xi,m)\int_{\mathcal{O}} |\dot{\boldsymbol{u}}| + |\dot{\boldsymbol{u}}|(|\boldsymbol{\sigma}_D|-1)_+ \,\mathrm{d}x \\ &\le C(\xi,m)(\|\boldsymbol{f}(t)\|_d + 1) + C(\xi,m)\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\|_2 \|(|\boldsymbol{\sigma}_D|-1)_+\|_2 \\ &\le C(\xi,m)(\|\boldsymbol{f}(t)\|_d + 1). \end{aligned}$$

Finally, using all estimates for I_1, I_2 and I_3, the ellipticity condition (2.3) and absorbing the term arising in I_1 by the left hand side, we find (noticing also that $\boldsymbol{f} = \operatorname{div} \boldsymbol{\sigma}^s$, which follows from (2.9) and the definition of $\mathcal{F}(t)$)

$$\begin{aligned} &\frac{d}{dt}\int_{\mathcal{O}} \mathbf{A}\nabla\boldsymbol{\sigma}(t)\cdot\nabla\boldsymbol{\sigma}(t)\xi^{2m}\,\mathrm{d}x + \mu^{-1}\int_{|\boldsymbol{\sigma}|>1} \frac{|\nabla|\boldsymbol{\sigma}_D||^2 + |\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)\xi^{2m}}{|\boldsymbol{\sigma}_D|}\,\mathrm{d}x \\ &\qquad \le C(\xi,m)\left(1 + \|\boldsymbol{\sigma}^s(t)\|_{2,\infty} + \int_{\mathcal{O}} \mathbf{A}\nabla\boldsymbol{\sigma}(t)\cdot\nabla\boldsymbol{\sigma}(t)\xi^{2m}\,\mathrm{d}x\right) \end{aligned}$$

and the Gronwall lemma and the assumptions on data directly lead to the estimate

$$\begin{aligned} &\sup_{t\in(0,T)}\int_B |\nabla\boldsymbol{\sigma}(t)|^2 + |\nabla \operatorname{div}\boldsymbol{u}(t)|^2\,\mathrm{d}x + \mu^{-1}\int_0^T\int_B \frac{|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\,\mathrm{d}x\,\mathrm{d}t \\ &\le C(B), \end{aligned} \tag{3.17}$$

for arbitrary open $B \subset \overline{B} \subset \mathcal{O}$. Thus $W^{1,2}_{loc}$ estimate in (2.12) is proven.

3.3. **Uniform $L^{d'+\varepsilon}$ estimates for $\dot{\boldsymbol{u}}$.** Finally we present a proof of the uniform, i.e., μ "independent", local in $\mathcal{O}$ estimates in any dimension $d \ge 2$. In addition, we would like to emphasize that contrary to [11] or [10] these estimates do not rely on the L^∞ control of $\boldsymbol{\sigma}$. These estimates, will be further extended to the whole $\mathcal{O}$ for the limit solution in next subsections. We again do not use here the superscript μ for denoting the solution of the μ-th approximation, but we shall clearly describe any dependence on μ in all estimates presented below.

Hence, following the above mentioned paper, we fix $x_0 \in \mathcal{O}$ and denote $R^* \le 1$ the largest number such that $B_{2R^*}(x_0) \subset \mathcal{O}$. Then for every $R \le R^*$ we can find $\tau \in \mathcal{D}(\mathcal{O})$ such that $\tau \equiv 1$ in $B_R(x_0)$, $\tau \equiv 0$ outside $B_{2R}(x_0)$ and $|\nabla\tau| \le R^{-1}$. Then it directly follows from the embedding theorem that

$$\left(\int_{\mathcal{O}} |\dot{\boldsymbol{u}}\tau|^{d'}\,\mathrm{d}x\right)^{\frac{1}{d'}} \le C\int_{\mathcal{O}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}\tau)|\,\mathrm{d}x \le C\int_{\mathcal{O}} |\dot{\boldsymbol{u}}||\nabla\tau| + |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})|\tau\,\mathrm{d}x$$

Hence, dividing the resulting inequality by R^{d-1} and using the properties of τ we find that

$$\left(\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}|^{d'}}{R^d}\,\mathrm{d}x\right)^{\frac{1}{d'}} \le C\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}|}{R^d}\,\mathrm{d}x + \int_{\mathcal{O}} \frac{|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})|\tau}{R^{d-1}}\,\mathrm{d}x. \tag{3.18}$$

Next, we focus on the last term on the right hand side. Using the identity (3.1), we see that

$$\begin{aligned}|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})| &\le C|\dot{\boldsymbol{\sigma}}| + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \le C|\dot{\boldsymbol{\sigma}}| + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot\frac{\boldsymbol{\sigma}_D}{1+(|\boldsymbol{\sigma}_D|-2)_+}\\ &\le C|\dot{\boldsymbol{\sigma}}| + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\frac{\boldsymbol{\sigma}_D}{1+(|\boldsymbol{\sigma}_D|-2)_+}.\end{aligned}$$

Consequently, we find that

$$\begin{aligned}&\int_{\mathcal{O}}\frac{|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})|\tau}{R^{d-1}}\,\mathrm{d}x \le CR\int_{\mathcal{O}}\frac{|\dot{\boldsymbol{\sigma}}|\tau}{R^d}\,\mathrm{d}x + \int_{\mathcal{O}}\frac{\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\cdot\boldsymbol{\sigma}_D\tau}{R^{d-1}(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x\\
&= CR\int_{\mathcal{O}}\frac{|\dot{\boldsymbol{\sigma}}|\tau}{R^d}\,\mathrm{d}x - \int_{\mathcal{O}}\frac{\dot{\boldsymbol{u}}\cdot\operatorname{div}\boldsymbol{\sigma}_D\tau + \boldsymbol{\sigma}_D\cdot(\dot{\boldsymbol{u}}\otimes\nabla\tau)}{R^{d-1}(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x\\
&\quad + \int_{\mathcal{O}}\frac{\boldsymbol{\sigma}_D\cdot(\dot{\boldsymbol{u}}\otimes\nabla(|\boldsymbol{\sigma}_D|-2)_+)\tau}{R^{d-1}(1+(|\boldsymbol{\sigma}_D|-2)_+)^2}\,\mathrm{d}x\\
&\le CR\int_{\mathcal{O}}\frac{|\dot{\boldsymbol{\sigma}}|\tau}{R^d} + \frac{|\dot{\boldsymbol{u}}||\nabla\boldsymbol{\sigma}_D|((|\boldsymbol{\sigma}_D|-1)_+)^{\frac12}\tau}{R^d|\boldsymbol{\sigma}_D|^{\frac12}}\,\mathrm{d}x + C\int_{B_{2R}(x_0)}\frac{|\dot{\boldsymbol{u}}|}{R^d}\,\mathrm{d}x\\
&\quad - \frac{1}{d}\int_{\mathcal{O}}\frac{d\dot{\boldsymbol{u}}\cdot\boldsymbol{f}\tau - \dot{\boldsymbol{u}}\cdot\nabla\operatorname{tr}\boldsymbol{\sigma}\tau}{R^{d-1}(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x\\
&\le CR\int_{\mathcal{O}}\frac{|\dot{\boldsymbol{\sigma}}|\tau + |\dot{\boldsymbol{u}}||\boldsymbol{f}|\tau}{R^d} + \frac{|\dot{\boldsymbol{u}}||\nabla\boldsymbol{\sigma}_D|((|\boldsymbol{\sigma}_D|-1)_+)^{\frac12}\tau}{R^d|\boldsymbol{\sigma}_D|^{\frac12}}\,\mathrm{d}x + C\int_{B_{2R}(x_0)}\frac{|\dot{\boldsymbol{u}}|}{R^d}\,\mathrm{d}x\\
&\quad + \frac{1}{d}\int_{\mathcal{O}}\frac{\dot{\boldsymbol{u}}\cdot\nabla(\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}})\tau}{R^{d-1}} - \frac{\dot{\boldsymbol{u}}\cdot\nabla\operatorname{tr}\boldsymbol{\sigma}\tau(|\boldsymbol{\sigma}_D|-2)_+\tau}{R^{d-1}(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x\\
&\le CR\int_{\mathcal{O}}\frac{(|\dot{\boldsymbol{\sigma}}| + |\dot{\boldsymbol{u}}||\boldsymbol{f}| + |\operatorname{div}\dot{\boldsymbol{u}}||\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|)\tau}{R^d} + \frac{|\dot{\boldsymbol{u}}||\nabla\boldsymbol{\sigma}_D|((|\boldsymbol{\sigma}_D|-1)_+)^{\frac12}}{R^d|\boldsymbol{\sigma}_D|^{\frac12}}\,\mathrm{d}x\\
&\quad + C\int_{B_{2R}(x_0)}\frac{|\dot{\boldsymbol{u}}|}{R^d} + \frac{|\dot{\boldsymbol{u}}||\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|}{R^d}\,\mathrm{d}x,\end{aligned}$$

where we denoted

$$\overline{\operatorname{tr}\boldsymbol{\sigma}} := \frac{1}{|B_{2R}(x_0)|}\int_{B_{2R}(x_0)}\operatorname{tr}\boldsymbol{\sigma}.$$

Next, we estimate the terms with $\overline{\operatorname{tr}\boldsymbol{\sigma}}$. First, using the Young and the Hölder inequalities, we find that for arbitrary $q \in (2, \frac{2d}{d-2})$ that

$$\begin{aligned}\int_{B_{2R}(x_0)}\frac{|\operatorname{div}\dot{\boldsymbol{u}}||\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|}{R^d}\,\mathrm{d}x &\le C(q)\int_{B_{2R}(x_0)}\frac{|\operatorname{div}\dot{\boldsymbol{u}}|^{q'} + |\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|^q}{R^d}\,\mathrm{d}x\\ &\le C(q)\int_{B_{2R}(x_0)}\frac{|\operatorname{div}\dot{\boldsymbol{u}}|^{q'} + |\boldsymbol{\sigma}|^q}{R^d}\,\mathrm{d}x.\end{aligned}$$

Finally, we focus on the estimate of the last term with $\overline{\operatorname{tr}\boldsymbol{\sigma}}$. For this purpose we use the Bogovskii operator, see [7, Theorem 10.11], and we can find $\boldsymbol{v} \in W_0^{1,1}(B_{2R}(x_0);\mathbb{R}^d)$ such that

$$\begin{aligned}&\operatorname{div}\boldsymbol{v} = \operatorname{sgn}\,(\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}})|\dot{\boldsymbol{u}}| - \overline{\operatorname{sgn}\,(\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}})|\dot{\boldsymbol{u}}|},\\ &\int_{B_{2R}(x_0)}R^p|\boldsymbol{v}|^p + |\nabla\boldsymbol{v}|^p\,\mathrm{d}x \le C(p)\int_{B_{2R}(x_0)}|\dot{\boldsymbol{u}}|^p\,\mathrm{d}x,\end{aligned}\tag{3.19}$$

for all $(p \in (1,\infty)$. Hence, we observe with the help of (3.16) that

$$
\begin{aligned}
&\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}||\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|}{R^d}\,\mathrm{d}x = \int_{B_{2R}(x_0)} \frac{\operatorname{div}\boldsymbol{v}(\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}})}{R^d}\,\mathrm{d}x \\
&\quad = -\int_{B_{2R}(x_0)} \frac{\boldsymbol{v}\cdot\nabla\operatorname{tr}\boldsymbol{\sigma}}{R^d}\,\mathrm{d}x = -d\int_{B_{2R}(x_0)} \frac{\boldsymbol{v}\cdot(\boldsymbol{f} - \operatorname{div}\boldsymbol{\sigma}_D)}{R^d}\,\mathrm{d}x \\
&\quad = -d\int_{B_{2R}(x_0)} \frac{\boldsymbol{v}\cdot\operatorname{div}\boldsymbol{\sigma}^s}{R^d} - \frac{\boldsymbol{v}\cdot\operatorname{div}\boldsymbol{\sigma}_D(|\boldsymbol{\sigma}_D|-2)_+}{R^d(1+(|\boldsymbol{\sigma}_D|-2)_+)} - \frac{\boldsymbol{v}\cdot\operatorname{div}\boldsymbol{\sigma}_D}{R^d(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x \\
&\quad = d\int_{B_{2R}(x_0)} \frac{\nabla\boldsymbol{v}\cdot\boldsymbol{\sigma}^s}{R^d} - \frac{\nabla\boldsymbol{v}\cdot\boldsymbol{\sigma}_D}{R^d(1+(|\boldsymbol{\sigma}_D|-2)_+)}\,\mathrm{d}x \\
&\qquad + d\int_{B_{2R}(x_0)} \frac{\boldsymbol{v}\cdot\operatorname{div}\boldsymbol{\sigma}_D(|\boldsymbol{\sigma}_D|-2)_+}{R^d(1+(|\boldsymbol{\sigma}_D|-2)_+)} + \frac{\boldsymbol{\sigma}_D\cdot(\boldsymbol{v}\otimes\nabla(|\boldsymbol{\sigma}_D|-2)_+)}{R^d(1+(|\boldsymbol{\sigma}_D|-2)_+)^2}\,\mathrm{d}x \\
&\quad \le C\int_{B_{2R}(x_0)} \frac{|\nabla\boldsymbol{v}|(|\boldsymbol{\sigma}^s|+1)}{R^d}\,\mathrm{d}x + C\int_{B_{2R}(x_0)} \frac{|\boldsymbol{v}||\nabla\boldsymbol{\sigma}_D|(|\boldsymbol{\sigma}_D|-1)_+^{\frac12}}{R^d|\boldsymbol{\sigma}_D|^{\frac12}}\,\mathrm{d}x.
\end{aligned}
$$

Consequently, using (3.19) and the assumption on $\boldsymbol{\sigma}^s$, we get for all $p,q \in (1,\infty)$

$$
\begin{aligned}
&\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}||\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}|}{R^d}\,\mathrm{d}x \le C\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}|^p}{R^d}\,\mathrm{d}x\right)^{\frac1p} \\
&+ CR\left(\int_{B_{2R}(x_0)} \frac{(\mu|\dot{\boldsymbol{u}}|^2)^{\frac q2}}{R^d}\right)^{\frac1q}\left(\int_{B_{2R}(x_0)} \left(\frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\right)^{\frac{q'}{2}} R^{-d}\,\mathrm{d}x\right)^{\frac{1}{q'}}.
\end{aligned}
$$

Hence, going back to (3.18) and substituting all terms and using the Young inequality, we have that for arbitrary $p \in (1,d')$ and arbitrary $q \in (2, \frac{2d}{d-2})$ and arbitrary $\delta > 0$ that

$$
\begin{aligned}
&\left(\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}|^{d'}}{R^d}\,\mathrm{d}x\right)^{\frac{1}{d'}} \le C(p)\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}|^p}{R^d}\,\mathrm{d}x\right)^{\frac1p} \\
&\quad + C(q)\int_{B_{2R}(x_0)} \frac{|\operatorname{div}\dot{\boldsymbol{u}}|^{q'} + |\boldsymbol{\sigma}|^q + |\dot{\boldsymbol{\sigma}}| + |\dot{\boldsymbol{u}}||\boldsymbol{f}|}{R^{d-1}}\,\mathrm{d}x \\
&\quad + C(q)\int_{B_{2R}(x_0)} \frac{\delta^{1-q}(\mu|\dot{\boldsymbol{u}}|^2)^{\frac q2} + \delta\left(\frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\right)^{\frac{q'}{2}}}{R^{d-1}}\,\mathrm{d}x\,.
\end{aligned}
\tag{3.20}
$$

In order to estimate last two terms, we find $\psi_{1,2} \in W_0^{1,1}(\mathcal{O})$ solving

$$
\begin{aligned}
\triangle\psi_1 &= |\operatorname{div}\dot{\boldsymbol{u}}|^{q'} + |\boldsymbol{\sigma}|^q + |\dot{\boldsymbol{\sigma}}| + |\dot{\boldsymbol{u}}||\boldsymbol{f}| && \text{in } \mathcal{O}, \\
\triangle\psi_2 &= \delta^{1-q}(\mu|\dot{\boldsymbol{u}}|^2)^{\frac q2} + \delta\left(\frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\right)^{\frac{q'}{2}}. && \text{in } \mathcal{O}
\end{aligned}
$$

Then, we find a nonnegative ξ being equal to one in $B_{2R}(x_0)$ and vanishing outside $B_{3R}(x_0)$ fulfilling $|\nabla\xi| \le CR^{-1}$ and deduce by integration by parts that

$$
\begin{aligned}
C\int_{B_{3R}(x_0)} \frac{|\nabla\psi_1|+|\nabla\psi_2|}{R^d}\,\mathrm{d}x &\ge \int_{B_{2R}(x_0)} \frac{|\operatorname{div}\dot{\boldsymbol{u}}|^{q'}+|\boldsymbol{\sigma}|^q+|\dot{\boldsymbol{\sigma}}|+|\dot{\boldsymbol{u}}||\boldsymbol{f}|}{R^{d-1}}\,\mathrm{d}x\\
&+\int_{B_{2R}(x_0)} \frac{\delta^{1-q}(\mu|\dot{\boldsymbol{u}}|^2)^{\frac{q}{2}}+\delta\left(\frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\right)^{\frac{q'}{2}}}{R^{d-1}}\,\mathrm{d}x\,.
\end{aligned}
$$

Using this estimate in (3.20), we obtain

$$
\begin{aligned}
\left(\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}|^{d'}}{R^d}\,\mathrm{d}x\right)^{\frac{1}{d'}} &\le C(p)\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}|^{p}}{R^d}\,\mathrm{d}x\right)^{\frac{1}{p}}\\
&\quad + C(q)\int_{B_{3R}(x_0)} \frac{|\nabla\psi_1|+|\nabla\psi_2|}{R^d}\,\mathrm{d}x.
\end{aligned}
$$

Therefore, using the Gehring lemma, we see that there exists $\frac{d}{d-2} > p_0 > d'$ depending only on d and $C(p)$ such that for all $r\in[d',p_0]$ we have

$$
\int_{\tilde{\mathcal{O}}} |\dot{\boldsymbol{u}}|^r\,\mathrm{d}x \le C(\tilde{\mathcal{O}})\left(\int_{\mathcal{O}} |\dot{\boldsymbol{u}}|^{d'}\,\mathrm{d}x\right)^{\frac{r}{d'}} + C(q,\hat{\mathcal{O}})\int_{\hat{\mathcal{O}}} |\nabla\psi_1|^r+|\nabla\psi_2|^r\,\mathrm{d}x \tag{3.21}
$$

for all open $\tilde{\mathcal{O}}$ and $\hat{\mathcal{O}}$ fulfilling $\tilde{\mathcal{O}}\subset\overline{\tilde{\mathcal{O}}}\subset\hat{\mathcal{O}}\subset\overline{\hat{\mathcal{O}}}\subset\mathcal{O}$. Consequently, using the theory for the Laplace equation we can estimate the last two terms as (recall that $r>d'$)

$$
\begin{aligned}
&\int_{\hat{\mathcal{O}}} |\nabla\psi_1|^r+|\nabla\psi_2|^r\,\mathrm{d}x \le C(\hat{\mathcal{O}})\left(\int_{\mathcal{O}} |\triangle\psi_1|^{\frac{dr}{d+r}}+|\triangle\psi_2|^{\frac{dr}{d+r}}\,\mathrm{d}x\right)^{\frac{d+r}{d}}\\
&\le C(\hat{\mathcal{O}},\boldsymbol{f})\|\operatorname{div}\dot{\boldsymbol{u}}\|^{rq'}_{\frac{q'dr}{d+r}}+\|\boldsymbol{\sigma}\|^{rq}_{\frac{qdr}{d+r}}+\|\dot{\boldsymbol{\sigma}}\|^{r}_{\frac{dr}{d+r}}+\|\dot{\boldsymbol{u}}\|^{r}_{\frac{dr}{d+r}}\\
&\quad + C(\hat{\mathcal{O}})\delta^{r(1-q)}\,\|\sqrt{\mu}\dot{\boldsymbol{u}}\|^{qr}_{\frac{qdr}{d+r}}+C(\hat{O})\delta^r\left\|\frac{\sqrt{\mu^{-1}}|\nabla\boldsymbol{\sigma}_D|\sqrt{(|\boldsymbol{\sigma}_D|-1)_+}}{\sqrt{|\boldsymbol{\sigma}_D|}}\right\|^{q'r}_{\frac{q'dr}{d+r}}.
\end{aligned}
$$

To bound the term on the right hand side, for given r we fix q such that

$$
q' = \frac{2(d+r)}{dr} \Longleftrightarrow q = \frac{2(d+r)}{2d+2r-dr}
$$

and recalling that $r\in(d',d/(d-2))$, the above estimate leads to

$$
\begin{aligned}
&\int_{\hat{\mathcal{O}}} |\nabla\psi_1|^r+|\nabla\psi_2|^r\,\mathrm{d}x \le C(\hat{\mathcal{O}})\left(\int_{\mathcal{O}} |\triangle\psi_1|^{\frac{dr}{d+r}}+|\triangle\psi_2|^{\frac{dr}{d+r}}\,\mathrm{d}x\right)^{\frac{d+r}{d}}\\
&\le C(\hat{\mathcal{O}},\boldsymbol{f})\|\operatorname{div}\dot{\boldsymbol{u}}\|_2^{\frac{2(d+r)}{d}}+\|\boldsymbol{\sigma}\|^{\frac{2r(d+r)}{2d+2r-dr}}_{W^{1,2}(\hat{\mathcal{O}})}+\|\dot{\boldsymbol{\sigma}}\|^r_{d'}+\|\dot{\boldsymbol{u}}\|^r_{d'}\\
&+C(\hat{\mathcal{O}})\delta^{r(1-q)}\,\|\sqrt{\mu}\dot{\boldsymbol{u}}\|^{\frac{2r(d+r)}{2d+2r-dr}}_{\frac{2dr}{2d+2r-dr}}+C(\hat{O})\delta^r\left(\int_{\mathcal{O}}\frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D|^2(|\boldsymbol{\sigma}_D|-1)_+}{|\boldsymbol{\sigma}_D|}\,\mathrm{d}x\right)^{\frac{d+r}{d}}
\end{aligned}
$$

Finally, using this estimate in (3.21) and combining it with the a priori bounds (3.12), (3.13) and (3.17), we see that for almost al time $t \in (0,T)$ there holds

$$\begin{aligned}\|\dot{\boldsymbol{u}}(t)\|_{L^r(\tilde{\mathcal{O}})}^{\frac{dr}{d+r}} &\le C(\tilde{\mathcal{O}}) + C(\hat{\mathcal{O}})\delta^{-\frac{dr}{d+r}\frac{dr}{2d+2r-dr}} \|\sqrt{\mu}\dot{\boldsymbol{u}}(t)\|_{\frac{2dr}{2d+2r-dr}}^{\frac{2dr}{2d+2r-dr}} \\ &\quad + C(\hat{\mathcal{O}})\delta^{\frac{dr}{d+r}} \int_{\mathcal{O}} \frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D(t)|^2(|\boldsymbol{\sigma}_D(t)|-1)_+}{|\boldsymbol{\sigma}_D(t)|}\,\mathrm{d}x\,.\end{aligned} \tag{3.22}$$

Further, since $r \in (d', d/(d-2))$ we have that

$$1 \le \frac{2dr}{2d+2r-dr} \le \frac{2d}{d-2}$$

and we can use the interpolation inequality

$$\|\sqrt{\mu}\dot{\boldsymbol{u}}(t)\|_{\frac{2dr}{2d+2r-dr}} \le \|\sqrt{\mu}\dot{\boldsymbol{u}}(t)\|_1^{\alpha}\left(\|\sqrt{\mu}\dot{\boldsymbol{u}}(t)\|_2 + \|\sqrt{\mu}\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}(t))\|_2\right)^{1-\alpha}$$

with $\alpha \in (0,1)$ given as

$$\frac{2d+2r-dr}{2dr} = \alpha + \frac{(1-\alpha)(d-2)}{2d} \qquad \Longrightarrow \qquad \frac{2d+4r-2dr}{r(d+2)} = \alpha. \tag{3.23}$$

Thus, going back to (3.22) and using (3.13), we see that

$$\begin{aligned}\|\dot{\boldsymbol{u}}(t)\|_{L^r(\tilde{\mathcal{O}})}^{\frac{dr}{d+r}} &\le C(\tilde{\mathcal{O}}) + C(\hat{\mathcal{O}})\delta^{-\frac{dr}{d+r}\frac{dr}{2d+2r-dr}} \mu^{\frac{dr\alpha}{2d+2r-dr}} \\ &\quad + C(\hat{\mathcal{O}})\delta^{\frac{dr}{d+r}} \int_{\mathcal{O}} \frac{\mu^{-1}|\nabla\boldsymbol{\sigma}_D(t)|^2(|\boldsymbol{\sigma}_D(t)|-1)_+}{|\boldsymbol{\sigma}_D(t)|}\,\mathrm{d}x.\end{aligned} \tag{3.24}$$

Thus, fixing now δ in such a way that

$$\delta^{-\frac{dr}{d+r}\frac{dr}{2d+2r-dr}} \mu^{\frac{dr\alpha}{2d+2r-dr}} = 1 \implies \delta^{\frac{dr}{d+r}} = \mu^{\alpha}$$

and integrating (3.24) over any time interval I and using (3.17) we deduce that for all $r \in (d', \min(p_0, d/(d-2))$ there exists $\alpha(r) > 0$ such that

$$\int_I \|\dot{\boldsymbol{u}}(t)\|_{L^r(\tilde{\mathcal{O}})}^{\frac{dr}{d+r}}\,\mathrm{d}t \le C(p_0, \tilde{\mathcal{O}})(|I| + \mu^{\alpha(r)}), \tag{3.25}$$

which leads to the uniform improvement of the spatial integrability of the velocity field stated in (2.12). Moreover, it is evident that once letting $\mu \to 0$ we can even deduce $L^{\infty}(0,T; L^r_{loc}(\mathcal{O}))$ bound for the limiting velocity field $\dot{\boldsymbol{u}}$, which is however not valid up to the boundary $\partial\mathcal{O}$, which will be improved later.

3.4. **Time regularity.** This subsection is devoted to improvement of the time regularity for $\dot{\boldsymbol{\sigma}}^{\mu}$, which will be uniform with respect to the approximative parameter μ. We again omit writing superscript μ in this subsection.

For arbitrary w, we denote its times shift as $\Delta_t^{\tau} w(t,x) := w(t+\tau,x) - w(t,x)$ and with the help of this notation, we take the scalar product of (3.1) with $-\Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)$ and integrate the result over $\mathcal{O}$ to get

$$\begin{aligned}-\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s) + \mu^{-1}(|\boldsymbol{\sigma}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)\,\mathrm{d}x \\ = \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s) - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s)\,\mathrm{d}x.\end{aligned}$$

The first term on the right hand side vanishes similarly as before and after moving the corresponding terms onto the right hand side and applying the Hölder inequality and (2.3), we deduce that

$$
\begin{aligned}
&-\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^\tau \dot{\boldsymbol{\sigma}} + \mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} \cdot \Delta_t^\tau \dot{\boldsymbol{\sigma}} \,\mathrm{d}x \\
&\le \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}^s} - \dot{\boldsymbol{\sigma}}) \,\mathrm{d}x \\
&\quad + C_2 \|\dot{\boldsymbol{\sigma}}\|_2 \|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_2 + \mu^{-1} \||(\boldsymbol{\sigma}_D|-1)_+\|_1 \|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_\infty.
\end{aligned} \tag{3.26}
$$

Using the algebraic identity (note that A is symmetric)

$$
-\mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^\tau \dot{\boldsymbol{\sigma}} = \frac{1}{2} \mathsf{A} \Delta_t^\tau \dot{\boldsymbol{\sigma}} \cdot \Delta_t^\tau \dot{\boldsymbol{\sigma}} - \frac{1}{2} \Delta_t^s \left(\mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} \right),
$$

we further observe with the help of (2.3) and the a piori estimate (3.12) that (we also use the fact that $\mathcal{O}$ is bounded)

$$
\begin{aligned}
&\int_{\mathcal{O}} C_1 |\Delta_t^\tau \dot{\boldsymbol{\sigma}}|^2 - 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} \cdot \Delta_t^\tau \dot{\boldsymbol{\sigma}} \,\mathrm{d}x \\
&\le \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}^s} - \dot{\boldsymbol{\sigma}}) + \Delta_t^\tau \left(\mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} \right) \,\mathrm{d}x + C \|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_\infty.
\end{aligned} \tag{3.27}
$$

Finally, we integrate the resulting inequality with respect to τ over the interval $(0,h)$ and with respect to t over the interval (t_1,t_2) with $0 \le t_1 < t_2 \le T-h$ to get

$$
\begin{aligned}
&C_1 \int_{t_1}^{t_2} \int_0^h \int_{\mathcal{O}} C_1 |\Delta_t^\tau \dot{\boldsymbol{\sigma}}|^2 \,\mathrm{d}\tau \,\mathrm{d}t \\
&- \int_{t_1}^{t_2} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+ \frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} \cdot \left(\int_0^h \Delta_t^\tau \dot{\boldsymbol{\sigma}} \,\mathrm{d}\tau \right) \mathrm{d}x \,\mathrm{d}t \\
&\le \int_{t_1}^{t_2} \int_0^h \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}^s} - \dot{\boldsymbol{\sigma}}) + \Delta_t^\tau \left(\mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} \right) \,\mathrm{d}x + C \|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_\infty \,\mathrm{d}\tau \,\mathrm{d}t.
\end{aligned} \tag{3.28}
$$

Next, we focus on the estimate on the right hand side. First, we have

$$
\begin{aligned}
\int_{t_1}^{t_2} \int_0^h \|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_\infty \,\mathrm{d}\tau \,\mathrm{d}t &= \int_0^h \tau \left(\int_{t_1}^{t_2} \frac{\|\Delta_t^\tau \dot{\boldsymbol{\sigma}^s}\|_\infty}{\tau} \,\mathrm{d}t \right) \mathrm{d}\tau \\
&\le \int_0^h \tau \int_0^T \|\ddot{\boldsymbol{\sigma}^s}\|_\infty \,\mathrm{d}t \,\mathrm{d}\tau \le ch^2,
\end{aligned} \tag{3.29}
$$

where for the last estimate we used the assumption on $\boldsymbol{\sigma}^s$. Second, using the substitution theorem and the Hölder inequality, we obatin

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_0^h\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot\Delta_t^\tau(\dot{\boldsymbol{\sigma}^s}-\dot{\boldsymbol{\sigma}})\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&=\int_0^h\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\dot{\boldsymbol{\sigma}^s}(t+\tau)-\dot{\boldsymbol{\sigma}}(t+\tau))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\quad-\int_0^h\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&=\int_0^h\int_{t_1+\tau}^{t_2+\tau}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau))\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\quad-\int_0^h\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&=\int_0^h\int_{t_1+\tau}^{t_2}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau)-\dot{\boldsymbol{u}}_0(t))\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\quad+\int_0^h\int_{t_2}^{t_2+\tau}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau))\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\quad-\int_0^h\int_{t_1}^{t_1+\tau}\int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))\cdot(\dot{\boldsymbol{\sigma}}^s(t)-\dot{\boldsymbol{\sigma}}(t))\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\leq C\int_0^h \tau\|\dot{\boldsymbol{\sigma}}^s-\dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)}\int_0^T\|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0)\|_2\,\mathrm{d}t\,\mathrm{d}\tau\\
&\quad+C\int_0^h \tau\tau\|\dot{\boldsymbol{\sigma}}^s-\dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)}\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_{L^\infty(0,T;L^2)}\,\mathrm{d}\tau\\
&\leq Ch^2,
\end{aligned}
\tag{3.30}
$$

where we used the assumptions on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}^s$ and the a priori estimate (3.12). Finally, the remaining term on the right hand side of (3.28) is estimated as follows

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_0^h\int_{\mathcal{O}}\Delta_t^\tau\left(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\right)\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&=\quad\int_0^h\int_{t_1}^{t_2}\int_{\mathcal{O}}\left(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\right)(t+\tau)-\left(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\right)(t)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&=\quad\int_0^h\int_{t_2}^{t_2+\tau}\int_{\mathcal{O}}\left(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\right)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau-\int_0^h\int_{t_1}^{t_1+\tau}\int_{\mathcal{O}}\left(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\right)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\tau\\
&\leq Ch^2\|\dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)}\leq Ch^2,
\end{aligned}
\tag{3.31}
$$

where the estimate (3.12) is used again. Substituting (3.29)–(3.30) into (3.28), we conclude

$$
\begin{aligned}
&C_1\int_{t_1}^{t_2}\int_0^h\int_{\mathcal{O}} C_1|\Delta_t^\tau\dot{\boldsymbol{\sigma}}|^2\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad-\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot\left(\int_0^h\Delta_t^\tau\dot{\boldsymbol{\sigma}}\,\mathrm{d}\tau\right)\,\mathrm{d}x\,\mathrm{d}t\leq Ch^2.
\end{aligned}
\tag{3.32}
$$

Thus, it remains to evaluate the second term on the left hand side. Using the convexity of $(|\boldsymbol{\sigma}_D|-1)_+^2$, we continue as follows

$$
\begin{aligned}
&\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot\left(\int_0^h \Delta_t^\tau \dot{\boldsymbol{\sigma}}\,\mathrm{d}\tau\right)\mathrm{d}x\,\mathrm{d}t\\
&\quad=\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot(\boldsymbol{\sigma}_D(t+h)-\boldsymbol{\sigma}_D(t)-h\dot{\boldsymbol{\sigma}}_D(t))\,\mathrm{d}x\,\mathrm{d}t\\
&\quad=\int_{t_1}^{t_2}\int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|}\cdot(\boldsymbol{\sigma}_D(t+h)-\boldsymbol{\sigma}_D(t)-h\dot{\boldsymbol{\sigma}}_D(t))\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\le\int_{t_1}^{t_2}\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t+h)|-1)_+^2-(|\boldsymbol{\sigma}_D(t)|-1)_+^2-h\mu^{-1}\frac{d}{dt}(|\boldsymbol{\sigma}_D(t)|-1)_+^2\,\mathrm{d}t\\
&\quad\le\int_0^h\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t_2+t)|-1)_+^2-(|\boldsymbol{\sigma}_D(t_2)|-1)_+^2\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\quad-\int_0^h\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t_1+t)|-1)_+^2-(|\boldsymbol{\sigma}_D(t_1)|-1)_+^2\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

Consequently, the relation (3.32) reduces to

$$
\begin{aligned}
&C_1\int_{t_1}^{t_2}\int_0^h \|\Delta_t^\tau\dot{\boldsymbol{\sigma}}\|_2^2\,\mathrm{d}\tau\,\mathrm{d}t\le Ch^2\\
&\quad+\int_0^h\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t_2+t)|-1)_+^2-(|\boldsymbol{\sigma}_D(t_2)|-1)_+^2\,\mathrm{d}x\,\mathrm{d}t\\
&\quad-\int_0^h\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t_1+t)|-1)_+^2-(|\boldsymbol{\sigma}_D(t_1)|-1)_+^2\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
\tag{3.33}
$$

Going now back to (3.4), we see that

$$
\begin{aligned}
\frac{d}{dt}\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+^2\,\mathrm{d}x&=-\int_{\mathcal{O}}\mathsf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\,\mathrm{d}x\\
&\quad+\int_{\mathcal{O}}(\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\dot{\boldsymbol{\sigma}}-\dot{\boldsymbol{\sigma}}^s)+\mathsf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}^s+\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\frac{\boldsymbol{\sigma}_D\cdot\dot{\boldsymbol{\sigma}}_D^s}{|\boldsymbol{\sigma}_D|}\,\mathrm{d}x
\end{aligned}
$$

and integrating this inequality over $(\tau,\tau+t)$ we get

$$
\begin{aligned}
&\left|\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\sigma}_D(t+\tau)|-1)_+^2-\mu^{-1}(|\boldsymbol{\sigma}_D(\tau)|-1)_+^2\,\mathrm{d}x\right|\le C_2\int_\tau^{t+\tau}\|\dot{\boldsymbol{\sigma}}\|_2^2\\
&\quad+C\int_\tau^{t+\tau}\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2\|\dot{\boldsymbol{\sigma}}-\dot{\boldsymbol{\sigma}}^s\|_2+\|\dot{\boldsymbol{\sigma}}\|_2\|\dot{\boldsymbol{\sigma}}^s\|_2+\|\mu^{-1}(|\boldsymbol{\sigma}_D|-1)_+\|_1\|\dot{\boldsymbol{\sigma}}_D^s\|_\infty\\
&\quad\le Ct,
\end{aligned}
\tag{3.34}
$$

where we used the assumptions on data and (3.12). Consequently, using this inequality in (3.33) we obtain the final inequality

$$
\frac{1}{h^2}\int_{t_1}^{t_2}\int_0^h\|\Delta_t^\tau\dot{\boldsymbol{\sigma}}\|_2^2\,\mathrm{d}\tau\,\mathrm{d}t\le C. \tag{3.35}
$$

Recalling the equivalence of norms in the Nikolskii spaces, see [8], we have that (3.35) implies that

$$h^{-1}\int_0^{T-h} \|\Delta_t^h \dot{\boldsymbol{\sigma}}\|_2^2 \,\mathrm{d}\,\mathrm{d}t \le C, \tag{3.36}$$

which is the last remaining part of the uniform estimate (2.12).

3.5. **Limit** $\mu \to 0_+$**.** In this subsection we let $\mu \to 0_+$ and show the the limit object solves the original problem, i.e., the the Prandtl–Reuss model. This part is rather standard, compare e.g. with [9, 17] but we shall need several identities stated in this part also in further subsections in order to identify the plastic strain and also for the improved integrability result for the velocity up to the boundary $\partial\mathcal{O}$.

Recalling the a priori estimates (3.12)–(3.13), we can find subsequences labeled by μ such that if letting $\mu \to 0_+$ we have

$$\boldsymbol{\sigma}^\mu \rightharpoonup^* \boldsymbol{\sigma} \qquad \text{weakly}^* \text{ in } W^{1,\infty}(0,T; L^2(\mathcal{O}, \mathbb{R}^{d\times d}_{sym})), \tag{3.37}$$

$$\mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ \frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|} \rightharpoonup^* \dot{\boldsymbol{e}}_{\mathrm{p}} \qquad \text{weakly}^* \text{ in } L^\infty(0,T; \mathcal{M}(\overline{\mathcal{O}}; \mathbb{R}^{d\times d}_{sym})), \tag{3.38}$$

$$\mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ \rightharpoonup^* \lambda \qquad \text{weakly}^* \text{ in } L^\infty(0,T; \mathcal{M}(\overline{\mathcal{O}})), \tag{3.39}$$

$$\boldsymbol{\varepsilon}(\boldsymbol{u}^\mu) \rightharpoonup^* \boldsymbol{\varepsilon}(\boldsymbol{u}) \qquad \text{weakly}^* \text{ in } W^{1,\infty}(0,T; \mathcal{M}(\overline{\mathcal{O}}; \mathbb{R}^{d\times d}_{sym})), \tag{3.40}$$

$$\operatorname{div} \boldsymbol{u}^\mu \rightharpoonup^* \operatorname{div} \boldsymbol{u} \qquad \text{weakly}^* \text{ in } W^{1,\infty}(0,T; L^2(\mathcal{O})), \tag{3.41}$$

$$\boldsymbol{u}^\mu \rightharpoonup^* \boldsymbol{u} \qquad \text{weakly}^* \text{ in } W^{1,\infty}(0,T; L^{d'}(\mathcal{O}; \mathbb{R}^d)). \tag{3.42}$$

Moreover, recalling (3.17) and (3.25) we also get that for some $\varepsilon > 0$

$$\boldsymbol{\sigma}^\mu \rightharpoonup^* \boldsymbol{\sigma} \qquad \text{weakly}^* \text{ in } L^\infty(0,T; W^{1,2}_{loc}(\mathcal{O}; \mathbb{R}^{d\times d}_{sym})), \tag{3.43}$$

$$\operatorname{div} \boldsymbol{u}^\mu \rightharpoonup^* \operatorname{div} \boldsymbol{u} \qquad \text{weakly}^* \text{ in } L^\infty(0,T; W^{1,2}_{loc}(\mathcal{O})), \tag{3.44}$$

$$\boldsymbol{u}^\mu \rightharpoonup \boldsymbol{u} \qquad \text{weakly in } L^1(0,T; L^{d'+\varepsilon}_{loc}(\mathcal{O}; \mathbb{R}^d)), \tag{3.45}$$

$$\boldsymbol{\sigma}^\mu \rightharpoonup^* \boldsymbol{\sigma} \qquad \text{weakly in } N^{\frac{3}{2},2}(0,T; L^2(\mathcal{O}; \mathbb{R}^{d\times d}_{sym})). \tag{3.46}$$

In addition, using the weak lower semicontinuity, we also have that $|\dot{\boldsymbol{e}}_{\mathrm{p}}|$ is absolutely continuous with respect to λ and

$$|\dot{\boldsymbol{e}}_{\mathrm{p}}(t)| \le \lambda(t) \tag{3.47}$$

in the sense of measures for almost all $t \in (0,T)$. Having these convergence results, it is not difficult to let $\mu \to 0_+$ in (3.1) and show that (2.14) holds (in fact it holds for almost all $t \in (0,T)$ in the sense of measures).

Moreover, it follows from (3.39) and from the fact that $\boldsymbol{\sigma}^\mu(t) \in \mathcal{F}_{el}(t)$ that for almost all $t \in (0,T)$ we have $\boldsymbol{\sigma}(t) \in \mathcal{F}(t)$. Next, multiplying (3.1) by $\boldsymbol{\sigma}^\mu - \boldsymbol{\sigma}$ and integrating the result over $\mathcal{O} \times (0,t)$ we can deduce with the help of the fact that

$|\boldsymbol{\sigma}_D| \le 1$ almost everywhere that

$$
\begin{aligned}
&\frac{C_1}{2}\|\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t)\|_2^2 + \int_0^t \|\mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D|-1)_+(|\boldsymbol{\sigma}^{\mu}_D(\tau)|-|\boldsymbol{\sigma}_D(\tau)|)\|_1\\
&\quad\le \frac{1}{2}\int_{\mathcal{O}} \mathsf{A}(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))\cdot(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))\,\mathrm{d}x\\
&\qquad+\int_0^t\int_{\mathcal{O}} \mu^{-1}\frac{(|\boldsymbol{\sigma}^{\mu}_D(\tau)|-1)_+}{|\boldsymbol{\sigma}^{\mu}_D(\tau)|}\boldsymbol{\sigma}^{\mu}_D(\tau)\cdot(\boldsymbol{\sigma}^{\mu}_D(\tau)-\boldsymbol{\sigma}_D(\tau))\,\mathrm{d}x\,\mathrm{d}\tau\\
&\quad=\int_0^t\int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol{\sigma}}^{\mu}(\tau)-\dot{\boldsymbol{\sigma}}(\tau))\cdot(\boldsymbol{\sigma}^{\mu}(\tau)-\boldsymbol{\sigma}(\tau))\,\mathrm{d}x\,\mathrm{d}\tau\\
&\qquad+\int_0^t\int_{\mathcal{O}} \mu^{-1}\frac{(|\boldsymbol{\sigma}^{\mu}_D(\tau)|-1)_+}{|\boldsymbol{\sigma}^{\mu}_D(\tau)|}\boldsymbol{\sigma}^{\mu}_D(\tau)\cdot(\boldsymbol{\sigma}^{\mu}_D(\tau)-\boldsymbol{\sigma}_D(\tau))\,\mathrm{d}x\,\mathrm{d}\tau\\
&\quad=-\int_0^t\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}}(\tau)\cdot(\boldsymbol{\sigma}^{\mu}(\tau)-\boldsymbol{\sigma}(\tau))-\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^{\mu}(\tau))\cdot(\boldsymbol{\sigma}^{\mu}(\tau)-\boldsymbol{\sigma}(\tau))\,\mathrm{d}x\,\mathrm{d}\tau\\
&\quad=-\int_0^t\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}}(\tau)\cdot(\boldsymbol{\sigma}^{\mu}(\tau)-\boldsymbol{\sigma}(\tau))-\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(\tau))\cdot(\boldsymbol{\sigma}^{\mu}(\tau)-\boldsymbol{\sigma}(\tau))\,\mathrm{d}x\,\mathrm{d}\tau,
\end{aligned}
$$

where for the last inequality we used the fact that $\dot{\boldsymbol{u}}^{\mu}-\dot{\boldsymbol{u}}_0 \in \mathcal{V}$. Consequently, using (3.37), we see that the right hand side vanishes as $\mu \to 0_+$ and consequently, we also get that

$$
\boldsymbol{\sigma}^{\mu} \to \boldsymbol{\sigma} \qquad \text{strongly in } \mathcal{C}([0,T];L^2(\mathcal{O},\mathbb{R}^{d\times d}_{sym})), \tag{3.48}
$$

$$
\mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D|-1)_+^2 \to 0 \qquad \text{strongly in } L^1(0,T;L^1(\mathcal{O})). \tag{3.49}
$$

Next, repeating the very similar procedure as above, we can also obtain the identity

$$
\begin{aligned}
&\int_{\mathcal{O}} \mu^{-1}\frac{(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+}{|\boldsymbol{\sigma}^{\mu}_D(t)|}\boldsymbol{\sigma}^{\mu}_D(t)\cdot(\boldsymbol{\sigma}^{\mu}_D(t)-\boldsymbol{\sigma}_D(t))\,\mathrm{d}x\\
&\quad=\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}}^{\mu}(t)\cdot(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))-\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))\cdot(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))\,\mathrm{d}x.
\end{aligned}
$$

Consequently, having the uniform convergence (3.48) and also (3.37) we see that for all time $t \in [0,T]$

$$
\begin{aligned}
&\lim_{\mu\to 0_+}\sup_{t\in(0,T)}\|\mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+^2\|_1\\
&\quad\le \lim_{\mu\to 0_+}\sup_{t\in(0,T)}\|\mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+(|\boldsymbol{\sigma}^{\mu}_D(t)|-|\boldsymbol{\sigma}_D(t)|)\|_1\\
&\quad\le \lim_{\mu\to 0_+}\sup_{t\in(0,T)}\int_{\mathcal{O}} \mu^{-1}\frac{(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+}{|\boldsymbol{\sigma}^{\mu}_D(t)|}\boldsymbol{\sigma}^{\mu}_D(t)\cdot(\boldsymbol{\sigma}^{\mu}_D(t)-\boldsymbol{\sigma}_D(t))\,\mathrm{d}x\\
&\quad= \lim_{\mu\to 0_+}\sup_{t\in(0,T)}\int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}}^{\mu}(t)\cdot(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))-\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))\cdot(\boldsymbol{\sigma}^{\mu}(t)-\boldsymbol{\sigma}(t))\,\mathrm{d}x=0\,.
\end{aligned}
\tag{3.50}
$$

In addition, for arbitrary $t \in (0,T)$ we denote

$$
K_{\varepsilon}(t) := \{x \in \mathcal{O};\ M|\boldsymbol{\sigma}_D(t)| \le 1-\varepsilon\},
$$

where M denotes the non-centred maximal function. Due to the properties of the maximal function, we see that this set is closed (hence $\lambda(t)$ measurable) and from

(3.39) and (3.50) we have

$$\begin{aligned}\lambda(t)(K_\varepsilon(t)) &= \lim_{\mu\to 0_+}\int_{K_\varepsilon(t)}\mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\,\mathrm{d}x\\ &\le \varepsilon^{-1}\lim_{\mu\to 0_+}\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+(|\boldsymbol{\sigma}^\mu_D(t)|-|\boldsymbol{\sigma}_D(t)|)\,\mathrm{d}x = 0.\end{aligned} \tag{3.51}$$

Consequently, we see that $\lambda(t)$ is supported on the set, where $M|\boldsymbol{\sigma}_D(t)|=1$. Finally, using all above convergence results and testing (3.1) by $\boldsymbol{\sigma}^\mu-\boldsymbol{\sigma}^s$, it is standard to show (2.5). This finishes the first part of the convergence results stated in Theorem 2.1.

3.6. **BMO estimates and improved integrability results.** This section contains two key novelties of the paper. Since the BMO property for the Cauchy stress is somehow independent of the structure of equation and is more related to the function spaces properties, we formulate it as a separate lemma. On purpose, we state it for a domain Ω to emphasize its independence of the Prandtl–Reuss model.

Lemma 3.1 (BMO estimates)**.** *Let $\Omega\subset\mathbb{R}^d$ be an open set. Assume that $\boldsymbol{\sigma}\in L^2(\Omega;\mathbb{R}^{d\times d})$ be such that $\operatorname{div}\boldsymbol{\sigma}\in L^d(\Omega;\mathbb{R}^d)$ and $\boldsymbol{\sigma}_D\in L^\infty(\Omega;\mathbb{R}^{d\times d})$. The for any ball $B_R(x_0)\subset\Omega$ there holds*

$$\int_{B_R(x_0)}|\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}}_{B_R(x_0)}|^2\,\mathrm{d}x\le C(d)R^d(\|\boldsymbol{\sigma}_D\|^2_\infty+\|\operatorname{div}\boldsymbol{\sigma}\|^2_d), \tag{3.52}$$

where the constant $C(d)$ depends only on the dimension d and $\overline{\boldsymbol{\sigma}}_{B_R(x_0)}$ denotes the mean value of $\boldsymbol{\sigma}$ over $B_R(x_0)$. In addition if Ω is Lipschitz, then there exists $R_0>0$ such that for all $R\in(0,R_0)$ and all $x_0\in\overline{\Omega}$ there holds

$$\int_{\Omega\cap B_R(x_0)}|\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}}_{\Omega\cap B_R(x_0)}|^2\,\mathrm{d}x\le C(d,\Omega)R^d(\|\boldsymbol{\sigma}_D\|^2_\infty+\|\operatorname{div}\boldsymbol{\sigma}\|^2_d). \tag{3.53}$$

Furthermore, $\boldsymbol{\sigma}$ can be extended onto the whole $\mathbb{R}^d$ by some $\tilde{\boldsymbol{\sigma}}$ such that for all $R>0$ and all $x_0\in\mathbb{R}^d$ there holds

$$\int_{B_R(x_0)}|\tilde{\boldsymbol{\sigma}}-\overline{\tilde{\boldsymbol{\sigma}}}_{B_R(x_0)}|^2\,\mathrm{d}x\le C(d,\Omega)R^d(\|\boldsymbol{\sigma}_D\|^2_{L^\infty(\Omega)}+\|\operatorname{div}\boldsymbol{\sigma}\|^2_{L^d(\Omega)}). \tag{3.54}$$

Moreover, if $\boldsymbol{\sigma}$ is symmetric then the extension $\tilde{\boldsymbol{\sigma}}$ is symmetric as well.

It is also worth mentioning that we can replace the assumption $\boldsymbol{\sigma}_D\in L^\infty$ by $\boldsymbol{\sigma}_D\in BMO$. Since such a generalization is trivial, it is left to the reader.

Proof of Lemma 3.1. We start the proof with (3.52). To simplify the notation, we consider only the ball $B_R(0)=:B_R$. Using the properties of the Bogovskii operator, we know that for any $u\in L^2(B_R)$ fulfilling

$$\int_{B_R}u\,\mathrm{d}x=0,$$

there exists $\mathbf{B}\in W^{1,2}_0(B_R;\mathbb{R}^d)$ satisfying

$$\operatorname{div}\mathbf{B}=u \text{ in } B_R,\qquad \|\nabla\mathbf{B}\|_2\le C(d)\|u\|_2. \tag{3.55}$$

Since $\boldsymbol{\sigma}_D$ is bounded, we see that to prove (3.52), it is enough to show that

$$\int_{B_R}|\operatorname{tr}\boldsymbol{\sigma}-\overline{\operatorname{tr}\boldsymbol{\sigma}}_{B_R}|^2\,\mathrm{d}x\le C(d)R^d(\|\boldsymbol{\sigma}_D\|^2_\infty+\|\operatorname{div}\boldsymbol{\sigma}\|^2_d). \tag{3.56}$$

Hence, setting $u := \operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}_{B_R}$ in (3.55) and using the density of smooth functions (to justify the following formal integration by parts) we get

$$
\begin{aligned}
&\int_{B_R} |\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}_{B_R}|^2 \,\mathrm{d}x = \int_{B_R} (\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}_{B_R}) \operatorname{div}\mathbf{B} \,\mathrm{d}x = -\int_{B_R} \nabla \operatorname{tr}\boldsymbol{\sigma} \cdot \mathbf{B} \,\mathrm{d}x \\
&= -d\int_{B_R} (\operatorname{div}\boldsymbol{\sigma} - \operatorname{div}\boldsymbol{\sigma}_D) \cdot \mathbf{B} \,\mathrm{d}x = -d\int_{B_R} \operatorname{div}\boldsymbol{\sigma} \cdot \mathbf{B} \,\mathrm{d}x - d\int_{B_R} \boldsymbol{\sigma}_D \cdot \nabla\mathbf{B} \,\mathrm{d}x \\
&\le C(d)(\|\operatorname{div}\boldsymbol{\sigma}\|_d \|\mathbf{B}\|_2 R^{\frac{d-2}{2}} + \|\boldsymbol{\sigma}_D\|_\infty \|\nabla\mathbf{B}\|_2 R^{\frac{d}{2}}) \\
&\le C(d) R^{\frac{d}{2}} (\|\operatorname{div}\boldsymbol{\sigma}\|_d + \|\boldsymbol{\sigma}_D\|_\infty) \|\nabla\mathbf{B}\|_2 \\
&\le C(d) R^{\frac{d}{2}} (\|\operatorname{div}\boldsymbol{\sigma}\|_d + \|\boldsymbol{\sigma}_D\|_\infty) \left(\int_{B_R} |\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}_{B_R}|^2 \,\mathrm{d}x\right)^{\frac{1}{2}},
\end{aligned}
$$

where we used the Poincaré inequality and (3.55). Consequently, from the above inequality the estimate (3.56) and consequently also (3.52) directly follow.

Next, in order to prove (3.53), we use the fact that Ω is Lipschitz. Therefore we can find $R_0 > 0$ such that for all $x_0 \in \partial\Omega$ and all $R \in (0, R_0)$ we have

$$
R^d \le C(\Omega, d) |\Omega \cap B_R(x_0)|. \tag{3.57}
$$

In addition, it also follows from the properties of the Bogovskii operator that for any $u \in L^2(\Omega \cap B_R(x_0))$ fulfilling $\int_{\Omega \cap B_R(x_0)} u = 0$ there exists $\mathbf{B} \in W_0^{1,2}(\Omega \cap B_R(x_0); \mathbb{R}^d)$ such that

$$
\operatorname{div}\mathbf{B} = u \text{ in } \Omega \cap B_R, \qquad \|\nabla\mathbf{B}\|_2 \le C(d, \Omega)\|u\|_2. \tag{3.58}
$$

Consequently, using (3.57) and (3.58), we can use exactly the same procedure as above and we deduce that for all $x_0 \in \partial\Omega$ and all $R \in (0, R)$ we have

$$
\int_{\Omega \cap B_R} |\operatorname{tr}\boldsymbol{\sigma} - \overline{\operatorname{tr}\boldsymbol{\sigma}}_{\Omega \cap B_R}|^2 \,\mathrm{d}x \le C(d, \Omega) R^d (\|\boldsymbol{\sigma}_D\|_\infty^2 + \|\operatorname{div}\boldsymbol{\sigma}\|_d^2) \tag{3.59}
$$

and additionally we also see that (3.53) holds for all $x_0 \in \partial\Omega$. Finally, let us consider arbitrary $x_0 \in \Omega$ and $R \in (0, R_0)$. If $B_R(x_0) \subset \Omega$, we already have the estimate (3.52). Hence, let us assume that $B_R(x_0) \cap \partial\Omega \neq \emptyset$. Then there exist $\tilde{x}_0 \in \partial\Omega$ such that $B_R(x_0) \subset B_{2R}(\tilde{x}_0)$. Thus using the triangle inequality, we deduce that

$$
\begin{aligned}
&\int_{B_R(x_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_R(x_0)\cap\Omega}|^2 \,\mathrm{d}x \\
&\le 2\int_{B_R(x_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)\cap\Omega}|^2 \,\mathrm{d}x + 2|B_R(x_0)\cap\Omega| |\overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)} - \overline{\boldsymbol{\sigma}}_{B_R(x_0)\cap\Omega}|^2 \\
&\le 2\int_{B_{2R}(\tilde{x}_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)\cap\Omega}|^2 \,\mathrm{d}x \\
&\quad + 2|B_R(x_0)\cap\Omega| \left| \frac{1}{|B_R(x_0)\cap\Omega|} \int_{B_R(x_0)\cap\Omega} (\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)}) \,\mathrm{d}x \right|^2 \\
&\le 2\int_{B_{2R}(\tilde{x}_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)\cap\Omega}|^2 \,\mathrm{d}x + 2\int_{B_R(x_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)})|^2 \,\mathrm{d}x \\
&\le 4\int_{B_{2R}(\tilde{x}_0)\cap\Omega} |\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_{B_{2R}(\tilde{x}_0)\cap\Omega}|^2 \,\mathrm{d}x \le C(d, \Omega) R^d (\|\boldsymbol{\sigma}_D\|_{L^\infty(\Omega)}^2 + \|\operatorname{div}\boldsymbol{\sigma}\|_{L^d(\Omega)}^2),
\end{aligned}
$$

where for the last inequality we used (3.53) with $\tilde{x}_0 \in \partial\Omega$. Consequently, we see that (3.53) holds for all $x_0 \in \overline{\Omega}$. □

Since $|\boldsymbol{\sigma}_D| \le 1$ and $\operatorname{div}\boldsymbol{\sigma} = \operatorname{div}\boldsymbol{\sigma}^s$, we can use the assumption on data and Lemma 3.1 to get the first part of (2.15), i.e., the fact that $\boldsymbol{\sigma} \in L^\infty(0,T;BMO(\mathcal{O}))$. Next, we will show also the second part of (2.15), i.e., the uniform improvement of the integrability of $\dot{\boldsymbol{u}}$ and consequently also on some deeper characterization of a possibly singular part of the measure λ. First, we can recall the estimate (3.25) and by using the weak lower semicontinuity we get that for all $\tilde{\mathcal{O}} \subset\subset \mathcal{O}$ and any interval $I \subset (0,T)$ we have

$$\int_I \int_{\tilde{\mathcal{O}}} |\dot{\boldsymbol{u}}|^{d'+\varepsilon} \,\mathrm{d}x\,\mathrm{d}t \le C(\tilde{\mathcal{O}})|I|.$$

Consequently we see that $\dot{\boldsymbol{u}} \in L^\infty(0,T;L^{d'+\varepsilon}_{loc}(\Omega;\mathbb{R}^d))$. Our goal is to get this information up to the boundary $\partial\mathcal{O}$ in case $\mathcal{O}$ is Lipschitz. We shall proceed almost exactly in the same manner as before, however, we focus here only on the limiting behaviour of $\boldsymbol{u}$. The starting point is to estimate the weak* limit of $\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)$ up to $\partial\mathcal{O}$. To do that, let us consider arbitrary smooth nonnegative $\tau \in \mathcal{C}^\infty(\overline{\mathcal{O}} \times [0,T])$. It follows from (3.1) that

$$\begin{aligned}
|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)| &\le C_2|\dot{\boldsymbol{\sigma}}^\mu| + \mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ \le C_2|\dot{\boldsymbol{\sigma}}^\mu| + \mu^{-1}(|\boldsymbol{\sigma}^\mu_D| - 1)_+ \frac{\boldsymbol{\sigma}^\mu_D}{|\boldsymbol{\sigma}^\mu_D|} \cdot \boldsymbol{\sigma}^\mu \\
&\le C_2|\dot{\boldsymbol{\sigma}}^\mu| + C_2|\dot{\boldsymbol{\sigma}}^\mu||\boldsymbol{\sigma}^\mu| + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma}^\mu \\
&= C_2|\dot{\boldsymbol{\sigma}}^\mu| + C_2|\dot{\boldsymbol{\sigma}}^\mu||\boldsymbol{\sigma}^\mu| + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma} + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot (\boldsymbol{\sigma}^\mu - \boldsymbol{\sigma}) \\
&= C_2|\dot{\boldsymbol{\sigma}}^\mu|(1 + |\boldsymbol{\sigma}^\mu|) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma} + (\mathbf{A}\dot{\boldsymbol{\sigma}}^\mu + \mu^{-1}\frac{(|\boldsymbol{\sigma}^\mu_D| - 1)_+}{|\boldsymbol{\sigma}^\mu_D|}\boldsymbol{\sigma}^\mu_D) \cdot (\boldsymbol{\sigma}^\mu - \boldsymbol{\sigma}).
\end{aligned}$$

Thus, multiplying the resulting inequality by τ (which is smooth), integrating over $(0,T) \times \mathcal{O}$, using (3.46), (3.48) and (3.50) and nonnegativity of τ we get that

$$\begin{aligned}
&\lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)|\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\le \lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma}\tau \,\mathrm{d}x\,\mathrm{d}t + C_2 \int_0^T \int_{\mathcal{O}} |\dot{\boldsymbol{\sigma}}|(1 + |\boldsymbol{\sigma}|)\tau \,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{3.60}$$

Next, using the safety load condition (2.9), we also get

$$\begin{aligned}
&\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma}^s \\
&= \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu - \dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu) \cdot \boldsymbol{\sigma}^s_D + \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}^s \operatorname{div}\dot{\boldsymbol{u}}^\mu \\
&\le \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu - \dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) + (1 - \delta)|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)| + \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}^s \operatorname{div}\dot{\boldsymbol{u}}^\mu
\end{aligned}$$

Hence, inserting this into (3.60) and using (3.41), and using integration by parts, we finally obtain

$$
\begin{aligned}
&\delta \lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)|\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\le \int_0^T \int_{\mathcal{O}} C_2|\dot{\boldsymbol{\sigma}}|(1+|\boldsymbol{\sigma}|)\tau + \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}^s \operatorname{div}\dot{\boldsymbol{u}}\tau + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s)\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu - \dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s)\tau \,\mathrm{d}x\,\mathrm{d}t \\
&= \int_0^T \int_{\mathcal{O}} C_2|\dot{\boldsymbol{\sigma}}|(1+|\boldsymbol{\sigma}|)\tau + \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}^s \operatorname{div}\dot{\boldsymbol{u}}\tau + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s)\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}^\mu - \dot{\boldsymbol{u}}_0)\otimes\nabla\tau)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s) \,\mathrm{d}x\,\mathrm{d}t \\
&= \int_0^T \int_{\mathcal{O}} C_2|\dot{\boldsymbol{\sigma}}|(1+|\boldsymbol{\sigma}|)\tau + \frac{1}{d}\operatorname{tr}\boldsymbol{\sigma}^s \operatorname{div}\dot{\boldsymbol{u}}\tau + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s)\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_0^T \int_{\mathcal{O}} ((\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\otimes\nabla\tau)\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^s) \,\mathrm{d}x\,\mathrm{d}t,
\end{aligned} \tag{3.61}
$$

where for the last equality we used the fact that $\boldsymbol{\sigma} \in L^\infty(0,T;L^p(\Omega;\mathbb{R}^{d\times d}))$ for any $p<\infty$ (which follows from the BMO property) and the weak convergence result (3.42). Moreover, using the assumptions on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}^s$, and the fact that $|\boldsymbol{\sigma}_D| \le 1$, we observe

$$
\begin{aligned}
&\delta \lim_{\mu\to 0_+} \int_0^T \int_{\mathcal{O}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu)|\tau \,\mathrm{d}x\,\mathrm{d}t \\
&\le C \int_0^T \int_{\mathcal{O}} (|\dot{\boldsymbol{\sigma}}| + |\operatorname{tr}\boldsymbol{\sigma}|)(1+|\operatorname{tr}\boldsymbol{\sigma}|)\tau + (|\dot{\boldsymbol{u}}|+1)|\nabla\tau| \,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \frac{1}{d}\int_0^T \int_{\mathcal{O}} (\operatorname{tr}\boldsymbol{\sigma} - \operatorname{tr}\boldsymbol{\sigma}^s)(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\cdot\nabla\tau \,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{3.62}
$$

Having this estimate, we now focus on the limiting inequality for $\dot{\boldsymbol{u}}$. Defining $\dot{\boldsymbol{u}}^\mu_h(t) := \frac{1}{h}\int_0^h \dot{\boldsymbol{u}}^\mu(t+s)\,\mathrm{d}s$, we know from (3.40) and (3.42) that for all $t\in[0,T-h]$

$$
\dot{\boldsymbol{u}}^\mu_h(t) \to \dot{\boldsymbol{u}}_h(t) \quad \text{strongly in } L^1(\mathcal{O};\mathbb{R}^d). \tag{3.63}
$$

Consequently, using the embedding theorem[5], the assumptions on $\boldsymbol{u}_0$, we have that for all $\xi \in \mathcal{C}^1(\overline{\mathcal{O}})$ and all $t\in[0,T-h]$

$$
\begin{aligned}
\|\dot{\boldsymbol{u}}^\mu_h(t)\xi\|_{d'} &\le \|(\dot{\boldsymbol{u}}^\mu_h - (\dot{\boldsymbol{u}}_0)_h(t))\xi\|_{d'} + C\|\xi\|_{d'} \\
&\le C(\|\xi\|_{d'} + \|\boldsymbol{\varepsilon}((\dot{\boldsymbol{u}}^\mu_h(t) - (\dot{\boldsymbol{u}}_0)_h(t))\xi)\|_1) \\
&\le C(\|\xi\|_{d'} + \|\nabla\xi\|_1 + \||\dot{\boldsymbol{u}}^\mu_h(t)||\nabla\xi\|_1 + \||\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu_h(t))|\xi\|_1) \\
&\le C(\|\xi\|_{d'} + \|\nabla\xi\|_1 + \||\dot{\boldsymbol{u}}^\mu_h(t)||\nabla\xi|\|_1) + \frac{C}{h}\int_t^{t+h}\int_{\mathcal{O}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu(s))||\xi| \,\mathrm{d}x\,\mathrm{d}s.
\end{aligned}
$$

[5]In case that we do not prescribe Dirichlet data for $\boldsymbol{u}$, we have to chose ξ such that it has zero trace on a part of the boundary.

Thus, setting $\tau := \chi_{[t,t+h]}\xi$ in (3.62), letting $\mu \to 0_+$, using (3.63) and weak lower semicontiuity, we observe

$$\begin{aligned}\|\dot{\boldsymbol{u}}_h(t)\xi\|_{d'} &\le C(\|\xi\|_{d'} + \|\nabla\xi\|_1 + \||\dot{\boldsymbol{u}}_h(t)||\nabla\xi|\|_1)\\ &\quad + \frac{C}{h}\int_t^{t+h}\int_{\mathcal{O}} (|\dot{\boldsymbol{\sigma}}| + |\operatorname{tr}\boldsymbol{\sigma}|)(1+|\operatorname{tr}\boldsymbol{\sigma}|)\xi + (|\dot{\boldsymbol{u}}|+1)|\nabla\xi| \,\mathrm{d}x\,\mathrm{d}s\\ &\quad + \frac{C}{h}\int_t^{t+h}\left|\int_{\mathcal{O}} (\operatorname{tr}\boldsymbol{\sigma} - \operatorname{tr}\boldsymbol{\sigma}^s)(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\cdot\nabla\xi\,\mathrm{d}x\right|\mathrm{d}s.\end{aligned}$$

Finally, we restrict ourselves only on the Lebesgue points $t \in (0,T)$ for all related quantities, we can let $h \to 0_+$ in the above inequality to get

$$\begin{aligned}\|\dot{\boldsymbol{u}}(t)\xi\|_{d'} &\le C(\|\xi\|_{d'} + \|\nabla\xi\|_1 + \||\dot{\boldsymbol{u}}(t)||\nabla\xi|\|_1)\\ &\quad + C\int_{\mathcal{O}} (|\dot{\boldsymbol{\sigma}}| + |\operatorname{tr}\boldsymbol{\sigma}|)(1+|\operatorname{tr}\boldsymbol{\sigma}|)\xi\,\mathrm{d}x\\ &\quad + C\left|\int_{\mathcal{O}} (\operatorname{tr}\boldsymbol{\sigma} - \operatorname{tr}\boldsymbol{\sigma}^s)(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\cdot\nabla\xi\,\mathrm{d}x\right|,\end{aligned} \tag{3.64}$$

which is the starting point for further investigation. First, we use the Bogovskii operator and find $\mathbf{B}$ such that for almost all time $t \in (0,T)$

$$\operatorname{div}\mathbf{B} = (|\dot{\boldsymbol{\sigma}}| + |\operatorname{tr}\boldsymbol{\sigma}|)(1+|\operatorname{tr}\boldsymbol{\sigma}|) - \overline{(|\dot{\boldsymbol{\sigma}}| + |\operatorname{tr}\boldsymbol{\sigma}|)(1+|\operatorname{tr}\boldsymbol{\sigma}|)}_{\mathcal{O}} \text{ in } \mathcal{O}.$$

Consequently, using (3.37) and the fact that $\boldsymbol{\sigma} \in BMO$, we have that for all $p \in [1,2)$

$$\mathbf{B} \in L^\infty(0,T; W_0^{1,p}(\Omega;\mathbb{R}^{d\times d\times d})). \tag{3.65}$$

For further purposes we can also find $\boldsymbol{v}$ fulfilling

$$\boldsymbol{v} \in L^\infty(0,T; W_0^{1,p}(\Omega;\mathbb{R}^{d})) \tag{3.66}$$

and solving

$$\operatorname{div}\boldsymbol{v} = \operatorname{div}(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) - \overline{\operatorname{div}(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)}_{\mathcal{O}} \text{ in } \mathcal{O}.$$

Then, we can rewrite (3.64) with the help of integration by parts as

$$\begin{aligned}\|\dot{\boldsymbol{u}}(t)\xi\|_{d'} &\le C(\|\xi\|_{d'} + \|\nabla\xi\|_1 + \|(|\dot{\boldsymbol{u}}(t)| + |\mathbf{B}(t)|)|\nabla\xi|\|_1)\\ &\quad + C\left|\int_{\mathcal{O}} (\operatorname{tr}\boldsymbol{\sigma}(t) - \operatorname{tr}\boldsymbol{\sigma}^s(t))(\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\cdot\nabla\xi\,\mathrm{d}x\right|,\end{aligned} \tag{3.67}$$

We start with the interior estimates. Hence for arbitrary $x_0 \in \mathcal{O}$ and $R > 0$ such that $B_{2R}(x_0) \subset \Omega$. We find smooth nonnegative ξ such that $\xi = 1$ in $B_R(x_0)$ and $x = 0$ in $\mathbb{R}^d \setminus B_{2R}(x_0)$ satisfying $|\nabla\xi| \le CR^{-1}$. Then, using the assumption on $\boldsymbol{\sigma}^s$, $\boldsymbol{u}_0$ and (3.65), the estimate (3.67) implies that for almost all $t \in (0,T)$

$$\begin{aligned}\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^{d'}}{R^d}\,\mathrm{d}x &\le C + C\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}(t)| + |\mathbf{B}(t)|}{R^d}\,\mathrm{d}x\right)^{d'}\\ &\quad + C\left|\int_{B_{2R}(x_0)} \frac{|\operatorname{tr}\boldsymbol{\sigma}(t) - \overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}||\dot{\boldsymbol{u}(t)}|}{R^d}\,\mathrm{d}x\right|^{d'}\\ &\quad + C\left|\int_{B_{2R}(x_0)} \frac{\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}(\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\cdot\nabla\xi}{R^{d-1}}\,\mathrm{d}x\right|^{d'}.\end{aligned} \tag{3.68}$$

Next, using the Hölder inequality and the BMO estimate for $\boldsymbol{\sigma}$ we get that that for any $q \in (1, d')$

$$\left|\int_{B_{2R}(x_0)} \frac{|\operatorname{tr}\boldsymbol{\sigma}(t) - \overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}||\dot{\boldsymbol{u}}(t)|}{R^d}\,\mathrm{d}x\right|^{d'} \le C\left|\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^q}{R^d}\,\mathrm{d}x\right|^{\frac{d'}{q}}. \tag{3.69}$$

For the last term in (3.68), we use integration by parts and the definition of $\boldsymbol{v}$ to get (using again the BMO estimate for $\boldsymbol{\sigma}$)

(3.70)

$$\begin{aligned}
&\left|\int_{B_{2R}(x_0)} \frac{\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}(\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\cdot\nabla\xi}{R^{d-1}}\,\mathrm{d}x\right|^{d'}\\
&= \left|\int_{B_{2R}(x_0)} \frac{\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}\operatorname{div}(\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\xi}{R^{d-1}}\,\mathrm{d}x\right|^{d'}\\
&\le C\left|\int_{B_{2R}(x_0)} \frac{|\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}|}{R^{d-1}}\,\mathrm{d}x\right|^{d'} + \left|\int_{B_{2R}(x_0)} \frac{\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}\operatorname{div}\boldsymbol{v}(t)\xi}{R^{d-1}}\,\mathrm{d}x\right|^{d'}\\
&\le C + \left|\int_{B_{2R}(x_0)} \frac{\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{B_{2R}(x_0)}\boldsymbol{v}(t)\cdot\nabla\xi}{R^{d-1}}\,\mathrm{d}x\right|^{d'}\\
&\le C + C\left(\int_{B_{2R}(x_0)} \frac{|\boldsymbol{\sigma}(t)|}{R^d}\,\mathrm{d}x\right)^{d'}\left(\int_{B_{2R}(x_0)} \frac{|\boldsymbol{v}(t)|}{R^d}\,\mathrm{d}x\right)^{d'}\\
&\le C + C\int_{B_{2R}(x_0)} \frac{|\boldsymbol{\sigma}(t)|^{zd'}}{R^d}\,\mathrm{d}x + C\int_{B_{2R}(x_0)} \frac{|\boldsymbol{v}(t)|^{z'd'}}{R^d}\,\mathrm{d}x,
\end{aligned}$$

where we chose z large enough so that

$$z'd' < \frac{2d}{d-2}.$$

Substituting (3.69)–(3.70) into (3.68), we finally deduce

(3.71)

$$\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^{d'}}{R^d}\,\mathrm{d}x \le C + C\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^q}{R^d}\,\mathrm{d}x\right)^{\frac{d'}{q}} + C\int_{B_{2R}(x_0)} \frac{|g(t)|}{R^d}\,\mathrm{d}x,$$

where

$$g(t) := |\boldsymbol{\sigma}(t)|^{zd'} + |\boldsymbol{v}(t)|^{z'd'} + |\mathsf{B}(t)|^{d'}.$$

Note that due to the BMO estimates for $\boldsymbol{\sigma}$, the properties (3.65)–(3.66) and the embedding theorem, we know that any $\varepsilon > 0$ find $z > 1$ such that[6]

$$g \in L^\infty(0, T; L^{\frac{2d-\varepsilon}{d-2}}(\mathcal{O})).$$

Thus, we have everything prepared for the reverse Hölder inequality in the interior of $\mathcal{O}$. To get it up to the boundary, let us now consider $x_0 \in \partial\mathcal{O}$ such that $B_{2R}(x_0) \cap \partial\mathcal{O}_N = \emptyset$. In this case we can proceed exactly as before, replacing only the mean values accordingly (using the Lipchitz continuity of $\mathcal{O}$ and the global

[6]In case $d = 2$ we get any L^p space. Similarly, in case of the Hencky model, we do not deal with the time derivative estimates and we can get any L^p estimates.

BMO estimates for $\boldsymbol{\sigma}$). Notice here that the integration by parts in (3.70) still can be used since $(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot \boldsymbol{n} = 0$ on $\partial\mathcal{O} \setminus \partial\mathcal{O}_N$ in the sense of traces.

Thus in what follows, we focus on the case when $x_0 \in \partial\mathcal{O}_N$ and $R > 0$ be such that $B_{2R}(x_0) \cap \partial\mathcal{O} \subset \partial\mathcal{O}_N$. In this case we can use the same procedure as above replacing the mean values accordingly except the integration by parts used in (3.70) since we do not control $(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot \boldsymbol{n}$ on $\partial\mathcal{O}_N$. Hence, instead of (3.71) we arrive at

$$\begin{aligned}\int_{B_R(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^{d'}}{R^d} \,\mathrm{d}x &\le C(1 + |\overline{\operatorname{tr}\boldsymbol{\sigma}(t)}_{\mathcal{O}\cap B_{2R}(x_0)}|) \left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^q}{R^d} \,\mathrm{d}x \right)^{\frac{d'}{q}} \\ &\quad + C \int_{B_{2R}(x_0)} \frac{1 + |g(t)|}{R^d} \,\mathrm{d}x. \end{aligned} \tag{3.72}$$

Next, the key difference from the interior estimates is that we need to provide a sufficiently good estimate for the mean value of $\operatorname{tr}\boldsymbol{\sigma}(t)$. Without loss of generality (since $\mathcal{O}$ is Lipschitz), assume that $x_0 = 0$ and that

$$\begin{aligned} V &:= \{x \in \mathbb{R}^d;\ x = (x', x_d), x' \in (-kR, kR)^{d-1},\ a(x') < x_d < a(x') + kR\} \\ &\supset \mathcal{O} \cap B_{2R}(x_0), \end{aligned}$$

where k is some constat depending on $\partial\mathcal{O}$ and a is a Lipschitz function and if $a(x') = x_d$ then $x \in \partial\mathcal{O}$. Next, since for any $x \in \partial\mathcal{O}_N$ we have in the sense of traces

$$\operatorname{tr}\boldsymbol{\sigma}(t) = \operatorname{tr}\boldsymbol{\sigma}(t)\boldsymbol{n}\cdot\boldsymbol{n} = d(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}_D(t))\boldsymbol{n}\cdot\boldsymbol{n} = d(\boldsymbol{\sigma}^s(t) - \boldsymbol{\sigma}_D(t)) \cdot (\boldsymbol{n} \otimes \boldsymbol{n}),$$

where we used the fact that $(\boldsymbol{\sigma} - \boldsymbol{\sigma}^s)\boldsymbol{n} = 0$ on $\partial\mathcal{O}_N$. Consequently since $|\boldsymbol{\sigma}_D| \le 1$ and $\boldsymbol{\sigma}^s$ is smooth we have that

$$\operatorname{tr}\boldsymbol{\sigma} \in L^\infty(0, T; L^\infty(\partial\mathcal{O}_N)).$$

Therefore, for arbitrary $x \in V$, we have (for a.a. t, which we do not write here)

$$\begin{aligned} \operatorname{tr}\boldsymbol{\sigma}(x) - \operatorname{tr}\boldsymbol{\sigma}(x', a(x')) &= \int_{a(x')}^{x_d} \frac{\partial \operatorname{tr}\boldsymbol{\sigma}(x', s)}{\partial s} \,\mathrm{d}s \\ &= d \int_{a(x')}^{x_d} \sum_{i=1}^{d} \frac{\boldsymbol{\sigma}^{dj}(x', s) - \boldsymbol{\sigma}_D^{dj}(x', s)}{\partial x_j} \,\mathrm{d}s \\ &= d \int_{a(x')}^{x_d} \sum_{i=1}^{d} \frac{(\boldsymbol{\sigma}^s)^{dj}(x', s) - \boldsymbol{\sigma}_D^{dj}(x', s)}{\partial x_j} \,\mathrm{d}s \\ &= d((\boldsymbol{\sigma}^s)^{dj}(x) - \boldsymbol{\sigma}_D^{dj}(x)) - d((\boldsymbol{\sigma}^s)^{dj}(x', a(x')) - \boldsymbol{\sigma}_D^{dj}(x', a(x')) \\ &\quad + d \int_{a(x')}^{x_d} \sum_{i=1}^{d-1} \frac{(\boldsymbol{\sigma}^s)^{dj}(x', s) - \boldsymbol{\sigma}_D^{dj}(x', s)}{\partial x'_j} \,\mathrm{d}s. \end{aligned}$$

Thus, using the bound for $\operatorname{tr}\boldsymbol{\sigma}$ on $\partial\mathcal{O}_N$ and the assumptions on $\boldsymbol{\sigma}^s$ we have

$$\begin{aligned} \left| \int_V \operatorname{tr}\boldsymbol{\sigma}(x) \,\mathrm{d}x \right| &\le C|V| \\ &+ d \left| \int_{(-kR,kR)^{d-1}} \int_{a(x')}^{a(x')+kR} \int_{a(x')}^{x_d} \sum_{i=1}^{d-1} \frac{(\boldsymbol{\sigma}^s)^{dj}(x', s) - \boldsymbol{\sigma}_D^{dj}(x', s)}{\partial x'_j} \,\mathrm{d}s \,\mathrm{d}x_d \,\mathrm{d}x' \right|. \end{aligned}$$

Next, using integration by parts, we deduce

$$\begin{aligned}
&\left|\int_{(-kR,kR)^{d-1}}\int_{a(x')}^{a(x')+kR}\int_{a(x')}^{x_d}\sum_{i=1}^{d-1}\frac{(\boldsymbol{\sigma}^s)^{dj}(x',s)-\boldsymbol{\sigma}_D^{dj}(x',s)}{\partial x_j'}\,\mathrm{d}s\,\mathrm{d}x_d\,\mathrm{d}x'\right|\\
&=\left|\int_{(-kR,kR)^{d-1}}\int_{a(x')}^{a(x')+kR}(x_d-a(x')-kR)\sum_{i=1}^{d-1}\frac{(\boldsymbol{\sigma}^s)^{dj}(x)-\boldsymbol{\sigma}_D^{dj}(x)}{\partial x_j'}\,\mathrm{d}x_d\,\mathrm{d}x'\right|\\
&=\left|\int_V(x_d-a(x')-kR)\sum_{i=1}^{d-1}\frac{(\boldsymbol{\sigma}^s)^{dj}(x)-\boldsymbol{\sigma}_D^{dj}(x)}{\partial x_j}\,\mathrm{d}x\right|\\
&\le\left|\int_{\partial V}\sum_{i=1}^{d-1}(x_d-a(x')-kR)((\boldsymbol{\sigma}^s)^{dj}(x)-\boldsymbol{\sigma}_D^{dj}(x))\boldsymbol{n}_j\,\mathrm{d}S\right|\\
&\quad+\left|\int_V\sum_{i=1}^{d-1}\frac{a(x')}{\partial x_j'}((\boldsymbol{\sigma}^s)^{dj}(x)-\boldsymbol{\sigma}_D^{dj}(x))\,\mathrm{d}x\right|\\
&\le C|V|+CR|\partial V|\le C|V|.
\end{aligned}$$

Thus, combining the estimates together, we have that

$$|\overline{\operatorname{tr}\boldsymbol{\sigma}}_V|\le C. \tag{3.73}$$

Hence, since the mean value is bounded, we can use it in (3.72) and continue as before. The rest of the proof, i.e., the proper estimates for any x_0 can be deduced from the above estimates and the triangle inequality. This finishes the proof of (2.15).

3.7. **Morrey condition for $\dot{\boldsymbol{e}}_{\mathbf{p}}$ and λ.** We again start we the point wise estimate for the approximation. Using the fact that $\boldsymbol{\sigma}^s$ satisfies the safety load condition, we have

$$\begin{aligned}
\delta\mu^{-1}(|\boldsymbol{\sigma}_D^\mu|-1)_+&\le\mu^{-1}(|\boldsymbol{\sigma}_D^\mu|-1)_+\frac{\boldsymbol{\sigma}_D^\mu}{|\boldsymbol{\sigma}_D^\mu|}\cdot(\boldsymbol{\sigma}^\mu-\boldsymbol{\sigma}^s)\\
&=\mathsf{A}\dot{\boldsymbol{\sigma}}^\mu\cdot(\boldsymbol{\sigma}^\mu-\boldsymbol{\sigma}^s)+\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu-\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}^\mu-\boldsymbol{\sigma}^s)+\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\boldsymbol{\sigma}^\mu-\boldsymbol{\sigma}^s).
\end{aligned}$$

Multiplying this inequality by arbitrary nonnegative $\tau\in\mathcal{C}^1(\overline{\Omega})$, integrating the result over $\mathcal{O}$, using integration by parts and letting $\mu\to 0_+$, (we omit details here, since the very similar step was carefully justified in the previous section) we get that for almost all $t\in(0,T)$

$$\delta\langle\lambda(t),\tau\rangle\le\int_{\mathcal{O}}C(1+|\dot{\boldsymbol{\sigma}}(t)|+|\boldsymbol{\sigma}(t)|)(1+|\boldsymbol{\sigma}(t)|)\tau-(\boldsymbol{\sigma}(t)-\boldsymbol{\sigma}^s(t))\cdot(\nabla\tau\otimes\dot{\boldsymbol{u}}(t))\,\mathrm{d}x. \tag{3.74}$$

Since $\boldsymbol{\sigma}\in L^\infty(0,T;BMO(\mathcal{O};\mathbb{R}^{d\times d}_{sym}))$ and $\dot{\boldsymbol{u}}\in L^\infty(0,T;L^{d'+\varepsilon}(\mathcal{O};\mathbb{R}^d))$ for some $\varepsilon>0$, we can find $\delta>0$ such that after using the Hölder inequality we have

$$\operatorname*{ess\,sup}_{t\in(0,T)}\langle\lambda(t),\tau\rangle\le C\|\tau\|_{1,d-\delta}. \tag{3.75}$$

Consequently, any Borel $\tilde{\mathcal{O}}$ having zero $(d-\delta)$-Sobolev capacity satisfies for almost all $t\in(0,T)$

$$\lambda(t)(\tilde{\mathcal{O}})=0, \tag{3.76}$$

In addition, taking for arbitrary $x_0 \in \overline{\mathcal{O}}$ and arbitrary $R > 0$ the function $\tau = 1$ in $B_R(x_0)$ and 0 outside the ball $B_{2R}(x_0)$ satisfying $\nabla\tau| \le CR^{-1}$, we get that there exists $\alpha > 0$ such that for all balls we have

$$\operatorname*{ess\,sup}_{t\in(0,T)} \lambda(t)(\overline{B_R(x_0)\cap\mathcal{O}}) \le CR^{\alpha},$$

which is the last part of the claim stated in (2.16).

Hence, we also see that λ is absolutely continuous with respect to α-dimensional Hausdorff measure. In addition, repeating step by step the procedure from the previous subsection, we can get for arbitrary $\delta > 0$ the estimate

$$\begin{aligned}\lambda(t)(\overline{B_R(x_0)\cap\mathcal{O}}) &\le C\int_{B_{2R}(x_0)} (1+|\dot{\boldsymbol{\sigma}}(t)|+|\boldsymbol{\sigma}(t)|+\operatorname{div}\dot{\boldsymbol{u}}(t)|)(1+|\boldsymbol{\sigma}(t)|)\,\mathrm{d}x\\ &\quad C(\delta)R^{d-1}\left(\int_{B_{2R}(x_0)} \frac{|\dot{\boldsymbol{u}}(t)|^{1+\delta}}{R^d}\right)^{\frac{1}{1+\delta}}.\end{aligned} \tag{3.77}$$

Consequently, we see that decomposing λ as

$$\lambda := \lambda^d + \lambda^s = \lambda^d + \lambda^{d-1} + \lambda^c,$$

where λ^d is the regular part of λ, i.e., continuous with respect to the Lebesgue measure, λ^{d-1} is continuous with respect to the $(d-1)$-Hausdorff measure and λ^c is the Cantor part, we immediately have that the Cantor part is supported only on the set where $M|\dot{\boldsymbol{u}}| = \infty$.

3.8. **Identification of $\dot{\boldsymbol{e}}_{\mathrm{p}}$.** This part will finish the proof of Theorem 2.1. First, we will recall here the standard procedure of identification of $\dot{\boldsymbol{e}}_{\mathrm{p}}$ (compare with (1.8) and (1.9) and also with e.g. [4]) and further we will show that in two dimensional setting, we have a sharper result. Namely, we shall show that

$$\dot{\boldsymbol{e}}_{\mathrm{p}}(t) = \lambda\boldsymbol{\sigma}_p(t),$$

where

$$\boldsymbol{\sigma}_p(t) = \boldsymbol{\sigma}_D(t) \qquad \lambda^d \text{ - almost everywhere}$$

and for almost all $t \in (0,T)$

$$\lim_{\varepsilon\to 0_+}\int_{\overline{\mathcal{O}}} |\boldsymbol{\sigma}_p(t) - \boldsymbol{\sigma}^{\varepsilon}_D(t)|\,\mathrm{d}\lambda^s = 0,$$

where $\boldsymbol{\sigma}^{\varepsilon}_D(t)$ is a proper mollification of $\boldsymbol{\sigma}_D$. Further, we show that if $d = 2$, we can simply set $\boldsymbol{\sigma}_p(t) := \boldsymbol{\sigma}_D(t)$ everywhere in $\mathcal{O}$.

To start with this plan, we first notice that for arbitrary $\tilde{\boldsymbol{\sigma}} \in L^1(0,T;\mathcal{C}(\overline{\mathcal{O}};\mathbb{R}^{d\times d}_{sym}))$ fulfilling $\|\tilde{\boldsymbol{\sigma}}_D(t)\|_\infty \le 1$, and almost all $t \in (0,T)$ we have

$$\begin{aligned}0 &\le \mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+\frac{\boldsymbol{\sigma}^{\mu}_D(t)}{|\boldsymbol{\sigma}^{\mu}_D(t)|}\cdot(\boldsymbol{\sigma}^{\mu}_D(t)-\tilde{\boldsymbol{\sigma}}_D(t))\\ &= \mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)^2_+ + \mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+ - \mu^{-1}(|\boldsymbol{\sigma}^{\mu}_D(t)|-1)_+\frac{\boldsymbol{\sigma}^{\mu}_D(t)}{|\boldsymbol{\sigma}^{\mu}_D(t)|}\cdot\tilde{\boldsymbol{\sigma}}_D\end{aligned}$$

Hence, using the convergence results (3.38)–(3.39) and (3.50), we deduce that for arbitrary nonnegative $\tau \in L^1(0,T;\mathcal{C}(\overline{\mathcal{O}}))$ we have

$$0 \le \int_0^T \langle\lambda(t),\tau(t)\rangle - \langle\dot{\boldsymbol{e}}_{\mathrm{p}}(t),\tilde{\boldsymbol{\sigma}}_D(t)\tau(t)\rangle\,\mathrm{d}t. \tag{3.78}$$

Consequently, since τ and $\tilde{\boldsymbol{\sigma}}$ are arbitrary, we also have for almost all $t \in (0,T)$

$$|\dot{\boldsymbol{e}}_{\mathrm{p}}(t)| \le \lambda(t) \quad \text{in sense of measures on } \overline{\mathcal{O}}. \tag{3.79}$$

In addition, for almost all all time $t \in (0,T)$ and arbitrary $\tau \in \mathcal{C}^1(\overline{\mathcal{O}})$ we have that (note that the limits below now denote the weak* limits in $L^\infty(0,T)$)

(3.80)

$$\begin{aligned}
&\langle \lambda(t), \tau\rangle - \langle \dot{\boldsymbol{e}}_{\mathrm{p}}(t), \boldsymbol{\sigma}^s_D(t)\tau\rangle \\
&= \lim_{\mu\to 0}\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\tau - \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot \boldsymbol{\sigma}^s_D(t)\tau \,\mathrm{d}x \\
&= \lim_{\mu\to 0}\int_{\mathcal{O}} -\mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)^2_+\tau + \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}^s_D(t)\tau \,\mathrm{d}x \\
&= \lim_{\mu\to 0}\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}_D(t))\tau \,\mathrm{d}x \\
&\quad + \lim_{\mu\to 0}\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}_D(t) - \boldsymbol{\sigma}^s_D(t))\tau \,\mathrm{d}x
\end{aligned}$$

To evaluate limits on the right hand side we use (3.1). First, since $|\boldsymbol{\sigma}_D| \le 1$, we have that

$$\begin{aligned}
&\left|\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}_D(t))\tau \,\mathrm{d}x\right| \\
&\quad \le \|\tau\|_\infty \int_{\mathcal{O}} \left|\mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}_D(t))\right| \mathrm{d}x \\
&= \|\tau\|_\infty \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}_D(t)) \,\mathrm{d}x \\
&= \|\tau\|_\infty \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}^\mu(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)))\cdot(\boldsymbol{\sigma}^\mu(t) - \boldsymbol{\sigma}(t)) \,\mathrm{d}x,
\end{aligned}$$

where the last equality follows from the integration by parts and the fact that $\operatorname{div}\boldsymbol{\sigma}^\mu = \operatorname{div}\boldsymbol{\sigma}$. Thus, employing (3.12) and (3.48), we see that

$$\lim_{\mu\to 0}\sup_{t\in(0,T)}\left|\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}^\mu_D(t) - \boldsymbol{\sigma}_D(t))\tau \,\mathrm{d}x\right| = 0.$$

Similarly, for the second term in (3.80) we have

$$\begin{aligned}
&\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}^\mu_D(t)|-1)_+\frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|}\cdot (\boldsymbol{\sigma}_D(t) - \boldsymbol{\sigma}^s_D(t))\tau \,\mathrm{d}x \\
&= \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}^\mu(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu(t) - \dot{\boldsymbol{u}}_0(t)) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)))\cdot(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t))\tau \,\mathrm{d}x \\
&= \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}^\mu(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)))\cdot(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t))\tau \,\mathrm{d}x \\
&\qquad - \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}^\mu(t) - \dot{\boldsymbol{u}}_0(t))\otimes\nabla\tau)\cdot(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t)) \,\mathrm{d}x \\
&\overset{\mu\to 0_+}{\to} \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)))\cdot(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t))\tau \,\mathrm{d}x \\
&\qquad - \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\otimes\nabla\tau)\cdot(\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t)) \,\mathrm{d}x\,,
\end{aligned}$$

where for the identification of the last weak limit we used (3.37) and (3.42). Thus, the relation (3.80) reduces for almost all $t \in (0,T)$ to

$$\begin{aligned}\langle \lambda(t), \tau \rangle - \langle \dot{\boldsymbol{e}}_{\mathrm{p}}(t), \boldsymbol{\sigma}^s_D(t)\tau \rangle = &\int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))) \cdot (\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t))\tau \,\mathrm{d}x \\ &- \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t)) \otimes \nabla\tau) \cdot (\boldsymbol{\sigma}(t) - \boldsymbol{\sigma}^s(t)) \,\mathrm{d}x \,. \end{aligned} \tag{3.81}$$

Finally, for arbitrary $\tilde{\boldsymbol{\sigma}} \in \mathcal{C}([0,T] \times \overline{\mathcal{O}}; \mathbb{R}^{d\times d})$ with $\operatorname{div}\tilde{\boldsymbol{\sigma}} \in L^1(0,T; L^d(\mathcal{O};\mathbb{R}^d))$ fulfilling $\tilde{\boldsymbol{\sigma}}\boldsymbol{n} = 0$ on $\partial\mathcal{O}_N$ and $(\tilde{\boldsymbol{\sigma}}\boldsymbol{n})_{\boldsymbol{\tau}} = 0$ on $\partial\mathcal{O}_M$ we get by using (3.38) and the identity (3.1) the following identification of weak limits (for almost all $t \in (0,T)$)

$$\begin{aligned}\langle \dot{\boldsymbol{e}}_{\mathrm{p}}(t), \tilde{\boldsymbol{\sigma}}_D(t)\tau \rangle &= \lim_{\mu\to 0} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D(t)| - 1)_+ \frac{\boldsymbol{\sigma}^\mu_D(t)}{|\boldsymbol{\sigma}^\mu_D(t)|} \cdot \tilde{\boldsymbol{\sigma}}(t)\tau \,\mathrm{d}x \\ &= \lim_{\mu\to 0} \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}^\mu(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu(t) - \dot{\boldsymbol{u}}_0(t))) \cdot \tilde{\boldsymbol{\sigma}}(t)\tau \,\mathrm{d}x \\ &= \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))) \cdot \tilde{\boldsymbol{\sigma}}(t)\tau \,\mathrm{d}x \\ &\quad - \lim_{\mu\to 0} \int_{\mathcal{O}} (\dot{\boldsymbol{u}}^\mu(t) - \dot{\boldsymbol{u}}_0(t)) \cdot \operatorname{div}(\tilde{\boldsymbol{\sigma}}(t)\tau) \,\mathrm{d}x \\ &= \int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))) \cdot \tilde{\boldsymbol{\sigma}}(t)\tau \,\mathrm{d}x \\ &\quad - \int_{\mathcal{O}} (\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t))\tau \cdot \operatorname{div}\tilde{\boldsymbol{\sigma}}(t) \,\mathrm{d}x - \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t)) \otimes \nabla\tau) \cdot \tilde{\boldsymbol{\sigma}}(t) \,\mathrm{d}x \,. \end{aligned} \tag{3.82}$$

Based on the above identity, we finally identify $\dot{\boldsymbol{e}}_{\mathrm{p}}$. We extend $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^s$ outside $\mathcal{O}$ and for any $\varepsilon > 0$ define a continuous function (w.r.t x)

$$\boldsymbol{\sigma}^\varepsilon(t,x) := \frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} \boldsymbol{\sigma}(t,y) \,\mathrm{d}y$$

and also similarly

$$\boldsymbol{\sigma}^{\varepsilon,s}(t,x) := \frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} \boldsymbol{\sigma}^s(t,y) \,\mathrm{d}y.$$

Due to this definition and the estimate (2.15), we have

$$\boldsymbol{\sigma}^\varepsilon \to \boldsymbol{\sigma} \qquad \text{strongly in } L^p((0,T) \times \mathcal{O}; \mathbb{R}^{d\times d}) \text{ for all } p \in [1,\infty), \tag{3.83}$$

$$\boldsymbol{\sigma}^{\varepsilon,s} \to \boldsymbol{\sigma}^s \qquad \text{strongly in } \mathcal{C}([0,T] \times \overline{\mathcal{O}}; \mathbb{R}^{d\times d}). \tag{3.84}$$

Next, assume that τ is compactly supported. Therefore, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ and for all x belonging to the support of τ we have $B_\varepsilon(x) \subset \mathcal{O}$. Consequently, we have

$$\operatorname{div}(\boldsymbol{\sigma}^\varepsilon - \boldsymbol{\sigma}^{\varepsilon,s}) = 0 \text{ in } \operatorname{supp}\tau$$

Therefore, setting $\tilde{\boldsymbol{\sigma}} := \boldsymbol{\sigma}^\varepsilon - \boldsymbol{\sigma}^{\varepsilon,s}$ (notice that although $\tilde{\boldsymbol{\sigma}}$ does not have zero trace on $\partial\mathcal{O}$ such choice is possible since τ has compact support) in (3.82) we obtain

$$\begin{aligned}\langle \dot{\boldsymbol{e}}_{\mathrm{p}}(t), (\boldsymbol{\sigma}^\varepsilon_D(t) - \boldsymbol{\sigma}^{\varepsilon,s}_D(t))\tau \rangle = &\int_{\mathcal{O}} (-\mathsf{A}\dot{\boldsymbol{\sigma}}(t) + \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t))) \cdot (\boldsymbol{\sigma}^\varepsilon(t) - \boldsymbol{\sigma}^{\varepsilon,s}(t))\tau \,\mathrm{d}x \\ &- \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}_0(t)) \otimes \nabla\tau) \cdot (\boldsymbol{\sigma}^\varepsilon(t) - \boldsymbol{\sigma}^{\varepsilon,s}(t)) \,\mathrm{d}x \,. \end{aligned}$$

Thus, letting $\varepsilon \to 0$ and comparing the result with (3.81), we obtain that for all $\varphi \in L^\infty(0,T)$

$$
\begin{aligned}
\int_0^T \langle \lambda, \tau\rangle \varphi \,\mathrm{d}t &= \int_0^T (\langle \lambda, \tau\rangle - \langle \dot{\boldsymbol{e}}_{\mathrm{p}}, \boldsymbol{\sigma}^s_D \tau\rangle + \langle \dot{\boldsymbol{e}}_{\mathrm{p}}, \boldsymbol{\sigma}^s_D \tau\rangle)\varphi \,\mathrm{d}t \\
&= \lim_{\varepsilon\to 0} \int_0^T \langle \dot{\boldsymbol{e}}_{\mathrm{p}}, (\boldsymbol{\sigma}^\varepsilon_D - \boldsymbol{\sigma}^{\varepsilon,s}_D)\tau\rangle\varphi + \langle \dot{\boldsymbol{e}}_{\mathrm{p}}, \boldsymbol{\sigma}^s_D\tau\rangle\varphi \,\mathrm{d}t \\
&= \lim_{\varepsilon\to 0} \int_0^T \langle \dot{\boldsymbol{e}}_{\mathrm{p}}, \boldsymbol{\sigma}^\varepsilon_D \tau\rangle \varphi \,\mathrm{d}t,
\end{aligned} \tag{3.85}
$$

where for the second equality we used (3.84). Since $|\boldsymbol{\sigma}_D| \le 1$ almost everywhere, we see that $|\boldsymbol{\sigma}^\varepsilon_D| \le 1$ everywhere and thus it follows from (3.78) and (3.85) that for almost all time $t \in (0,T)$

$$
\lambda(t) = |\dot{\boldsymbol{e}}_{\mathrm{p}}(t)| \quad \text{ in } \mathcal{M}(\mathcal{O}). \tag{3.86}
$$

Thus, denoting $\boldsymbol{\sigma}_p$ the Radon–Nykodým derivative of $\dot{\boldsymbol{e}}_{\mathrm{p}}$ we have that

$$
\dot{\boldsymbol{e}}_{\mathrm{p}}(t) = \lambda(t)\boldsymbol{\sigma}_p(t), \qquad |\boldsymbol{\sigma}_p(t)| = 1 \quad \lambda(t) \text{ almost everywhere in } \overline{\mathcal{O}}.
$$

Consequently, using the facts that $|\boldsymbol{\sigma}^\varepsilon_D| \le 1$ and $|\boldsymbol{\sigma}_p| \le 1$ λ-everywhere and (3.85) with $\varphi \equiv 1$ we deduce that

$$
\begin{aligned}
\int_0^T \int_{\mathcal{O}} |\boldsymbol{\sigma}^\varepsilon_D - \boldsymbol{\sigma}_p|^2 \tau \,\mathrm{d}\lambda \,\mathrm{d}t &= \int_0^T \int_{\mathcal{O}} (|\boldsymbol{\sigma}^\varepsilon_D|^2 + |\boldsymbol{\sigma}_p|^2)\tau - 2\boldsymbol{\sigma}^\varepsilon_D \cdot \boldsymbol{\sigma}_p \tau \,\mathrm{d}\lambda \,\mathrm{d}t \\
&\le 2\int_0^T \int_{\mathcal{O}} (1 - \boldsymbol{\sigma}^\varepsilon_D \cdot \boldsymbol{\sigma}_p)\tau \,\mathrm{d}\lambda \,\mathrm{d}t \\
&= 2\int_0^T \langle \lambda, \tau\rangle - 2\langle \dot{\boldsymbol{e}}_{\mathrm{p}}, \boldsymbol{\sigma}^\varepsilon_D \tau\rangle \,\mathrm{d}t \overset{\varepsilon\to 0}{\to} 0,
\end{aligned} \tag{3.87}
$$

which is the desired interior characterization of $\boldsymbol{\sigma}_p$.

Similarly, we shall proceed in a neighborhood of the Dirichlet boundary $\partial\mathcal{O}_D$. Hence let τ be compactly supported in a sufficiently small neighborhood of $\partial\mathcal{O}_D$ and let $\boldsymbol{n}$ denoted the outward unit vector at some fixed point. Then, we set in (3.82)

$$
\tilde{\boldsymbol{\sigma}}(x) := \boldsymbol{\sigma}^\varepsilon(x - \alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}^{\varepsilon,s}(x - \alpha\varepsilon\boldsymbol{n}),
$$

where α is sufficiently large (depending on the Lipschtz constant of $\partial\mathcal{O}$. Due to this definition, we still have that $\operatorname{div}\tilde{\boldsymbol{\sigma}} = 0$ in the support of τ. Thus, we can repeat the whole procedure again to get for almost all time $t \in (0,T)$

$$
\dot{\boldsymbol{e}}_{\mathrm{p}}(t) = \lambda\boldsymbol{\sigma}_p \text{ in } \mathcal{O} \cup \partial\mathcal{O}_D,
$$

and for arbitrary τ supported in a small neighborhood of $\partial\mathcal{O}_D$

$$
\lim_{\varepsilon\to 0_+} \int_{\overline{\mathcal{O}}} |\boldsymbol{\sigma}^\varepsilon_D(x - \alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}_p|^2 \tau(x) \,\mathrm{d}\lambda(t,x) = 0.
$$

Very similarly, we proceed also in a neighborhood of the Neumann boundary $\partial\mathcal{O}_N$, where instead of shifting outside, we shift the stress inside. Indeed, since $\boldsymbol{\sigma}\boldsymbol{n} = \boldsymbol{\sigma}^s\boldsymbol{n}$ on $\partial\mathcal{O}_N$ and $\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) = 0$ in $\mathcal{O}$, extending $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^s$ outside $\mathcal{O}$, we obtain that in the sense of distribution we have in a neighborhood of $\partial\mathcal{O}_N$ that $\operatorname{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^s) = 0$. Consequently, setting

$$
\tilde{\boldsymbol{\sigma}}(x) := \boldsymbol{\sigma}^\varepsilon(x + \alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}^{\varepsilon,s}(x + \alpha\varepsilon\boldsymbol{n}),
$$

we have $\operatorname{div}\tilde{\boldsymbol{\sigma}} = 0$ in $\mathcal{O}$ and eve more $\tilde{\boldsymbol{\sigma}} = 0$ on $\partial\mathcal{O}_N$. Therefore, repeating step by step the above we have for almost all time $t \in (0,T)$

$$\dot{\boldsymbol{e}}_{\mathrm{p}}(t) = \lambda(t)\boldsymbol{\sigma}_p(t) \text{ in } \mathcal{O} \cup \partial\mathcal{O}_D \cup \partial\mathcal{O}_N,$$

and for arbitrary τ supported in a small neighborhood of $\partial\mathcal{O}_N$

$$\lim_{\varepsilon\to 0_+} \int_{\overline{\mathcal{O}}} |\boldsymbol{\sigma}^\varepsilon_D(x+\alpha\varepsilon\boldsymbol{n}) - \overline{\boldsymbol{\sigma}}_D(x)|^2 \tau(x)\,\mathrm{d}\lambda(t,x) = 0.$$

In particular, due to the definition, we also have that

$$\begin{aligned}\int_{\partial\mathcal{O}_N} |\boldsymbol{\sigma}^s_D - \boldsymbol{\sigma}_p|^2\tau\,\mathrm{d}\lambda &= \lim_{\varepsilon\to 0_+}\int_{\partial\mathcal{O}_N} |\boldsymbol{\sigma}^{\varepsilon,s}_D(x+\alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}_p|^2\tau(x)\,\mathrm{d}\lambda(t,x)\\ &= \lim_{\varepsilon\to 0_+}\int_{\partial\mathcal{O}_N} |\boldsymbol{\sigma}^{\varepsilon}_D(x+\alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}_p|^2\tau(x)\,\mathrm{d}\lambda(t,x) = 0.\end{aligned}$$

But since $|\boldsymbol{\sigma}^s_D| < 1$ and $|\boldsymbol{\sigma}_p| = 1$ almost λ everywhere on $\partial\mathcal{O}$, we see that necessarily $\lambda(\partial\mathcal{O}_N) = 0$.

Finally, close to $\partial\mathcal{O}_M$ we do not mollify by using mean values but we rather mollify by using the convolution kernel and set (this means we shift the deviatoric part inside and the trace part outside)

$$\tilde{\boldsymbol{\sigma}}(x) := (\boldsymbol{\sigma}^\varepsilon_D - \boldsymbol{\sigma}^{\varepsilon,s}_D)(x+\alpha\varepsilon\boldsymbol{n}) + \frac{1}{d}\mathbf{I}(\operatorname{tr}\boldsymbol{\sigma}^\varepsilon - \operatorname{tr}\boldsymbol{\sigma}^{\varepsilon,s})(x-\alpha\varepsilon\boldsymbol{n}).$$

First, it is evident that $(\tilde{\boldsymbol{\sigma}}\boldsymbol{n})_{\boldsymbol{\tau}} = 0$ on $\partial\mathcal{O}_M$. Next, we also evaluate its divergence. Hence, for arbitrary τ vanishing in a neighborhood of $\partial\mathcal{O}_M$, we have

$$\begin{aligned}
&\int_{\mathcal{O}} (\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\tau\cdot\operatorname{div}\tilde{\boldsymbol{\sigma}}\,\mathrm{d}x\\
&= \int_{\mathcal{O}} (\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\cdot\operatorname{div}\left(\int_{\mathcal{O}} (\boldsymbol{\sigma}_D(y) - \boldsymbol{\sigma}^s_D(y))\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y)\,\mathrm{d}y\right)\mathrm{d}x\\
&+ \int_{\mathcal{O}} (\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\cdot\nabla\left(\int_{\mathcal{O}} \frac{1}{d}(\operatorname{tr}\boldsymbol{\sigma}(y) - \operatorname{tr}\boldsymbol{\sigma}^s(y))\eta_\varepsilon(x-\alpha\varepsilon\boldsymbol{n}-y)\,\mathrm{d}y\right)\mathrm{d}x\\
&= \int_{\mathcal{O}} (\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\cdot\operatorname{div}\left(\int_{\mathcal{O}} (\boldsymbol{\sigma}(y) - \boldsymbol{\sigma}^s(y))\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y)\,\mathrm{d}y\right)\mathrm{d}x\\
&+ \int_{\mathcal{O}} (\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\cdot\\
&\quad\cdot\nabla\left(\int_{\mathcal{O}} \frac{1}{d}(\operatorname{tr}\boldsymbol{\sigma}(y) - \operatorname{tr}\boldsymbol{\sigma}^s(y))(\eta_\varepsilon(x-\alpha\varepsilon\boldsymbol{n}-y) - \eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\,\mathrm{d}y\right)\mathrm{d}x\\
&= \int_{\mathcal{O}}\int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\otimes\nabla_x\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\cdot(\boldsymbol{\sigma}(y) - \boldsymbol{\sigma}^s(y))\,\mathrm{d}y\,\mathrm{d}x\\
&- \frac{1}{d}\int_{\mathcal{O}}\int_{\mathcal{O}} \operatorname{div}((\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x))(\operatorname{tr}\boldsymbol{\sigma}(y) - \operatorname{tr}\boldsymbol{\sigma}^s(y))\\
&\quad(\eta_\varepsilon(x-\alpha\varepsilon\boldsymbol{n}-y) - \eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\,\mathrm{d}y\,\mathrm{d}x\\
&= -\int_{\mathcal{O}}\int_{\mathcal{O}} ((\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x)\otimes\nabla_y\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\cdot(\boldsymbol{\sigma}(y) - \boldsymbol{\sigma}^s(y))\,\mathrm{d}y\,\mathrm{d}x\\
&- \frac{1}{d}\int_{\mathcal{O}} \operatorname{div}((\dot{\boldsymbol{u}}(x) - \dot{\boldsymbol{u}}_0(x))\tau(x))\\
&\quad\operatorname{tr}(\boldsymbol{\sigma}^\varepsilon(x-\alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}^\varepsilon(x+\alpha\varepsilon\boldsymbol{n}) - \boldsymbol{\sigma}^{\varepsilon,s}(x-\alpha\varepsilon\boldsymbol{n}) + \boldsymbol{\sigma}^{\varepsilon,s}(x+\alpha\varepsilon\boldsymbol{n}))\,\mathrm{d}x.
\end{aligned}$$

Therefore,

$$
\begin{aligned}
&\lim_{\varepsilon\to 0_+}\int_{\mathcal{O}} (\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\tau\cdot \operatorname{div}\tilde{\boldsymbol{\sigma}}\,\mathrm{d}x\\
&=-\lim_{\varepsilon\to 0_+}\int_{\mathcal{O}}\int_{\mathcal{O}}((\dot{\boldsymbol{u}}(x)-\dot{\boldsymbol{u}}_0(x))\tau(x)\otimes\nabla_y\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\cdot\\
&\quad\cdot(\boldsymbol{\sigma}(y)-\boldsymbol{\sigma}^s(y))\,\mathrm{d}y\,\mathrm{d}x\\
&=-\lim_{\varepsilon\to 0_+}\int_{\mathcal{O}}\int_{\partial\mathcal{O}_M}((\dot{\boldsymbol{u}}(x)-\dot{\boldsymbol{u}}_0(x))\tau(x)\cdot\\
&\quad\cdot(\boldsymbol{\sigma}(y)-\boldsymbol{\sigma}^s(y))\boldsymbol{n}\eta_\varepsilon(x+\alpha\varepsilon\boldsymbol{n}-y))\,\mathrm{d}S(y)\,\mathrm{d}x\\
&=-\lim_{\varepsilon\to 0_+}\int_{\partial\mathcal{O}_M}((\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\tau)^\varepsilon(y-\alpha\varepsilon\boldsymbol{n})\cdot\\
&\quad\cdot(\boldsymbol{\sigma}(y)-\boldsymbol{\sigma}^s(y))\boldsymbol{n}\,\mathrm{d}S(y)\\
&=-\lim_{\varepsilon\to 0_+}\int_{\partial\mathcal{O}_M}((\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\tau)^\varepsilon(y-\alpha\varepsilon\boldsymbol{n})\cdot\boldsymbol{n}((\boldsymbol{\sigma}(y)-\boldsymbol{\sigma}^s(y))\cdot\boldsymbol{n}\otimes\boldsymbol{n}))\,\mathrm{d}S(y)=0.
\end{aligned}
$$

Hence we get $\lambda(\partial\mathcal{O}_M)=0$. Note that close to boundaries of $\partial\mathcal{O}_D$, $\partial\mathcal{O}_N$ and $\partial\mathcal{O}_M$ we have to do it more carefully and we refer to [5] for more details, where very similar problem with the symmetric gradient is treated.

3.9. **Identification of plastic strain.** First, let us recall that $\lambda=0$ on $\partial\mathcal{O}_N\cup\partial\mathcal{O}_M$ an consequently $\dot{\boldsymbol{e}}_{\mathrm{p}}=0$ on this set. For the rest of the boundary, we can use (3.82) to directly deduce (letting $\tau\to\chi_{\partial\mathcal{O}_D}$) that

$$\dot{\boldsymbol{e}}_{\mathrm{p}}=\frac{1}{2}\left((\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\otimes\boldsymbol{n}+\boldsymbol{n}\otimes(\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0)\right)\,\mathrm{d}S$$

and consequently, we also have that

$$\lambda=\frac{|\dot{\boldsymbol{u}}-\dot{\boldsymbol{u}}_0|}{\sqrt{2}}\,\mathrm{d}S.$$

Finally, we investigate the behaviour of $\dot{\boldsymbol{e}}_{\mathrm{p}}$ in $\mathcal{O}$. Notice that since we know that for almost all time $t\in(0,T)$

$$\boldsymbol{\sigma}_D^\varepsilon\to\boldsymbol{\sigma}_p \qquad \text{strongly in } L^1((0,T)\times\mathcal{O};\mathbb{R}^{d\times d}\,\mathrm{d}\lambda)$$

then due to the Luzin theorem, we also know that $\boldsymbol{\sigma}_D=\boldsymbol{\sigma}_p$ almost everywhere (in the sense of Lebesgue measure) in $\mathcal{O}$. We know refine this result and show that the exceptional set where it is not equal has the Hausdorff dimension less or equal to $(d-2)$. Moreover, we identify such a set also as a set where the displacment blows-up, which in addition will lead in dimension two that we can set everywhere

$$\boldsymbol{\sigma}_p=\boldsymbol{\sigma}_D. \tag{3.88}$$

We now proceed slightly formally, and shall work on almost every time level $t\in(0,T)$. Let us define the reduced maximal function $M_p^\delta f$ for any $\delta>0$ and $p\le d$ as

$$M_p^\delta f(x):=\sup_{r\in(0,\delta)} r^{-p}\int_{B_r(x)}|f(y)|\,\mathrm{d}y. \tag{3.89}$$

Then we have for arbitrary $p \in (d-2, d)$ the estimate

$$
\begin{aligned}
|\boldsymbol{\sigma}^{\varepsilon}(x) - \boldsymbol{\sigma}(x)| &= \left| \int_0^{\varepsilon} \frac{d}{dr} \frac{1}{|B_r(x)|} \int_{B_r(x)} \boldsymbol{\sigma}(y) \, \mathrm{d}y \, \mathrm{d}r \right| \\
&= \left| \int_0^{\varepsilon} \frac{1}{r|B_r(x)|} \int_{B_r(x)} \nabla \boldsymbol{\sigma}(y)(x-y) \, \mathrm{d}y \, \mathrm{d}r \right| \\
&\le C(d) \int_0^{\varepsilon} r^{-d} \int_{B_r(x)} |\nabla \boldsymbol{\sigma}(y)| \, \mathrm{d}y \, \mathrm{d}r \\
&\le C(d) \int_0^{\varepsilon} r^{\frac{p-d}{2}} \left(\int_{B_r(x)} \frac{|\nabla \boldsymbol{\sigma}(y)|^2}{r^p} \, \mathrm{d}y \right)^{\frac{1}{2}} \mathrm{d}r \\
&\le \frac{2C(d)}{p-d+2} \left(M_p^{\varepsilon} |\nabla \boldsymbol{\sigma}|^2(x) \right)^{\frac{1}{2}} \varepsilon^{\frac{p-(d-2)}{2}}
\end{aligned} \tag{3.90}
$$

Let us take arbitrary closed $\mathcal{O}_0 \subset\subset \mathcal{O}$ and define $2\delta_0 := \operatorname{dist}(\partial \mathcal{O}, \partial \mathcal{O}_0)$. Then, for arbitrary $\delta \le \delta_0$ we set

$$
\mathcal{O}^{\delta}_{\boldsymbol{\sigma}}(t) := \{ x \in \mathcal{O}_0 : \ M^{\delta}_{d-2+\delta} |\nabla \boldsymbol{\sigma}(t)|^2(x) > \delta^{-1} \}.
$$

Further, we know that for any $\delta \in (0, \delta_0)$ the set $\mathcal{O}^{\delta}_{\boldsymbol{\sigma}}$ is open, and consequently $\mathcal{O}_0 \setminus \mathcal{O}^{\delta}$ is closed and so λ-measurable. Hence combining (3.87) and (3.90) we easily deduce that for all $x \in \mathcal{O}_0 \setminus \mathcal{O}^{\delta}_{\boldsymbol{\sigma}}$

$$
\boldsymbol{\sigma}_D(x) = \boldsymbol{\sigma}_p(x). \tag{3.91}
$$

Notice that from this relation follows that the set where (3.91) does not hold has the Hausdorff dimension at most $(d-2)$.

Next, we identify $\boldsymbol{\sigma}_p$ also on the set, where we control displacement in a sufficient manner. Let us define for arbitrary $\varepsilon > 0$ the set

$$
\mathcal{O}^{\varepsilon}_{\boldsymbol{v}}(t) := \left\{ x \in \mathcal{O}_0; \ M^{\varepsilon}_{d-1-\frac{\varepsilon}{2}} |\dot{\boldsymbol{u}}(t)|^{1+\varepsilon} > \varepsilon^{-1} \right\}.
$$

We shall show that for any $\varepsilon > 0$

$$
\lim_{\delta \to 0_+} \lambda(\mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}) = 0, \tag{3.92}
$$

from which one can deduce the validity of (3.91) λ-almost everywhere in $\mathcal{O}_0 \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}(t)$.

Indeed, first using the definition of $\mathcal{O}^{\delta}_{\boldsymbol{\sigma}}$ we can for any $x \in \mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}$ find R_x such that

$$
\mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}} \subset \bigcup_{x \in \mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}} B_{R_x}(x),
$$

where $R_x \in (0, \delta]$ and

$$
\int_{B_{R_x}(x)} |\nabla \boldsymbol{\sigma}|^2 \ge \delta^{-1} R_x^{d-2+\delta}.
$$

Hence, the Vitali covering theorem, we can find a countable union of balls $\{B_{R_i}(x_i)\}_{i=1}^{\infty} \subset \mathbb{R}^d$ with $\{x_i\}_{i=1}^{\infty} \subset \mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}$ such that

$$
\mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}} \subset \bigcup_{i=1}^{\infty} B_{R_i}(x_i), \qquad B_{R_i/4}(x_i) \cap B_{R_j/4}(x_j) = \emptyset, \qquad R_i \in (0, \delta],
$$

$$
\int_{B_{R_i}(x_i)} |\nabla \boldsymbol{\sigma}|^2 \ge R_i^{d-2+\delta} \delta^{-1}.
$$

Immediately, we observe that

$$\sum_{i=1}^{\infty} R_i^{d-2+\delta} \le C\delta \int_{\mathcal{O}_0} |\nabla \boldsymbol{\sigma}|^2 \le C(\mathcal{O}_0)\delta. \tag{3.93}$$

Then, using the estimate (3.77) we find that (denoting for simplicity $B_i := B_{R_i}(x_i)$)

$$\begin{aligned}
\lambda(\mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}) &\le \sum_{i=1}^{\infty} \lambda(B_{R_1}(x_i)) \\
&\le C \sum_{i=1}^{\infty} \int_{2B_i} (1 + |\dot{\boldsymbol{\sigma}}(t)| + |\boldsymbol{\sigma}(t)| + |\operatorname{div} \dot{\boldsymbol{u}}(t)|)(1 + |\boldsymbol{\sigma}(t)|)\, \mathrm{dx} \\
&\quad + C(\varepsilon) \sum_{i=1}^{\infty} R_i^{d-1} \left(\int_{2B_i} \frac{|\dot{\boldsymbol{u}}(t)|^{1+\varepsilon}}{R^d}\, \mathrm{dx} \right)^{\frac{1}{1+\varepsilon}} \\
&\le C \int_{\bigcup_{i=1}^{\infty} 2B_i} (1 + |\dot{\boldsymbol{\sigma}}(t)| + |\boldsymbol{\sigma}(t)| + |\operatorname{div} \dot{\boldsymbol{u}}(t)|)(1 + |\boldsymbol{\sigma}(t)|)\, \mathrm{dx} \\
&\quad + C(\varepsilon) \sum_{i=1}^{\infty} R_i^{d-1-\frac{1+\frac{\varepsilon}{2}}{1+\varepsilon}} \left(\int_{2B_i} \frac{|\dot{\boldsymbol{u}}(t)|^{1+\varepsilon}}{R^{d-1-\frac{\varepsilon}{2}}}\, \mathrm{dx} \right)^{\frac{1}{1+\varepsilon}},
\end{aligned} \tag{3.94}$$

where for the second inequality we used the properties of the Vitali covering. Next, due to (3.37), (3.41) and (3.43), we can estimate the first term on the right hand side of (3.94) as

$$\int_{\bigcup_{i=1}^{\infty} 2B_i} (1 + |\dot{\boldsymbol{\sigma}}(t)| + |\boldsymbol{\sigma}(t)| + |\operatorname{div} \dot{\boldsymbol{u}}(t)|)(1 + |\boldsymbol{\sigma}(t)|)\, \mathrm{dx} \le C \left| \bigcup_{i=1}^{\infty} 2B_i \right| \overset{\delta \to 0}{\to} 0,$$

where the last convergence result follows from (3.93). Next, for the remaining term on the right hand side of (3.94), we use the fact that for all $i \in \mathbb{N}$ we have that $x_i \notin \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}$. Consequently, for all $\delta \le \varepsilon/2(1+\varepsilon)$, we deduce with the help of the definition of $\mathcal{O}^{\varepsilon}_{\boldsymbol{v}}(t)$ that

$$\begin{aligned}
C(\varepsilon) &\sum_{i=1}^{\infty} R_i^{d-1-\frac{1+\frac{\varepsilon}{2}}{1+\varepsilon}} \left(\int_{2B_i} \frac{|\dot{\boldsymbol{u}}(t)|^{1+\varepsilon}}{R^{d-1-\frac{\varepsilon}{2}}}\, \mathrm{dx} \right)^{\frac{1}{1+\varepsilon}} \\
&\le C(\varepsilon) \sum_{i=1}^{\infty} R_i^{d-2+\frac{\frac{\varepsilon}{2}}{1+\varepsilon}} \left(M^{\delta}_{d-1-\frac{\varepsilon}{2}} |\dot{\boldsymbol{u}}(t)|^{1+\varepsilon} \right)^{\frac{1}{1+\varepsilon}} \\
&\le C(\varepsilon) \sum_{i=1}^{\infty} R_i^{d-2+\delta} \overset{\delta \to 0}{\to} 0,
\end{aligned}$$

where the last convergence result follows from (3.93). Hence, we deduce from (3.94) that

$$\lambda(\mathcal{O}^{\delta}_{\boldsymbol{\sigma}} \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}) \to 0 \qquad \text{as } \delta \to 0. \tag{3.95}$$

Since, we already know that $\boldsymbol{\sigma}_p = \boldsymbol{\sigma}_D$ in $\mathcal{O}_0 \setminus \mathcal{O}^{\delta}_{\boldsymbol{\sigma}}$ for all $\delta > 0$ then it follows from (3.95) that the same relation holds λ-everywhere in $\mathcal{O}_0 \setminus \mathcal{O}^{\varepsilon}_{\boldsymbol{v}}$ for all $\varepsilon > 0$. Consequently, we can identify the possible set, where the desired inequality is not true as follows

$$\{(t,y):\ \boldsymbol{\sigma}_p(t,y)\neq\boldsymbol{\sigma}_p(t,y)\}\subset$$
$$\left\{(t,y):\ \liminf_{\delta\to 0}\sup_{R\in(0,\delta)}\delta\int_{B_R(y)}R^{2-d-\delta}|\nabla\boldsymbol{\sigma}(t,x)|^2\,\mathrm{dx}\geq 1\right.$$
$$\left.\text{and }\liminf_{\varepsilon\to 0}\sup_{R\in(0,\varepsilon)}\varepsilon\int_{B_R(y)}R^{1-d-\varepsilon/2}|\dot{\boldsymbol{u}}(t,x)|^{1+\varepsilon}\,\mathrm{dx}\geq 1\right\}. \tag{3.96}$$

Finally, it is not difficult to observe by using the Hölder inequality that in case, when $\dot{\boldsymbol{u}}(t)\in L^q(\mathcal{O})$ for some $q>d$ then the identity (3.88) holds $\lambda(t)$-everywhere in $\mathcal{O}$, which completes the proof of Theorem 2.1.

References

[1] G. Anzellotti and M. Giaquinta. Existence of the Displacements Field for an Elasto-Plastic Body subject to Henck's Law and von Mises Yield Condition. *manuscripta mathematica*, 32:101–136, 1980.

[2] G. Anzellotti and M. Giaquinta. On the Existence of the Fields of Stresses and Displacements for an Elasto-Perfectly Plastic Body in Static Equilibrium. *Journal de Mathématiques Pures et Appliquées*, 61:219–244, 1982.

[3] Gabriele Anzellotti. On the existence of the rates of stress and displacement for Prandtl-Reuss plasticity. *Quart. Appl. Math.*, 41(2):181–208, 1983/84.

[4] Gabriele Anzellotti. On the extremal stress and displacement in Hencky plasticity. *Duke Math. J.*, 51(1):133–147, 1984.

[5] L. Beck, M. Bulíček, J. Málek, and E. Süli. On the existence of integrable solutions to nonlinear elliptic systems and variational problems with linear growth. Preprint MORE/2016/01, 2016.

[6] M. Bulíček, J. Frehse, and J. Málek. On boundary regularity for the stress in problems of linearized elasto-plasticity. *Int. J. Adv. Eng. Sci. Appl. Math.*, 1(4):141–156, 2009.

[7] E. Feireisl and A. Novotný. *Singular limits in thermodynamics of viscous fluids*. Advances in Mathematical Fluid Mechanics. Birkhäuser Verlag, Basel, 2009.

[8] Jens Frehse and Sebastian Schwarzacher. On regularity of the time derivative for degenerate parabolic systems. *SIAM J. Math. Anal.*, 47(5):3917–3943, 2015.

[9] Jens Frehse and Maria Specovius-Neugebauer. Fractional differentiability for the stress velocities to the solution of the Prandtl-Reuss problem. *ZAMM Z. Angew. Math. Mech.*, 92(2):113–123, 2012.

[10] R. M. Hardt and D. Kinderlehrer. Some regularity results in plasticity. In *Geometric measure theory and the calculus of variations (Arcata, Calif., 1984)*, volume 44 of *Proc. Sympos. Pure Math.*, pages 239–244. Amer. Math. Soc., Providence, RI, 1986.

[11] Robert Hardt and David Kinderlehrer. Elastic plastic deformation. *Appl. Math. Optim.*, 10(3):203–246, 1983.

[12] H. Hencky. Zur Theorie Plastischer Deformationen. *Z. Angew. Math. Mech.*, 4:323–334, 1924.

[13] J. Lubliner. *Plasticity Theory.* Macmilan Publishing Company, New York, 1990.

[14] L. Prandtl. Spannungsverteilung in plastischen körpen. In *Proceeding of the First Int. Congr. Appl. Mech.*, pages 43–54, Delft, 1924.

[15] A. Reuss. Berücksichtigung der elastischen formänderung in der plastizitätstheorie. *Z. Angenew. Math. Mech.*, 10:266–271, 1930.

[16] M. Steinhauer. *On analysis of some nonlinear systems of partial differential equations of continuum mechanics*. Bonner Mathematische Schriften [Bonn Mathematical Publications], 359. Universität Bonn Mathematisches Institut, Bonn, 2003. Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, 2002.

[17] R. Temam. *Mathematical Problems in Plasticity.* Gauthier-Villars/Bordas, Paris, 1985.

[18] Roger Temam. A generalized Norton-Hoff model and the Prandtl-Reuss law of plasticity. *Arch. Rational Mech. Anal.*, 95(2):137–183, 1986.

Mathematical Institute of Charles University, Sokolovská 83, 186 75 Prague, Czech Republic
E-mail address: mbul8060@karlin.mff.cuni.cz

Institute of Applied Mathematics, Endenicher Allee 60, D-53121 Bonn, Germany
E-mail address: erdbeere@iam.uni-bonn.de