**Proceedings of the ASME 2023
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE 2023
August 20–23, 2023, Boston, Massachusetts**

IDETC2023-111079

DATA-DRIVEN MULTIFIDELITY TOPOLOGY DESIGN WITH A LATENT CROSSOVER OPERATION

Taisei Kii[1], Kentaro Yaji[1,*], Kikuo Fujita[1], Zhenghui Sha[2], Carolyn C. Seepersad[2]

[1]Department of Mechanical Engineering, Osaka University, Osaka 565-0871, Japan
[2]Walker Department of Mechanical Engineering, University of Texas Austin, Austin, TX 78712

ABSTRACT

Topology optimization is one of the most flexible structural optimization methodologies. However, in exchange for its high degree 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 bad local optima. Its core is a data-driven multifidelity topology design (MFTD) method, in which design candidates generated by solving low-fidelity topology optimization problems are updated based on evolutionary algorithms (EAs) through high-fidelity evaluation. The key component of the data-driven MFTD is a deep generative model that compresses the dimension of the original data into a low-dimensional manifold, i.e., the latent space. In the original framework, convergence variability and premature convergence problems arise as the generative process is performed randomly in the latent space. Inspired by a popular crossover operation, we propose a data-driven MFTD framework incorporating 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 method. Furthermore, the optimized designs exhibit performance comparable to or better than that in conventional gradient-based topology optimization using the P-norm measure.

Keywords: Topology optimization; Deep generative model; Maximum stress minimization; Latent crossover

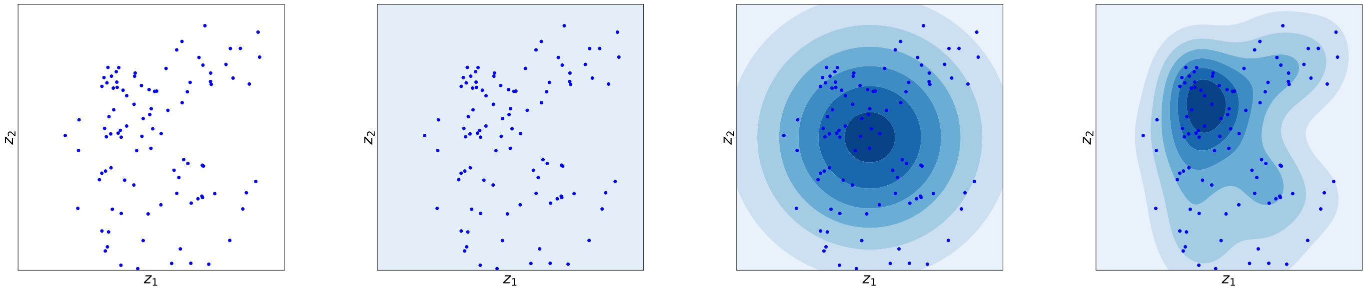
1. INTRODUCTION

Topology optimization, first proposed by Bendsøe and Kikuchi [1], enables the determination of an optimized material distribution for a structural optimization problem and offers a high degree 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. 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 a strongly nonlinear problem, 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 is evolutionary algorithms (EAs) since they are gradient-free [3]. 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 [4]. 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 [5–8] 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 [9] 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 dimen-

*Corresponding author: yaji@mech.eng.osaka-u.ac.jp



(a) Parent individuals **(b) Uniform random sampling** **(c) Normal random sampling** **(d) Latent crossover**

FIGURE 1: THE PROBABILITY DISTRIBUTION FOR GENERATING OFFSPRING IN 2D LATENT SPACE

sionality, some deep generative models can dramatically reduce the dimensionality of the topology optimization problem. In deep generative models such as variational autoencoders (VAEs) [10] and generative adversarial networks (GANs) [11], an encoder is built to compress high-dimensional data into a low-dimensional manifold, called latent space. In addition, a decoder reconstructs high-dimensional data from the latent space. As a review paper [12] mentioned, relevant studies on deep generative models for engineering design problems have increased dramatically in recent years. As pioneering work, Guo et al. [13] 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. [14] proposed a design framework that iteratively trains a GAN to generate a variety of designs. Kazemi et al. [15] proposed a method to generate conceptual designs using a GAN for multi-physics topology optimization problems.

Based on combining EAs and deep generative models, Yaji et al. [16] proposed a data-driven multifidelity topology design (MFTD) method that enables gradient-free topology optimization under a high degree of design freedom. 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 [17], 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 paper proposes a data-driven MFTD framework incorporating a particular crossover operation based on EAs, called *latent crossover*. Specifically, simplex crossover (SPX) [18]—a crossover operator of real-coded genetic algorithms (RCGAs) [19]—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 using the P-norm measure for the maximum stress minimization problem.

2. LATENT CROSSOVER

In data-driven MFTD [16], whose details are descrived in Section 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 RCGAs 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 [20] 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 following design guideline [21–23] for RCGA crossover operators was proposed focusing on the perspective of statistics to concretize the above theory. That is, 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. Fig. 1 shows an example of the probability distribution for generating offspring in a two-dimensional latent space. The darker areas have a higher probability of generating offspring. Assuming that the distribution of the parent population, as shown in Fig. 1a, is given, data-driven MFTD performs sampling by uniform random numbers 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. 1b. 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 data-driven MFTD, it remains only as an operation similar to crossover and cannot be considered strictly performing crossover because of random sampling. Since the input data follows a normal distribution in the latent space due to the nature of VAEs, generating offspring through sampling based on a normal distribution rather than a uniform distribution is also possible. However, as shown in Fig. 1c, 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. Based on EA's 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 above, it is impossible to strictly inherit the statistical characteristics of the parent population through random 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 [24]. Thus, applying latent crossover to the parent population in Fig. 1a, the probability distribution of generated offspring is expected to become the one shown in Fig. 1d. 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 multi-objective topology optimization problem:

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

Here, J_i ($i = 1,2,\ldots,r_o$) and G_j ($j = 1,2,\ldots,r_c$) are the objective and constraint functions, respectively. The optimization problem defined by Eq. (1) is a 0-1 optimization problem with $\boldsymbol{\gamma}$ 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 multifidelity topology design (MFTD) [25] and divide the problem of Eq. (1) into two procedures: low-fidelity optimization and high-fidelity evaluation, to solve the problem.

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 flowchart of the proposed data-driven MFTD with latent crossover is shown in Fig. 2, and the details of each step are explained below.

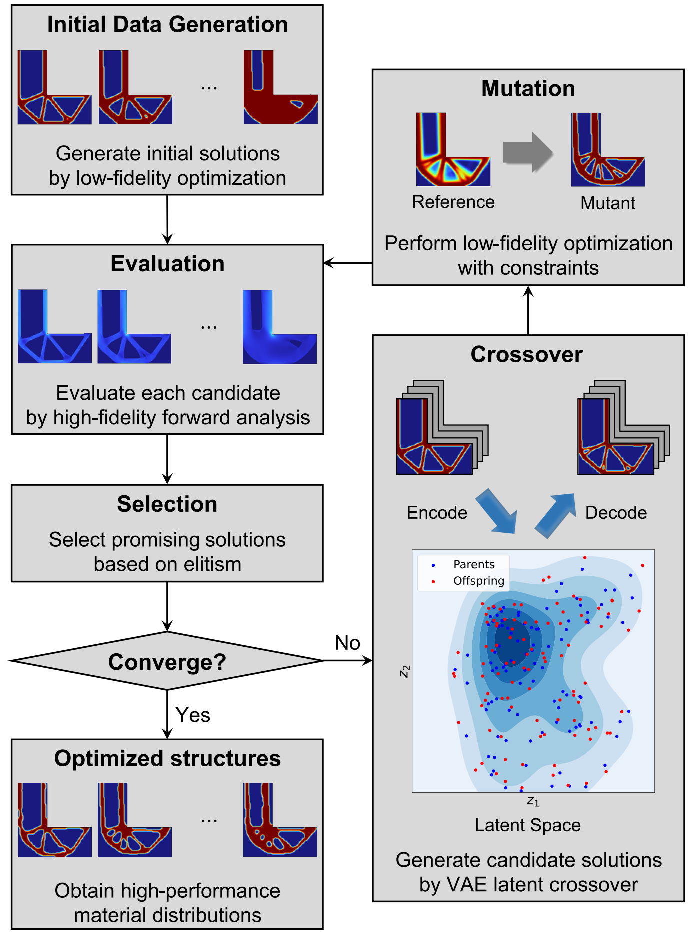


FIGURE 2: SCHEMATIC FLOWCHART OF DATA-DRIVEN MFTD WITH LATENT CROSSOVER

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:

$$\begin{aligned} \underset{\boldsymbol{\gamma}^{(k)}}{\text{minimize}} \quad & \widetilde{J}_i(\boldsymbol{\gamma}^{(k)}) \\ \text{subject to} \quad & \widetilde{G}_j(\boldsymbol{\gamma}^{(k)}, \mathbf{s}^{(k)}) \leq 0, \\ & \gamma_e^{(k)} \in [0,1],\ e = 1,2,\ldots,N \\ \text{for given} \quad & \mathbf{s}^{(k)},\ k = 1,2,\ldots,K. \end{aligned} \tag{2}$$

Here, $\widetilde{J}_i$ and $\widetilde{G}_j$ are the objective and constraint functions for the low-fidelity optimization problem, respectively, which can be easily computed by pseudo-functions. Additionally, $\mathbf{s} = [s_1, s_2, \ldots, s_{N_{sd}}]$ 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 a maximum limit of a constraint and optimization parameters such as a 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).

Selection As mentioned in Section 2, the 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 those to be preserved in the next generation. This paper uses the nondominated sorting genetic algorithm II (NSGA-II) [26] 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, which are not dominated by any other solutions, are selected from the population, 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. The 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 Sections 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 [27], is applied, its outcome is limited to a specific subspace against the whole solution space. This limitation is 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 Section 1.

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

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

where $m = 1, 2, \ldots, N_{\text{mut}}$ is the number of mutants, v_e is the elemental volume, $\widetilde{G}_{\text{mut}}^{\max}$ is a parameter that controls the degree of overlap between the reference material distribution $\boldsymbol{\gamma}^{\text{ref}(m)}$ and the design variable $\boldsymbol{\gamma}^{(m)}$, 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 paper 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.

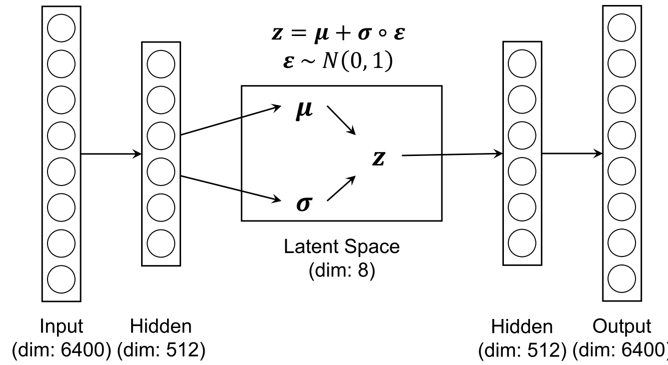


FIGURE 3: ARCHITECTURE OF VAE

3.2 Variational Autoencoder

Fig. 3 shows the architecture of the VAE used in the numerical examples in Section 4. 6400 input/output elements are combined into two 8-dimensional layers, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, through a hidden layer of 512 dimensions. $\boldsymbol{\mu}$ is the mean vector, and $\boldsymbol{\sigma}$ is the variance vector of the latent variables $\mathbf{z}$. The following equation defines the latent variable vector $\mathbf{z}$:

$$\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \circ \boldsymbol{\varepsilon}, \quad (4)$$

where $\circ$ is the operator that calculates the element-wise product, and $\boldsymbol{\varepsilon}$ is a vector of random numbers from the standard normal 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_{VAE} is used for the training:

$$L_{\text{VAE}} \coloneqq L_{\text{recon}} + \varrho L_{\text{KL}}, \quad (5)$$

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

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

Compared to simple dimensionality reduction using autoencoders (AEs), VAEs are trained by incorporating probabilistic variation through $\boldsymbol{\varepsilon}$, allowing for 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

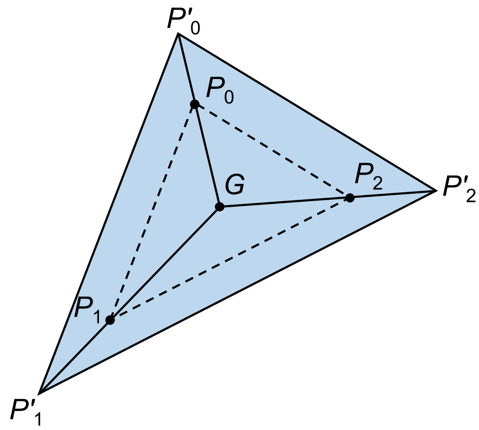


FIGURE 4: SPX OFFSPRING GENERATION AREA FOR 2D

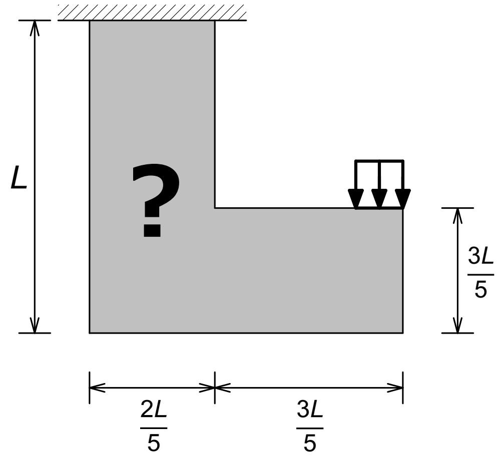


FIGURE 5: DESIGN PROBLEM OF L-BRACKET

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 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 Section 2, the generation of probability distribution shown in Fig. 1 can be problematic. In this paper, we perform latent crossover using the crossover operator explained in Section 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 [18, 28, 29] have been proposed to address this issue. This paper uses the simplex crossover (SPX) [18] for a latent crossover operator. SPX is one of the multi-parent crossover operators for RCGAs that generates offspring using three or more parent individuals and is consistent with the crossover design guidelines [21–23] 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 $\boldsymbol{P}_0, \boldsymbol{P}_1, \ldots, \boldsymbol{P}_n$ from the population.

(2) Calculate the centroid $\boldsymbol{G}$ of the parent individuals as follows:

$$\boldsymbol{G} = \sum_{i=0}^{n} \boldsymbol{P}_i. \tag{7}$$

(3) Calculate $\boldsymbol{x}_k$ and $\boldsymbol{C}_k$ for $k = 0, 1, \ldots, n$ as follows:

$$\boldsymbol{x}_k = \boldsymbol{G} + \varepsilon(\boldsymbol{P}_k - \boldsymbol{G}), \tag{8}$$

$$\boldsymbol{C}_k = \begin{cases} \boldsymbol{0} & (k = 0) \\ r_{k-1}(\boldsymbol{x}_{k-1} - \boldsymbol{x}_k + \boldsymbol{C}_{k-1}) & (k = 1, \ldots, n). \end{cases} \tag{9}$$

Here, ε is a parameter called the expansion rate, and $\sqrt{n+2}$ is the recommended value for inheriting population statistics [18]. r_k is obtained by transforming a uniform random number $u(0,1)$ in the interval $[0,1]$ as follows:

$$r_k = \begin{cases} 0 & (k < 0) \\ u(0,1)^{\frac{1}{k+1}} & (k = 1, \ldots, n-1) \\ 1 & (k \geq 1). \end{cases} \tag{10}$$

(4) Generate a child individual $\boldsymbol{C}$ as follows:

$$\boldsymbol{C} = \boldsymbol{x}_n + \boldsymbol{C}_n. \tag{11}$$

With these procedures, SPX generates offspring uniformly within the enclosed space of the ε-extended polytope $P'_0, P'_1, \ldots, P'_n$ centered at the centroid of the parent individuals $P_0, P_1, \ldots, P_n$, as shown in Fig. 4. Therefore, SPX is a crossover operator that achieves a balance between exploration and exploitation [30].

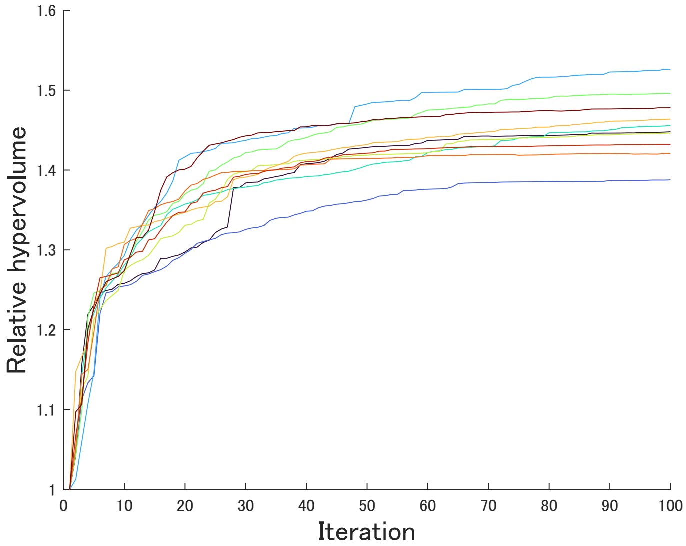
4. NUMERICAL EXAMPLES

4.1 Problem Setting

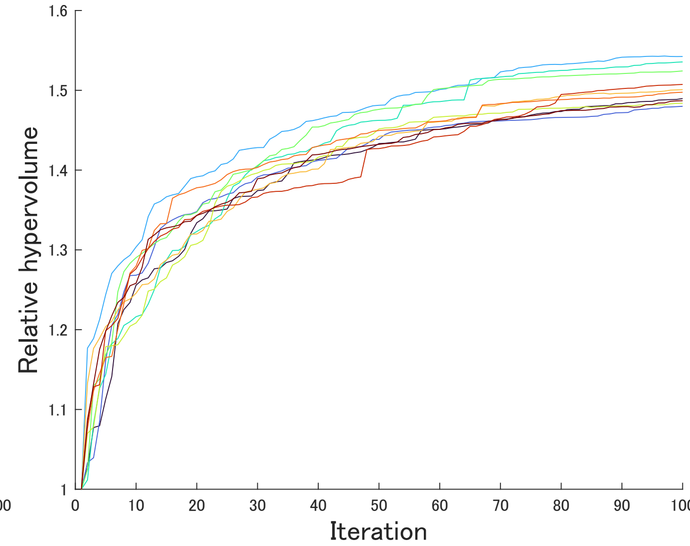
This study applies the proposed method to the design problem of a two-dimensional L-bracket. It is widely used as a benchmark for stress-based topology design [31–34] and is a minimax problem with its high nonlinearities caused by the stress singularity, called re-entrant corner, at the inner corner. It can be formulated as the following multi-objective optimization problem:

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

Here, σ_{vM} is the von Mises stress, the maximum of which is an objective function, and the volume is the other objective function.



(a) Random sampling



(b) Latent crossover

FIGURE 6: HYPERVOLUME FOR TEN TRIALS

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.

The design domain and boundary conditions for the L-bracket, as shown in Fig. 5, include fixing the upper end and applying a vertical downward distributed load at the top 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 1×10^{-9} instead of 0 to avoid the singular stiffness matrix, and Poisson's ratio is set to 0.3.

Under the assumption that a promising solution can be obtained even with stiffness maximization [32], we formulate the minimum compliance problem and use it as a low-fidelity optimization problem:

$$\begin{aligned} \underset{\gamma^{(k)}}{\text{minimize}} \quad & \widetilde{J}_1 = \mathbf{f}^{\mathsf{T}}\mathbf{u} \\ \text{subject to} \quad & \widetilde{J}_2 = \sum_{e=1}^{N} v_e \gamma_e^{(k)} \le s^{(k)}, \\ & \gamma_e^{(k)} \in [0,1],\ e = 1, 2, \ldots, N \\ \text{for given} \quad & s^{(k)}. \end{aligned} \tag{13}$$

Here, $\mathbf{f}$ and $\mathbf{u}$ are vectors in the equilibrium equation, namely, $\mathbf{Ku} = \mathbf{f}$, 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 ε-constraint method for the original optimization 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 a design variable filter [35] is applied to ensure the smoothness of $\boldsymbol{\gamma}$ in D.

As for the parameters related to the overall procedure, the number of initial data and Pareto solutions from the selection operation are both set to 100. Regarding the parameters related to the mutation operation, N_{mut} is set to 16 and $\widetilde{G}_{\text{mut}}$ is set to 0.01.

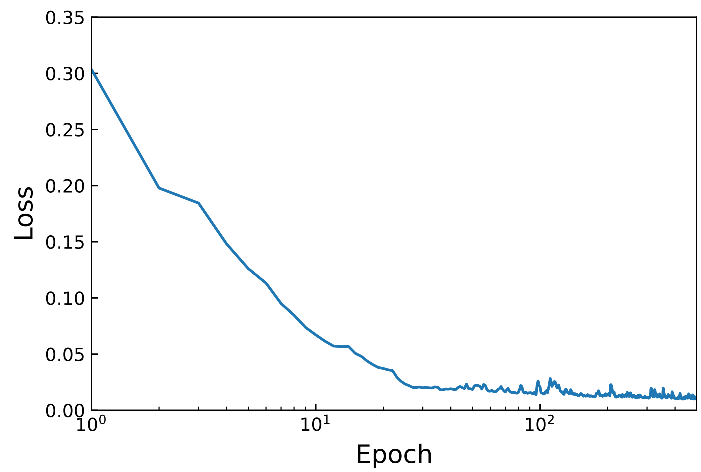


FIGURE 7: LEARNING HISTORY AT ITERATION 0 FOR THE ARCHITECTURE IN FIGURE 3, THE LOSS FUNCTION IN EQUATION 5, AND THE TRAINING DATA DISCRIVED IN SECTION 4.4

During the latent crossover, 9 parent individuals are used by the SPX method because the dimension of the VAE latent space is 8.

4.2 Verification of VAE Model

First, we verify the VAE model and parameters, which plays a central role in the 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_{VAE} of Eq. (5) is not improved in every iteration for a total of 50 iterations.

Fig. 7 shows the history of the loss function in Eq. (5) during training using the material distribution data at iteration 0 descrived in Section 4.4 as an example. The loss function converges smoothly, indicating that the VAE is appropriately trained

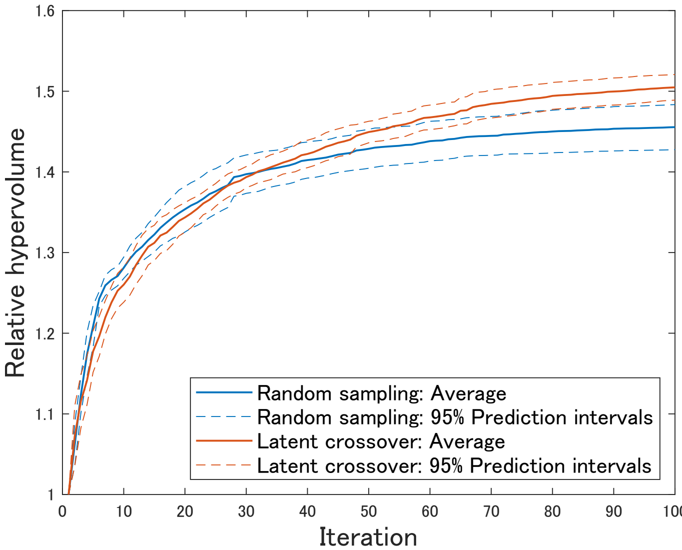


FIGURE 8: COMPARISON OF HYPERVOLUME

under the investigated condition.

4.3 Verification of Latent Crossover Effect

For the problem set up in Section 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 [36] over ten trials, which is normalized using the initial one. The hypervolume is a measure for 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. 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.

Fig. 6 shows the iteration history of the hypervolume indicator over ten trials. In terms of the value at 100 iterations, random sampling in Fig. 6a shows a considerable variation in the range from 1.38 to 1.52, while the latent crossover in Fig. 6b 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. 8. 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.

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. However, this approach maintains diversity and prevents premature convergence, which leads to a more advanced Pareto front in the final iteration (at iteration 100). This improvement can be explained based on the theory that the balance between exploration and exploitation [30], 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.

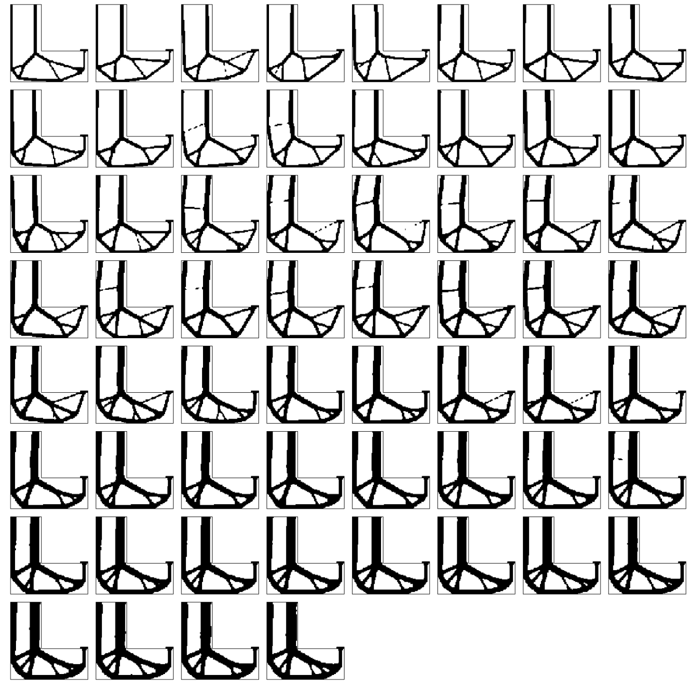


FIGURE 9: OPTIMIZED STRUCTURES BY GRADIENT-BASED TOPOLOGY OPTIMIZATION USING THE *P*-NORM MEASURE

4.4 Validity of Optimized Structure

Next, we compare structures obtained through data-driven MFTD with structures obtained through direct optimization using a gradient-based approach without relying on MFTD principles. However, deriving sensitivity analysis for the optimization problem defined by Eq. (12) is impossible because J_1 is the maximum value of the von Mises stress and γ_e is a discrete value $\{0, 1\}$. Therefore, we use the P-norm measure [37, 38], commonly used in stress-based topology optimization [32, 33], and relax γ_e to $[0, 1]$, as follows:

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

Here, P is the stress norm parameter, and J is called P-norm stress. For the multi-objective problem formulated in Eq. (12), the

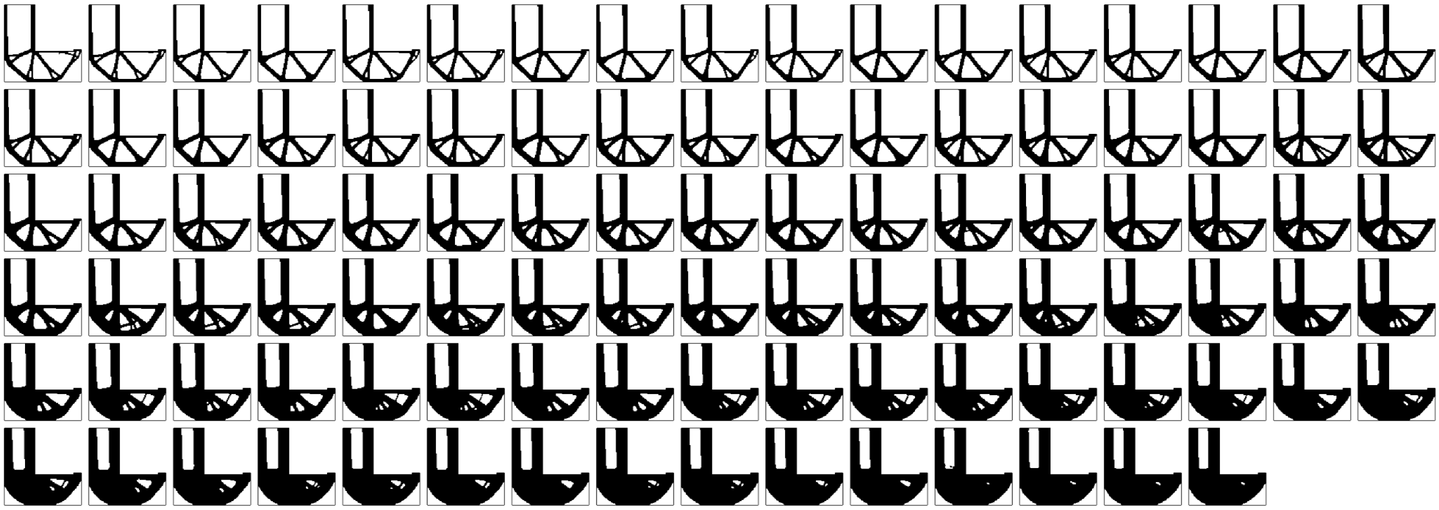


FIGURE 10: INITIAL DATA GENERATED BY SOLVING A MEAN COMPLIANCE MINIMIZATION PROBLEM UNDER VARIOUS VOLUME CONSTRAINT SETTINGS AS THE LOW-FIDELITY TOPOLOGY OPTIMIZATION PROBLEM

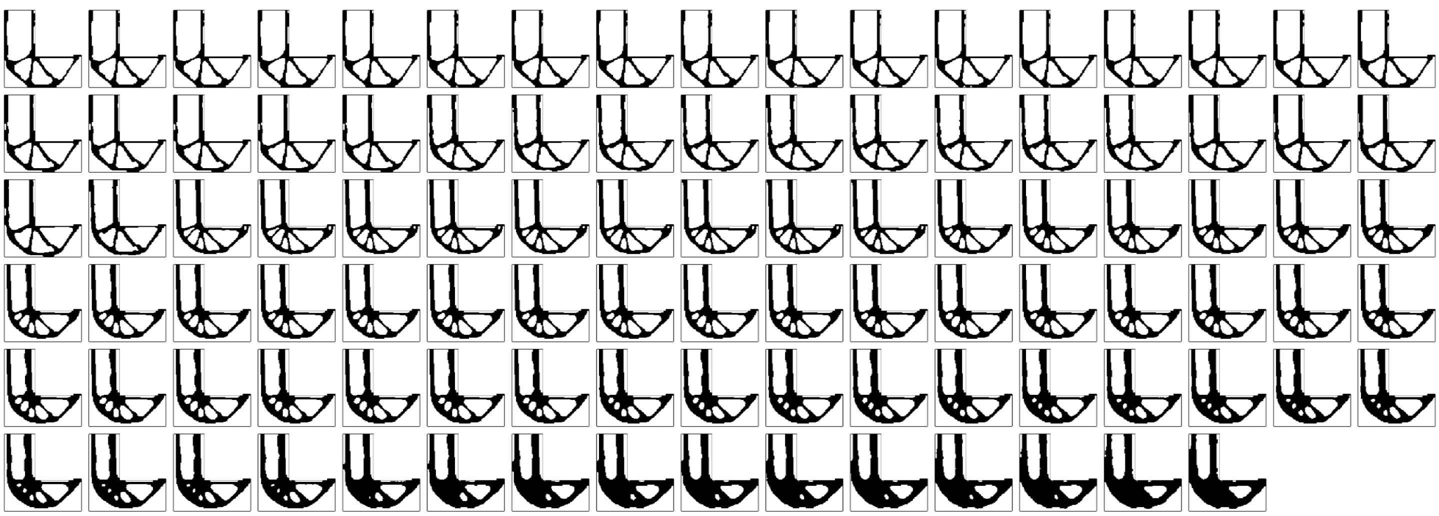


FIGURE 11: OPTIMIZED STRUCTURES BY DATA-DRIVEN MFTD

volume is set as the constraint function based on the ε-constraint method. When the stress norm parameter $P \to \infty$, the P-norm stress approaches the maximum stress value $\max(\sigma_{\mathrm{vM}})$, 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 [32, 34] have shown that $P = 8$ yields the most reasonable design, but in order to set an objective function closer to the maximum stress value, this paper uses a method that iteratively increases P from 8 to 16 and 32 using the continuation method [39]. This operation enables us to use a more rigorous approximation function while stably solving the optimization problem. We set the P-norm stress based on the result of the continuous approach as the objective function and use the method of moving asymptotes (MMA) [40] as the optimization method. Fig. 9 shows 60 directly solved designs while changing $V_{\max}$ from 20% to 50% in 0.5% increments.

Fig. 10 shows the initial dataset obtained by solving the low-fidelity optimization problem in Eq. (13), and Fig. 11 shows the optimized structures obtained by data-driven MFTD with 300 iterations. The initial dataset, which consists of compliance minimization designs, has structures that cause stress concentration at their re-entrant corners, whereas the structures obtained by data-driven MFTD have rounded shapes with their re-entrant corners smoothed out. The improved performance and reduced volume can be seen by comparing the plots of iteration 0 and iteration 300 in the objective function space shown in Fig. 12.

When comparing the performance of Pareto solutions obtained by data-driven MFTD and the solutions obtained by direct optimization in the objective space shown in Fig. 12, it can be confirmed that equivalent performance solutions are obtained in general. In particular, data-driven MFTD outperforms direct optimization for solutions in the volume range of 0.7 to 1.0. In addition, while the solutions for direct optimization are sparsely distributed in the objective space due to the instability of the objective function, the solutions for data-driven MFTD form an

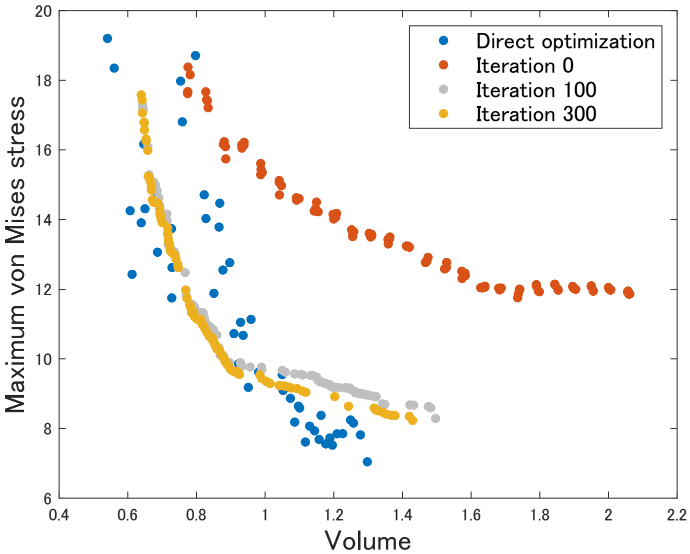


FIGURE 12: OBJECTIVE SPACE ON THE VOLUME AND MAXIMUM VON MISES STRESS

orderly Pareto front.

The designs shown in Fig. 9 obtained by direct optimization are structures that appear to be composed of straight members and often have triangular or rectangular voids. One advantage 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 generally curved structures with rounded appearances, as shown in Fig. 11, suggesting that stress concentration is avoided. In addition, it is found that various patterns are obtained due to the multimodality caused by the strong nonlinearity of the P-norm measure, while the final results of the proposed method in Fig. 11 relatively tend to have a similar feature in terms of their topology. Although we cannot prove that the obtained Pareto solutions are the global optima respectively, the result indicates that the data-driven MFTD method is likely to yield a unique set of Pareto solutions through an extensive search process.

5. CONCLUSION

This paper proposed a data-driven multifidelity topology design (MFTD) framework incorporating latent crossover that performs crossover in the latent space of the variational autoencoder (VAE). Since the latent space is constructed as continuous real numbers, this paper employed the simplex crossover (SPX) as a latent crossover operator based on the theoretical aspects of crossover in real-coded genetic algorithms (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 paper 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 mean compliance minimization 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 are 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.

To verify the efficacy of the proposed framework on different optimization problems, we plan to apply it to other problems involving strongly nonlinear physical phenomena, such as turbulence and geometric nonlinearity.

ACKNOWLEDGMENTS

The authors would like to thank Misato Kato, a graduate student at Osaka University, for providing the data used in Fig. 9. The second author was supported by JSPS KAKENHI Grant Numbers 20KK0329 and 20H02054.

REFERENCES

[1] Bendsøe, M. P. and Kikuchi, N. "Generating Optimal Topologies in Structural Design Using a Homogenization Method." *Computer Methods in Applied Mechanics and Engineering* Vol. 71 No. 2 (1988): pp. 197–224. DOI 10.1016/0045-7825(88)90086-2.

[2] Bendsøe, M. P. and Sigmund, O. *Topology Optimization: Theory, Methods, and Applications*. Springer Science & Business Media (2003).

[3] Mitchell, M. and Taylor, C. E. "Evolutionary Computation: An Overview." *Annual Review of Ecology and Systematics* Vol. 30 No. 1 (1999): pp. 593–616. DOI 10.1146/annurev.ecolsys.30.1.593.

[4] Goldberg, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison Wesley (1989).

[5] Wang, S. Y. and Tai, K. "Structural Topology Design Optimization Using Genetic Algorithms with a Bit-Array Representation." *Computer Methods in Applied Mechanics and Engineering* Vol. 194 No. 36-38 (2005): pp. 3749–3770. DOI 10.1016/j.cma.2004.09.003.

[6] Madeira, J. F. A., Pina, H. L. and Rodrigues, H. C. "GA Topology Optimization Using Random Keys for Tree Encoding of Structures." *Structural and Multidisciplinary Optimization* Vol. 40 (2010): pp. 227–240. DOI 10.1007/s00158-008-0353-1.

[7] Zhou, H. "Topology Optimization of Compliant Mechanisms Using Hybrid Discretization Model." *Journal of Mechanical Design* Vol. 132 No. 11 (2010): p. 111003. DOI 10.1115/1.4002663.

[8] Balamurugan, R., Ramakrishnan, C. V. and Swaminathan, N. "A Two Phase Approach Based on Skeleton Convergence and Geometric Variables for Topology Optimization Using Genetic Algorithm." *Structural and Multidisciplinary Optimization* Vol. 43 (2011): pp. 381–404. DOI 10.1007/s00158-010-0560-4.

[9] Sigmund, O. "On the Usefulness of Non-Gradient Approaches in Topology Optimization." *Structural and Multidisciplinary Optimization* Vol. 43 (2011): pp. 589–596. DOI 10.1007/s00158-011-0638-7.

[10] Kingma, D. P. and Welling, M. "Auto-Encoding Variational Bayes." arXiv preprint (2013). DOI 10.48550/arXiv.1312.6114.

[11] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. "Generative Adversarial Networks." arXiv preprint (2014). DOI 10.48550/arXiv.1406.2661.

[12] Regenwetter, L., Nobari, A. H. and Ahmed, F. "Deep Generative Models in Engineering Design: A Review." *Journal of Mechanical Design* Vol. 144 No. 7 (2022): p. 071704. DOI 10.1115/1.4053859.

[13] Guo, T., Lohan, D. J., Cang, R., Ren, M. Y. and Allison, J. T. "An Indirect Design Representation for Topology Optimization Using Variational Autoencoder and Style Transfer." *AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (2018): p. 0804. DOI 10.2514/6.2018-0804.

[14] Oh, S., Jung, Y., Kim, S., Lee, I. and Kang, N. "Deep Generative Design: Integration of Topology Optimization and Generative Models." *Journal of Mechanical Design* Vol. 141 No. 11 (2019): p. 111405. DOI 10.1115/1.4044229.

[15] Kazemi, H., Seepersad, C. C. and Kim, H. A. "Multiphysics Design Optimization via Generative Adversarial Networks." *Journal of Mechanical Design* Vol. 144 No. 12 (2022): p. 121702. DOI 10.1115/1.4055377.

[16] Yaji, K., Yamasaki, S. and Fujita, K. "Data-Driven Multifidelity Topology Design Using a Deep Generative Model: Application to Forced Convection Heat Transfer Problems." *Computer Methods in Applied Mechanics and Engineering* Vol. 388 (2022): p. 114284. DOI 10.1016/j.cma.2021.114284.

[17] Yamasaki, S., Yaji, K. and Fujita, K. "Data-Driven Topology Design Using a Deep Generative Model." *Structural and Multidisciplinary Optimization* Vol. 64 No. 3 (2021): pp. 1401–1420. DOI 10.1007/s00158-021-02926-y.

[18] Tsutsui, S., Yamamura, M. and Higuchi, T. "Multi-Parent Recombination with Simplex Crossover in Real Coded Genetic Algorithms." *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation* Vol. 1 (1999): pp. 657–664. URL https://dl.acm.org/doi/pdf/10.5555/2933923.2933986.

[19] Herrera, F., Lozano, M. and Verdegay, J. L. "Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis." *Artificial Intelligence Review* Vol. 12 (1998): pp. 265–319. DOI 10.1023/A:1006504901164.

[20] Kita, H. and Yamamura, M. "A Functional Specialization Hypothesis for Designing Genetic Algorithms." *IEEE International Conference on Systems, Man, and Cybernetics* Vol. 3 (1999): pp. 579–584. DOI 10.1109/ICSMC.1999.823277.

[21] Kita, H. and Yamamura, M. "Design Guidelines for Genetic Algorithms based on Function Specialization Hypothesis." *Journal of the Society of Instrument and Control Engineers* Vol. 38 No. 10 (1999): pp. 612–617. DOI 10.11499/sicejl1962.38.612.

[22] Kita, H., Ono, I. and Kobayasi, S. "Multi-Parental Extension of the Unimodal Normal Distribution Crossover for Real-Coded Genetic Algorithms." *Transactions of the Society of Instrument and Control Engineers* Vol. 36 No. 10 (2000): pp. 875–883. DOI 10.9746/sicetr1965.36.875.

[23] Beyer, H. G. and Deb, K. "On Aelf-Adaptive Features in Real-Parameter Evolutionary Algorithms." *IEEE Transactions on Evolutionary Computation* Vol. 5 No. 3 (2001): pp. 250–270. DOI 10.1109/4235.930314.

[24] Herrera, F., Lozano, M. and Sánchez, A. M. "A Taxonomy for the Crossover Operator for Real-Coded Genetic Algorithms: An Experimental Study." *International Journal of Intelligent Systems* Vol. 18 No. 3 (2003): pp. 309–338. DOI 10.1002/int.10091.

[25] Yaji, K., Yamasaki, S. and Fujita, K. "Multifidelity Design Guided by Topology Optimization." *Structural and Multidisciplinary Optimization* Vol. 61 (2020): pp. 1071–1085. DOI 10.1007/s00158-019-02406-4.

[26] Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* Vol. 6 No. 2 (2002): pp. 182–197. DOI 10.1109/4235.996017.

[27] Michalewicz, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer Berlin, Heidelberg (1996).

[28] Eshelman, L. J., Mathias, K. E. and Schaffer, J. D. "Crossover Operator Biases: Exploiting the Population Distribution." *Proceedings of the 7th International Conference on Genetic Algorithms* (1997): pp. 354–361.

[29] Ono, I., Kita, H. and Kobayashi, S. "A Real-Coded Genetic Algorithm Using the Unimodal Normal Distribution Crossover." *Advances in Evolutionary Computing: Theory and Applications* (2003): pp. 213–237.

[30] Črepinšek, M., Liu, S. H. and Mernik, M. "Exploration and Exploitation in Evolutionary Algorithms: A Survey." *ACM Computing Surveys* Vol. 45 No. 3 (2013). DOI 10.1145/2480741.2480752.

[31] Duysinx, P. and Bendsøe, M. P. "Topology Optimization of Continuum Structures with Local Stress Constraints." *International Journal for Numerical Methods in Engineering* Vol. 43 No. 8 (1998): pp. 1453–1478. DOI 10.1002/(SICI)1097-0207(19981230)43:8<1453::AID-NME480>3.0.CO;2-2.

[32] Le, C., Norato, J. A., Bruns, T., Ha, C. and Tortorelli, D. "Stress-Based Topology Optimization for Continua." *Structural and Multidisciplinary Optimization* Vol. 41 (2010): pp. 605–620. DOI 10.1007/s00158-009-0440-y.

[33] Holmberg, E., Torstenfelt, B. and Klarbring, A. "Stress Constrained Topology Optimization." *Structural and Multidisciplinary Optimization* Vol. 48 (2013): pp. 33–47. DOI 10.1007/s00158-012-0880-7.

[34] Norato, J. A., Smith, H. A., Deaton, J. D. and Kolonay, R. M. "A Maximum-Rectifier-Function Approach to Stress-Constrained Topology Optimization." *Structural and Multidisciplinary Optimization* Vol. 65 No. 10 (2022): p. 286. DOI 10.1007/s00158-022-03357-z.

[35] Bourdin, B. "Filters in Topology Optimization." *International Journal for Numerical Methods in Engineering* Vol. 50 No. 9 (2001): pp. 2143–2158. DOI 10.1002/nme.116.

[36] Shang, K., Ishibuchi, H., He, L. and Pang, L. M. "A Survey on the Hypervolume Indicator in Evolutionary Multi-objective Optimization." *IEEE Transactions on Evolutionary Computation* Vol. 25 No. 1 (2021): pp. 1–20. DOI 10.1109/TEVC.2020.3013290.

[37] Yang, R. J. and Chen, C. J. "Stress-Based Topology Optimization." *Structural Optimization* Vol. 12 (1996): pp. 98–105. DOI 10.1007/BF01196941.

[38] Duysinx, P. and Sigmund, O. "New Developments in Handling Stress Constraints in Optimal Material Distribution." *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization* (1998)DOI 10.2514/6.1998-4906.

[39] Li, L. and Khandelwal, K. "Volume Preserving Projection Filters and Continuation Methods in Topology Optimization." *Engineering Structures* Vol. 85 (2015): pp. 144–161. DOI 10.1016/j.engstruct.2014.10.052.

[40] Svanberg, K. "The method of moving asymptotes—a new method for structural optimization." *International Journal for Numerical Methods in Engineering* Vol. 24 No. 2 (1987): pp. 359–373. DOI 10.1002/nme.1620240207.