REGULARITY RESULTS FOR TWO STANDARD MODELS IN ELASTO-PERFECT-PLASTICITY THEORY WITH HARDENING

MIROSLAV BULÍČEK, JENS FREHSE, AND MARIA SPECOVIUS–NEUGEBAUER

Dedicated to Umberto Mosco

ABSTRACT. We consider two most studied standard models in the theory of elasto-plasticity with hardening in arbitrary dimension $d \geq 2$, namely, the kinematic hardening and the isotropic hardening problem. While the existence and uniqueness of the solution is very well known, the optimal regularity up to the boundary remains an open problem. Here, we show that in the interior we have Sobolev regularity for the stress and hardening while for their time derivatives we have the "half" derivative with the spatial and time variable. This was well known for the limiting problem but we show that these estimates are uniform and independent of the order of approximation. The main novelty consist of estimates near the boundary. We show that for the stress and the hardening parameter, we control tangential derivative in the Lebesgue space L^2, and for time derivative of the stress and the hardening we control the "half" time derivative and also spatial tangential derivative. Last, for the normal derivative, we show that the stress and the hardening have the 3/5 derivative with respect to the normal and for the time derivative of the stress and the hardening we show they have the 1/5 derivative with respect to the normal direction, provided we consider the kinematic hardening or near the Dirichlet boundary. These estimates are independent of dimension. In case, we consider the isotropic hardening near the Neumann boundary we shall obtain $W^{\alpha,2}$ regularity for the stress and the hardening with some $\alpha > 1/2$ depending on the dimension and $W^{\beta,2}$ with some $\beta > 1/6$ for the time derivative of the stress and the hardening. Finally, in case of kinematic hardening the same regularity estimate holds true also for the velocity gradient.

1. INTRODUCTION

In this paper we deal with the regularity estimates for solutions to some models of linearized elasto–plasticity with hardening. We have mainly two cases in mind, the isotropic hardening and the kinematic harding. Our main goal is to provide the uniform estimates on the Cauchy stress and the hardening parameter and their time derivatives in fractional Sobolev spaces up to the boundary and consequently by using an interpolation technique also to improve the available regularity results for the small strain tensor.

1.1. **Physical background.** 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, the current and the preferred (natural) configurations coincide and we can approximate the strain tensor by the linearized strain tensor $\boldsymbol{\varepsilon}(\boldsymbol{u})$, which is defined as

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

where $\boldsymbol{u} : (0, T) \times \mathcal{O} \to \mathbb{R}^d$ is the displacement and the interval $(0, T)$ represents the loading parameter, which we call "time" in what follows. 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

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

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 $\boldsymbol{\varepsilon}(\boldsymbol{u})$. Since we deal with elasto–plastic effects, 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.,

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

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 strictly convex function vanishing at zero and exploding at infinity and the elastic strain is related to the stress through

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

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

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

Furthermore, we require in the paper that there exists a constant fourth order tensor $\mathbf{A} \in \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^{d \times d}_{sym}$ such that for all $\boldsymbol{\sigma} \in \mathbb{R}^{d \times d}_{sym}$

$$(1.5) \qquad \mathbf{A} \equiv \frac{\partial^2 \psi(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma} \partial \boldsymbol{\sigma}}.$$

Then, evidently, the relation (1.4) can be rewritten as

$$(1.6) \qquad \boldsymbol{\sigma} = \mathbf{A}^{-1} \boldsymbol{e}_{\mathrm{el}} \quad \Leftrightarrow \quad \boldsymbol{e}_{\mathrm{el}} = \mathbf{A} \boldsymbol{\sigma}.$$

Note that thanks to (1.5), the tensor $\mathbf{A}$ is symmetric. In addition, since ψ is assumed to be convex, we certainly know that $\mathbf{A}$ is invertible and therefore (1.6) makes good sense.

Concerning the plastic strain, we assume that (usually $\boldsymbol{e}_{\mathrm{p}}$ is considered to be relevant to incompressible motion)

$$(1.7) \qquad \operatorname{tr} \boldsymbol{e}_{\mathrm{p}} = 0,$$

where tr denotes the trace of $\boldsymbol{e}_{\mathrm{p}}$. Further, we need to specify under which conditions it may appear and how is related to the "hardening". In the paper, we shall assume two cases, the **isotropic hardening**, which is described by a scalar function

$$\xi : [0, T] \times \mathcal{O} \to \mathbb{R},$$

which is related to the plastic strain by the following flow rule (here H is a given positive constant)

$$(1.8) \qquad H\dot{\xi} = |\dot{\boldsymbol{e}}_{\mathrm{p}}|.$$

Here, and also in what follows, we use the "dot" to abbreviate the partial derivative with respect to the variable t, i.e.,

$$\dot{z} := \frac{\partial z}{\partial t}$$

for any function z depending on t. Furthermore, we require the von Mises condition

$$|\boldsymbol{\sigma}_D| - \xi \le \kappa,$$

and we assume that there is no plastic strain if the above inequality is strict, i.e.,

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

On the other hand, if $|\boldsymbol{\sigma}_D| - \xi = \kappa$, we require that

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

To summarize, we can write these conditions in a more compact form

$$(1.9) \quad \dot{\boldsymbol{e}}_{\mathrm{p}} = \dot{\lambda}\frac{\boldsymbol{\sigma}_D}{|\boldsymbol{\sigma}_D|} \quad \text{with} \quad \dot{\lambda} \ge 0, \quad |\boldsymbol{\sigma}_D| - \xi \le \kappa \quad \text{and} \quad \dot{\lambda}\left(|\boldsymbol{\sigma}_D| - \xi - \kappa\right) = 0.$$

In a similar way, we shall define the **kinematic hardening**. Hence, we assume that the hardening parameter

$$\boldsymbol{\xi} : [0,T] \times \mathcal{O} \to \mathbb{R}_{sym}^{d \times d}$$

obeys the following flow rule (here H is the symmetric positively definite fourth order tensor)

$$(1.10) \qquad \mathsf{H}\dot{\boldsymbol{\xi}} = \dot{\boldsymbol{e}}_{\mathrm{p}}.$$

The related von Mises condition then reads as

$$|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| \le \kappa,$$

and we assume that there is no plastic strain if the above inequality is strict, i.e.,

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

On the other hand, if $|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| = \kappa$, we require that

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

Again, the above conditions can be equivalently rewritten as

$$(1.11)$$

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

We do not claim that two models introduced above are the only ones of physical and engineering interest. Indeed, there is a lot of models describing elastic and plastic deformation with memory effects, see for example the book [1], and each of them can be used in a specific situation. The models of kinematic and isotopic hardening are just two most prototypic examples. On the other hand, it seems that the models discussed in this paper usually have better regularity properties than the others, and therefore it was also our motivation to focus on them.

1.2. **Weak formulation of the problem and the main result.** To complete the problem, we have to specify the boundary and initial values. We start with the hardening variable and we assume that $\boldsymbol{\xi}(0, x) \equiv 0$ in case of kinematic hardening and that $\xi(0, x) = 0$ in case of isotropic hardening. Next, we assume that the boundary of the domain can be decomposed onto two smooth sets $\partial \mathcal{O}_D$ - the Dirichlet part, and $\partial \mathcal{O}_N$ - the Neumann part with the unit normal outward vector denoted by $\boldsymbol{n}$. Finally, we prescribe the data $\boldsymbol{u}_0 : [0, T] \times \mathcal{O} \to \mathbb{R}^d$ and $\boldsymbol{\sigma}_0 : [0, T] \times \mathcal{O} \to \mathbb{R}^{d \times d}_{sym}$, and we require that $\boldsymbol{u} = \boldsymbol{u}_0$ on $[0, T] \times \partial \mathcal{O}_D$ (Dirichlet data for the displacement - the displacement on the part $\partial \mathcal{O}_D$) and that $\boldsymbol{u}(0) = \boldsymbol{u}_0(0)$ (the initial displacement), and that $\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{\sigma}_0 \cdot \boldsymbol{n}$ on $[0, T] \times \mathcal{O}_N$ (the Neumann data - the traction on $\partial \mathcal{O}_N$) and $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0(0)$ (the initial value of the Cauchy stress). Note here, that there is a necessary compatibility condition $\boldsymbol{\varepsilon}(\boldsymbol{u}_0(0)) = \mathsf{A}\boldsymbol{\sigma}_0(0)$. To recall all data, we also have a given body forces $\boldsymbol{f} : [0, T] \times \mathcal{O} \to \mathbb{R}^d$ and the material parameters[1] $\kappa \geq C_1 > 0$ for some positive constant C_1, positively definite symmetric fourth order tensor A and the constant $H > 0$ in case of isotropic harding and the constant fourth order symmetric positively definite tensor H in case of kinematic hardening. Here, positively definite means that there exists a positive constant C_1 such that $C_1 \leq H$ and such that for all $\boldsymbol{\tau} \in \mathbb{R}^{d \times d}_{sym}$ there holds

$$(1.12) \qquad C_1 |\boldsymbol{\tau}|^2 \leq \mathsf{A}\boldsymbol{\tau} \cdot \boldsymbol{\tau} \leq C_1^{-1} |\boldsymbol{\tau}|^2 \quad \text{and} \quad C_1 |\boldsymbol{\tau}|^2 \leq \mathsf{H}\boldsymbol{\tau} \cdot \boldsymbol{\tau} \leq C_1^{-1} |\boldsymbol{\tau}|^2.$$

Finally, we can summarize the above description and formulate the problem of elasto-plastic hardening in the following classical way:

Kinematic hardening: We look for a quintuple $(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{u}, \boldsymbol{e}_{\mathrm{el}}, \boldsymbol{e}_{\mathrm{p}}) : [0, T] \times \mathcal{O} \to \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^d \times \mathbb{R}^{d \times d}_{sym} \times \mathbb{R}^{d \times d}_{sym}$ such that

$$
\begin{aligned}
-\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{f}, \quad \mathsf{H}\dot{\boldsymbol{\xi}} = \dot{\boldsymbol{e}}_{\mathrm{p}}, \quad \boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{e}_{\mathrm{el}} + \boldsymbol{e}_{\mathrm{p}}, \quad \boldsymbol{e}_{\mathrm{el}} = \mathsf{A}\boldsymbol{\sigma} \quad &\text{in } [0, T] \times \mathcal{O}, \\
\dot{\boldsymbol{e}}_{\mathrm{p}} = |\dot{\boldsymbol{e}}_{\mathrm{p}}| \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{\kappa} \quad &\text{in } [0, T] \times \mathcal{O}, \\
(1.13) \qquad \kappa \geq |\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| \text{ and } |\dot{\boldsymbol{e}}_{\mathrm{p}}|(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - \kappa) = 0 \quad &\text{in } [0, T] \times \mathcal{O}, \\
\boldsymbol{u} = \boldsymbol{u}_0 \quad &\text{on } [0, T] \times \partial \mathcal{O}_D, \\
\boldsymbol{\sigma}\boldsymbol{n} = \boldsymbol{\sigma}_0 \boldsymbol{n} \quad &\text{on } [0, T] \times \partial \mathcal{O}_N, \\
\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0, \quad \boldsymbol{\xi}(0) = 0, \quad \boldsymbol{u}(0) = \boldsymbol{u}_0(0) \quad &\text{in } \mathcal{O}.
\end{aligned}
$$

where $T > 0$ is the given length of the time[2] interval, the given threshold $\kappa > 0$ is a von Mises condition, $\boldsymbol{f} : [0, T] \times \mathcal{O} \to \mathbb{R}^d$ are the given volume forces, the stress $\boldsymbol{\sigma}_0 : [0, T] \times \mathcal{O} \to \mathbb{R}^{d \times d}_{sym}$ represents the initial value $\boldsymbol{\sigma}(0)$ and the traction $\boldsymbol{\sigma}\boldsymbol{n}$ and the prescribed displacement on the boundary $[0, T] \times \partial \mathcal{O}_D$ 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 $\partial \mathcal{O}$.

---

[1] For simplicity, we assume that the material parameters are constant. Nevertheless, we could allow them to be time dependent.

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

Isotropic hardening: We look for a quintuple $(\boldsymbol{\sigma}, \xi, \boldsymbol{u}, \boldsymbol{e}_{\mathrm{el}}, \boldsymbol{e}_{\mathrm{p}}) : [0,T] \times \mathcal{O} \to \mathbb{R}_{sym}^{d \times d} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}_{sym}^{d \times d} \times \mathbb{R}_{sym}^{d \times d}$ such that

(1.14)

$$-\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{f}, \quad H\dot{\xi} = |\dot{\boldsymbol{e}}_{\mathrm{p}}|, \quad \boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{e}_{\mathrm{el}} + \boldsymbol{e}_{\mathrm{p}}, \quad \boldsymbol{e}_{\mathrm{el}} = \mathbf{A}\boldsymbol{\sigma} \qquad \text{in } [0,T] \times \mathcal{O},$$

$$\dot{\boldsymbol{e}}_{\mathrm{p}} = |\dot{\boldsymbol{e}}_{\mathrm{p}}| \frac{\boldsymbol{\sigma}_D}{\kappa + \xi} \qquad\qquad\qquad \text{in } [0,T] \times \mathcal{O},$$

$$\kappa \geq |\boldsymbol{\sigma}_D| - \xi \ \text{ and } \ |\dot{\boldsymbol{e}}_{\mathrm{p}}|(|\boldsymbol{\sigma}_D| - \kappa - \xi) = 0 \qquad \text{in } [0,T] \times \mathcal{O},$$

$$\boldsymbol{u} = \boldsymbol{u}_0 \qquad\qquad\qquad \text{on } [0,T] \times \partial\mathcal{O}_D,$$

$$\boldsymbol{\sigma}\boldsymbol{n} = \boldsymbol{\sigma}_0\boldsymbol{n} \qquad\qquad\qquad \text{on } [0,T] \times \partial\mathcal{O}_N,$$

$$\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0, \quad \xi(0) = 0, \quad \boldsymbol{u}(0) = \boldsymbol{u}_0(0) \qquad \text{in } \mathcal{O}.$$

Furthermore, to simplify the notation, we require that (indeed, it is just simplification of a notation, in fact, to be able to solve the problem, the existence of $\boldsymbol{\sigma}_0$ fulfilling this equation is necessary)

$$\operatorname{div}\boldsymbol{\sigma}_0 = \boldsymbol{f} \quad \text{almost everywhere in } (0,T) \times \mathcal{O}.$$

Then, we choose a proper subspace of the Sobolev space $W^{1,2}(\mathcal{O}; \mathbb{R}^d)$, which will be used in what follows

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

and we define the set of admissible stresses as

$$\mathcal{F}_k(t) := \Big\{(\boldsymbol{\sigma}, \boldsymbol{\xi}) \in L^2(\Omega; \mathbb{R}_{sym}^{d \times d}) \times L^2(\Omega; \mathbb{R}_{sym}^{d \times d}); \ |\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| \leq \kappa,$$

$$\text{and for all } \boldsymbol{v} \in \mathcal{V} \text{ there holds } \int_{\mathcal{O}} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}) \, \mathrm{d}x = 0\Big\}$$

and

$$\mathcal{F}_i(t) := \Big\{(\boldsymbol{\sigma}, \xi) \in L^2(\Omega; \mathbb{R}_{sym}^{d \times d}) \times L^2(\Omega; \mathbb{R}); \ |\boldsymbol{\sigma}_D| \leq \kappa + \xi,$$

$$\text{and for all } \boldsymbol{v} \in \mathcal{V} \text{ there holds } \int_{\mathcal{O}} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}) \, \mathrm{d}x = 0\Big\}.$$

Notice here, that $\mathcal{F}_i$ corresponds to isotropic hardening while the set $\mathcal{F}_k$ is related to kinematic hardening and we can introduce the following definitions.

Definition 1.1 (Kinematic hardening). *Let $\mathcal{O} \subset \mathbb{R}^d$ be a Lipschitz domain. Assume that $\boldsymbol{\sigma}_0 \in W^{1,2}(0,T; L^2(\Omega; \mathbb{R}_{sym}^{d \times d}))$ and $\boldsymbol{u}_0 \in W^{1,2}(0,T; W^{1,2}(\mathcal{O}; \mathbb{R}^d))$. We say that $(\boldsymbol{\sigma}, \boldsymbol{\xi}) \in W^{1,2}(0,T; L^2(\mathcal{O}; \mathbb{R}_{sym}^{d \times d})) \times W^{1,2}(0,T; L^2(\mathcal{O}; \mathbb{R}_{sym}^{d \times d}))$ is a weak solution to (1.13) if $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0$, $\boldsymbol{\xi}(0) = 0$ and for almost all $t \in (0,T)$ there holds $(\boldsymbol{\sigma}(t), \boldsymbol{\xi}(t)) \in \mathcal{F}_k(t)$ and, in addition, we require that for almost all $t \in (0,T)$ and all $(\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\xi}}) \in \mathcal{F}_k(t)$ there holds*

(1.15)
$$\int_{\mathcal{O}} \mathbf{A}\dot{\boldsymbol{\sigma}}(t) \cdot (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}})) + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot (\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}}) \, \mathrm{d}x \leq \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}}) \, \mathrm{d}x.$$

In a very similar way, we can also introduce the notion of a weak solution to the isotropic model (1.14), where we shall replace $\mathcal{F}_k(t)$ by $\mathcal{F}_i(t)$ in a natural way.

Definition 1.2 (Isotropic hardening). *Let $\mathcal{O} \subset \mathbb{R}^d$ be a Lipschitz domain. Assume that $\boldsymbol{\sigma}_0 \in W^{1,2}(0,T; L^2(\Omega; \mathbb{R}_{sym}^{d \times d}))$ and $\boldsymbol{u}_0 \in W^{1,2}(0,T; W^{1,2}(\mathcal{O}; \mathbb{R}^d))$. We say that $(\boldsymbol{\sigma}, \xi) \in W^{1,2}(0,T; L^2(\mathcal{O}; \mathbb{R}_{sym}^{d \times d})) \times W^{1,2}(0,T; L^2(\mathcal{O}; \mathbb{R})$ is a weak solution to (1.14) if $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0$, $\xi(0) = 0$ and for almost all $t \in (0,T)$ there holds $(\boldsymbol{\sigma}(t), \xi(t)) \in \mathcal{F}_i(t)$*

and, in addition, we require that for almost all $t \in (0, T)$ and all $(\tilde{\boldsymbol{\sigma}}, \tilde{\xi}) \in \mathcal{F}_i(t)$ there holds

$$(1.16) \qquad \int_{\mathcal{O}} \mathbf{A}\dot{\boldsymbol{\sigma}}(t) \cdot (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}})) + H\dot{\xi}(\xi - \tilde{\xi}) \, \mathrm{d}x \leq \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma}(t) - \tilde{\boldsymbol{\sigma}}) \, \mathrm{d}x.$$

Before stating the main result of the paper, we introduce the safety load condition for the initial data $\boldsymbol{\sigma}_0(0)$, namely

$$(1.17) \qquad \qquad \qquad \|\boldsymbol{\sigma}_{0D}(0)\|_\infty < \kappa.$$

The existence of weak solution to kinematic or isotropic hardening problem in the sense of Definitions 1.1–1.2 is very standard. However, to be able to talk also about the displacement and to obtain regularity results, one usually needs to assume certain compatibility condition on data. In the available literature, the authors usually consider more restrictive assumption than (1.17), which is however more related to the classical problems of elasto–plasticity without hardening. Nevertheless, in our setting, the assumption (1.17) is sufficient, since it leads to the standard safety load condition. Indeed, for kinematic hardening we can set $\boldsymbol{\xi}_0(t) := \boldsymbol{\sigma}_0(t) - \boldsymbol{\sigma}_0(0)$ and then directly we also have

$$(1.18) \qquad (\boldsymbol{\sigma}_0(t), \boldsymbol{\xi}_0(t)) \in \mathcal{F}_k(t), \quad \sup_{t \in (0,T)} \|\boldsymbol{\sigma}_{D0}(t) - \boldsymbol{\xi}_{0D}(t)\|_\infty < \kappa,$$

provided (1.17) holds. Similarly, in the isotropic hardening case, we se $\xi_0(t) := |\boldsymbol{\sigma}_{0D}(t)| - |\boldsymbol{\sigma}_{0D}(0)|$ and we again have

$$(1.19) \qquad (\boldsymbol{\sigma}_0(t), \xi_0(t)) \in \mathcal{F}_i(t), \quad \sup_{t \in (0,T)} \||\boldsymbol{\sigma}_{0D}(t)| - \xi(t)\|_\infty < \kappa,$$

provided (1.17) holds.

Finally, we state the main results of the paper. We consider an approximated problem and show not only the convergence to the original problem but also regularity estimates that are uniform with respect to the approximation parameter. The first one is for the kinematic hardening model. For the approximation, we introduce a new class of admissible stresses as

$$\mathcal{F}_{el}(t) := \Big\{ (\boldsymbol{\sigma}, \boldsymbol{\xi}) \in L^2(\Omega; \mathbb{R}^{d \times d}_{sym}) \times L^2(\Omega; \mathbb{R}^{d \times d}_{sym});$$

$$\text{and for all } \boldsymbol{v} \in \mathcal{V} \text{ there holds } \int_{\mathcal{O}} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot \boldsymbol{\varepsilon}(\boldsymbol{v}) \, \mathrm{d}x = 0 \Big\}.$$

and our result for kinematic hardening reads as follows.

Theorem 1.1 (Kinematic hardening). *Let all assumptions of Definition 1.1 be satisfied. Then for all $\mu > 0$ there exists a unique triple $(\boldsymbol{\sigma}^\mu, \boldsymbol{\xi}^\mu, \boldsymbol{u}^\mu)$ such that $\boldsymbol{u}^\mu - \boldsymbol{u}_0 \in W^{1,2}(0, T; W^{1,2}_0(\mathcal{O}; \mathbb{R}^d))$, $\boldsymbol{\xi}^\mu(0) = 0$, $\boldsymbol{\sigma}^\mu = \boldsymbol{\sigma}_0(0)$ and*

$$(1.20) \qquad \left. \begin{aligned} \mathbf{A}\dot{\boldsymbol{\sigma}}^\mu + \mu^{-1}(|\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D| - \kappa)_+ \frac{\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D|} &= \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^\mu), \\ \mathbf{H}\dot{\boldsymbol{\xi}} = \mu^{-1}(|\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D| - \kappa)_+ \frac{\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}^\mu_D - \boldsymbol{\xi}_D|} \end{aligned} \right\} \quad \textit{a.e. in } (0, T) \times \mathcal{O}.$$

In addition, if the safety initial load condition (1.17) holds and $\boldsymbol{\sigma}_0$ and $\boldsymbol{u}_0$ satisfy

$$(1.21) \qquad \begin{aligned} \boldsymbol{\sigma}_0 &\in W^{2,\infty}(0, T; L^2(\mathcal{O}; \mathbb{R}^{d \times d}_{sym})) \cap W^{1,\infty}(0, T; W^{1,2}(\mathcal{O}; \mathbb{R}^{d \times d}_{sym})), \\ \boldsymbol{u}_0 &\in W^{2,\infty}(0, T; W^{1,2}(\mathcal{O}; \mathbb{R}^{d \times d}_{sym})) \cap W^{1,\infty}(0, T; W^{2,2}(\mathcal{O}; \mathbb{R}^{d \times d}_{sym})), \end{aligned}$$

then we have the following uniform estimates

$$(1.22) \qquad \sup_{t \in (0,T)} \left(\|\dot{\boldsymbol{\sigma}}^{\mu}(t)\|_2^2 + \|\dot{\boldsymbol{\xi}}^{\mu}(t)\|_2 + \|\dot{\boldsymbol{u}}(t)\|_{1,2}^2 \right) \le C(\boldsymbol{u}_0, \boldsymbol{\sigma}_0, \mathcal{O}, T)$$

and there exists a sequence that we do not relabel such that

$$(\boldsymbol{\sigma}^{\mu}, \boldsymbol{\xi}^{\mu}, \boldsymbol{u}^{\mu}) \rightharpoonup^* (\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{u})$$

in the topology induced by the estimate (1.22), where $(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{u})$ is a solution in a sense of Definition 1.1 and satisfies (1.13) almost everywhere.

Moreover, if $\mathcal{O} \in \mathcal{C}^{1,1}$ then for any compact $\tilde{\mathcal{O}} \subset \mathcal{O}$, any $\delta > 0$ and arbitrary nonnegative $\phi \in \mathcal{C}^{\infty}$ fulfilling $\operatorname{supp}\phi \cap \overline{\partial\mathcal{O}}_D \cap \overline{\partial\mathcal{O}}_N = \emptyset$, we have the following estimate

$$(1.23)$$

$$\sup_{t \in (0,T)} \left(\|\boldsymbol{\sigma}^{\mu}(t)\|_{W^{1,2}(\tilde{\mathcal{O}})} + \|\boldsymbol{\xi}^{\mu}(t)\|_{W^{1,2}(\tilde{\mathcal{O}})} + \|\boldsymbol{u}^{\mu}(t)\|_{W^{2,2}(\tilde{\mathcal{O}})} \right)$$

$$+ \sup_{t \in (0,T)} \left(\|\phi\boldsymbol{\sigma}^{\mu}(t)\|_{N_{\boldsymbol{\tau}}^{1,2}(\mathcal{O})} + \|\phi\boldsymbol{\xi}^{\mu}(t)\|_{N_{\boldsymbol{\tau}}^{1,2}(\mathcal{O})} + \|\phi\nabla\boldsymbol{u}^{\mu}(t)\|_{N_{\boldsymbol{\tau}}^{1,2}(\mathcal{O})} \right)$$

$$+ \sup_{t \in (0,T)} \left(\|\phi\boldsymbol{\sigma}^{\mu}(t)\|_{N_{\boldsymbol{n}}^{\frac{3}{5}-\delta,2}(\mathcal{O})} + \|\phi\boldsymbol{\xi}^{\mu}(t)\|_{N_{\boldsymbol{n}}^{\frac{3}{5}-\delta,2}(\mathcal{O})} + \|\phi\nabla\boldsymbol{u}^{\mu}(t)\|_{N_{\boldsymbol{n}}^{\frac{3}{5}-\delta,2}(\mathcal{O})} \right)$$

$$+ \|\dot{\boldsymbol{\sigma}}^{\mu}\|_{N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))} + \|\dot{\boldsymbol{\xi}}^{\mu}\|_{N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))} + \|\nabla\dot{\boldsymbol{u}}^{\mu}\|_{N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))}$$

$$+ \|\dot{\boldsymbol{\sigma}}^{\mu}\|_{L^2(0,T;N^{\frac{1}{2},2}(\tilde{\mathcal{O}}))} + \|\dot{\boldsymbol{\xi}}^{\mu}\|_{L^2(0,T;N^{\frac{1}{2},2}(\tilde{\mathcal{O}}))} + \|\nabla\dot{\boldsymbol{u}}^{\mu}\|_{L^2(0,T;N^{\frac{1}{2},2}(\tilde{\mathcal{O}}))}$$

$$+ \|\phi\dot{\boldsymbol{\sigma}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{\tau}}^{\frac{1}{2},2}(\mathcal{O}))} + \|\phi\dot{\boldsymbol{\xi}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{\tau}}^{\frac{1}{2},2}(\mathcal{O}))} + \|\phi\nabla\dot{\boldsymbol{u}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{\tau}}^{\frac{1}{2},2}(\mathcal{O}))}$$

$$+ \|\phi\dot{\boldsymbol{\sigma}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{n}}^{\frac{1}{5}-\delta,2}(\mathcal{O}))} + \|\phi\dot{\boldsymbol{\xi}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{n}}^{\frac{1}{5}-\delta,2}(\mathcal{O}))} + \|\phi\nabla\dot{\boldsymbol{u}}^{\mu}\|_{L^2(0,T;N_{\boldsymbol{n}}^{\frac{1}{5}-\delta,2}(\mathcal{O}))}$$

$$\le C(\phi, \delta, \tilde{\mathcal{O}}, \boldsymbol{u}_0, \boldsymbol{\sigma}_0),$$

which due to the weak lower semicontinuity holds also for the limit $(\boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{u})$.

Please notice here that we used the notations $N^{\alpha,p}$ for the standard Nikoloskii space, $N_{\boldsymbol{\tau}}^{\alpha,p}$ for the space, where we control the tangential differences, i.e., the space of functions whose fractional α-th tangential[3] derivatives belongs to the Lebesgue space L^p and similarly, $N_{\boldsymbol{n}}^{\alpha,p}$ for the space, where the α-th normal derivative belongs to L^p.

For the isotropic hardening, we have the following result.

Theorem 1.2 (Isotropic hardening). *Let all assumptions of Definition 1.1 be satisfied. Then for all $\mu > 0$ there exists a unique triple $(\boldsymbol{\sigma}^{\mu}, \xi^{\mu}, \boldsymbol{u}^{\mu})$ such that $\boldsymbol{u}^{\mu} - \boldsymbol{u}_0 \in W^{1,2}(0,T; W_0^{1,2}(\mathcal{O}; \mathbb{R}^d))$, $\xi^{\mu}(0) = 0$, $\boldsymbol{\sigma}^{\mu} = \boldsymbol{\sigma}_0(0)$ and*

$$(1.24) \qquad \left. \begin{aligned} \mathbf{A}\dot{\boldsymbol{\sigma}}^{\mu} + \mu^{-1}(|\boldsymbol{\sigma}_D^{\mu}| - \kappa - \xi)_+ \frac{\boldsymbol{\sigma}_D^{\mu}}{|\boldsymbol{\sigma}_D^{\mu}|} &= \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^{\mu}), \\ H\dot{\xi} &= \mu^{-1}(|\boldsymbol{\sigma}_D^{\mu}| - \kappa - \xi)_+ \end{aligned} \right\} \qquad a.e. \ in \ (0,T) \times \mathcal{O}.$$

---

[3]Tangential here means in the directions that are orthogonal to the normal vector at boundary $\partial\mathcal{O}$.

In addition, if the safety initial load condition (1.17) *holds and* $\boldsymbol{\sigma}_0$ *and* $\boldsymbol{u}_0$ *satisfy*

$$(1.25) \qquad \begin{aligned} &\boldsymbol{\sigma}_0 \in W^{2,\infty}(0,T;L^2(\mathcal{O};\mathbb{R}_{sym}^{d\times d})) \cap W^{1,\infty}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}_{sym}^{d\times d})), \\ &\boldsymbol{u}_0 \in W^{2,\infty}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}_{sym}^{d\times d})) \cap W^{1,\infty}(0,T;W^{2,2}(\mathcal{O};\mathbb{R}_{sym}^{d\times d})), \end{aligned}$$

then we have the following uniform estimates

$$(1.26) \qquad \sup_{t\in(0,T)} \left(\|\dot{\boldsymbol{\sigma}}^\mu(t)\|_2^2 + \|\dot{\xi}^\mu(t)\|_2 + \|\dot{\boldsymbol{u}}(t)\|_{1,2}^2 \right) \le C(\boldsymbol{u}_0,\boldsymbol{\sigma}_0,\mathcal{O},T)$$

and there exists a sequence that we do not relabel such that

$$(\boldsymbol{\sigma}^\mu, \xi^\mu, \boldsymbol{u}^\mu) \rightharpoonup^* (\boldsymbol{\sigma}, \xi, \boldsymbol{u})$$

in the topology induced by the estimate (1.22), *where* $(\boldsymbol{\sigma}, \xi, \boldsymbol{u})$ *is a solution in a sense of Definition 1.2 and satisfies* (1.14) *almost everywhere.*

    Moreover, defining

$$\alpha := \frac{2d - 7 + \sqrt{1 + 4d^2 + 20d}}{8(d-1)},$$

then for $\mathcal{O} \in \mathcal{C}^{1,1}$, *any compact* $\tilde{\mathcal{O}} \subset \mathcal{O}$, *any* $\delta > 0$ *and arbitrary nonnegative* $\phi \in \mathcal{C}^\infty$ *fulfilling* $\mathrm{supp}\,\phi \cap \overline{\partial\mathcal{O}}_D \cap \overline{\partial\mathcal{O}}_N = \emptyset$, *we have the following estimate*

$$(1.27) \qquad \begin{aligned} &\sup_{t\in(0,T)} \left(\|\boldsymbol{\sigma}^\mu(t)\|_{W^{1,2}(\tilde{\mathcal{O}})} + \|\xi^\mu(t)\|_{W^{1,2}(\tilde{\mathcal{O}})} \right) \\ &\sup_{t\in(0,T)} \left(\|\phi\boldsymbol{\sigma}^\mu(t)\|_{N_\tau^{1,2}(\mathcal{O})} + \|\phi\xi^\mu(t)\|_{N_\tau^{1,2}(\mathcal{O})} \right) \\ &\sup_{t\in(0,T)} \left(\|\phi\boldsymbol{\sigma}^\mu(t)\|_{N_n^{\alpha-\delta,2}(\mathcal{O})} + \|\phi\xi^\mu(t)\|_{N_n^{\alpha-\delta,2}(\mathcal{O})} \right) \\ &\quad + \|\dot{\boldsymbol{\sigma}}^\mu\|_{N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))} + \|\dot{\xi}^\mu\|_{N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))} \\ &\quad + \|\dot{\boldsymbol{\sigma}}^\mu\|_{L^2(0,T;N^{\frac{1}{2},2}(\tilde{\mathcal{O}}))} + \|\dot{\xi}^\mu\|_{L^2(0,T;N^{\frac{1}{2},2}(\tilde{\mathcal{O}}))} \\ &\quad + \|\phi\dot{\boldsymbol{\sigma}}^\mu\|_{L^2(0,T;N_\tau^{\frac{1}{2},2}(\mathcal{O}))} + \|\phi\dot{\xi}^\mu\|_{L^2(0,T;N_\tau^{\frac{1}{2},2}(\mathcal{O}))} \\ &\quad + \|\phi\dot{\boldsymbol{\sigma}}^\mu\|_{L^2(0,T;N_n^{\frac{\alpha}{3}-\delta,2}(\mathcal{O}))} + \|\phi\dot{\xi}^\mu\|_{L^2(0,T;N_n^{\frac{\alpha}{3}-\delta,2}(\mathcal{O}))} \\ &\le C(\phi,\delta,\tilde{\mathcal{O}},\boldsymbol{u}_0,\boldsymbol{\sigma}_0), \end{aligned}$$

which due to the weak lower semicontinuity holds also for the limit $(\boldsymbol{\sigma}, \xi, \boldsymbol{u})$. *In addition, for any nonnegative* $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ *that fulfils* $\mathrm{supp}\,\varphi \cap \partial\mathcal{O}_N$, *there holds*

$$(1.28) \qquad \begin{aligned} &\sup_{t\in(0,T)} \left(\|\varphi\boldsymbol{\sigma}^\mu(t)\|_{N_n^{\frac{3}{5}-\delta,2}(\mathcal{O})} + \|\varphi\xi^\mu(t)\|_{N_n^{\frac{3}{5}-\delta,2}(\mathcal{O})} \right) \\ &\quad + \|\varphi\dot{\boldsymbol{\sigma}}^\mu\|_{L^2(0,T;N_n^{\frac{1}{5}-\delta,2}(\mathcal{O}))} + \|\varphi\dot{\xi}^\mu\|_{L^2(0,T;N_n^{\frac{1}{5}-\delta,2}(\mathcal{O}))} \\ &\le C(\delta,\boldsymbol{u}_0,\boldsymbol{\sigma}_0,\varphi), \end{aligned}$$

    To end this part of the paper, we emphasize the essential novelties stated in Theorems 1.1–1.2. The existence of a solution was already established in [10, 11] and there is nothing new in the paper. Also the interior $W^{1,2}$ regularity for the stress has been proven in [14, 15] for various models. A key improvement concerning the interior regularity is due to [8] (see also the related paper [9] for problems without

hardening), where the authors showed[4] that $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ belongs to $L^2(0,T;N_{loc}^{\frac{1}{2},2}(\mathcal{O}))$ and $N^{\frac{1}{2},2}(0,T;L^2(\mathcal{O}))$ (see also [2], where a weaker result is obtained for similar problems). However, their estimates were true only for the limit, i.e., for the solution, but were not uniform with respect to approximation. Our result overcomes this weakness and we are able to obtain the uniform μ-independent estimates, see also [4], where the similar statement was proven for plasticity without hardening, or also [12] for up to the boundary or [3] for $W_{loc}^{1,2}$ results for various models in elasto-plasticity theory without hardening.

Concerning the results up to the boundary, the tangential regularity for $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$ was already obtained in [5, 6] and the authors also obtained that the solution belongs to $L^\infty(0,T;N_{\boldsymbol{n}}^{\frac{1}{2}+\delta,2}(\mathcal{O}))$. Hence, our result significantly improves this estimate since in case of kinematic hardening or in case of Dirichlet data we have $N_{\boldsymbol{n}}^{\frac{3}{5}-\delta,2}(\mathcal{O})$ independently of dimension. In addition for isotropic case and Neumann data, we can precisely trace the improvement as stated in (1.27). Finally, and this is the main improvement, we are able to obtain also the fractional $N_{loc}^{\frac{1}{2},2}$ and $N_{\boldsymbol{\tau}}^{\frac{1}{2},2}$ regularity for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ and even more we have an information $N_{\boldsymbol{n}}^{\frac{1}{5}-\delta,2}$ in normal direction, which is obtained by the cross interpolation. Note that in view of the counterexamples to $W^{1,2}$ regularity up to the boundary proven e.g. in [16, 13], such estimates seems to be optimal.

In the rest of the paper, we will focus only on the kinematic hardening and we shall just emphasize where are the differences. Obviously, in the kinematic hardening case, we can transfer the obtained regularity from $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ to $\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})$ just by using the equation. Then the regularity for $\nabla\dot{\boldsymbol{u}}$ just follows from the Korn inequality applied on $\mathcal{O}$ or its sub-domain and we do not provide details here. On the other hand, in case of isotropic hardening, we are not able to use such a procedure. The best, we can do is just to transfer better integrability to $\nabla\dot{\boldsymbol{u}}$, see also [5, 6], but this is also omitted here, since it is just direct consequence of the Korn inequality and embedding theorem. The second case is when we combine the isotropic hardening and the Neumann boundary conditions. The reason for that is that in such case we cannot use any version of anisotropic Korn inequality to transfer optimal anisotropic integrability from symmetric gradient to the full gradient. Also, we would like to emphasize that our restriction on the constant $\mathbf{A}$ and $\mathbf{H}$ is not necessary and the proof would remain almost exactly identical if they are Lipschitz functions of (t, x). Finally, to simplify the presentation, we consider only the flat boundary, however for $\mathcal{C}^{1,1}$ boundaries, it is just a technical difficulty to transform the problem with general boundary to the case of flat boundary. Finally, let us remark that we cannot avoid a possible singularity on $\overline{\partial\mathcal{O}}_D \cap \overline{\partial\mathcal{O}}_N$ from principal reasons, since even for linear elliptic problems one may observe a singularity.

2. Proof of Theorems 1.1 and 1.2

We focus here mainly on the kinematic hardening case, since the proof for the isotropic hardening is very similar. Only on certain places, we discuss the possible differences. Also to simplify notation, we set $\kappa \equiv 1$ in what follows. Finally having

---

[4]In fact they showed an estimate, from which one can deduce the result following the method invented in [7].

a priori estimates stated in Theorem 1.1 or in Theorem 1.2, it is very classical to pass to the limit $\mu \to 0_+$ and to obtain the solution to the original problem, i.e., to the kinematic hardening and the isotropic hardening problems. Therefore, we also skip the limiting passage in the proof. Finally, we do not discuss the problem of existence of a solution for fix $\mu > 0$ since it was already established by many authors, see e.g. [10, 11], but we rather focus on a priori estimates. Also in order to shorten the text, we omit writing superscripts to emphasize we deal with a solution to an approximative problem. Furthermore, we do not trace the dependence of all constants on $\mathcal{O}$ or C_1 and in what follows the constant C has a meaning of some universal generic constant that may vary line to line but is independent of μ. In case we want to emphasize the dependence of this constant on some parameter it is clearly denoted.

2.1. **First a priori uniform estimates.** Thus, we shall assume that for any $\mu > 0$ there exists a solution (1.20). The existence of such a σ can be shown e.g. by the Rothe approximation and we refer the interested reader to [8] or [17], where even a more difficult case of problem without hardening is treated, or to original papers [10, 11]. Hence, we assume that there is $\boldsymbol{u} \in W^{1,2}(0,T;W^{1,2}(\mathcal{O};\mathbb{R}^d))$ such that for all $t \in (0,t)$ $\boldsymbol{u} - \boldsymbol{u}_0 \in \mathcal{V}$ and $\boldsymbol{u}(0) = \boldsymbol{u}_0(0)$, and that there is $(\boldsymbol{\sigma},\boldsymbol{\xi}) \in \mathcal{F}_{el}$ fulfilling

$$
(2.1) \quad \left.\begin{aligned}
\mathbf{A}\dot{\boldsymbol{\sigma}} + \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} &= \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}), \\
\mathbf{H}\dot{\boldsymbol{\xi}} = \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} &
\end{aligned}\right\} \quad \text{in } (0,T) \times \mathcal{O}.
$$

The next step is to derive the uniform (μ independent estimates) for $(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\xi})$. We proceed here formally, since the estimates are known, see e.g. [18, 17, 8]. Taking the scalar product of the first equation in (2.1) with $\boldsymbol{\sigma} - \boldsymbol{\sigma}_0$, recall here that $\boldsymbol{\sigma}_0$ satisfies the compatibility condition (1.18), we deduce after integration over $\mathcal{O}$ that

$$
\begin{aligned}
(2.2) \quad &\int_{\mathcal{O}} \mathbf{A}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) + \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{(\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot (\boldsymbol{\sigma}_D - \boldsymbol{\sigma}_{0D})}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \, dx \\
&= \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \, dx + \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) - \mathbf{A}\dot{\boldsymbol{\sigma}}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \, dx.
\end{aligned}
$$

Second identity, we deduce from (2.1) by taking the scalar product with ξ_0. Thus, we have

$$
\begin{aligned}
(2.3) \quad &\int_{\mathcal{O}} \mathbf{H}(\dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\xi}}_0) \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \, dx \\
&= \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{(\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot (\boldsymbol{\xi}_D - \boldsymbol{\xi}_{0D})}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \, dx - \dot{\boldsymbol{\xi}}_0 \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \, dx.
\end{aligned}
$$

Since $(\boldsymbol{\sigma}(t),\boldsymbol{\xi}) \in \mathcal{F}_{el}$ and $(\boldsymbol{u} - \boldsymbol{u}_0) \in \mathcal{V}$, we see that the first term on the right hand side of (2.1) vanishes. In addition, since $(\boldsymbol{\sigma}_0,\boldsymbol{\xi}_0)$ satisfies the compatibility condition (1.18), we observe

$$
\begin{aligned}
&(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ (\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot (\boldsymbol{\sigma}_D - \boldsymbol{\sigma}_{0D} - \boldsymbol{\xi}_D + \boldsymbol{\xi}_{0D}) \\
&= (|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ (|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|^2 - (\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot (\boldsymbol{\sigma}_{0D} - \boldsymbol{\xi}_{0D})) \\
&\geq (|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ |\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - |\boldsymbol{\sigma}_{0D} - \boldsymbol{\xi}_{0D}|) \geq 0.
\end{aligned}
$$

Finally, we can sum (2.2) and (2.3) and by using the above inequality, the symmetry of $\mathbf{A}$ and $\mathbf{H}$ as well as the ellipticity (1.12) and also the boundedness of $\mathbf{A}$ and $\mathbf{H}$, we arrive to the inequality

$$
\begin{aligned}
\frac{d}{dt} & \int_{\mathcal{O}} \mathbf{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) + \mathbf{H}(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \, dx \\
& \leq C \left(\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2 + \|\dot{\boldsymbol{\sigma}}_0\|_2^2 + \|\dot{\boldsymbol{\xi}}_0\|_2^2 \right) \\
& \quad + C \int_{\mathcal{O}} \mathbf{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) + \mathbf{H}(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \, dx.
\end{aligned}
\tag{2.4}
$$

Consequently, by the Gronwall lemma, we have

$$
\begin{aligned}
\sup_{t \in (0,T)} & (\|\boldsymbol{\sigma}(t)\|_2^2 + \|\boldsymbol{\xi}(t)\|_2^2) \\
& \leq C \left(\|\boldsymbol{\sigma}_0(0)\|_2^2 + \int_0^T \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2 + \|\dot{\boldsymbol{\sigma}}_0\|_2^2 + \|\dot{\boldsymbol{\xi}}_0\|_2^2 \, dt \right) \leq C,
\end{aligned}
\tag{2.5}
$$

where the last inequality follows from the assumptions on data (namely on $\boldsymbol{\sigma}_0$, $\boldsymbol{u}_0$ and $\boldsymbol{\xi}_0$).

The next step is to test the first equation in (2.1) by $\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0$ and the second equation by $\dot{\boldsymbol{\xi}}$. Doing so, and summing the resulting identities, we get

$$
\begin{aligned}
\int_{\mathcal{O}} & \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} + \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{(\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot (\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \, dx \\
& = \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{(\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \cdot \dot{\boldsymbol{\sigma}}_{0D}}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \, dx \\
& \quad + \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}}_0 \, dx \\
& = \int_{\mathcal{O}} \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\sigma}}_{0D} + (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}}_0 \, dx,
\end{aligned}
\tag{2.6}
$$

where the last equality follows from $(2.1)_2$ and the fact that $(\boldsymbol{\sigma}, \boldsymbol{\xi}) \in \mathcal{F}_{el}$. Using the ellipticity (1.12), the Young inequality and the following identity

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

we see that it follows from (2.6) that

$$
\frac{d}{dt} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+^2 \, dx + C_1(\|\dot{\boldsymbol{\sigma}}\|_2^2 + \|\dot{\boldsymbol{\xi}}\|_2^2) \leq C(\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2 + \|\dot{\boldsymbol{\sigma}}_0\|_2^2).
$$

Thus, integrating with respect to $t \in (0, T)$ and using the fact that $\boldsymbol{\xi}(0) = 0$ and that $|\boldsymbol{\sigma}(0)| = |\boldsymbol{\sigma}_0(0)| \leq 0$, we get the uniform bound

$$
\begin{aligned}
\sup_{t \in (0,T)} & \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+^2 \, dx + \int_0^T \|\dot{\boldsymbol{\sigma}}\|_2^2 + \|\dot{\boldsymbol{\xi}}\|_2^2 \, dt \\
& \leq C \int_0^T \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2 + \|\dot{\boldsymbol{\sigma}}_0\|_2^2 \, dt \leq C,
\end{aligned}
\tag{2.7}
$$

where the last inequality (with C being independent of μ) follows from the assumptions on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}_0$. In addition, it follows from (2.1) that

$$
|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})| \leq C(|\dot{\boldsymbol{\sigma}}| + |\dot{\boldsymbol{\xi}}|)
$$

and consequently, (2.7), the Korn inequality and the assumptions on $\boldsymbol{u}_0$ leads to the uniform bound

$$(2.8) \qquad \int_0^T \|\dot{\boldsymbol{u}}\|_{1,2}^2 \, \mathrm{d}t \leq C.$$

The last step is L^∞ bound for the time derivative. We apply the time derivative to (2.1) and take the scalar product of the first equation with $\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0$ and the scalar product of the second equation with $\dot{\boldsymbol{\xi}}$. Summing these to identities we obtain

$$(2.9) \qquad \begin{aligned} &\int_{\mathcal{O}} \mathsf{A}(\ddot{\boldsymbol{\sigma}} - \ddot{\boldsymbol{\sigma}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathsf{H}\ddot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \, \mathrm{d}x \\ &\quad + \int_{\mathcal{O}} \mu^{-1} \frac{\partial(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|}}{\partial t} \cdot (\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D - \dot{\boldsymbol{\sigma}}_{0D}) \, \mathrm{d}x \\ &\quad = \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}} - \ddot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \, \mathrm{d}x + \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0) - \mathsf{A}\ddot{\boldsymbol{\sigma}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \, \mathrm{d}x. \end{aligned}$$

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

$$(2.10) \qquad \mu^{-1} \frac{\partial(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|}}{\partial t} \cdot \dot{\boldsymbol{\beta}} = \frac{\mu^{-1} \chi_{|\boldsymbol{\beta}|>1}}{|\boldsymbol{\beta}|} \left(|\dot{\boldsymbol{\beta}}|^2(|\boldsymbol{\beta}| - 1) + |\partial_t|\boldsymbol{\beta}||^2 \right) \geq 0.$$

Consequently, using the Hölder inequality and the above inequality, we see that (2.9) implies (using also the second identity in (2.1))

$$(2.11) \qquad \begin{aligned} &\frac{1}{2}\frac{d}{dt} \int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathsf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} - 2\mathsf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\sigma}}_{0D} \, \mathrm{d}x \\ &\qquad \leq C(\|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0)\|_2 + \|\ddot{\boldsymbol{\sigma}}_0\|_2)(1 + \|\dot{\boldsymbol{\xi}}\|_2 + \|\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}^s\|_2). \end{aligned}$$

Hence, adding the term

$$\frac{1}{2}\frac{d}{dt} \int_{\mathcal{O}} \mathsf{H}\dot{\boldsymbol{\sigma}}_{0D} \cdot \dot{\boldsymbol{\sigma}}_{0D} \, \mathrm{d}x$$

to both sides of (2.11), we deduce

$$(2.12) \qquad \begin{aligned} &\frac{1}{2}\frac{d}{dt} \int_{\mathcal{O}} \mathsf{A}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \cdot (\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathsf{H}(\dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\sigma}}_{0D}) \cdot (\dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\sigma}}_{0D}) \, \mathrm{d}x \\ &\qquad \leq C(\|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0)\|_2 + \|\ddot{\boldsymbol{\sigma}}_0\|_2)(1 + \|\dot{\boldsymbol{\sigma}}_{0D}\|_2 + \|\dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\sigma}}_{0D}\|_2 + \|\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_{0D}\|_2). \end{aligned}$$

Consequently, integration of this inequality and the ellipticity assumption (1.12) lead to the estimate

$$(2.13) \qquad \begin{aligned} &\sup_{t \in (0,T)} \left(\|\dot{\boldsymbol{\sigma}}(t)\|_2^2 + \|\dot{\boldsymbol{\xi}}(t)\|_2^2 \right) \\ &\qquad \leq C \left(1 + \|\dot{\boldsymbol{\sigma}}_0(0)\|_2^2 + \left(\int_0^T \|\ddot{\boldsymbol{u}}_0\|_{1,2} + \|\ddot{\boldsymbol{\sigma}}_0\|_2 \, \mathrm{d}t \right)^2 \right) \leq C. \end{aligned}$$

Furthermore, it follows from (2.1) and the Korn inequality that

$$(2.14) \qquad \sup_{t \in (0,T)} \|\dot{\boldsymbol{u}}(t)\|_{1,2}^2 \leq C.$$

2.2. Uniform $W^{1,2}$ estimates.

In this subsection, we derive the uniform interior estimates on $\nabla\boldsymbol{\sigma}$ and on $\nabla\boldsymbol{\xi}$ and estimates for tangential derivatives of $(\boldsymbol{\sigma},\boldsymbol{\xi})$ up to the boundary.

To simplify the presentation, we consider here only a flat boundary case but it can be straightforwardly extended to the general boundary. Also since the interior regularity is in fact easier to prove than the boundary regularity, we provide here only the estimates near the boundary for tangential derivatives. Hence to simplify the notation, we assume from now the most difficult case, i.e., we focus on a cube $(-1,1)^{d-1} \times (0,1) \subset \mathcal{O}$, where the Dirichlet and the Neumann parts are supposed to satisfy

$$(-1,1)^{d-2} \times (-1,0) \times \{0\} \subset \partial\mathcal{O}_D \qquad (-1,1)^{d-2} \times (0,1) \times \{0\} \subset \partial\mathcal{O}_N.$$

Our goal is to show that except the set $(-1,1)^{d-2} \times \{0\} \times \{0\}$ we have uniform estimates for $D_j\boldsymbol{\sigma}$ and $D_j\boldsymbol{\xi}$ in the space L^2 for all $j = 1,\ldots,d-1$, where D_j denotes the partial derivative with respect to x_j.

Thus, let $\phi \in \mathcal{D}(-1,1)^d$ be arbitrary nonnegative function satisfying $\phi \leq 1$. Furthermore, we require that for some $\varepsilon_0 > 0$, the function ϕ satisfies for all $x_1,\ldots,x_{d-1}$ and all $|s| + |t| \leq \varepsilon_0$ that $\tau(x_1,\ldots,x_{d-2},s,t) = 0$. Next, we fix arbitrary $j = 1,\ldots,d-1$ and apply the operator D_j to both equations in (2.1). Then we take the scalar product of the first equation with $D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2$ and the scalar product of the second equation with $D_j\boldsymbol{\xi}\phi^2$, sum the results and integrate over $\mathcal{O}$ to deduce the identity

$$\int_{\mathcal{O}} \mathbf{A}D_j(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 + \mathbf{H}D_j\dot{\boldsymbol{\xi}} \cdot D_j\boldsymbol{\xi}\phi^2 \,\mathrm{d}x$$

$$(2.15) \qquad + \int_{\mathcal{O}} \mu^{-1}D_j\left(\frac{(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|}\right) \cdot D_j(\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D)\phi^2 \,\mathrm{d}x$$

$$= \int_{\mathcal{O}} D_j\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 - \mathbf{A}D_j\dot{\boldsymbol{\sigma}}_0 \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 \,\mathrm{d}x.$$

Note that the second integral on the left hand side is nonnegative (see the same procedure as for the estimates for the first time derivatives) and can be neglected. Next, using the symmetry of $\mathbf{A}$ and $\mathbf{H}$, we can deduce that

$$\frac{1}{2}\frac{d}{dt}\int_{\mathcal{O}} \mathbf{A}D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 + \mathbf{H}D_j\boldsymbol{\xi} \cdot D_j\boldsymbol{\xi}\phi^2 \,\mathrm{d}x$$

$$\leq \int_{\mathcal{O}} \boldsymbol{\varepsilon}(D_j\dot{\boldsymbol{u}} - D_j\dot{\boldsymbol{u}}_0) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 \,\mathrm{d}x$$

$$+ \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(D_j\dot{\boldsymbol{u}}_0) - \mathbf{A}D_j\dot{\boldsymbol{\sigma}}_0) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2$$

$$(2.16) \qquad \leq \int_{\mathcal{O}} \nabla(\phi^2(D_j\dot{\boldsymbol{u}} - D_j\dot{\boldsymbol{u}}_0)) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \,\mathrm{d}x$$

$$- 2\int_{\mathcal{O}} \nabla\phi \otimes (D_j\dot{\boldsymbol{u}} - D_j\dot{\boldsymbol{u}}_0)) \cdot D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi \,\mathrm{d}x$$

$$+ C(\|\dot{\boldsymbol{u}}_0\|_{2,2} + \|\dot{\boldsymbol{\sigma}}\|_{1,2})\|D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2$$

$$\leq C(\|\dot{\boldsymbol{u}}_0\|_{2,2} + \|\dot{\boldsymbol{u}}\|_{1,2} + \|\dot{\boldsymbol{\sigma}}\|_{1,2})\|D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2,$$

where for the last inequality we used integration by parts and the fact that D_j is the tangential derivative and so $D_j(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)$ vanishes on $\partial\mathcal{O}_D$ and similarly

$D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\boldsymbol{n} = 0$ on $\partial\mathcal{O}_N$ and also thanks to the fact that $\mathrm{div}\, D_j(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) = 0$. Consequently, it directly follows from (2.16) that

$$(2.17) \qquad \sup_{t \in (0,T)} \left(\|D_j\boldsymbol{\sigma}(t)\phi\|_2 + \|\phi D_j\boldsymbol{\xi}(t)\|_2 \right) \leq C,$$

where we used the a priori estimate (2.13) and the assumptions on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}_0$.

2.3. **Fractional time regularity for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$.** In this section, we prove the first new result. Although the estimate is known for the solution of the original problem, see [8], it was not clear whether the estimate can be obtained uniformly with respect to the parameter μ. In addition, in [8], the $N^{\frac{1}{2}-\delta,2}$ regularity is proven, while here we obtain $1/2$ derivative estimate.

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 the first equation in (2.1) with $-\Delta_t^\tau(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)$, and the scalar product of the second equation in (2.1) with $-\Delta_t^\tau\dot{\boldsymbol{\xi}}$, sum the resulting equalities and finally integrate the result over $\mathcal{O}$ to get

$$-\int_{\mathcal{O}} \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^\tau\dot{\boldsymbol{\xi}}\, \mathrm{d}x$$
$$-\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\, \mathrm{d}x$$
$$= \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) - \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^\tau\dot{\boldsymbol{\sigma}}_{0D}\, \mathrm{d}x.$$

The first term on the right hand side vanishes and after the use of the Hölder inequality and the bound (1.12), and after reorganisation of all terms, we deduce that

$$-\int_{\mathcal{O}} \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^\tau\dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^\tau\dot{\boldsymbol{\xi}}\, \mathrm{d}x$$
$$(2.18) \qquad -\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\, \mathrm{d}x$$
$$\leq C(\|\dot{\boldsymbol{\sigma}}\|_2 + \|\dot{\boldsymbol{\xi}}\|_2)\|\Delta_t^\tau\dot{\boldsymbol{\sigma}}_0\|_2 + \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}})\, \mathrm{d}x.$$

Next, we focus on the first term on the left hand side. Due to the symmetry of $\mathbf{A}$ and $\mathbf{H}$, we have

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

Therefore, substituting these identities into (2.18) and using the ellipticity of $\mathbf{A}$ and $\mathbf{H}$ (see (1.12)), we further observe with the help of the a piori estimate (2.13) that

(2.19)
$$\int_{\mathcal{O}} C_1(|\Delta_t^\tau\dot{\boldsymbol{\sigma}}|^2 + |\Delta_t^\tau\dot{\boldsymbol{\xi}}|^2) - 2\mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\, \mathrm{d}x$$
$$\leq \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}) + \Delta_t^\tau\left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}}\right)\, \mathrm{d}x + C\|\Delta_t^\tau\dot{\boldsymbol{\sigma}}_0\|_2.$$

Next, we integrate the resulting inequality with respect to τ over the interval $(0, h)$ and with respect to t over the interval $(0, T - h)$ to get

(2.20)
$$\begin{aligned}
C_1 \int_0^{T-h} &\int_0^h \|\Delta_t^\tau \dot{\boldsymbol{\sigma}}\|_2^2 + \|\Delta_t^\tau \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t \\
&\leq \int_0^{T-h} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \times \\
&\qquad \times \left(\int_0^h \Delta_t^\tau (\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D) \, \mathrm{d}\tau \right) \, \mathrm{d}x \, \mathrm{d}t \\
&\quad + \int_0^{T-h} \int_0^h \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}) + \Delta_t^\tau \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right) \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t \\
&\quad + C \int_0^{T-h} \int_0^h \|\Delta_t^\tau \dot{\boldsymbol{\sigma}}_0\|_2 \, \mathrm{d}\tau \, \mathrm{d}t.
\end{aligned}$$

Our goal is to provide uniform estimates for all terms on the right hand side. We start with the easiest one. Hence, using the assumption on $\boldsymbol{\sigma}_0$, we have

(2.21)
$$\begin{aligned}
\int_0^{T-h} \int_0^h \|\Delta_t^\tau \dot{\boldsymbol{\sigma}}_0\|_2 \, \mathrm{d}\tau \, \mathrm{d}t &= \int_0^h \tau \left(\int_0^{T-h} \frac{\|\Delta_t^\tau \dot{\boldsymbol{\sigma}}_0\|_2}{\tau} \, \mathrm{d}t \right) \, \mathrm{d}\tau \\
&\leq \int_0^h \tau \int_0^T \|\ddot{\boldsymbol{\sigma}}_0\|_2 \, \mathrm{d}t \, \mathrm{d}\tau \leq Ch^2.
\end{aligned}$$

Next, we focus on the term with $\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)$. We simply shift the differences to $\boldsymbol{u}_0$ and then use the assumptions on $\boldsymbol{u}_0$ and alreday obtained estimates. In more details, we reorganise the first term as follows

$$\begin{aligned}
\int_0^{T-h} &\int_0^h \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}) \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t \\
&= \int_0^h \int_0^{T-h} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}}_0(t+\tau) - \dot{\boldsymbol{\sigma}}(t+\tau)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&\quad - \int_0^h \int_0^{T-h} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&= \int_0^h \int_\tau^{T-h+\tau} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau)) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&\quad - \int_0^h \int_0^{T-h} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&= \int_0^h \int_\tau^{T-h} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau) - \dot{\boldsymbol{u}}_0(t)) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&\quad + \int_0^h \int_{T-h}^{T-h+\tau} \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t-\tau)) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau \\
&\quad - \int_0^h \int_0^\tau \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0(t)) \cdot (\dot{\boldsymbol{\sigma}}_0(t) - \dot{\boldsymbol{\sigma}}(t)) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau
\end{aligned}$$

and then with the help of the Hölder inequality we obtain

$$\int_0^{T-h} \int_0^h \int_{\mathcal{O}} 2\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}) \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$

(2.22)
$$\leq C \int_0^h \tau \|\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)} \|\boldsymbol{\varepsilon}(\ddot{\boldsymbol{u}}_0)\|_{L^1(0,T;L^2)} \, \mathrm{d}\tau$$

$$+ C \int_0^h \tau \|\dot{\boldsymbol{\sigma}}_0 - \dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)} \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_{L^\infty(0,T;L^2)} \, \mathrm{d}\tau$$

$$\leq Ch^2,$$

where we used the assumptions on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}_0$ and the a priori estimate (2.13). The term involving the matrices $\mathbf{A}$ and $\mathbf{H}$ on the right hand side of (2.20) is estimated as follows

$$\int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_t^\tau \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right) \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$

$$= \int_0^h \int_0^{T-h} \int_{\mathcal{O}} \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right)(t+\tau) - \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right)(t) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau$$

(2.23)
$$= \int_0^h \int_{T-h}^{T-h+\tau} \int_{\mathcal{O}} \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau$$

$$- \int_0^h \int_0^\tau \int_{\mathcal{O}} \left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}} \right) \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}\tau$$

$$\leq Ch^2(\|\dot{\boldsymbol{\sigma}}\|_{L^\infty(0,T;L^2)} + \|\dot{\boldsymbol{\xi}}\|_{L^\infty(0,T;L^2)}) \leq Ch^2,$$

where the estimate (2.13) is used again. Thus, it remains to evaluate the first term on the right hand side of (2.20). To simplify formula, we set $\boldsymbol{\beta} := \boldsymbol{\sigma}_D - \boldsymbol{\xi}_D$ and using the convexity of $(|\boldsymbol{\beta}| - 1)_+^2$, we continue as follows

(2.24)

$$\int_0^{T-h} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \cdot \left(\int_0^h \Delta_t^\tau(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D) \, \mathrm{d}\tau \right) \, \mathrm{d}x \, \mathrm{d}t$$

$$= \int_0^{T-h} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\int_0^h \Delta_t^\tau \dot{\boldsymbol{\beta}} \, \mathrm{d}\tau \right) \, \mathrm{d}x \, \mathrm{d}t$$

$$= \int_0^{T-h} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\boldsymbol{\beta}(t+h) - \boldsymbol{\beta}(t) - h\dot{\boldsymbol{\beta}}(t) \right) \, \mathrm{d}x \, \mathrm{d}t$$

$$= \int_0^{T-h} \int_{\mathcal{O}} 2\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\boldsymbol{\beta}(t+h) - \boldsymbol{\beta}(t) - h\dot{\boldsymbol{\beta}}(t) \right) \, \mathrm{d}x \, \mathrm{d}t$$

$$\leq \int_0^{T-h} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}(t+h)| - 1)_+^2 - (|\boldsymbol{\beta}(t)| - 1)_+^2 - h\mu^{-1} \frac{d}{dt}(|\boldsymbol{\beta}(t)| - 1)_+^2 \, \mathrm{d}t$$

$$\leq \int_0^h \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}(T-h+t)| - 1)_+^2 - (|\boldsymbol{\beta}(T-h)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$- \int_0^h \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}(t)| - 1)_+^2 - (|\boldsymbol{\beta}(0)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t.$$

Note that we also used the fact that $|\boldsymbol{\beta}(0)| \leq 1$, which follows from the fact that $\boldsymbol{\xi}(0) = 0$ and (1.18). To estimate the right hand side, we notice that (2.6) leads to

the identity

$$(2.25) \quad \begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\beta}|-1)_+^2\,\mathrm{d}x \\ &= \int_{\mathcal{O}}\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\sigma}}_{0D}+(\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\cdot(\dot{\boldsymbol{\sigma}}-\dot{\boldsymbol{\sigma}}_0)+\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}_0-\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}-\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}}\,\mathrm{d}x, \end{aligned}$$

which after integration over arbitrary interval $(\tau,\tau+\alpha)\subset(0,T)$ and with the help of the a priori estimate (2.13) and the assumption on data $\boldsymbol{\sigma}_0$ and $\boldsymbol{u}_0$, leads to

$$(2.26) \quad \begin{aligned} &\left|\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\beta}(\alpha+\tau)|-1)_+^2-\mu^{-1}(|\boldsymbol{\beta}(\tau)|-1)_+^2\,\mathrm{d}x\right| \\ &\leq C\int_{\tau}^{\alpha+\tau}\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2^2+\|\dot{\boldsymbol{\sigma}}\|_2^2+\|\dot{\boldsymbol{\sigma}}_0\|_2^2+\|\dot{\boldsymbol{\xi}}\|_2^2\,\mathrm{d}t\leq C\alpha. \end{aligned}$$

Thus, using this estimate in (2.24), we see that

$$(2.27) \quad \begin{aligned} &\int_0^{T-h}\int_{\mathcal{O}}2\mu^{-1}(|\boldsymbol{\beta}|-1)_+\frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|}\cdot\left(\int_0^h\Delta_t^\tau\dot{\boldsymbol{\beta}}\,\mathrm{d}\tau\right)\,\mathrm{d}x\,\mathrm{d}t \\ &\leq C\int_0^h t\,\mathrm{d}t\leq Ch^2. \end{aligned}$$

Finally, we substitute the estimates (2.21), (2.22), (2.23) and (2.27) into (2.20) and finish this part with the uniform estimate

$$(2.28) \quad \frac{1}{h^2}\int_0^{T-h}\int_0^h\|\Delta_t^\tau\dot{\boldsymbol{\sigma}}\|_2^2+\|\Delta_t^\tau\dot{\boldsymbol{\xi}}\|_2^2\,\mathrm{d}\tau\,\mathrm{d}t\leq C.$$

Finally, up to small differences we mimic the procedure from [7], to deduce the proper estimate from (2.28). Indeed, we can compute

$$\begin{aligned} &h^{-1}\int_0^{T-2h}\|\Delta_t^h\boldsymbol{\beta}\|_2^2\,\mathrm{d}t \\ &= h^{-1}\int_0^{T-2h}\left\|\frac{1}{h}\int_0^h\boldsymbol{\beta}(t+h)-\boldsymbol{\beta}(t+h-\tau)+\boldsymbol{\beta}(t+h-\tau)-\boldsymbol{\beta}(t)\,\mathrm{d}\tau\right\|_2^2\,\mathrm{d}t \\ &\leq 2h^{-2}\int_0^{T-2h}\int_0^h\|\boldsymbol{\beta}(t+h)-\boldsymbol{\beta}(t+h-\tau)\|_2^2+\|\boldsymbol{\beta}(t+h-\tau)-\boldsymbol{\beta}(t)\|_2^2\,\mathrm{d}\tau\,\mathrm{d}t \\ &\leq 4h^{-2}\int_0^{T-h}\int_0^h\|\Delta_t^\tau\boldsymbol{\beta}\|_2^2\,\mathrm{d}\tau\,\mathrm{d}t. \end{aligned}$$

Consequently, it then follows from (2.28) that

$$(2.29) \quad h^{-1}\int_0^{T-2h}\|\dot{\boldsymbol{\sigma}}(t+h)-\dot{\boldsymbol{\sigma}}(t)\|_2^2+\|\dot{\boldsymbol{\xi}}(t+h)-\dot{\boldsymbol{\xi}}(t)\|_2^2\,\mathrm{d}t\leq C.$$

Note that there are only minor changes in the proof for isotropic hardening and therefore we do not provide it here.

2.4. **Fractional spatial regularity for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$.** The second main result is the spatial fractional regularity upto the boundary in tangential direction and also the interior fractional spatial regularity. It is again based of [8], but we do provide here the estimates independent of μ and extend them up to the boundary when dealing with tangential direction. We keep the notation from the previous section and focus

only on the flat boundary. Next, we introduce the space-time shift as follows. For arbitrary w, we denote $\Delta_{t,j}^{\tau,h} w(t,x) := w(t+\tau, x+he_j) - w(t,x)$, where e_j is the unite vector in the j-th direction and $j = 1, \ldots, d-1$. We take the scalar product of the first equation in (2.1) with $-\Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2$, of the second equation in (2.1) with $-\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}}\phi^2$, sum the resulting equalities and finally integrate the result over $\mathcal{O}$ to get

$$-\int_{\mathcal{O}} \mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}}\phi^2 \, \mathrm{d}x$$
$$-\int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\phi^2 \, \mathrm{d}x$$
$$= \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 - \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_{t,j}^{\tau,h}\boldsymbol{\sigma}_{0D}\phi^2 \, \mathrm{d}x.$$

Next, we focus on the first term on the left hand side. Similarly as before, we have

$$-\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}} = \frac{1}{2}\mathbf{A}\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}} \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}} - \frac{1}{2}\Delta_{t,j}^{\tau,h}\left(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}}\right),$$
$$-\mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}} = \frac{1}{2}\mathbf{H}\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}} \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}} - \frac{1}{2}\Delta_{t,j}^{\tau,h}\left(\mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}}\right).$$

Therefore, using this identities and also the ellipticity condition (1.12), we get after integration with respect to $t \in (0, T-h)$ and $\tau \in (0, h)$

$$C_1 \int_0^{T-h} \int_0^h \|\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}\phi\|_2^2 + \|\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\xi}}\phi\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t$$
$$\leq 2\int_0^{T-h} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \times$$
$$\times \left(\int_0^h \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D) \, \mathrm{d}\tau\right)\phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
(2.30)
$$+ \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_{t,j}^{\tau,h}(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}})\phi^2 + \Delta_{t,j}^{\tau,h}(\mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}})\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$+ 2\int_0^{T-h} \int_0^h \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$+ 2\int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) - \mathbf{H}\dot{\boldsymbol{\xi}} - \mathbf{A}\dot{\boldsymbol{\sigma}}) \cdot \Delta_{t,j}^{\tau,h}\boldsymbol{\sigma}_0\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t.$$

Our goal is to provide uniform estimates for all terms on the right hand side.

The fourth term in (2.30). We start with the last term. Using the Hölder inequality and the characterization of Sobolev spaces as well as the uniform bound (2.13)

and the assumption on $\boldsymbol{u}_0$, we have

$$
\begin{aligned}
(2.31) \quad & \int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) - \mathsf{H}\dot{\boldsymbol{\xi}} - \mathsf{A}\dot{\boldsymbol{\sigma}}) \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}_0\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \leq C \int_0^{T-h} \int_0^h (\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2 + \|\dot{\boldsymbol{\xi}}\|_2 + \|\dot{\boldsymbol{\sigma}}\|)\|\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}_0\phi^2\|_2 \, \mathrm{d}\tau\, \mathrm{d}t \\
& \leq C \int_0^{T-h} \int_0^h \|(\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}_0 - \Delta_j^h\dot{\boldsymbol{\sigma}}_0)\phi^2\|_2 + \|\Delta_j^h\dot{\boldsymbol{\sigma}}_0\phi^2\|_2 \, \mathrm{d}\tau\, \mathrm{d}t \\
& \leq C h \int_0^h \int_0^{T-h} \|\ddot{\boldsymbol{\sigma}}_0\|_2 + \|\nabla\dot{\boldsymbol{\sigma}}_0\|_2 \, \mathrm{d}\tau\, \mathrm{d}t \leq C h^2.
\end{aligned}
$$

The third term in (2.30)**.** In the third term in (2.30), we first split the time-space shift to the space shift and the time shift. Next, for the time shift, we again use the equation (2.1), while for the space shift we use the integration by parts (note that the boundary term vanishes since we have shifts only in tangential direction and we also know that $\mathrm{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) = 0$) and then we also move the shift from $\boldsymbol{\sigma}$ to other terms. In addition, we keep all shifts on $\boldsymbol{\sigma}_0$ since it has sufficient regularity. More precisely, we have

$$
\begin{aligned}
& \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& = \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)\phi^2 - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad + \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \cdot (\Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) - \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0))\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& = \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot (\Delta_t^{\tau}\dot{\boldsymbol{\sigma}} - \Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}})\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad \int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)) \cdot \Delta_t^{\tau}\dot{\boldsymbol{\sigma}}_0\phi^2 - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) \cdot \Delta_t^{\tau}\dot{\boldsymbol{\sigma}}\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad + \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \nabla(\phi^2(\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}})) \cdot (\Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) - \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)) \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad - \int_0^{T-h} \int_0^h \int_{\mathcal{O}} ((\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \otimes \nabla\phi^2) \cdot (\Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0) - \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}} - \dot{\boldsymbol{\sigma}}_0)) \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& = -\int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_j^{-h}(\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\phi^2) \cdot \dot{\boldsymbol{\sigma}}(t+\tau) \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad + \int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)) \cdot \Delta_t^{\tau}\dot{\boldsymbol{\sigma}}_0\phi^2 \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad - \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \mathsf{A}\dot{\boldsymbol{\sigma}} \cdot \Delta_t^{\tau}\dot{\boldsymbol{\sigma}}\phi^2 + \mathsf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^{\tau}\dot{\boldsymbol{\xi}}\phi^2 \, \mathrm{d}x + \mathsf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^{\tau}(\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\phi^2 \, \mathrm{d}\tau\, \mathrm{d}t \\
& \quad - \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_j^h((\dot{\boldsymbol{u}}_0 - \dot{\boldsymbol{u}}) \otimes \nabla\phi^2) \cdot ((\dot{\boldsymbol{\sigma}}(t+\tau) - \dot{\boldsymbol{\sigma}}_0(t+\tau))) \, \mathrm{d}x\, \mathrm{d}\tau\, \mathrm{d}t.
\end{aligned}
$$

Next, we apply the Hölder inequality, use the symmetry of $\mathbf{A}$ and $\mathbf{H}$, the assumption on $\boldsymbol{u}_0$ and $\boldsymbol{\sigma}_0$, the uniform bounds (2.13)–(2.14) and the time regularity estimate (2.28) to conclude

$$
\begin{aligned}
&\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\varepsilon(\dot{\boldsymbol{u}}_0-\dot{\boldsymbol{u}})\cdot\Delta_{t,j}^{\tau,h}(\dot{\boldsymbol{\sigma}}-\dot{\boldsymbol{\sigma}}_0)\phi^2-\varepsilon(\dot{\boldsymbol{u}}_0)\cdot\Delta_{t,j}^{\tau,h}\dot{\boldsymbol{\sigma}}\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&\leq\int_0^{T-h}\int_0^h h\|\varepsilon(\dot{\boldsymbol{u}}_0)\phi^2\|_{1,2}\|\dot{\boldsymbol{\sigma}}(t+\tau)\|_2+\tau\|\varepsilon(\dot{\boldsymbol{u}})-\varepsilon(\dot{\boldsymbol{u}}_0)\|_2\|\ddot{\boldsymbol{\sigma}}_0\|_2\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad+\frac{1}{2}\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\mathbf{A}\Delta_t^{\tau}\dot{\boldsymbol{\sigma}}\cdot\Delta_t^{\tau}\dot{\boldsymbol{\sigma}}\phi^2+\mathbf{H}\Delta_t^{\tau}\dot{\boldsymbol{\xi}}\cdot\Delta_t^{\tau}\dot{\boldsymbol{\xi}}\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad-\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\Delta_t^{\tau}(\dot{\boldsymbol{\sigma}}_D-\dot{\boldsymbol{\xi}}_D)\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad-\frac{1}{2}\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\Delta_t^{\tau}(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}})\phi^2+\Delta_t^{\tau}(\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}})\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad+\int_0^{T-h}\int_0^h h\|(\dot{\boldsymbol{u}}_0-\dot{\boldsymbol{u}})\otimes\nabla\phi^2\|_{1,2}\|\dot{\boldsymbol{\sigma}}(t+\tau)-\dot{\boldsymbol{\sigma}}_0(t+\tau)\|_2\,\mathrm{d}\tau\,\mathrm{d}t\\
&\leq Ch^2-\frac{1}{2}\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\Delta_t^{\tau}(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}})\phi^2+\Delta_t^{\tau}(\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}})\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad+\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\Delta_t^{\tau}(\dot{\boldsymbol{\sigma}}_D-\dot{\boldsymbol{\xi}}_D)\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t.
\end{aligned}
\tag{2.32}
$$

There are still remaining two terms on the right hand side. But for the first one we can use exactly the same computation as in (2.23) and observe

$$
\left|\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\Delta_t^{\tau}(\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}})\phi^2+\Delta_t^{\tau}(\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}})\phi^2\right|\leq Ch^2.
\tag{2.33}
$$

To estimate the remaining term in (2.32), we again use the abbreviation $\boldsymbol{\beta}:=\boldsymbol{\sigma}_D-\boldsymbol{\xi}_D$ and following the computation in (2.24) we deduce

$$
\begin{aligned}
&2\int_0^{T-h}\int_0^h\int_{\mathcal{O}}\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\Delta_t^{\tau}(\dot{\boldsymbol{\sigma}}_D-\dot{\boldsymbol{\xi}}_D)\phi^2\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}t\\
&=\int_0^{T-h}\int_{\mathcal{O}}2\mu^{-1}(|\boldsymbol{\beta}|-1)_+\frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|}\cdot\left(\int_0^h\Delta_t^{\tau}\dot{\boldsymbol{\beta}}\,\mathrm{d}\tau\right)\phi^2\,\mathrm{d}x\,\mathrm{d}t\\
&\leq\int_0^h\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\beta}(T-h+t)|-1)_+^2\phi^2-(|\boldsymbol{\beta}(T-h)|-1)_+^2\phi^2\,\mathrm{d}x\,\mathrm{d}t\\
&\quad-\int_0^h\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\beta}(t)|-1)_+^2\phi^2-(|\boldsymbol{\beta}(0)|-1)_+^2\phi^2\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
\tag{2.34}
$$

Similarly as before (compare with (2.25)), we also have the identity

$$
\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{O}}\mu^{-1}(|\boldsymbol{\beta}|-1)_+\phi^2\,\mathrm{d}x\\
&=\int_{\mathcal{O}}(\varepsilon(\dot{\boldsymbol{u}})\cdot\dot{\boldsymbol{\sigma}}\phi^2-\mathbf{A}\dot{\boldsymbol{\sigma}}\cdot\dot{\boldsymbol{\sigma}}\phi^2-\mathbf{H}\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}}\phi^2\,\mathrm{d}x,
\end{aligned}
\tag{2.35}
$$

which after integration over arbitrary interval $(\tau, \tau + \alpha) \subset (0, T)$ and with the help of the a priori estimate (2.13) and (2.14), leads to

$$(2.36) \qquad \left| \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}(\alpha + \tau)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(\tau)| - 1)_+^2 \, \mathrm{d}x \right| \leq C\alpha.$$

Thus, using this estimate in (2.34), we see that

$$(2.37) \qquad \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \Delta_t^\tau (\dot{\boldsymbol{\sigma}}_D - \dot{\boldsymbol{\xi}}_D)\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t \leq Ch^2.$$

The second term in (2.30). Again here, we split the time and space shift and then use (2.33) and the spatial regularity of ϕ as follows

$$(2.38)$$
$$\int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_{t,j}^{\tau,h}(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}})\phi^2 + \Delta_{t,j}^{\tau,h}(\mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}})\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$= \int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\Delta_{t,j}^{\tau,h}(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}}) - \Delta_t^\tau(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}}))\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$+ \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_t^\tau(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}})\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$= - \int_0^{T-h} \int_0^h \int_{\mathcal{O}} (\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}})(t + \tau)(\Delta_j^{-h}\phi^2) \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$+ \int_0^{T-h} \int_0^h \int_{\mathcal{O}} \Delta_t^\tau(\mathbf{A}\dot{\boldsymbol{\sigma}} \cdot \dot{\boldsymbol{\sigma}} + \mathbf{H}\dot{\boldsymbol{\xi}} \cdot \dot{\boldsymbol{\xi}})\phi^2 \, \mathrm{d}x \, \mathrm{d}\tau \, \mathrm{d}t$$
$$\leq Ch^2 + Ch \int_0^T \int_0^h \|\dot{\boldsymbol{\sigma}}\|_2^2 + \|\dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t \leq Ch^2.$$

The first term in (2.30). We again use the abbreviation $\boldsymbol{\beta} := \boldsymbol{\sigma}_D - \boldsymbol{\xi}_D$. The we estimate the first term in (2.30) very similarly as in (2.24) as follows. First, we rewrite it in the following way

$$2 \int_0^{T-h} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\int_0^h \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\beta}} \, \mathrm{d}\tau \right) \phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
$$= 2 \int_0^{T-h} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\Delta_{t,j}^{h,h}\boldsymbol{\beta} - \Delta_j^h\boldsymbol{\beta} - h\dot{\boldsymbol{\beta}} \right) \phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
$$= 2 \int_0^{T-h} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \Delta_{t,j}^{h,h}\boldsymbol{\beta}\phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
$$- 2 \int_0^{T-h} \int_{\mathcal{O}} \left(\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \right)(t, x + he_j) \cdot \Delta_j^h\boldsymbol{\beta}\phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
$$- h \int_0^{T-h} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}| - 1)_+^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t$$
$$+ 2 \int_0^{T-h} \int_{\mathcal{O}} \Delta_j^h \left(\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \right) \cdot \Delta_j^h\boldsymbol{\beta}\phi^2 \, \mathrm{d}x \, \mathrm{d}t$$

and then using the convexity and the fact $|\boldsymbol{\beta}(0)| \leq 1$, we can estimate it as

$$
\begin{aligned}
2 \int_0^{T-h} & \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \cdot \left(\int_0^h \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\beta}} \, \mathrm{d}\tau \right) \phi^2 \, \mathrm{d}x \, \mathrm{d}t \\
& \leq \int_0^{T-h} \int_{\mathcal{O}} (\Delta_{t,j}^{h,h} (\mu^{-1}(|\boldsymbol{\beta}| - 1)_+^2) - \Delta_j^h (\mu^{-1}(|\boldsymbol{\beta}| - 1)_+^2)) \phi^2 \, \mathrm{d}x \, \mathrm{d}t \\
& \quad - h \int_{\mathcal{O}} (\mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(0)| - 1)_+^2) \phi^2 \, \mathrm{d}x \\
& \quad + 2 \int_0^{T-h} \int_{\mathcal{O}} \Delta_j^h \left(\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \right) \cdot \Delta_j^h \boldsymbol{\beta} \phi^2 \, \mathrm{d}x \, \mathrm{d}t \\
& = \int_0^{T-h} \int_{\mathcal{O}} \Delta_t^h (\mu^{-1}(|\boldsymbol{\beta}| - 1)_+^2)) \phi^2(x - he_j) \, \mathrm{d}x \, \mathrm{d}t \\
& \quad - h \int_{\mathcal{O}} (\mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(0)| - 1)_+^2) \phi^2(x - he_j) \, \mathrm{d}x \\
& \quad + h \int_{\mathcal{O}} (\mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(0)| - 1)_+^2) \Delta_j^{-h} \phi^2 \, \mathrm{d}x \\
& \quad + 2 \int_0^{T-h} \int_{\mathcal{O}} \Delta_j^h \left(\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \right) \cdot \Delta_j^h \boldsymbol{\beta} \phi^2 \, \mathrm{d}x \, \mathrm{d}t.
\end{aligned}
\tag{2.39}
$$

Next, we estimate the remaining terms separately. First, using the very similar procedure as in (2.24) and consequent inequalities, we observe

$$
\begin{aligned}
\int_0^{T-h} & \int_{\mathcal{O}} \Delta_t^h (\mu^{-1}(|\boldsymbol{\beta}| - 1)_+^2)) \phi^2(x - he_j) \, \mathrm{d}x \, \mathrm{d}t \\
& - h \int_{\mathcal{O}} (\mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(0)| - 1)_+^2) \phi^2(x - he_j) \, \mathrm{d}x \leq Ch^2,
\end{aligned}
$$

where the constant C depends on data and on ϕ. Next, by uniform bound (2.26), we also have the estimate

$$
\begin{aligned}
h \int_{\mathcal{O}} & (\mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 - \mu^{-1}(|\boldsymbol{\beta}(0)| - 1)_+^2) \Delta_j^{-h} \phi^2 \, \mathrm{d}x \\
& \leq Ch^2 \|\nabla \phi^2\|_\infty \int_{\mathcal{O}} \mu^{-1}(|\boldsymbol{\beta}(T - h)| - 1)_+^2 \, \mathrm{d}x \leq Ch^2.
\end{aligned}
$$

Finally, for the remaining term, we apply the Δ_j^h to (2.1) and test by $\Delta_j^h \boldsymbol{\sigma}$ and by $\Delta_j^h \boldsymbol{\xi}$. Then we can repeat the same procedure as in estimating the tangential derivatives for $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$ and deduce again

$$
2 \int_0^{T-h} \int_{\mathcal{O}} \Delta_j^h \left(\mu^{-1}(|\boldsymbol{\beta}| - 1)_+ \frac{\boldsymbol{\beta}}{|\boldsymbol{\beta}|} \right) \cdot \Delta_j^h \boldsymbol{\beta} \phi^2 \, \mathrm{d}x \, \mathrm{d}t \leq Ch^2.
$$

Summarizing the estimates. Hence, if we use the above estimates in (2.30), we get

$$
\frac{1}{h^2} \int_0^{T-h} \int_0^h \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t \leq C
\tag{2.40}
$$

for arbitrary $j = 1, \ldots, d - 1$ and localization function ϕ. Then, by a simple inequality, we deduce

(2.41)
$$
\frac{1}{h} \int_0^{T-h} \|\phi \Delta_j^h \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi \Delta_j^h \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}t
$$
$$
= \frac{1}{h^2} \int_0^{T-h} \int_0^h \|\phi \Delta_j^h \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi \Delta_j^h \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t
$$
$$
\leq \frac{C}{h^2} \int_0^{T-h} \int_0^h \|\phi (\Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\sigma}} - \Delta_j^h \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi (\Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\xi}} - \Delta_j^h \dot{\boldsymbol{\xi}}\|_2^2
$$
$$
+ \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t
$$
$$
\leq \frac{C}{h^2} \int_0^{T-h} \int_0^h \|\Delta_t^\tau \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi (\Delta_t^\tau \dot{\boldsymbol{\xi}}\|_2^2 + \|\phi \Delta_{t,j}^{\tau,h} \dot{\boldsymbol{\xi}}\|_2^2 \, \mathrm{d}\tau \, \mathrm{d}t \leq C,
$$

where the last inequality follows from time regularity estimates and (2.40).

3. Normal derivatives estimates

In this final part, we derive the estimate for the normal derivative. Note that the estimate is again uniform with respect to μ. We also keep the notation near the boundary and use the function ϕ which is compactly supported in a cube $(-1 + h_0, 1 - h_0)^d$ and equal to one in a cube $(-1 + 2h_0, 1 - 2h_0)^d$. We start this part by using already proven time fractional regularity to transfer also the spatial regularity.

3.1. Estimate for for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ via time interpolation.
Here, we show how the fractional estimates in the d-direction for $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$ and the fractional estimates in the time direction for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ can improve the spatial regularity of $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$. The key estimate is formulated in the following.

Lemma 3.1. *Let ϕ be as above. Then for any $\delta \in (0, \frac{1}{3})$, the solution satisfies*

(3.1)
$$
\int_0^T \int_{\mathcal{O}} |\Delta_d^h \dot{\boldsymbol{\sigma}}|^2 \phi^2 + |\Delta_d^h \dot{\boldsymbol{\xi}}|^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t
$$
$$
\leq C(\delta) \left(\int_0^T \int_{\mathcal{O}} |\Delta_d^h \boldsymbol{\sigma}|^2 \phi^2 + |\Delta_d^h \boldsymbol{\xi}|^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{3} - \delta},
$$

where the constant $C(\delta)$ depends only on data and explodes as $\delta \to 0_+$.

Proof. The proof is based on the interpolation of Bochner-Sobolev spaces. We recall the classical Bochner-Sobolev interpolation between $W^{\alpha,2}$ and L^2 and also the Nikolskii-Sobolev embedding $N^{\frac{3}{2},2} \hookrightarrow W^{\alpha,2}$ valid for all $\alpha < \frac{3}{2}$ to get (we use any $\alpha \in (1, \frac{3}{2})$)

$$
\int_0^T \|\dot{f}\|_2^2 \, \mathrm{d}t \leq C \|f\|_{L^2(0,T;L^2(\mathcal{O}))}^{2 - \frac{2}{\alpha}} \|f\|_{W^{\alpha,2}(0,T;L^2(\mathcal{O}))}^{\frac{2}{\alpha}}
$$

(3.2)
$$
\leq C \|f\|_{L^2(0,T;L^2(\mathcal{O}))}^{2 - \frac{2}{\alpha}} \|f\|_{N^{\frac{3}{2},2}(0,T;L^2(\mathcal{O}))}^{\frac{2}{\alpha}}
$$
$$
\leq C \int_0^T \|f\|_2^2 \, \mathrm{d}t + C(\alpha) \left(\sup_{h \in (0,T)} \int_0^{T-h} \frac{\|\Delta_t^h \dot{f}\|_2^2}{h} \, \mathrm{d}t \right)^{\frac{1}{\alpha}} \left(\int_0^T \|f\|_2^2 \, \mathrm{d}t \right)^{1 - \frac{1}{\alpha}},
$$

where the constant $C(\alpha)$ explodes as $\alpha \to \frac{3}{2}$. Consequently, using the above estimate on $f := \Delta_d^h \boldsymbol{\sigma} \phi$ and $\Delta_d^h \boldsymbol{\xi} \phi$ and using the a priori bound (2.29), we see that

$$
\begin{aligned}
&\int_0^T \int_{\mathcal{O}} |\Delta_d^h \dot{\boldsymbol{\sigma}}|^2 \phi^2 + |\Delta_d^h \dot{\boldsymbol{\xi}}|^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad \leq C \int_0^T \int_{\mathcal{O}} |\Delta_d^h \boldsymbol{\sigma}|^2 \phi^2 + |\Delta_d^h \boldsymbol{\xi}|^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad\quad + C(\alpha) \left(\int_0^T \int_{\mathcal{O}} |\Delta_d^h \boldsymbol{\sigma}|^2 \phi^2 + |\Delta_d^h \boldsymbol{\xi}|^2 \phi^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{1 - \frac{1}{\alpha}}.
\end{aligned}
$$

(3.3)

Since $\alpha \in (1, \frac{3}{2})$ can be arbitrary and we have the control (2.13), the estimate (3.1) follows. $\qquad\square$

3.2. **First estimate for $\boldsymbol{\sigma}$ and $\boldsymbol{\xi}$.** Here, we start with an estimate, that directly leads to $\frac{1}{2}$ regularity of the stress and hardening, but it will also serve later for the bootstrap argument.

Lemma 3.2. *Let ϕ be chosen such that $\phi(x', s)$ is independent of s for all $s \in [0, h_0]$ with $h_0 > 0$. Then for all $h \in (0, h_0)$ the following estimate holds*

$$
\begin{aligned}
&\sup_{t \in (0,T)} \|\phi \Delta_d^h \boldsymbol{\sigma}(t)\|_2^2 + \|\phi \Delta_d^h \boldsymbol{\xi}(t)\|_2^2 \\
&\qquad \leq C h^{\frac{3}{2}} + C h \int_0^T \left(\int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \mathrm{d}t.
\end{aligned}
$$

(3.4)

In addition, if the $\operatorname{supp} \phi \cap \overline{\mathcal{O}}_N = \emptyset$ then we have

$$
\begin{aligned}
&\sup_{t \in (0,T)} \|\phi \Delta_d^h \boldsymbol{\sigma}(t)\|_2^2 + \|\phi \Delta_d^h \boldsymbol{\xi}(t)\|_2^2 \\
&\qquad \leq C h^{\frac{3}{2}} + C h \int_0^T \left(\int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) \phi|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \mathrm{d}t.
\end{aligned}
$$

(3.5)

Proof. We apply the operator Δ_d^h to both equations in (2.1) and and test by $\Delta_d^h \boldsymbol{\sigma} \phi^2$ and $\Delta_d^h \boldsymbol{\xi} \phi^2$ respectively. Note that since $x + h e_d \in \mathcal{O}$ such operation is well defined. Thus, doing so, we observe

$$
\begin{aligned}
&\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{O}} \mathbf{A}(\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0) \cdot (\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0) \phi^2 \, \mathrm{d}x \\
&\quad + \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{H}(\Delta_d^h \boldsymbol{\xi} - \Delta_d^h \boldsymbol{\xi}_0) \cdot (\Delta_d^h \boldsymbol{\xi} - \Delta_d^h \boldsymbol{\xi}_0) \phi^2 \, \mathrm{d}x \\
&\quad + \int_{\mathcal{O}} \Delta_d^j \left(\mu^{-1} (|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D| - 1)_+ \frac{\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D}{|\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D|} \right) \cdot \Delta_d^h (\boldsymbol{\sigma}_D - \boldsymbol{\xi}_D) \phi^2 \, \mathrm{d}x \\
&= \int_{\mathcal{O}} \Delta_d^h \varepsilon(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot \Delta_d^h (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \phi^2 \, \mathrm{d}x \\
&\quad + \int_{\mathcal{O}} \Delta_d^h \varepsilon(\dot{\boldsymbol{u}}_0) \cdot \Delta_d^h (\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) \phi^2 - \Delta_d^h (\mathbf{A} \dot{\boldsymbol{\sigma}}_0 + \mathbf{H} \dot{\boldsymbol{\xi}}_0) \cdot (\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0) \phi^2 \, \mathrm{d}x.
\end{aligned}
$$

(3.6)

The terms on the left hand side are those from which we read information. The second term on the right hand side can be easily estimated and prepared for the

Grownall lemma as follows

$$
\begin{aligned}
(3.7) \quad & \int_{\mathcal{O}} \Delta_d^h \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0) \cdot \Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 - \Delta_d^h(\mathbf{A}\dot{\boldsymbol{\sigma}}_0 + \mathbf{H}\dot{\boldsymbol{\xi}}_0) \cdot (\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0)\phi^2 \, \mathrm{d}x \\
& \leq \|(\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0)\phi\|_2 (\|\Delta_d^h \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}_0)\|_2 + \|\Delta_d^h(\mathbf{A}\dot{\boldsymbol{\sigma}}_0 + \mathbf{H}\dot{\boldsymbol{\xi}}_0)\|_2) \\
& \leq Ch(\|\nabla\dot{\boldsymbol{\sigma}}_0\|_2 + \|\nabla\dot{\boldsymbol{\xi}}_0\|_2)\|(\Delta_d^h \boldsymbol{\sigma} - \Delta_d^h \boldsymbol{\sigma}_0)\phi\|_2,
\end{aligned}
$$

where the last inequality follows from Sobolev characterization and the assumptions on data. Thus, we can now focus on the most critical term in (3.6), which is the first integral on the right hand side. To simplify the notation we use the abbreviation $x' := (x_1, \ldots, x_{d-1})$. Then using integration by parts, we observe

$$
\begin{aligned}
(3.8) \quad & \int_{\mathcal{O}} \Delta_d^h \boldsymbol{\varepsilon}(\boldsymbol{u} - \dot{\boldsymbol{u}}_0) \cdot \Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 \, \mathrm{d}x = \int_{\mathcal{O}} \nabla\Delta_d^h(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \cdot \Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi^2 \, \mathrm{d}x \\
& = -2\int_{\mathcal{O}} (\Delta_d^h(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0) \otimes \nabla\phi) \cdot \Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi \, \mathrm{d}x \\
& \quad + \sum_{i=1}^{d} \int_{\{x_d=0\}} \Delta_d^h(\dot{\boldsymbol{u}}_i - \dot{\boldsymbol{u}}_{0i})\Delta_d^h(\boldsymbol{\sigma}_{id} - \boldsymbol{\sigma}_{0id})\phi^2 \, \mathrm{d}x' \\
& \leq C\|\Delta_d^h(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\|_2\|\Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2 \\
& \quad + \sum_{i=1}^{d} \int_{\{x_d=0\}} \Delta_d^h(\dot{\boldsymbol{u}}_i - \dot{\boldsymbol{u}}_{0i})\Delta_d^h(\boldsymbol{\sigma}_{id} - \boldsymbol{\sigma}_{0id})\phi^2 \, \mathrm{d}x'
\end{aligned}
$$

The first term on the right hand side can be estimated by the use of characterization of Sobolev functions and the a priori bound (2.14) as

$$
\begin{aligned}
(3.9) \quad & \|\Delta_d^h(\dot{\boldsymbol{u}} - \dot{\boldsymbol{u}}_0)\|_2\|\Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2 \leq Ch(\|\nabla\dot{u}\|_2 + \|\nabla\dot{u}_0\|_2)\|\Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2 \\
& \leq Ch\|\Delta_d^h(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0)\phi\|_2.
\end{aligned}
$$

For the second term on the right hand side, we first replace the differences by the corresponding integral, then use the fact that $\mathrm{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_0) = 0$ (we also assume that $h \ll 1$ so that $\phi(x', 0) = \phi(x', x_d)$ for arbitrary $x_d \in (0, h)$)

$$
\begin{aligned}
& \int_{\{x_d=0\}} \Delta_d^h(\dot{\boldsymbol{u}}_i - \dot{\boldsymbol{u}}_{0i})\Delta_d^h(\boldsymbol{\sigma}_{id} - \boldsymbol{\sigma}_{0id})\phi^2 \, \mathrm{d}x' \\
& = \int_{\mathbb{R}^{d-1}} \left(\int_0^h D_d(\dot{\boldsymbol{u}}_i - \dot{\boldsymbol{u}}_{0i}) \, \mathrm{d}x_d\right) \left(\int_0^h D_d(\boldsymbol{\sigma}_{id} - \boldsymbol{\sigma}_{0id}) \, \mathrm{d}x_d\right) \phi^2(x', 0) \, \mathrm{d}x' \\
& \leq \int_{\mathbb{R}^{d-1}} \left(\int_0^h |D_d\dot{\boldsymbol{u}}\phi| + |\nabla(\dot{\boldsymbol{u}}_0\phi)| \, \mathrm{d}x_d\right) \left|\sum_{j=1}^{d-1} \int_0^h D_j(\boldsymbol{\sigma}_{ij} - \boldsymbol{\sigma}_{0ij}) \, \mathrm{d}x_d\right| \phi(x', 0) \, \mathrm{d}x'.
\end{aligned}
$$

Then, we apply the Hölder inequality to conclude

$$
\int_{\{x_d=0\}} \Delta_d^h(\dot{\boldsymbol{u}}_i - \dot{\boldsymbol{u}}_{0i})\Delta_d^h(\boldsymbol{\sigma}_{id} - \boldsymbol{\sigma}_{0id})\phi^2 \,\mathrm{d}x'
$$

$$
\leq Ch \sum_{j=1}^{d-1} \int_{\mathbb{R}^{d-1}} \left(\int_0^h (|D_d\dot{\boldsymbol{u}}\phi|^2 + |\nabla(\dot{\boldsymbol{u}}_0\phi)|^2)\,\mathrm{d}x_d \right)^{\frac{1}{2}} \times
$$

(3.10)
$$
\times \left(\int_0^h (|D_j\boldsymbol{\sigma}|^2 + |\nabla\boldsymbol{\sigma}_0|^2)\phi^2 \,\mathrm{d}x_d \right)^{\frac{1}{2}} \,\mathrm{d}x'
$$

$$
\leq Ch\left(\sum_{j=1}^{d-1} \|D_j\boldsymbol{\sigma}\phi\|_2 + \|\nabla\boldsymbol{\sigma}_0\|_2\right) \left(\int_{\mathcal{O}\cap\{x_d\in(0,h)\}} (|D_d\dot{\boldsymbol{u}}\phi|^2 + |\nabla(\dot{\boldsymbol{u}}_0\phi)|^2)\,\mathrm{d}x \right)^{\frac{1}{2}}
$$

$$
\leq Ch^{\frac{3}{2}} + Ch \left(\int_{\mathcal{O}\cap\{x_d\in(0,h)\}} |D_d\dot{\boldsymbol{u}}\phi|^2 \,\mathrm{d}x \right)^{\frac{1}{2}}
$$

Hence, using (3.7)–(3.10) in (3.6) and applying the Gronwall lemma, and using the fact that $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0$ and $\boldsymbol{\xi}(0) = \boldsymbol{\xi}_0$, and already proven a priori estimates, we deduce (3.4). In case that the support of ϕ is located just closed to the Dirichlet boundary, we may use the Korn inequality (or the trace theorem) and to replace $|D_d\dot{\boldsymbol{u}}\phi|$ by $|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\phi|$ in the above estimate and to conclude (3.5). $\qquad\square$

3.3. **Estimate for $\nabla\dot{\boldsymbol{u}}$ on a strip in terms of $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$.** In previous section, we deduce a uniform estimate for normal fractional derivatives in terms of $\nabla\dot{\boldsymbol{u}}$, the right hand side of (3.4)and (3.5), respectively. In this part, we show how this term can be estimated in terms of $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$. In fact, we prove two different estimates. The first one deals with the case that we have isotropic hardening and we are closed to the Neumann part of the boundary. The second case covers the kinematic hardening independently of Dirichlet or boundary data or the isotropic hardening in case of Dirichlet data.

Lemma 3.3. *For arbitrary $h \in (0, h_0)$ and $p > 2$, the solution satisfies*

$$
(3.11) \quad \int_0^T \left(\int_{\mathcal{O}\cap\{x_d\in(0,h)\}} |D_d\dot{\boldsymbol{u}}\phi|^2 \,\mathrm{d}x \right)^{\frac{1}{2}} \,\mathrm{d}t \leq Ch^{\frac{p-2}{2p}}(1 + \int_0^T \|\dot{\boldsymbol{\sigma}}\phi\|_p + \|\dot{\boldsymbol{\xi}}\phi\|_p \,\mathrm{d}t).
$$

In addition, if we consider the kinematic hardening, we have

(3.12)
$$
\int_0^T \left(\int_{\mathcal{O}\cap\{x_d\in(0,h)\}} |D_d\dot{\boldsymbol{u}}\phi|^2 \,\mathrm{d}x \right)^{\frac{1}{2}} \,\mathrm{d}t
$$

$$
\leq C(\delta)h^{\frac{p-2}{2p}} \left(1 + \sup_{s\in(0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi\Delta_d^s\boldsymbol{\sigma}|^2 + |\phi\Delta_d^s\boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta+\frac{p-2}{p})}} \,\mathrm{d}x\,\mathrm{d}t \right)^{\frac{1-3\delta}{6}}.
$$

Furthermore, for kinematic and isotropic hardening we also have

$$(3.13) \qquad \int_0^T \left(\fint_{\mathcal{O} \cap \{x_d \in (0,h)\}} |\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\phi|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \mathrm{d}t$$

$$\leq C(\delta) h^{\frac{p-2}{2p}} \left(1 + \sup_{s \in (0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \boldsymbol{\sigma}|^2 + |\phi \Delta_d^s \boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta + \frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1-3\delta}{6}}.$$

Proof. To prove the first case, we just use the Hölder and the Korn inequality as follows

$$(3.14)$$
$$\int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \leq C h^{\frac{p-2}{p}} \|D_d(\dot{\boldsymbol{u}}\phi)\|^2_{L^p(\mathcal{O})} \leq C h^{\frac{p-2}{p}} \|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}\phi)\|^2_{L^p(\mathcal{O})}$$

$$\leq C h^{\frac{p-2}{p}} (\|\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})\phi\|^2_{L^p(\mathcal{O})} + 1) \leq C h^{\frac{p-2}{p}} (1 + \|\dot{\boldsymbol{\sigma}}\phi\|^2_p + \|\dot{\boldsymbol{\xi}}\phi\|^2_p),$$

where for the last inequality we used the equation (1.20). Then we see that (3.11) directly follows.

Next, we focus on (3.12). First, we recall the fractional Sobolev embedding $W^{\frac{p-2}{2p},2}(0,1) \hookrightarrow L^p(0,1)$ valid for all $p \in [2,\infty)$. Then with the help of the Hölder inequality, we obtain for almost all x' and t that

$$\int_0^h |D_d \dot{\boldsymbol{u}} \phi(t,x',x_d)|^2 \, \mathrm{d}x_d \leq h^{\frac{p-2}{p}} \left(\int_0^{2h_0} |D_d \dot{\boldsymbol{u}} \phi(t,x',x_d)|^p \, \mathrm{d}x_d \right)^{\frac{2}{p}}$$

$$\leq C h^{\frac{p-2}{p}} \left(\int_0^{2h_0} |D_d \dot{\boldsymbol{u}} \phi(t,x',x_d)|^2 \, \mathrm{d}x_d \right.$$

$$\left. + \int_0^{2h_0} \int_0^{2h_0} \frac{|D_d \dot{\boldsymbol{u}} \phi(t,x',x_d) - D_d \dot{\boldsymbol{u}} \phi(t,x',y_d)|^2}{|x_d - y_d|^{1+\frac{p-2}{p}}} \, \mathrm{d}x_d \, \mathrm{d}y_d \right)$$

Hence, using the Fubini theorem and the Korn inequality, we can continue with the estimate of the full integral as follows

$$\int_0^T \int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \, \mathrm{d}t \leq C h^{\frac{p-2}{p}} \left(\int_0^T \int_{(-1,1)^d} \int_0^{2h_0} |D_d \dot{\boldsymbol{u}} \phi(t,x',x_d)|^2 \, \mathrm{d}x \, \mathrm{d}t \right.$$

$$\left. + \int_0^T \int_{(-1,1)^d} \int_0^{2h_0} \int_0^{2h_0} \frac{|D_d \dot{\boldsymbol{u}} \phi(t,x',x_d) - D_d \dot{\boldsymbol{u}} \phi(t,x',y_d)|^2}{|x_d - y_d|^{1+\frac{p-2}{p}}} \, \mathrm{d}x_d \, \mathrm{d}y_d \, \mathrm{d}x' \, \mathrm{d}t \right)$$

$$\leq C h^{\frac{p-2}{p}} \left(1 + \int_0^{4h_0} \int_0^T \int_{\mathcal{O}} \frac{|D_d \Delta_d^s(\dot{\boldsymbol{u}}\phi)|^2}{|s|^{1+\frac{p-2}{p}}} \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \right)$$

$$\leq C h^{\frac{p-2}{p}} \left(1 + \int_0^{4h_0} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_d^s \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}\phi)|^2}{|s|^{1+\frac{p-2}{p}}} \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \right)$$

$$\leq C h^{\frac{p-2}{p}} \left(1 + \int_0^{4h_0} \int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})|^2}{|s|^{1+\frac{p-2}{p}}} \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \right)$$

$$\leq C h^{\frac{p-2}{p}} \left(1 + \int_0^{4h_0} \int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \dot{\boldsymbol{\sigma}}|^2 + |\phi \Delta_d^s \dot{\boldsymbol{\xi}}|^2}{|s|^{1+\frac{p-2}{p}}} \, \mathrm{d}x \, \mathrm{d}t \, \mathrm{d}s \right),$$

where we used the equation to evaluate $\varepsilon(\dot{\boldsymbol{u}})$ in terms of $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$. Note that at this step we use the fact that we deal with the kinematic hardening. Now, we use Lemma 3.1 to replace time derivative on the right hand side. Hence, doing so, we observe that for all $\delta > 0$, we have

$$
\int_0^T \int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \, \mathrm{d}t
$$

$$
\leq C(\delta) h^{\frac{p-2}{p}} \left(1 + \int_0^{4h_0} \frac{1}{s^{1-\delta}} \left(\int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \boldsymbol{\sigma}|^2 + |\phi \Delta_d^s \boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta + \frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{3}-\delta} \mathrm{d}s \right)
$$

$$
\leq C(\delta) h^{\frac{p-2}{p}} \left(1 + \sup_{s \in (0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \boldsymbol{\sigma}|^2 + |\phi \Delta_d^s \boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta + \frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{3}-\delta}.
$$

Thus, using the Hölder inequality and the above estimate, we have

$$
\int_0^T \left(\int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \mathrm{d}t \leq C \left(\int_0^T \int_{\mathcal{O} \cap \{x_d \in (0,h)\}} |D_d \dot{\boldsymbol{u}} \phi|^2 \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}}
$$

$$
\leq C(\delta) h^{\frac{p-2}{2p}} \left(1 + \sup_{s \in (0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi \Delta_d^s \boldsymbol{\sigma}|^2 + |\phi \Delta_d^s \boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta + \frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{6}-\frac{\delta}{2}},
$$

which is (3.12). To obtain (3.13), we proceed similarly, with the only change that from the beginning we have the point-wise estimate $|\varepsilon(\dot{\boldsymbol{u}})| \leq C(|\dot{\boldsymbol{\sigma}}| + |\dot{\boldsymbol{\xi}}|)$, which follows from (1.20) and (1.24), respectively.

$\square$

3.4. Estimate for for $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\xi}}$ via anisotropic embedding.

Lemma 3.4. *Let $p \in (2, \frac{2(d-1)}{d-2})$ and $\beta > 0$ be given as*

$$
\beta := \frac{p-2}{4(d-1) - 2p(d-2)} \tag{3.15}
$$

Then for any solution and any $\delta > 0$ there holds
(3.16)

$$
\int_0^T \|\dot{\boldsymbol{\sigma}} \phi\|_p + \|\dot{\boldsymbol{\xi}} \phi\|_p \, \mathrm{d}t \leq \frac{C}{\delta} \left(1 + \sup_{h \in (0,1)} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_d^h \dot{\boldsymbol{\sigma}} \phi|^2 + |\Delta_d^h \dot{\boldsymbol{\xi}} \phi|^2}{h^{2(\beta + \delta)}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1}{2}}.
$$

Proof. We use the following version of anisotropic embedding. We assume that $f \in L^2(0, 1; L^2((-1, 1)^d))$ and denote $B := (-1, 1)^{d-1}$. We define $\lambda \in (0, 1)$ by the relation

$$
\lambda := \frac{1}{2\beta(d-1) + 1}
$$

and using the Sobolev embedding, we have (recall the definition of β in (3.15))

$$
W^{\lambda \beta, 2}(0, 1) \hookrightarrow L^p(0, 1),
$$
(3.17)
$$
W^{\frac{1-\lambda}{2}, 2}(B) \hookrightarrow L^p(B).
$$

Then, we can use the cross-interpolation in Sobolev-Bochner spaces and the above embedding to observe

$$
\|f\|_p^2 \leq C\|f\|_{W^{\lambda \beta, 2}(0,1; W^{\frac{1-\lambda}{2}, 2}(B))}^2 \leq C(\|f\|_{L^2(0,1; W^{\frac{\alpha}{2}, 2}(B))}^2 + \|f\|_{W^{\frac{\alpha \lambda \beta}{\alpha - 1 + \lambda}, 2}(0,1; L^2(B))}^2)
$$

for arbitrary $1 - \lambda < \alpha < 1$. Next, we use the above inequality for $\dot{\boldsymbol{\sigma}}\phi$ and integrate also over time $t \in (0, T)$. Thus, using also the definition of fractional Sobolev norm, we have (using also the properties of ϕ and the a priori estimate (2.13))

$$
\begin{aligned}
\int_0^T \|\dot{\boldsymbol{\sigma}}\phi\|_p^2 &\leq C \int_0^T \|\dot{\boldsymbol{\sigma}}\phi\|_{L^2(0,1;W^{\frac{\alpha}{2},2}(B))}^2 + \|\dot{\boldsymbol{\sigma}}\phi\|_{W^{\frac{\alpha\lambda\beta}{\alpha-1+\lambda},2}(0,1;L^2(B))}^2 \, \mathrm{d}t \\
&\leq C + C \int_0^T \int_0^1 \int_B \int_B \frac{|\dot{\boldsymbol{\sigma}}\phi(t,x',x_d) - \dot{\boldsymbol{\sigma}}\phi(t,y',x_d)|^2}{|x'-y'|^{d-1+\alpha}} \, \mathrm{d}x' \, \mathrm{d}y' \, \mathrm{d}x_d \, \mathrm{d}t \\
&\quad + C \int_0^T \int_0^1 \int_0^1 \int_B \frac{|\dot{\boldsymbol{\sigma}}\phi(t,x',x_d) - \dot{\boldsymbol{\sigma}}\phi(t,x',y_d)|^2}{|x_d - y_d|^{1+\frac{2\alpha\lambda\beta}{\alpha-1+\lambda}}} \, \mathrm{d}x' \, \mathrm{d}x_d \, \mathrm{d}y_d \, \mathrm{d}t \\
&\leq C + C \int_0^T \int_{\mathcal{O}} \int_{2B} \frac{|\dot{\boldsymbol{\sigma}}\phi(t,x'+z',x_d) - \dot{\boldsymbol{\sigma}}\phi(t,x',x_d)|^2}{|z'|^{d-1+\alpha}} \, \mathrm{d}z' \, \mathrm{d}x \, \mathrm{d}t \\
&\quad + C \int_0^T \int_{\mathcal{O}} \int_0^1 \frac{|\dot{\boldsymbol{\sigma}}\phi(t,x',x_d) - \dot{\boldsymbol{\sigma}}\phi(t,x',x_d+h)|^2}{h^{1+\frac{2\alpha\lambda\beta}{\alpha-1+\lambda}}} \, \mathrm{d}h \, \mathrm{d}x \, \mathrm{d}t \\
&\leq C + C \sup_{i=1,\ldots,d-1} \sup_{h\in(0,1)} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_i^h \dot{\boldsymbol{\sigma}}\phi|^2}{h} \, \mathrm{d}x \, \mathrm{d}t \int_{2B} \frac{1}{|z'|^{d-2+\alpha}} \, \mathrm{d}z' \\
&\quad + C \sup_{h\in(0,1)} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_d^h \dot{\boldsymbol{\sigma}}\phi|^2}{h^{2(\beta+\delta)}} \, \mathrm{d}x \, \mathrm{d}t \int_0^1 \frac{1}{s^{1-2(\beta+\delta)+\frac{2\alpha\lambda\beta}{\alpha-1+\lambda}}} \, \mathrm{d}s \\
&\leq \frac{C}{1-\alpha} + \frac{C}{2(\beta+\delta) - \frac{2\alpha\lambda\beta}{\alpha-1+\lambda}} \sup_{h\in(0,1)} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_d^h \dot{\boldsymbol{\sigma}}\phi|^2}{h^{2(\beta+\delta)}} \, \mathrm{d}x \, \mathrm{d}t.
\end{aligned}
$$

Consequently, for any $\delta > 0$, we can find $\alpha \in (0,1)$ such that

$$
2(\beta + \delta) > \frac{2\alpha\lambda\beta}{\alpha - 1 + \lambda}
$$

and (3.16) for $\dot{\boldsymbol{\sigma}}$ follows by Hölder inequality. The same scheme is used for the estimate for $\dot{\boldsymbol{\xi}}$.

$\square$

3.5. Final estimate for normal derivatives - the case of Dirichlet boundary or the case of kinematic hardening.
In this part we finish the proof of the main theorem. In particular, we focus on the normal derivative estimates stated in (1.23) and (1.27). We start with (3.4) to conclude the starting estimate

$$
(3.18) \qquad \sup_{t\in(0,T)} \sup_{h\in(0,1)} \frac{\|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2}{h} \leq C.
$$

Then, we use (3.4) (in case of kinematic hardening) or in (3.5) (in case of Dirichlet boundary condition), and the term on the right hand side is replaced by the corresponding estimates in (3.12) and (3.13), respectively. Thus we conclude in both cases that

$$
\sup_{t\in(0,T)} \|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2
$$

$$
\leq Ch^{\frac{3}{2}} + C(\delta)h^{1+\frac{p-2}{2p}} \left(1 + \sup_{s\in(0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi\Delta_d^s\boldsymbol{\sigma}|^2 + |\phi\Delta_d^s\boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta+\frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t \right)^{\frac{1-3\delta}{6}}.
$$

Hence, it follows that

(3.19)
$$\sup_{t\in(0,T)} \sup_{h\in(0,4h_0)} \int_{\mathcal{O}} \frac{|\phi\Delta_d^h\boldsymbol{\sigma}(t)|^2 + |\phi\Delta_d^h\boldsymbol{\xi}(t)|^2}{h^{1+\frac{p-2}{2p}}} \, \mathrm{d}x$$
$$\le C(\delta) \left(1 + \sup_{s\in(0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi\Delta_d^s\boldsymbol{\sigma}|^2 + |\phi\Delta_d^s\boldsymbol{\xi}|^2}{|s|^{\frac{3}{1-3\delta}(\delta+\frac{p-2}{p})}} \, \mathrm{d}x \, \mathrm{d}t\right)^{\frac{1-3\delta}{6}},$$

provided that the right hand side is finite. Hence, we can start the iteration with (3.18). In the first step, thanks to (3.18), we can set in (3.19) arbitrary $p < 3$ and find sufficiently small $\delta > 0$ such that

$$\frac{3}{1-3\delta}\left(\delta + \frac{p-2}{p}\right) < 1$$

and we immediately get an improvement of (3.18). Consequently, iterating such procedure is possible as long as

$$1 + \frac{p-2}{2p} > \frac{3(p-2)}{p} \qquad \Leftrightarrow \qquad \frac{1}{5} > \frac{p-2}{2p}.$$

Thus, it follows that we are able to obtain

(3.20)
$$\sup_{t\in(0,T)} \sup_{h\in(0,4h_0)} \int_{\mathcal{O}} \frac{|\phi\Delta_d^h\boldsymbol{\sigma}(t)|^2 + |\phi\Delta_d^h\boldsymbol{\xi}(t)|^2}{h^{\frac{6}{5}-\delta}} \, \mathrm{d}x \le C(\delta)$$

and by (3.1) we also have

(3.21)
$$\sup_{h\in(0,4h_0)} \int_0^T \int_{\mathcal{O}} \frac{|\phi\Delta_d^h\dot{\boldsymbol{\sigma}}|^2 + |\phi\Delta_d^h\dot{\boldsymbol{\xi}}|^2}{h^{\frac{2}{5}-\delta}} \, \mathrm{d}x \, \mathrm{d}t \le C(\delta),$$

which finishes the proof for kinematic hardening or the case of Dirichlet boundary conditions.

3.6. Final estimate for normal derivatives - the case of Neumann boundary and the isotropic hardening.
In this case we use the anisotropic embedding. We again start with (3.4) to conclude the starting estimate

(3.22)
$$\sup_{t\in(0,T)} \sup_{h\in(0,1)} \frac{\|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2}{h} \le C.$$

Then we start with iteration. Using (3.4) and (3.11), we observe that

(3.23)
$$\sup_{t\in(0,T)} \sup_{h\in(0,1)} \frac{\|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2}{h^{1+\frac{p-2}{2p}}} \le C\left(1 + \int_0^T \|\dot{\boldsymbol{\sigma}}\phi\|_p + \|\dot{\boldsymbol{\xi}}\phi\|_p \, \mathrm{d}t\right).$$

Next, recalling the definition of β, see (3.15),

$$\beta := \frac{p-2}{4(d-1) - 2p(d-2)},$$

and combining (3.16) and (3.1) and the Young inequality, we also have

(3.24)
$$\int_0^T \|\dot{\boldsymbol{\sigma}}\phi\|_p + \|\dot{\boldsymbol{\xi}}\phi\|_p \, \mathrm{d}t \le \frac{C}{\delta}\left(1 + \sup_{h\in(0,1)} \int_0^T \int_{\mathcal{O}} \frac{|\Delta_d^h\boldsymbol{\sigma}\phi|^2 + |\Delta_d^h\boldsymbol{\xi}\phi|^2}{h^{6(\beta+2\delta)}} \, \mathrm{d}x \, \mathrm{d}t\right),$$

where $\delta > 0$ is arbitrary. Thus, combining (3.23) and (3.24), we have

$$
\begin{aligned}
(3.25) \quad & \sup_{t\in(0,T)} \sup_{h\in(0,1)} \frac{\|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2}{h^{1+\frac{p-2}{2p}}} \\
& \leq \frac{C}{\delta}\left(1 + \sup_{h\in(0,1)}\int_0^T\int_{\mathcal{O}}\frac{|\Delta_d^h\boldsymbol{\sigma}\phi|^2 + |\Delta_d^h\boldsymbol{\xi}\phi|^2}{h^{6(\beta+2\delta)}}\,\mathrm{d}x\,\mathrm{d}t\right).
\end{aligned}
$$

Thus, we can again start with iteration, which is possible as long as

$$
1 + \frac{p-2}{2p} > 6\beta \quad \Leftrightarrow \quad \beta < \frac{2d-7+\sqrt{1+4d^2+20d}}{24(d-1)}.
$$

Consequently, we have the final estimate (here $\delta \in (0,1)$ is arbitrary)

$$
\sup_{t\in(0,T)} \sup_{h\in(0,1)} \frac{\|\phi\Delta_d^h\boldsymbol{\sigma}(t)\|_2^2 + \|\phi\Delta_d^h\boldsymbol{\xi}(t)\|_2^2}{h^{\frac{2d-7+\sqrt{1+4d^2+20d}}{4(d-1)}-\delta}} \leq \frac{C}{\delta}.
$$

and thanks to time interpolation (3.1), it leads to

$$
\sup_{h\in(0,1)}\int_0^T \frac{\|\phi\Delta_d^h\dot{\boldsymbol{\sigma}}\|_2^2 + \|\phi\Delta_d^h\dot{\boldsymbol{\xi}}\|_2^2}{h^{\frac{2d-7+\sqrt{1+4d^2+20d}}{12(d-1)}-\delta}}\,\mathrm{d}t \leq \frac{C}{\delta},
$$

where the constant $C(\delta)$ explodes as $\delta \to 0_+$. Hence, the proof is complete.

References

[1] H.-D. Alber. *Materials with memory*, volume 1682 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998. Initial-boundary value problems for constitutive equations with internal variables.

[2] H.-D. Alber and S. Nesenenko. Local H^1-regularity and $H^{1/3-\delta}$-regularity up to the boundary in time dependent viscoplasticity. *Asymptot. Anal.*, 63(3):151–187, 2009.

[3] J.-F. Babadjian and M. G. Mora. Stress regularity in quasi-static perfect plasticity with a pressure dependent yield criterion. *J. Differential Equations*, 264(8):5109–5151, 2018.

[4] M. Bulíček and J. Frehse. A revision of results for standard models in elasto-perfect-plasticity theory. *Calc. Var. Partial Differential Equations*, 57(2):Paper No. 50, 43, 2018.

[5] J. Frehse and D. Löbach. Regularity results for three-dimensional isotropic and kinematic hardening including boundary differentiability. *Math. Models Methods Appl. Sci.*, 19(12):2231–2262, 2009.

[6] J. Frehse and D. Löbach. Improved L^p-estimates for the strain velocities in hardening problems. *GAMM-Mitt.*, 34(1):10–20, 2011.

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

[8] J. Frehse and M. Specovius-Neugebauer. Fractional interior differentiability of the stress velocities to elastic plastic problems with hardening. *Boll. Unione Mat. Ital. (9)*, 5(3):469–494, 2012.

[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] C. Johnson. Existence theorems for plasticity problems. *J. Math. Pures Appl. (9)*, 55(4):431–444, 1976.

[11] C. Johnson. On plasticity with hardening. *J. Math. Anal. Appl.*, 62(2):325–336, 1978.

[12] D. Knees. Global regularity of the elastic fields of a power-law model on Lipschitz domains. *Math. Methods Appl. Sci.*, 29(12):1363–1391, 2006.

[13] D. Knees. On global spatial regularity and convergence rates for time-dependent elasto-plasticity. *Math. Models Methods Appl. Sci.*, 20(10):1823–1858, 2010.

[14] G. A. Seregin. Differential properties of the stress tensor in the Coulomb-Mohr theory of plasticity. *Algebra i Analiz*, 4(6):234–252, 1992.

[15] G. A. Seregin. Differential properties of solutions of evolution variational inequalities in the theory of plasticity. *J. Math. Sci.*, 72(6):3449–3458, 1994. Nonlinear boundary-value problems. Differential and pseudo-differential operators.

[16] G. A. Seregin. Remarks on regularity up to the boundary for solutions to variational problems in plasticity theory. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 233(Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funkts. 27):227–232, 258–259, 1996.

[17] 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.

[18] R. 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

Email address: mbul8060@karlin.mff.cuni.cz

INSTITUTE OF APPLIED MATHEMATICS, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, D-53121 BONN, GERMANY

Email address: aaa@iam.uni-bonn.de

FACHBEREICH 17, MATHEMATIK, UNIVERSITÄT KASSEL, HEINRICHPLETTSTR.40, D-34132 KASSEL, GERMANY

Email address: specovi@mathematik.uni-kassel.de