A Multifidelity Method for a Nonlocal Diffusion Model

Parisa Khodabakhshi[a,*], Karen E. Willcox[a], Max Gunzburger[a,b]

[a] *Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Texas 78712 USA*
[b] *Department of Scientific Computing, Florida State University, Tallahassee, Florida 32306 USA*

Abstract

Nonlocal models feature a finite length scale, referred to as the horizon, such that points separated by a distance smaller than the horizon interact with each other. Such models have proven to be useful in a variety of settings. However, due to the reduced sparsity of discretizations, they are also generally computationally more expensive compared to their local differential equation counterparts. We introduce a multifidelity Monte Carlo method that combines the high-fidelity nonlocal model of interest with surrogate models that use coarser grids and/or smaller horizons and thus have lower fidelities and lower costs. Using the multifidelity method, the overall computational cost of uncertainty quantification is reduced without compromising accuracy. It is shown for a one-dimensional nonlocal diffusion example that speedups of up to two orders of magnitude can be achieved using the multifidelity method to estimate the expectation of an output of interest.

Keywords: Nonlocal models, Multifidelity methods, Uncertainty quantification, Monte Carlo Methods, Nonlocal diffusion

1. A Nonlocal Diffusion Model

We consider a derivative-free integral equation model for diffusion that allows for pairs of points separated by a finite distance to interact with each other, in contrast to partial differential equation (PDE) models for which pairs of points interact only within the infinitesimal neighborhoods needed to define derivatives. Such nonlocal models allow for the nucleation and propagation of discontinuities in their solutions as well as other phenomena that cannot be adequately modeled by PDEs. As a result, nonlocal models have been used in many diverse settings such as subsurface flows, diffusion processes, fracture mechanics, and image processing, just to name a few. Detailed consideration on nonlocal models and of their use in applications can be found in, e.g., [1, 3, 4, 5] and the references cited therein. It is important to note that the nonlocal models we consider are not integral equation reformulation of PDEs, e.g., defined by using Green functions, but are instead based on a different modeling paradigm.

The main impediment preventing the even more widespread use of nonlocal models is caused by the reduced sparsity of discretizations compared to that for PDEs. Of course, that reduction is due to nonlocality, i.e., due to points being separated by a finite distance interacting with each other. As a result, both the assembly and solution costs of discrete approximations can be

*Corresponding author: parisa@austin.utexas.edu.

substantially greater compared to that for related PDEs. In multi-query applications (e.g., uncertainty quantification, optimization, and control) the increased computational costs can render nonlocal modeling to be infeasible.

Here, we show how lower-cost, lower-fidelity surrogate models can be used to obtain, at a much lower total cost and without compromising accuracy, a high-fidelity approximation. Specifically, we do so for the uncertainty quantification of an output of interest that depends on the solution of a nonlocal diffusion model. Our approach generalizes to the nonlocal setting the multifidelity methods introduced in [6, 7] for the local PDE setting. We emphasize that due to the reduced sparsity of nonlocal models compared to that of PDE models, the need for approaches such as the one we introduce in this paper is much greater and thus multifidelity methods have potential for greater impact for nonlocal models compared to local models.

A nonlocal diffusion model. In the remainder of this section we provide a brief review of the nonlocal model we consider; detailed presentations are given in, e.g., [1, 3, 4, 5]. To simplify the exposition, we confine ourselves to a one-dimensional setting (as a proof of concept); generalization to higher dimensions is a straightforward, albeit tedious, exercise.

Given a domain $\Omega = (0,\ L) \subset \mathbb{R}$ and a length scale δ (referred to as the *interaction radius* or *horizon*), we define the *interaction domain* $\Omega_{\mathcal{I}} = \{y \in \mathbb{R} \setminus \Omega\ :\ |y - x| \leq \delta \text{ for some } x \in \Omega\} = [-\delta, 0] \cup [L,\ L+\delta]$, i.e., $\Omega_{\mathcal{I}}$ consists of all points in $\mathbb{R} \setminus \Omega$ that interact with points in Ω. For given functions $b(x)$, $g(x)$, and $\gamma(x, x')$ defined on Ω, $\Omega_{\mathcal{I}}$, and $(\Omega \cup \Omega_{\mathcal{I}}) \times (\Omega \cup \Omega_{\mathcal{I}})$, respectively, the one-dimensional, steady-state nonlocal diffusion model we consider is given by

$$\begin{cases} -2 \displaystyle\int_{x-\delta}^{x+\delta} \left(u_\delta(x') - u_\delta(x)\right) \gamma(x, x')\, dx' = b(x) \quad \forall x \in \Omega = (0,\ L) & \text{(a)} \\ \qquad\qquad\qquad\qquad u_\delta(x) = g(x) \quad \forall x \in \Omega_{\mathcal{I}} = [-\delta,\ 0] \cup [L,\ L+\delta]\,. & \text{(b)} \end{cases} \tag{1}$$

We refer to (1b) as a *Dirichlet volume constraint* because it is applied on a domain having nonzero volume and refer to (1) as a *volume-constrained problem*, with the qualifier *Dirichlet* indicating that the constraint (1b) involves the specification of function values. The authors refer the reader to [2, 3] for detailed discussion on the well-posedness of the nonlocal diffusion problem.

The function $\gamma(x, x')$ in (1) is referred to as the *kernel*; the choice made for $\gamma(x, x')$ determines the properties of solutions including their smoothness properties. As a result of the flexibility available in the choice of $\gamma(x, x')$, nonlocal models can model a wide variety of observed behaviors. For example, for bounded kernels, the nonlocal problem admits solutions having jump discontinuities; of course, such solutions are not obtainable from second-order elliptic PDEs. Note that (1) is a nonlocal analog of the PDE Dirichlet boundary-value problem $-\nabla \cdot (a(x)\nabla u) = b(x)$ in Ω and $u = g(x)$ on the boundary of Ω. Indeed, as the horizon $\delta \to 0$, the nonlocal model, when properly formulated (i.e. scaling γ appropriately), reduces to such a PDE problem; see [2, 3, 4].

Nonlocal models such as (1) have been subject to discretization via the same approaches as those used for discretizing PDEs. Which of these approaches one selects is not germane to our goals. For the sake of concreteness, we use a finite difference method. Specifically, for a positive integer N, we set $h = L/N$, $N_\delta = \lceil \delta/h \rceil$ (where $\lceil \cdot \rceil$ denotes rounding to the nearest larger integer), $x_n = nh$ for $n = -(N_\delta - 1), \ldots, N + (N_\delta - 1)$, $x_n = -\delta$ for $n = -N_\delta$ and $x_n = L + \delta$ for $n = N + N_\delta$

and let $u_n \approx u_\delta(x_n)$. Then, u_n for $n = -N_\delta, \dots, N + N_\delta$ is determined from

$$\begin{cases} -2 \displaystyle\sum_{n'=n-N_\delta}^{n+N_\delta} (u_{n'} - u_n)\gamma(x_n, x_{n'})h = b(x_n) \quad \text{for} \quad n = 1, \dots, N-1 \\ u_n = g(x_n) \quad \text{for} \quad n = -N_\delta, \dots, 0, N, \dots, N + N_\delta. \end{cases} \tag{2}$$

The reduced sparsity of the discretizations is now easily explained. For fixed δ, as h is reduced, the number of nodes $2N_\delta + 1$ that interact with a given node increases and the same occurs if h is fixed and δ increases. This issue is even more prominent in 2D and 3D problems where the computational complexity increases more rapidly with δ.

2. Multifidelity Monte Carlo Methods

In this section, we provide a brief account of the multifidelity Monte Carlo (MFMC) method of [6] in which a much more detailed presentation is given.

The goal is to determine approximate statistical information about an output of interest (OoI) $f^{(1)}(\mathbf{z}) : \mathcal{Z} \to \mathcal{F}$, where $\mathbf{z}$ is a random input selected from the input domain $\mathcal{Z} \subset \mathbb{R}^d$ and $\mathcal{F} \subset \mathbb{R}$ denotes the corresponding output domain. For the sake of concreteness, we consider the (statistical) quantity of interest (QoI) to be the expected value $\mathbb{E}[f^{(1)}(\mathbf{z})]$; other QoIs such as higher moments of the OoI can be treated in a similar manner. We approximate the QoI via Monte Carlo (MC) sampling, i.e., for M_{MC} independent and identically distributed random samples $\mathbf{z}_m \in \mathcal{Z}$, $m = 1, \dots, M_{MC}$, drawn from a given probability density function, we have

$$f^{MC}_{M_{MC}} = \frac{1}{M_{MC}} \sum_{m=1}^{M_{MC}} f^{(1)}(\mathbf{z}_m) \approx \mathbb{E}\left[f^{(1)}(\mathbf{z})\right]. \tag{3}$$

If $M^{MC} \gg 1$ and $f^{(1)}(\cdot)$ is expensive to evaluate, (3) may be prohibitive in cost.

Suppose we have in hand a set of lower-fidelity OoIs $f^{(2)}, \dots, f^{(K)} : \mathcal{Z} \subset \mathcal{F}$ with corresponding costs w_k, $k = 2, \dots, K$. Then, for a given set of weights $\alpha_2, \dots, \alpha_K$ and a given set of sample sizes $0 < M_1 \le M_2 \le \cdots \le M_k$, the unbiased MFMC estimator is defined as

$$f^{MFMC}_{M_{MF}} = f^{(1)}_{M_1} + \sum_{k=2}^{K} \alpha_k \left(f^{(k)}_{M_k} - f^{(k)}_{M_{k-1}} \right) \approx \mathbb{E}\left[f^{(1)}(\mathbf{z})\right], \tag{4}$$

where $f^{(k)}_{M_k}$ and $f^{(k)}_{M_{k-1}}$ denote the MC estimators for the OoI $f^{(k)}(\cdot)$, respectively, using M_k and M_{k-1} input samples drawn from the input domain $\mathcal{Z}$, i.e.,

$$f^{(k)}_{M_k} = \frac{1}{M_k} \sum_{m=1}^{M_k} f^{(k)}(\mathbf{z}_m) \quad \text{and} \quad f^{(k)}_{M_{k-1}} = \frac{1}{M_{k-1}} \sum_{m=1}^{M_{k-1}} f^{(k)}(\mathbf{z}_m). \tag{5}$$

Note that the M_{k-1} samples taken in the second sum can be reused in the first sum so that the total number of evaluations of $f^{(k)}(\cdot)$ for each term in (4) is M_k. Thus, the cost incurred to determine the MFMC estimator (4) is given by $\sum_{k=1}^{K} w_k M_k$.

For $k = 1, \dots, K$, let $\sigma_k^2 = \mathrm{Var}(f^{(k)}(\mathbf{z}))$, $\rho_{1,k} = \mathrm{Cov}(f^{(k)}(\mathbf{z}), f^{(1)}(\mathbf{z}))/(\sigma_k \sigma_1)$ = Pearson correlation coefficient. Then, the optimal values for the weights α_k and the number of samples M_k are determined, for a given budget p, by *minimizing the variance of* $f^{MFMC}_{M_{MF}}$ *subject to the constraints* $\sum_{k=1}^{K} w_k M_k = p$ *and* $0 < M_1 \leq M_2 \leq \cdots \leq M_k$. If one assumes that

$$|\rho_{1,1}| > \cdots > |\rho_{1,K}| \qquad \text{and} \qquad \mu_k = \frac{w_{k-1}}{w_k} > \nu_k = \frac{\rho_{1,k-1}^2 - \rho_{1,k}^2}{\rho_{1,k}^2 - \rho_{1,k+1}^2} \quad \text{for} \quad k = 2, \ldots, K \tag{6}$$

with $\rho_{1,K+1} = 0$ hold, this optimization problem has the unique analytic solution (see [6])

$$\alpha_k = \frac{\rho_{1,k}\sigma_1}{\sigma_k} \quad \text{and} \quad M_k = M_1 r_k \quad \text{for} \quad k = 2, \ldots, K \qquad \text{with} \quad M_1 = \frac{p}{\sum_{k=1}^{K} w_k r_k}, \tag{7}$$

where $r_k = \left(w_1(\rho_{1,k}^2 - \rho_{1,k+1}^2)/w_k(1-\rho_{1,2}^2)\right)^{1/2}$ for $k = 1, \ldots, K$. We note that the constraint $M_1 > 0$ ensures that the accuracy of the MFMC estimator corresponds to that of the truth model $f^{(1)}(\cdot)$, i.e., the accuracy is not damaged by the use of surrogate models.

The first requirement in (6) is easily satisfied by re-ordering the OoIs $f^{(k)}(\cdot)$. The second requirement is violated whenever the decrease in the accuracy of the low-fidelity model k (in the order of decreasing correlation coefficient) is more significant that the reduction of its cost of evaluation. The models that violate the second requirement are then eliminated from the set of models used in the MFMC estimator. The second requirement is defined by comparing each model to the preceding one in the order of decreasing correlation coefficient. One then continues checking the second requirement and eliminating the models that violate this requirement until all the remaining models satisfy the second requirement.

Note that (7) does not, in general, yield integer values for the number of samples M_k. In this study, we have chosen to round up to the nearest integer. Also, the values of σ_k and $\rho_{1,k}$ are generally not known a priori, so that in practice they are estimated by computing a very few realizations of the models $f^{(k)}(\cdot)$ for $k = 1, \ldots, K$, where "very few" is relative to the very large number of samples one would have to use if one were to use a straightforward MC estimator for $\mathbb{E}[f^{(1)}(\cdot)]$. Such realizations can even be reused to estimate the costs w_k of evaluating $f^{(k)}(\cdot)$ should those costs not be known.

3. A Numerical Example

Consider the discretized nonlocal diffusion model (2) with $L = 1$, a random constant source term b for $x \in (0, 1)$, the volume constraints $u(x) = g$ with g a random constant for $x \in [-\delta, 0]$, and $u(x) = 0$ for $x \in [1, 1 + \delta]$, and the constant kernel function $\gamma(x, x') = 1/2\delta^3$ for $|x - x'| < \delta$. The random inputs are independently, identically, and uniformly distributed realizations of $\mathbf{z} = (b, g)$ within the input domain $\mathcal{Z} = [-1.1 \; -0.9] \times [0.9 \; 1.1] \subset \mathbb{R}^2$. Note that the dependence of the kernel on δ is a scaling factor that ensures that as $\delta \to 0$, i.e., as the extent of nonlocal interactions vanish, the nonlocal problem reduces to its local differential equation counterpart $-d^2u/dx^2 = b$.[1]

The *high-fidelity* (or *truth*) quantity of interest (QoI) is the expected value of the output of interest (OoI) $f^{(1,1)}(\mathbf{z}) = \left(\int_\Omega u(x, \delta_1, h_1; \mathbf{z})^2 dx\right)^{1/2}$ for $\delta_1 = 0.25$ and $h_1 = 1/N_1 = 2^{-10}$. An approximation of the truth value of the QoI $\mathbb{E}[f^{(1,1)}(\mathbf{z})]$ is defined as the Monte Carlo (MC) estimator (4) using $M_{MC} = 5 \times 10^8$ samples of $(b, g) \in \mathcal{Z}$.

[1]In this study we have chosen, for validation purposes, the nonlocal model such that it converges to the local model in the limit of $\delta \to 0$. Yet this is not a requirement for the proposed multifidelity framework.

We define eight lower-fidelity, lower-cost surrogate estimators by using, in (2), two smaller horizon values $\delta_2 = 0.25/2$ and $\delta_3 = 0.25/2^2$ and/or two larger grid length values $h_2 = 2^{-9}$ and $h_3 = 2^{-8}$. The models are indexed using a pair of indices (i, j) with the first index representing the δ-model, and the second index the h-model, e.g., $f^{(i,j)}(\mathbf{z}) = \left(\int_\Omega u(x, \delta_i, h_j; \mathbf{z})^2 dx\right)^{1/2}$. Including the high-fidelity estimator, we specifically solve (2) for samples $\mathbf{z}_m = (b_m, g_m) \in \mathcal{Z}$ to obtain $f^{(i,j)}(\mathbf{z}_m)$ for $i, j = 1, \ldots, 3$. Letting $k = j + 3(i-1)$ for $i, j = 1, \ldots, 3$, the costs w_k and correlation coefficients $\rho_{1,k}$ for the surrogates are approximated by the averages over 50 samples of the random inputs; cost measurements are defined as wall-clock times. Using 500 or even 1000 samples to estimate costs and correlation coefficients results in similar numbers as given below for 50 samples. Table 1 summarizes the information on the models used, as well as the correlation coefficient of models $f^{(i,j)}$ with the truth model $f^{(1,1)}$, and the cost of each model. Algorithm 1 summarizes the computational procedure for a nonlocal multifidelity Monte Carlo method, given the high-fidelity horizon $\delta_1 = 0.25$ and grid length $h_1 = 2^{-10}$.

model	$f^{(1,1)}$	$f^{(1,2)}$	$f^{(1,3)}$	$f^{(2,1)}$	$f^{(2,2)}$	$f^{(2,3)}$	$f^{(3,1)}$	$f^{(3,2)}$	$f^{(3,3)}$
k	1	2	3	4	5	6	7	8	9
δ	0.25	0.25	0.25	0.25/2	0.25/2	0.25/2	$0.25/2^2$	$0.25/2^2$	$0.25/2^2$
h	2^{-10}	2^{-9}	2^{-8}	2^{-10}	2^{-9}	2^{-8}	2^{-10}	2^{-9}	2^{-8}
$\rho_{1,k}$	1.00000	0.99999	0.99994	0.99520	0.99459	0.99328	0.98882	0.98720	0.98378
w_k	0.07555	0.01121	0.00182	0.02497	0.00401	0.00069	0.00819	0.00147	0.00032
w_k/w_1	1.0000	0.1483	0.0241	0.3305	0.0531	0.0092	0.1084	0.0194	0.0043

Table 1: Summary information about the eight surrogate models and the truth model used in this study. The models are indexed using a pair of indices (i, j), where the first index corresponds to the δ-model with 1 corresponding to the truth horizon, and 2 and 3 corresponding to the smaller surrogate horizon, and the second index corresponds to the mesh refinement where 1 represents the finest (truth) grid, and 2 and 3 represent the coarser surrogates. The costs and correlation coefficients are approximated using 50 random samples. The index k used in rows 4 – 6 is equivalent to $k = j + 3(i-1)$.

We consider four specific cases. The first is simply the MC estimator (3) with the number of samples M_{MC} limited by a given budget p, i.e., the number of samples M_{MC} is determined by dividing budget p by the cost of model $f^{(1,1)}$ in Table 1. The next two cases illustrate the separate effects due to the use of either h or δ-dependent surrogates for which known hierarchies of costs and fidelities are known a priori, i.e., we are in the realm of multilevel Monte Carlo methods. The last case is a multifidelity setting for which such a hierarchy is not known a priori.

Case 1: $\delta = 0.25$ and $h = 2^{-10}$, i.e., the high-fidelity, high-cost (truth) model, $f^{(1,1)}$.

Case 2: $h = 2^{-10}$ is fixed at the truth value and $\delta_1 = 0.25$, $\delta_2 = 0.25/2$, and $\delta_3 = 0.25/2^2$ so that the first of these gives the truth model $f^{(1,1)}$ and the other two define cheaper and lower-fidelity surrogates $f^{(2,1)}$ and $f^{(3,1)}$, respectively. These three models satisfy both requirements in (6) in their current order ($f^{(1,1)}$, $f^{(2,1)}$, $f^{(3,1)}$). We refer to this case as the "three-δ case."

Case 3: $\delta = 0.25$ is fixed at the truth value and $h_1 = 2^{-10}$, $h_2 = 2^{-9}$, and $h_3 = 2^{-8}$ so that the first of these gives the truth model $f^{(1,1)}$ and the other two defining cheaper and lower-fidelity surrogates $f^{(1,2)}$ and $f^{(1,3)}$, respectively. Similar to Case 2, these three models satisfy both requirements in (6) in their current order ($f^{(1,1)}$, $f^{(1,2)}$, $f^{(1,3)}$). We refer to this case as the "three-h case."

Case 4: $\delta_1 = 0.25$, $\delta_2 = 0.25/2$, and $\delta_3 = 0.25/2^2$ and $h_1 = 2^{-10}$, $h_2 = 2^{-9}$, and $h_3 = 2^{-8}$ so that the first pair $\delta_1 = 0.25, h_1 = 2^{-10}$ gives the truth estimator $f^{(1,1)}$ and the other eight δ_i, h_j pairs give the eight lower-fidelity, lower-cost surrogates $f^{(i,j)}$, which together with the truth choice, are

Algorithm 1 Nonlocal Multifidelity Monte Carlo method

procedure NL-MFMC(high-fidelity model $f^{(1,1)}$ with horizon δ_1 and grid length h_1, and budget p)

  For a given $s > 1$, generate sequences of models $f^{(i,j)}$ with $\delta_i = \delta_1/s^{i-1}$, and $h_j = h_1/s^{j-1}$ where $i = 1, \cdots, n_\delta$, and $j = 1, \cdots, n_h$.

  Let $1 \leq k = j + n_h\,(i-1) \leq n_\delta n_h$, and estimate costs w_k, correlation coefficients $\rho_{1,k}$, and variances σ_k with a few realizations of model $f^{(k)}$.

  Re-order models $f^{(k)}$ to satisfy the first assumption of (6).

  Check the second assumption of (6), and eliminate models as needed. Repeat this step until no further elimination is required.

  For the $K \leq n_\delta n_h$ remaining models from the elimination process, set $\rho_{1,K+1} = 0$, and compute r_k for $k = 1, \ldots, K$

$$r_k = \left(w_1(\rho_{1,k}^2 - \rho_{1,k+1}^2)/w_k(1-\rho_{1,2}^2)\right)^{1/2}$$

  Determine coefficients α_k and number of samples M_k from (7).

  Round up the number of samples to the next integer, and draw $\lceil M_K \rceil$ realizations of $\mathbf{z}_m$.

  Evaluate the MFMC estimator f_M^{MFMC} from (4).

  return f_M^{MFMC}

end procedure

arranged in the order $f^{(1,1)}$, $f^{(1,2)}$, $f^{(1,3)}$, $f^{(2,1)}$, $f^{(2,2)}$, $f^{(2,3)}$, $f^{(3,1)}$, $f^{(3,2)}$, $f^{(3,3)}$. With this ordering, the surrogates satisfy the first requirement in (6), so that no re-ordering is required. With the models in the order of descending correlation coefficient, the second requirement is checked. It is observed that in the first round, models $f^{(2,1)}$, and $f^{(3,1)}$ violate the second requirement. After eliminating these models, the second requirement is checked again, where it is found that models $f^{(2,2)}$, and $f^{(3,2)}$ violate the second condition. The four remaining surrogates ($f^{(1,2)}$, $f^{(1,3)}$, $f^{(2,3)}$, $f^{(3,3)}$) satisfy both requirements in (6). Therefore, the number of surrogates is whittled down from 8 to 4. We refer to this case as the "five-δ, h case."

Discussion of results. For each surrogate and for the truth model, the number of samples taken for each model is determined from (7). The number of samples M_k are only approximately determined because, as already noted, the costs w_k and correlation coefficients $\rho_{1,k}$ are themselves only approximations of the true values. As a result, the actual total costs are not exactly equivalent to the assigned budget. Figure 1 depicts the distribution of the number of samples between the different models used in each of the four MFMC cases. From the $\rho_{1,k}$, w_k, and w_k/w_1 rows of Table 1, we observe that the pair of three-h surrogates $f^{(1,2)}$ and $f^{(1,3)}$ are better correlated to the truth model $f^{(1,1)}$ than are the pair of three-δ surrogates $f^{(2,1)}$, and $f^{(3,1)}$. The correlation coefficients are $0.99999, 0.99994$ for the three-h surrogates whereas they are $0.99520, 0.98882$ for the three-δ surrogates. The higher correlation between the three-h surrogates and the truth model means that a fewer number of samples of the truth model are needed compared to that for the three-δ surrogates. Of course, correlation is not the only factor that influences the efficiency of the MFMC method; the costs associated with the surrogates also enter into the fray, i.e., the total cost for an MFMC method is $\sum_{k=1}^{K} w_k M_k$. Once again, in this regard, the three-h surrogates prove superior to the three-δ surrogates that have costs w_k/w_1 relative to the truth model given by $0.1483, 0.0241$ and $0.3305, 0.1084$, respectively; see Table 1 and Figure 1.

We now turn to a more transparent comparison of the effectiveness of the three MFMC esti-

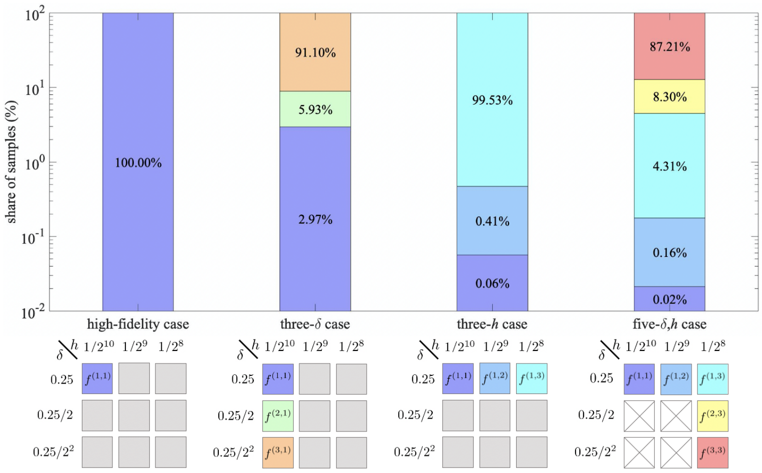

Figure 1: The figure shows the distribution of the total number of samples across different models used in each MFMC case. Note the logarithmic scale on the vertical axis. The gray boxes indicate surrogates that were not included for each particular case. The boxes with crosses indicate surrogates that were eliminated because they violated the second requirement of (6).

mators we have defined. For Figure 2, we consider the eight values $\log_{10} p = \{0,\ 0.5,\ 1,\ 1.5,\ 2,\ 2.5,\ 3,\ 3.5,\ 4,\ 4.5\}$ for the assigned budget. For each budget value, we determine, for all cases considered, an approximate mean-squared error (MSE) for the MFMC estimator $f^{MFMC}_{M_{MF}}$ given in (4), using the true value of the QoI as estimated by the Monte Carlo estimator $f^{MC}_{M_{MC}}$ given in (3) with $M_{MC} = 5 \times 10^8$ samples. Specifically, we estimate the MSE by averaging the square of the difference between the MC estimator and the MFMC estimators over ten independent runs of the latter.

Viewing the plots in Figure 2 vertically, we observe that for the same chosen budget value p, all three MFMC estimators, when compared to the use of only the high-fidelity estimator, result in better accuracy for the same costs, with the 3-δ estimators yielding about one order of magnitude and the 3-h and five-δ, h estimators yielding two or three orders of magnitude reductions in the error for the same cost. Viewing the plots horizontally, we observe that for a desired MSE, all three MFMC estimator result in lower costs, with again the 3-h and five-δ, h estimators yielding the largest gains.

We also observe from Figure 2 that the use of surrogate models defined for larger values of h outperforms the use of surrogate models defined for smaller values of δ. Most of the gain that the the five-δ, h achieves over the single-model case is due to the inclusion of different h surrogates relative to the gains affected by instead using different δ surrogates.

Future studies. Follow-up research take on several avenues. The relation between δ and h will be investigated, and, in particular, how to determine optimal choices for the sequences of h and δ so as to increase the gain from the use of multifidelity methods. Future work will also address higher dimensions in which the reduced sparsity from nonlocality has an even more significant effect on the computational complexity when compared to that for local PDE models. We note that this study is a proof of concept in which we investigate MFMCs in the context of the linear nonlocal diffusion model; more complex nonlinear nonlocal models such as the nonlocal Cahn-Hilliard model studied in [8] will also be the subject of the subsequent studies.

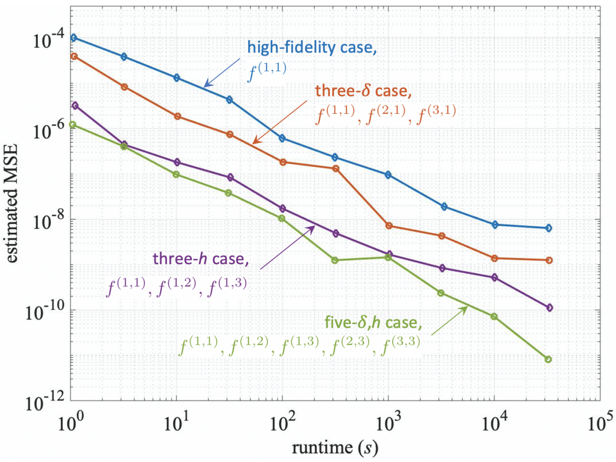

Figure 2: The estimated mean-squared errors vs. the budget p in wall-clock seconds for the three MFMC estimators and the MC estimator.

Acknowledgments. The authors The authors thankfully acknowledge Benjamin Peherstorfer for the advice and explanations he provided during research leading to this paper. The authors also thank Olena Burkovska for the advice she provided related to the coding aspects of the research. This work was supported by US Department of Energy grants DE-SC0019303 and DE-SC0021077 as part of the AEOLUS Multifaceted Mathematics Integrated Capability Center.

References

[1] B. Aksoylu, M. L. Parks, Variational theory and domain decomposition for nonlocal problems, *Applied Mathematics and Computation* 217 (2011) 6498–6515.

[2] M. Gunzburger, R. B. Lehoucq, A nonlocal vector calculus with application to nonlocal boundary value problems, *Multiscale Modeling & Simulation* 8 (2010) 1581–1598.

[3] Q. Du, M. Gunzburger, R. B. Lehoucq, K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints, *SIAM Review* 54 (2012) 667–696.

[4] Q. Du, M. Gunzburger, R. B. Lehoucq, K. Zhou, A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws, *Mathematical Models and Methods in Applied Sciences* 23 (2013) 493–540.

[5] M. D'Elia, Q. Du, C. Glusa, M. Gunzburger, X. Tian, Z. Zhou, Numerical methods for nonlocal and fractional models, *ACTA Numerica* 30, (2020)

[6] B. Peherstorfer, K. Willcox, M. Gunzburger, Optimal model management for multifidelity Monte Carlo estimation, *SIAM Journal on Scientific Computing* 38 (2016) A3163–A3194.

[7] B. Peherstorfer, K. Willcox, M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *SIAM Review* 60 (2018) 550–591.

[8] O. Burkovska, M. Gunzburger, On a nonlocal Cahn-Hilliard model permitting sharp interfaces, *arXiv:2004.14379* (2020).