

# STOCHASTIC DESIGN AND CONTROL IN RANDOM HETEROGENEOUS MATERIALS \*

RAPHAEL STERNFELS <sup>†</sup> AND PHAEDON-STELIOS KOUTSOURELAKIS <sup>‡</sup>

**Abstract.** The present paper discusses a sampling framework that enables the optimization of complex systems characterized by high-dimensional uncertainties and design variables. We are especially concerned with problems relating to random heterogeneous materials where uncertainties arise from the stochastic variability of their properties. In particular, we reformulate topology optimization problems to account for the design of truly random composites. In addition, we address the optimal prescription of input loads/excitations in order to achieve a target response by the random material system. The methodological advances proposed in this paper consist of an adaptive Sequential Monte Carlo scheme that economizes the number of runs of the forward solver and allows the analyst to identify several local maxima that provide important information with regards to the robustness of the design. We further propose a principled manner of introducing information from approximate models that can ultimately lead to further reductions in computational cost.

**Key words.**

**AMS subject classifications.**

**1. Introduction.** The analysis of materials which exhibit very small length scales of heterogeneity has attracted considerable attention in recent years [44]. This is because fine details in the microstructure can give rise to marked differences in the macroscale response. In reality, the majority of such materials exhibit randomness as local physical and mechanical properties fluctuate stochastically. In multiphase materials for example the distribution of the constituent phases in space does not follow a particular pattern and is characterized by disorder. It is therefore obvious that a probabilistic description is most appropriate and provides a sounder basis for their characterization and quantification of the performance of the systems where these appear.

While marked advances have been achieved in the context of modeling [48, 20, 18, 29] and uncertainty propagation [42, 6, 47, 45], the problem of design/optimization in the presence of randomness has received much less attention [28, 49]. We address two problems in the context of random heterogeneous materials. On one hand we examine the problem of designing of random heterogeneous materials. This can be seen as the fully stochastic counterpart of topology optimization, a deterministic tool for the systematic design of composite microstructures with desirable macroscopic properties [33, 32, 34, 39, 35, 40, 41, 2, 37, 36, 12, 14]. Rather than fully specifying the spatial distribution of the constitutive phases, we are interested in controlling statistics of the associated probability distribution (i.e. volume fractions, spatial correlations), while the resulting microstructure remains random. Such a capability could prove particularly useful in the fabrication of heterogeneous materials. For example in polycrystalline materials, it is known that macro-scale forming parameters such as forging rates, die shapes and preform shapes or heat treatment, do not uniquely define the final polycrystalline texture but rather its statistical features [43]. In naturally

---

\*This work was supported by the OSD/AFOSR MURI'09 award to Cornell University on uncertainty quantification

<sup>†</sup>School of Civil and Environmental Engineering, 372 Hollister Hall, Cornell University, Ithaca, NY 14853, (hrs63@cornell.edu)

<sup>‡</sup>School of Civil and Environmental Engineering & Center for Applied Mathematics, 369 Hollister Hall, Cornell University, Ithaca, NY 14853, (pk285@cornell.edu).

occurring random heterogeneous materials such as soils, remediation procedures such as solidification or stabilization allow us to alter the microstructure and its properties by altering the probabilistic characteristics of the existing phases and introducing randomly dispersed new materials.

The second problem of interest involves the optimization of the input in the presence of uncertainties. In particular we consider a random heterogeneous material, fully defined by the probability distribution of its properties and seek to identify the input such that the *expected* response is as close as possible to a target, desired output [49]. Both problems are examined in the context of heat conduction but can be readily extended to other physics as the methodology advocated makes non-intrusive use of forward solvers (e.g. Finite Elements).

The present papers advocates a simulation strategy that recasts the optimization problem as a sampling problem in the expanded space which apart from the random variables includes the design/control parameters [24]. High probability regions of the auxiliary target density correspond to maxima of the design objectives. To that end we employ adaptive *Sequential Monte Carlo* (SMC) methods that are well-suited in high dimensions, are directly parallelizable and capable of identifying multimodal densities which correspond to local maxima [10, 21]. In addition, they give rise to a population of estimates that provide valuable information with regards to the robustness of the optima. Despite their advantages, SMC-based samplers require multiple calls to deterministic simulators which might be impractical or infeasible for highly complex, nonlinear solvers. To that end we propose an adaptive sampling framework which aims at reducing the number of calls to the forward solver in order to attain a certain level of accuracy. Furthermore, we propose a hierarchical strategy where *approximate* forward models can be rigorously incorporated in order to give accurate estimates at a reduced computational cost. Such approximate solvers can be derived by using principled reduced-order approximations or simply by coarsening the discretization (e.g. in Finite Element analysis).

**2. Proposed Approach.** We consider systems characterized by a vector of uncertainties denoted by  $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^{n_\theta}$  distributed according to the density  $p(\boldsymbol{\theta})$ . This vector will characterize the spatial variability of material properties (e.g. conductivity at various locations) and in general its dimension will be large (i.e.  $n_\theta \gg 1$ ). We denote by  $\mathbf{d} \in \mathcal{D} \subset \mathbb{R}^{N_d}$  the vector of design/control variables. In the problems examined in this paper, these represent the input or the statistics of the distribution of  $\boldsymbol{\theta}$ , i.e.  $p(\boldsymbol{\theta} | \mathbf{d})$ . Our goal is to find the values of the design variables  $\mathbf{d}$  in the feasible domain  $\mathcal{D}$ , that maximize the *expected utility*  $\hat{U}(\mathbf{d})$ :

$$\hat{U}(\mathbf{d}) = \int_{\Theta} U(\boldsymbol{\theta}, \mathbf{d}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (2.1)$$

The utility function  $U(\boldsymbol{\theta}, \mathbf{d})$  depends on the output of interest (and is therefore a function of the system uncertainties and design variables) and in the present framework it suffices that it is non-negative  $\forall \boldsymbol{\theta}, \mathbf{d}$ <sup>1</sup>. The formulation above is quite general and can readily be adapted to cases of practical interest. For example if  $U(\boldsymbol{\theta}, \mathbf{d}) = \mathbf{1}_{\mathcal{A}}(\boldsymbol{\theta}, \mathbf{d})$  is the indicator function of an event  $\mathcal{A}$  of interest (e.g. failure, or exceedance of a response threshold) then maximizing  $\hat{U}(\mathbf{d})$  in Equation (2.1) is equivalent to the maximization of the probability associated with the event  $\mathcal{A}$  (similarly one can minimize

---

<sup>1</sup>even if  $U$  is negative, it suffices that  $U(\boldsymbol{\theta}, \mathbf{d}) \geq U_0 > -\infty$  in which case one can utilize  $U(\boldsymbol{\theta}, \mathbf{d}) - U_0 \geq 0$

the the probability of event  $\mathcal{A}$  by employing the indicator function of the complementary even  $\mathcal{A}^c$  in place of  $U$  in Equation (2.1)). Furthermore if  $\mathbf{u}(\boldsymbol{\theta}, \mathbf{d})$  denotes the output vector (i.e. displacements, stresses, temperatures etc) and  $\mathbf{u}_{target}$  a target/desired response, then for  $U(\boldsymbol{\theta}, \mathbf{d}) = \exp\{-\|\mathbf{u}(\boldsymbol{\theta}, \mathbf{d}) - \mathbf{u}_{target}\|\}$  in Equation (2.1), implies finding  $\mathbf{d}$  for which the response is, on average, as close to the target.

The systems of interest are considered complex in the sense that the utility function  $U(\boldsymbol{\theta}, \mathbf{d})$  is not known explicitly, and can only be computed for a particular  $(\boldsymbol{\theta}, \mathbf{d})$  with a call to , a potentially costly, deterministic solver, e.g. a finite element solver. Thus the number of calls to such a solver will dominate the total amount of computational work. It is clear that a brute force application of deterministic optimization procedures directly on  $\hat{U}(\mathbf{d})$  is impractical, as each evaluation of this function (and potentially its derivatives) will in general require a high-dimensional integration with respect to the uncertainties  $\boldsymbol{\theta}$ . Furthermore, discretizing the design space  $\mathcal{D}$  is inefficient or infeasible when the dimension  $n_d$  of  $\mathbf{d}$  is large.

For these reasons we advocate a *sampling* strategy, first proposed in [24]. This entails defining an (unnormalized) probability density  $\pi(\boldsymbol{\theta}, \mathbf{d})$  defined on the joint space  $\Theta \times \mathcal{D}$ :

$$\pi(\boldsymbol{\theta}, \mathbf{d}) \propto U(\boldsymbol{\theta}, \mathbf{d}) p(\boldsymbol{\theta}) 1_{\mathcal{D}}(\mathbf{d}) \quad (2.2)$$

where  $1_{\mathcal{D}}(\mathbf{d})$  is the indicator function of the feasible domain  $\mathcal{D}$ . The marginal  $\pi(\mathbf{d}) = \int \pi(\boldsymbol{\theta}, \mathbf{d}) d\boldsymbol{\theta}$  is clearly proportional to  $\hat{U}(\mathbf{d})$  and as a result samples drawn from the joint density  $\pi(\boldsymbol{\theta}, \mathbf{d})$  will be marginally distributed according to  $\hat{U}(\mathbf{d})$  and populate regions where this attains its maximum value(s). It is also important to point out that such an approach does not lead to point estimates for the maxima of the expected utility but also provides information about the variability of the latter with respect to  $\mathbf{d}$  and therefore the *robustness* of the select design  $\mathbf{d}$  [30, 38]. If the global maximum of  $\hat{U}(\mathbf{d})$  is desired, then this can be achieved within the same framework by artificially expanding the sampling space [16]. In particular, one employs the (unnormalized) density  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$ , where  $\boldsymbol{\theta}_{1:n} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_n)$ , which is defined on  $\underbrace{\Theta \times \dots \times \Theta}_{n \text{ times}} \times \mathcal{D}$

[24]:

$$\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d}) \propto \prod_{j=1}^n U(\boldsymbol{\theta}_j, \mathbf{d}) p(\boldsymbol{\theta}_j) 1_{\mathcal{D}}(\mathbf{d}) \quad (2.3)$$

It is clear again that the marginal with respect to  $\mathbf{d}$ , i.e.  $\pi(\mathbf{d}) = \int \pi(\boldsymbol{\theta}_{1:n}, \mathbf{d}) d\boldsymbol{\theta}_{1:n}$  is proportional to  $\hat{U}^n(\mathbf{d})$ . As a result, the  $\mathbf{d}$ -coordinates of samples drawn from  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$  will be more tightly concentrated around the maxima of  $\hat{U}(\mathbf{d})$ , increasingly so with  $n$ .

Despite its flexibility, such an approach requires sampling in the *joint* space of random and design variables. Its dimension is even higher when the augmented density of Equation (2.3) is employed. While Monte Carlo strategies offer the most general method for carrying out the sampling task, a naive implementation would be impractical as it would require a large number of evaluations of the utility function  $U(\boldsymbol{\theta}, \mathbf{d})$  and therefore a lot of calls to the forward solver. Furthermore, it might fail to identify all the modes of the distribution  $\pi(\boldsymbol{\theta}, \mathbf{d})$  in Equation (2.2) (or  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$  in Equation (2.3)) which correspond to local maxima of the expected utility  $\hat{U}(\mathbf{d})$ . These local maxima can be of considerable value in terms of their physical and engineering significance as they also reveal valuable features with regards to the sensitivity of the expected output to the design/control variables. Traditionally *Markov Chain Monte*

*Carlo* techniques (MCMC) have been employed which are based on building a Markov chain that asymptotically converges to the target density (in this case  $\pi$ ) by appropriately defining a transition kernel. While convergence can be assured under weak conditions [22, 25], the rate of convergence can be extremely slow and require a lot of utility function evaluations. Particularly in cases where the target density has multiple modes, very large *mixing times* might be required. For that purpose we advocate the use of Sequential Monte Carlo (SMC) procedures which have the capability of sampling from multi-modal distributions in high-dimensional spaces and discuss an *adaptive* scheme for reducing the computational cost. It is noted that SMC strategies have been employed in this framework and in the context of Bayesian optimal design in [1, 21] and for maximum likelihood estimation in latent variable models in [16].

**2.1. Adaptive Sequential Monte Carlo.** SMC strategies represent a set of flexible simulation-based methods for sampling from a *sequence* of probability distributions [23, 11]. As with Markov Chain Monte Carlo methods (MCMC), the target distribution(s) need only be known up to a constant as is the case in Equation (2.2) and Equation (2.3). They utilize a set of random samples (commonly referred to as *particles*), which are propagated using a combination of *importance sampling*, *resampling* and MCMC-based *rejuvenation* mechanisms [10, 9]. Each of these particles is associated with an *importance weight*. These weights are updated sequentially along with the particle locations. Hence if  $\{(\boldsymbol{\theta}^{(i)}, \mathbf{d}^{(i)}), w^{(i)}\}_{i=1}^N$  represent  $N$  such particles and associated weights for distribution  $\pi(\boldsymbol{\theta}, \mathbf{d})$  in Equation (2.2) then:

$$\pi(\boldsymbol{\theta}, \mathbf{d}) \approx \sum_{i=1}^N W^{(i)} \delta_{\boldsymbol{\theta}^{(i)}}(\boldsymbol{\theta}) \delta_{\mathbf{d}^{(i)}}(\mathbf{d}) \quad (2.4)$$

where  $W^{(i)} = w^{(i)} / \sum_{k=1}^N w^{(k)}$  are the normalized weights and  $\delta_{\mathbf{x}}(\cdot)$  is the Dirac function centered at  $\mathbf{x}$ . Furthermore, for any function  $h(\boldsymbol{\theta}, \mathbf{d})$  which is  $\pi$ -integrable [8, 7]:

$$\sum_{i=1}^N W^{(i)} h(\boldsymbol{\theta}^{(i)}, \mathbf{d}^{(i)}) \rightarrow \int h(\boldsymbol{\theta}, \mathbf{d}) \pi(\boldsymbol{\theta}, \mathbf{d}) \, d\boldsymbol{\theta} d\mathbf{d} \quad (\text{almost surely}) \quad (2.5)$$

The underlying idea in SMC strategies is to operate on a sequence of distributions, starting from one that can be accurately and easily sampled from, and gradually changing it until the target density is reached. In that respect, it is quite similar to simulated annealing employed in optimization but more general sequences of distributions can be adopted as it will be demonstrated in subsequent sections. Initially however we consider a sequence parameterized by  $\gamma \in [0, 1]$  which plays the role of reciprocal temperature such that:

$$\pi_{\gamma}(\boldsymbol{\theta}, \mathbf{d}) \propto U^{\gamma}(\boldsymbol{\theta}, \mathbf{d}) p(\boldsymbol{\theta}) 1_{\mathcal{D}}(\mathbf{d}) \quad (2.6)$$

It is easily seen that for  $\gamma = 0$  one recovers the distribution of  $p(\boldsymbol{\theta}) 1_{\mathcal{D}}(\mathbf{d})$  i.e. the random variables  $\boldsymbol{\theta}$  is distributed according to their density  $p(\boldsymbol{\theta})$  and the design variables  $\mathbf{d}$  uniformly in the feasible domain  $\mathcal{D}$ . It is implicitly assumed that generating samples from these distributions is tractable. For  $\gamma = 1$ , one recovers the target density of Equation (2.2). Starting with a particulate approximation for  $\pi_{\gamma=0}$  (which trivially involves drawing samples from the  $p(\boldsymbol{\theta})$  for  $\boldsymbol{\theta}$  and the uniform in  $\mathcal{D}$  for  $\mathbf{d}$  with weights  $w^{(i)} = 1$ ), the goal is to gradually update the importance weights and particle locations in order to approximate the target density. The role of these auxiliary

distributions  $\pi_\gamma$  for  $\gamma \in (0, 1)$  is to *bridge the gap between*  $\pi_{\gamma=0}$  and  $\pi_{\gamma=1}$  and provide a smooth transition path where importance sampling can be efficiently applied. It is easily understood that as the number of these distributions increases the accuracy increases since the transition becomes smoother, but at the same time so does the computational cost as more evaluations of the utility function would be needed. On the other hand too few intermediate distributions  $\pi_\gamma$  can adversely affect the overall accuracy of the approximation.

To that end we propose an *adaptive* SMC scheme that automatically determines the number of intermediate distributions needed [9, 19]. In this process we are guided by the Effective Sample Size (ESS, [22]). In particular, let  $S$  be the total number of intermediate distributions (which is unknown a priori) and  $\gamma_s$ ,  $s = 1, 2, \dots, S$  the associated reciprocal temperatures such that  $0 = \gamma_1 < \gamma_2 < \dots < \gamma_S = 1$ , which are also unknown a priori. Let also  $\{(\boldsymbol{\theta}_s^{(i)}, \mathbf{d}_s^{(i)}), W_s^{(i)}\}_{i=1}^N$  denote the particulate approximation of  $\pi_{\gamma_s}$  defined as in Equation (2.6) for  $\gamma = \gamma_s$ . The Effective Sample Size of these particles is then defined as  $ESS_s = 1/\sum_{i=1}^N (W_s^{(i)})^2$  and provides a measure of the population variance. One extreme, i.e. when  $ESS_s = 1$ , arises when a single particle has a unit normalized weight whereas the rest have zero weights and as a result provide no information. The other extreme, i.e.  $ESS_s = N$ , arises when all the particles are equally informative and have equal weights  $W_s^{(i)} = 1/N$ .

If the next bridging distribution  $\pi_{\gamma_{s+1}}$  is very similar to  $\pi_{\gamma_s}$  (ie.  $\gamma_{s+1} \approx \gamma_s$ ), then  $ESS_{s+1}$  should not be that much different from  $ESS_s$ . On the other hand if that difference is pronounced then  $ESS_{s+1}$  could drop dramatically. Hence in determining the next auxiliary distribution, we define an acceptable reduction in the ESS, i.e.  $ESS_{s+1} \geq \zeta ESS_s$  (where  $\zeta < 1$ ) and prescribe  $\gamma_{s+1}$  (Equation (2.6)) accordingly. The proposed adaptive SMC algorithm is summarized in Table 2.1.

It should be noted that unlike MCMC schemes, the particle perturbations in the *Rejuvenation* step do not require that the  $P_s(\cdot, \cdot)$  is *ergodic* [10]. It suffices that it is a  $\pi_{\gamma_s}$ -invariant kernel, which readily allows adaptively changing its parameters in order to achieve better mixing rates. A more detailed discussion on the kernels and adaptivity schemes used in the Rejuvenation step is deferred for section 3.

The same idea can be employed in sampling in the extended space with respect to the density  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$ ,  $n = 1, 2, \dots$  in Equation (2.3). Specifically, suppose  $\{(\boldsymbol{\theta}_{1:n-1}^{(i)}, \mathbf{d}^{(i)}), W^{(i)}\}_{i=1}^N$  represents a particulate approximation of the density  $\pi(\boldsymbol{\theta}_{1:n-1}, \mathbf{d})$  defined in Equation (2.3). In order to obtain samples from  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$  a new sequence of bridging distributions is built in the spirit of Equation (2.6) as follows:

$$\pi_{n-1, \gamma} \propto U^\gamma(\boldsymbol{\theta}_n, \mathbf{d}) p(\boldsymbol{\theta}_n) \prod_{j=1}^{n-1} U(\boldsymbol{\theta}_j, \mathbf{d}) p(\boldsymbol{\theta}_j) 1_{\mathcal{D}}(\mathbf{d}), \quad \gamma \in [0, 1] \quad (2.8)$$

It is immediately obvious that for  $\gamma = 0$  we recover  $\pi(\boldsymbol{\theta}_{1:n-1}, \mathbf{d})$  and for  $\gamma = 1$  the target  $\pi(\boldsymbol{\theta}_{1:n}, \mathbf{d})$ . The adaptive SMC scheme of Table 2.1 can be applied identically for the aforementioned sequence. An additional advantage of the proposed approach is that the state augmentation takes place sequentially which gives the analyst the opportunity to terminate the algorithm if sufficient information on the maximum (or maxima) of the expected utility  $\hat{U}(\mathbf{d})$  has been obtained.

**2.2. Using information from approximate models.** The proposed adaptive SMC sampling framework allows for great flexibility in selecting the sequence of dis-

TABLE 2.1  
Basic steps of the adaptive SMC algorithm proposed

**Adaptive SMC algorithm:**

1. **Initialization:** Set  $s = 1$  and  $\gamma_1 = 0$ . Initialize population  $\{(\boldsymbol{\theta}_1^{(i)}, \mathbf{d}_1^{(i)}), w_1^{(i)}\}_{i=1}^N$  where  $\boldsymbol{\theta}_1^{(i)}$  are i.i.d draws from  $p(\boldsymbol{\theta})$ ,  $\mathbf{d}_1^{(i)}$  are i.i.d draws from  $1_{\mathcal{D}}(\mathbf{d})$  and  $w_1^{(i)} = 1$  ( $ESS_1 = N$ ).

2. **While**  $\gamma_s < 1$ :

- a) Set  $s = s + 1$ .
- b) *Reweighting-Importance Sampling:* If:

$$w_s^{(i)}(\gamma_s) = w_{s-1}^{(i)} \frac{\pi_{\gamma_s}(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})}{\pi_{\gamma_{s-1}}(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})} = w_{s-1}^{(i)} U^{\gamma_s - \gamma_{s-1}}(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)}) \quad (2.7)$$

are the *updated* weights as a function of  $\gamma_s$ , then determine  $\gamma_s \in (\gamma_{s-1}, 1]$  so that the associated  $ESS_s = \zeta ESS_{s-1}$  (the value  $\zeta = 0.95$  was used in all the examples). Calculate  $w_s^{(i)}$  for this  $\gamma_s$ .

- c) *Resampling:* If  $ESS_s \leq ESS_{min}$  then resample (the value  $ESS_{min} = N/2$  was used in all the examples).
- d) *Rejuvenation:* Use an MCMC kernel  $P_s \left( (\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)}), (\boldsymbol{\theta}_s^{(i)}, \mathbf{d}_s^{(i)}) \right)$  that leaves  $\pi_{\gamma_s}$  invariant to perturb each particle  $(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)}) \rightarrow (\boldsymbol{\theta}_s^{(i)}, \mathbf{d}_s^{(i)})$
- e) The current population  $\{(\boldsymbol{\theta}_s^{(i)}, \mathbf{d}_s^{(i)}), w_s^{(i)}\}_{i=1}^N$  provides a particulate approximation of  $\pi_{\gamma_s}$  in the sense of Equations (2.4), (2.5).

3. **end while**

tributions which can be adapted to the specifics of the problem. In the following we present an alternative that can lead to significant computational savings.

It is clear that for cases of practical interest the most expensive part of the computations relates to the repeated evaluations of the utility function  $U(\boldsymbol{\theta}, \mathbf{d})$  and in particular in the Rejuvenation step of the algorithm in Table 2.1 (the utility functions values can be stored in memory and used in the Reweighting step). This is because each evaluation implies a run of the forward solver which can in many cases be costly. For that purpose we propose employing *approximate* computational models which might be less expensive but provide inaccurate evaluations of the utility function  $U(\boldsymbol{\theta}, \mathbf{d})$ . This idea has been successfully employed in [17, 19].

Such inexpensive, approximate models can be formally constructed using reduced-order modeling strategies [5] or less-rigorously by coarsening the spatio-temporal discretization of the governing PDEs or increasing the allowable error if iterative solvers are used for the solution of the system of algebraic equations. As it has also been demonstrated in [19], it is not important that the solutions of the approximate solver deviate significantly from the reference, but it suffices that they exhibit some sort of dependence in the sense to be explained in the sequence. For clarity of the presentation we assume that the approximate model consists of a coarsened discretization of the governing PDEs and the utility function evaluated by this solver is denoted by  $U_c(\boldsymbol{\theta}, \mathbf{d})$ . In contrast, the utility function of the reference/fine solver

with the desired discretization is denoted by  $U_f(\boldsymbol{\theta}, \mathbf{d})$ . The goal is to sample from  $\pi_f(\boldsymbol{\theta}, \mathbf{d}) \propto U_f(\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta})1_{\mathcal{D}}(\mathbf{d})$  as in Equation (2.2). The ideas presented can be readily generalized to a sequence of approximate models with increasing resolution, something that might be desired in practical cases where it is not a priori known what the appropriate resolution should be.

In order to make use of the information furnished by the approximate solver, we first define a sequence of distributions which employ  $U_c(\boldsymbol{\theta}, \mathbf{d})$  as follows:

$$\pi_{\hat{\gamma}}(\boldsymbol{\theta}, \mathbf{d}) \propto U_c^{\hat{\gamma}}(\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta})1_{\mathcal{D}}(\mathbf{d}), \quad \hat{\gamma} \in [0, 1] \quad (2.9)$$

The target density which is obtained for  $\hat{\gamma} = 1$  is  $\pi_c(\boldsymbol{\theta}, \mathbf{d}) \propto U_c(\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta})1_{\mathcal{D}}(\mathbf{d})$ . Sampling from this sequence of distribution can be achieved using the adaptive SMC scheme of Table 2.1. It is important to note that due the reduced cost associated which each evaluation of  $U_c$  the overall expense is generally much lower than trying to sample from  $\pi_f$ . Once a particulate approximation of  $\pi_c$  has been obtained, we propose operating on the following sequence:

$$\pi_{\tilde{\gamma}}(\boldsymbol{\theta}, \mathbf{d}) \propto U_c^{1-\tilde{\gamma}}(\boldsymbol{\theta}, \mathbf{d})U_f^{\tilde{\gamma}}(\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta})1_{\mathcal{D}}(\mathbf{d}), \quad \tilde{\gamma} \in [0, 1] \quad (2.10)$$

which for  $\tilde{\gamma} = 1$  recovers the target density  $\pi_f(\boldsymbol{\theta}, \mathbf{d})$ . The adaptive SMC scheme of Table 2.1 can be readily with a slight change in the Reweighting step where the updated weights of Equation (2.7) are now given by:

$$w_s^{(i)}(\tilde{\gamma}_s) = w_{s-1}^{(i)} \frac{\pi_{\tilde{\gamma}_s}(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})}{\pi_{\tilde{\gamma}_{s-1}}(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})} = w_{s-1}^{(i)} \left( \frac{U_f(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})}{U_c(\boldsymbol{\theta}_{s-1}^{(i)}, \mathbf{d}_{s-1}^{(i)})} \right)^{\Delta\tilde{\gamma}_s} \quad (2.11)$$

where  $\Delta\tilde{\gamma}_s = \tilde{\gamma}_s - \tilde{\gamma}_{s-1}$ . This reweighting has to be performed for all intermediate steps  $s = 1, \dots, S$  and since evaluations of the expensive utility  $U_f$  are required at each iteration (and each particle), the overall cost is proportional to  $S$ . The expression above implies that the second sequence of distributions is used to “correct” the inferences produced using the approximate solver. It is critically important to point out, that the correction required does not depend on the difference  $U_f - U_c$  but rather on the variability of the ratio  $\frac{U_f}{U_c}$  over the  $(\boldsymbol{\theta}, \mathbf{d})$  space. Hence if the regions of high-probability of  $\pi_c$  (i.e. the regions where  $U_c$  is high) coincide with the high-probability regions under  $U_f$  very few (potentially only one) iterations would be needed. The role of the approximate solver is to steer the sampling at regions of interest at a fraction of the cost. Good approximate solvers are therefore those for which the ratio  $\frac{U_f}{U_c}$  is as close to a constant as possible over the  $(\boldsymbol{\theta}, \mathbf{d})$  space. Note that this can be achieved even if  $U_f - U_c$  is large. We demonstrate in section 3 the reduction in computational cost that can be achieved. Furthermore, if a sequence of increasingly expensive solvers is utilized additional sequences of distributions as in Equation (2.10) can be defined. If the corresponding densities do not change, the analyst has the option of terminating the sampling. Finally in the case of state augmentation of Equation (2.3), the approximate solver(s) can be readily used at each  $n$  by defining two sequences as in Equation (2.9) and Equation (2.10).

**3. Numerical experiments.** Both problems examined in this work involve random heterogeneous materials in the context of steady-state heat diffusion with a governing PDE:

$$-\nabla \cdot (\lambda(\mathbf{x})\nabla T(\mathbf{x})) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega \quad (3.1)$$

The conductivity field  $\lambda(\mathbf{x})$  and/or input  $f(\mathbf{x})$  depend on the vector of uncertainties  $\boldsymbol{\theta}$  and design variables  $\mathbf{d}$ . As a result the solution field  $T(x)$  will also implicitly depend on  $(\boldsymbol{\theta}, \mathbf{d})$ . The finite element method is employed for the discretization of the aforementioned equations leading to the usual system of linear algebraic equations:

$$\mathbf{K}(\boldsymbol{\theta}, \mathbf{d}) \mathbf{T} = \mathbf{F}(\mathbf{d}) \quad (3.2)$$

The random variables  $\boldsymbol{\theta}$  parameterize the conductivities in a manner to be specialized in each of the examples below along with the boundary conditions.

Before embarking in the presentation of the results, it is worth providing some details on the MCMC kernels used in the Rejuvenation step of the adaptive SMC sampling scheme (Table 2.1). In particular we employ a Metropolis-within-Gibbs [25] or component-wise Hastings [4] sampler for the  $\boldsymbol{\theta}$  and  $\mathbf{d}$  coordinates separately i.e. we sample from the conditionals  $\pi_{\gamma_s}(\boldsymbol{\theta} \mid \mathbf{d}) \propto U^{\gamma_s}(\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta})$  and  $\pi_{\gamma_s}(\mathbf{d} \mid \boldsymbol{\theta}) \propto U^{\gamma_s}(\boldsymbol{\theta}, \mathbf{d})1_{\mathcal{D}}(\mathbf{d})$  respectively (from Equation (2.2)). Since the latter are analytically unavailable we employ a Metropolis-adjusted Langevin algorithm (MALA, [4]), which for updating the  $\boldsymbol{\theta}$ -coordinates of a particle  $i$ , from  $\boldsymbol{\theta}_{s-1}^{(i)}$  to  $\boldsymbol{\theta}_s^{(i)}$  requires:

$$\boldsymbol{\theta}_s^{(i)} = \boldsymbol{\theta}_{s-1}^{(i)} + \frac{\sigma_{\boldsymbol{\theta}}^2}{2} \nabla_{\boldsymbol{\theta}} \log \pi_{\gamma_s}(\boldsymbol{\theta}_{s-1}^{(i)} \mid \mathbf{d}_{s-1}^{(i)}) + \sigma_{\boldsymbol{\theta}} \mathbf{Z}_{s-1}^{(i)} \quad (3.3)$$

where  $\mathbf{Z}_{s-1}^{(i)}$  a vector of i.i.d Gaussian random variables  $\mathbf{Z}_{s-1}^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . The aforementioned proposal was augmented with the standard Metropolis accept/reject step [22]. Two observations are in order. Firstly, application of this scheme requires the calculation of derivatives of the utility function  $U$ . Due to its dependence on the solution vector of Equation (3.2) (in a manner to be made specific in the ensuing examples), this entails differentiation of  $\mathbf{T}(\boldsymbol{\theta}, \mathbf{d})$ . It is noted that such derivatives are also used in deterministic topology optimization schemes [15, 46]. Secondly, the parameter  $\sigma_{\boldsymbol{\theta}}$  controls the variance of the random noise. In the simulations performed its value was adjusted at each step of the SMC scheme so as to ensure an acceptance ratio between 50% and 80% [27]. Subsequently, and in order to update the  $\mathbf{d}$ -coordinates of a particle  $i$ , from  $\mathbf{d}_{s-1}^{(i)}$  to  $\mathbf{d}_s^{(i)}$ , we employ a MALA scheme:

$$\mathbf{d}_s^{(i)} = \mathbf{d}_{s-1}^{(i)} + \frac{\sigma_{\mathbf{d}}^2}{2} \nabla_{\mathbf{d}} \log \pi_{\gamma_s}(\mathbf{d}_{s-1}^{(i)} \mid \boldsymbol{\theta}_s^{(i)}) + \sigma_{\mathbf{d}} \tilde{\mathbf{Z}}_{s-1}^{(i)} \quad (3.4)$$

where  $\tilde{\mathbf{Z}}_{s-1}^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . The free parameter  $\sigma_{\mathbf{d}}$  is also adaptively adjusted so as to ensure an acceptance rate between 50% and 80%.

**3.1. Designing random heterogeneous materials.** This problem is inspired by deterministic topology optimization (most recently [46]). The problem domain  $\Omega$  is the unit square and Dirichlet and Neumann boundary conditions are prescribed as in Figure 3.1. The heat source density is assumed constant  $f(\mathbf{x}) = 1$  as in [46]. Typically in deterministic topology optimization the problem is posed as:

**Deterministic topology optimization:** *Given two materials with (significantly) different conductivities  $\lambda_1$  and  $\lambda_2$  (the values 1 and 0.01 were used in this study) which are to be used with volume fractions  $V_1$  and  $V_2$  (such that  $V_1 + V_2 = 1$ ), find the the spatial distribution of the two phases that minimize  $J = \int f(\mathbf{x})T(\mathbf{x}) d\Omega \approx \mathbf{F}^T \mathbf{T}$  (the latter objective is referred to as minimum compliance objective in solid mechanics [15]).*

In this work, we reformulate the problem for random heterogeneous materials. In particular we consider a composite made up of these two phases i.e.:

$$\lambda(\mathbf{x}) = \begin{cases} \lambda_1, & \mathbf{x} \in \text{phase 1} \\ \lambda_2, & \mathbf{x} \in \text{phase 2} \end{cases} \quad (3.5)$$

It is further assumed that  $\lambda(\mathbf{x})$  is a *random field* whose first-order distribution is fully defined by the volume fractions i.e.  $Pr[\lambda(\mathbf{x}) = \lambda_1] = V_1$  and  $Pr[\lambda(\mathbf{x}) = \lambda_2] = V_2 = 1 - V_1$ . Hence the conductivity at each  $\mathbf{x} \in \Omega$  (or each pixel in the discretized case) is a binary random variable. Clearly, optimizing with respect to  $\lambda(\mathbf{x})$  and using  $\int f(\mathbf{x})T(\mathbf{x}) d\Omega$  as an objective function is meaningless as they are both random. A viable set of design variables  $\mathbf{d}$  (given  $V_1$  and  $V_2$ ) consists of the higher-order statistics of the random field. More concretely, we consider conductivity random fields defined by a mapping from a zero-mean, unit variance Gaussian random field  $g(\mathbf{x})$  [3, 26] as follows:

$$\lambda(\mathbf{x}) = \lambda_1 + (\lambda_2 - \lambda_1)\Phi\left(\frac{g(\mathbf{x}) - \mu}{\epsilon}\right) \quad (3.6)$$

where  $\Phi(\cdot)$  is the standard normal cumulative distribution function<sup>2</sup>. The threshold  $\mu$  is selected in order to ensure the desired volume fractions. In this study, for  $V_1 = V_2 = 0.5$ , the value  $\mu = 0$  was used. The higher order statistics of the conductivity field are therefore determined by the covariance function of the Gaussian field  $g(\mathbf{x})$  [26, 44]. We consider a spectral representation of a statically homogeneous Gaussian random field [31] with respect to its power spectral density  $S_g(\omega_1, \omega_2)$  [31]<sup>3</sup>:

$$g(\mathbf{x} = (x_1, x_2)) = \sqrt{2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \left( A_{n_1, n_2} \cos(\omega_{1, n_1} x_1 + \omega_{2, n_2} x_2 + \phi_{n_1, n_2}) + \tilde{A}_{n_1, n_2} \cos(\omega_{1, n_1} x_1 - \omega_{2, n_2} x_2 + \tilde{\phi}_{n_1, n_2}) \right) \quad (3.7)$$

where:

- $\omega_{1, n_1} = n_1 \Delta\omega_1 = n_1 \frac{\omega_{1, max}}{N_1}$ ,  $\omega_{2, n_2} = n_2 \Delta\omega_2 = n_2 \frac{\omega_{2, max}}{N_2}$ <sup>4</sup>
-

$$A_{n_1, n_2} = \begin{cases} \sqrt{\frac{1}{2} S_g(\omega_{1, n_1}, \omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & n_1 = n_2 = 0 \\ \sqrt{S_g(\omega_{1, n_1}, \omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & n_1 = 0 \ \& \ n_2 > 0 \\ & , \text{ or } n_1 > 0 \ \& \ n_2 = 0 \\ \sqrt{2 S_g(\omega_{1, n_1}, \omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & \text{otherwise} \end{cases} \quad (3.8)$$

- 

$$\tilde{A}_{n_1, n_2} = \begin{cases} \sqrt{\frac{1}{2} S_g(\omega_{1, n_1}, -\omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & n_1 = n_2 = 0 \\ \sqrt{S_g(\omega_{1, n_1}, -\omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & n_1 = 0 \ \& \ n_2 > 0 \\ & , \text{ or } n_1 > 0 \ \& \ n_2 = 0 \\ \sqrt{2 S_g(\omega_{1, n_1}, -\omega_{2, n_2}) \Delta\omega_1 \Delta\omega_2}, & \text{otherwise} \end{cases} \quad (3.9)$$

- $\phi_{n_1, n_2}$ ,  $\tilde{\phi}_{n_1, n_2}$  are random phase angles uniformly distributed in  $[0, 2\pi]$ .

Given the expressions above, the following parameterization is adopted:

<sup>2</sup>a very small  $\epsilon$  is used so that  $\Phi$  approximates a Heaviside function at  $\mu$ . In this study  $\epsilon = 0.001$

<sup>3</sup>the power spectral density is the Fourier transform of the autocovariance function

<sup>4</sup>the upper-cutoff frequencies effectively determine the scale of heterogeneity

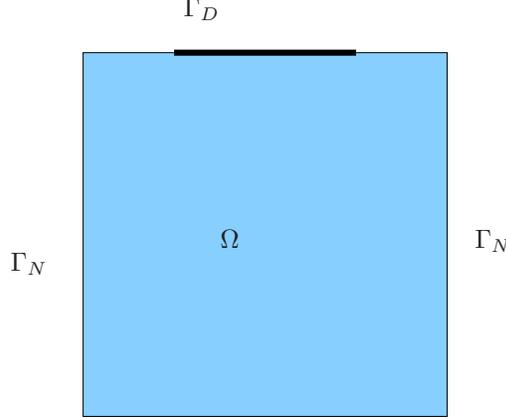


FIG. 3.1. Configuration for problem described in section 3.1 taken from [46]. Dirichlet boundary conditions specified on  $\Gamma_D = [0.25, 0.75]$  where  $T = 0$  and Neumann boundary conditions on  $\Gamma_N$  to be  $\lambda \frac{\partial T}{\partial n} = \mathbf{0}$ .

- random variables  $\boldsymbol{\theta} = \{\phi_{n_1, n_2}, \tilde{\phi}_{n_1, n_2}\}_{n_1, n_2=0}^{N_1-1, N_2-1}$  of dimension  $N_1 N_2$ , and with  $p(\boldsymbol{\theta}) = \left(\frac{1}{2\pi}\right)^{N_1 N_2}$ .
- design variables  $\mathbf{d} = \{S_g(\omega_{1, n_1}, \omega_{2, n_2})\}_{n_1=0, n_2=-(N_2-1)}^{N_1-1, N_2-1}$  of dimension  $(2N_2 - 1)N_1$ . It is noted that since we consider unit variance Gaussian fields, the power spectral density must integrate to 1. This imposes a constraint on the sum of the design variables. An additional constraint is that  $S_g(\omega_{1, n_1}, \omega_{2, n_2}) \geq 0, \forall n_1, n_2$ .

The objective function  $J(\boldsymbol{\theta}, \mathbf{d}) = \int f(\mathbf{x})T(\mathbf{x}) \, d\Omega \approx \mathbf{F}^T \mathbf{T}$  used in deterministic topology optimization formulations, will now be a random variable due to its dependence on the random, spatial distribution of conductivities. The two extreme values it attains are  $\sim 10^{-3}$  when  $\lambda(\mathbf{x}) = \lambda_1 = 1, \forall \mathbf{x} \in \Omega$  and  $\sim 10^{-1}$  when  $\lambda(\mathbf{x}) = \lambda_2 = 0.01 \forall \mathbf{x} \in \Omega$ . The utility function we employ is:

$$U(\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} 1 & \text{if } J(\boldsymbol{\theta}, \mathbf{d}) \leq J_0 = 2.5 \times 10^{-3} \\ e^{-c(J(\boldsymbol{\theta}, \mathbf{d}) - J_0)} & \text{if } J(\boldsymbol{\theta}, \mathbf{d}) > J_0 = 2.5 \times 10^{-3} \end{cases} \quad (3.10)$$

For a large  $c$  (the value  $c = 10$  was used in this study) the aforementioned utility function approximates the Heaviside function at  $J(\boldsymbol{\theta}, \mathbf{d}) = J_0$ . As a result the expected utility  $\hat{U}(\mathbf{d})$  (Equation (2.1)) represents the probability that  $J$  is less than the threshold  $J_0$ . Hence, the formulation of the problem in the stochastic topology optimization framework proposed is:

**Stochastic topology optimization:** *Given two materials with (significantly) different conductivities  $\lambda_1$  and  $\lambda_2$  (the values 1 and 0.01 were used in this study) which are to be randomly distributed with volume fractions  $V_1 = V_2 = 0.5$ , find the random spatial distribution as prescribed by the design variables  $\mathbf{d}$  above, that maximize the probability that  $J(\boldsymbol{\theta}, \mathbf{d})$  is less than a prescribed threshold  $J_0$ .*

In the simulations performed,  $N = 100$  particles were used and  $n = 5$  state augmentations (Equation (2.3)). Since the design variables  $\mathbf{d}$  represent the power spectral density  $S_g$ , the particulate representations (i.e. particle values  $\mathbf{d}^{(i)}$  and weights  $W^{(i)}$ )

were used to estimate the expected value of  $S_g$  as in Figure 3.2(a). The corresponding autocovariance function is depicted in Figure 3.2(b).

The result of the *deterministic topology optimization* is shown in Figure 3.3(a) [46]. The high-conductivity material is depicted with black ( $\lambda_1 = 1$ ) and mostly occupies the region under the boundary  $\Gamma_D$  where Dirichlet boundary conditions are specified. Thick or thin clusters of the phase 1 emanate from  $\Gamma_D$  over the whole domain in order for the energy introduced by the heat source  $f(\mathbf{x})$  (Equation (3.1)) to permeate throughout  $\Omega$ , ensuring a low value of the objective. Obviously such realizations (as in Figure 3.3(a)) are inconsistent with a statistically homogeneous random field  $\lambda(\mathbf{x})$  prescribed in the stochastic optimization framework. For example it is clear that the volume fraction of the black phase is close to 1 near  $\Gamma_D$  and drops to zero far away. The proposed framework however allows one to control microstructures by controlling the statistics of their random distribution (in this case through  $S_g$ ). Hence the composite is always random and its first-order distribution (as expressed by the volume fractions) is always the same.

Figures 3.3(b) depict three realizations of the random medium generated by the optimal  $S_g$  shown in Figure 3.2(a) through the mapping of Equation (3.6). It is noted that the average  $S_g$  corresponding to a uniform distribution on the design variables, gives rise to random checkerboards, i.e. realizations that exhibit zero correlation. The realizations of Figure 3.3(b) however exhibit strong correlation patterns. In particular it is noted that the black phase (high conductivity) exhibits connected paths, particularly in the vertical direction ( $x_2$ ). This is also verified in Figure 3.4 where the lineal-path functions of the black phase for various separations in the directions  $x_1$  (horizontal) and  $x_2$  (vertical) have been calculated. In the pixelized version, the lineal-path function  $L(\Delta x)$  [44] expresses the probability that a line of pixel-length  $\Delta x$  lies wholly on the black phase. It can be readily calculated using the spectral density  $S_g$  (Figure 3.2(a), or the autocovariance (Figure 3.2(b)) of the underlying Gaussian field in Equation (3.6). It can be clearly seen that the *optimal random microstructure* is statistically anisotropic. As it is perhaps expected from the result of the deterministic topology optimization (Figure 3.3(a)), the optimal random composite exhibits higher connectivity in the vertical direction  $x_2$  than in the horizontal  $x_1$ .

**3.2. Design/Control of random heterogeneous materials.** The goal of this problem is to optimally select the input in a random system described by a heterogeneous medium so as to maximize an expected utility related to the response. In particular we consider the rectangular domain  $\Omega = [-1, 1] \times [0, 1]$  of Figure 3.5, where  $T = 0$  on  $\Gamma_D$ ,  $-\lambda(\mathbf{x}) \frac{\partial T(\mathbf{x})}{\partial n} = 0$  on  $\Gamma_{N_1}$  and  $-\lambda(\mathbf{x}) \frac{\partial T(\mathbf{x})}{\partial n} = \mathbf{q}$  on  $\Gamma_{N_2}$ .

The design variables  $\mathbf{d}$  in this problem parameterize the imposed flux  $\mathbf{q}$  on the left boundary  $\Gamma_{N_2}$ . We consider several utility functions that relate to the temperature profile  $\mathbf{T}_\Sigma$  along the vertical line  $\Sigma$  located at  $x_1 = 0$  [49]. The random variables  $\boldsymbol{\theta}$  parameterize the random conductivity field  $\lambda(\mathbf{x})$ . In particular we consider a statistically homogeneous  $\lambda(\mathbf{x})$  defined through a zero-mean Gaussian field  $g(\mathbf{x})$  as:

$$\lambda(\mathbf{x}; \boldsymbol{\theta}) = e^{g(\mathbf{x})} \quad (3.11)$$

The autocovariance  $C(\Delta x_1, \Delta x_2)$  of  $g(\mathbf{x})$  is prescribed as:

$$C(\Delta x_1, \Delta x_2) = \exp\left\{-\frac{\Delta x_1^2 + \Delta x_2^2}{x_0}\right\} \quad (3.12)$$

The value of  $x_0 = 0.2$  for the correlation length was used in this study. In order

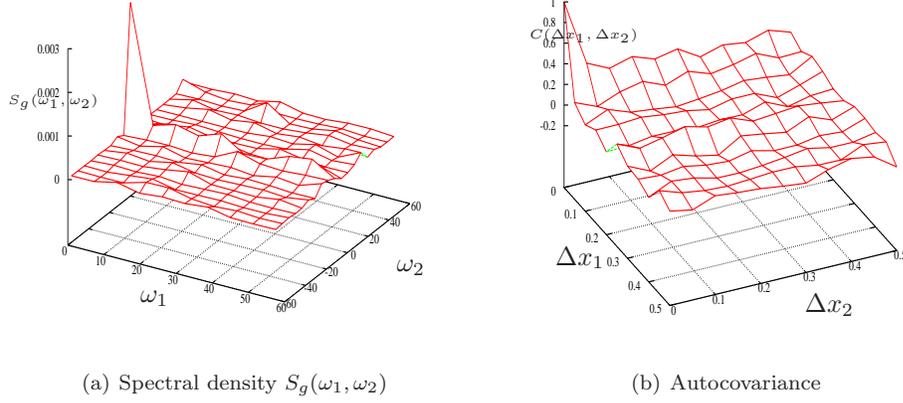


FIG. 3.2. *a) Average power spectral density  $S_g(\omega_1, \omega_2)$  found by averaging over the particles i.e.  $\sum_{i=1}^N W^{(i)} \mathbf{d}^{(i)}$  (Equation (3.7) for  $N_1 = N_2 = 10$  and  $\omega_{1,max} = \omega_{2,max} = 20\pi$ ). b) Corresponding autocovariance  $C(\Delta x_1, \Delta x_2)$  for various separations  $(\Delta x_1, \Delta x_2)$ . The latter is found by taking the Fourier transform of  $S_g(\omega_1, \omega_2)$  in a).*

to obtain a resolution-independent representation of  $g(\mathbf{x})$  we employed a Karhunen-Loève expansion:

$$g(\mathbf{x}) = \sum_{i=1}^{n_\theta} \theta_i \sqrt{\mu_i} \phi_i(\mathbf{x}) \quad (3.13)$$

where the eigenvalues  $\mu_i$  and eigenfunctions  $\phi_i(\mathbf{x})$  of the autocovariance can be calculated semi-analytically as in [13]. The series was truncated at  $n_\theta = 1,000$  which was found to represent 95% of the variance of  $g(\mathbf{x})$ . The corresponding 1,000 standard normal variates  $\theta_i$  represent the random variables  $\boldsymbol{\theta}$  in this problem. Figure 3.6 depicts a sample realization of  $\lambda(x)$  obtained from Equations (3.11) and (3.13).

**3.2.1. One design variable - Bimodal expected utility.** In this example we consider a constant flux on the left boundary  $\Gamma_{N_2}$  and therefore a single design variable. The utility function employed was:

$$U(\boldsymbol{\theta}, \mathbf{d}) = e^{-\frac{\|\mathbf{T}_\Sigma - \mathbf{T}_{target}^{(1)}\|^2}{2\sigma^2}} + 6e^{-\frac{\|\mathbf{T}_\Sigma - \mathbf{T}_{target}^{(2)}\|^2}{2\sigma^2}} \quad (3.14)$$

Each of the terms in the sum above provide a measure of the difference between the temperature profile  $\mathbf{T}_\Sigma$  along  $\Sigma$  (Figure 3.5) and some target values  $\mathbf{T}_{target}^{(1)}$  and  $\mathbf{T}_{target}^{(2)}$ . Hence we seek the imposed flux so that the temperatures  $\mathbf{T}_\Sigma$  are as close as possible to the prescribed targets. It is noted that  $\mathbf{T}_\Sigma$  depends on  $\boldsymbol{\theta}$  and  $\mathbf{d}$  but this has been omitted for notational economy. The reason two target profiles were selected is to assess whether the proposed scheme can correctly identify more than one local maxima of the expected utility. It is finally mentioned that the value  $\sigma = 5$

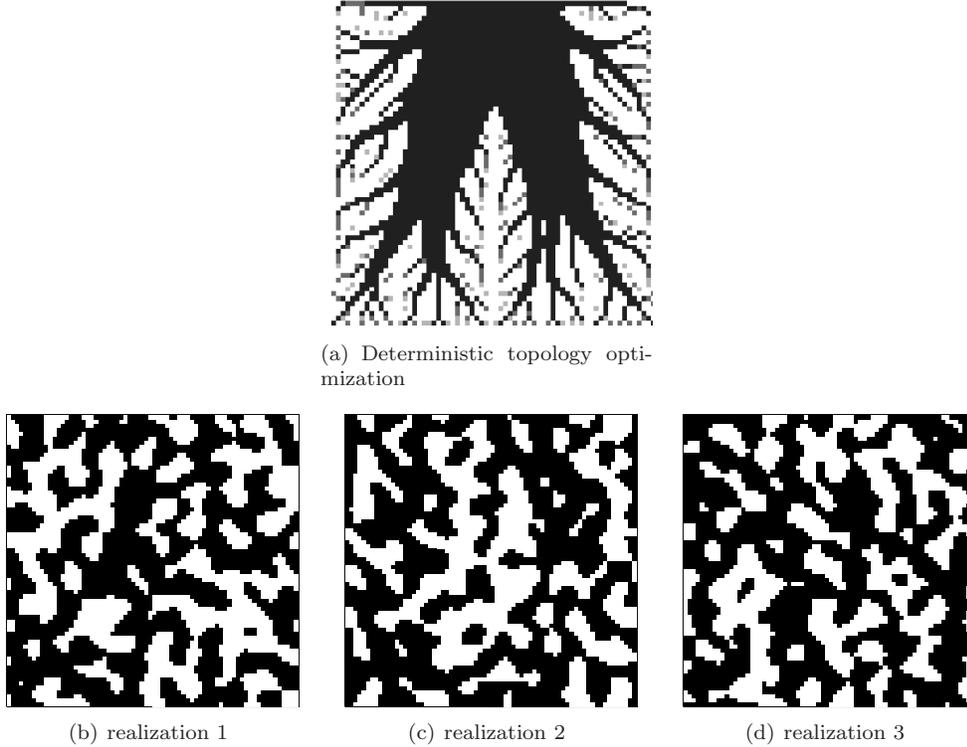


FIG. 3.3. *a) Result of deterministic optimization taken from [46] (copyright permission pending). b), c), d) Sample realizations obtained from the proposed stochastic topology optimization scheme. All results correspond to a  $70 \times 70$  grid. The high-conductivity material is depicted with black ( $\lambda_1 =$ ).*

was used and the target temperature profiles were constant along  $\Sigma$  and equal to 35 and 75 respectively.

Figure 3.7 depicts the  $\mathbf{d}$  coordinates of the particles and their histogram which is proportional to the expected utility  $\hat{U}(\mathbf{d})$  (Equation (2.1)). This simulation was performed with  $N = 1,000$  particles and entailed sampling in  $n_{\theta} + n_{\mathbf{d}} = 1,000 + 1 = 1,000$  dimensions. The algorithm can clearly identify and populate the two modes which correspond to two distinct local maxima of the expected utility, despite the high-dimensionality of the sampling space.

The same utility function was used and state augmentation was employed in order to test the capability of the algorithm to zoom in the global maximum. Figure 3.8 depicts particles and expected utilities with  $n = 1$ ,  $n = 3$  and  $n = 5$  state augmentations as in Equation (2.3). It is clearly seen that the global maximum is identified.

**3.2.2. Two design variables - Unimodal expected utility.** In this problem we increase the design variables to two i.e.  $\mathbf{d} = (d_1, d_2)$ . In particular it is assumed that  $d_1$  represents the flux on the upper half of  $\Gamma_{N_2}$  and  $d_2$  at the lower. The following

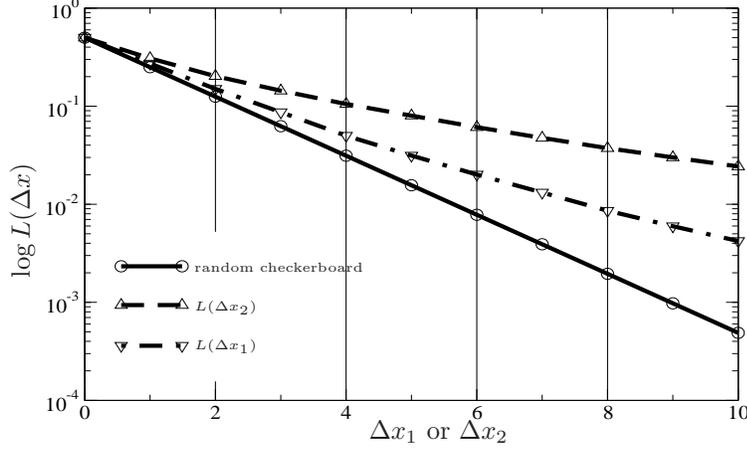


FIG. 3.4. Lineal path functions in the directions  $x_1$  and  $x_2$  for various separations  $\Delta x_1$  or  $\Delta x_2$ . These are compared with the linear path function of a random checkerboard (i.e. no correlation).

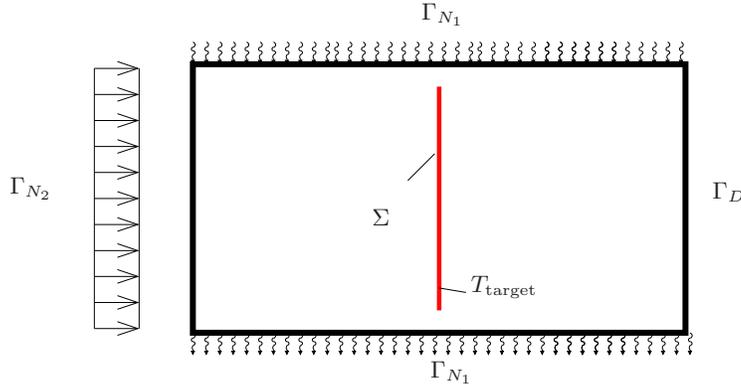


FIG. 3.5. Problem domain

utility function was used:

$$U(\boldsymbol{\theta}, \mathbf{d}) = e^{-\frac{\|\mathbf{T}_\Sigma - \mathbf{T}_{target}\|^2}{2\sigma^2}} \quad (3.15)$$

and the target temperature profile was taken constant and equal to 35. Figure 3.9 depicts the  $\mathbf{d}$ -coordinates of  $N = 100$  particles that were used to solve this problem without ( $n = 1$ ) and with state augmentation ( $n = 5$ ). These runs entailed sampling in 1,002 and 5,002 dimensions respectively. An interesting observation arising from these results is that the maxima of the expected utility seem to be attained for  $d_1 + d_2 = 20$ . This is more clear in Figure 3.9(b) and provides physical insight into the sensitivity of the random system to the input  $\mathbf{d}$ . In particular, it appears that the total flux (i.e.  $d_1 + d_2$ ) controls the temperature profile along  $\Sigma$ .

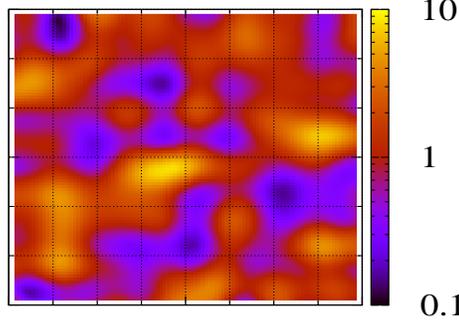


FIG. 3.6. Sample realization of the conductivity field  $\lambda(\mathbf{x})$  prescribed in Equations (3.11) and (3.13)

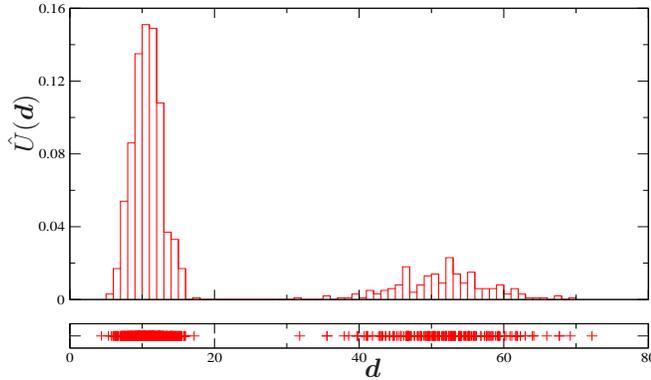


FIG. 3.7. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=1,000}$  and corresponding histogram. The latter is proportional to the expected utility  $\hat{U}(\mathbf{d})$  of the problem in section 3.2.1

**3.2.3. Multi-resolution analysis.** The goal of this last example is to illustrate the potential of significant computational savings by employing approximate models in the manner explained in section 2.2. In particular, we consider a two-resolution approach where the role of the approximate model is played by a finite element solver with a coarse resolution. Our coarse/approximate model consists of 200 triangular finite elements and our reference/fine solver of 800 finite elements (in both cases uniform meshes were used). The comparisons in terms computational cost are expressed in terms of the (equivalent) number of calls to the finer (most expensive) solver. In this problem, the cost of the coarse solver is much less and corresponds to  $\frac{1}{64}$  calls to the fine solver. Our goal is to use the former in order to expedite the sampling and reduce the overall number of calls to the latter solver.

We consider a single design variable  $\mathbf{d}$  (representing the flux on  $\Gamma_{N_2}$ ) and the utility function of Equation (3.15). The cost of the reference solution which employs

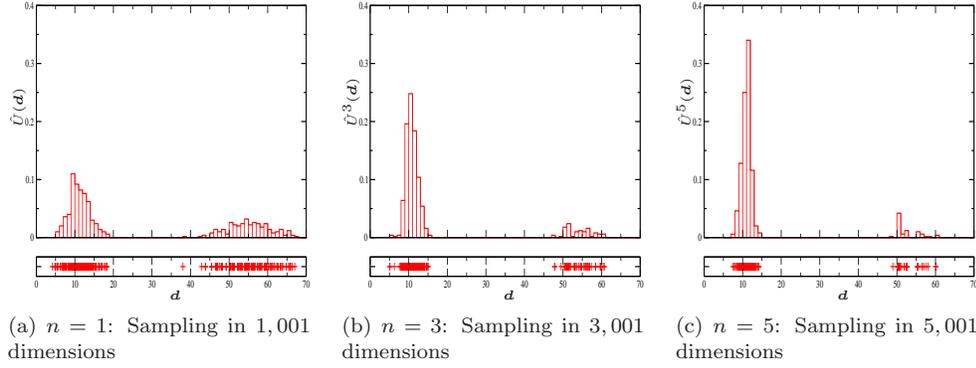


FIG. 3.8. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=1,000}$  and corresponding histograms for  $n = 1, 3, 5$  state augmentations for the utility function of Equation (3.14)

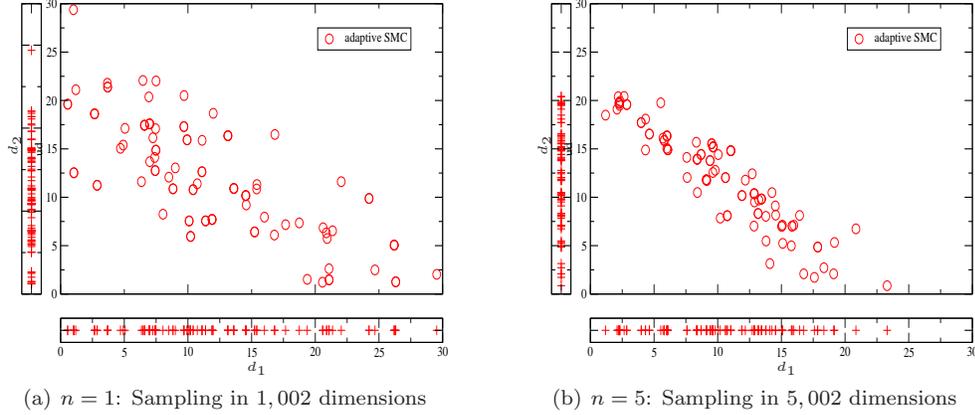


FIG. 3.9. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=100}$  for the utility function of Equation (3.15)

only the fine solver and the advocated adaptive SMC scheme is 7,200 calls. The result of this simulation in terms of the  $\mathbf{d}$ -coordinates of the  $N = 100$  particles used and the expected utility  $\hat{U}(\mathbf{d})$  are depicted in Figure 3.10.

Figure 3.11 compares the value of the utility function (Equation (3.15)) calculated with for the same  $\theta$  and  $\mathbf{d}$  values using the coarse i.e.  $U_c$  and fine, i.e.  $U_f$  solvers (section 2.2). It can be seen that the coarse model underestimates  $U_f$  and in absolute terms provides a poor approximation. Furthermore one observes a significant scatter which clearly implies that  $U_c$  cannot uniquely predict  $U_f$ .

In the proposed framework however, as explained in section 2.2, it suffices that the output  $U_c$  of the coarse/approximate model is correlated with  $U_f$  in order to achieve computational savings. We employed the two sequences of distributions as in Equation (2.9) and Equation (2.10). The first allows us to estimate the maxima of the

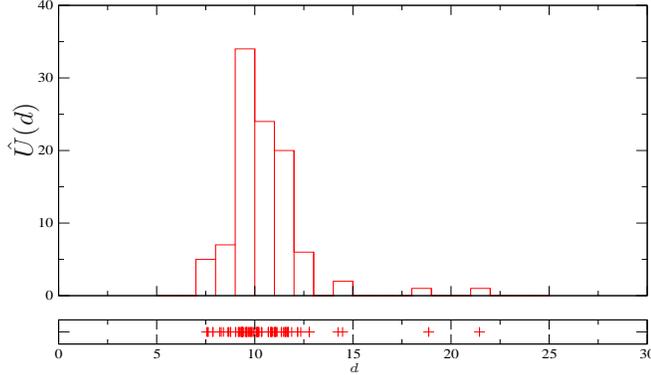


FIG. 3.10. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=100}$  and corresponding histogram. The latter is proportional to the expected utility  $\hat{U}(\mathbf{d})$  in section 3.2.3

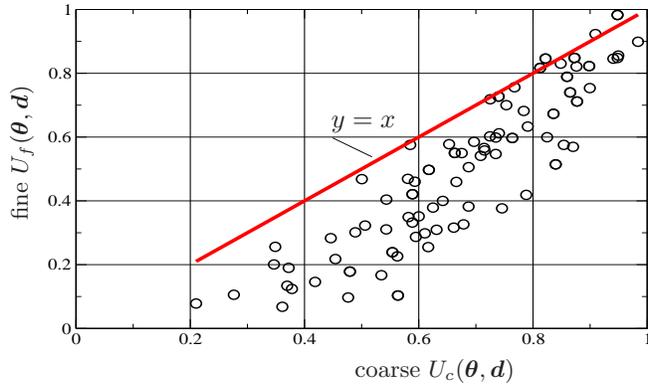


FIG. 3.11. Each circle corresponds to  $(U_c(\boldsymbol{\theta}, \mathbf{d}), U_f(\boldsymbol{\theta}, \mathbf{d}))$  (given by Equation (3.15)) evaluated using the coarse and fine solvers for the same  $(\boldsymbol{\theta}, \mathbf{d})$  values each time

coarse/approximate expected utility and requires only calls to the inexpensive/coarse solver. The  $\mathbf{d}$ -coordinates and the estimated expected approximate utility are depicted in Figure 3.12 (blue line - “coarse”). The cost of obtaining this result with  $N = 100$  particles was equivalent to 138 calls to the fine/expensive solver. Obviously the result differs from the expected fine/reference utility (red line - “fine” in Figure 3.12). Using this distribution as the starting point for sampling from the second sequence of distributions in Equation (2.10), which requires calls to the coarse and fine solvers, ultimately leads to the result depicted in Figure 3.12 (green line - “coarse+fine”). The cost of sampling from this second sequence was equivalent to 717 runs of the most expensive solver. Hence even though the total cost was  $138 + 717 = 955$  runs, i.e. a reduction by a factor of 7, the result obtained practically coincides with the reference solution.

The efficiency gain by utilizing the approximate model was also tested for the

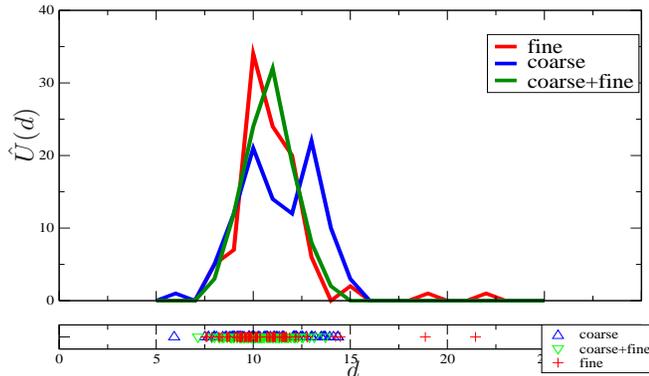


FIG. 3.12. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=100}$  and corresponding histograms obtained using: only the fine solver (red), only the coarse solver (blue) and, combination of coarse and fine solvers (green) as in the sequences of Equation (2.9) and Equation (2.10).

utility function of Equation (3.14). As it can be seen in Figure 3.13, which compares  $U_c$  and  $U_f$ , the coarse solver provides a very poor approximation of the output of the fine solver in terms of  $U_f$ . It is also noted that the quality of the approximation seems to deteriorate for large utility function values which are of interest.

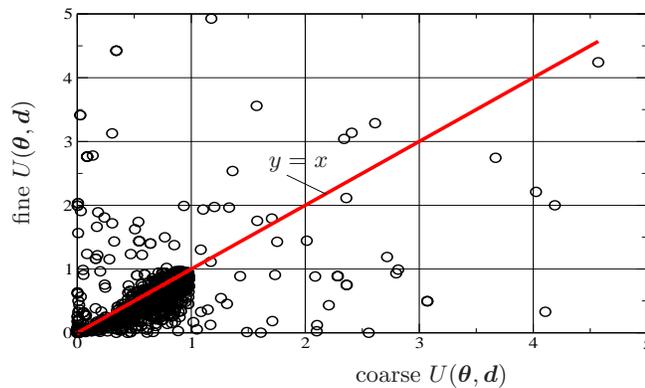


FIG. 3.13. Each circle corresponds to  $(U_c(\boldsymbol{\theta}, \mathbf{d}), U_f(\boldsymbol{\theta}, \mathbf{d}))$  (given by Equation (3.14)) evaluated using the coarse and fine solvers for the same  $(\boldsymbol{\theta}, \mathbf{d})$  values each time

Figure 3.14 compares the accuracy of the sampling with  $N = 1,000$  particles. It is noted that the cost of using exclusively the fine solver (red line - “fine”) is equivalent to 57,000 runs whereas the total cost of employing the coarse and fine solver in the manner explained in section 2.2 is 25,500 runs. Even though the computational savings achieved (a factor of 2) are not as striking as in the previous case, it is still significant, particularly when considering the poor correlation between the two outputs in Figure 3.13. The expected utilities estimated exhibit negligible differences (red and green lines in Figure 3.14) despite the presence of two modes.

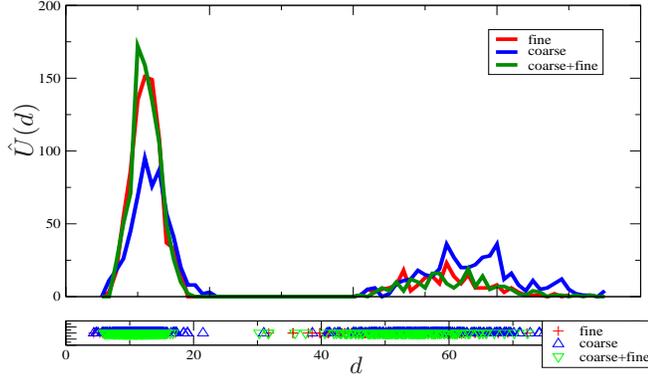


FIG. 3.14. Particles  $\{\mathbf{d}^{(i)}\}_{i=1}^{N=1,000}$  and corresponding histograms obtained using: only the fine solver (red), only the coarse solver (blue) and, combination of coarse and fine solvers (green) as in the sequences of Equation (2.9) and Equation (2.10).

**4. Conclusions.** A critical task in the context of random heterogeneous materials involves their design and the optimization of their response/behavior/performance. An embarrassingly parallelizable sampling scheme is discussed that is capable of dealing with systems with high-dimensional vectors of uncertainties and design variables. An efficient adaptive SMC scheme is proposed that can efficiently populate regions of the design space where the expected utility function attains its maxima. The proposed framework allows the principled utilization of approximate models in order to achieve further reduction in computational cost and enables the multi-resolution analysis of such problems. An extension currently investigated involves the utilization of statistical regression techniques to identify the relation between the approximate and the reference models that could lead to more efficiency gains. Moreover, we are currently examining the development of probabilistic reduced-order models that would be adaptively trained and be suited to the analysis objectives.

## REFERENCES

- [1] BA AMZAL, F. BOIS, E. PARENT, AND CP. ROBERT, *Bayesian-optimal design via interacting particle systems*, Journal of the American Statistical Association, 101 (2006), pp. 773–785.
- [2] M.P. BENDSOE AND O. SIGMUND, *Material interpolation schemes in topology optimization*, ARCHIVE OF APPLIED MECHANICS, 69 (1999), pp. 635 – 654.
- [3] N.F. BERK, *Scattering properties of leveled-wave model for random morphologies*, Physical Review A, 44 (1991), pp. 5069–5079.
- [4] J BESAG, *Markov chain Monte Carlo for statistical inference*, Center for Statistics and the Social Sciences, (2001).
- [5] JOHN BURKARDT, QIANG DU, AND MAX GUNZBURGER, *Reduced order modeling of complex systems Reduced-order modeling*, (2003), pp. 29–38.
- [6] D.C. CHARMPIS, G.I. SCHUELLER, AND M.F. PELLISSETTI, *The need for linking micromechanics of materials with stochastic finite elements: A challenge for materials science*, COMPUTATIONAL MATERIALS SCIENCE, 41 (2007).
- [7] N CHOPIN, *Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference*, ANNALS OF STATISTICS, 32 (2004), pp. 2385 – 2411.
- [8] P. DEL MORAL, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer New York, 2004.
- [9] P. DEL MORAL, A. DOUCET, AND A. JASRA, *Sequential Monte Carlo for Bayesian Computation (with discussion)*, in Bayesian Statistics 8, Oxford University Press, 2006.
- [10] ———, *Sequential Monte Carlo Samplers*, Journal of the Royal Statistical Society B, 68 (2006), pp. 411–436.
- [11] A. DOUCET, J. F. G. DE FREITAS, AND N. J. GORDON, eds., *Sequential Monte Carlo Methods in Practice*, Springer New York, 2001.
- [12] A. GERSBORG-HANSEN, M.P. BENDSOE, AND O. SIGMUND, *Topology optimization of heat conduction problems using the finite volume method*, STRUCTURAL AND MULTIDISCIPLINARY OPTIMIZATION, 31 (2006), pp. 251 – 259.
- [13] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, 1991.
- [14] L.V. GIBIANSKY AND O. SIGMUND, *Multiphase composites with extremal bulk modulus*, JOURNAL OF THE MECHANICS AND PHYSICS OF SOLIDS, 48 (2000), pp. 461 – 498.
- [15] J.K. GUEST AND T. IGUSA, *Structural optimization under uncertain loads and nodal locations*, COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING, 198 (2008), pp. 116 – 124.
- [16] M.A. JOHANSEN, A. DOUCET, AND M. DAVY, *Particle methods for maximum likelihood parameter estimation in latent variable models*, Statistics and Computing, 18 (2008), pp. 47–57.
- [17] MC KENNEDY AND A O’HAGAN, *Predicting the output from a complex computer code when fast approximations are available*, 87 (2000), pp. 1 – 13.
- [18] P.S. KOUTSOURELAKIS, *Probabilistic characterization and simulation of multi-phase random media*, Probabilistic Engineering Mechanics, 21 (2006).
- [19] ———, *Accurate uncertainty quantification using inaccurate models*, SIAM Journal of Scientific Computing, 31 (2009), pp. 3274–3300.
- [20] P.S. KOUTSOURELAKIS AND G. DEODATIS, *Simulation of binary random processes with applications to two-phase random media*, Journal of Engineering Mechanics ASCE, 131 (2005).
- [21] H. KÜCK, N. DE FREITAS, AND DOUCET. A., *SMC Samplers for Bayesian Optimal Nonlinear Design*, in Nonlinear Statistical Signal Processing Workshop (NSSPW), 2006.
- [22] J.S. LIU, *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics, Springer, 2001.
- [23] S.N. MACÉACHERN, M. CLYDE, AND J.S. LIU, *Sequential Importance Sampling for Nonparametric Bayes Models: The Next Generation*, The Canadian Journal of Statistics / La Revue Canadienne de Statistique, 27 (1998), pp. 251–267.
- [24] P. MÜLLER, *Simulation based optimal design*, in Bayesian Statistics 6, 1998.
- [25] C. P. ROBERT AND G. CASELLA, *Monte Carlo Statistical Methods*, Springer New York, 2nd ed., 2004.
- [26] A.P. ROBERTS AND M. TEUBNER, *Transport properties of heterogeneous materials derived from gaussian random fields: Bounds and simulation*, Physical Review E, 51 (1995), pp. 4141–4153.
- [27] G.O. ROBERTS AND J.S. ROSENTHAL, *Optimal scaling for various metropolis hastings algorithms*, Statist. Sci., 16 (2001), pp. 351–367.
- [28] A. ROMKES, J.T. ODEN, AND K. VEMAGANTI, *Multi-scale goal-oriented adaptive modeling of random heterogeneous materials*, MECHANICS OF MATERIALS, 38 (2006), pp. 859 –

- 872.
- [29] S. SANKARAN AND N. ZABARAS, *A maximum entropy approach for property prediction of random microstructures*, Acta Materialia, (2006), pp. 2265–2276.
  - [30] G.I. SCHUELLER AND H.A. JENSEN, *Computational methods in optimization considering uncertainties - an overview*, COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING, 198 (2008), pp. 2 – 13.
  - [31] M. SHINOZUKA AND G. DEODATIS, *Simulation of multi-dimensional stochastic processes by spectral representation*, ASME Applied Mechanics Reviews, 49 (1996), pp. 29–53.
  - [32] O. SIGMUND, *Materials with prescribed constitutive parameters - an inverse homogenization problem*, INTERNATIONAL JOURNAL OF SOLIDS AND STRUCTURES, 31 (1994), pp. 2313 – 2329.
  - [33] ———, *Tailoring materials for specific needs*, JOURNAL OF INTELLIGENT MATERIAL SYSTEMS AND STRUCTURES, 5 (1994), pp. 736 – 742.
  - [34] ———, *Tailoring materials with prescribed elastic properties*, MECHANICS OF MATERIALS, 20 (1995), pp. 351 – 368.
  - [35] ———, *On the design of compliant mechanisms using topology optimization*, MECHANICS OF STRUCTURES AND MACHINES, 25 (1997), pp. 493 – 524.
  - [36] ———, *A new class of extremal composites*, JOURNAL OF THE MECHANICS AND PHYSICS OF SOLIDS, 48 (2000), pp. 397 – 428.
  - [37] ———, *Topology optimization: a tool for the tailoring of structures and materials*, PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A-MATHEMATICAL PHYSICAL AND ENGINEERING SCIENCES, 358 (2000), pp. 211 – 227.
  - [38] ———, *Manufacturing tolerant topology optimization*, ACTA MECHANICA SINICA, 25 (2009), pp. 227 – 239.
  - [39] O. SIGMUND AND S. TORQUATO, *Composites with extremal thermal expansion coefficients*, APPLIED PHYSICS LETTERS, 69 (1996), pp. 3203 – 3205.
  - [40] ———, *Design of materials with extreme thermal expansion using a three-phase topology optimization method*, JOURNAL OF THE MECHANICS AND PHYSICS OF SOLIDS, 45 (1997), pp. 1037 – 1067.
  - [41] ———, *Design of smart composite materials using topology optimization*, SMART MATERIALS & STRUCTURES, 8 (1999), pp. 365 – 379.
  - [42] G. STEFANOY AND M. PAPADRAKAKIS, *Stochastic finite element analysis of shells with combined random material and geometric properties*, COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING, 193 (2004), pp. 139 – 160.
  - [43] V SUNDARARAGHAVAN AND N ZABARAS. A, *Multi-length scale sensitivity analysis for the control of texture-dependent properties in deformation processing.*, International Journal of Plasticity, 24 (2008), pp. 1581–1605.
  - [44] S. TORQUATO, *Random Heterogeneous Materials*, Springer-Verlag, 2002.
  - [45] S.S. VEL AND A.J. GOUPEE, *Multiscale thermoelastic analysis of random heterogeneous materials part i: Microstructure characterization and homogenization of material properties*, COMPUTATIONAL MATERIALS SCIENCE, 48 (2010), pp. 22 – 38.
  - [46] E. WADBRO AND M. BERGGREN, *Megapixel topology optimization on a graphics processing unit*, SIAM Review, (2009).
  - [47] X.F. XU, X. CHEN, AND L.H. SHEN, *A green-function-based multiscale method for uncertainty quantification of finite body random heterogeneous materials*, COMPUTERS & STRUCTURES, 87 (2009), pp. 1416 – 1426.
  - [48] C.L.Y. YEONG AND S. TORQUATO, *Reconstructing Random Media I and II*, Physical Review E, 58 (1998), pp. 224–233.
  - [49] N. ZABARAS AND B. GANAPATHYSUBRAMANIAN, *A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach*, Journal of Computational Physics, (2008), pp. 4697–4735.