Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 325 (2017) 463–487
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Stochastic sampling for deterministic structural topology optimization with many load cases: Density-based and ground structure approaches

Xiaojia Shelly Zhang[a], Eric de Sturler[b], Glaucio H. Paulino[a,*]

[a] School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA, 30332, USA
[b] Department of Mathematics, Virginia Tech, McBryde Hall, 225 Stanger Street, Blacksburg, VA, 24061, USA

Received 5 July 2016; received in revised form 29 June 2017; accepted 30 June 2017
Available online 6 July 2017

Highlights

- Randomized approach for deterministic topology optimization with many load cases.
- Approach solves a few linear systems per optimization step rather than hundreds.
- Approach leads to quality solutions with drastically reduced computational costs.
- Includes a damping scheme (for randomization) based on simulated annealing.
- Results show that the scheme is effective and leads to rapid convergence.

Abstract

We propose an efficient probabilistic method to solve a fully deterministic problem — we present a randomized optimization approach that drastically reduces the enormous computational cost of optimizing designs under many load cases for both continuum and truss topology optimization. Practical structural designs by deterministic topology optimization typically involve many load cases, possibly hundreds or more. The optimal design minimizes a, possibly weighted, average of the compliance under each load case (or some other objective). This means that, in each optimization step, a large finite element problem must be solved for each load case, leading to an enormous computational effort. On the contrary, the proposed randomized optimization method with stochastic sampling requires the solution of only a few (e.g., 5 or 6) finite element problems (large linear systems) per optimization step. Based on simulated annealing, we introduce a damping scheme for the randomized approach. Through numerical examples in two and three dimensions, we demonstrate that the randomization algorithm drastically reduces computational cost to obtain

* Corresponding author.
E-mail addresses: xzhang645@gatech.edu (X.S. Zhang), sturler@vt.edu (E. de Sturler), paulino@gatech.edu (G.H. Paulino).

http://dx.doi.org/10.1016/j.cma.2017.06.035

similar final topologies and results (e.g., compliance) to those of standard algorithms. The results indicate that the damping scheme is effective and leads to rapid convergence of the proposed algorithm.

Keywords: Topology optimization with many load cases; Stochastic sampling; Randomized algorithm; Trace estimator; Density-based method; Ground structure method

1. Introduction

Structural topology optimization is an important tool that, if properly used, can lead to significantly improved designs. In the field of structural topology optimization, designs accounting for many load cases are common practice [1]. Indeed, real-world structural designs, for example, high-rise buildings and long-span bridges, generally involve numerous load cases [1]. For the end-compliance minimization formulation, several methods that include many load cases have been published [2–4]. The main idea is to minimize a proper norm of the vector of compliances from all load cases. This paper concentrates on one popular method, the weighted sum formulation that minimizes a weighted sum of the vector of compliances from all load cases, which requires the solution of the structural equation for each load case. Another formulation that accounts for many load cases is the min–max formulation, which minimizes the worst-case compliance of all load cases, i.e., the infinity norm of the vector of compliances [4,5]. This formulation also requires the solution of a large system of linear equations for each load case. Therefore, in terms of computational cost, it is similar to the aforementioned weighted sum formulation. In this paper, we propose a randomized algorithm that drastically reduces the computational cost. Similar techniques have been applied to solve inverse problems – see [6–8] and references therein, and least squares problems – see [9] and references therein. More general randomization techniques, especially randomized linear algebra, such as trace estimation and matrix decomposition, are discussed and surveyed in [10–13]. In [14], Avron and Toledo discussed and derived bounds on several trace estimators and proved a bound on the number of samples required to guarantee a chosen accuracy. Their results were improved and extended by Roosta-Khorasani and Ascher [15]. More recently, Saibaba et al. [16] presented randomized algorithms for estimates of the trace and determinant of positive semidefinite Hermitian matrices. Their algorithms yield more accurate estimates; but the estimates are not unbiased. *We emphasize that the use of random samples in our algorithm does not reflect any uncertainty in the load cases but instead arises from a stochastic optimization approach that avoids solving for a large number of load cases in each optimization step. In our proposed algorithm, randomization is used to solve a deterministic optimization problem with deterministic loads that remain fixed throughout the optimization process.* Thus, our randomized algorithm is conceptually different from robust topology optimization approaches (see, e.g., [17–19]), in which the uncertainty of the loads is characterized by probability distributions.

The remainder of the paper is organized as follows. We finish this section with a brief review of the weighted sum formulation of the standard nested minimum end-compliance topology optimization with many load cases, using both the density-based method and the ground structure method (GSM). Section 2 reviews the stochastic sampling techniques we use to estimate the trace of a general matrix and proposes the randomized algorithm for minimum end-compliance topology optimization under many load cases. Section 3 discusses the algorithmic parameters and introduces a damping scheme for the randomized algorithm. Section 4 presents numerical examples in two- and three-dimensions highlighting the efficiency and effectiveness of the proposed algorithm, and Section 5 provides concluding remarks with suggestions for extending the work.

1.1. Density-based topology optimization formulation with many load cases

For a set of m given load cases $\boldsymbol{f}_i, i = 1, \ldots, m$, the standard weighted sum formulation for the minimum end-compliance design with many load cases using the density-based method can be written as [4],

$$\begin{aligned} \min_{\boldsymbol{\rho}} C(\boldsymbol{\rho}) &= \min_{\boldsymbol{\rho}} \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i{}^T \boldsymbol{u}_i(\boldsymbol{\rho}), \\ \text{s.t.} \quad & \sum_{e=1}^{M} v^{(e)} \bar{\rho}^{(e)} - V_{\max} \le 0, \\ & 0 < \rho^{\min} \le \rho^{(e)} \le 1,\ e = 1, \ldots, M, \\ & \text{with } \boldsymbol{u}_i(\bar{\boldsymbol{\rho}}) = \boldsymbol{K}(\boldsymbol{E}(\bar{\boldsymbol{\rho}}))^{-1} \boldsymbol{f}_i,\ i = 1, \ldots, m. \end{aligned} \tag{1}$$

In this minimization problem, the objective function C is the weighted-average compliance of the corresponding structure, $\boldsymbol{\rho} \in \mathbb{R}^M$ is the vector of design variables (the density field), and M is the number of elements in the finite element mesh. We define $\bar{\boldsymbol{\rho}}$ and $\boldsymbol{H}$ as the filtered physical density and the filter matrix such that $\bar{\boldsymbol{\rho}} = \boldsymbol{H}\boldsymbol{\rho}$ [20]. In order to ensure a positive definite stiffness matrix $\boldsymbol{K} \in \mathbb{R}^{dN \times dN}$, a lower bound $\rho^{\min}$ is prescribed on $\rho^{(e)}$, where N is the number of nodes and d is the dimension of the problem, so dN is the number of degrees of freedom. The volume (area) of element e is given by $v^{(e)}$, and $V_{\max}$ is the prescribed upper bound on the total volume. The weight and the displacement vector associated with load case $\boldsymbol{f}_i \in \mathbb{R}^{dN}$ are given by α_i ($\alpha_i > 0$, $\sum_{i=1}^{m} \alpha_i = 1$) and $\boldsymbol{u}_i \in \mathbb{R}^{dN}$, respectively. The Young's modulus $\boldsymbol{E}$ is defined by, for example, the Solid Isotropic Material with Penalization (SIMP) [21,22] approach. Other models, e.g., RAMP (Rational Approximation of Material Properties) [4,23], can be used and would not alter the conceptual presentation of the topic. For the SIMP approach with the density filter, we have $E^{(e)} = [\bar{\rho}^{(e)}]^p (E_0)$, where E_0 is the elastic modulus for solid material and p is a penalization factor. The gradient (sensitivity) of the objective function corresponds to a weighted sum of the sensitivities of each individual loading case, which can be expressed as

$$\nabla C_{\rho}^{(e)} = \frac{\partial C}{\partial \rho^{(e)}} = -\sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i{}^T \frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}} \boldsymbol{u}_i. \tag{2}$$

The formulation (1) is known to be convex when $p = 1$ [24]. Using $p > 1$ to obtain a solid-void solution, one makes the problem non-convex and, as expected, the solution obtained from the optimization algorithm may not be the global minimum.

1.2. Ground-structure based formulation with many load cases

The standard weighted sum formulation for the minimum end-compliance design with the ground structure method (GSM) under many load cases takes the following form,

$$\begin{aligned} \min_{\boldsymbol{x}} C(\boldsymbol{x}) &= \min_{\boldsymbol{x}} \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i{}^T \boldsymbol{u}_i(\boldsymbol{x}), \\ \text{s.t.} \quad & \sum_{e=1}^{M} L^{(e)} x^{(e)} - V_{\max} \le 0, \\ & 0 < x^{\min} \le x^{(e)} \le x^{\max},\ e = 1, \ldots, M, \\ & \text{with } \boldsymbol{u}_i(\boldsymbol{x}) = \boldsymbol{K}(\boldsymbol{x})^{-1} \boldsymbol{f}_i,\ i = 1, \ldots, m. \end{aligned} \tag{3}$$

The vector $\boldsymbol{x} \in \mathbb{R}^M$ is a vector of design variables, with component $x^{(e)}$ being the cross-sectional area of truss member e. It is subjected to lower bound $x^{\min}$ and upper bound $x^{\max}$. Furthermore, M is the number of truss members in the ground structure (GS), $L^{(e)}$ is the length of truss member e, and $V_{\max}$ is the prescribed upper bound on the total volume. For the ith load case $\boldsymbol{f}_i \in \mathbb{R}^{dN}$, α_i and $\boldsymbol{u}_i \in \mathbb{R}^{dN}$ are the corresponding weight factor and displacement vector. As in the density-based method, the sensitivity of the objective function in the GSM is the weighted sum of the sensitivities from each load case:

$$\nabla C_{x}^{(e)} = \frac{\partial C}{\partial x^{(e)}} = -\sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i{}^T \frac{\partial \boldsymbol{K}}{\partial x^{(e)}} \boldsymbol{u}_i. \tag{4}$$

The standard nested formulation of the end-compliance objective function with a single load case has been proven to be convex in [25] for a positive definite stiffness matrix and in [26] for a positive semi-definite stiffness matrix. The formulation (3) with multiple load cases is easily proven to be convex.

1.3. Synopsis

The optimization algorithm contains three main components: solving the structural equilibrium problem for a set of given design variables, computing the gradient of the objective function, and updating the design variables. In this paper, we use the density-based and GS methods — both methods utilize the Optimality Criterion (OC) algorithm [27] as the update scheme, which only requires gradient information of the objective function and the volume constraint. The convergence criterion for optimization used in formulations (1) and (3) is that the maximum change in design variables drops below a given tolerance, τ_{opt}.

The overall computational cost of the standard optimization formulations (1) and (3) can be estimated by the total number of structural equations (large linear system) solves, i.e., $m \times N_{\text{step}}$, where N_{step} is the number of optimization steps. For realistic three-dimensional (3D) problems, where the mesh size, the number of design variables, and the number of load cases are large, the associated computational cost is enormous. Thus, in this paper, we propose a randomized approach that reduces the problem with many load cases (m) to a sequence of optimization steps with only a few load cases ($n \ll m$) to solve per step and that yields results similar to those from the standard weighted sum approach.

2. Stochastic sampling and topology optimization

This section proposes a stochastic sampling approach for the minimum end-compliance topology optimization formulations with many load cases. First, we briefly review a stochastic sampling technique to estimate the trace of a general matrix. Then, the stochastic sampling technique is applied to the minimum end-compliance topology optimization with many load cases using both density-based and GS methods.

2.1. Stochastic sampling of matrices

For a general matrix $\boldsymbol{A} \in \mathbb{R}^{q\times q}$, stochastic sampling techniques can be used to estimate the trace of $\boldsymbol{A}$. We discuss one popular approach here, the Hutchinson trace estimator [10], but several alternatives exist, e.g., the Gaussian estimators and unit vector estimators—see [14,15] and references therein. Let $\boldsymbol{\xi} \in \mathbb{R}^q$ be a random vector containing entries that are independent and identically distributed (i.i.d.) with value ± 1 each with probability $1/2$. This distribution is known as the Rademacher distribution. It follows immediately that, for each entry, the expectation $\mathbb{E}(\xi_i) = 0$. Since the entries are independent, the expectation of $\xi_i \xi_j$ is given by,

$$\mathbb{E}\left(\xi_i \xi_j\right) = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \tag{5}$$

Now consider the expectation of the random variable $\boldsymbol{\xi}^T \boldsymbol{A} \boldsymbol{\xi}$,

$$\mathbb{E}\left(\boldsymbol{\xi}^T \boldsymbol{A} \boldsymbol{\xi}\right) = \mathbb{E}\left(\sum_{i=1}^{q}\sum_{j=1}^{q} \xi_i A_{ij} \xi_j\right) = \sum_{i=1}^{q}\sum_{j=1}^{q} A_{ij} \mathbb{E}\left(\xi_i \xi_j\right). \tag{6}$$

Since $\mathbb{E}\left(\xi_i \xi_j\right) = 1$ only when $i = j$, and 0 otherwise, we get

$$\mathbb{E}\left(\boldsymbol{\xi}^T \boldsymbol{A} \boldsymbol{\xi}\right) = \sum_{i=1}^{q} A_{ii} = \text{trace}\left(\boldsymbol{A}\right). \tag{7}$$

As a result, the trace of a given matrix $\boldsymbol{A}$ can be estimated using random samples. Here, we utilize the sample average approximation (SAA) technique, see, e.g., [28], which approximates the expected value by the average. The sample average or empirical mean for n samples (or realizations) $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$, of the random variable $\boldsymbol{\xi}$ is defined as

$$\mathbb{E}_S\left(\boldsymbol{\xi}^T \boldsymbol{A} \boldsymbol{\xi}\right) = \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{\xi}_k^T \boldsymbol{A} \boldsymbol{\xi}_k. \tag{8}$$

According to the Law of Large Numbers (LLN) [29], as the number of (independent) samples n approaches ∞, the sample average $\mathbb{E}_S\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)$ converges to the expectation $\mathbb{E}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)$, which is the trace of $\boldsymbol{A}$,

$$\mathbb{E}_S\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) = \frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\xi}_k^T\boldsymbol{A}\boldsymbol{\xi}_k \to \mathbb{E}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) = \text{trace}\left(\boldsymbol{A}\right). \tag{9}$$

The use of trace estimators can be highly efficient if the matrix $\boldsymbol{A}$ is not explicitly available and its computation is expensive. We also have that $\mathbb{E}_S\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)$ is an unbiased estimator of $\mathbb{E}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)$ [28]. Since the expected accuracy of such estimates is given by the variance, the goal is to obtain the estimates with the smallest variance. The Rademacher distribution, from which we choose the random vectors $\boldsymbol{\xi}$, was shown in [10] to be the distribution that yields the smallest variance. Other studies rank various trace estimators differently according to criteria other than the variance. For further information, readers are referred to, e.g., [14,15]. For a symmetric matrix $\boldsymbol{A}$, the variance of $\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}$ is derived as

$$\begin{aligned}
\text{Var}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) &= \mathbb{E}\left\{\left[\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi} - \mathbb{E}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)\right]^2\right\} = \mathbb{E}\left\{\left[\sum_{i=1}^{q}\sum_{j=1}^{q}\xi_i A_{ij}\xi_j - \sum_{i=1}^{q}A_{ii}\right]^2\right\} \\
&= \mathbb{E}\left\{\left[\sum_{\substack{i,j=1\\ i\neq j}}^{q}\xi_i A_{ij}\xi_j\right]^2\right\} = \mathbb{E}\left\{\sum_{\substack{i,j=1\\ i\neq j}}^{q}\sum_{\substack{k,l=1\\ k\neq l}}^{q}A_{ij}A_{kl}\xi_i\xi_j\xi_k\xi_l\right\} \\
&= \sum_{\substack{i,j=1\\ i\neq j}}^{q}\sum_{\substack{k,l=1\\ k\neq l}}^{q}A_{ij}A_{kl}\mathbb{E}\left\{\xi_i\xi_j\xi_k\xi_l\right\} = \sum_{\substack{i,j=1\\ i\neq j}}^{q}\sum_{\substack{k,l=1\\ k\neq l}}^{q}A_{ij}A_{kl}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}) \\
&= \sum_{\substack{i,j=1\\ i\neq j}}^{q}(A_{ij}A_{ij}+A_{ij}A_{ji}) = 2\sum_{\substack{i,j=1\\ i\neq j}}^{q}A_{ij}^2,
\end{aligned} \tag{10}$$

where δ_{ij} is the Kronecker delta function. In addition, the standard deviation of $\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}$ is given by

$$\text{Dev}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) = \sqrt{\text{Var}\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right)} = \sqrt{2\sum_{\substack{i,j=1\\ i\neq j}}^{q}A_{ij}^2}. \tag{11}$$

The variance and standard deviation can be estimated using a finite number of samples. Given the samples above, we define the sample variance and sample standard deviation as

$$\text{Var}_S\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) = \frac{1}{n}\sum_{k=1}^{n}\left[\boldsymbol{\xi}_k^T\boldsymbol{A}\boldsymbol{\xi}_k - \frac{1}{n}\sum_{\ell=1}^{n}\boldsymbol{\xi}_\ell^T\boldsymbol{A}\boldsymbol{\xi}_\ell\right]^2, \tag{12}$$

and

$$\text{Dev}_S\left(\boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{\xi}\right) = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[\boldsymbol{\xi}_k^T\boldsymbol{A}\boldsymbol{\xi}_k - \frac{1}{n}\sum_{\ell=1}^{n}\boldsymbol{\xi}_\ell^T\boldsymbol{A}\boldsymbol{\xi}_\ell\right]^2}. \tag{13}$$

Below, we use these derivations for estimating the objective function (the compliance), as well as its gradient or sensitivity. A key issue for efficiency is that both can be estimated with the same set of random vectors.

2.2. Randomized topology optimization with stochastic sampling

We apply the stochastic sampling technique to both the density-based method and the GSM. The basic idea is to replace the compliance and its gradient by stochastic estimates and to use these estimates in the optimization algorithm.

2.2.1. Density-based method

Consider the standard topology optimization formulation in (1) with m load cases $\boldsymbol{f}_i$, $i = 1, \ldots, m$, and corresponding weights α_i. We define a weighted load matrix $\boldsymbol{F} \in \mathbb{R}^{dN \times m}$ as $\boldsymbol{F} = \left[\sqrt{\alpha_1}\boldsymbol{f}_1, \ldots, \sqrt{\alpha_m}\,\boldsymbol{f}_m\right]$. In a similar fashion, we define the weighted displacement matrix $\boldsymbol{U} \in \mathbb{R}^{dN \times m} = \left[\sqrt{\alpha_1}\,\boldsymbol{u}_1, \ldots, \sqrt{\alpha_m}\,\boldsymbol{u}_m\right]$, whose columns are the corresponding displacement fields. The matrix $\boldsymbol{U}$ is defined by the equilibrium equation, $\boldsymbol{U} = \boldsymbol{K}^{-1}\boldsymbol{F}$. So, we can write the end-compliance and its sensitivities as traces of symmetric matrices

$$C(\boldsymbol{\rho}) = \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i^T \boldsymbol{u}_i = \text{trace}\left(\boldsymbol{F}^T\boldsymbol{U}\right) = \text{trace}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\right) \tag{14}$$

and

$$\nabla C_{\boldsymbol{\rho}}^{(e)} = -\sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i^T \frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}} \boldsymbol{u}_i = -\,\text{trace}\left(\boldsymbol{U}^T \frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}} \boldsymbol{U}\right) = -\,\text{trace}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\boldsymbol{F}\right). \tag{15}$$

With the random variable $\boldsymbol{\xi}$ as defined above, the end-compliance and the topology optimization problem can be expressed as

$$C(\boldsymbol{\rho}) = \text{trace}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\right) = \mathbb{E}\left(\boldsymbol{\xi}^T\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\boldsymbol{\xi}\right) = \mathbb{E}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right], \tag{16}$$

and

$$\begin{aligned} &\min_{\boldsymbol{\rho}} C(\boldsymbol{\rho}) = \min_{\boldsymbol{\rho}} \mathbb{E}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}(\boldsymbol{\rho})^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right], \\ &\text{s.t.} \quad \sum_{e=1}^{M} v^{(e)}\rho^{(e)} - V_{\max} \le 0, \\ &\qquad 0 < \rho^{\min} \le \rho^{(e)} \le 1,\ e = 1, \ldots, M. \end{aligned} \tag{17}$$

The sensitivity of the objective function can be expressed as

$$\nabla C_{\boldsymbol{\rho}}^{(e)} = -\,\text{trace}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\boldsymbol{F}\right) = -\mathbb{E}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right]. \tag{18}$$

Eqs. (17) and (18), stated as a stochastic programming problem, are equivalent to the standard formulation (1) and (2). We approximate the compliance and its gradient by replacing their expectation in the equations above by their sample average estimates. Given the i.i.d. random sample $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$ as n realizations of the random vector $\boldsymbol{\xi}$, we define

$$\widehat{C}^S(\boldsymbol{\rho}) = \frac{1}{n}\sum_{k=1}^{n} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right)^T\boldsymbol{K}(\boldsymbol{\rho})^{-1}\left(\boldsymbol{F}\boldsymbol{\xi}_k\right), \tag{19}$$

and

$$(\nabla\widehat{C}_{\boldsymbol{\rho}}^S)^{(e)} = \frac{\partial \widehat{C}^S}{\partial \rho^{(e)}} = -\frac{1}{n}\sum_{k=1}^{n} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right)^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\left(\boldsymbol{F}\boldsymbol{\xi}_k\right). \tag{20}$$

We remark that $\widehat{C}^S(\boldsymbol{\rho})$ and $\nabla\widehat{C}_{\boldsymbol{\rho}}^S$ are unbiased estimators for the compliance and its gradient. By the LLN, $\widehat{C}^S(\boldsymbol{\rho}) \to C(\boldsymbol{\rho})$ and $\nabla\widehat{C}_{\boldsymbol{\rho}}^S \to \nabla C_{\boldsymbol{\rho}}$ (with probability 1) for any feasible $\boldsymbol{\rho}$ as $n \to \infty$ [30].

Within each step of the optimization algorithm, we use the same random vectors to estimate the compliance and its gradient using (19) and (20). However, to avoid convergence for a specific random load case, a new set of n random vectors is selected at each optimization step. If $n \ll m$, our proposed algorithm reduces the computational cost from $m \times N_{\text{step}}$ to roughly $n \times N_{\text{step}}$ if the convergence of the optimization is not affected.

The idea to approximate the gradient and the objective function in a structural optimization problem is similar to the use of stochastic gradient-based methods in other areas, e.g., stochastic gradient descent (SGD) [31], and stochastic meta-descent (SMD) [32], which are optimization methods mainly for unconstrained optimization problems. Stochastic gradient methods use small sub-samples (also referred as mini-batches) to estimate the gradient and have been applied in other fields, such as large-scale machine learning [33,34]. For our application, the SAA gradient estimate always has a positive angle with the gradient, and hence the negative SAA gradient estimate is a descent direction for the unconstrained case. We demonstrate numerically how effective the approximation

is in Section 4. Because the optimization problem for the density-based method with penalization is non-convex, deterministic gradient-based methods may not converge to the global minimum (as expected). The randomized approach leads to solutions that are (roughly) as accurate in terms of the end-compliance as those from the standard weighted sum approach. This is demonstrated by the results in Section 4.2.

In general, for stochastic gradient methods, a damping or averaging scheme (also referred to as a decay schedule of the scalar gain or gain vector adaptation) is needed to achieve convergence [34]. Various damping or averaging methods for step sizes have been proposed for different types of problems [35]. For the structural optimization problems in this paper, we propose a damping scheme that adjusts the move limits (reminiscent of a trust region), which especially befits the structural optimization framework. The idea of the proposed damping scheme is similar to adjusting the size of the update in simulated annealing [36,37] and is discussed in detail in Section 3.1. Robbins and Monro [31] have given conditions for the convergence of stochastic gradient methods that use such damping schemes; see also [34]. A more recent (and accessible) paper discussing the convergence of stochastic gradient methods in detail, and considering more general cases than Robbins and Monro [31], is the one by Bottou, Curtis and Nocedal [38].

The numerical results (see Section 4) show that the convergence of the proposed approach is typically rapid if our damping scheme is properly used, roughly as fast as for the standard weighted sum approach and sometimes faster. Since the randomized approach solves only n linear systems at each optimization step, compared with m for the standard weighted sum approach, and $n \ll m$, the proposed randomized approach is computationally much more efficient. Moreover, randomized approaches with a proper damping scheme can be more robust in finding the global minimum for non-convex optimization problems than deterministic algorithms. We demonstrate this with numerical examples in Section 4. In general, such increased robustness comes at the price of slower convergence.

To analyze the effects of randomization on the proposed approaches, we consider the variance and sample variance for the compliance estimate and its gradient.

$$\mathrm{Var}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right] = 2\sum_{\substack{i,j=1\\ i\neq j}}^{m}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\right)_{ij}^2, \tag{21}$$

and

$$\mathrm{Var}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right] = 2\sum_{\substack{i,j=1\\ i\neq j}}^{m}\left(\boldsymbol{F}^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\boldsymbol{F}\right)_{ij}^2. \tag{22}$$

Similarly, the sample variance of the compliance and its gradient can be expressed as

$$\mathrm{Var}_S\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right] = \frac{1}{n}\sum_{k=1}^{n}\left(\boldsymbol{\xi}_k^T\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\boldsymbol{\xi}_k - \frac{1}{n}\sum_{\ell=1}^{n}\boldsymbol{\xi}_\ell^T\boldsymbol{F}^T\boldsymbol{K}^{-1}\boldsymbol{F}\boldsymbol{\xi}_\ell\right)^2, \tag{23}$$

and

$$\begin{aligned}&\mathrm{Var}_S\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right]\\ &\quad= \frac{1}{n}\sum_{k=1}^{n}\left(\boldsymbol{\xi}_k^T\boldsymbol{F}^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\boldsymbol{F}\boldsymbol{\xi}_k - \frac{1}{n}\sum_{\ell=1}^{n}\boldsymbol{\xi}_\ell^T\boldsymbol{F}^T\boldsymbol{K}^{-1}\frac{\partial \boldsymbol{K}}{\partial \rho^{(e)}}\boldsymbol{K}^{-1}\boldsymbol{F}\boldsymbol{\xi}_\ell\right)^2.\end{aligned} \tag{24}$$

2.2.2. *Ground structure method*

A randomized version of the optimization problem for the GSM with the weighted force matrix $\boldsymbol{F}$ is analogous to that for the density-based method (17). The minimization problem for the GSM is given by

$$\begin{aligned}&\min_{\boldsymbol{x}} C(\boldsymbol{x}) = \min_{\boldsymbol{x}} \mathbb{E}\left[(\boldsymbol{F}\boldsymbol{\xi})^T\boldsymbol{K}(\boldsymbol{x})^{-1}(\boldsymbol{F}\boldsymbol{\xi})\right],\\ &\text{s.t.}\quad \sum_{e=1}^{M} L^{(e)}x^{(e)} - V_{\max} \le 0,\\ &\qquad 0 < x^{\min} \le x^{(e)} \le x^{\max},\ e = 1,\ldots,M,\end{aligned} \tag{25}$$

and the sensitivity of the objective function can be expressed as,

$$\nabla C_{\boldsymbol{x}}^{(e)} = -\mathbb{E}\left[(\boldsymbol{F}\boldsymbol{\xi})^{T} \boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial x^{(e)}} \boldsymbol{K}^{-1} (\boldsymbol{F}\boldsymbol{\xi}) \right]. \tag{26}$$

Here too, we replace the objective function and its gradient by their sample average approximation. Using the same assumptions, for an i.i.d. random sample $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$, the compliance $C(\boldsymbol{x})$ and its gradient can be estimated by

$$\widehat{C}^{S}(\boldsymbol{x}) = \frac{1}{n} \sum_{k=1}^{n} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right)^{T} \boldsymbol{K}(\boldsymbol{x})^{-1} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right), \tag{27}$$

and

$$(\nabla \widehat{C}_{\boldsymbol{x}}^{S})^{(e)} = \frac{\partial \widehat{C}^{S}}{\partial x^{(e)}} = -\frac{1}{n} \sum_{k=1}^{n} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right)^{T} \boldsymbol{K}^{-1} \frac{\partial \boldsymbol{K}}{\partial x^{(e)}} \boldsymbol{K}^{-1} \left(\boldsymbol{F}\boldsymbol{\xi}_k\right), \tag{28}$$

where $\widehat{C}^{S}(\boldsymbol{x})$ and $(\nabla \widehat{C}_{\boldsymbol{x}}^{S})^{(e)}$ are the estimated compliance and estimated sensitivity of the corresponding structure analogous to (19) and (20). The variance and sample variance of the objective function and its sensitivity take the same form as (21)–(24), and therefore are not listed. Since the standard optimization formulation for the GSM is convex, it has a unique global minimum. Hence, we expect the approximated minima obtained with the randomized method to be equal to or larger than the one obtained from the standard formulation. This is confirmed by the results in Section 4.

We conclude this subsection by making a remark on the flexibility of our proposed randomized algorithm for topology optimization. This algorithm can be combined with many types of update schemes that are based on gradient information, such as the Optimality Criterion (OC) method [27] and the Method of Moving Asymptotes (MMA) [39]. In this work, because the estimated sensitivities in (20) and (28) are guaranteed to be non-positive, we adopt OC as the update scheme for both continuum and truss topology optimization. The OC method is a highly efficient update scheme that is tailored for self-adjoint problems and widely used in topology optimization: see, e.g. [40–42].

2.3. Discrete filter for ground structure method with stochastic sampling

In the GSM, a discrete filter similar to the filter proposed in [43] (see below) is applied to the truss topology optimization with both the standard and the randomized algorithms to extract valid structures out of ground structures. The use of the discrete filter reduces the number of redundant bars and the size of the structural problem within each solve, which reduces the cost of subsequent optimization steps. For the standard multiple load case optimization problem (3), the discrete filter can be expressed as

$$Filter\left(\boldsymbol{x}_k, \alpha_f\right) = \begin{cases} 0, & \textbf{if} \quad \dfrac{x_k^{(e)}}{\max\left(\boldsymbol{x}_k\right)} < \alpha_f < 1, \\ x^{(e)}, & \text{otherwise}, \end{cases} \tag{29}$$

where α_f is the prescribed filter value and $x_k^{(e)}$ is the design variable for truss member e at the kth step. The discrete filter is applied to the GS at each step and removes the truss members with normalized areas smaller than the filter parameter α_f.

For the randomized algorithm for the GSM, we define a discrete filter that slightly differs from the standard discrete filter. To avoid inadvertent removal of truss members due to an (occasional) poor estimate in the randomized algorithm, the discrete filter removes truss members only when their normalized areas have remained below the prescribed filter value for n_f steps, namely,

$$Filter^{S}\left(\boldsymbol{x}_k, \alpha_f\right) = \begin{cases} 0, & \text{if} \quad \max\left(\dfrac{x_{k-n_f+1}^{(e)}}{\max\left(\boldsymbol{x}_{k-n_f+1}\right)}, \ldots, \dfrac{x_k^{(e)}}{\max\left(\boldsymbol{x}_k\right)} \right) < \alpha_f < 1, \\ x^{(e)}, & \text{otherwise}, \end{cases} \tag{30}$$

where $Filter^{S}(\boldsymbol{x}_k, \alpha_f)$ denotes the filter for the randomized algorithm, and n_f is a chosen number of monitored steps (see the discussion of parameters below). The use of the filter for the randomized algorithm ($Filter^{S}$) leads to a more efficient GSM with stochastic sampling than if the filter were not used (see numerical examples section).

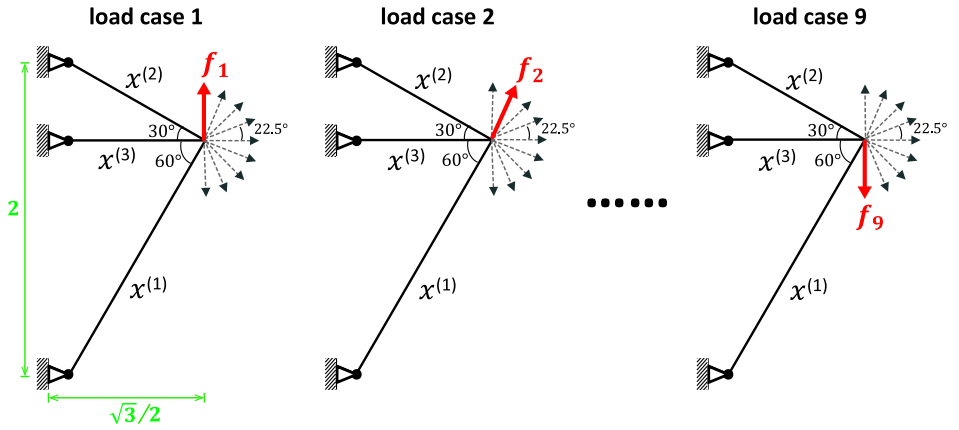

Fig. 1. Three-bar truss structure under 9 load cases: initial ground structure, load and boundary conditions.

3. A damping scheme and algorithmic parameters for randomized optimization

This section introduces a damping scheme to facilitate convergence and demonstrates the effectiveness of this scheme through a three-bar truss example. We further discuss the algorithmic parameters that are used in the proposed randomized optimization framework and comment on the range of values chosen for those parameters.

3.1. The proposed damping scheme: effective step ratio and step size reduction

In the randomized algorithm, the structure must be adjusted based on the random linear combination of load cases that changes at each optimization step. Therefore, the convergence criteria commonly used for the standard structural optimization framework are insufficient for the proposed framework. Based on ideas from simulated annealing [36,37], we propose a damping scheme that evaluates the average progress per step and reduces the move limit (similar to the scalar gain in SGD) accordingly. We define the effective step ratio R as follows (using the GSM notation):

$$R = \frac{\frac{1}{n_{\text{step}}} \| \left(\boldsymbol{x}_k - \boldsymbol{x}_{k-n_{\text{step}}+1} \right) \|}{\| \boldsymbol{x}_k - \boldsymbol{x}_{k-1} \|}, \tag{31}$$

where n_{step} is the sample window size (see below). The average step size over a sample window is divided by the current step length. This effective step ratio serves as an indicator of the optimizer's status, i.e., the ratio is relatively large when the optimizer is making progress; and relatively small (typically smaller than 0.1) when the step is not effective. Once R is below a prescribed tolerance τ_{step}, i.e., $R < \tau_{\text{step}}$, we reduce the allowable move limit of the optimizer by a prescribed scale factor γ. The move limit is not adjusted in the first n_{step} optimization steps.

The effectiveness of the damping scheme is illustrated through a simple numerical example for the GSM. This example is not representative for the effectiveness of the randomized algorithm; it is selected to emphasize the poor performance of the randomized optimization algorithm without a damping scheme. We consider a three-bar truss structure supported at their left ends and subjected to a set of 9 equal-weighted load cases, $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_9$, at their right joint, as shown in Fig. 1. The problem formulation is given as follows:

$$\begin{aligned} &\min_{\boldsymbol{x}} C(\boldsymbol{x}) = \sum_{i=1}^{9} \boldsymbol{f}_i{}^T \boldsymbol{u}_i(\boldsymbol{x}), \\ \text{s.t.} \quad &\sum_{e=1}^{3} L^{(e)} x^{(e)} - V_{\max} \le 0, \\ &0 < x^{\min} \le x^{(e)} \le x^{\max}, \; e = 1, \ldots, 3, \\ &\text{with } \boldsymbol{u}_i(\boldsymbol{x}) = \boldsymbol{K}(\boldsymbol{x})^{-1} \boldsymbol{f}_i, \; i = 1, \ldots, 9. \end{aligned} \tag{32}$$

Table 1
Solution of the 3-bar truss of Fig. 1 using the standard GSM and the GSM with stochastic sampling and damping.

	Standard GSM	Randomized GSM w/ damping
$x^{(1)}$	0.0344	0.0344
$x^{(2)}$	0.0290	0.0290
$x^{(3)}$	0.0132	0.0134
C^*	8.1666	8.1667

We set the volume constraint as $V_{\max} = 0.1$, and take initial guess as $x_0 = V_{\max}/\sum_e L^{(e)} = 0.0278$. In addition, we choose the lower and upper bounds to be $x^{\min} = 10^{-8}x_0$ and $x^{\max} = 10^4 x_0$, respectively. The initial move limit is taken as $move = 10^{-1}x_0$. The tolerance for convergence of the optimization is 10^{-8}. For the randomized algorithm, we choose the sample size as $n = 6$ (i.e., 6 sample load cases), and select a new sample at each iteration.

We use the proposed randomized algorithm with and without the damping scheme. The contour plots of the objective function with the optimization history of $x^{(1)}$ and $x^{(2)}$ ($x^{(3)}$ can be computed from the volume constraint because, in practice, the sum of volumes will always be equal to $V_{\max}$) for both cases are shown in Fig. 2(a) and 2(b). The optimization history from the standard weighted sum approach is plotted for reference. The standard approach obtains the global minimum of the given problem within 25 steps where $\boldsymbol{x}^* = \left[x^{(1)}, x^{(2)}, x^{(3)}\right]^T = [0.0344, 0.0290, 0.0132]^T$ and $C(\boldsymbol{x}^*) = 8.1666$. The randomized algorithm without damping does not converge, and the updates become ineffective after roughly 10 steps (Fig. 2a). The randomized algorithm with our damping scheme converges to $\widehat{\boldsymbol{x}}^S = [0.0343, 0.0290, 0.0134]^T$ with $C(\widehat{\boldsymbol{x}}^S) = 8.1667$. This is close to the optimal solution obtained by the standard algorithm. The results are summarized in Table 1. For this small problem, we use $n_{\text{step}} = 10$, $\tau_{\text{step}} = 0.05$, and $\gamma = 2$. Since the solution of this example is relatively trivial, the randomized algorithm with our damping scheme converges slower than the standard algorithm. For more complicated and realistic problems (e.g., examples in Section 4), the convergence rates for the randomized and standard algorithms are comparable.

To verify that the solution from the randomized algorithm with damping converges to a KKT (Karush–Kuhn–Tucker) point (optimal solution in case of the GSM), we examine the angle $\theta^{\mathcal{A}}$ between $\nabla C^{\mathcal{A}}$ (reduced gradient vector of the objective function) and $\boldsymbol{L}^{\mathcal{A}}$ (reduced gradient vector of the volume constraint) for the standard algorithm, and randomized algorithm with and without damping, as shown in Fig. 2(c). The solution is a KKT point if $\theta^{\mathcal{A}} = \pi$. For the standard algorithm, $\theta^{\mathcal{A}}$ converges quickly to π. For the randomized algorithm with damping, $\theta^{\mathcal{A}}$ gradually converges to π (the case without damping does not converge). Hence, we numerically show that the solution from the randomized algorithm with damping converges to the optimal solution in the truss optimization framework.

3.2. Overview of algorithmic parameters for randomized optimization

This subsection summarizes the important parameters that are used in the randomized optimization framework, with some comments on the possible range of values to be used in practice.

Sample size. In the proposed randomized optimization framework, the larger the number of sample load cases n, the more accurate the estimate of the compliance will be. However, the computational cost increases with the number of sample load cases, because in each optimization step we need to solve n systems of equations. Thus, we need to balance the accuracy of the estimates and the computational complexity. Typically, the results from the randomized algorithm are relatively insensitive to the number of sample load cases. Indeed, with a small number of sample load cases ($n \ll m$) we obtain solutions comparable to those from the standard algorithm in terms of topology and compliance value, and sometimes even better in terms of compliance value. It seems that for these problems the estimated gradient is fairly accurate even for small sample sizes. This is also demonstrated in the numerical examples in the next section. To provide some insight into the choice of sample size, a study is conducted using different sample sizes in Section 4.1. For the examples in the remainder of this paper, we choose the number of sample load cases $n = 6$, unless otherwise stated, to demonstrate the accuracy and computational efficiency of the randomized optimization framework.

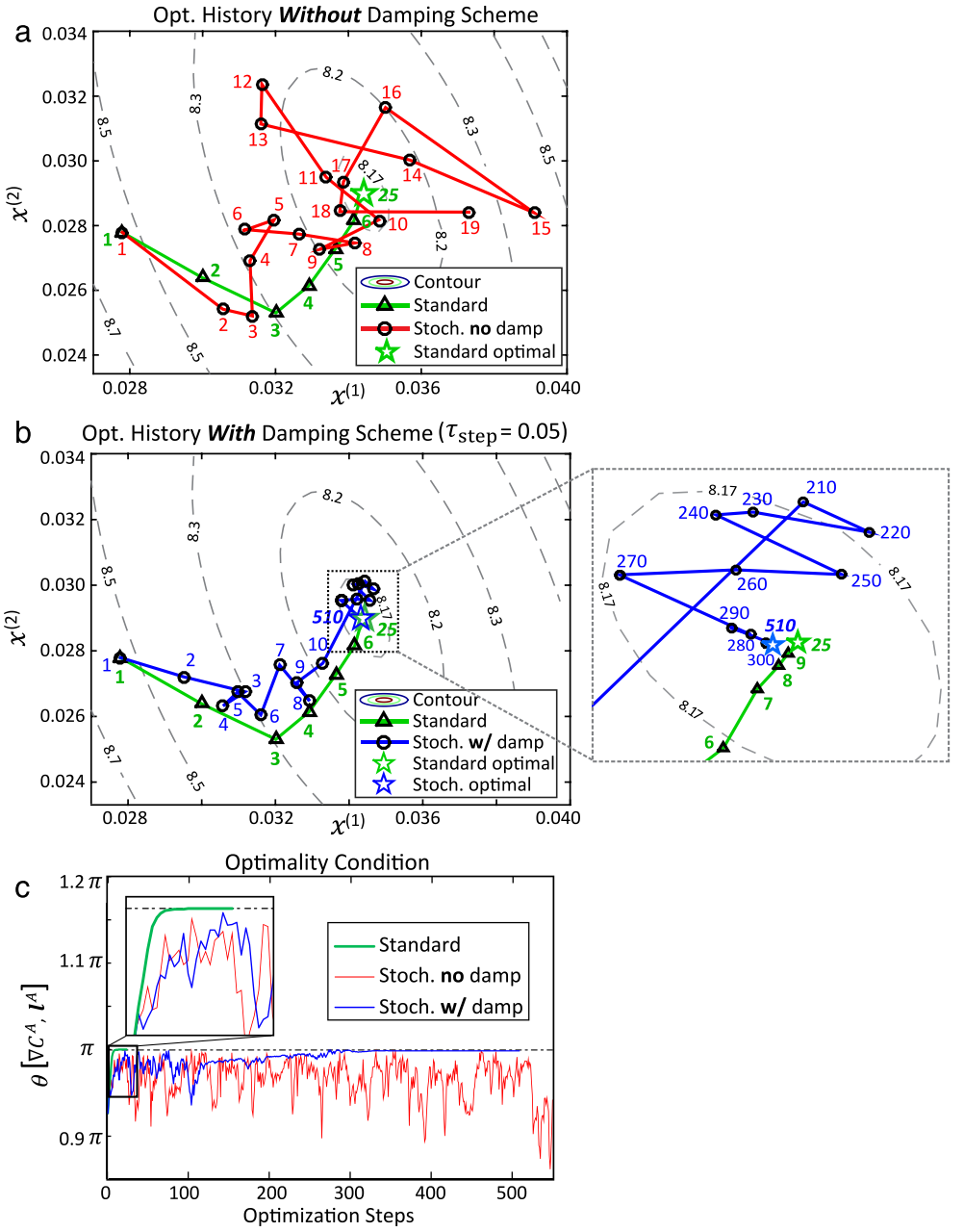

Fig. 2. Illustration of the damping scheme in the three-bar truss example: contour plots of the objective function with the optimization history of $x^{(1)}$ and $x^{(2)}$ for (a) the standard weighted sum approach and the randomized approach without a damping scheme; (b) the standard weighted sum approach and the randomized approach with the proposed damping scheme. (c) The angle between the reduced gradient vectors of the objective function and the volume constraint.

Frequency to select a new sample. The frequency to select a new sample, n_s, or the number of optimization steps with a fixed random sample, influences the convergence rate of the optimization. If the frequency is too low, optimization steps are influenced too much by specific sets of random loads, and convergence may be slow. In this work, we select a new random sample every step, i.e., $n_s = 1$.

Filter parameters. Applying the discrete filter in the GSM reduces the number of redundant bars, which drastically reduces the computational cost, for both standard and randomized algorithms. The discrete filter for the randomized

algorithm, as discussed in Section 2.3, also helps to limit the effects of the stochastic estimates. For the randomized algorithm, we choose a relatively small filter size $\alpha_f = 10^{-4}$ and remove truss members when their normalized cross-sectional areas have remained below α_f for $n_f = 10$ cumulative steps. Further details about the filter parameters can be found in the paper by Ramos Jr. and Paulino [43].

Parameters in the damping scheme. The damping scheme introduced in Section 3.1 has four parameters, the effective step ratio (R), the sample window size (n_{step}), the tolerance (τ_{step}), and the scale factor (γ). The parameter choices in this damping scheme are crucial to the quality of the final solution. We average over a moving window. As the window size n_{step} increases, we average over more steps. If n_{step} is too large, it slows down convergence because the algorithm adapts more slowly. In practice, we have found $n_{\text{step}} = 100$ is sufficient for problems containing more than one thousand design variables. At the start of the optimization, the number of steps after which we start to dampen the allowable move limit is typically chosen to be the same as the window size.

The tolerance for the effective step ratio, τ_{step}, serves as a threshold to determine when updates become ineffective. Therefore, the choice of τ_{step} affects the rate of convergence and sometimes the quality of the solutions. In general, a loose tolerance leads to faster convergence (reduces the move limit more frequently) at the expense of the quality of design (in terms of the compliance value). One must balance the quality of the results and the convergence rate. The GSM is more sensitive to τ_{step} than the density-based method. This is demonstrated in Section 4.1. Therefore, a stricter tolerance is needed for the GSM. In practice, we choose $\tau_{\text{step}} = 0.05$ for the GSM and $\tau_{\text{step}} = 0.1$ for the density-based method. For the move limit scale factor γ, we have found that $\gamma = 2$ is typically a good choice.

4. Numerical examples

We present several numerical examples in both two and three dimensions to demonstrate the effectiveness as well as the computational efficiency of the proposed randomized algorithm for topology optimization. Both density-based method and GSM are used. The first two examples (Section 4.1) investigate the sensitivity of the density-based method and the GSM to the tolerance τ_{step} (see Eq. (31)). Moreover, Example 1 (Section 4.1.1) shows the relation between sample size and quality of the optimized design. Example 2 (Section 4.1.2) illustrates the effect of the discrete filter in the GSM for the proposed randomized algorithm. The last two examples (Sections 4.2 and 4.3) in 3D demonstrate the capability of the proposed algorithm to create practical structural designs at greatly reduced computational cost.

To quantify the computational cost of the standard and randomized optimization algorithms, we define $N_{\text{solve}} = n \times N_{\text{step}}$ as the total number of linear systems of equations to solve in the optimization process, where N_{step} is the number of optimization steps. This is a measure of the computational efficiency of an optimization formulation. The optimization process is considered converged if the current step size (bounded by the move limit) is below a prescribed tolerance τ_{opt} for the optimization process, that is, $\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\| < \tau_{\text{opt}}$.

For the density-based method (continuum), we incorporate the proposed technique into the computer program PolyTop [44] in 2D and the topology optimization code in 3D [45]. For plotting 3D continuum results, we utilize TOPSlicer [46]. All the problems are initialized as follows. The initial guess is taken as $\rho_0 = V_{\max}/\sum_{e=1}^{M} v^{(e)}$, where $v^{(e)}$ is the volume (area) of element e. The convergence tolerance is $\tau_{\text{opt}} = 10^{-2}$; the initial move limit is chosen as $move = 0.05$; the damping factor for the OC update scheme is $\eta = 0.5$. We use a continuation scheme, in which the penalization factor starts at $p = 1$ and each time increases by 0.5 (or 1) until $p = 3$ (for 2D problems [44]) or 4 (for 3D problems [46]). For example, $p = [1, 1.5, 2, 2.5, 3]$.

For the GSM (truss-layout), we generate initial ground structures (without overlapped bars) using the collision zone technique from Refs. [47,48] and plot final topologies in 3D using the program GRAND3 [48]. The initial guess of the design variables is taken as $x_0 = V_{\max}/\sum_{e=1}^{M} L^{(e)}$; the convergence tolerance is $\tau_{\text{opt}} = 10^{-8}$; the initial move limit is chosen as $move = x_0 \times 10^4$; the damping factor for the OC update scheme is $\eta = 0.5$. When the discrete filter is used in the GSM, we use $n_f = 10$ and $\alpha_f = 10^{-4}$ during the optimization process, unless otherwise stated; the lower and upper bounds on the design variables are $x^{\min} = 0$ and $x^{\max} = 10^4 x_0$ (unbounded in practical terms). For the standard GSM (without the discrete filter), we apply a cut-off value 10^{-2} that defines the final structure at the end of the optimization [49]. The lower and upper bounds are defined by $x^{\min} = 10^{-2} x_0$ and $x^{\max} = 10^4 x_0$, respectively. For all results in the GSM, we remove unstable nodes and floating bars and then check the final topologies to ensure that they are at global equilibrium — a detailed explanation can be found in Ref. [50].

In both the continuous and truss-layout optimization, the randomized algorithm uses the following parameters. Unless otherwise stated, the sample size is chosen to be $n = 6$; in the damping scheme, the window size is $n_{\text{step}} = 100$,

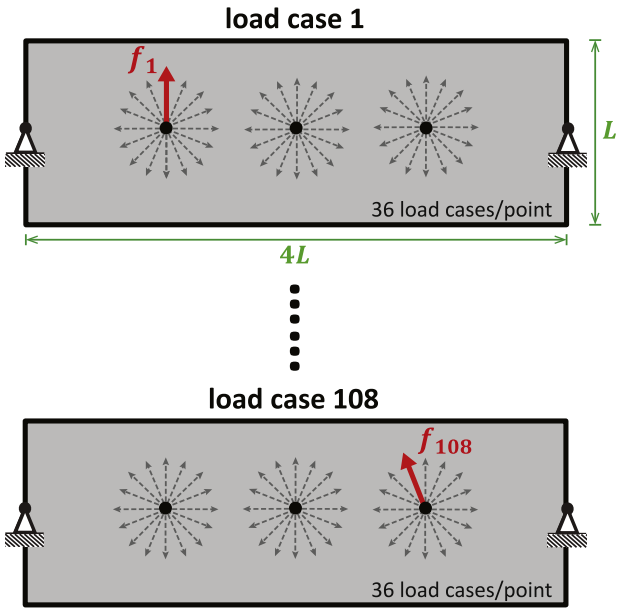

Fig. 3. Two-dimensional box domain with load and support conditions. A total of 108 equal-weighted load cases are applied at three given points with each point having 36 load cases applied from 0° to 350° (dotted arrows are the schematic illustrations of non-active load cases).

and we use step size reduction factor $\gamma = 2$. For the density-based method, we use a tolerance for the effective step ratio $\tau_{\text{step}} = 0.1$, and for the GSM we use $\tau_{\text{step}} = 0.05$. Let $\boldsymbol{\rho}^*$ and $\boldsymbol{x}^*$ represent the optimal solutions of the standard formulations in (1) and (3), $\widehat{\boldsymbol{\rho}}^S$ and $\widehat{\boldsymbol{x}}^S$ represent the optimal solutions obtained from the randomized algorithm for density-based and GS methods. To evaluate the quality of the solutions, in the case of the randomized algorithm, we present the *true values* of the objective function $C(\widehat{\boldsymbol{\rho}}^S)$ and $C(\widehat{\boldsymbol{x}}^S)$ at the approximated solutions $\widehat{\boldsymbol{\rho}}^S$ and $\widehat{\boldsymbol{x}}^S$ (instead of their estimators $\widehat{C}^S(\widehat{\boldsymbol{\rho}}^S)$ and $\widehat{C}^S(\widehat{\boldsymbol{x}}^S)$) and compare them with those obtained from the standard algorithm $C(\boldsymbol{\rho}^*)$ and $C(\boldsymbol{x}^*)$. The relative difference is defined as $\Delta C = \left(C(\widehat{\boldsymbol{x}}^S) - C(\boldsymbol{x}^*)\right)/C(\boldsymbol{x}^*)$.

4.1. Two-dimensional box domain with 108 load cases

We present a two-dimensional (2D) topology optimization problem whose design domain and boundary conditions are shown in Fig. 3. A total of 108 equal-weighted load cases are applied at three given points, with each point having 36 load cases applied along different angles (from 0° to 350°). In this section, both the density-based and the GS methods are used.

4.1.1. Example 1: Continuum topology optimization with density-based method

Using the density-based method, we demonstrate the reduction of the computational cost by means of the randomized algorithm. We further investigate the sensitivity of the final optimized topologies to the tolerance τ_{step} in the damping scheme and the sample size n. Since the final topology from the standard algorithm is symmetric both horizontally and vertically, we enforce symmetry of the topologies in the randomized case by enforcing horizontal and vertical symmetry of the distribution of the density field [44]. A total number of 25,600 quadrilateral elements are used to discretize the domain which gives 52,002 degrees of freedom (DOFs). The linear density filter that defines the solid-void boundary takes the radius of 1 (see Section 1.1). For comparison purposes, the final topology obtained from the standard topology optimization is shown in Fig. 4(a). The final topology has $C(\boldsymbol{\rho}^*) = 3.257$ and converges after 1048 steps. In each optimization step, we solve 108 linear systems (corresponding to 108 load cases), which leads to a total 113,184 solves. Since the continuation method is used for the penalization factor, p, the jumps in the compliance correspond to $p = 1.5, 2.0, 2.5, 3.0$ (initially $p = 1.0$) [44].

To investigate the sensitivity to τ_{step}, we consider $\tau_{\text{step}} = 0.1$ and $\tau_{\text{step}} = 0.05$, and the results are compared with that from the standard algorithm [4]. For both cases, the number of sample load cases used is $n = 6$. Fig. 4(b) and

Table 2
Results for Example 1 (density-based), averaged over 5 trials.

Density-based	$C(\boldsymbol{\rho}^*)$	$C(\widehat{\boldsymbol{\rho}}^S)$	ΔC	$Dev_S(C(\widehat{\boldsymbol{\rho}}^S))$	τ_{step}	$\cos\bar{\theta}$	n	N_{step}	N_{solve}
Standard	3.257	–	–	–	–	–	–	1048	113,184
Stoch. $\tau_{\text{step}} = 0.05$	–	3.315	1.79%	$2.00E-2$	0.05	0.971	6	1052	6312
Stoch. $\tau_{\text{step}} = 0.1$	–	**3.337**	**2.45%**	**$9.17E-3$**	**0.1**	**0.971**	**6**	**695**	**4,170**
Stoch. $n = 4$	–	3.389	4.05%	$2.15E-2$	0.1	0.959	4	618	2,472
Stoch. $n = 5$	–	3.356	3.05%	$3.28E-3$	0.1	0.967	5	684	3,420
Stoch. $n = 7$	–	3.334	2.36%	$2.84E-2$	0.1	0.976	7	684	4,788
Stoch. $n = 20$	–	3.291	1.05%	$3.33E-3$	0.1	0.991	20	838	16,760
Stoch. $n = 50$	–	3.278	0.65%	$8.13E-3$	0.1	0.996	50	850	42,500

(c) show the optimized topologies for the randomized algorithm for a single representative trial (one trial is one run of the numerical experiment) for each τ_{step}, and Fig. 4(d) shows the convergence histories of the objective function for all cases for the representative trial. Since the sample load cases are generated randomly at each step, the final optimized topology and its compliance vary slightly with each trial. Therefore, the associated results in Table 2 are averaged over 5 trials. For the randomized algorithm, we also report the standard deviations of the compliance for 5 trials — the standard deviations for all the cases are small, indicating that the randomized algorithm generates results with similar compliance in every instance. For this example, the standard algorithm leads to a lower compliance and simpler topology. However, the randomized algorithm for both tolerances uses substantially fewer solves for final topologies similar to the one obtained with the standard algorithm. For $\tau_{\text{step}} = 0.1$, the number of linear systems to solve is 27 times fewer than for the standard algorithm (113,184 solves vs. 4170 solves), and the convergence of the optimization is more rapid. The final topologies obtained for the two tolerances and the compliance for each case (see Fig. 4) suggest that the tolerance in the damping scheme has a minor influence on the final results in the density-based method. The relative differences between the compliance for the standard algorithm and those for the randomized algorithm are 1.79% for $\tau_{\text{step}} = 0.05$ and 2.45% for $\tau_{\text{step}} = 0.1$. It seems that smaller τ_{step} leads to slower convergence but a slightly better compliance: the optimization with $\tau_{\text{step}} = 0.05$ takes an average 1052 steps to converge while the one with $\tau_{\text{step}} = 0.1$ takes an average 695 steps. Therefore, the latter is more computationally efficient with only $N_{\text{solve}} = 4170$ compared with $N_{\text{solve}} = 6312$ for the former one.

Next, we check the quality of the estimated gradients by plotting the angle between the true gradient and the estimated gradient. The very small angles show that the estimated gradient is about as effective as the true gradient, and that the negative estimated gradient is a descent direction. As shown in Fig. 4(e) and (f), we plot the angle (θ), and the cosine of the angle ($\cos\theta$), between the gradient and the estimated gradient for each of the two tolerances for one representative trial. The moving averages (over 50 steps) of $\cos\theta$, in both cases, are close to 1.0.

The next study demonstrates the influence of sample sizes on optimization results and the reduction of the computational cost by using the randomized algorithm. In this study, $\tau_{\text{step}} = 0.1$. Fig. 5 gives the compliances for sample sizes $n = 4, 5, 6, 7, 20, 50$ and final topologies (from one representative trial). Each data point in Fig. 5 is obtained by averaging 5 trials, and the data are summarized in Table 2. Several observations can be made based on Fig. 5 and Table 2. For one, the randomized algorithm leads to similar optimal topologies and compliances compared with the standard algorithm. For $n = 4$, N_{solve} is significantly smaller than for the standard algorithm, by a factor of 45 on average. Moreover, the compliance improves as we increase n, indicating that larger n offers better estimation during the optimization and ultimately yields stiffer optimal structures. However, N_{solve} also increases as we use more sample load cases. From the optimized structures and compliances, it seems that $n = 6$ is sufficient for this problem with greatly reduced computational cost. The compliance differs from the standard algorithm by only 2.45%. Table 2 shows that the cosine of the average angle between the gradient and the estimated gradient, $\cos\bar{\theta}$, ranges from 0.959 and 0.996 for various sample sizes. Using the randomized algorithm, we can almost fully recover the original optimal results by either increasing the sample size (e.g., $n = 50$) or choosing smaller τ_{step}, as both methods lead to highly accurate designs.

4.1.2. Example 2: Truss topology optimization with ground structure method

As example 2, we study again the problem presented in Fig. 3, but this time using the GSM. We demonstrate that our approach greatly reduces the total number of linear solves. In addition, we investigate the sensitivity of the results

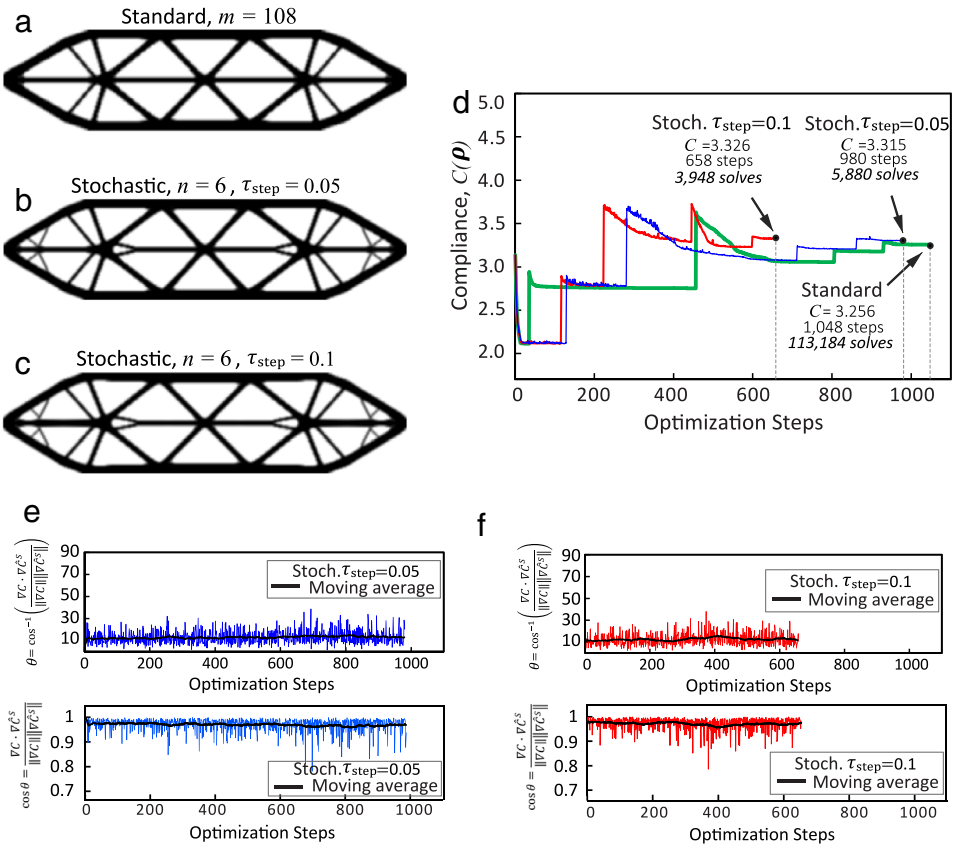

Fig. 4. Results for Example 1 using the density-based method with 25,600 quadrilateral elements and 52,002 degrees of freedom (DOFs). (a) The optimized topology obtained by the standard algorithm [4]; (b) the optimized topology obtained by the randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.05$ (one representative trial); (c) the optimized topology obtained by the randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.1$ (one representative trial); (d) the convergence of the compliance for above cases; (e) the angle, and the cosine of the angle, between the gradient ∇C_x and the estimated gradient $\nabla \widehat{C}_x^S$ for the randomized case with $\tau_{\text{step}} = 0.05$ demonstrate that the directions are aligned. (f) the angle, and the cosine of the angle, between ∇C_x and $\nabla \widehat{C}_x^S$ for the randomized case with $\tau_{\text{step}} = 0.1$ demonstrate that the directions are aligned.

to τ_{step} as well as the influence of the discrete filter on final solutions of the randomized algorithm. We use a full-level GS (16×4 grid) with 2196 non-overlapped bars to discretize the domain [47,48]. The optimal topology and the convergence of the compliance for the standard algorithm are shown in Fig. 6(a) and (d). The optimal compliance for the standard algorithm, $C(\boldsymbol{x}^*) = 4.219$, is obtained in 406 steps and $N_{\text{solve}} = 43,848$. For the randomized algorithm, we compare two cases using different tolerances in the damping scheme, $\tau_{\text{step}} = 0.05$ and $\tau_{\text{step}} = 0.1$. The results are summarized in Table 3, averaged over 5 trials. The optimized structures and the convergence of the compliance for a single representative trial are shown in Fig. 6(b)–(d).

In contrast to the results for the density-based method in Section 4.1.1, the results for the GSM indicate that the choice of τ_{step} has a significant impact on the optimized structure. Although $\tau_{\text{step}} = 0.05$ results in a larger number iterations/optimization steps and a larger number of linear system solves, N_{solve}, compared with $\tau_{\text{step}} = 0.1$, its final structure is simpler, has slightly lower compliance, and is similar to the structure obtained by the standard algorithm. This suggests that $\tau_{\text{step}} = 0.05$ is a better choice for this example. The convergence rate for the randomized algorithm is about the same as for the standard algorithm. Since the standard optimization problem for the GSM is convex, its solution is the global minimum; hence, the compliances obtained with the randomized algorithms will be larger than or equal to those obtained with the standard algorithm. However, the relative differences in compliance are very small, only 0.09% and 0.35% on average, and the randomized algorithm achieves these results with considerably less computational effort ($N_{\text{solve}} = 5627$ and 2322 on average for the randomized algorithm versus $N_{\text{solve}} = 43848$ for the standard algorithm). Symmetry was not enforced for the GSM in the present study. To show the accuracy of

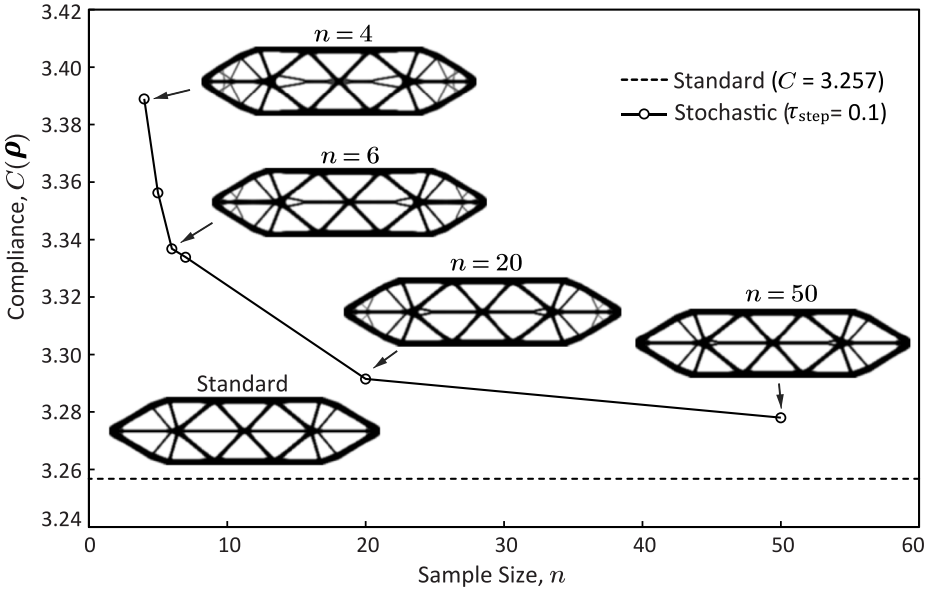

Fig. 5. Study of sample sizes ($n = 4, 5, 6, 7, 20, 50$) versus the resulting final compliance (or end-compliance) using the randomized algorithm in Example 1. The final topologies (from representative trials) are included.

Table 3
Results for Example 2 (GSM), averaged over 5 trials.

GSM	$C(\boldsymbol{x}^*)$	$C(\tilde{\boldsymbol{x}}^S)$	ΔC	$Dev_S(C(\tilde{\boldsymbol{x}}^S))$	τ_{step}	$\cos\bar{\theta}$	n	N_{step}	N_{solve}
Standard	4.219	–	–	–	–	–	–	406	43,848
Stoch. $\tau_{\text{step}} = 0.05$	–	4.222	0.09%	$6.84E-4$	0.05	0.911	6	938	5627
Stoch. $\tau_{\text{step}} = 0.1$	–	4.231	0.35%	$3.36E-3$	0.1	0.912	6	387	2322
Stoch. w/ Filter $\tau_{\text{step}} = 0.05$	–	**4.222**	**0.09%**	$\mathbf{9.93E-4}$	**0.05**	**0.978**	**6**	**907**	**5442**

estimated gradient of compliance for both tolerances, we plot $\cos\theta$ between the gradient and the estimated gradient for one representative trial in Fig. 6(e). The cosine of the average angle, $\cos\bar{\theta}$, for the two randomized cases are 0.911 and 0.912, indicating that the estimated gradients of both randomized cases are quite accurate and lead to effective optimization steps.

Next, we also include the discrete filter [43] in the randomized algorithm and study its influence on the optimized results. Based on previous observations, we choose $\tau_{\text{step}} = 0.05$ and $n = 6$. We use the following parameters for the discrete filter (see Section 2.3): $n_f = 10$ and $\alpha_f = 0.0001$. Fig. 7(a) shows the final topology obtained by the randomized algorithm with the discrete filter in one representative trial (cf. Fig. 6(b), which was obtained without the filter). The convergence of the compliance and $\cos\theta$ in each step are shown in Fig. 7(b)–(c). The results, averaged over five trials, are summarized in Table 3. Similar to Example 1, the standard deviations provided in the table for the randomized algorithm show that the compliance of the optimal structure for each trial is almost identical. The discrete filter combined with the randomized algorithm leads to a simpler final topology. From Fig. 7(c), the direction of the estimated gradient seems to be more accurate than the ones without the discrete filter (Fig. 6(e)), which is confirmed by $\cos\bar{\theta} = 0.978$. This indicates that the removal of some non-useful members helps to limit the collateral effects of the stochastic estimates. As compared to the standard algorithm, the discrete filter combined with the randomized algorithm not only leads to a reduction in the number of linear system solves ($N_{\text{solve}} = 5442$ versus $N_{\text{solve}} = 43848$), but the size of the linear systems also keeps decreasing due to the discrete filter, which further improves the computational efficiency.

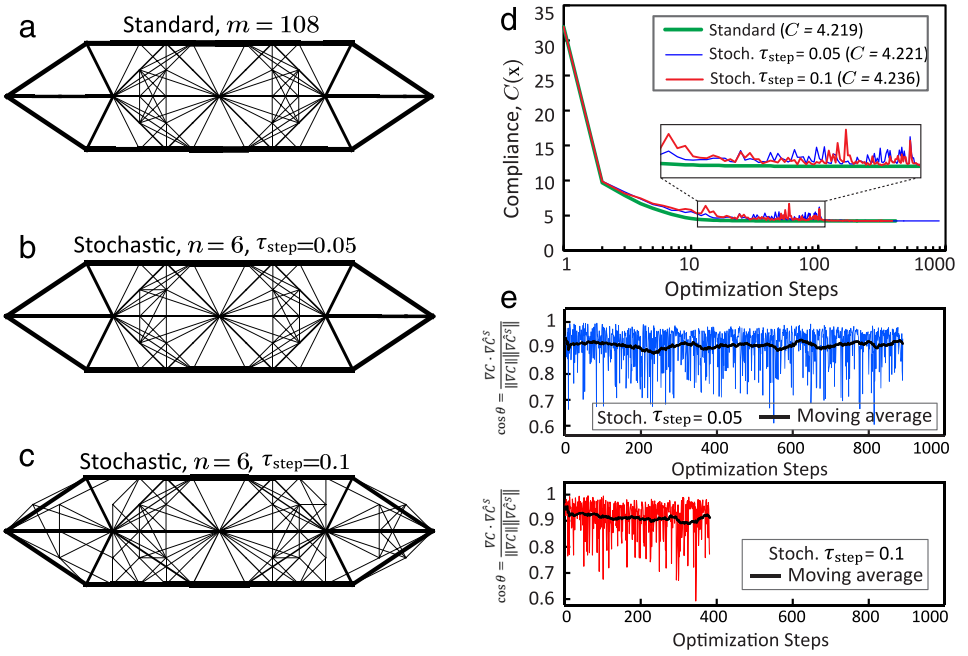

Fig. 6. Results for the GSM (without the discrete filter) for Example 2 with a full-level GS (16 × 4 grid) and 2196 non-overlapped bars. (a) The optimized structure obtained by the standard algorithm; (b) the optimized structure obtained by the randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.05$; (c) the optimized structure obtained by the randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.1$; (d) the convergence of the compliance for all above cases; (e) the cosine of θ between the gradient ∇C_x and the estimated gradient $\nabla \widehat{C}_x^S$ for the randomized algorithm ($\tau_{\text{step}} = 0.05$ and $\tau_{\text{step}} = 0.1$).

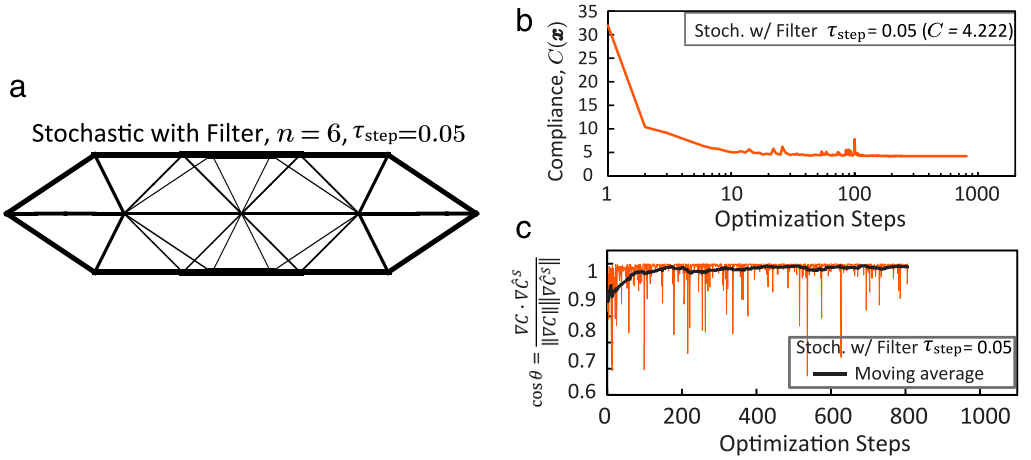

Fig. 7. Results for the GSM with the discrete filter for Example 2. (a) The optimized structure obtained from randomized algorithm with $n = 6$, $\tau_{\text{step}} = 0.05$, and $\alpha_f = 0.0001$; (b) the convergence of the compliance; (c) the cosine of θ between the gradient ∇C_x and the estimated gradient $\nabla \widehat{C}_x^S$ for the randomized case.

4.2. Example 3: Three-dimensional bridge design with density-based method

In this section, we demonstrate the quality of the design and the great reduction in computational work of the proposed randomized algorithm using a 3D bridge design in the non-convex, continuum topology optimization framework. The design domain, load, and boundary conditions are shown in Fig. 8(a). A total of 144 equal-weighted load cases are applied to the bridge deck (non-designable layer). Based on the structural symmetry, we optimize a

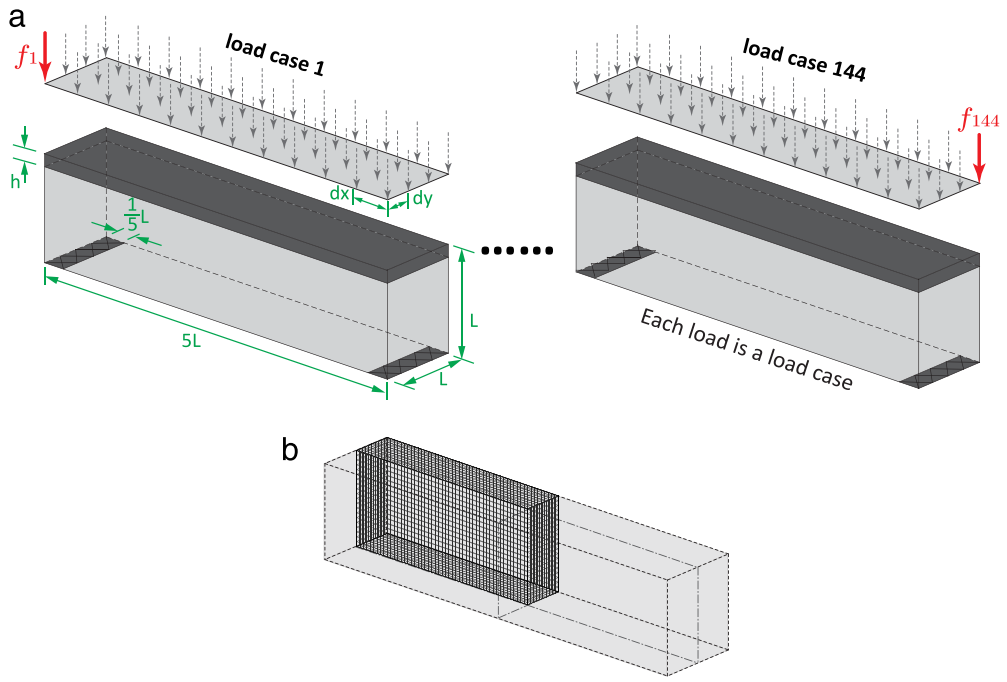

Fig. 8. Three-dimensional bridge design with (a) the geometry, load and boundary conditions; (b) the quarter domain is modeled by a mesh with 10,000 brick elements and 35,343 DOFs.

quarter of the domain, as shown in Fig. 8(b), which reduces the number of load cases to $m = 36$. We use 10,000 brick elements to discretize the quarter domain, resulting in 35,343 degrees of freedom (DOFs). We take $V_{\max} = 0.1 \times M$ (where $M = 10,000$) and use a quadratic density filter which takes the radius of 2.5 (see Section 1.1). The penalization factor for the continuation scheme takes values $p = 1, 2, 3, 4$ [46]. The randomized case uses the following parameter values: the sample size is chosen to be $n = 6$; the window size $n_{\text{step}} = 100$, $\gamma = 2$ and $\tau_{\text{step}} = 0.1$ (the effective step ratio is calculated in terms of $\boldsymbol{\rho}$). Fig. 9 shows the optimized structures obtained using the standard algorithm and the randomized algorithm. The results are summarized in Table 4.

The standard algorithm leads to final topology with $C(\boldsymbol{\rho}^*) = 542.1$ and converges in 968 steps. In each optimization step, we solve 36 linear systems (load cases), which leads to $N_{\text{solve}} = 34848$. Our randomized algorithm, while offering a nearly identical topology as the standard algorithm, drastically reduces the computational cost from 34,848 solves to 4662 solves and leads to even better compliance $C(\widehat{\boldsymbol{\rho}}^S) = 523.6$.

4.3. Example 4: Three-dimensional high-rise building design with ground structure method

To illustrate the effectiveness of the randomized algorithm combined with the discrete filter on a practical engineering design, we optimize a simplified 3D high-rise building (twisting tower) in the truss topology optimization framework. The domain of the twisting tower, given in Fig. 10, has 11 floors with the 1st floor fixed to the ground. The tower twists a full 90° from its base to its crown. To obtain constructible structures, we use a $4 \times 4 \times 11$ grid to discretize the domain followed by the generation of a level 7 initial GS, containing 3556 non-overlapping members [48]. The base mesh and the void zone are shown in Fig. 10(b). As shown Fig. 10(c), 77 equal-weighted load cases are applied to the building. Based on our previous studies, we choose $\tau_{\text{step}} = 0.05$ and $n = 6$. We apply the discrete filter for both the standard and the randomized algorithms; specifically, we choose $n_f = 10$, a small filter value ($\alpha_f = 0.0001$) used during optimization, and a larger filter value ($\alpha_f = 0.001$) in the final step to control the resolution of the final topology [43,51]. Fig. 11 shows the optimized structures obtained using the standard and randomized algorithms. A summary of the results is provided in Table 5.

The optimal compliance for the standard algorithm, $C(\boldsymbol{x}^*) = 4.388$, is obtained in 382 optimization steps and a total of $N_{\text{solve}} = 29414$ linear solves. With the randomized algorithm, we obtain a similar final structure as well as a

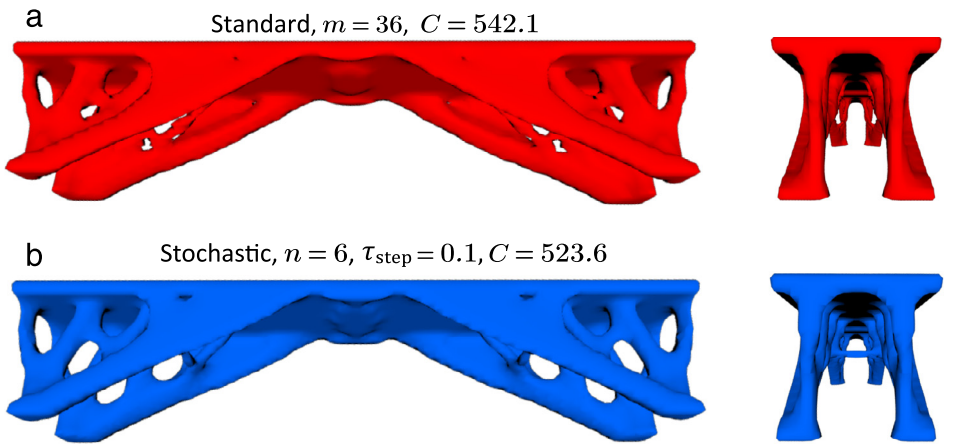

Fig. 9. Optimized structures of the 3D bridge design obtained from (a) the standard algorithm; (b) the randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.1$.

Table 4
Results for Example 3 (density-based bridge design).

Study 1	$C(\boldsymbol{\rho}^*)$	$C(\widehat{\boldsymbol{\rho}}^S)$	ΔC	τ_{step}	n	N_{step}	N_{solve}
Standard	542.1	–	–	–	–	968	34,848
Stochastic	–	523.6	−3.4%	0.1	6	777	4,662

Table 5
Results for Example 4 (tower design).

3D GSM	$C(\boldsymbol{x}^*)$	$C(\widehat{\boldsymbol{x}}^S)$	ΔC	τ_{step}	$\cos\bar{\theta}$	n	N_{step}	N_{solve}
Standard	4.388	–	–	–	–	–	382	29,414
Stochastic	–	4.408	0.46%	0.05	0.937	6	735	4,410

compliance that is only 0.46% higher than that obtained with the standard algorithm. These results are obtained at a drastically reduced computational cost, i.e., $N_{\text{solve}} = 4410$ linear solves. In addition to the reduction in the number of linear system solves obtained by the randomized algorithm, the discrete filter also reduces the size of the linear systems as the optimization proceeds, because the use of the filter scheme reduces the number of design variables, and hence the size of the stiffness matrices. This significantly decreases the CPU time and memory usage [51], contributing to great computational efficiency.

5. Concluding remarks and perspective

This paper proposes an efficient randomized optimization approach for topology optimization that drastically reduces the enormous computational cost of optimizing practical structural systems under many load cases while producing high-quality designs. We apply this approach to the nested minimum end-compliance topology optimization using both the density-based method (non-convex) and the ground structure method (convex) by minimizing a weighted sum/average of the compliance over many load cases. Minimizing the weighted average of the compliance over many load cases requires the solution of the state equations for each load case in every optimization step to compute the (weighted) compliance and its gradient. Because the objective function and its gradient can be defined as the traces of symmetric matrices, we use the Hutchinson trace estimator (which provides the lowest variance) combined with the sample average approximation technique, to estimate both quantities. This reduces the computational cost from $m \times N_{\text{step}}$ to roughly $n \times N_{\text{step}}$, where m is the number of load cases, and $n \ll m$ is the sample size. We further propose a damping scheme for the randomized algorithm, derived from simulated annealing, to obtain fast convergence. We discuss the algorithmic parameters for our scheme and provide some information on how to choose them.

The results for several generic examples and practical engineering designs demonstrate that the proposed randomized algorithm provides high-quality designs at a drastically reduced computational cost. Based on the limited

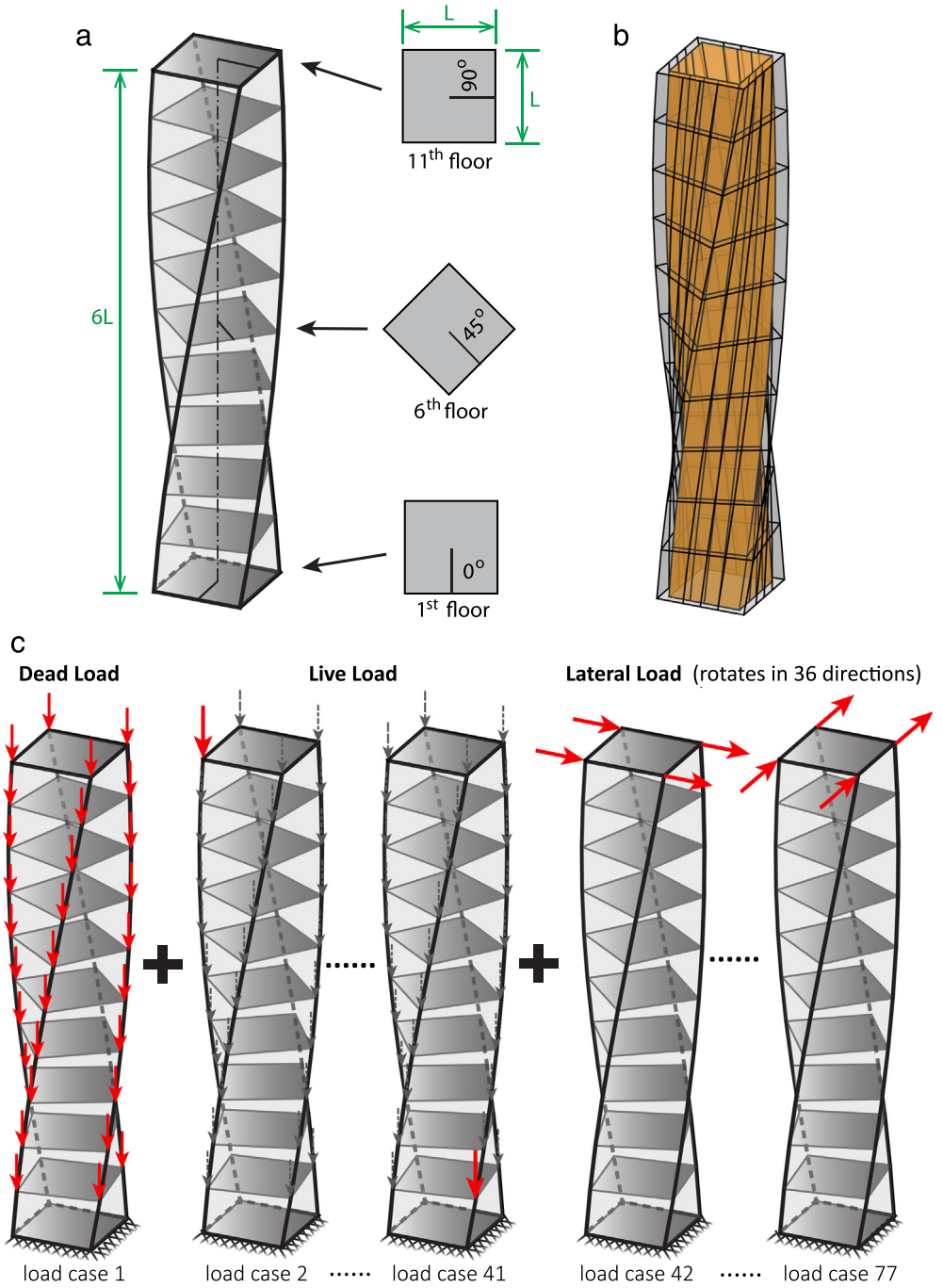

Fig. 10. Twisting tower (inspired by the Cayan tower [52]): (a) geometry; (b) base mesh with a void zone in the middle; (c) load and boundary conditions. One dead load case, 40 live load cases, and 36 lateral load cases (the lateral load is applied at 4 corners on the top floor and rotating in 36 directions).

number and size of examples investigated, we show that the proposed randomized algorithm substantially reduces the computational cost (e.g. by a factor of up to 45 in one of the examples), while obtaining a final compliance very close to that obtained by the standard algorithm. Our proposed damping scheme leads to fast convergence of the optimization scheme. In addition, the combination of the discrete filter with the randomized algorithm leads to fewer linear solves with smaller systems, resulting in great computational efficiency.

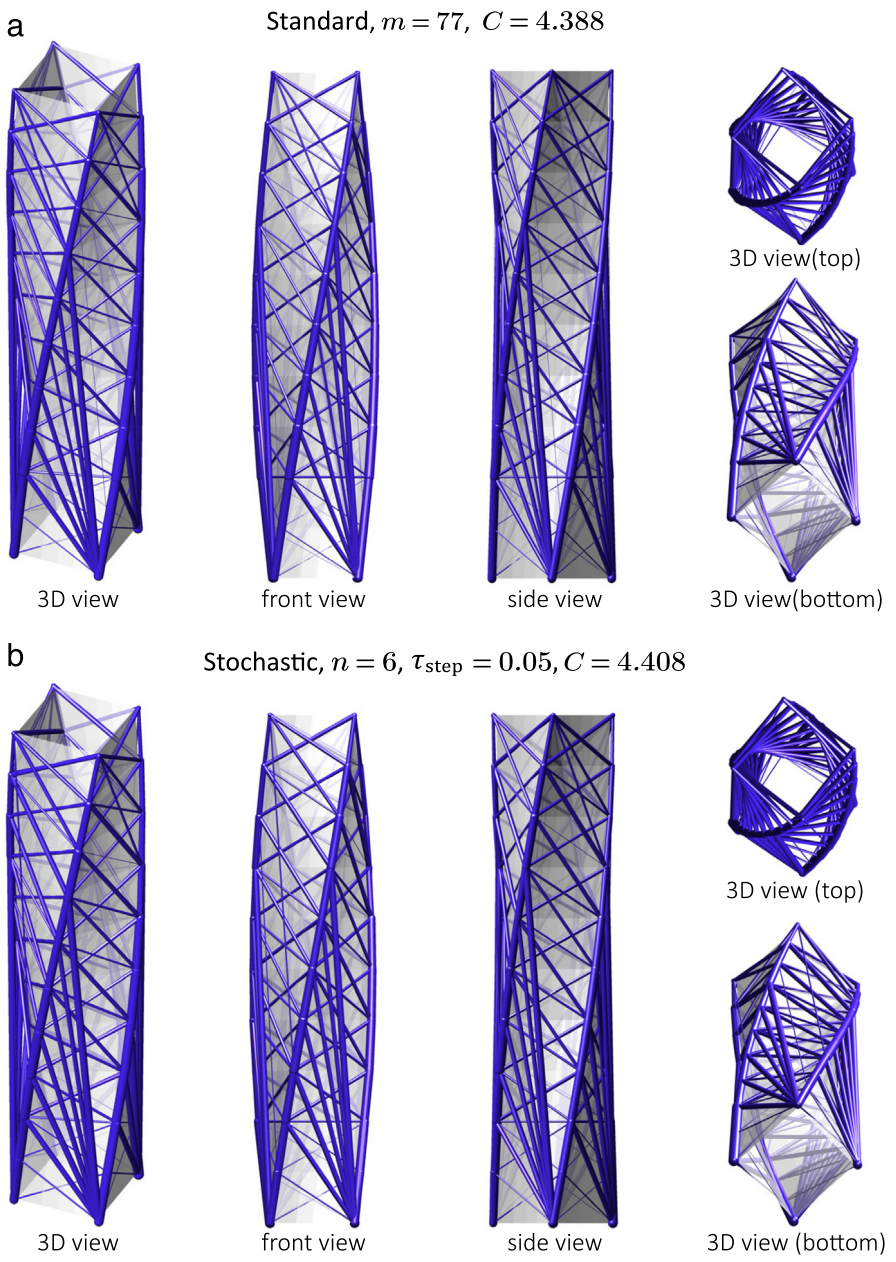

Fig. 11. Optimized structures of the 3D twisting tower obtained from: (a) the standard algorithm; (b) the proposed randomized algorithm with $n = 6$ and $\tau_{\text{step}} = 0.05$.

The proposed randomized algorithm is flexible and can be combined with any gradient-based update scheme. In this paper, we combine this technique with the optimality criteria, but other optimization methods can be used, such as the method of moving asymptotes (MMA)—see, for example, the book by Christensen and Klarbring [49].

There are several important directions for future research. Although we have provided an important proof-of-concept, many questions remain open. One important question concerns optimal choices of parameters for both overall computational cost and quality of design. In this case, we could consider dynamically varying the parameters. For example, if necessary, the quality of designs may be improved by increasing the sample size (n) as the minimum point is approached. Another important question concerns the choice among randomized optimization methods/approaches,

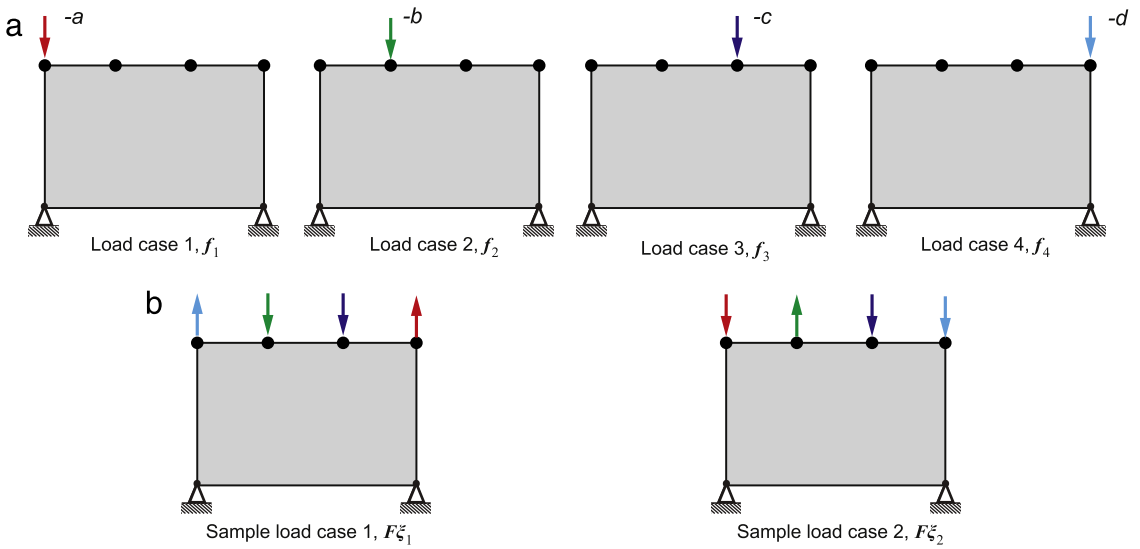

Fig. A.12. A simply supported rectangle with (a) 4 equally-weighed load cases, $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_4$, acting independently on different nodes; and (b) two sample load cases based on random vectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$.

both in terms of the overall methods as well as the choices in estimates and approximations. We need to further analyze what topology optimization problems can be solved efficiently using this approach. For non-self adjoint problems, instead of minimizing quadratic forms, we may use randomization for the minimization of similar problems, e.g. $\min_p \|\boldsymbol{C}^T(\boldsymbol{A}(p))^{-1}\boldsymbol{B} - \boldsymbol{D}\|$ where $\boldsymbol{A}(p)$ is not symmetric. This requires randomizing over both the columns of $\boldsymbol{B}$ and the columns of $\boldsymbol{C}$ [53]. Finally, it would be useful to prove convergence to either a local or global minimum of the topology optimization problem under appropriate conditions. Thus, we hope that the present paper will motivate further the use of stochastic sampling in various areas connected to (large scale) topology optimization.

Acknowledgments

The authors G.H. Paulino and X.S. Zhang acknowledge the financial support from the US National Science Foundation (NSF) under project #1559594 (formerly #1335160). The authors are also grateful for the endowment provided by the Raymond Allen Jones Chair at the Georgia Institute of Technology. The work by E. de Sturler was supported in part by the grant NSF DMS 1217156. The information provided in this paper is the sole opinion of the authors and does not necessarily reflect the view of the sponsoring agencies.

Appendix A. Stochastic sampling of load cases

In this appendix, we present an illustration of the stochastic sampling of load cases. As shown in Fig. A.12(a), we consider a simply supported rectangle with 4 equally-weighed load cases, $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_4$, acting independently. Numbering the 4 nodes on the top of the rectangle from left to right as 1 to 4, we can express the 4 independent load vectors and load matrix $\boldsymbol{F}$ as

$$\boldsymbol{f}_1 = \begin{bmatrix} 0 \\ -a \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{f}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -b \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -c \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{f}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -d \end{bmatrix}, \quad \boldsymbol{F} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -b & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -c & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -d \end{bmatrix}. \tag{A.1}$$

From the Rademacher distribution, we randomly select 2 vectors, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, with the following form

$$\boldsymbol{\xi}_1 = \begin{bmatrix} -1 & 1 & 1 & -1 \end{bmatrix}^T \text{ and } \boldsymbol{\xi}_2 = \begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}^T. \tag{A.2}$$

Then, according to Eq. (16), we obtain the corresponding sample load cases as

$$\boldsymbol{F}\boldsymbol{\xi}_1 = \begin{bmatrix} 0 & a & 0 & -b & 0 & -c & 0 & d \end{bmatrix}^T \text{ and } \boldsymbol{F}\boldsymbol{\xi}_2 = \begin{bmatrix} 0 & -a & 0 & b & 0 & -c & 0 & -d \end{bmatrix}^T. \tag{A.3}$$

The two sample load cases are shown in Fig. A.12(b). We can see that in both sample load cases, all the load cases $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_4$ appear simultaneously but may act in opposing directions.

Appendix B. Nomenclature

n_f	Number of monitored steps for the discrete filter in the randomized algorithm
n_s	Frequency to select a new sample
n_{step}	Sample window size in the damping scheme
τ_{step}	Tolerance for the damping scheme
γ	Scale factor in the damping scheme
R	Effective step ratio in the damping scheme
α_{Top}	Resolution of the structural topology
α_i	Weighting factor for the ith load case
α_f	Discrete filter value
$\boldsymbol{f}_i$	External force vector for the ith load case
$\boldsymbol{u}_i$	Displacement vector for the ith load case
$\boldsymbol{\xi}$	Random vector with its entries drawn from the Rademacher distribution
$\boldsymbol{\rho}^*$	Optimal solution obtained by the standard algorithm in the density-based method
$\boldsymbol{x}^*$	Optimal solution obtained by the standard algorithm in the GSM
$\hat{\boldsymbol{\rho}}^S$	Optimal solution obtained by the randomized algorithm in the density-based method
$\hat{\boldsymbol{x}}^S$	Optimal solution obtained by the randomized algorithm in the GSM
N_{solve}	Total number of linear system solves in the optimization process
$\boldsymbol{H}$	Filter matrix for density
$\overline{\boldsymbol{\rho}}$	Filtered density vector
p	Penalization parameter in the density-based method
d	Number of dimensions
E_0	Young's modulus of the solid material
$\boldsymbol{F}$	Weighted external force matrix
C	Weighted compliance (objective function)
$\hat{C}$	Estimated objective function by the randomized algorithm
$\boldsymbol{K}$	Global stiffness matrix
$\rho^{(e)}$	Density of element e (eth design variable in the density-based method)
$L^{(e)}$	Length of truss member e
M	Number of elements in the mesh/initial ground structure
m	Number of load cases
N	Number of nodes in the mesh/initial ground structure
n	Sample size
τ_{opt}	Tolerance for the optimization process
$\boldsymbol{U}$	Weighted displacement matrix
$v^{(e)}$	Volume of element e
$V_{\max}$	Prescribed maximum volume
$x^{(e)}$	Cross-sectional area of member e (eth design variable in the GSM)
$x^{\min}$, $x^{\max}$	Lower and upper bounds for the cross-sectional area of the member in the GSM
$\rho^{\min}$	Lower bound for density in the density-based method

References

[1] M. Sarkisian, Designing Tall Buildings: Structure as Architecture, second ed., Routledge, 2016.
[2] A. Diaz, M. Bendsøe, Shape optimization of structures for multiple loading conditions using a homogenization method, Struct. Optim. 4 (1) (1992) 17–22.

[3] M.P. Bendsøe, A. Ben-Tal, J. Zowe, Optimization methods for truss geometry and topology design, Struct. Optim. 7 (3) (1994) 141–159.
[4] M.P. Bendsøe, O. Sigmund, Topology Optimization: Theory, Methods, and Applications, in: Engineering, Springer, Berlin, Germany, 2003, p. 370.
[5] W. Achtziger, Multiple-load truss topology and sizing optimization: Some properties of minimax compliance, J. Optim. Theory Appl. 98 (2) (1998) 255–280.
[6] E. Haber, M. Chung, F. Herrmann, An effective method for parameter estimation with PDE constraints with multiple right-hand sides, SIAM J. Optim. 22 (3) (2012) 739–757.
[7] F. Roosta-Khorasani, K. van den Doel, U. Ascher, Stochastic algorithms for inverse problems involving PDEs and many measurements, SIAM J. Sci. Comput. 36 (5) (2014) S3–S22.
[8] R. Neelamani, C.E. Krohn, J.R. Krebs, J.K. Romberg, M. Deffenbaugh, J.E. Anderson, Efficient seismic forward modeling using simultaneous random sources and sparsity, Geophysics 75 (6) (2010) WB15–WB27.
[9] H. Avron, P. Maymounkov, S. Toledo, Blendenpik: Supercharging LAPACK's least-squares solver, SIAM J. Sci. Comput. 32 (3) (2010) 1217–1236.
[10] M. Hutchinson, A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines, Commun. Stat.-Simul. Comput. 18 (3) (1989) 1059–1076.
[11] J.A. Tropp, User-friendly Tools for Random Matrices: An Introduction. Tech. Rep.,, DTIC Document, 2012.
[12] N. Halko, P. Martinsson, J.A. Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, SIAM Rev. 53 (2) (2011) 217–288.
[13] P. Drineas, M.W. Mahoney, RandNLA: randomized numerical linear algebra, Commun. ACM 59 (6) (2016) 80–90.
[14] H. Avron, S. Toledo, Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix, J. ACM 58 (2) (2011) Article 8.
[15] F. Roosta-Khorasani, U. Ascher, Improved bounds on sample size for implicit matrix trace estimators, Found. Comput. Math. 15 (5) (2015) 1187–1212.
[16] A.K. Saibaba, A. Alexanderian, I.C. Ipsen, Randomized Matrix-free Trace and Log-Determinant Estimators, 2016. ArXiv preprint arXiv: 1605.04893.
[17] M. Carrasco, B. Ivorra, R. Lecaros, Á. Ramos del Olmo, An expected compliance model based on topology optimization for designing structures submitted to random loads, Differ. Equ. Appl. 4 (1) (2012) 111–120.
[18] P.D. Dunning, H.A. Kim, Robust topology optimization: Minimization of expected and variance of compliance, AIAA J. 51 (11) (2013) 2656–2664.
[19] M. Tootkaboni, A. Asadpoure, J.K. Guest, Topology optimization of continuum structures under uncertainty–a polynomial chaos approach, Comput. Methods Appl. Mech. Engrg. 201 (2012) 263–275.
[20] B. Bourdin, Filters in topology optimization, Internat. J. Numer. Methods Engrg. 50 (9) (2001) 2143–2158.
[21] M.P. Bendsøe, Optimal shape design as a material distribution problem, Struct. Optim. 1 (4) (1989) 193–202.
[22] M. Zhou, G. Rozvany, The COC algorithm, Part II: Topological, geometrical and generalized shape optimization, Comput. Methods Appl. Mech. Engrg. 89 (1–3) (1991) 309–336.
[23] M. Stolpe, K. Svanberg, An alternative interpolation scheme for minimum compliance topology optimization, Struct. Multidiscip. Optim. 22 (2) (2001) 116–124.
[24] J. Petersson, A finite element analysis of optimal variable thickness sheets, SIAM J. Numer. Anal. 36 (6) (1999) 1759–1778.
[25] K. Svanberg, On local and global minima in structural optimization, in: E. Gallhager, R.H. Ragsdell, O.C. Zienkiewicz (Eds.), New Directions in Optimum Structural Design, John Wiley and Sons, Chichester, 1984, pp. 327–341.
[26] W. Achtziger, Topology optimization in structural mechanics, in: Topology Optimization of Discrete Structures, in: International Centre for Mechanical Sciences, vol. 374, Springer Vienna, 1997, pp. 57–100.
[27] A.A. Groenwold, L. Etman, On the equivalence of optimality criterion and sequential approximate optimization methods in the classical topology layout problem, Internat. J. Numer. Methods Engrg. 73 (3) (2008) 297–316.
[28] A. Shapiro, D. Dentcheva, A. Ruszczyński, Lectures on Stochastic Programming: Modeling and Theory, in: MPS-SIAM Series on Optimization, Philadelphia, 2009.
[29] C.M. Grinstead, J.L. Snell, Introduction to Probability, second ed., American Mathematical Society, 2006.
[30] A.J. Kleywegt, A. Shapiro, T. Homem-de Mello, The sample average approximation method for stochastic discrete optimization, SIAM J. Optim. 12 (2) (2001) 479–502.
[31] H. Robbins, S. Monro, A stochastic approximation method, Ann. Math. Stat. 22 (3) (1951) 400–407.
[32] N.N. Schraudolph, Local gain adaptation in stochastic gradient descent, in: Intl. Conf. Artificial Neural Networks, vol. 2, IEE, Edinburgh, Scotland, 1999, pp. 569–574.
[33] S.V.N. Vishwanathan, N.N. Schraudolph, M.W. Schmidt, K.P. Murphy, Accelerated training of conditional random fields with stochastic gradient methods, in: Proceedings of the 23rd International Conference on Machine Learning, ACM, Pittsburgh, PA, 2006, pp. 969–976.
[34] N.N. Schraudolph, J. Yu, S. Günter, A stochastic quasi-Newton method for online convex optimization, in: 11th Intl. Conf. on Artificial Intelligence and Statistics (AIstats), Soc. for Artificial Intelligence and Statistics, 2007, pp. 433–440.
[35] A. Nemirovski, A. Juditsky, G. Lan, A. Shapiro, Robust stochastic approximation approach to stochastic programming, SIAM J. Optim. 19 (4) (2009) 1574–1609.
[36] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.
[37] P. Salamon, P. Sibani, R. Frost, Facts, Conjectures, and Improvements for Simulated Annealing, in: SIAM Monographs on Mathematical Modeling and Computation, SIAM, Philadelphia, 2002.
[38] L. Bottou, F.E. Curtis, J. Nocedal, Optimization methods for large-scale machine learning, 2016. ArXiv preprint arXiv:1606.04838.

[39] K. Svanberg, The method of moving asymptotes- a new method for structural optimization, Internat. J. Numer. Methods Engrg. 24 (2) (1987) 359–373.
[40] H. Nguyen-Xuan, A polytree-based adaptive polygonal finite element method for topology optimization, Internat. J. Numer. Methods Engrg. 110 (10) (2017) 972–1000.
[41] S. Zargham, T.A. Ward, R. Ramli, I.A. Badruddin, Topology optimization: a review for structural designs under vibration problems, Struct. Multidiscip. Optim. 53 (6) (2016) 1157–1177.
[42] J. Lin, Y. Guan, G. Zhao, H. Naceur, P. Lu, Topology optimization of plane structures using smoothed particle hydrodynamics method, Internat. J. Numer. Methods Engrg. 110 (8) (2017) 726–744.
[43] A.S. Ramos Jr., G.H. Paulino, Filtering structures out of ground structures –a discrete filtering tool for structural design optimization, Struct. Multidiscip. Optim. 54 (2016) 95–116.
[44] C. Talischi, G.H. Paulino, A. Pereira, I.F.M. Menezes, PolyTop: A Matlab implementation of a general topology optimization framework using unstructured polygonal finite element meshes, Struct. Multidiscip. Optim. 45 (3) (2012) 329–357.
[45] K. Liu, A. Tovar, An efficient 3D topology optimization code written in MATLAB, Struct. Multidiscip. Optim. 50 (6) (2014) 1175–1196.
[46] T. Zegard, G.H. Paulino, Bridging topology optimization and additive manufacturing, Struct. Multidiscip. Optim. 53 (1) (2016) 175–192.
[47] T. Zegard, G.H. Paulino, GRAND–Ground structure based topology optimization for arbitrary 2D domains using MATLAB, Struct. Multidiscip. Optim. 50 (5) (2014) 861–882.
[48] T. Zegard, G.H. Paulino, GRAND3–Ground structure based topology optimization for arbitrary 3D domains using MATLAB, Struct. Multidiscip. Optim. 52 (6) (2015) 1161–1184.
[49] P. Christensen, A. Klarbring, An Introduction to Structural Optimization, Vol. 153, Springer Science & Business Media, 2009.
[50] X. Zhang, S. Maheshwari, A.S. Ramos Jr., G.H. Paulino, Macroelement and macropatch approaches to structural topology optimization using the ground structure method, ASCE J. Struct. Eng. 142 (11) (2016) 04016090.
[51] X. Zhang, A.S. Ramos Jr., G.H. Paulino, Material nonlinear topology optimization using the ground structure method with a discrete filtering scheme, Struct. Multidiscip. Optim. 55 (6) (2017) 2045–2072.
[52] Cayan Tower (Infinity Tower). URL: http://cayan.net.
[53] S.S. Aslan, E. de Sturler, M.E. Kilmer, Randomized approach to nonlinear inversion combining simultaneous random and optimized sources and detectors, arXiv:1706.05586.