Received: 7 January 2022 | Accepted: 8 March 2022

DOI: 10.1002/nme.6971

RESEARCH ARTICLE

WILEY

Gurson–Tvergaard–Needleman model guided fracture-resistant structural designs under finite deformations

Guodong Zhang | **Kapil Khandelwal**

Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, Indiana, USA

Correspondence
Department of Civil and Environmental Engineering and Earth Sciences
University of Notre Dame, 156 Fitzpatrick Hall, Notre Dame, IN 46556, USA.
Email: kapil.khandelwal@nd.edu

Funding information
National Science Foundation, Grant/Award Number: CMMI-1762277

Abstract

A finite strain shear modified Gurson–Tvergaard–Needleman (GTN) model based on multiplicative elastoplasticity, together with its implementation details, is presented. This GTN model which can simulate the loss of load-carrying capacity of porous metals through nucleation, growth, shearing, and coalescence of voids is incorporated in an optimization framework for designing structures with optimal plastic energy dissipation capacity while satisfying prescribed constraints on material usage and damage. An adjoint method-based analytical path-dependent sensitivity analysis is presented that can be used with gradient-based optimization algorithms. Using the proposed topology optimization formulations, the optimized designs exhibit well-constrained fracture at the design displacements. Ultimate performance analyses of the optimized designs demonstrate that the fracture-resistant designs can have higher ductility, ultimate strength as well as improved energy dissipation, as compared to the designs guided by von Mises plasticity where no fracture mechanisms are modeled. Moreover, due to finite deformations, a fracture can initiate at multiple locations and critical fracture locations can change during the loading process. In addition, different failure modes are revealed from different optimized designs under large deformations.

KEYWORDS

adjoint method, ductile damage, finite strain plasticity, shear modified GTN model, topology optimization

1 | INTRODUCTION

Design of efficient structures toward prescribed objectives, such as optimal energy dissipation, stiffness, etc., with well-controlled damage or fracture behavior is an important task in many engineering applications. In many cases, due to the complex physics of material damage and the presence of multiple conflicting design constraints, traditional experimental and/or empirical design approaches are not suitable for this task. With the development of advanced computational and design optimization methods, however, this design goal can be now accomplished and requires a careful formulation of the inverse design problem that involves modeling of the underlying damage mechanisms. Among the three main categories of structural optimization—that is, sizing, shape, and topology—the topology optimization considers both the connectivity of structural features as well as the size and shape of the features, and thus offers greater

flexibility and design freedom. Topology optimization has also undergone significant advancements following the pioneering work by Bendsøe and Kikuchi,[1] for instance, see References 2 and 3 for an overview of this area. In a density-based topology optimization approach, which is the focus of this study, a structure is parameterized by discretized density field via finite elements, which is then optimized toward minimizing a prescribed objective function while satisfying a set of predefined constraints. For instance, to prevent material failure due to yielding or brittle fracture, studies have considered stress-based topology optimization,[4,5] where linear elastic material behavior is assumed with a failure criterion based on von Mises[4,5] and Drucker–Prager models,[6] among others. Although promising, this design approach is confined to elastic material behavior, and therefore, cannot account for inelastic processes in ductile materials, such as elastoplastic energy dissipation and ductile damage accumulation.

Consideration of inelastic material response in topology optimization toward the development of fracture-resistant designs is challenging, and entails path-dependent sensitivity analysis,[7,8] and also leads to related issues such as convergence in finite element analysis due to material damage[9] and complex material interpolation schemes for inelastic constitutive models.[10,11] In the past, fracture-resistant structural designs have been explored with various continuum damage mechanics (CDM) based models, for example, ductile designs with Lemaitre and nonlocal CDM models,[9,11,12] brittle fracture mitigated designs with nonlocal CDM models[10,13] and phase-field models.[14–16] For instance, Kato et al.[10] designed nonlinear elastoplastic reinforced composites by maximizing the total work given an upper limit on material volume, while Amir and Sigmund[13] investigated the stiffness design of reinforced concrete structures with constrained material volume, both with nonlocal inelastic-damage model for concrete while considering different models for reinforcement and concrete-reinforcement interface. Kai and Waisman[12] and Russ and Waisman[14] considered topological design for minimizing material usage with stiffness and damage constraints where material damage is modeled with nonlocal damage and phase-field fracture models, respectively; Li et al.[9,11] employed topology optimization for maximizing plastic energy dissipation with constrained material volume and damage using local and nonlocal damage models. In a recent study, Li and Khandelwal[17] investigated elastoplastic energy-absorbing designs with controlled damage with a micromechanics-based Gurson–Tvergaard–Needleman porous plasticity model. Studies have also exploited uncoupled damage models, for example, the Johnson–Cook failure criterion and Crach-FEM fracture model for damage-proof designs.[18–20] Irrespective of these advancements, all the previous works on the fracture-resistant topology optimization are confined to the small strain regime. However, fracture initiation for many ductile metals (e.g., structural grade steels ASTM A997, A36), is often preceded by large (plastic) deformation, and the small strain assumption might not be applicable. While a similar design philosophy can be used with finite deformation-based ductile fracture models, such an extension is yet not available and is the focus of this study.

Ductile damage models for metals under finite deformations consider the progressive loss of load-carrying capacity due to nucleation, shearing, growth, and coalescence of micro voids.[21–23] For micromechanically motived damage approaches, along with the earlier works,[24,25] Gurson[26] first established a yield criterion for porous solids based on homogenization of an existing spherical void embedded in a von Mises plastic solid. With a pre-existing void that is assumed to remain spherical during loading, the homogenized yield surface is pressure-sensitive and shrinks as the voids grow. Later, modifications to the Gurson model were made by Chu and Needleman,[27] Tvergaard,[28] and Tvergaard and Needleman[29] to account for void nucleation and coalescence. The resulted model is generally referred to as the Gurson–Tvergaard–Needleman (GTN) model and is mostly used for modeling ductile fracture under high-stress triaxiality. Under low-stress triaxiality, the influence of void shape change (elongation) due to shear deformations was investigated in fracture tests,[30,31] where the influence of the third deviatoric stress invariant (or Lode parameter) on the void shape evolution and coalescence was demonstrated. Based on these results, various shear modifications have been proposed, for example, Xue,[32] Nahshon and Hutchinson,[33] Malcher et al.,[34] Zhou et al.,[35] and Dæhli et al.[36] A review of the porous plasticity for the ductile damage modeling can be found in Reference 21. Despite this progress, most GTN-type model implementations in the literature are confined to the small strain regime.[37–39] As ductile damage often accompanies large deformation, the formulation and implementation of ductile fracture models in the finite strain regime are crucial. More importantly, the predictive capabilities of these advanced finite strain-based damage models can be employed in an *inverse topology design* framework, wherein the goal is to design fracture/damage resistant elastoplastic energy-dissipating systems under large deformations. In essence, the design task will entail designing a structural topology guided by the physics of finite strain elastoplastic damage such that the designed structures have desirable performance measured in terms of fracture-resistance and energy dissipation.

The contribution of this study is to formulate the shear-modified GTN model proposed by Nielsen and Tvergaard[40] in a consistent finite strain framework with numerical implementation details and to extend the previous work on the small

strain GTN guided fracture-resistant designs[17] to finite strain applications. Specifically, a finite-strain shear-modified GTN model is first developed, which is formulated based on multiplicative elasto-plasticity and combines different features, that is, shear modifications[33,40] and numerical improvements,[41] from the previous studies. A series of 2D and 3D numerical simulations are carried out to demonstrate the performance of the implemented model. This finite strain based ductile fracture model is then incorporated in a density-based topology optimization framework. To this end, based on this model, an analytical path-dependent sensitivity calculation using an adjoint method is derived. With the incorporation of techniques such as adaptive linear energy interpolation for addressing mesh distortion[42] and p-norm approximation of maximum damage field,[5] the design of fracture-resistant energy-dissipating structures is explored under finite strains. Ultimate postperformance analysis of the optimized designs is also carried out to investigate different critical fracture mechanisms.

The rest of the article is organized as follows. In Section 2, a finite strain shear modified GTN model is presented. A brief review of F-bar element formulation that is used in the numerical implementation is presented in Section 3. A density-based topology optimization framework for the fracture-resistant design task is developed in Section 4. Section 5 gives the derivation of the path-dependent sensitivity analysis. In Section 6, a series of numerical examples are carried out to verify model implementation and to demonstrate the model performance. In Section 7, the path-dependent sensitivity calculation is verified together with two design optimization examples. Finally, concluding remarks are given Section 8.

2 | FINITE STRAIN SHEAR MODIFIED GTN MODEL

A finite strain shear modified GTN model based on the multiplicative elastoplastic formulation is presented in this section.

2.1 | Kinematics

Let $\Omega_0 \in \mathbb{R}^3$ be the reference configuration of a continuum body of interest with $\boldsymbol{X}$ representing arbitrary material point position in Ω_0, that is, $\boldsymbol{X} \in \Omega_0$. A motion of the body that carries it from its reference configuration to its current configuration $\Omega_t \in \mathbb{R}^3$ can be described by a smooth one-to-one mapping $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t)$ where $t \in \mathbb{R}_+$ denotes time. The associated local deformation gradient is defined by $\boldsymbol{F} := \boldsymbol{\nabla}_{\boldsymbol{X}}\boldsymbol{\varphi}$ with $\det \boldsymbol{F} > 0$. With the multiplicative decomposition assumption,[43] the deformation gradient is split into elastic and inelastic parts by

$$\boldsymbol{F} = \boldsymbol{F}^e.\boldsymbol{F}^p, \tag{1}$$

where $\boldsymbol{F}^p$ represents a local inelastic deformation that can be attributed to dislocations of the crystal lattice from a micromechanics viewpoint and $\boldsymbol{F}^e$ represents elastic distortions. Other strain measures that will be useful in the following model descriptions are plastic right Cauchy–Green tensor $\boldsymbol{C}^p$ and left elastic Cauchy–Green tensor (also called elastic Finger tensor) $\boldsymbol{b}^e$ that are defined, respectively, as.

$$\boldsymbol{C}^p = \boldsymbol{F}^{pT}.\boldsymbol{F}^p \; and \; \boldsymbol{b}^e = \boldsymbol{F}^e.\boldsymbol{F}^{eT} \tag{2}$$

which are related to each other by pull-back and push-forward operations.

$$\boldsymbol{C}^{p-1} = \boldsymbol{F}^{-1}.\boldsymbol{b}^e.\boldsymbol{F}^{-T} \; and \; \boldsymbol{b}^e = \boldsymbol{F}.\boldsymbol{C}^{p-1}.\boldsymbol{F}^T. \tag{3}$$

The material time derivative of the elastic Finger tensor $\dot{\boldsymbol{b}}^e$ can be expressed as.

$$\dot{\boldsymbol{b}}^e = \mathcal{L}_v\left[\boldsymbol{b}^e\right] + \boldsymbol{l}.\boldsymbol{b}^e + \boldsymbol{b}^e.\boldsymbol{l}^T \; with \; \mathcal{L}_v\left[\boldsymbol{b}^e\right] = \boldsymbol{F}.\dot{\overline{\boldsymbol{C}^{p-1}}}.\boldsymbol{F}^T, \tag{4}$$

where $\mathcal{L}_v\left[\boldsymbol{b}^e\right]$ denotes Lie derivative, $\boldsymbol{l}$ the spatial velocity gradient given by $\boldsymbol{l} = \dot{\boldsymbol{F}}.\boldsymbol{F}^{-1}$ and $\dot{\overline{\boldsymbol{C}^{p-1}}}$ the material time derivative of the inverse of $\boldsymbol{C}^p$.

2.2 | Free energy

The Helmholtz free energy for isotropic elastic–plastic solids is assumed to take the form

$$\psi = \psi^e\left(\boldsymbol{b}^e\right) + \psi^p(\alpha), \tag{5}$$

where $\psi^e\left(\boldsymbol{b}^e\right)$ denotes elastic strain energy at macroscale while $\psi^p(\alpha)$ accounts for isotropic hardening, and α is the microscopic effective plastic strain of the undamaged matrix material. It is noted that kinematic hardening is not included in the current framework. The elastic strain energy function used in this study is defined as.[41,44]

$$\begin{aligned}
\psi^e\left(\boldsymbol{b}^e\right) &= \psi^e_{vol}\left(J^e\right) + \psi^e_{iso}\left(\widehat{\boldsymbol{b}}^e\right),\\
\psi^e_{vol}\left(J^e\right) &= \frac{1}{2}\kappa\left(\ln J^e\right)^2,\\
\psi^e_{iso}\left(\widehat{\boldsymbol{b}}^e\right) &= \frac{1}{4}\mu\left[\boldsymbol{I} : \left(ln\widehat{\boldsymbol{b}}^e\right)^2\right],
\end{aligned} \tag{6}$$

where $J^e := \det \boldsymbol{F}^e$ represents elastic volumetric deformation and $\widehat{\boldsymbol{b}}^e := J^{e-\frac{2}{3}}\boldsymbol{b}^e$ quantifies isochoric deformations. Here, κ and μ are the bulk and shear modulus. The use of logarithmic Hencky strains $\ln J^e$ and $\ln \widehat{\boldsymbol{b}}^e$ in expressing the stored energy is advantageous for numerical implementations, as shown in Appendix A. The energy due to plastic hardening $\psi^p(\alpha)$ is not explicitly specified here, instead, its derivative $\partial\psi^p/\partial\alpha$ is prescribed in the next section, motivated from a thermodynamics perspective.

2.3 | Thermodynamics

Thermodynamics require energy dissipation D_{int} to be nonnegative during any deformation, that is,

$$D_{\text{int}} = \boldsymbol{\tau} : \boldsymbol{d} - \dot{\psi} = \left(\boldsymbol{\tau} - 2\frac{\partial\psi^e}{\partial\boldsymbol{b}^e}.\boldsymbol{b}^e\right) : \boldsymbol{d} - \left(2\frac{\partial\psi^e}{\partial\boldsymbol{b}^e}.\boldsymbol{b}^e\right) : \left(\frac{1}{2}\mathcal{L}_v\left[\boldsymbol{b}^e\right].\boldsymbol{b}^{e-1}\right) - \frac{\partial\psi^p}{\partial\alpha}\dot{\alpha} \geq 0, \tag{7}$$

where $\boldsymbol{d} = \frac{1}{2}\left(\boldsymbol{l} + \boldsymbol{l}^T\right)$ is the rate of deformation tensor. Equation (7) utilizes the decomposition of $\dot{\boldsymbol{b}}^e$ in Equation (4) and the fact that $\partial\psi^e/\partial\boldsymbol{b}^e$ commutes with $\boldsymbol{b}^e$ due to isotropy. Following the standard Coleman–Noll procedure,[45] the following relationships can be established

$$\boldsymbol{\tau} = 2\frac{\partial\psi^e}{\partial\boldsymbol{b}^e}.\boldsymbol{b}^e, \tag{8}$$

$$D_{int} = \boldsymbol{\tau} : \left(-\frac{1}{2}\mathcal{L}_v\left[\boldsymbol{b}^e\right].\boldsymbol{b}^{e-1}\right) + h\dot{\alpha} \geq 0, \tag{9}$$

where

$$h := -\frac{\partial\psi^p}{\partial\alpha}, \tag{10}$$

specifies isotropic hardening rule, where h is the thermodynamic variable conjugate to α. A function with combined linear and saturation hardening terms are used to describe this variable, that is,

$$h(\alpha) = k_1\alpha + k_2(1 - exp(-\delta\alpha)), \tag{11}$$

where k_1 and k_2 control the magnitudes of the linear hardening and saturated hardening, respectively, and δ controls the saturation rate.

2.4 | Yield criterion

The yield criterion for porous metals was first proposed by Gurson,[26] which was later modified by Chu and Needleman[27] for void nucleation, Tvergaard and Needleman[29] for void coalescence, Nahshon and Hutchinson,[33] and Nielsen and

Tvergaard[40] for void shearing. As suggested by Mahnken,[41] the yield function is formulated in a linear form, as opposed to quadratic form, as

$$\phi = \|\boldsymbol{\tau}_{dev}\| - \sqrt{\frac{2}{3}}\, sign(z)\sqrt{|z|}\, \zeta J \tag{12}$$

with

$$z = 1 + q_3 f^{*2} - 2q_1 f^* Cosh[M], \tag{13}$$

$$\zeta(\alpha) = \sigma_{y0} + h(\alpha), \tag{14}$$

$$f^* = \begin{cases} f & \text{if } f \le f_C \\ f_C + K(f - f_C) & \text{if } f > f_C \end{cases} \quad \text{with } K = \frac{f_U^* - f_C}{f_F - f_C}, \tag{15}$$

$$Cosh[M] := \begin{cases} \cosh M & \text{if } M \le M_c \\ \cosh M_c + \sinh M_c\,(M - M_c) + \frac{1}{2}\cosh M_c (M - M_c)^2 & \text{if } M > M_c \end{cases} \tag{16}$$

$$M = \tfrac{3}{2} q_2 \tfrac{1}{\zeta} \tfrac{\tau_m}{J},$$

where $\|.\|$ is the Euclidean norm that results in isotropic yield criterion, and extension to anisotropic yield criterion can be found in References 39 and 46. Here, $\tau_m = \mathrm{tr}(\boldsymbol{\tau})/3$ is the Kirchhoff pressure and $\boldsymbol{\tau}_{dev} = \boldsymbol{\tau} - \tau_m \boldsymbol{I}$ is the deviatoric part of $\boldsymbol{\tau}$; $J = \det \boldsymbol{F}$ is the Jacobian; ζ is the microscopic effective stress of the undamaged matrix material with σ_{y0} the initial yield stress (the relationship between ζ and α is characterized as a uniaxial stress–strain curve); f represents void volume fraction; q_1, q_2, and q_3 are model parameters; parameter f_C denotes the critical void volume fraction beyond which accelerated (by coefficient $K > 1$) void volume fraction growth is expected due to voids coalescence; parameter f_F represents the void volume fraction at *final fracture* ($f = f_F$); f_U^* stands for the *ultimate value* of the modified void volume fraction ($f^* = f_U^*$) at which macroscopic stress carrying capacity vanishes. As a result, the coefficient K is determined by $f^*(f_F) = f_U^*$. The value of f_U^* can be determined through $\lim_{\boldsymbol{\tau}\to\boldsymbol{0}} \phi(f^*) = 0$, which gives $f_U^* = 1/q_1$ when $q_3 = q_1^2$ (chosen in this study). The modification of $\cosh(\cdot)$ to $Cosh(\cdot)$ is suggested in Reference 41 for numerical stability, where M_c is a threshold parameter.

2.5 | Flow rule and void volume fraction evolution law

The flow rule for the macroscopic plastic strain ($\boldsymbol{F}^p$ or $\boldsymbol{C}^p$ or equivalently $\boldsymbol{b}^e$) is derived following the principle of maximum dissipation,[47] that is,

$$-\frac{1}{2}\mathcal{L}_v\left[\boldsymbol{b}^e\right].\boldsymbol{b}^{e-1} = \gamma \boldsymbol{A} \;\text{ with }\; \boldsymbol{A} := \frac{\partial \phi}{\partial \boldsymbol{\tau}}, \tag{17}$$

where γ is the plastic multiplier. The flow rule for the microscopic effective plastic strain is obtained by enforcing the condition that the plastic work in the microscopic matrix material is equal to the macroscopic plastic work, that is, $(1-f)\zeta\dot{\alpha} = \frac{1}{J}\boldsymbol{\tau} : \gamma \boldsymbol{A}$, which yields

$$\dot{\alpha} = \frac{\gamma}{(1-f)\zeta J}\boldsymbol{\tau} : \boldsymbol{A}. \tag{18}$$

The flow rule for the void volume fraction $\dot{f}$ is considered to be the sum of contributions from void growth ($\dot{f}_g$), nucleation ($\dot{f}_n$) and shearing ($\dot{f}_s$)

$$\dot{f} = \dot{f}_g + \mathscr{k}\dot{f}_n + \dot{f}_s \;\text{ with }\; \mathscr{k}(\tau_m) = \begin{cases} 1 \text{ if } \tau_m \ge 0 \\ 0 \text{ if } \tau_m < 0 \end{cases}, \tag{19}$$

where the parameter $\mathscr{k}$ is used to control the void nucleation rate such that nucleation only happens under tension, that is, positive hydrostatic stress. In Equation (19), the void growth rate is governed by the conservation of mass

$$\dot{f}_g = \gamma(1-f)\boldsymbol{A} : \boldsymbol{I} \tag{20}$$

and the void nucleation and shearing are postulated, respectively, as[27]

$$\dot{f}_n = D\dot{\alpha} = \frac{\gamma}{(1-f)\zeta J} D\boldsymbol{\tau} : \boldsymbol{A} \text{ with } D := \frac{f_N}{s_N\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\alpha-\alpha_N}{s_N}\right)^2} \tag{21}$$

and[33,40]

$$\dot{f}_s = \gamma k_\omega f \omega_0 \frac{\boldsymbol{\tau}_{dev}}{\tau_e} : \boldsymbol{A} \text{ with } \tau_e = \sqrt{\frac{3}{2}}\|\boldsymbol{\tau}_{dev}\| \text{ and } \omega_0 = \omega(\boldsymbol{\tau})\Omega(T) \tag{22}$$

with

$$\omega(\boldsymbol{\tau}) = 1 - \left(\frac{27}{2}\frac{\check{J}_3}{\tau_e^3}\right)^2 \text{ with } \check{J}_3 = det\, \boldsymbol{\tau}_{dev} \tag{23}$$

$$\Omega(T) = \begin{cases} 1 & \text{for } T < T_1 \\ \frac{(T-T_2)^2}{\Delta T^2}\left[3 + \frac{2(T-T_2)}{\Delta T}\right] & \text{for } T_1 \le T \le T_2 \text{ with } \Delta T = T_2 - T_1, \\ 0 & \text{for } T > T_2 \end{cases} \tag{24}$$

$$T := \frac{\tau_m}{\tau_e}$$

where in the void nucleation term in Equation (21) suggested by Chu and Needleman,[27] f_N denotes the void volume fraction to be nucleated, s_N the standard deviation and α_N the mean nucleation strain. The shear modification in Equation (22) introduced by Nahshon and Hutchinson[33] is to account for softening and fracture under shear deformation which was later modified by adding function $\Omega(T)$ by Nielsen and Tvergaard[40] for eliminating shear modifications at high-stress triaxiality. The motivation of function $\omega(\boldsymbol{\tau}) \in [0,1]$ is to distinguish axisymmetric stress states ($\omega = 0$) and all states consisting of a pure shear stress plus a hydrostatic pressure ($\omega = 1$).[33] However, as discussed in Nielsen and Tvergaard,[40] for some stress states high triaxiality can be along with nonzero ω value. To control the effect of the shear modification term with the stress triaxiality T, function $\Omega(T)$ is introduced by Nielsen and Tvergaard.[40] Here, k_ω controls the magnitude of damage growth rate in pure shear states, T_1 and T_2 are chosen parameters ($T_1 < T_2$) such that the shear modification by Nahshon and Hutchinson[33] is used when $T \le T_1$ and is not present when $T \ge T_2$.[40] The integration of these rate equations in Equations (17)–(19) is given in Appendix A.

2.6 | KKT and consistency conditions

The plastic loading and unloading processes are described by Karush–Kuhn–Tucker (KKT) conditions

$$\gamma \ge 0, \phi \le 0, \gamma\phi = 0 \tag{25}$$

and the consistency condition

$$\gamma\dot{\phi} = 0. \tag{26}$$

2.7 | Remarks

The GTN model will degenerate to a finite strain J_2 plasticity model if the parameters

$$f_0 = q_1 = q_3 = f_N = k_\omega = 0 \tag{27}$$

are chosen.

The numerical implementation details of the shear modified GTN model are given in Appendix A.

3 | FINITE ELEMENT ANALYSIS WITH F-BAR FORMULATION

In the finite element simulations with the GTN model, both 3-D and 2-D plane strain conditions are considered, and F-bar (4-node quadrilateral for 2-D and 8-node brick for 3-D) elements are used. Although the plastic flow in GTN model is compressible and volumetric locking might not be severe, F-bar element formulation[48] is still used in this study in order to handle locking in the case of J_2 plasticity. For completeness, a brief review of F-bar formulation is given here. First, a modified deformation gradient $\overline{\boldsymbol{F}}$ is defined as

$$\overline{\boldsymbol{F}} = \begin{cases} r^{1/2}\boldsymbol{F}\,(\text{in}-\text{plane part}) & \text{for } 2-\text{D plane strain} \\ r^{1/3}\boldsymbol{F} & \text{for } 3-\text{D} \end{cases}, \quad \text{with } r = \frac{det\boldsymbol{F}_0}{det\boldsymbol{F}} \tag{28}$$

where $\boldsymbol{F}$ and $\boldsymbol{F}_0$ are the deformation gradients evaluated at the current quadrature point and the centroid of the element, respectively. The first Piola–Kirchhoff (PK) stress $\boldsymbol{P}$ is then computed as

$$\boldsymbol{P} = \begin{cases} r^{-1/2}\overline{\boldsymbol{P}}\,(\text{in}-\text{plane part}) & \text{for } 2-\text{D plane strain} \\ r^{-2/3}\overline{\boldsymbol{P}} & \text{for } 3-\text{D} \end{cases}, \tag{29}$$

where $\overline{\boldsymbol{P}}$ is the first PK stress evaluated by any (elastic or inelastic) constitutive model with the deformation gradient input as $\overline{\boldsymbol{F}}$.

The finite element discretization of the weak form of balance of linear momentum gives the global equilibrium equations

$$\begin{aligned} &\boldsymbol{R}(\boldsymbol{u}) = \boldsymbol{F}_{int}(\boldsymbol{u}) - \boldsymbol{F}_{ext} = \boldsymbol{0}, \\ &\boldsymbol{F}_{int}(\boldsymbol{u}) = \mathop{\mathcal{A}}_{e=1}^{n_{ele}} \boldsymbol{F}^e_{int}(\boldsymbol{u}_e),\ \boldsymbol{F}^e_{int}(\boldsymbol{u}_e) = \int_{\Omega^e_0} \boldsymbol{B}^T\boldsymbol{P}\, dV, \end{aligned} \tag{30}$$

where $\boldsymbol{B}$ denotes the derivative of shape functions and $\boldsymbol{F}_{ext}$ is the external force vector evaluated in the standard way.

4 | DENSITY-BASED TOPOLOGY OPTIMIZATION

In the density-based topology optimization,[2] a design is described by an element-wise constant density field $\rho(\boldsymbol{X})$ that indicates the presence ($\rho = 1$) or absence ($\rho = 0$) of the material in an element. To accommodate gradient-based optimizers, the discrete density variables are relaxed to continuous values, that is, $\rho \in [0, 1]$, where $0 \le \rho \le 1$ represents the mixture of void and solid phases. This density-based approach for structural topology optimization is illustrated in Figure 1, where the design domain is described by a FE discretized density field that is optimized for predefined objective and constraint functions. During the optimization process, FE analysis and sensitivity analysis are carried out at each optimization iteration to examine the optimality conditions and to provide gradient information for the next design update. The conceptual design is provided by the final optimized topology but with potential gray areas, which is further postprocessed via B-splines fitting to obtain the discrete design that can be readily manufactured.

4.1 | Topology optimization formulation

In this study, the objective is to maximize the plastic work together with constraints on material damage (f) and material usage. To this end, the mathematical optimization formulation reads

$$\begin{aligned} &\min_{\boldsymbol{x}} f_0(\boldsymbol{x}) = -W^p \\ &\text{s.t. } \boldsymbol{R}(\boldsymbol{u}) = \boldsymbol{0} \end{aligned} \tag{31}$$

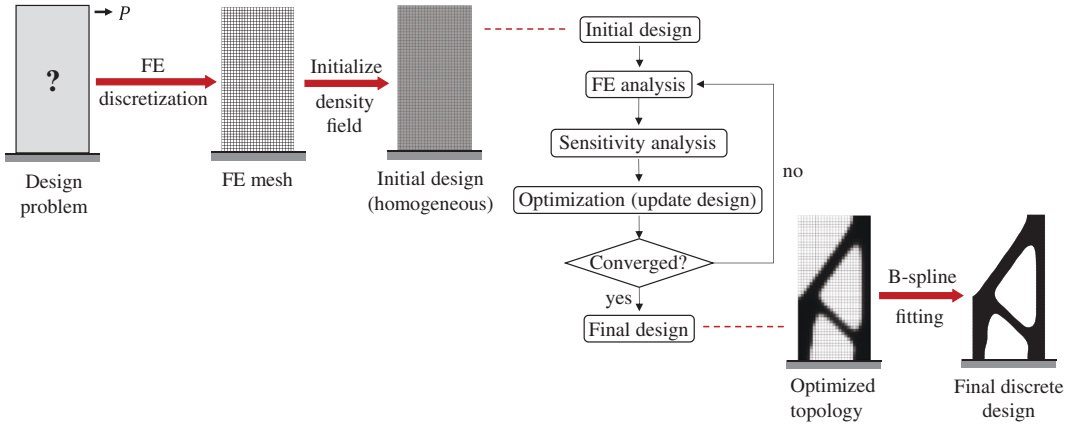


FIGURE 1 Flow chart of the structural topology optimization

$$\begin{aligned} \boldsymbol{H}\left(\boldsymbol{C}^p, \alpha, f, \ldots\right) &= \boldsymbol{0} \\ f_1(\boldsymbol{x}) &= \frac{D_{max}}{\overline{D}} - 1 \leq 0 \\ f_2(\boldsymbol{x}) &= \frac{1}{V}\sum_{e=1}^{n_{ele}} \rho(\boldsymbol{x}) v_e - V_f \leq 0 \\ 0 \leq x_e \leq 1, e &= 1, \ldots, n_{ele}, \end{aligned}$$

where the equality constraints $\boldsymbol{R}(\boldsymbol{u}) = \boldsymbol{0}$ and $\boldsymbol{H}\left(\boldsymbol{C}^p, \alpha, f, \ldots\right) = \boldsymbol{0}$ that represents the structural global equilibrium and local constitutive equations are not enforced by the optimization algorithm, but instead are satisfied by solving the structural responses (displacement $\boldsymbol{u}$ and internal variables e.g., $\boldsymbol{C}^p$, $\alpha, f, \ldots$) that in turn govern the objective and constraint functions values. This nested solution approach is commonly used in large-scale topology optimization.[49] Here x_e $(e = 1, \ldots, n_{ele})$ denote the design variables that are mapped to the density variables ρ_e $(e = 1, \ldots, n_{ele})$ through the density filter.[50]

$$\boldsymbol{\rho} = \boldsymbol{W}\boldsymbol{x} \text{ with } W_{ij} = \frac{w_{ij} v_j}{\sum_{k=1}^{n_{ele}} w_{ik} v_k} \text{ and } w_{ij} = max\left(r_{\min} - \left\|\boldsymbol{X}_i - \boldsymbol{X}_j\right\|, 0\right) \tag{32}$$

in which $r_{\min}$ is the filter radius, eth element and $\boldsymbol{X}_i$ the coordinates of the centroid of the ith element. The density filter is used to address mesh dependency and checkerboard pattern of the optimized topology.[51]

The plastic work, W^p, during the loading process $t \in [0, T]$ is calculated by

$$W^p = \int_T \int_{\Omega_0} (\boldsymbol{\tau} : \gamma \boldsymbol{A})\, dV \approx \sum_{k=1}^{n} \sum_{e=1}^{n_{ele}} \sum_{s=1}^{n_{ipt}} \rho_e^{p_w} \left(\boldsymbol{\tau}_{e_s}^k : \Delta\gamma_{e_s}^k \boldsymbol{A}_{e_s}^k\right) w_{e_s}, \tag{33}$$

where the subscripts s, e, and k denote the indices for integration point, element, and loading step, respectively. Here, w_{e_s} is the product of Jacobian and quadrature weight at sth integration point in eth element. A penalization on the intermediate densities $(\rho < 1)$ with power $p_w > 1$ is introduced for optimization purposes for suppressing intermediate densities.

The maximum damage $D_{\max}$ is approximated through a p-norm function of the void volume fraction by

$$D_{\max} = \left(\sum_{e=1}^{n_{ele}} \left(\sum_{s=1}^{n_{ipt}} \left(f_{e_s}^n\right)^{p_d}\right)\right)^{1/p_d}, \tag{34}$$

where a large value is usually used for the power p_d to approximate the maximum value of void volume fraction f, which is the measure of damage in the domain, at the last step n.

4.2 | Material interpolation by density variable

The dependence of the objective and constraint functions on the design variables $\boldsymbol{x}$ is through the interpolation of the material constitutive parameters by the density variables $\boldsymbol{\rho}$ as well as penalizations in Equations (33) and (34). Table 3 lists the model parameters that are used in the shear modified GTN model for topology optimization, which are taken as the model parameters for the solid elements. For intermediate densities $\rho < 1$, the model parameters are interpolated as

$$\begin{aligned} E(\rho) &= \left[\varepsilon_1 + (1-\varepsilon_1)\,\rho^{p_1}\right] E, \\ \sigma_{y0}(\rho) &= \left[\varepsilon_2 + (1-\varepsilon_2)\,\rho^{p_2}\right] \sigma_{y0}, \\ k_1(\rho) &= \left[\varepsilon_1 + (1-\varepsilon_1)\,\rho^{p_2}\right] k_1, \\ k_2(\rho) &= \left[\varepsilon_1 + (1-\varepsilon_1)\,\rho^{p_2}\right] k_2, \end{aligned} \tag{35}$$

while other parameters remain constant irrespective of the density values. Here, ε_1 and ε_2 are lower bounds, p_1 and p_2 are penalization powers.

To ensure the convergence of finite element analysis, especially during early stage of optimization, where many intermediate density elements are present, a reduction of the damage evolution in the intermediate densities is introduced by replacing Equation (19) with

$$\dot{f} = \rho^{p_f}\left(\dot{f}_g + \cancel{k}\dot{f}_n + \dot{f}_s\right), \tag{36}$$

where penalization power p_f changes from greater than zero to zero during the optimization via continuation. Note that Equation (19) is recovered when $p_f = 0$.

4.3 | Adaptive linear energy interpolation for mesh distortion

In the density-based topology optimization, due to the material interpolation by density (see Section 4.2), elements with low density have low stiffness. As a result, during the loading process, these low-stiffness elements can undergo large distortions, which can hinder the convergence of finite element analysis. To address that, a linear energy interpolation scheme that is first proposed by Wang et al.[52] and later extended to an adaptive scheme in Zhang et al.[42] is utilized. The basic idea is to split the element deformation energy into two parts, one formulated with small deformation theory and the other one with finite deformation theory. The interpolation of the element energy is based on its density variable ρ such that when $\rho = 0$ small deformation theory is used, while when $\rho = 1$ finite deformation theory is used. To this end, the deformation gradient $\boldsymbol{F}$ is interpolated as

$$\boldsymbol{F} = \boldsymbol{I} + \eta \nabla_{\boldsymbol{X}} \boldsymbol{u} \;\text{ with }\; \eta(\rho) = \frac{exp(\beta\rho)}{exp(c\beta) + exp(\beta\rho)}, \tag{37}$$

where c and β are interpolation parameters chosen as $\beta = 120$ and $c = 0.08$ with c value adaptively updated (if needed) using the scheme proposed in Zhang et al.[42] It should be noted that the F-bar formulation in Section 3 is based on the interpolated deformation gradient in Equation (37). The element internal force in Equation (30) is accordingly modified to

$$\boldsymbol{F}^e_{int} = \int_{\Omega^e_0} \eta \boldsymbol{B}^T \boldsymbol{P} dV + \int_{\Omega^e_0} \left(1-\eta^2\right) \boldsymbol{B}^T_L [\mathbb{C} : \boldsymbol{\varepsilon}] dV, \tag{38}$$

where $\boldsymbol{B}_L$ denotes the matrix of shape functions derivatives that gives the vector form of the small strain measure when applied to displacement, that is, $[\boldsymbol{\varepsilon}] = \boldsymbol{B}_L \boldsymbol{u}$ with $\boldsymbol{\varepsilon} := \nabla^s_{\boldsymbol{X}} \boldsymbol{u}$, and $\mathbb{C}$ is the linear isotropic elastic moduli determined by the interpolated Young's modulus $E(\rho) = [\varepsilon_1 + (1-\varepsilon_1)\,\rho^{p_L}]\,E$ and a constant Poisson's ratio ν (E and ν are the same as that of the GTN model) where p_L is the penalization power.

4.4 | Parameter settings

In the topology optimization, continuation scheme is used to slowly increase the penalization powers p_1, p_2, p_w, and p_L and gradually decrease relaxation power p_f to avoid analysis failure during early optimization iterations due to large deformations, and to relax the nonconvexity of the original optimization problem. Specifically, p_1 is raised from 1 to 5 by an increment 0.2 per 25 iterations; p_2 and p_w are raised from 1 to 3 by 0.1 per 25 iterations; p_f is reduced from 3 to 0 by -0.1 every 25 iterations; p_L is increased from 4 to 6 by 0.1 per 25 iterations. Here, the power p_d in the p-norm function in Equation (34) is fixed to be 15. The lower bounds are chosen as $\varepsilon_1 = 10^{-8}$ and $\varepsilon_2 = 10^{-3}$ in Equation (35). Other optimization parameters, for example, r_{min} in Equation (32), $\overline{D}$ and V_f in Equation (31), are specified later for different problem settings. It should be noted that the parameters should be carefully chosen for successful optimization. Following is the rationale for the selection of some of the above parameters: (a) p_1 is chosen to be greater than p_2 to reduce both energy dissipation and damage in the low-density elements to penalize intermediate densities and relax design space; (b) p_L is chosen to be greater than p_1 to discourage using low-density elements to exploit small deformation kinematics[42]; and (c) the power p_d in the p-norm function is chosen to be 15 to balance the nonconvexity and accuracy of the maximum damage approximation.[17]

5 | SENSITIVITY ANALYSIS

Due to the path-dependent behavior of the elastoplastic GTN model, the sensitivity analysis is also path-dependent. The adjoint method is adopted for sensitivity analysis, as the number of design variables far exceeds the number of objective and constraint functions. The calculations follow the framework proposed by Michaleris et al.,[53] which is expanded in the work of Alberdi et al.[8] for inelastic topology optimization problems.

5.1 | Adjoint formulation

The adjoint function is constructed as

$$\widehat{F} = F\left(\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n, \boldsymbol{v}^1, \ldots, \boldsymbol{v}^n, \boldsymbol{\rho}\right) + \sum_{k=1}^{n} {\lambda^k}^T \boldsymbol{R}^k\left(\boldsymbol{u}^k, \boldsymbol{u}^{k-1}, \boldsymbol{v}^k, \boldsymbol{v}^{k-1}, \boldsymbol{\rho}\right) + \sum_{k=1}^{n} {\boldsymbol{\mu}^k}^T \boldsymbol{H}^k\left(\boldsymbol{u}^k, \boldsymbol{u}^{k-1}, \boldsymbol{v}^k, \boldsymbol{v}^{k-1}, \boldsymbol{\rho}\right), \tag{39}$$

where F represents the objective (f_0) or constraint (f_1 or f_2) function, $\boldsymbol{u}^k$ and $\boldsymbol{v}^k$ are the solution and auxiliary variables at step k and are determined by the corresponding global equilibrium ($\boldsymbol{R}^k = \boldsymbol{0}$) and local constitutive equations ($\boldsymbol{H}^k = \boldsymbol{0}$), and λ^k and $\boldsymbol{\mu}^k$ are the corresponding adjoint variables. The goal here is to compute the derivative $dF/d\boldsymbol{x}$ which can be immediately obtained from the derivative $dF/d\boldsymbol{\rho}$ through simple chain using Equation (32). Hence, the main effort is to compute the derivatives $dF/d\boldsymbol{\rho}$.

Because the equilibrium and constitutive equations are always satisfied irrespective of the density variables $\boldsymbol{\rho}$, it is clear that $d\widehat{F}/d\boldsymbol{\rho} \equiv dF/d\boldsymbol{\rho}$. Taking derivatives of $\widehat{F}$ with respect to $\boldsymbol{\rho}$ and eliminating all terms that contain the implicit derivatives $d\boldsymbol{u}^k/d\boldsymbol{\rho}$ and $d\boldsymbol{v}^k/d\boldsymbol{\rho}$ yields

$$\frac{d\widehat{F}}{d\boldsymbol{\rho}} = \frac{\partial \widehat{F}}{\partial \boldsymbol{\rho}} + \sum_{k=1}^{n}\left({\lambda^k}^T \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{\rho}} + {\boldsymbol{\mu}^k}^T \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{\rho}}\right), \tag{40}$$

where the adjoint variables λ^k and $\boldsymbol{\mu}^k$ are calculated in backward order from the following system of equations.

$$\begin{gathered}
\text{step } n: \\
\begin{cases} \frac{\partial F}{\partial \boldsymbol{u}^n} + {\lambda^n}^T \frac{\partial \boldsymbol{R}^n}{\partial \boldsymbol{u}^n} + {\boldsymbol{\mu}^n}^T \frac{\partial \boldsymbol{H}^n}{\partial \boldsymbol{u}^n} = \boldsymbol{0}, \\ \frac{\partial F}{\partial \boldsymbol{v}^n} + {\lambda^n}^T \frac{\partial \boldsymbol{R}^n}{\partial \boldsymbol{v}^n} + {\boldsymbol{\mu}^n}^T \frac{\partial \boldsymbol{H}^n}{\partial \boldsymbol{v}^n} = \boldsymbol{0}, \end{cases} \\
\text{step } k\ (k = n-1, \ldots, 2, 1): \\
\begin{cases} \frac{\partial F}{\partial \boldsymbol{u}^k} + {\lambda^{k+1}}^T \frac{\partial \boldsymbol{R}^{k+1}}{\partial \boldsymbol{u}^k} + {\boldsymbol{\mu}^{k+1}}^T \frac{\partial \boldsymbol{H}^{k+1}}{\partial \boldsymbol{u}^k} + {\lambda^k}^T \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{u}^k} + {\boldsymbol{\mu}^k}^T \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{u}^k} = \boldsymbol{0}, \\ \frac{\partial F}{\partial \boldsymbol{v}^k} + {\lambda^{k+1}}^T \frac{\partial \boldsymbol{R}^{k+1}}{\partial \boldsymbol{v}^k} + {\boldsymbol{\mu}^{k+1}}^T \frac{\partial \boldsymbol{H}^{k+1}}{\partial \boldsymbol{v}^k} + {\lambda^k}^T \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{v}^k} + {\boldsymbol{\mu}^k}^T \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{v}^k} = \boldsymbol{0}. \end{cases}
\end{gathered} \tag{41}$$

Finally, all the explicit derivatives needed to complete the adjoint sensitivity calculation in Equation (40) are

$$\begin{aligned}
&\text{For } F: \frac{\partial F}{\partial \boldsymbol{\rho}}, \frac{\partial F}{\partial \boldsymbol{u}^k}, \frac{\partial F}{\partial \boldsymbol{v}^k},\\
&\text{For } \boldsymbol{R}^k: \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{\rho}}, \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{u}^k}, \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{u}^{k-1}}, \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{v}^k}, \frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{v}^{k-1}},\\
&\text{For } \boldsymbol{H}^k: \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{\rho}}, \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{u}^k}, \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{u}^{k-1}}, \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{v}^k}, \frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{v}^{k-1}}.
\end{aligned} \tag{42}$$

In this study, the solution variables ($\boldsymbol{u}$) are chosen as the displacement field, whereas the auxiliary variables ($\boldsymbol{v}$) are chosen to be $\begin{bmatrix}\Delta\gamma & \alpha & f & \boldsymbol{b}^e\end{bmatrix}^T$ at each quadrature point. The corresponding constraints are the global equilibrium equations for $\boldsymbol{u}$ and local constitutive equations for $\boldsymbol{v}$, which are formulated as

$$\boldsymbol{R}^k = \underset{e=1}{\overset{n_{ele}}{\mathcal{A}}} \boldsymbol{R}_e^k = \boldsymbol{0} \text{ with } \boldsymbol{R}_e^k = \sum_{s=1}^{n_{ipt}} \left[\eta \boldsymbol{B}_{e_s}^T \boldsymbol{P}_{e_s}^k + \left(1-\eta^2\right) \boldsymbol{B}_{L,e_s}^T \left[\mathbb{C} : \boldsymbol{\varepsilon}_{e_s}^k\right]\right] w_{e_s} - \boldsymbol{F}_{ext}^e, \tag{43}$$

$$\boldsymbol{H}^k = \begin{bmatrix} \boldsymbol{H}_1^k \\ \vdots \\ \boldsymbol{H}_e^k \\ \vdots \\ \boldsymbol{H}_{n_{ele}}^k \end{bmatrix} = \boldsymbol{0} \text{ with } \boldsymbol{H}_e^k = \begin{bmatrix} \boldsymbol{H}_{e_1}^k \\ \boldsymbol{H}_{e_2}^k \\ \boldsymbol{H}_{e_3}^k \\ \boldsymbol{H}_{e_4}^k \end{bmatrix} \text{ and}$$

for elastic step (i.e., $\Delta\gamma = 0$ at sth quadrature point)

$$\boldsymbol{H}_{e_s}^k = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} \Delta\gamma \\ \alpha - \alpha_{k-1} \\ f - f_{k-1} \\ \boldsymbol{b}^e - \boldsymbol{b}^{e^{tr}} \end{bmatrix},$$

while for plastic step (i.e., $\Delta\gamma > 0$ at sth quadrature point)

$$\boldsymbol{H}_{e_s}^k = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} \tau_d - \sqrt{\frac{2}{3}}\, sign(z)\sqrt{|z|}\, \zeta J \\ \alpha - \alpha_{k-1} - \frac{\Delta\gamma}{(1-f)\zeta J}\left(\tau_d + C_0\tau_m\right) \\ f - f_{k-1} - \rho^{p_f}\Delta\gamma\left((1-f)C_0 + \cancel{\kappa}\left(\tau_m\right)\frac{\mathcal{D}}{(1-f)\zeta J}\left(\tau_d + C_0\tau_m\right) + \sqrt{\frac{2}{3}}k_\omega f\omega_0\right) \\ exp(2\Delta\gamma \boldsymbol{A}).\boldsymbol{b}^e - \boldsymbol{b}^{e^{tr}} \end{bmatrix}, \tag{44}$$

where the subscript "e_s" is used to denote the term evaluated at the sth integration point in eth element. It should be noted that the subscript "e" in $\boldsymbol{R}_e^k$ and $\boldsymbol{H}_e^k$ denotes the element number, whereas the superscript "e" in $\boldsymbol{b}^e$ denotes the *elastic* Finger tensor. The element external force $\boldsymbol{F}_{ext}^e$ is assumed to be configuration independent. The definition of C_0 is given in Appendix A. All the required derivatives are given in Appendix B.

6 | NUMERICAL EXAMPLES: MODEL PERFORMANCE

To verify the implementation of the shear modified GTN model given in Appendix A as well as to show the performance of the model as compared to other models such as the original Gurson model (without void nucleation, shearing terms, and growth acceleration due to voids coalescence), GTN model (without shear modification) and J_2 plasticity model, some of the benchmark problems including 2D plane strain localization, necking of 3D cylindrical bar and other high and low-stress triaxiality controlled fractures are examined in this section. In the nonlinear finite element analysis, the

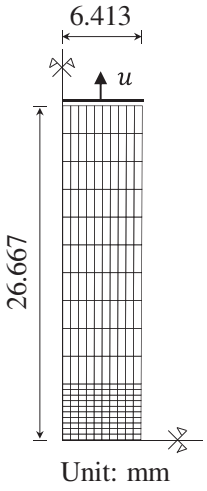


FIGURE 2 Sketch of the plane strain localization problem

TABLE 1 Model parameters for the plane strain localization problem

Model	Parameters	
J_2 plasticity	$f_0 = q_1 = q_3 = f_N = k_\omega = 0$	$E = 206.9$ GPa, $\nu = 0.29$, $\sigma_{y0} = 0.45$ GPa, $k_1 = -0.012924$ GPa,
Gurson	$f_0 = 0.01, f_N = k_\omega = 0, q_1 = q_3 = 1$	$k_2 = 0.265$ GPa, $\delta = 16.93$
Shear modified GTN	$f_0 = 0.01$	

Newton–Raphson scheme with an adaptive step-size strategy is used, and convergence is assumed when the energy residual drops below 1.0E-12.[54] All the numerical computations are carried out in a matlab-based in-house finite element library CPSSL-FEA developed at the University of Notre Dame.

6.1 | Plane strain localization

The 2D plane strain localization problem has been examined in the previous studies.[44,48,55] As shown in Figure 2, a quarter of the specimen with a width of 12.826 mm and height of 53.334 mm is discretized by a 10×20 finite element (FE) mesh with 4-node F-bar element and is subjected to prescribed displacements ($u = 5$ mm) at the top surface. The strain localization is triggered by a width reduction of 1.8% at the center of the bar. Here, three constitutive models are considered—shear modified GTN, original Gurson, and J_2 plasticity models—by choosing a different set of model parameters, see Table 1, where all the unspecified parameters are chosen from Table 3. In addition, $f^* \equiv f$ is considered in the original Gurson model. The load–displacement curves for the three models are shown in Figure 3, while the strain localizations can be seen in Figure 4. It can be seen from Figure 3 that due to the added nucleation and shear modified terms, the damage in the shear modified GTN model develops much faster than that in the original Gurson model, while both introduce more softening effects compared to the J_2 plasticity model. In addition, the final displacement of $u = 5$ cannot be reached with the GTN model and the specimen fails around the displacement of $u = 3.63$. This example also verifies the model and plane strain F-bar element implementation.

6.2 | Necking of a cylindrical bar

To demonstrate the model performance in a 3D case, the necking of a cylindrical bar under tension[48,55] is examined. The cylindrical bar is of radius 6.413 mm and length 53.334 mm. Due to symmetry, only the symmetric octant section, as shown in Figure 5, is modeled and discretized by 860 8-node 3D F-bar elements. A geometric imperfection of 1.8% of the

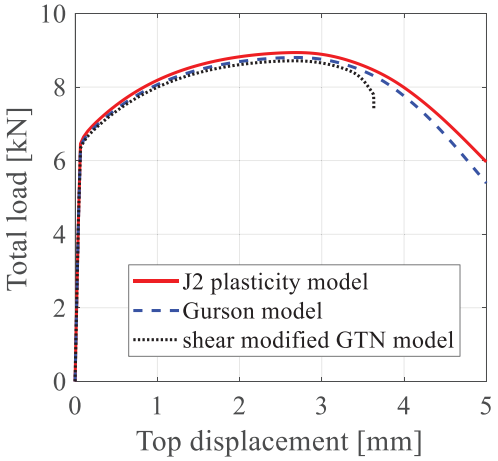


FIGURE 3 Load–displacement curves of the plane strain localization problem with three models

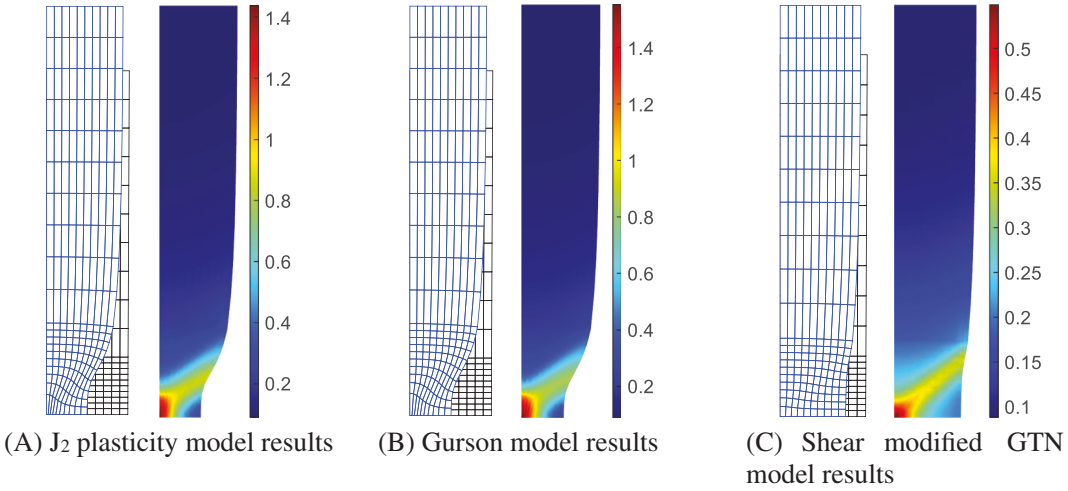


FIGURE 4 Deformed meshes and equivalent plastic strain (α) distributions of the plane strain bar at the final loading point from three models. (A) J_2 plasticity model results, (B) Gurson model results, (C) shear modified GTN model results

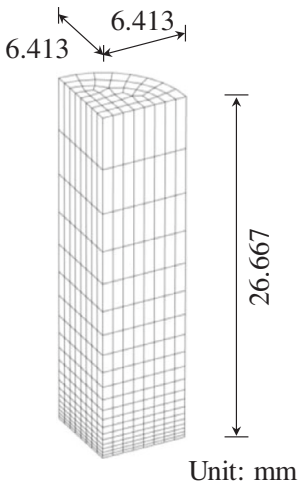


FIGURE 5 FE mesh of the symmetric octant part of 3D cylindrical bar

TABLE 2 Model parameters for the 3D cylindrical bar necking problem

Model	Parameters	
J_2 plasticity	$f_0 = q_1 = q_3 = f_N = k_\omega = 0$	$E = 206.9$ GPa, $\nu = 0.29$, $\sigma_{y0} = 0.45$ GPa, $k_1 = 0.12924$ GPa,
Shear modified GTN	$f_0 = 0, f_C = 0.05, f_F = 0.3, f_U^* = 0.667, f_N = 0.01, s_N = 0.1, \alpha_N = 0.35, k_\omega = 3, T_1 = 0, T_2 = 0.5, M_c = 15, q_3 = q_1^2 = 2.25, q_2 = 1$	$k_2 = 0.265$ GPa, $\delta = 16.93$

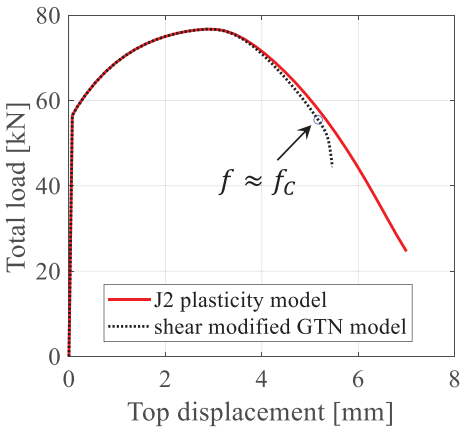


FIGURE 6 Load–displacement curves of the cylindrical bar necking problem with two models

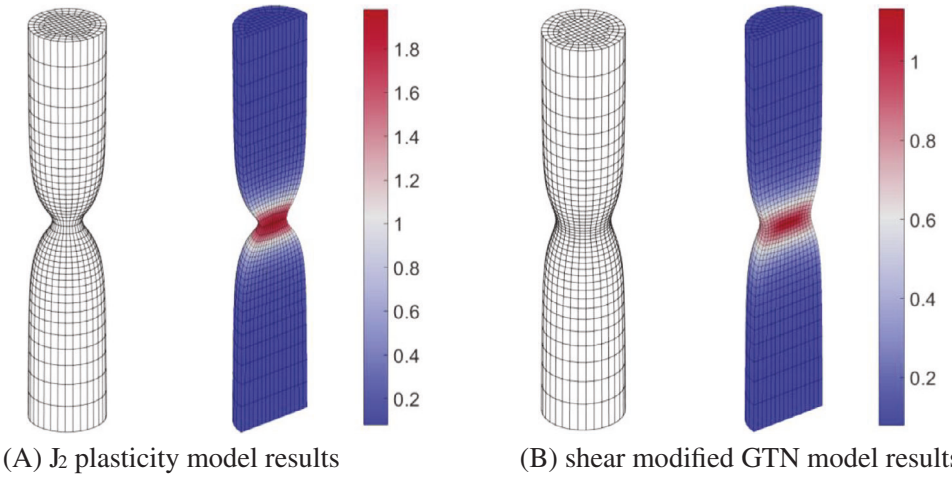
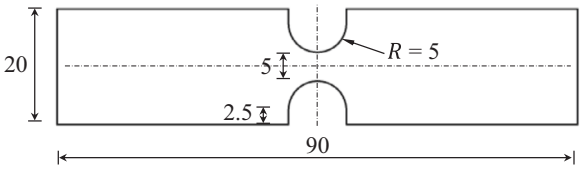
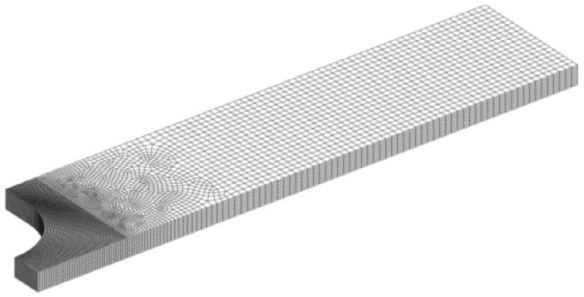


FIGURE 7 Deformed meshes and equivalent plastic strain (α) distributions of the 3D cylindrical bar at the final loading point from two models

radius is introduced at the center to trigger the necking, where the radius reduction is linear from the top to the center. The bar is under tension, with a displacement $u = 7$ mm applied at the top surface. The J_2 plasticity and shear modified GTN models are considered with the material parameter settings given in Table 2. Figure 6 shows the load–displacement curves, where a close match can be observed before the start of void coalescence after f_C. For visualization purposes, the deformation of the full specimen together with the lateral cross-sectional plastic strain (α) distribution are plotted in Figure 7 for the two models. The results further verify the model implementation and show the influences of added void growth, nucleation, and shearing terms.



(A) Specimen geometry in x-y plane (thickness in z-direction is 1.5 mm)



(B) FE mesh (quarter part) with 78255 elements

FIGURE 8 Geometry and FE mesh of the notched plate

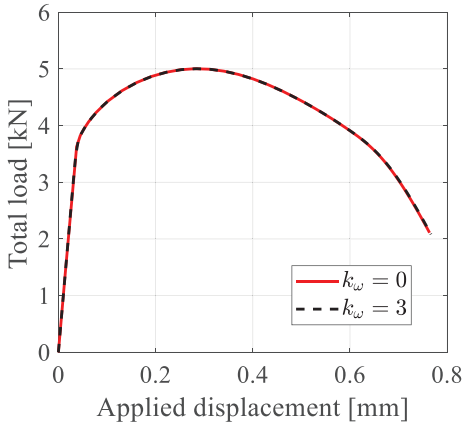


FIGURE 9 Load–displacement curves of the notched plate problem with GTN ($k_\omega = 0$) model and shear modified GTN model with $k_\omega = 3$

6.3 | Notched plate

To show stress triaxiality controlled fracture induced by nucleation and void growth, a notched plate shown in Figure 8A with one end fixed and the other end under displacement control (tension) is considered.[56] The geometry of the specimen as well as the FE mesh of its symmetric quarter section is shown in Figure 8. The material parameters are the same as those for the shear modified GTN model in Table 2, except for k_ω, which is chosen to be either 0 or 3 in this example. It is noted that the original GTN model without shear modification is recovered when $k_\omega = 0$. The load–displacement curves of the two cases ($k_\omega = 0$ and $k_\omega = 3$) are plotted in Figure 9, where no significant differences can be observed. This implies that the shear modification term (f_s) in Equation (19) has a negligible effect on the fracture in this example. This is further confirmed by the similar void volume fraction (f) distributions in the two cases in Figure 10. This behavior can be explained using Equations (22) and (24), where $\dot{f}_s$ is vanishing when the triaxiality is high, that is, $T > T_2$ in the fracture region, see Figure 11 where $T > 1.5$ at the fracture tip. Figure 12 shows the contribution of each fracture mechanism—void nucleation (f_n), growth (f_g) and shearing (f_s)—in the elements of the three layers from the center of the plate, where negligible f_s in Figure 12B can be observed and the final fracture is largely attributed to the void growth, as expected.

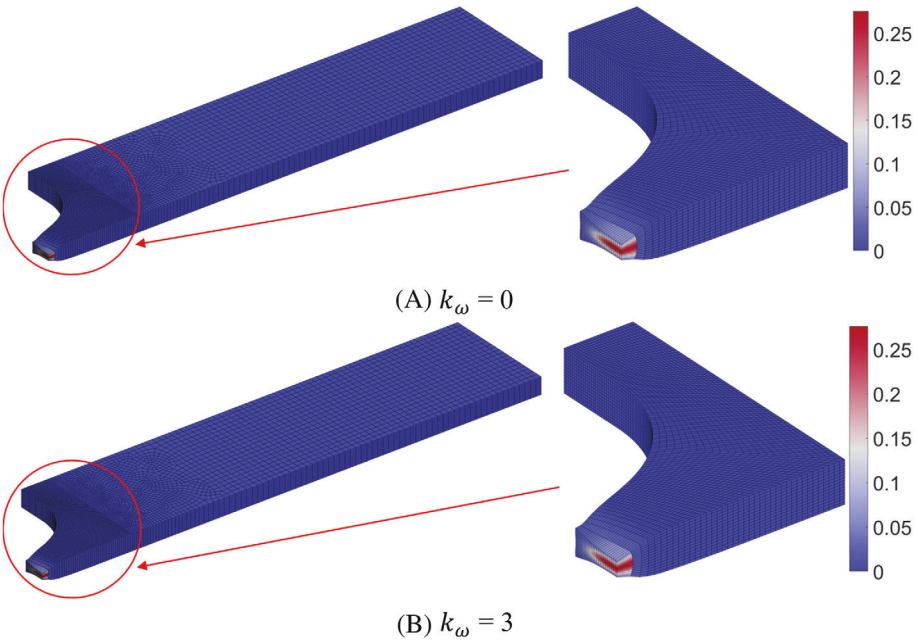


FIGURE 10 Void volume fraction (f) distributions on the deformed meshes of the notched plate at final converged loading points with different k_ω values

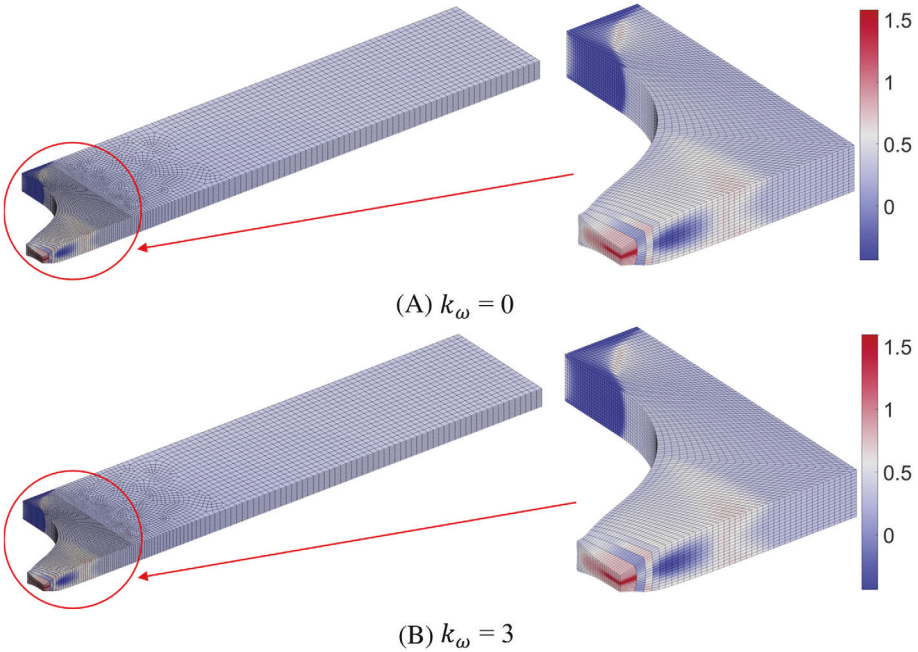


FIGURE 11 Stress triaxiality (T) distributions on the deformed meshes of the notched plate at final converged loading points with different k_ω values

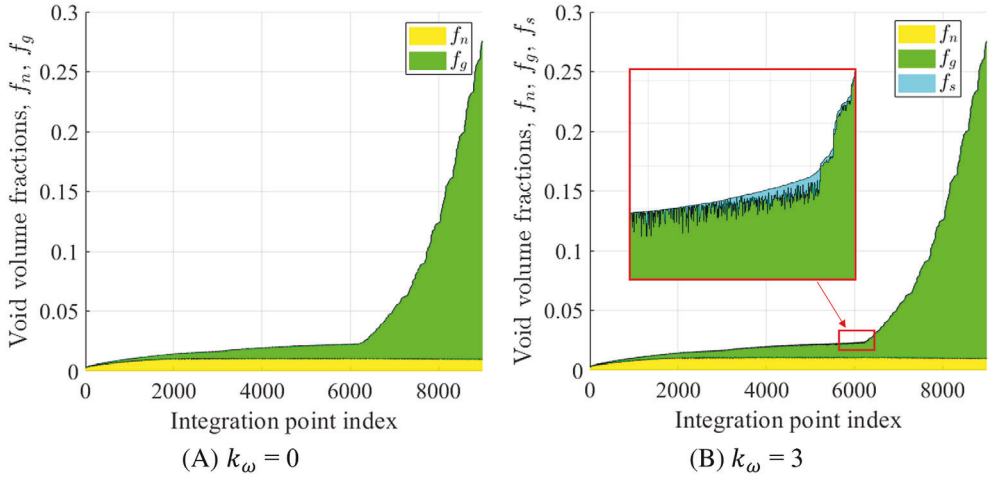
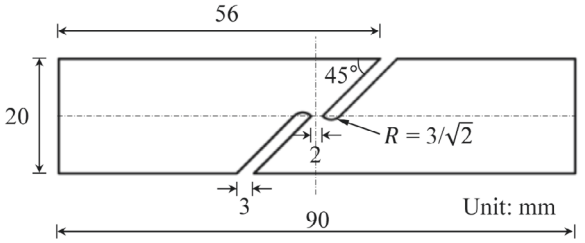
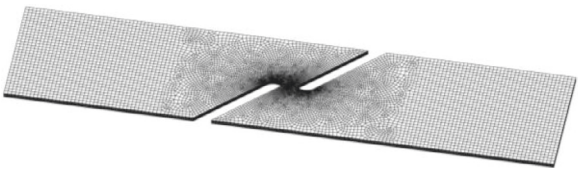


FIGURE 12 Different contributions (f_n, f_g, and f_s) to the void volume fraction f at each integration point in the critical fracture region (three layers of elements from the center of the plate) of the notched plate at final converged loading points with different k_ω values ()*Note*: the integration point index is numbered by the order of the total f value



(A) Specimen geometry in x-y plane (thickness in z-direction is 1.5 mm)



(B) FE mesh (half thickness) with 66920 elements

FIGURE 13 Geometry and FE mesh of the shear plate

6.4 | Fracture under shear

This example intends to show how the fracture is controlled by void shearing mechanism when the stress triaxiality is low. To this end, a notched plate of thickness 15 mm shown in Figure 13A is considered. The plate is fixed at one end and under uniform displacement (tension) at the other end. Due to the out-of-plane symmetry condition, only half-thickness of the specimen is considered in the FE model, see Figure 13B. The specimen of similar geometry has been experimentally investigated in Reference 56. The material parameters are again the same as the shear modified GTN row in Table 2, except k_ω which is varied, that is, $k_\omega \in \{0, 1, 2 \text{ and } 3\}$, to study the effect of the shear modification term in the model. The load–displacement curves obtained for these cases are plotted in Figure 14, where it can be seen that the mechanical response highly depends on the shear modification term—that is, higher k_ω value results in an earlier fracture. It is noted that the case with $k_\omega = 0$ does not reach the final failure up to the applied displacement, which is 2.5 mm (see Figure 14). To show the difference brought by the shear modification in terms of the parameter k_ω, Figure 15 and Figure 16 plot

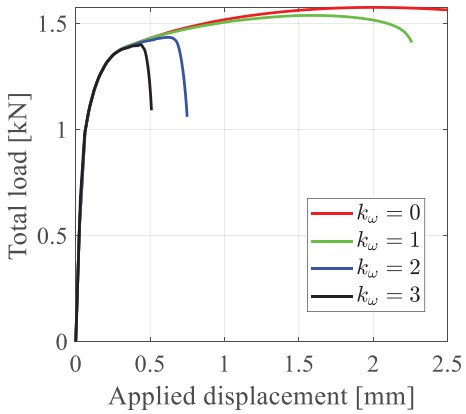


FIGURE 14 Load–displacement curves of the shear plate problem with different k_ω values

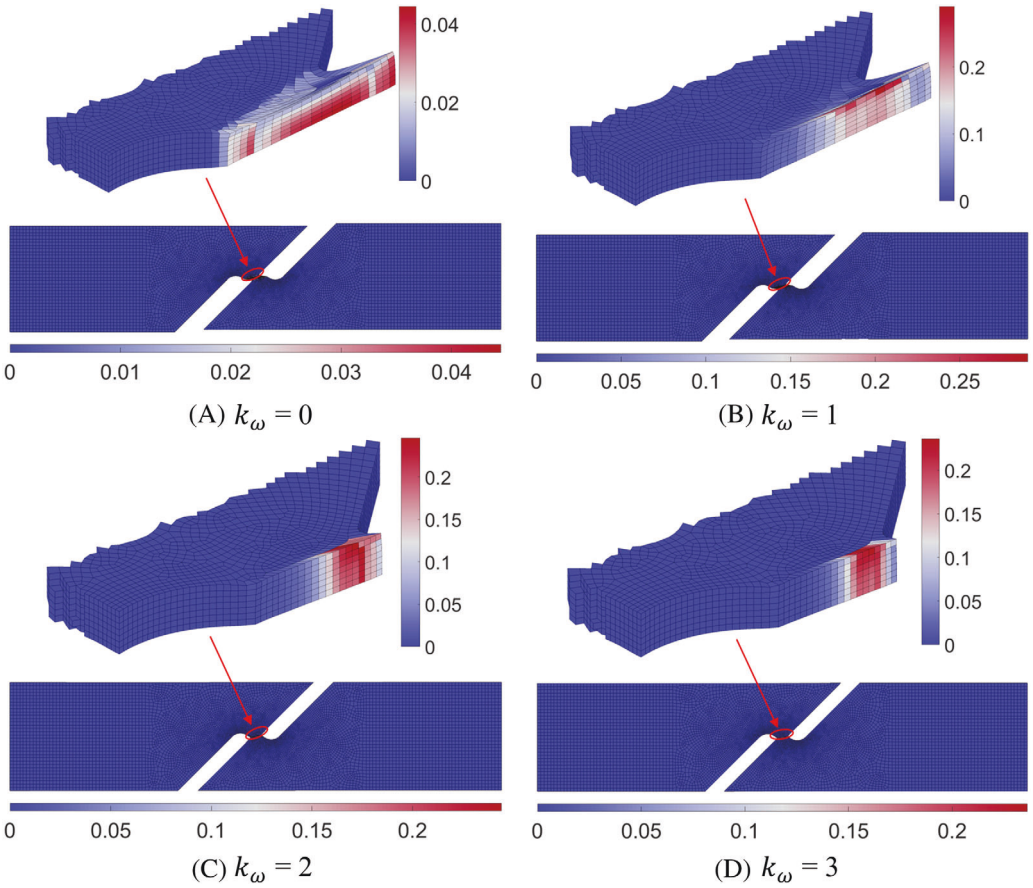


FIGURE 15 Void volume fraction (f) distributions on the deformed meshes of the shear plate at final converged loading points with different k_ω values

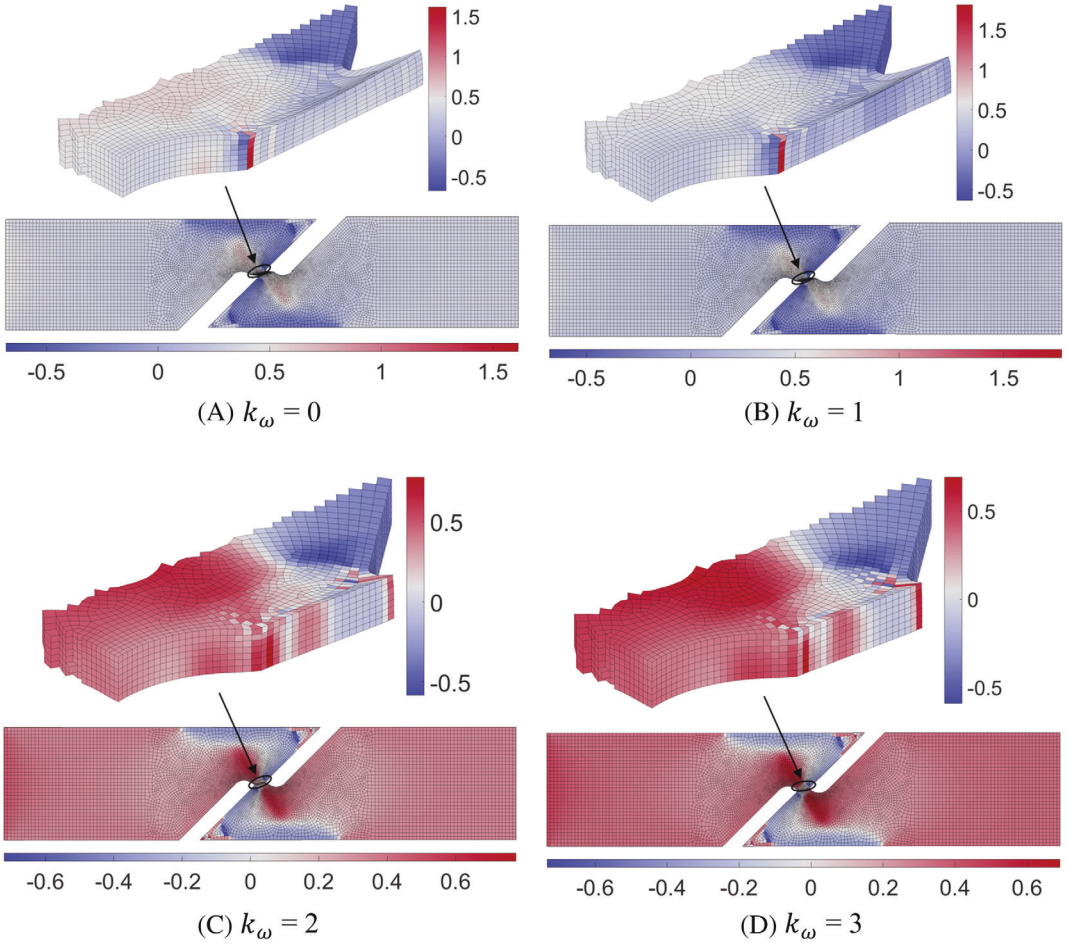


FIGURE 16 Stress triaxiality (T) distributions on the deformed meshes of the shear plate at final converged loading points with different k_ω values

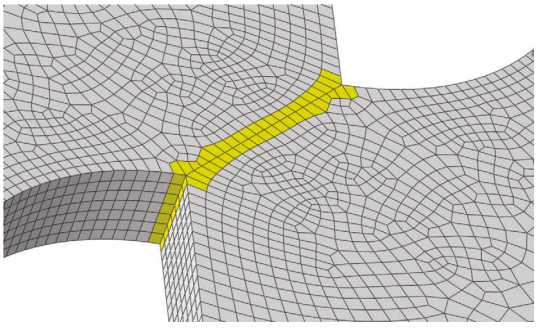


FIGURE 17 Critical fracture/shear region with highlighted elements

the void volume fraction (f) and stress triaxiality (T) distributions, respectively, in the critical region. This shows that in the neighborhood of the highest fracture (f) value in Figure 15, for $k_\omega = 0$ case, $T \leq 0.5$ while for $k_\omega = 1$, 2, and 3 cases, $T \leq 0$ and continuously decreases as k_ω increases (see Figure 16). As a result, from Equations (22)–(24), the contribution from void shearing can be high and this is further verified by plotting the proportions of f_n, f_g and f_s in the critical region (highlighted in Figure 17) in Figure 18. It can be observed from Figures 18B,C that the void growth f_g can be negative due to negative pressure (compression). In addition, for $k_\omega = 1$, 2, and 3, the final fracture is largely attributed to void shearing. Hence, the shear modification is of critical importance in capturing the shear fracture mechanism in this example.

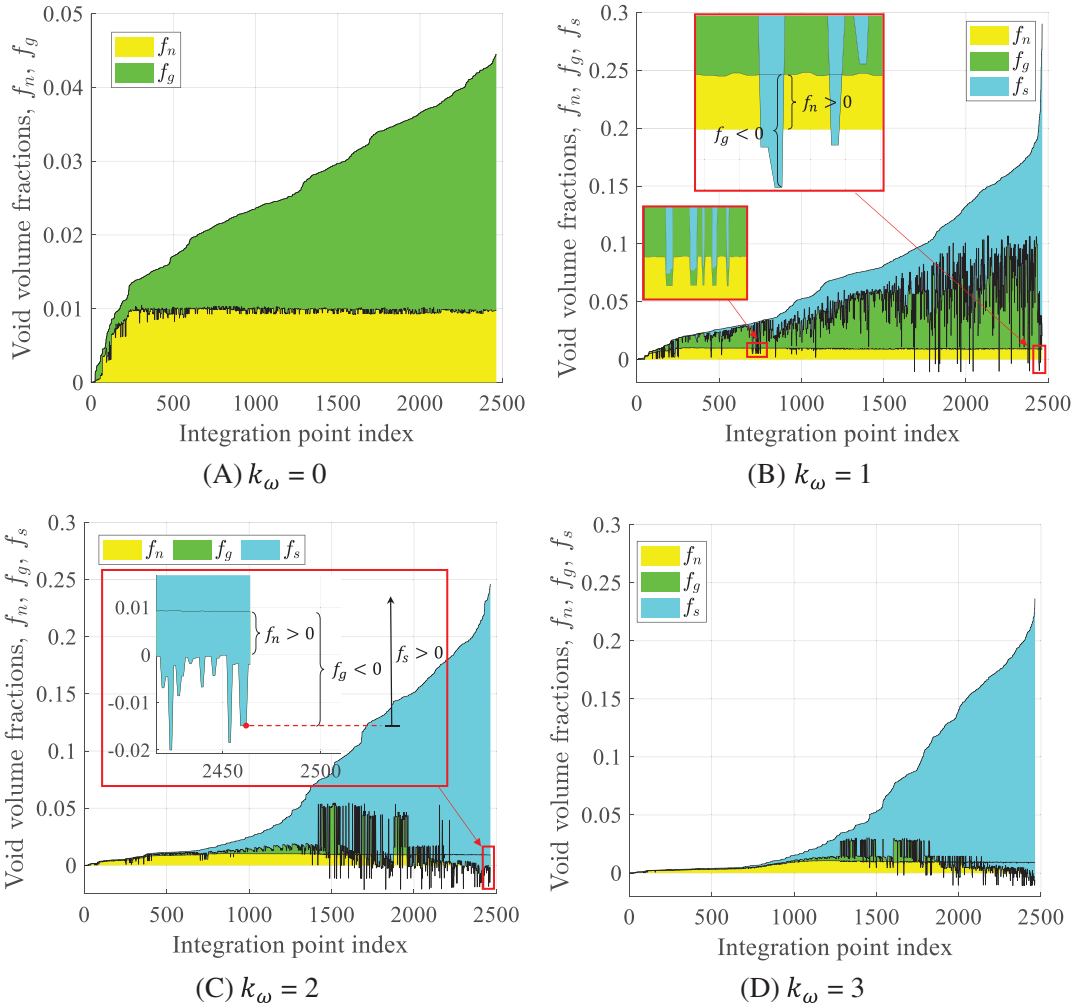


FIGURE 18 Different contributions (f_n, f_g, and f_s) to the void volume fraction f at each integration point in the critical fracture region of the shear plate at final converged loading points with different k_ω values (*Note*: the integration point index is numbered by the order of the total f value)

7 | TOPOLOGY OPTIMIZATION RESULTS

In this section, topology optimization is carried out using the shear-modified GTN model. In the following optimization problems, the shear modified GTN model is used with the parameters given in Table 3, which are adapted from Kiran and Khandelwal[57] for ASTM A992 steels considering mesh dependency, and the mesh size in all the numerical examples is chosen to be around 5 mm. The density filter radius is chosen as $r_{min} = 12$ mm. As displacement control is considered in the optimization examples, padding blocks are added at the loading area to avoid strain localization. The padding blocks are considered to be hyperelastic by choosing the material parameters as $E = 4$ GPa, $\sigma_{y0} = 1.0\text{E+}9$ together with those in Equation (27) and Table 3. Since the damage constraint value $\overline{D}$ in this study is chosen to be below the coalescence threshold f_c, f^* is replaced by f (Equation (15)) in FE analysis during the optimization process. This helps to improve FE convergence, as fast void evolution after f_c can happen for some intermediate designs. However, the f^* function in Equation (15) is used for all the postperformance analyses after optimization. The method of moving asymptotes[58] is used as the optimization algorithm with default parameter settings. Initial designs consist of a homogeneous distribution of density values, that is, $\rho = V_f$, where V_f is the threshold value in the volume fraction constraint in Equation (31). Unless otherwise stated, the optimization examples are terminated after 800 iterations, where clean topology has emerged. Finally, in accordance with the GTN model parameters in Table 3, the element size in all the B-spline fitted designs is controlled to be around 5 mm. The B-spline fitting of the optimized design after optimization is carried out using Rhino[59] with a level-set value of 0.5.

TABLE 3 Shear modified GTN model parameters for topology optimization studies

Parameter	E	ν	σ_{y0}	k_1	k_2	δ	f_N	α_N	s_N	f_0
	200 GPa	0.3	376 MPa	460 MPa	120 MPa	22.6	0.02	0.35	0.05	0
	f_C	f_F	f_U^*	q_1	q_2	q_3	T_1	T_2	M_c	k_ω
	0.08	0.1	0.6667	1.5	1.0	2.25	0	0.5	15	3

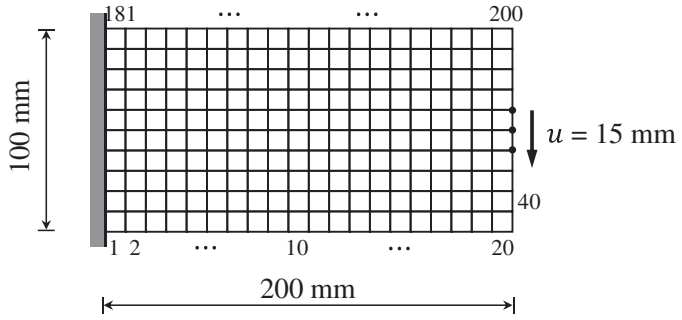
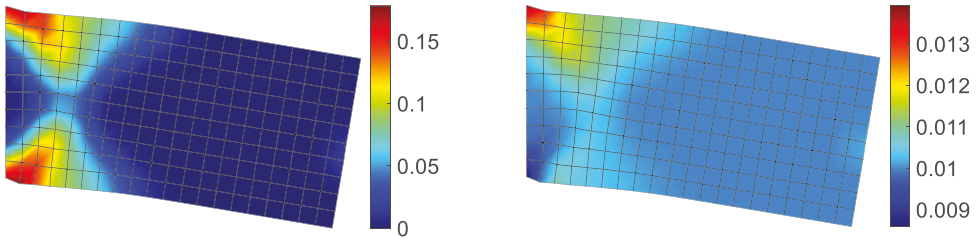


FIGURE 19 Cantilever with homogeneous density field $\rho = 0.8$ for sensitivity verification (20 ×10 FE mesh with numbering)



(A) Equivalent plastic strain (α) distribution (B) Damage (f) distribution

FIGURE 20 Results of cantilever with homogeneous density field $\rho = 0.8$

7.1 | Sensitivity verification

The adjoint sensitivity calculation proposed in Section 5.1 and Appendix B is verified using the central difference method on a cantilever problem sketched in Figure 19. The cantilever is discretized by a 20 × 10 FE mesh with a homogeneous density field $\rho = 0.8$. The material parameters for the shear modified GTN model are chosen from Table 3 except the initial void volume fraction, which is set as $f_0 = 0.01$. The optimization parameters in Equation (31) are set as $\overline{D} = 0.02$ and $V_f = 0.8$. The equivalent plastic strain (α) and damage (f) distributions are shown in Figure 20. In the central difference method, a perturbation of $\Delta\rho = 10^{-4}$ is used for each element's density. The sensitivities, that is, $df_0/d\rho$ and $df_1/d\rho$, using the two methods are compared in Figure 21 where close matches can be observed with relative errors around $10^{-5} \sim 10^{-10}$. Thus, the adjoint sensitivity calculation is correct and can be used in the gradient-based optimization algorithms in the topology optimization problems.

7.2 | Cantilever problem

The cantilever problem sketched in Figure 22 is considered with the design domain discretized by a FE mesh 120 × 60. The volume constraint in Equation (31) is considered with $V_f = 0.65$. Here, two cases listed in Table 4 are considered, where optimization with J_2 plasticity in case-1 is realized by choosing the parameters in Equation (27) and removing the constraint f_1 in Equation (31). The optimized topologies are shown in Figure 23, and the corresponding optimization

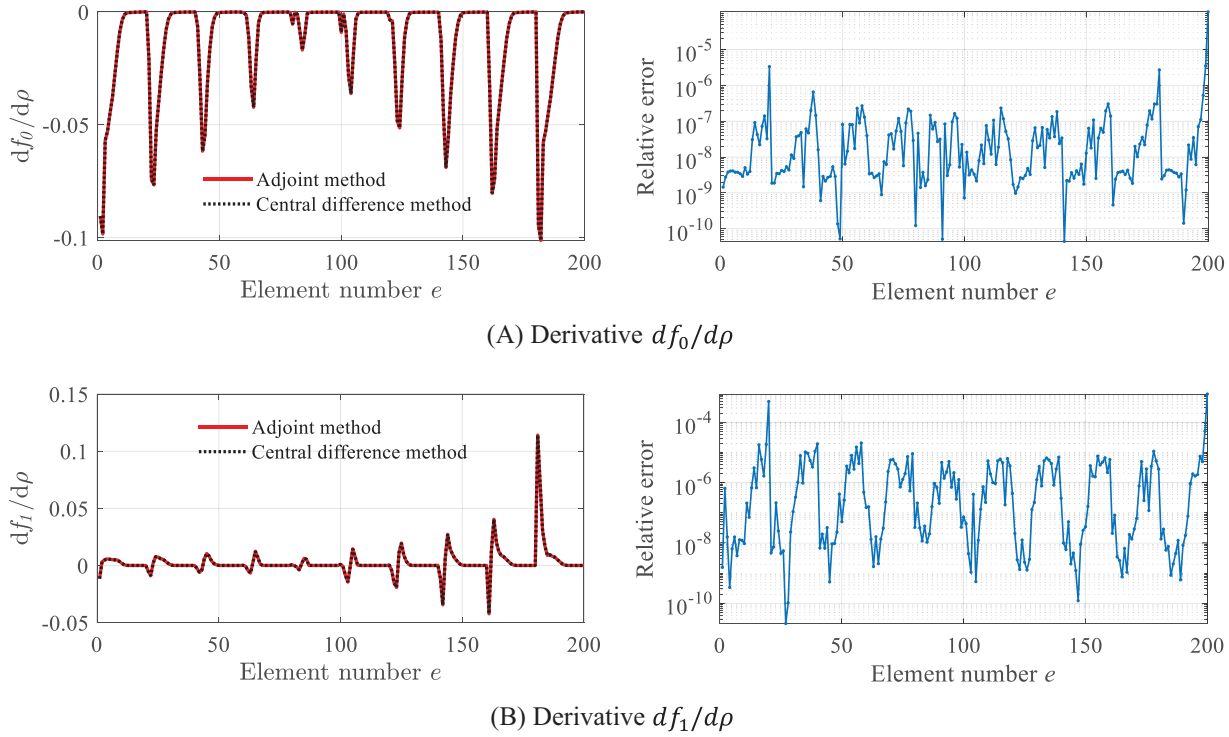


FIGURE 21 Comparison of sensitivity analysis results from adjoint method and central difference method. Left: match between the two methods; Right: relative error of adjoint method w.r.t. central difference method

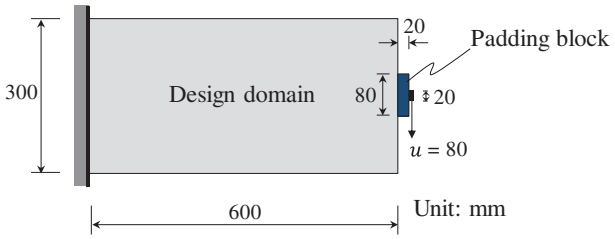
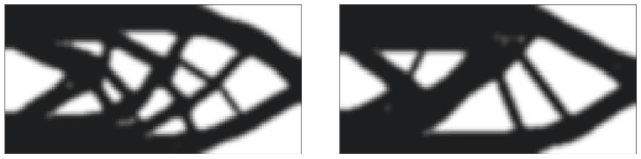


FIGURE 22 Sketch of the cantilever problem

TABLE 4 Two optimization cases

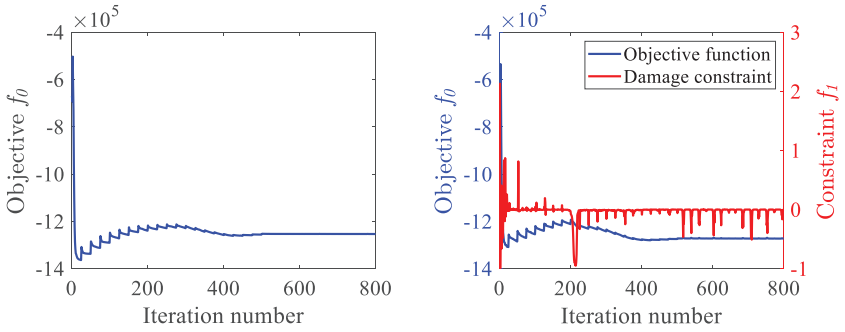
Case-1	J_2 plasticity design with no damage constraint
Case-2	GTN design with $\overline{D} = 2 \times 10^{-3}$

histories of objective and damage constraint function values are plotted in Figure 24 where smooth convergence can be observed. The two optimized designs are then both analyzed with the same shear modified GTN model and the results are shown in Figures 25 and 26. As is shown in Figure 25, the overall load–displacement curves of the two designs do not differ much up to the target displacement (80 mm). However, a further check on the energy dissipation and void volume fraction distributions in Figure 26 reveals that without modeling and constraining fracture in J_2 model-based topology optimization, the resulted fracture at the target displacement in the optimized design can be high, that is, $f_{\max} = 0.0353$. On the contrary, the maximum fracture in the shear modified GTN design is constrained below the target threshold value, which is 2×10^{-3}. As a result, the plastic work is less concentrated (more spread out) in the solid components in the GTN design, as the plastic strain that leads to the evolution of void volume fraction is constrained due to the fracture constraint, see Figure 26A,C.



(A) J_2 plasticity design (B) GTN design with $\overline{D} = 2\times10^{-3}$

FIGURE 23 Optimized topologies for the two cases in Table 4



(A) J_2 plasticity design (B) GTN design

FIGURE 24 Optimization histories of objective and damage constraint function values for the two cases in Table 4

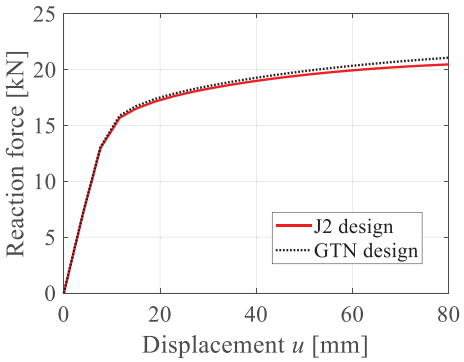
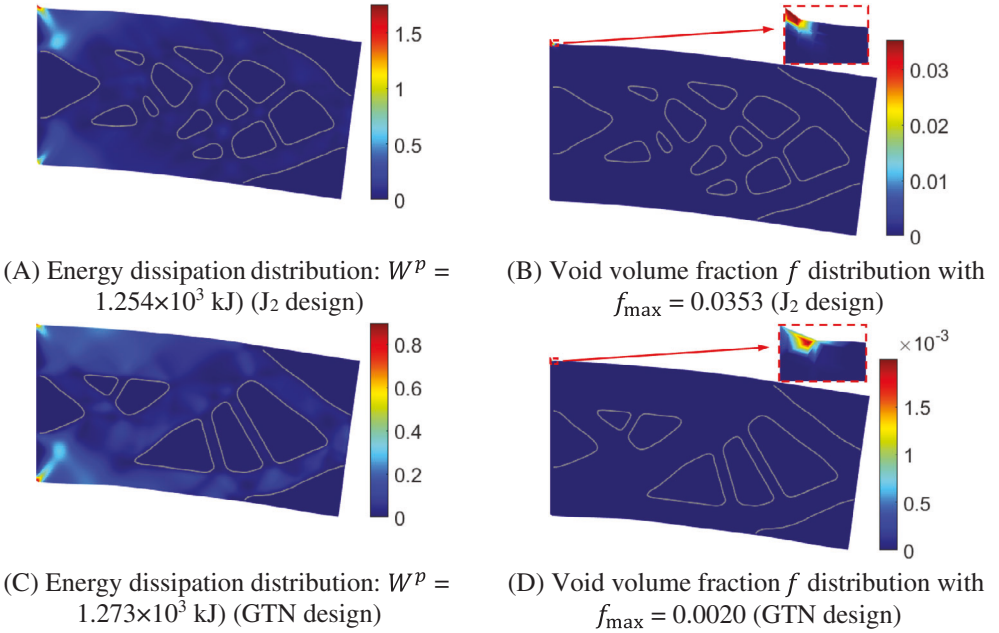


FIGURE 25 Load–displacement curves from analyses of the two designs in Figure 23 using shear modified GTN model

To study the ultimate performances including the final fracture mechanisms of the two optimized designs, the two designs in Figure 23 are fitted using B-splines to get rid of the void and intermediate densities, and the corresponding FE models are shown in Figure 27. The two fitted designs are then loaded and analyzed under the displacement control until complete fracture, that is, $f_{\max} \approx f_F$, is reached. The load–displacement curves and some performance measures are shown in Figure 28, where the ductility factor is defined as the ratio between the ultimate displacement and the yield displacement, that is, u_u/u_y.[60] For example, the yield displacement is $u_y = 9.5\,\text{mm}$ for both designs, while the ultimate displacement $u_u = 106.2\,\text{mm}$ for J_2 design and $u_u = 199.3\,\text{mm}$ for GTN design, see Figure 28. As can be seen, compared to the J_2 design, the GTN design has higher ductility, load-carrying capacity, and energy-absorbing capacity. Figures 29 and 30 show the distributions of energy dissipation, plastic strain (α), and void volume fraction and its different components at the failure point of the two designs, respectively. Again, plastic strain and plastic work distributions are more spread out in the GTN design in Figure 30, as compared to the J_2 design in Figure 29. The fracture patterns in the two designs are similar, both with damage initiated at two locations—the top and bottom sides of the tensile bar close to the support area. The final fracture happens at the top corner and is dominated by the nucleation and growth of voids, while the contribution of void volume fraction due to shear is negligible.



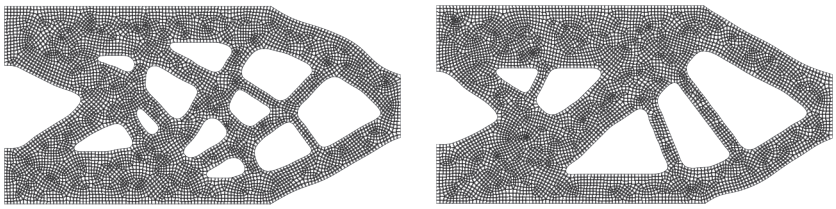
(A) Energy dissipation distribution: W^p = 1.254×10^3 kJ) (J$_2$ design)

(B) Void volume fraction f distribution with $f_{\max}$ = 0.0353 (J$_2$ design)

(C) Energy dissipation distribution: W^p = 1.273×10^3 kJ) (GTN design)

(D) Void volume fraction f distribution with $f_{\max}$ = 0.0020 (GTN design)

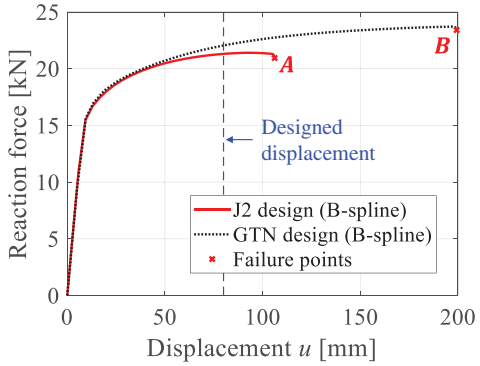
FIGURE 26 Results of analyses with GTN model at the target displacement (u = 80 mm) for the two designs in Figure 23



(A) J$_2$ plasticity design

(B) GTN design

FIGURE 27 B-spline fitted topologies for the two cases in Table 4



Designs	J$_2$ design	GTN design
Ductility factor	11.18	20.98
Load carrying capacity [kN]	21.40	23.72
Energy dissipation W^p [kJ]	1.916×10^3	4.108×10^3

FIGURE 28 Ultimate performances of the two designs in Figure 27 using the GTN model

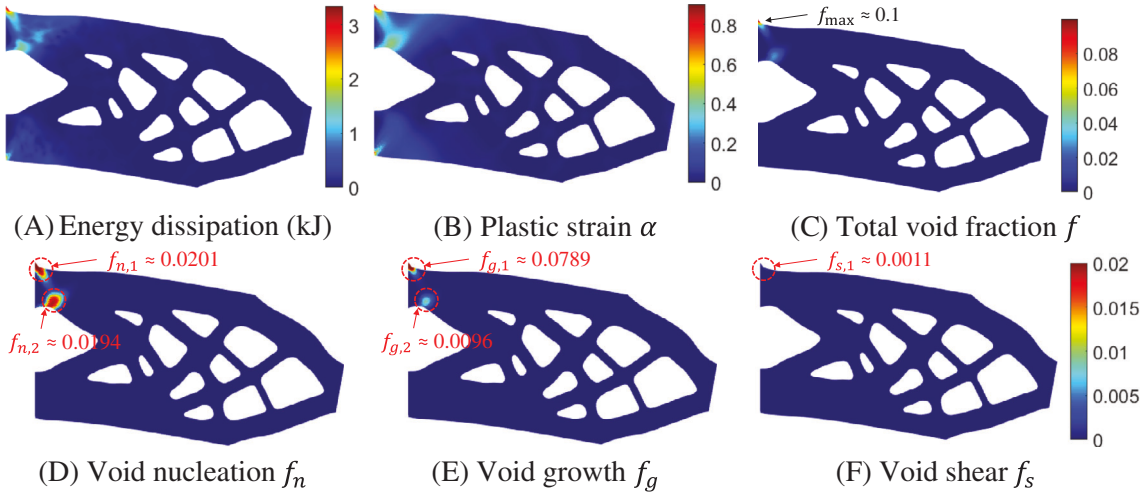


FIGURE 29 Ultimate performance of the J_2 design at the failure point $\boldsymbol{A}$ ($u = 106.2$ mm) in Figure 28 (color maps in (D) ~ (F) are adjusted to be the same shown in the color bar of (F) for a better comparison)

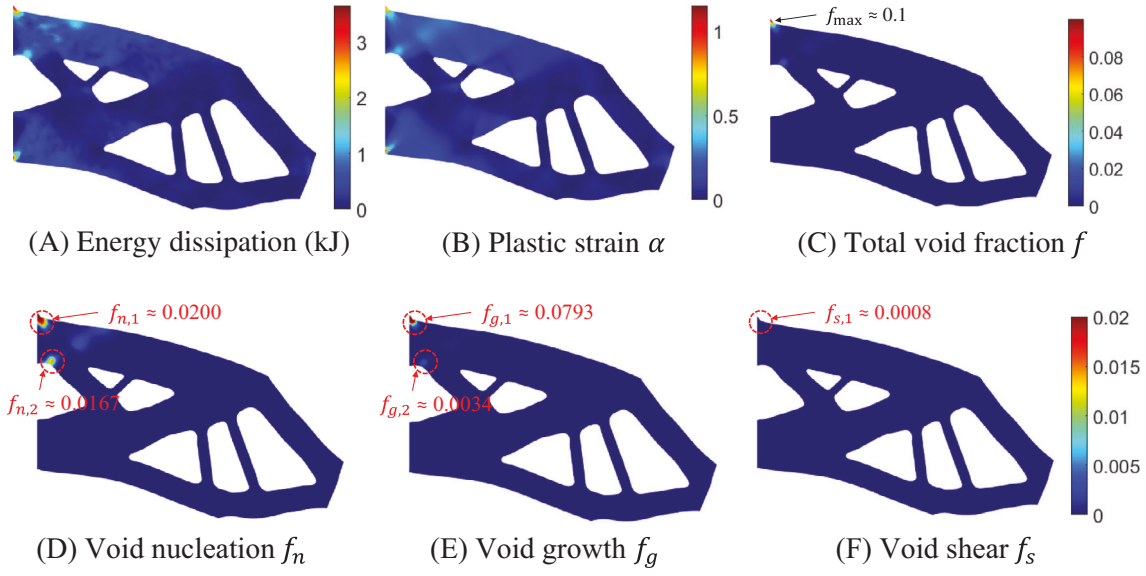


FIGURE 30 Ultimate performance of the GTN design at the failure point $\boldsymbol{B}$ ($u = 199.3$ mm) in Figure 28 (color maps in (D) ~ (F) are adjusted to be the same shown in the color bar of (F) for a better comparison)

7.3 | L-shape problem

The second example considers an L-shape problem shown in Figure 31, where a pre-existing column with the top side fixed at the boundary is not designed and is incorporated to enhance overall stability. The L-shape domain is discretized by 7500 F-bar quadrilateral elements. With material volume fraction constraint $V_f = 0.5$, two different damage constraints are considered, that is, $\overline{D} = 1 \times 10^{-4}$ and $\overline{D} = 5 \times 10^{-4}$, see Equation (31). Due to slow convergence, the optimization is terminated after 2000 iterations and the optimized topologies are shown in Figure 32. No distinguishable difference can be observed from the load–displacement curves in Figure 33 of the two designs, although the damage (f) in the design with $\overline{D} = 1 \times 10^{-4}$ is five times smaller than that of the design with $\overline{D} = 5 \times 10^{-4}$, see Figure 34B,D. Next, B-spline fitted designs are generated (Figure 35) to study the ultimate performance of the two designs. The load–displacement curves up to the failure point of the two fitted designs are shown in Figure 36, together with some performance measures. It can be seen, for the design with $\overline{D} = 1 \times 10^{-4}$ both the ductility and energy dissipation are higher while the

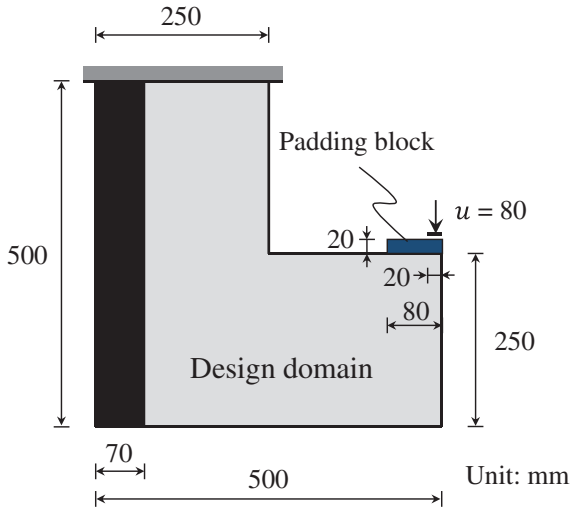
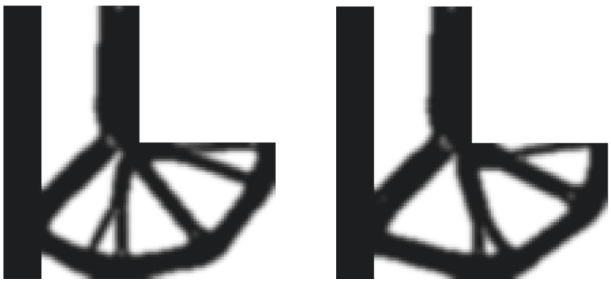


FIGURE 31 Sketch of L-shape problem



(A) GTN design with $\overline{D} = 1\times10^{-4}$ (B) GTN design with $\overline{D} = 5\times10^{-4}$

FIGURE 32 Optimized topologies for different damage constraints

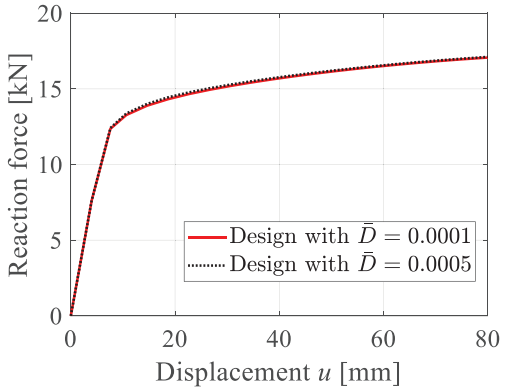
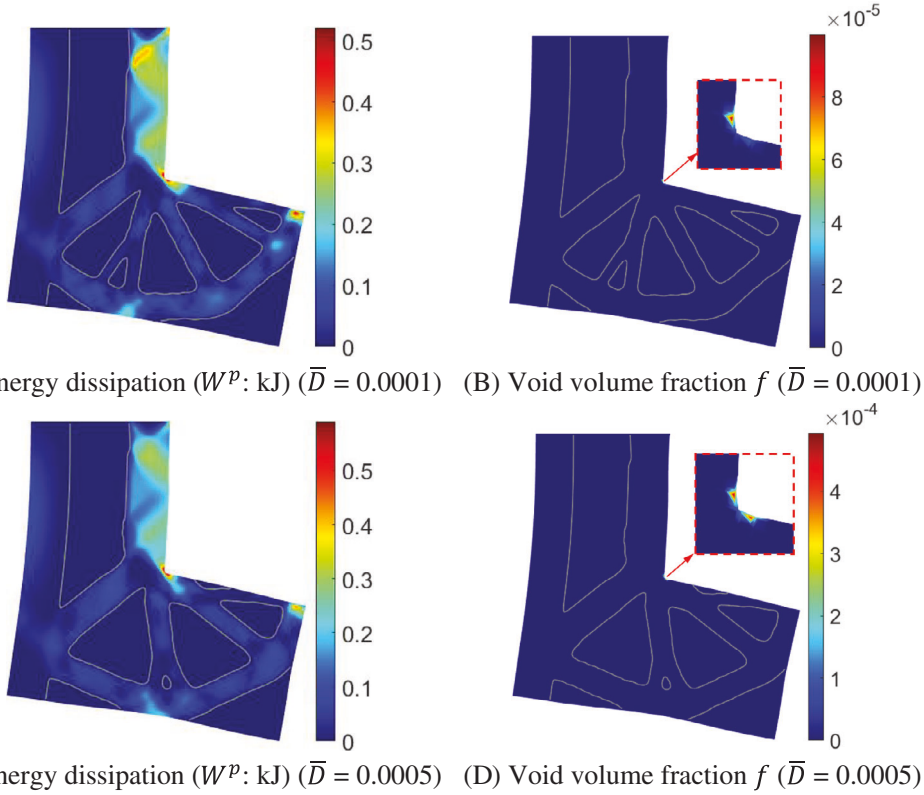


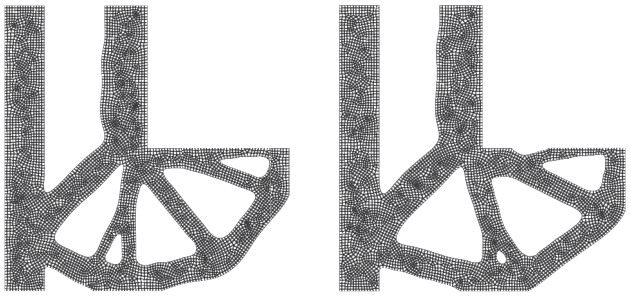
FIGURE 33 Load–displacement curves of the two designs in Figure 32



(A) Energy dissipation (W^p: kJ) ($\overline{D} = 0.0001$) (B) Void volume fraction f ($\overline{D} = 0.0001$)

(C) Energy dissipation (W^p: kJ) ($\overline{D} = 0.0005$) (D) Void volume fraction f ($\overline{D} = 0.0005$)

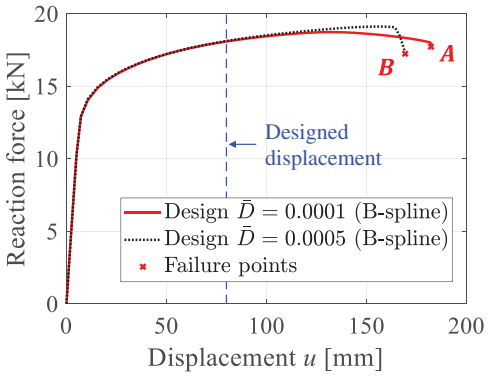
FIGURE 34 Analysis results at target displacement $u = 80$ mm for the two topologies in Figure 32



(A) Design with $\overline{D} = 1\times10^{-4}$ (B) Design with $\overline{D} = 5\times10^{-4}$

FIGURE 35 B-spline fitted designs for the two topologies in Figure 32

load carrying capacity is lower, as compared to the counterparts of the design with $\overline{D} = 5\times10^{-4}$. Next, from the damage (f) distributions of the two designs at the designed displacement (Figure 37) as well as at each ultimate failure point (Figures 38C and 39C), it is interesting to observe that the fracture locations change as the applied load increases. As a result, the two designs exhibit two different final fracture patterns, see Figures 38C and 39C. In the design with $\overline{D} = 1\times10^{-4}$, damage initiates at two locations during loading and eventually localizes to one that leads to the final failure, see Figure 38. Similarly, in the design with $\overline{D} = 5\times10^{-4}$, see Figure 39, damage initiates at four locations at different loading stages, before the final fracture localization. In both designs, void nucleation and growth still play the main role in the final failure, while shear-related damage is negligible. It is remarked that the performances of the design before and after B-spline fitting can be different, see Figures 34 and 37, where the damage is about five times higher after B-spline fitting. This is due to the effect of gray areas removal and geometric variations during B-spline fitting.



GTN designs	$\bar{D} = 1\times10^{-4}$	$\bar{D} = 5\times10^{-4}$
Ductility factor	25.14	23.37
Load carrying capacity [kN]	18.73	19.12
Energy dissipation W^p [kJ]	3.038×10^3	2.810×10^3

FIGURE 36 Ultimate performances of the two designs in Figure 35

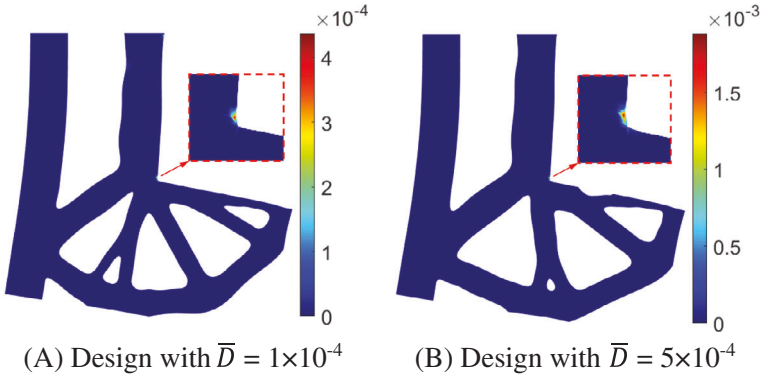


FIGURE 37 Void volume fraction (f) distributions of the two designs in Figure 35 at designed displacement ($u = 80$ mm)

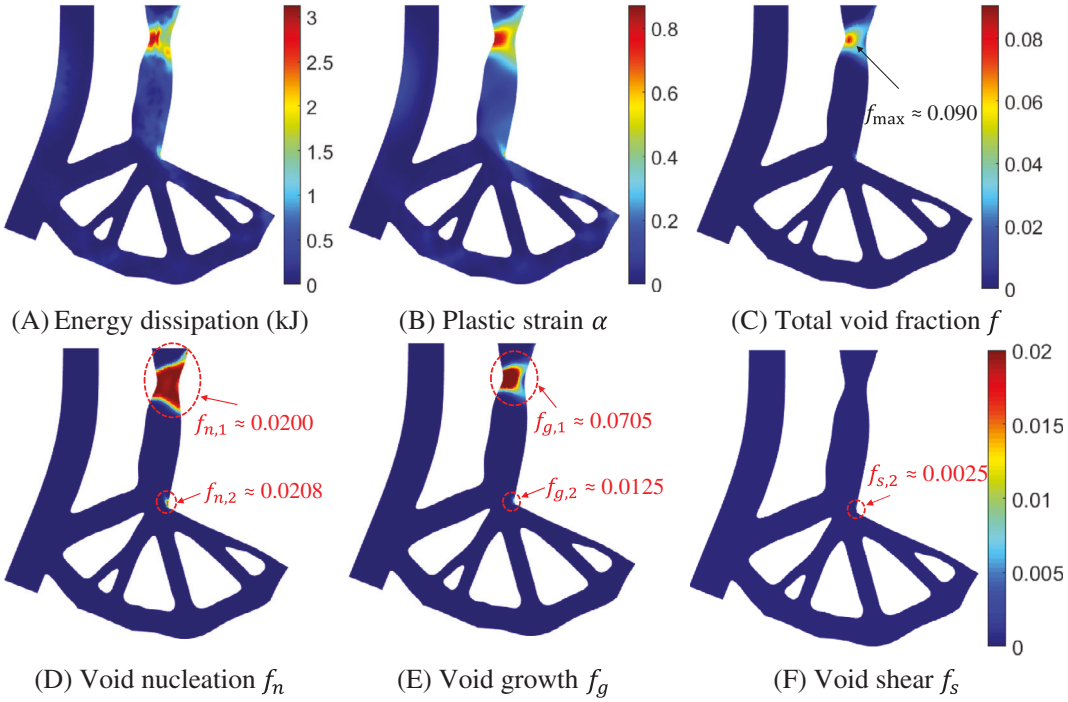


FIGURE 38 Ultimate performance of the design with $\bar{D} = 0.0001$ at the failure point $\boldsymbol{A}$ ($u = 182.3$ mm) in Figure 36 (color maps in (D) ~ (F) are adjusted to be the same shown in the color bar of (F) for a better comparison)

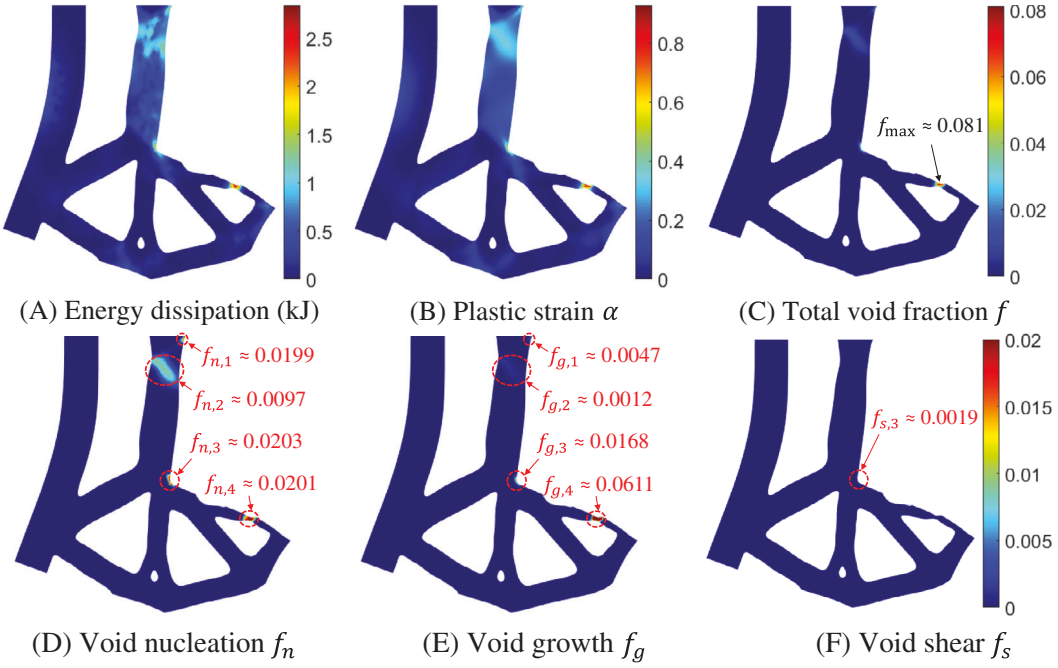


FIGURE 39 Ultimate performance of the design with $\overline{D} = 0.0005$ at the failure point $\boldsymbol{B}$ ($u = 169.4$ mm) in Figure 36 (color maps in (D) ~ (F) are adjusted to be the same shown in the color bar of (F) for a better comparison)

8 | CONCLUSIONS

This study presents a finite strain shear modified GTN model based on the multiplicative decomposition of the deformation gradient. The model implementation details are provided and verified with multiple examples. The performance of the model is demonstrated via different 2D and 3D examples, where the shear modification term, which is active only under low-stress triaxiality, is shown to be capable of capturing the damage due to shear. More importantly, this study takes the first attempt to incorporate this model in a finite strain regime into topology optimization for fracture-resistant energy-dissipating structural designs. Using a density-based method, the optimization framework is realized by appropriate material interpolation schemes for design parameterization, adaptive linear energy interpolation for low-density elements distortion under large deformations, and consistent path-dependent sensitivity calculations for the gradient-based optimization algorithm. No assumptions on the magnitude of strains are made and the ultimate performance of the optimized topologies is evaluated and confirmed by a postanalysis on the corresponding B-spline fitted designs, which can be potentially manufactured for further experimental investigations. It is remarked that the structural performance of optimized designs can be different after B-spline fitting due to the removal of gray areas and the variations of geometric boundaries. However, this difference can be further reduced by using higher mesh density together with projection Schemes.[61]

Compared to the maximum energy-absorbing designs with the J_2 plasticity model, it is shown that the GTN model assisted design with fracture constraint may lead to higher ductility, energy-absorbing capacity, and ultimate load-carrying capacity. More interestingly, it is observed that fracture can initiate at multiple locations during the large deformations and the critical fracture location can shift from one location to another during the loading process due to stress redistributions. Moreover, it is important to note that as the designs are obtained for target design displacements, the structural performance (e.g., energy dissipation and damage) is only optimized up to that design point. Hence, the optimality of the *ultimate* performance at failure is not guaranteed, in general, although promising results have been obtained for the cantilever problem in Section 7.2.

In future work, experimental work will be carried out to validate the proposed finite strain formulation of the shear modified GTN model together with model parameter calibration. To this end, the presented sensitivity analysis can be extended to carry out sensitivity analysis w.r.t. to geometric and material parameters. Moreover, in the future work on the fracture-resistant energy-dissipating designs, cyclic loading scenarios with kinematic hardening models[46,62] and structural stability constraints[63] in the optimization formulation can be explored.

In addition, high-performance computing resources can be employed to carry out topology optimization in 3D design domains, which require additional computational resources. These issues will be addressed in our future work.

ACKNOWLEDGMENTS

The presented work is supported in part by the US National Science Foundation through grant CMMI-1762277. Any opinions, findings, conclusions, and recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the sponsors.

ORCID

Guodong Zhang https://orcid.org/0000-0003-1999-3033
Kapil Khandelwal https://orcid.org/0000-0002-5748-6019

REFERENCES

1. Bendsøe MP, Kikuchi N. Generating optimal topologies in structural design using a homogenization method. *Comput Methods Appl Mech Eng*. 1988;71(2):197-224.
2. Sigmund O, Maute K. Topology optimization approaches. *Struct Multidiscip Optim*. 2013;48(6):1031-1055.
3. Deaton JD, Grandhi RV. A survey of structural and multidisciplinary continuum topology optimization: post 2000. *Struct Multidiscip Optim*. 2014;49(1):1-38.
4. Duysinx P, Bendsøe MP. Topology optimization of continuum structures with local stress constraints. *Int J Numer Methods Eng*. 1998;43(8):1453-1478.
5. Le C et al. Stress-based topology optimization for continua. *Struct Multidiscip Optim*. 2010;41(4):605-620.
6. Luo Y, Kang Z. Topology optimization of continuum structures with Drucker–Prager yield stress constraints. *Comput Struct*. 2012;90-91:65-75.
7. Kato J, Hoshiba H, Takase S, Terada K, Kyoya T. Analytical sensitivity in topology optimization for elastoplastic composites. *Struct Multidiscip Optim*. 2015;52(3):507-526.
8. Alberdi R, Zhang G, Li L, Khandelwal K. A unified framework for nonlinear path-dependent sensitivity analysis in topology optimization. *Int J Numer Methods Eng*. 2018;115(1):1-56.
9. Li L, Zhang G, Khandelwal K. Topology optimization of energy absorbing structures with maximum damage constraint. *Int J Numer Methods Eng*. 2017;112(7):737-775.
10. Kato J, Lipka A, Ramm E. Multiphase material optimization for fiber reinforced composites with strain softening. *Struct Multidiscip Optim*. 2008;39(1):63-81.
11. Li L, Zhang G, Khandelwal K. Failure resistant topology optimization of structures using nonlocal elastoplastic-damage model. *Struct Multidiscip Optim*. 2018;58(4):1589-1618.
12. James KA, Waisman H. Failure mitigation in optimal topology design using a coupled nonlinear continuum damage model. *Comput Methods Appl Mech Eng*. 2014;268:614-631.
13. Amir O, Sigmund O. Reinforcement layout design for concrete structures based on continuum damage and truss topology optimization. *Struct Multidiscip Optim*. 2013;47(2):157-174.
14. Russ JB, Waisman H. Topology optimization for brittle fracture resistance. *Comput Methods Appl Mech Eng*. 2019;347:238-263.
15. Russ JB, Waisman H. A novel topology optimization formulation for enhancing fracture resistance with a single quasi-brittle material. *Int J Numer Methods Eng*. 2020;121(13):2827-2856.
16. Wu C, Fang J, Zhou S, et al. Level-set topology optimization for maximizing fracture resistance of brittle materials using phase-field fracture model. *Int J Numer Methods Eng*. 2020;121(13):2929-2945.
17. Li L, Khandelwal K. Topology optimization of energy-dissipating plastic structures with shear modified Gurson-Tvergaard-Needleman model. *J Struct Eng*. 2020;146(11):04020229.
18. Li L, Khandelwal K. Design of fracture resistant energy absorbing structures using elastoplastic topology optimization. *Struct Multidiscip Optim*. 2017;56(6):1447-1475.
19. Alberdi R, Khandelwal K. Topology optimization of pressure dependent elastoplastic energy absorbing structures with material damage constraints. *Finite Elem Anal Des*. 2017;133:42-61.
20. Russ JB, Waisman H. A novel elastoplastic topology optimization formulation for enhanced failure resistance via local ductile failure constraints and linear buckling analysis. *Comput Methods Appl Mech Eng*. 2021;373:113478.
21. Benzerga AA, Leblond JB, Needleman A, Tvergaard V. Ductile failure modeling. *Int J Fract*. 2016;201(1):29-80.
22. Tekoğlu C, Hutchinson JW, Pardoen T. On localization and void coalescence as a precursor to ductile fracture. *Philos Trans R Soc A Math Phys Eng Sci*. 2015;373(2038):20140121.
23. Rousselier G. Porous plasticity revisited: macroscopic and multiscale modeling. *Int J Plast*. 2021;136:102881.
24. McClintock FA. A criterion for ductile fracture by the growth of holes. *J Appl Mech*. 1968;35(2):363-371.
25. Rice JR, Tracey DM. On the ductile enlargement of voids in triaxial stress fields*. *J Mech Phys Solids*. 1969;17(3):201-217.

26. Gurson AL. Continuum theory of ductile rupture by void nucleation and growth: part I—Yield criteria and flow rules for porous ductile media. *J Eng Mater Technol*. 1977;99(1):2-15.
27. Chu CC, Needleman A. Void nucleation effects in biaxially stretched sheets. *J Eng Mater Technol*. 1980;102(3):249-256.
28. Tvergaard V. Influence of voids on shear band instabilities under plane strain conditions. *Int J Fract*. 1981;17(4):389-407.
29. Tvergaard V, Needleman A. Analysis of the cup-cone fracture in a round tensile bar. *Acta Metall*. 1984;32(1):157-169.
30. Bao Y, Wierzbicki T. On fracture locus in the equivalent strain and stress triaxiality space. *Int J Mech Sci*. 2004;46(1):81-98.
31. Barsoum I, Faleskog J. Rupture mechanisms in combined tension and shear—Experiments. *Int J Solids Struct*. 2007;44(6):1768-1786.
32. Xue L. Constitutive modeling of void shearing effect in ductile fracture of porous materials. *Eng Fract Mech*. 2008;75(11):3343-3366.
33. Nahshon K, Hutchinson JW. Modification of the Gurson model for shear failure. *Eur J Mech A/Solids*. 2008;27(1):1-17.
34. Malcher L, Andrade Pires FM, César de Sá JMA. An extended GTN model for ductile fracture under high and low stress triaxiality. *Int J Plast*. 2014;54:193-228.
35. Zhou J, Gao X, Sobotka JC, Webler BA, Cockeram BV. On the extension of the Gurson-type porous plasticity models for prediction of ductile fracture under shear-dominated conditions. *Int J Solids Struct*. 2014;51(18):3273-3291.
36. Dæhli LE, Morin D, Børvik T, Hopperstad OS. A lode-dependent Gurson model motivated by unit cell analyses. *Eng Fract Mech*. 2018;190:299-318.
37. Bettaieb MB, Lemoine X, Duchêne L, Habraken AM. On the numerical integration of an advanced Gurson model. *Int J Numer Methods Eng*. 2011;85(8):1049-1072.
38. Cao T-S, Montmitonnet P, Bouchard P-O. A detailed description of the Gurson–Tvergaard–Needleman model within a mixed velocity–pressure finite element formulation. *Int J Numer Methods Eng*. 2013;96(9):561-583.
39. Shahzamanian MM. Anisotropic Gurson-Tvergaard-Needleman plasticity and damage model for finite element analysis of elastic-plastic problems. *Int J Numer Methods Eng*. 2018;115(13):1527-1551.
40. Nielsen KL, Tvergaard V. Ductile shear failure or plug failure of spot welds modelled by modified Gurson model. *Eng Fract Mech*. 2010;77(7):1031-1047.
41. Mahnken R. Aspects on the finite-element implementation of the Gurson model including parameter identification. *Int J Plast*. 1999;15(11):1111-1137.
42. Zhang G, Alberdi R, Khandelwal K. Topology optimization with incompressible materials under small and finite deformations using mixed u/p elements. *Int J Numer Methods Eng*. 2018;115(8):1015-1052.
43. Gurtin ME, Fried E, Anand L. *The Mechanics and Thermodynamics of Continua*. Cambridge University Press; 2010.
44. Steinmann P, Miehe C, Stein E. Comparison of different finite deformation inelastic damage models within multiplicative elastoplasticity for ductile materials. *Comput Mech*. 1994;13(6):458-474.
45. Coleman BD, Noll W. The thermodynamics of elastic materials with heat conduction and viscosity. *Arch Ration Mech Anal*. 1963;13(1):167-178.
46. Miehe C, Apel N, Lambrecht M. Anisotropic additive plasticity in the logarithmic strain space: modular kinematic formulation and implementation based on incremental minimization principles for standard materials. *Comput Methods Appl Mech Eng*. 2002;191(47):5383-5425.
47. de Souza Neto EA, Peric D, Owen DR. *Computational Methods for Plasticity: Theory and Applications*. John Wiley & Sons; 2011.
48. de Souza Neto EA, Perić D, Dutko M, Owen DRJ. Design of simple low order finite elements for large strain analysis of nearly incompressible solids. *Int J Solids Struct*. 1996;33(20):3277-3296.
49. Christensen PW, Klarbring A. *An Introduction to Structural Optimization*. Vol 153. Springer; 2009.
50. Bourdin B. Filters in topology optimization. *Int J Numer Methods Eng*. 2001;50(9):2143-2158.
51. Sigmund O, Petersson J. Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct Optim*. 1998;16(1):68-75.
52. Wang F, Lazarov BS, Sigmund O, Jensen JS. Interpolation scheme for fictitious domain techniques and topology optimization of finite strain elastic problems. *Comput Methods Appl Mech Eng*. 2014;276:453-472.
53. Michaleris P, Tortorelli DA, Vidal CA. Tangent operators and design sensitivity formulations for transient non-linear coupled problems with applications to elastoplasticity. *Int J Numer Methods Eng*. 1994;37(14):2471-2499.
54. Crisfield MA. *Non-Linear Finite Element Analysis of Solids and Structures, Essentials*. Wiley; 1996.
55. Simo JC, Armero F. Geometrically non-linear enhanced strain mixed methods and the method of incompatible modes. *Int J Numer Methods Eng*. 1992;33(7):1413-1449.
56. He Z, Zhu H, Hu Y. An improved shear modified GTN model for ductile fracture of aluminium alloys under different stress states and its parameters identification. *Int J Mech Sci*. 2021;192:106081.
57. Kiran R, Khandelwal K. Gurson model parameters for ductile fracture simulation in ASTM A992 steels. *Fatigue Fract Eng Mater Struct*. 2014;37(2):171-183.
58. Svanberg K. The method of moving asymptotes—A new method for structural optimization. *Int J Numer Methods Eng*. 1987;24(2):359-373.
59. McNeel RAO. *Rhinoceros 3D, Version 6.0, in mcneel-rhinoceros*. Robert McNeel \& Associates; 2010.
60. Park R. Evaluation of ductility of structures and structural assemblages from laboratory testing. *Bull N Z Soc Earthq Eng*. 1989;22(3):155-166.
61. Wang F, Lazarov BS, Sigmund O. On projection methods, convergence and robust formulations in topology optimization. *Struct Multidiscip Optim*. 2011;43(6):767-784.

62. Besson J, Guillemer-Neel C. An extension of the green and Gurson models to kinematic hardening. *Mech Mater*. 2003;35(1):1-18.
63. Wriggers P. *Nonlinear Finite Element Methods*. Springer Science & Business Media; 2008.

How to cite this article: Zhang G, Khandelwal K. Gurson–Tvergaard–Needleman model guided fracture-resistant structural designs under finite deformations. *Int J Numer Methods Eng*. 2022;1-45. doi: 10.1002/nme.6971

APPENDIX A. IMPLEMENTATION DETAILS OF FINITE STRAIN SHEAR MODIFIED GTN MODEL

This appendix gives details on the implementation of the finite strain shear modified GTN model. Some notations useful in the following derivations are introduced first, which are

$$\widetilde{\boldsymbol{n}} \coloneqq \frac{\boldsymbol{\tau}_{dev}}{\|\boldsymbol{\tau}_{dev}\|} \text{ or } \boldsymbol{\tau}_{dev} = \|\boldsymbol{\tau}_{dev}\|\,\widetilde{\boldsymbol{n}},$$
$$Sinh[M] \coloneqq \frac{\partial\, Cosh[M]}{\partial M},$$
$$C_0 \coloneqq \sqrt{\tfrac{3}{2}} q_1 q_2 f^* \frac{1}{\sqrt{|z|}} Sinh[M],$$
$$\boldsymbol{A} \coloneqq \frac{\partial \phi}{\partial \boldsymbol{\tau}} = \widetilde{\boldsymbol{n}} + \frac{C_0}{3}\boldsymbol{I}.$$

Due to isotropy, the Kirchhoff stress $\boldsymbol{\tau}$ and elastic Finger tensor $\boldsymbol{b}^e$ are coaxial (see Equation (8)). The coaxiality of $\boldsymbol{\tau}$, $\boldsymbol{\tau}_{dev}$ and $\boldsymbol{\tau}_{vol}$ is obvious. The eigen decompositions of these tensors are.

$$\widetilde{\boldsymbol{n}} = \sum_{a=1}^{3} \widetilde{n}_a \boldsymbol{n}_a \otimes \boldsymbol{n}_a \;\text{ with }\; \boldsymbol{n}_i.\boldsymbol{n}_j = \delta_{ij}, \tag{A2}$$

$$\boldsymbol{\tau} = \sum_{a=1}^{3} \tau_a \boldsymbol{n}_a \otimes \boldsymbol{n}_a, \boldsymbol{\tau}_{vol} = \tau_m \boldsymbol{I}, \boldsymbol{\tau}_{dev} = \sum_{a=1}^{3} (\tau_a - \tau_m)\, \boldsymbol{n}_a \otimes \boldsymbol{n}_a, \tag{A3}$$

$$\boldsymbol{b}^e = \sum_{a=1}^{3} {\lambda_a^e}^2 \boldsymbol{n}_a \otimes \boldsymbol{n}_a, \tag{A4}$$

where $\boldsymbol{n}_a$ ($a = 1,2,3$) are orthonormal bases and λ_a^e ($a = 1,2,3$) the square roots of the eigenvalues of $\boldsymbol{b}^e$. Here the Kirchhoff pressure $\tau_m = (\tau_1 + \tau_2 + \tau_3)/3$. A bracket is used to denote the vector representation of tensors in eigen spaces, that is,

$$[\widetilde{\boldsymbol{n}}] = \begin{bmatrix} \widetilde{n}_1 \\ \widetilde{n}_2 \\ \widetilde{n}_3 \end{bmatrix}, \tag{A5}$$

$$[\boldsymbol{\tau}] = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}, [\boldsymbol{\tau}_{vol}] = \begin{bmatrix} \tau_m \\ \tau_m \\ \tau_m \end{bmatrix} = [\boldsymbol{P}_{vol}]\,[\boldsymbol{\tau}], [\boldsymbol{\tau}_{dev}] = \begin{bmatrix} \tau_{dev,1} \\ \tau_{dev,2} \\ \tau_{dev,3} \end{bmatrix} = [\boldsymbol{P}_{dev}]\,[\boldsymbol{\tau}], \tag{A6}$$

where the volumetric and deviatoric projection matrices $[\boldsymbol{P}_{vol}]$ and $[\boldsymbol{P}_{dev}]$ are defined as

$$[\boldsymbol{P}_{vol}] \coloneqq \frac{1}{3}[\boldsymbol{1}][\boldsymbol{1}]^T = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

$$[\boldsymbol{P}_{dev}] \coloneqq [\boldsymbol{I}] - [\boldsymbol{P}_{vol}] = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}, \tag{A7}$$

in which $[\boldsymbol{I}]$ denotes (3×3) identity matrix and $[\mathbf{1}] \coloneqq \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$.

In this appendix, the current step k is omitted in the subscript while the last step is denoted by a subscript $(k-1)$.

A.1. Integration of rate equations

There are three rate equations that need to be integrated for solving for the corresponding internal variables $\boldsymbol{b}^e$ (or $\boldsymbol{C}^p$ or $\boldsymbol{F}^p$), α, and f.

The first rate equation in Equation (17) can be derived, with the help of Equation $(4)_2$, further to.

$$\dot{\boldsymbol{C}}^i = -2\gamma \boldsymbol{F}^{-1}.\boldsymbol{A}.\boldsymbol{F}.\boldsymbol{C}^i \text{ with } \boldsymbol{C}^i \coloneqq \boldsymbol{C}^{p-1} \tag{A8}$$

which is then integrated using exponential map integrator that gives.

$$\boldsymbol{C}^i = exp\left(-2\gamma \Delta t \left(\boldsymbol{F}^{-1}.\boldsymbol{A}.\boldsymbol{F}\right)\right).\boldsymbol{C}^i_{k-1} \tag{A9}$$

which, by a straightforward manipulation, gives.

$$\boldsymbol{b}^{e^{tr}} = exp(2\Delta\gamma \boldsymbol{A}).\boldsymbol{b}^e \text{ with } \boldsymbol{b}^{e^{tr}} \coloneqq \boldsymbol{F}.\boldsymbol{C}^i_{k-1}.\boldsymbol{F}^T \text{ and } \Delta\gamma \coloneqq \gamma \Delta t. \tag{A10}$$

Taking the logarithm of Equation (A10) and expressing the results in eigen space gives.

$$\sum_{a=1}^{3} \varepsilon_a^{e^{tr}} \boldsymbol{n}_a^{tr} \otimes \boldsymbol{n}_a^{tr} = \Delta\gamma \widetilde{\boldsymbol{n}} + \Delta\gamma \frac{C_0}{3} \boldsymbol{I} + \sum_{a=1}^{3} \varepsilon_a^e \boldsymbol{n}_a \otimes \boldsymbol{n}_a, \tag{A11}$$

where the eigenvalues and eigenvectors of $\boldsymbol{b}^{e\,tr}$ are denoted by $\left(\lambda_a^{e\,tr}\right)^2$ and $\boldsymbol{n}_a^{tr}$ $(a = 1,2,3)$, respectively, with $\varepsilon_a^{e^{tr}} \coloneqq \ln \lambda_a^{e\,tr}$. Here Equation (A4) is used with $\varepsilon_a^e \coloneqq \ln \lambda_a^e$. Finally, Equation (A11) leads to.

$$\left[\boldsymbol{\varepsilon}^e\right] = \left[\boldsymbol{\varepsilon}^{e^{tr}}\right] - \Delta\gamma[\widetilde{\boldsymbol{n}}] - \Delta\gamma \frac{C_0}{3}[\mathbf{1}] \text{ and } \boldsymbol{n}_a = \boldsymbol{n}_a^{tr}, a = 1, 2, 3, \tag{A12}$$

where $[\boldsymbol{\varepsilon}^e] = \begin{bmatrix} \varepsilon_1^e & \varepsilon_2^e & \varepsilon_3^e \end{bmatrix}^T$ and $\left[\boldsymbol{\varepsilon}^{e^{tr}}\right] = \begin{bmatrix} \varepsilon_1^{e^{tr}} & \varepsilon_2^{e^{tr}} & \varepsilon_3^{e^{tr}} \end{bmatrix}^T$.

The second rate equation in Equation (18) for α can be straightforwardly integrated using backward Euler rule as.

$$\alpha = \alpha_{k-1} + \frac{\Delta\gamma}{(1-f)\zeta J} \left(\|\boldsymbol{\tau}_{dev}\| + C_0 \tau_m\right). \tag{A13}$$

Similarly, the third rate equation in Equation (19) for f is integrated as.

$$f = f_{k-1} + \Delta\gamma \left((1-f)C_0 + \cancel{k} \frac{\mathcal{D}}{(1-f)\zeta J} \left(\|\boldsymbol{\tau}_{dev}\| + C_0 \tau_m\right) + \sqrt{\frac{2}{3}} k_\omega f \omega_0 \right). \tag{A14}$$

A.2. Stress tensor

Due to isotropy, the elastic strain energy at macroscale ψ^e only depends on the eigenvalues of $\boldsymbol{b}^e$, see Equation (6). As a result, the Kirchhoff stress $\boldsymbol{\tau}$ defined in Equation (8) can be expressed in the eigen space as.

$$\boldsymbol{\tau} = \sum_{a=1}^{3} \lambda_a^e \frac{\partial \psi^e}{\partial \lambda_a^e} \boldsymbol{n}_a \otimes \boldsymbol{n}_a = \sum_{a=1}^{3} \frac{\partial \psi^e}{\partial \varepsilon_a^e} \boldsymbol{n}_a \otimes \boldsymbol{n}_a \equiv \frac{\partial \psi^e}{\partial \boldsymbol{\varepsilon}^e}. \tag{A15}$$

On the other hand, the strain energy ψ^e in Equation (6) can be rephrased, after a straightforward manipulation, as

$$\psi^e_{vol}\left(J^e\right) = \frac{1}{2}\kappa\left(\varepsilon^e_1 + \varepsilon^e_2 + \varepsilon^e_3\right)^2,$$
$$\psi^e_{iso}\left(\widehat{\boldsymbol{b}}^e\right) = \mu\left[\boldsymbol{\varepsilon}^e_{dev}\right]^T\left[\boldsymbol{\varepsilon}^e_{dev}\right] \quad \text{with} \quad \left[\boldsymbol{\varepsilon}^e_{dev}\right] = \left[\boldsymbol{P}_{dev}\right]\left[\boldsymbol{\varepsilon}^e\right]. \tag{A16}$$

Hence, with the calculated eigen vectors $\boldsymbol{n}_a = \boldsymbol{n}^{tr}_a$ $(a = 1,2,3)$, the eigenvalues of the Kirchhoff stress $\boldsymbol{\tau}$ can be computed by

$$\left[\boldsymbol{\tau}_{vol}\right] = \frac{\partial\psi^e_{vol}\left(J^e\right)}{\partial\left[\boldsymbol{\varepsilon}^e\right]} = 3\kappa\left[\boldsymbol{P}_{vol}\right]\left[\boldsymbol{\varepsilon}^e\right],$$
$$\left[\boldsymbol{\tau}_{dev}\right] = \frac{\partial\psi^e_{iso}\left(\widehat{\boldsymbol{b}}^e\right)}{\partial\left[\boldsymbol{\varepsilon}^e\right]} = 2\mu\left[\boldsymbol{P}_{dev}\right]\left[\boldsymbol{\varepsilon}^e\right]. \tag{A17}$$

Therefore, if elastic trial step satisfies the yield criterion $\phi \le 0$ in Equation (12), that is, $\boldsymbol{\varepsilon}^e = \boldsymbol{\varepsilon}^{e^{tr}}$ and $\Delta\gamma = 0$, $\alpha = \alpha_{k-1}$ and $f = f_{k-1}$, the stress can be updated by Equation (A17). Otherwise, the plastic flow is not zero at the current step. For this case, the calculation of stress tensor requires the updated internal variables. To that end, the substitution of Equation (A12) in Equation (A17) leads to

$$\tau_m = \tau^{tr}_m - \kappa\Delta\gamma C_0 \ \text{ with } \ \tau^{tr}_m := \kappa[\mathbf{1}]^T\left[\boldsymbol{\varepsilon}^{e^{tr}}\right],$$
$$\tau_d = \tau^{tr}_d - 2\mu\Delta\gamma \ \text{ with } \ \tau_d := \|\boldsymbol{\tau}_{dev}\|, \tau^{tr}_d := \left\|\boldsymbol{\tau}^{tr}_{dev}\right\|, \left[\boldsymbol{\tau}^{tr}_{dev}\right] := 2\mu\left[\boldsymbol{P}_{dev}\right]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right], \tag{A18}$$

where note that $[\boldsymbol{\tau}_{dev}] = \tau_d[\widetilde{\boldsymbol{n}}]$ and $\left[\boldsymbol{\tau}^{tr}_{dev}\right] = \tau^{tr}_d[\widetilde{\boldsymbol{n}}]$ and $[\widetilde{\boldsymbol{n}}]$ can be calculated by.

$$[\widetilde{\boldsymbol{n}}] = \left[\widetilde{\boldsymbol{n}}^{tr}\right] = \frac{\left[\boldsymbol{\tau}^{tr}_{dev}\right]}{\left\|\boldsymbol{\tau}^{tr}_{dev}\right\|} = \frac{\left[\boldsymbol{P}_{dev}\right]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}{\sqrt{\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]^T\left[\boldsymbol{P}_{dev}\right]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}}. \tag{A19}$$

Hence, the stress together with the internal variables ($\tau_m, \tau_d, \Delta\gamma, \alpha$, and f) are obtained by solving a system of nonlinear equations

$$\boldsymbol{R}_{\text{mat}} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \end{bmatrix} = \mathbf{0} \ \text{ with}$$

$$R_1 = \tau_d - \sqrt{\frac{2}{3}}\text{sign}(z)\sqrt{|z|}\zeta J,$$
$$R_2 = \alpha - \alpha_{k-1} - \frac{\Delta\gamma}{(1-f)\zeta J}\left(\tau_d + C_0\tau_m\right), \tag{A20}$$
$$R_3 = f - f_{k-1} - \Delta\gamma\left((1-f)C_0 + \mathscr{K}\frac{\mathcal{D}}{(1-f)\zeta J}\left(\tau_d + C_0\tau_m\right) + \sqrt{\frac{2}{3}}k_\omega f\omega_0\right),$$
$$R_4 = \tau_m - \tau^{tr}_m + \kappa\Delta\gamma C_0,$$
$$R_5 = \tau_d - \tau^{tr}_d + 2\mu\Delta\gamma,$$

using Newton–Raphson (NR) solver. The calculation of the Jacobian matrix $\partial\boldsymbol{R}_{\text{mat}}/\partial\boldsymbol{\chi}$ with $\boldsymbol{\chi} = \begin{bmatrix}\Delta\gamma & \alpha & f & \tau_m & \tau_d\end{bmatrix}^T$, can be straightforwardly carried out.

A.3. Consistent tangent moduli

The consistent tangent moduli $\mathbb{A} := \partial\boldsymbol{P}/\partial\boldsymbol{F}$ is needed for guaranteeing quadratic convergence rate in NR solver, where $\boldsymbol{P}$ is the first PK stress with $\boldsymbol{P} = \boldsymbol{\tau}.\boldsymbol{F}^{-T}$. Expressed in index form.

$$\mathbb{A}_{ijkl} = JF^{-1}_{jp}\mathbb{a}_{ipkq}F^{-1}_{lq} \text{ with } J\mathbb{a}_{ijkl} = \frac{\partial \tau_{ij}}{\partial F_{kq}}F_{lq} - \tau_{il}\delta_{jk}, \tag{A21}$$

where the only term that needs to be calculated is $\partial\boldsymbol{\tau}/\partial\boldsymbol{F}$ and is computed by chain rule.

$$\frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{F}} = \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\varepsilon}^{e^{tr}}} : \frac{\partial \boldsymbol{\varepsilon}^{e^{tr}}}{\partial \boldsymbol{b}^{e^{tr}}} : \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{F}} + \frac{\partial \boldsymbol{\tau}}{\partial J} \otimes \frac{\partial J}{\partial \boldsymbol{F}}. \tag{A22}$$

Since the tensors $\boldsymbol{\tau}$, $\boldsymbol{\varepsilon}^{e^{tr}}$ and $\boldsymbol{b}^{e^{tr}}$ are coaxial, following the procedures given in Appendix A in de Souza Neto et al.,[47] the calculation of the terms $\partial\boldsymbol{\tau}/\partial\boldsymbol{\varepsilon}^{e^{tr}}$ and $\partial\boldsymbol{\varepsilon}^{e^{tr}}/\partial\boldsymbol{b}^{e^{tr}}$ can be carried out in the eigen space and requires only $\partial[\boldsymbol{\tau}]/\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]$ and $\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]/\partial\left[\boldsymbol{b}^{e^{tr}}\right]$.

A.3.1. Plastic step

For plastic step, from Equation (A20), it can be seen that the driving force for the stress tensor $\boldsymbol{\tau}$ comes from $\boldsymbol{\varepsilon}^{e^{tr}}$ and J. Thus, taking the total differentiation of $\boldsymbol{R}_{\text{mat}}$ in Equation (A20) at the solution point gives.

$$d\boldsymbol{R}_{mat} = \frac{\partial \boldsymbol{R}_{mat}}{\partial \chi}d\chi + \frac{\partial \boldsymbol{R}_{mat}}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}d\left[\boldsymbol{\varepsilon}^{e^{tr}}\right] + \frac{\partial \boldsymbol{R}_{mat}}{\partial J}dJ = \boldsymbol{0}, \tag{A23}$$

where, again, the tedious but straightforward derivation of the derivative terms $\partial\boldsymbol{R}_{\text{mat}}/\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]$ and $\partial\boldsymbol{R}_{\text{mat}}/\partial J$ is omitted here for brevity. From Equation (A23), it can be seen that

$$\begin{aligned}
\frac{\partial \chi}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} &= \left[\frac{\partial \boldsymbol{R}_{\text{mat}}}{\partial \chi}\right]^{-1}\frac{\partial \boldsymbol{R}_{\text{mat}}}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}, \\
\frac{\partial \chi}{\partial J} &= \left[\frac{\partial \boldsymbol{R}_{mat}}{\partial \chi}\right]^{-1}\frac{\partial \boldsymbol{R}_{mat}}{\partial J}
\end{aligned} \tag{A24}$$

which gives the derivatives $\partial\tau_m/\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]$, $\partial\tau_d/\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]$, $\partial\tau_m/\partial J$ and $\partial\tau_d/\partial J$.

With that, the term $\partial[\boldsymbol{\tau}]/\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]$ is calculated as

$$\begin{aligned}
\frac{\partial[\boldsymbol{\tau}]}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} &= [\mathbf{1}]\frac{\partial \tau_m}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} + \frac{\partial\left[\boldsymbol{\tau}_{dev}\right]}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} \\
\frac{\partial\left[\boldsymbol{\tau}_{dev}\right]}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} &= \left[\widetilde{\boldsymbol{n}}^{tr}\right]\frac{\partial \tau_d}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} + \tau_d\frac{\partial\left[\widetilde{\boldsymbol{n}}^{tr}\right]}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} \\
\frac{\partial\left[\widetilde{\boldsymbol{n}}^{tr}\right]}{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} &= \frac{[\boldsymbol{P}_{dev}]}{\sqrt{\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]^T[\boldsymbol{P}_{dev}]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}} - \frac{[\boldsymbol{P}_{dev}]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}{\left(\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]^T[\boldsymbol{P}_{dev}]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]\right)^{\frac{3}{2}}}\left([\boldsymbol{P}_{dev}]\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]\right)^T
\end{aligned} \tag{A25}$$

and the term $\partial\boldsymbol{\tau}/\partial J$ is calculated as.

$$\frac{\partial \boldsymbol{\tau}}{\partial J} = \frac{\partial \tau_m}{\partial J}\boldsymbol{I} + \frac{\partial \tau_d}{\partial J}\widetilde{\boldsymbol{n}}. \tag{A26}$$

The derivation of the term $\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]/\partial\left[\boldsymbol{b}^{e^{tr}}\right]$ is straightforward, that is,

$$\frac{\partial\left[\boldsymbol{\varepsilon}^{e^{tr}}\right]}{\partial\left[\boldsymbol{b}^{e^{tr}}\right]} = \begin{bmatrix} \frac{1}{2\lambda_1^{e^{tr}2}} & 0 & 0 \\ 0 & \frac{1}{2\lambda_2^{e^{tr}2}} & 0 \\ 0 & 0 & \frac{1}{2\lambda_3^{e^{tr}2}} \end{bmatrix}. \tag{A27}$$

Finally, the remaining terms in Equation (A22), that is, $\partial \boldsymbol{b}^{e^{tr}}/\partial \boldsymbol{F}$ and $\partial J/\partial \boldsymbol{F}$, are derived as.

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{F}} = \boldsymbol{I} \boxtimes \left(\boldsymbol{F}.\boldsymbol{C}_{k-1}^{i}\right) + \left(\boldsymbol{F}.\boldsymbol{C}_{k-1}^{i}\right) \boxdot \boldsymbol{I}, \tag{A28}$$

$$\frac{\partial J}{\partial \boldsymbol{F}} = J\boldsymbol{F}^{-T}, \tag{A29}$$

where the operations $\boxtimes$ and $\boxdot$ are defined such that

$$\begin{aligned} \boldsymbol{A} \boxtimes \boldsymbol{B} &:= A_{ik}B_{jl}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l, \\ \boldsymbol{A} \boxdot \boldsymbol{B} &:= A_{il}B_{jk}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l, \end{aligned} \tag{A30}$$

for any second-order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$.

A.3.2. Elastic step

For elastic step, the computations of $\partial \boldsymbol{\varepsilon}^{e^{tr}}/\partial \boldsymbol{b}^{e^{tr}}$ and $\partial \boldsymbol{b}^{e^{tr}}/\partial \boldsymbol{F}$ still follow Equations (A27) and (A28), while the other terms in Equation (A22) are calculated as

$$\begin{aligned} &\frac{\partial [\boldsymbol{\tau}]}{\partial \left[\boldsymbol{\varepsilon}^{e^{tr}}\right]} = \frac{\partial [\boldsymbol{\tau}]}{\partial [\boldsymbol{\varepsilon}^{e}]} = 3\kappa \left[\boldsymbol{P}_{vol}\right] + 2\mu \left[\boldsymbol{P}_{dev}\right], \\ &\frac{\partial \boldsymbol{\tau}}{\partial J} = \boldsymbol{0}. \end{aligned} \tag{A31}$$

APPENDIX B. DERIVATIVES REQUIRED FOR THE SENSITIVITY ANALYSIS

This appendix gives the required derivatives in Equation (42) for the sensitivity analysis. For the ease of presentation, tensor notations are also used besides matrix forms. It is remarked here that the constitutive model (GTN) evaluation is based on the modified deformation gradient $\overline{\boldsymbol{F}}$ in Equation (*28*). Hence, all the terms inside the model (e.g., $\boldsymbol{F}^{e}, \boldsymbol{F}^{p}, \boldsymbol{b}^{e}, \boldsymbol{\tau}, \boldsymbol{P}$, J, f, α) are based on the input $\overline{\boldsymbol{F}}$ instead of $\boldsymbol{F}$. Due to high computational cost of 3D nonlinear topology optimization, only 2D plane strain problems are considered for the fracture-resistant design optimization. Therefore, the derivation in this section is based on 2D plane strain assumption. It can, however, be extended to 3D case straightforwardly. For notational convenience in the derivation, the dependence on $\overline{\boldsymbol{F}}$ is revealed only on the Kirchhoff and first PK stress tensors by using $\overline{\boldsymbol{\tau}}$ and $\overline{\boldsymbol{P}}$ while overbar is not used on the others. Note that the step index—k for the current step and $(k-1)$ for the previous step—will be present in the super/subscript when needed for clarification purposes. Besides, super/subscripts that include load step, element, and integration point indices are omitted for simplicity when there is no confusion. Some useful derivatives are given below for reference.

Derivatives w.r.t. ρ.

$$\frac{\partial \boldsymbol{F}}{\partial \rho} = \frac{\partial \eta}{\partial \rho}\nabla_{\boldsymbol{X}}\boldsymbol{u}, \frac{\partial \boldsymbol{F}_0}{\partial \rho} = \frac{\partial \eta}{\partial \rho}\nabla_{\boldsymbol{X}}^{0}\boldsymbol{u}, \tag{B1}$$

$$\frac{\partial r}{\partial \rho} = \frac{\partial r}{\partial \boldsymbol{F}} : \frac{\partial \boldsymbol{F}}{\partial \rho} + \frac{\partial r}{\partial \boldsymbol{F}_0} : \frac{\partial \boldsymbol{F}_0}{\partial \rho} \text{ with } \frac{\partial r}{\partial \boldsymbol{F}} = -r\boldsymbol{F}^{-T} \text{ and } \frac{\partial r}{\partial \boldsymbol{F}_0} = r\boldsymbol{F}_0^{-T}, \tag{B2}$$

$$\frac{\partial \overline{\boldsymbol{F}}}{\partial \rho} = \frac{1}{2}r^{-1/2}\frac{\partial r}{\partial \rho}\boldsymbol{F} + r^{1/2}\frac{\partial \boldsymbol{F}}{\partial \rho}, \tag{B3}$$

$$\frac{\partial J}{\partial \rho} = J\boldsymbol{F}_0^{-T} : \frac{\partial \boldsymbol{F}_0}{\partial \rho}, \tag{B4}$$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \rho} = \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_k} : \frac{\partial \overline{\boldsymbol{F}}_k}{\partial \rho} + \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_{k-1}} : \frac{\partial \overline{\boldsymbol{F}}_{k-1}}{\partial \rho}, \tag{B5}$$

$$\frac{\partial \overline{\boldsymbol{\tau}}}{\partial \rho} = \mathcal{G}\overline{\boldsymbol{\tau}}, \frac{\partial \tau_d}{\partial \rho} = \mathcal{G}\tau_d, \frac{\partial \tau_m}{\partial \rho} = \mathcal{G}\tau_m \text{ with } \mathcal{G} = \frac{p_1\left(1-\varepsilon_1\right)\rho^{p_1-1}}{\left[\varepsilon_1 + \left(1-\varepsilon_1\right)\rho^{p_1}\right]}, \tag{B6}$$

$$\frac{\partial \zeta}{\partial \rho} = \frac{\partial \sigma_{y0}}{\partial \rho} + \frac{\partial k_1}{\partial \rho}\alpha + \frac{\partial k_2}{\partial \rho}(1 - exp(-\delta\alpha)), \tag{B7}$$

$$\frac{\partial M}{\partial \rho} = \frac{3}{2} q_2 \frac{1}{\zeta J} \left(\frac{\partial \tau_m}{\partial \rho} - \frac{\tau_m}{\zeta} \frac{\partial \zeta}{\partial \rho} - \frac{\tau_m}{J} \frac{\partial J}{\partial \rho} \right), \tag{B8}$$

$$\frac{\partial z}{\partial \rho} = \frac{\partial z}{\partial M} \frac{\partial M}{\partial \rho}, \tag{B9}$$

$$\frac{\partial C_0}{\partial \rho} = \sqrt{\frac{3}{2}} q_1 q_2 f^* \left(-\frac{1}{2} \frac{1}{|z|^{3/2}} sign(z) \frac{\partial z}{\partial \rho} Sinh[M] + \frac{1}{\sqrt{|z|}} \frac{\partial Sinh[M]}{M} \frac{\partial M}{\partial \rho} \right), \tag{B10}$$

$$\frac{\partial \boldsymbol{A}}{\partial \rho} = \frac{1}{3} \frac{\partial C_0}{\partial \rho} \boldsymbol{I}, \tag{B11}$$

where the Equation (B6) is due to the choice of $\boldsymbol{b}^e$ as auxiliary variable and the dependency of $\overline{\boldsymbol{\tau}}$ on $\boldsymbol{b}^e$ through Equation (8). With that, the derivative $\partial \widetilde{\boldsymbol{n}} / \partial \rho = \boldsymbol{0}$, which leads to the derivative in Equation (B11).

Derivatives w.r.t. $\boldsymbol{u}$

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}} = \eta \boldsymbol{B}, \quad \frac{\partial \boldsymbol{F}_0}{\partial \boldsymbol{u}} = \eta \boldsymbol{B}_0, \tag{B12}$$

$$\frac{\partial r}{\partial \boldsymbol{u}} = r\eta \left(\boldsymbol{F}_0^{-T} : \boldsymbol{B}_0 - \boldsymbol{F}^{-T} : \boldsymbol{B} \right), \tag{B13}$$

$$\frac{\partial \overline{\boldsymbol{F}}}{\partial \boldsymbol{u}} = r^{1/2} \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}} + \frac{1}{2} r^{-1/2} \boldsymbol{F} \otimes \frac{\partial r}{\partial \boldsymbol{u}}, \tag{B14}$$

$$\frac{\partial J}{\partial \boldsymbol{u}} = \frac{\partial det \boldsymbol{F}_0}{\partial \boldsymbol{u}} = J\eta \boldsymbol{F}_0^{-T} : \boldsymbol{B}_0, \tag{B15}$$

$$\frac{\partial M}{\partial \boldsymbol{u}} = \frac{\partial M}{\partial J} \frac{\partial J}{\partial \boldsymbol{u}}, \tag{B16}$$

$$\frac{\partial z}{\partial \boldsymbol{u}} = \frac{\partial z}{\partial M} \frac{\partial M}{\partial \boldsymbol{u}}, \tag{B17}$$

$$\frac{\partial C_0}{\partial \boldsymbol{u}} = \frac{\partial C_0}{\partial M} \frac{\partial M}{\partial \boldsymbol{u}}, \tag{B18}$$

$$\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{u}} = \frac{1}{3} \boldsymbol{I} \otimes \frac{\partial C_0}{\partial \boldsymbol{u}}, \tag{B19}$$

where simple derivatives such as $\partial M / \partial J$, $\partial z / \partial M$, and $\partial C_0 / \partial M$ are omitted.

Derivatives of $\boldsymbol{b}^{e^{tr}}$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_k} = \boldsymbol{I} \odot \left(\overline{\boldsymbol{F}}_{k-1}^{-1} . \boldsymbol{b}_{k-1}^e . \overline{\boldsymbol{F}}_\Delta^T \right) + \left(\overline{\boldsymbol{F}}_\Delta . \boldsymbol{b}_{k-1}^e . \overline{\boldsymbol{F}}_{k-1}^{-T} \right) \boxdot \boldsymbol{I}, \tag{B20}$$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_{k-1}} = -\overline{\boldsymbol{F}}_\Delta \odot \left(\overline{\boldsymbol{F}}_{k-1}^{-1} . \boldsymbol{b}_{k-1}^e . \overline{\boldsymbol{F}}_\Delta^T \right) - \left(\overline{\boldsymbol{F}}_\Delta . \boldsymbol{b}_{k-1}^e . \overline{\boldsymbol{F}}_{k-1}^{-T} \right) \boxdot \overline{\boldsymbol{F}}_\Delta, \tag{B21}$$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{b}_{k-1}^e} = \frac{1}{2} \left(\boldsymbol{F}_\Delta \boxtimes \boldsymbol{F}_\Delta + \boldsymbol{F}_\Delta \boxdot \boldsymbol{F}_\Delta \right), \tag{B22}$$

where $\overline{\boldsymbol{F}}_\Delta := \overline{\boldsymbol{F}}_k . \overline{\boldsymbol{F}}_{k-1}^{-1}$, $\boldsymbol{b}^{e^{tr}} = \overline{\boldsymbol{F}}_\Delta . \boldsymbol{b}_{k-1}^e . \overline{\boldsymbol{F}}_\Delta^T$, and again the step number is put as subscript in $\boldsymbol{b}^e$ and $\overline{\boldsymbol{F}}$. The operations $\boxtimes$ and $\boxdot$ are defined in Equation (A30), while the operation $\odot$ is defined as

$$\boldsymbol{A} \odot \boldsymbol{B} := A_{ik} B_{lj} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l, \tag{B23}$$

Derivatives w.r.t. $\overline{\boldsymbol{\tau}}$

$$\frac{\partial \tau_m}{\partial \overline{\boldsymbol{\tau}}} = \frac{1}{3} \boldsymbol{I}, \quad \frac{\partial \tau_d}{\partial \overline{\boldsymbol{\tau}}} = \widetilde{\boldsymbol{n}}, \tag{B24}$$

$$\frac{\partial \widetilde{\boldsymbol{n}}}{\partial \overline{\boldsymbol{\tau}}} = \frac{\partial}{\partial \overline{\boldsymbol{\tau}}} \left(\frac{\overline{\boldsymbol{\tau}}_{dev}}{\left\| \overline{\boldsymbol{\tau}}_{dev} \right\|} \right) = \frac{1}{\tau_d} \mathbb{P}^s_{dev} - \frac{1}{\tau_d^2} \overline{\boldsymbol{\tau}}_{dev} \otimes \widetilde{\boldsymbol{n}} = \frac{1}{\tau_d} \left(\mathbb{P}^s_{dev} - \widetilde{\boldsymbol{n}} \otimes \widetilde{\boldsymbol{n}} \right), \tag{B25}$$

$$\frac{\partial T}{\partial \overline{\boldsymbol{\tau}}} = \frac{1}{3\tau_e}\boldsymbol{I} - \sqrt{\frac{3}{2}}\frac{\tau_m}{\tau_e^2}\widetilde{\boldsymbol{n}}, \tag{B26}$$

$$\frac{\partial \check{J}_3}{\partial \overline{\boldsymbol{\tau}}} = \check{J}_3 \overline{\boldsymbol{\tau}}_{dev}^{-T} : \mathbb{P}^s_{dev}, \tag{B27}$$

$$\frac{\partial \omega}{\partial \overline{\boldsymbol{\tau}}} = -\frac{27^2}{2}\frac{\check{J}_3}{\tau_e^6}\left(\frac{\partial \check{J}_3}{\partial \overline{\boldsymbol{\tau}}} - 3\sqrt{\frac{3}{2}}\frac{\check{J}_3}{\tau_e}\frac{\partial \tau_d}{\partial \overline{\boldsymbol{\tau}}}\right), \tag{B28}$$

$$\frac{\partial M}{\partial \overline{\boldsymbol{\tau}}} = \frac{1}{3}\frac{\partial M}{\partial \tau_m}\boldsymbol{I}, \tag{B29}$$

$$\frac{\partial z}{\partial \overline{\boldsymbol{\tau}}} = \frac{\partial z}{\partial M}\frac{\partial M}{\partial \overline{\boldsymbol{\tau}}}, \tag{B30}$$

$$\frac{\partial \boldsymbol{A}}{\partial \overline{\boldsymbol{\tau}}} = \frac{\partial \widetilde{\boldsymbol{n}}}{\partial \overline{\boldsymbol{\tau}}} + \frac{1}{\sqrt{6}}q_1 q_2 f^* \boldsymbol{I} \otimes \left(-\frac{1}{2}\frac{1}{|z|^{\frac{3}{2}}} sign(z) Sinh[M]\frac{\partial z}{\partial \overline{\boldsymbol{\tau}}} + \frac{1}{\sqrt{|z|}}\frac{\partial Sinh[M]}{M}\frac{\partial M}{\partial \overline{\boldsymbol{\tau}}}\right), \tag{B31}$$

where the straightforward derivatives such as $\partial M/\partial \tau_m$ and $\partial z/\partial M$ are omitted.

Derivatives of $\exp(2\Delta\gamma \boldsymbol{A})$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \Delta\gamma} = 2\boldsymbol{A}.exp(2\Delta\gamma \boldsymbol{A}), \tag{B32}$$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \alpha} = 2\Delta\gamma exp(2\Delta\gamma \boldsymbol{A}).\frac{\partial \boldsymbol{A}}{\partial \alpha}, \tag{B33}$$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial f} = 2\Delta\gamma exp(2\Delta\gamma \boldsymbol{A}).\frac{\partial \boldsymbol{A}}{\partial f}, \tag{B34}$$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{A}} = \frac{\partial [exp(2\Delta\gamma \boldsymbol{A})]}{\partial [\boldsymbol{A}]} = \sum_{i=1}^{3}\sum_{j=1}^{3}\left(2\Delta\gamma\delta_{ij}e^{2\Delta\gamma A_i}\right)\boldsymbol{e}_i \otimes \boldsymbol{e}_j, \tag{B35}$$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{b}^e} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{A}} : \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{b}^e}, \tag{B36}$$

where a bracket [■] outside a tensor denotes the eigenvalue vector of the tensor. The derivative $\partial \exp(2\Delta\gamma \boldsymbol{A})/\partial \boldsymbol{A}$ in Equation (B35) is carried out in the eigen space following the approach in Reference 47. The derivative $\partial \boldsymbol{A}/\partial \boldsymbol{b}^e$ is computed by chain rule with the term $\partial \boldsymbol{A}/\partial \overline{\boldsymbol{\tau}}$ in Equation (B31) and the term $\partial \overline{\boldsymbol{\tau}}/\partial \boldsymbol{b}^e$ in Equation (B37).

Derivatives w.r.t. $\boldsymbol{b}^e$

The implicit dependence on $\boldsymbol{b}^e$ comes exclusively from the Kirchhoff stress $\overline{\boldsymbol{\tau}}$. The derivative $\partial \overline{\boldsymbol{\tau}}/\partial \boldsymbol{b}^e$ can be simply carried out using chain rule as

$$\frac{\partial \overline{\boldsymbol{\tau}}}{\partial \boldsymbol{b}^e} = \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \boldsymbol{\varepsilon}^e} : \frac{\partial \boldsymbol{\varepsilon}^e}{\partial \boldsymbol{b}^e}, \tag{B37}$$

where the derivatives $\partial \overline{\boldsymbol{\tau}}/\partial \boldsymbol{\varepsilon}^e$ and $\partial \boldsymbol{\varepsilon}^e/\partial \boldsymbol{b}^e$ can be straightforwardly carried out in the eigen space with derivatives of the eigenvalues $\partial[\overline{\boldsymbol{\tau}}]/\partial[\boldsymbol{\varepsilon}^e]$ and $\partial[\boldsymbol{\varepsilon}^e]/\partial\left[\boldsymbol{b}^e\right]$ computed as

$$\begin{aligned}
\frac{\partial[\overline{\boldsymbol{\tau}}]}{\partial[\boldsymbol{\varepsilon}^e]} &= 3\kappa[\boldsymbol{P}_{vol}] + 2\mu[\boldsymbol{P}_{dev}], \\
\frac{\partial[\boldsymbol{\varepsilon}^e]}{\partial\left[\boldsymbol{b}^e\right]} &= \begin{bmatrix} \frac{1}{2\lambda_1^{e2}} & 0 & 0 \\ 0 & \frac{1}{2\lambda_2^{e2}} & 0 \\ 0 & 0 & \frac{1}{2\lambda_3^{e2}} \end{bmatrix},
\end{aligned} \tag{B38}$$

where $[\boldsymbol{P}_{vol}]$ and $[\boldsymbol{P}_{dev}]$ are given in Appendix A, and λ_i^e $(i = 1,2,3)$ are square root of the eigenvalues of $\boldsymbol{b}^e$.

B.1 Derivatives of W^p

Since W^p is different by only a negative sign from the objective function, this section gives the derivatives of W^p, that is, taking F as W^p in Equation $(42)_1$. By the notations introduced in Appendix A, Equation (33) can be equivalently rephrased as

$$W^p = \sum_{k=1}^{n}\sum_{e=1}^{n_{ele}}\sum_{s=1}^{n_{ipt}} \rho_e^{p_w}\Delta\gamma_{e_s}^k(\tau_d + C_0\tau_m)_{e_s}^k w_{e_s}, \tag{B39}$$

B.1.1. *Derivative $\partial W^p/\partial\rho$*

$$\frac{\partial W^p}{\partial \boldsymbol{\rho}} = \begin{bmatrix} \frac{\partial W^p}{\partial \rho_1} & \frac{\partial W^p}{\partial \rho_2} & \cdots & \frac{\partial W^p}{\partial \rho_{n_{ele}}} \end{bmatrix}, \tag{B40}$$

where for $e = 1, \ldots, n_{ele}$,

$$\frac{\partial W^p}{\partial \rho_e} = \sum_{k=1}^{n}\sum_{s=1}^{n_{ipt}} \left[\rho_e^{p_w}\Delta\gamma_{e_s}^k \left(\frac{\partial \tau_d}{\partial \rho} + \frac{\partial C_0}{\partial \rho}\tau_m + C_0\frac{\partial \tau_m}{\partial \rho} \right)_{e_s}^k w_{e_s} + p_w \rho_e^{p_w - 1}\Delta\gamma_{e_s}^k(\tau_d + C_0\tau_m)_{e_s}^k w_{e_s} \right], \tag{B41}$$

in which the terms $\partial\tau_d/\partial\rho$, $\partial C_0/\partial\rho$ and $\partial\tau_m/\partial\rho$ are given in Equations (B6) and (B10).

B.1.2. *Derivative $\partial W^p/\partial \boldsymbol{u}^k$*

$$\frac{\partial W^p}{\partial \boldsymbol{u}^k} = \mathop{\mathcal{A}}_{e=1}^{n_{ele}} \frac{\partial W^p}{\partial \boldsymbol{u}_e^k} \quad \text{with} \quad \frac{\partial W^p}{\partial \boldsymbol{u}_e^k} = \sum_{s=1}^{n_{ipt}} \rho_e^{p_w}\Delta\gamma_{e_s}^k \left(\tau_m \frac{\partial C_0}{\partial \boldsymbol{u}} \right)_{e_s}^k w_{e_s}, \tag{B42}$$

in which $\partial C_0/\partial \boldsymbol{u}$ at each quadrature point is given in Equation (B18).

B.1.3. *Derivative $\partial W^p/\partial \boldsymbol{v}^k$*

$$\frac{\partial W^p}{\partial \boldsymbol{v}^k} = \begin{bmatrix} \frac{\partial W^p}{\partial \boldsymbol{v}_1^k} & \frac{\partial W^p}{\partial \boldsymbol{v}_2^k} & \cdots & \frac{\partial W^p}{\partial \boldsymbol{v}_{n_{ele}}^k} \end{bmatrix} \text{ with}$$

$$\frac{\partial W^p}{\partial \boldsymbol{v}_e^k} = \begin{bmatrix} \frac{\partial W^p}{\partial \boldsymbol{v}_{e_1}^k} & \frac{\partial W^p}{\partial \boldsymbol{v}_{e_2}^k} & \cdots & \frac{\partial W^p}{\partial \boldsymbol{v}_{e_{(n_{ipt})}}^k} \end{bmatrix}, e = 1, \ldots, n_{ele}, \tag{B43}$$

where

$$\frac{\partial W^p}{\partial \boldsymbol{v}_{e_s}^k} = \begin{bmatrix} \frac{\partial W^p}{\partial \Delta\gamma_{e_s}^k} & \frac{\partial W^p}{\partial \alpha_{e_s}^k} & \frac{\partial W^p}{\partial f_{e_s}^k} & \frac{\partial W^p}{\partial (\boldsymbol{b}^e)_{e_s}^k} \end{bmatrix}, s = 1, \ldots, n_{ipt} \tag{B44}$$

with

$$\frac{\partial W^p}{\partial \Delta\gamma_{e_s}^k} = \rho_e^{p_w}(\tau_d + C_0\tau_m)_{e_s}^k w_{e_s}, \tag{B45}$$

$$\frac{\partial W^p}{\partial \alpha_{e_s}^k} = \rho_e^{p_w}\Delta\gamma_{e_s}^k \left(\frac{\partial C_0}{\partial \alpha}\tau_m \right)_{e_s}^k w_{e_s}, \tag{B46}$$

$$\frac{\partial W^p}{\partial f_{e_s}^k} = \rho_e^{p_w}\Delta\gamma_{e_s}^k \left(\frac{\partial C_0}{\partial f}\tau_m \right)_{e_s}^k w_{e_s}, \tag{B47}$$

$$\frac{\partial W^p}{\partial (\boldsymbol{b}^e)_{e_s}^k} = \rho_e^{p_w}\Delta\gamma_{e_s}^k \left(\frac{\partial \tau_d}{\partial \boldsymbol{b}^e} + \frac{\partial C_0}{\partial \boldsymbol{b}^e}\tau_m + C_0\frac{\partial \tau_m}{\partial \boldsymbol{b}^e} \right)_{e_s}^k w_{e_s}, \tag{B48}$$

in which

$$\frac{\partial C_0}{\partial \alpha} = \sqrt{\frac{3}{2}} q_1 q_2 f^* \left(-\frac{1}{2} \frac{1}{|z|^{3/2}} sign(z) \frac{\partial z}{\partial \alpha} Sinh[M] + \frac{1}{\sqrt{|z|}} \frac{\partial Sinh[M]}{\partial M} \frac{\partial M}{\partial \alpha} \right), \tag{B49}$$

$$\frac{\partial M}{\partial \alpha} = -\frac{3}{2} q_2 \frac{\tau_m}{\zeta^2 J} \frac{\partial \zeta}{\partial \alpha}, \tag{B50}$$

$$\frac{\partial z}{\partial \alpha} = -2 q_1 f^* Sinh[M] \frac{\partial M}{\partial \alpha}, \tag{B51}$$

$$\frac{\partial C_0}{\partial f} = \sqrt{\frac{3}{2}} q_1 q_2 Sinh[M] \left(\frac{1}{\sqrt{|z|}} \frac{\partial f^*}{\partial f} - \frac{1}{2} \frac{1}{|z|^{3/2}} sign(z) f^* \frac{\partial z}{\partial f} \right), \tag{B52}$$

$$\frac{\partial C_0}{\partial \boldsymbol{b}^e} = \sqrt{\frac{3}{2}} q_1 q_2 f^* \left(\frac{1}{\sqrt{|z|}} \frac{\partial Sinh[M]}{\partial M} \frac{\partial M}{\partial \boldsymbol{b}^e} - \frac{1}{2} \frac{1}{|z|^{3/2}} sign(z) Sinh[M] \frac{\partial z}{\partial \boldsymbol{b}^e} \right), \tag{B53}$$

$$\frac{\partial M}{\partial \boldsymbol{b}^e} = \frac{\partial M}{\partial \tau_m} \frac{\partial \tau_m}{\partial \overline{\boldsymbol{\tau}}} : \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \boldsymbol{b}^e}, \tag{B54}$$

$$\frac{\partial z}{\partial \boldsymbol{b}^e} = \frac{\partial z}{\partial \tau_m} \frac{\partial \tau_m}{\partial \overline{\boldsymbol{\tau}}} : \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \boldsymbol{b}^e}, \tag{B55}$$

where the terms $\partial \tau_m / \partial \overline{\boldsymbol{\tau}}$, $\partial \overline{\boldsymbol{\tau}} / \partial \boldsymbol{b}^e$ can be found in Equations (B24) and (B37), while other terms such as $\partial \zeta / \partial \alpha$, $\partial z / \partial f$, $\partial f^* / \partial f$, $\partial M / \partial \tau_m$, and $\partial z / \partial \tau_m$ can be straightforwardly carried out and are omitted therefore.

B.2 Derivatives of $D_{\max}$

For the damage constraint f_1 in Equation (31), the derivatives of D_{max} are required in Equation $(42)_1$ with F being taken as f_1.

B.2.1. *Derivative* $\partial D_{\max} / \partial \boldsymbol{\rho}$

$$\frac{\partial D_{max}}{\partial \boldsymbol{\rho}} = \boldsymbol{0}. \tag{B56}$$

B.2.2. Derivative $\partial D_{\max} / \partial \boldsymbol{u}^k$

$$\frac{\partial D_{max}}{\partial \boldsymbol{u}^k} = \boldsymbol{0}. \tag{B57}$$

B.2.3. Derivative $\partial D_{max} / \partial \boldsymbol{v}^k$

By definition, it is clear that $\partial D_{max} / \partial \boldsymbol{v}^k = \boldsymbol{0}$ for $k \neq n$, while for $k = n$,

$$\frac{\partial D_{max}}{\partial \boldsymbol{v}^n} = \left[\frac{\partial D_{max}}{\partial \boldsymbol{v}_1^n} \quad \frac{\partial D_{max}}{\partial \boldsymbol{v}_2^n} \quad \cdots \quad \frac{\partial D_{max}}{\partial \boldsymbol{v}_{n_{ele}}^n} \right] \text{ with}$$

$$\frac{\partial D_{max}}{\partial \boldsymbol{v}_e^n} = \left[\frac{\partial D_{max}}{\partial \boldsymbol{v}_{e_1}^n} \quad \frac{\partial D_{max}}{\partial \boldsymbol{v}_{e_2}^n} \quad \cdots \quad \frac{\partial D_{max}}{\partial \boldsymbol{v}_{e_{(n_{ipt})}}^n} \right], e = 1, \ldots, n_{ele}, \tag{B58}$$

where for $s = 1, \ldots, n_{ipt}$,

$$\frac{\partial D_{max}}{\partial \boldsymbol{v}_{e_s}^n} = \left[\frac{\partial D_{max}}{\partial \Delta \gamma_{e_s}^n} \quad \frac{\partial D_{max}}{\partial \alpha_{e_s}^n} \quad \frac{\partial D_{max}}{\partial f_{e_s}^n} \quad \frac{\partial D_{max}}{\partial \left(\boldsymbol{b}^e \right)_{e_s}^n} \right] \tag{B59}$$

with

$$\frac{\partial D_{max}}{\partial \Delta \gamma_{e_s}^n} = 0, \frac{\partial D_{max}}{\partial \alpha_{e_s}^n} = 0, \frac{\partial D_{max}}{\partial \left(\boldsymbol{b}^e \right)_{e_s}^n} = \boldsymbol{0} \tag{B60}$$

and

$$\frac{\partial D_{max}}{\partial f_{e_s}^n} = \frac{1}{p_d}\left(\sum_{e=1}^{n_{ele}}\left(\sum_{s=1}^{n_{ipt}}\left(f_{e_s}^n\right)^{p_d}\right)\right)^{\frac{1}{p_d}-1}\cdot\left(p_d\left(f_{e_s}^n\right)^{p_d-1}\right). \tag{B61}$$

B.3 Derivatives of $\boldsymbol{R^k}$

B.3.1. *Derivative* $\partial \boldsymbol{R^k}/\partial\boldsymbol{\rho}$

The derivative $\partial \boldsymbol{R}^k/\partial\boldsymbol{\rho}$ is a sparse matrix that needs assembly of the element term $\partial \boldsymbol{R}_e^k/\partial\rho_e$ $(e = 1, \dots, n_{ele})$ with the understanding of $\partial \boldsymbol{R}_e^k/\partial\rho_j = \boldsymbol{0}$ for $e \neq j$,

$$\frac{\partial \boldsymbol{R}_e^k}{\partial \rho_e} = \sum_{s=1}^{n_{ipt}}\left(\frac{\partial \eta_e}{\partial \rho_e}\boldsymbol{B}_{e_s}^T\boldsymbol{P}_{e_s}^k + \eta_e\boldsymbol{B}_{e_s}^T\frac{\partial \boldsymbol{P}_{e_s}^k}{\partial \rho_e}\right)w_{e_s} + \sum_{s=1}^{n_{ipt}}\left(\left(1-\eta_e^2\right)\boldsymbol{B}_{L,e_s}^T\left[\frac{\partial \mathbb{C}}{\partial \rho_e} : \boldsymbol{\varepsilon}_{e_s}^k\right] - 2\eta_e\frac{\partial \eta_e}{\partial \rho_e}\boldsymbol{B}_{L,e_s}^T\left[\mathbb{C} : \boldsymbol{\varepsilon}_{e_s}^k\right]\right)w_{e_s} \tag{B62}$$

with

$$\frac{\partial \boldsymbol{P}}{\partial \rho} = -\frac{1}{2}r^{-3/2}\frac{\partial r}{\partial \rho}\overline{\boldsymbol{P}} + r^{-1/2}\frac{\partial \overline{\boldsymbol{P}}}{\partial \rho} \text{ and } \frac{\partial \overline{\boldsymbol{P}}}{\partial \rho} = \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \rho}.\overline{\boldsymbol{F}}^{-T} + \overline{\boldsymbol{\tau}}.\frac{\partial \overline{\boldsymbol{F}}^{-T}}{\partial \rho} \tag{B63}$$

which are understood as being evaluated at sth quadrature point in eth element. Here, an overbar is used to denote terms that are evaluated based on $\overline{\boldsymbol{F}}$ (see Section 3).

B.3.2. *Derivatives* $\partial \boldsymbol{R^k}/\partial \boldsymbol{u^k}$ *and* $\partial \boldsymbol{R^k}/\partial \boldsymbol{u^{k-1}}$

It is clear that

$$\frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{u}^{k-1}} = \boldsymbol{0}. \tag{B64}$$

For $\partial \boldsymbol{R}^k/\partial \boldsymbol{u}^k$, it takes standard finite element assembly with the element term evaluated as

$$\frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{u}_e^k} = \sum_{s=1}^{n_{ipt}}\eta_e\boldsymbol{B}_{e_s}^T\frac{\partial \boldsymbol{P}_{e_s}^k}{\partial \boldsymbol{u}_e^k}w_{e_s} + \sum_{s=1}^{n_{ipt}}\left(1-\eta_e^2\right)\boldsymbol{B}_{L,e_s}^T[\mathbb{C}]\boldsymbol{B}_{L,e_s}w_{e_s} \tag{B65}$$

with

$$\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{u}_e} = -\frac{1}{2}r^{-\frac{3}{2}}\overline{\boldsymbol{P}}\otimes\frac{\partial r}{\partial \boldsymbol{u}_e} + r^{-\frac{1}{2}}\frac{\partial \overline{\boldsymbol{P}}}{\partial \boldsymbol{u}_e}, \tag{B66}$$

$$\frac{\partial \overline{\boldsymbol{P}}}{\partial \boldsymbol{u}_e} = \left.\frac{\partial \overline{\boldsymbol{P}}}{\partial \overline{\boldsymbol{F}}}\right|_{\overline{\boldsymbol{\tau}}\ fixed} : \frac{\partial \overline{\boldsymbol{F}}}{\partial \boldsymbol{u}_e} \text{ with } \left.\frac{\partial \overline{\boldsymbol{P}}}{\partial \overline{\boldsymbol{F}}}\right|_{\overline{\boldsymbol{\tau}}\ fixed} = -\left(\overline{\boldsymbol{\tau}}.\overline{\boldsymbol{F}}^{-T}\right)\boxdot\overline{\boldsymbol{F}}^{-1}, \tag{B67}$$

where, again, subscripts/superscript are omitted for brevity. Here $\overline{\boldsymbol{\tau}}$ depends exclusively on $\boldsymbol{b}^e$ and ρ, see Equations (6) and (8).

B.3.3. *Derivatives* $\partial \boldsymbol{R^k}/\partial \boldsymbol{v^k}$ *and* $\partial \boldsymbol{R^k}/\partial \boldsymbol{v^{k-1}}$

First, obviously

$$\frac{\partial \boldsymbol{R}^k}{\partial \boldsymbol{v}^{k-1}} = \boldsymbol{0}. \tag{B68}$$

Second, the derivative $\partial \boldsymbol{R}^k/\partial \boldsymbol{v}^k$ is a sparse matrix that requires assembly of the element term $\partial \boldsymbol{R}_e^k/\partial \boldsymbol{v}_e^k$ $(e = 1, \dots, n_{ele})$ while $\partial \boldsymbol{R}_e^k/\partial \boldsymbol{v}_j^k = \boldsymbol{0}$ for $e \neq j$. The nonszero term is derived as

$$\frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{v}_e^k} = \left[\frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{v}_{e_1}^k} \quad \frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{v}_{e_2}^k} \quad \cdots \quad \frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{v}_{e_{(n_{ipt})}}^k}\right],$$

$$\frac{\partial \boldsymbol{R}_e^k}{\partial \boldsymbol{v}_{e_s}^k} = \left[\mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad \frac{\partial \boldsymbol{R}_e^k}{\partial (\boldsymbol{b}^e)_{e_s}^k}\right], s = 1, \ldots, n_{ipt},$$

$$\frac{\partial \boldsymbol{R}_e^k}{\partial (\boldsymbol{b}^e)_{e_s}^k} = \eta_e \left(r_{e_s}^k\right)^{-1/2} \boldsymbol{B}_{e_s}^T \frac{\partial \overline{\boldsymbol{P}}_{e_s}^k}{\partial (\boldsymbol{b}^e)_{e_s}^k} w_{e_s}, \tag{B69}$$

$$\frac{\partial \overline{\boldsymbol{P}}}{\partial \boldsymbol{b}^e} = \frac{\partial \overline{\boldsymbol{P}}}{\partial \overline{\boldsymbol{\tau}}} : \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \boldsymbol{b}^e} \text{ with } \frac{\partial \overline{\boldsymbol{P}}}{\partial \overline{\boldsymbol{\tau}}} = \frac{1}{2}\left(\boldsymbol{I} \boxtimes \overline{\boldsymbol{F}}^{-1} + \boldsymbol{I} \boxdot \overline{\boldsymbol{F}}^{-1}\right).$$

B.4 Derivatives of H^k

This section gives the derivatives only for the plastic step, see (44), while for the elastic step, the calculation of the derivatives is straightforward.

B.4.1. *Derivative* $\partial H^k / \partial \rho$

$$\frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{\rho}} = \begin{bmatrix} \frac{\partial \boldsymbol{H}_1^k}{\partial \rho_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\partial \boldsymbol{H}_{n_{ele}}^k}{\partial \rho_{n_{ele}}} \end{bmatrix} \text{ with } \frac{\partial \boldsymbol{H}_e^k}{\partial \rho_e} = \begin{bmatrix} \frac{\partial \boldsymbol{H}_{e_1}^k}{\partial \rho_e} \\ \frac{\partial \boldsymbol{H}_{e_2}^k}{\partial \rho_e} \\ \vdots \\ \frac{\partial \boldsymbol{H}_e^k(n_{ipt})}{\partial \rho_e} \end{bmatrix}, \frac{\partial \boldsymbol{H}_{e_s}^k}{\partial \rho_e} = \begin{bmatrix} \frac{\partial h_1}{\partial \rho_e} \\ \frac{\partial h_2}{\partial \rho_e} \\ \frac{\partial h_3}{\partial \rho_e} \\ \frac{\partial \boldsymbol{h}_4}{\partial \rho_e} \end{bmatrix}, s = 1, \ldots, n_{ipt} \tag{B70}$$

with

$$\frac{\partial h_1}{\partial \rho_e} = \frac{\partial \tau_d}{\partial \rho_e} - \sqrt{\frac{2}{3}} sign(z) \left(\frac{1}{2} \frac{1}{\sqrt{|z|}} sign(z) \frac{\partial z}{\partial \rho_e} \zeta J + \sqrt{|z|} J \frac{\partial \zeta}{\partial \rho_e} + \sqrt{|z|} \zeta \frac{\partial J}{\partial \rho_e}\right), \tag{B71}$$

$$\frac{\partial h_2}{\partial \rho_e} = \frac{\Delta\gamma}{(1-f)\zeta J}(\tau_d + C_0 \tau_m)\left(\frac{1}{\zeta}\frac{\partial \zeta}{\partial \rho_e} + \frac{1}{J}\frac{\partial J}{\partial \rho_e}\right) - \frac{\Delta\gamma}{(1-f)\zeta J}\left(\frac{\partial \tau_d}{\partial \rho_e} + \frac{\partial C_0}{\partial \rho_e}\tau_m + C_0 \frac{\partial \tau_m}{\partial \rho_e}\right), \tag{B72}$$

$$\begin{aligned} \frac{\partial h_3}{\partial \rho_e} &= -\rho_e^{p_f} \Delta\gamma \left((1-f)\frac{\partial C_0}{\partial \rho_e} + \cancel{k} \frac{D}{(1-f)\zeta J}\left(\frac{\partial \tau_d}{\partial \rho_e} + \frac{\partial C_0}{\partial \rho_e}\tau_m + C_0 \frac{\partial \tau_m}{\partial \rho_e}\right) - \cancel{k} \frac{D}{(1-f)\zeta J}(\tau_d + C_0 \tau_m)\left(\frac{1}{\zeta}\frac{\partial \zeta}{\partial \rho_e} + \frac{1}{J}\frac{\partial J}{\partial \rho_e}\right) + \sqrt{\frac{2}{3}} k_\omega f \frac{\partial \omega_0}{\partial \rho_e}\right) \\ &\quad - p_f \rho_e^{p_f - 1} \Delta\gamma \left((1-f)C_0 + \cancel{k} \frac{D}{(1-f)\zeta J}(\tau_d + C_0 \tau_m) + \sqrt{\frac{2}{3}} k_\omega f \omega_0\right), \end{aligned} \tag{B73}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \rho_e} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \rho_e} . \boldsymbol{b}^e - \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \rho_e}, \tag{B74}$$

where the terms $\partial\tau_d/\partial\rho_e, \partial\tau_m/\partial\rho_e, \partial z/\partial\rho_e, \partial\zeta/\partial\rho_e, \partial J/\partial\rho_e, \partial C_0/\partial\rho_e$ are given in Equations (B6), (B9), (B7), (B4) and (B10), while $\partial\omega_0/\partial\rho_e$, $\partial \exp(2\Delta\gamma \boldsymbol{A})/\partial\rho_e$ and $\partial \boldsymbol{b}^{e^{tr}}/\partial\rho_e$ can be computed as

$$\frac{\partial \omega_0}{\partial \rho_e} = \left(\Omega \frac{\partial \omega}{\partial \overline{\boldsymbol{\tau}}} + \omega \frac{\partial \Omega}{\partial T}\frac{\partial T}{\partial \overline{\boldsymbol{\tau}}}\right) : \frac{\partial \overline{\boldsymbol{\tau}}}{\partial \rho_e}, \tag{B75}$$

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \rho_e} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{A}} : \frac{\partial \boldsymbol{A}}{\partial \rho_e}, \tag{B76}$$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \rho_e} = \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_k} : \frac{\partial \overline{\boldsymbol{F}}_k}{\partial \rho_e} + \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_{k-1}} : \frac{\partial \overline{\boldsymbol{F}}_{k-1}}{\partial \rho_e} \tag{B77}$$

with the help of Equations (B26), (B28), (B6), (B35), (B11), (B20), (B21), and (B3).

B.4.2. *Derivatives $\partial \boldsymbol{H}^k/\partial \boldsymbol{u}^k$ and $\partial \boldsymbol{H}^k/\partial \boldsymbol{u}^{k-1}$*

The derivative matrix $\partial \boldsymbol{H}^k/\partial \boldsymbol{u}^k$ is sparse and can be assembled by the element term $\partial \boldsymbol{H}_e^k/\partial \boldsymbol{u}_e^k$ ($e = 1, \dots, n_{ele}$)

$$\frac{\partial \boldsymbol{H}_e^k}{\partial \boldsymbol{u}_e^k} = \begin{bmatrix} \frac{\partial \boldsymbol{H}_{e_1}^k}{\partial \boldsymbol{u}_e^k} \\ \vdots \\ \frac{\partial \boldsymbol{H}_e^k{}_{(n_{ipt})}}{\partial \boldsymbol{u}_e^k} \end{bmatrix} \text{ with } \frac{\partial \boldsymbol{H}_{e_s}^k}{\partial \boldsymbol{u}_e^k} = \begin{bmatrix} \frac{\partial h_1}{\partial \boldsymbol{u}_e^k} \\ \frac{\partial h_2}{\partial \boldsymbol{u}_e^k} \\ \frac{\partial h_3}{\partial \boldsymbol{u}_e^k} \\ \frac{\partial \boldsymbol{h}_4}{\partial \boldsymbol{u}_e^k} \end{bmatrix}, s = 1, \dots, n_{ipt}, \tag{B78}$$

where

$$\frac{\partial h_1}{\partial \boldsymbol{u}_e^k} = -\sqrt{\frac{2}{3}}\, sign(z)\zeta \left(\frac{1}{2} sign(z) \frac{1}{\sqrt{|z|}} \frac{\partial z}{\partial \boldsymbol{u}_e^k} J + \sqrt{|z|} \frac{\partial J}{\partial \boldsymbol{u}_e^k} \right), \tag{B79}$$

$$\frac{\partial h_2}{\partial \boldsymbol{u}_e^k} = \frac{\Delta\gamma}{(1-f)\zeta J^2} (\tau_d + C_0\tau_m) \frac{\partial J}{\partial \boldsymbol{u}_e^k} - \frac{\Delta\gamma\tau_m}{(1-f)\zeta J} \frac{\partial C_0}{\partial \boldsymbol{u}_e^k}, \tag{B80}$$

$$\frac{\partial h_3}{\partial \boldsymbol{u}_e^k} = -\rho^{p_f}\Delta\gamma \left((1-f)\frac{\partial C_0}{\partial \boldsymbol{u}_e^k} - \frac{\hbar D}{(1-f)\zeta J^2} (\tau_d + C_0\tau_m) \frac{\partial J}{\partial \boldsymbol{u}_e^k} + \frac{\hbar D\tau_m}{(1-f)\zeta J} \frac{\partial C_0}{\partial \boldsymbol{u}_e^k} \right), \tag{B81}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \boldsymbol{u}_e^k} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{u}_e^k} \circledcirc \boldsymbol{b}^e - \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{u}_e^k}, \tag{B82}$$

in which $\partial z/\partial \boldsymbol{u}_e^k$, $\partial J/\partial \boldsymbol{u}_e^k$ and $\partial C_0/\partial \boldsymbol{u}_e^k$, are given in Equations (B17), (B15) and (B18), respectively, and the operation $\circledcirc$ is defined such that for any third-order tensor $\boldsymbol{\mathcal{A}}$ and second-order tensor $\boldsymbol{B}$,

$$\boldsymbol{\mathcal{A}} \circledcirc \boldsymbol{B} := \mathcal{A}_{imk} B_{mj} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k. \tag{B83}$$

The derivatives $\partial \exp(2\Delta\gamma \boldsymbol{A})/\partial \boldsymbol{u}_e^k$ and $\partial \boldsymbol{b}^{e^{tr}}/\partial \boldsymbol{u}_e^k$ in Equation (B82) are obtained using chain rule by

$$\frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{u}_e^k} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \boldsymbol{A}} : \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{u}_e^k}, \tag{B84}$$

$$\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{u}_e^k} = \frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_k} : \frac{\partial \overline{\boldsymbol{F}}_k}{\partial \boldsymbol{u}_e^k} \tag{B85}$$

with the help of Equations (B35), (B19), (B20), and (B14).

Similarly, the sparse derivative matrix $\partial \boldsymbol{H}^k/\partial \boldsymbol{u}^{k-1}$ is assembled from the elementary term $\partial \boldsymbol{H}_{e_s}^k/\partial \boldsymbol{u}_e^{k-1}$ ($e = 1, \dots, n_{ele}$ and $s = 1, \dots, n_{ipt}$)

$$\frac{\partial \boldsymbol{H}_{e_s}^k}{\partial \boldsymbol{u}_e^{k-1}} = \begin{bmatrix} \frac{\partial h_1}{\partial \boldsymbol{u}_e^{k-1}} \\ \frac{\partial h_2}{\partial \boldsymbol{u}_e^{k-1}} \\ \frac{\partial h_3}{\partial \boldsymbol{u}_e^{k-1}} \\ \frac{\partial \boldsymbol{h}_4}{\partial \boldsymbol{u}_e^{k-1}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \frac{\partial \boldsymbol{h}_4}{\partial \boldsymbol{u}_e^{k-1}} \end{bmatrix} \text{ with } \frac{\partial \boldsymbol{h}_4}{\partial \boldsymbol{u}_e^{k-1}} = -\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \boldsymbol{u}_e^{k-1}} = -\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \overline{\boldsymbol{F}}_{k-1}} : \frac{\partial \overline{\boldsymbol{F}}_{k-1}}{\partial \boldsymbol{u}_e^{k-1}}, \tag{B86}$$

where the terms $\partial \boldsymbol{b}^{e^{tr}}/\partial \overline{\boldsymbol{F}}_{k-1}$ and $\partial \overline{\boldsymbol{F}}_{k-1}/\partial \boldsymbol{u}_e^{k-1}$ can be found in Equations (B21) and (B14).

B.4.3. *Derivatives $\partial \boldsymbol{H}^k/\partial \boldsymbol{v}^k$ and $\partial \boldsymbol{H}^k/\partial \boldsymbol{v}^{k-1}$*

First, the derivative matrix $\partial \boldsymbol{H}^k/\partial \boldsymbol{v}^k$ is calculated as

$$\frac{\partial \boldsymbol{H}^k}{\partial \boldsymbol{v}^k} = \begin{bmatrix} \frac{\partial \boldsymbol{H}_1^k}{\partial \boldsymbol{v}_1^k} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \frac{\partial \boldsymbol{H}_{n_{ele}}^k}{\partial \boldsymbol{v}_{n_{ele}}^k} \end{bmatrix} \text{ with } \frac{\partial \boldsymbol{H}_e^k}{\partial \boldsymbol{v}_e^k} = \begin{bmatrix} \frac{\partial \boldsymbol{H}_{e_1}^k}{\partial \boldsymbol{v}_{e_1}^k} & & & \\ & \frac{\partial \boldsymbol{H}_{e_2}^k}{\partial \boldsymbol{v}_{e_2}^k} & & \\ & & \ddots & \\ & & & \frac{\partial \boldsymbol{H}_e^k{}_{(n_{ipt})}}{\partial \boldsymbol{v}_e^k{}_{(n_{ipt})}} \end{bmatrix}, \tag{B87}$$

for $e = 1, \ldots, n_{ele}$, and

$$\frac{\partial \boldsymbol{H}_{e_s}^k}{\partial \boldsymbol{v}_{e_s}^k} = \begin{bmatrix} \frac{\partial h_1}{\partial \Delta\gamma_{e_s}^k} & \frac{\partial h_1}{\partial \alpha_{e_s}^k} & \frac{\partial h_1}{\partial f_{e_s}^k} & \frac{\partial h_1}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \\ \frac{\partial h_2}{\partial \Delta\gamma_{e_s}^k} & \frac{\partial h_2}{\partial \alpha_{e_s}^k} & \frac{\partial h_2}{\partial f_{e_s}^k} & \frac{\partial h_2}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \\ \frac{\partial h_3}{\partial \Delta\gamma_{e_s}^k} & \frac{\partial h_3}{\partial \alpha_{e_s}^k} & \frac{\partial h_3}{\partial f_{e_s}^k} & \frac{\partial h_3}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \\ \frac{\partial \boldsymbol{h}_4}{\partial \Delta\gamma_{e_s}^k} & \frac{\partial \boldsymbol{h}_4}{\partial \alpha_{e_s}^k} & \frac{\partial \boldsymbol{h}_4}{\partial f_{e_s}^k} & \frac{\partial \boldsymbol{h}_4}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \end{bmatrix}, s = 1, \ldots, n_{ipt} \tag{B88}$$

with

$$\frac{\partial h_1}{\partial \Delta\gamma_{e_s}^k} = 0, \tag{B89}$$

$$\frac{\partial h_1}{\partial \alpha_{e_s}^k} = -\frac{1}{\sqrt{6}} \frac{1}{\sqrt{|z|}} \frac{\partial z}{\partial \alpha_{e_s}^k} \zeta J - \sqrt{\frac{2}{3}}\, sign(z) \sqrt{|z|}\, J \frac{\partial \zeta}{\partial \alpha_{e_s}^k}, \tag{B90}$$

$$\frac{\partial h_1}{\partial f_{e_s}^k} = -\frac{1}{\sqrt{6}} \frac{1}{\sqrt{|z|}} \zeta J \frac{\partial z}{\partial f_{e_s}^k}, \tag{B91}$$

$$\frac{\partial h_1}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} = \frac{\partial \tau_d}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} - \frac{1}{\sqrt{6}} \frac{1}{\sqrt{|z|}} \zeta J \frac{\partial z}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k}, \tag{B92}$$

$$\frac{\partial h_2}{\partial \Delta\gamma_{e_s}^k} = -\frac{1}{(1-f)\zeta J} \left(\tau_d + C_0 \tau_m\right), \tag{B93}$$

$$\frac{\partial h_2}{\partial \alpha_{e_s}^k} = 1 + \frac{\Delta\gamma}{(1-f)\zeta^2 J} \left(\tau_d + C_0 \tau_m\right) \frac{\partial \zeta}{\partial \alpha_{e_s}^k} - \frac{\Delta\gamma \tau_m}{(1-f)\zeta J} \frac{\partial C_0}{\partial \alpha_{e_s}^k}, \tag{B94}$$

$$\frac{\partial h_2}{\partial f_{e_s}^k} = -\frac{\Delta\gamma}{(1-f)^2 \zeta J} \left(\tau_d + C_0 \tau_m\right) - \frac{\Delta\gamma \tau_m}{(1-f)\zeta J} \frac{\partial C_0}{\partial f_{e_s}^k}, \tag{B95}$$

$$\frac{\partial h_2}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} = -\frac{\Delta\gamma}{(1-f)\zeta J} \left(\frac{\partial \tau_d}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} + \frac{\partial C_0}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \tau_m + C_0 \frac{\partial \tau_m}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \right), \tag{B96}$$

$$\frac{\partial h_3}{\partial \Delta\gamma_{e_s}^k} = -\rho^{p_f} \left((1-f) C_0 + \kappa \frac{\mathcal{D}}{(1-f)\zeta J} \left(\tau_d + C_0 \tau_m\right) + \sqrt{\frac{2}{3}} k_\omega f \omega_0 \right), \tag{B97}$$

$$\frac{\partial h_3}{\partial \alpha_{e_s}^k} = -\rho^{p_f} \Delta\gamma \left((1-f) \frac{\partial C_0}{\partial \alpha_{e_s}^k} + \kappa \frac{1}{(1-f)\zeta J} \left(\tau_d + C_0 \tau_m\right) \frac{\partial \mathcal{D}}{\partial \alpha_{e_s}^k} - \kappa \frac{\mathcal{D}}{(1-f)\zeta^2 J} \left(\tau_d + C_0 \tau_m\right) \frac{\partial \zeta}{\partial \alpha_{e_s}^k} + \kappa \frac{\mathcal{D}}{(1-f)\zeta J} \tau_m \frac{\partial C_0}{\partial \alpha_{e_s}^k} \right), \tag{B98}$$

$$\frac{\partial h_3}{\partial f_{e_s}^k} = 1 - \rho^{p_f} \Delta\gamma \left(-C_0 + (1-f) \frac{\partial C_0}{\partial f_{e_s}^k} + \kappa \frac{\mathcal{D}}{(1-f)^2 \zeta J} \left(\tau_d + C_0 \tau_m\right) + \kappa \frac{\mathcal{D}}{(1-f)\zeta J} \tau_m \frac{\partial C_0}{\partial f_{e_s}^k} + \sqrt{\frac{2}{3}} k_\omega \omega_0 \right), \tag{B99}$$

$$\frac{\partial h_3}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} = -\rho^{p_f} \Delta\gamma \left((1-f) \frac{\partial C_0}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} + \frac{\kappa \mathcal{D}}{(1-f)\zeta J} \left(\frac{\partial \tau_d}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} + \frac{\partial C_0}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \tau_m + C_0 \frac{\partial \tau_m}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \right) + \sqrt{\frac{2}{3}} k_\omega f \frac{\partial \omega_0}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \right), \tag{B100}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \Delta\gamma_{e_s}^k} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \Delta\gamma_{e_s}^k} . \boldsymbol{b}^e, \tag{B101}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \alpha_{e_s}^k} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \alpha_{e_s}^k} . \boldsymbol{b}^e, \tag{B102}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial f_{e_s}^k} = \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial f_{e_s}^k}.\boldsymbol{b}^e, \tag{B103}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} = \frac{1}{2}[exp(2\Delta\gamma \boldsymbol{A}) \boxtimes \boldsymbol{I} + exp(2\Delta\gamma \boldsymbol{A}) \boxdot \boldsymbol{I}] + \frac{\partial exp(2\Delta\gamma \boldsymbol{A})}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^k} \boxplus \left(\boldsymbol{b}^e\right)_{e_s}^k, \tag{B104}$$

where the operation $\boxplus$ is defined such that for any fourth-order tensor $\mathbb{A}$ and second-order tensor $\boldsymbol{B}$,

$$\mathbb{A} \boxplus \boldsymbol{B} := \mathbb{A}_{imkl} B_{mj} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k \otimes \boldsymbol{e}_l \tag{B105}$$

Second, the derivative matrix $\partial \boldsymbol{H}^k / \partial \boldsymbol{v}^{k-1}$ is calculated in a similar manner as in Equation (B87) but with

$$\frac{\partial \boldsymbol{H}_{e_s}^k}{\partial \boldsymbol{v}_{e_s}^k} = \begin{bmatrix} \frac{\partial h_1}{\partial \Delta\gamma_{e_s}^{k-1}} & \frac{\partial h_1}{\partial \alpha_{e_s}^{k-1}} & \frac{\partial h_1}{\partial f_{e_s}^{k-1}} & \frac{\partial h_1}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} \\ \frac{\partial h_2}{\partial \Delta\gamma_{e_s}^{k-1}} & \frac{\partial h_2}{\partial \alpha_{e_s}^{k-1}} & \frac{\partial h_2}{\partial f_{e_s}^{k-1}} & \frac{\partial h_2}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} \\ \frac{\partial h_3}{\partial \Delta\gamma_{e_s}^{k-1}} & \frac{\partial h_3}{\partial \alpha_{e_s}^{k-1}} & \frac{\partial h_3}{\partial f_{e_s}^{k-1}} & \frac{\partial h_3}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} \\ \frac{\partial \boldsymbol{h}_4}{\partial \Delta\gamma_{e_s}^{k-1}} & \frac{\partial \boldsymbol{h}_4}{\partial \alpha_{e_s}^{k-1}} & \frac{\partial \boldsymbol{h}_4}{\partial f_{e_s}^{k-1}} & \frac{\partial \boldsymbol{h}_4}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} \end{bmatrix}, \tag{B106}$$

where

$$\frac{\partial h_1}{\partial \Delta\gamma_{e_s}^{k-1}} = 0, \frac{\partial h_1}{\partial \alpha_{e_s}^{k-1}} = 0, \frac{\partial h_1}{\partial f_{e_s}^{k-1}} = 0, \frac{\partial h_1}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} = \boldsymbol{0}, \tag{B107}$$

$$\frac{\partial h_2}{\partial \Delta\gamma_{e_s}^{k-1}} = 0, \frac{\partial h_2}{\partial \alpha_{e_s}^{k-1}} = -1, \frac{\partial h_2}{\partial f_{e_s}^{k-1}} = 0, \frac{\partial h_2}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} = \boldsymbol{0} \tag{B108}$$

$$\frac{\partial h_3}{\partial \Delta\gamma_{e_s}^{k-1}} = 0, \frac{\partial h_3}{\partial \alpha_{e_s}^{k-1}} = 0, \frac{\partial h_3}{\partial f_{e_s}^{k-1}} = -1, \frac{\partial h_3}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} = \boldsymbol{0}, \tag{B109}$$

$$\frac{\partial \boldsymbol{h}_4}{\partial \Delta\gamma_{e_s}^{k-1}} = \boldsymbol{0}, \frac{\partial \boldsymbol{h}_4}{\partial \alpha_{e_s}^{k-1}} = \boldsymbol{0}, \frac{\partial \boldsymbol{h}_4}{\partial f_{e_s}^{k-1}} = \boldsymbol{0}, \frac{\partial \boldsymbol{h}_4}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}} = -\frac{\partial \boldsymbol{b}^{e^{tr}}}{\partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}}, \tag{B110}$$

in which the term $\partial \boldsymbol{b}^{e^{tr}} / \partial \left(\boldsymbol{b}^e\right)_{e_s}^{k-1}$ is given in Equation (B22).