Asymptotic Bayesian Optimization: A Markov Sampling-Based Framework for Design Optimization

D. J. Jerez[1], H. A. Jensen[2], M. Beer[1,3,4], and J. Chen[5]

[1] *Institute for Risk and Reliability, Leibniz Universität Hannover, 30167 Hannover, Germany.*

[2] *Departamento de Obras Civiles, Universidad Técnica Federico Santa María, 2390302 Valparaiso, Chile*

[3] *International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 200092 Shanghai, China.*

[4] *Institute for Risk and Uncertainty and School of Engineering, University of Liverpool, L69 7ZF Liverpool, United Kingdom.*

[5] *State Key Laboratory of Disaster Reduction in Civil Engineering & College of Civil Engineering, Tongji University, 200092 Shanghai, China.*

Abstract

This paper presents a Markov sampling-based framework, called Asymptotic Bayesian Optimization, for solving a class of constrained design optimization problems. The optimization problem is converted into a unified two-phase sample generation problem which is solved by an effective Markov chain Monte Carlo simulation scheme. First, an exploration phase generates designs distributed over the feasible design space. Based on this information, an exploitation phase obtains a set of designs lying in the vicinity of the optimal solution set. The proposed formulation can handle continuous, discrete, or mixed discrete-continuous design variables. Appropriate adaptive proposal distributions for the continuous and discrete design variables are suggested. The set of optimal solutions provides valuable sensitivity information of the different quantities involved in the problem with respect to the design variables. Representative examples including an analytical problem involving nonlinear benchmark functions, a classical engineering design problem, and a performance-based design optimization problem of a structural system under stochastic excitation are presented to show the effectiveness and potentiality of the proposed optimization scheme. Validation calculations show that the scheme is a flexible, efficient and competitive choice for solving a wide range of classical and complex engineering design problems.

Keywords: Discrete-continuous optimization, Dynamic systems, Markov sampling method, Metropolis-Hastings algorithm, Performance-based design, Proposal distributions, Stochastic

*H.A. Jensen, hector.jensen@usm.cl

optimization.

1. Introduction

Constrained optimization problems originate from a large number of involved engineering design processes. The problem is generally formulated in terms of minimizing a cost function or maximizing a utility function subject to multiple inequality constraints. In addition, and due to manufacturing limitations, some design variables cannot be considered as continuous but should be treated as discrete in many cases. Due to the significance of this type of problems, the development of efficient and robust constrained optimization algorithms has been an important area of research in engineering design [1, 2, 3].

One class of optimization schemes for solving constrained optimization problems is based on traditional mathematical optimization algorithms [1, 4, 5, 6]. This type of schemes has been extensively used in a large number of engineering design problems. Recently, stochastic-based search algorithms have been also proposed for constrained optimization. This class of algorithms can be classified into three main groups: evolution-based, physics-based, and swarm-based methods. Evolution-based methods, which are inspired by the laws of natural evolution, include Genetic Algorithms (GA) [7], Evolution Strategies (ES) [8], Genetic Programming (GP) [9], etc. On the other hand, physics-based methods such as Simulated Annealing (SA) [10], Gravitational Search Algorithm (GSA) [11], Subset Simulation-based algorithms (SuS) [12, 13], and Ray Optimization (RO) [14], replicate physical rules. Finally, swarm-based techniques that imitate the social behavior of different groups include Particle Swarm Optimization (PSO) [15], Ant Colony Optimization (ACO) [16], Harmony Search (HS) [17], Artificial Bee Colony (ABC) [18], etc. Some of the advantages of using stochastic search algorithms include their simplicity, flexibility, and local optima avoidance. In this context, one important issue related to constrained optimization is constraint-handling [19, 20, 21]. In this regard, a number of strategies have been suggested in the context of specific stochastic optimization algorithms such as evolutionary algorithms [22], simulated annealing [23], particle swarm optimization [24, 25], and subset simulation-based algorithm [13]. The previous stochastic optimization algorithms have been applied in a number of constrained optimization problems with different levels of efficiency and robustness.

Considering that design optimization of complex systems is a challenging problem, and the fact that there is no single method capable of solving all types of constrained optimization problems,

there is still room for further developments in this area. This motivates the attempt to develop an effective and flexible framework for solving complex constrained optimization problems, including problems with mixed discrete-continuous design variables. In the proposed scheme, which is called Asymptotic Bayesian Optimization (ABO), the optimization problem is converted into a problem of successively generating samples according to a sequence of probability distributions with supports increasingly concentrated in a vicinity of the optimum solution set. The samples are generated by a unified two-phase approach based on an efficient Markov chain Monte Carlo technique [26, 27]. The first phase corresponds to an exploration state which generates designs uniformly distributed over the feasible design space, while the second phase, which is an exploitation state, generates a set of designs lying in the vicinity of the optimal solution set. The proposed constraint-handling approach is direct and does not require special constraint-handling techniques. In fact, the same framework for obtaining samples in the vicinity of the optimal solution set is used for finding designs in the feasible space. The solution scheme can efficiently explore the sensitivity of the objective function and constraints with respect to the design variables in the feasible design space as well as in the neighborhood of the optimal solution set. In this context, appropriate adaptive proposal distributions are suggested for the continuous and discrete design variables. Moreover, the optimization algorithm can be implemented with few control parameters.

In summary, it is the objective of this contribution to propose a unified Markov sampling-based framework for solving a class of constrained optimization problems with application to design optimization. The contribution can be viewed as an extension and generalization of the work presented in [28, 29, 30, 31] in the sense that the same formulation can be used for a wide range of engineering design problems involving continuous, discrete, or mixed discrete-continuous design variables. The structure of the paper is as follows. In Section 2, the general formulation of the problem is presented. The relationship between the optimization and the sample generation problem is explained in Section 3. Section 4 outlines the sample generation scheme to be implemented. The exploration and exploitation phases are discussed in Sections 5 and 6, respectively. Some advantages of the proposed optimization scheme are highlighted in Section 7. The performance and capabilities of the proposed algorithm are demonstrated in Section 8 by means of three example problems. The paper closes with some final remarks.

2. Problem Formulation

Consider the constrained optimization problem formulated as

$$\begin{aligned} \text{Min}_{\mathbf{x}} \;\; & c(\mathbf{x}) \\ \text{s.t.} \;\; & g_j(\mathbf{x}) \leq 0 \;\; , \;\; j = 1, \ldots, n_g \\ & \mathbf{x} \in \mathbf{X} \end{aligned} \tag{1}$$

where $c(\mathbf{x})$ is the objective function, $g_j(\mathbf{x})$ is the j^{th} constraint function, n_g is the number of inequality constraints, $\mathbf{x}$ represents the set of design variables, and $\mathbf{X}$ is the search space. The set of design variables is defined as $\mathbf{x} = \langle \mathbf{x}_c^T, \mathbf{x}_d^T \rangle^T \in \mathbf{X} = \mathbf{X}_c \times \mathbf{X}_d \subset R^{n_c+n_d}$, where $\mathbf{x}_c(x_{ci}, i = 1, \ldots, n_c) \in \mathbf{X}_c \subset R^{n_c}$ denotes the set of continuous design variables, n_c is the number of continuous design variables, $\mathbf{x}_d(x_{di}, i = 1, \ldots, n_d) \in \mathbf{X}_d \subset R^{n_d}$ denotes the set of discrete design variables, and n_d is the number of discrete design variables. The side constraints for the continuous design variables are given by $x_{ci}^l \leq x_{ci} \leq x_{ci}^u, i = 1, \ldots, n_c$, where x_{ci}^l and x_{ci}^u are the corresponding lower and upper bounds. Finally, the set of available discrete values $\mathbf{X}_{di}$ for the i^{th} discrete design variable is written as $x_{di} \in \mathbf{X}_{di} = \{x_{di(j)}, j = 1, \ldots, n_{di}\}, i = 1, \ldots, n_d$. For convenience, the available discrete values are listed in an ascending order. The objective function $c(\mathbf{x})$ can be defined in terms of general cost functions, while the design constraints $g_j(\mathbf{x}) \leq 0, j = 1, \ldots, n_g$ can be given in terms of different design specifications.

3. Relationship Between Optimization Problem and Sample Generation Problem

Based on the connection between statistical mechanics and combinatorial optimization proposed in [10], with the idea of simulated annealing, the optimization problem in Eq. (1) can be converted into a problem of generating sample points or designs according to a specially devised distribution. To examine this connection, it is first observed that finding the minimum of the objective function $c(\mathbf{x})$ is equivalent to find the maximum of the function $\exp(-c(\mathbf{x})/K)$, for any given value of $K > 0$ [10]. The parameter K is usually called *temperature* by analogy with the Boltzmann-Gibbs distribution in statistical mechanics [32]. Next, artificially treating the design variables as random variables distributed over the feasible design space $\mathbf{X}_{\text{feasible}}$, where

$$\mathbf{X}_{\text{feasible}} = \left\{ \mathbf{x} = \langle \mathbf{x}_c^T, \mathbf{x}_d^T \rangle^T : \mathbf{x}_c \in \mathbf{X}_c \wedge \mathbf{x}_d \in \mathbf{X}_d \wedge g_j(\mathbf{x}) \leq 0, \, j = 1, \ldots, n_g \right\}, \tag{2}$$

consider the non-normalized distribution

$$f_K(\mathbf{x}) \propto \exp\left(-\frac{c(\mathbf{x})}{K}\right) \; I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}) \tag{3}$$

where $I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x})$ is the indicator function of the feasible design space $\mathbf{X}_{\text{feasible}}$, that is, $I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}) = 1$, for $\mathbf{x} \in \mathbf{X}_{\text{feasible}}$, and $I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}) = 0$, otherwise. The distribution $f_K(\mathbf{x})$ becomes proportional to $I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x})$ as $K \to \infty$, and it becomes spikier as $K \to 0$. The previous results correspond to the concept of annealing which indicates that as K decreases, the distribution $f_K(\mathbf{x})$ puts more and more of its probability mass into the set of feasible designs that maximize the function $\exp(-c(\mathbf{x})/K)$. Thus, a sample or design drawn from $f_K(\mathbf{x})$ will be in a vicinity of the optimal solution set $\mathbf{X}_c^*$ with very high probability when $K \to 0$ [30, 31]. Then, if a number of samples (designs) following the distribution $f_K(\mathbf{x})$ as $K \to 0$ can be generated, the sample points with the smallest value of $c(\mathbf{x})$ among the generated designs can provide a good approximation for the optimal solution set of the problem.

4. Sample Generation

As previously pointed out, the optimization problem can be converted into a problem of generating sample points (designs) according to the non-normalized distribution $f_K(\mathbf{x})$ with $K \to 0$. The generation of the required samples can be carried out by Markov chain Monte Carlo techniques [33]. This is a family of stochastic simulation algorithms for sampling from arbitrary probability density distributions. They are based on constructing a Markov chain whose state probability distribution converges to any desired target distribution as its stationary distribution. In this context, a number of standard algorithms may be used, including the independent Metropolis-Hastings algorithm [34], the random walk Metropolis-Hastings algorithm [35, 36], the asymptotically independent Markov sampling scheme [37], etc. In the present formulation, a highly effective Markov chain Monte Carlo simulation technique called the transitional Markov chain Monte Carlo (TMCMC) method is employed [26, 27]. It is noted that the treatment of the design variables as random variables is just a tool in the present formulation for setting the optimization problem into a sample generation problem [30, 31, 38]. In the framework of the TMCMC method, define a series of non-normalized intermediate distributions of the form

$$f_{K_0}(\mathbf{x}) \propto I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}) \quad , \quad f_{K_j}(\mathbf{x}) \propto \exp\left(-\frac{c(\mathbf{x})}{K_j}\right) \; I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}) \quad , \quad j = 1, 2, \ldots \tag{4}$$

where $\infty = K_0 > K_1 > \ldots > K_j > \ldots$ is a sequence of monotonically decreasing parameters with $K_j \to 0$ as $j \to \infty$. These parameters are constructed adaptively in such a way that the distributions $f_{K_j}(\mathbf{x})$ and $f_{K_{j+1}}(\mathbf{x})$ be similar [26, 34, 39]. This small change of the shape between consecutive distributions allows to efficiently obtain samples from $f_{K_{j+1}}(\mathbf{x})$ based on the samples from $f_{K_j}(\mathbf{x})$. To this end, different criteria can be used. In particular, a criterion based on the effective sample size technique [34, 39] is considered in the present implementation. According to this criterion, the value of K_{j+1}, given K_j, is chosen to satisfy the equation

$$\frac{\sum_{i=1}^{n} \exp\left(-2c(\mathbf{x}_i^j)\Delta K_j\right)}{\left(\sum_{i=1}^{n} \exp\left(-c(\mathbf{x}_i^j)\Delta K_j\right)\right)^2} = \frac{1}{\nu n} \tag{5}$$

where $\Delta K_j = 1/K_{j+1} - 1/K_j$, and $\nu \in (0,1)$ is a user-defined parameter. At the zeroth level, $j = 0$, uniformly distributed samples $\mathbf{x}_1^0, \ldots, \mathbf{x}_n^0$ are generated over the feasible design space $\mathbf{X}_{\text{feasible}}$, where n is the number of samples per stage. The samples at stage $j+1$, i.e., $\mathbf{x}_1^{j+1}, \ldots, \mathbf{x}_n^{j+1}$, $j = 0, 1, \ldots$, which are approximately distributed according to $f_{K_{j+1}}(\mathbf{x})$, are obtained by generating Markov chains using the Metropolis-Hastings algorithm [35, 36]. The lead sample of each chain, $\tilde{\mathbf{x}}^{j+1}$, is a sample from the previous stage, $\mathbf{x}_i^j$, drawn with probability equal to its normalized importance weight, $\bar{w}_i^j, i = 1, \ldots, n$ (see Appendix A). Each candidate design $\mathbf{x}^\star = \langle \mathbf{x}_c^{\star T}, \mathbf{x}_d^{\star T} \rangle^T$ is generated from an adaptive proposal distribution with independent continuous and discrete components [29] as described in appendices B and C, respectively. The candidate design is then accepted or rejected according to the procedure described in Appendix D. The procedure is repeated until the required number of samples has been obtained. It is noted that when the number of samples $n \to \infty$, the samples generated by the previous procedure are actually distributed according to the non-normalized intermediate distributions $f_{K_j}(\mathbf{x}), j = 1, 2, \ldots$ which ultimately are densely concentrated near the optimum solution set. Furthermore, the target non-normalized distribution $f_K(\mathbf{x}) \propto \exp(-c(\mathbf{x})/K) I_{X_{\text{feasible}}}(\mathbf{x}), K \to 0$, can be viewed as the posterior distribution of a Bayesian model updating problem where $\exp(-c(\mathbf{x})/K), K \to 0$, takes the role of the likelihood function and $I_{X_{\text{feasible}}}(\mathbf{x})$ of the prior distribution. These features are the reasons for naming the proposed optimization scheme "Asymptotic Bayesian Optimization". The generation of samples at stage $j+1$, i.e., $\mathbf{x}_1^{j+1}, \ldots, \mathbf{x}_n^{j+1}$, based on the samples generated at level j, is schematically represented in Fig. 1. A detailed implementation of the TMCMC method can be found in

[26, 27].

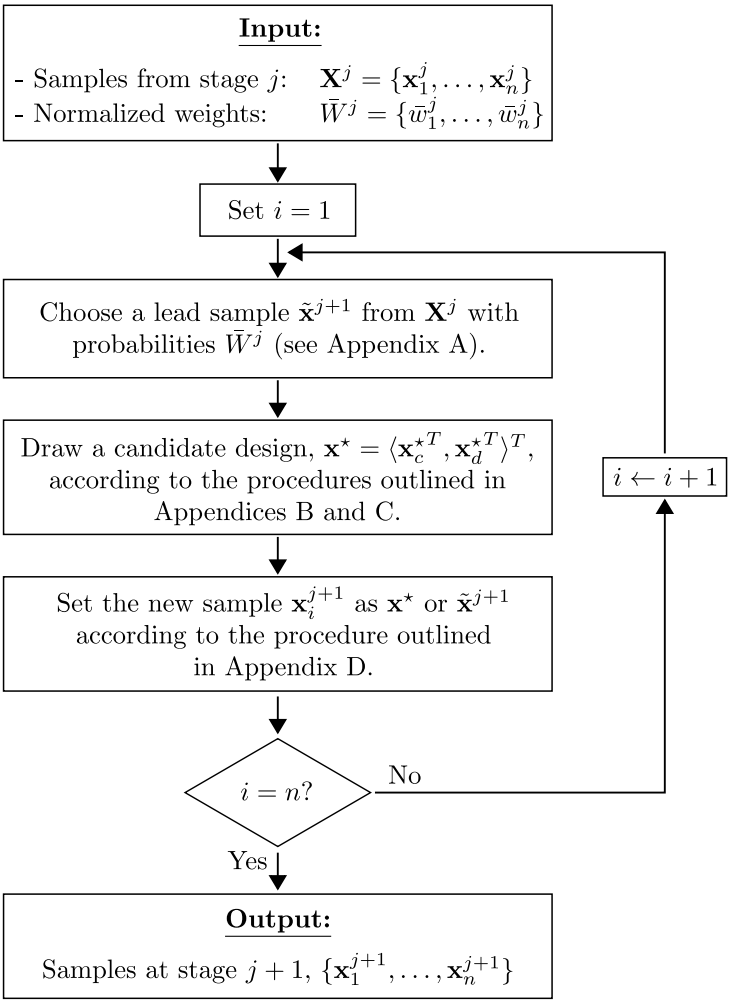

Figure 1: Sample generation process (Flowchart 1).

5. Exploration Phase

The first step of the proposed solution scheme requires the generation of a set of samples uniformly distributed over the feasible design space $\mathbf{X}_{\text{feasible}}$. To this end, an exploration phase that investigates the feasible domain of the search space is introduced. Following some of the ideas suggested in [28], define the auxiliary unconstrained optimization problem

$$\begin{aligned} \text{Min}_{\mathbf{x}} \quad & h(\mathbf{x}) = \text{Max}\{0, \text{Max}_{j=1,\ldots,n_g} \{g_j(\mathbf{x})\}\} \\ \text{s.t.} \quad & \mathbf{x} \in \mathbf{X} \end{aligned} \tag{6}$$

where all terms have been previously defined. Note that the optimal solution set $\mathbf{X}_h^*$ of the unconstrained optimization problem is given by

$$\mathbf{X}_h^* = \left\{ \mathbf{x} = \langle \mathbf{x}_c^T, \mathbf{x}_d^T \rangle^T : \mathbf{x}_c \in \mathbf{X}_c \wedge \mathbf{x}_d \in \mathbf{X}_d \wedge g_j(\mathbf{x}) \leq 0 \ , \ j = 1, \ldots, n_g \right\} \tag{7}$$

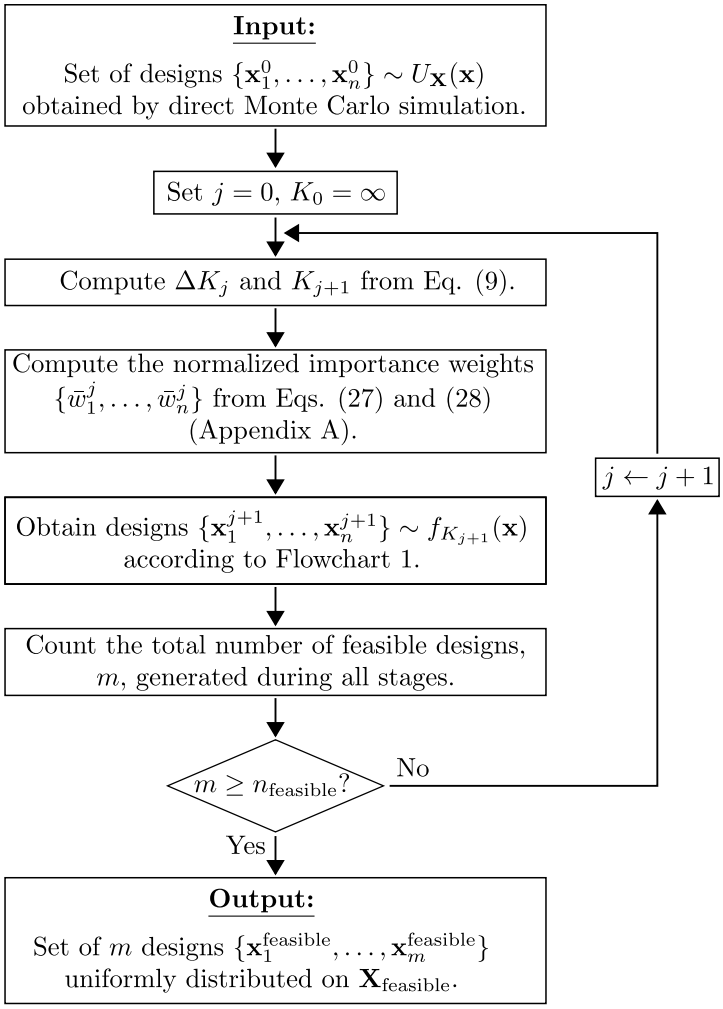

Figure 2: Exploration phase of ABO (Flowchart 2).

with minimum objective function value equal to 0. Thus, the optimal solution set $\mathbf{X}_h^*$ of the unconstrained optimization problem given in Eq. (6) coincides with the feasible design space defined in Eq. (2), i.e., $\mathbf{X}_{\text{feasible}} = \mathbf{X}_h^*$. Note that the problem is unconstrained in the sense that only side constraints on the design variables are considered. The previous optimization problem can be solved as indicated in the previous section, that is, by means of the TMCMC method. In this case, define the sequence of non-normalized intermediate distributions

$$f_{K_0}(\mathbf{x}) = U_{\mathbf{X}}(\mathbf{x}) \ , \ f_{K_j}(\mathbf{x}) \propto \exp\left(-\frac{h(\mathbf{x})}{K_j}\right) U_{\mathbf{X}}(\mathbf{x}) \ , \ j = 1, 2, \ldots \tag{8}$$

where $U_{\mathbf{X}}(\mathbf{x})$ is the uniform distribution defined over the set that characterizes the side constraints, i.e., $\mathbf{X} = \{\mathbf{x} = \langle \mathbf{x}_c^T, \mathbf{x}_d^T \rangle^T : \mathbf{x}_c \in \mathbf{X}_c \wedge \mathbf{x}_d \in \mathbf{X}_d\}$. Therefore, the samples at the first phase ($K_0 = \infty$) can be generated efficiently by direct Monte Carlo simulation. Furthermore, according to the sample generation scheme described in the previous section, the samples at the last stage of the process ($K_j \to 0$) represent designs with objective function value $h(\mathbf{x}) = 0$. In this phase, the parameter K_{j+1} satisfies the equation

$$\frac{\sum_{i=1}^{n} \exp\left(-2h(\mathbf{x}_i^j)\Delta K_j\right)}{\left(\sum_{i=1}^{n} \exp\left(-h(\mathbf{x}_i^j)\Delta K_j\right)\right)^2} = \frac{1}{\nu n} \tag{9}$$

according to the effective sample size technique [34, 39]. It can be shown that all feasible designs obtained during the different stages of the exploration phase are uniformly distributed over $\mathbf{X}_{\text{feasible}}$ [28]. Therefore, a possible stopping criterion is to obtain a sufficient amount of feasible designs during all stages. In the proposed implementation, the sampling process stops when $m \geq n_{\text{feasible}}$, where m is the total number of feasible designs obtained during the entire simulation process, and n_{feasible} is a user-defined target value. The procedure to generate samples in the feasible design space is schematically shown in Fig. 2.

6. Exploitation Phase

6.1. Samples in the Optimal Solution Set

The exploitation phase uses the samples generated in the exploration phase to obtain a set of designs lying in the vicinity of the optimal solution set $\mathbf{X}_c^*$. In the context of the approach presented in Section 4, the samples uniformly distributed over the feasible design space $\mathbf{X}_{\text{feasible}}$ are the ones obtained from the exploration phase. In addition, as $K_j \to 0$, the distribution $f_{K_j}(\mathbf{x})$ converges to a uniform distribution over the optimal solution set $\mathbf{X}_c^*$ and, therefore, the generated samples become more and more concentrated around $\mathbf{X}_c^*$ as the iterations progress. Clearly, for numerical implementation the algorithm should stop based on any suitable stopping rule. In the present formulation, the sampling procedure stops if a user-defined number of stages, $N_{\max}$, are completed or if the sample coefficient of variation (c.o.v.) of the objective function is below a

certain threshold. In particular, the optimization process stops at stage $j = 0, 1, \ldots$ if $j+1 = N_{\max}$ or, alternatively, $\delta_{j+1} < \gamma\delta_0$, where $\gamma \in (0, 1)$ is a user-defined parameter, and

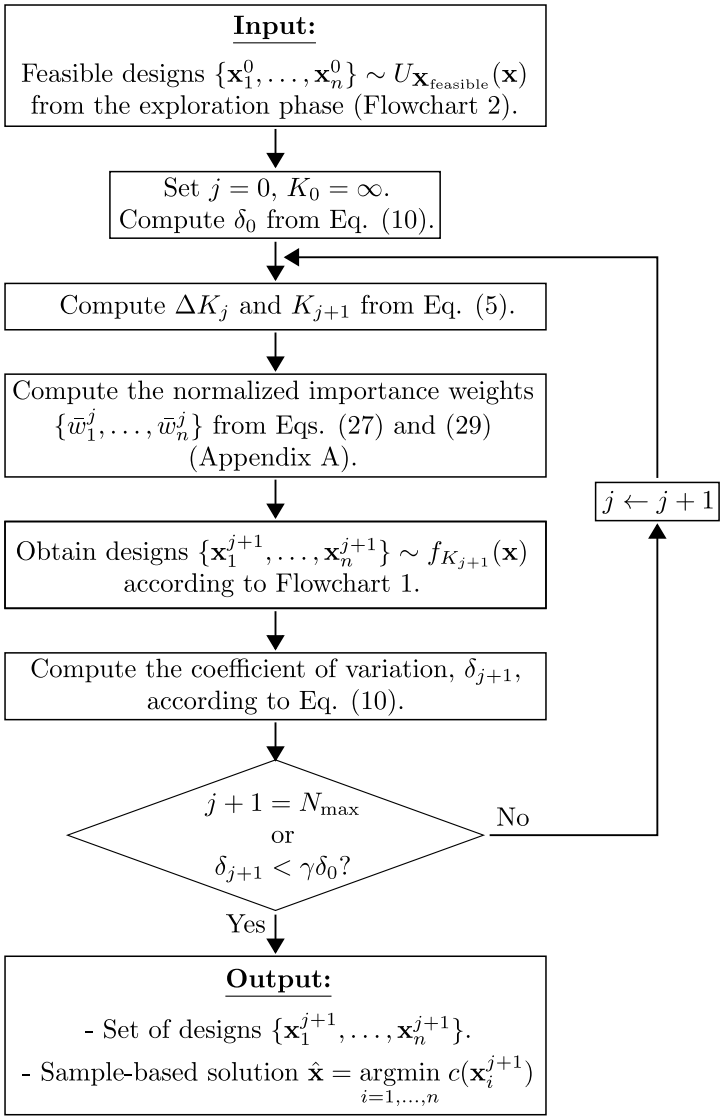

Figure 3: Exploitation phase of ABO (Flowchart 3).

$$\delta_j = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(c(\mathbf{x}_i^j) - \left[\frac{1}{n}\sum_{l=1}^{n} c(\mathbf{x}_l^j)\right]\right)^2} \Bigg/ \left(\frac{1}{n}\sum_{l=1}^{n} c(\mathbf{x}_l^j)\right) \quad , \quad j = 0, 1, \ldots \tag{10}$$

is the sample c.o.v. of the objective function during stage j. In other words, the algorithm runs until a prescribed number of stages are completed or until δ_{j+1} becomes less than some fraction γ of the initial sample c.o.v. of the objective function, δ_0. Smaller values for γ correspond to better approximations of the optimal solution set. It is noted that alternative stopping criteria can be implemented as well. The corresponding procedure to generate a set of candidate designs

is illustrated in Fig. 3.

6.2. Additional Remarks

Due to the annealing property of the approach, candidate designs with objective function values larger than those of the corresponding lead samples can be still accepted during the exploitation phase (see Appendix D). Thus, the probability of accepting a worse candidate solution during the initial stages is not negligible, but it decreases as the temperature parameter approaches zero. In this manner, sub-optimal regions in the feasible design space are potentially visited during the initial stages of the sampling process. The final stages, on the other hand, are almost completely focused on improving the objective function values of the designs. This feature is beneficial towards avoiding local optima and improving the robustness of the overall optimization scheme.

7. Practical Observations

Some practical benefits of using the proposed optimization scheme can be summarized as follows.

Adequacy for high performance computing. The sampling simulation technique under consideration, i.e., the TMCMC method, is very-well suited for parallel implementation in a computer cluster. In fact, the first level of the exploration phase, which corresponds to direct Monte Carlo simulation, can be fully scheduled in parallel. In addition, each of the subsequent levels of the exploration and exploitation phases produces a set of Markov chains that are perfectly parallel. Thus, a number of computer workers can handle the generation of samples corresponding to the different chains. This is very important when dealing with optimization problems involving expensive function evaluations.

Improved flexibility for decision making. Asymptotic Bayesian Optimization produces a set of nearly optimal solutions instead of a single optimal solution as well as a number of designs uniformly distributed over the feasible design space. In this manner, sensitivity information about the feasible designs and the final design can be obtained directly. This type of information can be advantageous in many practical cases where additional considerations or alternative criteria can be taken into account to select the appropriate final design. Thus, the approach provides flexibility to the decision-making process.

Robustness and effectiveness. Due to the theoretical basis of the formulation and the properties of the TMCMC method, the Asymptotic Bayesian Optimization scheme has high chances to reach

a vicinity of the global optimum in an effective manner, even in presence of multiple local optima, complex feasible design spaces, and problems with multiple discontinuous sub-feasible regions. Moreover, no restrictions are imposed on the number of constraints. Furthermore, due to the generality and flexibility of the formulation it can handle, in principle, different types of optimization problems. From the structural optimization point of view, these problems may include discrete-continuous design variables, complex linear and nonlinear systems, and performance-based design problems.

Implementation simplicity. Special constraint-handling techniques, such as penalty function methods or other approaches, are not necessary within the context of the proposed two-phase scheme. In fact, the same framework for obtaining samples in the vicinity of the optimal solution set (exploitation phase) is used for finding designs in the feasible set (exploration phase). In addition, the proposed approach requires the definition of few control parameters. These features represent an advantage from a practical viewpoint.

8. Examples

Due to the generality of the proposed optimization scheme, a wide range of optimization problems can be considered for evaluating its effectiveness. For clarity and conciseness, three representative numerical examples are chosen and presented in this section. First, a test problem involving highly nonlinear benchmark functions with continuous design variables is examined to illustrate the capabilities of the proposed method in detail and, in addition, to evaluate the effect of the algorithm parameters on its performance. Then, the effectiveness of the Asymptotic Bayesian Optimization scheme is demonstrated by two design optimization problems: a classical engineering design problem including mixed discrete-continuous design variables, and a performance-based discrete-design optimization problem of a structural system under stochastic excitation. As previously pointed out, a number of additional or alternative engineering design problems can be considered as well.

8.1. Example No. 1: Benchmark Functions

8.1.1. Optimization Problem

The constrained optimization problem of interest is stated as

$$\begin{aligned} \text{Min}_{\mathbf{x}} \;\; & c(x_1, x_2) \\ \text{s.t.} \;\; & g(x_1, x_2) \leq 0 \\ & -3.0 \leq x_i \leq 3.0\,, \; i = 1, 2 \end{aligned} \tag{11}$$

where the objective function $c(x_1, x_2)$ is the so-called *six-hump camel back* function given by

$$c(x_1, x_2) = 4.0x_1^2 - 2.1x_1^4 + x_1^6/3.0 + x_1x_2 - 4.0x_2^2 + 4.0x_2^4 \tag{12}$$

and the constraint function $g(x_1, x_2)$ is defined in terms of the *Schaffer function N.2* as

$$g(x_1, x_2) = 0.1 + \frac{\sin^2(x_1^2 - x_2^2) - 0.5}{[1 + 0.001(x_1^2 - x_2^2)]^2} \tag{13}$$

where x_1 and x_2 are the design variables which are treated as continuous variables. For illustration purposes, the objective and constraint functions are shown in Fig. 4. The left figure shows the objective function in the entire design space, while the right figure depicts the constraint function which is quite involved with abrupt variations in the design space. The corresponding feasible design space is illustrated in Fig. 5, where the two optimum solutions and some contours of the objective function are also shown. It is seen that the feasible design space is rather complex, as it involves several disconnected regions and some of them represent a small portion of the search space. In addition, the objective contours indicate that this example problem involves several disconnected sub-optimal regions. The optimal solutions of the optimization problem are given by $\mathbf{x}^* = \langle 0.0898, -0.7126 \rangle^T$ and $\mathbf{x}^* = \langle -0.0898, 0.7126 \rangle^T$, with optimum objective function $c(\mathbf{x}^*) = -1.0316$.

8.1.2. Exploration Phase

The following parameter values of the proposed approach are considered for the numerical implementation of the test problem: $\gamma = 0.05$ (stopping criterion parameter); $\nu = 0.5$ (effective sample size parameter); $n = 1000$ (number of samples per stage); and $n_{\text{feasible}} = 5000$ (target feasible sample size). In addition, the scaling parameter β, associated with the proposal distribution (see Appendix B), is determined by an adaptive scheme that monitors the acceptance rate of the updating process [40]. Figure 6 shows the evolution of the samples during the different stages of

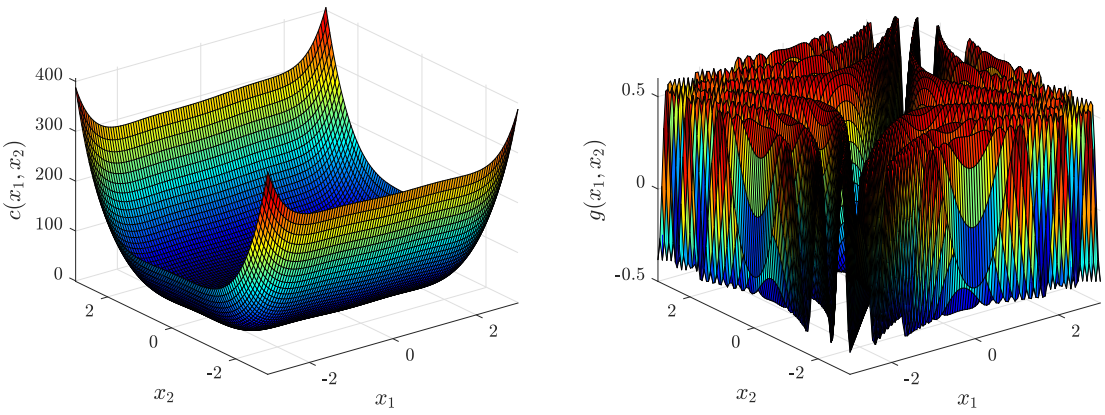

Figure 4: Left: Objective function in the entire design space. Right: Constraint function. Example No. 1.

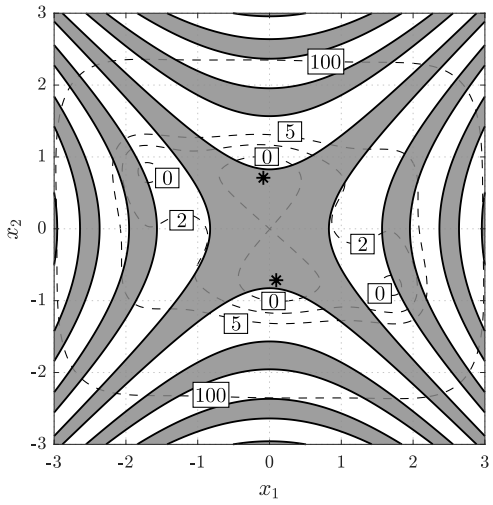

Figure 5: Sketch of the feasible design space (gray area), optimal solutions (*), and contours of the objective function (dashed lines). Example No. 1.

the exploration phase. The process stops at stage $j = 5$. That is, a total of six stages are required to verify the stopping criterion of the exploration phase. Thus, the samples generated at the final stage correspond to stage 5. It is observed that the samples tend to populate the feasible design space more effectively as the number of stages increases. In fact, the shape of the set of samples at the final stage is very similar to the feasible design space shown in Fig. 5. During the different stages, almost 5800 feasible samples are obtained. Thus, it is clear that the method generates a set of samples uniformly distributed over the feasible design space in an effective manner for this example.

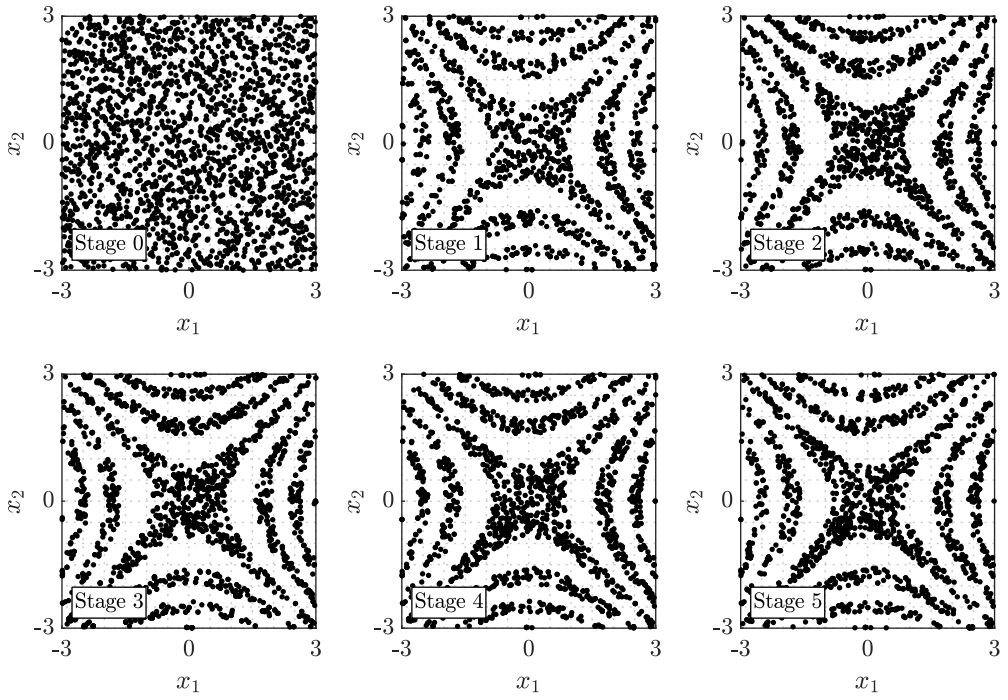

Figure 6: Samples generated during the different stages of the exploration phase. Example No. 1.

8.1.3. Exploitation Phase

Based on the set of feasible designs, the exploitation phase aims to obtain samples in a vicinity of the optimal solution set. Figure 7 shows the evolution of the samples obtained during the different stages of the exploitation phase. Note that the samples at the initial stage of the exploitation phase (stage 0 in the figure) correspond to the approximately 5800 feasible designs obtained during the exploration phase. It is seen that the intermediate distributions of the samples tend to be more and more concentrated near the two optimum solutions as the iterations progress. Two clusters can be clearly seen at the last stage of the exploitation phase, whose mean and optimum solutions are given in Table 1. The previous results indicate that the method effectively populates a vicinity of the optimal solution set during the final stages for this example, which is consistent with the theoretical foundations of the proposed approach. In summary, the preceding results show the applicability and effectiveness of the proposed approach in a rather complex optimization problem involving a disconnected feasible design space with multiple global optima and several disconnected sub-optimal regions.

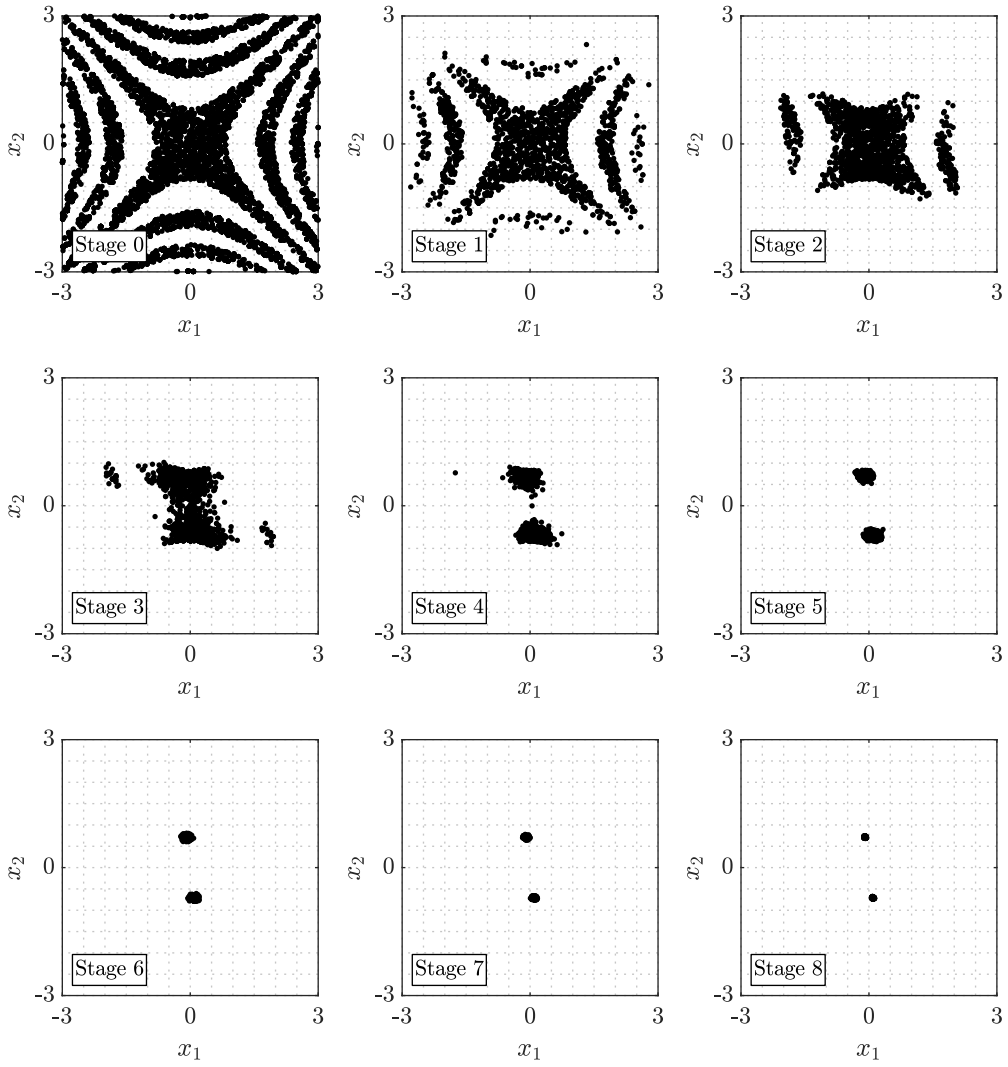

Figure 7: Evolution of samples obtained during the different stages of the exploitation phase. Example No. 1.

Design	Upper cluster	Lower cluster
Sample-based mean optimum $\bar{\mathbf{x}}$	$\bar{x}_1 = -0.0905$, $\bar{x}_2 = 0.7124$	$\bar{x}_1 = 0.0876$, $\bar{x}_2 = -0.7124$
Objective value $c(\bar{\mathbf{x}})$	-1.0316	-1.0316
Sample-based optimum $\hat{\mathbf{x}}^*$	$\hat{x}_1^* = -0.0901$, $\hat{x}_2^* = 0.7126$	$\hat{x}_1^* = 0.0896$, $\hat{x}_2^* = -0.7126$
Objective value $c(\hat{\mathbf{x}}^*)$	-1.0316	-1.0316
Actual optimum $\mathbf{x}^*$	$x_1^* = -0.0898$, $x_2^* = 0.7126$	$x_1^* = 0.0898$, $x_2^* = -0.7126$
Objective value $c(\mathbf{x}^*)$	-1.0316	-1.0316

Table 1: Sample-based mean, sample-based optimum, and actual optimum for each cluster. Example No. 1.

8.1.4. Statistical Performance

Preliminary validation calculations show that the number of samples per stage, n, plays an important role on the performance of the proposed approach. In this regard, a statistical analysis is carried out to study the influence of this parameter on the quality of the results. In particular, a total of 30 independent optimization runs are conducted for different numbers of samples per stage. They range from 50 to 1000. For comparison purposes, the target number of feasible samples is taken as $n_{\text{feasible}} = 2n$ and the total number of stages during the exploitation phase is limited to 8. The effective sample size parameter is taken as before, that is, $\nu = 0.5$. The statistical analysis is performed as follows. For the r^{th} independent run ($r = 1, \ldots, N_r = 30$), the sample-based optimum cost c_{opt}^r is obtained. In this framework, the optimum cost refers to the smallest objective function value found in each independent run. Based on these values, four statistical parameters are computed, namely, the best optimum cost $c_{\text{opt}}^{\text{best}}$, the worst optimum cost $c_{\text{opt}}^{\text{worst}}$, the average optimum cost $c_{\text{opt}}^{\text{avg}}$, and the coefficient of variation of the optimum cost $c_{\text{opt}}^{\text{c.o.v.}}$. The statistical parameters are formally defined as

$$
\begin{aligned}
c_{\text{opt}}^{\text{best}} &= \text{Min}_{r=1,\ldots,N_r}\ c_{\text{opt}}^r \quad , \quad c_{\text{opt}}^{\text{worst}} = \text{Max}_{r=1,\ldots,N_r}\ c_{\text{opt}}^r \\
c_{\text{opt}}^{\text{avg}} &= \frac{1}{N_r}\sum_{r=1}^{N_r} c_{\text{opt}}^r \quad , \quad c_{\text{opt}}^{\text{c.o.v.}} = \frac{\sqrt{\frac{1}{N_r-1}\sum_{r=1}^{N_r}(c_{\text{opt}}^r - c_{\text{opt}}^{\text{avg}})^2}}{\mid c_{\text{opt}}^{\text{avg}} \mid}
\end{aligned}
\tag{14}
$$

Table 2 shows the performance of the method in terms of the number of samples per stage. It is seen that the best optimum cost coincides with the reference value (-1.03163) even for a small number of samples per stage. On the other hand, the best, average, and worst optimum costs remain almost invariant, from the practical viewpoint, when the number of samples per stage is greater than 200. Moreover, as expected, the corresponding c.o.v. of the optimum cost reduces as n increases. The previous results indicate that the scheme is able to explore the design space in a very effective manner, even with a relatively small number of samples per stage. Another interpretation of these results is that the proposed method, which is based on Markov chain Monte Carlo simulation, exhibits a good performance in terms of its ergodicity in this particular problem. Although the appropriate value of n is problem-dependent, 200 samples per stage seem to be suitable for this example.

Finally, the influence of the effective sample size parameter, ν, on the quality of the optimization results is examined. Table 3 shows the best optimum cost, the worst optimum cost, the

optimal cost				
	best	mean	worst	c.o.v. (%)
n	($c_{\text{opt}}^{\text{best}}$)	($c_{\text{opt}}^{\text{avg}}$)	($c_{\text{opt}}^{\text{worst}}$)	($c_{\text{opt}}^{\text{c.o.v.}}$)
50	-1.03163	-1.03003	-1.01513	3.5×10^{-3}
100	-1.03163	-1.03154	-1.03103	1.5×10^{-4}
150	-1.03163	-1.03157	-1.03096	1.2×10^{-4}
200	-1.03163	-1.03158	-1.03135	7.0×10^{-5}
500	-1.03163	-1.03162	-1.03158	8.0×10^{-6}
1000	-1.03163	-1.03163	-1.03162	2.0×10^{-6}

Table 2: Statistical performance of the proposed scheme (ABO) in terms of the number of samples per stage (n). Example No. 1.

average optimum cost, and the c.o.v. of the optimum cost for different values of ν. The number of samples per stage is set equal to 200. It is seen that the best and average optimum costs are relatively similar to the reference solution, except for higher values of ν. In addition, the smallest difference between the best and worst optimum costs is obtained for intermediate values of the effective sample size parameter, that is, $\nu = 0.5$. Similarly, the c.o.v. of the optimum cost tends to decrease for intermediate values of ν. These results are reasonable due to the role that ν plays in the optimization process. On the one hand, consecutive intermediate distributions become more similar between each other for higher values of ν and, as a result, more stages are required to obtain a distribution that is densely concentrated in a vicinity of the optimum solution set. This slower converge leads to a higher variability of the sample-based optimum cost, since the number of stages is limited to eight. On the other hand, smaller values for ν allow more abrupt changes in the shape of consecutive intermediate distributions. This, in turn, can be detrimental to the effectiveness of the Metropolis-Hastings method and, eventually, diminish the accuracy of the sample-based optimum solution. Thus, intermediate values for ν (around 0.5) should be preferred in this example to reduce the variability of the optimum solutions for a fixed computational cost.

A similar behavior of the method performance with respect to ν is observed for alternative sample sizes. Analogously, the effect of n on the statistical performance of the method remains similar for different values of the effective sample size parameter. Such results are not presented here for conciseness and brevity. As indicated before, a number of samples per stage of around 200 and values for the effective sample size parameter roughly between 0.4 and 0.6 provide a

optimum cost				
	best	mean	worst	c.o.v. (%)
ν	($c_{\text{opt}}^{\text{best}}$)	($c_{\text{opt}}^{\text{avg}}$)	($c_{\text{opt}}^{\text{worst}}$)	($c_{\text{opt}}^{\text{c.o.v.}}$)
0.3	-1.03163	-1.03160	-1.03124	8.5×10^{-5}
0.5	-1.03163	-1.03158	-1.03135	7.0×10^{-5}
0.7	-1.03163	-1.03124	-1.02899	5.2×10^{-4}
0.9	-1.03091	-1.02093	-0.99980	9.1×10^{-3}

Table 3: Statistical performance of the proposed scheme (ABO) in terms of the effective sample size parameter (ν). Example No. 1.

reasonable tradeoff between efficiency and accuracy for this problem.

8.2. Example No. 2: A Classical Engineering Design Problem

8.2.1. Speed Reducer Design

The objective of this problem is to minimize the weight of the speed reducer shown in Fig. 8. A number of requirements associated with gear and shaft design practices, including bending and clamp constraints, strength conditions on gear shafts, permissible magnitude of deflection, etc., must be satisfied [41]. Seven design variables are involved in the optimization problem (see Fig. 8): width of the gear face (x_1), teeth module (x_2), number of pinion teeth (x_3), length between bearings of shafts 1 and 2 (x_4 and x_5, respectively), and diameter of shafts 1 and 2 (x_6 and x_7, respectively). The problem can be formulated as

$$\begin{aligned} \text{Min}_{\mathbf{x}} \quad & c(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0 \ , \ \ j = 1, \ldots, 11 \\ & \mathbf{x} \in \mathbf{X} \end{aligned} \tag{15}$$

where the cost function is given by

$$\begin{aligned} c(\mathbf{x}) = & \, 0.7854 x_1 x_2^2 (3.3333 x_3^2 + 14.9334 x_3 - 43.0934) \\ & - 1.508 x_1 (x_6^2 + x_7^2) + 7.4777 (x_6^3 + x_7^3) + 0.7854 (x_4 x_6^2 + x_5 x_7^2), \end{aligned} \tag{16}$$

the constraint functions are defined as

$$g_1(\mathbf{x}) = \frac{27}{x_1 x_2^2 x_3} - 1 \;\;,\;\; g_2(\mathbf{x}) = \frac{397.5}{x_1 x_2^2 x_3^2} - 1 \;\;,\;\; g_3(\mathbf{x}) = \frac{1.93 x_4^3}{x_2 x_6^4 x_3} - 1$$

$$g_4(\mathbf{x}) = \frac{1.93 x_5^3}{x_2 x_7^4 x_3} - 1 \;\;,\;\; g_5(\mathbf{x}) = \frac{\sqrt{\left(\frac{745 x_4}{x_2 x_3}\right)^2 + 16.9 \times 10^6}}{110.0 x_6^3} - 1$$

$$g_6(\mathbf{x}) = \frac{\sqrt{\left(\frac{745 x_5}{x_2 x_3}\right)^2 + 157.5 \times 10^6}}{85.0 x_7^3} - 1 \;\;,\;\; g_7(\mathbf{x}) = \frac{x_2 x_3}{40} - 1 \;\;,\;\; g_8(\mathbf{x}) = \frac{5 x_2}{x_1} - 1$$

$$g_9(\mathbf{x}) = \frac{x_1}{12 x_2} - 1 \;\;,\;\; g_{10}(\mathbf{x}) = \frac{1.5 x_6 + 1.9}{x_4} - 1 \;\;,\;\; g_{11}(\mathbf{x}) = \frac{1.1 x_7 + 1.9}{x_5} - 1 \tag{17}$$

and the side constraints on the design variables are $2.6 \leq x_1 \leq 3.6$, $0.7 \leq x_2 \leq 0.8$, $x_3 \in \{17, 18, \ldots, 28\}$, $7.3 \leq x_4 \leq 8.3$, $7.3 \leq x_5 \leq 8.3$, $2.9 \leq x_6 \leq 3.9$, and $5.0 \leq x_7 \leq 5.5$. Note that the number of pinion teeth (x_3) is an integer quantity, whereas the rest of design variables are continuous. Thus, this is a mixed discrete-continuous optimization problem. A thorough description of the objective and constraint functions can be found in [41].

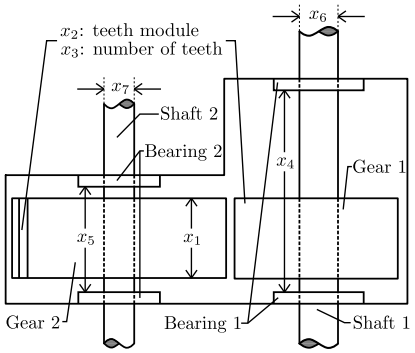

Figure 8: Schematic of the speed reducer design problem. Example No. 2.

The proposed approach is implemented to solve the speed reducer design problem. The effective sample size parameter is taken as $\nu = 0.4$ and $n = 500$ samples per stage are considered. The exploration phase stops after obtaining $n_{\text{feasible}} = 1000$ feasible designs. For illustration purposes, 40 exploitation stages are considered in this case. In addition, the parameters of the discrete proposal distribution (see Appendix C) are defined as $\lambda^* = 2$ and $\tau = 0.05$ for the exploration phase, and as $\lambda^* = 1$ and $\tau = 0$ for the exploitation phase. Note that the parameter λ^* is updated at the beginning of each exploitation stage according to an adaptive scheme that reuses information gathered during previous stages (see Appendix C). Preliminary validation calculations indicate

that the previous parameter values are appropriate in the context of this mixed discrete-continuous optimization problem.

First, an exploration phase is performed to generate samples uniformly distributed over the feasible set. After seven stages, a total of 1304 feasible designs are obtained. These designs are shown in Fig. 9 by means of two-dimensional projections and marginal histograms. It is noted that the feasible supports of variables x_1 and x_2 present the largest reduction when compared with the initial search space. This gives an insight on the sensitivity of the constraints with respect to the design variables.

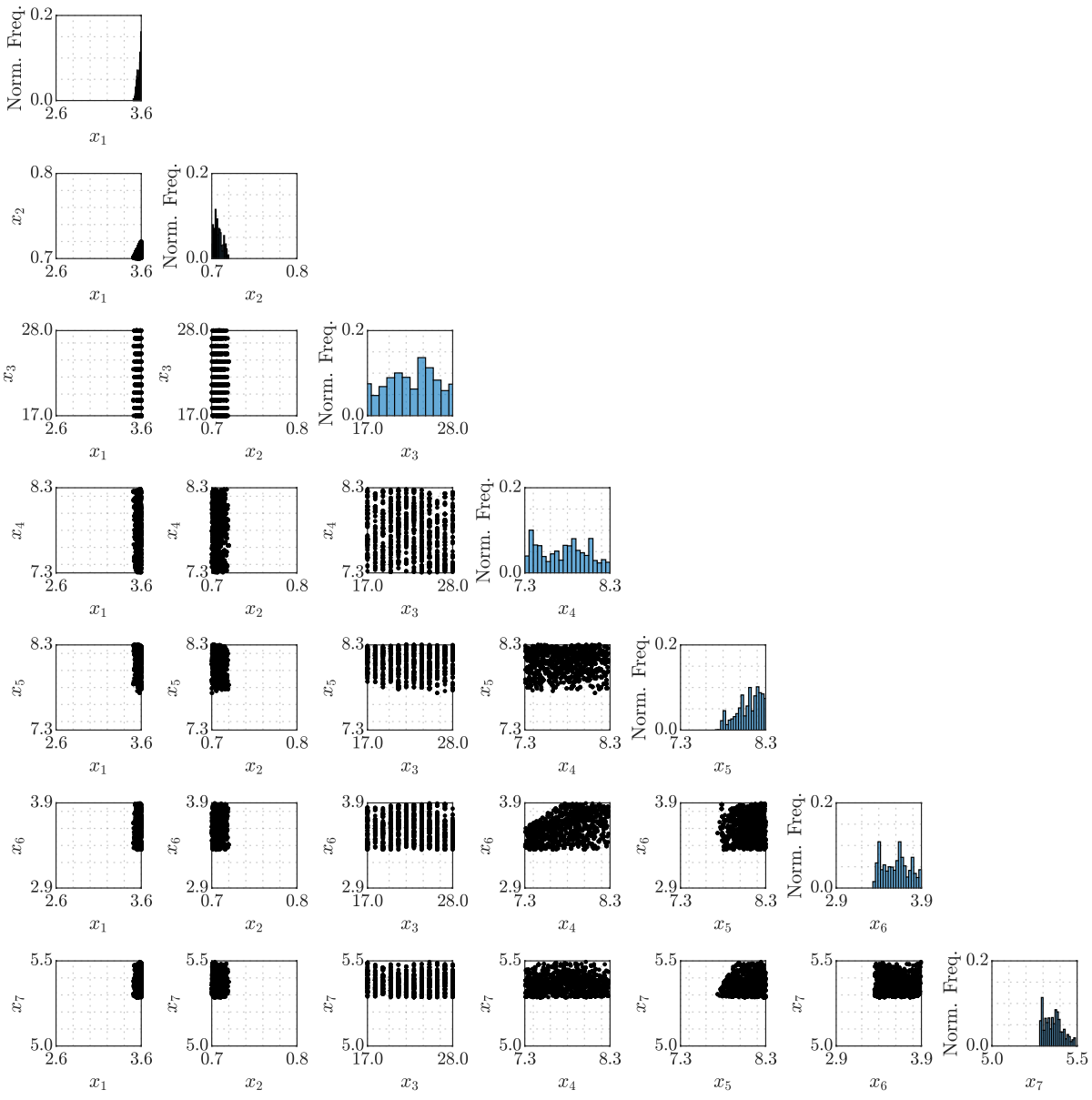

Figure 9: Samples uniformly distributed over the feasible design set. Example No. 2.

The samples shown in Fig. 9 are used as the initial set of feasible designs for the exploitation phase. During this phase, samples increasingly concentrated near the set that minimizes the objective function are iteratively generated. To illustrate this, Fig. 10 presents the marginal histograms obtained during stages $j = 0, 5, 10, 15$ and 39 (last stage). Note that the support of x_3 is reduced significantly during the initial stages. In addition, the vicinity of the optimum solution set becomes more densely populated as the optimization process continues, as expected. Correspondingly, the range of the objective function is also reduced during the different stages of the proposed approach. The corresponding sample-based optimum obtained is $\mathbf{x}^* = \langle 3.50000, 0.70000, 17, 7.30001, 7.71533, 3.35022, 5.28665 \rangle^T$ with $c(\mathbf{x}^*) = 2994.4727$.

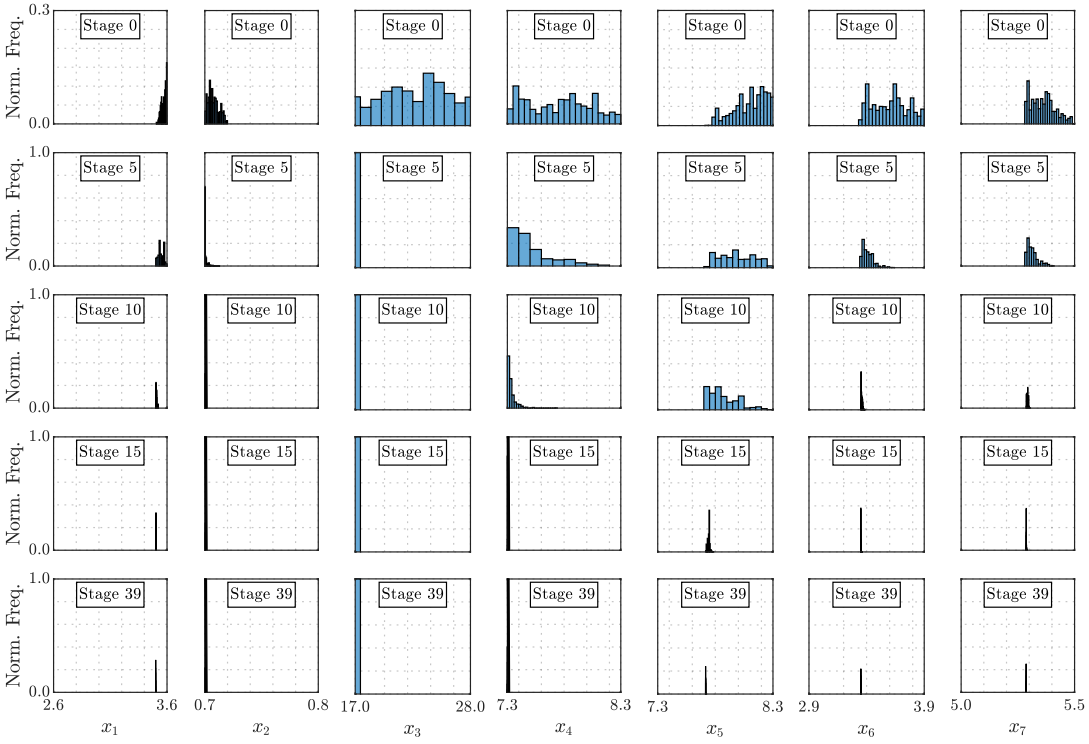

Figure 10: Marginal histograms obtained during different stages of the exploitation phase. Example No. 2.

Finally, Table 4 presents the best solution obtained with the proposed approach across 30 independent runs. In addition, the table also shows the solution of this problem obtained by other stochastic search techniques reported in several references. They include Social Behavior inspired Optimization technique (SBO) [42], Particle Swarm Optimization with Differential Evolution (PSO-DE) [43], Differential Evolution with Level Comparison (DELC) [44], and Mine Blast Algorithm (MBA) [45]. For reference purposes, the number of function calls corresponding to each solution is also reported in the table. It is seen that the solution obtained by the proposed

approach for this application example is competitive with respect to the ones obtained by other stochastic search techniques. Thus, the proposed framework can be an efficient choice for this problem.

Design variables	Optimal values				
	SBO [42]	PSO-DE [43]	DELC [44]	MBA [45]	ABO
x_1	3.500 000	3.500 000	3.500 000	3.500 000	3.500 000
x_2	0.700 000	0.700 000	0.700 000	0.700 000	0.700 000
x_3	17	17	17	17	17
x_4	7.300 000	7.300 000	7.300 000	7.300 033	7.300 004
x_5	7.800 000	7.800 000	7.715 319	7.715 772	7.715 321
x_6	3.350 215	3.350 215	3.350 240	3.350 218	3.350 215
x_7	5.286 683	5.286 683	5.286 654	5.286 654	5.286 655
Weight (lb)	2996.232 157	2996.348 165	2994.471 066	2994.482 453	2994.471 550
Function calls	70000	54350	30000	25000	24000

Table 4: Best solutions reported by different algorithms. Example No. 2.

8.3. Example No. 3: A Performance-Based Optimization Problem

8.3.1. Description of the Model

The structural model under stochastic excitation shown in Figure 11, which has been borrowed from [29], is considered in this example. Each floor is supported by 48 columns as shown in Fig. 11. The columns on axes A, C, D, and F contribute to the horizontal resistance of the floors in the x direction, while those on axes B and E work primarily in the y direction. In addition, a bracing system consisting of tubular steel brace elements is placed in axes A, C, D and F acting in the x direction, and in axes 1, 2, 7 and 8 acting in the y direction. A typical configuration of the brace elements is shown in Fig. 12. Thus, a total of 128 brace elements are used in the model, with Young's modulus $E = 2.1 \times 10^{11}$ N/m^2 and weight density $\rho = 7.42$ ton/m^3. All floors have a constant height equal to 3.2 m, leading to a total height of 12.8 m. For a given floor, all columns are assumed to be equal and their specifications are given in Table 5 [46]. It is assumed that each floor may be represented with sufficient accuracy as rigid within the $x-y$ plane when compared to the flexibility of the horizontal resistant elements. Hence, each floor can be represented by three

degrees of freedom, i.e., two translational displacements along the x and y axes, and a rotational displacement about the z axis. The associated masses $m_x = m_y$ and m_z are taken as constant for all floors and equal to 5.98×10^5 kg and 1.10×10^8 kg m^2, respectively. In addition, a 2% of critical damping is assumed in the model. It is noted that no attempt has been made to consider a more detailed structural model since the objective of this example is to evaluate the effectiveness of the proposed optimization scheme in a performance-based optimization problem involving a structural dynamical system under stochastic excitation.

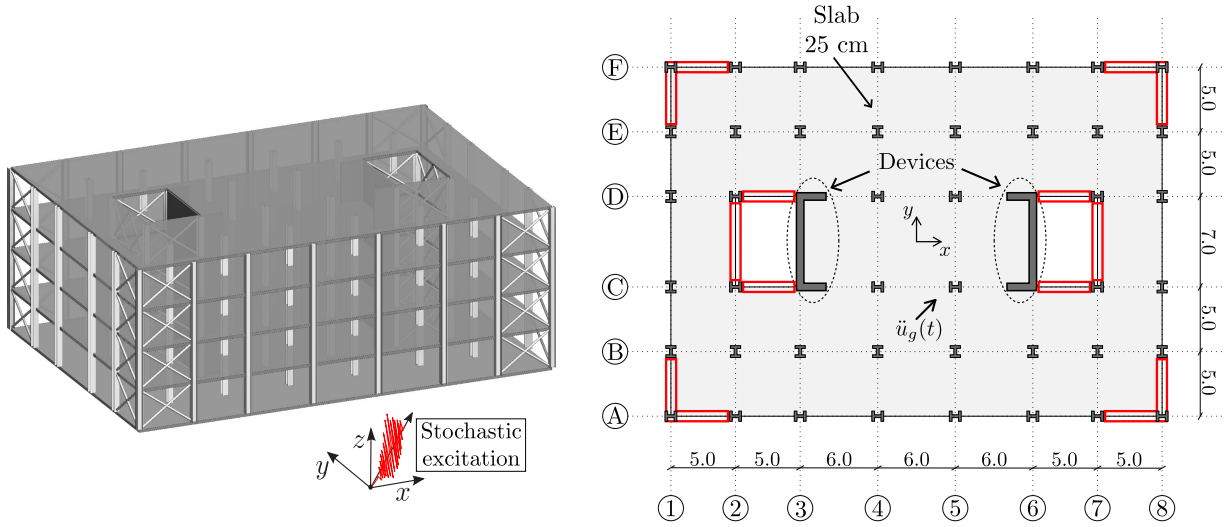

Figure 11: Isometric (left) and plan (right) view of the structural model. Example No. 3.

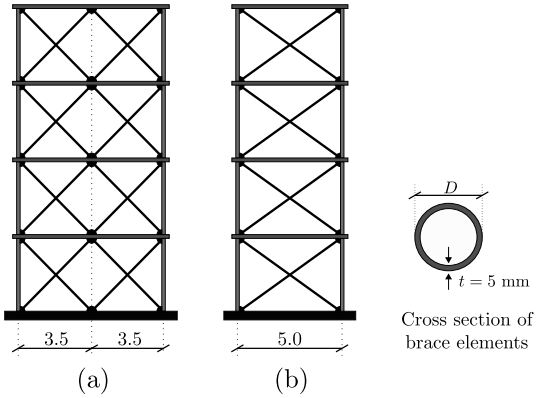

Figure 12: Typical configuration of brace elements. (a) Brace system in axes 2 and 7. (b) Brace system in axes A, C, D, F, 1 and 8. Example No. 3.

The system is subjected to a base acceleration modeled as a non-stationary stochastic process. In particular, a stochastic model based on a point-source model is considered [47, 48]. The model is characterized by a white noise sequence and a series of parameters such as radiation pattern, shear

Floor	Type of section
1	W24 × 131
2	W24 × 131
3	W24 × 104
4	W24 × 104

Table 5: Specification of column elements. Example No. 3.

wave velocity in the vicinity of the source, corner frequencies, local site conditions, velocity pulse parameters, and additional seismicity parameters such as the moment magnitude and rupture distance. Details of the procedure can be found in [47, 48, 49, 50]. The excitation is applied at 45° with respect to the x axis and its duration is taken as $T = 15$ s with a sampling interval equal to $\Delta t = 0.01$ s. Based on these values and according to the stochastic excitation model under consideration, it can be shown that more than 1500 random variables are involved in the generation of base acceleration samples [47]. Thus, a high-dimensional uncertain parameter space is involved in this problem. For illustration purposes, Fig. 13 shows a synthetic ground motion sample corresponding to the stochastic point-source model.

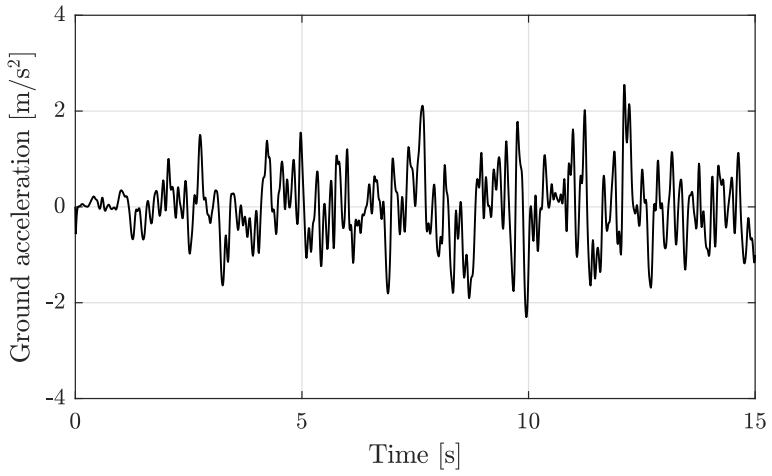

Figure 13: Synthetic ground acceleration sample from the stochastic point-source model. Example No. 3.

For improved seismic behavior, the model is reinforced with nonlinear devices at each floor. At each floor, six devices are implemented as shown in the floor plan of the model (see Fig. 11). Specifically, four devices in the x direction and two devices in the y direction are considered. These elements provide additional resistance against relative displacements between floors. The devices follow the inter-story restoring force law $\kappa(t) = k_d\Big(\delta(t) - \gamma^1(t) + \gamma^2(t)\Big)$, where k_d denotes the

initial stiffness of the device, $\delta(t)$ is the relative displacement between floors, and $\gamma^1(t)$ and $\gamma^2(t)$ denote the plastic elongations of the device. Using the auxiliary variable $\mu(t) = \delta(t) - \gamma^1(t) + \gamma^2(t)$, the plastic elongations $\gamma^1(t)$ and $\gamma^2(t)$ are specified by the differential equations [51]

$$\dot{\gamma}^i(t) = \varsigma(i)\dot{\delta}(t)H(\varsigma(i)\dot{\delta}(t))\times \left[H(\varsigma(i)\mu(t) - \mu_y)\frac{\varsigma(i)\mu(t) - \mu_y}{\mu_p - \mu_y}H(\mu_p - \varsigma(i)\mu(t)) + H(\varsigma(i)\mu(t) - \mu_p)\right] \quad , \quad i = 1, 2 \tag{18}$$

where $H(\cdot)$ denotes the Heaviside step function, $\varsigma(1) = 1$, $\varsigma(2) = -1$, μ_y is a parameter specifying the onset of yielding, and k_d μ_p is the maximum restoring force of the device. The values $\mu_p = 6.0 \times 10^{-3}$ m, $\mu_y = 0.7u_p$, and $k_d = 6.0 \times 10^8$ N/m are used for the nonlinear elements. A typical displacement-restoring force curve of the nonlinear devices is shown in Fig. 14. Note that, because of yielding, energy dissipation due to hysteresis is introduced in the structural response.

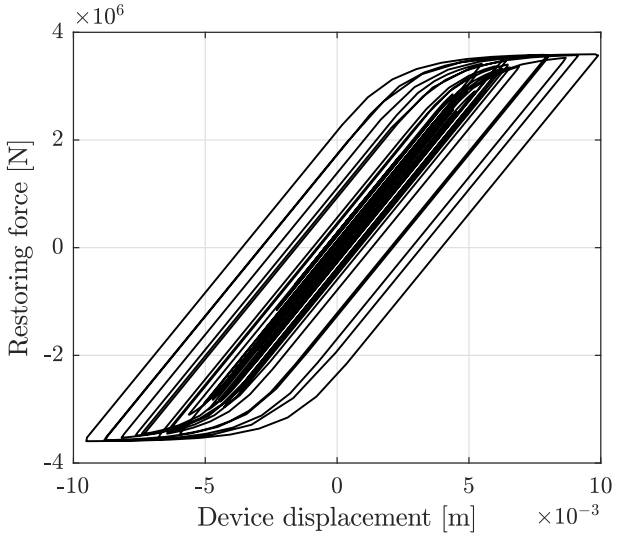

Figure 14: Typical displacement-restoring force curve of the nonlinear device. Example No. 3.

8.3.2. Design Problem

The objective function for the optimization problem is defined in terms of the expected value of the root-mean-square (RMS) of the displacement response at the top floor, while the design constraints are given in terms of cost limitations, geometric requirements and availability of section sizes. The vector of design variables $\mathbf{x}$ is defined in terms of the areas of the cross-sections of the steel brace elements. For illustration purposes, the brace elements located every two floors are linked to two design variables, one associated with the x direction and one with the y direction.

This leads to four design variables in total. The axes and stories corresponding to each design variable are given in Table 6. In this setting, x_1 and x_3 are associated with the brace elements pointing in the x direction, while x_3 and x_4 with those pointing in the y direction. The values for the different design variables (areas of tubular cross-sections) must be selected from the discrete set of available member sizes presented in Table 7. Thus, each design variable can be chosen from a discrete set of 48 tubular elements and, therefore, more than 5×10^6 different configurations for the bracing system can be devised.

Design variables	Stories	Axes
x_1	1-2	A C D F
x_2	1-2	1 2 7 8
x_3	3-4	A C D F
x_4	3-4	1 2 7 8

Table 6: Design variables. Example No. 3.

The design problem is written as

$$\begin{aligned} \text{Min}_{\mathbf{x}} \quad & c(\mathbf{x}) \\ \text{s.t.} \quad & g_1(\mathbf{x}) = \bar{v}(\mathbf{x}) - 1 \leq 0 \\ & g_2(\mathbf{x}) = x_3/x_1 - 1 \leq 0 \\ & g_3(\mathbf{x}) = x_4/x_2 - 1 \leq 0 \\ & x_i \in X \quad , \quad i = 1, \ldots, 4 \end{aligned} \tag{19}$$

where $\mathbf{x} = \langle x_1, x_2, x_3, x_4 \rangle^T$ is the vector of design variables, $c(\mathbf{x})$ is the objective function, $\bar{v}(\mathbf{x})$ is a normalized cost function associated with the total volume of the bracing system, and X represents the set of available discrete values for the design variables given in Table 7. It is noted that $g_1(\mathbf{x})$ is associated with cost limitations, whereas $g_2(\mathbf{x})$ and $g_3(\mathbf{x})$ impose geometric conditions on the final design. As previously pointed out, the objective function corresponds to the expected value of the root-mean-square of the displacement response at the top floor, that is, $c(\mathbf{x}) = E_{\boldsymbol{\theta}}[RMS_d(\mathbf{x}, \boldsymbol{\theta})]$ where $\boldsymbol{\theta}$ is the vector of uncertain parameters involved in the stochastic excitation model, $E_{\boldsymbol{\theta}}(\cdot)$ denotes the expectation with respect to the distribution of $\boldsymbol{\theta}$, and

D (in)	A (mm^2)	D (in)	A (mm^2)	D (in)	A (mm^2)
2	719	4	1517	6	2315
2 1/8	769	4 1/8	1567	6 1/8	2365
2 1/4	819	4 1/4	1617	6 1/4	2415
2 3/8	869	4 3/8	1667	6 3/8	2465
2 1/2	919	4 1/2	1717	6 1/2	2515
2 5/8	969	4 5/8	1767	6 5/8	2565
2 3/4	1019	4 3/4	1817	6 3/4	2615
2 7/8	1069	4 7/8	1866	6 7/8	2664
3	1118	5	1916	7	2714
3 1/8	1168	5 1/8	1966	7 1/8	2764
3 1/4	1218	5 1/4	2016	7 1/4	2814
3 3/8	1268	5 3/8	2066	7 3/8	2864
3 1/2	1318	5 1/2	2116	7 1/2	2914
3 5/8	1368	5 5/8	2166	7 5/8	2964
3 3/4	1418	5 3/4	2216	7 3/4	3014
3 7/8	1468	5 7/8	2265	7 7/8	3063

Table 7: Available values for the design variables. Example No. 3.

$$RMS_d(\mathbf{x},\boldsymbol{\theta}) = \frac{1}{N_T}\sqrt{\sum_{k=1}^{N_T}\{u_x^2(t_k,\mathbf{x},\boldsymbol{\theta}) + u_y^2(t_k,\mathbf{x},\boldsymbol{\theta})\}} \tag{20}$$

where $u_x(t_k,\mathbf{x},\boldsymbol{\theta})$ and $u_y(t_k,\mathbf{x},\boldsymbol{\theta})$ are the x and y components, respectively, of the roof displacement at time instant $t_k, k = 1,\ldots,N_T, N_T = 1500$, for a given realization of $\boldsymbol{\theta}$. The estimate of $c(\mathbf{x})$ is evaluated by means of Monte Carlo simulation [33]. In particular, 2000 samples are considered, that is, the evaluation of the objective function at each design involves 2000 dynamic analyses in this case. In addition, the normalized cost function is defined as $\bar{v}(\mathbf{x}) = \sum_{i=1}^{4} v_i^* x_i$, with normalized constants $v_1^* = v_3^* = 0.9402 \times 10^{-4}$ and $v_2^* = v_4^* = 1.2363 \times 10^{-4}$.

8.3.3. Results

The proposed optimization framework is implemented to handle this performance-based design problem considering $n = 200$ samples per stage and $\nu = 0.5$ for the effective sample size parameter. In addition, the stopping criteria for the exploration and exploitation phases are to obtain $n_{\text{feasible}} = 400$ feasible designs and to verify $\delta_{j+1} < \gamma\delta_0$ with $\gamma = 0.01$ (see Section 6.1), respectively. Moreover, the proposal distributions for the discrete design variables (see Appendix C) are defined by $\lambda_l^* = 3$, $\tau_l = 0.05$, $l = 1,\ldots,4$, for the exploration phase, and by $\lambda_l^* = 1$, $\tau_l = 0$, $l = 1,\ldots,4$, for the exploitation phase. As in the previous example, the parameters λ_l^* are updated during the different stages of the exploitation phase according to the strategy presented in Appendix C. Additional validation computations indicate that the previous implementation details are adequate in the context of this application.

Feasible designs are first obtained during an exploration phase. In this case, a total of 544 designs uniformly distributed over the feasible region are obtained after four TMCMC stages, which are shown in Fig. 15 by means of two-dimensional projections and marginal histograms. Note that the projections $x_1 - x_3$ and $x_2 - x_4$ show some interaction between the design variables, which is associated with the effect of the geometric constraints. In this regard, the results of the exploration phase provide an insight on the sensitivity of the constraints with respect to the design variables. Starting with this set of feasible designs, an exploitation phase is then carried out to generate a set of designs lying in a vicinity of the optimal solution set. For illustration purposes, Fig. 16 shows the minimum and maximum objective function values obtained during the different stages of the exploitation phase. Note that the difference between both values reduces

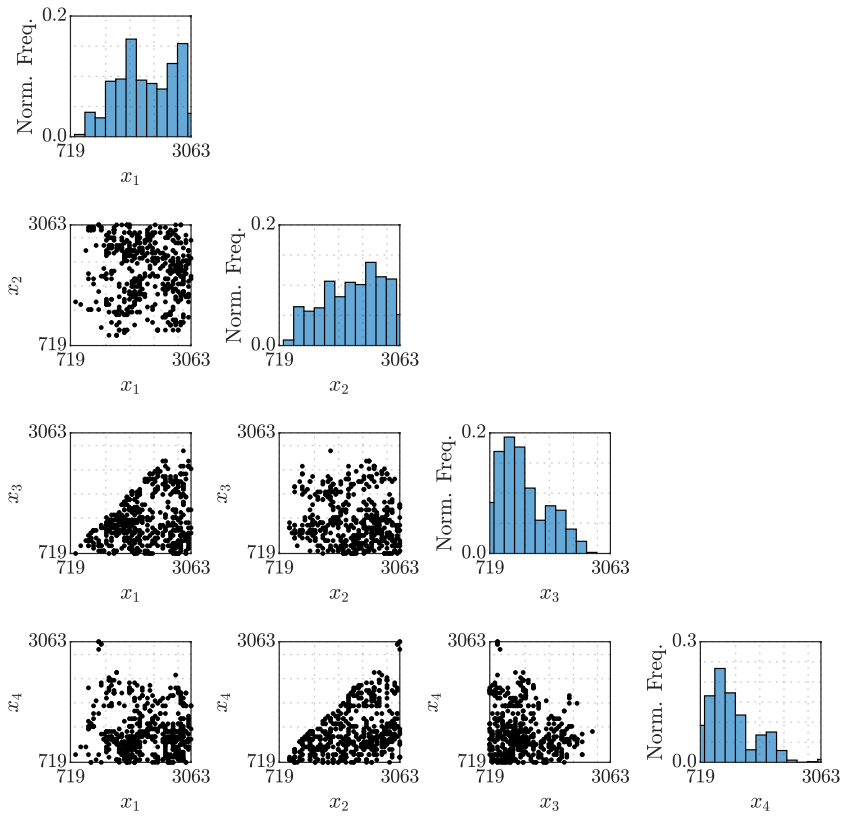

Figure 15: Two-dimensional sample projections and marginal histograms of the feasible samples. Example No. 3.

as the optimization process continues, which is consistent with the theoretical foundations of the proposed approach.

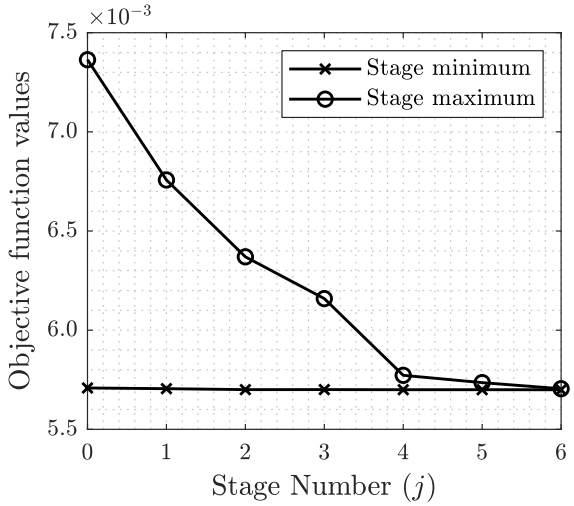

Figure 16: Maximum and minimum objective function values obtained during the different stages of the exploitation phase. Example No. 3.

Figure 17 shows the samples, in terms of marginal histograms, obtained during exploitation

stages $j = 0$, 2, 4, and 6 (final stage). It is seen that the designs are increasingly concentrated near the set that minimizes the objective function, as expected. The reduction in the RMS of the displacement at the top floor is attained by selecting higher values for x_1 and x_2, and intermediate values for x_3 and x_4. In other words, larger section sizes are preferred for the braces allocated in the lower stories, which is reasonable from the structural viewpoint. This provides valuable information on the sensitivity of the objective function with respect to the design variables. Furthermore, the last stage of the optimization procedure provides several designs with very similar objective function values (see Fig. 16) which, due to the variability in the estimation of this quantity, can be regarded as equivalent from the objective function viewpoint. Therefore, the final design can be selected by considering alternative criteria. This highlights one of the advantages of the proposed optimization framework, that is, a set of candidate designs is obtained instead of a single solution, which provides additional flexibility for the overall decision-making process.

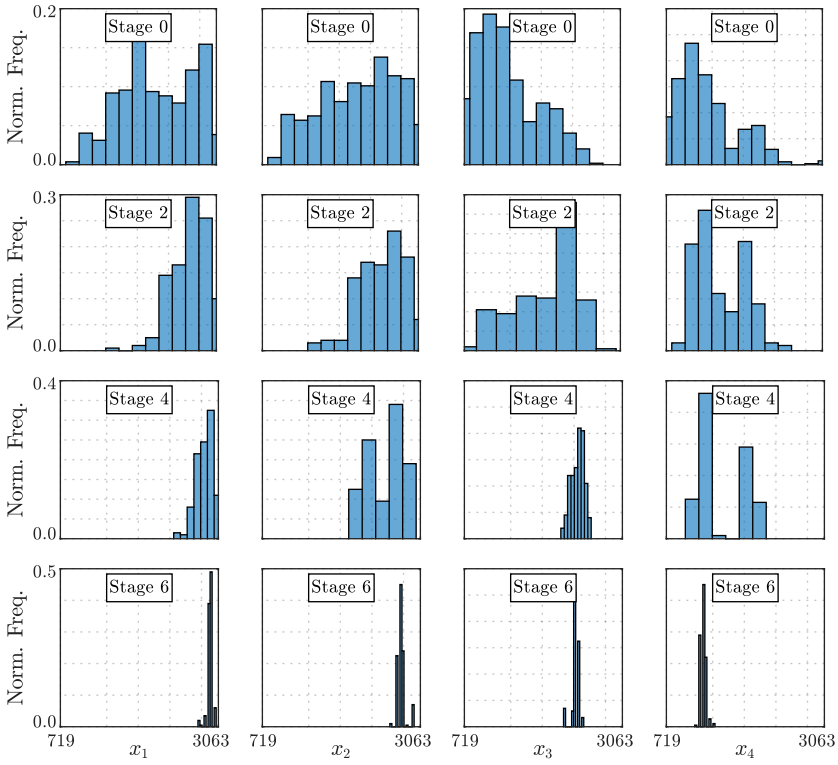

Figure 17: Marginal histograms obtained during different stages of the exploitation phase. Example No. 3.

For reference and comparison purposes, Table 8 shows the best solution obtained by the proposed approach. In addition, the optimum design obtained by means of Genetic Algorithms (GA) [7], which considers a population size of 200 individuals is also reported in the table. It is seen

that both solutions are qualitatively similar and, taking into account the variability in the estimation of the objective function value, they can be regarded as equivalent from the optimization viewpoint. However, it is observed that the cost constraint is slightly violated in the GA-based solution. Finally, in terms of numerical efforts, the total number of function calls (expected value of the root-mean-square of the displacement response at the top floor) is in the order of 2000 for both cases.

Design variables	Optimal values	
	ABO	GA
x_1	3063	2963
x_2	2714	2864
x_3	2265	2116
x_4	1318	1368
$E_{\boldsymbol{\theta}}[RMS_d(\mathbf{x}, \boldsymbol{\theta})]$	5.70×10^{-3}	5.70×10^{-3}
$g_1(\mathbf{x})$	-5.8×10^{-4}	8.3×10^{-4}
$g_2(\mathbf{x})$	-2.61	-0.37
$g_3(\mathbf{x})$	-0.51	-0.68
Function calls	2000	2138

Table 8: Best solutions reported by the proposed method (ABO) and Genetic Algorithms (GA). Example No. 3.

The results presented in this section and additional validation calculations illustrate that the proposed optimization framework is an effective tool to handle complex constrained optimization problems, such as those involving performance-based measures, nonlinear structural systems under stochastic excitation, and discrete design variables. Thus, the proposed approach is a competitive, general and flexible choice for dealing with a general class of constrained design optimization problems.

9. Conclusions

A general Markov sampling-based framework for solving a class of constrained design optimization problems has been presented. The design problem is reformulated as the equivalent task of obtaining samples uniformly distributed over the optimum solution set. To generate such samples, a sequence of distributions increasingly concentrated around the optimum solution set

is introduced. Furthermore, a unified two-phase strategy is developed. An exploration phase generates an initial set of designs uniformly distributed over the feasible domain, and then an exploitation phase generates designs increasingly concentrated in a vicinity of the optimum solution set. The transitional Markov chain Monte Carlo method is adapted to generate the required samples during both phases, and appropriate adaptive proposal distributions are implemented for the continuous and discrete design variables. The capabilities of the proposed approach, which is quite general, have been demonstrated in different types of representative design optimization problems. Numerical results have shown some of the advantages and benefits of the proposed Asymptotic Bayesian Optimization scheme.

- First, the proposed method can handle complex feasible design spaces. In fact, the exploration phase is able to deal with linear and nonlinear constraint functions, and feasible designs can be efficiently generated in cases involving non-trivial geometries for the feasible design space such as disconnected regions.

- Second, the exploitation phase successfully generates designs in a vicinity of the optimum solution set for relatively complex objective functions. Moreover, the method can also handle problems involving single and multiple optima. In general, few stages are required to identify a region that lies close to the optimum solution set.

- Third, the same framework can be used for continuous, discrete, or mixed discrete-continuous design variables. In this context, suitable adaptive proposal distributions for the continuous and discrete design variables are suggested.

- Fourth, the approach provides benefits from the practical viewpoint. In this regard, relatively few user-defined parameters are involved in the proposed approach. Numerical results indicate that sample sizes in the range of 200-500 and intermediate values for the effective sample size parameter, e.g. 0.5, yield a good tradeoff between efficiency and accuracy for the design problems investigated in this work. In addition, constraints are handled directly during the sampling process. Therefore, special constraint-handling techniques are not required by the proposed optimization scheme.

- Fifth, information on the sensitivity of the constraint and objective functions with respect to the design variables is provided as a by-product of the sampling process. This feature is

particularly valuable in situations where problem functions are not known explicitly. Thus, additional insight for design purposes is obtained without any additional computational effort and, as a result, improved flexibility can be accomplished throughout the overall decision making process.

- Sixth, the same stochastic sampling technique is used to explore the feasible and optimum sets, which is advantageous from the implementation viewpoint.

- Finally, the example problems considered in this contribution and additional validation calculations indicate that the method is very competitive with respect to other state-of-the-art stochastic search techniques. Overall, the proposed Markov sampling-based framework is a valuable tool to deal with a wide range of constrained design optimization problems. In this regard, it is stressed that the proposed Asymptotic Bayesian Optimization scheme is quite general. In other words, it is not customized to a particular class of engineering design problems.

Future research efforts aim to extend the proposed framework to more complex structural optimization problems, including general performance-based design optimization problems. In this context, suitable strategies based on parametric reduced-order models can be integrated to increase the overall efficiency of the optimization procedure. The treatment of multiobjective optimization problems by means of the proposed approach is an additional subject for future work. Another research direction involves the implementation and assessment of alternative sampling schemes within the two-phase framework. Some of these topics are currently under consideration.

Acknowledgments

The research reported here was supported in part by ANID (National Agency for Research and Development, Chile) under grant number 1200087. Also, this research has been supported by ANID and DAAD (German Academic Exchange Service) under CONICYT-PFCHA/Doctorado Acuerdo Bilateral DAAD Becas Chile/2018-62180007. These supports are gratefully acknowledged by the authors.

10. Appendix A

The lead sample in terms of the continuous and discrete design variables, that is, $\tilde{\mathbf{x}}^{j+1} = \langle \tilde{\mathbf{x}}_c^{j+1^T}, \tilde{\mathbf{x}}_d^{j+1^T} \rangle^T$, is a sample from stage j that is selected according to a probability equal to the normalized importance weight $\bar{w}_i^j$ of the samples $\mathbf{x}_i^j, i = 1, \ldots, n$, that is,

$$\bar{w}_i^j = \frac{w_i^j}{\sum_{l=1}^n w_l^j}, \quad w_i^j = \frac{f_{K_{j+1}}(\mathbf{x}_i^j)}{f_{K_j}(\mathbf{x}_i^j)}, \quad i = 1, \ldots, n \tag{21}$$

where $w_i^j, i = 1, \ldots, n$ are the importance weights given by

$$w_i^j = \exp\left(-h(\mathbf{x}_i^j)\left[\frac{1}{K_{j+1}} - \frac{1}{K_j}\right]\right) \tag{22}$$

for the exploration phase, and

$$w_i^j = \exp\left(-c(\mathbf{x}_i^j)\left[\frac{1}{K_{j+1}} - \frac{1}{K_j}\right]\right) \tag{23}$$

for the exploitation phase. If a sample $\mathbf{x}_i^j$ has been already drawn, then the last sample of its corresponding Markov chain is selected as the lead sample.

11. Appendix B

A symmetric local proposal distribution p_c is considered for the continuous design variables. The proposal is taken as a Gaussian distribution centered at the lead sample, say $\tilde{\mathbf{x}}_c^{j+1}$, whose covariance matrix $\mathbf{\Sigma}_j$ is equal to a scaled version of the estimate covariance matrix of the continuous design variables following the intermediate distribution $f_{K_j}(\mathbf{x})$. That is,

$$\mathbf{\Sigma}_j = \beta^2 \sum_{i=1}^n \bar{w}_i^j \left(\mathbf{x}_{ci}^j - \bar{\mathbf{x}}_c^j\right)\left(\mathbf{x}_{ci}^j - \bar{\mathbf{x}}_c^j\right)^T \tag{24}$$

where β is a scaling parameter, $\mathbf{x}_{ci}^j, i = 1, \ldots, n$ are the samples at stage j, $\bar{w}_i^j, i = 1, \ldots, n$ are the normalized importance weights, and $\bar{\mathbf{x}}_c^j = \sum_{i=1}^n \bar{w}_i^j \mathbf{x}_{ci}^j$ is the estimate mean of the continuous design variables following $f_{K_{j+1}}(\mathbf{x})$. The scaling parameter β can be defined directly by the user, or it can be determined by an adaptive scheme that monitors the acceptance rate in the context of the Metropolis-Hastings algorithm [27, 40]. It is noted that alternative sampling schemes in the so-called underlying normal space [27] can be also considered in the context of the ABO method.

12. Appendix C

The proposal distribution p_d for the vector of discrete design variables $\mathbf{x}_d$ is defined in terms of independent proposal distributions corresponding to each discrete variable, that is,

$$p_d(\mathbf{x}_d^\star|\tilde{\mathbf{x}}_d^{j+1}) = \prod_{l=1}^{n_d} p_{dl}(x_{dl}^\star|\tilde{x}_{dl}^{j+1}) \tag{25}$$

where $\mathbf{x}_d^\star = \langle x_{d1}^\star, \ldots, x_{dn_d}^\star \rangle^T$ is the candidate sample, $\tilde{\mathbf{x}}_d^{j+1} = \langle \tilde{x}_{d1}^{j+1}, \ldots, \tilde{x}_{dn_d}^{j+1} \rangle^T$ is the lead sample of the discrete design variables, and $p_{dl}(x_{dl}^\star|\tilde{x}_{dl}^{j+1})$ is the local proposal distribution for the discrete design variable $x_{dl}^\star$. The candidate sample $x_{dl}^\star$ is selected from the set of discrete values adjacent to $\tilde{x}_{dl}^{j+1}$, i.e., $Adj(\tilde{x}_{dl}^{j+1})$, which is defined as

$$Adj(\tilde{x}_{dl}^{j+1}) = \{x_{dl(i)_{i=1,\ldots,n_{di}}} : \lambda(\tilde{x}_{dl}^{j+1}, x_{dl(i)}) \leq \lambda_l^*\} \tag{26}$$

where $\lambda(\tilde{x}_{dl}^{j+1}, x_{dl(i)})$ is the *distance* between the lead sample $\tilde{x}_{dl}^{j+1}$ and the sample $x_{dl(i)}$. For example, if the lead sample corresponds to the s^{th} available value of the discrete variable x_{dl}, that is, $\tilde{x}_{dl}^{j+1} = x_{dl(s)}$, the distance between these two samples is given by $\lambda(\tilde{x}_{dl}^{j+1}, x_{dl(i)}) = \lambda(x_{dl(s)}, x_{dl(i)}) = |s - i|$. The proposal distribution for the l^{th} discrete design variable is defined as [29]

$$p_{dl}(x_{dl}^\star|\tilde{x}_{dl}^{j+1}) = \begin{cases} \frac{1-\tau}{Card(Adj(\tilde{x}_{dl}^{j+1}))} & \text{if } x_{dl}^\star \in Adj(\tilde{x}_{dl}^{j+1}) \\ \frac{\tau}{Card(Adj^c(\tilde{x}_{dl}^{j+1}))} & \text{if } x_{dl}^\star \notin Adj(\tilde{x}_{dl}^{j+1}) \end{cases} \tag{27}$$

where τ represents a small probability of randomly selecting a nonadjacent value of the lead sample $\tilde{x}_{dl}^{j+1}$, $Adj^c(\tilde{x}_{dl}^{j+1})$ is the complement set of $Adj(\tilde{x}_{dl}^{j+1})$, and $Card(\cdot)$ is the number of discrete available values in the corresponding set.

It is noted that the values for the distribution parameters λ_l^* and τ are problem-dependent. Such parameters can be selected directly by the user or adaptively tuned during the sampling process. In this contribution, an adaptive strategy is implemented. At the beginning of each exploitation stage, the parameter λ_l^* is updated by reusing information obtained during the previous stage. The procedure is carried out as follows. First, the elements of $\mathbf{X}_{dl}$ that were observed during the previous stage are identified. Then, the *maximum number of consecutive values* observed during the previous stage, denoted by η, is obtained. Based on this number, the parameter λ_l^* is updated as

$$\lambda_l^* = \min\{\lambda_l^*, m\} \tag{28}$$

where m is the largest integer such that $m \leq (\eta - 1)/2$. The process is repeated for $l = 1, \ldots, n_d$. This scheme tends to decrease the value of λ_l^* for advanced exploitation stages, which can improve the efficiency of the proposed approach. Certainly, alternative strategies can be implemented as well.

13. Appendix D

The implementation of the acceptance/rejection test, in the context of the Metropolis-Hastings algorithm, is as follows. The candidate sample $\mathbf{x}^\star = \langle \mathbf{x}_c^{\star T}, \mathbf{x}_d^{\star T} \rangle^T$, is accepted with probability $\rho^\star$, where

$$\rho^\star = \text{Min}\left\{1, I_{\mathbf{X}}(\mathbf{x}^\star) \frac{\exp\left(-h(\mathbf{x}^\star)/K_{j+1}\right)}{\exp\left(-h(\tilde{\mathbf{x}}^{j+1})/K_{j+1}\right)} \frac{p_d(\tilde{\mathbf{x}}_d^{j+1}|\mathbf{x}_d^\star)}{p_d(\mathbf{x}_d^\star|\tilde{\mathbf{x}}_d^{j+1})}\right\} \tag{29}$$

for the exploration phase, where $I_{\mathbf{X}}(\mathbf{x}^\star) = 1$ if $\mathbf{x}^\star \in \mathbf{X}$ and $I_{\mathbf{X}}(\mathbf{x}^\star) = 0$ otherwise, and

$$\rho^\star = \text{Min}\left\{1, I_{\mathbf{X}_{\text{feasible}}}(\mathbf{x}^\star) \frac{\exp\left(-c(\mathbf{x}^\star)/K_{j+1}\right)}{\exp\left(-c(\tilde{\mathbf{x}}^{j+1})/K_{j+1}\right)} \frac{p_d(\tilde{\mathbf{x}}_d^{j+1}|\mathbf{x}_d^\star)}{p_d(\mathbf{x}_d^\star|\tilde{\mathbf{x}}_d^{j+1})}\right\} \tag{30}$$

for the exploitation phase. According to the definition of the proposal distribution p_d, the ratio $p_d(\tilde{\mathbf{x}}_d^{j+1}|\mathbf{x}_d^\star)/p_d(\mathbf{x}_d^\star|\tilde{\mathbf{x}}_d^{j+1})$ is given by

$$\frac{p_d(\tilde{\mathbf{x}}_d^{j+1}|\mathbf{x}_d^\star)}{p_d(\mathbf{x}_d^\star|\tilde{\mathbf{x}}_d^{j+1})} = \prod_{l=1}^{n_d} \begin{cases} \frac{Card(Adj(\tilde{x}_{dl}^{j+1}))}{Card(Adj(x_{dl}^\star))} & \text{if } x_{dl}^\star \in Adj(\tilde{x}_{dl}^{j+1}) \\ \frac{Card(Adj^c(\tilde{x}_{dl}^{j+1}))}{Card(Adj^c(x_{dl}^\star))} & \text{if } x_{dl}^\star \notin Adj(\tilde{x}_{dl}^{j+1}) \end{cases} \tag{31}$$

If the candidate state $\mathbf{x}^\star$ is rejected, the lead sample $\tilde{\mathbf{x}}^{j+1}$ is repeated. More information about the Metropolis-Hastings algorithm can be found in [35, 36].

14. References

References

[1] R. T. Haftka and Z. Gürdal. *Elements of Structural Optimization, Third edition.* Springer Netherlands, 1992.

[2] J. S. Arora. *Introduction to optimum design, Fourth edition.* Elsevier, 2017.

[3] H.A. Jensen and J. Sepulveda. Structural optimization of uncertain dynamical structural systems considering mixed-design variables. *Probabilistic Engineering Mechanics*, 26(2):269-280, 2011.

[4] M. W. Huang and J. S. Aurora. Optimal design with discrete variables: some numerical experiments. *International Journal for Numerical Methods in Engineering*, 40(1):165–188, 1997.

[5] J.A. Tomlin. *Branch-and-Bound Methods for Integer and Non-convex Programming.* J. Abadie (ed.), Elsevier Publishing Co., New York, 437-450, 1970.

[6] A. D. Belegundu and T. R. Chandrupatla. *Optimization Concepts and Applications in Engineering, Second edition.* Cambridge University Press, New York, USA, 2011.

[7] T. H. Holland. *Adaptation in Natural and Artificial Systems.* MIT Press, Cambridge, MA, USA, 1992.

[8] H.-G. Beyer. *The Theory of Evolution Strategies.* Springer Berlin Heidelberg, 2001.

[9] W. Banzhaf, P. Nordin, and R. E. Keller. *Genetic Programming: An Introduction.* Morgan Kaufmann Publ. Inc., 1997.

[10] S. Kirkpatrick, C. D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4589):165–172, 1983.

[11] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi. GSA: A Gravitational Search Algorithm. *Information Sciences*, 179(13):2232–2248, 2009.

[12] S.-K. Au and J.L. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263-277, 2001.

[13] H.-S. Li and S.-K. Au. Design optimization using Subset Simulation algorithm. *Structural Safety*, 32(6):384–392, 2010.

[14] A. Kaveh and M. Khayatazad. A new meta-heuristic method: Ray Optimization. *Computers and Structures*, 112-113:283–294, 2012.

[15] J. Kennedy and R. Eberhart. Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4:1942–1948, 1995.

[16] M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):29–41, 1996.

[17] Z. W. Geem, J. H. Kim, and G. V. Loganathan. A new heuristic optimization algorithm: Harmony Search. *Simulation*, 76(2):60–68, 2001.

[18] D. Karaboga. *An idea based on honey bee swarm for numerical optimization.* Computer Engineering Department, Erciyes University, Turkey, Technical report TR06, 2005.

[19] C. A. Coello Coello. Use of a self-adaptive penalty approach for engineering optimization problems. *Computers in Industry*, 41(2):113–127, 2000.

[20] Y. Dong, J. Tang, B. Xu, and D. Wang. An application of swarm optimization to nonlinear programming. *Computers & Mathematics with Applications*, 49(11-12):1655–1668, 2005.

[21] E. Mezura-Montes and C. A. Coello Coello. Use of multi-objective optimization concepts to handle constraints in genetic algorithms. *In: Abraham A, Jain L, Goldberg R (eds.). Evolutionary multi-objective optimization.* Springer, Berlin, 2005.

[22] Z. Michalewicz. Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary Computation*, 4(1):1—32, 1996.

[23] A. Hedar and M. Fukushima. Derivative-free filter simulated annealing method for constrained continuous global optimization. *Journal of Global Optimization*, 35(4):521-–549, 2006.

[24] L. J. Li, Z. Huang, F. Liu, and Q. H. Wu. A heuristic particle swarm optimizer for optimization of pin connected structures. *Computers and Structures*, 85(7–8):340-–349, 2007.

[25] Q. He and L. Wang. A hybrid particle swarm optimization with a feasibility-based rule for constrained optimization. *Applied Mathematics and Computation*, 186(2):1407-–1422, 2007.

[26] J. Ching and Y.C. Chen. Transitional Markov chain Monte Carlo method for Bayesian updating, model class selection, and model averaging. *Journal of Engineering Mechanics*, 133:816–832, 2007.

[27] W. Betz, I. Papaioannou, and D. Straub. Transitional Markov chain Monte Carlo: observations and improvements. *Journal of Engineering Mechanics*, 142(5):04016016, 2016.

[28] H. Jensen, D. Jerez, and M. Beer. A general two-phase Markov chain Monte Carlo approach for constrained design optimization: application to stochastic structural optimization. *Computer Methods in Applied Mechanics and Engineering*, 373:113487, 2021.

[29] H. Jensen, D. Jerez, and M. Beer. Structural synthesis considering mixed discrete-continuous design variables: a Bayesian framework. *Mechanical Systems and Signal Processing*, 162:108042, 2022.

[30] K. M. Zuev and J. L. Beck. Global optimization using the asymptotically independent Markov sampling method. *Computers and Structures*, 126:107–119, 2013.

[31] J. Wang and L. S. Katafygiotis. Reliability-based optimal design of linear structures subjected to stochastic excitations. *Structural Safety*, 47:29–38, 2014.

[32] V. Černý. Thermodynamical approach to the traveling salesman problem: An efficient sim-

ulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985.

[33] C. Robert and G. Casella. *Monte Carlo Statistical Methods.* Springer New York, 2010.

[34] J. S. Liu. Metropolized independent sampling with comparison to rejection sampling and importance sampling. *Statistics and Computing*, 6(2):113–119, 1996.

[35] N. Metropolis, A. Resenbluth, M. Resenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[36] W. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[37] J. L. Beck and K. M. Zuev. Asymptotically independent Markov sampling: a new Markov chain Monte Carlo scheme for Bayesian inference. *International Journal for Uncertainty Quantification*, 3(5):445–474, 2013.

[38] H. A. Jensen, D. J. Jerez, and M. Valdebenito. An adaptive scheme for reliability-based global design optimization: a Markov chain Monte Carlo approach. *Mechanical Systems and Signal Processing*, 143:106836, 2020.

[39] A. Kong, J. S. Liu and W. H. Wong. Sequential imputations and Bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994.

[40] H. A. Jensen, C. Esse, V. Araya, and C. Papadimitriou. Implementation of an adaptive meta-model for Bayesian finite element model updating in time domain. *Reliability Engineering & System Safety*, 160:174–190, 2017.

[41] J. Golinsky. An adaptive optimization system applied to machine synthesis. *Mechanism and Machine Theory*, 8:419–436, 1973.

[42] T. Ray and K. M. Liew. Society and civilization: an optimization algorithm based on the simulation of social behavior. *IEEE Transactions on Evolutionary Computation*, 7(4):386–396, 2003.

[43] H. Liu, Z. Cai, and Y. Wang. Hybridizing particle swarm optimization with differential evolution for constrained numerical and engineering optimization. *Applied Soft Computing*, 10(2):629–640, 2010.

[44] L. Wang and L.-P. Li. An effective differential evolution with level comparison for constrained engineering design. *Structural and Multidisciplinary Optimization*, 41(6):947–963, 2009.

[45] A. Sadollah, A. Bahreininejad, H. Eskandar, and M. Hamdi. Mine blast algorithm: a new

population based algorithm for solving constrained engineering optimization problems. *Applied Soft Computing*, 13(5):2592–2612, 2013.

[46] American Institute of Steel Construction. *AISC manual of steel construction: load and resistance factor design (3rd ed.)*, 2001.

[47] D. M. Moore. Simulation of ground motion using the stochastic method. *Pure and Applied Geophysics*, 160:635–676, 2003.

[48] G. M. Atkinson and W. Silva. Stochastic modeling of California ground motions. *Bulletin of the Seismological Society of America*, 90(2):255–274, 2000.

[49] J. G. Anderson and S. E. Hough. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America*, 74(5):1969–1993, 1984.

[50] G. P. Mavroeidis. A mathematical representation of near-fault ground motions. *Bulletin of the Seismological Society of America*, 93(3):1099–1131, 2003.

[51] H. J. Pradlwarter, G. I. Schuëller and U. Dorka. Reliability of MDOF-systems with hysteretic devices. *Engineering Structures*, 20(8):685–691, 1998.