Journal of the Operations Research Society of Japan
Vol. 63, No. 3, July 2020, pp. 78–92

A NOTE ON ACCELERATED PROXIMAL GRADIENT METHOD FOR ELASTOPLASTIC ANALYSIS WITH TRESCA YIELD CRITERION

Wataru Shimizu
IHI Corporation

Yoshihiro Kanno
The University of Tokyo

(Received September 17, 2019; Revised December 24, 2019)

Abstract Diverse accelerated first-order methods have recently received considerable attention for solving large-scale convex optimization problems. This short paper shows that an exiting accelerated proximal gradient method for solving quasi-static incremental elastoplastic problems with the von Mises yield criterion can be naturally extended to the Tresca yield criterion.

Keywords: Optimization, mathematical modeling, semidefinite programming, accelerated first-order method, proximal gradient method, computational plasticity

1. Introduction

It has been long recognized that plasticity is closely linked to optimization. Particularly, in this decade, *second-order cone programming* (SOCP) and *semidefinite programming* (SDP) have been extensively applied to diverse problems in plasticity [3, 4, 11, 12, 15, 26].

Roughly speaking, solids and structures subjected to the static external load undergo only *elastic deformation* when the magnitude of the load is sufficiently small. When the magnitude of the stress (i.e., the internal force per unit area) exceeds a certain criterion, *plastic deformation* also takes place at the points where excess of the stress criterion occurs. Elastic deformation is reversible, while plastic deformation is irreversible. Therefore, elastoplastic behavior depends in general on the deformation history. A stress constraint describing the elastic limit is referred to as the *yield criterion*. Each metallic material has a specific yield criterion. For the Tresca and Mohr–Coulomb yield criteria, Bisbos [3] presented SDP formulations for elastoplastic shakedown analysis problems. Also, SDP approaches have been proposed for plastic limit analysis [4, 12, 15] as well as quasi-static elastoplastic analysis [11]. In this short paper, we address quasi-static elastoplastic analysis with the Tresca yield criterion. This analysis consists of solving a sequence of the *incremental problems*, which are obtained by applying the backward Euler time dicretization to the equation of motion under the assumption that the inertia term is negligible (i.e., the variation of the structural deformation is sufficiently slow).

Recently, *accelerated proximal gradient methods* have been proposed for the incremental problems of elastoplastic truss structures [8] and elastoplastic continua with the von Mises yield criterion [24]. It is worth noting that the truss problem can be recast as a *(convex) quadratic programming* (QP) problem [8, 13], while the von Mises problem can be recast as an SOCP problem [7, 26]. The numerical experiments suggest that the accelerated proximal gradient methods outperform standard optimization solvers implementing primal-dual interior-point methods for QP and SOCP [8, 24]. Unlike approaches using interior-point methods, these accelerated proximal gradient methods solve an unconstrained nonsmooth

convex optimization problem, in which only the incremental plastic strains and the incremental nodal displacements are treated as explicit optimization variables. The accelerated proximal gradient method is one of *accelerated gradient methods* (a.k.a. *optimal first-order methods*) that are popularly used for solving large-scale convex optimization problems arising in data science [2, 19, 20]. Applications of accelerated first-order methods to contact mechanics can be found in Mazhar *et al.* [16] and Kanno [9]; numerical experiments in [17] demonstrate that the method proposed by Mazhar *et al.* [16] is faster than conventional second-order methods in computational contact mechanics.

In this short paper, we show that the accelerated gradient method for incremental elastoplastic analysis with the von Mises yield criterion [24] can naturally be extended to the Tresca yield criterion. It is worth noting that the Tresca yield criterion means that plastic deformation begins when the maximum shear stress attains at a critical value. This assumption is consistent with the microscopic physical law of plastic yielding of crystal metals. Namely, in a metal crystal, it is widely recognized that plastic deformation occurs when the shear stress along a slip system (which is a pair of a particular plane and its associated direction, specified by the type of crystal lattice) attains at a threshold called the critical resolved shear stress [1]. Since most of conventional metal materials have polycrystalline structures, their mechanical properties are considered isotropic, which naturally leads to the Tresca yield criterion where the plasticity threshold is linked to the maximum shear stress among on any planes and in any directions. In physical experiments, it is usually observed that the yielding of metals occurs between the von Mises and Tresca criteria, and, particularly, for ductile metals the Tresca criterion matches the experimental results better [6]. However, in numerical simulation, the von Mises criterion is used much more often, because computation with the Tresca criterion is more difficult due to nonsmoothness of its yield surface [6]. Therefore, development of efficient and easily-implementable numerical methods with the Tresca yield criterion is in demand, which motivates us to study an accelerated proximal gradient method. In this paper, we also perform preliminary numerical experiments to examine efficiency of the proposed method, compared with a standard primal-dual interior-point method for SDP.

In our notation, $^\top$ denotes the transpose of a vector or a matrix. For $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$, we use $\|\boldsymbol{x}\|_2$ and $\|\boldsymbol{x}\|_1$, respectively, to denote its Euclidean norm and ℓ_1-norm, i.e., $\|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$ and $\|\boldsymbol{x}\|_1 = |x_1| + \cdots + |x_n|$. We use $\mathrm{diag}(\boldsymbol{x})$ to denote a diagonal matrix, the vector of diagonal entries of which is $\boldsymbol{x}$. Let $\mathcal{S}^n$ denote the set of $n \times n$ real symmetric matrices. We also write $\boldsymbol{\alpha} \in \mathcal{S}^3$ if $\boldsymbol{\alpha}$ is a second-order symmetric tensor with the dimension three, because in this paper we always consider the Cartesian coordinate system. Following the conventional notation in continuum mechanics, we use $X : Y$ to denote the scalar product (i.e., the inner product) of $X \in \mathcal{S}^n$ and $Y \in \mathcal{S}^n$. Similarly, we use $\boldsymbol{p} \cdot \boldsymbol{q}$ to denote the scalar product of $\boldsymbol{p} \in \mathbb{R}^n$ and $\boldsymbol{q} \in \mathbb{R}^n$. For $X \in \mathcal{S}^n$, we use $\|X\|_{\mathrm{F}}$ and $\|X\|_*$ to denote its Frobenius norm and the nuclear norm, respectively, i.e., $\|X\|_{\mathrm{F}} = \sqrt{X : X}$ and $\|X\|_* = \|\boldsymbol{\lambda}(X)\|_1$, where $\boldsymbol{\lambda}(X) \in \mathbb{R}^n$ is a vector of the eigenvalues of X in nonincreasing order, i.e., $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$. For $X \in \mathcal{S}^3$, we use $\mathrm{tr}(X)$ and $\mathrm{dev}(X)$ to denote its trace and deviator (i.e., $\mathrm{dev}(X) = X - (\mathrm{tr}(X)/3)I$), respectively. For closed convex functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $F : \mathcal{S}^n \to \mathbb{R} \cup \{+\infty\}$, define the proximal mappings of f and F,

respectively, by

$$\mathbf{prox}_f(\boldsymbol{x}) = \operatorname*{arg\,min}_{\boldsymbol{z}\in\mathbb{R}^n}\Big\{ f(\boldsymbol{z}) + \frac{1}{2}\|\boldsymbol{z}-\boldsymbol{x}\|_2^2\Big\},$$
$$\mathbf{prox}_F(X) = \operatorname*{arg\,min}_{Z\in\mathcal{S}^n}\Big\{ F(Z) + \frac{1}{2}\|Z-X\|_{\mathrm{F}}^2\Big\}.$$

Also, we use $\partial f(\boldsymbol{x})$ to denote the subdifferential of f at $\boldsymbol{x}\in\mathbb{R}^n$.

2. Overview of Accelerated Proximal Gradient Methods for Elastoplastic Incremental Problems

2.1. Existing method for the von Mises yield criterion

Shimizu and Kanno [24] proposed an accelerated proximal gradient method for the von Mises yield criterion with the strain hardening. We here summarize its essentials by restricting ourselves, for simplicity, to the perfect plasticity.*

Consider an elastoplastic body discretized according to the conventional procedure of the finite element method with the Gauss quadrature. Throughout the paper, we assume small deformation. Let d and r denote the number of degrees of freedom of the nodal displacements and the number of the Gauss evaluation points, respectively. We use $\dot{\boldsymbol{u}}\in\mathbb{R}^d$ and $\boldsymbol{q}\in\mathbb{R}^d$ to denote the incremental nodal displacement and the nodal external force, respectively.† At Gauss evaluation point l ($l=1,\dots,r$), the compatibility relation is written as $\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}=B_l\dot{\boldsymbol{u}}$, where $\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}\in\mathcal{S}^3$ and $\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\in\mathcal{S}^3$ are, respectively, the elastic and plastic parts of the incremental strain, and B_l is a linear operator. The incremental problem can be formulated as the following optimization problem [24]:‡

$$\text{Minimize}\quad \sum_{l=1}^{r}\Big(\frac{1}{2}\mathbf{C}_l\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\boldsymbol{\sigma}_{0l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\sqrt{\tfrac{2}{3}}R_{0l}\|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}\Big)-\boldsymbol{q}\cdot\dot{\boldsymbol{u}} \tag{1a}$$

$$\text{subject to}\quad \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\operatorname{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l})=B_l\dot{\boldsymbol{u}},\quad l=1,\dots,r. \tag{1b}$$

Here, $\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}$, $\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}$ ($l=1,\dots,r$), and $\dot{\boldsymbol{u}}$ are variables to be optimized, while $\mathbf{C}_l$, $\boldsymbol{\sigma}_{0l}$, and R_{0l} correspond to the (fourth-order) elasticity tensor, the stress tensor of the present state, and the radius of the yield surface. It is worth noting that (1a) is the sum of the increments of the elastic strain energy function, the (plastic) dissipation function, and the external work. We can eliminate $\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}$ from (1a) by using (1b), which yields the unconstrained form:

$$\begin{aligned}\text{Minimize}\quad f(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}},\dot{\boldsymbol{u}}) :=& \sum_{l=1}^{r}\frac{1}{2}\mathbf{C}_l(B_l\dot{\boldsymbol{u}}-\operatorname{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l})):(B_l\dot{\boldsymbol{u}}-\operatorname{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}))\\ &+\sum_{l=1}^{r}\boldsymbol{\sigma}_{0l}:(B_l\dot{\boldsymbol{u}}-\operatorname{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}))+\sum_{l=1}^{r}\sqrt{\tfrac{2}{3}}R_{0l}\|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}-\boldsymbol{q}\cdot\dot{\boldsymbol{u}}.\end{aligned}$$

*The strain hardening is a model of time-evolution of parameters (e.g., R_{0l} in the notation below) involved in the yield criterion. The perfect plasticity means the assumption that the strain hardening does not take place (i.e., R_{0l} is assumed to be a constant).

†It should be clear that $\dot{\boldsymbol{u}}$, as well as the related values appearing below, is not the rate value but the incremental value, though we use this notation instead of, e.g., $\Delta\boldsymbol{u}$, for notational simplicity.

‡Although the effect of strain hardening was addressed in [24], we here consider the perfect plasticity for simplifying the presentation and clarifying the connection with the formulations with the Tresca yield function presented in section 2.2 and section 2.3. It is worth noting that (1) corresponds to (17) in [24] when we put $H_l=0$ and $\boldsymbol{\beta}_{0l}=\boldsymbol{o}$.

Define $g : (\mathcal{S}^3)^r \times \mathbb{R}^d \to \mathbb{R}$ and $h_l^{\mathrm{m}} : \mathcal{S}^3 \to \mathbb{R}$ $(l = 1, \dots, r)$ by

$$g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) = f(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) - \sum_{l=1}^{r} h_l^{\mathrm{m}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}),$$
$$h_l^{\mathrm{m}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = \sqrt{\tfrac{2}{3}} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}.$$

The proximal gradient method [24] updates the incumbent solution, denoted $\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}1}^{(k)}, \dots, \dot{\boldsymbol{\varepsilon}}_{\mathrm{p}r}^{(k)}$ and $\dot{\boldsymbol{u}}^{(k)}$, as

$$\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}^{(k+1)} := \mathbf{prox}_{\alpha h_l^{\mathrm{m}}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}^{(k)} - \alpha \nabla_{\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}} g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}^{(k)}, \dot{\boldsymbol{u}}^{(k)})), \quad l = 1, \dots, r, \tag{2}$$
$$\dot{\boldsymbol{u}}^{(k+1)} := \dot{\boldsymbol{u}}^{(k)} - \alpha \nabla_{\dot{\boldsymbol{u}}} g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}^{(k)}, \dot{\boldsymbol{u}}^{(k)}), \tag{3}$$

where $\alpha > 0$ is the step length. It is worth noting that the computation in (2) can be done via a simple analytical formula. Also, the computation in (3) is certainly simple because g is a convex quadratic function. By applying the Nesterov acceleration [2, 18] and its adaptive restart [20], we obtain Algorithm 1 in Shimizu and Kanno [24, section 4.2].

2.2. Perfect plasticity with the Tresca yield criterion

For the perfectly plastic case, the dissipation function for the Tresca yield criterion is given by $\frac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_*$ with the incompressibility condition of the incremental plastic strain [22, 23]; see also [14]. With referring to problem (1), we see that the corresponding incremental problem is formulated as follows:[§]

$$\text{Minimize} \quad \sum_{l=1}^{r} \Big(\frac{1}{2} \mathbf{C}_l \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \boldsymbol{\sigma}_{0l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \frac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_* \Big) - \boldsymbol{q} \cdot \dot{\boldsymbol{u}} \tag{4a}$$
$$\text{subject to} \quad \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \mathrm{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = B_l \dot{\boldsymbol{u}}, \quad l = 1, \dots, r, \tag{4b}$$
$$\mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = 0, \quad l = 1, \dots, r. \tag{4c}$$

Substitution of (4b) into (4a) yields the following equivalent form:

$$\text{Minimize} \quad \sum_{l=1}^{r} \frac{1}{2} \mathbf{C}_l (B_l \dot{\boldsymbol{u}} - \mathrm{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l})) : (B_l \dot{\boldsymbol{u}} - \mathrm{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}))$$
$$+ \sum_{l=1}^{r} \boldsymbol{\sigma}_{0l} : (B_l \dot{\boldsymbol{u}} - \mathrm{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l})) + \sum_{l=1}^{r} \frac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_* - \boldsymbol{q} \cdot \dot{\boldsymbol{u}} \tag{5a}$$
$$\text{subject to} \quad \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = 0, \quad l = 1, \dots, r. \tag{5b}$$

Here, $\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}1}, \dots, \dot{\boldsymbol{\varepsilon}}_{\mathrm{p}r} \in \mathcal{S}^3$ and $\dot{\boldsymbol{u}} \in \mathbb{R}^d$ are optimization variables.

Define $h_l^{\mathrm{t}} : \mathcal{S}^3 \to \mathbb{R} \cup \{+\infty\}$ $(l = 1, \dots, r)$ by

$$h_l^{\mathrm{t}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = \begin{cases} \frac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_* & \text{if } \mathrm{tr}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = 0, \\ +\infty & \text{otherwise} \end{cases} \tag{6}$$

[§]With the von Mises yield criterion, we can easily confirm that condition (4c) is satisfied by any optimal solution of problem (1); see [26].

to see that problem (5) is embedded into the form

$$\text{minimize} \quad g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) + \sum_{l=1}^{r} h_l^{\mathrm{t}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}). \tag{7}$$

The proximal gradient method in section 2.1 can be extended to problem (7) as

$$\begin{aligned}
\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}^{(k+1)} &:= \mathbf{prox}_{\alpha h_l^{\mathrm{t}}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}^{(k)} - \alpha \nabla_{\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}} g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}^{(k)}, \dot{\boldsymbol{u}}^{(k)})), \quad l = 1, \dots, r, \\
\dot{\boldsymbol{u}}^{(k+1)} &:= \dot{\boldsymbol{u}}^{(k)} - \alpha \nabla_{\dot{\boldsymbol{u}}} g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}^{(k)}, \dot{\boldsymbol{u}}^{(k)}).
\end{aligned}$$

The only difference is that, for a given $X \in \mathcal{S}^3$, we are now required to compute

$$\mathbf{prox}_{\alpha h_l^{\mathrm{t}}}(X) = \operatorname*{arg\,min}_{Z \in \mathcal{S}^3} \left\{ \frac{1}{2} R_{0l} \|Z\|_* + \frac{1}{2} \|Z - X\|_{\mathrm{F}}^2 \;\middle|\; \operatorname{tr}(Z) = 0 \right\} \tag{8}$$

efficiently; the other computations can be performed in the same manner as Algorithm 1 in Shimizu and Kanno [24]. In section 3.1, we give an analytical formula for computing (8).

2.3. Strain hardening with the Tresca yield criterion

We are now in position to consider the Tresca yield criterion with the strain hardening. We assume combination of the linear isotropic hardening and the linear kinematic hardening with the Prager model. By adding the corresponding dissipation function, calculated from the maximum dissipation law,[¶] to the objective function of problem (4), we obtain the following optimization problem for the incremental problem:

$$\begin{aligned}
\text{Minimize} \quad & \sum_{l=1}^{r} \left(\frac{1}{2} \mathbf{C}_l \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \boldsymbol{\sigma}_{0l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \frac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_* \right) \\
& \qquad + \sum_{l=1}^{r} \left(\boldsymbol{\beta}_{0l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l} + \frac{1}{2} H_{\mathrm{k}l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}^2 + \frac{1}{8} H_{\mathrm{i}l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_*^2 \right) - \boldsymbol{q} \cdot \dot{\boldsymbol{u}} && \text{(9a)} \\
\text{subject to} \quad & \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l} + \operatorname{dev}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = B_l \dot{\boldsymbol{u}}, \quad l = 1, \dots, r, && \text{(9b)} \\
& \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = 0, \quad l = 1, \dots, r. && \text{(9c)}
\end{aligned}$$

Here, $\boldsymbol{\beta}_0 \in \mathcal{S}^3$ corresponds to the back stress, and $H_{\mathrm{i}l} > 0$ and $H_{\mathrm{k}l} > 0$ are constant isotropic and hardening moduli, respectively, at the lth evaluation point of the Gauss quadrature.

Define $\hat{g} : (\mathcal{S}^3)^r \times \mathbb{R}^d \to \mathbb{R}$ and $\hat{h}_l^{\mathrm{t}} : \mathcal{S}^3 \to \mathbb{R} \cup \{+\infty\}$ $(l = 1, \dots, r)$ by

$$\hat{g}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) = g(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) + \sum_{l=1}^{r} \left(\boldsymbol{\beta}_{0l} : \dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l} + \frac{1}{2} H_{\mathrm{k}l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}^2 \right), \tag{10}$$

$$\hat{h}_l^{\mathrm{t}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = \begin{cases} \dfrac{1}{2} R_{0l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_* + \dfrac{1}{8} H_{\mathrm{i}l} \|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_*^2 & \text{if } \operatorname{tr}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}) = 0, \\ +\infty & \text{otherwise} \end{cases} \tag{11}$$

to see that problem (9) is embedded into the form

$$\text{minimize} \quad \hat{g}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}, \dot{\boldsymbol{u}}) + \sum_{l=1}^{r} \hat{h}_l^{\mathrm{t}}(\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}). \tag{12}$$

[¶]This calculation is standard in the plasticity theory; see, e.g., [5].

The proximal gradient method can be applied to this problem in a manner similar to section 2.2. Since $\hat{g}$ is a convex quadratic function, we can compute its gradient in a straightforward manner. Therefore, what we need is to compute

$$\mathbf{prox}_{\alpha\hat{h}_l^{\mathrm{t}}}(X) = \operatorname*{arg\,min}_{Z\in\mathcal{S}^3}\Big\{\frac{1}{2}R_{0l}\|Z\|_* + \frac{1}{8}H_{il}\|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_*^2 + \frac{1}{2}\|Z-X\|_{\mathrm{F}}^2 \;\Big|\; \operatorname{tr}(Z)=0\Big\} \tag{13}$$

for a given $X \in \mathcal{S}^3$. We give an explicit formula for this in section 3.2.

3. Computing Proximal Mapping

3.1. Proximal mapping for perfect plasticity

In this section, we present an explicit form of (8).

We say that $F : \mathcal{S}^n \to \mathbb{R}\cup\{+\infty\}$ is a spectral function if it satisfies $F(X) = F(UXU^\top)$ for any $X \in \mathcal{S}^n$ and any orthogonal matrix $U \in \mathbb{R}^{n\times n}$. For each spectral function F, there exists a symmetric function $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ satisfying $F(X) = f(\boldsymbol{\lambda}(X))$ $(\forall X \in \mathcal{S}^n)$. Then we have

$$\mathbf{prox}_F(X) = U_X \operatorname{diag}(\mathbf{prox}_f(\boldsymbol{\lambda}(X)))U_X^\top,$$

where U_X is an orthogonal matrix satisfying $X = U_X \operatorname{diag}(\boldsymbol{\lambda}(X))U_X^\top$ [21, section 6.7.2]. It is easy to show that h_l^{t} defined in (6) is a spectral function. Also, defining $h_l^{\mathrm{te}} : \mathbb{R}^3 \to \mathbb{R}\cup\{+\infty\}$ by

$$h_l^{\mathrm{te}}(\boldsymbol{x}) = \begin{cases} \dfrac{1}{2}R_{0l}\|\boldsymbol{x}\|_1 & \text{if } x_1+x_2+x_3=0,\\ +\infty & \text{otherwise,}\end{cases}$$

we see that $h_l^{\mathrm{t}}(X) = h_l^{\mathrm{te}}(\boldsymbol{\lambda}(X))$ $(\forall X \in \mathcal{S}^3)$ holds. Accordingly, we can compute (8) straightforwardly if we can explicitly evaluate

$$\mathbf{prox}_{h_l^{\mathrm{te}}}(\boldsymbol{x}) = \operatorname*{arg\,min}_{\boldsymbol{z}\in\mathbb{R}^3}\Big\{c_l\|\boldsymbol{z}\|_1 + \frac{1}{2}\|\boldsymbol{z}-\boldsymbol{x}\|_2^2 \;\Big|\; z_1+z_2+z_3=0\Big\}, \tag{14}$$

where we put $c_l = R_{0l}/2$ (>0) for notational simplicity, and we assume without loss of generality that

$$x_1 \geq x_2 \geq x_3. \tag{15}$$

It is worth noting that, for $X \in \mathcal{S}^3$, we can obtain $\boldsymbol{\lambda}(X)$ and U_X analytically [10]; this analytical approach is actually adopted in the numerical experiments presented in section 4.

Let $P = \{\boldsymbol{z} \in \mathbb{R}^3 \mid z_1+z_2+z_3=0\}$. The projection of point $\boldsymbol{x} \in \mathbb{R}^3$ onto P is defined by

$$\Pi_P(\boldsymbol{x}) = \operatorname*{arg\,min}_{\boldsymbol{z}\in\mathbb{R}^3}\{\|\boldsymbol{z}-\boldsymbol{x}\|_2 \mid \boldsymbol{z}\in P\}.$$

Since $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ defined by

$$\boldsymbol{v}_1 = \begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\\ 0\end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix}1/\sqrt{6}\\ 1/\sqrt{6}\\ -2/\sqrt{6}\end{bmatrix}$$

form an orthonormal basis of P, we see that $\boldsymbol{z} \in P$ if and only if there exist ζ_1, $\zeta_2 \in \mathbb{R}$ satisfying $\boldsymbol{z} = \zeta_1 \boldsymbol{v}_1 + \zeta_2 \boldsymbol{v}_2$. Therefore, $\Pi_P(\boldsymbol{x})$ can be explicitly given as

$$\Pi_P(\boldsymbol{x}) = \xi_1 \boldsymbol{v}_1 + \xi_2 \boldsymbol{v}_2$$

with

$$(\xi_1, \xi_2) = \Big(\frac{-x_1 + x_2}{\sqrt{2}}, \frac{x_1 + x_2 - 2x_3}{\sqrt{6}} \Big),$$

where (15) is equivalent to

$$-\sqrt{3}\xi_2 \leq \xi_1 \leq 0. \tag{16}$$

Since we have $\|\boldsymbol{z} - \boldsymbol{x}\|_2^2 = \|\boldsymbol{z} - \Pi_P(\boldsymbol{x})\|_2^2 + \|\boldsymbol{x} - \Pi_P(\boldsymbol{x})\|_2^2$ (i.e., the Pythagorean theorem) for $\boldsymbol{z} \in P$, (14) can be reduced to

$$\mathbf{prox}_{h_l^{\mathrm{te}}}(\boldsymbol{x}) = \operatorname*{arg\,min}_{\boldsymbol{z} \in \mathbb{R}^3} \Big\{ c_l \|\boldsymbol{z}\|_1 + \frac{1}{2} \|\boldsymbol{z} - \Pi_P(\boldsymbol{x})\|_2^2 \;\Big|\; \boldsymbol{z} \in P \Big\}.$$

Furthermore, by again using the fact that $\boldsymbol{z} \in P$ is equivalent to the existence of ζ_1, $\zeta_2 \in \mathbb{R}$ satisfying $\boldsymbol{z} = \zeta_1 \boldsymbol{v}_1 + \zeta_2 \boldsymbol{v}_2$, we can evaluate (14) by computing the optimal solution of the following optimization problem for given ξ_1 and ξ_2 satisfying (16):

$$\operatorname*{Minimize}_{\boldsymbol{\zeta} = (\zeta_1, \zeta_2) \in \mathbb{R}^2} \quad \theta(\boldsymbol{\zeta}) := c_l \|\zeta_1 \boldsymbol{v}_1 + \zeta_2 \boldsymbol{v}_2\|_1 + \frac{1}{2} \|\boldsymbol{\zeta} - \boldsymbol{\xi}\|_2^2. \tag{17}$$

Since problem (17) has only two optimization variables, we can solve it analytically as follows:

Theorem 1. *For ξ_1 and ξ_2 satisfying* (16) *and $c_l > 0$, the optimal solution of problem* (17)*, denoted (ζ_1^*, ζ_2^*), is*

$$(\zeta_1^*, \zeta_2^*) = \begin{cases} (0, 0) & \text{if } \sqrt{3}\xi_2 \leq \xi_1 + 2\sqrt{2}c_l, \\ (\xi_1 + \sqrt{2}c_l, \xi_2 - \sqrt{2/3}c_l) & \text{if } \sqrt{3}\xi_2 \leq -3\xi_1 - 2\sqrt{2}c_l, \\ (\xi_1, \xi_2 - 2\sqrt{2/3}c_l) & \text{if } \sqrt{3}\xi_2 \geq -3\xi_1 + 2\sqrt{2}c_l, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise.} \end{cases}$$

A proof of Theorem 1 appears in appendix A.

3.2. Proximal mapping with strain hardening

This section provides an explicit form of (13).

We can easily verify that $\hat{h}_l^{\mathrm{t}}$ defined in (11) is a spectral function. Indeed, if we define $\hat{h}_l^{\mathrm{te}} : \mathbb{R}^3 \to \mathbb{R} \cup \{+\infty\}$ by

$$\hat{h}_l^{\mathrm{te}}(\boldsymbol{x}) = \begin{cases} \dfrac{1}{2} R_{0l} \|\boldsymbol{x}\|_1 + \dfrac{1}{8} H_{il} \|\boldsymbol{x}\|_1^2 & \text{if } \boldsymbol{x} \in P, \\ +\infty & \text{otherwise,} \end{cases}$$

then we have $\hat{h}_l^{\mathrm{t}}(X) = \hat{h}_l^{\mathrm{te}}(\boldsymbol{\lambda}(X))$ $(\forall X \in \mathcal{S}^3)$. Accordingly, to obtain (13) explicitly, it is sufficient to evaluate

$$\begin{aligned} \mathbf{prox}_{\hat{h}_l^{\mathrm{te}}}(\boldsymbol{x}) &= \operatorname*{arg\,min}_{\boldsymbol{z} \in \mathbb{R}^3} \Big\{ c_{l1} \|\boldsymbol{z}\|_1 + c_{l2} \|\boldsymbol{z}\|_1^2 + \frac{1}{2} \|\boldsymbol{z} - \boldsymbol{x}\|_2^2 \;\Big|\; \boldsymbol{z} \in P \Big\} \\ &= \operatorname*{arg\,min}_{\boldsymbol{z} \in \mathbb{R}^3} \Big\{ c_{l1} \|\boldsymbol{z}\|_1 + c_{l2} \|\boldsymbol{z}\|_1^2 + \frac{1}{2} \|\boldsymbol{z} - \Pi_P(\boldsymbol{x})\|_2^2 \;\Big|\; \boldsymbol{z} \in P \Big\}, \end{aligned} \tag{18}$$

where we put $c_{l1} = R_{0l}/2$ (> 0) and $c_{l2} = H_{il}/8$ (> 0) for notational simplicity. In the same manner as section 3.1, we can write $\Pi_P(\boldsymbol{x}) = \xi_1\boldsymbol{v}_1 + \xi_2\boldsymbol{v}_2$ $(-\sqrt{3}\xi_2 \leq \xi_1 \leq 0)$ and $\boldsymbol{z} = \zeta_1\boldsymbol{v}_1 + \zeta_2\boldsymbol{v}_2$. Accordingly, the optimization problem in (18) is reduced to

$$\underset{\boldsymbol{\zeta}\in\mathbb{R}^2}{\text{minimize}} \quad c_{l1}\|\zeta_1\boldsymbol{v}_1 + \zeta_2\boldsymbol{v}_2\|_1 + c_{l2}\|\zeta_1\boldsymbol{v}_1 + \zeta_2\boldsymbol{v}_2\|_1^2 + \frac{1}{2}\|\boldsymbol{\zeta} - \boldsymbol{\xi}\|_2^2. \tag{19}$$

In a manner analogous to Theorem 1, we can obtain the optimal solution of problem (19) as follows:

$$(\zeta_1^*, \zeta_2^*) = \begin{cases} (0,0) \\ \qquad \text{if } \sqrt{3}\xi_2 \leq \xi_1 + 2\sqrt{2}c_{l1}, \\ \left(\dfrac{(3+4c_{l2})\xi_1 + 4\sqrt{3}c_{l2}\xi_2 + 3\sqrt{2}c_{l1}}{3+16c_{l2}}, \dfrac{4\sqrt{3}c_{l2}\xi_1 + 3(1+4c_{l2})\xi_2 - \sqrt{6}c_{l1}}{3+16c_{l2}}\right) \\ \qquad \text{if } \sqrt{3}(1+8c_{l2})\xi_2 \leq -(3+8c_{l2})\xi_1 - 2\sqrt{2}c_{l1}, \\ \left(\xi_1, \dfrac{3\xi_2 - 2\sqrt{6}c_{l1}}{3+16c_{l2}}\right) \\ \qquad \text{if } \sqrt{3}\xi_2 \geq (-3+16c_{l2})\xi_1 + 2\sqrt{2}c_{l1}, \\ \left(\dfrac{\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_{l1}}{4(1+4c_{l2})}, \dfrac{-\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_{l1}}{4(1+4c_{l2})}\right) \\ \qquad \text{otherwise.} \end{cases}$$

4. Preliminary Numerical Experiments

In this section, we perform preliminary numerical experiments to examine the effiency of the method presented in this paper. The main part of the method solving problem (12) was implemented in MATLAB ver. 7.11. Computation of the proximal mapping in section 3 was implemented in C++. Computation was carried out on a 2.4 GHz Intel Core i7 processor with 16 GB RAM.

We solved the elastoplastic problem in Shimizu and Kanno [24, section 5.2] with, this time, the Tresca yield criterion. The finite element mesh, the boundary condition, and the problem size are detailed in [24]. The initial point for the proposed method is $\mathbf{0}$. The step length, α, is determined in the same manner as in [24]. We terminated the algorithm when the Euclidean norm of the update of the solution vector becomes less than the threshold, ϵ, where we consider two cases, $\epsilon = 10^{-8}$ and 10^{-10}. Figure 1 shows a typical iteration history of the proposed method. It also reports the results of the accelerated gradient method without restart and the (unaccelerated) proximal gradient method. We can observe that the acceleration with restart drastically speeds up the convergence.

For comparison of efficiency, we also solved problem (12) with an SDP approach. Namely,

by using the formulations in [11, 15], problem (9) can be recast as follows:

$$\begin{aligned}
\text{Minimize} \quad & \sum_{l=1}^{r}\Big(\frac{1}{2}\mathbf{C}_l\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\boldsymbol{\sigma}_{0l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+R_{0l}\gamma_l\Big) \\
& \qquad +\sum_{l=1}^{r}\Big(\boldsymbol{\beta}_{0l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}+\frac{1}{2}H_{\mathrm{k}l}\|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}^2+\frac{1}{2}H_{\mathrm{i}l}\gamma_l^2\Big)-\boldsymbol{q}\cdot\dot{\boldsymbol{u}} && \text{(20a)}\\
\text{subject to} \quad & \dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}+\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}=B_l\dot{\boldsymbol{u}}, \quad l=1,\dots,r, && \text{(20b)}\\
& \dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}=\dot{\boldsymbol{e}}_l^{+}-\dot{\boldsymbol{e}}_l^{-}, \quad l=1,\dots,r, && \text{(20c)}\\
& \operatorname{tr}\dot{\boldsymbol{e}}_l^{+}=\gamma_l, \quad l=1,\dots,r, && \text{(20d)}\\
& \operatorname{tr}\dot{\boldsymbol{e}}_l^{-}=\gamma_l, \quad l=1,\dots,r, && \text{(20e)}\\
& \dot{\boldsymbol{e}}_l^{+}\succeq\boldsymbol{o}, \quad l=1,\dots,r, && \text{(20f)}\\
& \dot{\boldsymbol{e}}_l^{-}\succeq\boldsymbol{o}, \quad l=1,\dots,r. && \text{(20g)}
\end{aligned}$$

Here, $\dot{\boldsymbol{e}}_l^{+}\in\mathcal{S}^3$, $\dot{\boldsymbol{e}}_l^{-}\in\mathcal{S}^3$, and $\gamma_l\in\mathbb{R}$ $(l=1,\dots,r)$ are additional optimization variables, and the constraints in (20f) amd (20g) mean that $\dot{\boldsymbol{e}}_l^{+}$ and $\dot{\boldsymbol{e}}_l^{-}$ are positive semidefinite. Furthermore, the quadratic term $\sum_{l=1}^{r}\frac{1}{2}H_{\mathrm{i}l}\gamma_l^2$ in (20a) can be converted to minimization of an additional variable $w\in\mathbb{R}$ under the second-order cone constraint

$$w\geq\sum_{l=1}^{r}\frac{1}{2}H_{\mathrm{i}l}\gamma_l^2 \quad\Leftrightarrow\quad w+\frac{1}{2}\geq\left\|\begin{bmatrix} w-\frac{1}{2}\\ \sqrt{H_{\mathrm{i}1}}\gamma_1\\ \vdots\\ \sqrt{H_{\mathrm{i}r}}\gamma_r\end{bmatrix}\right\|_2.$$

Similarly, $\sum_{l=1}^{r}\frac{1}{2}\mathbf{C}_l\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}:\dot{\boldsymbol{\varepsilon}}_{\mathrm{e}l}$ and $\sum_{l=1}^{r}\frac{1}{2}H_{\mathrm{k}l}\|\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}l}\|_{\mathrm{F}}^2$ in (20a) are handled by means of second-order cone constraints. Accordingly, we obtain an SDP problem. We solved this SDP problem by using SDPT3 ver. 4 [25], a primal-dual interior-point method, with the default setting.

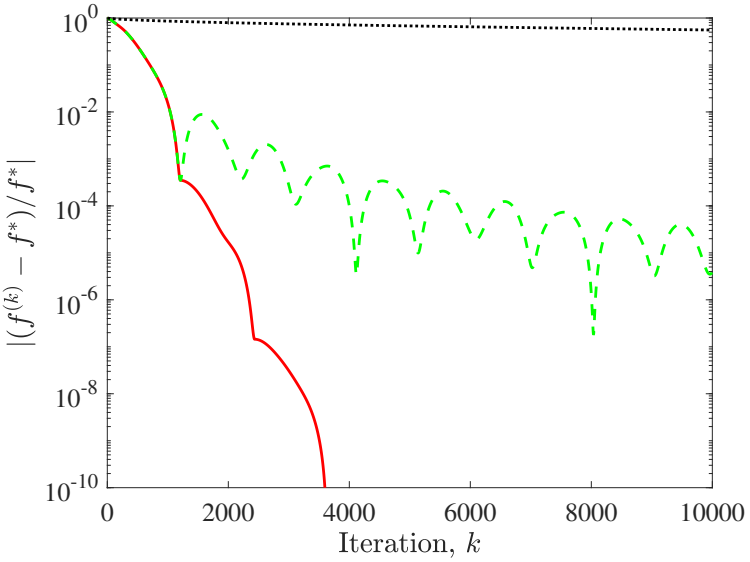

Figure 1: Convergence history of the objective value. "*Solid line*" The proposed method; "*dashed line*" the accelerated proximal gradient method without restart; and "*dotted line*" the proximal gradient method

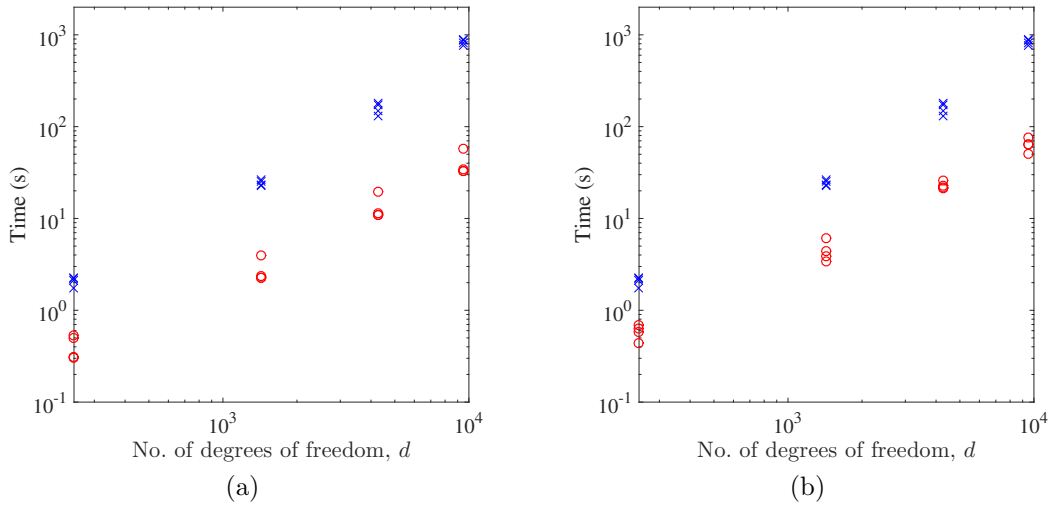

Figure 2: Computational time. "○" The proposed method; and "×" a primal-dual interior-point method. (a) $\epsilon = 10^{-8}$; and (b) $\epsilon = 10^{-10}$ in the proposed method

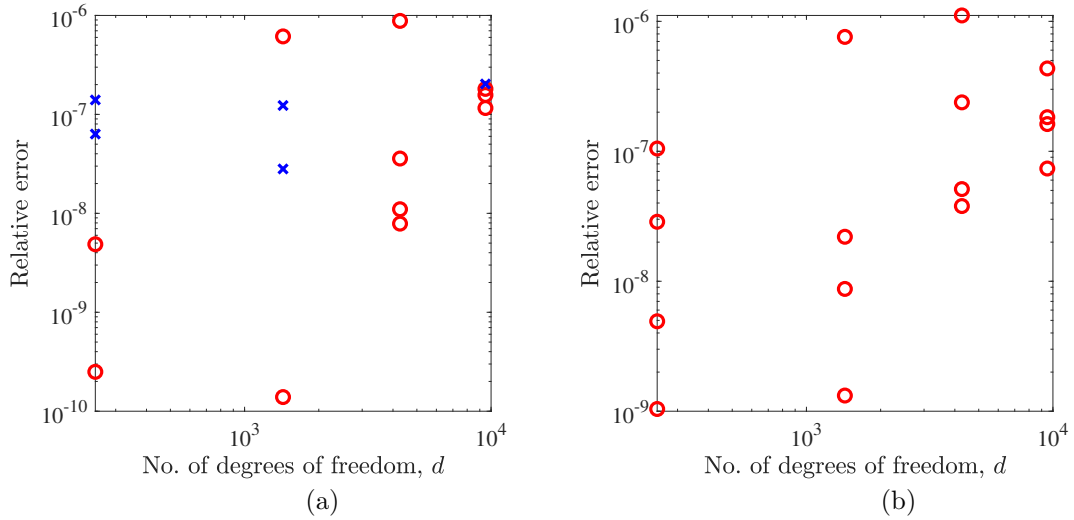

Figure 3: Relative difference of the objective values obtained by the proposed method and a primal-dual interior-point method. (a) $\epsilon = 10^{-8}$; and (b) $\epsilon = 10^{-10}$ in the proposed method. "○" The objective value obtained by the proposed method is smaller; and "×" the one by a primal-dual interior-point method is smaller

Figure 2 compares the computational time of the two methods. In every case, the proposed method is faster than SDPT3. Figure 3 compares the accuracy of the obtained solutions. The relative error is defined by $(\check{f} - f^*)/f^*$, where $\check{f}$ and f^* are the objective values of the solutions obtained by the proposed method and SDPT3, respectively. A smaller objective value means higher accuracy. We can confirm that, with $\epsilon = 10^{-8}$, the solution obtained by the proposed method agrees very well with the one obtained by SDPT3. Also, the proposed method with $\epsilon = 10^{-10}$ finds a solution with higher accuracy than SDPT3 in every case. Figure 4 reports the number of iterations required by the proposed method. We can observe that the number of iterations increases gradually as the problem size increases.

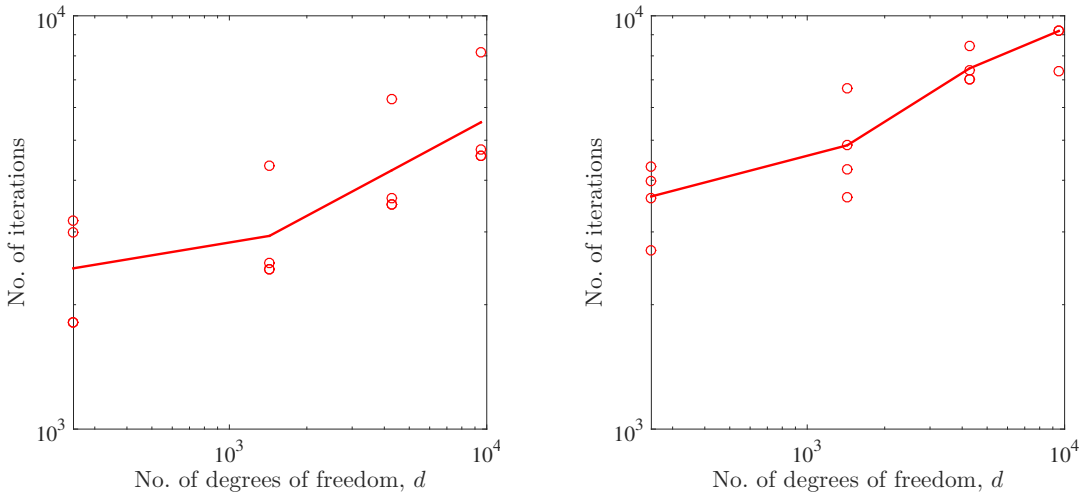

Figure 4: The number of iteration of the proposed method. (a) $\epsilon = 10^{-8}$; and (b) $\epsilon = 10^{-10}$

5. Conclusions

In this paper, we have extended an accelerated proximal gradient method for solving quasi-static incremental elastoplastic problems with the von Mises yield criterion to the ones with Tresca yield criterion. The proposed method solves an unconstrained nonsmooth convex optimization problem, which has a form similar to the low-rank matrix approximation problem using the nuclear norm of a matrix. The preliminary numerical experiments suggest that the proposed method outperforms a standard primal-dual interior-point method for semidefinite programming.

A. Proof of Theorem 1

Theorem 1 can be proved as follows.

A necessary condition for optimality is that there exists $\mu \leq 0$ satisfying

$$\mu(\boldsymbol{\zeta} - \boldsymbol{\xi}) \in \partial \|\zeta_1 \boldsymbol{v}_1 + \zeta_2 \boldsymbol{v}_2\|_1.$$

The set of $\boldsymbol{\zeta}$ satisfying this necessary condition corresponds to the two line segments indicated by dotted lines in Figure 5. The optimal solution can be found by exploring these two line segments as follows.

We first investigate the case in Figure 5a, i.e., $-\xi_1 \leq \sqrt{3}\xi_2 \leq -3\xi_1$. The dotted line segments in Figure 5a are written explicitly as

$$l_1 = \left\{ (\zeta_1, \zeta_2) \;\middle|\; \sqrt{3}\zeta_1 + \zeta_2 = 0,\ \zeta_1 \in \left[-\frac{1}{2}\xi_1 - \frac{\sqrt{3}}{2}\xi_2, 0 \right] \right\},$$
$$l_2 = \left\{ (\zeta_1, \zeta_2) \;\middle|\; \zeta_1 + \sqrt{3}\zeta_2 = \xi_1 + \sqrt{3}\xi_2,\ \zeta_1 \in \left[\xi_1, -\frac{1}{2}\xi_1 - \frac{\sqrt{3}}{2}\xi_2 \right] \right\}.$$

If $\boldsymbol{\zeta} \in l_1$, then the objective function is reduced to

$$\theta(\zeta_1, \zeta_2(\zeta_1)) = 2\zeta_1^2 - (\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l)\zeta_1 + \frac{1}{2}\|\boldsymbol{\xi}\|_2^2. \tag{21}$$

A direct calculation yields

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_1\} = \begin{cases} (0,0) \\ \qquad \text{if } \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l) \geq 0, \\ l_1 \cap l_2 \\ \qquad \text{if } \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l) \leq -\dfrac{1}{2}\xi_1 - \dfrac{\sqrt{3}}{2}\xi_2, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) \\ \qquad \text{otherwise,} \end{cases}$$

namely,

$$\begin{aligned} &\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_1\} \\ &\quad = \begin{cases} (0,0) & \text{if } 2\sqrt{2}c_l \geq -\xi_1 + \sqrt{3}\xi_2, \\ l_1 \cap l_2 & \text{if } 2\sqrt{2}c_l \leq -3\xi_1 - \sqrt{3}\xi_2, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise.} \end{cases} \end{aligned} \tag{22}$$

On the other hand, if $\boldsymbol{\zeta} \in l_2$, then the objective function is reduced to

$$\theta(\zeta_1, \zeta_2(\zeta_1)) = \frac{2}{3}\zeta_1^2 - \frac{4}{3}(\xi_1 + \sqrt{2}c_l)\zeta_1 + \frac{\sqrt{2}}{3}(\sqrt{2}\xi_1^2 + c_l\xi_1 + \sqrt{3}c_l\xi_2)$$

A direct calculation yields

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_2\} = \begin{cases} l_1 \cap l_2 & \text{if } \xi_1 + \sqrt{2}c_l \geq -\dfrac{1}{2}\xi_1 - \dfrac{\sqrt{3}}{2}\xi_2, \\ (\xi_1 + \sqrt{2}c_l, \xi_2 - \sqrt{2/3}c_l) & \text{if } \xi_1 + \sqrt{2}c_l \leq -\dfrac{1}{2}\xi_1 - \dfrac{\sqrt{3}}{2}\xi_2, \end{cases}$$

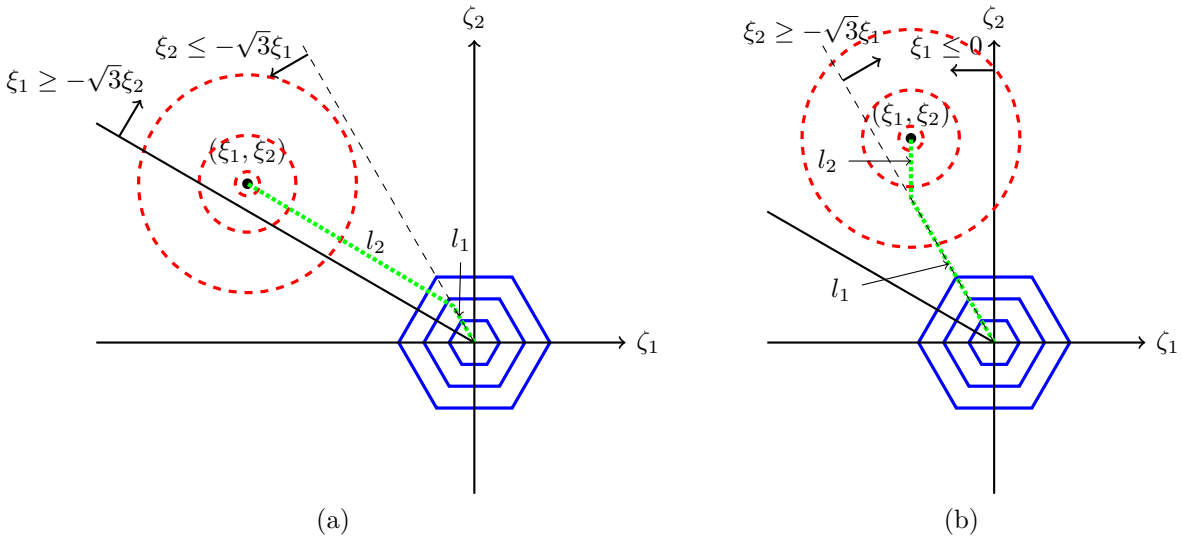

Figure 5: Schematic of problem (17). "*Solid thick lines*" Contours of $c_l\|\zeta_1\boldsymbol{v}_1 + \zeta_2\boldsymbol{v}_2\|_1$ (each contour forms a regular hexagon); and "*dashed thick lines*" contours of $\frac{1}{2}\|\boldsymbol{\zeta} - \boldsymbol{\xi}\|_2^2$. The optimal solution exists on the union of two line segments, $l_1 \cup l_2$, indicated by dotted lines. (a) $-\xi_1 \leq \sqrt{3}\xi_2 \leq -3\xi_1$; and (b) $0 \leq -\sqrt{3}\xi_1 \leq \xi_2$

namely,

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_2\} = \begin{cases} l_1 \cap l_2 & \text{if } 2\sqrt{2}c_l \geq -3\xi_1 - \sqrt{3}\xi_2, \\ (\xi_1 + \sqrt{2}c_l, \xi_2 - \sqrt{2/3}c_l) & \text{if } 2\sqrt{2}c_l \leq -3\xi_1 - \sqrt{3}\xi_2. \end{cases} \tag{23}$$

From (22) and (23), we obtain

$$(\zeta_1^*, \zeta_2^*) = \begin{cases} (0,0) & \text{if } 2\sqrt{2}c_l \geq -\xi_1 + \sqrt{3}\xi_2, \\ (\xi_1 + \sqrt{2}c_l, \xi_2 - \sqrt{2/3}c_l) & \text{if } 2\sqrt{2}c_l \leq -3\xi_1 - \sqrt{3}\xi_2, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise.} \end{cases} \tag{24}$$

We next investigate the case in Figure 5b, i.e., $0 \leq -\sqrt{3}\xi_1 \leq \xi_2$. The dotted line segments in Figure 5b are written as

$$l_1 = \left\{(\zeta_1, \zeta_2) \;\middle|\; \sqrt{3}\zeta_1 + \zeta_2 = 0,\ \zeta_1 \in [\xi_1, 0]\right\},$$
$$l_2 = \left\{(\zeta_1, \zeta_2) \;\middle|\; \zeta_1 = \xi_1,\ \zeta_2 \in \left[-\sqrt{3}\xi_1, \xi_2\right]\right\}.$$

If $\boldsymbol{\zeta} \in l_1$, then the objective function is reduced to (21), which yields

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_1\} = \begin{cases} (0,0) & \text{if } \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l) \geq 0, \\ l_1 \cap l_2 & \text{if } \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l) \leq \xi_1, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise,} \end{cases}$$

namely,

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_1\} = \begin{cases} (0,0) & \text{if } 2\sqrt{2}c_l \geq -\xi_1 + \sqrt{3}\xi_2, \\ l_1 \cap l_2 & \text{if } 2\sqrt{2}c_l \leq 3\xi_1 + \sqrt{3}\xi_2, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise.} \end{cases} \tag{25}$$

Alternatively, if $\boldsymbol{\zeta} \in l_2$, then the objective function is reduced to

$$\theta(\zeta_1(\zeta_2), \zeta_2) = \frac{1}{2}\zeta_2^2 - (\xi_2 - 2\sqrt{2/3}c_l)\zeta_2 + \frac{1}{2}\xi_2^2.$$

A direct calculation yields

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_2\} = \begin{cases} l_1 \cap l_2 & \text{if } \xi_2 - 2\sqrt{2/3}c_l \leq -\sqrt{3}\xi_1, \\ (\xi_1, \xi_2 - 2\sqrt{2/3}c_l) & \text{if } \xi_2 - 2\sqrt{2/3}c_l \geq -\sqrt{3}\xi_1, \end{cases}$$

namely,

$$\arg\min\{\theta(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in l_2\} = \begin{cases} l_1 \cap l_2 & \text{if } 2\sqrt{2}c_l \geq 3\xi_1 + \sqrt{3}\xi_2, \\ (\xi_1, \xi_2 - 2\sqrt{2/3}c_l) & \text{if } 2\sqrt{2}c_l \leq 3\xi_1 + \sqrt{3}\xi_2. \end{cases} \tag{26}$$

From (25) and (26), we obtain

$$(\zeta_1^*, \zeta_2^*) = \begin{cases} (0,0) & \text{if } 2\sqrt{2}c_l \geq -\xi_1 + \sqrt{3}\xi_2, \\ (\xi_1, \xi_2 - 2\sqrt{2/3}c_l) & \text{if } 2\sqrt{2}c_l \leq 3\xi_1 + \sqrt{3}\xi_2, \\ \dfrac{1}{4}(\xi_1 - \sqrt{3}\xi_2 + 2\sqrt{2}c_l, -\sqrt{3}\xi_1 + 3\xi_2 - 2\sqrt{6}c_l) & \text{otherwise.} \end{cases} \tag{27}$$

Thus, the optimal solution is obtained as (24) and (27).

Acknowledgments

The work of the second author is partially supported by JSPS KAKENHI 17K06633 and JST CREST Grant Number JPMJCR1911, Japan.

References

[1] C.R. Barrett, W.D. Nix, and A.S. Tetelman: *The Principles of Engineering Materials* (Prentice-Hall, Englewood Cliffs, 1973).

[2] A. Beck and M. Teboulle: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, **2** (2009), 183–202.

[3] C.D. Bisbos: Semidefinite optimization models for limit and shakedown analysis problems involving matrix spreads. *Optimization Letters*, **1** (2007), 101–109.

[4] M. Gueguin, G. Hassen, and P. de Buhan: Numerical assessment of the macroscopic strength criterion of reinforced soils using semidefinite programming. *International Journal for Numerical Methods in Engineering*, **99** (2014), 522–541.

[5] W. Han and B.D. Reddy: *Plasticity (2nd ed.)* (Springer, New York, 2013).

[6] J. Lemaitre and J.L. Chaboche: *Mehcanics of Solid Materials* (Cambridge University Press, Cambridge, 1990).

[7] Y. Kanno: *Nonsmooth Mechanics and Convex Optimization* (CRC Press, Boca Raton, 2011).

[8] Y. Kanno: A fast first-order optimization approach to elastoplastic analysis of skeletal structures. *Optimization and Engineering*, **17** (2016), 861–896.

[9] Y. Kanno: An accelerated Uzawa method for application to frictionless contact problem. *Optimization Letters*, to appear. DOI:10.1007/s11590-019-01481-2

[10] J. Kopp: Efficient numerical diagonalization of Hermitian 3×3 matrices. *International Journal of Modern Physics, C*, **19** (2008), 523–548.

[11] K. Krabbenhøft, A.V. Lyamin, and S.W. Sloan: Formulation and solution of some plasticity problems as conic programs. *International Journal of Solids and Structures*, **44** (2007), 1533–1549.

[12] K. Krabbenhøft, A.V. Lyamin, and S.W. Sloan: Three-dimensional Mohr–Coulomb limit analysis using semidefinite programming. *Communications in Numerical Methods in Engineering*, **24** (2008), 1107–1119.

[13] G. Maier: A quadratic programming approach for certain classes of non-linear structural problems. *Meccanica*, **3** (1968), 121–130.

[14] A. Makrodimopoulovs: Remarks on some properties of conic yield restrictions in limit analysis. *International Journal for Numerical Methods in Biomedical Engineering*, **26** (2010), 1449–1461.

[15] C.M. Martin and A. Makrodimopoulovs: Finite-element limit analysis of Mohr–Coulomb materials in 3D using semidefinite programming. *Journal of Engineering Mechanics (ASCE)*, **134** (2008), 339–347.

[16] H. Mazhar, T. Heyn, D. Negrut, and A. Tasora: Using Nesterov's method to accelerate multibody dynamics with friction and contact. *ACM Transactions on Graphics*, **34** (2015), Article No. 32.

[17] D. Melanz, L. Fang, P. Jayakumar, and D. Negrut: A comparison of numerical methods for solving multibody dynamics problems with frictional contact modeled via differential variational inequalities. *Computer Methods in Applied Mechanics and Engineering*, **320** (2017), 668–693.

[18] Y. Nesterov: A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, **27** (1983), 372–376.

[19] Y. Nesterov: *Introductory Lectures on Convex Optimization: A Basic Course* (Kluwer Academic Publishers, Dordrecht, 2004).

[20] B. O'Donoghue and E. Candès: Adaptive restart for accelerated gradient schemes. *Foundations of Computational Mathematics*, **15** (2015), 715–732.

[21] N. Parikh and S. Boyd: Proximal algorithms. *Foundations and Trends in Optimization*, **1** (2013), 123–231.

[22] J. Salençon: An introduction to the yield design theory and its applications to soil mechanics. *European Journal of Mechanics, A/Solids*, **9** (1990), 477–500.

[23] R.T. Shield and D.C. Drucker: The application of limit analysis to punch-indentation problems. *Journal of Applied Mechanics (ASME)*, **20** (1953), 453–460.

[24] W. Shimizu and Y. Kanno: Accelerated proximal gradient method for elastoplastic analysis with von Mises yield criterion. *Japan Journal of Industrial and Applied Mathematics*, **35** (2018), 1–32.

[25] R.H. Tütüncü, K.C. Toh, and M.J. Todd: Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, **B95** (2003), 189–217.

[26] K. Yonekura and Y. Kanno: Second-order cone programming with warm start for elastoplastic analysis with von Mises yield criterion. *Optimization and Engineering*, **13** (2012), 181–218.

Yoshihiro Kanno
Mathematics and Informatics Center,
The University of Tokyo, Hongo 7-3-1,
Tokyo 113-8656, Japan
E-mail: kanno@mist.i.u-tokyo.ac.jp