

Efficient Propagation of Uncertainty via Reordering Monte Carlo Samples

Danial Khatamsaz *

Dept. of Mechanical Engineering.
Texas A&M University
College Station, TX, 77843
Email: khatamsaz@tamu.edu

Vahid Attari

Dept. of Materials Science and Engineering.
Texas A&M University
College Station, TX, 77843
Email: attari.v@tamu.edu

Raymundo Arróyave

Dept. of Materials Science and Engineering.
Texas A&M University
College Station, TX, 77843
Email: rarroyave@tamu.edu

Douglas L. Allaire

Dept. of Mechanical Engineering.
Texas A&M University
College Station, TX, 77843
Email: dallaire@tamu.edu

Uncertainty analysis in the outcomes of model predictions is a key element in decision-based material design to establish confidence in the models and evaluate the fidelity of models. Uncertainty Propagation (UP) is a technique to determine model output uncertainties based on the uncertainty in its input variables. The most common and simplest approach to propagate the uncertainty from a model inputs to its outputs is by feeding a large number of samples to the model, known as Monte Carlo (MC) simulation which requires exhaustive sampling from the input variable distributions. However, MC simulations are impractical when models are computationally expensive. In this work, we investigate the hypothesis that while all samples are useful on average, some samples must be more useful than others. Thus, reordering MC samples and propagating more useful samples can lead to enhanced convergence in statistics of interest earlier and thus, reducing the computational burden of UP process. Here, we introduce a methodology to adaptively reorder MC samples and show how it results in reduction of computational expense of UP processes.

1 INTRODUCTION

In many engineering applications, decision-making processes rely on numerical simulation models. Most often, inputs to numerical models have some sort of uncertainty that induce uncertainty in model outputs. Thus, characterization, propagation, and analysis of uncertainty is a crucial step in any model development task. Understanding uncertainties enables providing a confidence measure to evaluate the applicability of different computational models for decision-making. Uncertainty quantification (UQ) and uncertainty propagation (UP) are recognized as essential components in many engineering applications where UQ refers to understanding uncertainty sources and UP refers to determining output uncertainty of a model due to uncertainties of input variables.

The most common and simplest approach to propagate the uncertainty from input to output is by feeding a large number of inputs to numerical models, known as Monte Carlo (MC) simulation. Based on the strong law of large

*Corresponding author.

numbers and the central limit theorem, convergence in the distribution of a quantity of interest is expected. MC integration methods are known as the gold standard approach to carry out UP [1, 2]. However, the computational expense associated with MC simulations makes such methods prohibitive and impractical in many engineering applications. To mitigate the computational burden of MC simulations, other methods have been developed such as importance sampling [3] and adaptive sampling [4]. Other approaches to carry out UP are local expansion-based methods [5] that are weak against large variability of inputs, functional expansion-based methods [6], and numerical integration-based methods [7]. Change of probability measure from a desired input distribution is another technique to handle UP problems [8, 9, 10]. There are different ways to transfer a proposal measure to a target measure and one widely used method is the use of Radon-Nikodym (R-N) derivative [11, 12]. A change of measure based on R-N theory is performed by calculating importance weights using the density ratio of target to proposal densities. Note that although the density ratio cannot be calculated via a closed-form expression when underlying probability distributions are unknown, the R-N theory still applies. Accordingly, a sample-based approach has been proposed in Ref.[9]. The idea is to generate a large number of hypercubes different in size all over the input space. The density ratio of target and proposal samples is calculated by counting samples inside each hypercube. Next, a system of linear equations, one equation per hypercube, is solved to obtain the weights. Via sampling from weighted proposal samples, the empirical distribution of target samples is approximated. Another approach is proposed in Ref. [13] that works with determinable empirical distribution functions. They calculate importance weights by minimizing the L_2 -norm between a weighted proposal empirical distribution and a target distribution function. Although this approach claims to be effective in high-dimensional and large-scale problems, as many samples occupy the boundaries of high-dimensional spaces, numerical ill-conditioning eventually happens. Although implementing a change of measure method enables efficiency gains by skipping the propagation of target samples to computational models, it requires the availability of previously simulated data using the same model on identical input-output spaces. In scenarios where no such set of data or proposal samples exist, there is no choice but directly propagating target samples through computational models.

In this study, we propose an efficient approach to mitigate the computational burden of MC simulation methods for uncertainty propagation purposes. Assume that there exist a large set of samples yet to be propagated through a computational model to obtain the empirical distribution of the model’s outputs. While all samples are important on average, the hypothesis here is some samples can be more useful in representing the empirical distribution of all samples. In other words, the Addition or elimination of a particular sample has an impact on the empirical distribution of all the samples, but this impact is not similar for every sample. Herein, the goal is to determine the importance of samples based on their role in defining the empirical distribution of all the samples. Therefore, it is possible to re-order samples based on their importance to be propagated through a model sequentially. Our approach suggests an efficient use of resources by picking the most informative samples when evaluation of all samples is not practical.

The rest of the paper proceeds as follows. In Sec. 2, we introduce the proposed framework to reorder samples of a given set based on their importance in representing the empirical distribution of all samples. Next, in Sec. 3, the application is demonstrated on an engineering problem that requires running a computationally expensive simulation model. Finally, in Sec. 4, we provide concluding remarks and discuss avenues of future works.

2 METHODOLOGY

In this section, we discuss our proposed method in detail and provide algorithms for easy implementation of the sequentially optimal sampling concept. The method can be applied to any set of samples regardless of the dimensionality and distribution of samples.

In algorithm 1, different steps of the method are stated. Assuming that we have available a set of samples S yet to be propagated through a model. We are interested to determine the importance of each sample to re-order samples accordingly. Thus, by sequentially propagating them through a model of interest, we assure once the computational resources are exhausted, we obtain the empirical distribution of a quantity of interest with the most similarity to the case where all samples had been propagated. The algorithm starts by initializing sets P and R to represent sets of sequentially picked samples and the samples yet to be picked respectively. At every iteration, samples from the set R are temporarily augmented to the set P one by one. The goal is to find the sample that minimizes the dissimilarity between empirical distributions of the temporarily updated set of picked samples and set S . Here, we use the Wasserstein metric for this purpose where $\mathbf{W} = [w_1, w_2, \dots, w_d]$ is the vector that entry w_i indicates Wasserstein distance

between two empirical distributions in i^{t^h} dimension of a d dimensional space. At every iteration, the minimizer of $\|\mathbf{W}\|_1 := \sum_{i=1}^d w_i$ is picked to be added to the set P and to be removed from the set R. In algorithm 1, function "Wass" takes samples from both sets and calculates Wasserstein distance. We suggest using the Manhattan distance of Wasserstein metric (L_1 -norm) to calculate Wasserstein distance to avoid the dominance of large Wasserstein distance of a single dimension which may cause diminishing reductions in Wasserstein distances in other dimensions. We call this technique as "Adaptive Sampling Method". By implementing algorithm 1, assuming the goal is to re-order n

Algorithm 1 Adaptive sampling to re-order planned Monte Carlo samples

given:
sample set $S=\{s_1, s_2, \dots, s_n\}$
set of sequentially picked samples $P=\{\}$
 $R \leftarrow S - P$
while $R \neq \emptyset$ **do**
 $s_{\text{picked}} = \operatorname{argmin}_{s_i \in R} \|\text{Wass}(S, P+s_i)\|_1$
 $P \leftarrow P + s_{\text{picked}}$
 $R \leftarrow S - P$
end while

samples, the algorithm has to complete $\frac{n(n+1)}{2} - 1$ iterations which exponentially increases with the number of samples. In such cases, instead of identifying the best sample at each iteration, it is suggested to look for the best batch of samples to update the set P. Algorithm 2 shows different steps in the batch setting. Considering b as the batch size, k different batches of samples are randomly generated by picking b random samples from the set R. Then, instead of augmenting a single sample, a batch of samples is temporarily augmented to the set P to calculate the Wasserstein distance between the sets S and P. The best batch of samples is determined to update the set P and to be removed from the set R accordingly.

Algorithm 2 Adaptive sampling to re-order planned Monte Carlo samples in batch setting

given:
sample set $S=\{s_1, s_2, \dots, s_n\}$
set of sequentially picked samples $P=\{\}$
batch size b
number of batches to generate k
 $R \leftarrow S - P$
while $R \neq \emptyset$ **do**
 $b_{\text{picked}} = \operatorname{argmin}_{b_i \subseteq R}^{i=1:k} \|\text{Wass}(S, P+b_i)\|_1$
 $P \leftarrow P + b_{\text{picked}}$
 $R \leftarrow S - P$
end while

3 DEMONSTRATION

A moving boundary problem for the study of interface evolution during spinodal decomposition in alloys is used to demonstrate the framework developed in this study. The model is based on a free energy model for heterogeneous medium accounting for bulk and interfacial free energies,

$$F^{tot}(c, \nabla c) = \int_V [f_{bulk} + \frac{\kappa}{2}(\nabla c)^2] dV \quad (1)$$

where c is alloy composition, κ is gradient energy coefficient, and f_{bulk} is the bulk free energy function given as,

$$f_{bulk} = W(c - c_\alpha)(c - c_\beta) \quad (2)$$

where W is the barrier height of phase transformation, and c_α and c_β are the equilibrium composition of the product phases that are set to 0.3 and 0.7, respectively. Through high-throughput phase-field simulations, time series of synthetic microstructures will be generated for the investigation of parameter space on the microstructure landscape of a hypothetical alloy during isothermal thermal annealing. The boundary value problem follows:

$$\begin{aligned} \frac{\partial c}{\partial t} &= \nabla \cdot \left\{ M \nabla \left(\frac{\partial f_{bulk}}{\partial c} - \kappa \nabla^2 c \right) \right\} & 0 < x, y < L_x, L_y \\ & & 0 < t < t^* \\ \text{BC:} & c(0, y, t) = c(L_x, y, t) & c(x, 0, t) = c(x, L_y, t) \\ \text{IC:} & c(x, y, 0) = c^* + A\zeta \end{aligned} \quad (3)$$

where M is the inherently positive effective atomic mobility of the species. The lengths of the simulation domain are set to $L_x = L_y = 200$ with the grid size of 256×256 and t^* is the final model run time. BC and IC denote the used boundary and initial conditions, respectively. c^* is the initial average value of the alloy composition perturbed by a constant noise magnitude A , and ζ is a Gaussian random number with the interval of $[-1, +1]$. The simulations were carried out using combinations of $[c^*, W, \kappa, M]$ parameter sets. The material properties, such as barrier height of transformation, mobility, and gradient energy coefficient for a given alloy with composition (c^*) are often highly uncertain or not available. As a result, a prior distribution with a certain physical range is necessary to be taken into account. Our assumed distributions for these parameters are shown in Fig. 1.

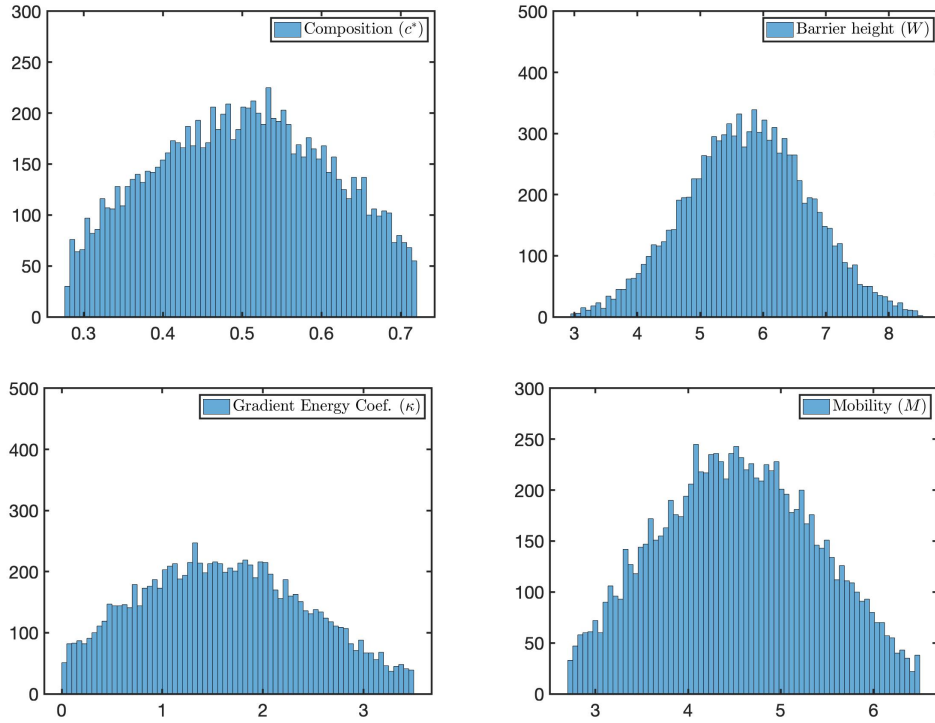


Fig. 1. Proposed distributions for input parameters of the phase-field model (i.e., $[c^*, W, \kappa, M]$)

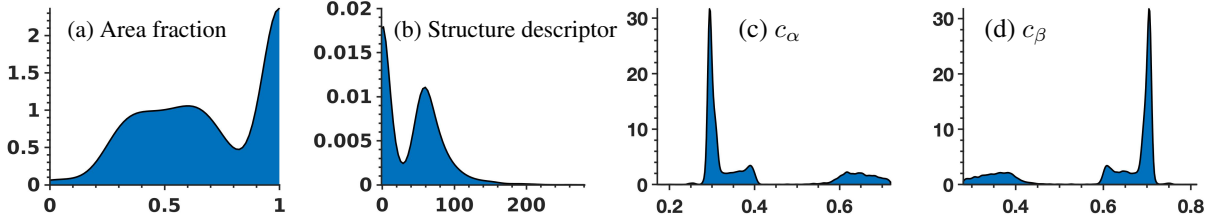


Fig. 2. Probability density functions of Quantities of Interest (QoI) extracted from microstructure images generated by the phase-field model. (a) Area fraction of phases, (b) radially averaged FFT structure descriptor and composition of (c) phase α and (d) phase β in the simulation domain

In this phase-field model, the direct output is time-series images of microstructures, each with a dimension of 256×256 . For further evaluation of the microstructures, these images are often condensed into a reduced set of physical and non-physical Quantities of Interest (QoI). The conventional reduction of image information is often a one-way transfer without the possibility of inverse transfer from QoI to microstructure image. Due to this condensation, materials' properties and performance are subject to significant uncertainty. Our study determines the area fraction of the phases, the composition of each phase, and the characteristic length scale of the microstructure from radially averaged Fast Fourier Transform (FFT) spectra [14]. The probability density functions for these QoIs are extracted for a constant heat treatment duration (i.e., t^*) and are shown in Fig. 2. These posterior distributions show a diverse range of values for quantities of interest. For instance, the sharp peaks in Fig. 2(c) and (d) show the equilibrium composition of 0.3 and 0.7 for the two product phases as dictated by the free energy. Some simulations, however, could also produce non-equilibrium composition values due to an uncertain set of kinetic and thermodynamic parameters. There-

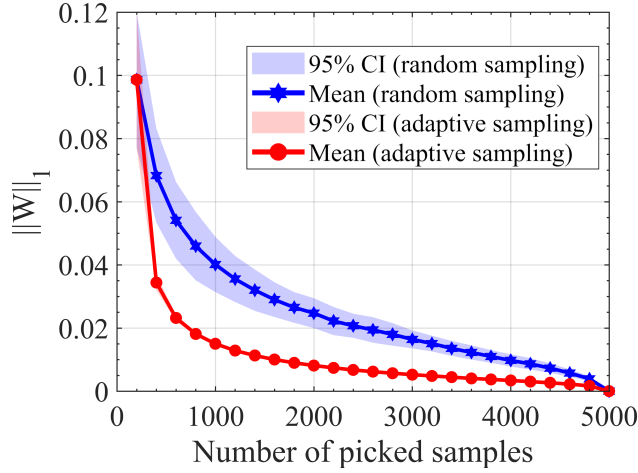


Fig. 3. Manhattan distance of Wasserstein metric between empirical distributions of sequentially picked samples (adaptive and random sampling) and all samples

fore, due to inherent parametric uncertainties, nonlinearity, and difficulties in the post-processing of microstructure images, phase-field simulations are computationally expensive to run. Moreover, advanced phase-field models often combine several order parameters and multiphysics interactions (e.g., thermal, electrical, mechanical, magnetic). This increased degree of complexity in the numerical and parametric calculation of these models often results in the uncertain evaluation of the material's property and performance space during modeling real-world processes (e.g., additive manufacturing [15], memristive materials for brain-like (neuromorphic) computing [16], electrodeposition reaction kinetics in battery materials [17], solder interconnect joint formation [18] and electromigration [19], microstructure evolution in thermoelectric materials for energy conversion [20], to name a few). It is important to note that Eq. 3

(i.e., Cahn-Hilliard equation) and some of its variants are also relevant to phenomena other than phase separation in materials. For instance, tumor growth [21], population dynamics [22], image processing [23] and even the irregular structure in Saturn’s rings [24] are some noteworthy examples.

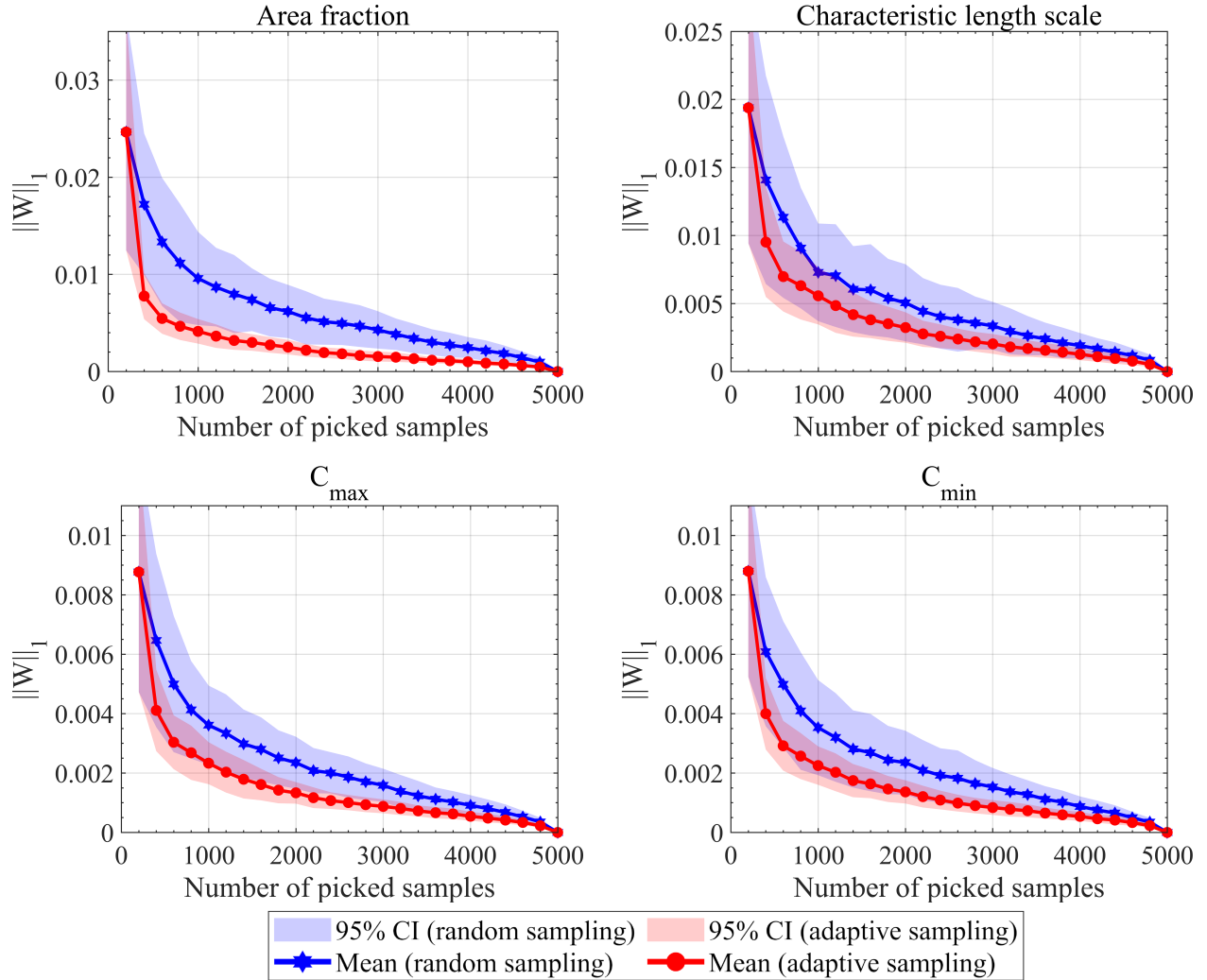


Fig. 4. Mean and confidence interval of Wasserstein distance between empirical distributions of sequentially picked samples (adaptive and random sampling) and all samples in the output space

We seek to enhance the traditional Monte Carlo sampling methods to enable their use when faced with computationally expensive phase-field models. We, therefore, consider the problem of enhancing the convergence rate of Monte Carlo simulations by creating algorithms that ensure optimal convergence of a sequentially sampled input vectors. Here, we have available a set of 5000 samples, and we implement the adaptive sampling method in the batch setting to re-order the samples. We then sequentially propagate them through the model. The Batch size is set to 200 and at every iteration, 20,000 different batches are generated (note that in the last iteration, only 200 samples remain in the set R , so it can be skipped by simply augmenting the last batch to the set P). The simulations are replicated 100 times. The Manhattan distance of the Wasserstein metric (L^1 -norm) is plotted in Fig. 3. For comparison purposes, the result of a random sampling policy is also illustrated. There are two key points in Fig. 3: first, using the adaptive

sampling policy to pick the best samples, the Wasserstein distance between the sets of picked samples and all samples is significantly smaller compared to the random sampling policy. This means the same measure of similarity between 2 sets is achieved using a much less number of samples. This emphasizes the fact that some samples are more useful (more informative about the distribution). Thus, based on these results, the same inference about distribution is made using less number of samples if they are picked optimally. The second key point is the narrow confidence interval of the adaptive sampling method. This essentially indicates that almost the same set of samples is consistently recognized at different replications. Note that, multiple replications only apply to the batch setting as at every replication, different sets of batches are generated whereas if the samples are picked one by one, all replication essentially return exactly the same order of samples.

In the next step, samples are propagated through the model to obtain the empirical distributions of all quantities of interest. Figure 4 illustrates Wasserstein distances comparing the adaptive sampling and random sampling policies. Here, the confidence intervals for both policies are wider since the distances between two samples in the input space and output space are different. However, still significant efficiency gains are observed comparing the required number of propagated samples to achieve the same Wasserstein distance in adaptive and random sampling policies.

As mentioned earlier, one can search for the most useful batch of samples to pick among different generated batches. To investigate the impact of batch size on the performance of the framework, we have performed adaptive sampling using different batch sizes. The results are depicted in Fig. 5. As we reduce the batch size (increasing reso-

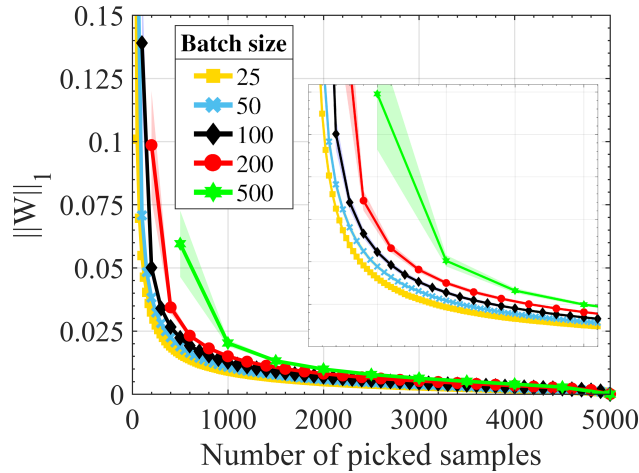


Fig. 5. Comparison of Wasserstein distances between sequentially picked and all samples at different batch sizes

lution), it improves the similarity between the empirical distributions of sequentially picked samples and all samples. However, note that even with the largest batch size, after one iteration, the difference is minimized since the framework recognizes and picks a batch with the most useful samples anyway. Therefore, if the goal is to pick the smallest number of samples possible, a smaller batch size is beneficial while for more relaxed conditions, a larger batch size can also do the job. The best result is achieved when samples are picked sequentially one by one. The trade-off here is smaller batch sizes require more iterations to complete the process. In our problem, since the number of samples is not drastically large (5,000 samples), Wasserstein metric calculations take almost the same computational time at any iteration, thus the wall-time increases almost linearly with respect to the number of iterations.

4 CONCLUSIONS AND FUTURE WORK

Although MC simulations suggest a simple approach to propagate uncertainty from a model inputs to its outputs, running thousands of simulations is impractical in many engineering applications. In this work, we introduced the concept of re-ordering MC samples based on their usefulness in representing the empirical distribution of all samples. In this sense, while all samples are important on average, some samples are more informative. We proposed to

determine the importance of samples based on their impact on the Wasserstein distance between the sets of all and sequentially picked samples. The more informative a sample is, the more reduction in Wasserstein distance is observed. After re-ordering all samples, they are sequentially propagated through a computational model. The results show significant efficiency gains in comparison to random sample propagation. We also provided the method in the batch setting to decrease computational time by recognizing informative batches of samples instead of testing samples one by one. The results of simulations using different batch sizes suggest that using smaller batch sizes increases efficiency by effectively picking only highly informative samples. However, the batch size effect will be diminished after a few iterations as the framework will pick batches with important samples quickly.

In this work, the assumption is that samples are generated in the first place and then, we re-order samples before propagating them to computational models. The subject of future work is to propose methods to sample from a distribution consciously instead of generating random samples from input variable distributions. Therefore, efficiency gains are expected in scenarios where sampling from distributions can be computationally demanding.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the National Science Foundation through Grant No. CDSE-2001333 and DMR-1905325, as well as ARPA-E through contract DE-AR0001427. Calculations were carried out at the Texas A&M High-Performance Research Computing (HPRC) Facility.

REFERENCES

- [1] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., 2004, “Bayesian data analysis chapman & hall,” *CRC Texts in Statistical Science*.
- [2] Swiler, L. P., 2006, Bayesian methods in engineering design problems. Tech. rep., Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA . . .
- [3] Melchers, R., 1989, “Importance sampling in structural systems,” *Structural safety*, **6**(1), pp. 3–10.
- [4] Bucher, C. G., 1988, “Adaptive sampling—an iterative fast monte carlo procedure,” *Structural safety*, **5**(2), pp. 119–126.
- [5] Thoft-Cristensen, P., and Baker, M. J., 2012, *Structural reliability theory and its applications* Springer Science & Business Media.
- [6] Xiu, D., and Karniadakis, G. E., 2002, “Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos,” *Computer methods in applied mechanics and engineering*, **191**(43), pp. 4927–4948.
- [7] Evans, D. H., 1967, “An application of numerical integration techniques to statistical tolerancing,” *Technometrics*, **9**(3), pp. 441–456.
- [8] Amaral, S., Allaire, D., and Willcox, K., 2017, “Optimal L_2 L_2 -norm empirical importance weights for the change of probability measure,” *Statistics and Computing*, **27**(3), pp. 625–643.
- [9] Sanghvi, M., Honarmandi, P., Attari, V., Duong, T., Arroyave, R., and Allaire, D. L., 2019, “Uncertainty propagation via probability measure optimized importance weights with application to parametric materials models,” In AIAA Scitech 2019 forum, p. 0967.
- [10] Kloek, T., and Van Dijk, H. K., 1978, “Bayesian estimates of equation system parameters: an application of integration by monte carlo,” *Econometrica: Journal of the Econometric Society*, pp. 1–19.
- [11] Bourgin, R. D., 2006, *Geometric aspects of convex sets with the Radon-Nikodym property*, Vol. 993 Springer.
- [12] Shepp, L. A., 1966, “Radon-nikodym derivatives of gaussian measures,” *The Annals of Mathematical Statistics*, pp. 321–354.
- [13] Amaral, S., Allaire, D., and Willcox, K., 2017, “Optimal L_2 -norm empirical importance weights for the change of probability measure,” *Statistics and Computing*, **27**(3), may, p. 625–643.
- [14] Honarmandi, P., Attari, V., and Arroyave, R., 2022, “Accelerated materials design using batch bayesian optimization: A case study for solving the inverse problem from materials microstructure to process specification,” *Computational Materials Science*, **210**, p. 111417.
- [15] Karayagiz, K., Johnson, L., Seede, R., Attari, V., Zhang, B., Huang, X., Ghosh, S., Duong, T., Karaman, I., Elwany, A., et al., 2020, “Finite interface dissipation phase field modeling of Ni–Nb under additive manufacturing conditions,” *Acta Materialia*, **185**, pp. 320–339.

- [16] Shi, Y., and Chen, L.-Q., 2018, "Phase-field model of insulator-to-metal transition in VO_2 under an electric field," *Physical Review Materials*, **2**(5), p. 053803.
- [17] Davidson, R., Verma, A., Santos, D., Hao, F., Fincher, C. D., Zhao, D., Attari, V., Schofield, P., Van Buskirk, J., Fraticelli-Cartagena, A., et al., 2020, "Mapping mechanisms and growth regimes of magnesium electrodeposition at high current densities," *Materials Horizons*, **7**(3), pp. 843–854.
- [18] Attari, V., and Arroyave, R., 2016, "Phase field modeling of joint formation during isothermal solidification in 3DIC micro packaging," *Journal of Phase Equilibria and Diffusion*, **37**(4), pp. 469–480.
- [19] Attari, V., Ghosh, S., Duong, T., and Arroyave, R., 2018, "On the interfacial phase growth and vacancy evolution during accelerated electromigration in Cu/Sn/Cu microjoints," *Acta Materialia*, **160**, pp. 185–198.
- [20] Yi, S.-i., Attari, V., Jeong, M., Jian, J., Wang, H., Arroyave, R., and Yu, C., 2018, "Strain-induced suppression of the miscibility gap in nanostructured Mg_2Si – Mg_2Sn solid solutions," *Journal of Materials Chemistry A*.
- [21] Khain, E., and Sander, L. M., 2008, "Generalized cahn-hilliard equation for biological applications," *Physical review E*, **77**(5), p. 051129.
- [22] Cohen, D. S., and Murray, J. D., 1981, "A generalized diffusion model for growth and dispersal in a population," *Journal of Mathematical Biology*, **12**(2), pp. 237–249.
- [23] Bertozzi, A. L., Esedoglu, S., and Gillette, A., 2006, "Inpainting of binary images using the cahn–hilliard equation," *IEEE Transactions on image processing*, **16**(1), pp. 285–291.
- [24] Tremaine, S., 2003, "On the origin of irregular structure in saturn's rings," *The Astronomical Journal*, **125**(2), p. 894.