Probabilistic Engineering Mechanics 69 (2022) 103291

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

On information fusion for reliability estimation with multifidelity models

Carsten Proppe *, Jonas Kaupp

Chair of Engineering Mechanics, Karlsruhe Institute of Technology, Kaiserstr. 10, Bdg. 10.23, 76131 Karlsruhe, Germany

ARTICLE INFO

Keywords:
Multifidelity
Model hierarchy
Information fusion
Reliability estimation
Moving particles algorithm
Importance sampling

ABSTRACT

Multifidelity models attempt to reduce the computational effort by combining simulation models of different approximation quality and from different sources. Information fusion combines outputs from a model hierarchy in order to obtain efficient estimators for a quantity of interest. In this paper, information fusion is applied to reliability estimation. To this end, efficient multifidelity estimators for the probability of failure are developed by combining additive and multiplicative information fusion with importance sampling and importance splitting (notably the moving particles method). Importance sampling and importance splitting based multifidelity reliability estimators are compared focusing on relative error and coefficient of variation.

1. Introduction

Reliability analysis in engineering is concerned with the determination of the probability that the performance function of the system (e.g. the difference between resistance and load of a structure) becomes negative. For most engineering applications, the failure probability is rather small, which precludes the application of direct Monte Carlo simulation for its evaluation. Instead, importance sampling and importance splitting have been widely applied to estimate the occurrence of rare events [1].

Importance sampling estimates the occurrence of rare events by generating samples from an alternative distribution and correcting for the bias by the introduction of weights. The success of this method relies on the quality of the importance sampling density, which is therefore often constructed in an adaptive way, e.g. by the cross-entropy method [2].

Importance splitting allows estimating small failure probabilities efficiently, even for problems that involve a high-dimensional vector of input random variables [1]. It is based on a multiplicative decomposition of the failure probability in larger conditional probabilities that are estimated by means of Markov chain Monte Carlo simulation methods, cf. e.g. subset simulation [3]. Recently, the moving particles method has been proposed for reliability estimation [4], which is an importance splitting method that associates a threshold to each sample, moves samples to new positions in the design space and counts the number of moves of the initial samples to reach the failure domain. This quantity yields an estimator for the failure probability, which is of comparable accuracy and efficiency as the subset simulation estimator [5]. The advantage of this algorithm is the description of the number of moves by a Poisson distribution and thus the application of Poisson process theory to importance splitting.

In addition, the introduction of sets of models instead of a single model offers a great potential for increasing the efficiency of reliability computations. In multilevel and multifidelity methods, sets of models are introduced that comprise in general a computational expensive high-fidelity model and one or several less expensive low-fidelity models. In multilevel methods [6], models are ordered by computational accuracy [7], e.g. by means of a discretization parameter (a mesh parameter or a time step) that is linked to the approximation error. In multifidelity methods the output of low-fidelity models is leveraged to increase the computational efficiency for an estimator of a quantity of interest while maintaining its accuracy [8]. The low-fidelity models can come from various sources including e.g. different discretization methods, different mathematical models or even experimental models. The relation to the high-fidelity model is in general established by means of a statistical parameter such as the Pearson correlation, dependence-based measures or information-based measures [9].

The quantity of interest computed by means of the different models can be combined by serial application (i.e. information fusion) and parallel application (i.e. information filtering) of the different models [10]. Information fusion and information filtering can also be combined [11]. For information fusion, additive information fusion based on a telescoping sum seems to dominate the literature, cf. e.g. [6]. In addition, multiplicative information fusion by means of metamodeling techniques based on maximum likelihood estimation or regression has been proposed as well, especially in the context of co-Kriging [12–14].

For reliability estimation, information filtering has been well established in recent years. The general idea is to apply a selective refinement strategy, such that realizations far away from the boundary between the failure domain and the safe domain are solved by a lower accuracy than those close to the boundary, which further reduces the computational effort.

* Corresponding author.
E-mail addresses: proppe@kit.edu (C. Proppe), jonas.kaupp@kit.edu (J. Kaupp).

https://doi.org/10.1016/j.probengmech.2022.103291
Received 29 October 2021; Received in revised form 3 March 2022; Accepted 26 April 2022
Available online 14 May 2022

This concept has been combined with direct Monte Carlo simulation [15,16] and importance sampling [17]. A multilevel method based on information filtering has been combined with importance splitting in [18]. In [19], low-fidelity models of increasing accuracy have been employed as preconditioners to determine an importance sampling density based on cross-entropy minimization. Concerning information fusion, applications to reliability estimation are less frequent. Direct Monte Carlo simulation and co-Kriging have been proposed for reliability estimation in [13]. A multilevel Monte Carlo method has been proposed that is based on additive information fusion and information filtering [11]. Moreover, additive information fusion has been combined with cross-entropy based importance sampling in [20].

The aim of this contribution is to study the combination of information fusion and efficient simulation of rare events by importance sampling and importance splitting in a systematic manner with the aim of further increasing the efficiency of the latter methods by introducing a set of models instead of the application of a single high-fidelity model. To this end, both additive and multiplicative information fusion are combined with cross-entropy based importance sampling and the moving particles importance splitting algorithm. Besides the already existing methods for reliability estimation with additive information fusion, new methods based on multiplicative information fusion are developed and introduced. They combine cross-entropy based importance sampling with a regression approach and the moving particles method with multivariate estimation for Poisson distributions. This yields in total four different prototypes of multifidelity rare event simulation techniques that are established and critically compared.

This paper is organized as follows: in the next section, information fusion for the combination of the model output is briefly introduced and the concepts of additive and multiplicative information fusion are outlined. The following section introduces cross-entropy based importance sampling and the moving particles algorithm for importance splitting. After that, importance sampling and importance splitting are combined with additive and multiplicative information fusion. The proposed algorithms are tested on examples comprising a simple cumulative distribution function, a system of stochastic ordinary differential equations and a nonlinear stochastic partial differential equation. Finally, conclusions are drawn.

2. Information fusion

The application of a set of models instead of a single model offers great potential to develop computationally efficient simulation models. While less computational demanding low-fidelity models help to reduce the computational effort, the high-fidelity model is necessary to maintain the accuracy of the simulations.

In principle, the information obtained from a set of models can be evaluated in two different ways. Either the models are applied in parallel and an appropriate model out of the set of models hierarchy is selected to carry out a specific simulation task — or the model hierarchy is applied in series and the information of each model is fused such that the overall computational effort is less than that of using solely the high-fidelity model. According to [10], the former approach is called information filtering, while the latter approach is termed information fusion. As mentioned in the introduction, information filtering and information fusion are not necessarily competing techniques, but can be combined. To this end, starting from a set of models, the combination of model outputs in parallel by information filtering leads to metamodels whose output is then combined in series by information fusion.

Concerning information fusion, a first method that combines a high- and a low-fidelity model in a multiplicative manner has been presented as "global–local approximation" in [21]. Both additive and multiplicative combination of output quantities obtained from a high- and a low-fidelity model has been investigated for the analysis of a crack in a stiffened composite panel in [22]. The generalization of these ideas to sets of models leads to information fusion based on telescoping or weighted sums [8] and information fusion based on telescoping products [14], respectively.

The former method associates costs to the evaluation of the different models and determines an optimal number of simulation runs for a fixed computational budget. For reliability estimation, this method has been combined with information filtering in [11].

The possibility to link models by telescoping products of conditional probabilities for the output variables has been mentioned in [14], although the investigations are limited to the case of two models. These methods require the calibration of one or several metamodels that allow to replace high-fidelity simulations by computationally less expensive simulations of lower fidelity. They do not optimize costs.

2.1. Additive information fusion

For additive information fusion, denote by Q_l a quantity of interest related to the lth model. An estimator for the quantity of interest of the high-fidelity model L that combines the information from the set of models in an additive manner is obtained from the telescoping sum [6,23]

$$E[Q_L] = E[Q_0] + \sum_{l=1}^{L} E[Q_l - Q_{l-1}]. \tag{1}$$

The aim is to compute each of the estimates on the right-hand side of this equation individually by Monte Carlo simulation. A reduction of the overall computational effort can be expected if the variance of the differences decreases with increasing index l and thus, for a given coefficient of variation, Monte Carlo estimates for the differences of highly accurate models will require less samples. The fact that there is a nested sequence of approximations might also be beneficial in a similar manner as for multigrid methods.

It is important to note that for the estimation of $E[Q_l - Q_{l-1}]$ both Q_l and Q_{l-1} are evaluated for the same samples. If the dimension of the random vector depends on the selected model of the model hierarchy, it is possible to generate the samples for the model with higher fidelity and to obtain the corresponding samples for the model with lower fidelity by coarse-graining.

If models from different sources are considered, such that a joint input vector cannot be established, an additive combination of the model outputs is still possible. For this purpose, in analogy to the method of control variates, the estimator

$$\hat{Q} = \hat{Q}_L + \sum_{l=1}^{L-1} \alpha_l (\hat{Q}_l - \hat{\mu}_l) \tag{2}$$

is considered, cf. [8], where

$$\hat{Q}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} Q_{l,i} \text{ and } \hat{\mu}_l = \frac{1}{N_{l+1}} \sum_{i=1}^{N_{l+1}} Q_{l,i}, \text{ with } N_l > N_{l+1}. \tag{3}$$

Then, optimal weights α_l, $l = 1, 2, \dots, L-1$ and sample sizes N_l, $l = 1, 2, \dots, L$ are determined such that the variance of the estimator is minimal for a fixed computational budget. As for control variates, the efficiency of the estimator depends on the Pearson correlation between the high-fidelity and the low-fidelity models [24].

2.2. Multiplicative information fusion

Multiplicative information fusion relies on training samples of the model with higher and lower fidelity and a statistical metamodel that links the quantity of interest of both models. The metamodel can be based on the joint probability density function or, by means of the decomposition

$$\begin{aligned} p(Q_L) &= \int \dots \int p(Q_L, \dots, Q_l, \dots, Q_0) \mathrm{d}Q_L \dots \mathrm{d}Q_l \dots \mathrm{d}Q_0 \\ &= \int p(Q_L|Q_{L-1}) \dots \int p(Q_l|Q_{l-1}) \dots \\ &\int p(Q_1|Q_0) p(Q_0) \mathrm{d}Q_{L-1} \dots \mathrm{d}Q_{l-1} \dots \mathrm{d}Q_0 \end{aligned} \tag{4}$$

on the conditional probability density function [14]. The metamodel can comprise linear or non-linear relationships and is calibrated by means of Bayesian regression or maximum likelihood estimation. It allows to replace high-fidelity computations by low-fidelity ones. Depending on the choice of the metamodel, dependence-based measures or information-based measures between the high-fidelity model and the low-fidelity models may provide more meaningful information than the Pearson correlation.

3. Reliability estimation

Reliability estimation deals with the evaluation of the failure probability

$$P_F = \int_F p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = \int I_{g<0}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}, \tag{5}$$

where $F = \{\boldsymbol{\theta} \in \mathbb{R}^n | g(\boldsymbol{\theta}) < 0\}$ denotes the failure domain,

$$I_{g<0}(\boldsymbol{\theta}) = \begin{cases} 1, & \text{if } g(\boldsymbol{\theta}) < 0, \\ 0, & \text{if } g(\boldsymbol{\theta}) \geq 0 \end{cases} \tag{6}$$

is the indicator function and $p(\boldsymbol{\theta})$ the joint probability density function of the random vector $\boldsymbol{\Theta}$. In general, the performance function $g(\boldsymbol{\theta})$ is not known exactly, but is computed by numerical approximation.

3.1. Importance sampling

The aim of importance sampling is to reduce the variance of Monte Carlo simulation by sampling from an alternative density, the importance sampling density. To this end, the importance sampling density $p_{IS}(\boldsymbol{\theta})$ is introduced into Eq. (5):

$$P_F = \int I_{g<0}(\boldsymbol{\theta}) \frac{p(\boldsymbol{\theta})}{p_{IS}(\boldsymbol{\theta})} p_{IS}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}. \tag{7}$$

The importance sampling estimate is then computed from the weighted average

$$\hat{P}_{F,IS} = \frac{1}{N}\sum_{i=1}^{N} I_{g<0}(\boldsymbol{\theta}^i) \frac{p(\boldsymbol{\theta}^i)}{p_{IS}(\boldsymbol{\theta}^i)}, \tag{8}$$

where the N samples $\boldsymbol{\theta}^i$, $i = 1, \ldots, N$, are drawn from $p_{IS}(\boldsymbol{\theta})$. The optimal importance sampling density is given by

$$p_{IS}(\boldsymbol{\theta}) = \frac{I_{g<0}(\boldsymbol{\theta})p(\boldsymbol{\theta})}{P_F} \tag{9}$$

which however requires already the knowledge of P_F and is thus unfeasible.

In the cross-entropy method, an optimal importance sampling density is computed within a family of densities $\tilde{p}(\boldsymbol{\theta}, \boldsymbol{\nu})$ with parameter vector $\boldsymbol{\nu}$ by minimizing the Kullback–Leibler divergence (i.e. the cross-entropy)

$$\int p_{IS}(\boldsymbol{\theta}) \ln p_{IS}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} - \int p_{IS}(\boldsymbol{\theta}) \ln \tilde{p}(\boldsymbol{\theta}, \boldsymbol{\nu})\mathrm{d}\boldsymbol{\theta} \tag{10}$$

to the optimal importance sampling density [25]. This amounts to maximizing the expression

$$\int I_{g<0}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) \ln \tilde{p}(\boldsymbol{\theta}, \boldsymbol{\nu})\mathrm{d}\boldsymbol{\theta}. \tag{11}$$

The estimation of this quantity requires again an importance sampling procedure, which can be based on the same family of densities $\tilde{p}(\boldsymbol{\theta}, \boldsymbol{\nu})$. This leads to an iterative scheme for the determination of the optimal parameter $\boldsymbol{\nu}$, cf. [2].

3.2. Importance splitting

The idea of importance splitting in reliability estimation is to discard samples that are far away from the failure domain and to split at least some of the remaining ones [26]. This procedure is repeated gradually by introducing thresholds and retaining only those samples that pass the threshold. Importance splitting can also be interpreted as a special adaptive importance sampling procedure, where the final samples define the importance sampling density.

One of the most prominent methods that applies importance splitting is subset simulation [3], where the thresholds are intermediate levels of the performance function, Markov chain Monte Carlo simulation is applied to split the remaining samples and the failure probability is computed as a telescoping product of successive conditional probabilities.

The moving particles algorithm [4] can be considered as a special case of subset simulation with a maximum number of subsets. This algorithm has the advantage that the quantity to be estimated, the number of moves of the particles, is a Poisson distributed random variable. Thus, properties of the Poisson distribution can be used to combine information fusion with efficient reliability calculation.

The moving particles algorithm yields an estimate of P_F. It starts with an initial Monte Carlo simulation with N samples. These initial samples – the particles – are then moved to the failure domain by the following procedure: The values $g(\boldsymbol{\theta}_j)$, $j = 1, \ldots, N$, of the N samples are ranked. The sample with the maximum value of the performance function is moved to a new position in the sample space with reduced value of the performance function by the following procedure:

A Markov chain Monte Carlo simulation (MCMC) is carried out starting from one of the remaining samples and the next state of the Markov chain is accepted, if the value of the performance function is reduced. The Markov chain can be generated e.g. by application of the Metropolis–Hastings algorithm or by direct sampling from a normal transition kernel.

For each initial sample, the number M of moves until the sample reaches the failure domain is count. As has been shown in [4], the number of moves to get an initial sample into the failure domain follows a Poisson distribution with parameter $\lambda = -\log P_F$. The estimator for the parameter of the Poisson distribution is obtained from $\lambda = E[M]$ as

$$\hat{\lambda} = \frac{\sum_{j=1}^{N} M_j}{N}, \tag{12}$$

where M_j, $j = 1, \ldots, N$ denotes the number of moves until the initial sample j reaches the failure state.

In order to obtain an unbiased estimate, it is mandatory that the trajectories of the Poisson process generated from the initial samples remain independent until the samples finally reach the failure domain. In [27], two means are proposed to maintain the independence:

- Burn-in: The Markov chain Monte Carlo simulation is carried out with a burn-in period. The burn-in should ensure the independence of the candidate and the starting point of the Markov chain.
- Seed avoidance: Repeated use of the same starting point for the Markov chain should be avoided. Once a sample has been used as starting point, the sample and its offspring should not be used as starting point again.

The coefficient of variation for the failure probability estimated with the moving particles algorithm is given by

$$\delta_{mp} = \sqrt{\frac{-\log P_F}{N}}, \tag{13}$$

cf. [4], and the average number of function evaluations is

$$N_{mp} = N(1 - T \log P_F), \tag{14}$$

where the first term accounts for the initial Monte Carlo simulation and the second term for the Markov chain samples (with burn-in period T). A burn-in period of $T = 5$ and $N = 1000$ initial samples were used in the examples.

Table 1
Relative error and coefficient of variation (c.o.v.), single level method.

Order	$r = 1$		$r = 3$		$r = 5$	
	Rel. error [%]	C.o.v.	Rel. error [%]	C.o.v.	Rel. error [%]	C. o. v.
$n = 1$	20	0.0013	499	0.002	21510	0.022
$n = 2$	5	0.0015	59	0.006	291	0.017
$n = 3$	2	0.0015	23	0.006	81	0.010
$n = 4$	1	0.0018	12	0.005	39	0.006
$n = 5$	1	0.0017	8	0.005	23	0.007

4. Information fusion and importance sampling

Both information fusion by additive combination and multiplicative combination of the model outputs can be combined with importance sampling for reliability estimation. For additive information fusion, the quantity of interest is the indicator function $I_{g_l<0}(\boldsymbol{\theta})$, $l = 1, \dots, L$, while for multiplicative information fusion, the quantity of interest is the value $g_l(\boldsymbol{\theta})$, $l = 1, \dots, L$, of the performance function, which is interpreted as a random variable.

4.1. Additive information fusion

For additive combination of the model outputs according to Eq. (1), the importance sampling density for each expectation has to focus on the differences of the performance functions of neighbouring models. An ordered set of performance functions $g_l(\boldsymbol{\theta})$, $l = 0, \dots, L$ is needed, where $l = 0$ denotes the model with lowest fidelity and $l = L$ the high-fidelity model. Alternatively, the importance sampling estimators might be computed separately for each model and then the differences of the estimators are weighted, cf. [8].

Example 1, additive multifidelity method: Consider the circular performance function $g(\theta_1, \theta_2) = r - \sqrt{\theta_1^2 + \theta_2^2}$ with parameter $r > 0$ and a standard normal distribution for the random variables θ_1, θ_2. The limit state function $g(\theta_1, \theta_2) = 0$ is approximated by a regular convex polygon with $n + 3$ facets and orbiting radius r and the corresponding performance functions $g_n(\theta_1, \theta_2)$ represent the oriented Euclidean distance of a sample from the approximated limit state function.

Table 1 displays the convergence of the single level approximations with approximation order n. As can be seen, the relative error decreases quickly. For lower failure probabilities, a high approximation order is necessary, e.g. for $r = 5$, an octogonal approximation is not sufficient to reduce the relative error below 10%.

For the multifidelity method with additive information fusion, the expectations for the differences of the failure probabilities are computed by an adaptive cross-entropy based importance sampling scheme that utilizes a Gaussian mixture model to represent the importance sampling density, cf. [28]. As Table 2 reveals, the relative error of the multifidelity method is similar to the single level method for $r = 1$, lower for $r = 3$ and $n > 2$ and higher for $r = 5$. The coefficient of variation of the estimator for the differences $\Delta_{n+1n} = E[g_{n+1}(\theta_1, \theta_2) - g_n(\theta_1, \theta_2)]$ is lower than that of the single level estimator only for $r = 5$ and low approximation order ($n < 4$). In contrast to multilevel methods, the coefficient of variation *increases* with increasing approximation order n. Thus, the multifidelity method is not efficient in this case.

The reason for these results of the multifidelity method can be seen from Fig. 1 that displays the samples obtained from the importance sampling density for the differences Δ_{21} and Δ_{43} for $r = 1$. As can be seen, the importance sampling density focus in both cases on the domain where the approximated performance functions differ in sign; however, for Δ_{43}, these domains are so small that many samples are still outside of these domains. For increasing approximation order n, the number of separated domains that have to be covered by the importance sampling density increases and the domains become smaller. Thus, for higher n, it is more difficult to cover these domains very well.

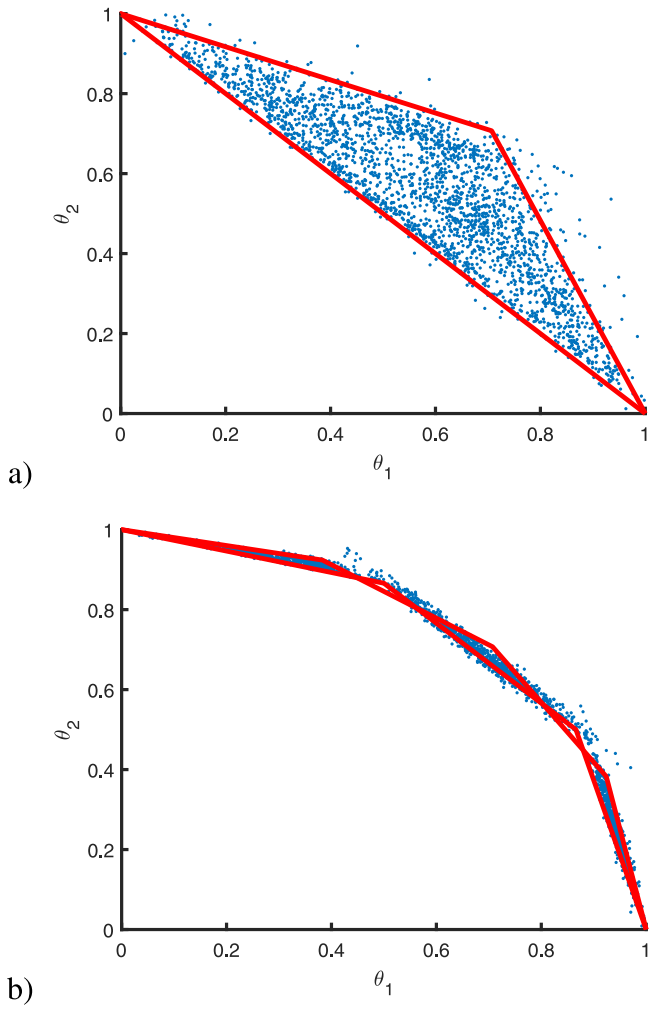

Fig. 1. Samples from the importance sampling density for differences between approximation orders. (a) Δ_{21}, (b) Δ_{43}.

This leads to the increase of the coefficient of variation and also to the increase of the relative error for larger values of r.

Next, the weighted multifidelity method proposed in [20] is considered, where a weighted sum of importance sampling estimates is computed. This approach requires unbiased importance sampling estimates. Thus, the importance sampling densities are calibrated by means of the different approximations of the performance function $g_i(\boldsymbol{\theta})$, $i = 1, \dots, n$; however, the samples generated with the different importance sampling densities are evaluated with respect to the performance function $g_n(\boldsymbol{\theta})$ of highest fidelity. This is different from the other methods presented in this paper, where the generated samples are either evaluated by the performance function of the same or the next higher level. In order to be efficient, a reduction of the coefficient of variation for the weighted multifidelity estimator compared to the single level approach is therefore mandatory, such that in total, less samples are evaluated by the high-fidelity performance function.

Table 3 indicates that the coefficient of variation of the weighted multifidelity method is indeed smaller than that of the single level approach. It decreases with increasing high-fidelity approximation order n. For $n = 5$, the coefficient of variation of the single level approach is almost twice as large as that of the weighted multifidelity method. This is due to the fact that with increasing n the approach weights more estimators, namely a total of n, one for each level.

Table 2
Relative error, mean value and coefficient of variation for the differences between approximation orders.

Order	$r = 1$			$r = 3$			$r = 5$		
	Err. [%]	Mean	C.o.v.	Err. [%]	Mean	C.o.v.	Err. [%]	Mean	C.o.v.
Δ_{21}	5	0.0914	0.003	60	0.0489	0.0029	337	7.97E−04	0.0099
Δ_{32}	2	0.0195	0.007	20	0.0045	0.0075	108	8.52E−06	0.0083
Δ_{43}	1	0.0079	0.0151	6	0.0015	0.0135	57	1.93E−06	0.0227
Δ_{54}	0	0.0042	0.0244	0	7.53E−04	0.0214	34	8.39E−07	0.0318

Table 3
Relative error and coefficient of variation, weighted multifidelity method.

Order	$r = 1$		$r = 3$		$r = 5$	
	Rel. error [%]	C.o.v.	Rel. error [%]	C.o.v.	Rel. error [%]	C. o. v.
$n = 2$	5	0.0014	59	0.0043	293	0.02
$n = 3$	2	0.001	23	0.0038	81	0.0095
$n = 4$	1	9.6E−4	12	0.0031	39	0.0051
$n = 5$	1	8.3E−4	8	0.0027	23	0.0046

4.2. Multiplicative information fusion

For multiplicative information fusion, the importance sampling density is based on the model with lower fidelity and pairs of output quantities for neighbouring models are obtained for samples generated by means of the importance sampling density. These pairs of samples are then utilized to find a functional expression by regression that relates the high-fidelity model output to the low-fidelity model output.

Example 1, multiplicative bifidelity method: Consider the same performance function and its approximation as before and the following bifidelity method: For the lower approximation order (n) the importance sampling density is calibrated by the same cross-entropy based importance sampling algorithm applied before. After that, 100 samples (out of 50000 samples) with highest weights are identified and only for these samples, the performance function of approximation order $n + 1$ is evaluated. A linear relation between the 100 values of the low-fidelity performance function and the high-fidelity performance function is then calibrated by regression. By means of this linear relation, approximations of the high-fidelity performance function are computed for the remaining 49900 samples.

Table 4 summarizes the relative error and the coefficient of variation obtained for the bifidelity method. It can be seen that the relative error of the bifidelity method is comparable to that of the higher approximation order $n + 1$. Thus, the bifidelity method considerably improves the approximations of the failure probability. The coefficient of variation of the bifidelity estimator is higher than that of the single level estimator. It scales with the coefficient of determination R^2 and thus might be attributed to the regression error. For higher approximation order n and for lower values of the radius r, the coefficient of determination increases because the differences between two successive levels are smaller in these cases. Thus, the correlation between the values of the performance functions is higher which reduces the coefficient of variation.

Comparing the results for example 1 as summarized in Tables 2–4 the following conclusions can be drawn: The combination of additive information fusion using telescoping sums with importance sampling requires an importance sampling density that focuses on differences between approximations of successive order; a reduction of the coefficient of variation is then difficult to achieve (Table 2). Implementing additive information fusion by a weighted sum of importance sampling estimators leads to a reduction of the coefficient of variation (Table 3), but requires all samples generated by the different importance sampling densities to be evaluated by the high-fidelity performance function. The combination of multiplicative information fusion with importance sampling leads to mean square errors that are comparable to those of the next higher level (Table 4). A reduction of the coefficient of variation is achieved, if the regression error is small. For linear regression, this is the case if the approximations of successive order are highly correlated.

5. Information fusion and importance splitting

For importance splitting based on the moving particles algorithm, the number of moves of the particles are the quantity of interest for both additive and multiplicative information fusion. For multiplicative information fusion, the adequate statistical related to the moving particles method is a bivariate (or more general a multivariate) Poisson distribution, whose parameters are estimated from a rather large number of particles for which the moves are obtained with the low-fidelity model and few particles for which both the low- and the high-fidelity model must be evaluated.

In order to obtain a high variance reduction by the multifidelity estimator it is important that the correlation of the number of moves between models of different fidelity is high. Therefore, it is necessary to closely relate the Markov chains obtained with different models that are employed in each move of a particle. However, the stationary distributions that the Markov chains approach are different due to application of different models when evaluating the performance function.

To solve this issue, two approaches have been proposed in the literature. In [29], the random input vector of a high- and a low-fidelity model is divided into a coarse part containing the joint input parameters of both models and a fine part comprising additional input parameters of the high-fidelity model. The Markov chain is advanced first on the coarse part of the input vector and after that, the fine part is generated. Although this approach is useful in many situations, it cannot be applied in the context of the moving particles algorithm, as the final criterion for the acceptance of a candidate – the reduction of the value of the performance function – is different for the high- and the low-fidelity model. Acceptance by the low-fidelity model does not imply acceptance by the high-fidelity model and vice versa. Therefore, fixing the coarse part of the input parameters by means of the low fidelity model introduces a bias when the high fidelity model is employed in the moving particles algorithm, as a part of the intermediate failure domains cannot be reached. An alternative approach is the method proposed recently in [30], in which only the initial values and the random numbers used during the Markov chain simulation are the same, but not the intermediate states. This method can be successfully applied to the moving particles algorithm.

5.1. Additive information fusion

For the moving particles algorithm, denote by M_l the number of moves when the numerically computed performance function $g_l(\boldsymbol{\theta})$ is applied. For additive information fusion with a telescoping sum, the estimator for the Poisson distribution parameter at level L and thus for $-\log P_{F_L}$, where P_{F_L} denotes the failure probability computed with the approximated performance function $g_L(\boldsymbol{\theta})$ at approximation order L, is obtained from

$$E[M_L] = E[M_0] + \sum_{l=1}^{L} E[M_l - M_{l-1}] \tag{15}$$

and reads

$$\hat{M}_L^M = \frac{1}{N_0}\sum_{i=1}^{N_0} M_0^{(i)} + \sum_{l=1}^{L}\frac{1}{N_l}\sum_{i=1}^{N_l}(M_l^{(i)} - M_{l-1}^{(i)}). \tag{16}$$

The corresponding single level estimator is

$$\hat{M}_L^S = \frac{1}{N_S}\sum_{i=1}^{N_S} M_L^{(i)}. \tag{17}$$

Table 4
Relative error, coefficient of variation and coefficient of determination for the bifidelity method.

Order	$r = 1$			$r = 3$			$r = 5$		
	Err. [%]	C.o.v.	R^2 [%]	Err. [%]	C.o.v.	R^2 [%]	Err. [%]	C.o.v.	R^2 [%]
$n = 1$	4	0.0226	85.31	42	0.3854	35.61	68	2.2314	24.50
$n = 2$	2	0.0059	98.78	10	0.1067	75.68	30	0.4132	60.46
$n = 3$	1	0.0035	99.67	7	0.0461	93.48	12	0.1184	86.75
$n = 4$	1	0.0024	99.89	5	0.023	97.42	7	0.0974	90.81
$n = 5$	1	0.002	99.95	5	0.0164	98.80	6	0.0513	93.74

The mean square error for the single level estimator is

$$E[(\hat{M}_L^S - E[M])^2] = V(\hat{M}_L^S) + \left(E[\hat{M}_L^S] - E[M]\right)^2 = N_S^{-1} V(M_L) + \left(E[M_L] - E[M]\right)^2 \tag{18}$$

where $V(.)$ denotes the variance operator. The corresponding error for the estimator (16) is

$$E[(\hat{M}_L^M - E[M])^2] = \sum_{l=0}^{L} N_l^{-1} V(Y_l) + \left(E[M_L] - E[M]\right)^2, \tag{19}$$

where $Y_0 = M_0$ and $Y_l = M_l - M_{l-1}$, $l > 0$. In both expressions, the last term is the numerical approximation error, while the first term is the statistical error.

In the following, we analyse the multifidelity version of the algorithm, i.e. for a model hierarchy that is based on a series of decreasing numerical approximation errors h_l, $l = 0, \dots, L$. For a given overall precision $\epsilon > 0$ such that the mean square error is less than ϵ^2, we wish to achieve a statistical as well as an approximation error less than $\epsilon^2/2$. For the approximation error, we have

$$\left(E[M_L] - E[M]\right)^2 < \frac{\epsilon^2}{2}. \tag{20}$$

If

$$\left|\log P_F - \log P_{F_l}\right| = \mathcal{O}(h_l), \tag{21}$$

we obtain the condition $h_l = \mathcal{O}(\epsilon)$ for the approximation error. For the statistical error of the single level estimator, we impose

$$N_S^{-1} V(M_L) = \mathcal{O}(\epsilon^2), \tag{22}$$

which leads, if $V(M_L)$ is approximately constant, to $N_S = \mathcal{O}(\epsilon^{-2})$ for the number of samples, the total number of function estimations being thus $N_S(1 - T \log P_{F_L})$, cf. Eq. (14). Suppose that the computational cost $C(M_l^{(i)}) = (1 - T \log P_{F_l})$ to compute a single sample is of order $\mathcal{O}(h_l^{-r})$ for some $r > 0$. Then the computational cost for the single level estimator is of order

$$C(M_L^S) = \mathcal{O}(N_S h_L^{-r}) = \mathcal{O}(\epsilon^{-2-r}). \tag{23}$$

For the multifidelity estimator, the total computational cost is

$$C(M_L^M) = \sum_{l=0}^{L} N_l C(Y_l^{(i)}). \tag{24}$$

Fixing the total computational cost, the statistical error becomes minimal, if $N_l = \lambda \sqrt{V(Y_l)/C(Y_l^{(i)})}$, with

$$\lambda = \epsilon^{-2} \sum_{l=0}^{L} \sqrt{V(Y_l) C(Y_l^{(i)})} \tag{25}$$

if the statistical error should be equal to ϵ^2. In this case, the total computational cost of the multifidelity algorithm is

$$C(M_L^M) = \epsilon^{-2} \left(\sum_{l=0}^{L} \sqrt{V(Y_l) C(Y_l^{(i)})} \right)^2. \tag{26}$$

If Eq. (21) holds and

$$\left|\log P_F^2 - \log P_{F_l}^2\right| = \mathcal{O}(h_l), \tag{27}$$

then from the fact that M_l follows a Poisson distribution, the variance of Y_l is of $\mathcal{O}(h_l)$. Depending on the increase of the costs $C(M_l^{(i)})$ with h_l, which is described by the exponent r, the first or the last term in the sum in (26) will dominate. In fact, if $0 < r < 1$, the costs increase less than the variance will decrease and thus the sum in (26) is dominated by $\sqrt{V(Y_0) C(Y_0^{(i)})}$, so that $C(M_L^M) = \mathcal{O}(\epsilon^{-2})$. On the other hand, if $r > 1$, then the sum is dominated by $\sqrt{V(Y_L) C(Y_L^{(i)})}$ and $C(M_L^M) = \mathcal{O}(\epsilon^{-2} h_L h_L^{-r}) = \mathcal{O}(\epsilon^{-1-r})$.

Thus, compared with the computational cost of order $\mathcal{O}(\epsilon^{-2-r})$ for the single level estimator, considerable savings are obtained with the multifidelity estimator. This can be attributed to the fact that the order of magnitude for the cost of the single level estimator involves the product of the variance $V(Y_0)$ and the cost $C(Y_L^{(i)})$.

Conditions (21) and (27) depend directly on the approximation of the performance function. For instance, if information fusion and information filtering are combined, such that $|g_h(x) - g(x)| < |g_h(x)|$ or $|g_h(x) - g(x)| \leq h$, one has $\left|P_F - P_{F_h}\right| < h$ ([11], Lemma 3.4). From this result, a direct calculation shows that $\left|\log P_F - \log P_{F_h}\right| < C_1 h$ and $\left|\log P_F^2 - \log P_{F_h}^2\right| < C_2 h$, cf. [9].

Example 2: Stochastic heat equation with random heat source

Consider the linear stochastic partial differential equation

$$\mathrm{d}u = \frac{\partial^2 u}{\partial x^2} \mathrm{d}t + \theta u \mathrm{d}t + \sigma \mathrm{d}W_t \tag{28}$$

with parameters θ and σ, where $\mathrm{d}W$ are the increments of a standard cylindrical Wiener process

$$W_t = \sum_{k=1}^{\infty} W_t^k \sin(k\pi x), \tag{29}$$

and W_t^k are standard independent Brownian motions.

The boundary conditions are $u(x = 0, t) = u(x = 1, t) = 0$ and the initial condition reads

$$u(x, 0) = \sqrt{2} \sum_{k=1}^{\infty} \sin(k\pi x), \tag{30}$$

so that the solution of the stochastic partial differential equation decouples and can be written as

$$u(x, t) = \sqrt{2} \sum_{k=1}^{\infty} u_k(t) \sin(k\pi x), \tag{31}$$

where the coefficients $u_k(t)$ are described by the linear stochastic ordinary differential equations

$$\mathrm{d}u_k = (-\pi^2 k^2 + \theta) u_k \mathrm{d}t + \sigma \mathrm{d}W_t^k. \tag{32}$$

Their solutions at time t are

$$u_k(t) = \exp((-\pi^2 k^2 + \theta)t) + \xi_t^k, \tag{33}$$

where ξ_t^k is a normal random variable with mean zero and variance

$$\frac{\sigma^2 (1 - \exp(2(-\pi^2 k^2 + \theta)t))}{2(\pi^2 k^2 - \theta)}. \tag{34}$$

Failure is assumed to occur if $u(x = 0.5, t = T)$ is larger than a given threshold. In the following, $\theta = 1$, $\sigma = 1$, $T = 1/50$ were set.

The solution of the stochastic partial differential Eq. (28) is approximated by truncating the solution at $k = k_{max}$. In addition, time integration is applied in order to solve the stochastic ordinary differential Eqs. (32). The approximation parameters for this problem are the

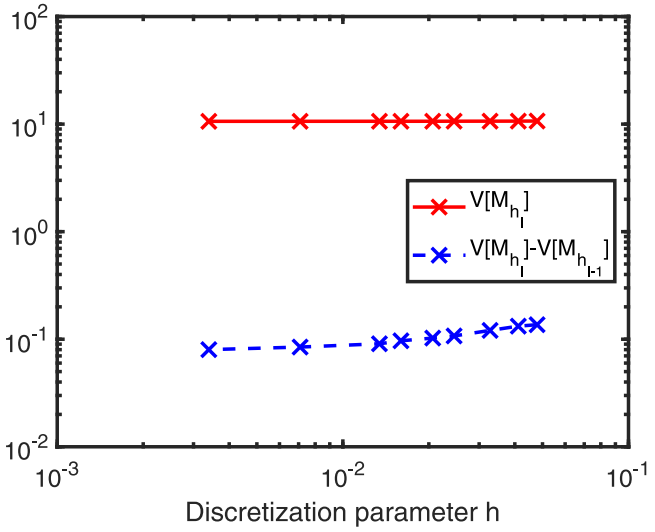

Fig. 2. Variance of M_{h_l} and $M_{h_l} - M_{h_{l-1}}$.

truncation parameter k_{max} and the time step. In order to relate them to a single discretization parameter h, a reference failure probability has been computed from the solution given by Eqs. (33) and (34) and a sufficiently high value of k_{max}. With this reference value and given values for the two approximation parameters, the single discretization parameter h is obtained as relative error with respect to a direct Monte Carlo simulation, where the number of samples was so high that the estimation error was negligible.

Fig. 2 displays the variance of the single level and the multilevel estimator for the number of moves in dependence of the approximation parameter h is computed by comparing the exact and the approximate failure probability. The threshold value for failure has been set to 1.5. It can be seen that the variance of the multilevel estimator is smaller than that of the single level estimator and that it decreases with decreasing discretization parameter h, while the variance of the single level estimator is nearly independent of h. As the number of moves is a Poisson distributed random variable, the same result holds for the mean values.

5.2. Multiplicative information fusion

For information fusion with multiplicative combination of the model outputs, consider again without loss of generality the bifidelity situation with a high-fidelity and a low-fidelity model. M_h and M_l denote the random variables for the moves of the particles related to the high-fidelity and the low-fidelity model, respectively. As the low-fidelity model is assumed to be computationally more efficient, suppose that samples M_{l_i}, $i = 1, \dots, N$, are generated for particles with the low-fidelity models and pairs of samples (M_{l_i}, M_{h_i}), $i = 1, \dots, M$ that were obtained with both the low- and the high-fidelity model, where $M \ll N$. The objective is to estimate the parameters λ_l and λ_h of the two dependent Poisson distributed random variables M_h and M_l. To this end, three independent Poisson distributed random variables are introduced and the parameters of the three Poisson distributions are estimated by maximum likelihood estimation via the expectation maximization algorithm. The corresponding iteration scheme is described in [31].

Example 3: Consider the same problem as described in example 2. In this example only the truncation order k_{max} has been varied. The moving particles method has been applied with 1000 initial samples (particles). For the bifidelity method described above, only ten particles were considered, for which the number of moves to reach the failure domain were computed with both the low- and high-fidelity model. Fig. 3 summarizes the results for the single level and the bifidelity method. For low truncation order, the multiplicative bifidelity method reduces the relative error compared to the single level method. Moreover, due to the maximum likelihood estimation via expectation maximization, also the standard deviation of the estimator is considerably reduced. This demonstrates the advantages of the bifidelity method over the single level moving particles algorithm.

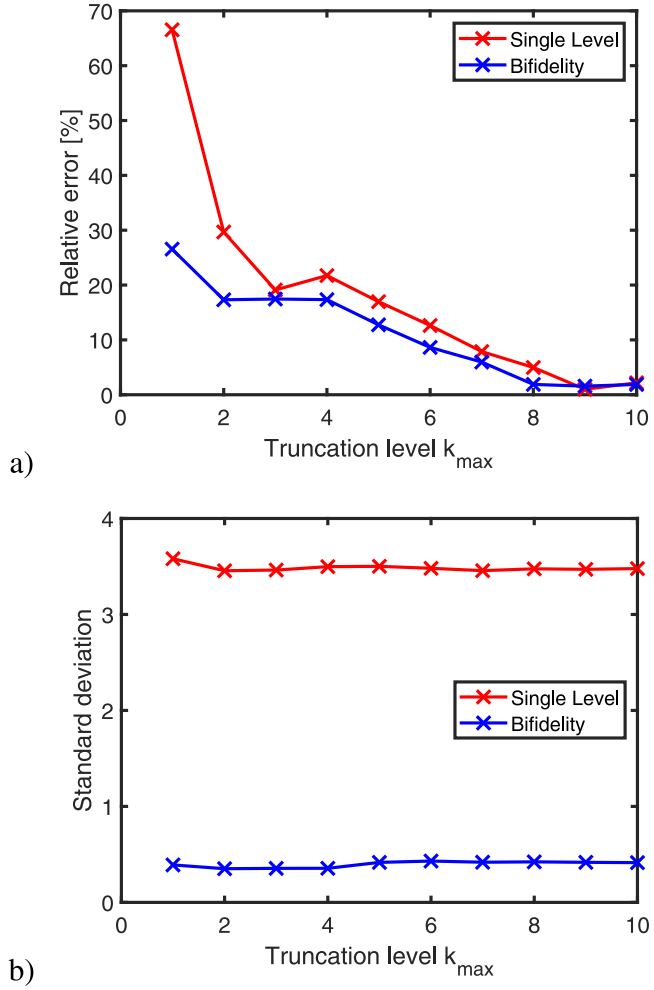

Fig. 3. Relative error and standard deviation for the single level estimator and the bifidelity method with multiplicative information fusion.

Example 4: Burgers' equation

Finally, additive and multiplicative information fusion combined with importance splitting is applied to a nonlinear partial differential equation, namely Burgers' equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \alpha(x,\theta)\frac{\partial^2 u}{\partial x^2}, \tag{35}$$

where the stochastic viscosity $\alpha(x,\theta)$ is given by the truncated Karhunen–Loève expansion

$$\alpha(x,\theta) = 1 + \sum_{i=1}^{M} \sqrt{\lambda_i} f_i(x)\xi_i(\theta), \tag{36}$$

where $\xi(\theta) = [\xi_1(\theta), \xi_2(\theta), \dots, \xi_M(\theta)]$ are independent standard normal random variables and λ_i, $f_i(x)$, $i = 1, \dots, M$, are the eigenvalues and eigenvectors of a Fredholm integral equation of 2nd kind for the exponential covariance kernel

$$Cov(x_1, x_2) = \sigma^2 \exp(-\|x_1 - x_2\|), \tag{37}$$

where $\sigma = 1$.

Table 5
Relative error and standard deviation for the single level, the additive multifidelity and the multiplicative bifidelity method.

Grid points	Single level		Additive multifidelity		Multiplicative bifidelity	
	Rel. error [%]	Std. dev.	Rel. error [%]	Std. dev.	Rel. error [%]	Std. dev.
$n_{max} = 40$	15.32	2.06	15.32	2.06	8.81	0.37
$n_{max} = 42$	11.57	2.04	11.83	0.2	7.90	0.37
$n_{max} = 44$	6.75	2.02	8.87	0.18	6.95	0.37
$n_{max} = 46$	4.74	2.01	6.26	0.16	5.84	0.37
$n_{max} = 48$	2.72	2.01	3.98	0.15	4.65	0.37

The boundary conditions are $u(x = 0, t) = 0$ and $u(x = \ell, t) = 0$, the initial condition is $u(x, t = 0) = \sin(\frac{\pi x}{\ell})$.

The solution is represented by a polynomial chaos expansion

$$u(x, t, \xi) = \sum_{i=0}^{P} u_i(x, t)\psi_i(\xi), \tag{38}$$

cf. [32], and the partial differential equation that governs the expansion coefficients reads as follows:

$$\sum_{i=0}^{P} \frac{\partial u_i}{\partial t}\psi_i(\xi) + \sum_{i=0}^{P} u_i\psi_i(\xi) \sum_{j=0}^{P} \frac{\partial u_j}{\partial x}\psi_j = (1 + \sum_{i=1}^{M} \sqrt{\lambda_i} f_i(x)\xi_i(\theta))(\sum_{i=0}^{P} \frac{\partial^2 u_i}{\partial x^2}\psi_i(\xi)). \tag{39}$$

Application of a Galerkin scheme with respect to the polynomial chaos and a finite difference approximation in space and time leads to the equation

$$\begin{aligned} &\frac{1}{\Delta t} \sum_{i=0}^{P} \sum_{j=0}^{P} b_{ij}(u_j^{(m,n+1)} - u_j^{(m,n)}) \\ &+ \frac{1}{2(\Delta x)} \sum_{i=0}^{P} \sum_{j=0}^{P} \sum_{k=0}^{P} c_{ijk} u_j^{(m,n)}(u_k^{(m+1,n)} - u_k^{(m-1,n)}) \\ &= \frac{1}{(\Delta x)^2} \sum_{i=0}^{P} \sum_{j=0}^{P} \sum_{k=1}^{M} d_{ijk}\sqrt{\lambda_k} f_k(x)(u_j^{(m+1,n)} - 2u_j^{(m,n)} + u_j^{(m-1,n)}) \\ &+ \frac{1}{(\Delta x)^2} \sum_{i=0}^{P} \sum_{j=0}^{P} b_{ij}(u_j^{(m+1,n)} - 2u_j^{(m,n)} + u_j^{(m-1,n)}) \end{aligned} \tag{40}$$

for the discretized expansion coefficients $u_i^{(m,n+1)}$, cf. [33], where

$$\begin{aligned} b_{ij} &= \int \psi_i(x)\psi_j(x)\mathrm{d}P_G(x), \\ c_{ijk} &= \int x_k\psi_i(x)\psi_j(x)\psi_k(x)\mathrm{d}P_G(x), \\ d_{ijk} &= \int x_k\psi_i(x)\psi_j(x)\mathrm{d}P_G(x) \end{aligned} \tag{41}$$

and $\mathrm{d}P_G(x)$ denotes the M-dimensional standard normal measure.

Failure is assumed to occur if $u(x = \ell/2, t = T)$ is larger than a given threshold (0.42). Table 5 summarizes the results obtained with the single level method, the multifidelity method with additive information fusion and the bifidelity method with multiplicative information fusion. Only the time step has been varied by adapting the number n_{max} of steps until the final time T is reached, while the other parameters were fixed.

For all methods, the relative error decreases rapidly if the approximation parameter n_{max} is increased. However, this decrease is less pronounced for the bifidelity method with multiplicative information fusion.

With regard to the standard deviation of the estimator, Table 5 demonstrates again that the single level method leads to the highest standard deviation and that the standard deviation for the multifidelity method decreases with increasing values of the approximation parameter n_{max}. For the bifidelity method, the standard deviation is nearly independent of n_{max} and nearly twice as large as the standard deviation of the multifidelity method after the first level. The standard deviation of the bifidelity estimator can be further reduced, if the number of high-fidelity samples is increased.

6. Conclusions

This paper combines additive and multiplicative information fusion with importance sampling and importance splitting in order to efficiently estimate the probability of failure.

Obtaining a good importance sampling density is still a challenging task, especially in high dimensions. This is also the case, if the importance sampling density should focus on level differences. It therefore appears to be more advantageous to combine different importance sampling estimators by a weighted sum, even if this approach requires unbiased estimators and thus the evaluation of all samples by the high-fidelity performance function. Multiplicative information fusion might have advantages compared to additive information fusion as it reduces the relative error and requires less samples to be evaluated by the high-fidelity performance function. However, multiplicative information fusion relies on establishing a relationship between the values of the performance function obtained with the model of lower and of higher fidelity.

Importance splitting is a rather robust method to yield accurate reliability estimates. Taking advantage of Poisson process theory, both additive and multiplicative information fusion can be combined with the moving particles method in order to further increase the efficiency of the moving particles method.

The proposed methods help to balance the approximation error and the statistical error by information fusion. It is demonstrated that these methods may lead to a considerable increase in efficiency. The approaches can be extended by taking the data and model error in a Bayesian setting into account. It is also noted that the methods presented in the paper can be applied to sensitivity analysis and Bayesian inference.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Germany – project no. 428469142 – as part of the priority programme "Polymorphic uncertainty modelling for the numerical design of structures" (SPP 1886) which is gratefully acknowledged by the authors.

References

[1] G. Schuëller, H. Pradlwarter, P. Koutsourelakis, A critical appraisal of reliability estimation procedures for high dimensions, Probab. Eng. Mech. 19 (4) (2004) 463–474.
[2] N. Kurtz, J. Song, Cross-entropy-based adaptive importance sampling using Gaussian mixture, Struct. Saf. 42 (2013) 35–44.
[3] S.K. Au, J. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Probab. Eng. Mech. 16 (2001) 263–277.
[4] A. Guyader, N. Hengartner, E. Matzner-Løber, Simulation and estimation of extreme quantiles and extreme probabilities, Appl. Math. Optim. 64 (2011) 171–196.

[5] C. Proppe, Markov chain methods for reliability estimation, in: Proceedings International Conference on Structural Safety an Reliability, ICOSSAR, Vienna, 2017.
[6] M.B. Giles, Multilevel Monte Carlo path simulation, Oper. Res. 56 (2008) 337–361.
[7] F. Müller, D.W. Meyer, P. Jenny, Solver-based vs. grid-based multilevel Monte Carlo for two phase flow and transport in random heterogeneous porous media, Comput. Phys. 268 (1) (2014) 39–50.
[8] B. Peherstorfer, K. Willcox, M. Gunzburger, Optimal model management for multifidelity Monte Carlo estimation, SIAM J. Sci. Comput. 38 (5) (2016) A3163–A3194.
[9] C. Proppe, Reliability estimation with multi-fidelity simulation methods, in: M. Beer, E. Zio (Eds.), Proc. 29th European Safety and Reliability Conference, Hannover, Germany, 22-26 September, 2019.
[10] B. Peherstorfer, K. Willcox, M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, SIAM Rev. 60 (3) (2018) 550–591.
[11] D. Elfverson, F. Hellman, A. Målqvist, A multilevel Monte Carlo method for computing failure probabilities, SIAM/ASA J. Uncertain. Quantif. 4 (1) (2016) 312–330.
[12] L. Le Gratiet, Multi-fidelity Gaussian Process Regression for Computer Experiments, (Ph.D. thesis), Université Paris-Diderot, 2013.
[13] P. Perdikaris, D. Venturi, J.O. Royset, G.E. Karniadakis, Multi-fidelity modelling via recursive co-kriging and Gaussian-Markov random fields, Proc. Roy. Soc. A 471 (2015) 20150018.
[14] J. Biehler, Efficient Uncertainty Quantification for Large-Scale Biomechanical Models Using a Bayesian Multi-Fidelity Approach, (Ph.D. thesis), TU Munich, 2016.
[15] J. Li, D. Xiu, Evaluation of failure probability via surrogate models, J. Comput. Phys. 229 (2010) 8966–8980.
[16] P. Chen, A. Quarteroni, Accurate and efficient evaluation of failure probability for partial different equations with random input data, Comput. Methods Appl. Mech. Engrg. 267 (2013) 233–260.
[17] J. Li, J. Li, D. Xiu, An efficient surrogate-based method for computing rare failure probability, J. Comput. Phys. 230 (2011) 8683–8697.
[18] E. Ullmann, I. Papaioannou, Multilevel estimation of rare events, SIAM/ASA J. Uncertain. Quantif. 3 (1) (2015) 922–953.
[19] B. Peherstorfer, B. Kramer, K. Willcox, Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation, SIAM/ASA J. Uncertain. Quantif. 6 (2) (2018) 737–761.
[20] B. Kramer, A. Marques, B. Peherstorfer, U. Villa, K. Willcox, Multifidelity probability estimation via fusion of estimators, J. Comput. Phys. 392 (2019) 385–402.
[21] R.T. Haftka, Combining global and local approximations, AIAA J. 29 (9) (1991) 1523–1525.
[22] R. Vitali, R.T. Haftka, B.V. Sankar, Multi-fidelity design of stiffened composite panel with a crack, Struct. Multidiscip. Optim. 23 (5) (2002) 347–356.
[23] S. Heinrich, Multilevel Monte Carlo methods, in: Proceedings of the 3rd International Conference on Large-Scale Scientific Computing, LSSC 01, Sozopol, Bulgaria, 2001, pp. 58–67.
[24] A. Gorodetsky, G. Geraci, M. Eldred, J.D. Jakeman, A generalized approximate control variate framework for multifidelity uncertainty quantification, J. Comput. Phys. 408 (2020) 109257.
[25] R.Y. Rubinstein, D.P. Kroese, Simulation and the Monte Carlo Method, Third ed., Springer, 2017.
[26] H. Kahn, T.E. Harris, Estimation of particle transmission by random sampling, Natl. Bur. Stand. Appl. Math. Ser. 12 (1951) 27–30.
[27] C. Walter, Moving particles: A parallel optimal multilevel splitting method with application in quantiles estimation and meta-model based algorithms, Struct. Saf. 55 (2015) 10–25.
[28] S. Geyer, I. Papaioannou, D. Straub, Cross entropy-based importance sampling using Gaussian densities revisited, Struct. Saf. 76 (2019) 15–27.
[29] T.J. Dodwell, C. Ketelsen, R. Scheichl, A.L. Teckentrup, Multilevel Markov chain Monte Carlo, SIAM/ASA J. Uncertain. Quantif. 3 (2015) 1075–1108.
[30] A. Jasra, K.J.H. Law, Y. Xu, Markov chain simulation for multilevel Monte Carlo, Found. Data Sci. 3 (1) (2021) 27–47.
[31] K. Adamids, S. Loukas, ML estimation in the bivariate Poisson distribution in the presence of missing values via the EM algorithm, J. Stat. Comput. Simul. 50 (3–4) (1994) 163–172.
[32] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements - a Spectral Approach, Springer, Berlin, Heidelberg, 1991.
[33] O.H. Galal, A proposed stochastic finite difference approach based on homogeneous chaos expansion, J. Appl. Math. (2013) ID 950469.