



Title	Latent Crossover for Data-Driven Multifidelity Topology Design
Author(s)	Kii, Taisei; Yaji, Kentaro; Fujita, Kikuo et al.
Citation	Journal of Mechanical Design. 2024, 146(5), p. 051713
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/95639">https://hdl.handle.net/11094/95639</a>
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/378717189>

# Latent Crossover for Data-Driven Multifidelity Topology Design

Article in *Journal of Mechanical Design* · March 2024

DOI: 10.1115/1.4064979

CITATIONS

0

READS

42

5 authors, including:



Kentaro Yaji  
Osaka University

95 PUBLICATIONS 910 CITATIONS

[SEE PROFILE](#)



Kikuo Fujita  
Osaka University

350 PUBLICATIONS 1,857 CITATIONS

[SEE PROFILE](#)



## Taisei Kii

Department of Mechanical Engineering,  
 Osaka University,  
 Osaka 565-0871, Japan  
 e-mail: kii@syd.mech.eng.osaka-u.ac.jp

## Kentaro Yaji<sup>1</sup>

Department of Mechanical Engineering,  
 Osaka University,  
 Osaka 565-0871, Japan  
 e-mail: yaji@mech.eng.osaka-u.ac.jp

## Kikuo Fujita

Department of Mechanical Engineering,  
 Osaka University,  
 Osaka 565-0871, Japan  
 e-mail: fujita@mech.eng.osaka-u.ac.jp

## Zhenghui Sha

Walker Department of Mechanical Engineering,  
 University of Texas Austin,  
 Austin, TX 78712  
 e-mail: zsha@austin.utexas.edu

## Carolyn Conner Seepersad

George W. Woodruff School of Mechanical  
 Engineering,  
 Georgia Institute of Technology,  
 Atlanta, GA 30318  
 e-mail: carolyn.seepersad@me.gatech.edu

# Latent Crossover for Data-Driven Multifidelity Topology Design

*Topology optimization is one of the most flexible structural optimization methodologies. However, in exchange for its high level of design freedom, typical topology optimization cannot avoid multimodality, where multiple local optima exist. This study focuses on developing a gradient-free topology optimization framework to avoid being trapped in undesirable local optima. Its core is a data-driven multifidelity topology design (MFTD) method, in which the design candidates generated by solving low-fidelity topology optimization problems are updated through a deep generative model and high-fidelity evaluation. As its key component, the deep generative model compresses the original data into a low-dimensional manifold, i.e., the latent space, and randomly arranges new design candidates over the space. Although the original framework is gradient free, its randomness may lead to convergence variability and premature convergence. Inspired by a popular crossover operation of evolutionary algorithms (EAs), this study merges the data-driven MFTD framework and proposes a new crossover operation called latent crossover. We apply the proposed method to a maximum stress minimization problem in 2D structural mechanics. The results demonstrate that the latent crossover improves convergence stability compared to the original data-driven MFTD method. Furthermore, the optimized designs exhibit performance comparable to or better than that in conventional gradient-based topology optimization using the  $P$ -norm measure. [DOI: 10.1115/1.4064979]*

**Keywords:** topology optimization, deep generative model, latent crossover, maximum stress minimization, data-driven design, machine learning, multiobjective optimization

## 1 Introduction

Topology optimization, first proposed by Bendsoe and Kikuchi [1], enables the determination of an optimized material distribution for a structural optimization problem and offers a high level of design freedom [2]. While this attractive feature makes it applicable to various structural design problems, topology optimization faces challenges with multimodality, where multiple local optima exist in the solution space [3]. That is, gradient-based optimizers used in conventional topology optimization methods may fall into low-performance local optima. This intractable characteristic is often seen in strongly nonlinear problems, e.g., minimax problems; thus, it is challenging to obtain structures that exhibit high levels of performance.

One of the standard ways to overcome the problem of multimodality in engineering optimization applications is evolutionary algorithms (EAs) since they are gradient-free [4]. An EA, such as the genetic algorithm, mimics the evolutionary mechanisms of living organisms, and solutions are represented as strings of

genes. The solution search is performed by applying three basic genetic operations: selection, crossover, and mutation, to a population of individuals. Each iteration of these genetic operations is referred to as a generation. The selection is an operation that retains individuals with relatively better objective function values in the population for the next generation. The crossover is an operation that partially exchanges genes between selected individuals to generate new individuals (offspring) that inherit traits from old ones (parents). However, if some individuals in the population have significantly higher fitness than others in the early stages of the search, they may weed out others by selection and crossover, leading to a loss of diversity and a high probability of premature convergence [5]. The mutation is an operation that introduces new genes into the population by changing a portion of the genes of selected individuals, which helps maintain diversity in the population. Several methods [6–9] have been proposed to solve topology optimization problems using EAs, taking advantage of their gradient-free nature. While they can perform a global search for strongly nonlinear problems, Sigmund [10] has pointed out issues with EA-based topology optimization. That is, topology optimization problems often require a large number of design variables, and the computational cost of the EA increases exponentially with the number of design variables due to the so-called *curse of dimensionality*.

As a potentially promising way to avoid the curse of dimensionality, some deep generative models can dramatically reduce the

<sup>1</sup>This article is a revised version of Paper No. DETC2023-111079 published in the ASME IDETC Design Automation Conference.

<sup>1</sup>Corresponding author.

Contributed by the Design Automation Committee of ASME for publication in the JOURNAL OF MECHANICAL DESIGN. Manuscript received July 14, 2023; final manuscript received February 17, 2024; published online March 28, 2024. Assoc. Editor: Shikui Chen.

dimensionality of the topology optimization problem. Variational autoencoders (VAEs) [11] and generative adversarial networks (GANs) [12] are popular deep generative models. In a VAE, an encoder is built to compress high-dimensional data into a low-dimensional manifold, called latent space, and maps it to a probability distribution, while a decoder reconstructs high-dimensional data from the latent space. In a GAN, a generator creates new data samples by starting from random noise and trying to produce data that are indistinguishable from real data. A discriminator, on the other hand, assesses these generated samples and tries to distinguish them from real data. As a review paper [13] mentioned, relevant studies on deep generative models for engineering design problems have increased dramatically in recent years. As pioneering work, Guo et al. [14] proposed a data-driven indirect design representation for high-dimensional design problems, which iteratively optimizes the latent space of a VAE as the design variable field. Oh et al. [15] proposed a design framework that iteratively trains a GAN to generate a variety of designs. Kazemi et al. [16] proposed a method to generate conceptual designs using a GAN for multi-physics topology optimization problems.

On the basis of combining EAs and deep generative models, Yaji et al. [17] proposed a data-driven multifidelity topology design (MFTD) method that enables gradient-free topology optimization. The basic idea of data-driven MFTD is that design candidates, generated by solving low-fidelity topology optimization problems, are iteratively updated using an EA that guides queries to a high-fidelity analysis model. The key to this framework builds upon data-driven topology design [18], incorporating a VAE as a crossover-like operation for each optimization step. The effectiveness of the framework was demonstrated for topology optimization problems that are hard to solve directly with conventional methods, such as minimax and turbulent flow problems. However, since the generative process in a VAE is based on a uniform random sampling in the latent space, it is expected that the effectiveness of the approach can be improved if the crossover operation is adopted based on EAs.

This article proposes a particular crossover operation based on EAs, called *latent crossover*, for the data-driven MFTD framework. Specifically, simplex crossover (SPX) [19]—a crossover operator of real-coded genetic algorithms (RCGAs) [20]—is used for latent crossover. We apply the proposed method to a maximum stress minimization problem of an L-bracket and verify the effectiveness of latent crossover, comparing it with the original data-driven MFTD. We also discuss its usefulness by comparing the results of the proposed method with those of gradient-based topology optimization (GTO) using the  $P$ -norm measure for the maximum stress minimization problem.

## 2 Latent Crossover

In data-driven MFTD [17], whose details are described in Sec. 3, the high-dimensional material distribution data of the design candidates are encoded by a VAE into low-dimensional real-valued

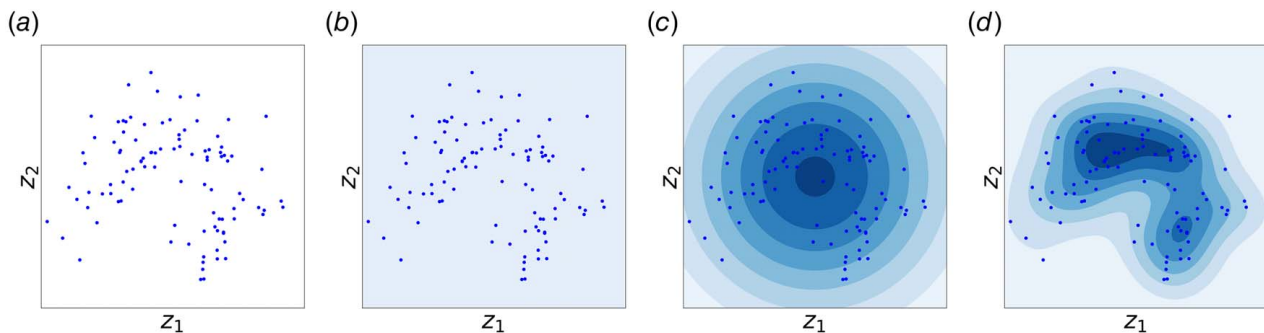
latent variables that correspond to EA genes, making the framework similar to the RCGA among EAs. Its high representation flexibility makes crossover more important in the RCGA than in the binary GA, and it has been the subject of various studies. For example, Kita and Yamamura [21] proposed a theory called the function specialization hypothesis concerning the selection and crossover operators in RCGAs, which includes the following ideas:

- The selection operator eliminates individuals with low fitness and, meanwhile, selects and replicates those with high fitness. Therefore, it is designed to narrow the population distribution gradually.
- The crossover operator transforms the distribution by combining parent individuals to generate offspring and is designed to retain the ability to generate new offspring for a finite population, but not to change the population distribution.

The design guideline for RCGA crossover operators uses statistics to concretize the aforementioned theory [22–24]. Specifically, the crossover operator should be designed to inherit statistics such as the mean vector and variance/covariance matrix of the population.

In data-driven MFTD, candidate solutions are generated through random sampling from the latent space of a VAE, so in terms of the genetic distribution and statistics of the population, we consider the probability distribution of the generated offspring. Figure 1 shows an example of the probability distribution for generating offspring in a two-dimensional latent space in the range from  $-2$  to  $+2$  for each dimension. The darker areas have a higher probability of generating offspring. Assuming that the distribution of the parent population, as shown in Fig. 1(a), is given, data-driven MFTD performs sampling based on a uniform distribution in the latent space, regardless of the distribution of the parent population. The resulting probability distribution of the generated offspring becomes the one shown in Fig. 1(b). It cannot be said that the statistics of the parent population are inherited. Although the use of a VAE as a deep generative model enables a crossover-like operation in the original data-driven MFTD, it is similar to crossover but cannot be considered strictly performing crossover because of random sampling. Since the input data follow a normal distribution in the latent space due to the nature of VAEs [11], generating offspring through sampling based on a normal distribution rather than a uniform distribution can be considered reasonable. However, as shown in Fig. 1(c), the probability of generated offspring does not follow the distribution of the parent population; therefore, the statistics of the parent population are not inherited in this case either. Based on the EA concept, preserving the diversity of the population helps prevent premature convergence, but crossover-like sampling from the latent space using random sampling can lead to an early loss of diversity in the population. This results in fluctuation in convergence and, in the worst case, failures to perform a global search, leading to the possibility of getting stuck in local optima.

As mentioned earlier, it is impossible to strictly inherit the statistical characteristics of the parent population through random



**Fig. 1** Probability distribution for generating offspring in 2D latent space. For the dots that are encoded parents, the darker shaded areas have a higher probability of generating offspring: (a) Parent individuals, (b) uniform random sampling, (c) normal random sampling, and (d) latent crossover.

sampling. According to its nature, a crossover operation generates offspring by targeting small areas for parents who are close together and large areas for those who are far apart [25]. Thus, applying latent crossover to the parent population in Fig. 1(a), the probability distribution of generated offspring is expected to become the one shown in Fig. 1(d). Therefore, it can be said that a crossover operation in the latent space, i.e., the latent crossover, is promising.

### 3 Framework

**3.1 Data-Driven MFTD With Latent Crossover.** Data-driven MFTD focuses on solving the following general multiobjective topology optimization problem:

$$\begin{aligned} & \underset{\gamma}{\text{minimize}} && [J_1(\gamma), J_2(\gamma), \dots, J_{r_0}(\gamma)] \\ & \text{subject to} && G_j(\gamma) \leq 0, \\ & && \gamma_e \in \{0, 1\}, e = 1, 2, \dots, N \end{aligned} \quad (1)$$

Here,  $J_i$  ( $i = 1, 2, \dots, r_0$ ) and  $G_j$  ( $j = 1, 2, \dots, r_c$ ) are the objective and constraint functions, respectively. The optimization problem defined by Eq. (1) is a 0-1 optimization problem with  $\gamma$  composed of  $N$  design variables. Since such a problem is a nonlinear mathematical optimization problem with a massive number of design variables, we adopt the concept of MFTD [26] and divide the problem of Eq. (1) into two procedures: low-fidelity optimization and high-fidelity evaluation. The low-fidelity optimization is formulated as an easily solvable pseudo-problem and is used to generate a variety of candidate solution structures by employing artificial parameters. On the other hand, the high-fidelity evaluation is used to evaluate the performance of candidate solutions using objective and constraint functions formulated by Eq. (1).

By using the MFTD approach and a deep generative model, data-driven MFTD iteratively updates solution candidates in a gradient-free manner similar to EAs. Note that the latent space is updated at

every optimization step. The schematic flow of the proposed data-driven MFTD with latent crossover is shown in Fig. 2, and the details of each step are explained here.

**Initial Data Generation:** For the original optimization problem of Eq. (1), we solve a low-fidelity optimization problem formulated as follows, which can be easily solved as a simple pseudo-problem:

$$\begin{aligned} & \underset{\gamma^{(k)}}{\text{minimize}} && \tilde{J}_i(\gamma^{(k)}) \\ & \text{subject to} && \tilde{G}_j(\gamma^{(k)}, \mathbf{s}^{(k)}) \leq 0, \\ & && \gamma_e^{(k)} \in [0, 1], e = 1, 2, \dots, N \end{aligned} \quad (2)$$

for given  $\mathbf{s}^{(k)}, k = 1, 2, \dots, K$

Here,  $\tilde{J}_i$  and  $\tilde{G}_j$  are the objective and constraint functions for the low-fidelity optimization problem, respectively, which can be easily computed by pseudo-functions. In addition,  $\mathbf{s} = [s_1, s_2, \dots, s_{N_{sd}}]$  represents the set of  $N_{sd}$  types of artificial design parameters called seeding parameters, and  $\mathbf{s}^{(k)}$  is the sample point of  $\mathbf{s}$ . For instance, the seeding parameters are defined as the maximum limit of constraints and optimization parameters such as the filter radius. By solving the relaxed low-fidelity optimization problem of Eq. (2) under various seeding parameter settings, where  $\gamma_e^{(k)}$  is relaxed to  $[0, 1]$ ,  $K$  kinds of promising and diverse material distributions are prepared as initial solutions.

**Evaluation:** The performance of candidate solutions is evaluated using a high-fidelity analysis model, which is used to compute the original multiple objective functions  $J_i$  and  $G_j$  in Eq. (1) with discrete  $\gamma_e$  binarized to  $\{0, 1\}$ .

**Selection:** As mentioned in Sec. 2, selection is a critical genetic operation in RCGAs. For problems as in Eq. (1), it is necessary to evaluate solutions using multiple objective functions and select

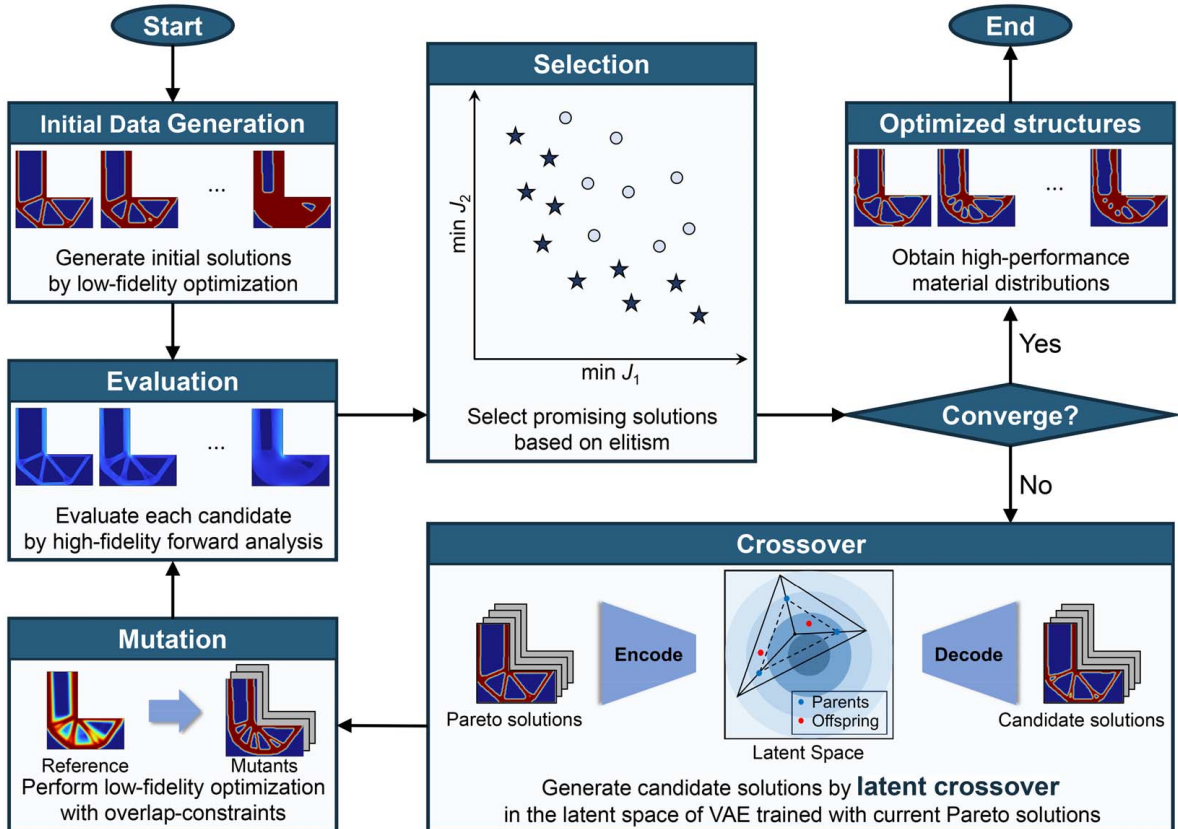


Fig. 2 Schematic illustration of data-driven MFTD with latent crossover



those to be preserved in the next generation. This article uses the nondominated sorting genetic algorithm II (NSGA-II) [27] strategy as a selection algorithm, which selects candidates in a multi-objective manner by ranking them based on the Pareto dominance relation using distances in the objective function space. The nondominated candidate solutions are selected based on performance evaluation values from the high-fidelity model, and then a set of Pareto solutions is constructed.

**Crossover:** A VAE is trained with the Pareto solution set as input to construct a latent space, where high-dimensional material distributions are encoded into low-dimensional latent variables. Here, it is important to note that the learning data are not accumulated iteratively but rather, a fixed number of data to be selected is predetermined, and a VAE is trained anew in each iteration. Latent crossover is performed using these latent variables to generate offspring in the latent space. Decoding the offspring generated by latent crossover yields new material distributions that inherit the characteristics of the input data, and candidate solutions are generated. The details of the VAE and the latent crossover operation are described in Secs. 3.2 and 3.3, respectively.

**Mutation:** The latent space of the VAE is constructed using the Pareto solution set of the current generation and corresponds to a subspace in which the solutions are distributed. Even if the mutation method of RCGAs, such as the nonuniform mutation operator [28], is applied in the latent space, its outcome is limited to a specific subspace against the whole solution space. This limitation exists because such a mutation only performs a local search in the subspace around the solutions distributed in the whole solution space. Thus, it cannot be expected to maintain the diversity of the population and prevent premature convergence, as discussed in Sec. 1.

Therefore, under the following constraint function, the low-fidelity optimization problem is solved using the same method as when generating initial data:

$$\tilde{G}_{\text{mut}}(\gamma^{(m)}) = \sum_{e=1}^N v_e \gamma_e^{(m)} \gamma_e^{\text{ref}(m)} \leq \tilde{G}_{\text{mut}}^{\text{max}} |D| \quad (3)$$

where  $m = 1, 2, \dots, N_{\text{mut}}$  is the number of mutants,  $v_e$  is the elemental volume,  $\tilde{G}_{\text{mut}}^{\text{max}}$  is a parameter that controls the degree of overlap between the reference material distribution  $\gamma^{\text{ref}(m)}$  and the design variable  $\gamma^{(m)}$ , and  $|D| = \sum_{e=1}^N v_e$  is the volume of  $D$ . In brief, the role of the constraint of Eq. (3) is to generate a different material distribution from  $\gamma_e^{\text{ref}(m)}$ .

This article uses the average value of material distributions in a given generation as a reference structure. This average distribution can be considered to be representative of the material distributions of the population. By solving the low-fidelity optimization problem with the constraint function of Eq. (3) and the reference structure, promising candidate solutions can be generated with unique features that are not present in the population. This approach enables a mutation-like operation, similar to the mutation in EAs, to maintain diversity and prevent premature convergence. It should be noted that the mutants added to the population through this operation are still limited to a specific subspace and may not search the whole solution space comprehensively.

**3.2 Variational Autoencoder.** Figure 3 shows the architecture of the VAE used in the numerical examples in Sec. 4. A total of 6400 input/output elements are combined into two eight-dimensional layers,  $\mu$  and  $\sigma$ , through a hidden layer of 512 dimensions.  $\mu$  is the mean vector, and  $\sigma$  is the variance vector of the latent variables  $\mathbf{z}$ . The following equation defines the latent variable vector  $\mathbf{z}$ :

$$\mathbf{z} = \mu + \sigma \odot \epsilon \quad (4)$$

where  $\odot$  is the operator that calculates the element-wise product and  $\epsilon$  is a vector of random numbers from the standard normal

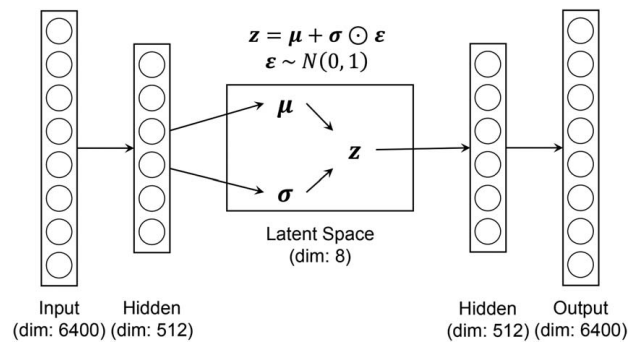


Fig. 3 Architecture of VAE

distribution. In VAEs, unsupervised learning is performed using the same dataset for both input and output, constructing the latent space. The following loss function  $L_{\text{VAE}}$  is used for the training:

$$L_{\text{VAE}} = L_{\text{recon}} + \rho L_{\text{KL}} \quad (5)$$

$$L_{\text{KL}} = -\frac{1}{2} \sum_{i=1}^{N_{\text{lt}}} (1 + \log(\sigma_i^2) - \mu_i^2 - \sigma_i^2) \quad (6)$$

where  $N_{\text{lt}}$  is the dimension of the latent space and  $\mu_i$  and  $\sigma_i$  are the  $i$ th elements of  $\mu$  and  $\sigma$ , respectively.  $L_{\text{recon}}$  is a reconstruction loss using mean squared error, and  $L_{\text{KL}}$  is known as the Kullback–Leibler (KL) divergence.  $\rho$  is the weight parameter that controls the influence of the KL divergence to regularize the latent space to the standard normal distribution.

Here, the VAE trained with the architecture shown in Fig. 3 and the loss function of Eq. (5) constructs a latent space following a single standard normal distribution. In contrast, there are advanced generative models such as Gaussian mixture VAEs [29] whose latent space follows multiple distributions. For instance, on the basis of this idea, Tsumoto et al. [30] have proposed a clustering method for solutions obtained through topology optimization. Due to the search mechanism of evolutionary algorithms, data-driven MFTD could involve the training data being distributed into several clusters, and Gaussian mixture VAEs might provide better learning accuracy compared to the standard VAEs in such cases. However, as mentioned in Sec. 3, since VAEs are trained anew at each iteration in the optimization process, this study employs the aforementioned standard VAEs in terms of computational cost and learning stability.

Compared to simple dimensionality reduction using autoencoders, VAEs are trained by incorporating probabilistic variation through  $\epsilon$ , allowing for the estimation of the given dataset distribution, and can be used as a deep generative model for continuous data generation. When using material distributions as a dataset for topology optimization, essential features within the dataset are extracted by compressing them into dramatically smaller latent variables. According to the standard normal distribution, latent variables do not take extremely large or small values. To represent all material distributions without excessive randomness, original data-driven MFTD [17] generates offspring by sampling uniform random numbers in  $[-4, 4]$ , which covers 99.7% of the data within  $\pm 4\sigma$ , for each latent variable. However, as mentioned in Sec. 2, generating offspring with a uniform probability distribution in the latent space, as shown in Fig. 1(b), regardless of the distribution of parent individuals, can be problematic. In this article, we perform latent crossover using the crossover operator explained in Sec. 3.3.

**3.3 Simplex Crossover.** Due to the high degree of freedom of representing genes as real-valued vectors, the RCGA has limited offspring that can be generated from selected parent individuals using crossover operators, such as the single-point crossover

commonly used in binary evolutionary algorithms. Several crossover operators for RCGAs [19,31,32] have been proposed to address this issue. This article uses the simplex crossover (SPX) [19] for a latent crossover operator. SPX is one of the multiparent crossover operators for RCGAs that generates offspring using three or more parent individuals and is consistent with the crossover design guidelines [22–24] as it inherits the average value and covariance matrix of the population.

When the search space is defined as the real  $n$ -dimensional space  $\mathbb{R}^n$ , where individuals are represented as vectors of real numbers, the algorithm for SPX is as follows:

- (1) Randomly select  $(n + 1)$  parent individuals  $P_0, P_1, \dots, P_n$  from the population.
- (2) Calculate the centroid  $G$  of the parent individuals as follows:

$$G = \frac{1}{n} \sum_{i=0}^n P_i \quad (7)$$

- (3) Calculate variables  $x_k$  and  $C_k$  for  $k = 0, 1, \dots, n$  as follows:

$$x_k = G + \varepsilon(P_k - G) \quad (8)$$

$$C_k = \begin{cases} 0 & (k = 0) \\ r_{k-1}(x_{k-1} - x_k + C_{k-1}) & (k = 1, \dots, n) \end{cases} \quad (9)$$

Here,  $\varepsilon$  is the expansion rate parameter, and  $\sqrt{n+2}$  is the recommended value for inheriting population statistics [19].  $r_k$  is obtained by transforming a uniform random number  $u(0, 1)$  in the interval  $[0, 1]$  as follows:

$$r_k = \begin{cases} u(0, 1)^{\frac{1}{k+1}} & (k = 0, \dots, n-1) \\ 1 & (k = n) \end{cases} \quad (10)$$

- (4) Generate a child individual  $C$  as follows:

$$C = x_n + C_n \quad (11)$$

With these procedures, SPX generates offspring uniformly within the enclosed space of the  $\varepsilon$ -extended polytope  $P'_0, P'_1, \dots, P'_n$  centered at the centroid of the parent individuals  $P_0, P_1, \dots, P_n$ , as shown in Fig. 4. Therefore, SPX is a crossover operator that achieves a balance between exploration and exploitation [33].

## 4 Numerical Examples

**4.1 Problem Setting.** Data-driven MFTD, as mentioned in Sec. 3.1, is a framework for multimodal optimization problems with high nonlinearity and targets problems where the low-fidelity

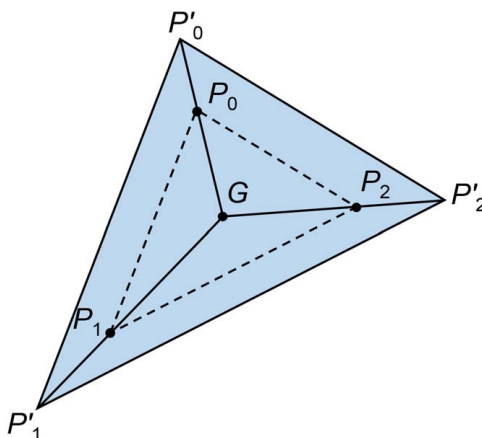


Fig. 4 SPX offspring generation area for 2D

optimization problem is formulated as an easily solvable pseudo-problem for the original one to be solved.

This study applies the proposed method to the design problem of a two-dimensional L-bracket. It is widely used as a benchmark for stress-based topology design [34–37] and is a minimax problem with its high nonlinearity caused by the stress singularity at the reentrant, inner corner. It is formulated as the following multi-objective optimization problem:

$$\begin{aligned} &\text{minimize}_{\gamma} \quad J_1 = \max(\sigma_{vM}), \\ &\quad \quad \quad J_2 = \sum_{e=1}^N v_e \gamma_e \\ &\text{subject to} \quad \gamma_e \in \{0, 1\}, e = 1, 2, \dots, N \end{aligned} \quad (12)$$

Here,  $\sigma_{vM}$  is the von Mises stress, the maximum of which is an objective function, and the volume is the other objective function. Note that the design variables are defined as discrete values, 0 or 1, to deal with the ideal topology optimization problem with high-fidelity evaluation, and  $v_e$  is the elemental volume.

The design domain and boundary conditions for the L-bracket, as shown in Fig. 5, include fixing the upper edge and applying a vertical downward distributed load at the top right corner to avoid stress concentration. The length of the bracket is set to  $L = 2$ , and the design domain is divided into 6400 square elements ( $N = 6400$ ). Young's modulus of the structural material is set to 1, one of the voids is set to  $1 \times 10^{-9}$  instead of 0 to avoid the singular stiffness matrix, and Poisson's ratio is set to 0.3.

It is necessary to formulate the low-fidelity optimization problem as a simple problem that can be easily solvable. In previous studies [17,26], the focus was on the fidelity of physical phenomena, and the governing equations of the flow model were modified. In this study, we also refer to this method and formulate the minimum compliance problem as a low-fidelity optimization problem under the assumption that a promising solution can be obtained even with stiffness maximization [35]:

$$\begin{aligned} &\text{minimize}_{\gamma^{(k)}} \quad \tilde{J}_1 = \mathbf{f}^T \mathbf{u} \\ &\text{subject to} \quad \tilde{J}_2 = \sum_{e=1}^N v_e \gamma_e^{(k)} \leq s^{(k)}, \\ &\quad \quad \quad \gamma_e^{(k)} \in [0, 1], e = 1, 2, \dots, N \\ &\text{for given} \quad s^{(k)} \end{aligned} \quad (13)$$

Here,  $\mathbf{f}$  and  $\mathbf{u}$  are vectors in the equilibrium equation, namely,  $\mathbf{K}\mathbf{u} = \mathbf{f}$ , with the global stiffness matrix  $\mathbf{K}$ . In Eq. (13), the volume is converted from an objective function to a constraint function based on the  $\varepsilon$ -constraint method for the original optimization

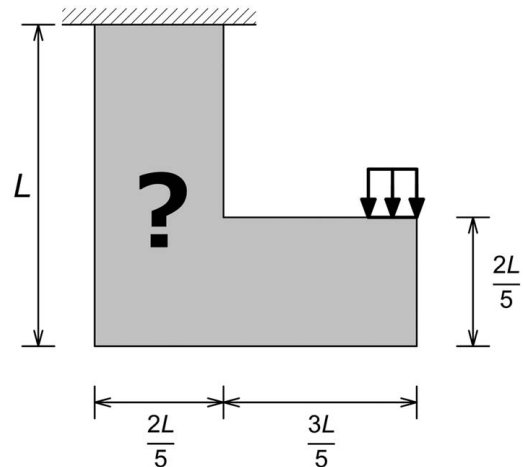


Fig. 5 Design problem of L-bracket

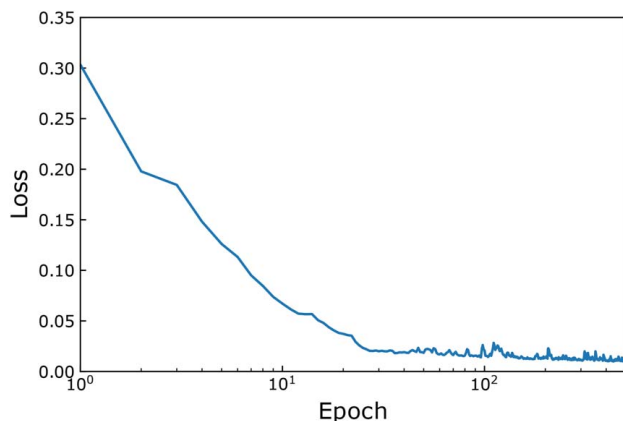
problem of Eq. (12), and since  $\gamma_e^{(k)}$  is relaxed to  $[0, 1]$ , this problem can be easily solved using the density-based method [2]. Note that the density filter [38,39] is applied to ensure the smoothness of  $\gamma$  in  $D$ , and we use the method of moving asymptotes (MMA) [40] as the gradient-based optimizer.

As for the parameters set based on preliminary studies related to the overall procedure, the number of initial data and Pareto solutions from the selection operation are set to 100 and 300, respectively. Regarding the parameters related to the mutation operation,  $N_{\text{mut}}$  is set to 16, and  $G_{\text{mut}}^{\text{max}}$  is set to 0.01. During the latent crossover, nine parent individuals are used by the SPX method because the dimension of the VAE latent space is 8.

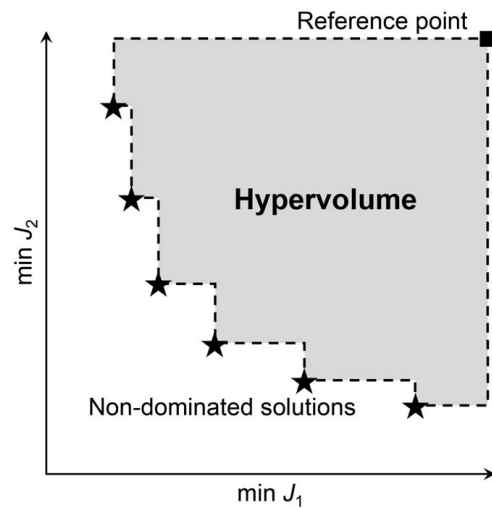
**4.2 Verification of Variational Autoencoder Model.** First, we verify the VAE model and parameters, which play a central role in data-driven MFTD. After preliminary studies on the hyperparameters, we establish the VAE architecture as shown in Fig. 3. The VAE is trained with 100 material distribution samples with 500 epochs, a batch size of 20, and a learning rate of 0.001. The training is terminated if the loss function  $L_{\text{VAE}}$  of Eq. (5) is not improved in every iteration for a total of 50 iterations.

Figure 6 shows the history of the loss function in Eq. (5) during training using the material distribution data at iteration 0 described in Sec. 4.4 as an example. The number of epochs is represented on a logarithmic scale to highlight the areas with significant changes in the loss function. The loss function converges smoothly, indicating that the VAE is appropriately trained under the investigated condition.

**4.3 Verification of Latent Crossover Effect.** For the problem setup in Sec. 4.1, we compare the original and proposed data-driven MFTD frameworks. Since both methods involve random effects, we evaluate and compare them using the hypervolume indicator [41] over ten trials, which is normalized using the initial one. The hypervolume is a measure of the convergence performance of multi-objective optimization. In the case of two objectives, it is represented by the area formed by the reference point and the Pareto front in the objective space as shown in Fig. 7, so a larger hypervolume value means that the Pareto front has progressed. Although mutation is usually performed at regular intervals of iterations, we confirmed that in the case of this design problem, the mutants are selected only once at the beginning, and no mutants are selected as elite solutions thereafter. Therefore, we used the initial data composed of the mutants and initial solutions to compare them with the search performance by crossover without mutation. As this validation involves multiple computations due to the inclusion of randomness, the number of Pareto solutions created through selection has been set to 100 for computational efficiency.



**Fig. 6** Learning history at iteration 0 of our VAE, where the architecture is shown in Fig. 3, the loss function is defined as Eq. (5), and the training data is described in Sec. 4.4



**Fig. 7** Illustration of hypervolume in the case of two-objective optimization

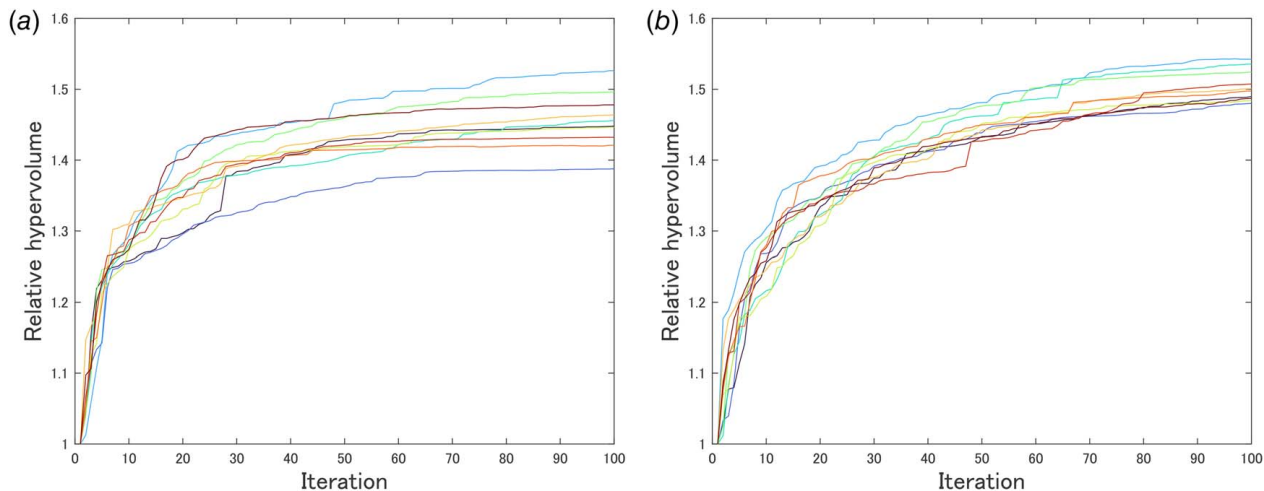
Figure 8 shows the iteration history of the hypervolume indicator over ten trials. Note that its value of each iteration is relative hypervolume normalized by the initial one. In terms of the value at 100 iterations, random sampling in Fig. 8(a) shows a considerable variation in the range from 1.38 to 1.52, while the latent crossover in Fig. 8(b) remains stable in the range from 1.48 to 1.54. The average values of each hypervolume indicator in the ten trials are plotted in Fig. 9. Up to iteration 30, the value of random sampling is higher than that of latent crossover. However, after iteration 30, this relationship is reversed, and at iteration 100, the average value of random sampling is 1.45, while that of latent crossover is 1.50, indicating a difference of 5%. In addition, at iteration 100, the lower limit of the 95% prediction intervals for the latent crossover case exceeds the upper limit for the random sampling case. A t-test was performed on the hypervolume values at iteration 100, and the p-value was 0.00180, which is less than 0.05. Therefore, it can be considered statistically significant that the latent crossover outperforms the random sampling.

In addition, we compare the performance of the best and worst cases among the ten trials shown in Fig. 9 in terms of the relative hypervolume value. Figure 10 presents a comparison of their performance. It is evident from Fig. 10 that the best case with latent crossover achieved the most advanced Pareto front. Even in the worst case with latent crossover, the Pareto front exhibits a spread in the objective function space, whereas in the worst case with random sampling, the Pareto front is highly contracted and fails to maintain diversity. This issue could be serious regarding the nature of EAs [5], as there is an increased risk that the optimized structures are local optima with poor performance.

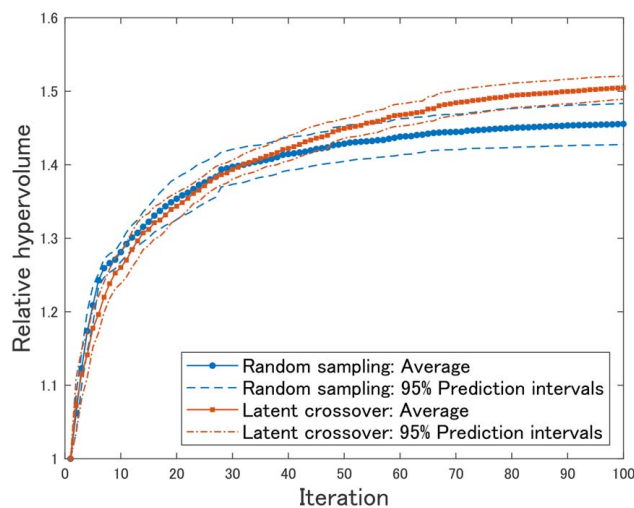
The SPX operator used as the latent crossover operator gradually changes the population distribution while inheriting the statistics, so the increase in hypervolume is slower in the early stages of the search (up to iteration 30) compared to the random sampling. Therefore, this approach maintains diversity and prevents premature convergence, which leads to a more advanced Pareto front in the final iteration (at iteration 100) in Fig. 10. This improvement can be explained based on the theory that the balance between exploration and exploitation [33], i.e., expanding the Pareto front and advancing it, respectively, is significant in EAs. From these results and discussions, it can be concluded that data-driven MFTD achieved stable and high search performance with the latent crossover based on the theory of RCGAs.

**4.4 Validity of Optimized Structure.** Next, we compare the structures obtained through data-driven MFTD with structures





**Fig. 8** Hypervolume for ten trials of random sampling versus latent crossover operations for data-driven MFTD: (a) random sampling and (b) latent crossover

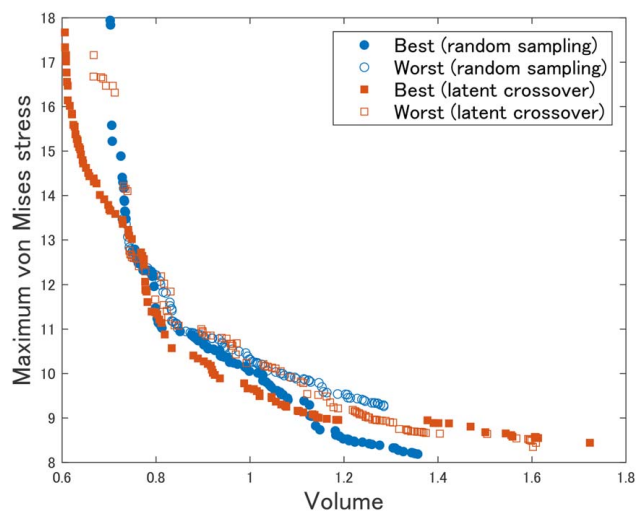


**Fig. 9** Comparison of hypervolume for random sampling versus latent crossover operations for data-driven MFTD

obtained through direct optimization using a gradient-based approach without relying on MFTD principles. Despite only solving the mean compliance minimization problem of Eq. (13) as the low-fidelity optimization problem, we investigate how closely the structures obtained by data-driven MFTD can approach the performance of structures obtained by conventional gradient-based optimization. In addition, we examine the differences between these structures.

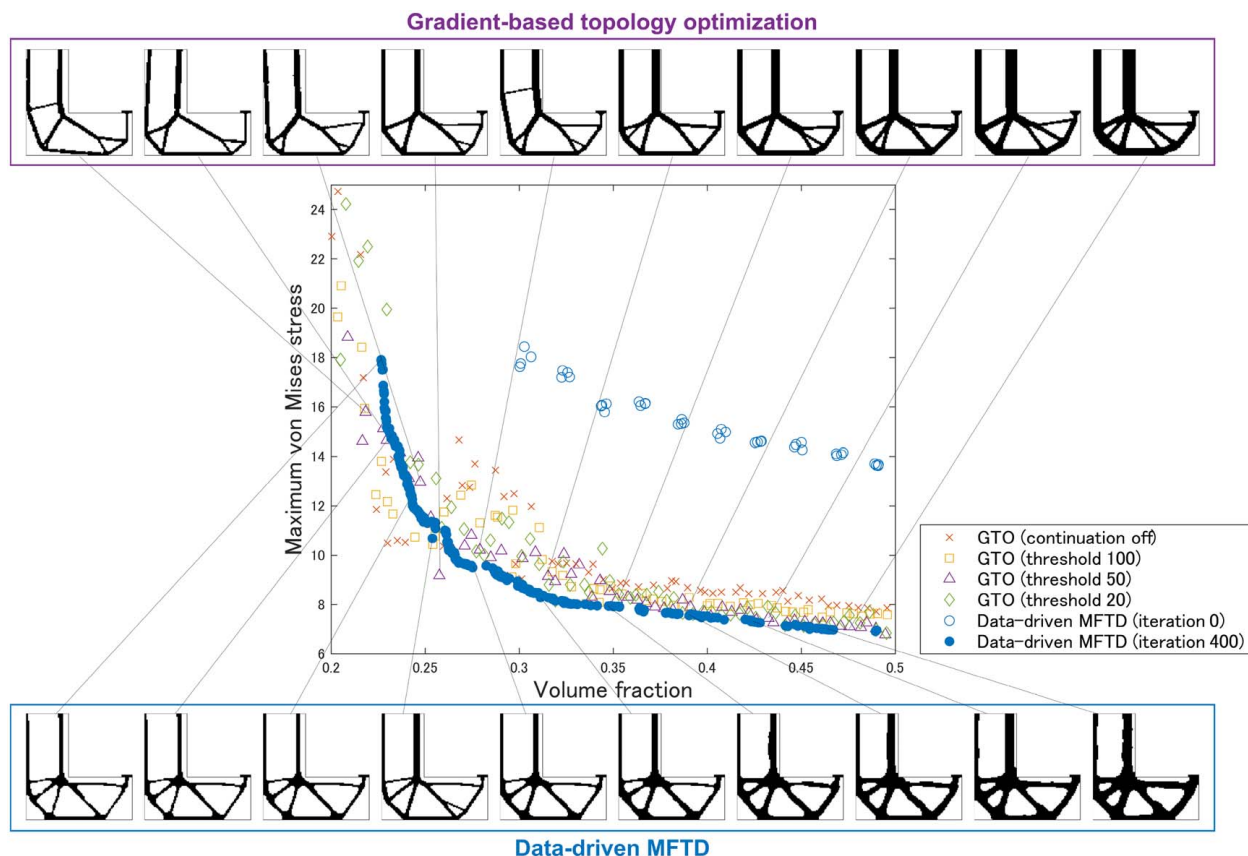
Here, we set various conditions for gradient-based topology optimization. First, given that  $J_1$  represents the maximum value of von Mises stress and  $\gamma_e$  takes a discrete value in  $\{0, 1\}$ , sensitivity analysis becomes impractical for the formulation of the optimization problem in Eq. (12). Therefore, we use the  $P$ -norm measure [42,43], commonly used in stress-based topology optimization [35,36], and relax  $\gamma_e$  to  $[0, 1]$ , as follows:

$$\begin{aligned} &\underset{\gamma}{\text{minimize}} && J = \left( \frac{1}{N} \sum_{e=1}^N (\sigma_{vM})^P \right)^{\frac{1}{P}} \\ &\text{subject to} && G = \sum_{e=1}^N v_e \gamma_e \leq V_{\max} |D|, \\ &&& \gamma_e \in [0, 1], e = 1, 2, \dots, N \end{aligned} \quad (14)$$



**Fig. 10** Objective space represented as volume versus maximum von Mises stress for random sampling versus latent crossover for data-driven MFTD

Here,  $P$  is the stress norm parameter, and  $J$  is called  $P$ -norm stress. For the multi-objective problem formulated in Eq. (12), the volume is set as the constraint function based on the  $\epsilon$ -constraint method. When the stress norm parameter  $P \rightarrow \infty$ , the  $P$ -norm stress approaches the maximum stress value  $\max(\sigma_{vM})$ , but the smoothness is lost. On the other hand, when  $P = 1$ , the smoothness is maintained, but it approaches the average stress value, resulting in an optimized structure closer to the compliance minimum design. Previous studies [35,37] have shown that  $P = 8$  yields the most reasonable designs, and we also use this value in this study. The optimization problem formulated in Eq. (14) is solved using the density-based method [2] with the density filter [38,39], following the commonly employed gradient-based topology optimization approach. A filter radius is set to 0.05, which corresponds to 2.5 element sizes. In addition, in order to binarize  $\gamma_e$  and translate the solution as the original optimization problem of Eq. (12), the heaviside projection [44] is applied to remove the grayscale generated by the density filter. A threshold parameter  $\eta$  is set to 0.5, and the sharpness parameter  $\beta$  is doubled at each constant step, employing a continuation approach. The final result and convergence behavior can be influenced by the continuation threshold. Hence, multiple thresholds of 100, 50, and 20 are used, including the method without the continuation. We use the MMA [40] as the gradient-based optimizer and the move limit is



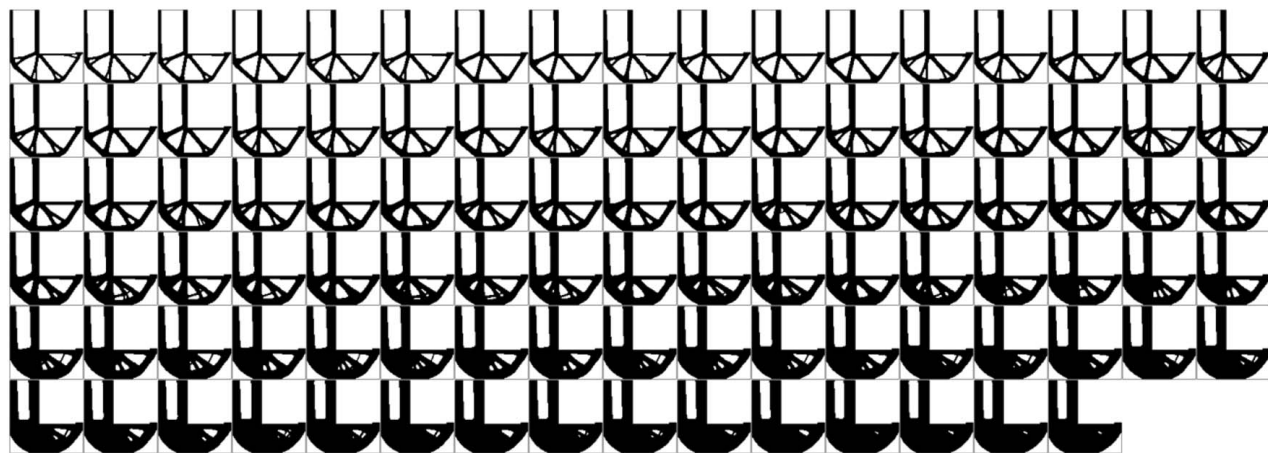
**Fig. 11 Objective space represented as volume versus maximum von Mises stress and corresponding optimized structures for GTO versus data-driven MFTD (for GTO, an excerpt showcasing the optimized designs with a threshold of 50 is provided.)**

set to 0.05. The initial value for  $\gamma_e$  is set to match the volume fraction  $V_{\max}$  used as the constraint function in Eq. (14). For example, for a volume constraint of 30%, the initial value will be set to 0.3. Varying  $V_{\max}$  from 0.2 to 0.5 in increments of 0.005 to generate multiple solutions, binarizing  $\gamma_e$  at 0.5 to create a Pareto solution set for the original problem of Eq. (12), and then we compare it with the solutions obtained using the proposed method. The maximum number of steps is set to 300.

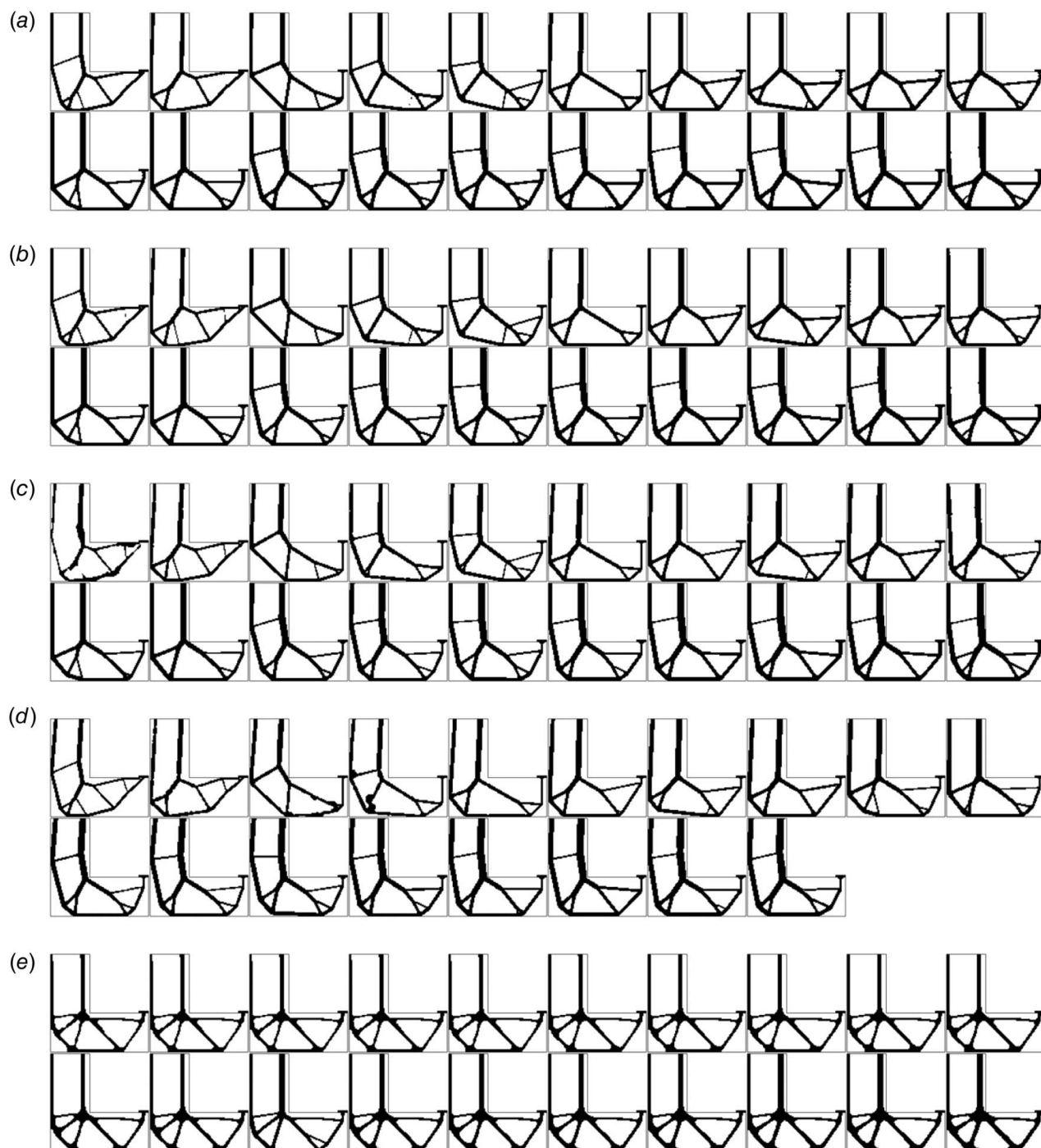
Figure 11 illustrates the structures and performance comparison of results obtained through GTO and data-driven MFTD. First, we discuss the optimization results of data-driven MFTD.

Figure 12 shows the initial dataset obtained by solving the low-fidelity topology optimization problem in Eq. (13). The initial dataset, which consists of compliance minimization designs, has structures that cause stress concentration at their reentrant corners, whereas the optimized structures shown in Fig. 11 have rounded shapes with their reentrant corners smoothed out. The improved performance and reduced volume can be seen by comparing the plots of iteration 0 and iteration 400 in the objective function space shown in Fig. 11.

When comparing the optimization results of GTO and data-driven MFTD in Fig. 11, it can be confirmed that the solutions



**Fig. 12 Initial data generated by solving a mean compliance minimization problem under various volume constraint settings as the low-fidelity topology optimization problem**



**Fig. 13 Optimized structures with volume fractions of 0.2 to 0.3: (a) GTO (continuation off), (b) GTO (threshold 100), (c) GTO (threshold 50), (d) GTO (threshold 20), and (e) data-driven MFTD**

obtained from data-driven MFTD exhibit performance comparable to or better than those from GTO. This is particularly notable in the volume fraction range of 0.3 to 0.5. In the range of lower volume fractions from 0.2 to 0.3, GTO exhibits significant variations in structural performance due to the parameter of continuation thresholds. This suggests that it might be getting trapped in local minima with poor structural performance, likely due to the multimodality caused by the strong nonlinearity of the objective function in Eq. (14). In addition, even with the application of the Heaviside projection, complete removal of the grayscale is not achievable, and especially for low-volume structures, there is a tendency for discontinuities, leading to significant changes in maximum stress values

before and after the binarization of  $\gamma_e$ , as pointed out by Kato et al. [45]. These effects result in the solutions obtained by GTO having a sparse distribution in the objective space. On the other hand, as described in Sec. 3, data-driven MFTD employs an evolutionary algorithm, enabling gradient-free solution updates. This means it is less affected by the multimodality of the objective function. In addition, using Eq. (12) for high-fidelity evaluation of the maximum stress itself with discrete  $\gamma_e$ , rather than using the  $P$ -norm stress with continuous  $\gamma_e$  in Eq. (14), allows the obtained solutions to form an orderly Pareto front. Here, the poor performance of the data-driven MFTD solutions in the range of volume fractions from 0.2 to 0.3 may be attributed to the mutation method.



As described in Sec. 3.1, in data-driven MFTD, we introduce an overlap constraint as a mutation method to solve the LF optimization problem, generating promising structures different from the reference design. The parameter  $G_{mut}^{max}$ , which controls the degree of overlap, uses a constant value independent of the volume. Therefore, while larger structures may be effectively mutated, smaller structures might face challenges in obtaining valid solutions. Due to the reduced effect of the mutation in low-volume regions, it is speculated that the method has led to a kind of local optimum. This suggests that there is room for improvement in the mutation strategy.

Comparing the optimized structures in Fig. 11, the designs obtained by GTO successfully avoid stress concentration at their reentrant corners. However, they consist of straight members and often have triangular or rectangular voids. One of the advantages of data-driven MFTD is that material distributions are represented as vectors and updated using a VAE, eliminating the need for sensitivity analysis. Therefore, as in Eq. (12), the maximum stress can be used directly as the objective function. This feature leads to overall curved structures with rounded appearances at their reentrant corners and elsewhere, as shown in Fig. 11, suggesting that stress concentration is further avoided. In addition, the optimized designs obtained through GTO exhibit various patterns, suggesting entrapment in local minima due to the multimodality of the  $P$ -norm stress in Eq. (14). On the other hand, the optimized designs obtained through data-driven MFTD exhibit nearly identical topologies regardless of volume, differing mainly in member thickness. Compared to GTO, data-driven MFTD achieves global search and appears to reach a promising structural topology. Optimized structures with volume fractions of 0.2–0.3, where these trends are clearly reflected, are shown in Fig. 13. In the case of GTO, it is evident that regardless of continuation thresholds, structures differ significantly even with only a 0.005 difference in volume fraction constraint. This confirms that solutions obtained through GTO are merely local solutions due to multimodality. On the other hand, the optimized structure obtained through data-driven MFTD in Fig. 13(e) maintains a consistent topology regardless of volume. This demonstrates an effective optimization, even for low-volume structures, where conventional GTO struggles, indicating resilience against the influence of multimodality.

As described earlier, it has been demonstrated that the data-driven MFTD framework can address the complex problem of maximum stress minimization by solving the simple problem of mean compliance minimization as a low-fidelity optimization problem. Compared to the solutions by conventional gradient-based optimization, the obtained structures exhibit comparable or better performance and have similar characteristics in terms of avoiding the stress concentration at reentrant corners. This finding suggests that data-driven MFTD may be capable of deriving promising solutions in a gradient-free manner, even in cases of strong multimodal problems where gradient-based optimization is more challenging or potentially infeasible. Note that using multiple initial values in gradient-based topology optimization might yield the optimized structures similar to or better than those obtained with data-driven MFTD. However, it is unclear which initial values should be employed, or whether better solutions exist in the first place. Compared to conventional gradient-based topology optimization, the result indicates that the data-driven MFTD method is likely to yield a unique set of Pareto solutions through an extensive search process.

To generate the data in Fig. 11, we run both data-driven MFTD and GTO codes over a 2.7 GHz AMD Ryzen Threadripper PRO 3995WX 64-Cores CPU. The VAE code for data-driven MFTD was run on a NVIDIA RTX A6000 GPU. The time required to generate the optimized structures in Fig. 11 was 33.7 min for GTO, while data-driven MFTD took 6.8 h. It should be noted that there are potential future improvements to accelerate data-driven MFTD, such as training a VAE every fixed iteration instead of every iteration and utilizing surrogate models for structural performance evaluation.

## 5 Conclusion

This article proposed a *latent crossover* strategy that performs crossover in the latent space of the VAE for the data-driven MFTD framework. Since the latent space is constructed with continuous real numbers, this article employed the SPX as a latent crossover operator based on the theoretical aspects of crossover in RCGAs. The results showed that the proposed method improves the search performance compared to the original method, which performs random sampling in the latent space. As an interesting aspect, this article confirms that the proposed method achieves almost the same performance as that of gradient-based topology optimization using the  $P$ -norm measure for the maximum stress minimization problem, despite only solving the mean compliance minimization problem as the low-fidelity topology optimization problem. Furthermore, it was found that the final results of the proposed method tend to achieve a similar topology, while the optimized results of the gradient-based method exhibit various patterns due to the multimodality caused by the strong nonlinearity of the  $P$ -norm measure. Hence, the data-driven MFTD approach is expected to yield a unique set of Pareto solutions through gradient-free searching.

The concept of latent crossover enables the integration of evolutionary algorithms and machine learning methods. In our future work, we plan to incorporate various types of evolutionary algorithms other than RCGAs, as well as VAE-based advanced machine learning methods into the proposed framework. In addition, to verify the efficacy of the proposed framework on different optimization problems, we consider developing a systematic formulation method for the low-fidelity optimization problem and plan to apply it to other multimodal problems involving strongly nonlinear physical phenomena.

## Acknowledgment

The second author was supported by JSPS KAKENHI (Grant Nos. 20KK0329, 20H02054, and 23H03799).

## Conflict of Interest

There are no conflicts of interest.

## Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

## References

- [1] Bendsoe, M. P., and Kikuchi, N., 1988, "Generating Optimal Topologies in Structural Design Using a Homogenization Method," *Comput. Methods Appl. Mech. Eng.*, **71**(2), pp. 197–224.
- [2] Bendsoe, M. P., and Sigmund, O., 2003, *Topology Optimization: Theory, Methods, and Applications*, Springer Science & Business Media, Berlin.
- [3] Sigmund, O., and Petersson, J., 1998, "Numerical Instabilities in Topology Optimization: A Survey on Procedures Dealing With Checkboards, Mesh-Dependencies and Local Minima," *Struct. Optim.*, **16**(1), pp. 68–75.
- [4] Mitchell, M., and Taylor, C. E., 1999, "Evolutionary Computation: An Overview," *Annu. Rev. Ecol. Syst.*, **30**(1), pp. 593–616.
- [5] Goldberg, D. E., 1989, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley, New York.
- [6] Wang, S. Y., and Tai, K., 2005, "Structural Topology Design Optimization Using Genetic Algorithms With a Bit-Array Representation," *Comput. Methods Appl. Mech. Eng.*, **194**(36–38), pp. 3749–3770.
- [7] Madeira, J. F. A., Pina, H. L., and Rodrigues, H. C., 2010, "GA Topology Optimization Using Random Keys for Tree Encoding of Structures," *Struct. Multidiscipl. Optim.*, **40**(1), pp. 227–240.
- [8] Zhou, H., 2010, "Topology Optimization of Compliant Mechanisms Using Hybrid Discretization Model," *J. Mech. Des.*, **132**(11), p. 111003.
- [9] Balamurugan, R., Ramakrishnan, C. V., and Swaminathan, N., 2011, "A Two Phase Approach Based on Skeleton Convergence and Geometric Variables for



- Topology Optimization Using Genetic Algorithm,” *Struct. Multidiscipl. Optim.*, **43**(3), pp. 381–404.
- [10] Sigmund, O., 2011, “On the Usefulness of Non-Gradient Approaches in Topology Optimization,” *Struct. Multidiscipl. Optim.*, **43**(5), pp. 589–596.
- [11] Kingma, D. P., and Welling, M., 2013, “Auto-Encoding Variational Bayes,” *arXiv preprint*.
- [12] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014, “Generative Adversarial Networks,” *arXiv preprint*.
- [13] Regenwetter, L., Nobari, A. H., and Ahmed, F., 2022, “Deep Generative Models in Engineering Design: A Review,” *J. Mech. Des.*, **144**(7), p. 071704.
- [14] Guo, T., Lohan, D. J., Cang, R., Ren, M. Y., and Allison, J. T., 2018, “An Indirect Design Representation for Topology Optimization Using Variational Autoencoder and Style Transfer,” 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Kissimmee, FL, Jan. 8–12, p. 0804.
- [15] Oh, S., Jung, Y., Kim, S., Lee, I., and Kang, N., 2019, “Deep Generative Design: Integration of Topology Optimization and Generative Models,” *ASME J. Mech. Des.*, **141**(11), p. 111405.
- [16] Kazemi, H., Seepersad, C. C., and Kim, H. A., 2022, “Multiphysics Design Optimization via Generative Adversarial Networks,” *ASME J. Mech. Des.*, **144**(12), p. 121702.
- [17] Yaji, K., Yamasaki, S., and Fujita, K., 2022, “Data-Driven Multifidelity Topology Design Using a Deep Generative Model: Application to Forced Convection Heat Transfer Problems,” *Comput. Methods Appl. Mech. Eng.*, **388**, p. 114284.
- [18] Yamasaki, S., Yaji, K., and Fujita, K., 2021, “Data-Driven Topology Design Using a Deep Generative Model,” *Struct. Multidiscipl. Optim.*, **64**(3), pp. 1401–1420.
- [19] Tsutsui, S., Yamamura, M., and Higuchi, T., 1999, “Multi-Parent Recombination With Simplex Crossover in Real Coded Genetic Algorithms,” Genetic and Evolutionary Computation Conference, Orlando, FL, July 13–17, Vol. 1, pp. 657–664.
- [20] Herrera, F., Lozano, M., and Verdegay, J. L., 1998, “Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis,” *Artif. Intell. Rev.*, **12**(4), pp. 265–319.
- [21] Kita, H., and Yamamura, M., 1999, “A Functional Specialization Hypothesis for Designing Genetic Algorithms,” 1999 IEEE Conference on Systems, Man, and Cybernetics, Tokyo, Japan, Oct. 12–15, Vol. 3, pp. 579–584.
- [22] Kita, H., and Yamamura, M., 1999, “Design Guidelines for Genetic Algorithms Based on Function Specialization Hypothesis,” *J. SCIE*, **38**(10), pp. 612–617.
- [23] Kita, H., Ono, I., and Kobayashi, S., 2000, “Multi-Parental Extension of the Unimodal Normal Distribution Crossover for Real-Coded Genetic Algorithms,” *Trans. SCIE*, **36**(10), pp. 875–883.
- [24] Beyer, H. G., and Deb, K., 2001, “On Self-Adaptive Features in Real-Parameter Evolutionary Algorithms,” *IEEE Trans. Evol. Comput.*, **5**(3), pp. 250–270.
- [25] Herrera, F., Lozano, M., and Sánchez, A. M., 2003, “A Taxonomy for the Crossover Operator for Real-Coded Genetic Algorithms: An Experimental Study,” *Int. J. Intell. Syst.*, **18**(3), pp. 309–338.
- [26] Yaji, K., Yamasaki, S., and Fujita, K., 2020, “Multifidelity Design Guided by Topology Optimization,” *Struct. Multidiscipl. Optim.*, **61**(3), pp. 1071–1085.
- [27] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T., 2002, “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, **6**(2), pp. 182–197.
- [28] Michalewicz, Z., 1996, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer Berlin, Heidelberg.
- [29] Dilokthanakul, N., Mediano, P. A. M., Garnelo, M., Lee, M. C. H., Salimbeni, H., Arulkumaran, K., and Shanahan, M., 2016, “Deep Unsupervised Clustering with Gaussian Mixture Variational Autoencoders,” *arXiv preprint*.
- [30] Tsumoto, R., Fujita, K., Nomaguchi, Y., Yamasaki, S., and Yaji, K., 2022, “Classification-Directed Conceptual Structure Design Based on Topology Optimization, Deep Clustering, and Logistic Regression,” ASME 2022 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, St. Louis, MO, Aug. 14–17, p. V03AT03A012.
- [31] Eshelman, L. J., Mathias, K. E., and Schaffer, J. D., 1997, “Crossover Operator Biases: Exploiting the Population Distribution,” 1997 International Conference on Genetic Algorithms, East Lansing, MI, July 19–23, pp. 354–361.
- [32] Ghosh, A., and Tsutsui, S., 2003, “A Real-Coded Genetic Algorithm Using the Unimodal Normal Distribution Crossover,” *Advances in Evolutionary Computing: Theory and Applications*, Springer Berlin, Heidelberg, pp. 213–237.
- [33] Črepinšek, M., Liu, S. H., and Mernik, M., 2013, “Exploration and Exploitation in Evolutionary Algorithms: A Survey,” *ACM Comput. Surv.*, **45**(3), pp. 1–33.
- [34] Duysinx, P., and Bendsøe, M. P., 1998, “Topology Optimization of Continuum Structures With Local Stress Constraints,” *Int. J. Numer. Methods Eng.*, **43**(8), pp. 1453–1478.
- [35] Le, C., Norato, J. A., Bruns, T., Ha, C., and Tortorelli, D., 2010, “Stress-Based Topology Optimization for Continua,” *Struct. Multidiscipl. Optim.*, **41**(4), pp. 605–620.
- [36] Holmberg, E., Torstenfelt, B., and Klarbring, A., 2013, “Stress Constrained Topology Optimization,” *Struct. Multidiscipl. Optim.*, **48**(1), pp. 33–47.
- [37] Norato, J. A., Smith, H. A., Deaton, J. D., and Kolonay, R. M., 2022, “A Maximum-Rectifier-Function Approach to Stress-Constrained Topology Optimization,” *Struct. Multidiscipl. Optim.*, **65**(10), p. 286.
- [38] Bruns, T. E., and Tortorelli, D. A., 2001, “Topology Optimization of Non-Linear Elastic Structures and Compliant Mechanisms,” *Comput. Methods Appl. Mech. Eng.*, **190**(26), pp. 3443–3459.
- [39] Bourdin, B., 2001, “Filters in Topology Optimization,” *Int. J. Numer. Methods Eng.*, **50**(9), pp. 2143–2158.
- [40] Svanberg, K., 1987, “The Method of Moving Asymptote—New Method for Structural Optimization,” *Int. J. Numer. Methods Eng.*, **24**(2), pp. 359–373.
- [41] Shang, K., Ishibuchi, H., He, L., and Pang, L. M., 2021, “A Survey on the Hypervolume Indicator in Evolutionary Multiobjective Optimization,” *IEEE Trans. Evol. Comput.*, **25**(1), pp. 1–20.
- [42] Yang, R. J., and Chen, C. J., 1996, “Stress-Based Topology Optimization,” *Struct. Optim.*, **12**(2), pp. 98–105.
- [43] Duysinx, P., and Sigmund, O., 1998, “New Developments in Handling Stress Constraints in Optimal Material Distribution,” AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, MO, Sept. 2–4.
- [44] Wang, F., Lazarov, B. S., and Sigmund, O., 2011, “On Projection Methods, Convergence and Robust Formulations in Topology Optimization,” *Struct. Multidiscipl. Optim.*, **43**(6), pp. 767–784.
- [45] Kato, M., Kii, T., Yaji, K., and Fujita, K., 2023, “Tackling an Exact Maximum Stress Minimization Problem With Gradient-Free Topology Optimization Incorporating a Deep Generative Model,” ASME 2023 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Boston, MA, Aug. 20–23, p. V03BT03A008.