Critical Fractile Optimization Method Using Truncated Halton Sequence with Application to SAW Filter Design

Kiyoharu Tagawa(✉)

Kindai University, Higashi-Osaka 577-8502, Japan
tagawa@info.kindai.ac.jp

Abstract. This paper proposes an efficient optimization method to solve the Chance Constrained Problem (CCP) described as the critical fractile formula. To approximate the Cumulative Distribution Function (CDF) in CCP with an improved empirical CDF, the truncated Halton sequence is proposed. A sample saving technique is also contrived to solve CCP by using Differential Evolution efficiently. The proposed method is applied to a practical engineering problem, namely the design of SAW filter.

Keywords: Chance Constrained Problem · Empirical distribution

1 Introduction

In real-world optimization problems, various uncertainties have to be taken into account. Traditionally, there are two kinds of problem formulations for handling uncertainties in the optimization [11], namely the deterministic one and the stochastic one. Chance Constrained Problem (CCP) [13] is one of the possible formulation of the stochastic optimization problem. Since the balance between optimality and reliability can be taken with a probability in CCP, a number of real-world optimization problems have been formulated as CCPs [7,9].

CCP has been studied in the field of stochastic programming [13]. If the chance constraint is linear, CCP can be transformed to a deterministic optimization problem. Otherwise, CCP is so hard to solve because the time-consuming Monte Carlo simulation is needed to calculate the empirical probability that the chance constraint is satisfied. For solving CCP with the optimization methods of nonlinear programming, the stochastic programming assumes that the chance constraint is differentiable and convex. Even though Evolutionary Algorithms (EAs) are also reported to solve CCP [8,12], they use Monte Carlo simulations to evaluate the feasibility of every solution in the process of optimization.

In our previous paper [16], an optimization method based on Differential Evolution (DE) [14] was given to solve CCP without the Monte Carlo simulation. Specifically, CCP is described by using the Cumulative Distribution Function (CDF) of uncertain function value. In order to approximate CDF from samples,

A. Auger et al. (Eds.): PPSN 2018, LNCS 11101, pp. 464–475, 2018.
https://doi.org/10.1007/978-3-319-99253-2_37

an extended version of Empirical CDF (ECDF) [10], which is called Weighted ECDF (W_ECDF) [15], was employed. Thereby, for solving the CCP formulated with CDF, an Adaptive DE (ADE) combined with W_ECDF was used.

This paper focuses on a specific CCP known as the critical fractile formula [4] and improves the previous method [16] by introducing two new techniques. Firstly, the truncated Halton sequence is proposed to approximate CDF with W_ECDF more efficiently. Secondly, a new ADE equipped with a sample saving technique is proposed. The improved method is applied to the structural design of Surface Acoustic Wave (SAW) filters [2], which are widely used in the radio frequency circuits of mobile communication systems such as cellular phones.

2 Background and Problem Formulation

As stated above, there are two problem formulations for handling uncertainties. Robust optimization problem is a deterministic problem formulation [3]. Let $\boldsymbol{x} = (x_1, \cdots, x_D) \in \boldsymbol{X} \subseteq \Re^D$, $\boldsymbol{X} = [\underline{x}_j, \overline{x}_j]^D$, $j = 1, \cdots, D$ be a vector of decision variables, or a solution. The uncertainty is given by a vector of random variables $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_K) \in \boldsymbol{\Xi}$ with a support $\boldsymbol{\Xi} \subseteq \Re^K$. Robust optimization problem is defined with a measurable function $g : \boldsymbol{X} \times \boldsymbol{\Xi} \to \Re$ as

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} \gamma \quad \text{s.t. } \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} : g(\boldsymbol{x}, \boldsymbol{\xi}) \le \gamma. \tag{1}$$

The feasible solution $\boldsymbol{x} \in \boldsymbol{X}$ of the robust optimization problem in (1) has to satisfy the constraint $g(\boldsymbol{x}, \boldsymbol{\xi}) \le \gamma$ absolutely with 100% probability. Therefore, it seems to be too conservative from an engineering perspective.

CCP is a stochastic problem formulation [13]. By introducing any required sufficiency level $\alpha \in (0, 1)$ into an infinite number of constraints in (1), CCP reduces the conservativism of the robust optimization problem as

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} \gamma \quad \text{s.t. } \Pr(g(\boldsymbol{x}, \boldsymbol{\xi}) \le \gamma) \ge \alpha \tag{2}$$

where $\Pr(A)$ denotes the probability that an event A will occur.

Actually, CCP may have more than one constraint. Besides, there are two types of CCPs, namely separate CCP and joint CCP [13]. In this paper, separate CCP having only one chance constraint is considered as shown in (2).

The presence of the uncertainty in CCP leads to different results for repeated evaluations of the same solution $\boldsymbol{x} \in \boldsymbol{X}$. Since $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ is a vector of random variables, the function value $g(\boldsymbol{x}, \boldsymbol{\xi}) \in \Re$ in (2) becomes a random variable too. The CDF of $g(\boldsymbol{x}, \boldsymbol{\xi})$ depending on the solution $\boldsymbol{x} \in \boldsymbol{X}$ is defined as

$$F(\boldsymbol{x}, \gamma) = \Pr(g(\boldsymbol{x}, \boldsymbol{\xi}) \le \gamma). \tag{3}$$

By using the inverse CDF of $g(\boldsymbol{x}, \boldsymbol{\xi})$, an alternative formulation of the CCP in (2), which is known as the critical fractile formula [4], is written as

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} \gamma(\boldsymbol{x}) = F^{-1}(\boldsymbol{x}, \alpha) \tag{4}$$

where $\gamma(\boldsymbol{x})$ denotes the critical fractile $\gamma = \gamma(\boldsymbol{x})$ achieved by $\boldsymbol{x} \in \boldsymbol{X}$.

The probability distribution of $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ in CCP is usually known [13]. If the probability distribution of $g(\boldsymbol{x}, \boldsymbol{\xi}) \in \Re$ is also known or the inverse CDF of $g(\boldsymbol{x}, \boldsymbol{\xi})$ can be derived analytically, the CCP in (4) can be transformed into a deterministic optimization problem [4,13]. Otherwise, for solving the original CCP in (2), the probability $\Pr(g(\boldsymbol{x}, \boldsymbol{\xi}) \le \gamma)$ in (2) has to be evaluated repeatedly with the Monte Carlo simulation by changing the value of $\gamma \in \Re$.

3 Approximation of CDF

3.1 Empirical CDF (ECDF)

In real-world optimization problems, $g(\boldsymbol{x}, \boldsymbol{\xi})$ in (3) is too complex to derive its CDF analytically. Therefore, an approximation of the CDF is composed from samples. Let $g(\boldsymbol{x}, \boldsymbol{\xi}^n) \in \Re$, $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$, $n = 1, \cdots, N$ be a set of random samples of the function value $g(\boldsymbol{x}, \boldsymbol{\xi})$ in (3). The indicator function is defined as

$$\mathbb{1}(g(\boldsymbol{x}, \boldsymbol{\xi}^n) \le \gamma) = \begin{cases} 1 & \text{if } g(\boldsymbol{x}, \boldsymbol{\xi}^n) \le \gamma \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

From the samples $g(\boldsymbol{x}, \boldsymbol{\xi}^n)$, $n = 1, \cdots, N$, ECDF [10] is composed as

$$\mathbb{F}(\boldsymbol{x}, \gamma) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}(g(\boldsymbol{x}, \boldsymbol{\xi}^n) \le \gamma). \tag{6}$$

Let $\tilde{\mathbb{F}}(\boldsymbol{x}, \gamma)$ be a smoothed ECDF. The CDF of $g(\boldsymbol{x}, \boldsymbol{\xi})$ is approximated by $\tilde{\mathbb{F}}(\boldsymbol{x}, \gamma)$. Since $\tilde{\mathbb{F}}(\boldsymbol{x}, \gamma)$ is a monotone increasing function, we can get the inverse CDF value, or the critical fractile in (4), numerically as $\gamma = \tilde{\mathbb{F}}^{-1}(\boldsymbol{x}, \alpha)$.

As a drawback of ECDF, many samples are required to approximate CDF accurately because the samples $g(\boldsymbol{x}, \boldsymbol{\xi}^n)$, $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$ taken from the tail part of the probability distribution on $\boldsymbol{\Xi} \subseteq \Re^K$ are relatively few in number.

3.2 Weighted Empirical CDF (W_ECDF)

W_ECDF [15] is an improved ECDF to approximate CDF in (3). In order to take samples $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$ from $\boldsymbol{\Xi} \subseteq \Re^K$ uniformly, K-dimensional Halton Sequence (HS) is used instead of the random sampling. HS is a low-discrepancy sequence [5]. Let $\boldsymbol{\theta}^n \in \boldsymbol{\Theta} \subseteq \Re^K$, $n = 1, \cdots, N$ be a set of points generated as HS. Considering the support $\boldsymbol{\Xi} \subseteq \Re^K$, the region $\boldsymbol{\Theta} \subseteq \Re^K$ of HS is chosen as $\boldsymbol{\Theta} \supseteq \boldsymbol{\Xi}$.

Let $f : \boldsymbol{\Xi} \to [0, \infty)$ be the Probability Density Function (PDF) of $\boldsymbol{\xi} \in \boldsymbol{\Xi}$. Each of the points $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$ of HS is weighted by the PDF of $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ as $f(\boldsymbol{\theta}^n)$. Thereby, W_ECDF is composed from $g(\boldsymbol{x}, \boldsymbol{\theta}^n)$, $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$, $n = 1, \cdots, N$ as

$$\mathbb{F}(\boldsymbol{x}, \gamma) = \frac{1}{W} \sum_{n=1}^{N} f(\boldsymbol{\theta}^n)\, \mathbb{1}(g(\boldsymbol{x}, \boldsymbol{\theta}^n) \le \gamma) \tag{7}$$

where $W = f(\boldsymbol{\theta}^1) + \cdots + f(\boldsymbol{\theta}^n) + \cdots + f(\boldsymbol{\theta}^N)$.

By using a smoothed W_ECDF $\tilde{\mathbb{F}}(\boldsymbol{x}, \gamma)$, we can obtain $\gamma = \tilde{\mathbb{F}}^{-1}(\boldsymbol{x}, \gamma)$.

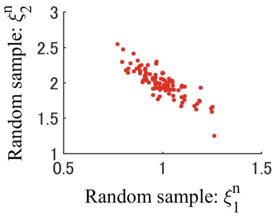


Fig. 1. RS: $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$

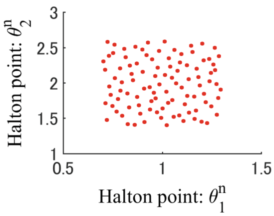


Fig. 2. HS: $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$

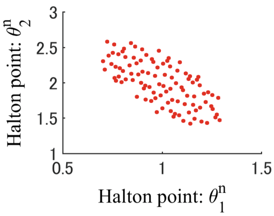


Fig. 3. THS: $\boldsymbol{\theta}^n \in \boldsymbol{S}$

3.3 Truncated Halton Sequence (THS)

In our previous paper [16], we supposed that all of the random variables $\xi_j \in \Re$, $j = 1, \cdots, K$ are mutually independent. Besides, for composing W_ECDF in (7) from $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$, the region $\boldsymbol{\Theta} \subseteq \Re^K$ of HS was given by a hyper-cube.

In this paper, Truncated HS (THS) is proposed to compose W_ECDF more efficiently. The region $\boldsymbol{S} \subseteq \boldsymbol{\Theta}$ of THS is defined with $\boldsymbol{\Theta} \subseteq \Re^K$ as

$$\boldsymbol{S} = \{\boldsymbol{\theta}^n \in \boldsymbol{\Theta} \mid f(\boldsymbol{\theta}^n) \geq f_{\min}\} \tag{8}$$

where the minimum PDF value $f_{\min}$ is a parameter given in advance.

By using the points $\boldsymbol{\theta}^n \in \boldsymbol{S}$, $n = 1, \cdots, N$ of THS for composing W_ECDF, we can eliminate futile points $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$ such as $f(\boldsymbol{\theta}^n) \approx 0$. The correlation between two random variables ξ_i and ξ_j, $i \neq j$ is also reflected in $\boldsymbol{\theta}^n \in \boldsymbol{S}$ naturally.

Example of W_ECDF with THS. Let's consider a stochastic function:

$$g(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{x}\,\boldsymbol{\xi}^T = x_1\,\xi_1 + x_2\,\xi_2 \tag{9}$$

where $\boldsymbol{\xi} \in \boldsymbol{\Xi} \subseteq \Re^2$ is following a 2-dimensional normal distribution such as

$$\boldsymbol{\xi} = (\xi_1, \xi_2) \sim \mathcal{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \mathcal{N}_2(1, 2, 0.1^2, 0.2^2, -0.8) \tag{10}$$

where ρ denotes the correlation coefficient between ξ_1 and ξ_2.

Figure 1 shows $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$, $N = 100$ generated by the Random Sampling (RS) of $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ in (10). Figure 2 shows $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$, $N = 100$. Figure 3 shows $\boldsymbol{\theta}^n \in \boldsymbol{S}$, $N = 100$. Since HS [5] is deterministic, the randomized HS [19] is used in this paper.

From the theory of probability [1], the PDF of $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ in (10) is

$$f(\boldsymbol{\xi}) = \frac{1}{2\,\pi\,\sqrt{|\boldsymbol{\Sigma}|}}\,\exp\left(-\frac{1}{2}\,(\boldsymbol{\xi} - \boldsymbol{\mu})\,\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{\xi} - \boldsymbol{\mu})^T\right) \tag{11}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and the covariance matrix $\boldsymbol{\Sigma}$ is given as

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_1\,\sigma_2\,\rho \\ \sigma_1\,\sigma_2\,\rho & \sigma_2^2 \end{pmatrix}. \tag{12}$$

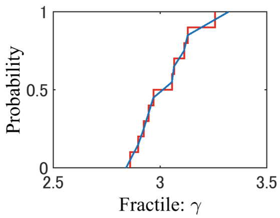


Fig. 4. ECDF in (6)

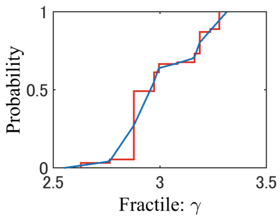


Fig. 5. W_ECDF in (7)

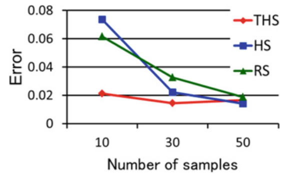


Fig. 6. Estimation error

From the linearity of the normal distribution, the value of $g(\boldsymbol{x}, \boldsymbol{\xi})$ in (9) also follows a normal distribution with mean $\mu_g(\boldsymbol{x})$ and variance $\sigma_g^2(\boldsymbol{x})$ as

$$g(\boldsymbol{x},\, \boldsymbol{\xi}) \sim \mathcal{N}(\mu_g(\boldsymbol{x}),\, \sigma_g^2(\boldsymbol{x})) = \mathcal{N}(\boldsymbol{x}\,\boldsymbol{\mu}^T,\, \boldsymbol{x}\,\boldsymbol{\Sigma}\,\boldsymbol{x}^T). \tag{13}$$

From (13), the CDF of $g(\boldsymbol{x}, \boldsymbol{\xi})$ in (9) can be derived exactly as

$$F(\boldsymbol{x},\, \gamma) = \Pr\left(\frac{g(\boldsymbol{x},\, \boldsymbol{\xi}) - \mu_g(\boldsymbol{x})}{\sigma_g(\boldsymbol{x})} \leq \frac{\gamma - \mu_g(\boldsymbol{x})}{\sigma_g(\boldsymbol{x})}\right) = \Phi\left(\frac{\gamma - \mu_g(\boldsymbol{x})}{\sigma_g(\boldsymbol{x})}\right) \tag{14}$$

where Φ denotes the CDF of the standard normal distribution [1].

ECDF and W_ECDF are used to approximate $F(\hat{\boldsymbol{x}}, \gamma)$ in (14) for a solution $\hat{\boldsymbol{x}} = (1,\, 1)$. Figure 4 shows an example of the step function of ECDF and its smoothed one. ECDF is composed from $N = 10$ samples $g(\hat{\boldsymbol{x}},\, \boldsymbol{\xi}^n)$, $\boldsymbol{\xi}^n \in \boldsymbol{\Xi}$. Similarly, Fig. 5 shows W_ECDF and its smoothed one composed from $N = 10$ samples $g(\hat{\boldsymbol{x}},\, \boldsymbol{\theta}^n)$, $\boldsymbol{\theta}^n \in \boldsymbol{S}$. From Figs. 4 and 5, the samples $g(\hat{\boldsymbol{x}},\, \boldsymbol{\theta}^n)$ for W_ECDF are distributed wider than the samples $g(\hat{\boldsymbol{x}},\, \boldsymbol{\xi}^n)$ for ECDF.

From (14), the critical fractile $\hat{\gamma} = F^{-1}(\hat{\boldsymbol{x}},\, \alpha) \approx 3.17$ is obtained exactly for $\alpha = 0.9$. Figure 6 compares between THS, HS, and RS in the estimation error $|\tilde{\mathbb{F}}^{-1}(\hat{\boldsymbol{x}},\, \alpha) - \hat{\gamma}|$ averaged over 10 runs. For generating $\boldsymbol{\theta}^n \in \boldsymbol{S}$ from $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$, $f_{\min} = 0.01$ is used in (8) and about 40% of $\boldsymbol{\theta}^n \in \boldsymbol{\Theta}$ are dumped. From Fig. 6, the estimation error with THS is small even if the sample size N is small.

4 Critical Fractile Optimization Method

4.1 Differential Evolution with Sample Saving Technique

By using the smoothed W_ECDF composed of N samples and a correction level $\beta \geq \alpha$, the CCP in (4), namely the critical fractile formula, is written as

$$\min_{\boldsymbol{x} \in \boldsymbol{X}} \;\gamma(\boldsymbol{x}) = \tilde{\mathbb{F}}^{-1}(\boldsymbol{x},\, \beta) \tag{15}$$

where the correction level is initialized as $\beta := \alpha$ and regulated in the procedure of the proposed optimization method as noted below if it is necessary.

The original versions of many EAs including DE have been developed to solve unconstrained optimization problems. Therefore, they can be applied directly to the CCP in (15). In this paper, one of the most successful ADE, namely JADE without archive [20], is used. As well as DE, JADE has a set of solutions $\boldsymbol{x}_i \in \boldsymbol{P}_t$, $i = 1, \cdots, N_P$ called population. An initial population $\boldsymbol{P}_0 \subseteq \boldsymbol{X}$ is generated randomly. Then every solution $\boldsymbol{x}_i \in \boldsymbol{P}_0$ is evaluated N times and the objective function $\gamma(\boldsymbol{x}_i)$ in (15) is estimated from $g(\boldsymbol{x}_i, \boldsymbol{\theta}^n)$, $\boldsymbol{\theta}^n \in \boldsymbol{S}$, $n = 1, \cdots, N$ as stated above. At each generation t, $\boldsymbol{x}_i \in \boldsymbol{P}_t$, $i = 1, \cdots, N_P$ is assigned to a parent in turn. By using the strategy named "DE/current-to-$p$best/1/bin" [20], a child $\boldsymbol{u}_i \in \boldsymbol{X}$ is generated from the parent $\boldsymbol{x}_i \in \boldsymbol{P}_t$ and evaluated N times. If $\gamma(\boldsymbol{u}_i) \le \gamma(\boldsymbol{x}_i)$ holds, the parent $\boldsymbol{x}_i \in \boldsymbol{P}_t$ is replaced by the child $\boldsymbol{u}_i \in \boldsymbol{X}$.

JADE applied to a real-world optimization problem spends most of time to evaluate children. The proposed sample saving technique called "pretest" can find and eliminate fruitless children with a few samples. When a newborn child $\boldsymbol{u}_i \in \boldsymbol{X}$ is compared with its parent $\boldsymbol{x}_i \in \boldsymbol{P}_t$, the pretest takes its samples $g(\boldsymbol{u}_i, \boldsymbol{\theta}^n)$ one by one. Let $m \le N$ be the number of samples obtained so far. From these samples, the empirical probability is calculated with weights as

$$\widehat{\Pr}(\gamma(\boldsymbol{u}_i) > \gamma(\boldsymbol{x}_i)) = \frac{1}{W} \sum_{n=1}^{m} f(\boldsymbol{\theta}^n)\, \mathbb{1}(g(\boldsymbol{u}_i, \boldsymbol{\theta}^n) > \gamma(\boldsymbol{x}_i)) \tag{16}$$

where $\widehat{\Pr}(A)$ denotes the predicted value of $\Pr(A)$ through observations.

If $\widehat{\Pr}(\gamma(\boldsymbol{u}_i) > \gamma(\boldsymbol{x}_i)) > 2\,(1-\beta)$ holds on the way, $\boldsymbol{u}_i \in \boldsymbol{X}$ is regarded as worse than $\boldsymbol{x}_i \in \boldsymbol{P}_t$ and discarded without evaluating $\gamma(\boldsymbol{u}_i) = \tilde{\mathbb{F}}^{-1}(\boldsymbol{u}_i, \beta)$.

JADE combined with Pretest is named JADEP. In the global optimization process of JADEP, the pretest is used locally in the competition between parent and child. Therefore, the pretest doesn't degrade the performance of JADE.

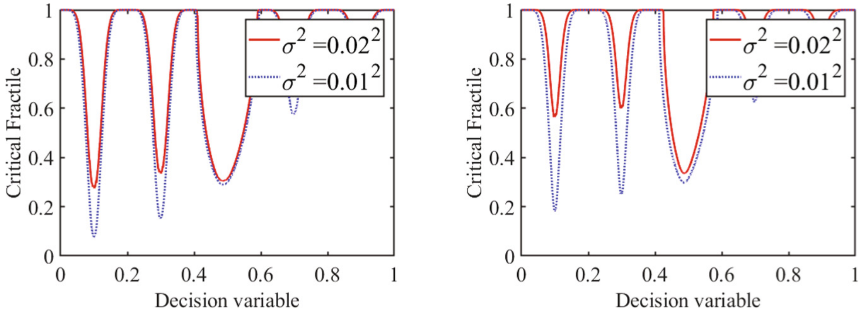
4.2 Verification of Solution Using Monte Carlo Simulation

We verify the feasibility of the solution $\boldsymbol{x}_b \in \boldsymbol{X}$ obtained by JADEP for the CCP in (15). Specifically, by using a huge number of random samples $g(\boldsymbol{x}_b, \boldsymbol{\xi}^n)$, $\boldsymbol{\xi}^n \in \boldsymbol{\varXi}$, $n = 1, \cdots, \hat{N}$, we calculate the empirical probability that the chance constraint of the CCP in (2) is satisfied with the solution $\boldsymbol{x}_b \in \boldsymbol{X}$ as

$$\widehat{\Pr}(g(\boldsymbol{x}_b, \boldsymbol{\xi}) \le \gamma(\boldsymbol{x}_b)) = \frac{1}{\hat{N}} \sum_{n=1}^{\hat{N}} \mathbb{1}(g(\boldsymbol{x}_b, \boldsymbol{\xi}^n) \le \gamma(\boldsymbol{x}_b)). \tag{17}$$

If $\widehat{\Pr}(g(\boldsymbol{x}_b, \boldsymbol{\xi}) \le \gamma(\boldsymbol{x}_b)) \ge \alpha$ holds, we regard that $\boldsymbol{x}_b \in \boldsymbol{X}$ is a feasible solution of the CCP in (2). Otherwise, we increase the value of the correction level β just a little and apply JADEP to the CCP in (15) again.

The sample size $\hat{N}$ in (17) is determined as follows. Let $\boldsymbol{x}^\star \in \boldsymbol{X}$ be the optimum solution of the CCP in (2) and $y_n = \mathbb{1}(g(\boldsymbol{x}^\star, \boldsymbol{\xi}^n) \le \gamma)$. Therefore, $\Pr(y_n = 1) = \alpha$ and $\Pr(y_n = 0) = (1-\alpha)$ hold. Let $\hat{y}$ be the sample mean of y_n, $n = 1, \cdots, \hat{N}$. From the central limit theorem [1], the confidence interval of the sample mean $\hat{y}$ is obtained for a confidence level $q \in (0, 1)$ as



(a) Sufficiency level $\alpha = 0.7$ (b) Sufficiency level $\alpha = 0.9$

Fig. 7. Landscapes of the critical fractile $\gamma(x)$ of $h(x, \xi)$ in (21)

$$\Pr(|\hat{y} - \alpha| \leq \varepsilon) = \Pr\left(|\hat{y} - \alpha| \leq z_{q/2}\sqrt{\frac{\alpha\,(1-\alpha)}{\hat{N}}}\right) \geq 1 - q \tag{18}$$

where ε is a margin of error and $z_{q/2}$ is the z-score for $q/2 \in (0,\, 0.5]$.

From desired ε and q in (18), the sample size $\hat{N}$ is determined as

$$\hat{N} = \left(\frac{z_{q/2}}{\varepsilon}\right)^2 \alpha\,(1-\alpha). \tag{19}$$

In this paper, $\varepsilon = 10^{-3}$ and $q = 0.01$ are chosen in (18). Therefore, if $\alpha = 0.9$ is given by the CCP in (2), we have $\hat{N} = 597,128$ from (19).

5 Numerical Experiment on Test Problem

5.1 Test Problem of CCP

The following function $h(x)$, $x \in [0,\, 1]$ has five unequal valleys [18].

$$h(x) = \begin{cases} 1 - e(x)\,|\sin(5\,\pi\,x)|^{0.5} & \text{if } 0.4 < x \leq 0.6 \\ 1 - e(x)\,\sin(5\,\pi\,x)^6 & \text{otherwise} \end{cases} \tag{20}$$

where $e(x) = \exp(-2\,\log_2((x - 0.1)/0.8)^2)$.

A random variable $\xi \in \Re$ is added to the function $h(x)$ in (20) as

$$h(x,\, \xi) = h(x + \xi),\ \ \xi \sim \mathcal{N}(0,\, \sigma^2). \tag{21}$$

Figure 7 illustrates the landscapes of the critical fractiles $\gamma(x) = F^{-1}(x,\, \alpha)$ evaluated from the CDF of $h(x,\, \xi)$ in (21). From Fig. 7, the value of $\gamma(x)$ depends not only on the sufficiency level α but also on the variance σ^2 in (21).

As an instance of the CCP in (2), $g(\boldsymbol{x},\, \boldsymbol{\xi})$ is defined as

$$g(\boldsymbol{x},\, \boldsymbol{\xi}) = \sqrt{h(x_1,\, \xi_1)\, h(x_2,\, \xi_2)} \tag{22}$$

where $h(x_j,\, \xi_j)$ is given by (21). ξ_1 and ξ_2 are mutually independent.

Table 1. Comparison of JADE and JADEP on the CCP defined by (22)

α	σ^2	JADE			JADEP			Rate
		$\gamma(\boldsymbol{x}_b)$	$\widehat{\Pr}(A)$	β	$\gamma(\boldsymbol{x}_b)$	$\widehat{\Pr}(A)$	β	
		0.305	0.718	0.805	0.305	0.718	0.809	0.139
0.7	0.02^2	(0.003)	(0.002)	(0.013)	(0.003)	(0.001)	(0.003)	(0.034)
		0.083	0.719	0.810	0.082	0.716	0.807	0.192
0.7	0.01^2	(0.000)	(0.002)	(0.000)	(0.001)	(0.005)	(0.004)	(0.029)
		0.340	0.908	0.950	0.340	0.908	0.949	0.385
0.9	0.02^2	(0.001)	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.048)
		0.213	0.909	0.949	0.196	0.909	0.943	0.407
0.9	0.01^2	(0.035)	(0.002)	(0.002)	(0.008)	(0.006)	(0.004)	(0.049)

5.2 Comparison Between JADEP and JADE

JADEP is compared with JADE on the CCP defined by $g(\boldsymbol{x}, \boldsymbol{\xi})$ in (22). They are coded by MATLAB. The population size $N_P = 20$ is used. The maximum number of generations is fixed to $G_{\max} = 100$. The sample size $N = 30$ is used to compose W_ECDF. JADEP and JADE are run 20 times in each case.

Table 1 shows the result of experiment averaged over 20 runs. In Table 1, $\gamma(\boldsymbol{x}_b)$ is the critical fractile attained with the best solution $\boldsymbol{x}_b$. The feasibility of $\boldsymbol{x}_b$ is ensured by the empirical probability $\widehat{\Pr}(A)$ as stated above. The rate denotes the percentage of children eliminated by the pretest of JADEP.

From the rate in Table 1, the pruning effect of the pretest depends on the case, but its works in all cases. From the result of Wilcoxon test about the value of $\gamma(\boldsymbol{x}_b)$, it is confirmed that there is no difference between JADE and JADEP in all cases. Consequently, the proposed pretest can reduce the number of the children examined N times without spoiling the quality of obtained solution.

6 Application to SAW Filter Design

6.1 Structure and Mechanism of SAW Filer

A SAW filter consists of some electrodes and reflectors, namely Inter Digital Transducers (IDTs) and Shorted Metal Strip Arrays (SMSAs), fabricated on a piezoelectric substrate. Figure 8 shows the symmetric structure of a resonator type SAW filter. The input-port of SAW filter is connected to two transmitter IDTs (IDT-Ts). The output-port is connected to a receiver IDT (IDT-R).

IDT-T converts electric input signals into acoustic signals. The acoustic signal of a specific frequency resonates between two SMSAs. The resonant frequency depends on the geometrical structure of SAW filter. Then IDT-R reconverts the enhanced acoustic signal to electric output signal. As a result, the resonator type SAW filter in Fig. 8 works as an electro-mechanical band-pass filter.

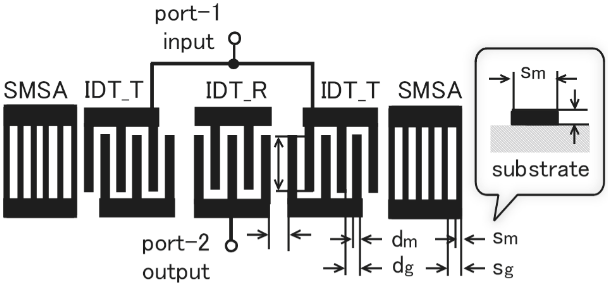


Fig. 8. Symmetric structure of resonator type SAW filter

Table 2. Design parameters of SAW filter

x_j	e_j	$[\underline{x}_j,\ \overline{x}_j]$	Description
x_1	—	[0.25, 0.35]	Thickness of electrode
x_2	—	[0.45, 0.55]	Metallization ratio of IDT: d_m/d_g
x_3	—	[0.45, 0.55]	Metallization ratio of SMSA: s_m/s_g
x_4	—	[1.0, 1.1]	Pitch ratio of SMSA: d_g/s_g
x_5	—	[1.0, 1.1]	Gap between IDT_R and IDT_T
x_6	—	[250.0, 350.0]	Overlap between electrodes
x_7	5.0	[50, 200]	Number of strips of SMSA
x_8	1.0	[10.5, 30.5]	Number of finger-pairs of IDT_R
x_9	0.5	[10, 30]	Number of finger-pairs of IDT_T

Table 3. JADEP

Parameter	Value
N_P	100
$G_{\max}$	200
N	100

6.2 Design of SAW Filer Under Uncertainty

In order to describe the structure of SAW filter in Fig. 8, design parameters, or decision variables $\boldsymbol{x} = (x_1, \cdots, x_9)$, are chosen as shown in Table 2. Each design parameter takes either a continuous value $x_j \in \Re$ or a discrete value at $e_j \in \Re$ interval. In the procedure of JADEP, a decision variable $x_j \in \Re$ is rounded to the nearest discrete value if it has to take a discrete value. Figure 8 also illustrates graphically the design parameters of SAW filter listed in Table 2.

We consider processing errors $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) \in \Re^3$ for the thickness of electrode x_1 and the metallization ratios of IDT and SMSA x_j, $j = 2, 3$ as

$$x_1\,(1 + \xi_1),\ \ \xi_1 \sim \mathcal{E}_{\mathcal{XP}}(\lambda) = \mathcal{E}_{\mathcal{XP}}(100) \tag{23}$$

where $\mathcal{E}_{\mathcal{XP}}(\lambda)$ denotes the exponential distribution with mean $1/\lambda$ and

$$x_j + \xi_j,\ \ j = 2, 3 \tag{24}$$

where $(\xi_2, \xi_3) \sim \mathcal{N}_2(\mu_2, \mu_3, \sigma_2^2, \sigma_3^2, \rho) = \mathcal{N}_2(0, 0, 0.01^2, 0.01^2, 0.5)$.

Each of IDT and SMSA can be modeled by an elemental circuit. Therefore, the equivalent circuit model of SAW filter is built up from the elemental circuits of IDT and SMSA [6,17], and then transformed to a network model as

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{25}$$

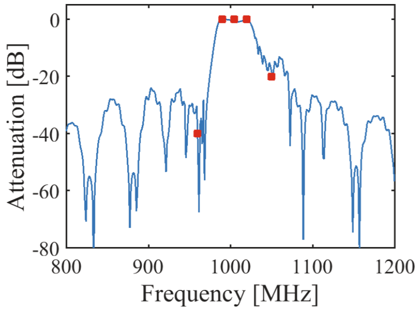


Fig. 9. Reference points $R(\omega_k)$ in (27)

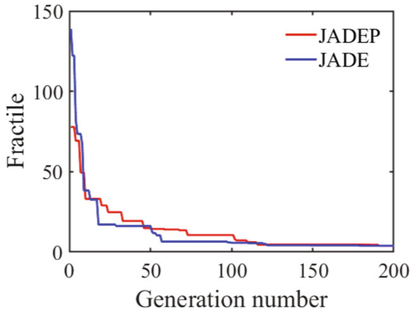


Fig. 10. Convergence plots

where a_p, $p = 1, 2$ denotes the input signal at port-p, while b_p denotes the output signal at port-p. Scattering parameter s_{pq} gives the transition characteristic from port-q to port-p, while s_{pp} gives the reflection characteristic at port-p.

From (25), the attenuation of SAW filter is defined as

$$L(\boldsymbol{x}, \boldsymbol{\xi}, \omega) = 20 \log_{10}(|s_{21}(\boldsymbol{x}, \boldsymbol{\xi}, \omega)|) \tag{26}$$

where s_{21} depends on $\boldsymbol{x} \in \boldsymbol{X}$, $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, and frequency $\omega \in \Re$.

For the attenuation in (26), some reference points $R(\omega_k)$ and weights c_k are specified at frequencies ω_k, $k = 1, \cdots, M$. Thereby, the design of SAW filter is formulated as the CCP in (2) by using the following function:

$$g(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{k=1}^{M} c_k \left(L(\boldsymbol{x}, \boldsymbol{\xi}, \omega_k) - R(\omega_k)\right)^2 \tag{27}$$

where $M = 5$ reference points are specified as shown in Fig. 9. Three points are given in the pass-bound and two points are given in the stop-bound.

6.3 Result of Experiment and Discussion

JADEP is compared with JADE on the above design problem of SAW filter. Table 3 shows the values of the parameters of JADEP. The same parameter values are used for JADE. Thereby, JADE and JADEP are run on a personal desktop computer (CPU: Intel Core i7@3.40GHz, OS: Windows 7).

Figure 10 shows a typical example of the convergence plots of JADEP and JADE which start from the same initial population. Figure 11 compares JADEP with JADE in the critical fractiles $\gamma(\boldsymbol{x}_b)$ of the obtained solutions $\boldsymbol{x}_b \in \boldsymbol{X}$ for some sufficiency levels α. From Fig. 11, we can confirm the trade-off between the values of $\gamma(\boldsymbol{x}_b)$ and α. From Figs. 10 and 11, there is no significant difference between JADEP and JADE in $\gamma(\boldsymbol{x}_b)$, namely the quality of solution.

Since each sample $g(\boldsymbol{u}_i, \boldsymbol{\theta}^n)$ has to be evaluated through the simulation of SAW filter, the efficiency of JADEP is much higher than JADE. In order to obtain a solution of the CCP in (15), JADEP spent 642 [sec] on average except

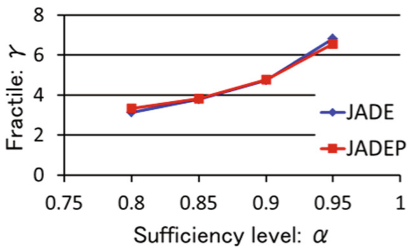


Fig. 11. Trade-off between γ and α

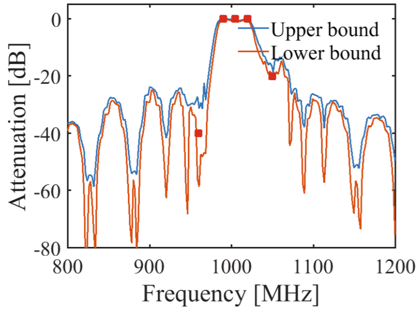


Fig. 12. Prediction interval in (28)

the verification of solution using the Monte Carlo simulation, while JADE spent 1,464 [sec]. The pretest of JADEP discarded more than 70% of the children.

The prediction interval of the attenuation in (26) is defined as

$$\Pr(\underline{L}(\boldsymbol{x},\,\omega) \le L(\boldsymbol{x},\,\boldsymbol{\xi},\,\omega) \le \overline{L}(\boldsymbol{x},\,\omega)) = (1-q). \tag{28}$$

From the inverse CDF of $L(\boldsymbol{x},\,\boldsymbol{\xi},\,\omega)$, the upper and lower bounds are

$$\begin{cases} \overline{L}(\boldsymbol{x},\,\omega) = F^{-1}(\boldsymbol{x},\,\omega,(1-q/2)) \\ \underline{L}(\boldsymbol{x},\,\omega) = F^{-1}(\boldsymbol{x},\,\omega,\,q/2). \end{cases} \tag{29}$$

By approximating the CDF in (29) with W_ECDF, the prediction interval in (28) can be estimated for a solution $\boldsymbol{x}_b \in \boldsymbol{X}$ found by JADEP. Figure 12 shows an example the prediction interval of $L(\boldsymbol{x}_b,\,\boldsymbol{\xi},\,\omega)$ estimated for $q = 0.1$. The reference points in Fig. 9 exist within the prediction interval in Fig. 12.

By using Figs. 11 and 12, which are provided by the proposed method, we can guarantee the performance of SAW filter under uncertainties.

7 Conclusion

For solving CCP efficiently, two new techniques were contrived to improve the optimization method based on JADEP and W_ECDF. Firstly, THS was used to compose W_ECDF from fewer samples. Secondly, the sample saving technique called Pretest was introduced into JADE. Finally, the contribution of this paper was demonstrated on the design of SAW filter formulated as CCP.

In this paper, an appropriate value of $f_{\min}$ in (8) was decided empirically considering the range of PDF and the number of points $\boldsymbol{\theta}^n \in \boldsymbol{S}$. Future work includes how to decide the value of $f_{\min}$ theoretically for generating THS.

Acknowledgment. This work was supported by JSPS (17K06508).

References

1. Ash, R.B.: Basic Probability Theory. Dover, Downers Grove (2008)
2. Bauer, T., Eggs, C., Wagner, K., Hagn, P.: A bright outlook for acoustic filtering. IEEE Microwave Mag. **16**(7), 73–81 (2015)
3. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
4. Geoffrio, A.M.: Stochastic programming with aspiration or fractile criteria. Manag. Sci. **13**(9), 672–679 (1967)
5. Halton, J.H.: On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. **2**(1), 84–90 (1960)
6. Hashimoto, K.: Surface Acoustic Wave Devices in Telecommunications - Modeling and Simulation. Springer, Heidelberg (2000). https://doi.org/10.1007/978-3-662-04223-6
7. Jiekang, W., Jianquan, Z., Guotong, C., Hongliang, Z.: A hybrid method for optimal scheduling of short-term electric power generation of cascaded hydroelectric plants based on particle swarm optimization and chance-constrained programming. IEEE Trans. Power Syst. **23**(4), 1570–1579 (2008)
8. Liu, B., Zhang, Q., Fernández, F.V., Gielen, G.G.E.: An efficient evolutionary algorithm for chance-constrained bi-objective stochastic optimization. IEEE Trans. Evol. Comput. **17**(6), 786–796 (2013)
9. Lubin, M., Dvorkin, Y., Backhaus, S.: A robust approach to chance constrained optimal power flow with renewable generation. IEEE Trans. Power Syst. **31**(5), 3840–3849 (2016)
10. Martinez, A.R., Martinez, W.L.: Computational Statistics Handbook with MATLAB ®, 2nd edn. Chapman & Hall/CRC, Boca Raton (2008)
11. Parkinson, A., Sorensen, C., Pourhassan, N.: A general approach for robust optimal design. J. Mech. Des. **115**(1), 74–80 (1993)
12. Poojari, C.A., Varghese, B.: Genetic algorithm based technique for solving chance constrained problems. Eur. J. Oper. Res. **185**, 1128–1154 (2008)
13. Prékopa, A.: Stochastic Programming. Kluwer Academic Publishers, Alphen aan den Rijn (1995)
14. Price, K.V., Storn, R.M., Lampinen, J.A.: Differential Evolution - A Practical Approach to Global Optimization. Springer, Heidelberg (2005). https://doi.org/10.1007/3-540-31306-0
15. Tagawa, K.: A statistical sensitivity analysis method using weighted empirical distribution function. In: Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing, pp. 79–84 (2016)
16. Tagawa, K., Miyanaga, S.: Weighted empirical distribution based approach to chance constrained optimization problems using differential evolution. In: Proceedings of IEEE CEC2017, pp. 97–104 (2017)
17. Tagawa, K., Sasaki, Y., Nakamura, H.: Optimum design of balanced SAW filters using multi-objective differential evolution. In: Dep, K., et al. (eds.) SEAL 2010. LNCS, vol. 6457, pp. 466–475. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-17298-4_50
18. Tsutsui, S.: A comparative study on the effects of adding perturbations to phenotypic parameters in genetic algorithms with a robust solution searching scheme. In: Proceedings of IEEE SMC, pp. 12–15 (1999)
19. Wang, X.: Randomized Halton sequences. Math. Comput. Model. **32**, 887–899 (2000)
20. Zhang, J., Sanderson, A.C.: JADE: adaptive differential evolution with optional external archive. IEEE Trans. Evol. Comput. **13**(5), 945–958 (2009)