

# Multifidelity Monte Carlo estimation for efficient uncertainty quantification in climate-related modeling

Anthony Gruber<sup>1,2</sup>, Max Gunzburger<sup>2,3</sup>, Lili Ju<sup>4</sup>, Rihui Lan<sup>5</sup>, and Zhu Wang<sup>4</sup>

<sup>1</sup>Center for Computing Research, Sandia National Laboratories, 1515 Eubank SE, Albuquerque, NM 87123, USA

<sup>2</sup>Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, USA

<sup>3</sup>Oden Institute for Engineering and Sciences, University of Texas, Austin, TX 78712, USA

<sup>4</sup>Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

<sup>5</sup>School of Mathematical Sciences, Ocean University of China, Qingdao, Shandong 266100, P.R. China

Correspondence: Anthony Gruber (adgrube@sandia.gov)

Received: 16 August 2022 – Discussion started: 14 September 2022 Revised: 17 November 2022 – Accepted: 20 January 2023 – Published: 21 February 2023

Abstract. Uncertainties in an output of interest that depends on the solution of a complex system (e.g., of partial differential equations with random inputs) are often, if not nearly ubiquitously, determined in practice using Monte Carlo (MC) estimation. While simple to implement, MC estimation fails to provide reliable information about statistical quantities (such as the expected value of the output of interest) in application settings such as climate modeling, for which obtaining a single realization of the output of interest is a costly endeavor. Specifically, the dilemma encountered is that many samples of the output of interest have to be collected in order to obtain an MC estimator that has sufficient accuracy - so many, in fact, that the available computational budget is not large enough to effect the number of samples needed. To circumvent this dilemma, we consider using multifidelity Monte Carlo (MFMC) estimation which leverages the use of less costly and less accurate surrogate models (such as coarser grids, reduced-order models, simplified physics, and/or interpolants) to achieve, for the same computational budget, higher accuracy compared to that obtained by an MC estimator - or, looking at it another way, an MFMC estimator obtains the same accuracy as the MC estimator at lower computational cost. The key to the efficacy of MFMC estimation is the fact that most of the required computational budget is loaded onto the less costly surrogate models so that very few samples are taken of the more expensive model of interest. We first provide a more detailed discussion about the need to consider an alternative to MC estimation for uncertainty quantification. Subsequently, we present a review,

in an abstract setting, of the MFMC approach along with its application to three climate-related benchmark problems as a proof-of-concept exercise.

*Copyright statement.* This article has been co-authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under contract no. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns right, title, and interest in and to the article and is responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https: //www.energy.gov/downloads/doe-public-access-plan (last access: 26 January 2023).

# 1 Introduction

In many application settings – climate modeling being a prominent one – large computational costs are incurred when solutions to a given model are approximated to within an acceptable accuracy tolerance. In fact, this cost can be prohibitively large when one has to obtain the results of multiple simulations, as is the case for, e.g., uncertainty quantification, control, and optimization, to name a few. Thus, there is often a need for compromise between the accuracy of simulation

algorithms and the number of simulations needed to obtain, say, in the uncertainty quantification setting, accurate statistical information.

For example, consider the following case, which represents the focus of this paper. Suppose one has a complex system, say, a system of discretized partial differential equations, for which the input data depends on a vector of randomly distributed parameters z. Letting u(z) denote the solution of the discretized partial differential equations, we define an output of interest F(u(z)) = F(z) that depends on u(z) so that, of course, it also depends on the choice of z. Next, suppose we wish to use Monte Carlo sampling to estimate the expected value  $Q = \mathbb{E}(F(z))$  of the output of interest F(z), i.e., we have the Monte Carlo estimator of Q given by

$$Q^{\mathrm{MC}} = \frac{1}{M} \sum_{m=1}^{M} F(z_m) = \frac{1}{M} \sum_{m=1}^{M} F(\boldsymbol{u}(z_m)) \approx \mathbb{E}(F(z))$$
$$= Q, \qquad (1)$$

where  $\{z_m\}_{m=1}^M$  denotes a set of M independent and identically distributed samples of the random vector z. This is a common task in uncertainty quantification, which provides useful statistical information in both predictive and inferential applications.

On the other hand, Monte Carlo estimation is not without its drawbacks. Consider the following scenario. Let  $\delta < 1$ denote a measure of the spatial grid size used to discretize the partial differential equation (PDE) system in question, and suppose that  $\delta$  is normalized by the diameter of the domain in which that system is posed. Then, we have it that the error in the approximate solution is of  $\mathcal{O}(\delta^{\nu_d})$  for some  $v_d > 0$  and that the cost (e.g., measured in seconds, days, or weeks) of a simulation which obtains the approximate solution is of  $\mathcal{O}(\delta^{-\nu_c})$  for some  $\nu_c > 0$ . Neglecting any additional costs connected with the evaluation of F(z) = F(u(z)) (for any given z) once u(z) is obtained, it is clear that the cost C incurred when obtaining the Monte Carlo estimator  $Q^{MC}$  is of  $\mathcal{O}(M\delta^{-\nu_c})$ . Of course, it is well known that the accuracy of a Monte Carlo estimator is of  $\mathcal{O}(1/\sqrt{M})$  so that in order for the accuracy of the estimator to be commensurate with the discretization error, we must have that  $1/\sqrt{M} \propto \mathcal{O}(\delta^{\nu_d})$ . Therefore, the needed number of samples is proportional to  $M \propto \delta^{-2\nu_{\rm d}}$ , and the total cost CM incurred when determining the Monte Carlo estimator  $O^{\text{MC}}$  is of  $\mathcal{O}(\delta^{-2\nu_{\rm d}-\nu_{\rm c}})$ .

The cost CM can quickly get out of control when dealing with large-scale problems. For example, suppose that the approximate solution of the PDE system is second-order accurate, i.e.,  $v_d = 2$ , and that a single simulation incurs a cost of  $\mathcal{O}(\delta^{-3})$  (i.e.,  $v_c = 3$ ), which is the best-case scenario in three dimensions. We then have that the number of samples needed is  $M \propto \delta^{-4}$ , and the total cost CM is  $\propto \delta^{-7}$ . Thus, with even a modest grid size of  $\delta = 0.01$ , we have it that the number of samples needed is  $M \propto 10^8$  and that the total cost is CM  $\propto 10^{14}$ , which is too high for practical use. Turning things around, suppose instead that the available computational budget allows for at most 10 thousand simulations, i.e.,  $M = 10^4$ , so that the resulting accuracy of the Monte Carlo simulator is of  $\mathcal{O}(10^{-2})$ . Here, the estimator error and approximation error are already commensurate when  $\delta = 0.1$ , making the choice of a smaller delta redundant if not infeasible. In fact, in this case, there is no way to achieve the four digits of accuracy sought by choosing  $\delta = 0.01$ , since the best we can do with the available budget is on the order of  $\mathcal{O}(10^{-2})$ . Note that the situation becomes even worse when the PDE system in question is time dependent, as a single simulation incurs an even larger cost when one accounts for the number of time steps used in the simulation.

Given that the cost of MC estimation is at times prohibitively expensive, it comes as no surprise that many alternatives or run-arounds to such estimation have been proposed. One approach in this direction has led to the development of many different random-parameter sampling schemes (e.g., quasi-Monte Carlo sampling, sparse-grid sampling, importance sampling, Latin hypercube sampling, lattice sampling, and compressed sensing, to name just a few) for which the estimation error is guaranteed to be smaller than its Monte Carlo equivalent; see, e.g., Addcock et al. (2022), Evans and Swartz (2000), Gunzburger et al. (2014), Nierderreiter (1992), Sloan (1994), and Smith (2013). On the other hand, some of these alternate approaches require smoothness of solutions to achieve better accuracy. Furthermore, most, if not all, of these methods are superior to Monte Carlo sampling only for the moderate dimension of the parameter vector z; see the references just cited.

A second approach towards reducing the cost of Monte Carlo (and, for that matter, for any type of) uncertainty estimation is to use approximate solutions of the PDE system that are less costly to obtain compared to the cost of obtaining the approximation of actual interest. For example, using simulations obtained using coarser grids or using reduced-order models such as reduced-basis or properorthogonal-decomposition methods are less costly, as are interpolation and support vector machine approximations; see, e.g., Cristianini and Shawe-Taylor (2000), Fritzen and Ryckelynck (2019), Keiper et al. (2018), Quarteroni and Rozza (2014), Quarteroni et al. (2016), and Steinwart and Christmann (2008). However, such approaches, by definition, result in less accurate approximations compared to the accuracy that one wants to achieve.

In this paper, we do not consider any of the possible alternate sampling schemes, nor do we exclusively consider using less costly and less accurate approximate solutions of the PDE system. Instead, because of the near ubiquity of its use in practice, our goal is to outperform traditional Monte Carlo estimation by using a nontraditional Monte Carlo sampling strategy and, in so doing, to refrain from incurring any loss of accuracy. To meet this goal, we invoke multifidelity Monte Carlo estimation which, in addition to the expensive and accurate PDE system approximation of interest (hereafter referred to as the "truth" approximation), also uses cheaper-

#### A. Gruber et al.: MFMC estimation for efficient UQ in climate modeling

to-obtain and less accurate approximations (which are referred to as the "surrogates"). The bottom line is that multifidelity Monte Carlo estimation meets our goal by leveraging increased sampling of the less accurate and less costly approximations alongside low sampling of the more expensive and more accurate truth approximation. The multifidelity Monte Carlo algorithm systematically determines the number of samples taken from each surrogate (i.e., there is no guess-work involved) and systematically (i.e., again, there is no guess-work involved) combines the samples of the surrogates to obtain the desired estimator. We note that multifidelity Monte Carlo estimation has already been shown to outperform Monte Carlo estimation in a variety of application settings; see, e.g., Clare et al. (2022), Dimarco et al. (2023), Konrad (2019), Law et al. (2022), Modderman (2021), Quick et al. (2019), Rezaeiravesh et al. (2020), Romer et al. (2020), and Valero et al. (2022) for some examples.

In Sect. 2, we review how the Monte Carlo and multifidelity Monte Carlo estimators are constructed in an abstract setting; here we follow the expositions in Peherstorfer et al. (2016) and also in Gruber et al. (2022a). Then, in Sects. 3 and 4, we demonstrate the effectiveness of multifidelity Monte Carlo estimation using three well-known climate-related benchmarks. Section 3.1 and 3.2 are respectively devoted to single-layer shallow-water equation models based on the SOMA test case of Wolfram et al. (2015) and Test Case 5 of Williamson et al. (1992). Section 4 is devoted to a benchmark case for the first-order model for ice sheets; see Blatter (1995), Pattyn (2003), and Perego et al. (2012), Tezaur et al. (2015). We close by providing some concluding remarks in Sect. 5.

# 2 Monte Carlo and multifidelity Monte Carlo estimators

An abstraction of the specific settings considered in Sects. 3 and 4 is given as follows:

- It involves having in hand a (discretized) partial differential equation (PDE) system for which the solution u(z) depends on a random vector of parameters  $z \in \Gamma$ , where  $\Gamma$  denotes a parameter domain.

Note that the input data to this PDE system (e.g., forcing terms, initial conditions, and coefficients) could depend on one or more of the components of the random vector z.

- We are interested in situations whereby, for any  $z \in \Gamma$ , obtaining u(z) is a costly endeavor.
- We define a scalar output of interest (OoI)  $F^1(z) = F^1(u(z))$  that depends on the solution u(z) of the (discretized) PDE system; of course, if obtaining u(z) is costly, then so is obtaining  $F^1(z)$ .

While they are not considered in this work, vectorvalued outputs of interest can also be treated with multifidelity Monte Carlo techniques. OoIs could be, e.g., averages or extremal values of the energy associated with the solution u(z).

- Having defined an OoI  $F^1(z)$ , we let  $Q_1 = \mathbb{E}(F^1)$  denote the quantity of interest (QoI) corresponding to the  $F^1(z)$ , where  $\mathbb{E}(\cdot)$  denotes the expected value with respect to *z*.
- Because the estimation of  $Q_1 = \mathbb{E}(F^1)$  is the central goal of this paper, we refer to  $F^1(z)$  as the "truth" output of interest.
- Commonly, even ubiquitously, a Monte Carlo (MC) sampling method is used to (approximately) quantify the uncertainty in the chosen OoI  $F^1(z)$ .
- Specifically, MC sampling is used to estimate the quantity of interest  $Q_1 = \mathbb{E}(F^1)$  corresponding to  $F^1(z)$ , i.e., we have it that the MC estimator of the QoI  $Q_1 = \mathbb{E}(F^1)$  is given by

$$Q_1^{\rm MC} = \frac{1}{M_1} \sum_{m=1}^{M_1} F^1(z_m) \approx Q_1 = \mathbb{E}\left(F^1\right),\tag{2}$$

where  $\{z_m\}_{m=1}^{M_1}$  denotes a set of  $M_1$  randomly selected points in the parameter domain  $\Gamma$ .

- We are then faced with the following dilemma: on the one hand, obtaining an acceptably accurate MC estimator  $Q_1^{\text{MC}}$  of the QoI  $Q_1$  requires obtaining the solution u(z) of the discretized PDE system at many randomly selected points  $z \in \Gamma$ ; on the other hand, each of those approximate solutions u(z) are computationally costly to obtain.

Quantifying uncertainties in climate system settings are victimized by this two-headed dilemma to the extent that, e.g., accurate long-time integrations often cannot be realized in practice.

Due to the issue of prohibitive computational cost, we turn to multifidelity Monte Carlo (MFMC) methods for uncertainty quantification. MFMC methods leverage the availability of surrogate outputs of interest  $F^k(z)$ , k = 2, ..., K, which have smaller computational complexity compared to that of the truth OoI  $F^1(z)$ . As mentioned in Sect. 1, there are many types of surrogates that can be used for this purpose, e.g., discretized PDEs with coarser spatial grids and larger time steps, reduced-basis and proper-orthogonal-decompositionbased reduced-order models, and interpolants of  $F^1(z)$ , to name just a few. For the truth and the K - 1 surrogate OoIs, Monte Carlo (MC) sampling yields the K unbiased estimators

$$Q_k^{\text{MC}} = \frac{1}{M_k} \sum_{m=1}^{M_k} F^k(z_m) \approx \mathbb{E}\left(F^k\right)$$
$$= Q_k \quad \text{for } k = 1, \dots, K,$$
(3)

where  $z_1, ..., z_{M_k}$  denote  $M_k$  independent and identically distributed (i.i.d.) samples of  $z \in \Gamma$ . The goal of an MFMC method is to use an appropriate linear combination of the *K* MC estimators  $Q_k^{\text{MC}}$  in Eq. (3) to estimate the QoI  $Q_1 = \mathbb{E}(F^1)$ . Specifically, we define the MFMC estimator as

$$Q^{\text{MFMC}} = Q_1^{\text{MC}} + \sum_{k=2}^{K} \alpha_k \left( Q_k^{\text{MC}} - Q_{k-1}^{\text{MC}} \right) \approx Q_1$$
$$= \mathbb{E} \left( F^1 \right), \tag{4}$$

where  $\{\alpha_k\}_{k=2}^K$  denotes a sequence of scalar weights, and  $\{M_k\}_{k=1}^K$  denotes a nondecreasing sequence of integers defining the number of samples.

Letting  $C_k$  denote the cost of evaluating the *k*th output of interest  $F^k$ , the costs of computing the respective MC and MFMC estimators are given as

$$C_k^{\text{MC}} = C_k M_k$$
 for  $k = 1, ..., K$  and  
 $C^{\text{MFMC}} = \sum_{k=1}^K C_k^{\text{MC}} = \mathbf{C} \cdot \mathbf{M},$ 

where  $\mathbf{C} = \{C_k\}_{k=1}^K$  and  $\mathbf{M} = \{M_k\}_{k=1}^K$  denote *K* vectors formed from the previously introduced sequences.

The variance  $\sigma_k^2$  of the *k*th output of interest  $F^k$  and the correlation  $\zeta_{k,k'}$  between the outputs of interest  $F^k$  and  $F^{k'}$  are given by, for k', k = 1, ..., K,

$$\sigma_k^2 = \operatorname{Var}\left[F^k(z)\right] \quad \text{and} \quad \zeta_{k,k'} = \frac{\operatorname{Cov}\left[F^k(z), F^{k'}(z)\right]}{\sigma_k \sigma_{k'}},$$

respectively, where  $\text{Cov}[\cdot, \cdot]$  denotes the statistical covariance. We then have that the mean-squared error (MSE) incurred by the MC estimator  $Q_k^{\text{MC}}$  of the QoI  $Q_k$  is given by

$$e(Q_k^{\mathrm{MC}}) = \mathbb{E}\left(\left(Q_k - Q_k^{\mathrm{MC}}\right)^2\right)$$
$$= \mathbb{E}\left(\left(\mathbb{E}\left(F^k(z)\right) - Q_k^{\mathrm{MC}}\right)^2\right)$$
$$= \mathbb{E}\left(F^k(z)\right)^2 - \mathbb{E}\left(F^k(z)^2\right)$$
$$= \frac{\sigma_k^2}{M_k} \quad \text{for } k = 1, ..., K,$$
(5)

whereas the MSE incurred by the  $Q^{\text{MFMC}}$  estimator of the QoI,  $Q_1$  is given by (see Peherstorfer et al., 2016)

$$e\left(Q^{\text{MFMC}}\right) = \frac{\sigma_1^2}{M_1} + \sum_{k=2}^{K} \left(\frac{1}{M_{k-1}} - \frac{1}{M_k}\right) \\ \left(\alpha_k^2 \sigma_k^2 - 2\alpha_k \sigma_k \sigma_1 \zeta_{1,k}\right).$$
(6)

Note that it can be shown (see, e.g., Peherstorfer et al., 2016) that the MSE for  $Q^{\text{MFMC}}$  is lower than that for  $Q_1^{\text{MC}}$  if and only if

$$\sqrt{\frac{e(Q^{\text{MFMC}})}{e(Q_1^{\text{MC}})}} = \sum_{k=1}^{K-1} \sqrt{\frac{C_k}{C_1} \left(\zeta_{1,k}^2 - \zeta_{1,k+1}^2\right)} < 1.$$
(7)

Given a fixed computational budget *B*, MFMC aims to construct an optimal sampling strategy  $\mathbf{M} = \{M_k\}_{k=1}^K$  along with an optimal set of weights  $\boldsymbol{\alpha} = \{\alpha_k\}_{k=2}^K$  so that the MSE (Eq. 6) of the multifidelity estimator  $Q^{\text{MFMC}}$  is lower than the MSE (Eq. 5) of the Monte Carlo estimator  $Q_1^{\text{MC}}$ . Viewed differently, this means that an appropriate estimator  $Q^{\text{MFMC}}$  can achieve a fixed MSE  $\varepsilon > 0$  at a smaller computational cost compared to that incurred for the Monte Carlo estimator  $Q_1^{\text{MC}}$  achieving MSE  $\varepsilon$ .

In Peherstorfer et al. (2016, 2018), unique optimal values of  $\mathbf{M} = \{M_k\}_{k=1}^K$  and  $\boldsymbol{\alpha} = \{\alpha_k\}_{k=2}^K$  are analytically obtained by minimizing the MSE (Eq. 6) of the  $Q^{\text{MFMC}}$  estimator (Eq. 4) over the real numbers. However, the values  $\mathbf{M}$  must certainly be integers for practical use, and the heuristic of Peherstorfer et al. (2016) which determines the optimal values can result in either a biased estimator of the expectation or (with naïve modification) a violation of the given budget *B*. For "small" computational budgets, the consequences of this can be quite severe, as illustrated in Gruber et al. (2022a).

Because "small" computational budgets are of high interest for climate modeling, here, instead of using the MFMC method of Peherstorfer et al. (2016, 2018), we use the modified MFMC method of Gruber et al. (2022a), which guarantees that the optimal sampling numbers are integers and that the computational budget is not exceeded. In that method, instead of simply minimizing the MSE  $e(Q^{\text{MFMC}})$  defined in Eq. (6), the modified MFMC estimator is determined through the use of at least some of the sequential minimization problems given by

for 
$$k = 1, 2, ..., K$$
 and, if  $k > 1$ ,  
for given  $M_{1,1}, ..., M_{k-1,k-1}$ ,  
minimize the functional

$$\mathcal{L}^{k}(\mathbf{M}_{k}, \boldsymbol{\alpha}_{k}; \lambda_{k}, \boldsymbol{\mu}_{k}, \xi_{k}) = e\left(\mathcal{Q}^{\text{MFMC}}\right) + \lambda_{k}\left(\sum_{k'=k}^{K} C_{k'}M_{k',k} - \left(B - \sum_{k'=1}^{k-1} C_{k'}\right)\right) + \sum_{k'=k+1}^{K} \mu_{k',k}\left(M_{k',k} - M_{k'-1,k}\right) - \xi_{k}M_{k,k},$$
(8)

#### https://doi.org/10.5194/gmd-16-1213-2023

Geosci. Model Dev., 16, 1213–1229, 2023

#### A. Gruber et al.: MFMC estimation for efficient UQ in climate modeling

where  $\mu_{k',k}$  for k' = k+1, ..., K,  $\lambda_k$ , and  $\xi_k > 0$  are Lagrange multipliers.

In Eq. (8), the first term added to the MSE  $e(Q^{\text{MFMC}})$  enforces the budget constraint – or, more precisely, for each k, enforces that the remaining available budget suffices to not exceed the given budget B after that budget has been depleted by the sampling already effected for the OoIs  $F^{k'}$ , k' = 1, ..., k - 1, where, of course, for k = 1 the whole budget B is available. The second term added in Eq. (8) enforces the monotone non-decrease in the number of samples  $M_{k',k}$ for k' > k. The third term added to the MSE enforces the positivity of  $M_{k,k}$ .

Repeating the arguments made in Peherstorfer et al. (2016) concerning optimal choices for the sample numbers and weights, the results given in Box 1 below are proved in Gruber et al. (2022a).

It is remarkable that, due to Eq. (11), the weights  $\alpha_{k',k}^*$  depend only on input data, specifically on the variances and correlations of the OoIs  $\{F^{k'}\}_{k'=1}^{K}$ . As such, the value of  $\alpha_{k',k}^*$  is independent of k so that

$$\alpha_{k',k}^* = \frac{\zeta_{1,k'}\sigma_1}{\sigma_{k'}} = \alpha_{k',1}^* \quad \text{for all } k = 1, \dots, K$$
  
and all  $k' = k + 1, \dots, K$ .

Thus, the set of weights  $\{\alpha_{k',1}^*\}_{k'=2}^K$  determined from the minimization of  $\mathcal{L}^1$  suffices to determine  $\alpha_{k',k}^*$  for all such k' and k; i.e., all  $\alpha_{k',k}^*$  for all k' and k are determined once and for all by the minimization of  $\mathcal{L}^1$ .

Unfortunately, as was the case in Peherstorfer et al. (2016, 2018), the results given in Box 1 do not immediately lead to a practical MFMC method because the sample numbers  $M_{k'k}^*$  given in Eqs. (10) and (11) are not, in general, integers. Moreover, as was also the case in Peherstorfer et al. (2016, 2018), the first and perhaps even the first few sampling numbers at stage k' = 1 may be < 1 in addition to being non-integer. In those papers, a rounding procedure is implemented; i.e., the sample numbers  $M_{k',1}$  are replaced by integers, though this is unsuitable for scenarios where  $M_{1,1} < 1$ as the choice  $M_{1,1} = 0$  leads to a biased estimator while the choice  $M_{1,1} = 1$  exceeds the computational budget. On the other hand, the modified MFMC method in Box 2 below is constructed to avoid these issues while preserving the optimality of the original solution (up to some amount of rounding). Note that in that box and elsewhere,  $|\cdot|$  denotes rounding downwards to the nearest integer.

# 3 Tests for the single-layer rotating shallow-water equations

Consider the single-layer rotating shallow-water equations (RSWEs) posed on the domain  $\Gamma \times [0, T]$  and given by (see Vallis, 2012)

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\boldsymbol{u}) = 0,$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{k} \cdot \nabla \times \boldsymbol{u} + f)(\boldsymbol{k} \times \boldsymbol{u})$$

$$+ \nabla \left(\frac{|\boldsymbol{u}|^2}{2} + g(h + h_{\rm b})\right) = G(h, \boldsymbol{u}),$$
(12)

where  $\Gamma$  denotes the surface of a sphere or a subset of that surface, [0, T] denotes a time interval; k denotes a unit vector perpendicular to the surface of the sphere; h(x, t) denotes the fluid thickness; u(x, t) denotes the vector velocity field tangential to the surface of the sphere; G(h, u) denotes the a forcing term that depends on the specific setting; f denotes the Coriolis parameter;  $h_b(x, t)$  denotes the bottom topography; g denotes the (constant) acceleration due to gravity; and  $\nabla$  denotes the tangential gradient – i.e.,  $\nabla f = Df - (Df \cdot k)k$ where D is the derivative operator of  $\mathbb{R}^3$ .

Note that the expression  $(\mathbf{k} \cdot \nabla \times \mathbf{u} + f)/h$  that appears in the velocity equation is known as the potential vorticity; see Vallis (2012). Supplementing Eq. (12) are the initial conditions  $h = h_0$  and  $\mathbf{u} = \mathbf{u}_0$  at t = 0; and, if  $\Gamma$  is a strict subset of the surface of the sphere, the boundary condition  $\mathbf{u} \cdot \mathbf{n} = 0$ on the boundary  $\partial \Gamma$  of  $\Gamma$ , where  $\mathbf{n}$  denotes the unit vector tangent to the surface of the sphere and also (outwardly) perpendicular to  $\partial \Omega$ .

The RSWEs represent a useful simplification of the primitive equations (Vallis, 2012) which are commonly used in oceanic and atmospheric modeling and which are obtained by assuming a small ratio between the vertical and horizontal length scales. In this way, the single-layer RSWEs describe the motion of a thin layer of fluid which lies on a rigid surface, yielding conditions used extensively in the modeling of oceanic and atmospheric flows.

Spatial discretization of the system (Eq. 12) is effected using the TRiSK scheme (Ringler et al., 2010; Thuburn et al., 2009), which is a staggered C-grid mimetic finite difference– finite volume scheme that preserves desirable physical properties, including conservation laws for mass, energy, and potential vorticity. TRiSK discretization involves the approximation  $h_{\ell}$  of the thickness h at the center of a grid cell  $P_{\ell}$ , the approximation  $u_e$  of the normal to the edge component  $u \cdot n$  of velocity at the centers of the edges of  $P_{\ell}$ , and the approximation of the potential vorticity  $(\mathbf{k} \cdot \nabla \times \mathbf{u} + f)/h$  at the vertices of  $P_{\ell}$ . The necessary meshing is done using spherical centroidal Voronoi tessellation (SCVT) grids (Jacobsen et al., 2013; Ringler et al., 2008; Yang et al., 2018); examples of such grids are given in Sect. 3.1 and 3.2 for the specific settings of those sections.

Temporal discretization of the system (Eq. 12) is effected using an explicit fourth-order Runge–Kutta method, although many alternative time-stepping schemes have also been proposed for this purpose; see, e.g., Leng et al. (2019), Meng et al. (2020), and Trahan and Dawson (2012). We de-

- Let  $\{F^k\}_{k=1}^K$  denote the set of computational outputs of interest with correlation coefficients  $\{\zeta_{1,k}\}_{k=2}^K$  and computational costs  $\{C_k\}_{k=1}^K$  that respectively satisfy

$$|\zeta_{1,k-1}| > |\zeta_{1,k}|$$
 and  $\frac{C_{k-1}}{C_k} > \frac{\zeta_{1,k-1}^2 - \zeta_{1,k}^2}{\zeta_{1,k}^2 - \zeta_{1,k+1}^2}, \quad k = 2, \dots, K.$  (9)

The first requirement is easily satisfied by a reordering of the surrogate outputs of interest  $F^2, \ldots, F^K$ . If, after that reordering, the second requirement is not satisfied for some  $k \in \{2, \ldots, K\}$ , then that  $F^k$  is removed from the list of surrogates.

- Let  $M_{1,1}, \ldots, M_{k-1,k-1}$  be given; these are all positive by construction.
- Also let  $\mathbf{M}_{k}^{*} = \{M_{k',k}^{*}\}_{k'=k}^{K}$  and  $\boldsymbol{\alpha}_{k}^{*} = \{\alpha_{k',k}^{*}\}_{k'=k+1}^{K}$ , and let

$$r_{k',k}^* = \frac{M_{k',k}^*}{M_{k,k}^*} = \sqrt{\frac{C_k}{C_{k'}} \left(\frac{\zeta_{1,k'}^2 - \zeta_{1,k'+1}^2}{\zeta_{1,k}^2 - \zeta_{1,k+1}^2}\right)}, \quad k' = k, \dots, K.$$

- Then, for each k = 1, ..., K, the unique global minimizer  $(\mathbf{M}_{k}^{*}, \boldsymbol{\alpha}_{k}^{*})$  of the functional  $\mathcal{L}^{k}$  is given by

$$M_{k,k}^{*} = \frac{B - \sum_{k'=1}^{k-1} C_{k'}}{\sum_{k'=k}^{K} C_{k'} r_{k',k}},\tag{10}$$

and

$$M_{k',k}^* = M_{k,k}^* r_{k',k}^* \quad \text{and} \quad \alpha_{k',k}^* = \frac{\zeta_{1,k'} \sigma_1}{\sigma_{k'}}, \quad k' = k+1, \dots, K.$$
(11)

# Box 1. Optimal non-integer sampling numbers for the modified MFMC method; see Gruber et al. (2022a).

For k = 1, 2, ...,

Ia. if the minimization of the Lagrangian functional  $\mathcal{L}^k$  constrained by  $M_{1,1} = M_{2,2} = \cdots = M_{k-1,k-1} = 1$  results in all K - k + 1 members of the set  $\{M_{k',k}\}_{k'=k}^K$  satisfying  $M_{k',k} \ge 1, k' = k, \dots, K$ , then

Ib. simply rounding downwards to the nearest integer produces integer sample numbers  $M_{k',k}^* \to \lfloor M_{k',k}^* \rfloor$ , k' = k, ..., K, so that the final set  $\{M_{k',k'} = 1\}_{k'=1}^{k-1} \cup \{\lfloor M_{k',k}^* \rfloor\}_{k'=k}^K$  of sample numbers are optimal under the constraint  $M_{1,1} = \cdots = M_{k',k'} = 1$  and *preserve the computational budget*. We then set  $\widehat{K} = k - 1$  and exit to step IV. Otherwise,

IIa. if  $M_{k,k}^* < 1$ , then it is rounded upwards  $M_{k,k}^* \to 1$ , and the function  $\mathcal{L}^{k+1}$  is minimized to obtain the remaining K-k components of  $M_{k',k}^*$ , k' = k + 1, ..., K. This yields

IIb. a new set  $\{M_{1,1}^* = 1, M_{2,2}^* = 1, \dots, M_{k,k}^* = 1\} \cup \{M_{k',k}^*\}_{k'=k+1}^K\}$  which may or may not contain entries less than 1. If it does not, the model evaluation vector becomes  $\mathbf{M}^* = \{M_{k',k'} = 1\}_{k'=1}^k \cup \{\lfloor M_{k',k}^* \rfloor\}_{k'=k+1}^K$ . We then set  $\hat{K} = k$  and exit to step IV. Otherwise,

- III. we increment  $k \rightarrow k + 1$  and return to step Ia.
- IV. The process terminates, and the final set of sample numbers is given by  $\{M_{1,1}^* = 1, M_{2,2}^* = 1, \dots, M_{\widehat{K},\widehat{K}}^* = 1\} \cup \{M_{k',\widehat{K}+1}^*\}_{k'=\widehat{K}+1}^K$ .

Box 2. Practical near-optimal integer sampling numbers for the modified MFMC method; see Gruber et al. (2022a).

 Table 1. Data corresponding to the three SCVT meshes used in the wind-driven gyre experiment.

Grid resolution		Number of cells	Number of edges	Number of vertices
8 km	$\Rightarrow$	120 953	364 124	243 172
16 km	$\Rightarrow$	30 217	91 285	61 069
32 km	$\Rightarrow$	8521	25 898	17 378

note by  $\{t_n\}_{n=0}^{N_t}$  with  $t_0 = 0$  and  $t_{N_t} = T$  the time instants used for temporal discretization.

#### 3.1 A wind-driven gyre test case for the RSWE system

The specific setting considered here is a modification of the benchmark test case referred to as "simulating ocean mesoscale activity" (SOMA; Wolfram et al., 2015), which involves a geodesic basin with radius 1250 km centered at latitude–longitude  $(35^\circ, 0^\circ)$  on the surface of the Earth. Inside the basin, the depth of fluid varies from 2500 m at the center to 100 m on the coastal shelf, creating a realistic topography which yields interesting dynamical behaviors which are useful for studying the propagation of fronts and eddies. The present experiment considers this same setting with a constant density  $\rho$ , leading to a computational model for a wind-driven barotropic gyre.

In Eq. (12),

$$G(h, \boldsymbol{u}) = d_{\text{bottom}}(h, \boldsymbol{u}) + d_{\text{wind}}(h)$$
  
with 
$$\begin{cases} d_{\text{bottom}}(h, \boldsymbol{u}) = -c_{\text{bottom}} \frac{|\boldsymbol{u}|\boldsymbol{u}|}{\rho h} \\ d_{\text{wind}}(h) = \frac{\tau_{\text{wind}}}{\rho h}, \end{cases}$$

with  $f_{\text{bottom}}(h, \boldsymbol{u})$  denoting a forcing term due to the bottom drag and  $f_{\text{wind}}(h)$  denoting a forcing term due to wind drag. Here,  $c_{\text{bottom}}$  denotes the bottom-drag coefficient chosen to be  $10^{-3}$ ,  $\rho$  denotes the density, and  $\tau_{\text{wind}}$  denotes the surface wind stress. See, e.g., Wolfram et al. (2015) for a discussion of this choice for  $G(h, \boldsymbol{u})$ .

For constructing the MC and MFMC estimators, we use three SCVT grids of the SOMA domain  $\Gamma$  listed in Table 1. A representative SCVT meshing of  $\Gamma$  is illustrated in Fig. 1. Note that for the SCVT grids used by the TRiSK scheme, the cells are almost all hexagonal, with a few pentagons and heptagons thrown in.

The output of interest we consider is the maximum fluid layer thickness, which is a common benchmark in oceanic RSWE simulations and which is relevant to, e.g., the detection of phenomena such as flooding. In particular, we simulate the RSWE until the final time of T = 3 d using the finest grid (8 km resolution) of 120 953 cells, and then we choose the OoI  $F_{gyre}^1$  given by

$$F_{\text{gyre}}^{1} = \max_{\ell=1,\dots,120953} h_{\ell},\tag{13}$$

where  $h_{\ell}$  denotes the fluid thickness at the center of the cell  $P_{\ell}$  at the final time T = 3.

The quantity of interest we choose considers the effect that perturbations of the initial velocity  $u_0 = u(0)$  have on  $F_{gyre}^1$  and how that effect can be quantified using MC and MFMC estimation. To this end, consider random dilations  $(1 + z)u_0$  of the initial velocity depending on the i.i.d. random variable z that is uniformly distributed over the interval [-0.5, 0.5]. Then, the OoI defined in Eq. (13) depends on the choice of z in that interval; i.e., we have it that  $F_{gyre}^1 = F_{gyre}^1(z)$ . In particular, we choose the QoI to be the expectation  $Q_{1,gyre} = \mathbb{E}\left(F_{gyre}^1\right)$  of the output of interest  $F_{gyre}^1(z)$  defined in Eq. (13).

Note that in practical RSWE simulations, as is the case for more sophisticated models such as the primitive equations, an approximation of the initial data  $u_0$  is often obtained from a pre-processing procedure in which the RSWE system is spun-up from rest, i.e., from a zero initial condition for u, up to some specified time; see Anderson et al. (1975) and Bleck and Boudra (1986). This procedure is invoked so as to eliminate transient artifacts which are not present in the current ocean or atmosphere and produces an initial configuration which is closer to observed oceanic and atmospheric data. The outcome of the spin-up calculation over the spin-up time frame of 15 d is illustrated in Fig. 2. The initial conditions  $u_0$  and  $h_0$  that supplement Eq. (12) are then simply set to the outcome of this pre-processing step, i.e., to the fluid thickness and velocity obtained at the end of the spin-up calculation.

# 3.1.1 MC and MFMC estimators

The MC estimator  $Q_{1,gyre}^{MC}$  of  $Q_{1,gyre} = \mathbb{E}\left(F_{gyre}^{1}\right)$  is given by Eq. (2). Unfortunately, obtaining acceptable accuracy using an MC estimator suffers from a double shortcoming. First, for any given *z*, obtaining  $F_{gyre}^{1}$  is a costly endeavor because it requires the solution of the discretized RSWE system to obtain the necessary 120 953 values of  $h_{\ell}(z)$ . Second, to obtain an MC estimator that is of acceptable accuracy requires obtaining  $F_{gyre}^{1}$  for many sample values of *z*.

Naturally, to mitigate this double shortcoming, we turn to the MFMC estimator described in Box 2 of Sect. 2. To do so, we define three surrogate OoIs,  $F_{gyre}^2(z)$ ,  $F_{gyre}^3(z)$ , and  $F_{gyre}^4(z)$ , all three of which are less costly to obtain compared to that of  $F_{gyre}^1(z)$ . Here,  $F_{gyre}^2(z)$  and  $F_{gyre}^3(z)$  are simply based on solving the discretized RSWE system using the coarser 16 and 32 km grids, respectively. The third surrogate,  $F_{gyre}^4(z)$ , is the piecewise-linear interpolant based on the values of  $F_{gyre}^1(z)$ , which are exact at the three points  $z = \{-0.5, 0, 0.5\}$ . These four OoIs constitute a reasonable multifidelity ensemble of models that are available during computational ocean studies, even more realistic ones such as primitive equation models. Moreover, all four OoIs can



Figure 1. The 32 km grid used for the wind-driven gyre test case and a zoom-in on a portion of that mesh.



Figure 2. For the setting of Sect. 3.1, the RSWE truth model solution (8 km resolution) with thickness h (a) and velocity field u (b) after integration of the system from rest for T = 15 d.

be leveraged for MFMC estimation, whereas more traditional MC and multilevel MC estimators only make use of the first or the first three OoIs, respectively. Approximations of the costs and correlations for the four OoIs  $\{F_{gyre}^k(z)\}_{k=1}^4$ are obtained by considering 100 uniform i.i.d. samples of  $z \in [-0.5, 0.5]$ . These are computed to be

$$\mathbf{C} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \begin{pmatrix} 101.1 \\ 12.83 \\ 1.714 \\ 0.05 \end{pmatrix}$$
  
and 
$$\boldsymbol{\zeta}_1 = \begin{pmatrix} \zeta_{1,1} \\ \zeta_{1,2} \\ \zeta_{1,3} \\ \zeta_{1,4} \end{pmatrix} = \begin{pmatrix} 1.00000000 \\ 0.99975045 \\ 0.99975920 \\ 0.99974835 \end{pmatrix}, \quad (14)$$

where  $C_k$  for k = 1, 2, 3 denotes the average computation time (in wall-clock seconds) necessary to advance the relevant discretized RSWE system by one time step, computed using 500 time steps for the simulation parameter z = 0.001. Note that the cost  $C_4$  is assigned arbitrarily because the cost of evaluating the interpolant is negligible. These costcorrelation pairs are ideal for MFMC estimation; i.e., the surrogates are very well correlated with  $F_{gyre}^1(z)$  and are much less costly to obtain compared to  $F_{gyre}^1(z)$ . Note that the 100 samples of z used to determine the data in Eq. (14) can be reused when determining the MC and MFMC estimators.

From Eq. (14), we observe that the more expensive surrogate  $F_{gyre}^2(z)$  is slightly less correlated to  $F_{gyre}^1(z)$  than the cheaper surrogate  $F_{gyre}^3(z)$  is. So, to satisfy the first criteria in Eq. (9), the correlations should be ordered in decreasing order:  $\zeta_{1,1}, \zeta_{1,3}, \zeta_{1,2}$ , and  $\zeta_{1,4}$ . However, it turns out that the second criteria in Eq. (9) for k = 2 is then not satisfied; thus,  $F_{gyre}^2(z)$  is removed from the list of surrogates when computing the MFMC estimator. That estimator is given by Eq. (4) with K = 4 and where k = 2 is excluded from the sum. This is an example which illustrates, as discussed in Sect. 2, that not all surrogates one chooses contribute to the efficiency of MFMC estimation and that, on the other hand, their omission does no harm to the accuracy of that estimator.

As already alluded to, the goal is to compare, for the same budget, the MSEs of the approximations  $Q_{gyre}^{MC}$  and  $Q_{gyre}^{MFMC}$ to the exact QoI  $Q_{1,gyre} = \mathbb{E}(F^1)$  produced by the MC and MFMC estimators. However, there is still an additional challenge to be dealt with, namely that it is unclear how to choose a reference QoI that one can use to determine these MSEs, as only limited data (200 high-fidelity samples) are available, and an MC approximation  $Q_{gyre}^{MC}$  with 200 samples is not sufficient for an accurate estimate of the expectation  $\mathbb{E}(F^1)$ . Therefore, we choose  $Q_{\text{gyre}}^{\text{ref}}$  to denote the average of the MFMC approximation  $Q_{\text{gyre}}^{\text{MFMC}}$  at the highest budget  $B = 128C_1$  taken over 250 runs using samples z that are independently drawn from the interval [-0, 5, 0.5], except for the fact that the first  $M_3$  (recall that  $M_2$  is omitted) samples of each run are drawn from the pre-collected sampling set of size 200. The use of MFMC in defining  $Q_{gyre}^{ref}$  makes use of the fact that the MFMC method is at least as accurate as MC estimation, which can be verified through the inequality (Eq. 7). With this choice, the MSEs in the MFMC and MC estimators can be measured with respect to the "exact quantity"  $Q_{gyre}^{ref}$ .

The results of this procedure are given in Table 2 and Fig. 3, from which it is evident that MFMC produces a much more precise estimate compared to that of MC. In addition to MSE  $e(Q_{gyre}^{MFMC})$ , we report the relative MSE defined as  $e_{rel}(Q_{gyre}^{MFMC}) = e(Q_{gyre}^{MFMC})/(Q_{gyre}^{ref})^2$ . It is interesting that most of the computational budget is loaded onto the very crude piecewise-linear interpolant approximation  $F_{gyre}^4(z)$ , allowing MFMC to achieve not only a lower MSE but also an estimate with much smaller variance from run to run. Especially telling is the bottom plot in Fig. 3, from which it is obvious that, for the same budget, MFMC estimation results in greater accuracy of 2 or 3 orders of magnitude compared to that MC estimation – or, looking at it another way, for the same relative MSE, MFMC estimation requires a much smaller budget compared to MC estimation.

#### **3.2** Test Case 5 for the RSWE system

We again consider the RSWE system (Eq. 12), but now  $\Omega$  denotes the whole surface of the sphere. Specifically, we consider the configuration of Test Case 5 defined in Williamson et al. (1992), which is widely used as a benchmark and employed as a stepping stone towards more realistic atmospheric models; see, e.g., Leng et al. (2019), Meng et al. (2020), Ringler et al. (2010), and Williamson et al. (1992).

**Table 2.** Results of the wind-driven barotropic gyre test with perturbed initial velocities for budgets  $B = 2^k C_1$  equivalent to  $2^k$  highest-fidelity runs.

	MC				
	k	$e(Q_{\rm gyre}^{\rm MC})$	$\times 10^{-5}$	$e_{\rm rel}(Q_{\rm gyre}^{\rm MC}) \times$	10 <sup>-3</sup>
	2		22.69		25.89
	4		11.08		12.65
	8		6.086		6.946
	16		2.539		2.898
	32		1.167		1.331
	64		0.4156	0	.4743
	128		0.1306	0	.1491
Modi	fied MFM	2			
k	Numbe	r of samples	taken of	$e(Q_{gyre}^{MFMC})$	$e_{\rm rel}(Q_{\rm gyre}^{\rm MFMC})$
	$F_{\rm gyre}^1(z)$	$F_{\rm gyre}^3(z)$	$F_{\rm gyre}^4(z)$	$\times 10^{-8}$	×10 <sup>-6</sup>
2	1	1	1968	43.09	49.18
4	1	3	4016	32.88	37.52
8	3	6	8032	10.50	11.98
16	7	12	16064	5.949	6.790
32	15	25	32 128	2.483	2.833
54	31	51	64 256	1.107	1.263
128	62	102	128 512	0.5733	0.6542

Note that, for this example, there are no additional forces acting on the system so that G(h, u) = 0 in Eq. (12).

Test Case 5 considers the flow over an isolated mountain centered at longitude  $\lambda_c = \frac{3\pi}{2}$  and latitude  $\theta_c = \frac{\pi}{6}$  with height

$$h_{\mathbf{b}}(z_1) = z_1 \left( 1 - \frac{r}{a} \right),\tag{15}$$

where  $a = \frac{\pi}{9}$ ,  $r^2 = \min\{a^2, (\lambda - \lambda_c)^2 + (\theta - \theta_c)^2\}$ , and  $\lambda, \theta$  denote longitude and latitude, respectively. In Eq. (15),  $z_1$  denotes a random variable that is uniformly distributed over the interval [1 km, 3 km].

The initial tangential (to the sphere) velocity in the longitudinal and latitudinal directions is chosen to be  $u_0(z_2) = (z_2 \cos \theta, 0)$ , where  $z_2$  denotes a random variable that is uniformly distributed over the interval [15 m s<sup>-1</sup>, 25 m s<sup>-1</sup>]. The initial fluid thickness *h* is chosen as

$$h_0(z_2) = \hat{h} - \frac{1}{g} \left( R\omega z_2 + \frac{z_2^2}{2} \right) \sin^2\theta,$$

where  $\hat{h} = 5.96$  km, R = 6371.22 km, and  $\omega = 7.292 \times 10^{-5}$  s<sup>-1</sup>. With this, the solutions u(z) and h(z) of the RSWE system (Eq. 12) depend on the random vector  $z = (z_1, z_2) \in [1 \text{ km}, 3 \text{ km}] \times [15 \text{ m s}^{-1}, 25 \text{ m s}^{-1}]$ . In Fig. 4, we provide an example of the initial thickness  $h_0(z_2)$  and the thickness  $h(z_1, z_2)$  after 10 d for specific values of  $z_1$  and  $z_2$ .

For the simulation results given in Fig. 4 and for other results in this subsection, we use the TRiSK scheme for spatial discretization and a fourth-order explicit Runge–Kutta



**Figure 3.** Results for the wind-driven barotropic gyre test with output of interest (Eq. 13) averaged over 250 applications of MC and MFMC estimation. (a) The quantity of interest Q as a function of the budget for the MC and MFMC estimators. (b) A zoom-in of the top figure. (c) The relative MSE of the MC and MFMC estimators as a function of the budget; shaded regions represent the standard deviations of the MC and MFMC predictions when compared to their averages over the 250 runs. Note that the melon region is contained in the blue region, indicating that the MFMC estimator has a smaller standard deviation than the MC estimator.



Figure 4. The initial thickness  $h_0(z_2)$  with  $z_2 = 23.82865$  m (a) and the thickness h after 10 d (b) and the mountain height  $h_b(z_1)$  with  $z_1 = 2023.78$  m.

method for temporal discretization. SCVT uniform gridding is employed as in Sect. 3.1, although now the grid covers the whole surface of the sphere. A comparison of two such SCVT grids with 480 and 240 km resolutions is provided in Fig. 5. In constructing the MC and MFMC estimators, we use the three globally refined SCVT meshes of the whole sphere according to the data in Table 3.

For the output of interest, we choose

$$F_{\text{test5}}^{1}(z) = \frac{1}{N_{\ell}^{1}} \max_{\ell=1,\dots,N_{\ell}^{1}} |u_{e,\ell}(z)|,$$

where  $N_{\ell}^1$  denotes the number of cell edges for the finest resolution case of 120 km, and  $u_{e,\ell}(z)$  denotes the value of the normal component of velocity  $u_e$  at the  $\ell$ th cell edge for any

**Table 3.** Data corresponding to the three SCVT meshes used in the

 Test Case 5 experiment.

Grid resolution		Number of cells	Number of edges	Number of vertices
120 km	$\Rightarrow$	40 962	122 880	81 920
240 km	$\Rightarrow$	10 242	30720	20 4 80
480 km	$\Rightarrow$	2562	7680	5120

choice of the random vector  $z = (z_1, z_2) \in [1 \text{ km}, 3 \text{ km}] \times [15 \text{ m s}^{-1}, 25 \text{ m s}^{-1}]$ . We then have it that the quantity of interest is given by



**Figure 5.** Two global SCVT meshes of the sphere surface with different grid resolutions: 480 km (**a**) and 240 km (**b**).

$$Q_{1,\text{test5}} = \mathbb{E}\left(F_{\text{test5}}^1(z)\right). \tag{16}$$

#### 3.2.1 MC and MFMC estimators

Similarly to before, the goal is now to construct and compare, for the same computational budget, MC and MFMC estimators of the QoI defined in Eq. (16). The MC estimator  $Q_{1,\text{test5}}^{MC}$ of the QoI  $Q_{1,\text{test5}} = \mathbb{E}(F_{\text{test5}}^1)$  is given by Eq. (2). The MFMC estimator  $Q_{1,\text{test5}}^{MFMC}$  of  $Q_{1,\text{test5}} = \mathbb{E}(F_{\text{test5}}^1)$  makes use not only of the MC estimator  $Q_{1,\text{test5}}^{MC}$  for the finest 120 km grid but also of the MC estimators  $Q_{2,\text{test5}}^{MC} \approx \mathbb{E}(F_{\text{test5}}^1)$ and  $Q_{3,\text{test5}}^{MC} \approx \mathbb{E}(F_{\text{test5}}^1)$ , corresponding the coarser 240 and 480 km grids, respectively.

In this case, 4500 uniform i.i.d. realizations of z are drawn for pre-computation, of which all models share 1200 and the two lower-fidelity surrogates share the entire 4500. Using 100 samples of the random variable z, the costs and correlation coefficients of these models are respectively approximated by

$$\mathbf{C} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 434.8 \\ 126.9 \\ 58.04 \end{pmatrix}$$
  
and  $\boldsymbol{\zeta}_1 = \begin{pmatrix} \zeta_{1,1} \\ \zeta_{1,2} \\ \zeta_{1,3} \end{pmatrix} = \begin{pmatrix} 1.00000000 \\ 0.99986604 \\ 0.99925882 \end{pmatrix},$ 

where the costs  $C_1$ ,  $C_2$ , and  $C_3$  denote the computation time (in wall-clock seconds) necessary to advance the relevant discretized RSWE system by one time step. As was the case for the wind-driven gyre experiment, these cost-correlation pairs are ideal for MFMC estimation; i.e., the surrogates OoIs  $F_{test5}^2(z)$  and  $F_{test5}^3(z)$  are very well correlated with  $F_{test5}^1(z)$ and are much less costly to obtain compared to  $F_{test5}^1(z)$ . Note that the 100 samples of z used to determine the data in Eq. (14) can be reused when determining the MC and MFMC estimators. From this point on, all other details about the construction and use of the MC and MFMC estimators

**Table 4.** Results of Test Case 5 with perturbed initial velocities for budgets  $B = 2^k C_1$  equivalent to  $2^k$  high-fidelity runs.

MC		
k	$e(Q_{\text{test5}}^{\text{MC}})$	$e_{\rm rel}(Q_{\rm test5}^{\rm MC}) \times 10^{-3}$
2	26.95	18.51
4	14.54	9.987
8	6.884	4.729
16	3.689	2.534
32	1.609	1.105
64	0.7501	0.5153
128	0.2846	0.1955
256	0.1048	0.07202

Modified	MFMC	

k	Number of	of samples	taken of	$e(Q_{\text{test5}}^{\text{MFMC}})$	$e_{\rm rel}(Q_{\rm test5}^{\rm MFMC})$
	$F_{\text{test5}}^1(z)$	$F_{\text{test}}^2(z)$	$F_{\text{test}}^3(z)$	$\times 10^{0}$	$\times 10^{-3}$
2	1	1	5	11.49	7.891
4	1	1	20	2.838	1.950
8	1	1	49	1.147	0.7879
16	1	2	106	0.6137	0.4216
32	1	5	218	0.2927	0.2011
64	2	10	437	0.09758	0.06704
128	5	20	874	0.03788	0.02603
256	10	41	1748	0.01169	0.008034

are the same as for the wind-driven barotropic gyre test case of Sect. 3.1.

The results provided in Table 4 and Fig. 6 show that MFMC estimation produces a much more precise estimate compared to MC estimation. As was true for the wind-driven gyre test case, especially telling is the bottom plot in Fig. 6, from which it is obvious that, for the same budget, MFMC estimation results in an order of magnitude greater accuracy compared to that MC estimation – or, looking at it another way, for the same relative MSE, MFMC estimation requires a much smaller budget compared to MC estimation.

## 4 First-order ice sheet model

The next experiment we consider illustrates the effectiveness of MFMC estimation on a QoI important for the realistic modeling of ice sheets such as those found near, e.g., Greenland, Antarctica, and various glaciers.

#### 4.1 The first-order model for ice sheets

The dynamical behavior of ice sheets is commonly modeled by what is referred to as the first-order model or the Blatter– Pattyn model. Here, we provide a short review of that model; detailed descriptions are given in, e.g., Blatter (1995), Pattyn (2003), Perego et al. (2012), and Tezaur et al. (2015).



**Figure 6.** Results over 250 runs of the RSWE system for the Test Case 5 experiment with quantity of interest given by Eq.(16). Shaded regions represent the variance in the MC and MFMC predictions over the 250 runs, which are independent except for the fact that all random sampling employs the same pre-collected (random) set of parameters. Note that the melon region is contained in the blue region, indicating that the MFMC estimator has a smaller standard deviation than the MC estimator.



**Figure 7.** An  $(x_1, x_3)$  cross section of the three-dimensional domain  $\Omega$  occupied by an ice sheet and the boundary segments defined in Eq. (17).

Let  $\Omega$  denote the three-dimensional domain occupied by the ice sheet with the boundary  $\Gamma = \Gamma_s \cup \Gamma_b \cup \Gamma_\ell$ , given by

$$\Gamma_{s} \leftarrow x_{3} = s(x_{1}, x_{2}) \implies \text{top surface of the ice sheet}$$
  

$$\Gamma_{b} \leftarrow x_{3} = b(x_{1}, x_{2}) \implies \text{bottom (or basal) surface}$$
  
of the ice sheet  

$$\Gamma_{\ell} \leftarrow \ell(x_{1}, x_{2}) = 0 \implies \text{lateral boundary.}$$
(17)

Figure 7 provides an illustration of the boundary segments defined in Eq. (17).

Also, let  $u_1$  and  $u_2$  denote the  $x_1$  and  $x_2$  components of the velocity vector  $\boldsymbol{u} = (u_1 \ u_2 \ u_3)^{\mathsf{T}}$ . Then, the first-order model equations for ice sheets is given by the following partial differential equations:

$$\begin{cases} -\nabla \cdot (2\mu\epsilon_1) + \rho g \frac{\partial s}{\partial x_1} = 0\\ -\nabla \cdot (2\mu\epsilon_2) + \rho g \frac{\partial s}{\partial x_2} = 0 \end{cases} \quad \text{for } \boldsymbol{x} = (x_1 \ x_2 \ x_3)^{\mathsf{T}} \in \Omega, \quad (18)$$

where  $\mu$  denotes the viscosity coefficient, g denotes the gravitational acceleration, and  $\rho$  denotes the density. In Eq. (18), the strain-rate tensor ( $\epsilon_1$ ,  $\epsilon_2$ ) is given by

$$\boldsymbol{\epsilon}_{1} = \begin{pmatrix} 2\frac{\partial u_{1}}{\partial x_{1}} + \frac{\partial u_{2}}{\partial x_{2}} \\ \frac{1}{2} \left( \frac{\partial u_{1}}{\partial x_{2}} + \frac{\partial u_{2}}{\partial x_{1}} \right) \\ \frac{1}{2} \frac{\partial u_{1}}{\partial x_{3}} \end{pmatrix} \text{ and } \boldsymbol{\epsilon}_{2} = \begin{pmatrix} \frac{1}{2} \left( \frac{\partial u_{1}}{\partial x_{2}} + \frac{\partial u_{2}}{\partial x_{1}} \right) \\ \frac{\partial u_{1}}{\partial x_{1}} + 2\frac{\partial u_{2}}{\partial x_{2}} \\ \frac{1}{2} \frac{\partial u_{2}}{\partial x_{3}} \end{pmatrix}$$

and the nonlinear viscosity coefficient  $\mu$  is given by the Glen flow law

$$\mu = \frac{1}{2} A^{-\frac{1}{n}} \boldsymbol{\epsilon}_{\mathrm{e}}^{\frac{1}{n}-1}$$

with n = 3 being the usual choice. The effective strain rate  $\epsilon_e$  is given by

$$\begin{aligned} \boldsymbol{\epsilon}_{\mathsf{e}}^{2} &= \left(\frac{\partial u_{1}}{\partial x_{1}}\right)^{2} + \left(\frac{\partial u_{2}}{\partial x_{2}}\right)^{2} + \frac{\partial u_{1}}{\partial x_{1}}\frac{\partial u_{2}}{\partial x_{2}} + \frac{1}{4}\left(\frac{\partial u_{1}}{\partial x_{2}} + \frac{\partial u_{2}}{\partial x_{1}}\right)^{2} \\ &+ \frac{1}{4}\left(\frac{\partial u_{1}}{\partial x_{3}}\right)^{2} + \frac{1}{4}\left(\frac{\partial u_{2}}{\partial x_{3}}\right)^{2}, \end{aligned}$$

and A is often chosen to obey the Arrhenius relation

$$A = A(T) = a \exp(-Q/RT),$$

where T denotes the absolute temperature measured in degrees kelvin, R denotes the universal gas constant, Q denotes the activation energy for creep, and a is an empirical flow constant often used as a tuning parameter.

The system (Eq. 18) is supplemented by the boundary conditions

$$\begin{aligned} \boldsymbol{\epsilon}_1 \cdot \boldsymbol{n} &= 0 & \text{and} & \boldsymbol{\epsilon}_2 \cdot \boldsymbol{n} &= 0 & \text{on } \Gamma_{\mathrm{s}}, \\ 2\mu \boldsymbol{\epsilon}_1 \cdot \boldsymbol{n} + \beta \boldsymbol{u} &= 0 & \text{and} & 2\mu \boldsymbol{\epsilon}_2 \cdot \boldsymbol{n} + \beta \boldsymbol{v} &= 0 & \text{on } \Gamma_{\mathrm{b}}, \\ \boldsymbol{u} &= 0 & \text{and} & \boldsymbol{v} &= 0 & \text{on } \Gamma_{\ell}, \end{aligned}$$

$$(19)$$

where  $\beta$  denotes a basal-friction parameter.

Once the horizontal components  $u_1$  and  $u_2$  of the velocity are determined, the vertical-velocity component w is determined by enforcing incompressibility; i.e., we have it that

$$\frac{\partial w}{\partial x_3} = -\frac{\partial u_1}{\partial x_1} - \frac{\partial u_2}{\partial x_2} \quad \text{for } \boldsymbol{x} \in \Omega.$$
(20)

Because the right-hand side is known, this is an ordinary differential equation for w. Of paramount interest in the modeling of ice sheets is the monitoring of the temporal evolution of the ice sheet domain  $\Omega$ . However, one notices that there are no time derivatives of  $u_1$  and  $u_2$  appearing in Eq. (18); i.e., that system is a static one. The reasoning behind this curiosity is twofold. Firstly, we have it that the system (Eq. 18) is coupled to an equation for the temporal evolution of the temperature T within the ice sheet and is also coupled to an equation for the temporal evolution of the ice sheet, which engenders changes in the ice sheet domain  $\Omega$ . Secondly, the timescale of changes in the temperature is much shorter (e.g., hourly) when compared to the timescale (e.g., at least many days) of changes in the ice sheet domain  $\Omega$ .

Thus, in a computational model of ice sheet dynamics, determination of the domain  $\Omega$  is coupled to that of the velocity  $\boldsymbol{u}$ . Precisely, the temperature and top surface are first advanced from a given domain-velocity pair  $(\Omega, \boldsymbol{u})$  over several time steps, producing an evolution that prescribes a new domain-temperature pair  $(\Omega, T)$ . Following this, the pair  $(\Omega, T)$  are used to generate a new velocity  $\boldsymbol{u}$  by solving Eqs. (18) and (20). While both stages of this process are important, the present experiment focuses on the computation of  $\boldsymbol{u}$  from  $\Omega$  and T.

# 4.2 MC and MFMC estimation

The specific application setting we consider is based on Experiment C of the benchmark examples in Leng et al. (2012) and Pattyn et al. (2008). Here, the ice domain is a rectangular parallelepiped that has a square base with side length L and thickness H and which lies on a slanted bed with slope parameter  $\theta$ . The upper surface boundary is given as

 $z_{\rm s}(x_1, x_2) = -x_1 \tan \theta,$ 

and the basal topography is given as

$$z_{\rm b}(x_1, x_2) = z_{\rm s}(x_1, x_2) - H.$$

For constructing the MC and MFMC estimators, we set the length L = 80 km and height H = 1 km and then use uniform tetrahedral grids of the ice domain  $\Omega_t$  with 120, 60, and 30 grid intervals in each horizontal direction and, correspondingly, 20, 10, and 5 intervals in the vertical direction. As a result, we have that number of vertices, horizontal triangles, and tetrahedra are determined according to the data in Table 5.

The first-order ice sheet model is discretized using the stabilized P1–P1 finite elements given in Zhang et al. (2011). To solve the resulting nonlinear system of discrete equations, 15 Picard iteration steps are carried out, after which a switch is made to a Newton iteration, up to a maximum of a total of 40 iterations. However, if the residual error decrease is less than 75 % relative to the residual error of the previous step, the nonlinear iteration is switched back to a Picard iteration. An example illustration of the discrete solution of this ice

grid resolution		number of vertices	number of horizontal triangles	number of tetrahedra
$120 \times 120 \times 20$	$\Rightarrow$	307 461	28 800	1 728 000
$60 \times 60 \times 10$	$\Rightarrow$	40 931	7200	216 000
$30 \times 30 \times 5$	$\Rightarrow$	5766	1800	27 000

sheet model for specific values of the parameters  $\theta$  and  $\beta$  is provided in Fig. 8 using the highest-resolution grid.

Because the cracking and melting of ice sheets is an important indicator of climate change (Clark et al., 1999; Hanna et al., 2013), we consider the the output of interest to be

$$F_{\rm ice}^1(z) = \frac{1}{2N_1} \sum_{\ell=1}^{N_1} |\boldsymbol{u}_n(z)|^2,$$

where  $N_1 = 307461$  is the number of vertices of the finest  $120 \times 120 \times 20$  grid, and  $u_n(z)$  denotes the value of the discrete velocity u at the *n*th vertex for any choice of the i.i.d. random vector  $z = (\theta, \beta) \in [0.2, 0.8] \times [800 \text{ m}, 1200 \text{ m}]$ . This OoI provides a measurement of how energetic the ice sheet is depending on its slope and friction coefficient and can be loosely related to how vigorously the ice will deform given a configuration specified by z. We then choose the quantity of interest, given by

$$Q_{1,\text{ice}} = \mathbb{E}\left(F_{\text{ice}}^1(z)\right). \tag{21}$$

Two surrogates for the ice sheet model,

$$F_{\text{ice}}^{2}(z) = \frac{1}{2N_{2}} \sum_{n=1}^{N_{2}} |\boldsymbol{u}_{n}(z)|^{2} \text{ and}$$
$$F_{\text{ice}}^{3}(z) = \frac{1}{2N_{3}} \sum_{n=1}^{N_{3}} |\boldsymbol{u}_{n}(z)|^{2},$$

are defined respectively using the coarser  $60 \times 60 \times 10$  grid with  $N_2 = 40931$  and the even coarser  $30 \times 30 \times 5$  grid with  $N_3 = 5766$ . Then, from Eq. (3), we have the corresponding three Monte Carlo estimators { $Q_{1,ice}^{MC}$ ,  $Q_{2,ice}^{MC}$ ,  $Q_{3,ice}^{MC}$ }, and from Eq. (4), we have the MFMC estimator  $Q_{ice}^{MFMC}$ , which makes use of these three MC estimators.

At this point, we proceed as was done in Sect. 3.2. For example, we now have that the approximate costs (measured in wall-clock seconds and averaged over 100 random samples) for computing  $F_{ice}^1(z)$ ,  $F_{ice}^2(z)$ , and  $F_{ice}^3(z)$ , as well as the approximate correlations for the three OoIs, are given by

$$\mathbf{C} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 285.5 \\ 23.05 \\ 2.690 \end{pmatrix}$$
  
and  $\boldsymbol{\zeta}_1 = \begin{pmatrix} \zeta_{1,1} \\ \zeta_{1,2} \\ \zeta_{1,3} \end{pmatrix} = \begin{pmatrix} 1.00000000 \\ 0.99999796 \\ 0.99996691 \end{pmatrix}.$ 



**Figure 8.** A simulation of the ice sheet model with  $\theta = 0.7387$  and  $\beta = 901.5232$ , discretized at the highest grid resolution. The top two plots display contours of the velocity components  $u_1$  and  $u_2$ , respectively, on the top surface  $\Gamma_s$ . The third plot displays a contour plot of velocity component  $u_3$ . The fourth plot displays a cutaway plot of the velocity magnitude  $|\mathbf{u}|$  inside the sheet.

Mo

k

The remaining experimental details are nearly identical to those of Sect. 3.2. As was the case for the oceanic simulations, the computational expense of these models precludes the collection of unlimited data, so a total of 8000 uniform i.i.d. realizations of z are drawn, of which all models share 200 and the lower-fidelity surrogates share 1000. The results of carrying out MFMC and MC estimation over 250 "independent" runs using this common data set are provided in Table 6 and Fig. 9. Here, it is especially apparent that the budget-preserving modifications to MFMC that led to the algorithm in Box 2 were necessary, as  $B \ll \mathbf{C} \cdot \mathbf{R}$  and because only one high-fidelity run is ever selected by the algorithm.

Despite this, it is clear from Table 6 and the plots in Fig. 9 that, within these small budgets, MFMC estimation produces an estimator which is much more accurate and precise than that of MC estimation. This is not surprising because MFMC estimation has the flexibility to load most of its budget onto the cheaper lower-fidelity surrogate models, evaluating them many times in order to bring down the overall variance of their estimates. Again, this provides empirical validation for the use of MFMC estimation over MC estimation when estimating model statistics, even in practical cases where data availability is low.

Table 6. Results of the	ice sheet e	experiment	for budgets	equivalent
to $2^k$ high-fidelity runs			C C	•

MC			
k	$e(Q_{\rm ice}^{\rm MC})\times 10^6$	$e_{\rm rel}(Q_{\rm ice}^{\rm MC}) \times$	$10^{-2}$
2	37.26		11.79
4	20.08	(	5.354
8	11.62		3.676
16	5.296		1.676
32	2.457	0.	7774
64	1.018	0.	3221
ied MFN	IC		
Number	of samples taken of	$e(Q_{ice}^{MFMC})$	$e_{\rm rel}(Q_{\rm i}^{\rm l})$
$F_{\rm ice}^1(z)$	$F_{\rm ice}^2(z) = F_{\rm ice}^3(z)$	$\times 10^4$	×10
-			

	$I_{\text{ice}}(z)$	$I_{\text{ice}}(z)$	$I_{\text{ice}}(z)$	~10	×10	
2	1	1	97	69.67	22.05	
4	1	1	309	16.81	5.323	
8	1	1	726	3.793	1.200	
16	1	4	1555	1.370	0.4335	
32	1	8	3215	1.123	0.3554	
64	1	17	6506	0.3393	0.1074	



**Figure 9.** Results over 250 runs of the ice sheet experiment with quantity of interest (Eq. 21). Shaded regions represent variance in the MC and MFMC predictions over the 250 runs. Note that the melon region is contained in the blue region, indicating that the MFMC estimator has a smaller standard deviation than the MC estimator.

#### 5 Concluding remarks

This paper serves to introduce multifidelity Monte Carlo estimation as an alternative to standard Monte Carlo estimation for quantifying uncertainties in the outputs of climate system models, albeit in very simplified settings. Specifically, we consider benchmark problems for the single-layer shallowwater equations relevant to ocean and atmosphere dynamics, and we also consider a benchmark problem for the firstorder model of ice sheet dynamics. The computational results presented here are promising in that they amply demonstrate the superiority of MFMC estimation when compared to MC estimation on these examples. Furthermore, the use of MFMC as an estimation method will surely be even more efficacious when quantifying uncertainties in more realistic climate-modeling settings for which the simulation costs are prohibitively large, e.g., for long-term climate simulations. Thus, our next goal is to apply MFMC estimation to more useful models of climate dynamics (such as the primitive equations for ocean and atmosphere and the Stokes model for ice sheets) that are also coupled to the dynamics of other climate system components and also to passive and active tracer equations.

*Code and data availability.* The climate simulation data used in this work along with Python code for reproducing the relevant experiments can be found at the GitHub repository (https://github.com/agrubertx/Multifidelity-Monte-Carlo, last access: 10 February 2023) with permanent identifier https://doi.org/10.5281/zenodo.7071646 (Gruber et al., 2022b).

*Author contributions.* The idea for this work was conceived by MG, and the experiments presented were designed jointly by all authors. Software development, data collection, and figure generation were carried out by AG and RL, while post-experiment analysis

was conducted by all authors. The resulting paper was written by MG and AG with editorial contributions from all authors.

*Competing interests.* The contact author has declared that none of the authors has any competing interests.

*Disclaimer.* This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract no. DE-NA0003525.

*Financial support.* This material is based upon work partially supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Mathematical Multifaceted Integrated Capability Centers (MMICCS) program, under Field Work Proposal no. 22025291 and the Multifaceted Mathematics for Predictive Digital Twins (M2dt) project. This work is further supported by U.S. Department of Energy Office of Science under grant nos. DE-SC0020270, DE-SC0022254, DE-SC0020418, and DE-SC0021077.

*Review statement.* This paper was edited by Dan Lu and reviewed by Huai Zhang and one anonymous referee.

# References

- Addcock, B., Brugiapaglia, S., and Webster, C.: Sparse Polynomial Approximation of High-Dimensional Functions, SIAM, 1–310, https://doi.org/10.1137/1.9781611976885, 2022.
- Anderson, D., David, L., and Gill, A.: Spin-up of a stratified ocean, with applications to upwelling, Deep-Sea Res. Ocean. Abstr., 22, 583–596, https://doi.org/10.1016/0011-7471(75)90046-7, 1975.
- Blatter, H.: Velocity and stress fields in grounded glaciers: A simple algorithm for including deviatoric stress gradients, J. Glaciol., 41, 333–344, https://doi.org/10.3189/S002214300001621X, 1995.
- Bleck, R. and Boudra, D.: Wind-driven spin-up in eddyresolving ocean models formulated in isopycnic and isobaric coordinates, J. Geophys. Res.-Oceans, 91, 7611–7621, https://doi.org/10.1029/JC091iC06p07611, 1986.
- Clare, M. C. A., Leijnse, T. W. B., McCall, R. T., Diermanse, F. L. M., Cotter, C. J., and Piggott, M. D.: Multilevel multifidelity Monte Carlo methods for assessing uncertainty in coastal flooding, Nat. Hazards Earth Syst. Sci., 22, 2491–2515, https://doi.org/10.5194/nhess-22-2491-2022, 2022.
- Clark, P., Alley, R., and Pollard, D.: Northern Hemisphere ice-sheet influences on global climate change, Science, 286, 1104–1111, https://doi.org/10.1126/science.286.5442.1104, 1999.
- Cristianini, N. and Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernelbased Learning Methods, Cambridge University Press, https://doi.org/10.1017/CBO9780511801389, 2000.
- Dimarco, G., Liu, L., Pareschi, L., and Zhu, X.: Multi-fidelity methods for uncertainty propagation in kinetic equations, Panoramas et synthèses, in press, 2023.
- Du, Q., Faber, V., and Gunzburger, M.: Centroidal Voronoi tessellations: Applications and algorithms, SIAM Rev., 41, 637–676, https://doi.org/10.1137/S0036144599352836, 1999.
- Du, Q., Gunzburger, M., and Ju, L.: Constrained centroidal Voronoi tessellations for surfaces, SIAM J. Sci. Comput., 24, 1488–1506, https://doi.org/10.1137/S1064827501391576, 2003.
- Evans, M. and Swartz, T.: Approximating integrals via Monte Carlo and deterministic methods, Vol. 20, Oxford University Press, Oxford, 2000.
- Fritzen, F. and Ryckelynck, D.: Machine Learning, Low-Rank Approximations and Reduced Order Modeling in Computational Mechanics, MDPI, https://doi.org/10.3390/books978-3-03921-410-5, 2019.
- Gruber, A., Gunzburger, M., Ju, L., and Wang, Z.: A multifidelity Monte Carlo method for realistic computational budgets, arXiv [preprint], https://doi.org/10.48550/arXiv.2206.07572, 2022a.
- Gruber, A., Gunzburger, M., Ju, L., and Wang, Z.: Code and data for multifidelity Monte Carlo simulation, Zenodo [code], https://doi.org/10.5281/zenodo.7071646, 2022b.
- Gunzburger, M., Webster, C., and Zhang, G.: Stochastic finite element methods for partial differential equations with random input data, Acta Numer., 23, 521–650, https://doi.org/10.1017/S0962492914000075, 2014.

- Hanna, E., Navarro, F., Pattyn, F., Domingues, C., Fettweis, X., Ivins, E., Nicholls, R., Ritz, C., Smith, B., Tulaczyk, S., Whitehouse, P. L., and Zwally, H. J.: Icesheet mass balance and climate change, Nature, 498, 51–59, https://doi.org/10.1038/nature12238, 2013.
- Jacobsen, D. W., Gunzburger, M., Ringler, T., Burkardt, J., and Peterson, J.: Parallel algorithms for planar and spherical Delaunay construction with an application to centroidal Voronoi tessellations, Geosci. Model Dev., 6, 1353–1365, https://doi.org/10.5194/gmd-6-1353-2013, 2013.
- Keiper, W., Milde, A., and Volkwein, S.: Reduced-Order Modeling (ROM) for Simulation and Optimization: Powerful Algorithms as Key Enablers for Scientific Computing, Springer, ISBN-13 9783030091996, 2018.
- Konrad, J.: Multifidelity Monte Carlo Sampling in Plasma Microturbulence Analysis, Bachelor's Thesis, Technical University of Munich, https://mediatum.ub.tum.de/doc/1522015/1522015.pdf (last access: 8 February 2022), 2019.
- Law, F., Cerfon, A., and Peherstorfer, B.: Accelerating the estimation of collisionless energetic particle confinement statistics in stellarators using multifidelity Monte Carlo, Nucl. Fusion, 62, 7, https://doi.org/10.1088/1741-4326/ac4777, 2022.
- Leng, W., Ju, L., Gunzburger, M., Price, S., and Ringler, T.: A parallel high-order accurate finite element nonlinear Stokes ice sheet model and benchmark experiments, J. Geophys. Res.-Earth, 117, F01001, https://doi.org/10.1029/2011JF001962, 2012.
- Leng, W., Ju, L., Wang, Z., and Pieper, K.: Conservative explicit local time-stepping schemes for the shallow water equations, J. Comput. Phys., 382, 152–176, https://doi.org/10.1016/j.jcp.2019.01.006, 2019.
- Meng, X., Hoang, T., Wang, Z., and Ju, L.: Localized exponential time differencing method for shallow water equations: Algorithms and numerical study, Commun. Comp. Phys., 29, 80–110, https://doi.org/10.4208/cicp.OA-2019-0214, 2020.
- Modderman, J.: Exploratory research on application multilevel multi-fidelity Monte Carlo in fluid dynamics topics: study on flow past a porous cylinder, PhD Thesis, Delft University of Technology, http://resolver.tudelft.nl/uuid: 293e9730-092f-4c88-97e1-092ab9abdce3 (last access: 8 February 2022), 2021.
- Nierderreiter, H.: Random Number Generation and quasi-Monte Carlo Methods, SIAM, 1–241, https://doi.org/10.1137/1.9781611970081, 1992.
- Nye, J.: The distribution of stress and velocity in glaciers and ice-sheets, Proc. R. Soc. Lond. A, 239, 113–133, https://doi.org/10.1098/rspa.1957.0026, 1957.
- Paterson, W.: Physics of Glaciers, Butterworth-Heinemann, ISBN-13 9780750647427, 1994.
- Pattyn, F.: A new three-dimensional higher-order thermomechanical ice-sheet model: basic sensitivity, ice stream development, and ice flow across subglacial lakes, J. Geophys. Res., 108, 1– 15, https://doi.org/10.1029/2002JB002329, 2003.
- Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A. J., Pollard, D., Price, S., Rückamp, M., Saito, F., Souček, O., Sugiyama, S., and Zwinger, T.: Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–

## A. Gruber et al.: MFMC estimation for efficient UQ in climate modeling

HOM), The Cryosphere, 2, 95–108, https://doi.org/10.5194/tc-2-95-2008, 2008.

- Peherstorfer, B., Willcox, K., and Gunzburger M.: Optimal model management for multifidelity Monte Carlo estimation, SIAM J. Sci. Comput., 38, A3163–A3194, https://doi.org/10.1137/15M1046472, 2016.
- Peherstorfer, B., Gunzburger, M., and Willcox, K.: Convergence analysis of multifidelity Monte Carlo estimation, Numer. Math., 139, 683–707, https://doi.org/10.1007/s00211-018-0945-7, 2018.
- Perego, M., Gunzburger, M., and Burkardt, J.: Parallel finiteelement implementation for higher-order ice-sheet models, J. Glaciol., 58, 76–88, https://doi.org/10.3189/2012JoG11J063, 2012.
- Quarteroni, A. and Rozza, G. (Eds.): Reduced Order Methods for Modeling and Computational Reduction, Springer, https://doi.org/10.1007/978-3-319-02090-7, 2014.
- Quarteroni, A., Manzoni, A., and Negri, F. (Eds.): Reduced Basis Methods for Partial Differential Equations: An Introduction, Springer, https://doi.org/10.1007/978-3-319-15431-2, 2016.
- Quick, J., Hamlington, P. E., King, R., and Sprague, M. A.: Multifidelity uncertainty quantification with applications in wind turbine aerodynamics, AIAA 2019-0542, AIAA Scitech 2019 Forum, 7–11 January 2019, San Diego, California, https://doi.org/10.2514/6.2019-0542, 2019.
- Rezaeiravesh, S., Vinuesa, R., and Schlatter P.: Towards multifidelity models with calibration for turbulent flows, in: WCCM-ECCOMAS2020, https://doi.org/10.23967/wccmeccomas.2020.348, 2020.
- Romer, U., Zafar, S., and Fezans, N.: A mutifidelity approach for uncertainty propagation in flight dynamics, in Fundamentals of High Lift for Future Civil Aircraft, Springer International Publishing, 463–478, https://doi.org/10.1007/978-3-030-52429-6\_28, 2020.
- Ringler, T., Ju, L., and Gunzburger, M.: A multi-resolution method for climate system modeling: Application of spherical centroidal Voronoi tessellations, Ocean Dynam., 58, 475–498, https://doi.org/10.1007/s10236-008-0157-2, 2008.
- Ringler, T., Thuburn, J., Klemp, J., and Skamarock, W.: A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured C-grids, J. Comput. Phys., 229, 3065– 3090, https://doi.org/10.1016/j.jcp.2009.12.007, 2010.
- Sloan, I. H.: Lattice Methods for Multiple Integrations, J. Comput. Appl. Math., 12–13, 131–143, https://doi.org/10.1016/0377-0427(85)90012-3, 1985.

- Smith, R. C.: Uncertainty Quantification: Theory, Implementation, and Applications, SIAM, ISBN-13 9781611973211, 2013.
- Steinwart, I. and Christmann, A.: Support Vector Machines, WIREs Comp. Stat., 1, 283–289, https://doi.org/10.1002/wics.49, 2008.
- Tezaur, I. K., Perego, M., Salinger, A. G., Tuminaro, R. S., and Price, S. F.: *Albany/FELIX*: a parallel, scalable and robust, finite element, first-order Stokes approximation ice sheet solver built for advanced analysis, Geosci. Model Dev., 8, 1197–1220, https://doi.org/10.5194/gmd-8-1197-2015, 2015.
- Thuburn, J., Ringler, T., Skamarock, W., and Klemp, J.: Numerical representation of geostrophic modes on arbitrarily structured C-grids, J. Comput. Phys., 228, 8321–8335, https://doi.org/10.1016/j.jcp.2009.08.006, 2009.
- Trahan, C. J. and Dawson, C.: Local time-stepping in Runge-Kutta discontinuous Galerkin finite element methods applied to the shallow-water equations, Comput. Methods Appl. Mech. Eng., 217–220, 139–152, https://doi.org/10.1016/j.cma.2012.01.002, 2012.
- Valero, M., Jofre, L., and Torres, R.: Multifidelity prediction in wildfire spread simulation: Modeling, uncertainty quantification and sensitivity analysis, Environ. Modell. Softw., 141, 105050, https://doi.org/10.1016/j.envsoft.2021.105050, 2021.
- Vallis, G. K., Atmospheric and Oceanic Fluid Dynamics, Cambridge University Press, ISBN-13 9781107588417, 2006.
- Williamson, D., Drake, J., Hack, J., Jacob, R., and Swarztrauber, P.: A standard test set for numerical approximations to the shallow water equations in spherical geometry, J. Comput. Phys., 102, 211–224, https://doi.org/10.1016/S0021-9991(05)80016-6, 1992.
- Wolfram, P., Ringler, T., Maltrud, M., Jacobsen, D., and Petersen, M.: Diagnosing isopycnal diffusivity in an eddying, idealized midlatitude ocean basin via Lagrangian, in situ, global, highperformance particle tracking (light), J. Phys. Oceanogr., 45, 2114–2133, https://doi.org/10.1175/JPO-D-14-0260.1, 2015.
- Yang, H., Gunzburger, M., and Ju, L.: Fast spherical centroidal Voronoi mesh generation: A Lloyd-preconditioned LBFGS method in parallel, J. Comput. Phys., 367, 235–252, https://doi.org/10.1016/j.jcp.2018.04.034, 2018.
- Zhang, H., Ju, L., Gunzburger, M., Ringler, T., and Price, S.: Coupled models and parallel simulations for three-dimensional full-Stokes ice sheet modeling, Numer. Math-Theory ME, 4, 396– 418, https://doi.org/10.1017/S1004897900000416, 2011.