Technische Universität München

Fakultät für Mathematik
Professur für Wissenschaftliches Rechnen
(Prof. Dr. Elisabeth Ullmann)

On efficient methods and error bounds for rare event estimation

Fabian Wagner, M.Sc.

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Oliver Junge

Prüfer der Dissertation:

1. Prof. Dr. Elisabeth Ullmann

2. Prof. Dr. Claudia Schillings (Universität Mannheim)

3. Prof. Dr. Jinglai Li (University of Birmingham)

Die Dissertation wurde am 24.06.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 02.11.2021 angenommen.

Zusammenfassung

Diese Dissertation widmet sich der Schätzung der Wahrscheinlichkeit von seltenen Ereignissen.

Wir präsentieren zwei neue Sampling-basierte Methoden, die eine effiziente Schätzung von Wahrscheinlichkeiten seltener Ereignisse ermöglichen. Die erste Methode ist ein Mehrgitterverfahren, welches eine Hierarchie von Diskretisierungsebenen verwendet. Diese Diskretisierungsebenen werden in einer adaptiven Weise kombiniert, um die Anzahl an Modellauswertungen auf einem feinen, rechenintensiven Gitter zu verringern. Die zweite Methode baut auf dem Ensemble Kalman Filter auf, der ursprünglich für Datenassimilationsprobleme entwickelt wurde. Hierfür formulieren wir das Problem der seltenen Ereignisse als ein inverses Problem und wenden den Ensemble Kalman Filter an, um Samples des seltenen Ereignis zu erzeugen. Neben der Beschreibung des Algorithmus untersuchen wir Eigenschaften des Partikelflusses des Ensemble Kalman Filters.

Ein weiterer Teil dieser Dissertation ist die Analyse des Approximationsfehlers der Wahrscheinlichkeit seltener Ereignisse. Dabei untersuchen wir den Approximationsfehler, der von der diskretisierten Lösung einer elliptischen Differentialgleichung stammt. Wir leiten eine obere Schranke des absoluten Fehlers her, welche sich verhält wie der Diskretisierungsfehler des zugrundeliegenden Modells multipliziert mit einer Approximation der Wahrscheinlichkeit des seltenen Ereignis.

Abstract

This dissertation is devoted to the estimation of the probability of rare events.

We present two novel sampling-based methods, which enable an efficient estimation of rare event probabilities. The first method is a multilevel algorithm, which employs a hierarchy of discretization levels. These discretization levels are combined in an adaptive way to decrease the number of high-level computationally intensive model evaluations. The second method builds on the ensemble Kalman filter, which has been originally developed for data assimilation problems. Thereby, we reformulate the rare event problem as an inverse problem and apply the ensemble Kalman filter to generate samples of the rare event. In addition to describing the algorithm, we investigate properties of the particle flow of the ensemble Kalman filter.

Another part of this dissertation is the analysis of the approximation error of the probability of rare events. Therein, we study the approximation error that stems from the discretized solution of an elliptic diffusion equation. We derive an upper bound of the absolute error, which behaves as the discretization error of the underlying model multiplied by an approximation of the rare event probability.

Acknowledgements

List of contributed articles

Core articles as principal author

I Fabian Wagner, Jonas Latz, Iason Papaioannou, Elisabeth Ullmann.
Multilevel sequential importance sampling for rare event estimation.
SIAM Journal on Scientific Computing, 42.4 (2020), pp. A2062–A2087.
`https://doi.org/10.1137/19M1289601`
(See also article [119] in the bibliography)

II Fabian Wagner, Jonas Latz, Iason Papaioannou, Elisabeth Ullmann.
Error analysis for probabilities of rare events with approximate models.
Accepted for publication in SIAM Journal on Numerical Analysis (2021).
arXiv preprint: `https://arxiv.org/abs/2008.06368`
(See also article [118] in the bibliography)

Further articles as principal author

III Fabian Wagner, Iason Papaioannou, Elisabeth Ullmann.
The ensemble Kalman filter for rare event estimation.
Submitted for publication in SIAM/ASA Journal on Uncertainty Quantification
(2021).
arXiv preprint: `https://arxiv.org/abs/2106.10062`
(See also article [120] in the bibliography)

I, Fabian Wagner, am the principal author of these articles.

Contents

1. Introduction

In many applications, the reliability of an engineering system is of high importance. Failure of the system can have high societal consequences, including financial losses and loss of life. Therefore, the determination of the probability that failure occurs is of high relevance. For instance, in choosing the geographic position of a radioactive waste disposal [23, 89], the porosity of the soil has a decisive influence. A potential risk of the disposal is the outflow of radioactive particles and the pollution of groundwater. Since the porosity of the soil is at best only available on a finite number of measurement points, the hydraulic conductivity is modelled as a random field. Therefore, numerical simulation and sampling methods are combined to determine the probability of failure. However, as the simulation of the flow of the radioactive particles requires the numerical solution of a *partial differential equation* (PDE), smart algorithms have to be chosen to obtain a computationally tractable estimate of the probability of failure. Further occurrences of rare events are examined in settings such as structural reliability [72, 84, 97, 99], financial risk [2], aerospace [86], and in steam generator tubing [22]. In this dissertation, we investigate two novel sampling-based methods and study the approximation error of the probability of failure, which is introduced by the numerical simulation method.

Formally, failure and safe events are distinguished by a so-called *limit-state function* (LSF). It is common to define states which lead to negative LSF outcomes as failure states, while safe states yield positive outcomes. The collection of all failure states is defined as the *failure domain* and the probability mass of the failure domain is the *probability of failure*. Indeed, the denomination *probability of rare events* is another commonly used term, which implies that failure occurs very infrequently.

In the following, we give an overview of existing methods and we categorize our novel approaches. We begin with deterministic approximation methods, which are based on an approximation of the failure domain and do not require sampling approaches. In the *first order reliability method* (FORM) [28, 52], the LSF is linearized at the *most likely failure point* (MLFP) and the FORM estimate is equal to the probability mass of the failure domain with respect to the linearized LSF. Similarly, the *second order reliability method* (SORM) [28, 81] approximates the LSF at the MLFP by a second order Taylor expansion.

Besides deterministic approximation methods, there are sampling-based methods. *Monte Carlo Sampling* (MCS) [41, 61, 105] can be easily applied to estimate the probability of failure. The MCS estimate is equal to the portion of failure samples with respect to the total number of samples. MCS gives an unbiased estimator, but a large number of samples is required to obtain an estimator with a small coefficient of variation if the failure probability is small. Thus, MCS is usually infeasible if the LSF consists of a computationally expensive model evaluation. Starting from MCS, several variance reduction techniques have been developed to increase efficiency. In *Importance Sampling* (IS) [1, 105], samples are drawn from a certain IS density and are reweighted to estimate the probability of failure. The IS density should admit large values in the failure domain to increase the number of failure samples. The best possible choice, the so-called *optimal IS density*, is in general not available [94]. Therefore, sequential approaches have been developed to approximate the optimal

IS density by a sequence of certain IS densities. In *Subset Simulation* (SuS) [4, 6, 8, 122], these densities have compact support and satisfy the *nestedness* property, i.e., their supports are nested and determine a sequence of more probable failure events. Similar to SuS, *Sequential Monte Carlo* (SMC) [18], *Moving Particle* [121], and *generalized splitting methods* [12, 48] have been developed, which also approximate the failure domain by a sequence of more probable failure events. In *Sequential Importance Sampling* (SIS) [94], the support of all IS densities is the whole parametric space, which induces nestedness automatically. The densities are approximated by samples which are sequentially updated by a *Markov chain Monte Carlo* (MCMC) algorithm [24, 92, 103]. In *cross-entropy-based IS* [45, 93], the sequence of densities is determined by minimizing the Kullback–Leibler divergence within a parametric family of densities. In *Line Sampling* [3, 67, 98], the probability of failure is estimated by samples which are distributed on the line perpendicular to the failure surface. In this dissertation, we extend the collection of sampling-based methods by the *ensemble Kalman filter* (EnKF) [120].

For part of the above methods, multilevel or multifidelity approaches have been developed to further improve the efficiency. Therein, sample evaluations from low accuracy models are combined with sample evaluations from high accuracy models to estimate an expectation with respect to the highest accuracy model. This yields computational benefits since evaluations of low accuracy models are usually less cost intensive. Applying the multilevel idea leads to *Multilevel Monte Carlo* (MLMC) [36, 47], *Multilevel Subset Simulation* (MLSuS) [114], *Multilevel Sequential Monte Carlo* [10, 27], and *multifidelity cross entropy based IS* [95]. In this dissertation, we extend the collection of multilevel methods by *Multilevel Sequential Importance Sampling* (MLSIS) [119].

Another way to improve the efficiency is the application of surrogate models. Recently, tools from Machine Learning [49, 87] are widely used to construct surrogate models in various fields. For instance, in [76] a hierarchy of neural network approximations is employed. In [91], SuS is combined with a Neural Network, while [59, 75] employs the polynomial chaos expansion as a surrogate model. Alternatives are *response surface* approaches [16, 96] and *active learning* approaches that are based on Gaussian process regression [35, 108] or polynomial chaos expansion [82]. We note that surrogate models are not considered further in this dissertation.

If the evaluation of the LSF requires the solution of a PDE, the exact LSF is in general not available. The LSF approximation involves a discretization scheme, such as *Finite Differences* [50, 73], *Finite Volumes* [40, 74] or *Finite Elements* (FEs) [14, 15, 21]. However, these discretization schemes introduce a PDE discretization error in the evaluation of the LSF, which leads to an erroneous probability of failure estimate. The authors of [36] derive an upper bound of the absolute approximation error of the probability of failure, which behaves as the PDE discretization error. In this dissertation, we derive, under certain assumptions, a novel upper bound for the absolute error, which behaves as the PDE discretization error multiplied by the FORM estimate [118]. Moreover, we derive an upper bound for the relative error of the FORM estimate. These bounds can be used to construct efficient multilevel methods as in [36].

1.1. Summary of results

This dissertation contains three articles, which deal with the estimation of rare event probabilities. Article I and III propose novel sampling-based methods to estimate the probability of rare events, while Article II performs an error analysis. In more detail, Article I develops a multilevel strategy for SIS [94], which is based on *Multilevel Sequential*^{2} *Monte Carlo* (MLS^{2}MC) given in [70]. In Article II, we perform an error analysis for the approximation error of the probability of rare events, which is induced by the PDE discretization error. In Article III, we apply the EnKF to estimate the probability of rare events and derive theoretical properties of the particle flow of the EnKF particles.

Core articles as principal author

- *Article I [119] in Appendix A:*
  Multilevel sequential importance sampling for rare event estimation
  In this article, we propose a novel multilevel method to estimate the probability of rare events. Our method is based on SIS [94] and requires an approximating sequence of the LSF with increasing accuracy. We approximate the optimal IS density with a twofold adaptive algorithm. On the one hand, we sequentially enhance the approximation of the indicator function, which is contained in the optimal IS density. On the other hand, we increase the discretization level, which gives a better approximation of the LSF. We apply the scheme of [70] to combine both sequential approaches in a computational cost-efficient way. The samples are sequentially updated by an MCMC algorithm such that the samples are distributed according to the desired target densities. For the MCMC algorithm, we consider *adaptive conditional sampling* (aCS) [92] and sampling from the *von Mises–Fisher–Nakagami* (vMFN) distribution model [93], which are applicable even in high-dimensional parameter settings. Indeed, the application of the vMFN distribution model as an independent proposal density in the MCMC step is another contribution of this article. In contrast to MLSuS, MLSIS does not have the nestedness issue, since the support of the IS densities is the whole parametric domain. Indeed, the nestedness issue of MLSuS is the main motivation to implement MLSIS. In numerical experiments, we compare the performance of MLSIS with SIS, SuS and MLSuS.

- *Article II [118] in Appendix B:*
  Error analysis for probabilities of rare events with approximate models
  In this article, we study the PDE approximation error of the probability of failure estimate. We assume that the evaluation of the LSF requires the solution of an elliptic PDE with a stochastic diffusion coefficient. Since the exact solution of the PDE is in general not available, it is approximated by an FE approximation. The mesh size of the FEs determines the accuracy of the approximation. A fine resolution of the mesh yields an accurate PDE approximation and, thus, an accurate approximation of the probability of failure. It is well known that the PDE discretization error has a certain convergence order with respect to the mesh size. In this article, we prove that, under certain assumptions, the induced approximation error of the probability

of failure behaves as the PDE discretization error multiplied by the FORM estimate. One of these assumptions is the convexity of the failure domains. For general failure domains, we prove that the relative error of the FORM estimate behaves as the PDE discretization error. Moreover, if the LSF is affine linear, we prove that the relative error of the probability of failure estimate behaves as the PDE discretization error.

Our provided error bounds require that the stochastic diffusion coefficient is uniformly elliptic and bounded. We conjecture and observe in numerical experiments that similar error bounds hold for diffusion coefficients which are only pathwise elliptic and bounded.

Further articles as principal author

- *Article III [120] in Appendix C:*
  The ensemble Kalman filter for rare event estimation
  In this article, we reformulate the rare event problem as an inverse problem and apply the EnKF for inverse problems [57, 106]. With this reformulation, the EnKF generates samples which are contained in the failure domain or in proximity to the failure domain. Since the EnKF generates approximate samples from a so-called analysis variable [37], the distribution of the EnKF particles is in general unknown. Therefore, we fit a distribution model with the generated EnKF particles and apply IS with respect to the fitted parametric density to estimate the probability of failure. Moreover, we apply the adaptive approach of SIS to determine the tempering parameters in the EnKF, which yields a more efficient algorithm [58]. To handle multi-modal failure domains, we combine a clustering approach with the localized covariance approach of [102].
  In addition, we derive theoretical properties of the EnKF in the rare event estimation context. We derive the continuous-time limit of the EnKF update, which results in a coupled system of *stochastic differential equations* (SDEs). If the LSF is affine linear, we prove for the continuous-time limit, infinite ensemble, and infinite time limit that the mean of the EnKF particles converges to a convex combination of the MLFP and the mean of the optimal IS density. In numerical experiments, we show that the EnKF yields the same level of accuracy as SIS. In the case where the failure domain is unimodal, the EnKF requires less computational costs than SIS for a fixed level of accuracy.

1.2. Outline of this dissertation

This dissertation continues with the problem setting of estimating the probability of failure in Chapter 2. In addition, we recap some of the sampling-based methods and their multilevel algorithms, as well as the FORM estimate. Chapter 3 discusses the model problems which are considered in the numerical experiments of the contributed articles. In Chapter 4, we discuss the MLSIS algorithm, which is based on Article I. Chapter 5 presents the analysis results of Article II with respect to the induced approximation error of the probability of failure. Chapter 6 presents the EnKF for rare event estimation given in Article III. In Chapter 7, we give a brief conclusion and outlook. Article I, II, and III are given in Appendix A, B, and C, respectively.

2. The probability of rare events

We start by defining the probability of failure in a formal way and discuss its computational challenges. Moreover, we discuss existing methods which form the basis of our contributions in the contributed articles.

2.1. Problem setting

We consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. As mentioned in the introduction, failure is determined by the outcome of a limit-state function $G : \mathbb{R}^n \to \mathbb{R}$. The LSF depends on a parameter $u \in \mathbb{R}^n$, which is the realisation of an n-variate random vector $U : \Omega \to \mathbb{R}^n$. Failure is defined as the event $G(U(\omega)) \leq 0$, while $G(U(\omega)) > 0$ is the safe event. We assume that U is distributed according to the n-variate standard normal distribution $\mathrm{N}(0, \mathrm{Id}_n)$, where $\mathrm{Id}_n \in \mathbb{R}^{n \times n}$ denotes the identity matrix. The *probability density function* (PDF) of U is denoted by $\varphi_n : \mathbb{R}^n \to [0, \infty[$.

Remark 2.1. *We require the Gaussian assumption in the proofs of Article II and III. In addition, the Gaussian assumption yields computational advantages in the implementations of the MCMC algorithm. Indeed, adaptive conditional sampling requires that U is Gaussian distributed [92]. However, we note that the Gaussian assumption is not very restrictive. Under mild assumptions, a non-Gaussian random vector $\widetilde{U}$ can be transformed to a Gaussian random variable U. The Nataf transform [30] can be applied if $\widetilde{U}$ can be modelled by a Gaussian copula. A second transform is the Rosenblatt transform [56], which can be applied if the conditional distributions of $\widetilde{U}_{i+1}$ given $\widetilde{U}_1, \ldots, \widetilde{U}_i$ are known for $i = 1, \ldots, n-1$, where $\widetilde{U}_i$ denotes the ith entry of $\widetilde{U}$.*

The *probability of failure* is determined by integrating the PDF φ_n over all failure states. Formally, the probability of failure is denoted by P_f and is defined by

$$
\begin{aligned}
P_f :=& \mathbb{P}(\{\omega \in \Omega : G(U(\omega)) \leq 0\}) \\
=& \int_{u \in \mathbb{R}^n} I(G(u) \leq 0)\varphi_n(u)\mathrm{d}u = \mathbb{E}_{\varphi_n}[I(G(U) \leq 0)],
\end{aligned}
\tag{2.1}
$$

where I is the indicator function with outcomes $I(\text{true}) = 1$ and $I(\text{false}) = 0$. The expression $\mathbb{E}_{\varphi_n}[\cdot]$ denotes the expectation assuming that U is distributed according to φ_n. The collection of all failure states is called the *failure domain* and is denoted by $A := \{u \in \mathbb{R}^n : G(u) \leq 0\}$. The boundary $\partial A := \{u \in \mathbb{R}^n : G(u) = 0\}$ is called the *limit-state surface* and the complement $A^C := \{u \in \mathbb{R}^n : G(u) > 0\}$ is the *safe domain*.

We consider two computational challenges which occur in rare event estimation. First, it is a priori unknown which parameter values belong to the failure domain. Hence, the failure domain has to be determined or approximated to evaluate the integral in (2.1). Secondly, the evaluation of G might be cost intensive or even not feasible, since only an approximation G_h is available, where $h > 0$ is a discretization parameter. In this case, the approximate probability of failure $P_{f,h}$ is defined by

$$
P_{f,h} := \mathbb{P}(\{\omega \in \Omega : G_h(U(\omega)) \leq 0\}) = \mathbb{E}_{\varphi_n}[I(G_h(U) \leq 0)],
$$

and the approximate failure domain is denoted by $A_h := \{u \in \mathbb{R}^n : G_h(u) \leq 0\}$. For instance, if the evaluation of G requires the solution of a PDE, an approximation G_h can be obtained by FEs, where h is the FE mesh size. If a small FE mesh size is employed, the evaluation of G_h is accurate but cost demanding.

In this dissertation, we perform an error analysis in Article II of the error $|P_f - P_{f,h}|$ if an approximation G_h is given. Moreover, we propose two novel sampling-based methods to estimate the probability of failure P_f accurately and efficient. The first method is the MLSIS estimator of Article I, which is a novel multilevel method. The second approach is the EnKF algorithm of Article III, which is a novel single level variance reduction technique for rare event estimation. Before we consider our novel approaches, we recap the relevant methods which form the basis of our approaches.

2.2. Sampling-based methods

We give a brief introduction to Monte Carlo Sampling, Importance Sampling, Sequential Importance Sampling, and Subset Simulation as well as to the multilevel algorithms Multilevel Monte Carlo and Multilevel Subset Simulation. As seen in the introduction, there are more sampling-based methods. However, these approaches are not discussed in detail as they are not relevant for our novel approaches. Since SIS, MLSIS, SuS, and MLSuS require a Markov Chain Monte Carlo algorithm, we will also introduce the *Metropolis–Hastings* sampler of [53, 85].

2.2.1. Single level methods

We note that the following estimators are defined with respect to the exact LSF G. By substituting G with G_h, the estimators for the approximate probability of failure $P_{f,h}$ are obtained.

Monte Carlo Sampling

An easy implementable and unbiased estimator is obtained by crude Monte Carlo Sampling [41, 61, 105]. Using $J \in \mathbb{N}$ samples $\{u^{(j)}\}_{j=1}^J$, which are independently distributed according to φ_n, the MCS estimator of the expectation in (2.1) is

$$\widehat{P}_f^{\mathrm{MCS}} := \frac{1}{J} \sum_{j=1}^J I(G(u^{(j)}) \leq 0).$$

By [105, Example 4.1], the *coefficient of variation* $\delta_{\widehat{P}_f^{\mathrm{MCS}}}$ of the MCS estimator is given by

$$\delta_{\widehat{P}_f^{\mathrm{MCS}}} := \frac{\mathrm{StD}[\widehat{P}_f^{\mathrm{MCS}}]}{\mathbb{E}[\widehat{P}_f^{\mathrm{MCS}}]} = \sqrt{\frac{1 - P_f}{P_f J}}, \tag{2.2}$$

where StD denotes the *standard deviation*. We note that the coefficient of variation is also a measure for the relative error of the MCS estimator [105]. From (2.2), it follows that a large number of samples is required to achieve a small coefficient of variation when P_f is small. This makes MCS infeasible if the evaluation of G is

cost demanding. Indeed, the MCS estimator forms a starting point and points to the main idea of sampling-based methods. The goal of these methods is to reduce the variance of MCS. Hence, they are termed *variance reduction techniques.* They achieve this through generating samples which are in the failure domain or, at least, in its proximity. We emphasize that the generation of failure samples is the main motivation of sampling-based methods to achieve an efficient estimator of the probability of failure. The following algorithms differ in the way failure samples are generated.

Importance Sampling

The disadvantage of crude MCS is that only a small fraction of the samples is contained in the failure domain and, thus, computational costs are wasted for many non-failure samples. In Importance Sampling [1, 105], we assume that a certain IS density $p^{\mathrm{IS}} : \mathbb{R}^n \to [0, \infty[$ is available, which admits larger values in the failure domain than φ_n. Thus, the integral in (2.1) is expressed in terms of the IS density p^{IS}, which gives

$$P_f = \int_{u \in \mathbb{R}^n} I(G(u) \leq 0) w(u) p^{\mathrm{IS}}(u) \mathrm{d}u = \mathbb{E}_{p^{\mathrm{IS}}}[I(G(U) \leq 0) w(U)], \qquad (2.3)$$

where $w(u) := \varphi_n(u)/p^{\mathrm{IS}}(u)$ denotes the *importance weight.* Again, MCS is applied to the expectation in (2.3), which yields the IS estimator

$$\widehat{P}_f^{\mathrm{IS}} := \frac{1}{J} \sum_{j=1}^{J} I(G(u^{(j)}) \leq 0) w(u^{(j)}),$$

where the samples $\{u^{(j)}\}_{j=1}^{J}$ are independently distributed according to the IS density p^{IS}. The choice of p^{IS} is crucial and influences the efficiency of the IS estimator. If the support of p^{IS} contains the failure domain A, $\widehat{P}_f^{\mathrm{IS}}$ is an unbiased estimator for the probability of failure. By [94], the *optimal importance sampling density p_{opt}* is given by

$$p_{\mathrm{opt}}(u) := \frac{1}{P_f} I(G(u) \leq 0) \varphi_n(u). \qquad (2.4)$$

Indeed, p_{opt} leads to a zero-variance estimator and one sample of p_{opt} is sufficient to estimate P_f exactly. However, as P_f and the failure domain are a priori inaccessible, p_{opt} is not applicable. A feasible choice for the IS density is discussed in [5], where the authors propose to apply the input distribution centred at the MLFP. Alternatively, the authors of [34] employ a surrogate model as an approximation to the LSF to approximate the optimal IS density. In this dissertation, we consider SIS as a sequential approach to derive an approximation of the optimal IS density.

Sequential Importance Sampling

In Sequential Importance Sampling [94], the optimal IS density p_{opt} is succesively approximated by a sequence of IS densities. This sequence is determined by an

approximation of the indicator function in (2.4). We denote by $\Phi : \mathbb{R} \to [0,1]$ the *cumulative distribution function* (CDF) of the univariate standard normal distribution. The CDF approximates the indicator function by

$$I(G(u) \leq 0) = \lim_{\sigma \downarrow 0} \Phi\left(-\frac{G(u)}{\sigma}\right), \quad \text{for } G(u) \neq 0.$$

A visualization of this approximation is given in Figure 1 of Article I. Alternative to the Gaussian CDF, the authors of [115] employ the standard logistic function to approximate the indicator function and to determine a parametric family of densities in cross-entropy-based IS. In Article III and in Chapter 6, we will see a novel smooth approximation of the indicator function, which arises in the application of the EnKF to rare event estimation.

To define a smooth transition to the optimal IS density, we denote by $\{\sigma_k\}_{k=0}^{N_T}$ a finite sequence of *temperatures*, which satisfies $\infty = \sigma_0 > \sigma_1 > \cdots > \sigma_{N_T} > 0$. The number N_T determines the number of temperatures and the number of IS densities. The sequence of IS densities is defined by

$$p_0^{\mathrm{SIS}}(u) := \varphi_n(u), \tag{2.5}$$

$$p_k^{\mathrm{SIS}}(u) := \frac{1}{P_k^{\mathrm{SIS}}} \Phi\left(-\frac{G(u)}{\sigma_k}\right) \varphi_n(u), \quad \text{for } k = 1, \ldots, N_T,$$

where P_k^{SIS} are normalizing constants, which are not directly available. However, this yields no problem for SIS as we see in the following. To determine the SIS estimator for the probability of failure, we use the fact that P_f can be expressed by

$$P_f = P_0^{\mathrm{SIS}} \frac{P_1^{\mathrm{SIS}}}{P_0^{\mathrm{SIS}}} \frac{P_2^{\mathrm{SIS}}}{P_1^{\mathrm{SIS}}} \cdots \frac{P_{N_T}^{\mathrm{SIS}}}{P_{N_T-1}^{\mathrm{SIS}}} \frac{P_f}{P_{N_T}^{\mathrm{SIS}}}, \tag{2.6}$$

where $P_0^{\mathrm{SIS}} = 1$ follows from (2.5). As shown in [26, 94], the fraction $S_k := P_k^{\mathrm{SIS}}/P_{k-1}^{\mathrm{SIS}}$ can be estimated by IS. It holds that

$$P_k^{\mathrm{SIS}} = \int_{u \in \mathbb{R}^n} \Phi\left(-G(u)/\sigma_k\right) \varphi_n(u) \mathrm{d}u = P_{k-1}^{\mathrm{SIS}} \int_{u \in \mathbb{R}^n} \frac{\Phi\left(-G(u)/\sigma_k\right)}{\Phi\left(-G(u)/\sigma_{k-1}\right)} p_{k-1}^{\mathrm{SIS}}(u) \mathrm{d}u$$

$$= P_{k-1}^{\mathrm{SIS}} \int_{u \in \mathbb{R}^n} w_k(u) p_{k-1}^{\mathrm{SIS}}(u) \mathrm{d}u, \tag{2.7}$$

where w_k are the importance weights, which are proportional to the ratios of two subsequent densities

$$w_1(u) = \Phi(-G(u)/\sigma_1), \quad w_k(u) = \frac{\Phi\left(-G(u)/\sigma_k\right)}{\Phi\left(-G(u)/\sigma_{k-1}\right)}, \quad \text{for } k > 1. \tag{2.8}$$

From (2.7), it follows that the ratio of two subsequent normalizing constants is given by $S_k = \mathbb{E}_{p_{k-1}^{\mathrm{SIS}}}[w_k(U)]$. This expectation is estimated by a sample mean similar to the expectations in (2.1) and (2.3). Using J samples $\{u_{k-1}^{(j)}\}_{j=1}^{J}$, which are distributed according to p_{k-1}^{SIS}, we obtain the estimator

$$\widehat{S}_k := \frac{1}{J} \sum_{j=1}^{J} w_k(u_{k-1}^{(j)}). \tag{2.9}$$

The final factor $P_f/P_{N_T}^{\text{SIS}}$ in (2.6) is determined by the importance weight of the optimal IS density p_{opt} with respect to the density $p_{N_T}^{\text{SIS}}$, which gives the estimator

$$\frac{\widehat{P_f}}{P_{N_T}^{\text{SIS}}} := \frac{1}{J} \sum_{j=1}^{J} \frac{I(G(u_{N_T}^{(j)}) \leq 0)}{\Phi(-G(u_{N_T}^{(j)})/\sigma_{N_T})}, \tag{2.10}$$

where the samples $\{u_{N_T}^{(j)}\}_{j=1}^{J}$ are distributed according to $p_{N_T}^{\text{SIS}}$. Multiplying the estimators in (2.9) and (2.10) yields the SIS estimator of the probability of failure

$$\widehat{P}_f^{\text{SIS}} := \frac{\widehat{P_f}}{P_{N_T}^{\text{SIS}}} \prod_{k=1}^{N_T} \widehat{S}_k.$$

Having derived the SIS estimator, we discuss now the adaptive procedure of determining the temperatures σ_k. We choose these temperatures in a way that two consecutive densities p_{k-1}^{SIS} and p_k^{SIS} are not too different. This discrepancy is measured in terms of the coefficient of variation of the importance weights w_k in (2.8). With the user specified *target coefficient of variation* $\delta_{\text{target}} > 0$, we determine σ_k adaptively by

$$\sigma_k = \underset{\sigma \in (0, \sigma_{k-1})}{\operatorname{argmin}} \left(\delta_{w_k} - \delta_{\text{target}} \right)^2, \quad \text{where } \delta_{w_k} := \frac{\operatorname{StD}_{p_{k-1}^{\text{SIS}}}[w_k(U)]}{\mathbb{E}_{p_{k-1}^{\text{SIS}}}[w_k(U)]}.$$

Requiring that the coefficient of variation is equal to δ_{target} is equivalent to requiring that the *effective sample size* is equal to a target value [70, Section 3.4]. Tempering is finished, if the discrepancy between the current density p_k^{SIS} and the optimal IS density p_{opt} is small. Again, the discrepancy is measured in terms of the coefficient of variation of the weights

$$w_{\text{opt},k}(u) := \frac{I(G(u) \leq 0)}{\Phi(-G(u)/\sigma_k)}.$$

If $\delta_{w_{\text{opt},k}} \leq \delta_{\text{target}}$, the final tempering level $k = N_T$ is reached. We note that $\delta_{w_{\text{opt},k}}$ has to be estimated in each tempering step using the samples $\{u_k^{(j)}\}_{j=1}^{J}$. If δ_{target} is small, a large number of tempering updates is performed and the coefficient of variation of the SIS estimator is small.

It is crucial that the samples in (2.9) are distributed according to the density p_{k-1}^{SIS}. In each tempering update, we apply a resample-move scheme to update the samples $\{u_{k-1}^{(j)}\}_{j=1}^{J}$, which are distributed approximately according to the density p_{k-1}^{SIS}, to samples $\{u_k^{(j)}\}_{j=1}^{J}$, which are distributed approximately according to the density p_k^{SIS}. The resampling is performed based on the weights w_k in (2.8) and the samples are moved by an MCMC algorithm. The initial samples $\{u_0^{(j)}\}_{j=1}^{J}$ are distributed according to the standard normal density φ_n. We note that due to the adaptive choice of σ_k, the samples $\{u_k^{(j)}\}_{j=1}^{J}$ are only distributed approximately according to p_k^{SIS} and the SIS estimator is biased [9].

Subset Simulation

Subset Simulation [4, 6] is commonly not described as a technique to approximate the optimal IS density. However, we explain SuS in this way to display similarities and differences to SIS as in [94, Section 2.6]. In SuS, the optimal IS density is approximated by a sequence of densities which have compact supports. These densities are defined as the optimal IS densities with respect to a more probable failure criterion. Therefore, the failure event $\{G(U(\omega)) \leq 0\}$ is sequentially approximated by the event $\{G(U(\omega)) \leq c_k\}$, where $\infty = c_0 > c_1 > \cdots > c_{N_T} = 0$. The sequence $\{c_k\}_{k=1}^{N_T}$ is adaptively chosen such that the sequence of domains $B_k := \{\omega \in \Omega : G(U(\omega)) \leq c_k\}$ satisfies

$$\mathbb{P}(B_k \mid B_{k-1}) = \widehat{c} \in (0, 1), \quad \text{for all } k = 1, \ldots, N_{T-1},$$

where $\mathbb{P}(B_k \mid B_{k-1}) = \mathbb{P}(B_k \cap B_{k-1})/\mathbb{P}(B_{k-1})$ is the conditional probability of the event $\omega \in B_k$ given that $\omega \in B_{k-1}$ if $\mathbb{P}(B_{k-1}) > 0$. The parameter $\widehat{c}$ is chosen larger than P_f and such that $\widehat{c} \cdot J \in \mathbb{N}$, typically $\widehat{c} = 0.1$. We note that $B_0 = \Omega$, while B_{N_T} is the failure domain. Based on the definition of B_k, it follows that the sequence of domains $B_0, B_1, \ldots, B_{N_T}$ is *nested*, i.e., $B_k \subset B_{k-1}$ for $k = 1, \ldots, N_T$. The sequence of IS densities in SuS is given by

$$p_k^{\mathrm{SuS}}(u) := \frac{1}{P_k^{\mathrm{SuS}}} I(G(u) \leq c_k)\varphi_n(u),$$

where P_k^{SuS} is chosen such that p_k^{SuS} is a PDF. We observe that $p_{N_T}^{\mathrm{SuS}}$ is the optimal IS density. The density p_k^{SuS} is equal to the density of conditioning U to the failure domain B_k, which we denote by $\varphi_n(\cdot \mid B_k)$. To determine the values c_k, the densities $\varphi_n(\cdot \mid B_k)$ are approximated by the collection of samples $\{u_k^{(j)}\}_{j=1}^J$. Similar to SIS, an MCMC algorithm is applied to transfer the samples $\{u_{k-1}^{(j)}\}_{j=1}^J$ into the samples $\{u_k^{(j)}\}_{j=1}^J$. The value c_k is determined such that the portion of samples satisfying $G(u_{k-1}^{(j)}) \leq c_k$ is equal to $\widehat{c}$. The samples which satisfy this condition are chosen as seeds in the MCMC algorithm. Thus, unlike SIS, no resampling step is required in SuS. Due to the nestedness of the domains B_k, the seeds are distributed approximately according to the target density $\varphi_n(\cdot \mid B_k)$; hence, no burn-in is applied in the MCMC run. We note that the seeds are only distributed approximately according to $\varphi_n(\cdot \mid B_k)$, since c_k is chosen adaptively based on the samples $\{u_{k-1}^{(j)}\}_{j=1}^J$ [12].

The number of update steps is equal to $N_T = \lceil \log_{1/\widehat{c}}(\widehat{c}/P_f) \rceil + 1$, where $\lceil \cdot \rceil$ denotes the ceiling function. Due to the nestedness, the probability of failure P_f can be expressed as the product of conditional probabilities

$$P_f = \mathbb{P}(B_1) \prod_{k=2}^{N_T} \mathbb{P}(B_k \mid B_{k-1}). \tag{2.11}$$

The probabilities $\mathbb{P}(B_1)$ and $\mathbb{P}(B_k \mid B_{k-1})$, for $k = 2, \ldots, N_T - 1$, are estimated by $\widehat{c}$, while the conditional probability $\mathbb{P}(B_{N_T} \mid B_{N_T-1})$ is estimated by

$$\widehat{P}_{B_{N_T}|B_{N_T-1}} := \frac{1}{J} \sum_{j=1}^J I\left(G\left(u_{N_T-1}^{(j)}\right) \leq 0\right). \tag{2.12}$$

Combining (2.11) and (2.12) yields the SuS estimator

$$\widehat{P}_f^{\mathrm{SuS}} := \widehat{c}^{\,N_T-1} \widehat{P}_{B_{N_T}|B_{N_T-1}},$$

which is a biased estimator of the probability of failure as shown in [4].

Remark 2.2. *We remark on the similarities and differences of SIS and SuS. Both methods can be described as a sequential approximation of the optimal IS density. The number of tempering steps is determined by δ_{target} in SIS and by $\widehat{c}$ in SuS. Due to the adaptive choices of σ_k and c_k, both methods yield a biased estimate [9] and the samples are only distributed approximately according to the target densities. To generate samples from these target densities, an MCMC algorithm is applied in each step of SIS and SuS. Unlike SIS, resampling is not required in SuS. The support of p_k^{SIS} is the whole parametric domain, while the support of $p_k^{\mathrm{SuS}} = \varphi_n(\cdot \mid B_k)$ is compact. Both methods satisfy the nestedness property, i.e., the support of p_k is a subset of the support of p_{k-1}.*

Markov chain Monte Carlo

We have seen that SIS and SuS require a Markov chain Monte Carlo algorithm to generate samples from the target densities. In the following, we discuss the famous *Metropolis–Hastings* sampler [53, 85]. The goal is to transfer samples $\{u_{k-1}^{(j)}\}_{j=1}^J$, which are distributed according to the density p_{k-1}, into samples from p_k. We denote by $J_{\mathrm{seeds}} \in \mathbb{N}$ the number of seeds and by $J_{\mathrm{burn}} \in \mathbb{N}$ the burn-in length. The number of seeds is chosen such that $J_{\mathrm{seeds}} := \widehat{s} \cdot J$, where $\widehat{s} \in (0, 1]$ is chosen such that $1/\widehat{s} \in \mathbb{N}$. Thus, J_{seeds} is the number of simulated Markov chains and $J_{\mathrm{burn}} + 1/\widehat{s}$ are their lengths.

We denote by $\{u_{k-1}^{(j_i)}\}_{i=1}^{J_{\mathrm{seeds}}}$ the set of seeds. In SIS, the seeds are determined by multinomial resampling with respect to the importance weights w_k given in (2.8). We note that the authors of [44] study the convergence behaviour of various resampling schemes. In SuS, a natural choice is $\widehat{s} = \widehat{c}$ and to choose the samples which lie in the subsequent failure domain B_k as seeds.

Starting from a seed $u_0 \in \{u_{k-1}^{(j_i)}\}_{i=1}^{J_{\mathrm{seeds}}}$, a *proposal* $\overline{u} \in \mathbb{R}^n$ is generated according to the *proposal density* $q(\cdot \mid u_0) : \mathbb{R}^n \to [0, \infty)$. If q does not depend on the seed u_0, we speak of an independent proposal density. The proposal is accepted with the probability determined by the *acceptance function* $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to [0, 1]$. For SIS [94], the acceptance function is given by

$$\alpha^{\mathrm{SIS}}(u_0, \overline{u}) := \min\left\{1, \frac{\Phi\left(-G(\overline{u})/\sigma_k\right)\varphi_n(\overline{u})q(u_0 \mid \overline{u})}{\Phi\left(-G(u_0)/\sigma_k\right)\varphi_n(u_0)q(\overline{u} \mid u_0)}\right\}. \tag{2.13}$$

For SuS [92], the acceptance function reads as

$$\alpha^{\mathrm{SuS}}(u_0, \overline{u}) := \min\left\{1, I(G(\overline{u}) \le c_k)\frac{\varphi_n(\overline{u})q(u_0 \mid \overline{u})}{\varphi_n(u_0)q(\overline{u} \mid u_0)}\right\}.$$

If $\overline{u}$ is accepted, it is attached to the Markov chain. In the next step, $\overline{u}$ is considered as the seed. If $\overline{u}$ is not accepted, u_0 is attached to the chain and u_0 is again considered

as the seed in the next iteration. This procedure is applied $J_{\mathrm{burn}} + 1/\widehat{s}$ times for each seed. At the end, the first J_{burn} states of each Markov chain are rejected. Due to the choice of the acceptance function, which determines the transition kernel of the Markov chain, the simulated Markov chain is stationary with respect to the target density p_k^{SIS} and p_k^{SuS}, respectively. If in addition the Markov chain is ergodic, the generated states will be distributed according to the target density after sufficiently many steps [103, Chapter 7].

We note that the proposal density has a high influence in the performance of the MCMC algorithm. If the proposal density leads to a small acceptance rate, many samples are rejected, while a large acceptance rate is an indicator for strongly dependent states. In Chapter 3, we consider the vMFN distribution model as a novel independent proposal density, which achieves high acceptance rates even in high-dimensions.

Remark 2.3. *We remark that the application of an MCMC algorithm usually requires expertise in tuning the relevant parameters and the proposal density. Our novel EnKF estimator proposed in Article III does not require an MCMC algorithm. From this perspective, the EnKF is easier to apply since less parameters have to be tuned.*

2.2.2. Multilevel methods

For the multilevel methods, we consider the case that the evaluation of G requires the solution of a PDE, and we assume that an approximate LSF G_{h_ℓ} is given on the discretization levels $\ell = 1, \ldots, L$. The first level $\ell = 1$ represents the smallest and the final level $\ell = L \in \mathbb{N}$ represents the highest discretization level, which employs the finest FE mesh size. Commonly, the sequence of FE mesh sizes is defined by $h_\ell = 2^{-\ell}$ for $\ell = 1, \ldots, L$. The approximating sequence of the LSFs is denoted by G_{h_ℓ} for $\ell = 1, \ldots, L$. For smaller h, the approximation G_h is more accurate but the evaluation is more cost intensive. For piecewise linear FEs, the cost of evaluating G_h is ideally

$$\mathrm{Cost}(G_h) = \mathcal{O}(h^{-d}),$$

where $d \in \mathbb{N}$ is the dimension of the computational domain. The overall goal of multilevel methods is the reduction of high-level LSF evaluations to save computational costs while keeping the achieved accuracy fixed.

Multilevel Monte Carlo

Since the probability of failure can be expressed as an expectation, the telescopic sum approach of [47, 54] can be applied. The probability of failure with respect to the final discretization level L can be expressed as

$$P_{f,h_L} = \mathbb{E}_{\varphi_n}[I(G_{h_1}(U) \le 0)] + \sum_{\ell=2}^{L} \mathbb{E}_{\varphi_n}[I(G_{h_\ell}(U) \le 0)] - \mathbb{E}_{\varphi_n}[I(G_{h_{\ell-1}}(U) \le 0)].$$

Since the difference between the function values of $G_{h_{\ell-1}}$ and G_{h_ℓ} is small, the variance of the estimator of the subproblem

$$Q_\ell := \mathbb{E}_{\varphi_n}[I(G_{h_\ell}(U) \leq 0)] - \mathbb{E}_{\varphi_n}[I(G_{h_{\ell-1}}(U) \leq 0)] = P_{f,h_\ell} - P_{f,h_{\ell-1}}$$

is small and, usually, a small number of samples is required to estimate Q_ℓ accurately. As the variance of the estimators Q_ℓ and the computational costs of G_{h_ℓ} are different, we consider a level dependent number of samples $J_\ell \in \mathbb{N}$. The Multilevel Monte Carlo estimator for the probability of failure is given by

$$\widehat{P}_{f,h_L}^{\mathrm{MLMC}} := \frac{1}{J_1} \sum_{j=1}^{J_1} I(G_{h_1}(u_1^{(j)}) \leq 0)$$
$$+ \sum_{\ell=2}^{L} \frac{1}{J_\ell} \sum_{j=1}^{J_\ell} I(G_{h_\ell}(u_\ell^{(j)}) \leq 0) - I(G_{h_{\ell-1}}(u_\ell^{(j)}) \leq 0),$$

where the samples $\{u_\ell^{(j)}\}_{j=1}^{J_\ell}$ are independently distributed according to φ_n for $\ell = 1, \ldots, L$. The number of samples J_ℓ has to be chosen such that the variance of the estimators Q_ℓ and the overall computational costs are minimized. The authors of [36] study the variance of the MLMC estimator and derive optimal values for J_ℓ.

Remark 2.4. *Since Q_ℓ is the difference of possible small approximate failure probabilities P_{f,h_ℓ} and $P_{f,h_{\ell-1}}$, Q_ℓ might be small as well. Thus, for small failure probabilities, the MLMC estimator suffers from the same issues as the MCS estimator since a large number of samples is required to estimate Q_ℓ accurately.*

Multilevel subset simulation

For the following, we assume, for simplicity, that the number of tempering steps N_T is the same as the number of discretization levels L. Thus, one level update represents one tempering update. In Multilevel Subset Simulation, the sequence of domains is defined as $B_\ell := \{\omega \in \Omega : G_{h_\ell}(U(\omega)) \leq c_\ell\}$, where $\infty = c_0 > c_1 > \cdots > c_L = 0$. Therefore, the sequence of domains is in general not nested, i.e., $B_\ell \not\subset B_{\ell-1}$ since the domains are defined with respect to different LSFs G_{h_ℓ}. Thus, we cannot use the formula derived in (2.11) to express the probability of failure. The authors of [114] derive a similar formula for general non-nested domains. They show that the probability of failure can be expressed as

$$P_{f,h_L} = \mathbb{P}(B_1) \prod_{\ell=2}^{L} \frac{\mathbb{P}(B_\ell \mid B_{\ell-1})}{\mathbb{P}(B_{\ell-1} \mid B_\ell)}. \tag{2.14}$$

As in SuS, the parameter values c_ℓ are chosen adaptively such that the numerator $\mathbb{P}(B_\ell \mid B_{\ell-1})$ is equal to $\widehat{c}$ for $\ell = 1, \ldots, L-1$, where $B_0 = \Omega$. However, the denominators in (2.14) need to be estimated. We note that the denominators would be equal to one, if the domains B_ℓ were nested. To generate failure samples, we apply an MCMC algorithm to shift the samples $\{u_{\ell-1}^{(j)}\}_{j=1}^{J}$, which are distributed according to $\varphi_n(\cdot \mid B_{\ell-1})$, into samples $\{u_\ell^{(j)}\}_{j=1}^{J}$, which are distributed according to

$\varphi_n(\cdot \mid B_\ell)$. To reduce correlations of the estimators for $\mathbb{P}(B_\ell \mid B_{\ell-1})$ and $\mathbb{P}(B_{\ell-1} \mid B_\ell)$, the authors of [114] apply a second MCMC run in each iteration to determine the estimators with a different collection of samples. With the generated samples, the conditional probability $\mathbb{P}(B_{\ell-1} \mid B_\ell)$ is estimated by

$$\widehat{P}_{B_{\ell-1}|B_\ell} = \frac{1}{J} \sum_{j=1}^{J} I\left(G_{h_{\ell-1}}(u_\ell^{(j)}) \leq c_{\ell-1}\right). \tag{2.15}$$

Combining (2.14) and (2.15) yields the MLSuS estimator

$$\widehat{P}_{f,h_L}^{\mathrm{MLSuS}} := \frac{\widehat{c}^{L-1} \widehat{P}_{B_{N_L}|B_{N_L-1}}}{\prod_{\ell=2}^{L} \widehat{P}_{B_{\ell-1}|B_\ell}},$$

where the final estimator $\widehat{P}_{B_{N_L}|B_{N_L-1}}$ is given by

$$\widehat{P}_{B_{N_L}|B_{N_L-1}} = \frac{1}{J} \sum_{j=1}^{J} I\left(G_{h_L}(u_{L-1}^{(j)}) \leq 0\right),$$

where the samples $\{u_{L-1}^{(j)}\}_{j=1}^{J}$ are distributed according to $\varphi_n(\cdot \mid B_{L-1})$. We note that due to the non-nestedness of the domains B_ℓ, the samples which are chosen as seeds in the MCMC step are no longer distributed according to the target density. Thus, burn-in is employed, which increases the computational costs.

Remark 2.5. *We note that the nestedness issue of MLSuS is our main motivation to implement the MLSIS estimator, which is explained in detail in Chapter 4. In MLSIS, the support of the employed densities is the whole parametric domain, which ensures nestedness automatically.*

2.3. The first order reliability method

In contrast to the sampling-based methods, the first order reliability method is an approximation method. In FORM [28, 52], the LSF is approximated by its linearization at the most likely failure point, which we denote by u^{MLFP}. As the name suggests, the MLFP is the failure point which admits the largest value with respect to φ_n or, equivalently, which has the smallest distance to the origin. Thus, the MLFP is a solution of the minimization problem

$$u^{\mathrm{MLFP}} = \underset{u \in \mathbb{R}^n}{\mathrm{argmin}} \quad \frac{1}{2}\|u\|_2^2, \quad \text{such that } G(u) = 0. \tag{2.16}$$

The FORM estimate P_f^{FORM} is equal to the probability mass of the failure domain with respect to the linearization of G. The linearization of G at the MLFP is

$$G^{\mathrm{lin}}(u) = \nabla_u G(u^{\mathrm{MLFP}})^T (u - u^{\mathrm{MLFP}}).$$

If $G(0) > 0$, the FORM estimate is given by

$$P_f^{\mathrm{FORM}} = \mathbb{P}(G^{\mathrm{lin}}(U) \leq 0) = \Phi(-\|u^{\mathrm{MLFP}}\|_2),$$

where Φ is the CDF of the univariate standard normal distribution. If $G(0) < 0$, we get $P_f^{\mathrm{FORM}} = 1 - \Phi(-\|u^{\mathrm{MLFP}}\|_2)$. The approximate FORM estimate $P_{f,h}^{\mathrm{FORM}}$ is determined with respect to the approximate MLFP u_h^{MLFP}, which is the solution of (2.16) with respect to G_h. Figure 1 visualizes the FORM estimate and the MLFP for an artificial example.
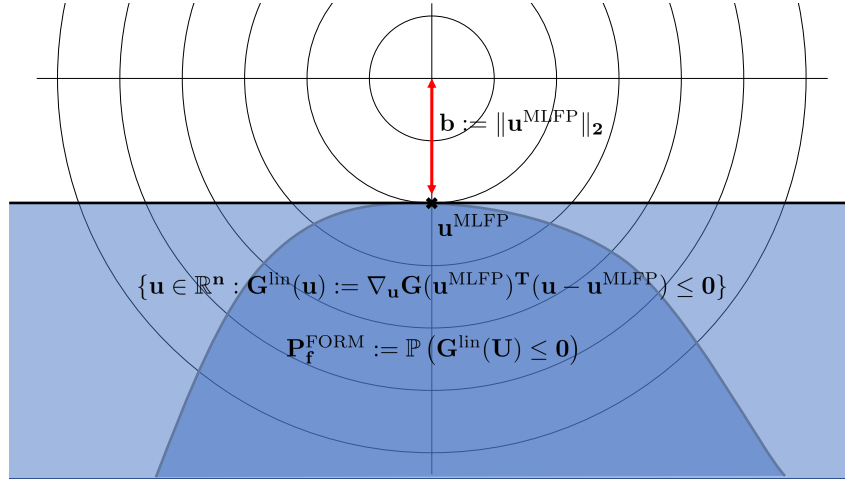

Figure 1: The FORM approximation of the probability of failure. The dark blue area shows the failure domain A. The light blue area shows the failure domain with respect to the linearization G^{lin} at u^{MLFP}.

We note that the crucial part is the determination of the MLFP. Since the parametric dimension n might be large and the LSF G highly nonlinear, the optimization problem in (2.16) is not trivially solvable. However, efficient algorithms [78] have been developed to determine (approximately) the MLFP. One of these is the Hasofer–Lind–Rackwitz–Fiessler algorithm [99]. In general, FORM is less cost intensive but less accurate than the sampling-based methods in Section 2.2. However, due to the exponential decrease of the Gaussian measure, P_f^{FORM} is usually accurate for small failure probabilities since the probability mass of the failure domain is mostly concentrated around the MLFP. We note that this is only valid for unimodal failure domains. For multi-modal failure domains, the FORM estimate is biased since it is an approximation of the probability mass of a single failure mode. Alternatively, the authors of [29] propose to determine for each failure mode an individual design point, which is the MLFP restricted to a single failure mode.

For the theoretical parts of Article II and III, we will use two properties of the FORM estimate which are related to the exact probability of failure. First, if the failure domain is a convex set, the FORM estimate is an upper bound for the probability of failure, i.e., $P_f \leq P_f^{\mathrm{FORM}}$. This property can be observed from Figure 1. If the failure domain A is convex, A is a subset of $A^{\mathrm{lin}} := \{u \in \mathbb{R}^n : G^{\mathrm{lin}}(u) \leq 0\}$. For our provided error bound in Theorem 2.12 of Article II, we require that the failure domains A and A_h are convex. However, we derive another error bound related to the FORM probability estimate in Proposition 2.11 of Article II, which does not require convexity. Secondly, if the LSF is affine linear, the FORM estimate is equal to the probability of failure, i.e., $P_f^{\mathrm{FORM}} = P_f$. In Theorem 4.7 of Article III, we derive the

limiting properties of the EnKF particles for the affine linear case. In Theorem 3.2 of Article II, we state an error bound for the relative error of the probability of failure if the LSF is affine linear.

3. Model problems

In this chapter, we define the LSFs which are discussed in the numerical experiments of Article I, II and III. We distinguish between low- and high-dimensional parameter spaces. In the low-dimensional settings, the LSF is analytically given and the probability of failure can be determined cheaply by MCS. In contrast, in the high-dimensional settings, only an approximation of the exact LSF is available and a reference estimate for the exact probability is determined on a high discretization level.

3.1. Low-dimensional parameter space

We begin with three LSFs which consist of a two-dimensional parameter space, i.e., $U : \Omega \to \mathbb{R}^2$ is a two-variate Gaussian random vector. These LSFs are analytically given and do not require the solution of an *ordinary differential equation* (ODE) or PDE. The respective failure domains have different shapes and different numbers of failure modes. The LSFs are considered in Article III, which applies the EnKF to rare event estimation.

Convex limit-state function

In [65], the following convex LSF is proposed

$$G^{(1)}(U(\omega)) = 0.1(U_1(\omega) - U_2(\omega))^2 - \frac{1}{\sqrt{2}}(U_1(\omega) + U_2(\omega)) + 2.5. \qquad (3.1)$$

The probability of failure is equal to $P_f \approx 4.21 \cdot 10^{-3}$ and the failure domain is convex and consists of a single mode [94]. Figure 2 of Article III illustrates the failure domain.

Parabolic limit-state function

In [29], the following parabolic LSF is proposed

$$G^{(2)}(U(\omega)) = 5 - U_2(\omega) - \frac{1}{2}(U_1(\omega) - 0.1)^2. \qquad (3.2)$$

The probability of failure is equal to $P_f \approx 3.01 \cdot 10^{-3}$ and the failure domain is concave and consists of two distinct areas with high probability mass [94]. Figure 4 of Article III illustrates the failure domain.

Series system reliability problem

In the third example, we consider a series system reliability problem given in [117], which is defined by the LSF

$$G^{(3)}(U(\omega)) = \min \left\{ \begin{array}{c} 0.1(U_1(\omega) - U_2(\omega))^2 - (U_1(\omega) + U_2(\omega))/\sqrt{2} + 3 \\ 0.1(U_1(\omega) - U_2(\omega))^2 + (U_1(\omega) + U_2(\omega))/\sqrt{2} + 3 \\ U_1(\omega) - U_2(\omega) + 7/\sqrt{2} \\ U_2(\omega) - U_1(\omega) + 7/\sqrt{2} \end{array} \right\}. \qquad (3.3)$$

The probability of failure is equal to $P_f \approx 2.2 \cdot 10^{-3}$ and the failure domain consists of four distinct modes [94]. Figure 6 of Article III illustrates the failure domain.

The following two LSFs require the solution of an ODE or PDE. The exact LSF is given; however, we apply discretization schemes to determine an approximation of the LSF to test the novel error bounds in Article II, which considers the approximation error of the probability of failure. Moreover, we provide the orders of convergence of the discretization schemes with respect to h since these orders are relevant for the error bounds.

ODE, 1-dimensional parameter space

In this example, the LSF depends on the solution of an ODE with a one-dimensional Gaussian random parameter. We seek the solution $y : [0, 1] \times \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-*almost every* (a.e.) $\omega \in \Omega$

$$\frac{\partial y(t, \omega)}{\partial t} = -U(\omega) y(t, \omega), t \in (0, 1), \text{ with initial condition } y(0, \omega) = 1, \qquad (3.4)$$

where $U \sim \mathrm{N}(0, 1)$ is a standard normal random variable. The exact solution of (3.4) is $y(t, \omega) = \exp(-U(\omega)t)$. The solution $y(\cdot, \omega)$ leads to failure if it is lager than $y_{\max} = 40$ at $t = 1$. Thus, the LSF is expressed as

$$G^{(4)}(U(\omega)) = 40 - \exp(-U(\omega)). \qquad (3.5)$$

The exact probability of failure is equal to $P_f = \Phi(-\log(y_{\max})) \approx 1.13 \cdot 10^{-4}$ and the MLFP is $u^{\mathrm{MLFP}} = -\log(y_{\max})$. We use the explicit Euler scheme and the Crank–Nicolson scheme to determine an approximation y_h. By [73, Section 6.3], the explicit Euler scheme is convergent of order one, i.e., $|y(t, \omega) - y_h(t, \omega)| = \mathcal{O}(h)$ for a fixed $\omega \in \Omega$. The Crank–Nicolson scheme is convergent of order $\mathcal{O}(h^2)$ [73, Chapter 9].

PDE, 2-dimensional parameter space

The following example is based on an elliptic boundary value problem. This problem is also considered in [37, 43]. On the unit interval $D = (0, 1)$, we seek a solution $y : \overline{D} \times \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$-\frac{\partial}{\partial x} \left(\exp\left(\frac{U_1(\omega)}{3} - 3 \right) \frac{\partial}{\partial x} y(x, \omega) \right) = 1 - x, \quad \text{for } x \in (0, 1),$$
$$\text{such that } y(0, \omega) = 0 \text{ and } y(1, \omega) = U_2(\omega),$$

where U_1 and U_2 are independent and standard normally distributed. The solution of this problem is

$$y(x, \omega) = U_2(\omega)x + \exp\left(-U_1(\omega)/3 + 3\right)\left(x^3/6 - x^2/2 + x/3\right).$$

We define failure as the event that $y(\cdot, \omega)$ is smaller than $y_{\min} = -1/3$ at $\widehat{x} = 1/3$. Thus, the LSF is expressed as

$$G^{(5)}(U(\omega)) = y(1/3, \omega) + 1/3. \qquad (3.6)$$

The probability of failure is equal to $P_f \approx 1.71 \cdot 10^{-4}$. The failure domain is convex and consists of a single failure mode as shown in Section 5.2 of Article II. We apply FEs with either piecewise linear or piecewise quadratic basis functions. For linear basis functions, the FE error of the point evaluation is of order $\mathcal{O}(h^2)$, while for quadratic basis functions, the error is of order $\mathcal{O}(h^3)$ [32, Theorem 1.1].

3.2. High-dimensional parameter space

We consider two settings with a high-dimensional parameter space. The first setting is the elliptic diffusion equation, which is the required problem setting of the provided error bounds in Article II. Therefore, we introduce this setting in a general form before we concretely state the considered LSFs. The second setting is the flow cell in a two-dimensional computational domain, which is considered in Article I. In both settings, a random diffusion coefficient is employed, which is approximated by its truncated *Karhunen–Loève* (KL) expansion [63, 79].

3.2.1. Elliptic diffusion equation

Before we explicitly state the LSFs, we discuss the general problem setting of the elliptic diffusion equation, which is discussed in detail in [38, Chapter 6]. We consider a computational domain $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, which is open, bounded, convex and polygonal. Given is a real-valued random field $a : D \times \Omega \to \mathbb{R}$ and a real-valued function $f \in L^2(D)$. We seek a random field $y : \overline{D} \times \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$-\nabla_x \cdot (a(x, \omega)\nabla_x y(x, \omega)) = f(x) \quad \forall x \in D, \qquad (3.7)$$

$$y(x, \omega) = 0 \quad \forall x \in \partial D. \qquad (3.8)$$

Since a might be non-differentiable, a solution to this strong formulation might not exist. Therefore, we consider the *pathwise variational formulation* of (3.7), (3.8). For $V := H_0^1(D)$, we seek $y(\cdot, \omega) \in V$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$\int_D a(x, \omega)\nabla_x y(x, \omega) \cdot \nabla_x v(x)\mathrm{d}x = \int_D f(x)v(x)\mathrm{d}x \quad \forall v \in V. \qquad (3.9)$$

For the discretization parameter $h > 0$, we define the *discretized pathwise variational formulation*. In the finite-dimensional vector space $V_h \subset V$, we seek $y_h \in V_h$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$\int_D a(x, \omega)\nabla_x y_h(x, \omega) \cdot \nabla_x v_h(x)\mathrm{d}x = \int_D f(x)v_h(x)\mathrm{d}x \quad \forall v_h \in V_h. \qquad (3.10)$$

By [80, Theorem 9.9], existence and uniqueness of a solution for (3.9) and (3.10) is ensured if $a(\cdot, \omega) \in L^\infty(D)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $a(x, \omega)$ is *pathwise elliptic and bounded*, i.e., there exists random variables $a_{\min}, a_{\max} : \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$0 < a_{\min}(\omega) \le a(x, \omega) \le a_{\max}(\omega) < \infty, \quad \text{for a.e. } x \in D. \qquad (3.11)$$

Indeed, our provided error bounds in Article II require a stricter assumption. We require that the diffusion coefficient $a(x, \omega)$ is *uniformly elliptic and bounded*, i.e., $a_{\min}, a_{\max} > 0$ in (3.11) are independent of $\omega \in \Omega$. For the pathwise case, we conjecture in Remark 2.13 of Article II that a similar error bound holds.

The stochastic dependence of the solution y is given by the diffusion coefficient $a(x, \omega)$. To satisfy the Gaussian assumption in Chapter 2, we require that $a(x, \omega)$ is a measurable function of an n-variate standard Gaussian random vector $U : \Omega \to \mathbb{R}^n$. In the following examples, $a(x, \omega)$ is approximated by its truncated KL expansion $a_n(x, \omega)$, which implicitly depends on U.

In the numerical experiments of the contributed articles, the diffusion coefficient $a(x, \omega) = \exp(Z(x, \omega))$ is a log-normal random field, i.e., $Z(x, \omega)$ is a Gaussian random field with constant mean μ_Z and variance ζ_Z^2. The covariance function of Z is $c(x_1, x_2) = \zeta_Z^2 \exp\left(-\|x_1 - x_2\|_1 / \lambda\right)$ for $x_1, x_2 \in D$, where $\lambda > 0$ denotes the correlation length. We note that log-normal random fields are only pathwise elliptic and bounded. Thus, the assumptions of the error bounds in Article II are not satisfied. However, our conjecture given in Remark 2.13 of Article II is observed in the numerical experiments. The dependence on the n-variate Gaussian random variable U is given by truncating the KL expansion of the random field Z after the n leading KL terms, which gives the truncated KL expansion

$$Z_n(x, \omega) = \mu_Z + \zeta_Z \sum_{i=1}^{n} \sqrt{\nu_i} \theta_i(x) U_i(\omega),$$

where (ν_i, θ_i) are the KL eigenpairs and $\{U_i\}_{i=1}^{n}$ are independent standard normal Gaussian random variables. The determination of the eigenpairs is given in [46, Section 2.3.3]. The discretized diffusion coefficient is denoted as $a_n(x, \omega) := \exp(Z_n(x, \omega))$. Due to the truncation, $a_n(x, \omega)$ is a smooth function with respect to x and the outcomes $U(\omega)$. We note that in the numerical experiments of the contributed articles, the truncation order n is fixed, i.e., the limit $n \to \infty$ is not considered.

Having considered the relation between the PDE solution and the Gaussian random variable U, we can now define the LSF. The solution y and the LSF G are related via a linear and bounded operator $\mathcal{F} : H_0^1(D) \to \mathbb{R}$ such that

$$G(U(\omega)) = y_{\max} - \mathcal{F} y(\cdot, \omega),$$

where $y_{\max} \in \mathbb{R}$ is a threshold. Hence, failure is defined as the event that $\mathcal{F} y(\cdot, \omega) \geq y_{\max}$. In a similar way, we define the approximate LSF by

$$G_h(U(\omega)) = y_{\max} - \mathcal{F}_h y_h(\cdot, \omega),$$

where $\mathcal{F}_h : V_h \to \mathbb{R}$ is the induced observation operator. For instance, if the solution $y(\cdot, \omega)$ is continuous, $\mathcal{F}$ can be chosen as the point evaluation of y at $\widehat{x} \in \overline{D}$.

In Article I and III, we consider (3.7) and (3.8) in the unit interval $D = (0, 1) \subset \mathbb{R}$ with $f(x) = 1$ for all $x \in D$. The boundary conditions are mixed with $y(0, \omega) = 0$ and $y_x(1, \omega) = 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. We define failure as the event that the solution $y(\cdot, \omega)$ is larger than $y_{\max} = 0.535$ at $\widehat{x} = 1$. Thus, our sixth LSF is given by

$$G^{(6)}(U(\omega)) := 0.535 - y(1, \omega) \leq 0. \tag{3.12}$$

We apply piecewise linear, continuous FEs with mesh size $h > 0$ to derive the approximate LSF G_h. By [32, Theorem 1.1], the PDE discretization error behaves as $\mathcal{O}(h^2)$. By MCS, we estimate a reference for the exact probability of failure. On the grid with mesh size $h = 1/512$, we generate $J = 2 \cdot 10^8$ independent samples of U and estimate the probability of failure as $P_f \approx 1.68 \cdot 10^{-4}$.

In Article II, we consider an LSF which is proposed in [111]. We define $f(x) = 0$ for all $x \in D$ and we consider the boundary conditions $y(0, \omega) = 1$ and $y(1, \omega) = 0$. By the Sobolev embedding theorem [51, Theorem 6.48], the solution $y(\cdot, \omega)$ is continuously differentiable and the evaluation of $y_x(\cdot, \omega)$ is well-defined. Failure is defined as the event that the flow rate

$$q(x, \omega) := -a_n(x, \omega) \frac{\partial y(x, \omega)}{\partial x}$$

is larger than a threshold $q_{\max}$ at $\widehat{x} = 1$. Thus, the LSF is defined as

$$G^{(7)}(U(\omega)) = q_{\max} - q(1, \omega). \tag{3.13}$$

Again, piecewise linear, continuous FEs are employed with mesh size $h > 0$. By [110, Section 1.6], the PDE discretization order behaves as $\mathcal{O}(h)$. In Article II, we consider $\lambda_1 = 0.3$ and $\lambda_2 = 0.1$ and set the truncation order to $n_1 = 10$ and $n_2 = 50$. A reference for the probability of failure is obtained by the average of 100 SIS simulations with $J = 10^4$ samples per level on a grid with mesh size $h = 2^{-12}$. For λ_1, we set $q_{\max,1} = 1.7$ and obtain $P_f \approx 3.38 \cdot 10^{-4}$. For λ_2, we set $q_{\max,2} = 1.5$ and obtain $P_f \approx 7.18 \cdot 10^{-5}$.

3.2.2. 2D flow cell

The following LSF is considered in Article I and in [114, Section 6.1]. The LSF is based on the travel time of a radioactive particle within a two-dimensional flow cell $D = (0, 1) \times (0, 1)$. The LSF describes, in a simplified way, the rare event arising in planning a radioactive waste repository, which is described in the introduction. For more details on dynamics in porous media, we refer to [7]. We seek the hydrostatic pressure $y : D \times \Omega \to \mathbb{R}$ and the Darcy velocity $q : D \times \Omega \to \mathbb{R}^2$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$q(x, \omega) = -a(x, \omega) \nabla_x y(x, \omega) \quad \text{for } x \in D,$$
$$\nabla_x \cdot q(x, \omega) = 0 \quad \text{for } x \in D.$$

Moreover, y and q satisfy the boundary conditions

$$\vec{\nu} \cdot q(x, \omega) = 0 \quad \text{for } x \in (0, 1) \times \{0, 1\},$$
$$y(x, \omega) = 1 \quad \text{for } x \in \{0\} \times (0, 1),$$
$$y(x, \omega) = 0 \quad \text{for } x \in \{1\} \times (0, 1),$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$, where $\vec{\nu} \in \mathbb{R}^2$ denotes the direction of the outer normal on the boundary. The diffusion coefficient $a(x, \omega)$ describes the hydraulic conductivity of the porous medium. As in Section 3.2.1, $a(x, \omega) = \exp(Z(x, \omega))$ is modelled as a

log-normal random field. The mean μ_Z and variance ζ_Z of Z are constant. We obtain Z by the truncated eigendecomposition of the covariance function $c(x_1, x_2) = \zeta_Z^2 \exp\left(-\|x_1 - x_2\|_1/\lambda\right)$, where $x_1, x_2 \in \mathbb{R}^2$.

As motivated in the introduction, pollution of groundwater through a radioactive particle is a failure event. Thus, the travel time of the particles is crucial since short travel times lead to many radioactive particles. We assume that the particle path starts at $x_0 = (0, 0.5)^T$. The travel time is the time a particle requires to move from x_0 to any other point on the boundary ∂D. We define failure as the event that the travel time is smaller than the threshold $\tau_{\min} = 0.03$. The travel time is determined by the velocity of the particle, which is equal to the Darcy velocity $q(x, \omega)$ divided by the porosity $p(x, \omega)$. In this experiment, we assume, for simplicity, that $p(x, \omega) = 1$ is constant. Thus, we determine the particle path $x(t, \omega) \in \mathbb{R}^2$ by solving the ODE

$$\frac{\partial}{\partial t} x(t, \omega) = q(x(t, \omega), \omega), \quad x(0, \omega) = x_0.$$

With $x(t, \omega)$, the travel time is determined by

$$\tau(\omega) = \underset{t > 0}{\operatorname{argmin}} \; x(t, \omega) \in \partial D.$$

Thus, the LSF is given by

$$G^{(8)}(U(\omega)) = \tau(\omega) - 0.03. \tag{3.14}$$

Since the exact solutions for y and q are not available, we employ lowest order Raviart–Thomas mixed FEs [101] to discretize the Darcy velocity q, while piecewise constant elements are employed to discretize the pressure y. The computational domain D is discretized by $2 \cdot 1/h^2$ uniform triangles, where h denotes the length of the legs of the triangles. After determining y_h and q_h, we apply the explicit Euler scheme to determine the approximate particle path as

$$x_h(t + \Delta t, \omega) = x_h(t, \omega) + \Delta t q_h(x_h(t, \omega), \omega), \; \text{where } \Delta t = \frac{h}{2\|q_h(x_h(t, \omega), \omega)\|_2}.$$

Thus, the travel time $\tau_h(\omega) \in [0, \infty)$ is given by

$$\tau_h(\omega) = \underset{t > 0}{\operatorname{argmin}} \; x_h(t, \omega) \in \partial D.$$

The reference solution of the probability of failure is $P_f \approx 4.67 \cdot 10^{-7}$ and is the estimated mean probability of failure over 100 realizations of SuS with $J = 10^4$ samples per level, mesh size $h = 1/128$, $\widehat{c} = 0.1$, and adaptive conditional sampling as the MCMC method without burn-in. We briefly discuss aCS in Section 4.3 as an MCMC algorithm which performs well even in high dimensions.

4. Multilevel Sequential Importance Sampling

In this chapter, we derive the Multilevel Sequential Importance Sampling algorithm of Article I, which is the combination of the SIS algorithm in [94] and the multilevel approach of the MLS^{2}MC algorithm in [70]. Therein, the level updating idea of [66] is employed, which is also known as *bridging*. We note that this multilevel approach is different to the telescoping sum approach of [47, 54]. The MLSIS algorithm extends the collection of multilevel methods for rare event estimation. The motivation of the MLSIS algorithm is the nestedness issue of MLSuS, which leads to additional costs as we have observed in Section 2.2.2. However, MLSuS still requires less computational costs than SuS as observed in [114]. In MLSIS, nestedness is automatically satisfied since the support of the MLSIS densities is the whole parametric domain.

In the following, we discuss the MLSIS algorithm, which consists of a smart heuristic of level and tempering updates. Moreover, we discuss the vMFN distribution model as an independent proposal density in the MCMC algorithm. Indeed, this is another contribution of Article I. This chapter is concluded with the discussion of the numerical experiments of Article I, where the performance of MLSIS is compared with SIS, SuS and MLSuS.

4.1. MLSIS densities

We consider the sequence of discretization levels $\ell = 1, \ldots, L$. On each discretization level, an approximate LSF G_{h_ℓ} is given, where $h_\ell > 0$ is the discretization parameter, e.g. FE mesh size. The goal of MLSIS is the estimation of the probability of failure $P_{f,h_L} = \mathbb{P}(G_{h_L}(U) \leq 0)$ with respect to the finest discretization level L. We note that P_{f,h_L} yields an approximation of the exact probability of failure P_f that is determined with respect to the exact LSF G. The approximation error of P_{f,h_L} is studied in Chapter 5.

As in SIS [94], MLSIS is based on an approximation of the optimal IS density at level L, which is given by

$$p_{\mathrm{opt},L}(u) := \frac{1}{P_{f,h_L}} I(G_{h_L}(u) \leq 0)\varphi_n(u). \tag{4.1}$$

In MLSIS, we approximate the indicator function in (4.1) by the approximate LSF G_{h_ℓ} and the CDF of the univariate normal distribution, which yields

$$I(G_{h_L}(u) \leq 0) = \lim_{h_\ell \downarrow h_L} \lim_{\sigma \downarrow 0} \Phi\left(-\frac{G_{h_\ell}(u)}{\sigma}\right), \quad \text{for } G_{h_L}(u) \neq 0. \tag{4.2}$$

From (4.2), we observe that the optimal IS density is reached in a twofold sequential way. In the MLSIS algorithm, we sequentially decrease the temperature σ and decrease the mesh size h_ℓ. As in SIS, we denote by $\{\sigma_k\}_{k=0}^{N_T}$ the sequence of temperatures, which satisfies $\infty = \sigma_0 > \sigma_1 > \cdots > \sigma_{N_T} > 0$. In the tempering update from

σ_{k-1} to σ_k, for $k = 2, \ldots, N_T$, we consider the densities

$$p_{k-1,\ell}^{\text{MLSIS}}(u) := \frac{1}{P_{k-1,\ell}^{\text{MLSIS}}} \Phi\left(-\frac{G_{h_\ell}(u)}{\sigma_{k-1}}\right) \varphi_n(u),$$

$$p_{k,\ell}^{\text{MLSIS}}(u) := \frac{1}{P_{k,\ell}^{\text{MLSIS}}} \Phi\left(-\frac{G_{h_\ell}(u)}{\sigma_k}\right) \varphi_n(u),$$

where $P_{k-1,\ell}^{\text{MLSIS}}$ and $P_{k,\ell}^{\text{MLSIS}}$ are normalizing constants. The starting density $p_{0,1}^{\text{MLSIS}}(u) = \varphi_n(u)$ is the PDF of the n-variate standard normal distribution on the coarsest level $\ell = 1$. The determination of the sequence of temperatures $\{\sigma_k\}_{k=1}^{N_T}$ is similar to SIS. Given the temperature σ_{k-1} and the target coefficient of variation δ_{target}, the temperature σ_k is determined adaptively by

$$\sigma_k = \operatorname*{argmin}_{\sigma \in (0, \sigma_{k-1})} \left(\delta_{w_{k,\ell}} - \delta_{\text{target}}\right)^2, \quad \text{where } \delta_{w_{k,\ell}} := \frac{\text{StD}_{p_{k-1,\ell}^{\text{MLSIS}}}[w_{k,\ell}(U)]}{\mathbb{E}_{p_{k-1,\ell}^{\text{MLSIS}}}[w_{k,\ell}(U)]}$$

and the weights are given by

$$w_{1,1}(u) = \Phi(-G_{h_1}(u)/\sigma_1), \qquad w_{k,\ell}(u) = \frac{\Phi(-G_{h_\ell}(u)/\sigma_k)}{\Phi(-G_{h_\ell}(u)/\sigma_{k-1})}, \quad \text{for } k > 1.$$

As in SIS, the densities $p_{k,\ell}^{\text{MLSIS}}$ are approximated by a collection of samples $\{u_{k,\ell}^{(j)}\}_{j=1}^J$. During the tempering update, the samples $\{u_{k-1,\ell}^{(j)}\}_{j=1}^J$ are transformed to the samples $\{u_{k,\ell}^{(j)}\}_{j=1}^J$ with an MCMC algorithm. The acceptance function given in (2.13) is adjusted by replacing G with G_{h_ℓ}. With the samples $\{u_{k-1,\ell}^{(j)}\}_{j=1}^J$, the fraction $S_{k,\ell} := P_{k,\ell}^{\text{MLSIS}}/P_{k-1,\ell}^{\text{MLSIS}}$ of two consecutive normalizing constants is estimated by the sample mean

$$\widehat{S_{k,\ell}} := \frac{1}{J} \sum_{j=1}^J w_{k,\ell}(u_{k-1,\ell}^{(j)}). \tag{4.3}$$

Having considered the tempering updates, we discuss now the level update. For the level update from level ℓ to level $\ell + 1$, for $\ell = 1, \ldots, L - 1$, the MLSIS densities of these levels are given by

$$p_{k,\ell}^{\text{MLSIS}}(u) := \frac{1}{P_{k,\ell}^{\text{MLSIS}}} \Phi\left(-\frac{G_{h_\ell}(u)}{\sigma_k}\right) \varphi_n(u), \tag{4.4}$$

$$p_{k,\ell+1}^{\text{MLSIS}}(u) := \frac{1}{P_{k,\ell+1}^{\text{MLSIS}}} \Phi\left(-\frac{G_{h_{\ell+1}}(u)}{\sigma_k}\right) \varphi_n(u). \tag{4.5}$$

We note that a level update can only be carried out after the first temperature σ_1 has been determined. The level update from ℓ to $\ell + 1$ is performed in $N_{B_\ell} > 0$ bridging steps. During these bridging steps, tempering is not performed. We denote the sequence of bridging temperatures by $0 = \beta_0 < \beta_1 < \cdots < \beta_{N_{B_\ell}} = 1$ to define the

bridging densities between $p_{k,\ell}^{\mathrm{MLSIS}}$ and $p_{k,\ell+1}^{\mathrm{MLSIS}}$. According to [66, 70], these bridging densities are given by

$$p_{k,\ell,t}^{\mathrm{MLSIS}}(u) := \frac{1}{P_{k,\ell,t}^{\mathrm{MLSIS}}} \Phi\left(-\frac{G_{h_{\ell+1}}(u)}{\sigma_k}\right)^{\beta_t} \Phi\left(-\frac{G_{h_\ell}(u)}{\sigma_k}\right)^{1-\beta_t} \varphi_n(u),$$

for $t = 0, \ldots, N_{B_\ell}$, where $P_{k,\ell,t}^{\mathrm{MLSIS}}$ are normalizing constants. We observe that $p_{k,\ell,0}^{\mathrm{MLSIS}}$ yields the MLSIS density on the level ℓ, which is given in (4.4), while $p_{k,\ell,N_{b_\ell}}^{\mathrm{MLSIS}}$ yields the MLSIS density on the level $\ell+1$, which is given in (4.5). The determination of the bridging temperatures $\{\beta_t\}_{t=1}^{N_{B_\ell}}$ is similar to the determination of the temperatures $\{\sigma_k\}_{k=1}^{N_T}$. Given the bridging temperature β_{t-1}, the bridging temperature β_t is determined adaptively by

$$\beta_t = \operatorname*{argmin}_{\beta \in (\beta_{t-1}, 1]} \left(\delta_{w_{k,\ell,t}} - \delta_{\mathrm{target}}\right)^2, \quad \text{where } \delta_{w_{k,\ell,t}} := \frac{\mathrm{StD}_{p_{k,\ell,t-1}^{\mathrm{MLSIS}}}[w_{k,\ell,t}(U)]}{\mathbb{E}_{p_{k,\ell,t-1}^{\mathrm{MLSIS}}}[w_{k,\ell,t}(U)]}$$

and the weights are given by

$$w_{k,\ell,t}(u) = \frac{\Phi(-G_{h_{\ell+1}}(u)/\sigma_k)^{\beta_t}\Phi(-G_{h_\ell}(u)/\sigma_k)^{1-\beta_t}}{\Phi(-G_{h_{\ell+1}}(u)/\sigma_k)^{\beta_{t-1}}\Phi(-G_{h_\ell}(u)/\sigma_k)^{1-\beta_{t-1}}}. \tag{4.6}$$

During the bridging update, the samples $\{u_{k,\ell,t-1}^{(j)}\}_{j=1}^J$ are transformed to the samples $\{u_{k,\ell,t}^{(j)}\}_{j=1}^J$ with an MCMC algorithm. To apply the MCMC algorithm of Section 2.2.1 to a bridging update, the acceptance function in (2.13) has to be replaced by $\alpha^{\mathrm{MLSIS}}(u_0, \overline{u}) = \min\{1, \alpha^{\mathrm{B}}(u_0, \overline{u})\}$, where

$$\alpha^{\mathrm{B}}(u_0, \overline{u}) := \frac{\Phi(-G_{h_{\ell+1}}(\overline{u})/\sigma_k)^{\beta_t}\Phi(-G_{h_\ell}(\overline{u})/\sigma_k)^{1-\beta_t}\varphi_n(\overline{u})q(u_0 \mid \overline{u})}{\Phi(-G_{h_{\ell+1}}(u_0)/\sigma_k)^{\beta_t}\Phi(-G_{h_\ell}(u_0)/\sigma_k)^{1-\beta_t}\varphi_n(u_0)q(\overline{u} \mid u_0)}. \tag{4.7}$$

With the samples $\{u_{k,\ell,t-1}^{(j)}\}_{j=1}^J$, the fraction $S_{k,\ell,t} := P_{k,\ell,t}^{\mathrm{MLSIS}}/P_{k,\ell,t-1}^{\mathrm{MLSIS}}$ of two consecutive normalizing constants is estimated by the sample mean

$$\widehat{S_{k,\ell,t}} := \frac{1}{J} \sum_{j=1}^J w_{k,\ell,t}(u_{k,\ell,t-1}^{(j)}).$$

Thus, the estimator of the fraction $S_{k,\ell}^{\ell+1} := P_{k,\ell+1}^{\mathrm{MLSIS}}/P_{k,\ell}^{\mathrm{MLSIS}}$ is given by

$$\widehat{S_{k,\ell}^{\ell+1}} := \prod_{t=1}^{N_{B_\ell}} \widehat{S_{k,\ell,t}}. \tag{4.8}$$

The MLSIS estimator for the probability of failure is given by the combination of the estimators in (4.3) and (4.8). Starting from the initial density $p_{0,1}^{\mathrm{MLSIS}}(0) = \varphi_n(u)$, we transform samples of $\varphi_n(u)$ into samples of the final density $p_{N_T,L}^{\mathrm{MLSIS}}$. The way from the starting density to the final density is determined by the multilevel update scheme, which is described in the next section. An example of this scheme is shown

in Figure 2. To state the MLSIS estimator, we denote by (k, ℓ_k), for $k = 1, \ldots, N_T$, the indices of tempering updates defined on the level ℓ_k. Moreover, we denote by (k_ℓ, ℓ), for $\ell = 1, \ldots L - 1$, the indices of level updates defined on the temperature k_ℓ. Thus, the MLSIS estimator is given by

$$\widehat{P_{f,h_L}^{\mathrm{MLSIS}}} := \left(\prod_{k=1}^{N_T} \widehat{S_{k,\ell_k}} \right) \left(\prod_{\ell=1}^{L-1} \widehat{S_{k_\ell,\ell}^{\ell+1}} \right) \left(\frac{1}{J} \sum_{j=1}^{J} \frac{I(G_{h_L}(u_{N_T,L}^{(j)}) \leq 0)}{\Phi(-G_{h_L}(u_{N_T,L}^{(j)})/\sigma_{N_T})} \right), \qquad (4.9)$$

where the samples $\{u_{N_T,L}^{(j)}\}_{j=1}^{J}$ are distributed approximately according to $p_{N_T,L}^{\mathrm{MLSIS}}$ and the sum in (4.9) represents the last tempering step from $p_{N_T,L}^{\mathrm{MLSIS}}$ to the optimal IS density $p_{\mathrm{opt},L}$, which is given in (4.1). As in SIS, the generated samples $\{u_{k,\ell,t}^{(j)}\}_{j=1}^{J}$ are only distributed approximately according to the target densities $p_{k,\ell,t}^{\mathrm{MLSIS}}$ due to the adaptive choice of σ_k and β_t [9]. Thus the MLSIS estimator is biased.
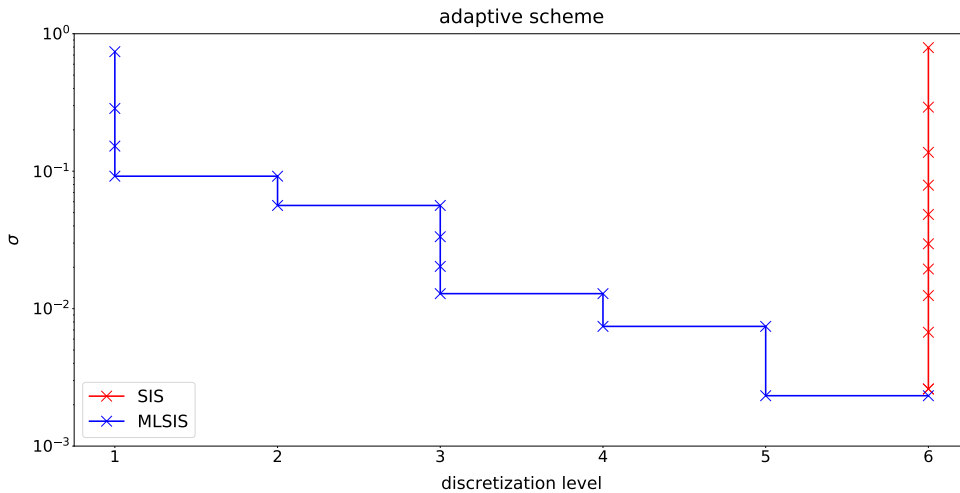

Figure 2: Example path of the multilevel update scheme of MLSIS and the tempering updates of SIS.

4.2. Multilevel update scheme

In this section, we discuss the heuristic of the multilevel update scheme, which has been developed in [70]. The scheme achieves that many tempering updates are performed on a low discretization level.

The scheme starts with the initial temperature $\sigma_0 = \infty$ and initial level $\ell = 1$. The initial samples $\{u_{0,1}^{(j)}\}_{j=1}^{J}$ are distributed according to φ_n. The first iteration of MLSIS is always a tempering update, where the first temperature σ_1 is determined. In the subsequent step, we determine if tempering or a level update is performed. Since the goal of multilevel methods is the reduction of high-level LSF evaluations, we update the level if the discrepancy between evaluations of two consecutive levels is too large. Therefore, we randomly select without replacement a small subset of samples $\{u_{k,\ell}^{(j_i)}\}_{i=1}^{J_{\mathrm{small}}}$ with $J_{\mathrm{small}} < J$ and we update the level for this subset through one bridging step. In the numerical experiments of Article I, we set $J_{\mathrm{small}} = 0.1 \cdot J$.

The discrepancy of the levels is measured by the coefficient of variation $\delta_{w_{k,\ell+1}^{J_{\text{small}}}}$ of the weights

$$w_{k,\ell+1}^{J_{\text{small}}}(u_{k,\ell}^{(j_i)}) = \frac{\Phi(-G_{h_{\ell+1}}(u_{k,\ell}^{(j_i)})/\sigma_k)}{\Phi(-G_{h_\ell}(u_{k,\ell}^{(j_i)})/\sigma_k)}, \quad \text{for } i = 1, \ldots, J_{\text{small}}. \tag{4.10}$$

In the case that $\delta_{w_{k,\ell+1}^{J_{\text{small}}}} > \delta_{\text{target}}$, we perform a level update since the discrepancy between the levels is high. In the other case, i.e., $\delta_{w_{k,\ell+1}^{J_{\text{small}}}} \leq \delta_{\text{target}}$, we perform tempering. We note that we always perform a level update if tempering has already finished. Moreover, tempering is always performed after a level update, if tempering has not already finished. To decide if tempering is finished, the coefficient of variation $\delta_{w_{\text{opt},k,\ell}}$ of the weights

$$w_{\text{opt},k,\ell}(u) = \frac{I(G_{h_\ell}(u))}{\Phi(-G_{h_\ell}(u)/\sigma_k)}$$

has to be calculated after each tempering and level update with respect to the samples $\{u_{k,\ell}^{(j)}\}_{j=1}^J$. If $\delta_{w_{\text{opt},k,\ell}} \leq \delta_{\text{target}}$, tempering is finished. The MLSIS iteration is finished if tempering is finished and the final discretization level $\ell = L$ is reached. The MLSIS algorithm is visualized in a diagrammatic representation in Figure 3 and a pseudo-code is shown in Algorithm 3.2 of Article I.
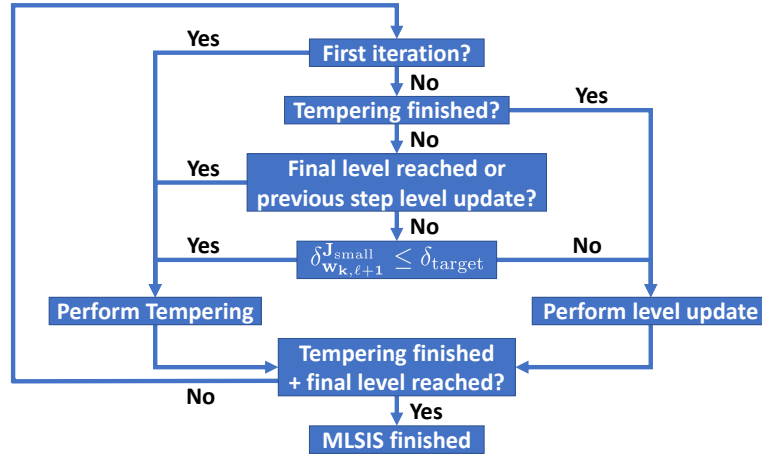

Figure 3: Diagrammic representation of the MLSIS algorithm.

Remark 4.1. *If the LSF depends on a random field that is approximated by a truncated Karhunen–Loève expansion, the authors of [114] propose a level dependent parameter dimension for the MLSuS estimator. This is due to the fact that highly oscillating KL terms cannot be accurately discretized on coarse grids, which leads to inaccurate evaluations. Employing a level dependent parameter dimension in the MLSIS algorithm reduces the discrepancy of LSF evaluations of two consecutive discretization levels, i.e., the coefficient of variation of the weights $w_{k,\ell+1}^{J_{\text{small}}}$ given in (4.10) is smaller. This leads to more tempering updates on coarse grids, which decreases the overall computational costs.*

To employ the level dependent parameter dimension for the MLSIS algorithm, we have to adjust the procedure of a level update. We denote the input dimensions by $n_\ell \in \mathbb{N}$ on the level $\ell = 1, \ldots, L$. To evaluate the weights in (4.6), the evaluations of G_{h_ℓ} and $G_{h_{\ell+1}}$ are required for the state $u \in \mathbb{R}^{n_\ell}$. Therefore, we denote by $\Delta n_\ell := n_{\ell+1} - n_\ell$ the difference of the input dimensions and we generate $\psi \in \mathbb{R}^{\Delta n_\ell}$ from the Δn_ℓ-variate standard normal distribution. Then, we define $\widetilde{u} = (u^T, \psi^T)^T \in \mathbb{R}^{n_{\ell+1}}$ and we evaluate G_{h_ℓ} for u and $G_{h_{\ell+1}}$ for $\widetilde{u}$. For the evaluation of the acceptance function in (4.7), we apply the same procedure for u_0 and $\overline{u}$.

4.3. MCMC with von Mises–Fisher–Nakagami distribution

In this section, we discuss the von Mises–Fisher–Nakagami distribution model as an independent proposal density in the MCMC algorithm. We note that random walk samplers, such as the random walk Metropolis–Hastings algorithm [53, 85], suffer from the curse of dimensionality. This is reflected by a small acceptance rate in high dimensions. In [92], the authors propose adaptive conditional sampling as an alternative method. The aCS method is a dependent MCMC algorithm, i.e., the proposal density depends on the current state of the Markov chain. In detail, the proposal density of aCS is given by $q(\overline{u} \mid u_0) = \varphi_n(\overline{u} \mid \rho u_0, (1 - \rho)^2 \mathrm{Id}_n)$, i.e., the PDF of the normal distribution with mean vector ρu_0 and covariance matrix $(1 - \rho)^2 \mathrm{Id}_n$. During the aCS iterations, the parameter $\rho \in [0, 1]$ is adaptively adjusted to ensure an acceptance rate around 44%, which has been suggested as an optimal rate by [104]. Due to the choice of the proposal density, aCS can be viewed as an adaptive version of the preconditioned Crank–Nicolson algorithm [24]. Contrary, in an independent MCMC algorithm, the proposal density does not depend on the current state u_0. Thus, the simulated states are less correlated and the final estimator of the probability of failure has usually a lower variance. According to [20, 94], we choose a parametric distribution model as the proposal density. Therefore, the parameters of the model have to be fitted in advance to give an approximation of the target density.

An example of a distribution model is the *Gaussian model* and the *Gaussian mixture model* [42, Section 1]. However, Gaussian densities suffer from the curse of dimensionality, since the Gaussian measure is concentrated around the hypersphere with norm equal to $\sqrt{n}$ in high dimensions [64, 93]. Therefore, we propose the vMFN distribution model as an alternative proposal density in Article I. This distribution model is proposed in [93] as a parametric model within cross-entropy-based IS. To apply this parametric model, we represent the parameter vector $u \in \mathbb{R}^n$ in its polar coordinate representation $u = r \cdot \vec{u}$, where $r = \|u\|_2$ denotes the norm of u and $\vec{u} = u/\|u\|_2 \in \mathbb{R}^n$ its normalized direction. In the vMFN distribution model, the direction $\vec{u}$ and radius r are sampled from different distribution models such that the vMFN distribution model is not affected by the concentration of measure issue in high dimensions. Given the direction $\vec{u}$ and radius r, the vMFN distribution model is defined as

$$p_{\mathrm{vMFN}}(r, \vec{u} \mid \mu, \kappa, \chi, \gamma) = p_{\mathrm{N}}(r \mid \chi, \gamma) \cdot p_{\mathrm{vMF}}(\vec{u} \mid \mu, \kappa).$$

The vMF distribution [123] $p_{\mathrm{vMF}}(\cdot \mid \mu, \kappa)$ defines the distribution of the direction on the n-dimensional hypersphere $\mathbb{S}^{n-1} := \{u \in \mathbb{R}^n : \|u\|_2 = 1\}$, where $\mu \in \mathbb{S}^{n-1}$ is a

mean direction and $\kappa \geq 0$ characterizes the concentration around μ. The Nakagami distribution [88] $p_{\mathrm{N}}(\cdot \mid \chi, \gamma)$ specifies the distribution of the radius, where $\chi \geq 0.5$ is a shape parameter and $\gamma > 0$ a spread parameter. In Figure 2 of Article I an illustration of p_{N} and p_{vMF} is shown for certain parameter values.

To apply the vMFN distribution model as a proposal density in the MLSIS algorithm, the parameters μ, κ, χ and γ have to be fitted in advance by maximizing the weighted log-likelihood. For μ and γ, explicit expressions are available as shown in [93]. However, for κ and χ, we use approximations as derived in [13, 123].

Remark 4.2. *For multi-modal failure domains, we propose to apply the vMFN mixture distribution model to capture the domains of the distinct failure regions. For $M > 1$ individual mixture components, the vMFN mixture distribution reads as*

$$p_{\mathrm{vMFN}}(r, \vec{u} \mid \boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{\chi}, \boldsymbol{\gamma}) = \sum_{m=1}^{M} \pi_m p_{\mathrm{vMFN}}(r, \vec{u} \mid \mu_m, \kappa_m, \chi_m, \gamma_m),$$

where the weights π_m represent the probability of each mode and $\sum_{m=1}^{M} \pi_m = 1$.

For the vMFN mixture distribution model, the samples have to be assigned to the modes. Therefore, the required parameters of the vMFN mixture model are estimated iteratively by the expectation-maximization algorithm [83]. We note that the number of parameters of the vMFN mixture distribution model scales as $\mathcal{O}(n)$. This yields computational benefits compared to the Gaussian mixture model, which requires $\mathcal{O}(n^2)$ parameters as discussed in [93].

4.4. Numerical experiments

In numerical experiments, we compare the performance of MLSIS with SIS, SuS and MLSuS in terms of the required computational costs and the achieved *relative root mean square error* (relRMSE), which is defined as

$$\mathrm{relRMSE} := \frac{\left(\mathbb{E}[(\widehat{P}_{f,h_L} - P_{f,h_L})^2] \right)^{1/2}}{P_{f,h_L}}, \tag{4.11}$$

where $\widehat{P}_{f,h_L}$ is the estimated probability of failure on the finest level L. We assume that the computational costs of one evaluation of G_{h_ℓ} are given by

$$\mathrm{Cost}(G_{h_\ell}) = 2^{-d(L-\ell)},$$

where d is the dimension of the computational domain D. Thus, one evaluation of G_{h_L} requires one unit of computational costs. We consider the LSF $G^{(6)}$ given in (3.12) and the LSF $G^{(8)}$ given in (3.14). We remind that the LSF $G^{(6)}$ requires the solution of the elliptic diffusion equation in the unit interval in 1D space given in (3.7) and we define failure as the event that the solution $y(x, \omega)$ is larger than $y_{\max} = 0.535$ at $\widehat{x} = 1$. The LSF $G^{(8)}$ considers the flow cell in 2D space and we define failure as the event that the travel time $\tau(\omega)$ of a radioactive particle is smaller than $\tau_{\min} = 0.03$. We apply SIS and MLSIS with varying number of samples J and

with different target coefficients of variations δ_{target}. Moreover, we consider aCS and the vMFN distribution model in the MCMC algorithm. SuS and MLSuS are also applied with a varying number of samples J and with $\widehat{c} = 0.1$ and $\widehat{c} = 0.25$. For SuS and MLSuS, we consider aCS as the MCMC algorithm.

For the LSF $G^{(6)}$, we observe in Figure 5 of Article I that SIS and MLSIS yield a similar relRMSE. However, MLSIS requires around 61% less computational costs than SIS for a fixed level of accuracy. Moreover, we observe that the vMFN distribution model in the MCMC algorithm leads to a smaller relRMSE than aCS. Figure 6 of Article I shows that MLSIS with the vMFN distribution model requires less computational costs than MLSuS with aCS for a fixed level of accuracy. However, the performance of MLSIS with aCS is similar to the performance of MLSuS with aCS. Indeed, MLSuS performs slightly better. Thus, we conclude for this experiment that sampling from the vMFN distribution model yields benefits compared to aCS and the performance of MLSIS is similar to the performance of MLSuS if both methods are applied with the same MCMC algorithm.

In the second experiment, we consider the LSF $G^{(8)}$. In Figure 9 of Article I, we observe that SIS and MLSIS reach the same level of accuracy. As in the first experiment, MLSIS requires less computational costs than SIS and the vMFN distribution model yields a smaller relRMSE than aCS. From Figure 10 of Article I, we observe that MLSIS requires less computational costs than MLSuS for a fixed level of accuracy. This holds true for both MCMC algorithms. For the second experiment, we conclude that MLSIS yields a better performance than MLSuS.

5. Error analysis for probabilities of rare events

This chapter is devoted to the error analysis given in Article II. We consider LSFs which require the solution of an elliptic PDE with stochastic diffusion coefficient. This problem setting is already discussed in Section 3.2.1.

The discretization of the PDE leads to an error in the evaluation of the LSF. This error leads to an approximation error of the probability of rare events. We note that PDE based LSFs arise frequently in rare event estimation and multilevel methods [36, 114, 119] have been developed as efficient estimators. However, these articles mostly deal with the sampling error. The approximation error has been considered so far only in [36]. Therein, the authors derive an upper bound of the absolute error, which scales as the PDE discretization error. The authors apply this bound to determine the optimal number of samples per level for the Multilevel Monte Carlo estimator. However, a systematic error analysis of the induced approximation error has been missing. Article II closes this gap, where we use the results of [36] as a starting point. Indeed, our motivation is the derivation of an upper bound of the relative error of the probability of failure.

We start with the relevant assumptions which we require to perform the error analysis. Thereafter, we discuss the main statements of Article II and outline the proof. In the proof, we observe an interesting connection to results from PDE based optimal control [100, 116]. Finally, we discuss the numerical experiments which are performed in Article II.

5.1. Relevant assumptions

In the following, we describe the problem setting and the relevant assumptions of Article II to perform the error analysis. The analysis is performed for LSFs which require the solution of an elliptic PDE with stochastic diffusion coefficient $a(x, \omega)$ and Dirichlet boundary condition. This is the problem setting of Section 3.2.1.

To apply the results from optimal control, we require that the solution y of the weak formulation (3.9) satisfies $y(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$. The regularity of the solution y is based on the properties of the computational domain D, the diffusion coefficient a and the right hand side f.

Assumption 5.1 (Assumption 2.3 of Article II)**.** *We assume that*

(i) the computational domain D is open, bounded, convex and polygonal,

(ii) $a_{\min}(\omega) \geq 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $1/a_{\min} \in L^p(\Omega)$ for all $p \in (0, \infty)$,

(iii) $a \in L^p(\Omega, C^1(\bar{D}))$ for all $p \in (0, \infty)$, i.e., the realisations $a(\cdot, \omega)$ are continuously differentiable,

(iv) $f \in L^2(D)$.

Under the above assumption, [113, Theorem 2.1] shows that $y(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$ is satisfied if $d = 2$. If $d = 3$ and D is convex, the authors state in [113, Remark 5.2 (c)] that the same property holds. If Assumption 5.1 (i) is replaced by requiring that D is open, bounded, and has a C^2 boundary, $y(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$ is also

satisfied by [19, Theorem 3.4]. We note that Assumption 5.1 (ii) is automatically satisfied for uniformly elliptic and bounded diffusion coefficients.

As in the previous chapters, we require that the LSF depends on the outcomes of an n-variate Gaussian random vector U. As seen in Section 3.2.1, this dependence is usually achieved by the truncated KL expansion of the stochastic diffusion coefficient. The following assumption is more general.

Assumption 5.2 (Assumption 2.4 of Article II).   *(i) The diffusion coefficient $a(x, \omega)$ is a measurable function of an n-variate random vector $U : \Omega \to \mathbb{R}^n$, where U follows the n-variate independent standard normal distribution. This means, there is a function $\widehat{a} : D \times \mathbb{R}^n \to \mathbb{R}$ with $a(x, \omega) = \widehat{a}(x, U(\omega))$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*

(ii) The diffusion coefficient $a(x, \omega)$ is three times continuously differentiable with respect to outcomes $u \in \mathbb{R}^n$ of U for all $x \in D$.

Assumption 5.2 (ii) is another regularity assumption which we require to apply the optimal control results of [100, 116].

The main goal of Article II is the derivation of an error bound for the approximation error of the probability of failure, which behaves as the PDE discretization error. For the latter, we assume that the approximation error of the LSFs $G(U(\omega)) = y_{\max} - \mathcal{F}y(\cdot, \omega)$ and $G_h(U(\omega)) = y_{\max} - \mathcal{F}_h y_h(\cdot, \omega)$ is uniformly bounded and behaves has $\mathcal{O}(h^s)$.

Assumption 5.3 (Assumption 2.5 of Article II). *The operator $\mathcal{F}$ is linear and bounded and there exists constants $C_{\mathrm{FE}} > 0$ and $s > 0$ independent of h such that the discretization error with respect to the solution of (3.9) and (3.10) satisfies for $\mathbb{P}$-a.e. $\omega \in \Omega$*

$$|G(U(\omega)) - G_h(U(\omega))| = |\mathcal{F}y(\cdot, \omega) - \mathcal{F}_h y_h(\cdot, \omega)| \le C_{\mathrm{FE}} h^s. \tag{5.1}$$

Since we require the uniform bound in (5.1), the diffusion coefficient has to be uniformly elliptic and bounded. Thus, our provided error bounds are not directly applicable for log-normal diffusion coefficients, since these are only pathwise elliptic and bounded. However, in Remark 2.13 of Article II, we outline an approach to handle the pathwise case – see also Remark 5.10.

Another assumption for the main statements is the requirement that the CDFs of the random variables $G(U)$ and $G_h(U)$ are Lipschitz continuous. We note that Assumption 5.3 and 5.4 are also relevant for the bound of the absolute error in [36, Lemma 3.4 M1], which forms our starting point.

Assumption 5.4 (Assumption 2.6 of Article II). *The CDFs of the random variables $G(U)$ and $G_h(U)$ are local Lipschitz continuous with Lipschitz constants $C_L > 0$ and $C_{L,h} > 0$, i.e., for a, b with $a < b$ it holds*

$$\mathbb{P}[G(U) \in]a, b]] = \mathbb{P}[G(U) \le b] - \mathbb{P}[G(U) \le a] \le C_L|a - b|,$$
$$\mathbb{P}[G_h(U) \in]a, b]] = \mathbb{P}[G_h(U) \le b] - \mathbb{P}[G_h(U) \le a] \le C_{L,h}|a - b|.$$

As discussed in Section 2.3, the FORM estimate is an upper bound of the probability of failure if the failure domain is convex. In the proof outline, we will see that the limit-state surface of a convex failure domain can be translated to the limit-state surface of the linearized failure domain. The convexity assumption is summarized in the following assumption.

Assumption 5.5 (Assumption 2.9 of Article II). *The failure domains A and A_h are unbounded, convex sets.*

The next assumption states that the gradients $\nabla_u G$ and $\nabla_u G_h$ are not degenerated at the limit-state surface. Moreover, we require that the direction from a point $u \in \partial A$ to its nearest neighbour $u_h \in \partial A_h$ is not orthogonal to $\nabla_u G(u)$.

Assumption 5.6 (Assumption 2.10 of Article II). *For all $h > 0$ there exists $\nu_h > 0$ such that for almost every $u \in \partial A$ it holds $\nabla_u G(u) \neq 0$, $\nabla_u G_h(u) \neq 0$ and $|\cos(\lhd(u - u_h, \nabla_u G(u)))| \geq \nu_h$, where $u_h \in \partial A_h$ is the point that has minimal distance to u and $\lhd(\cdot, \cdot)$ denotes the angle between two vectors.*

5.2. Error bounds

In the following, we state the main results of Article II. Proposition 5.7 shows that the relative error of the FORM estimates behaves as the PDE discretization error. We note that this proposition does not require convexity of the failure domains. Thus, this proposition is in general more applicable since a priori knowledge of the geometric properties of the failure domains is not required.

Proposition 5.7 (Proposition 2.11 of Article II). *Let $a(x, \omega)$ be a uniformly elliptic and bounded diffusion coefficient and let Assumptions 5.1, 5.2, 5.3, 5.4, and 5.6 hold. Then for $h > 0$ sufficiently small, the relative error of the FORM estimates is upper bounded by*

$$\frac{|P_f^{\mathrm{FORM}} - P_{f,h}^{\mathrm{FORM}}|}{P_f^{\mathrm{FORM}}} \leq \widehat{C}^{\mathrm{FORM}} h^s. \tag{5.2}$$

In contrast to the above proposition, Theorem 5.8 requires convexity of the failure domains. In this case, the absolute error of the probability of failure behaves as the PDE discretization error multiplied by the FORM estimate.

Theorem 5.8 (Theorem 2.12 of Article II). *Let $a(x, \omega)$ be a uniformly elliptic and bounded diffusion coefficient and let Assumptions 5.1, 5.2, 5.3, 5.4, 5.5 and 5.6 hold. Then for $h > 0$ sufficiently small, the error of the exact and approximate probability of failure is upper bounded by*

$$|P_f - P_{f,h}| \leq \widehat{C} h^s P_{f,h}^{\mathrm{FORM}}. \tag{5.3}$$

As we show in Article II, the constants $\widehat{C}^{\mathrm{FORM}} > 0$ in (5.2) and $\widehat{C} > 0$ in (5.3), respectively, depend on C_{FE}, h, n, $\|u^{\mathrm{MLFP}}\|_2$, and $\|u_h^{\mathrm{MLFP}}\|_2$.

We note that Proposition 5.7 and Theorem 5.8 require the problem setting of Subsection 3.2.1, i.e., the LSF depends on the solution of an elliptic diffusion equation.

The following theorem holds for general affine linear LSFs of the form $G(U) = a^T U + b$, where $a \in \mathbb{R}^n$, $b > 0$ and $U \sim \mathrm{N}(0, \mathrm{Id}_n)$. For affine linear LSFs, we show that the relative error of the probability of failure behaves as the PDE discretization error, where the derived constant $\widehat{C}^{\text{linear}}$ in (5.4) depends on b, a, C_{FE} and h.

Theorem 5.9 (Theorem 3.2 of Article II). *Let Assumptions 5.3 and 5.4 hold and $G(U) = a^T U + b$ be an affine linear LSF. The discretization parameter is chosen such that $-b + C_{\text{FE}} h^s < 0$. Then, the relative approximation error of the probability of failure is bounded by*

$$\frac{|P_f - P_{f,h}|}{P_f} \le \widehat{C}^{\text{linear}} h^s. \tag{5.4}$$

Remark 5.10. *We remark that the above statements require that the approximation error of the LSF is uniformly bounded. This requirement is in general not satisfied for diffusion coefficients which are only pathwise elliptic and bounded. For instance, log-normal diffusion coefficients are among this class. In Remark 2.13 of Article II, we conjecture that similar error bounds hold for the pathwise case. We claim that the random variable U can be restricted to a bounded domain $B_R := \{u \in \mathbb{R}^n : \|u\|_2 \le R\}$ and the random variable $C_{\text{FE}}(\omega)$ is uniformly bounded within B_R. Thus, the analysis is performed after restricting U to B_R. This approach can be seen as truncating the tails of U, which yields a uniformly elliptic and bounded diffusion coefficient.*

5.3. Proof outline

In the following, we give an outline of the proofs of the above statements. This outline is based on the steps (P1)–(P5) of Article II. The complete proofs are given in Article II. We visualize the steps of the proof by considering an artificial example. Figure 4 shows an illustration of an exact and approximate failure domain A and A_h, respectively, which are defined on a two-dimensional parameter space. We note that there is no mathematical model behind the LSF.
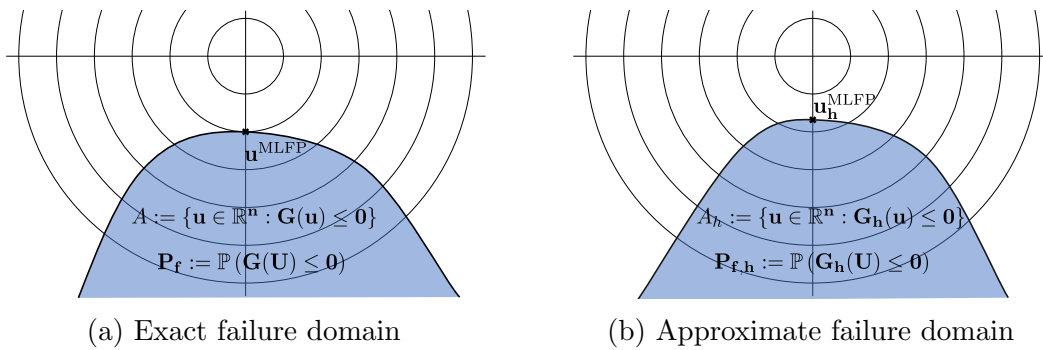

(a) Exact failure domain   (b) Approximate failure domain

Figure 4: The left plot shows the exact failure domain A and the right plot shows the approximate failure domain A_h of the artificial LSFs G and G_h.

In (P1), we show that the absolute error $|P_f - P_{f,h}|$ behaves as h^s, which is the result of [36, Lemma 3.4 M1]. For this result, we require Assumption 5.3 and 5.4. We start

by showing that the absolute error of the probability of failure is upper bounded by the Gaussian measure of the symmetric difference $A \triangle A_h$. This result is expressed as

$$|P_f - P_{f,h}| \leq \mathbb{P}\left(\{G(U) \leq 0\} \cap \{G_h(U) > 0\}\right) + \mathbb{P}\left(\{G(U) > 0\} \cap \{G_h(U) \leq 0\}\right)$$
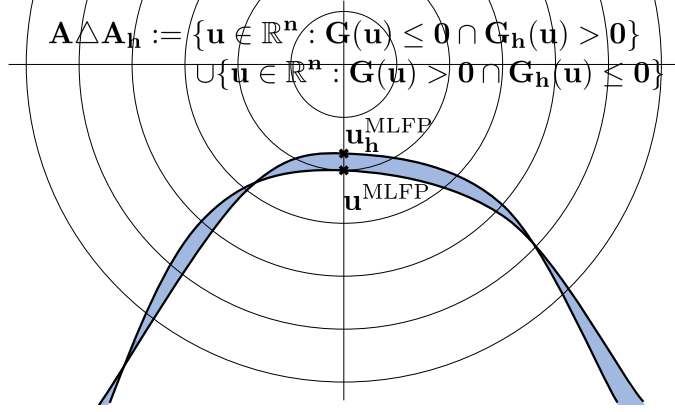$$= \mathbb{P}\left(A \triangle A_h\right).$$

Figure 5: Symmetric difference of the failure domains.

The symmetric difference is visualized in Figure 5. We observe that the absolute error is based on the Gaussian measure of the collection of states $u \in \mathbb{R}^n$, where the sign of G and G_h differs. By Assumption 5.3, this is only possible if $G(U) \in]-C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]$ and we conclude

$$|P_f - P_{f,h}| \leq \mathbb{P}\left(G(U) \in]-C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]\right) \leq 2C_L C_{\mathrm{FE}}h^s, \qquad (5.5)$$

where the final inequality follows from the Lipschitz continuity of the CDF of $G(U)$ given in Assumption 5.4. This concludes (P1).

In (P2), we show that C_L behaves as the probability of failure P_f if G is affine linear. Indeed, this proves Theorem 5.9. We assume that the LSF is given by $G(U) = a^T U + b$, where $a \in \mathbb{R}^n$ and $b > 0$ such that $-b + C_{\mathrm{FE}}h^s < 0$. Under this assumption, we upper bound the local Lipschitz constant in the intervals $G(U) \in]-C_{\mathrm{FE}}h^s, 0]$ and $G(U) \in [0, C_{\mathrm{FE}}h^s]$ separately. In the first interval, we show that

$$C_L \leq \left(\frac{b}{\zeta^2} + \frac{1}{b} + \frac{1}{b\zeta^2} + \frac{1}{b^3}\right) P_f, \qquad (5.6)$$

where $\zeta^2 := \|a\|_2^2$. For the second interval, we derive a similar result. Under the assumption that $-b + C_{\mathrm{FE}}h^s < 0$, we show that the local Lipschitz constant in the second interval is upper bounded by

$$C_L \leq \left(b + \frac{1}{b}\right)\left(\frac{1}{\zeta^2} + \frac{1}{(b - C_{\mathrm{FE}}h^s)^2}\right) \exp\left(\frac{2bC_{\mathrm{FE}}h^s - C_{\mathrm{FE}}^2 h^{2s}}{2\zeta^2}\right) P_f. \qquad (5.7)$$

Indeed, (5.6) and (5.7) require that U is a Gaussian random variable. Combining (5.6) and (5.7) yields an upper bound of the form $2C_L C_{\mathrm{FE}} \leq C_2\left(b, \zeta, h^s, C_{\mathrm{FE}}\right) P_f$. Plugging this expression into (5.5) concludes (P2) and proves Theorem 5.9.

In (P1), we have observed that the symmetric difference has a significant role in the error analysis. The symmetric difference can be expressed by the distances of the limit-state surfaces ∂A and ∂A_h. In (P3), we prove that the distance of ∂A and ∂A_h behaves as h^s. Indeed, we show that for all $u \in \partial A$ it holds

$$\text{dist}(u, \partial A_h) \leq C_3 h^s. \tag{5.8}$$

In (P3), we apply an a priori error bound which has been developed in [116] for optimal control problems. In particular, for $u^* \in \partial A$, we investigate the optimal control problem

$$\min_{u \in \mathbb{R}^n} \quad J_\rho(u) := \frac{1}{2} G(u)^2 + \frac{\rho}{2} \|u - u^*\|_2^2, \tag{5.9}$$

where $\rho > 0$ is a regularizing parameter. Indeed, the point u^* is a stable solution of (5.9). By [116, Theorem 3.4.1 and 3.4.2], there exists $\widehat{u}_h$ in a neighbourhood of u^*, where $\widehat{u}_h$ is a stable solution of the discretized optimal control problem

$$\min_{u \in \mathbb{R}^n} \quad J_{h,\rho}(u) := \frac{1}{2} G_h(u)^2 + \frac{\rho}{2} \|u - u^*\|_2^2.$$

Moreover, with Assumption 5.6, we derive that $\|u^* - \widehat{u}_h\|_2 = \mathcal{O}(h^s)$. Considering the limit $\rho \to 0$, we conclude that $\widehat{u}_h \to \partial A_h$ which shows (P3), i.e., the distance of the exact and approximate limit-state surface is upper bounded by $C_3 h^s$. This statement is visualized in Figure 6.
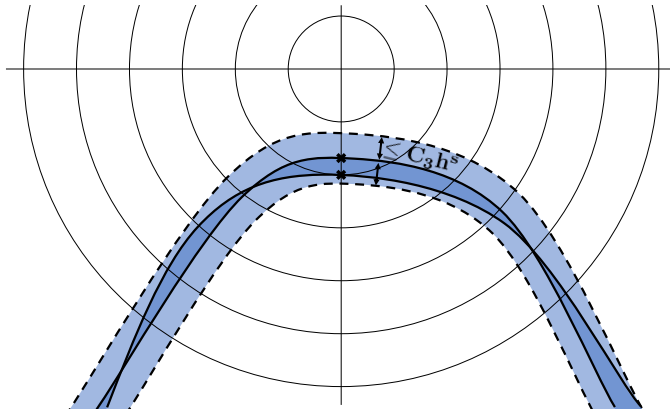

Figure 6: The distance between the exact and approximate limit-state surface behaves as $C_3 h^s$.

With (P3), we prove Proposition 5.7. We denote by $b := \|u^{\text{MLFP}}\|_2$ and $b_h := \|u_h^{\text{MLFP}}\|_2$ the distance of the MLFPs to the origin. From (P3), we derive that the distance of the MLFPs behaves as $|b - b_h| \leq C_3 h^s$. By definition, it holds that $P_f^{\text{FORM}} = \mathbb{P}(U_1 \leq -b)$ and $P_{f,h}^{\text{FORM}} = \mathbb{P}(U_1 \leq -b_h)$, where U_1 is a univariate standard normal random variable. Thus, we can show that the absolute error of the FORM estimate is upper bounded by

$$|P_f^{\text{FORM}} - P_{f,h}^{\text{FORM}}| \leq \mathbb{P}\left(U_1 \in]-b - C_3 h^s, -b + C_3 h^s]\right). \tag{5.10}$$

We define the affine linear LSF $\widetilde{G}(U_1) = U_1 + b$. We assume that $\widetilde{G}$ satisfies Assumption 5.3 with $\widetilde{C}_{\mathrm{FE}} = C_3$. Applying (P2) to $\widetilde{G}$ yields

$$\mathbb{P}\left(\widetilde{G}(U_1) \in] - C_3 h^s, C_3 h^s]\right) \le C_2(b_h, 1, h^s, C_3) h^s P_f^{\mathrm{FORM}}. \tag{5.11}$$

Since the right hand side of (5.10) is equal to the left hand side of (5.11), we conclude the proof of Proposition 5.7.

In (P4), we require the upper bound (5.8) of (P3). For convex failure domains, we prove that the Gaussian measure of the symmetric difference $A \triangle A_h$ behaves as the Gaussian measure of an interval in 1D. In particular, we derive the bound

$$\mathbb{P}(A \triangle A_h) \le C_4 \mathbb{P}(U_1 \in] - b_h - C_3 h^s, -b_h + C_3 h^s]), \tag{5.12}$$

where $U_1 \sim \mathrm{N}(0,1)$ and $b_h = \|u_h^{\mathrm{MLFP}}\|_2$. This bound follows from a geometrical consideration, which is visualized in Figure 7.
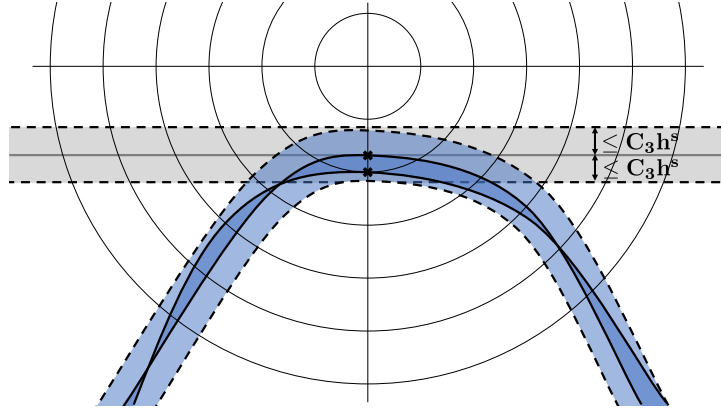

Figure 7: Translation of the difference of the failure domains to an interval in 1D.

In (P5), we define the affine linear LSF $\widetilde{G}_h(U_1) = U_1 + b_h$. We assume that $\widetilde{G}_h$ satisfies Assumption 5.3 with $\widetilde{C}_{\mathrm{FE}} = C_3$. Applying (P2) to $\widetilde{G}_h$ yields

$$\mathbb{P}\left(\widetilde{G}_h(U_1) \in] - C_3 h^s, C_3 h^s]\right) \le C_2(b_h, 1, h^s, C_3) h^s P_{f,h}^{\mathrm{FORM}}. \tag{5.13}$$

Finally, combining (5.12) and (5.13) concludes the proof of Theorem 5.8.

5.4. Numerical experiments

In numerical experiments, we investigate the convergence behaviour of the probability of failure for several LSFs. In particular, we consider the LSFs $G^{(4)}$, $G^{(5)}$, and $G^{(7)}$, which are defined in Chapter 2. Indeed, these LSFs do not satisfy all assumptions of Proposition 5.7 and Theorem 5.8, since the bound in Assumption 5.3 does not hold uniformly for $\mathbb{P}$-a.e. $\omega \in \Omega$. However, as noted in Remark 5.10, we expect a similar behaviour of the error bounds in this case.

The LSF $G^{(4)}$ given in (3.5) requires the solution of an ODE in 1D physical space, while the parametric space is one-dimensional. Failure occurs if the solution $y(t, \omega)$ is larger than $y_{\mathrm{max}} = 40$ at $t = 1$. The explicit Euler and the Crank–Nicolson

discretization schemes are applied to generate approximate LSFs. All quantities of interest can be calculated analytically and no sampling error is involved. We observe in Figure 5 of Article II that the distance $\|u^{\mathrm{MLFP}} - u_h^{\mathrm{MLFP}}\|_2$ behaves as $\mathcal{O}(h)$ for the explicit Euler and $\mathcal{O}(h^2)$ for the Crank-Nicolson scheme. Since the failure domains are half-rays, the distance of the MLFPs is equal to the distance of the failure domains and it holds that $P_f = P_f^{\mathrm{FORM}}$. Thus, we expect by Proposition 5.7 and Theorem 5.8 that the relative error of the probability of failure also behaves as $\mathcal{O}(h)$ and $\mathcal{O}(h^2)$, respectively. In Figure 5 of Article II, we observe that this conjecture indeed holds.

The LSF $G^{(5)}$ given in (3.6) requires the solution of an elliptic PDE in the unit interval $D = (0,1) \subset \mathbb{R}$. Failure occurs if the solution $y(x,\omega)$ is smaller than $y_{\min} = -1/3$ at $\widehat{x} = 1/3$. We note that the setting of $G^{(5)}$ is not as the setting described in Section 3.2.1, since the boundary condition is not of Dirichlet form. However, we expect that our provided error bounds also hold in this setting. Again, all quantities can be calculated analytically, and, in Section 5.2 of Article II, we show that the failure domains are convex. FEs are employed with either piecewise linear or quadratic shape functions. The linear shape functions lead to a discretization error of order $\mathcal{O}(h^2)$, while quadratic shape functions lead to $\mathcal{O}(h^3)$. In Figure 7 of Article II, we observe that the relative error of the FORM estimates scales as the FE discretization error. The same holds true for the error bound in Theorem 5.8.

The LSF $G^{(7)}$ given in (3.13) requires the solution of an elliptic PDE with log-normal diffusion coefficient. Failure occurs if the flow rare $q(x,\omega)$ is larger than $q_{\max,1} = 1.7$ at $\widehat{x} = 1$. In another setting, we consider $q_{\max,2} = 1.5$. We consider two different correlation lengths, $\lambda_1 = 0.3$ and $\lambda_2 = 0.1$. In the first case, the KL expansion is truncated after $n_1 = 10$ leading KL terms, while in the second case, the KL expansion is truncated after $n_2 = 50$ leading KL terms. For both cases, FEs are employed with piecewise linear shape functions. The probability of failure is estimated as the average of 100 SIS simulations with 10^4 samples per level. We apply sampling from the vMFN distribution model in the MCMC step as discussed in Section 4.3. In Figure 8 and Figure 9 of Article II, we observe for both correlation lengths that the relative error of the FORM estimates and the error bound in Theorem 5.8 behave as the expected convergence order $\mathcal{O}(h)$. We note that the LSF requires the evaluation of the derivative of the PDE solution. Thus, the PDE discretization error is of order $\mathcal{O}(h)$ [110, Section 1.6]. For $\lambda_2 = 0.1$, we observe a preasymptotic behaviour. For large h, the error bounds have a plateau behaviour since $P_{f,h}$ is much larger than P_f. For $h \approx 2 \cdot 10^{-2}$, we observe a very small error since $P_{f,h}$ is approximately P_f. For $h < 10^{-2}$, we observe the expected convergence order $\mathcal{O}(h)$.

6. The ensemble Kalman filter for rare event estimation

In this chapter, we discuss the ensemble Kalman filter for rare event estimation, which is the topic of Article III. Before we begin with the mathematical formulation, we give a brief overview on the development of the EnKF. The EnKF builds on the Kalman filter [62], which has been developed for data assimilation problems. The Kalman filter is exact if the dynamic and observation operator are linear and the distributions of the noise and initial state are Gaussian. The extended Kalman filter [71, Section 4.2.2] has been proposed for nonlinear dynamics and observations. Therein, the derivative of the dynamic is required. In the EnKF [39], the derivative is approximated via an *ensemble* of particles, which makes the EnKF applicable for a broad range of problems. Compared to other sampling-based methods, the main advantage of the EnKF is its simple implementation since the EnKF does not require an MCMC algorithm. In recent years, the EnKF has been developed for inverse problems [57, 106]. Thereby, the EnKF is applied to generate samples from the posterior distribution. However, the generated samples are not distributed according to the posterior distribution if the underlying model is nonlinear. According to [37], the samples are distributed according to a so-called *analysis variable*, which distribution is in general different to the posterior distribution. Due to this fact, we combine the EnKF with an Importance Sampling approach to estimate the probability of failure.

In more detail, we reformulate the rare event problem as an inverse problem and apply the EnKF algorithm of [106] to generate samples from the optimal IS density, i.e., failure samples. Since the generated samples are not distributed according to the optimal IS density, according to [37], we fit a distribution model with the generated samples and estimate the probability of failure by IS. This yields an unbiased estimator of the probability of failure if the support of the fitted distribution contains the failure domain.

If the LSF is affine linear, we prove, under certain assumptions, that the mean of the EnKF particles converges to a convex combination of the MLFP and the mean of the optimal IS density. Due to this property, we motivate the EnKF as an algorithm to generate samples which are in proximity of the MLFP and approximate the curvature of the limit-state surface.

6.1. The EnKF algorithm

We begin with the reformulation of the rare event problem as an inverse problem. We apply the *rectified linear unit* (ReLU) to the outcome of G and define the auxiliary LSF $\widetilde{G} : \mathbb{R}^n \to \mathbb{R}$ by

$$\widetilde{G}(u) := \max\{0, G(u)\}.$$

Thus, the optimal IS density is rewritten as

$$p_{\text{opt}}(u) = \frac{1}{P_f} I(\widetilde{G}(u) = 0)\varphi_n(u). \tag{6.1}$$

The goal is to apply the EnKF algorithm of [106] to generate samples from the optimal IS density, i.e., failure samples. We note that the algorithm of [106] is motivated by the SMC algorithm for Bayesian inverse problems [9, 31, 60]. To consider the rare event problem as an inverse problem, we define the artificial data as $y^{\dagger} = 0$ and we seek $u^{\dagger} \in \mathbb{R}^n$ such that

$$y^{\dagger} = \widetilde{G}(u^{\dagger}) + \eta, \tag{6.2}$$

where η is observational noise which is distributed according to $\mathrm{N}(0, 1)$. In this way, (6.2) has the form of a Bayesian inverse problem. Following the Bayesian perspective in [25, 112], we model (u, y) as a realisation from a jointly varying random variable (U, Y). In this Bayesian viewpoint, the optimal IS density p_{opt} is the posterior density, the indicator function represents the likelihood function, φ_n is the prior density, and P_f the evidence. This relationship is also observed in [115]. Under mild assumptions on the forward model, noise distribution and prior distribution, the inverse problem (6.2) is well-posed [69, 112].

In the EnKF, the optimal IS density is approximated in an adaptive sequential manner similar to SIS. We define the sequence of EnKF densities by

$$p_0^{\mathrm{EnKF}}(u) := \varphi_n(u),$$
$$p_k^{\mathrm{EnKF}}(u) := \frac{1}{P_k^{\mathrm{EnKF}}} \exp\left(-\frac{1}{2\sigma_k}\widetilde{G}(u)^2\right) \varphi_n(u), \quad \text{for } k = 1, \ldots, N_T,$$

where P_k^{EnKF} is a normalizing constant. The sequence $\infty = \sigma_0 > \sigma_1 > \cdots > \sigma_{N_T} > 0$ defines the sequence of temperatures. For $\sigma_k = 0$, the optimal IS density (6.1) is obtained. In Figure 1 of Article III, the approximation of the indicator function by the EnKF and SIS densities is shown. We observe that the EnKF densities are equal to the indicator function for $\{G \leq 0\}$, while the SIS densities are symmetric around $\{G = 0\}$. By the sequential definition of p_k^{EnKF}, it holds that

$$p_k^{\mathrm{EnKF}}(u) = \frac{P_{k-1}^{\mathrm{EnKF}}}{P_k^{\mathrm{EnKF}}} \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_{k-1}}\right)\widetilde{G}(u)^2\right) p_{k-1}^{\mathrm{EnKF}}(u).$$

Thus, we set $\Delta\sigma_k := 1/\sigma_k - 1/\sigma_{k-1}$ as the step size of the EnKF update. We note that in [106] a constant step size $\Delta\sigma_k = \Delta\sigma$ is employed. However, we employ an adaptive procedure similar to SIS to reduce the number of update steps. This approach is also applied in [58]. With the target coefficient of variation δ_{target}, we determine σ_k adaptively by

$$\sigma_k = \operatorname*{argmin}_{\sigma \in (0, \sigma_{k-1})} \left(\delta_{w_k} - \delta_{\mathrm{target}}\right)^2, \quad \text{where } \delta_{w_k} := \frac{\mathrm{StD}_{p_{k-1}^{\mathrm{EnKF}}}[w_k(U)]}{\mathbb{E}_{p_{k-1}^{\mathrm{EnKF}}}[w_k(U)]}$$

and the weights are given by

$$w_1(u) = \exp\left(-\frac{1}{2\sigma_1}\widetilde{G}(u)^2\right),$$
$$w_k(u) = \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_{k-1}}\right)\widetilde{G}(u)^2\right), \quad \text{for } k > 1.$$

Each density p_k^{EnKF} is approximated by an ensemble of equally weighted particles $\mathbf{u_k} = \{u_k^{(j)}\}_{j=1}^J$. In one EnKF update step, the ensemble $\mathbf{u_{k-1}}$ is updated to the ensemble $\mathbf{u_k}$ via

$$u_k^{(j)} = u_{k-1}^{(j)} + C_{\text{up}}(\mathbf{u_{k-1}}) \left(C_{\text{pp}}(\mathbf{u_{k-1}}) + \frac{1}{\Delta\sigma_k} \right)^{-1} \left(y_k^{(j)} - \widetilde{G}(u_{k-1}^{(j)}) \right),$$

for $j = 1, \ldots, J$, where the data $y^\dagger$ is perturbed by additive Gaussian noise $\xi_k^{(j)} \sim N(0, \Delta\sigma_k^{-1})$, i.e.,

$$y_k^{(j)} = y^\dagger + \xi_k^{(j)}. \tag{6.3}$$

The empirical cross-covariances $C_{\text{pp}} \in \mathbb{R}$ and $C_{\text{up}} \in \mathbb{R}^n$ are given by

$$C_{\text{pp}}(\mathbf{u_{k-1}}) := \frac{1}{J} \sum_{j=1}^J \left(\widetilde{G}(u_{k-1}^{(j)}) - \overline{G} \right) \left(\widetilde{G}(u_{k-1}^{(j)}) - \overline{G} \right),$$

$$C_{\text{up}}(\mathbf{u_{k-1}}) := \frac{1}{J} \sum_{j=1}^J \left(\widetilde{G}(u_{k-1}^{(j)}) - \overline{G} \right) \left(u_{k-1}^{(j)} - \overline{u} \right),$$

where the quantities $\overline{u} \in \mathbb{R}^n$ and $\overline{G} \in \mathbb{R}$ are the empirical means

$$\overline{u} := \frac{1}{J} \sum_{j=1}^J u_{k-1}^{(j)}, \qquad \overline{G} := \frac{1}{J} \sum_{j=1}^J \widetilde{G}(u_{k-1}^{(j)}).$$

The EnKF stops if the coefficient of variation $\delta_{w_{\text{opt},k}}$ is smaller than δ_{target}. At the kth iteration, the weights $w_{\text{opt},k}$ are defined by

$$w_{\text{opt},k}(u) := I(\widetilde{G}(u) = 0) \exp\left(\frac{1}{\sigma_k} \widetilde{G}(u)^2 \right).$$

We estimate $\delta_{w_{\text{opt},k}}$ after each EnKF update with the ensemble $\mathbf{u_k}$. If $\delta_{w_{\text{opt},k}} \leq \delta_{\text{target}}$, the iteration stops and we set $N_T := k$ as the number of EnKF updates.

Since $\widetilde{G}$ is always nonlinear, the EnKF does not generate samples which are distributed according to the EnKF densities p_k^{EnKF}. By [37], the samples are distributed according to the so-called *analysis variable*, whose distribution differs from the one of the EnKF densities. Thus, we cannot use the generated samples to estimate the fractions $P_k^{\text{EnKF}}/P_{k-1}^{\text{EnKF}}$ as in SIS. However, we motivate the EnKF as an algorithm to generate samples which are in proximity of the MLFP and approximate the curvature of the limit-state surface. In Theorem 6.3, we prove that this property holds if the LSF G is affine linear. With the generated samples $\mathbf{u_{N_T}}$, we fit a distribution model and generate J samples $\{\widehat{u}^{(j)}\}_{j=1}^J$, which are distributed according to the fitted density $\widehat{p}$. Finally, we apply IS to estimate P_f by

$$\widehat{P}_f^{\text{EnKF}} := \frac{1}{J} \sum_{j=1}^J I(G(\widehat{u}^{(j)}) \leq 0) \frac{\varphi_n(\widehat{u}^{(j)})}{\widehat{p}(\widehat{u}^{(j)})}. \tag{6.4}$$

We note that this procedure is similar to [68, Algorithm 3.1]. By the properties of IS, the EnKF estimator is unbiased if the support of $\widehat{p}$ contains the failure domain A. In Article III, we consider the Gaussian mixture model [42, Section 1] and the vMFN distribution model [93] for $\widehat{p}$. The EnKF algorithm for rare event estimation is shown in Algorithm 1 of Article III. In the numerical experiments of Article III, we observe that the EnKF estimator is skewed. This is due to the fact that the likelihood weights in (6.4) might be large in infrequent simulation runs due to the choice of the parametric model of $\widehat{p}$, which leads to infrequent outliers. This observation in also discussed in [90, Section 9.3].

6.2. Theoretical properties

In the following, we prove theoretical properties of the EnKF if the LSF G is affine linear. Our starting points are the investigated properties of the EnKF for inverse problems. The authors of [57] show that the EnKF particles stay within the subspace spanned by the initial ensemble. This property is also known as the *subspace property* and also holds true for the EnKF for rare event estimation since we apply the EnKF to the inverse problem (6.2). The continuous-time limit $\Delta\sigma_k \to 0$ is studied in [11, 106, 107], which yields a coupled system of stochastic differential equations. Based on these results, we derive the continuous-time limit of the EnKF for rare event estimation in Theorem 6.2. In [17, 43, 55], the mean and covariance of the EnKF particles are studied for an infinite ensemble in the linear Gaussian setting. In Theorem 6.3, we prove that the mean of the EnKF particles converges to a convex combination of the MLFP and the mean of the optimal IS density for $\sigma_k \to 0$.

For the remainder of this chapter, we assume that the EnKF is applied without noise, i.e., $\xi_k^{(j)} = 0$ in (6.3), for $j = 1, \ldots, J$ and $k = 1, \ldots, N_T$. Moreover, we add the limit-state surface $\{G = 0\}$ to the safe domain. Thus, failure is defined as the event that $G < 0$. This does not influence the probability of failure since the limit-state surface is a set with Lebesgue measure equal to zero. For the noise free case, it is useful to consider safe and failure particles separately. This is due to the fact that failure particles do not move and remain failure particles. Moreover, safe particles remain safe particles for all iteration as we see in Theorem 6.2. As another assumption, we require that the LSF is affine linear and the initial particles are independently standard Gaussian distributed. All assumptions are stated in the following assumption, which is also given in Article III.

Assumption 6.1 (Assumption 4.3 of Article III). *We assume that*

(i) $G(u) = a^T u + b$ with $a = (1, 0, \ldots, 0)^T \in \mathbb{R}^n$ and $b > 0$,

(ii) the initial particles $\{u_0^{(j)}\}_{j=1}^J$ are independently standard Gaussian distributed,

(iii) the EnKF is applied without noise, i.e., $y_k^{(j)} = y^\dagger$.

We note that the assumption $a = (1, 0, \ldots, 0)^T \in \mathbb{R}^n$ holds without loss of generality by the rotation invariance of the Gaussian measure.

In Theorem 6.2, we consider the continuous-time limit $\Delta\sigma_k \to 0$ of the EnKF update. We show that this limit yields a coupled system of SDEs. This result is

based on the derived continuous-time limit in [106]. We note that the SDE is written in terms of the derivative with respect to the time t. This is the usual notation for stating SDEs. The time t and the temperature σ are related via $t = 1/\sigma$. Thus, the limit $\sigma \to 0$ refers to the infinite time limit $t \to \infty$.

Theorem 6.2 (Theorem 4.1 of Article III). *Let Assumption 6.1 hold. Denote by $S = \{k \in [J] : G(u^{(k)}) \geq 0\}$ and $F = \{k \in [J] : G(u^{(k)}) < 0\}$ the index sets of the safe and failure particles. Then, the safe particles satisfy the flow*

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C_{\mathrm{uu}}(\mathbf{u})D_u\left(\frac{1}{2}G(u^{(j)})^2\right) + \frac{G(u^{(j)})}{J}\sum_{k\in F}G(u^{(k)})(u^{(k)} - \bar{u}), \qquad (6.5)$$

while failure particles do not move. The matrix $C_{\mathrm{uu}} \in \mathbb{R}^{n \times n}$ denotes the empirical covariance matrix of the ensemble $\{u^{(j)}\}_{j=1}^{J}$.

From the first summand in (6.5), we observe that the particles move in the direction of the negative gradient of G. Thus, this term pulls the particles to the limit-state surface. From the second summand in (6.5), we observe that the direction $G(u^{(k)})(u^{(k)} - \bar{u})$ points away from the limit-state surface if the mean $\bar{u}$ is in the safe domain since $G(u^{(k)}) < 0$ for $k \in F$. Indeed, in the proof of Theorem 6.3, we observe that the safe particles will always move to the limit-state surface, i.e., the first summand in (6.5) dominates the flow. If a particle reaches the limit-state surface, it remains there since $G(u^{(j)}) = 0$ and, thus, (6.5) is equal to zero.

Proof outline of Theorem 6.2. By [106], the particle flow satisfies the SDE

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = \frac{1}{J}\sum_{k=1}^{J}\langle\widetilde{G}(u^{(k)}) - \overline{G}, -\widetilde{G}(u^{(j)})\rangle(u^{(k)} - \bar{u}), \qquad (6.6)$$

where we have used that the EnKF is applied without noise. By splitting the sum in (6.6) into the safe and failure particles and rearranging the terms yields the SDE (6.5) for safe particles. For failure particles, it holds that $\widetilde{G}(u^{(j)}) = 0$. Thus, we obtain from (6.6) that these particles do not move. $\square$

The next theorem states that for $t \to \infty$ and $J \to \infty$, the ensemble mean converges to a convex combination of the MLFP and the mean of the optimal IS density. The mean of the optimal IS density is denoted by $u^{\mathrm{opt}} \in \mathbb{R}^n$. We consider an infinite ensemble $\{u^{(j)}\}_{j=1}^{\infty}$ and we denote by $U(t)$ the random variable which is distributed according to the particle density of this infinite ensemble. We denote the mean of the particles as $m(t) := \mathbb{E}[U(t)]$ and the covariance as $C(t) := \mathrm{Cov}[U(t)]$. Moreover, we split the infinite ensemble into safe and failure particles. We denote by $U_S(t)$ the random variable which is distributed according to the particle density of the safe particles $\{u^{(j)} : G(u^{(j)}) \geq 0\}$, while $U_F(t)$ denotes the random variable which is distributed according to the particle density of the failure particles $\{u^{(j)} : G(u^{(j)}) < 0\}$. Since failure particles remain failure particles and safe particles remain safe particles, it holds that $\mathbb{P}(U(t) = U_S(t)) = (1 - P_f)$ and $\mathbb{P}(U(t) = U_F(t)) = P_f$.

Theorem 6.3 (Theorem 4.7 of Article III). *Let Assumption 6.1 hold. For the large particle limit $J \to \infty$, the ensemble mean satisfies*

$$\lim_{t \to \infty} m(t) = (1 - P_f)u^{\mathrm{MLFP}} + P_f u^{\mathrm{opt}}.$$

From Theorem 6.3, we observe that the mean of the particles is in proximity of the MLFP and the safe particles are on the limit-state surface. For nonlinear LSFs, we expect a similar behaviour. We expect that the mean of the particles is in proximity of the MLFP and the safe particles move to the limit-state surface and remain there. Moreover, we expect that the covariance of the safe particles stays constant in direction parallel to the limit-state surface and decreases in direction perpendicular to the limit-state surface. Thus, the safe particles approximate the curvature and the shape of the limit-state surface.

Proof outline of Theorem 6.3. By splitting the ensemble in safe and failure particles, the mean $m(t)$ can be written as

$$m(t) = (1 - P_f)m_S(t) + P_f m_F(t).$$

Thus, it is sufficient to prove that $\lim_{t \to \infty} m_F(t) = u^{\mathrm{opt}}$ and $\lim_{t \to \infty} m_S(t) = u^{\mathrm{MLFP}}$. Since failure particles do not move, it holds that $m_F(t) := \mathbb{E}[U_F(t)] = u^{\mathrm{opt}}$ for all $t \geq 0$, which concludes the first part.

In the next step, we show in Lemma 4.8 of Article III that the mean of the safe particles $m_S(t) := \mathbb{E}[U_S(t)]$ satisfies

$$\frac{\mathrm{d}m_S(t)}{\mathrm{d}t} = -G(m_S(t))\left(C(t) - P_f\left(1 - m(t)(u^{\mathrm{opt}} + b)\right)\right), \tag{6.7}$$

where we use the fact that safe particles move only in direction perpendicular to the limit-state surface by the linearity of G. Thus, we can consider the dynamic for $n = 1$. To conclude the proof, we show with Lemma 4.10 of Article III that

$$C(t) - P_f\left(1 - m(t)(u^{\mathrm{opt}} + b)\right) > 0, \quad \text{for all } t \geq 0.$$

This implies that the right hand side of (6.7) is always negative as long as $G(m_S(t)) > 0$. Thus, the mean of the safe particles $m_S(t)$ moves to the MLFP $u^{\mathrm{MLFP}} = -b$, i.e., $\lim_{t \to \infty} m_S(t) = u^{\mathrm{MLFP}}$, which concludes the proof. $\qquad \square$

6.3. The EnKF for multi-modal failure domains

In the previous section, we have seen that the EnKF particles move to a single mean value. Thus, the EnKF algorithm of Section 6.1 is not useful to generate samples of separated failure modes. In [33, 77, 109], the authors propose to fit a GM distribution in each EnKF update step and to update the particles belonging to each mixture term separately. However, we propose the approach of [102] as a localization technique to determine localized covariance matrices. Therein, C_{pp} and C_{up} are localized for each particle of the ensemble. This is achieved by the weight

matrix $W \in \mathbb{R}^{J \times J}$, which represents the distances of the particles. The entries of the weight matrix are given by

$$W_{i,j} = \exp\left(-\frac{1}{2\zeta}\|u_{k-1}^{(i)} - u_{k-1}^{(j)}\|_2^2\right), \quad \text{for } i, j = 1, \ldots, J, \tag{6.8}$$

where $\zeta > 0$ is a parameter chosen by the user. From (6.8), we observe that the entries of W can be interpreted as the value of the PDF $\varphi_n(u_{k-1}^{(i)} \mid u_{k-1}^{(j)}, \zeta \cdot \mathrm{Id}_n)$, which is the Gaussian density with mean vector $u_{k-1}^{(j)}$ and covariance matrix $\zeta \cdot \mathrm{Id}_n$. With this insight, we propose a novel algorithm, where the covariance matrix is chosen adaptively. We propose to split the ensemble u_{k-1} into M clusters and we determine for each cluster the empirical covariance matrix $C_m \in \mathbb{R}^{n \times n}$ for $m = 1, \ldots, M$. For the particle $u_{k-1}^{(j)}$ belonging to the cluster m, we determine the jth column of the weight matrix by

$$W_{i,j} = \exp\left(-\frac{1}{2}\|C_m^{-1/2}(u_{k-1}^{(i)} - u_{k-1}^{(j)})\|_2^2\right), \quad \text{for } i = 1, \ldots, J.$$

After calculating all entries of W, it is normalized such that each column sums up to one. The normalized weight matrix is denoted by $\overline{W}$. The localized means are

$$\overline{u}_{\mathrm{loc}}^{(j)} = \sum_{i=1}^{J} \overline{W}_{i,j} u_{k-1}^{(i)}, \qquad \overline{G}_{\mathrm{loc}}^{(j)} = \sum_{i=1}^{J} \overline{W}_{i,j} \widetilde{G}(u_{k-1}^{(i)}).$$

With these localized means, we determine the localized cross-covariances by

$$C_{\mathrm{loc,pp}}(u_{k-1}^{(j)}) = \sum_{i=1}^{J} \overline{W}_{i,j}\left(\widetilde{G}(u_{k-1}^{(i)}) - \overline{G}_{\mathrm{loc}}^{(j)}\right)\left(\widetilde{G}(u_{k-1}^{(i)}) - \overline{G}_{\mathrm{loc}}^{(j)}\right),$$

$$C_{\mathrm{loc,up}}(u_{k-1}^{(j)}) = \sum_{i=1}^{J} \overline{W}_{i,j}\left(\widetilde{G}(u_{k-1}^{(i)}) - \overline{G}_{\mathrm{loc}}^{(j)}\right)\left(u_{k-1}^{(i)} - \overline{u}_{\mathrm{loc}}^{(j)}\right).$$

With the localized covariances, one update of the EnKF algorithm for multi-modal failure domains is given by

$$u_k^{(j)} = u_{k-1}^{(j)} + C_{\mathrm{loc,up}}(u_{k-1}^{(j)})\left(C_{\mathrm{loc,pp}}(u_{k-1}^{(j)}) + \frac{1}{\Delta\sigma_k}\right)^{-1}\left(y_k^{(j)} - \widetilde{G}(u_{k-1}^{(j)})\right). \tag{6.9}$$

We note that for equal weights $\overline{W}_{i,j} = 1/J$ for $i, j = 1, \ldots, J$, the usual EnKF algorithm of Section 6.1 is obtained. Since each particle moves according to localized mean and localized covariances, which are based on their nearest neighbourhood, the EnKF update (6.9) enables that particles move to separate failure modes.

6.4. Numerical experiments

In numerical experiments, we compare the performance of the EnKF with SIS in terms of the required computational costs and the achieved relative root mean square error, which is defined in (4.11). Since all LSFs are considered on a single

discretization level, one LSF evaluation represents one unit of computational cost. Both methods are applied with a varying number of samples J and with different target coefficients of variations δ_{target}. The EnKF is always employed with noise, which yields a more robust algorithm since failure particles and particles on the limit-state surface are moving.

We consider four LSFs which are defined in Chapter 3. The first three LSFs are $G^{(1)}$, $G^{(2)}$, and $G^{(3)}$, which are analytically given and are defined for a two-dimensional parameter space. Therefore, we can visualize how the particles are evolving during the iterations. The fourth example is the LSF $G^{(6)}$ (3.12), which requires the solution of an elliptic diffusion equation in 1D space with log-normal diffusion coefficient. We remind that we define failure as the event that the solution $y(x, \omega)$ is larger than $y_{\max} = 0.535$ at $\widehat{x} = 1$. We set $n = 150$ as the truncation order of the KL expansion.

For the first LSF $G^{(1)}$ (3.1), we observe in Figure 2 of Article III that the EnKF particles are more spread along the limit-state surface than the SIS particles. We expect this behaviour of the EnKF particles by the proof of Theorem 6.3, where we show that the variance of the particles in direction parallel to the limit-state surface does not change. In Figure 3 of Article III, we compare the performance of the methods. We observe that the EnKF requires less computational costs than SIS for a fixed level of accuracy. Indeed, the achieved relRMSE by the EnKF does not vary considerably if δ_{target} increases. For SIS, we observe that the error increases as δ_{target} increases.

The second LSF $G^{(2)}$ (3.2) leads to a failure domain with two distinct failure modes. Thus, we apply the EnKF with the multi-modal strategy of Section 6.3 with $\zeta = 2$ in (6.8). In Figure 5 of Article III, we see that the EnKF requires less computational costs than SIS for a fixed level of accuracy. However, a larger target coefficient of variation yields a larger error for the EnKF.

The third experiment investigates the LSF $G^{(3)}$ (3.3). The respective failure domain consists of four distinct failure modes. The EnKF is applied for different values for ζ. In Figure 6 of Article III, we observe that for small ζ, the ensemble splits more clearly into the four failure modes. However, for very small ζ, the ensemble is too concentrated. Therefore, we choose $\zeta = 0.25$ and investigate the performance more particularly. In Figure 7 of Article III, we observe that SIS requires less computational costs than the EnKF for a fixed level of accuracy. As another multi-modal strategy, we consider the adaptive strategy discussed in Section 6.3. In Figure 8 of Article III, we observe that for $\delta_{\text{target}} = 5.00$, the EnKF requires less computational costs than SIS for a fixed level of accuracy.

In the fourth experiment, the EnKF is applied to $G^{(6)}$ (3.12), which depends on a high-dimensional parameter space. For this LSF, we expect a unimodal failure domain. In Figure 9 of Article III, we observe that the performance of the EnKF is similar as in the first experiment. The reached relRMSE stays nearly constant while varying δ_{target}. Again, the EnKF requires less computational costs than SIS for a fixed level of accuracy.

We conclude that the EnKF yields a better performance as SIS for unimodal failure domains. For multi-modal failure domains, the multi-modal strategy of the EnKF requires a careful choice of its algorithmic parameters.

7. Conclusion and outlook

In this chapter, we briefly summarize the contributions of this dissertation and we give an outlook of possible directions for future research.

We have introduced the problem setting of rare event estimation and we have discussed several sampling based methods, which have been developed to estimate the probability of rare events. Among them are Monte Carlo Sampling, Importance Sampling, Sequential Importance Sampling, and Subset Simulation. Furthermore, we have introduced Multilevel Monte Carlo and Multilevel Subset Simulation, which enable usage of coarse grid limit-state function evaluations. These methods are applicable for LSFs which require the solution of a partial differential equation. In addition, we have considered the first order reliability method as a deterministic approximation method.

Based on SIS, we have developed Multilevel Sequential Importance Sampling as a novel multilevel algorithm for computing rare event probabilities. MLSIS combines the tempering approach of SIS with the multilevel approach of the Multilevel Sequential[2] Monte Carlo algorithm given in [70]. The MLSIS estimator overcomes the nestedness issue of MLSuS since the support of the employed densities is the whole parametric domain. In numerical experiments, we have observed that MLSIS requires less computational costs than SIS for a fixed level of accuracy and the performance of MLSIS is similar to the performance of MLSuS. Moreover, sampling from the von Mises–Fisher–Nakagami distribution model increases the accuracy for a fixed computational budget compared to adaptive conditional sampling.

We note that the bridging approach of the MLSIS estimator yields a flexible way of combining general LSF approximations. Thus, a future research direction could be the implementation of the MLSIS estimator for general multifidelity approximations.

In addition to the MLSIS estimator, we have implemented the ensemble Kalman filter as a novel single level variance reduction technique for rare event estimation. Thereby, we formulate the rare event problem as an inverse problem and apply the EnKF algorithm of [57, 106] to generate failure samples. After the EnKF iteration is finished, we fit a distribution model with the generated samples and apply IS to estimate the probability of failure. In addition to describing algorithmic properties, we derive theoretical properties of the EnKF tailored to rare event estimation. For affine linear LSFs, we derive the continuous-time limit of the EnKF update if the EnKF is applied without noise. If in addition an infinite ensemble size is considered, we show that the mean of the EnKF particles converges to a convex combination of the most likely failure point and the mean of the optimal IS density. In numerical experiments, we have observed that the EnKF requires less computational costs than SIS for a fixed level of accuracy if the failure domain is unimodal. For multi-modal failure domains, the EnKF requires a careful choice of its algorithmic parameters to achieve a similar performance as SIS.

This dissertation can be used as a starting point to derive further theoretical properties of the EnKF tailored to rare event estimation. Indeed, the theoretical investigation for nonlinear LSFs is missing. Moreover, we have only considered the case that the EnKF is applied without noise. As the EnKF is more robust in numerical experiments if it is applied with noise, the theoretical properties of the

noisy EnKF would be of higher importance. Motivated by the MLSIS algorithm, another research direction is the combination of the EnKF with a multilevel approach to enable usage of coarse grid LSF evaluations in PDE settings. Another direction of future research is the investigation of different reformulations of the rare event problem. In the present dissertation, we have applied the rectified linear unit to the original LSF to obtain an inverse problem formulation. Besides the ReLU function, one could consider alternative functions and study theoretical properties with respect to these alternative formulations.

Besides MLSIS and the EnKF, another part of this dissertation is the analysis of the approximation error of the probability of failure. If the LSF requires the solution of a PDE which relies on a uniformly elliptic and bounded diffusion coefficient, we derive, under further assumptions, that the absolute error of the probability of failure behaves as the PDE discretization error multiplied by the FORM estimate. One of these further assumptions is the convexity of the failure domains. For general failure domains, we show that the relative error of the FORM estimate behaves as the PDE discretization error. In numerical experiments, we have observed that our derived error bounds hold in settings, where the diffusion coefficient is only pathwise elliptic and bounded.

This dissertation opens several possible ways about future research directions concerning the approximation error of the probability of failure. One of these is the derivation of a proof that our derived error bounds hold if the diffusion coefficient is only pathwise elliptic and bounded. Moreover, in the given error analysis, we have considered the case that the dimension of the input random variable is fixed and finite. However, if the stochastic input is a random field, there is another contribution to the approximation error which stems from truncating the Karhunen–Loève expansion of the input random field. Thus, the investigation of the approximation error which is induced by truncating the KL expansion is another research direction. In this case, we expect that the error bound of the probability of failure behaves as the truncation error of the KL expansion. Finally, the derived error bounds could be used to construct efficient multilevel methods as in [36], which would combine the approaches of Article I and III.

List of abbreviations

aCS adaptive conditional sampling

a.e. almost every

CDF cumulative distribution function

EnKF ensemble Kalman filter

FE Finite Element

FORM first order reliability method

IS Importance Sampling

KL Karhunen–Loève

LSF limit-state function

MCMC Markov chain Monte Carlo

MCS Monte Carlo Sampling

MLFP most likely failure point

MLMC Multilevel Monte Carlo

MLSIS Multilevel Sequential Importance Sampling

MLS^{2}MC Multilevel Sequential^2 Monte Carlo

MLSuS Multilevel Subset Simulation

ODE ordinary differential equation

PDE partial differential equation

PDF probability density function

relRMSE relative root mean square error

ReLU rectified linear unit

SDE stochastic differential equation

SIS Sequential Importance Sampling

SMC Sequential Monte Carlo

SORM second order reliability method

StD standard deviation

SuS Subset Simulation

vMFN von Mises–Fisher–Nakagami

Bibliography

[1] S. Agapiou, O. Papaspiliopoulos, D. Sanz–Alonso, and A. M. Stuart. "Importance sampling: intrinsic dimension and computational cost". In: *Statistical Science* 32.3 (2017), pp. 405–431. DOI: 10.1214/17-STS611.

[2] A. Agarwal, S. de Marco, E. Gobet, and G. Liu. "Rare event simulation related to financial risks: efficient estimation and sensitivity analysis". In: *HAL* (2017). Working paper. URL: https://hal-polytechnique.archives-ouvertes.fr/hal-01219616.

[3] M. de Angelis, E. Patelli, and M. Beer. "Advanced line sampling for efficient robust reliability analysis". In: *Structural Safety* 52.Part B (2015), pp. 170–182. DOI: 10.1016/j.strusafe.2014.10.002.

[4] S.-K. Au and J. L. Beck. "Estimation of small failure probabilities in high dimensions by subset simulation". In: *Probabilistic Engineering Mechanics* 16.4 (2001), pp. 263–277. DOI: 10.1016/S0266-8920(01)00019-4.

[5] S.-K. Au and J. L. Beck. "Important sampling in high dimensions". In: *Structural Safety* 25.2 (2003), pp. 139–163. DOI: 10.1016/S0167-4730(02)00047-4.

[6] S.-K. Au and Y. Wang. *Engineering Risk Assessment with Subset Simulation.* John Wiley & Sons, Ltd, 2014. DOI: 10.1002/9781118398050.

[7] J. Bear. *Dynamics of fluids in porous media.* Courier Corporation, 2013.

[8] J. Bect, L. Li, and E. Vazquez. "Bayesian subset simulation". In: *SIAM/ASA Journal on Uncertainty Quantification* 5.1 (2017), pp. 762–786. DOI: 10.1137/16M1078276.

[9] A. Beskos, A. Jasra, N. Kantas, and A. Thiery. "On the convergence of adaptive sequential Monte Carlo methods". In: *The Annals of Applied Probability* 26.2 (2016), pp. 1111–1146. DOI: 10.1214/15-AAP1113.

[10] A. Beskos, A. Jasra, K. Law, R. Tempone, and Y. Zhou. "Multilevel sequential Monte Carlo samplers". In: *Stochastic Processes and their Applications* 127.5 (2017), pp. 1417–1440. DOI: 10.1016/j.spa.2016.08.004.

[11] D. Blömker, C. Schillings, P. Wacker, and S. Weissmann. "Well posedness and convergence analysis of the ensemble Kalman inversion". In: *Inverse Problems* 35.8 (2019), p. 085007. DOI: 10.1088/1361-6420/ab149c.

[12] Z. Botev and D. Kroese. "Efficient Monte Carlo simulation via the generalized splitting method". In: *Statistics and Computing* 22.1 (2012), pp. 1–16. DOI: 10.1007/s11222-010-9201-4.

[13] N. Bouhlel and A. Dziri. "Maximum likelihood parameter estimation of Nakagami–Gamma shadowed fading channels". In: *IEEE Communications Letters* 19.4 (2015), pp. 685–688. DOI: 10.1109/LCOMM.2015.2394293.

[14] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* 3rd ed. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511618635.

[15]   S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. 3rd ed. Texts in Applied Mathematics. Springer, New York, NY, 2008. DOI: 10.1007/978-0-387-75934-0.

[16]   C. G. Bucher and U. Bourgund. "A fast and efficient response surface approach for structural reliability problems". In: *Structural Safety* 7.1 (1990), pp. 57–66. DOI: 10.1016/0167-4730(90)90012-E.

[17]   L. Bungert and P. Wacker. "Complete dynamics and spectral decomposition of the ensemble Kalman inversion". In: *arXiv* (2021). URL: https://arxiv.org/abs/2104.13281.

[18]   F. Cérou, P. Del Moral, T. Furon, and A. Guyader. "Sequential Monte Carlo for rare event estimation". In: *Statistics and Computing* 22.3 (2012), pp. 795–808. DOI: 10.1007/s11222-011-9231-6.

[19]   J. Charrier, R. Scheichl, and A. L. Teckentrup. "Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods". In: *SIAM Journal on Numerical Analysis* 51.1 (2013), pp. 322–352. DOI: 10.1137/110853054.

[20]   N. Chopin. "A sequential particle filter method for static models". In: *Biometrika* 89.3 (2002), pp. 539–551. DOI: 10.1093/biomet/89.3.539.

[21]   P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics, 2002. DOI: 10.1137/1.9780898719208.

[22]   L. Cizelj, B. Mavko, and H. Riesch–Oppermann. "Application of first and second order reliability methods in the safety assessment of cracked steam generator tubing". In: *Nuclear Engineering and Design* 147.3 (1994), pp. 359–368. DOI: 10.1016/0029-5493(94)90218-6.

[23]   F. J. Cornaton, Y.-J. Park, S. D. Normani, E. A. Sudicky, and J. F. Sykes. "Use of groundwater lifetime expectancy for the performance assessment of a deep geologic waste repository: 1. theory, illustrations, and implications". In: *Water Resources Research* 44.4 (2008). DOI: 10.1029/2007WR006208.

[24]   S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. "MCMC methods for functions: modifying old algorithms to make them faster". In: *Statistical Science* 28.3 (2013), pp. 424–446. DOI: 10.1214/13-STS421.

[25]   M. Dashti and A. M. Stuart. "The Bayesian approach to inverse problems". In: *Handbook of Uncertainty Quantification*. Ed. by R. Ghanem, D. Higdon, and H. Owhadi. Springer, Cham, 2017, pp. 311–428. DOI: 10.1007/978-3-319-12385-1_7.

[26]   P. Del Moral, A. Doucet, and A. Jasra. "Sequential Monte Carlo samplers". In: *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 68.3 (2006), pp. 411–436. DOI: 10.1111/j.1467-9868.2006.00553.x.

[27]   P. Del Moral, A. Jasra, K. Law, and Y. Zhou. "Multilevel sequential Monte Carlo samplers for normalizing constants". In: *ACM Transactions on Modeling and Computer Simulation* 27.3 (2017), 20:1–20:22. DOI: 10.1145/3092841.

[28] A. Der Kiureghian. "First- and second-order reliability methods". In: *Engineering Design Reliability Handbook*. Ed. by E. Nikolaidis, D. M. Ghiocel, and S. Singhal. 1st ed. CRC Press, 2004. Chap. 14. DOI: 10.1201/9780203483930.

[29] A. Der Kiureghian and T. Dakessian. "Multiple design points in first and second-order reliability". In: *Structural Safety* 20.1 (1998), pp. 37–49. DOI: 10.1016/S0167-4730(97)00026-X.

[30] A. Der Kiureghian and P.-L. Liu. "Structural reliability under incomplete probability information". In: *Journal of Engineering Mechanics* 112.1 (1986), pp. 85–104. DOI: 10.1061/(ASCE)0733-9399(1986)112:1(85).

[31] A. Doucet and A. M. Johansen. "A tutorial on particle filtering and smoothing: fifteen years later". In: *The Oxford Handbook of Nonlinear Filtering*. Ed. by D. Crisan and B. Rozovskii. Oxford University Press, Oxford, 2011, pp. 656–704.

[32] J. Douglas, T. Dupont, and L. Wahlbin. "Optimal L_∞ error estimates for Galerkin approximations to solutions of two-point boundary value problems". In: *Mathematics of Computation* 29 (1975), pp. 475–483. DOI: 10.1090/S0025-5718-1975-0371077-0.

[33] L. Dovera and E. Della Rossa. "Multimodal ensemble Kalman filtering using Gaussian mixture models". In: *Computational Geosciences* 15 (2011), pp. 307–323. DOI: 10.1007/s10596-010-9205-3.

[34] V. Dubourg, B. Sudret, and F. Deheeger. "Metamodel-based importance sampling for structural reliability analysis". In: *Probabilistic Engineering Mechanics* 33 (2013), pp. 47–57. DOI: 10.1016/j.probengmech.2013.02.002.

[35] B. Echard, N. Gayton, and M. Lemaire. "AK-MCS: an active learning reliability method combining kriging and Monte Carlo simulation". In: *Structural Safety* 33.2 (2011), pp. 145–154. DOI: 10.1016/j.strusafe.2011.01.002.

[36] D. Elfverson, F. Hellman, and A. Målqvist. "A multilevel Monte Carlo method for computing failure probabilities". In: *SIAM/ASA Journal on Uncertainty Quantification* 4.1 (2016), pp. 312–330. DOI: 10.1137/140984294.

[37] O. G. Ernst, B. Sprungk, and H.-J. Starkloff. "Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 823–851. DOI: 10.1137/140981319.

[38] L. C. Evans. *Partial Differential Equations*. 2nd ed. Graduate Studies in Mathematics, Volume 19. American Mathematical Society, 2010.

[39] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. 2nd ed. Springer, Berlin, Heidelberg, 2006. DOI: 10.1007/978-3-642-03711-5.

[40] R. Eymard, T. Gallouët, and R. Herbin. "Finite volume methods". In: *Solution of Equation in $\mathbb{R}^n$ (Part 3), Techniques of Scientific Computing (Part 3)*. Vol. 7. Handbook of Numerical Analysis. Elsevier, 2000, pp. 713–1018. DOI: 10.1016/S1570-8659(00)07005-8.

[41] G. S. Fishman. *Monte Carlo: Concepts, Algorithms and Applications*. 1st ed. Springer Series in Operations Research. Springer, New York, NY, 1996. DOI: `10.1007/978-1-4757-2553-7`.

[42] S. Frühwirth–Schnatter. *Finite Mixture and Markov Switching Models*. 1st ed. Springer Series in Statistics. Springer, New York, NY, 2006. DOI: `10.1007/978-0-387-35768-3`.

[43] A. Garbuno–Inigo, F. Hoffmann, W. Li, and A. M. Stuart. "Interacting Langevin diffusions: gradient structure and ensemble Kalman sampler". In: *SIAM Journal on Applied Dynamical Systems* 19.1 (2020), pp. 412–441. DOI: `10.1137/19M1251655`.

[44] M. Gerber, N. Chopin, and N. Whiteley. "Negative association, ordering and convergence of resampling methods". In: *The Annals of Statistics* 47.4 (2019), pp. 2236–2260. DOI: `10.1214/18-AOS1746`.

[45] S. Geyer, I. Papaioannou, and D. Straub. "Cross entropy-based importance sampling using Gaussian densities revisited". In: *Structural Safety* 76 (2019), pp. 15–27. DOI: `10.1016/j.strusafe.2018.07.001`.

[46] R. Ghanem and P. Spanos. *Stochastic Finite Elements: A Spectral Approach*. 1st ed. Springer, New York, NY, 1991. DOI: `10.1007/978-1-4612-3094-6`.

[47] M. B. Giles. "Multilevel Monte Carlo methods". In: *Acta Numerica* 24 (2015), pp. 259–328. DOI: `10.1017/S096249291500001X`.

[48] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. "Multilevel splitting for estimating rare event probabilities". In: *Operations Research* 47.4 (1999), pp. 585–600. DOI: `10.1287/opre.47.4.585`.

[49] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. URL: `http://www.deeplearningbook.org`.

[50] D. F. Griffiths, J. W. Dold, and D. J. Silvester. *Essential Partial Differential Equations: Analytical and Computational Aspects*. 1st ed. Springer Undergraduate Mathematics Series. Springer, Cham, 2015. DOI: `10.1007/978-3-319-22569-2`.

[51] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. 2nd ed. Springer Series in Computational Mathematics. Springer, Berlin, Heidelberg, 2017. DOI: `10.1007/978-3-662-54961-2`.

[52] A. M. Hasofer and N. C. Lind. "An exact and invariant first order reliability format". In: *Journal of Engineering Mechanics* 100 (1974), pp. 111–121.

[53] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (1970), pp. 97–109. DOI: `10.2307/2334940`.

[54] S. Heinrich. "Multilevel Monte Carlo methods". In: *Large-Scale Scientific Computing*. Ed. by S. Margenov, J. Waśniewski, and P. Yalamov. Vol. 2179. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, 2001, pp. 58–67. DOI: `10.1007/3-540-45346-6_5`.

[55] M. Herty and G. Visconti. "Kinetic methods for inverse problems". In: *Kinetic & Related Models* 12.5 (2019), pp. 1109–1130. DOI: 10.3934/krm.2019042.

[56] M. Hohenbichler and R. Rackwitz. "Non-normal dependent vectors in structural safety". In: *Journal of the Engineering Mechanics Division* 107.6 (1981), pp. 1227–1238. DOI: 10.1061/JMCEA3.0002777.

[57] M. A. Iglesias, K. Law, and A. M. Stuart. "Ensemble Kalman methods for inverse problems". In: *Inverse Problems* 29.4 (2013), p. 045001. DOI: 10.1088/0266-5611/29/4/045001.

[58] M. A. Iglesias, M. Park, and M. V. Tretyakov. "Bayesian inversion in resin transfer molding". In: *Inverse Problems* 34.10 (2018), p. 105002. DOI: 10.1088/1361-6420/aad1cc.

[59] E. Janouchová, J. Sýkora, and A. Kučerová. "Polynomial chaos in evaluating failure probability: a comparative study". In: *Applications of Mathematics* 63.6 (2018), pp. 713–737. DOI: 10.21136/AM.2018.0335-17.

[60] A. Jasra, D. A. Stephens, A. Doucet, and T. Tsagaris. "Inference for Lévy-driven stochastic volatility models via adaptive sequential Monte Carlo". In: *Scandinavian Journal of Statistics* 38.1 (2010), pp. 1–22. DOI: 10.1111/j.1467-9469.2010.00723.x.

[61] H. Kahn and A. W. Marshall. "Methods of reducing sample size in Monte Carlo computations". In: *Journal of the Operations Research Society of America* 1.5 (1953), pp. 263–278. DOI: 10.1287/opre.1.5.263.

[62] R. E. Kalman. "A new approach to linear filtering and prediction problems". In: *Journal of Basic Engineering* 82.1 (1960), pp. 35–45. DOI: 10.1115/1.3662552.

[63] K. Karhunen. "Über lineare Methoden in der Wahrscheinlichkeitsrechnung". In: *Annales Academiae Scientiarum Fennicae: Ser. A. I. Mathematica - Physica* 37 (1947), pp. 1–79.

[64] L. S. Katafygiotis and K. M. Zuev. "Geometric insight into the challenges of solving high-dimensional reliability problems". In: *Probabilistic Engineering Mechanics* 23.2 (2008), pp. 208–218. DOI: 10.1016/j.probengmech.2007.12.026.

[65] S. Katsuki and D. M. Frangopol. "Hyperspace division method for structural reliability". In: *Journal of Engineering Mechanics* 120.11 (1994), pp. 2405–2427. DOI: 10.1061/(ASCE)0733-9399(1994)120:11(2405).

[66] P. S. Koutsourelakis. "A multi-resolution, non-parametric, Bayesian framework for identification of spatially-varying model parameters". In: *Journal of Computational Physics* 228.17 (2009), pp. 6184–6211. DOI: 10.1016/j.jcp.2009.05.016.

[67] P. S. Koutsourelakis, H. J. Pradlwarter, and G. I. Schuëller. "Reliability of structures in high dimensions, part I: algorithms and applications". In: *Probabilistic Engineering Mechanics* 19.4 (2004), pp. 409–417. DOI: 10.1016/j.probengmech.2004.05.001.

[68] D. P. Kroese, R. Y. Rubinstein, and P. W. Glynn. "Chapter 2 - The cross-entropy method for estimation". In: *Handbook of Statistics*. Ed. by C. R. Rao and V. Govindaraj. Vol. 31. Handbook of Statistics. Elsevier, 2013, pp. 19–34. DOI: 10.1016/B978-0-444-53859-8.00002-3.

[69] J. Latz. "On the well-posedness of Bayesian inverse problems". In: *SIAM/ASA Journal on Uncertainty Quantification* 8.1 (2020), pp. 451–482. DOI: 10.1137/19M1247176.

[70] J. Latz, I. Papaioannou, and E. Ullmann. "Multilevel sequential Monte Carlo for Bayesian inverse problems". In: *Journal of Computational Physics* 368 (2018), pp. 154–178. DOI: 10.1016/j.jcp.2018.04.014.

[71] K. Law, A. M. Stuart, and K. Zygalakis. *Data Assimilation: A Mathematical Introduction*. 1st ed. Texts in Applied Mathematics. Springer, Cham, 2015. DOI: 10.1007/978-3-319-20325-6.

[72] M. Lemaire, A. Chateauneuf, and J.-C. Mitteau. *Structural Reliability*. John Wiley & Sons, Ltd, 2009. DOI: 10.1002/9780470611708.

[73] R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. Society for Industrial and Applied Mathematics, 2007. DOI: 10.1137/1.9780898717839.

[74] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002. DOI: 10.1017/CBO9780511791253.

[75] J. Li and D. Xiu. "Evaluation of failure probability via surrogate models". In: *Journal of Computational Physics* 229.23 (2010), pp. 8966–8980. DOI: 10.1016/j.jcp.2010.08.022.

[76] K. Li, K. Tang, J. Li, T. Wu, and Q. Liao. "A hierarchical neural hybrid method for failure probability estimation". In: *IEEE Access* 7 (2019), pp. 112087–112096. DOI: 10.1109/ACCESS.2019.2934980.

[77] R. Li, V. Prasad, and B. Huang. "Gaussian mixture model-based ensemble Kalman filtering for state and parameter estimation for a PMMA process". In: *Processes* 4.2.9 (2016). DOI: 10.3390/pr4020009.

[78] P.-L. Liu and A. Der Kiureghian. "Optimization algorithms for structural reliability". In: *Structural Safety* 9.3 (1991), pp. 161–177. DOI: 10.1016/0167-4730(91)90041-7.

[79] M. Loève. *Probability Theory II*. 4th ed. Graduate Texts in Mathematics, Volume 46. Springer-Verlag, New York, 1978.

[80] G. J. Lord, C. E. Powell, and T. Shardlow. *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2014. DOI: 10.1017/CBO9781139017329.

[81] H. O. Madsen, S. Krenk, and N. C. Lind. *Methods of Structural Safety*. Prentice-Hall Series in Computational Mathematics. Prentice-Hall, 1985.

[82] S. Marelli and B. Sudret. "An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis". In: *Structural Safety* 75 (2018), pp. 67–74. DOI: 10.1016/j.strusafe.2018.06.003.

[83] G. McLachlan and D. Peel. "ML fitting of mixture models". In: *Finite Mixture Models*. John Wiley & Sons, Ltd, 2000. Chap. 2, pp. 40–80. DOI: 10.1002/0471721182.ch2.

[84] R. E. Melchers and A. T. Beck. *Structural Reliability Analysis and Prediction*. 3rd ed. John Wiley & Sons, Ltd, 2017. DOI: 10.1002/9781119266105.

[85] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. "Equations of state calculations by fast computing machine". In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[86] J. Morio and M. Balesdent. *Estimation of Rare Event Probabilities in Complex Aerospace and other Systems*. 1st ed. Woodhead Publishing, 2015. DOI: 10.1016/C2014-0-02344-1.

[87] K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. Adaptive Computation and Machine Learning series. The MIT Press, 2012. URL: https://mitpress.mit.edu/books/machine-learning-1.

[88] M. Nakagami. "The m-distribution, a general formula of intensity distribution of rapid fading". In: *Statistical Methods in Radio Wave Propagation*. Ed. by W. Hoffman. Pergamon, 1960, pp. 3–36. DOI: 10.1016/B978-0-08-009306-2.50005-4.

[89] U. Noseck et al. *Assessment of the long-term safety of repositories*. Gesellschaft für Anlage und Reaktorsicherheit (GRS) mbH, 2008. URL: https://www.grs.de/sites/default/files/pdf/GRS-237.pdf.

[90] A. B. Owen. *Monte Carlo theory, methods and examples*. 2013. URL: https://statweb.stanford.edu/~owen/mc/.

[91] V. Papadopoulos, D. G. Giovanis, N. D. Lagaros, and M. Papadrakakis. "Accelerated subset simulation with neural networks for reliability analysis". In: *Computer Methods in Applied Mechanics and Engineering* 223–224 (2012), pp. 70–80. DOI: 10.1016/j.cma.2012.02.013.

[92] I. Papaioannou, W. Betz, K. Zwirglmaier, and D. Straub. "MCMC algorithms for subset simulation". In: *Probabilistic Engineering Mechanics* 41 (2015), pp. 89–103. DOI: 10.1016/j.probengmech.2015.06.006.

[93] I. Papaioannou, S. Geyer, and D. Straub. "Improved cross entropy-based importance sampling with a flexible mixture model". In: *Reliability Engineering & System Safety* 191 (2019), p. 106564. DOI: 10.1016/j.ress.2019.106564.

[94] I. Papaioannou, C. Papadimitriou, and D. Straub. "Sequential importance sampling for structural reliability analysis". In: *Structural Safety* 62 (2016), pp. 66–75. DOI: 10.1016/j.strusafe.2016.06.002.

[95]   B. Peherstorfer, B. Kramer, and K. Willcox. "Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation". In: *SIAM/ASA Journal on Uncertainty Quantification* 6.2 (2018), pp. 737–761. DOI: 10.1137/17M1122992.

[96]   C. Proppe. "Estimation of failure probabilities by local approximation of the limit state function". In: *Structural Safety* 30.4 (2008), pp. 277–290. DOI: 10.1016/j.strusafe.2007.04.001.

[97]   D. Proske. *Bridge Collapse Frequencies versus Failure Probabilities*. 1st ed. Risk Engineering. Springer, Cham, 2018. DOI: 10.1007/978-3-319-73833-8.

[98]   R. Rackwitz. "Reliability analysis—a review and some perspectives". In: *Structural Safety* 23.4 (2001), pp. 365–395. DOI: 10.1016/S0167-4730(02)00009-7.

[99]   R. Rackwitz and B. Flessler. "Structural reliability under combined random load sequences". In: *Computers & Structures* 9.5 (1978), pp. 489–494. DOI: 10.1016/0045-7949(78)90046-9.

[100] R. Rannacher and B. Vexler. "A priori error estimates for the finite element discretization of elliptic parameter identification problems with pointwise measurements". In: *SIAM Journal on Control and Optimization* 44.5 (2005), pp. 1844–1863. DOI: 10.1137/040611100.

[101] P. A. Raviart and J. M. Thomas. "A mixed finite element method for 2-nd order elliptic problems". In: *Mathematical Aspects of Finite Element Methods*. Ed. by I. Galligani and E. Magenes. Vol. 606. Lecture Notes in Mathematics. Springer, Berlin, Heidelberg, 1977, pp. 292–315. DOI: 10.1007/BFb0064470.

[102] S. Reich and S. Weissmann. "Fokker–Planck particle systems for Bayesian inference: computational approaches". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.2 (2021), pp. 446–482. DOI: 10.1137/19M1303162.

[103] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. 2nd ed. Springer Texts in Statistics. Springer, New York, NY, 2004. DOI: 10.1007/978-1-4757-4145-2.

[104] G. O. Roberts and J. S. Rosenthal. "Optimal scaling for various Metropolis–Hastings algorithms". In: *Statistical Science* 16.4 (2001), pp. 351–367. DOI: 10.1214/ss/1015346320.

[105] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo Method*. 3rd ed. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd, 2016. DOI: 10.1002/9781118631980.

[106] C. Schillings and A. M. Stuart. "Analysis of the ensemble Kalman filter for inverse problems". In: *SIAM Journal on Numerical Analysis* 55.3 (2017), pp. 1264–1290. DOI: 10.1137/16M105959X.

[107] C. Schillings and A. M. Stuart. "Convergence analysis of ensemble Kalman inversion: the linear, noisy case". In: *Applicable Analysis* 97.1 (2018), pp. 107–123. DOI: 10.1080/00036811.2017.1386784.

[108] R. Schöbi, B. Sudret, and S. Marelli. "Rare event estimation using polynomial-chaos kriging". In: *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 3.2 (2017), p. D4016002. DOI: 10.1061/AJRUA6.0000870.

[109] K. W. Smith. "Cluster ensemble Kalman filter". In: *Tellus A: Dynamic Meteorology and Oceanography* 59.5 (2007), pp. 749–757. DOI: 10.1111/j.1600-0870.2007.00246.x.

[110] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method.* 2nd ed. Wellesley-Cambridge Press, 1997.

[111] D. Straub, I. Papaioannou, and W. Betz. "Bayesian analysis of rare events". In: *Journal of Computational Physics* 314 (2016), pp. 538–556. DOI: 10.1016/j.jcp.2016.03.018.

[112] A. M. Stuart. "Inverse problems: a Bayesian perspective". In: *Acta Numerica* 19 (2010), pp. 451–559. DOI: 10.1017/S0962492910000061.

[113] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann. "Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients". In: *Numerische Mathematik* 125 (2013), pp. 569–600. DOI: 10.1007/s00211-013-0546-4.

[114] E. Ullmann and I. Papaioannou. "Multilevel estimation of rare events". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 922–953. DOI: 10.1137/140992953.

[115] F. Uribe, I. Papaioannou, Y. M. Marzouk, and D. Straub. "Cross-entropy-based importance sampling with failure informed dimension reduction for rare event estimation". In: *arXiv* (2020). URL: https://arxiv.org/abs/2006.05496.

[116] B. Vexler. "Adaptive Finite Element Methods for Parameter Identification Problems". PhD thesis. University of Heidelberg, 2004. DOI: 10.11588/heidok.00004603.

[117] P. H. Waarts. *Structural reliability using finite element methods. An appraisal for DARS: Directional Adaptive Response surface Sampling.* Delft University Press, 2000.

[118] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann. "Error analysis for probabilities of rare events with approximate models". In: *arXiv* (2021). URL: https://arxiv.org/abs/2008.06368.

[119] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann. "Multilevel sequential importance sampling for rare event estimation". In: *SIAM Journal on Scientific Computing* 42.4 (2020), A2062–A2087. DOI: 10.1137/19M1289601.

[120] F. Wagner, I. Papaioannou, and E. Ullmann. "The ensemble Kalman filter for rare event estimation". In: *arXiv* (2021). URL: https://arxiv.org/abs/2106.10062.

[121] C. Walter. "Moving particles: a parallel optimal multilevel splitting method with application in quantiles estimation and meta-model based algorithms". In: *Structural Safety* 55 (2015), pp. 10–25. DOI: 10.1016/j.strusafe.2015.02.002.

[122] Z. Wang, M. Broccardo, and J. Song. "Hamiltonian Monte Carlo methods for subset simulation in reliability analysis". In: *Structural Safety* 76 (2019), pp. 51–67. DOI: 10.1016/j.strusafe.2018.05.005.

[123] Z. Wang and J. Song. "Cross-entropy-based adaptive importance sampling using von Mises–Fisher mixture for high dimensional reliability analysis". In: *Structural Safety* 59 (2016), pp. 42–52. DOI: 10.1016/j.strusafe.2015.11.002.

A. Core Article: Multilevel sequential importance sampling for rare event estimation

Summary

In this article, we develop Multilevel Sequential Importance Sampling (MLSIS) as a novel sampling-based estimator for the probability of rare events. We focus on limit-state functions (LSFs) which require the solution of a partial differential equation. These LSFs are expensive to handle, which often makes the estimation of failure events by simple Monte Carlo methods computationally intractable. Our novel estimator is based on Sequential Importance Sampling (SIS) [94] and combines the evaluation of LSFs with different discretization accuracies, e.g. a finite element mesh size. We employ the two-fold adaptive algorithm of [70], which ensures that we obtain an estimate based on the desired discretization accuracy. In the MLSIS algorithm, we sequentially enhance the approximation of the optimal Importance Sampling density. On the one hand, we increase the discretization level. On the other hand, we apply tempering to increase the approximation of the indicator function, which is part of the optimal Importance Sampling density.

Our estimator overcomes the nestedness problem associated with the Multilevel Subset Simulation (MLSuS) estimator, which has been developed in [114]. MLSIS automatically satisfies nestedness since the support of the employed densities is the whole parametric domain. As suggested in [114] for MLSuS, we employ a level dependent parameter dimension for the MLSIS estimator to decrease the discrepancy of two consecutive discretization levels. This yields computational benefits since more tempering updates are performed on coarse levels.

Another contribution of this article is a novel Markov chain Monte Carlo (MCMC) kernel based on independent proposals from an adaptively constructed von Mises–Fisher–Nakagami (vMFN) distribution. The vMFN distribution model yields benefits compared with a Gaussian mixture model, since the vMFN distribution model is applicable even in high-dimensional parametric spaces. In numerical experiments, we compare the performance of MLSIS with SIS, SuS and MLSuS. We observe that MLSIS requires less computational costs than SIS for a fixed level of accuracy, while the performance of MLSIS is similar to the performance of MLSuS. Moreover, sampling from the vMFN distribution model decreases the variance of the MLSIS estimator compared to adaptive conditional sampling.

Statement of individual contribution

Elisabeth Ullmann and Jonas Latz assigned me with the task of implementing and analysing the multilevel idea of [70] for SIS to develop a novel multilevel algorithm for rare event estimation. Iason Papaionnou suggested the application of the vMFN distribution model as an independent proposal density in the MCMC algorithm.

I was fully responsible for the implementation and execution of numerical experiments. I was supported by insightful discussions with Jonas Latz, Iason Papaionnou and Elisabeth Ullmann concerning technical details. Moreover, I was fully responsible for writing and preparing this manuscript. Jonas Latz, Iason Papaioannou, Elisabeth

Ullmann, and I proofread and polished the article together.

Permission to include

On the following page, a copy of the first page of the consent to publish agreement by SIAM may be found. This page includes the author's rights. A digital version of the consent to publish may be found at

    `https://www.siam.org/publications/journals/about-siam-journals/`
                        `information-for-authors`

(Accessed on 21 June 2021)

Consent to Publish

SIAM ("Publisher") requires Authors of articles in SIAM publications to provide a formal written Consent to Publish. The Author must sign the agreement except, in the case of "work-for-hire", when the Author's employer may sign as the party that has the right to grant rights to the Publisher. If there are multiple Authors of the material governed by this document, the term "Author" as used here refers to each and all of them, jointly and severally. [1]

Title of Contribution ("Work"): Multilevel Sequential Importance Sampling for Rare Event Estimation
Authors: Fabian Wagner, Jonas Latz, Iason Papaioannou, and Elisabeth Ullmann
Name of Journal: SIAM Journal on Scientific Computing
Manuscript Number: M128960

1. Author's Warranty

By signing this Consent, the Author warrants all of the following: The Work has not been published before in any form except as a preprint, unless explicitly noted as a footnote to the title. The Work is not being concurrently submitted to and is not under consideration by another publisher. The names listed above as authors appear in the manuscript itself, no author entitled to credit has been omitted, and there are no unnamed authors. The Author has the right to make the grants made to the Publisher complete and unencumbered. The Author also warrants that the Work does not libel anyone, violate anyone's privacy or publicity rights, infringe anyone's copyright, trademark, or trade secrets, or otherwise violate anyone's statutory or common law rights.

2. Author's Rights

A1. The Author may reproduce and distribute the Work (including derivative works) in connection with the Author's teaching, technical collaborations, conference presentations, lectures, or other scholarly works and professional activities as well as to the extent the fair use provisions of the U.S. Copyright Act permit. If the copyright is granted to the Publisher, then the proper notice of the Publisher's copyright should be provided.
A2. The Author may post the final draft of the Work, as it exists immediately prior to editing and production by the Publisher, on noncommercial pre-print servers such as arXiv.org.
A3. The Author may post the final published version of the Work on the Author's personal web site and on the web server of the Author's institution, provided that proper notice of the Publisher's copyright is included and that no separate or additional fees are collected for access to or distribution of the work.

3. Publisher's Rights

Even if the Author does not transfer Copyright to the Publisher, the Author grants the Publisher the following rights in perpetuity.

P1. The Publisher has unlimited rights throughout the world to publish and distribute the final version of the Work in any form and in all media now known or hereafter discovered.
P2. The Publisher has unlimited rights throughout the world to translate the final version of the Work and exercise all rights in all media in the resulting

Notice of publication and copyright

MULTILEVEL SEQUENTIAL IMPORTANCE SAMPLING FOR RARE EVENT ESTIMATION[*]

F. WAGNER[†], J. LATZ[‡], I. PAPAIOANNOU[§], AND E. ULLMANN[†]

Abstract. The estimation of the probability of rare events is an important task in reliability and risk assessment. We consider failure events that are expressed in terms of a limit state function, which depends on the solution of a partial differential equation (PDE). Since numerical evaluations of PDEs are computationally expensive, estimating such probabilities of failure by Monte Carlo sampling is intractable. We develop a novel estimator based on a sequential importance sampler using discretizations of PDE-based limit state functions with different accuracies. A twofold adaptive algorithm ensures that we obtain an estimate based on the desired discretization accuracy. Moreover, we suggest and study the choice of the Markov chain Monte Carlo kernel for use with sequential importance sampling. Instead of the popular adaptive conditional sampling method, we propose a new algorithm that uses independent proposals from an adaptively constructed von Mises–Fisher–Nakagami distribution.

Key words. reliability analysis, importance sampling, multilevel Monte Carlo, subset simulation, Markov chain Monte Carlo

AMS subject classifications. 35R60, 65C40, 65C60, 65N30

DOI. 10.1137/19M1289601

1. Introduction. Estimating the probability of rare events is crucial in reliability analysis and risk management and arises in applications in many fields. For instance, the authors in [3] examine rare events arising in financial risk settings while [41] studies the probability of collision between space debris and satellites. In planning a radioactive waste repository [12, 43], one is interested in the probability that radioactive particles leave the repository and pollute the groundwater in a long time horizon. The particle flow can be simulated by a finite element (FEM) [10] approximation of the groundwater flow and transport equation. Since the subsurface properties of the whole domain of interest are uncertain or only measurable at finitely many points, the soil is modelled as a random field. The particle transport has to be simulated for various realizations of the random field to estimate the probability that the radioactive particles come back to the human environment, which is a rare event.

All applications have in common that the probabilities of the events are small ($< 10^{-4}$) and the *limit state function* (LSF) underlies a computationally demanding model which depends on the discretization of the domain. If the discretization level is high, i.e., the mesh size is small, the FEM approximation is accurate but also cost intensive. These issues complicate the estimation of the probability of failure.

Before we introduce our novel approach, we give a brief overview of existing algorithms. On the one hand, there are deterministic approximation methods, such as the *first* and *second order reliability methods* (FORM, SORM) [39], which aim at approx-

imating the domain of parameters which lead to failure events. On the other hand, there are sampling based methods, which approximate the probability of failure events. Unlike approximation methods, sampling approaches are based on sample estimates and are usually more robust in terms of the complexity of the LSF. Since our novel approach is a sampling method, we focus on this category and give a larger overview.

Monte Carlo sampling [21, 52] can be easily applied to estimate the probability of failure and yields an unbiased estimator. However, due to the mentioned issues of rare event settings, the Monte Carlo estimator becomes intractable, since hardly any sample contributes to the rare event and each sample requires a cost intensive function evaluation. Therefore, variance reduction techniques have been developed to reduce the number of samples for obtaining an accurate estimate. For instance, the idea of *multilevel splitting* [8, 27] and *subset simulation* (SuS) [4, 5] is to decompose the rare event into a sequence of nested events. This enables expressing the probability of the rare event as a product of conditional probabilities of more frequent events. These methods require sampling from a sequence of probability density functions which is achieved with *Markov chain Monte Carlo* (MCMC) methods [44, 54].

Importance sampling (IS) methods employ an alternative sampling density, which if chosen properly can reduce considerably the variance of the standard Monte Carlo estimator [31]. The optimal choice of the sampling density is the density of the input variables conditional on the failure domain. However, direct sampling from the optimal density is not feasible, because the location of the failure domain is unknown prior to performing the simulation. As in multilevel splitting or SuS, a sequential approach can be applied to approximate the optimal IS density in a sequential manner. This leads to *sequential importance sampling* (SIS) [45, 47] or *sequential Monte Carlo* (SMC) [14] for the estimation of rare events. In our novel approach, we consider an adaptive methodology similar to adaptive SMC [7, 19, 30]. Another approach to estimate the optimal sampling density sequentially is the *cross-entropy* method [23], where the sampling density minimizes the Kullback–Leibler divergence to the optimal density within a family of parametrized densities. IS can also be applied to a hyperplane that is perpendicular to an important direction, a method known as *line sampling* [16, 35, 49].

The previous algorithms have the drawback that all evaluations have to be performed with respect to the same LSF. The evaluation of the LSF could require the solution of a discretized PDE, which depends on the mesh size of the computational domain. Since computational costs increase with decreasing mesh size, we wish to construct a method wherein the discretized PDE is solved on fine meshes only for very few realizations. Therefore, we apply a multilevel approach that uses a hierarchy of discretization levels. The authors in [20] use the telescoping sum approach of [26] to estimate the probability of failure. Applying the multilevel idea to the previously described methods gives *multilevel subset simulation* (MLSuS) [53] and *multilevel sequential Monte Carlo* [6, 17]. Moreover, a multifidelity approach combined with the cross-entropy method is investigated in [48]. Furthermore, the work in [37] develops the *multilevel sequential*^{2} *Monte Carlo* (MLS^{2}MC) estimator, which is a twofold sequential algorithm for Bayesian inverse problems.

In this paper, we consider SuS and SIS as well as their multilevel versions. In more detail, an MCMC algorithm [13, 28] is applied within SuS to gradually shift samples into consecutive domains, which are defined by the sequence of nested events. By the *nestedness property* [44], the simulated Markov chains do not require a burn-in period, since seeds are already distributed approximately according to the target distribution. Therefore, SuS is an efficient but slightly biased estimator [4]. The MLSuS

method, given in [53], employs a hierarchy of discretization levels and enables the usage of coarse grid function evaluations. MLSuS saves significant computational costs compared to SuS if the failure domains between discretization levels are still nested. However, nestedness is no longer guaranteed in the multilevel setting since the sequence of consecutive domains is based on LSFs with different accuracies. Therefore, a second MCMC step has to be performed. Additionally, a burn-in period is proposed since seeds are no longer distributed (approximately) according to the target distribution. Both issues increase the computational costs of the MLSuS estimator and thus decrease its efficiency. However, a level dependent parameter dimension can be applied to reduce variances between two accuracy levels of the LSF and approximately satisfy the nestedness property.

The nestedness issue of MLSuS is our main motivation to implement the MLS^{2}MC algorithm for rare event estimation. Nestedness is not an issue for MLS^{2}MC; the method samples a sequence of nonzero densities with IS and chooses each IS density to be close to each target density in the sequence. The idea of the MLS^{2}MC method is combined with the SIS approach and yields a *multilevel sequential importance sampling* (MLSIS) estimator for rare events. Note that both MLSIS as well as MLSuS are not based on the telescoping sum approach. To achieve an even more efficient algorithm, we apply the level dependent parameter dimension approach of [53]. Like SIS, the MLSIS method requires an MCMC algorithm to shift samples into consecutive target distributions. We consider an independent MCMC sampler that uses the *von Mises–Fisher Nakagami* (vMFN) distribution model fitted with the available weighted samples at each sampling level as the proposal distribution. The vMFN distribution is applied in [46] as a parametrized family of probability distributions for the cross-entropy method, which yields an efficient algorithm even in high dimensions. Employing the vMFN distribution as a proposal density is another main contribution of our work.

The paper is structured as follows. In section 2, the problem setting of estimating the probability of failure is defined and SIS as well as SuS are explained. The MLSIS estimator is described in section 3. In section 4, two MCMC algorithms are studied which are applied within SIS and MLSIS. In section 5, the studied estimators are applied to one- and two-dimensional test problems and the MLSIS estimator is compared with SIS as well as SuS and MLSuS. In section 6, a summary of the discussion and an outlook are given.

2. Background.

2.1. Problem setting. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $U : \Omega \to \mathbb{R}^n$. By [18, 29] it is assumed, without loss of generality, that U is distributed according to the n-variate standard normal distribution with density function φ_n. If a non-Gaussian random variable $\tilde{U}$ is used, an isoprobabilistic transformation $U = T(\tilde{U})$ is applied. Failure is defined in terms of an LSF $G : \mathbb{R}^n \to \mathbb{R}$ such that $G(U(\omega)) \leq 0$ for $\omega \in \Omega$. In many applications, the LSF G is not analytically given. We can only evaluate an approximation G_ℓ, where ℓ represents the discretization level. Increasing ℓ leads to a more accurate approximation. In the numerical examples presented in this paper, G_ℓ requires the solution of a PDE and ℓ specifies the mesh size of an FEM approximation. The *probability of failure* is defined as the measure of the *failure domain* $A := \{\omega \in \Omega : G(U(\omega)) \leq 0\}$, which is expressed as

$$(2.1) \qquad P_f := \mathbb{P}[A] = \mathbb{P}[G(U) \leq 0] = \int_{G(u) \leq 0} \varphi_n(u) \mathrm{d}u.$$

Using G_ℓ instead of G in (2.1) gives the approximation $P_{f,\ell}$, which includes numerical errors due to approximating the exact LSF G. Convergence is expected for increasing the level ℓ, i.e., decreasing the finite element mesh size.

The probability of failure can be estimated by crude Monte Carlo sampling [21]. By evaluating G_ℓ on the discretization level ℓ for $N \in \mathbb{N}$ independent samples distributed according to φ_n, we obtain the (single-level) Monte Carlo estimator $\hat{P}_{f,\ell}^{\mathrm{MC}}$ for $P_{f,\ell}$

$$(2.2) \qquad \hat{P}_{f,\ell}^{\mathrm{MC}} = \frac{1}{N} \sum_{k=1}^{N} I\left(G_\ell(u_k) \leq 0\right),$$

where I denotes the indicator function; i.e., $I(\text{true}) = 1$ and $I(\text{false}) = 0$. $\hat{P}_{f,\ell}^{\mathrm{MC}}$ is an unbiased estimator and easy to implement. Since the coefficient of variation of $\hat{P}_{f,\ell}^{\mathrm{MC}}$ is inversely proportional to the probability of failure $P_{f,\ell}$ (see [47]), a large number of samples is required if $P_{f,\ell}$ is small and a small coefficient of variation should be achieved. Hence, huge computational costs are required if G_ℓ is a cost demanding evaluation. This makes crude Monte Carlo sampling impractical for the estimation of rare failure probabilities.

2.2. Subset simulation and multilevel subset simulation. SuS and MLSuS are alternative approaches where the failure probability is estimated by a product of conditional probabilities. Consider the sequence of domains $B_0, B_1, \ldots, B_S$, where $B_S = A$ is the failure domain. In both approaches, the sequence is constructed such that

$$(2.3) \qquad \mathbb{P}[B_j \mid B_{j-1}] = \hat{p}_0 \in (0, 1),$$

while $\hat{p}_0$ is chosen to ensure that samples of B_j can be easily generated from samples of B_{j-1} [4]. In SuS, the sequence of domains is nested, i.e., $B_j \subset B_{j-1}$ for $j = 1, \ldots, S$, since the discretization level is fixed. Hence, the SuS estimator is given as

$$\hat{P}_{f,\ell}^{\mathrm{SuS}} := \hat{P}_{B_1} \prod_{j=2}^{S} \hat{P}_{B_j \mid B_{j-1}},$$

where $\hat{P}_{B_j \mid B_{j-1}}$ is an estimator for $\mathbb{P}[B_j \mid B_{j-1}]$. It has been shown in [47] that SuS is a special case of SIS, where the IS densities $p_{j,\ell}$ are chosen as the optimal IS density with respect to the domain B_j. In MLSuS [53], the sequence of domains is no longer nested since the domains B_j are defined for different LSFs G_ℓ, in the case of a level update. To overcome this problem, the conditional probability $\mathbb{P}[B_{j-1} \mid B_j]$ has to be estimated. This leads to the MLSuS estimator

$$(2.4) \qquad \hat{P}_{f,\ell}^{\mathrm{MLSuS}} := \hat{P}_{B_1} \prod_{j=2}^{S} \frac{\hat{P}_{B_j \mid B_{j-1}}}{\hat{P}_{B_{j-1} \mid B_j}}.$$

Moreover, samples which are taken as seeds in the MCMC step are not distributed according to the target distribution. Therefore, a burn-in is required. Both issues lead to increasing computational costs. Note that if the domains B_j for $j = 1, \ldots, S$ were nested, then the denominator in (2.4) is equal to one and no estimator for the denominator is required. To increase the denominator in (2.4), the authors in [53] apply a level dependent parameter dimension. This reduces the variance of two

consecutive levels and makes the MLSuS algorithm more robust. In section 3.3 of this work, we consider a level dependent parameter dimension for the MLSIS algorithm to reduce the variance of consecutive levels.

2.3. Importance sampling. IS is a variance reduction technique [2, 52], where the integral in (2.1) is calculated with respect to a certain *IS density* p_ℓ. If p_ℓ takes on large values in the failure domain, many samples following p_ℓ represent failure events. Therefore, fewer samples are required to estimate the probability of failure accurately. By [45] the failure probability $P_{f,\ell}$ is expressed as

$$P_{f,\ell} = \int_{\mathbb{R}^n} I\left(G_\ell(u) \leq 0\right) w_\ell(u) p_\ell(u) \mathrm{d}u = \mathbb{E}_{p_\ell}[I\left(G_\ell(u) \leq 0\right) w_\ell(u)],$$

where the *importance weight* is defined as $w_\ell(u) := \varphi_n(u)/p_\ell(u)$. Again, crude Monte Carlo sampling is applied, which yields the estimator

$$\hat{P}_{f,\ell}^{\mathrm{IS}} = \frac{1}{N} \sum_{k=1}^{N} I\left(G_\ell(u_k) \leq 0\right) w_\ell(u_k),$$

where the samples $\{u_k\}_{k=1}^N$ are distributed according to the IS density p_ℓ. $\hat{P}_{f,\ell}^{\mathrm{IS}}$ is an unbiased estimator for $P_{f,\ell}$ if the support of p_ℓ contains the failure domain $A_\ell := \{\omega \in \Omega : G_\ell(U(\omega)) \leq 0\}$. The *optimal IS density* is given by

$$(2.5) \qquad p_{\mathrm{opt},\ell}(u) := \frac{1}{P_{f,\ell}} I(G_\ell(u) \leq 0)\varphi_n(u),$$

which leads to a zero-variance estimator. Since $P_{f,\ell}$ and A_ℓ are unknown, $p_{\mathrm{opt},\ell}$ cannot be used in practice. In contrast, SIS achieves an approximation of $p_{\mathrm{opt},\ell}$ by approximating the optimal IS distribution in a sequential manner while starting from a known *prior density* p_0.

2.4. Sequential importance sampling. According to [47], the sequence of IS densities is determined from a smooth approximation of the indicator function. The *cumulative distribution function* (cdf) of the standard normal distribution is one possibility to approximate the indicator function. For $G_\ell(u) \neq 0$ we achieve pointwise convergence

$$I(G_\ell(u) \leq 0) = \lim_{\sigma \downarrow 0} \Phi\left(-\frac{G_\ell(u)}{\sigma}\right),$$

while for $G_\ell(u) = 0$ and $\forall \sigma > 0$ it holds that $\Phi\left(-G_\ell(u)/\sigma\right) = 1/2 \neq I(G_\ell(u) \leq 0)$, as visualized in Figure 1. Further approximation functions are examined in [36] with an additional sensitivity analysis.

With the preceding consideration, the sequence of IS densities $\{p_{j,\ell} : j = 0, \ldots, N_T\}$ is defined as

$$p_{j,\ell}(u) := \frac{1}{P_{j,\ell}} \Phi\left(-\frac{G_\ell(u)}{\sigma_j}\right) \varphi_n(u) = \frac{1}{P_{j,\ell}} \eta_{j,\ell}(u) \text{ for } j = 1, \ldots, N_T,$$
$$p_0(u) := \varphi_n(u),$$

where $\infty > \sigma_1 > \cdots > \sigma_{N_T} > 0$ represent a strictly decreasing sequence of temperatures or bandwidths and $P_{j,\ell}$ is a normalizing constant such that $p_{j,\ell}$ is a well-defined
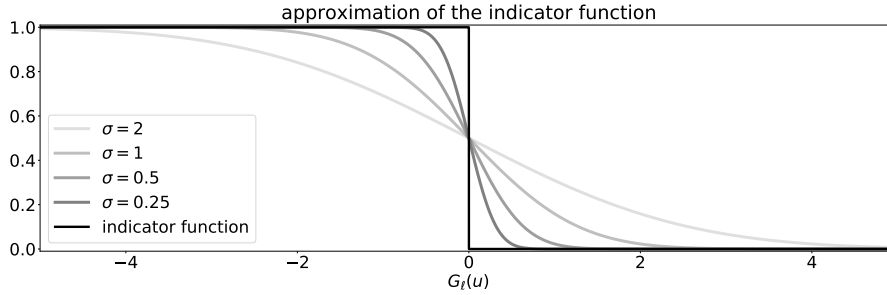
FIG. 1. *Approximation of the indicator function $I(G_\ell(u) \leq 0)$ by the cdf of the standard normal distribution $\Phi(-G_\ell(u)/\sigma)$.*

density function. The denomination "temperatures" and their use is motivated by the temperature of the Boltzmann distribution [25, Chapter VIII]. The number N_T of tempering steps is a priori unknown and specifies the number of tempering steps to approximate the optimal IS density sufficiently accurately. Applying the IS approach, $P_{j,\ell}$ is determined by sampling from the density $p_{j-1,\ell}$

$$(2.6) \quad P_{j,\ell} = \int_{\mathbb{R}^n} \eta_{j,\ell}(u)\mathrm{d}u = P_{j-1,\ell} \int_{\mathbb{R}^n} w_{j,\ell}(u)p_{j-1,\ell}(u)\mathrm{d}u = P_{j-1,\ell}\mathbb{E}_{p_{j-1,\ell}}[w_{j,\ell}(u)],$$

where $w_{j,\ell}(u) := \eta_{j,\ell}(u)/\eta_{j-1,\ell}(u)$. Hence, the fraction of consecutive normalizing constants $S_{j,\ell} = P_{j,\ell}/P_{j-1,\ell}$ is estimated by

$$(2.7) \qquad \hat{S}_{j,\ell} := \hat{\mathrm{E}}_{p_{j-1,\ell}}[w_{j,\ell}(u)] = \frac{1}{N} \sum_{k=1}^{N} w_{j,\ell}(u_k),$$

where the samples $\{u_k\}_{k=1}^{N}$ are distributed according to $p_{j-1,\ell}$. Using the definition of $\eta_{j,\ell}$ and $\eta_{j-1,\ell}$, the weights $w_{j,\ell}(u_k)$ for $k = 1, \ldots, N$ are given by

$$(2.8) \qquad w_{j,\ell}(u_k) = \frac{\Phi\left(-G_\ell(u_k)/\sigma_j\right)}{\Phi\left(-G_\ell(u_k)/\sigma_{j-1}\right)} \text{ for } j > 1,$$
$$w_{1,\ell}(u_k) = \Phi(-G_\ell(u_k)/\sigma_1).$$

To obtain an accurate estimator $\hat{S}_{j,\ell}$, the parameters σ_j are adaptively determined such that consecutive densities differ only slightly. This goal is achieved by requiring that the coefficient of variation of the weights $w_{j,\ell}$ is close to the target value δ_{target}, which is specified by the user. This leads to the following minimization problem:

$$(2.9) \qquad \sigma_j = \operatorname*{argmin}_{\sigma \in (0, \sigma_{j-1})} \left\| \delta_{w_{j,\ell}} - \delta_{\mathrm{target}} \right\|_2^2,$$

where $\delta_{w_{j,\ell}}$ is the *coefficient of variation* of the weights (2.8). This adaptive procedure is similar to the adaptive tempering in [7, 37] and is equivalent to requiring that the *effective sample size* takes a target value [37]. Note that the solution of the minimization problem in (2.9) does not require further evaluations of the LSF. Hence, its costs are negligible compared to the overall computational costs. The tempering iteration is finished if the coefficient of variation $\delta_{w_{\mathrm{opt},\ell}}$ of the weights with respect to

the optimal IS density

$$(2.10) \qquad w_{\mathrm{opt},\ell}(u_k) := I(G_\ell(u_k) \le 0) \frac{\varphi_n(u_k)}{\eta_{j,\ell}(u_k)}$$

is smaller than δ_{target} and, hence, the optimal IS density is approximated sufficiently well. According to [47], the SIS estimator of the probability of failure is defined as follows:

$$(2.11) \qquad \hat{P}_{f,\ell}^{\mathrm{SIS}} = \left(\prod_{j=1}^{N_T} \hat{S}_{j,\ell} \right) \frac{1}{N} \sum_{k=1}^{N} w_{\mathrm{opt},\ell}(u_k),$$

where the weights $w_{\mathrm{opt},\ell}$ are defined in (2.10) with $j = N_T$. The sum over the weights $w_{\mathrm{opt},\ell}(u_k)$ in (2.11) represents the last tempering step from the IS density $p_{N_T,\ell}$ to the optimal IS density $p_{\mathrm{opt},\ell}$ given in (2.5). It corresponds to the estimator of the ratio $P_{f,\ell}/P_{N_T,\ell}$ since $P_{f,\ell}$ is the normalizing constant of the optimal IS density.

During the iteration, MCMC sampling is applied to transfer samples distributed according to $p_{j-1,\ell}$ to samples distributed according to $p_{j,\ell}$ for $j = 1, \ldots, N_T$. Section 4 explains MCMC sampling in more detail. Algorithm 2.1 summarizes the procedure of one tempering step for sampling from $p_{j,\ell}$ and estimating $S_{j,\ell}$ starting from samples from $p_{j-1,\ell}$.

Algorithm 2.1 Tempering algorithm (N samples from $p_{j-1,\ell}$, σ_{j-1}, δ_{target}, G_ℓ).

1: determine σ_j from the optimization problem (2.9)
2: evaluate the weights $w_{j,\ell}$ as in (2.8) for the current set of samples
3: evaluate the estimator $\hat{S}_{j,\ell}$ as in (2.7)
4: resample the samples of $p_{j-1,\ell}$ based on their weights $w_{j,\ell}$
5: move the samples with MCMC to generate N samples from the density $p_{j,\ell}$
6: **return** N samples from $p_{j,\ell}$, σ_j, $\hat{S}_{j,\ell}$

Remark 2.1. We remark that nestedness, which is a prerequisite for SuS, is not an issue for SIS. This is because the intermediate sampling densities are smooth approximations of the optimal IS density and they all have supports in the whole outcome space. The proximity of two consecutive densities is ensured by (2.9). This property of SIS motivates the development of MLSIS in the following section. We note that MLSuS does not satisfy nestedness, which leads to the denominators in the estimator (2.4).

3. Multilevel sequential importance sampling. SIS and SuS have the drawback that all PDE solves are performed with the same discretization accuracy. This can lead to huge computational costs if the discretization level is high or the number of required tempering steps is large. Simply decreasing the level ℓ can lead to a bias in the estimated probability of failure, since the accuracy of the LSF decreases if the discretization level decreases. Therefore, the work in [37] develops the MLS^{2}MC method, where computations are performed on a sequence of increasing discretization levels while achieving an improvement in terms of computational costs. Originally, this method has been developed for *Bayesian inverse problems* [15]. In this section, we show how we can reformulate the MLS^{2}MC method as an MLSIS estimator for the probability of failure.

3.1. Bridging. Consider the sequence of discretization levels $\ell \in \{1, \ldots, L\}$, where $\ell = 1$, represents the smallest and $\ell = L \in \mathbb{N}$ the highest discretization level, i.e., finest element mesh size. Throughout this paper, it is assumed that the computational costs of evaluating G_ℓ are given by

$$
\text{(3.1)} \qquad \qquad \text{Cost}_\ell = \mathcal{O}(2^{-d(L-\ell)}),
$$

where $d \in \mathbb{N}$ is the dimension of the computational domain. In order to use a hierarchy of discretization levels, *bridging* is applied to transfer samples following a distribution on a coarse grid to samples following a distribution on the next finer grid. The level update is defined as proposed in [34]. The density $p_{j,\ell}$ of the coarse grid is transformed to the density $p_{j,\ell+1}$ of the next finer grid by the sequence

$$
\text{(3.2)} \qquad p_{j,\ell}^t(u) := \frac{1}{P_{j,\ell}^t} \Phi\left(-\frac{G_{\ell+1}(u)}{\sigma_j}\right)^{\beta_t} \Phi\left(-\frac{G_\ell(u)}{\sigma_j}\right)^{1-\beta_t} \varphi_n(u)
$$

for $t = 0, \ldots, N_{B_\ell}$, where $0 = \beta_0 < \beta_1 < \cdots < \beta_{N_{B_\ell}} = 1$, i.e., $p_{j,\ell}^0 = p_{j,\ell}$ and $p_{j,\ell}^{N_{B_\ell}} = p_{j,\ell+1}$. The number $N_{B_\ell} \in \mathbb{N}$ of intermediate bridging densities is a priori unknown. As in (2.6), the quantity $P_{j,\ell}^t$ in (3.2) can be calculated using samples distributed according to $p_{j,\ell}^{t-1}$. Similarly, the fraction of consecutive normalizing constants $S_{j,\ell}^t = P_{j,\ell}^t / P_{j,\ell}^{t-1}$ is estimated by

$$
\text{(3.3)} \qquad \hat{S}_{j,\ell}^t := \hat{\mathrm{E}}_{p_{j,\ell}^{t-1}}[w_{j,\ell}^t(u)] = \frac{1}{N} \sum_{k=1}^{N} w_{j,\ell}^t(u_k),
$$

where the samples $\{u_k\}_{k=1}^N$ are distributed according to $p_{j,\ell}^{t-1}$ and the weights are given by

$$
\text{(3.4)} \qquad w_{j,\ell}^t(u_k) := \frac{\Phi\left(-G_{\ell+1}(u_k)/\sigma_j\right)^{\beta_t} \Phi\left(-G_\ell(u_k)/\sigma_j\right)^{1-\beta_t}}{\Phi\left(-G_{\ell+1}(u_k)/\sigma_j\right)^{\beta_{t-1}} \Phi\left(-G_\ell(u_k)/\sigma_j\right)^{1-\beta_{t-1}}}
$$

for $k = 1, \ldots, N$. The bridging temperatures β_t are adaptively determined by solving the minimization problem

$$
\text{(3.5)} \qquad \beta_t = \operatorname*{argmin}_{\beta \in (\beta_{t-1}, 1]} \left\| \delta_{w_{j,\ell}^t} - \delta_{\text{target}} \right\|_2^2,
$$

where $\delta_{w_{j,\ell}^t}$ is the coefficient of variation of the weights. As in [37], we set the target coefficient of variation within the bridging steps to the same value as in the tempering steps. Within one level update, the bridging sequence is finished if $\beta_t = 1$ holds. Note that each level update requires a sequence of bridging densities and tempering is not performed during level updates. As in the tempering steps, MCMC sampling is applied to transfer samples between two consecutive bridging densities. By combining all estimators $\hat{S}$ of the tempering and bridging sequences given in (2.7) and (3.3), respectively, the MLSIS estimator for the probability of failure is given as

$$
\text{(3.6)} \qquad \hat{P}_f^{\text{MLSIS}} = \left(\prod_{j=1}^{N_T} \prod_{\ell=1}^{L} \prod_{t=1}^{N_{B_\ell}} \hat{S}_{j,\ell}^t \right) \frac{1}{N} \sum_{k=1}^{N} w_{\text{opt},L}(u_k),
$$

where the weights $w_{\text{opt},L}$ are defined in (2.10) with $j = N_T$ and represent the last tempering step from the IS density $p_{N_T,L}$ to the optimal IS density $p_{\text{opt},L}$ given in (2.5). Algorithm 3.1 summarizes the procedure of one level update.

Algorithm 3.1 Bridging algorithm (N samples from $p_{j,\ell}$, σ_j, δ_{target}, G_ℓ, $G_{\ell+1}$).

1: $t \leftarrow 0$
2: $\beta_t \leftarrow 0$
3: **while** $\beta_t < 1$ **do**
4:     $t \leftarrow t + 1$
5:     determine β_t from the optimization problem (3.5)
6:     evaluate the weights $w_{j,\ell}^t$ as in (3.4) for the current set of samples
7:     evaluate the estimator $\hat{S}_{j,\ell}^t$ as in (3.3)
8:     resample the samples of $p_{j,\ell}^{t-1}$ based on their weights $w_{j,\ell}^t$
9:     move the samples with MCMC to generate N samples from the density $p_{j,\ell}^t$
10: **end while**
11: **return** N samples from $p_{j,\ell+1}$, $\hat{S}_{j,\ell}^t$

3.2. Update scheme. The crucial part of the MLSIS method is to combine the adaptive tempering and bridging sequences and to provide a heuristic idea when to perform bridging or tempering. Initially, the samples $\{u_k\}_{k=1}^N$ are distributed according to the n-variate standard normal distribution φ_n, i.e., $\sigma_0 = \infty$. The LSF is evaluated on the smallest discretization level $\ell = 1$. Tempering is always performed in the first step in order to determine σ_1 to approximate the indicator function. The tempering finishes if the coefficient of variation $\delta_{w_{\text{opt},\ell}}$ of the weights with respect to the optimal IS density (2.10) is smaller than δ_{target}. The bridging finishes if the highest discretization level $\ell = L$ is reached. The combination of tempering and bridging determines the costs and accuracy of the method. The authors in [37] analyze the efficiency of different decision schemes, which leads to the following approach. The scheme should perform as many tempering steps as possible on small discretization levels while level updates are performed if the discrepancy between evaluations of two consecutive levels is too large. To measure this occurrence, a small subset of samples $\{u_{j_k}\}_{k=1}^{N_s}$ with $N_s < N$ is randomly selected without replacement. A level update is performed for this subset through one bridging step and the resulting coefficient of variation $\delta_{w^{N_s}}$ of the weights

$$w_{j,\ell}^{N_s}(u_{j_k}) = \frac{\Phi\left(-G_{\ell+1}(u_{j_k})/\sigma_j\right)}{\Phi\left(-G_\ell(u_{j_k})/\sigma_j\right)} \text{ for } k = 1, \ldots, N_s$$

is estimated. Depending on the estimated value $\delta_{w^{N_s}}$, two cases occur:
  (1) either $\delta_{w^{N_s}} > \delta_{\text{target}}$, bridging is performed since the accuracy is small, i.e., the difference between levels is high,
  (2) or $\delta_{w^{N_s}} \leq \delta_{\text{target}}$, tempering is performed since the accuracy is high, i.e., the difference between levels is small.
If case (1) occurs, the evaluations with respect to $G_{\ell+1}$ can be stored and reused in the bridging step and invested costs are not wasted. Whereas in case (2), these evaluations are no longer required and invested costs are wasted. Calculating $\delta_{w^{N_s}}$ for the sample subset is redundant if tempering has already finished. Then, bridging is always performed to reach the final discretization level. Moreover, as proposed in [37], tempering is performed after each level update, if the tempering has not already finished. In this case, calculating $\delta_{w^{N_s}}$ is redundant, too. Note that $\delta_{w_{\text{opt},\ell}}$ has to be calculated after each tempering and level update, to decide if tempering is finished. Finally, the MLSIS method is finished if both tempering and bridging are finished. The procedure is described in Algorithm 3.2.

Algorithm 3.2 MLSIS algorithm (N, n, L, δ_{target}, N_s, G_ℓ).

1: Generate N samples from the n-variate standard normal distribution φ_n
2: $\ell \leftarrow 1$
3: Perform Tempering
4: **while** Tempering is not finished **or** Bridging is not finished **do**
5:    **if** Tempering is finished **then**
6:       Perform Bridging
7:       $\ell \leftarrow \ell + 1$
8:    **else if** Bridging is finished **or** last step was a Bridging step **then**
9:       Perform Tempering
10:    **else**
11:       Perform Bridging in one step with a random subset of N_s samples
12:       Calculate $\delta_{w^{N_s}}$
13:       **if** $\delta_{w^{N_s}} < \delta_{\text{target}}$ **then**
14:          Perform Tempering
15:       **else**
16:          Perform Bridging
17:          $\ell \leftarrow \ell + 1$
18:       **end if**
19:    **end if**
20:    Calculate $\delta_{w_{\text{opt},\ell}}$
21:    **if** $\delta_{w_{\text{opt},\ell}} \leq \delta_{\text{target}}$ **then**
22:       Tempering is finished
23:    **end if**
24:    **if** $\ell = L$ **then**
25:       Bridging is finished
26:    **end if**
27: **end while**
28: **return** Probability of failure estimate

Remark 3.1. We note that, according to [37], the finest discretization level L can be chosen adaptively based on the coefficient of variation $\delta_{w^{N_s}}$ between two consecutive discretization levels. Bridging is finished if $\delta_{w^{N_s}}$ is smaller than a given bound which is much smaller than δ_{target}.

3.3. Level dependent dimension. As mentioned in section 2.2, the nestedness problem of MLSuS motivates [53] to study a level dependent parameter dimension. This approach can also be applied in MLSIS to reduce variances between level updates and, hence, increase the number of tempering updates on coarse grids. For this purpose, it is assumed that the LSF G depends on a random field that is approximated by a truncated *Karhunen-Loève* (KL) expansion. This setting occurs in many relevant applications as well as in numerical experiments presented in section 5. Since high order KL terms are highly oscillating, they cannot be accurately discretized on coarse grids, which leads to noisy evaluations and higher variances. By reducing the number of KL terms on coarse grids, the variance between consecutive LSF evaluations is reduced. Therefore, the coefficient of variation $\delta_{w^{N_s}}$ is smaller and case (2) in section 3.2 is more likely. Hence, more tempering steps are performed on small discretization levels, which decreases the computational costs for MLSIS.

4. Markov chain Monte Carlo. The goal of SIS and MLSIS is to transform samples from the prior density $p_0 = \varphi_n$ into samples of the optimal IS density p_{opt}. Thereby, a sequence of densities is defined which converges to the optimal one. MCMC is applied to transform samples into consecutive densities of the tempering and bridging steps.

Consider the tempering step from $p_{j-1,\ell}$ to $p_{j,\ell}$. The samples $\{u_k\}_{k=1}^N$ are distributed as $p_{j-1,\ell}$ and have to be transformed into samples that are distributed as $p_{j,\ell}$. To define the number of seeds of the MCMC algorithm, we choose a parameter

$$(4.1) \qquad c \in (0,1] \text{ such that } \frac{1}{c} \in \mathbb{N} \text{ and } c \cdot N \in \mathbb{N}.$$

Then, $N_c := c \cdot N$ seeds are randomly selected with replacement from the set $\{u_k\}_{k=1}^N$ according to their weights $\{w_{j,\ell}(u_k)\}_{k=1}^N$ given in (2.8). The set of seeds is denoted by $\{u_{k_j}\}_{j=1}^{N_c}$. In this procedure, which corresponds to *multinomial resampling*, samples with high weights are copied multiple times and samples with low weights are discarded. There are also other resampling methods, such as *stratified resampling* or *systematic resampling*, which can be applied. A study on their convergence behavior is given in [22]. The burn-in length is denoted by $N_b \in \mathbb{N}$. Starting with the seed $u_0 \in \{u_{k_j}\}_{j=1}^{N_c}$, a Markov chain of length $N_b + 1/c$ is simulated that has $p_{j,\ell}$ as its stationary distribution. The first N_b states are rejected after the simulation. Algorithm 4.1 states the MCMC procedure that employs the *Metropolis–Hastings* sampler [28, 40]. During the algorithm, a *proposal* $\bar{u}$ is generated according to the *proposal density q*. Moreover, the acceptance function $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ is given by

$$\alpha_T(u_0, \bar{u}) := \frac{\Phi\left(-G(\bar{u})/\sigma_j\right) \varphi_n(\bar{u}) q(u_0 \mid \bar{u})}{\Phi\left(-G(u_0)/\sigma_j\right) \varphi_n(u_0) q(\bar{u} \mid u_0)},$$

which is the ratio of the target density $p_{j,\ell}$ with respect to the current state of the chain u_0 and candidate $\bar{u}$. For a bridging step, the seeds are selected from samples distributed according to $p_{j,\ell}^t$ and the target density is $p_{j,\ell}^{t+1}$. The weights are given by $\{w_{j,\ell}^{t+1}(u_k)\}_{k=1}^N$ (see (3.4)), and the acceptance function α must be replaced by

$$\alpha_B(u_0, \bar{u}) = \frac{\Phi\left(-G_{\ell+1}(\bar{u})/\sigma_j\right)^{\beta_{t+1}} \Phi\left(-G_\ell(\bar{u})/\sigma_j\right)^{1-\beta_{t+1}} \varphi_n(\bar{u}) q(u_0 \mid \bar{u})}{\Phi\left(-G_{\ell+1}(u_0)/\sigma_j\right)^{\beta_{t+1}} \Phi\left(-G_\ell(u_0)/\sigma_j\right)^{1-\beta_{t+1}} \varphi_n(u_0) q(\bar{u} \mid u_0)}.$$

Remark 4.1. Since consecutive densities within SIS and MLSIS are constructed in a way that they are not too different and samples are weighted according to the target distribution, the burn-in length can be small or even negligible within SIS and MLSIS [47]. Note that for SuS and MLSuS, the $N \cdot \hat{p}_0$ samples with the lowest LSF values are selected as seeds.

4.1. Adaptive conditional sampling. The random walk Metropolis–Hastings algorithm [28, 40] is a classical MCMC algorithm. However, random walk samplers suffer from the curse of dimensionality, i.e., the acceptance rate is small in high-dimensions (see [44]). Since high-dimensional parameter spaces are considered in the numerical experiments in section 5, *adaptive conditional sampling* (aCS) is proposed, where the chain correlation is adapted to ensure a high acceptance rate. aCS is a dependent MCMC algorithm, i.e., the proposal density depends on the current seed u_0. More formally, the proposal q is defined as the conditional multivariate normal density with mean vector ρu_0 and covariance matrix $\Sigma = (1 - \rho^2)I_n$ with I_n denoting

---

Algorithm 4.1 MCMC algorithm $(u_0, q(\cdot \mid \cdot), \alpha(\cdot, \cdot), c, N_b)$.

1: Chain $= \emptyset$
2: **while** $i \leq N_b + 1/c$ **do**
3:     Generate a candidate $\bar{u}$ from the proposal density $q(\cdot \mid u_0)$
4:     Evaluate $\alpha(u_0, \bar{u})$
5:     Accept the candidate $\bar{u}$ with probability $\min\{1, \alpha(u_0, \bar{u})\}$
6:     **if** $\bar{u}$ is accepted **then**
7:         $u_0 \leftarrow \bar{u}$
8:     **end if**
9:     Chain $\leftarrow$ Chain $\cup\, u_0$
10:    $i \leftarrow i + 1$
11: **end while**
12: Discard the first N_b elements of Chain
13: **return** Chain

---

the identity matrix. During the iterations, $\rho \in [0, 1]$ is adaptively adjusted such that the acceptance rate is around 44% [51]. This value leads to an optimal value in terms of the minimum autocorrelation criterion. By the structure of the proposal, the acceptance functions read as

$$\alpha_T(u_0, \bar{u}) = \frac{\Phi\left(-G(\bar{u})/\sigma_j\right)}{\Phi\left(-G(u_0)/\sigma_j\right)},$$

$$\alpha_B(u_0, \bar{u}) = \frac{\Phi\left(-G_{\ell+1}(\bar{u})/\sigma_j\right)^{\beta_{t+1}} \Phi\left(-G_\ell(\bar{u})/\sigma_j\right)^{1-\beta_{t+1}}}{\Phi\left(-G_{\ell+1}(u_0)/\sigma_j\right)^{\beta_{t+1}} \Phi\left(-G_\ell(u_0)/\sigma_j\right)^{1-\beta_{t+1}}}.$$

A more detailed description of the algorithm with the adaptive adjustment of the correlation parameter is given in [44]. The aCS algorithm can be viewed as an adaptive version of the *preconditioned Crank–Nicolson* sampler [13] tailored to application within SIS.

4.2. Independent sampler with von Mises–Fisher Nakagami proposal distribution. Since aCS is a dependent MCMC algorithm, the states of the chains are correlated, which can lead to a higher variance of the estimated ratio of normalizing constants $\hat{S}_{j,\ell}$ and $\hat{S}_{j,\ell}^t$ given in (2.7) and (3.3), respectively. Hence, this leads to a higher variance of the estimated probability of failure (3.6). An independent MCMC algorithm overcomes this problem through using a proposal density that does not depend on the current state. The dependence on the current state enters in the acceptance probability. If the proposal density is chosen close to the target density, the acceptance probability will be close to one and the samples will be approximately independent. In the context of SIS and MLSIS, the available samples and corresponding weights of each previous density can be used to fit a distribution model to be used as the proposal density in the MCMC step [11, 47]. For instance, *Gaussian mixture* models can be used as a proposal density [47]. A drawback of Gaussian densities in high dimensions is the concentration of measure around the hypersphere with norm equal to $\sqrt{n}$ (see [32, 46]). Therefore, only the direction of the samples is of importance. Furthermore, the Gaussian mixture model with K densities has $Kn(n+3)/2 + (K-1)$ parameters, which have to be estimated. Both issues motivate the vMFN distribution. Therein, the direction is sampled from the *von Mises–Fisher* (vMF) distribution [55] while the radius is sampled from the *Nakagami* distribution
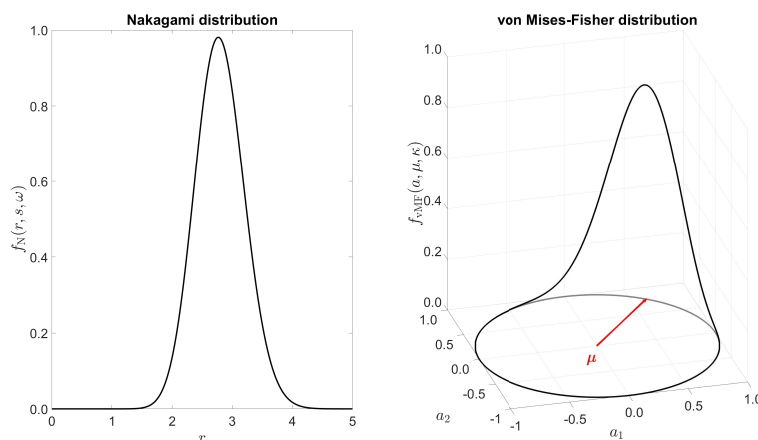
FIG. 2. *Illustration of the Nakagami distribution (left) and von Mises–Fisher distribution (right) in two dimensions. The parameters are defined as $\nu = (0.6, 0.75)^T$, $\kappa = 11$, $s = 12$, and $\gamma = 8$.*

(see [42]). The vMFN mixture model has only $K(n+3) + (K-1)$ parameters, which scales linearly in the dimension n. Note that for the Gaussian mixture, the number of parameters of the distribution model scales quadratically in the dimension n. To apply the vMFN distribution as the proposal density in Algorithm 4.1, the parameters of the distribution model have to be fitted in advance.

It is assumed that all samples $u \in \mathbb{R}^n$ are given in their polar coordinate representation $u = r \cdot a$, where $r = \|u\|_2 \in \mathbb{R}_+$ is the norm of u and $a = u/\|u\|_2 \in \mathbb{R}^n$ its direction. For $u = r \cdot a$ the vMFN distribution is defined as the product of the von Mises–Fisher and the Nakagami distribution, that is,

$$f_{\text{vMFN}}(r, a \mid \nu, \kappa, s, \gamma) = f_{\text{N}}(r \mid s, \gamma) \cdot f_{\text{vMF}}(a \mid \nu, \kappa).$$

The vMF distribution f_{vMF} defines the distribution of the direction on the n-dimensional hypersphere $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ and is given by

$$f_{\text{vMF}}(a \mid \nu, \kappa) = \frac{\kappa^{n/2-1}}{(2\pi)^{n/2} \mathcal{I}_{n/2-1}(\kappa)} \exp(\kappa \nu^T a),$$

where $\nu \in \mathbb{S}^{n-1}$ is a mean direction and $\kappa \geq 0$ characterizes the concentration around ν. $\mathcal{I}_{n/2-1}$ denotes the *modified Bessel function* of the first kind and order $n/2 - 1$ [1, Chapter 9]. On the contrary, the Nakagami distribution f_{N} specifies the distribution of the radius and is defined by

$$f_{\text{N}}(r \mid s, \gamma) := \frac{2s^s}{\Gamma(s)\gamma^s} r^{2s-1} \exp\left(-\frac{s}{\gamma} r^2\right),$$

where $\Gamma(s)$ is the *Gamma function*, $s \geq 0.5$ is a shape parameter, and $\gamma > 0$ a spread parameter. Figure 2 shows an illustration of f_{N} and f_{vMF} for certain parameter values.

Remark 4.2. We have defined the vMFN distribution for the radius and direction (r, a) on $[0, \infty) \times \mathbb{S}^{n-1}$. However, we actually approximate a distribution on $\mathbb{R}^n$, which defines the distribution of $u = r \cdot a \in \mathbb{R}^n$. By [33, Theorem 1.101], the distribution of

$u \in \mathbb{R}^n$, which is the product distribution of r and a, is given as

$$f_U(u) = \int_0^\infty f_N(r \mid s, \gamma) f_{vMF}\left(\frac{u}{r}\Big|\nu, \kappa\right)\frac{1}{r^n}\mathrm{d}r,$$

$$\propto \int_0^\infty r^{2s-1-n}\exp\left(-\frac{s}{\gamma}r^2 + \frac{\kappa\nu^T u}{r}\right)\mathrm{d}r.$$

Since we can easily separate u into r and a, we usually work with f_{vMFN} rather than f_U.

To define the vMFN distribution as a proposal density for Algorithm 4.1, the parameters ν, κ, s, and γ have to be fitted using the current set of samples $\{u_k = r_k \cdot a_k\}_{k=1}^N$ and their weights $\{w_k\}_{k=1}^N$, which are given by (2.8) or (3.4) for a tempering or bridging step, respectively. The parameters are determined by maximizing the weighted log-likelihood

$$\max_{\nu,\kappa,s,\gamma} \sum_{k=1}^N w_k \ln(f_{vMFN}(r_k, a_k \mid \nu, \kappa, s, \gamma)).$$

Differentiating this expression with respect to the parameters and setting the derivatives equal to zero yields the optimal parameters for the fitting [46]. However, for the concentration κ and shape parameter s we use an approximation since the derivatives require the solutions of nonlinear equations which arise from the Gamma function and modified Bessel function [9, 55]. The fitted mean direction $\hat{\nu}$ and concentration $\hat{\kappa}$ are given by

(4.2)

$$\hat{\nu} = \frac{\sum_{k=1}^N w_k \cdot a_k}{\|\sum_{k=1}^N w_k \cdot a_k\|_2}, \qquad \hat{\kappa} = \frac{\chi \cdot n - \chi^3}{1 - \chi^2}, \text{ where } \chi = \min\left\{\frac{\|\sum_{k=1}^N w_k \cdot a_k\|_2}{\sum_{k=1}^N w_k}, 0.95\right\}.$$

The upper bound of 0.95 in (4.2) is chosen to ensure numerical stability of the algorithm. If χ converges to 1, the vMFN distribution would converge to a point density [46]. Moreover for the Nakagami distribution, the fitted spread $\hat{\gamma}$ and shape parameter $\hat{s}$ are given by

$$\hat{\gamma} = \frac{\sum_{k=1}^N w_k \cdot r_k^2}{\sum_{k=1}^N w_k}, \qquad \hat{s} = \frac{\hat{\gamma}^2}{\nu_4 - \hat{\gamma}^2}, \text{ where } \nu_4 = \frac{\sum_{k=1}^N w_k \cdot r_k^4}{\sum_{k=1}^N w_k}.$$

To apply Algorithm 4.1 with respect to the vMFN distribution, the proposal $q(\cdot \mid u_0)$ is replaced by $f_{vMFN}(\cdot, \cdot, \hat{\nu}, \hat{\kappa}, \hat{s}, \hat{\gamma})$ with the fitted parameters.

Remark 4.3. If a mixture of vMFN distributions is considered with $K > 1$ individual vMFN densities, the vMFN mixture distribution reads as

$$f_{vMFNM}(r, a \mid \boldsymbol{\nu}, \boldsymbol{\kappa}, \boldsymbol{s}, \boldsymbol{\gamma}) = \sum_{j=1}^K \pi_j f_{vMFN}(r, a \mid \nu_j, \kappa_j, s_j, \gamma_j),$$

where the weights π_j represent the probability of each mode and $\sum_{j=1}^K \pi_j = 1$. In this case, the assignments of the samples to the modes is unknown and this assignment has to be estimated in addition. Therefore, the required parameters cannot be estimated in one iteration. For instance, the *Expectation-Maximization* algorithm [38] can be applied to estimate the parameters iteratively. The resulting formulas are given in

[46]. The usage of mixtures is motivated by multimodal failure domains. In the numerical experiments in section 5 only $K = 1$ is considered.

4.3. MCMC for a level dependent dimension. In the case that MLSIS or MLSuS are applied with a level dependent parameter dimension, the procedure of a level update has to be adjusted. Consider the level update from level ℓ to $\ell + 1$ and assume that the corresponding LSFs are defined as $G_\ell : \mathbb{R}^{n_\ell} \to \mathbb{R}$ and $G_{\ell+1} : \mathbb{R}^{n_{\ell+1}} \to \mathbb{R}$, respectively, where $n_\ell < n_{\ell+1}$. Before the first MCMC step of the level update is carried out, the weights $w_{j,\ell}^1(u_k)$ for $k = 1, \ldots, N$ (see (3.4)) have to be evaluated. However, these evaluations require the evaluation of $G_{\ell+1}$ with respect to the current samples $\{u_k\}_{k=1}^N$, which are defined on $\mathbb{R}^{n_\ell}$. In the beginning of MLSIS or MLSuS the samples u_k are initialized from the standard normal density φ_{n_1}. Therefore, it is natural to sample the missing dimensions $\Delta n_{\ell+1} = n_{\ell+1} - n_\ell$ from the standard normal distribution $\varphi_{\Delta n_{\ell+1}}$. Hence, for each $k = 1, \ldots, N$ we sample $\Delta n_{\ell+1}$ independent standard normal random variables $\psi_k \in \mathbb{R}^{\Delta n_{\ell+1}}$ and stack u_k and ψ_k together, i.e., $\tilde{u}_k = [u_k, \psi_k] \in \mathbb{R}^{n_{\ell+1}}$. In order to evaluate the weights $w_{j,\ell}^1(u_k)$, the LSF $G_{\ell+1}$ is evaluated for $\tilde{u}_k$ and G_ℓ for u_k. The seeds for the MCMC step are chosen based on these weights. Subsequently, Algorithm 4.1 is performed. Within the MCMC algorithm, a proposal $\bar{u} \in \mathbb{R}^{n_{\ell+1}}$ is sampled from $q(\cdot \mid u_0)$, which is suitable for the evaluations of $G_{\ell+1}$. For the LSF G_ℓ the first n_ℓ entries of $\bar{u}$ are taken as input.

5. Numerical experiments. In the following examples, all probability of failure estimates are obtained with respect to the same, finest discretization level, i.e., the multilevel methods iterate until this level is reached and the single-level methods are based on this level. Therefore, the obtained errors involve only sampling errors while discretization errors are not included.

5.1. 1D diffusion equation. We begin with Example 2 in [53], which considers the diffusion equation in the one-dimensional domain $D = [0, 1]$. In particular, the corresponding stochastic differential equation is given by

$$(5.1) \qquad -\frac{\partial}{\partial x}\left(a(x, \omega) \frac{\partial}{\partial x} v(x, \omega) \right) = 1 \text{ for } 0 \leq x \leq 1,$$

$$\text{such that } v(0, \omega) = 0 \text{ and } v'(1, \omega) = 0$$

for almost every (a.e.) $\omega \in \Omega$. Failure is defined as the event that the solution v is larger than 0.535 at $x = 1$, i.e., $G(\omega) := 0.535 - v(1, \omega) \leq 0$. The solution v is approximated by a piecewise linear, continuous FEM approximation v_h on a uniform grid with mesh size $h > 0$. Hence, the approximated LSF is given by $G_\ell(\omega) = 0.535 - v_{h_\ell}(1, \omega)$, where $\ell \in \mathbb{N}$ defines the discretization level. By crude Monte Carlo sampling (2.2) with $N = 10^7$ samples on a grid with mesh size $h = 1/512$, the probability of failure is estimated to be $P_f = 1.524 \cdot 10^{-4}$. In the following, this value is referred to as the reference solution. Figure 3 shows the mean of 10^5 realizations of solutions $v_h(\cdot, \omega)$ plus/minus the standard deviation for $h = 1/512$. Additionally, the respective histogram of their LSF values is presented. We see that very few realizations are larger than 0.535 at $x = 1$.

The coefficient function $a(x, \omega) = \exp(Z(x, \omega))$ in (5.1) is a log-normal random field with constant mean function $\mathbb{E}[a(x, \cdot)] = 1$ and standard deviation $\text{Std}[a(x, \cdot)] = 0.1$. That is, Z is a Gaussian random field with constant mean function $\mu_Z = \log(\mathbb{E}[a(x, \cdot)]) - \zeta_Z^2/2$ and variance $\zeta_Z^2 = \log\left((\text{Std}[a(x, \cdot)]^2 + \mathbb{E}[a(x, \cdot)]^2)/\mathbb{E}[a(x, \cdot)]^2\right)$.
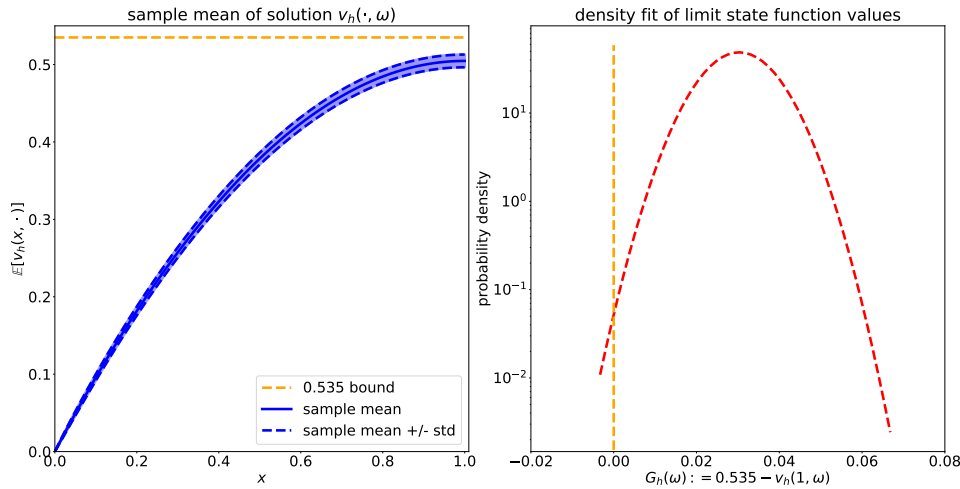
FIG. 3. *Mean over 10^5 realizations of solutions $v_h(\cdot, \omega)$ plus/minus the standard deviation (left). Fit of the respective LSF values $G_h(\omega)$ for $h = 1/512$ (right). Note that the probabilities are shown on a log-scale.*

Moreover, Z has an exponential type covariance function which is given by $c(x, y) = \zeta_Z^2 \exp(-|x - y|/\lambda)$, where $\lambda = 0.01$ denotes the correlation length. The infinite-dimensional log-normal random field a is discretized by the truncated KL expansion, of Z

$$Z(x, \omega) = \mu_Z + \zeta_Z \sum_{m=1}^{M} \sqrt{\nu_m} \theta_m(x) U_m(\omega),$$

where (ν_m, θ_m) are the KL eigenpairs and $\{U_m\}_{m=1}^{M}$ are independent standard normal Gaussian random variables. The eigenpairs can be analytically calculated as explained in [24, Section 2.3.3].

The probability of failure is estimated by SIS, MLSIS, SuS, and MLSuS. For all methods, the estimation is performed for $N = 250, 500, 1000, 2000$ samples and $N_s = 0.1 \cdot N$ samples are considered for the small sample subset to decide if either bridging or tempering is performed in the update scheme of section 3.2. For each parameter setting, the estimation is repeated 100 times. For the multilevel methods, the sequence of mesh sizes is $h_\ell = 2^{-\ell-1}$ for $\ell = 1, \ldots, 8$, i.e., the coarsest mesh size is $h_1 = 1/4$ and the finest $h_8 = 1/512$. If a level dependent dimension is considered, the parameter dimensions of the KL expansions are $n_1 = 10$, $n_2 = 20$, $n_3 = 40$, $n_4 = 80$, and $n_5 = n_6 = n_7 = n_8 = 150$ as proposed in [53]. For a fixed parameter dimension, the dimension is $n = 150$ for all discretization levels. This captures 87% of the variability of $\log(a)$ [53]. SIS and MLSIS are performed for target coefficient of variations $\delta_{\text{target}} = 0.25$ and $\delta_{\text{target}} = 0.50$, which is considered in (2.9), (3.5). aCS and the independent sampler with the vMFN distribution and one mixture term are considered as the MCMC methods without a burn-in. The parameter c to define the number of seeds of the MCMC algorithm in (4.1) is $c = 0.1$ or $c = 1$. For SuS and MLSuS, aCS is considered as the MCMC method without a burn-in and the parameter $\hat{p}_0$ in (2.3) is $\hat{p}_0 = 0.1$ or $\hat{p}_0 = 0.25$.

5.1.1. Results. Figure 4 shows the estimated mean probability of failure by SIS and MLSIS plus/minus its standard deviation. The estimates of the means are in accordance with the reference solution for all settings. As expected, the bias and
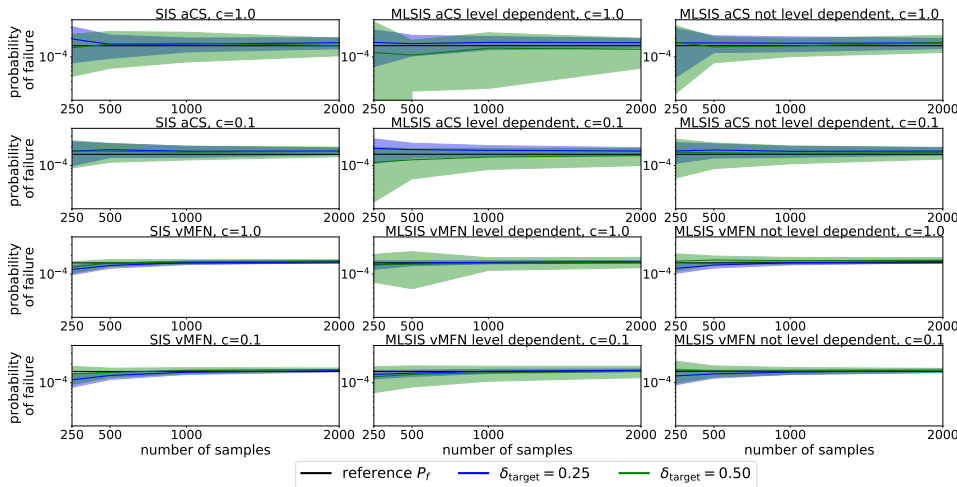
FIG. 4. *Estimated probability of failure by SIS and MLSIS averaged over* 100 *runs for* 250, 500, 1000, *and* 2000 *samples and* $\delta_{\text{target}} \in \{0.25, 0.50\}$. *The colored areas show the standard deviation of the estimates. The black lines show the reference estimate by Monte Carlo sampling. First row: aCS,* $c = 1.0$; *second row: aCS,* $c = 0.1$; *third row: vMFN,* $c = 1.0$; *fourth row: vMFN,* $c = 0.1$; *first column: SIS; second column: MLSIS with level dependent dimension; third column: MLSIS without level dependent dimension.*

standard deviation decrease with an increasing number of samples. Furthermore, the standard deviation is smaller for a smaller target coefficient of variation. We observe that sampling from the vMFN distribution with independent MCMC yields a smaller bias and smaller standard deviation than applying aCS. Additionally, we observe that $c = 0.1$ yields also a smaller standard deviation than $c = 1$. Comparing the MLSIS results with the SIS results for $\delta_{\text{target}} = 0.50$, we see that SIS reaches a smaller standard deviation. For $\delta_{\text{target}} = 0.25$ the results are similar. For MLSIS and $\delta_{\text{target}} = 0.50$, a level dependent parameter dimension leads to a higher standard deviation than a fixed parameter dimension. However, for $\delta_{\text{target}} = 0.25$, the results are similar. We summarize that $\delta_{\text{target}} = 0.25$ yields for all settings a similar bias and standard deviation. Only the MCMC algorithm has a larger influence on the standard deviation in this setting.

Figure 5 shows the *relative root mean square error* (RMSE) on the horizontal axis and the computational costs on the vertical axis of the SIS and MLSIS estimators. The relative RMSE is defined as

$$\text{relRMSE} := \frac{\left(\mathbb{E}\left[\left(\hat{P}_f - P_f \right)^2 \right] \right)^{\frac{1}{2}}}{P_f},$$

where $\hat{P}_f$ denotes the estimated probability of failure. The costs are calculated based on the formula given in (3.1) for $L = 8$ and $d = 1$. SIS and MLSIS yield a similar range of the relative RMSE but the computational costs are lower for MLSIS. Comparing the computational costs shown in Figure 5, we can save around 61% of the computational costs if we apply MLSIS for the estimation. This shows the achievement of the main goal of the MLSIS algorithm, that is, to save computational costs by employing a hierarchy of discretization levels. Furthermore, we observe that sampling from the
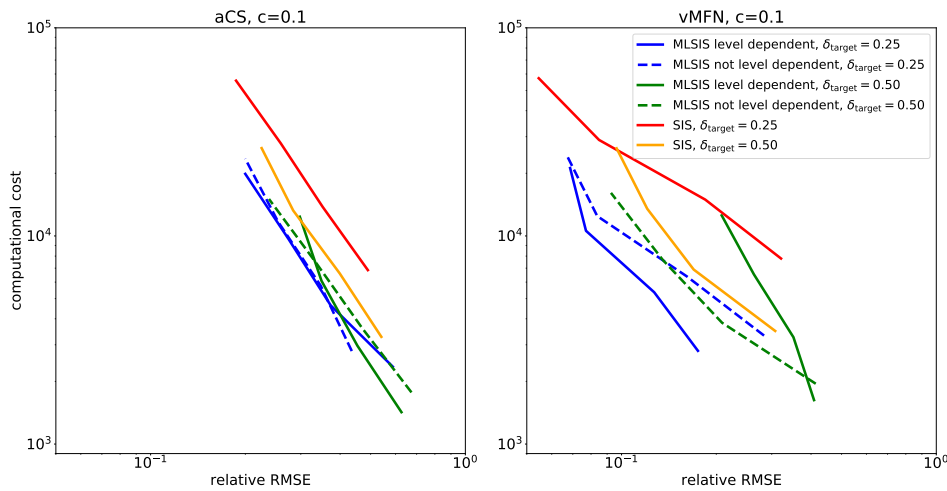
FIG. 5. *Computational costs and relative RMSE of SIS and MLSIS averaged over* 100 *runs for* 250, 500, 1000, *and* 2000 *samples and* $\delta_{target} \in \{0.25, 0.50\}$. *First column: aCS,* $c = 0.1$; *second column: vMFN,* $c = 0.1$.

vMFN distribution yields a lower relative RMSE than applying aCS. In the case of $\delta_{target} = 0.25$, a level dependent dimension yields a smaller relative RMSE and lower computational costs than a fixed parameter dimension. This was expected since variances between level updates are smaller and, therefore, more tempering steps are performed on coarse levels. However, MLSIS with a level dependent dimension, $\delta_{target} = 0.50$, and sampling from the vMFN distribution yields a higher relative RMSE than applying a fixed parameter dimension for the same computational cost.

Figure 6 shows the relative RMSE and computational costs of SIS, MLSIS, SuS, and MLSuS. We observe that SuS yields the same relative RMSE as SIS with aCS. However, SuS requires less computational costs. If we consider SIS with vMFN, the relative RMSE is smaller compared to SuS but the computational costs are higher for SIS. For the multilevel methods with a level dependent dimension, we observe that MLSuS and MLSIS with aCS yield a similar relative RMSE but MLSIS requires more computational costs. However, the savings with MLSuS are smaller compared to the single-level estimators. MLSIS with vMFN and $\delta_{target} = 0.25$ yields a much smaller relative RMSE than all other estimators and computational costs can be saved compared to MLSuS. Theses results are similar to the multilevel methods without a level dependent dimension. In this case, we can observe that MLSuS with $\hat{p}_0 = 0.1$ yields a large relative RMSE which is due to the nestedness issue of MLSuS.

5.2. 2D flow cell. We consider the 2D application in [53, section 6.1], which is a simplified setting of the rare event arising in planning a radioactive waste repository (see section 1). The probability of failure is based on the travel time of a particle within a 2D flow cell. Therein, the following PDE system has to be satisfied in the unit square domain $D = (0, 1) \times (0, 1)$:

$$q(x, \omega) = -a(x, \omega)\nabla v(x, \omega) \text{ for } x \in D,$$
$$\nabla \cdot q(x, \omega) = 0 \text{ for } x \in D$$

for a.e. $\omega \in \Omega$, where q is the Darcy velocity, v is the hydrostatic pressure, and a is the permeability of the porous medium, which is modelled as a log-normal random field.
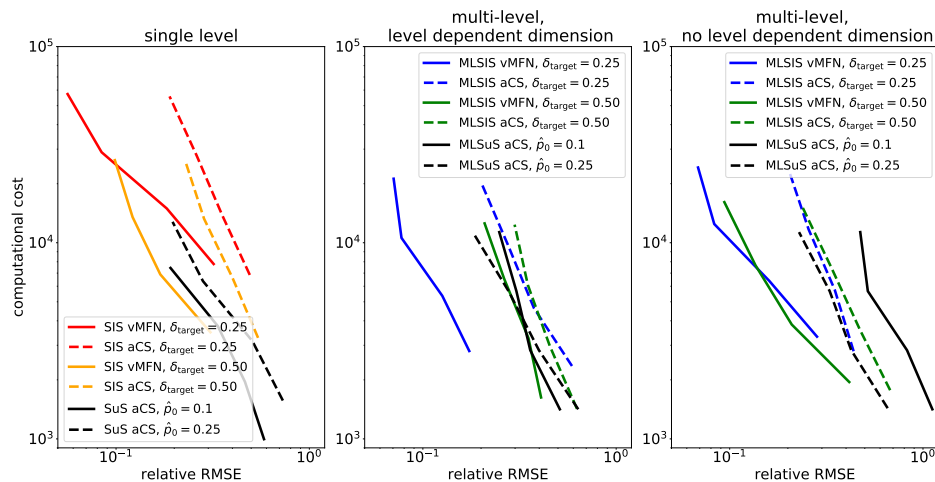
FIG. 6. *Computational costs and relative RMSE of SIS, MLSIS, SuS, and MLSuS averaged over 100 runs for 250, 500, 1000, and 2000 samples. SIS and MLSIS are considered with aCS and vMFN, c = 0.1, and $\delta_{\mathrm{target}} \in \{0.25, 0.50\}$. SuS and MLSuS are considered with aCS and $\hat{p}_0 = 0.1$ or $\hat{p}_0 = 0.25$. First column: single level; second column: multilevel with level dependent dimension; third column: multilevel without level dependent dimension.*

More precisely, $\log(a)$ is a Gaussian random field with mean $\mu_Z = 0$ and constant variance $\zeta_Z^2 = 1$. Moreover, Z has an exponential type covariance function

$$c(x, y) = \zeta_Z^2 \exp\left(-\frac{\|x - y\|_1}{\lambda}\right),$$

where $\lambda = 0.5$ denotes the correlation length. Again, the random field Z is discretized by its KL expansion. The PDE system is coupled with the following boundary conditions:

$$(5.2) \qquad\qquad \nu \cdot q(x, \omega) = 0 \text{ for } x \in (0, 1) \times \{0, 1\},$$

$$(5.3) \qquad\qquad v(x, \omega) = 1 \text{ for } x \in \{0\} \times (0, 1),$$

$$(5.4) \qquad\qquad v(x, \omega) = 0 \text{ for } x \in \{1\} \times (0, 1)$$

for a.e. $\omega \in \Omega$, where ν denotes the derivative with respect to the normal direction on the boundary. Equation (5.2) imposes that there is no flow across the horizontal boundaries and (5.3), (5.4) impose that there is inflow at the western boundary and outflow at the eastern boundary, respectively. The Darcy velocity q is discretized by lowest order Raviart–Thomas mixed FEs (see [50]). The pressure v is discretized by piecewise constant elements. The grid is determined by the mesh size h and consists of $2 \cdot 1/h^2$ uniform triangles.

The failure event is based on the time that a particle requires to travel from the initial point $x_0 = (0, 0.5)^T$ to any other point on the boundary ∂D. Given the Darcy velocity $q_{h_\ell}(x, \omega)$, the particle path $x(t, \omega)$ has to satisfy the following ordinary differential equation:

$$\frac{\partial}{\partial t} x(t, \omega) = q_{h_\ell}(x(t, \omega), \omega), \ x(0, \omega) = x_0.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We approximate the particle path with the forward Euler discretization

$$x_{h_\ell}(t + \Delta t, \omega) = x(t, \omega) + \Delta t q_{h_\ell}(x(t, \omega), \omega), \text{ where } \Delta t = \frac{h_\ell}{2\|q_{h_\ell}(x(t, \omega), \omega)\|_2}.$$

The travel time $\tau_{h_\ell}(\omega) \in [0, \infty)$ is defined as

$$\tau_{h_\ell}(\omega) = \operatorname*{argmin}_{t>0} x_{h_\ell}(t, \omega) \in \partial D.$$

The approximation of the particle path is different to the procedure in [53] and, therefore, the estimated probability of failures differs slightly. Failure is defined as the event that τ_{h_ℓ} is smaller than the threshold $\tau_0 = 0.03$. Hence, the respective LSF is defined as $G_\ell(\omega) := \tau_{h_\ell}(\omega) - \tau_0$. The reference solution of the probability of failure is $4.6730 \cdot 10^{-7}$ and is the estimated mean probability of failure over 100 realizations of SuS with $N = 10^4$ samples, mesh size $h = 1/128$, $\hat{p}_0 = 0.1$, and aCS as the MCMC method without burn-in. We note that SuS is a biased estimator and the relative bias scales as $\mathcal{O}(1/N)$ while the coefficient of variation scales as $\mathcal{O}(1/\sqrt{N})$ [4]. The coefficient of variation of the 100 probability of failure estimates is roughly 15%. Hence, we expect that the relative bias of the reference estimate is of order 10^{-2}.

Figure 7 shows a realization of a nonfailure event and of a failure event. The figure displays the permeability a and the respective solutions of the Darcy velocity q_h for $h = 1/128$ and shows the particle paths which start at x_0 and their respective travel times.

The probability of failure is estimated by SIS, MLSIS, SuS, and MLSuS. For all methods, the estimation is performed for $N = 250, 500, 1000$ samples and $N_s = 0.1 \cdot N$ samples are considered for the small sample subset to decide if either bridging or tempering is performed in the update scheme of section 3.2. For each parameter setting, the estimation is repeated 100 times. For the multilevel methods, the sequence of mesh sizes is $h_\ell = 2^{-\ell-1}$ for $\ell = 1, \ldots, 6$, i.e., the coarsest mesh size is $h_1 = 1/4$ and the finest $h_6 = 1/128$. The multilevel methods are applied with a level dependent dimension, where the parameter dimensions of the KL expansions are $n_1 = 10$, $n_2 = 20$, $n_3 = 40$, $n_4 = 80$, and $n_5 = n_6 = 150$. SIS and MLSIS are performed for target coefficient of variations equal to 0.50 and 1.00. aCS and the vMFN distributions are considered as the MCMC methods without a burn-in. The parameter c to define the number of seeds of the MCMC algorithm in (4.1) is $c = 0.1$. For SuS and MLSuS, aCS is considered as the MCMC method without a burn-in and the parameter $\hat{p}_0$ in (2.3) is $\hat{p}_0 = 0.1$ or $\hat{p}_0 = 0.25$.

5.2.1. Results. Figure 8 shows the estimated mean probability of failure calculated by SIS and MLSIS plus/minus its standard deviation. The estimates of the means are in accordance with the reference solution and the bias and standard deviation decrease with an increasing number of samples. The standard deviation is smaller for a smaller target coefficient of variation. As for the 1D problem, we observe that applying the independent sampler with the vMFN distribution yields a smaller standard deviation than applying aCS.

Figure 9 shows the relative RMSE on the horizontal axis and the computational costs on the vertical axis for the SIS and MLSIS estimator. The costs are calculated based on the formula given in (3.1) for $L = 6$ and $d = 2$. Again, SIS and MLSIS yield the same range of the relative RMSE but the computational costs are lower for MLSIS. Considering the computational costs shown in Figure 9, we can save around 61% of the computational costs if we apply MLSIS for the estimation. This is the same level
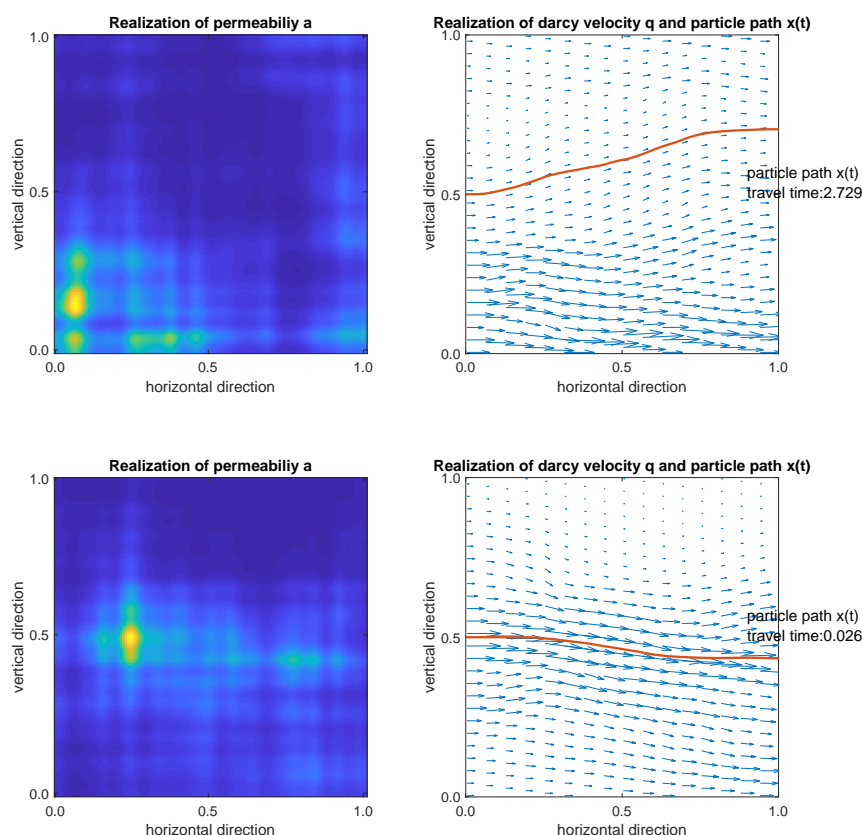
FIG. 7. *Realization of the permeability $a(\cdot, \omega)$ and the respective solution of the Darcy velocity $q_h(\cdot, \omega)$ and the particle path $x(t, \omega)$ for $h = 1/128$.*

of savings as in the 1D problem. However in the 2D problem, fewer level updates have to be performed than in the 1D problem setting. We expect that even more computational costs can be saved with MLSIS if we increase the highest discretization level.

Figure 10 shows the relative RMSE and computational costs of SIS, MLSIS, SuS, and MLSuS. We observe that SuS yields the same relative RMSE as SIS with aCS. However, SuS requires less computational cost. If we consider SIS with vMFN, the relative RMSE is smaller compared to SuS but the computational cost is higher for SIS. For the multilevel methods, we observe that MLSIS with $\delta_{\text{target}} = 1.00$ yields a smaller relative RMSE and requires less computational cost than MLSuS. This observation holds for both MCMC algorithms. In the 1D problem, we only observe that MLSIS with sampling from the vMFN distribution yields a more efficient estimator than MLSuS. For SuS and MLSuS, $\hat{p}_0 = 0.25$ yields higher computational cost and a slightly smaller relative RMSE than $\hat{p}_0 = 0.1$ since more intermediate failure domains are considered for $\hat{p}_0 = 0.25$.

6. Conclusion and outlook. Motivated by the nestedness issue of multilevel subset simulation, we implement multilevel sequential importance sampling to estimate the probability of rare events. We assume that the underlying limit state function depends on a discretization parameter ℓ. MLSIS samples a sequence of non-
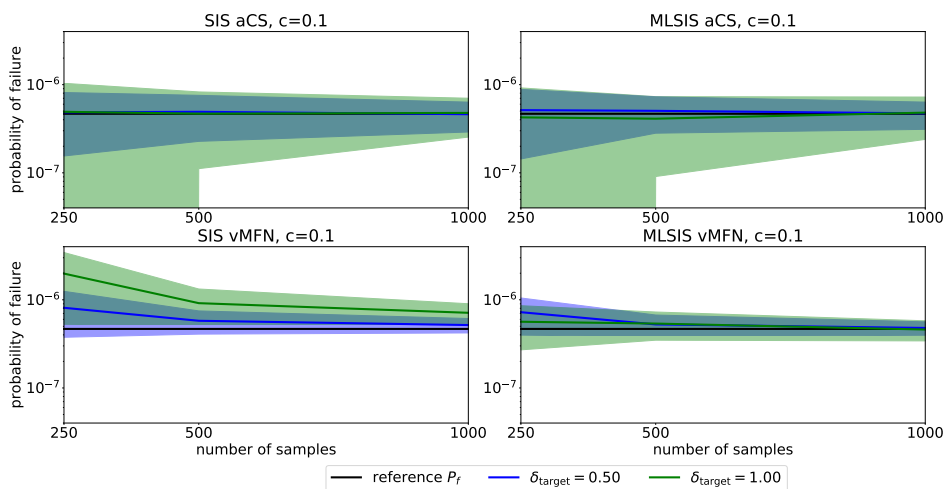
FIG. 8. *Estimated probability of failure by SIS and MLSIS averaged over* 100 *runs for* 250, 500, *and* 1000 *samples and* $\delta_{\mathrm{target}} \in \{0.50, 1.00\}$. *The colored areas show the standard deviation of the estimates. The black lines show the reference estimate by Monte Carlo sampling. First row: aCS,* $c = 0.1$; *second row: vMFN,* $c = 0.1$; *first column: SIS; second column: MLSIS with level dependent dimension.*
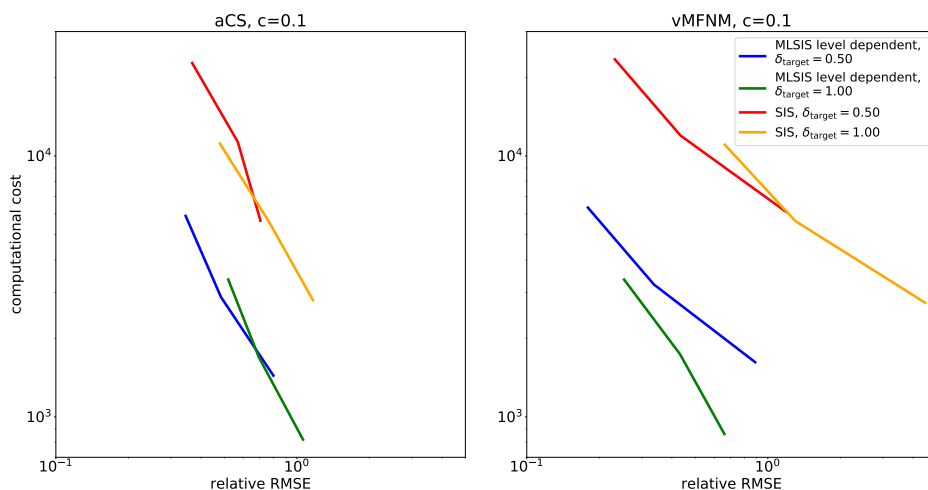
FIG. 9. *Computational costs and relative RMSE of SIS and MLSIS averaged over* 100 *runs for* 250, 500, *and* 1000 *samples and* $\delta_{\mathrm{target}} \in \{0.50, 1.00\}$. *First column: aCS,* $c = 0.1$; *second column: vMFN,* $c = 0.1$.

zero density functions that are adaptively chosen such that each pair of subsequent densities are only slightly different. Therefore, nestedness is not an issue for MLSIS. We combine the smoothing approach of the indicator function in [47] and the multilevel idea in [37]. This yields a twofold adaptive algorithm which combines tempering and bridging sequences in a clever way to reduce computational costs. Moreover, we apply the level dependent dimension approach of [53] to reduce variances between consecutive accuracy levels of the limit state function. This leads to more tempering updates on coarse levels and reduces computational costs. Another contribution of
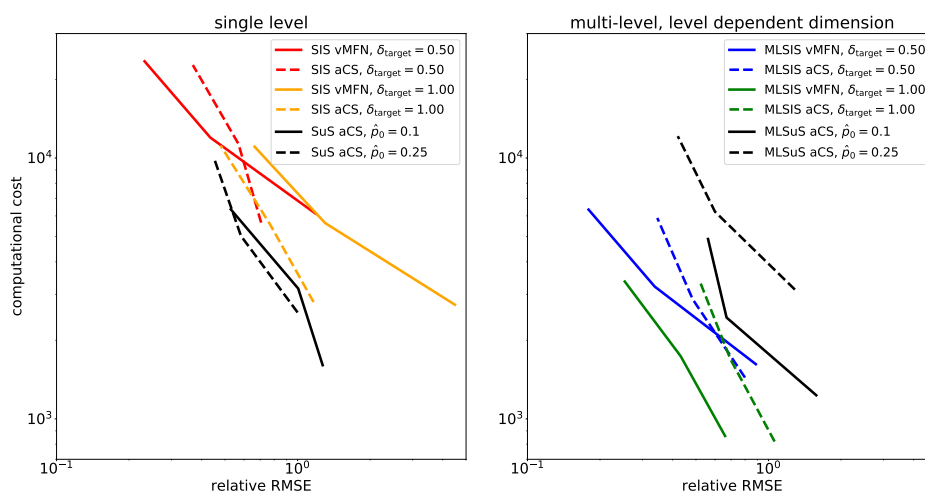
FIG. 10. *Computational costs and relative RMSE of SIS, MLSIS, SuS, and MLSuS averaged over* 100 *runs for* 250, 500, *and* 1000 *samples. SIS and MLSIS are considered with aCS and vMFN,* $c = 0.1$, *and* $\delta_{\text{target}} \in \{0.50, 1.00\}$. *SuS and MLSuS are considered with aCS and* $\hat{p}_0 = 0.1$ *or* $\hat{p}_0 = 0.25$. *First column: single level; second column: multilevel with level dependent dimension.*

our work is the vMFN distribution as a proposal density in an independent MCMC algorithm. This leads to an efficient MCMC algorithm even in high dimensions.

In numerical experiments in 1D and 2D space, we show for our experiments that MLSIS has a lower computational cost than SIS at any given error tolerance. For both experiments, MLSIS with the vMFN distribution leads to lower computational cost than multilevel subset simulation for the same accuracy. However, MLSIS with adaptive conditional sampling leads only for the 2D experiment to lower computational cost than multilevel subset simulation for the same accuracy. The results also show that applying the vMFN distribution as a proposal density in the MCMC algorithm reduces the bias and coefficient of variation of the MLSIS estimator compared to applying adaptive conditional sampling as the MCMC algorithm.

The bridging approach can also handle more general assumptions on the approximation sequence of the limit state function. For instance, the approximation sequence can arise within a multifidelity setting. Therein, bridging is applied to transfer samples between a low fidelity model and a high fidelity model.

Instead of using SIS to shift samples to the failure region, we plan to apply the ensemble Kalman filter for inverse problems as a particle based estimator for the probability of failure. In this case, the reliability problem is formulated as an inverse problem.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Department of Commerce, National Bureau of Standards, 1964.

[2] S. AGAPIOU, O. PAPASPILIOPOULOS, D. SANZ-ALONSO, AND A. M. STUART, *Importance sampling: Intrinsic dimension and computational cost*, Statist. Sci., 32 (2017), pp. 405–431, https://doi.org/10.1214/17-STS611.

[3] A. AGARWAL, S. DE MARCO, E. GOBET, AND G. LIU, *Rare event simulation related to financial risks: Efficient estimation and sensitivity analysis*, Working paper, https://hal-polytechnique.archives-ouvertes.fr/hal-01219616.

[4] S.-K. AU AND J. L. BECK, *Estimation of small failure probabilities in high dimensions by subset simulation*, Probab. Eng. Mech., 16 (2001), pp. 263–277, https://doi.org/10.1016/S0266-8920(01)00019-4.

[5] S.-K. AU AND Y. WANG, *Engineering Risk Assessment with Subset Simulation*, John Wiley, New York, 2014, https://doi.org/10.1002/9781118398050.

[6] A. BESKOS, A. JASRA, K. LAW, R. TEMPONE, AND Y. ZHOU, *Multilevel sequential Monte Carlo samplers*, Stochastic Process Appl., 127 (2017), pp. 1417–1440, https://doi.org/10.1016/j.spa.2016.08.004.

[7] A. BESKOS, A. JASRA, AND A. THIERY, *On the convergence of adaptive sequential Monte Carlo methods*, Ann. Appl. Probab., 26 (2013), pp. 1111–1146, https://doi.org/10.1214/15-AAP1113.

[8] Z. BOTEV AND D. KROESE, *Efficient Monte Carlo simulation via the generalized splitting method*, Statist. Comput., 22 (2012), pp. 1–16, https://doi.org/10.1007/s11222-010-9201-4.

[9] N. BOUHLEL AND A. DZIRI, *Maximum likelihood parameter estimation of Nakagami-gamma shadowed fading channels*, IEEE Commun. Lett., 19 (2015), pp. 685–688, https://doi.org/10.1109/LCOMM.2015.2394293.

[10] D. BRAESS, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, 3rd ed., Cambridge University Press, Cambridge, 2007, https://doi.org/10.1017/CBO9780511618635.

[11] N. CHOPIN, *A sequential particle filter method for static models*, Biometrika, 89 (2002), pp. 539–551, https://doi.org/10.1093/biomet/89.3.539.

[12] F. J. CORNATON, Y. PARK, S. D. NORMANI, E. A. SUDICKY, AND J. F. SYKES, *Use of groundwater lifetime expectancy for the performance assessment of a deep geologic waste repository*, Water Resour. Res., 44 (2008), https://doi.org/10.1029/2007WR006208.

[13] S. L. COTTER, G. O. ROBERTS, A. M. STUART, AND D. WHITE, *MCMC methods for functions: Modifying old algorithms to make them faster*, Statist. Sci., 28 (2013), pp. 424–446, https://doi.org/10.1214/13-STS421.

[14] F. CÉROU, P. DEL MORAL, T. FURON, AND A. GUYADER, *Sequential Monte Carlo for rare event estimation*, Stat. Comput., 22 (2012), pp. 795–808, https://doi.org/10.1007/s11222-011-9231-6.

[15] M. DASHTI AND A. M. STUART, *The Bayesian Approach to Inverse Problems*, Springer, New York, 2017, pp. 311–428, https://doi.org/10.1007/978-3-319-12385-1_7.

[16] M. DE ANGELIS, E. PATELLI, AND M. BEER, *Advanced line sampling for efficient robust reliability analysis*, Struct. Saf., 52 (2015), pp. 170–182, https://doi.org/10.1016/j.strusafe.2014.10.002.

[17] P. DEL MORAL, A. JASRA, K. LAW, AND Y. ZHOU, *Multilevel sequential Monte Carlo samplers for normalizing constants*, ACM Trans. Model. Comput. Simul., 27 (2017), 20, https://doi.org/10.1145/3092841.

[18] A. DER KIUREGHIAN AND P.-L. LIU, *Structural reliability under incomplete probability information*, J. Eng. Mech., 112 (1986), pp. 85–104, https://doi.org/10.1061/(ASCE)0733-9399(1986)112:1(85).

[19] A. DOUCET AND A. JOHANSEN, *A tutorial on particle filtering and smoothing: Fifteen years later*, Handb. Nonlinear Filtering, 12 (2009), pp. 656–704.

[20] D. ELFVERSON, F. HELLMAN, AND A. MÅLQVIST, *A multilevel Monte Carlo method for computing failure probabilities*, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 312–330, https://doi.org/10.1137/140984294.

[21] G. S. FISHMAN, *Monte Carlo: Concepts, Algorithms and Applications*, Springer, New York, 1996, https://doi.org/10.1007/978-1-4757-2553-7.

[22] M. GERBER, N. CHOPIN, AND N. WHITELEY, *Negative association, ordering and convergence of resampling methods*, Ann. Statist., 47 (2019), pp. 2236–2260, https://doi.org/10.1214/18-AOS1746.

[23] S. GEYER, I. PAPAIOANNOU, AND D. STRAUB, *Cross entropy-based importance sampling using Gaussian densities revisited*, Struct. Saf., 76 (2019), pp. 15–27, https://doi.org/10.1016/j.strusafe.2018.07.001.

[24] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, 1991, https://doi.org/10.1007/978-1-4612-3094-6.

[25] J. W. GIBBS, *Elementary Principles in Statistical Mechanics*, Charles Scribner's Sons, 1902.

[26] M. B. GILES, *Multilevel Monte Carlo methods*, Acta Numer., 24 (2015), pp. 259–328, https://doi.org/10.1017/S096249291500001X.

[27] P. GLASSERMAN, P. HEIDELBERGER, P. SHAHABUDDIN, AND T. ZAJIC, *Multilevel splitting for estimating rare event probabilities*, Oper. Res., 47 (1999), pp. 585–600, https://doi.org/10.1287/opre.47.4.585.

[28] W. K. HASTINGS, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57 (1970), pp. 97–109, https://doi.org/10.2307/2334940.

[29] M. HOHENBICHLER AND R. RACKWITZ, *Non-normal dependent vectors in structural safety*, J. Eng. Mech. Div., 107 (1981), pp. 1227–1238.

[30] A. JASRA, D. A. STEPHENS, A. DOUCET, AND T. TSAGARIS, *Inference for Lévy-driven stochastic volatility models via adaptive sequential Monte Carlo*, Scand. J. Stat., 38 (2011), pp. 1–22, https://doi.org/10.1111/j.1467-9469.2010.00723.x.

[31] H. KAHN AND A. W. MARSHALL, *Methods of reducing sample size in Monte Carlo computations*, J. Oper. Res. Soc. Amer., 1 (1953), pp. 263–278, https://doi.org/10.1287/opre.1.5.263.

[32] L. S. KATAFYGIOTIS AND K. M. ZUEV, *Geometric insight into the challenges of solving high-dimensional reliability problems*, Probab. Engrg. Mech., 23 (2008), pp. 208–218, https://doi.org/10.1016/j.probengmech.2007.12.026.

[33] A. KLENKE, *Probability Theory: A Comprehensive Course*, Springer, New York, 2014.

[34] P. S. KOUTSOURELAKIS, *A multi-resolution, non-parametric, Bayesian framework for identification of spatially-varying model parameters*, J. Comput. Phys., 228 (2009), pp. 6184–6211, https://doi.org/10.1016/j.jcp.2009.05.016.

[35] P. S. KOUTSOURELAKIS, H. J. PRADLWARTER, AND G. I. SCHUËLLER, *Reliability of structures in high dimensions, part* I: *Algorithms and applications*, Probab. Engrg. Mech., 19 (2004), pp. 409–417, https://doi.org/10.1016/j.probengmech.2004.05.001.

[36] S. LACAZE, L. BREVAULT, S. MISSOUM, AND M. BALESDENT, *Probability of failure sensitivity with respect to decision variables*, Struct. Multidiscip. Optim., 52 (2015), pp. 375–381, https://doi.org/10.1007/s00158-015-1232-1.

[37] J. LATZ, I. PAPAIOANNOU, AND E. ULLMANN, *Multilevel* Sequential^2 *Monte Carlo for Bayesian inverse problems*, J. Comput. Phys., 368 (2018), pp. 154–178, https://doi.org/10.1016/j.jcp.2018.04.014.

[38] G. MCLACHLAN AND D. PEEL, *ML fitting of mixture models*, in Finite Mixtrue Models, John Wiley and Sons, New York, 2005, pp. 40–80, https://doi.org/10.1002/0471721182.ch2.

[39] R. E. MELCHERS AND A. T. BECK, *Structural Reliability Analysis and Prediction*, 3rd ed., John Wiley and Sons, New York, 2018, https://doi.org/10.1002/9781119266105.

[40] H. METROPOLIS, A. ROSENBLUTH, M. ROSENBLUTH, A. TELLER, AND E. TELLER, *Equations of state calculations by fast computing machine*, J. Chem. Phys., 21 (1953), pp. 1087–1092, https://doi.org/10.1063/1.1699114.

[41] J. MORIO AND M. BALESDENT, *Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems*, 1st ed., Woodhead Publishing, Sawston, UK, 2015, https://doi.org/10.1016/C2014-0-02344-1.

[42] M. NAKAGAMI, *The m-distribution, a general formula of intensity distribution of rapid fading*, in Statistical Methods in Radio Wave Propagation, W. Hoffman, ed., Pergamon, Oxford, 1960, pp. 3–36, https://doi.org/10.1016/B978-0-08-009306-2.50005-4.

[43] U. NOSECK, D. BECKER, C. FAHRENHOLZ, E. FEIN, J. FLÜGGE, K.-P. KRÖHN, J. MÖNIG, I. MÜLLER-LYDA, T. ROTHFUCHS, A. RÜBEL, AND J. WOLF, *Assessment of the long-term safety of repositories*, Gesellschaft für Anlage und Reaktorsicherheit (GRS) mbH, 2008, https://www.grs.de/sites/default/files/pdf/GRS-237.pdf.

[44] I. PAPAIOANNOU, W. BETZ, K. ZWIRGLMAIER, AND D. STRAUB, *MCMC algorithms for subset simulation*, Probab. Engrg. Mech., 41 (2015), pp. 89–103, https://doi.org/10.1016/j.probengmech.2015.06.006.

[45] I. PAPAIOANNOU, K. BREITUNG, AND D. STRAUB, *Reliability sensitivity estimation with sequential importance sampling*, Struct. Saf., 75 (2018), pp. 24–34, https://doi.org/10.1016/j.strusafe.2018.05.003.

[46] I. PAPAIOANNOU, S. GEYER, AND D. STRAUB, *Improved cross entropy-based importance sampling with a flexible mixture model*, Reliab. Eng. Syst. Saf., 191 (2019), 106564, https://doi.org/10.1016/j.ress.2019.106564.

[47] I. PAPAIOANNOU, C. PAPADIMITRIOU, AND D. STRAUB, *Sequential importance sampling for structural reliability analysis*, Struct. Saf., 62 (2016), pp. 66–75, https://doi.org/10.1016/j.strusafe.2016.06.002.

[48] B. PEHERSTORFER, B. KRAMER, AND K. WILLCOX, *Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 737–761, https://doi.org/10.1137/17M1122992.

[49] R. RACKWITZ, *Reliability analysis—A review and some perspectives*, Struct. Saf., 23 (2001), pp. 365–395, https://doi.org/10.1016/S0167-4730(02)00009-7.

[50] P. A. RAVIART AND J. M. THOMAS, *A Mixed Finite Element Method for Second Order Elliptic Problems*, Math. Asp. Finite Elem. Methods 606, Springer, New York, 1977, pp. 292–315, https://doi.org/10.1007/BFb0064470.

[51] G. O. ROBERTS AND J. S. ROSENTHAL, *Optimal scaling for various Metropolis-Hastings algorithms*, Statist. Sci., 16 (2001), pp. 351–367, https://doi.org/10.1214/ss/1015346320.

[52] R. Y. RUBINSTEIN AND D. P. KROESE, *Simulation and the Monte Carlo Method: Concepts, Algorithms and Applications*, 3rd ed., John Wiley and Sons, New York, 2016, https://doi.org/10.1002/9781118631980.

[53] E. ULLMANN AND I. PAPAIOANNOU, *Multilevel estimation of rare events*, SIAM/ASA J. Uncertain. Quantif., 3 (2015), pp. 922–953, https://doi.org/10.1137/140992953.

[54] Z. WANG, M. BROCCARDO, AND J. SONG, *Hamiltonian Monte Carlo methods for subset simulation in reliability analysis*, Struct. Saf., 76 (2019), pp. 51–67, https://doi.org/10.1016/j.strusafe.2018.05.005.

[55] Z. WANG AND J. SONG, *Cross-entropy-based adaptive importance sampling using von Mises-Fisher mixture for high dimensional reliability analysis*, Struct. Saf., 59 (2016), pp. 42–52, https://doi.org/10.1016/j.strusafe.2015.11.002.

B. Core Article: Error analysis for probabilities of rare events with approximate models

Summary

In this article, we derive novel error bounds for the approximation error of the probability of rare events in settings, where the limit-state function (LSF) requires the solution of an elliptic partial differential equation (PDE) with stochastic diffusion coefficient. The discretization error associated with the PDE and its impact on the overall error of the probability of failure has not been studied to date.

The existing work of [36] considers the absolute approximation error of rare event probabilities. The authors of [36] show that the absolute error behaves as the PDE discretization error. Since rare event probabilities are usually very small, the absolute error is not of high relevance. In contrast, our analysis gives an upper bound for the relative error if the LSF is affine linear in the stochastic parameters. In the non-affine case, we use a linearization based on the first order reliability method (FORM). This enables to handle LSFs which require the solution of an elliptic PDE with a stochastic diffusion coefficient. We show that the relative error of the FORM estimates behaves as the PDE discretization error. Moreover, in the case that the failure domain is convex, we derive an upper bound for the absolute error of the probability of failure, which behaves as the discretization accuracy of the PDE multiplied by the FORM estimate. For all error bounds, we require that the diffusion coefficient is uniformly bounded.

We split the proof into five parts. First, we show that the absolute error is bounded by the PDE discretization error multiplied by the local Lipschitz constant associated with the cumulative distribution function of the LSF. This result is similar to the result of [36, Lemma 3.4 M1]. In the next step, we show that for affine linear LSFs, the upper bound of the local Lipschitz constant depends linearly on the probability of rare events, which gives an upper bound for the relative error. If the LSF is based on the solution of an elliptic PDE with stochastic diffusion coefficient, we show that the distance of the exact and approximate failure domains behaves as the PDE discretization error. Next, we show that the Gaussian measure of the symmetric difference of the failure domains can be bounded by the Gaussian measure of an interval in 1D. Finally, we linearize the LSF at the most likely failure point and apply the upper bound of the linear case which yields our derived error bound.

Statement of individual contribution

Elisabeth Ullmann and Jonas Latz assigned me with the task of analysing the approximation error of the probability of failure with respect to the PDE discretization error. The goal was the derivation of an error bound of the relative error, which behaves as the PDE discretization error.

In a discussion with Elisabeth Ullman and Jonas Latz, we have observed an interesting connection of the FORM minimization problem to optimal control problems. In another discussion with Daniel Walter, we got further insights about a-priori error estimates for optimal control.

I was fully responsible for the implementation and execution of numerical experiments and for writing and preparing this manuscript. Moreover, I was in charge to work out the technical steps in the proofs of our main statements. I was supported by insightful discussions with Jonas Latz, Iason Papaionnou and Elisabeth Ullmann concerning technical details. Jonas Latz, Iason Papaioannou, Elisabeth Ullmann, and I proofread and polished the article together.

Permission to include

On the following page, a copy of the E-Mail by SIAM may be found with the confirmation that the manuscript "Error analysis for probabilities of rare events with approximate models" may be included in my dissertation. Moreover, a digital version of the consent to publish may be found at

```
https://www.siam.org/publications/journals/about-siam-journals/
                     information-for-authors
```

(Accessed on 21 June 2021)

Fabian Wagner

Dear Mr. Wagner:

SIAM is happy to give permission to reuse material from your article in your thesis. Please credit the original article, if possible using the complete bibliographic information. I estimate that it will be published in about 2 months.

Sincerely,

Kelly Thomas
Managing Editor
Society for Industrial and Applied Mathematics
3600 Market Street - 6th Floor
Philadelphia, PA 19104
thomas@siam.org / (267) 350-6387

From: SINUM <sinum@siam.org>
Sent: Tuesday, May 25, 2021 11:29 AM
To: Karl W. Durst <Durst@siam.org>
Subject: M135980R (Error Analysis for Probabilities of Rare...)

Dear Sirs and Madams,

My name is Fabian Wagner, I am one of the authors of the article "Error Analysis for Probabilities of Rare Events with Approximate Models" (M135980R) that is accepted for publication in the SIAM Journal on Numerical Analysis.
Since I am pursuing a cumulative doctoral thesis, it is necessary by the rules of my university (TU Munich, Germany) that I provide a "letter of approval" by the publisher to all my articles that I want to include in my thesis. I hereby kindly ask for such a confirmation (by Email or weblink to your terms and condition page stating such a regulation), allowing me to use the article mentioned above in my doctoral thesis.

If you have any questions beforehand, do not hesitate to ask.

Best wishes,
Fabian Wagner

The first two books in the SIAM Data Science Book Series have now published! Learn more.

Notice of publication and copyright

The manuscript "Error analysis for probabilities of rare events with approximate models" is accepted for publication in SIAM Journal on Numerical Analysis (2021), published by the Society for Industrial and Applied Mathematics (SIAM).

On the following page, a copy of the E-Mail by SIAM may be found with the confirmation that the manuscript "Error analysis for probabilities of rare events with approximate models" is accepted for publication in SIAM Journal on Numerical Analysis.

Fabian Wagner

Dear Mr. Wagner,

I am pleased to accept "Error analysis for probabilities of rare events with approximate models" for publication.

No additional revision is needed. SIAM will send a separate request for your TeX file and you will also receive information about the publication process.

In the absence of information from you to the contrary, you confirm that all authors listed on the manuscript concur with publication in its final accepted form and that neither this manuscript nor any of its essential components have been published previously or submitted to another journal.

Thanks again for submitting your contribution to SINUM and for your support.

Best regards,

Robert Scheichl
Associate Editor
SIAM Journal on Numerical Analysis

Associate Editor (Remarks to the Author):

Please when preparing your final manuscript take the final comment by Referee 2 into account.

Referee #2 (Remarks to the Author):

The authors have prepared a revised version of the manuscript which clarifies all of the issues raised in my original report.

One final small remark: [30, Theorem 2.1] only applies to two-dimensional polygonal domains so doesn't cover the case d=3.

ERROR ANALYSIS FOR PROBABILITIES OF RARE EVENTS WITH APPROXIMATE MODELS *

F. WAGNER[†], J. LATZ[‡], I. PAPAIOANNOU[§], AND E. ULLMANN[†]

Abstract. The estimation of the probability of rare events is an important task in reliability and risk assessment. We consider failure events that are expressed in terms of a limit-state function, which depends on the solution of a partial differential equation (PDE). In many applications, the PDE cannot be solved analytically. We can only evaluate an approximation of the exact PDE solution. Therefore, the probability of rare events is estimated with respect to an approximation of the limit-state function. This leads to an approximation error in the estimate of the probability of rare events. Indeed, we prove an error bound for the approximation error of the probability of failure, which behaves like the discretization accuracy of the PDE multiplied by an approximation of the probability of failure, the first order reliability method (FORM) estimate. This bound requires convexity of the failure domain. For non-convex failure domains, we prove an error bound for the relative error of the FORM estimate. Hence, we derive a relationship between the required accuracy of the probability of rare events estimate and the PDE discretization level. This relationship can be used to guide practicable reliability analyses and, for instance, multilevel methods.

Key words. Uncertainty quantification, stochastic finite elements, error analysis, reliability analysis

AMS subject classifications. 35R60, 41A25, 65N15

1. Introduction. The distinction of safe and failure events is a crucial topic in reliability analysis and risk management. The occurrence of failure events cannot always be avoided; therefore, the estimation of the probability of such occurrences is of high significance. Indeed, failure probabilities are usually small; hence, the denomination *probability of rare events* is commonly used. The failure event can be defined in terms of a *limit-state function* (LSF). The LSF is a function of a set of uncertain parameters and the failure event is defined by the collection of parameter values for which the LSF takes non-positive values.

In this work, we consider settings where the LSF is based on the solution of an elliptic partial differential equation (PDE) with random diffusion coefficient. These situations frequently arise in engineering risk settings. For example, the authors in [5, 22] consider radioactive waste repositories. Therein, the departure of radioactive particles and their travel paths through the subsurface are of high relevance. The goal is to determine the probability that radioactive particles come back to the human environment. Since the exact subsurface properties and exact travel paths of the particles are unknown, the hydraulic conductivity of the soil is modelled as a random field while the particle flow is simulated by a finite element method (FEM) approximation of the groundwater flow and transport equation.

The application of discretization schemes, such as Finite Differences [20], Finite Vol-

---

[†]Department of Mathematics, Technical University of Munich, Boltzmannstraße 3, 85748 Garching (fabian.wagner@ma.tum.de, elisabeth.ullmann@ma.tum.de).

[‡]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA (jl2160@cam.ac.uk).

[§]Engineering Risk Analysis Group, Technical University of Munich, Theresienstraße 90, 80333 Munich (iason.papaioannou@tum.de).

umes [11] or FEM [3], introduce a PDE discretization error in the evaluation of the LSF. Consequently, this leads to an approximation error of the probability of rare events. The accuracy of the approximation depends on the discretization size of the spatial or temporal domain. In previous years, many methods have been developed for rare event estimation, which are based on a sequence of discretization levels with increasing accuracy. Examples are Multilevel Subset Simulation [31], Multilevel Monte Carlo [8], Multilevel Sequential Importance Sampling [33] or a multifidelity approach in [25]. However, only the sampling error of the methods has been considered so far. In contrast, we do not consider the sampling error and focus on the PDE discretization error.

There is a large amount of literature available, which derive error bounds for the PDE discretization error. However, there are few publications which consider the induced approximation error of rare event probabilities. The authors in [8] derive an upper bound for the absolute approximation error of the probability of failure which behaves as the PDE discretization error. However, the absolute error is of limited interest since failure probabilities are usually small. This manuscript closes this gap for LSFs which are based on elliptic PDEs. Indeed, the work of [8] forms the starting point for our contributions.

We derive an error bound for the probability of rare events which behaves as the PDE discretization error multiplied by the *first order reliability method* (FORM) [17, 19] estimate. The FORM estimate is determined by the minimum distance of the failure domain to the origin of an independent standard Gaussian input space. If the failure domain is a convex set, the FORM estimate is an upper bound for the probability of the rare event. We use this condition as an assumption for the derived error bound. Indeed, if the FORM estimate is equal to the probability of failure, our error bound gives an upper bound for the relative error. An example for this case is an LSF which is affine linear with respect to Gaussian stochastic parameters. Moreover, we provide an error bound of the relative error with respect to the FORM estimates. This bound is more generally applicable since convexity of the failure domains is not required.

The manuscript is structured as follows. In Section 2, we present the main theorem, underlying setting and relevant assumptions. The proof of the theorem is given in the subsequent sections. First, we show that the absolute error is bounded by the PDE discretization error multiplied by the local Lipschitz constant associated with the *cumulative distribution function* (CDF) of the LSF. In Section 3, we show that for affine linear LSFs, the upper bound of the local Lipschitz constant depends linearly on the probability of failure. This gives an upper bound for the relative error. If the LSF is based on the solution of an elliptic PDE with stochastic diffusion coefficient, we show in Section 4 that the distance of the exact and approximate failure domains behaves as the PDE discretization error. Thereafter, we show that the Gaussian measure of the symmetric difference of the failure domains can be bounded by the Gaussian measure of an interval in 1D if the failure domains are convex. Using this fact, the LSF is linearized around the *most likely failure point* (MLFP) and the results for affine linear LSFs are applied. This proves the main theorem. In Section 5, we consider three numerical examples, two in low dimensions with analytical solutions, and one high-dimensional example.

2. Problem Setting and Main Result. Even though many of our results are applicable to more general settings, we focus in this work on failure events that are based on the solution of an elliptic PDE with random diffusion coefficient and Dirichlet boundary conditions. In the following subsections, we first introduce this setting more

particularly. Then, we briefly describe the FORM approach to the estimation of rare event probabilities. The FORM estimate is part of the error bound that forms the main result of this work and that we summarise in Section 2.3.

2.1. Elliptic PDE, failure events, and their approximation. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, be an open, bounded, convex, polygonal domain. The given quantities are a real-valued random field $a : D \times \Omega \to \mathbb{R}$ and a real valued function $f \in L^2(D)$. We seek a random field $y : \overline{D} \times \Omega \to \mathbb{R}$, such that for $\mathbb{P}$-almost every (a.e.) $\omega \in \Omega$ it holds

$$(2.1) \qquad -\nabla_x \cdot (a(x, \omega) \nabla_x y(x, \omega)) = f(x) \quad \forall x \in D,$$

$$(2.2) \qquad y(x, \omega) = 0 \quad \forall x \in \partial D.$$

In practice, we employ a FEM discretization to solve (2.1). Thus, we consider the weak or variational form of the PDE. A random field $y : \overline{D} \times \Omega \to \mathbb{R}$ satisfies the *pathwise variational formulation*, if for a fixed $\omega \in \Omega$ it holds that $y(\cdot, \omega) \in V$ and

$$(2.3) \qquad \int_D a(x, \omega) \nabla_x y(x, \omega) \cdot \nabla_x v(x) \mathrm{d}x = \int_D f(x) v(x) \mathrm{d}x \quad \forall v \in V,$$

where $V := H_0^1(D)$. Let $h > 0$ denote a discretization parameter, typically the mesh size. We define the *discretized pathwise variational formulation* for $y_h \in V_h$ as

$$(2.4) \qquad \int_D a(x, \omega) \nabla_x y_h(x, \omega) \cdot \nabla_x v_h(x) \mathrm{d}x = \int_D f(x) v_h(x) \mathrm{d}x \quad \forall v_h \in V_h.$$

Here, $V_h \subset V$ is a finite dimensional vector space. In this manuscript, we consider two types of diffusion coefficients.

DEFINITION 2.1 (Ellipticity and boundedness of the diffusion coefficient).
 (I) *The diffusion coefficient a is uniformly elliptic and bounded if there exists $a_{\min}, a_{\max} > 0$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$*

$$(2.5) \qquad 0 < a_{\min} \le a(x, \omega) \le a_{\max} < \infty, \quad \text{for a.e. } x \in D.$$

 (II) *The diffusion coefficient a is pathwise elliptic and bounded if there exists real-valued random variables $a_{\min}, a_{\max} : \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$*

$$(2.6) \qquad 0 < a_{\min}(\omega) \le a(x, \omega) \le a_{\max}(\omega) < \infty, \quad \text{for a.e. } x \in D,$$

 where $a_{\min}(\omega) := \min_{x \in \overline{D}} a(x, \omega)$ and $a_{\max}(\omega) := \max_{x \in \overline{D}} a(x, \omega)$.

Remark 2.2. We note that our proved error bounds in Proposition 2.11 and Theorem 2.12 require that the approximation error of an observation operator acting on y and y_h is uniformly bounded; see Assumption 2.5. This assumption is in general violated for diffusion coefficients which are only pathwise elliptic and bounded, i.e., they do not satisfy (2.5). Therefore, Proposition 2.11 and Theorem 2.12 are only valid for uniformly elliptic and bounded diffusion coefficients. In Remark 2.7 and 2.13, we will discuss in which way our derived error bounds are useful for diffusion coefficients which only satisfy (2.6).

Under Assumption 2.3, existence and uniqueness of a solution for (2.3) and (2.4) is ensured by [21, Theorem 9.9]. Moreover, under Assumption 2.3 and for $d = 2$, the authors of [30, Theorem 2.1] show that the solution y of (2.3) satisfies $y(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$, which we require in the proof of Theorem 4.2. For $d = 3$, the authors state in [30, Remark 5.2 (c)] that the same property holds if D is convex.

Assumption 2.3 (Regularity of the diffusion coefficient). We assume that
  (i) the computational domain D is open, bounded, convex and polygonal,
  (ii) $a_{\min}(\omega) \geq 0$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $1/a_{\min} \in L^p(\Omega)$ for all $p \in (0, \infty)$,
  (iii) $a \in L^p(\Omega, C^1(\bar{D}))$ for all $p \in (0, \infty)$, i.e., the realisations $a(\cdot, \omega)$ are continuously differentiable,
  (iv) $f \in L^2(D)$.

We note that Assumption 2.3 (ii) is automatically satisfied for uniformly elliptic and bounded diffusion coefficients. Moreover, $y(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$ is still ensured if Assumption 2.3 (i) is replaced by requiring that D is open, bounded and has a C^2 boundary [4, Theorem 3.4].

Having considered the spatial regularity of the diffusion coefficient, we specify the parametric regularity of a in the following assumption. Moreover, we require that a depends on a Gaussian random variable.

Assumption 2.4 (Parametric form and parametric regularity of the diffusion coefficient).
  (i) The diffusion coefficient $a(x, \omega)$ is a measurable function of an n-variate random vector $U : \Omega \to \mathbb{R}^n$, where U follows the n-variate independent standard normal distribution. This means, there is a function $\hat{a} : D \times \mathbb{R}^n \to \mathbb{R}$ with $a(x, \omega) = \hat{a}(x, U(\omega))$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.
  (ii) The diffusion coefficient $a(x, \omega)$ is three times continuously differentiable with respect to outcomes $u \in \mathbb{R}^n$ of U for all $x \in D$.

Assumption 2.4 (i) implies that $a(x, \omega)$ can be viewed as a function in space depending on an n-dimensional parameter given by the outcomes $u \in \mathbb{R}^n$ of U. Thus, a can be viewed as finite dimensional noise [21, Definition 9.38]. We note that Assumption 2.4 (i) is not a strong restriction. Under mild assumptions, a non-Gaussian random variable $\widetilde{U}$ can be transformed via an isoprobabilistic transformation $U = T(\widetilde{U})$ to a Gaussian random variable U. For instance, if $\widetilde{U}$ can be modelled by a Gaussian copula, the Nataf transform [6] can be applied to express it as a function of a standard normal random variable. If the conditional distributions of $\widetilde{U}_{k+1}$ given $\widetilde{U}_1, \ldots, \widetilde{U}_k$ are known for $k = 1, \ldots, n-1$, the Rosenblatt transform can be applied [18].

Based on the elliptic PDE, we now define the LSF, the failure event, the failure probability, and their approximations. Failure is defined in terms of an LSF $G : \mathbb{R}^n \to \mathbb{R}$ such that $G(U(\omega)) \leq 0$ for $\omega \in \Omega$. Furthermore, we assume that the LSF G and the PDE solution y are related via a linear and bounded operator $\mathcal{F} : V \to \mathbb{R}$

$$(2.7) \qquad G(U(\omega)) := y_{\max} - \mathcal{F}y(\cdot, \omega),$$

where $y_{\max} \in \mathbb{R}$ is a constant. Analogously, we define the discretized LSF $G_h : \mathbb{R}^n \to \mathbb{R}$ as

$$(2.8) \qquad G_h(U(\omega)) := y_{\max} - \mathcal{F}_h y_h(\cdot, \omega),$$

where $\mathcal{F}_h : V_h \to \mathbb{R}$ is the induced discretization of $\mathcal{F}$. With the operator $\mathcal{F}$, we define the dual problem, where we seek the solution $z(\cdot, \omega) \in H_0^1(D)$ such that

$$(2.9) \qquad \int_D a(x, \omega) \nabla_x z(x, \omega) \cdot \nabla_x v(x) \mathrm{d}x = \mathcal{F}(v) \quad \forall v \in H_0^1(D).$$

Since $\mathcal{F}$ is linear and bounded, existence and uniqueness of a solution of the dual problem (2.9) is ensured by the Lax–Milgram theorem [10, Section 6.2.1]. By Assumption 2.3, it follows that $z(\cdot, \omega) \in H^2(D) \cap H_0^1(D)$, which we require in the proof of Theorem 4.2.

Our analysis is performed for the *probability of failure*. This quantity is defined as the probability mass of the *failure domain* $A := \{u \in \mathbb{R}^n : G(u) \leq 0\}$, which is expressed as

$$(2.10) \qquad P_f := \mathbb{P}[A] = \mathbb{P}[G(U) \leq 0] = \int_{u \in \mathbb{R}^n} I(G(u) \leq 0)\varphi_n(u)\mathrm{d}u,$$

where I denotes the indicator function; i.e., $I(\text{true}) = 1$ and $I(\text{false}) = 0$. The function $\varphi_n : \mathbb{R}^n \to \mathbb{R}$ denotes the *probability density function* (PDF) of the n-variate independent standard normal distribution, which we denote by $\mathrm{N}(0, \mathrm{Id}_n)$. Replacing G by G_h in (2.10) gives the approximation

$$(2.11) \qquad P_{f,h} := \mathbb{P}[A_h] = \mathbb{P}[G_h(U) \leq 0] = \int_{u \in \mathbb{R}^n} I(G_h(u) \leq 0)\varphi_n(u)\mathrm{d}u,$$

where $A_h = \{u \in \mathbb{R}^n : G_h(u) \leq 0\}$. Since $P_{f,h}$ includes numerical errors due to approximating the exact LSF G, we cannot expect equality of P_f and $P_{f,h}$. The main contribution of this work is the derivation of an upper bound for the error

$$(2.12) \qquad |P_f - P_{f,h}| \leq \mathbb{P}[A \triangle A_h],$$

where

$$A \triangle A_h := (A \setminus A_h) \cup (A_h \setminus A) = \{u \in \mathbb{R}^n : \text{ either } u \in A \text{ or } u \in A_h\}$$

is the *symmetric difference* of A and A_h. This upper bound behaves as a certain approximation to the rare event probability with the approximate model multiplied with the discretization error of $|\mathcal{F}y - \mathcal{F}_h y_h|$. For the latter, we assume the following approximation property.

Assumption 2.5 (Approximation error of the LSF). The operator $\mathcal{F}$ is linear and bounded and there exists constants $C_{\mathrm{FE}} > 0$ and $s > 0$ independent of h, such that the discretization error with respect to the solution of (2.3) and (2.4) satisfies for $\mathbb{P}$-a.e. $\omega \in \Omega$

$$(2.13) \qquad |G(U(\omega)) - G_h(U(\omega))| = |\mathcal{F}y(\cdot, \omega) - \mathcal{F}_h y_h(\cdot, \omega)| \leq C_{\mathrm{FE}} h^s.$$

Moreover, we require Lipschitz continuity of the CDFs of the random variables $G(U)$ and $G_h(U)$.

Assumption 2.6 (Regularity of the CDFs of $G(U)$, $G_h(U)$). The CDFs of the random variables $G(U)$ and $G_h(U)$ are local Lipschitz continuous with Lipschitz constants $C_L > 0$ and $C_{L,h} > 0$, i.e., for a, b with $a < b$ it holds

$$\mathbb{P}[G(U) \in]a, b]] = \mathbb{P}[G(U) \leq b] - \mathbb{P}[G(U) \leq a] \leq C_L|a - b|,$$
$$\mathbb{P}[G_h(U) \in]a, b]] = \mathbb{P}[G_h(U) \leq b] - \mathbb{P}[G_h(U) \leq a] \leq C_{L,h}|a - b|.$$

Remark 2.7. Note that the uniform bound in (2.13) might not be valid for diffusion coefficients which are only pathwise elliptic and bounded, i.e., are of type (II) in Definition 2.1. In this case, C_{FE} is a random variable and depends on $\omega \in \Omega$ such that

$$|G(U(\omega)) - G_h(U(\omega))| = |\mathcal{F}y(\cdot, \omega) - \mathcal{F}_h y_h(\cdot, \omega)| \leq C_{\mathrm{FE}}(\omega) h^s.$$

One idea to handle such cases is the restriction of the random variable U to a bounded domain $B_R = \{u \in \mathbb{R}^n : \|u\|_2 \leq R\}$, where $R > 0$. Thus, the random variable $C_{\mathrm{FE}}(\omega)$ is bounded uniformly within $\{\omega \in \Omega : U(\omega) \in B_R\}$. This idea can be seen as truncating the tails of the n-variate normal distribution. For log-normal random fields, the truncation yields a uniformly elliptic and bounded diffusion coefficient which satisfies (2.13). In Remark 2.13, we further discuss this idea and investigate our provided error bounds for pathwise elliptic and bounded diffusion coefficients.

In fact, the following example considers a case, where C_{FE} depends on $\omega \in \Omega$. The diffusion coefficient a_{II} is a log-normal random field and, thus, not uniformly elliptic and bounded.

Example 2.8. We consider the model problem (2.1) and (2.2) with $D = (0,1)$, $f(x) = 0$ for all $x \in D$ and boundary conditions $y(0, \omega) = 1$ and $y(1, \omega) = 0$. A similar problem is considered in [29]. We consider two examples for the diffusion coefficient

$$a_{\mathrm{I}}(x, \omega) := 2 + \tanh(Z(x, \omega)), \qquad a_{\mathrm{II}}(x, \omega) := \exp(Z(x, \omega)).$$

In both examples, Z is a Gaussian random field. We note that a_{I} is uniformly elliptic and bounded, i.e. (2.5) is satisfied, since $1 \leq a_{\mathrm{I}}(x, \omega) \leq 3$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ and all $x \in D$. The authors of [12] employ a similar random field model to describe a geotechnical material parameter. The diffusion coefficient a_{II} is only pathwise elliptic and bounded, i.e. (2.6) is satisfied. We assume that Z has constant mean μ_Z and constant variance σ_Z^2, while the covariance function is of exponential type. It is well known that the exponential covariance kernel $c(x, y) := \exp(\|x - y\|_1/\lambda)$ produces realisations which are not continuously differentiable [27, Chapter 4]. The parameter λ denotes the correlation length. The random field Z can be approximated via its truncated Karhunen–Loève expansion (KLE)

$$Z(x, \omega) \approx Z_n(x, \omega) := \mu_Z + \sigma_Z \sum_{m=1}^{n} \sqrt{\nu_m} z_m(x) U_m(\omega),$$

where (ν_m, z_m) are the KL eigenpairs of the correlation operator. For the exponential covariance kernel, a derivation of the eigenpairs is given in [14, Section 2.3.3]. We note that the approximation error introduced by the truncation of the KLE is not part of our analysis and the truncation order n is fixed.
The random variables $\{U_m\}_{m=1}^{n}$ are independent and standard normally distributed. Since we consider finitely many KL terms and the eigenfunctions $z_m(\cdot)$ are smooth, the realisations $a_{\mathrm{I},n}(\cdot, \omega) := 2 + \tanh(Z_n(\cdot, \omega))$ and $a_{\mathrm{II},n}(\cdot, \omega) := \exp(Z_n(\cdot, \omega))$ are sufficiently smooth and Assumptions 2.3 and 2.4 are satisfied. By the Sobolev embedding theorem [16, Theorem 6.48], the solution $y(\cdot, \omega)$ is continuously differentiable. Failure occurs if the flow rate

$$(2.14) \qquad q(x, \omega) := -a_{\mathrm{I/II},n}(x, \omega) \frac{\partial y(x, \omega)}{\partial x}$$

is larger than $q_{\max}$ at $x = 1$. Hence, the operator $\mathcal{F}$ is given as the point evaluation of the flow rate q at $x = 1$, which yields the LSF $G(U(\omega)) = q_{\max} - q(1, \omega)$. Linear FEs are applied to derive a discretization. By [28, Section 1.6], it follows that

$$\|y(\cdot, \omega) - y_h(\cdot, \omega)\|_{W^{1,\infty}} \leq C\|y(\cdot, \omega)\|_{W^{2,\infty}} h.$$

For the diffusion coefficient a_I, $C\|y(\cdot,\omega)\|_{W^{2,\infty}} \leq C_\mathrm{FE}$ can be uniformly bounded and Assumption 2.5 is satisfied for $s = 1$. However, for the diffusion coefficient a_II, $C_\mathrm{FE}(\omega) = C\|y(\cdot,\omega)\|_{W^{2,\infty}}$ is a random variable. In Section 5.3, we consider again this example with the log-normal diffusion coefficient a_II.

2.2. FORM probability of failure. We derive an upper bound for the error given in (2.12) which depends on the PDE discretization error and on an approximation of the probability of failure. This approximation will be given by the FORM estimate of the probability of failure; see [17, 19] for details. We now briefly introduce the FORM method.

We define the MLFP $u^\mathrm{MLFP} \in \mathbb{R}^n$ as the solution of the minimization problem

$$\min_{u \in \mathbb{R}^n} \quad \frac{1}{2}\|u\|_2^2, \quad \text{such that} \quad G(u) = 0.$$

Hence, u^MLFP is the element of the set $\{G = 0\}$ that has smallest distance to the origin and, thus, maximizes the Gaussian density φ_n. Accordingly, we denote the MLFP with respect to the discretization G_h as u_h^MLFP. We require that $G(0) > 0$ and $G_h(0) > 0$, since we are generally interested in estimating failure probabilities which are in the tail of the densities. Using the MLFPs, we obtain an estimate for the probability of failure via

$$P_f^\mathrm{FORM} = \Phi(-\|u^\mathrm{MLFP}\|_2) \quad \text{and} \quad P_{f,h}^\mathrm{FORM} = \Phi\left(-\|u_h^\mathrm{MLFP}\|_2\right),$$

where Φ is the CDF of the one-dimensional standard normal distribution. The FORM estimate is equal to the probability mass of the half-space which is defined through the hyperplane at the MLFP with direction perpendicular to the surface of the failure domain. Thus, the FORM estimate is an upper bound for the probability of failure, if the failure domain is convex. We use the convexity of the failure domains as an assumption on the LSF for Theorem 2.12. We state the convexity assumption in Assumption 2.9. In Assumption 2.10, we state an assumption on the gradient of the LSF, which is relevant for the proof of Theorem 4.2.

Assumption 2.9 (Geometry of the failure domains). The failure domains A and A_h are unbounded, convex sets.

Assumption 2.10 (Non-degeneracy of $\nabla_u G$ and $\nabla_u G_h$ at the limit-state surface). For all $h > 0$ there exists $\nu_h > 0$ such that for almost every $u \in \partial A$ it holds $\nabla_u G(u) \neq 0$, $\nabla_u G_h(u) \neq 0$ and $|\cos\left(\sphericalangle\left(u - u_h, \nabla_u G(u)\right)\right)| \geq \nu_h$, where $u_h \in \partial A_h$ is the point that has minimal distance to u and $\sphericalangle(\cdot,\cdot)$ denotes the angle between two vectors.

Assumption 2.10 states that the direction from a point $u \in \partial A$ to its nearest neighbour $u_h \in \partial A_h$ is not orthogonal to the gradient $\nabla_u G(u)$.

2.3. Error bound for the probability of failure. The following proposition and theorem are the main statements of this manuscript. Proposition 2.11 states an error bound of the relative error with respect to the FORM estimates P_f^FORM and $P_{f,h}^\mathrm{FORM}$. This bound is applicable in the general case where the geometries of the failure domains are unknown. In particular, the convexity of the failure domains is not required. In Theorem 2.12, we require convexity of the failure domains. Hence, the situation is more restrictive as compared with Proposition 2.11. In this case, we derive an error bound of the absolute error $|P_f - P_{f,h}|$ in dependence of the discretized FORM estimate $P_{f,h}^\mathrm{FORM}$. Subsequent to Theorem 2.12, we give a remark which discusses

the error bounds if the approximation error of G and G_h in Assumption 2.5 is not uniformly bounded. This remark builds on Remark 2.2 and 2.7.

PROPOSITION 2.11. *Let $a(x,\omega)$ be a uniformly elliptic and bounded diffusion coefficient and let Assumptions 2.3, 2.4, 2.5, 2.6, and 2.10 hold. Then for $h > 0$ sufficiently small, the relative error of the FORM estimates is upper bounded by*

$$(2.15) \qquad \frac{|P_f^{\mathrm{FORM}} - P_{f,h}^{\mathrm{FORM}}|}{P_f^{\mathrm{FORM}}} \leq \widehat{C}^{\mathrm{FORM}} h^s.$$

THEOREM 2.12. *Let $a(x,\omega)$ be a uniformly elliptic and bounded diffusion coefficient and let Assumptions 2.3, 2.4, 2.5, 2.6, 2.9 and 2.10 hold. Then for $h > 0$ sufficiently small, the error of the exact and approximate probability of failure is upper bounded by*

$$(2.16) \qquad |P_f - P_{f,h}| \leq \widehat{C} h^s P_{f,h}^{\mathrm{FORM}}.$$

We note that the constants $\widehat{C}^{\mathrm{FORM}} > 0$ and $\widehat{C} > 0$ in (2.15) and (2.16), respectively, depend on C_{FE}, h, n, $\|u^{\mathrm{MLFP}}\|_2$, and $\|u_h^{\mathrm{MLFP}}\|_2$. We will discuss the behaviour of $\widehat{C}^{\mathrm{FORM}}$ and $\widehat{C}$ with respect to their dependencies in the following sections. The outline of the proof of Theorem 2.12 is as follows:

(P1) For the absolute error, we derive the bound

$$|P_f - P_{f,h}| \leq C_1 h^s,$$

where C_1 depends on the local Lipschitz constant C_L of the CDF of $G(U)$.

(P2) Under the assumption that G is affine linear with respect to U, we derive an upper bound for the local Lipschitz constant of the CDF of $G(U)$ which depends linearly on P_f. This yields an upper bound for the relative error

$$|P_f - P_{f,h}|/P_f \leq C_2 h^s.$$

(P3) For LSFs of the form (2.7) and (2.8), we prove that the distance between the exact *limit-state surface* $\partial A := \{u \in \mathbb{R}^n : G(u) = 0\}$ and its approximation $\partial A_h := \{u \in \mathbb{R}^n : G_h(u) = 0\}$ has order $\mathcal{O}(h^s)$ of convergence for $h > 0$ sufficiently small, i.e., for all $u \in \partial A$ it holds

$$\mathrm{dist}(u, \partial A_h) \leq C_3 h^s.$$

(P4) Under the assumption that $\mathrm{dist}(u, \partial A_h) \leq C_3 h^s$, we derive an upper bound for the Gaussian measure of the symmetric difference $A \triangle A_h$ in the form

$$\mathbb{P}[A \triangle A_h] \leq C_4 \mathbb{P}[U_1 \in] - b_h - C_3 h^s, -b_h + C_3 h^s]],$$

where U_1 is distributed according to $\mathrm{N}(0,1)$ and $b_h := \|u_h^{\mathrm{MLFP}}\|_2$.

(P5) We define the affine linear function $\widetilde{G} := U_1 + b_h$ and apply the derived bound of the linear case (P2) to $\widetilde{G}$ to prove (2.16).

In the following sections, we provide full details of the steps (P1)–(P5). The proof of Proposition 2.11 requires (P1)–(P3) and a similar form of (P5) but does not require the upper bound of the Gaussian measure of $A \triangle A_h$ in (P4). Indeed, convexity of the failure domain is only required to prove (P4).

354   *Remark* 2.13. Given the idea from Remark 2.5, we conjecture that a similar error
355 bound as in (2.16) also holds in the case of pathwise ellipticity. Given an error
356 tolerance $\epsilon > 0$, we choose R such that $\mathbb{P}[U \notin B_R] \leq \epsilon/2$ and we bound the random
357 variable $C_{\mathrm{FE}}(\omega)$ in B_R. We propose to choose $\epsilon \ll P_f$ to ensure that the truncated
358 tails do not contain a large probability mass of the failure domain. We define the
359 quantities $P_f^\epsilon := \mathbb{P}[G(U) \leq 0 \cap U \in B_R]$ and $P_{f,h}^\epsilon := \mathbb{P}[G_h(U) \leq 0 \cap U \in B_R]$. By the
360 triangle inequality it holds that

361
362
$$|P_f - P_{f,h}| \leq |P_f - P_f^\epsilon| + |P_f^\epsilon - P_{f,h}^\epsilon| + |P_{f,h} - P_{f,h}^\epsilon| \leq |P_f^\epsilon - P_{f,h}^\epsilon| + \epsilon.$$

363 Thus, the restriction to B_R leads to an ϵ-error for the absolute error of the probability
364 of failure estimates. The proof of Theorem 2.12 can be used as a starting point to
365 derive a similar error bound for the absolute error $|P_f^\epsilon - P_{f,h}^\epsilon|$. The same idea can
366 be used to derive an error bound for the relative error with respect to the FORM
367 estimates in Proposition 2.11 in the case of pathwise ellipticity. However, providing
368 a complete proof of these bounds is out of the scope of this paper. We note that ϵ
369 and R are chosen with respect to the probability of failure P_f and do not depend on
370 $C_{\mathrm{FE}}(\omega)$. Choosing a small ϵ, requires a large radius R and a large upper bound for
371 $C_{\mathrm{FE}}(\omega)$ within B_R. If the user specified error tolerance ϵ is chosen, the upper bound
372 of $C_{\mathrm{FE}}(\omega)$ is constant for the whole analysis and does not blow up.
373 We conclude that the error bounds are useful also in cases where the diffusion co-
374 efficient is only pathwise elliptic and bounded, i.e., satisfies (2.6). In the numerical
375 experiments, we will only consider such settings. Indeed, the numerical results give
376 evidence for our conjecture.

377   *Remark* 2.14. We note that the approximation property of the LSF given in As-
378 sumption 2.5 determines the approximation property of the probability of failure. If
379 the approximation error of the LSF behaves in a more general form, (P1) and (P2)
380 can be directly applied to show that the approximation error of the probability of
381 failure behaves in the same manner. However, (P3) is only applicable for LSFs stem-
382 ming from an elliptic PDE and satisfying the regularity assumptions. Indeed, if it
383 is possible to show that (P3) holds for more general approximation properties of the
384 LSF, then (P4) and (P5) are directly applicable.
385 In Section 5.1, we consider an LSF which involves an ordinary differential equation
386 (ODE). For this example, we show that the distance of the failure domains behaves
387 as the convergence order of the applied time stepping scheme. Thus, (P3) is also valid
388 and our error bounds are applicable in this setting.

389   **2.4. Absolute error bound.** Under Assumption 2.5 and 2.6, we prove that
390 the upper bound for the absolute error of the probability of failure behaves as the
391 approximation error of the LSF, which proves (P1). This result and proof technique
392 are similar to [8, Lemma 3.4]. Considering equations (2.10) and (2.11), the approxi-
393 mation error is based on the symmetric difference $A \triangle A_h$. The following lemma gives
394 an upper bound for the absolute error.

395   LEMMA 2.15. *Let Assumptions 2.5 and 2.6 hold. Then, the absolute approxima-*
396 *tion error of the probability of failure is bounded in the following way:*

397 (2.17)
$$|P_f - P_{f,h}| \leq \mathbb{P}[G(U) \in] - C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]]$$
398
399 (2.18)
$$\leq 2C_L C_{\mathrm{FE}} h^s =: C_1 h^s.$$

Proof. Inserting the definitions of P_f and $P_{f,h}$ given in (2.10) and (2.11) in the left hand side of (2.17) we get

$$|P_f - P_{f,h}| = \left| \int_{\mathbb{R}^n} \left(I(G(u) \leq 0) - I(G_h(u) \leq 0) \right) \varphi_n(u) \mathrm{d}u \right|$$

$$= \left| \int_{\mathbb{R}^n} \left(I(G(u) \leq 0 \wedge G_h(u) > 0) - I(G(u) > 0 \wedge G_h(u) \leq 0) \right) \varphi_n(u) \mathrm{d}u \right|$$

$$(2.19) \qquad \leq \mathbb{P}[\{G(U) \leq 0\} \cap \{G_h(U) > 0\}] + \mathbb{P}[\{G(U) > 0\} \cap \{G_h(U) \leq 0\}]$$

$$= \mathbb{P}[A \triangle A_h],$$

where (2.19) follows from the triangle inequality. Using Assumption 2.5, we know that the case $(G(u) \leq 0 \wedge G_h(u) > 0)$ only occurs if $G(u) \in]-C_{\mathrm{FE}}h^s, 0] \cap G_h(u) \in]0, C_{\mathrm{FE}}h^s]$. Similarly, the case $(G_h(u) \leq 0 \wedge G(u) > 0)$ only occurs if $G_h(u) \in]-C_{\mathrm{FE}}h^s, 0]$ and $G(u) \in]0, C_{\mathrm{FE}}h^s]$. Therefore, the absolute approximation error is bounded by

$$|P_f - P_{f,h}| \leq \mathbb{P}[G(U) \in]-C_{\mathrm{FE}}h^s, 0] \cap G_h(U) \in]0, C_{\mathrm{FE}}h^s]]$$
$$+ \mathbb{P}[G_h(U) \in]-C_{\mathrm{FE}}h^s, 0] \cap G(U) \in]0, C_{\mathrm{FE}}h^s]].$$

Applying the multiplication rule $\mathbb{P}[B_1 \cap B_2] = \mathbb{P}[B_1 \mid B_2]\mathbb{P}[B_2]$ for $B_1, B_2 \in \mathcal{A}$ we get that

$$|P_f - P_{f,h}| \leq \mathbb{P}[G_h(U) \in]0, C_{\mathrm{FE}}h^s] \mid G(U) \in]-C_{\mathrm{FE}}h^s, 0]] \cdot \mathbb{P}[G(U) \in]-C_{\mathrm{FE}}h^s, 0]]$$
$$+ \mathbb{P}[G_h(U) \in]-C_{\mathrm{FE}}h^s, 0] \mid G(U) \in]0, C_{\mathrm{FE}}h^s]] \cdot \mathbb{P}[G(U) \in]0, C_{\mathrm{FE}}h^s]]$$
$$\leq \mathbb{P}[G(U) \in]-C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]],$$

where the last step follows from the fact that probabilities are always bounded by one. This proves inequality (2.17). To prove (2.18), we use the assumption on the local Lipschitz continuity of $\mathbb{P}[G(U) \leq \cdot]$

$$\mathbb{P}[G(U) \in]-C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]] \leq 2C_L C_{\mathrm{FE}}h^s. \qquad \square$$

Remark 2.16. By switching the roles of G by G_h in Lemma 2.15, we obtain

$$|P_f - P_{f,h}| \leq \mathbb{P}[G_h(U) \in]-C_{\mathrm{FE}}h^s, C_{\mathrm{FE}}h^s]] \leq 2C_{L,h}C_{\mathrm{FE}}h^s.$$

3. Affine linear limit-state function. The next step of the proofs of Proposition 2.11 and Theorem 2.12 is (P2). Here, we need to find an upper bound for the local Lipschitz constant C_L of Assumption 2.6 around the limit-state surface ∂A for the case where G is affine linear with respect to U.

Assumption 3.1. The LSF $G : \mathbb{R}^n \to \mathbb{R}$ is affine linear in the Gaussian random variable U, i.e., $G(U) = \alpha^T U + \beta$ where $\alpha \in \mathbb{R}^n$, $\beta > 0$ and $U \sim \mathrm{N}(0, \mathrm{Id}_n)$. Therefore, the probability of failure is

$$P_f = \mathbb{P}[G(U) \leq 0] = \mathbb{P}[\alpha^T U \leq -\beta] = \int_{-\infty}^{-\beta} \frac{1}{\sqrt{2\pi\|\alpha\|_2^2}} \exp\left(-\frac{u^2}{2\|\alpha\|_2^2} \right) \mathrm{d}u = F_W(-\beta),$$

where $W := \alpha^T U$ is distributed according to $\mathrm{N}(0, \|\alpha\|_2^2)$ and $F_W(\cdot)$ is the CDF of W.

The shift parameter β is assumed to be positive which yields that $G(0) > 0$. Moreover, we require that $-\beta + C_{\mathrm{FE}}h^s < 0$ since the lower and upper bounds in (3.1) are only defined for negative inputs and we evaluate these bounds at $w = -\beta + C_{\mathrm{FE}}h^s$. Together with Assumptions 2.5, 2.6 and 3.1, we prove statement (P2).

THEOREM 3.2. *Let Assumptions 2.5, 2.6 and 3.1 hold and $-\beta + C_{\mathrm{FE}} h^s < 0$. Then the relative approximation error of the probability of failure is bounded by*

$$\frac{|P_f - P_{f,h}|}{P_f} \leq C_2(\beta, \sigma, h^s, C_{\mathrm{FE}}) \cdot h^s,$$

where $\sigma^2 = \|\alpha\|_2^2$.

By Assumption 3.1, the probability of failure is directly given in terms of the CDF F_W. The goal is to derive an upper bound for the local Lipschitz constant C_L in Assumption 2.6, which depends linearly on the probability of failure. By Lemma 2.15, this is equivalent to deriving an upper bound for the local Lipschitz constant of $F_W(\cdot)$ in the interval $] - \beta - C_{\mathrm{FE}} h^s, -\beta + C_{\mathrm{FE}} h^s]$. We distinguish two cases in the proof. First, we assume that the approximation G_h is one-sided, i.e., $G(U) \leq G_h(U)$ almost surely. Secondly, we consider the non-one-sided case.

3.1. One-sided approximation.

Assumption 3.3. The approximation G_h of the LSF is one-sided with respect to the exact LSF G, that means

$$G(u) \leq G_h(u),$$

for all $u \in \mathbb{R}^n$ and $h > 0$.

Under Assumption 3.3, it follows that $P_{f,h} \leq P_f$ and $\mathbb{P}[G_h(U) \in] - C_{\mathrm{FE}} h^s, 0] \mid G(U) \in]0, C_{\mathrm{FE}} h^s]] = 0$. Hence, the bound for the absolute approximation error in Lemma 2.15 simplifies to

$$|P_f - P_{f,h}| \leq \mathbb{P}[G(U) \in] - C_{\mathrm{FE}} h^s, 0]] \leq C_L C_{\mathrm{FE}} h^s.$$

Therefore, it suffices to derive an upper bound for the local Lipschitz constant of $F_W(\cdot)$ within the interval $] - \beta - C_{\mathrm{FE}} h^s, -\beta]$. Observe that the derivative of the CDF $F_W(w)$ with respect to w is the PDF of the normal distribution with mean 0 and variance σ^2, denoted by $\varphi_W(w) = \exp\left(-w^2/(2\sigma^2)\right)/\sqrt{2\pi\sigma^2}$, which is strictly increasing for $w \in] - \infty, 0[$. Therefore, the local Lipschitz constant of F_W on the interval $] - \beta - C_{\mathrm{FE}} h^s, -\beta]$ is given by $\varphi_W(-\beta)$. The goal is to derive an upper bound for C_L in the form $C_L C_{\mathrm{FE}} \leq C_2 P_f$ for $C_2 > 0$. In order to derive this result, we consider the following bounds for the CDF.

PROPOSITION 3.4. *An upper and lower bound for the CDF $F_W(\cdot)$ on the interval $w \in] - \infty, 0[$ are given by*

$$(3.1) \qquad F_l(w) := -\frac{\sigma \exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{2\pi}} \frac{w}{w^2 + 1} \leq F_W(w) \leq -\frac{\sigma \exp\left(-\frac{w^2}{2\sigma^2}\right)}{w\sqrt{2\pi}} =: F_u(w).$$

The derivation of these bounds is given in [15].

With these bounds, we derive an upper bound for the local Lipschitz constant C_L having the desired form.

LEMMA 3.5. *Under Assumptions 2.6, 3.1 and 3.3, it holds that the local Lipschitz constant C_L within the interval $[-\beta - C_{\mathrm{FE}} h^s, -\beta]$ is bounded by*

$$C_L \leq \left(\frac{\beta}{\sigma^2} + \frac{1}{\beta} + \frac{1}{\beta\sigma^2} + \frac{1}{\beta^3}\right) P_f.$$

Proof. 1. The derivative of F_u is given by

$$F'_u(w) = \frac{\exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma} + \frac{\sigma \exp\left(-\frac{w^2}{2\sigma^2}\right)}{w^2\sqrt{2\pi}} = \frac{\exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}\left(1 + \frac{\sigma^2}{w^2}\right).$$

Hence, the derivative F'_u is also an upper bound for the PDF $\varphi_W(w)$. Moreover, $F'_u(w)$ is an increasing function for $w < 0$ since

$$F''_u(w) = \frac{-w\exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma^3} - \frac{\exp\left(-\frac{w^2}{2\sigma^2}\right)}{w\sqrt{2\pi}\sigma} - \frac{2\sigma\exp\left(-\frac{w^2}{2\sigma^2}\right)}{w^3\sqrt{2\pi}} > 0, \quad \text{for all } w < 0.$$

Therefore, the derivative F'_u at $w = -\beta$ gives us an upper bound for the local Lipschitz constant C_L.

2. F'_u can be written in terms of F_u:

$$(3.2) \qquad F'_u(-\beta) = -\frac{\sigma\exp\left(-\frac{\beta^2}{2\sigma^2}\right)}{\beta\sqrt{2\pi}}\left(-\frac{\beta}{\sigma^2} - \frac{1}{\beta}\right) = F_u(-\beta)\left(\frac{\beta}{\sigma^2} + \frac{1}{\beta}\right).$$

Since F_u is an upper bound for F_W, we know that $F_u(-\beta) \geq F_W(-\beta) = P_f$. Since F_l is a lower bound for F_W, we know that $F_l(-\beta) \leq P_f$. Combining these two statements, we get

$$(3.3) \qquad 1 \leq \frac{F_u(-\beta)}{P_f} \leq \frac{F_u(-\beta)}{F_l(-\beta)} = \frac{\beta^2 + 1}{\beta^2},$$

which yields $F_u(-\beta) \leq \left(1 + 1/\beta^2\right)P_f$.

Given (3.2) and (3.3), we conclude that

$$C_L = \varphi_W(-\beta) \leq F'_u(-\beta) = F_u(-\beta)\left(\frac{\beta}{\sigma^2} + \frac{1}{\beta}\right) \leq \left(\frac{\beta}{\sigma^2} + \frac{1}{\beta} + \frac{1}{\beta\sigma^2} + \frac{1}{\beta^3}\right)P_f. \quad \square$$

Combining the statements of Lemma 2.15 and Lemma 3.5, we conclude the proof of Theorem 3.2 with the constant

$$(3.4) \qquad C_{2,1}(\beta, \sigma) := \left(\frac{\beta}{\sigma^2} + \frac{1}{\beta} + \frac{1}{\beta\sigma^2} + \frac{1}{\beta^3}\right)C_{\text{FE}}.$$

Remark 3.6 (Sharper bounds). Applying the bounds of [1, 7.1.13], we can derive the following sharper upper and lower bounds of $F_W(w)$ for $w \in\,]-\infty, 0]$

$$(3.5) \qquad \sqrt{\frac{2}{\pi}}\frac{\exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{4 + w^2/\sigma^2} - w/\sigma} \leq F_W(w) \leq \sqrt{\frac{2}{\pi}}\frac{\exp\left(-\frac{w^2}{2\sigma^2}\right)}{\sqrt{2 + w^2/\sigma^2} - w/\sigma}.$$

Using (3.5), we derive a sharper bound for the Lipschitz constant C_L of the form

$$C_L C_{\text{FE}} \leq \left(\frac{\sqrt{4 + \beta^2/\sigma^2} + \beta/\sigma}{\sqrt{2 + \beta^2/\sigma^2} + \beta/\sigma}\right)\left(\frac{\beta}{\sigma^2} + \frac{\beta/\sqrt{2 + \beta^2/\sigma^2} + \sigma}{\sigma^2\sqrt{2 + \beta^2/\sigma^2} + \beta\sigma}\right)C_{\text{FE}}P_f$$

$$=: \widehat{C}_{2,1}(\beta, \sigma)P_f.$$

The proof works in a similar way as the proof of Lemma 3.5.

Figure 1 shows the constants $C_{2,1}$ and $\widehat{C}_{2,1}$ for varying β and $\sigma^2 \in \{0.1, 1.0, 10.0\}$ and $C_{\mathrm{FE}} = 1$. The x-axis shows the probability of failure $P_f = F_W(-\beta)$. We observe that the constants are large for small variances σ^2. Since $F_u(0)$ is not defined and $F_l(0) = 0$, the constant $C_{2,1}$ explodes to $+\infty$ for large probability of failures, i.e., $\beta \to 0$ while $\widehat{C}_{2,1}$ yields a small constant in this case. In both plots, we infer that the variance σ^2 has the main influence on the behaviour of the constants. The probability of failure has only a small influence. For the FORM estimate, the variance of the linearized LSF is the square of the norm of the gradient at the MLFP.
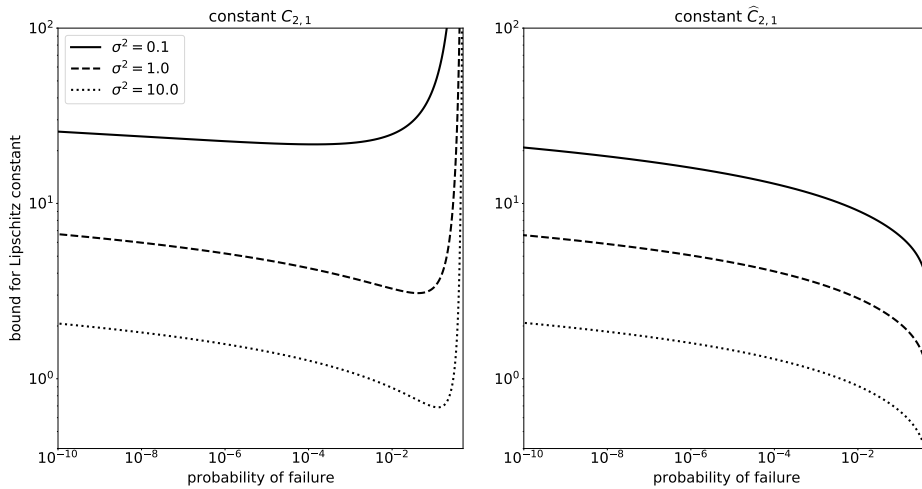

FIG. 1. *On the left: behaviour of $C_{2,1}$ for varying β and $\sigma^2 \in \{0.1, 1.0, 10.0\}$, where we assume that $G(U)$ is distributed according to $\mathrm{N}(\beta, \sigma^2)$ and $C_{\mathrm{FE}} = 1$. On the right: behaviour of $\widehat{C}_{2,1}$ for varying β and $\sigma^2 \in \{0.1, 1.0, 10.0\}$. The x-axis shows the respective probability of failure $P_f = F_W(-\beta)$ which lies in the interval $[10^{-10}, 1/2]$.*

3.2. Non-one-sided approximation. In general, we do not know if the LSF is one-sided, i.e., if $G(u) \leq G_h(u)$ for all $u \in \mathbb{R}^n$. Hence, we are required to bound the local Lipschitz constant in Assumption 2.6 within the interval $[-\beta - C_{\mathrm{FE}}h^s, -\beta + C_{\mathrm{FE}}h^s]$. Due to the fact that the bounds in (3.1) are not defined for $w \geq 0$, h must be chosen sufficiently small such that $-\beta + C_{\mathrm{FE}}h^s < 0$. Since we have already derived an upper bound for the local Lipschitz constant within the first half $[-\beta - C_{\mathrm{FE}}h^s, -\beta]$, we derive an upper bound for the second half $[-\beta, -\beta + C_{\mathrm{FE}}h^s]$. In this case, the local Lipschitz constant can be expressed by the derivative of the CDF $\varphi_W(z)$ at $z = -\beta + C_{\mathrm{FE}}h^s$. Applying the same steps as for the one-sided case, yields the following bound

$$C_L C_{\mathrm{FE}} \leq F_u'(-\beta + C_{\mathrm{FE}}h^s)C_{\mathrm{FE}}$$

$$\leq \left(\beta + \frac{1}{\beta}\right)\left(\frac{1}{\sigma^2} + \frac{1}{(\beta - C_{\mathrm{FE}}h^s)^2}\right)\exp\left(\frac{2\beta C_{\mathrm{FE}}h^s - C_{\mathrm{FE}}^2 h^{2s}}{2\sigma^2}\right)C_{\mathrm{FE}}P_f$$

$$(3.6) \qquad =: C_{2,2}(\beta, \sigma, h^s, C_{\mathrm{FE}})P_f.$$

Unfortunately, $C_{2,2}$ depends on the error bound $C_{\mathrm{FE}}h^s$ of the LSF approximation. We observe that $C_{2,2}$ converges to $C_{2,1}$ for $h \to 0$. We illustrate this in the following.

543 Figure 2 displays the constant $C_{2,2}/C_{\mathrm{FE}}$ while varying the step size h. We divide by
544 C_{FE} to eliminate the linear dependence in $C_{2,2}$. The order of convergence is either
545 $s = 1$ or $s = 2$. Moreover, different values for C_{FE} are considered. The variance is
546 $\sigma^2 = 1$ and $\beta = 4$. This yields a probability of failure of order 10^{-5}. We observe that
547 the constant $C_{2,2}/C_{\mathrm{FE}}$ is large for large values of C_{FE}. For the convergence order
548 $s = 1$, the step size h should be smaller than 10^{-2} to ensure a small constant even for
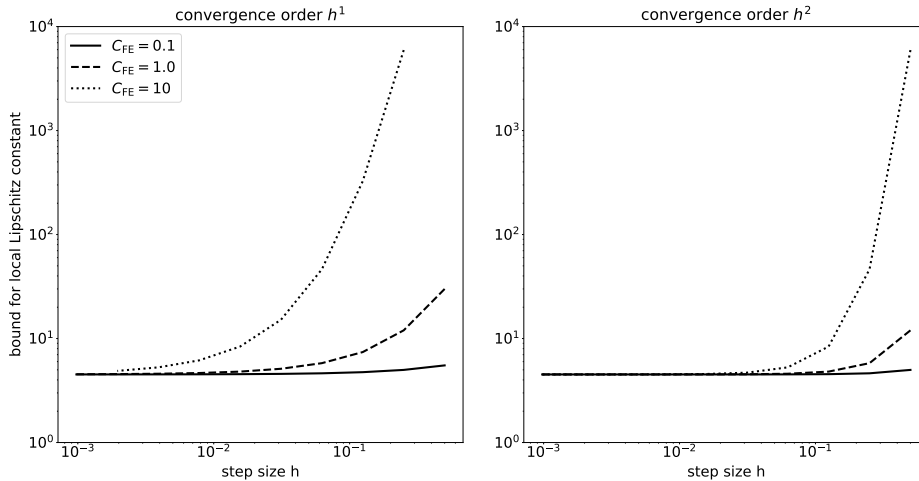large C_{FE}. For $s = 2$, the step size should be smaller than 10^{-1}.

FIG. 2. *On the left: behaviour of $C_{2,2}/C_{\mathrm{FE}}$ for varying the step size h and $C_{\mathrm{FE}} \in \{0.1, 1.0, 10.0\}$. The convergence order is $s = 1$. On the right: behaviour of $C_{2,2}/C_{\mathrm{FE}}$ for varying the step size h and $C_{\mathrm{FE}} \in \{0.1, 1.0, 10.0\}$. The convergence order is $s = 2$. In both plots, $\sigma^2 = 1$ and $\beta = 4$.*

549

550   *Remark* 3.7. Following the same steps to derive the constant $C_{2,2}$, one could also
551 derive a constant $\widehat{C}_{2,2}$ which is based on the sharper CDF bounds (3.5). This constant
552 also depends on $C_{\mathrm{FE}} h^s$.

553   Finally, we note that the result above allows us to prove Theorem 3.2. Combining
554 the constants in (3.4) and (3.6), we obtain the asserted inequality with

$$C_2(\beta, \sigma, h^s, C_{\mathrm{FE}}) := C_{2,1}(\beta, \sigma) + C_{2,2}(\beta, \sigma, h^s, C_{\mathrm{FE}}).$$

557   **4. Error Analysis with Optimal Control and FORM.** In this section, we
558 prove statements (P3)–(P5) which will conclude the proofs of Proposition 2.11 and
559 Theorem 2.12. In the affine linear case, we know that the error of the LSF in As-
560 sumption 2.5 is directly related to the distance between the exact and approximate
561 limit-state surface. However, this direct relation is not obvious in a more general
562 setting.
563 In the subsequent step, we derive an upper bound for the distance between the ex-
564 act and approximate limit-state surface ∂A and ∂A_h, respectively, in the case where
565 the LSF satisfies the assumptions of Proposition 2.11 and Theorem 2.12. This up-
566 per bound is based on results from optimal control theory. Finally, we estimate the
567 Gaussian measure of the symmetric difference of the failure domains A and A_h by the
568 FORM approximation $P_{f,h}^{\mathrm{FORM}}$.

4.1. Theoretical results from optimal control. We will formulate item (P3) in Theorem 4.2. For the proof of this result, we require several concepts from the theory of optimal control. We discuss those results below; based on the works [26, 32].

We commence with the unconstrained optimal control problem. It is given by

$$(4.1) \qquad \min_{u \in \mathbb{R}^n} \quad J(u) := \frac{1}{2}\|\widehat{G}(u) - \overline{g}\|_2^2 + \frac{\alpha}{2}\|u - \overline{u}\|_2^2,$$

where $u \in \mathbb{R}^n$ is the unknown parameter and $\widehat{G} : \mathbb{R}^n \to \mathbb{R}^{n_m}$ is the observation operator. The observation operator $\widehat{G}(u) = \mathcal{B}y(\cdot, u)$ implicitly depends on the solution $y(\cdot, u) \in V = H_0^1(D)$ of the weak formulation of an elliptic PDE

$$(4.2) \qquad \int_D a(x, u)\nabla_x y(x, u) \cdot \nabla_x v(x)\mathrm{d}x = \int_D f(x)v(x)\mathrm{d}x \quad \forall v \in V,$$

where $u \in \mathbb{R}^n$ is a fixed parameter, $f \in L^2(D)$, and $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, is an open, bounded, convex polygonal domain. The function $a(x, \cdot) : \mathbb{R}^n \to \mathbb{R}$ is assumed to be three times continuously differentiable for all $x \in D$ and $a(\cdot, u) \in L^\infty(D)$ for all $u \in \mathbb{R}^n$. The discretized unconstrained optimal control problem is denoted as

$$(4.3) \qquad \min_{u \in \mathbb{R}^n} \quad J_h(u) := \frac{1}{2}\|\widehat{G}_h(u) - \overline{g}\|_2^2 + \frac{\alpha}{2}\|u - \overline{u}\|_2^2,$$

where $\widehat{G}_h(u) = \mathcal{B}_h y_h(\cdot, u)$ is the discretized observation operator. Additionally, we assume that $\mathcal{B}$ and $\mathcal{B}_h$ are linear and bounded with respect to y and y_h, respectively. The parameters $\alpha \geq 0$ and $\overline{u}$ are regularizing parameters, while $\overline{g} \in \mathbb{R}^{n_m}$ is a given vector of, e.g., measurements.

In the following, we derive the necessary and sufficient optimality conditions for (4.1). The *first-order necessary optimality condition* is given as

$$\nabla_u J(u) = 0 \iff (\mathcal{D}\widehat{G}(u))^T\widehat{G}(u) + \alpha u = (\mathcal{D}\widehat{G}(u))^T\overline{g} + \alpha\overline{u},$$

where $\mathcal{D}\widehat{G}$ is the Jacobian of the observation operator $\widehat{G}$ with respect to u. The *second-order optimality condition* is satisfied, if some *coercivity parameter* $\gamma > 0$ exists, with

$$z^T\nabla_u^2 J(u^*)z \geq \gamma\|z\|_2^2 \quad \forall z \in \mathbb{R}^n,$$

where $\nabla_u^2 J(u^*)$ denotes the Hessian of J with respect to u. We call $u^* \in \mathbb{R}^n$ a *stable solution of* (4.1), if it satisfies both, the first and second-order optimality conditions.

THEOREM 4.1. *Let Assumption 2.3, 2.4 (ii) and 2.5 hold for the weak formulation* (4.2) *and the operator $\mathcal{B}$. Let $u^* \in \mathbb{R}^n$ be a stable solution of* (4.1) *such that $\nabla_u^2 J(u)$ is coercive with parameter $\gamma > 0$. Then for $h > 0$ sufficiently small, there exists a stable solution $u_h^* \in \mathbb{R}^n$ of the discrete problem* (4.3) *such that the following a priori error estimate holds*

$$\|u^* - u_h^*\|_2 \leq \frac{1}{\gamma}\|\nabla_u J(u^*) - \nabla_u J_h(u^*)\|_2.$$

Proof. The proof is given in [32, Theorem 3.4.1 and 3.4.2]. □

4.2. Back to rare event estimation. We now apply Theorem 4.1 to prove statement (P3).

THEOREM 4.2. *Let Assumptions 2.3, 2.4 (ii), 2.5 and 2.10 hold. Then for all* $u \in \partial A$ *it holds that*

$$\mathrm{dist}(u, \partial A_h) \leq C_3(h) \cdot h^s,$$

for $h > 0$ sufficiently small.

Proof. We apply Theorem 4.1 to an appropriate optimal control problem with $\alpha > 0$ to show that for $u^* \in \partial A$ exists a $\widehat{u}_h \in \mathbb{R}^n$ such that the distance between these points behaves as $\mathcal{O}(h^s)$. Then, we consider the limit $\alpha \to 0$.

Let $u^* \in \partial A$ be a point on the exact limit-state surface, i.e., $G(u^*) = 0$. We investigate the following optimal control problem

$$(4.4) \qquad \min_{u \in \mathbb{R}^n} \quad J_\alpha(u) := \frac{1}{2} G(u)^2 + \frac{\alpha}{2} \|u - u^*\|_2^2,$$

where we set $\overline{g} = 0$ and $\alpha > 0$. For $\alpha = 0$, there is no stable solution since for all $u \in \partial A$ it holds that $J_0(u) = 0$. Hence, the second order sufficient condition is violated. For $\alpha > 0$, the gradient and Hessian matrix of J_α are given as

$$\nabla_u J_\alpha(u) = G(u) \nabla_u G(u) + \alpha(u - u^*),$$
$$\nabla_u^2 J_\alpha(u) = G(u) \nabla_u^2 G(u) + \nabla_u G(u)(\nabla_u G(u))^T + \alpha I.$$

The point u^* is the unique global minimizer of (4.4) since $J_\alpha(u^*) = 0$ and $J_\alpha(u) > 0$ for all $u \in \mathbb{R}^n \setminus \{u^*\}$. Hence, u^* is a stable solution. For the discretization parameter $h > 0$, we define the discretized version of (4.4) as

$$(4.5) \qquad \min_{u \in \mathbb{R}^n} \quad J_{h,\alpha}(u) := \frac{1}{2} G_h(u)^2 + \frac{\alpha}{2} \|u - u^*\|_2^2.$$

By Theorem 4.1, there exists a point $\widehat{u}_h$ in a neighborhood of u^* such that $\widehat{u}_h$ is a stable solution of (4.5). Thus, we know that $\nabla_u J_{h,\alpha}(\widehat{u}_h) = 0$, which yields

$$G_h(\widehat{u}_h) \nabla_u G_h(\widehat{u}_h) + \alpha(\widehat{u}_h - u^*) = 0.$$

If $G_h(\widehat{u}_h) = 0$ we get that $\widehat{u}_h = u^*$ and the claim is proved. Now, we consider the case $G_h(\widehat{u}_h) \neq 0$ and we denote $u_h^* \in \partial A_h$ as the point on ∂A_h that has minimal distance to the point u^*. Moreover, we define the set of points

$$E := \{ p \in \mathbb{R}^n : |\cos(\sphericalangle(p, \nabla_u G(u^*)))| \geq |\cos(\sphericalangle(u^* - \widehat{u}_h, \nabla_u G(u^*)))| \},$$

where $\sphericalangle(a, b) := \arccos\left(a^T b / (\|a\|_2 \|b\|_2)\right)$, for $a, b \in \mathbb{R}^n$. The set E contains all directions $p \in \mathbb{R}^n$, which admit a smaller or equal angle with $\nabla_u G(u^*)$ than the direction pointing from u^* to $\widehat{u}_h$. For all $p \in E$, we conclude that

$$p^T \nabla_u^2 J_\alpha(u^*) p = p^T \nabla_u G(u^*)(\nabla_u G(u^*))^T p + \alpha \|p\|_2^2$$
$$\geq \left(\cos^2(\sphericalangle(u^* - \widehat{u}_h, \nabla_u G(u^*))) \|\nabla_u G(u^*)\|_2^2 + \alpha\right) \|p\|_2^2.$$

Thus, the Hessian matrix $\nabla_u^2 J_\alpha(u^*)$ is coercive for all $p \in E$ with parameter

$$\gamma = \cos^2(\sphericalangle(u^* - \widehat{u}_h, \nabla_u G(u^*))) \|\nabla_u G(u^*)\|_2^2 + \alpha.$$

Since $u^* - \widehat{u} \in E$, we can perform the proof of Theorem 4.1 only for directions $p \in E$ which yields that

$$\|u^* - \widehat{u}_h\|_2 \leq \frac{\|\nabla_u J_\alpha(u^*) - \nabla_u J_{h,\alpha}(u^*)\|_2}{\cos^2\left(\sphericalangle(u^* - \widehat{u}_h, \nabla_u G(u^*))\right)\|\nabla_u G(u^*)\|_2^2 + \alpha}$$

$$\leq \frac{\|\nabla_u G_h(u^*)\|_2}{\cos^2\left(\sphericalangle(u^* - \widehat{u}_h, \nabla_u G(u^*))\right)\|\nabla_u G(u^*)\|_2^2 + \alpha} C_{\mathrm{FE}} h^s,$$

since it holds $\|\nabla_u J_{h,\alpha}(u^*)\|_2 = |G_h(u^*)|\|\nabla_u G_h(u^*)\|_2 \leq C_{\mathrm{FE}} h^s \|\nabla_u G_h(u^*)\|_2$ and $\nabla_u J_\alpha(u^*) = 0$. Applying the limit, $\alpha \to 0$ we conclude that $G_h(\widehat{u}_h) \to 0$ and, thus, $\widehat{u}_h \to u_h^* \in \partial A_h$. By Assumption 2.10, it holds that $\cos^2\left(\sphericalangle(u^* - u_h^*, \nabla_u G(u^*))\right) \geq \nu_h^2$ and we conclude that

$$\|u^* - u_h^*\| \leq \frac{\|\nabla_u G_h(u^*)\|_2}{\nu_h^2 \|\nabla_u G(u^*)\|_2^2} C_{\mathrm{FE}} h^s.$$

Since this holds true for all $u^* \in \partial A$, we define

$$C_3(h) := C_{\mathrm{FE}} \cdot \sup_{u \in \partial A} \frac{\|\nabla_u G_h(u)\|_2}{\nu_h^2 \|\nabla_u G(u)\|_2^2}$$

and the desired statement is proved. $\square$

Remark 4.3. For the limit $h \to 0$, it holds that $G_h(u) \to G(u)$ for all $u \in \mathbb{R}^n$. Therefore, the limit-state surface ∂A_h converges to ∂A and $u_h \to u$. Hence, it holds $|\cos\left(\sphericalangle(u - u_h, \nabla_u G(u))\right)| \to 1$, where $u_h \in \partial A_h$ is the point on ∂A_h that has smallest distance to $u \in \partial A$. Thus, $\nu_h \to 1$, as $h \to 0$. If in addition $\nabla_u G_h(u)$ converges uniformly to $\nabla_u G(u)$, it yields that

$$\lim_{h \to 0} C_3(h) = C_{\mathrm{FE}} \cdot \sup_{u \in \partial A} 1/\|\nabla_u G(u)\|_2.$$

With (P1)–(P3), we can now give the proof of Proposition 2.11.

Proof of Proposition 2.11. We denote the distances of the MLFPs to the origin as $b = \|u^{\mathrm{MLFP}}\|_2$ and $b_h = \|u_h^{\mathrm{MLFP}}\|_2$. By definition, we know that $P_f^{\mathrm{FORM}} = \Phi(-b) = \mathbb{P}[U_1 \leq -b]$, where U_1 is a one-dimensional random variable distributed according to $N(0,1)$. Similar, $P_{f,h}^{\mathrm{FORM}} = \Phi(-b_h) = \mathbb{P}[U_1 \leq -b_h]$. From Theorem 4.2, we know that the distance between ∂A and ∂A_h is bounded from above by $C_3(h) h^s$. Thus, $|b - b_h| \leq C_3(h) h^s$. This yields that the absolute error $|P_f^{\mathrm{FORM}} - P_{f,h}^{\mathrm{FORM}}|$ is bounded from above by

(4.6) $|P_f^{\mathrm{FORM}} - P_{f,h}^{\mathrm{FORM}}| = |\mathbb{P}[U_1 \leq -b] - \mathbb{P}[U_1 \leq -b_h]|$

(4.7) $\leq \mathbb{P}[U_1 \in] - b - C_3(h) h^s, -b + C_3(h) h^s]],$

where we apply similar steps as in the proof of Lemma 2.15. The probability term in (4.7) is equal to $\mathbb{P}[\widetilde{G}(U_1) \in] - C_3(h) h^s, C_3(h) h^s]]$ where the LSF $\widetilde{G}(U_1) := U_1 + b$ satisfies (2.13) in Assumption 2.5 with $C_{\mathrm{FE}} = C_3(h)$. By definition it holds $P_f^{\mathrm{FORM}} = \mathbb{P}[\widetilde{G}(U_1) \leq 0]$. Since $\widetilde{G}$ is affine linear, we apply Theorem 3.2 to the LSF $\widetilde{G}$ with $\sigma = 1$ and $\beta = b$ which yields

(4.8) $\mathbb{P}[\widetilde{G}(U_1) \in] - C_3(h) h^s, C_3(h) h^s]] \leq C_2(b, 1, h^s, C_3(h)) h^s P_f^{\mathrm{FORM}}.$

Finally, combining (4.7) and (4.8) we conclude that

$$|P_f^{\mathrm{FORM}} - P_{f,h}^{\mathrm{FORM}}| \leq C_2(b, 1, h^s, C_3(h)) h^s P_f^{\mathrm{FORM}} =: \widehat{C}^{\mathrm{FORM}} h^s P_f^{\mathrm{FORM}}.   \square$$

The proof of Theorem 2.12 works in a similar way as the proof of Proposition 2.11. However, the inequalities in (4.6) and (4.7) do not hold directly for the absolute error $|P_f - P_{f,h}|$, which is upper bounded by the Gaussian measure of the set $S := A \triangle A_h$. The following theorem provides an upper bound of the Gaussian measure of S which is similar to (4.7). In this theorem, the convexity of the failure domains is required, i.e., we assume that Assumption 2.9 holds. We switch the roles of P_f and $P_{f,h}$, which yields that the derived error bound depends on $P_{f,h}^{\text{FORM}}$ and not on P_f^{FORM}.

THEOREM 4.4. *Let Assumption 2.4 (i) and 2.9 hold. Moreover, we assume that for all $u_h \in \partial A_h$ it holds that*

(4.9)
$$\text{dist}(u_h, \partial A) \leq \widetilde{C} h^s.$$

The distance between the origin and ∂A_h is denoted as b_h. Then, an upper bound for the Gaussian measure of the symmetric difference of A and A_h is given by

$$\mathbb{P}[U \in A \triangle A_h] \leq C_4(n)\mathbb{P}[U_1 \in] - b_h - \widetilde{C} h^s, -b_h + \widetilde{C} h^s]],$$

where U_1 is distributed according to $\mathrm{N}(0, 1)$. The constant C_4 is given by

$$C_4(n) = 1 + \pi^{1/2}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)},$$

where $\Gamma(\cdot)$ is the Gamma function.

Proof. For $n = 1$, the statement directly follows from assumption (4.9) with $C_4(n) = 1$. Consider $n > 1$. By the rotation invariance of the Gaussian measure $\mathbb{P}[U \in \cdot] = \mathrm{N}(0, \mathrm{Id}_n)$, we assume, without loss of generality, that the point with smallest distance to the origin is given by $b_h^* = (0, \dots, 0, -b_h)^T \in \mathbb{R}^n$, thus, $u_h^{\text{MLFP}} = b_h^*$. We denote the sets of interest by $D_1 := A \triangle A_h$ and $D_2 := \{u \in \mathbb{R}^n : u_n \in] - b_h - \widetilde{C} h^s, -b_h + \widetilde{C} h^s]\}$. First, we consider the closed ball around the origin $B_R := \{u \in \mathbb{R}^n : \|u\| \leq R\}$ and we show that $\mathbb{P}[U \in D_1 \cap B_R] \leq C_4(n)\mathbb{P}[U \in D_2 \cap B_R]$. Afterwards, we consider the limit $R \to \infty$.
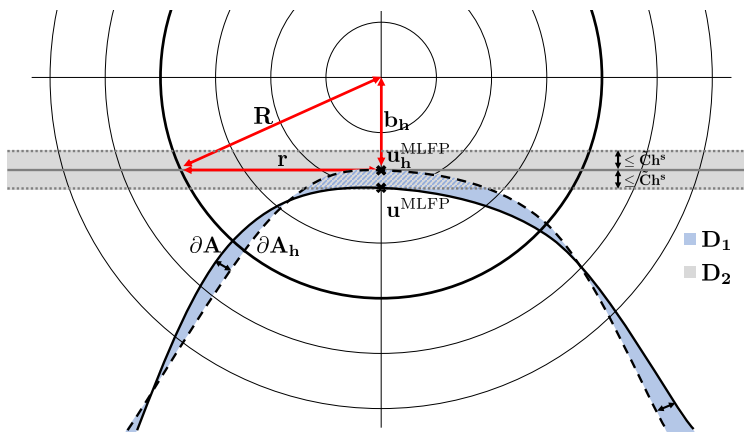
FIG. 3. *Illustration of the limit-state surfaces ∂A and ∂A_h in 2D. The illustrations shows that the Gaussian measure of $A \triangle A_h$ can be translated to a Gaussian measure at the MLFP u_h^{MLFP}.*

For $R \leq b_h$, it follows that $D_1 \cap B_R = \emptyset$, and the statement is valid. Therefore, we consider $R > b_h$. First, we determine the scaling factor of the $n-1$ dimensional Lebesgue measure of the transformation of a convex curve to a hyperplane. Therefore, we consider the sets $E_1 = B_R \cap \partial A_h$ and $E_2 = B_R \cap \{u \in \mathbb{R}^n : u_n = -b_h\}$. Note that E_2 is an $n-1$-dimensional ball centred in b_h^*. We visualize this in Figure 3. We call the radius of this ball $r := \sqrt{R^2 - b_h^2}$. Thus, the $n-1$-dimensional Lebesgue measure of E_2 is equal to the volume of the ball B_r in $n-1$ dimensions, which is given by

$$\lambda_{n-1}(E_2) = \frac{\pi^{(n-1)/2} r^{n-1}}{\Gamma((n+1)/2)}.$$

By the convexity of A_h, we conclude that an upper bound for the $n-1$ dimensional Lebesgue measure of E_1 is given by the surface measure of the set $B_R \cap \{u \in \mathbb{R}^n : u_n \leq -b_h\}$. This surface measure is bounded from above by the sum of the volume of the ball B_r in $n-1$ dimensions and $1/2$ of the surface of the ball B_r in n dimensions which yields

$$\lambda_{n-1}(E_1) \leq \frac{\pi^{(n-1)/2} r^{n-1}}{\Gamma((n+1)/2)} + \frac{\pi^{n/2} r^{n-1}}{\Gamma(n/2)}.$$

Hence, the fraction of the two $n-1$ dimensional Lebesgue measures is bounded by

$$(4.10) \qquad \frac{\lambda_{n-1}(E_1)}{\lambda_{n-1}(E_2)} \leq 1 + \pi^{1/2} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} = C_4(n).$$

The formulas for the volume and surface of a ball in n dimensions are given in [23, 5.19(iii)]. Inequality (4.10) bounds the scaling factor of the length of the curve E_1 with respect to the hyperplane E_2. Applying this result, we can transform the set D_1 into D_2. The probability of interest is given by

$$\mathbb{P}[U \in D_1 \cap B_R] = \int_{u \in \mathbb{R}^n} I(u \in D_1 \cap B_R) \varphi_n(u) \mathrm{d}u$$

$$= \frac{1}{(2\pi)^{n/2}} \int_{B_R} I(u \in D_1) \exp\left(-\frac{\|u\|_2^2}{2}\right) \mathrm{d}u$$

$$= \frac{1}{(2\pi)^{n/2}} \int_{r=0}^{r=R} \int_{S_{n-1}(r)} I(u \in D_1) \exp\left(-\frac{r^2}{2}\right) \mathrm{d}s \mathrm{d}r,$$

where $S_{n-1}(r) = \{u \in \mathbb{R}^n : \|u\|_2 = r\}$ is the surface of B_r. Since the distance of A and A_h is always smaller than $\widetilde{C} h^s$, the $n-1$ dimensional Lebesgue measure of the intersection $D_1 \cap S_{n-1}(r)$ is smaller or equal than the $n-1$ dimensional Lebesgue measure of the intersection $D_2 \cap S_{n-1}(r)$. Hence, since the standard Gaussian density is constant on $S_{n-1}(r)$ for all $r \geq 0$ and applying the transformation of D_1 to D_2 it follows that

$$\mathbb{P}[U \in D_1 \cap B_R] \leq C_4(n) \frac{1}{(2\pi)^{n/2}} \int_{r=0}^{r=R} \int_{S_{n-1}(r)} I(u \in D_2) \exp\left(-\frac{r^2}{2}\right) \mathrm{d}s \mathrm{d}r$$

$$= C_4(n) \mathbb{P}[U \in D_2 \cap B_R]$$

$$\leq C_4(n) \mathbb{P}[U \in D_2]$$

$$= C_4(n) \mathbb{P}[U_1 \in] - b_h - \widetilde{C} h^s, -b_h + \widetilde{C} h^s]].$$

Taking the limit $R \to \infty$ we get the desired statement, due to the continuity of measures. ☐

Remark 4.5. Unfortunately, taking the limit $n \to \infty$ yields $C_4(n) \to \infty$. Thus, this result does not easily generalise with respect to infinite-dimensional settings. However, the growth of $C_4(n)$ is $\mathcal{O}(n^{1/2})$ as visualised in Figure 4. Hence, even in high dimensions, the constant is reasonably small.
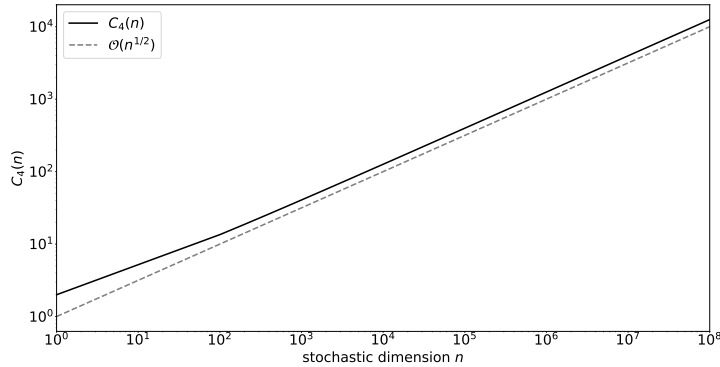

FIG. 4. *Behaviour of $C_4(n)$ for varying the stochastic dimension n.*

Having collected the proofs for (P1)–(P4), we can now give the proof of (P5) which proves Theorem 2.12.

Proof of Theorem 2.12. We apply similar steps as in the proof of Proposition 2.11, but we switch the roles of P_f and $P_{f,h}$. By Theorem 4.2 and applying Theorem 4.4 with $\widetilde{C} = C_3(h)$, we know that the absolute error of the probability of failure is bounded by

$$(4.11) \qquad |P_f - P_{f,h}| \leq C_4(n)\mathbb{P}[U_1 \in] - b_h - C_3(h)h^s, -b_h + C_3(h)h^s]].$$

Defining the LSF $\widetilde{G}_h(U_1) := U_1 + b_h$ and assuming that $\widetilde{G}$ satisfies (2.13) in Assumption 2.5 with $\widetilde{C}_{\mathrm{FE}} = C_3(h)$ yields

$$(4.12) \qquad \mathbb{P}[\widetilde{G}_h(U_1) \in] - C_3(h)h^s, C_3(h)h^s]] \leq C_2(b_h, 1, h^s, C_3(h))h^s P_{f,h}^{\mathrm{FORM}}.$$

Finally, combining (4.11) and (4.12) we conclude the proof of Theorem 2.12 with

$$|P_f - P_{f,h}| \leq C_2(b_h, 1, h^s, C_3(h)) \cdot C_4(n)h^s P_{f,h}^{\mathrm{FORM}} =: \widehat{C}h^s P_{f,h}^{\mathrm{FORM}}. \qquad ☐$$

Remark 4.6. We note that the assumptions on the regularity of the diffusion coefficient, as given in Assumption 2.3 and 2.4, are only relevant to prove Theorem 4.1 and 4.2, respectively. If the approximation error bound of the LSF behaves in another manner as in Assumption 2.5 and if it is possible to show that the distance between ∂A and ∂A_h behaves in the same manner, then the error bounds in Proposition 2.11 and Theorem 2.12 hold for this error bound.

5. Numerical Experiments. We now illustrate our results in several numerical experiments. We start with a one-dimensional parameter space example where the LSF involves an ODE, not a PDE. Then, we consider rare events that depend on

elliptic PDEs with stochastic spaces of different dimensions. In all experiments, the approximation error of G and G_h is not uniformly bounded as required in Assumption 2.5. We consider these settings to test the conjecture in Remark 2.13. Indeed, our provided error bounds in Proposition 2.11 and Theorem 2.12 are also observed in these non-uniform cases. In the first experiment, we observe that the distances of the failure surfaces behave as the discretization error. Following Remark 2.14, this behaviour enables to consider only (P4) and (P5) and we expect that the provided error bounds hold in this setting.

5.1. ODE, 1-dimensional parameter space. In the following example, which is also considered in [31], the LSF depends on the solution of an ODE with a one-dimensional Gaussian random parameter. Hence, this example does not actually depend on an elliptic PDE. We study it, since we can compute all quantities of interest analytically.

Let $y : [0,1] \times \Omega \to \mathbb{R}$ such that for $\mathbb{P}$-a.e. $\omega \in \Omega$, we have

$$\frac{\partial y(t,\omega)}{\partial t} = -U(\omega)y(t,\omega), t \in (0,1), \text{ with initial condition } y(0,\omega) = 1,$$

where $U \sim N(0,1)$ is a stochastic parameter. The exact solution to this ODE is given by $y(t,\omega) = \exp(-U(\omega)t)$ – a log-normal stochastic process. Failure is defined as the event that the solution $y(\cdot,\omega)$ is larger than $y_{\max} = 40$ at $t = 1$, which can be written in terms of the LSF $G(u) = y_{\max} - \exp(-u) \leq 0$. Hence, failure occurs if $u \leq -\log(y_{\max})$ and the exact probability of failure is equal to $P_f = \Phi(-\log(y_{\max})) \approx 1.13 \cdot 10^{-4}$. The MLFP is given by $u^{\text{MLFP}} = -\log(y_{\max})$.

Using the explicit Euler scheme to solve the ODE numerically, we derive the approximate solution $y_h(t = 1,\omega) = (1 - U(\omega)h)^{1/h}$, where $h > 0$ is the time step size. The explicit Euler scheme is convergent of order one, see [20, Section 6.3], i.e., $|y(t,\omega) - y_h(t,\omega)| = \mathcal{O}(h)$ for a fixed $\omega \in \Omega$. The approximate LSF is $G_h(u) = y_{\max} - (1 - uh)^{1/h}$. Hence, failure occurs if $u \leq (1 - y_{\max}^h)/h$ and the approximate probability of failure is equal to $P_{f,h} = \Phi((1 - y_{\max}^h)/h)$. The approximate MLFP is given by $u_h^{\text{MLFP}} = (1 - y_{\max}^h)/h$.

Since the space of the stochastic parameter space is one-dimensional and the exact and approximate failure domains are half-rays, it holds $P_f = P_f^{\text{FORM}}$ and $P_{f,h} = P_{f,h}^{\text{FORM}}$. Thus, the error bound of Proposition 2.11 and Theorem 2.12 yield a bound for the relative error $|P_f - P_{f,h}|/P_f$. As mentioned, the setting of this example is different to the setting of Proposition 2.11 and Theorem 2.12. Moreover, the approximation error of the LSF is not uniformly bounded. Figure 5 shows that the distance of the failure domains scales as $\mathcal{O}(h)$. Thus, we expect that the relative error of the probability of failure approximations has order $\mathcal{O}(h)$ of convergence for $h > 0$ sufficiently small.

As a second time stepping method, we consider the Crank–Nicolson scheme as given in [20, Chapter 9]. Applying the discretization rule, we get the approximate solution

$$\widetilde{y}_h(t = 1,\omega) = \left(\frac{1 - hU(\omega)/2}{1 + hU(\omega)/2}\right)^{1/h}.$$

Hence, the approximate probability of failure is given by $\widetilde{P}_{f,h} = \Phi(2h^{-1}(1-y_{\max}^h)/(1+y_{\max}^h))$. Since the Crank–Nicolson scheme is convergent of order 2, we expect that the relative error of the probability of failure approximations has order $\mathcal{O}(h^2)$ of convergence for $h > 0$ sufficiently small.

Figure 5 shows the approximate probability of failure by the explicit Euler and the

Crank–Nicolson scheme for the step sizes $h_\ell = 1/2^\ell$, for $\ell = 0, \ldots, 9$. We observe that the approximations computed with both these methods approach the exact probability of failure as h decreases. Moreover, we observe that the distance between the exact and approximate MLFPs converges in the same order as the PDE discretization error. Hence, the statement of Theorem 4.2 is also valid and we conclude that Proposition 2.11 and Theorem 2.12 are also applicable for this setting. On the right plot, we observe that the relative error of the explicit Euler approximations has order $\mathcal{O}(h)$ of convergence while the relative error of the Crank–Nicolson approximations has order $\mathcal{O}(h^2)$ of convergence. These are exactly the bounds which we get from theoretical discussions. For large h, we observe a plateau behaviour for the explicit Euler scheme. This is due to the fact that the Euler approximation $P_{f,h}$ is much smaller than P_f for large h. Hence, the relative error is nearly equal to one for large time step sizes h, until the convergence sets in.
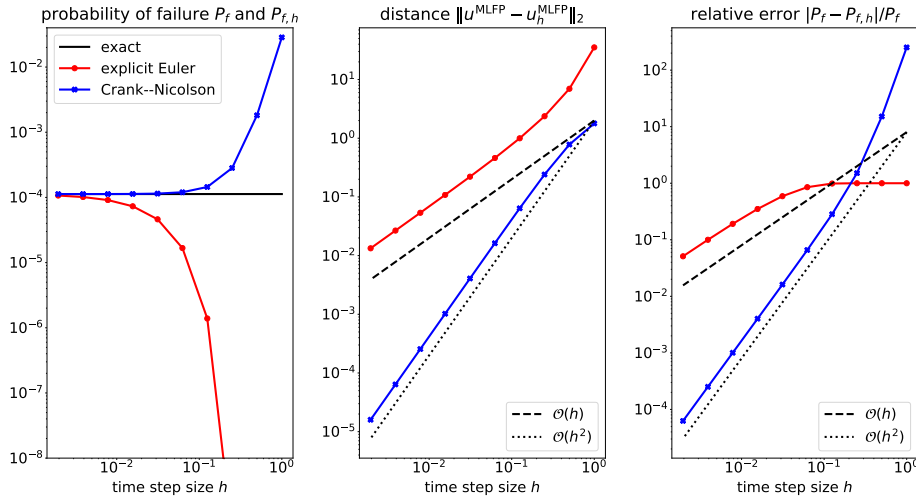

FIG. 5. *On the left: approximate probability of failure by the explicit Euler and Crank–Nicolson scheme for varying the time step size h. The black line shows the exact probability of failure. In the middle: distance between the exact and approximate MLFPs. The black lines show the order of convergence. On the right: relative error of the approximations with respect to the exact probability of failure.*

5.2. 2-Dimensional parameter space. The following example considers an LSF which depends on a two-dimensional stochastic parameter and is also considered in [9, 13]. In this case, the FORM estimate is not equal to the exact probability of failure. However, we can still derive analytical expressions for the exact and approximate LSF as well as for the exact and approximate limit-state surfaces. On the domain $D = (0, 1)$, we seek a solution $y : \overline{D} \times \Omega \to \mathbb{R}$ that solves the following elliptic boundary value problem

$$(5.1) \qquad -\frac{\partial}{\partial x}\left(\exp\left(\frac{U_1(\omega)}{3} - 3\right)\frac{\partial}{\partial x}y(x, \omega)\right) = 1 - x, \quad \text{for } x \in (0, 1),$$

$$(5.2) \qquad \text{such that } y(0, \omega) = 0 \text{ and } y(1, \omega) = U_2(\omega)$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$. The random variables U_1 and U_2 are independent and standard normally distributed. The exact solution of this problem is

$$y(x, \omega) = U_2(\omega)x + \exp\left(-U_1(\omega)/3 + 3\right)\left(x^3/6 - x^2/2 + x/3\right).$$

Failure is defined as the event that the solution $y(\cdot, \omega)$ is smaller than $y_{\max} = -1/3$ at $\widehat{x} = 1/3$. Hence, we express the LSF as $G(U_1(\omega), U_2(\omega)) = y(1/3, \omega) - y_{\max}$. Applying linear FEs with mesh size parameter $h > 0$, we compute the approximate solution to (5.1) and (5.2) which we denote by $y_h : \overline{D} \times \Omega \to \mathbb{R}$. Accordingly, the approximate LSF is given by $G_h(U_1(\omega), U_2(\omega)) = y_h(1/3, \omega) - y_{\max}$.
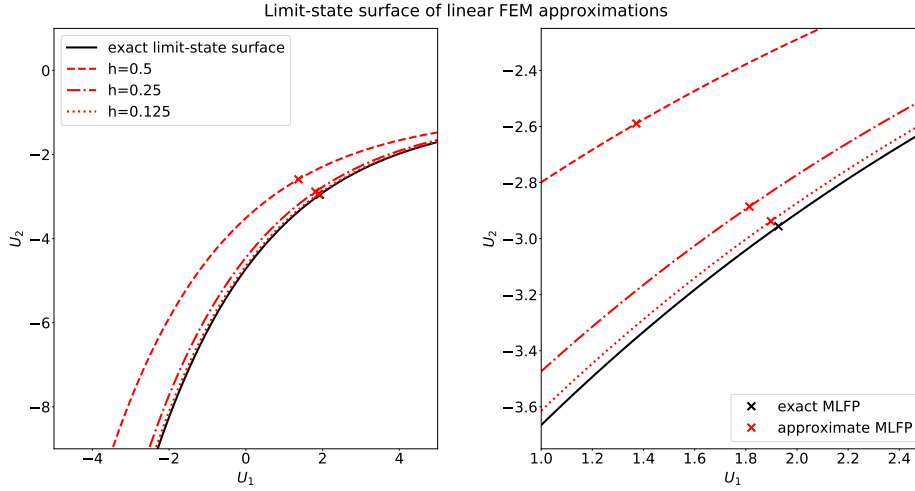

FIG. 6. *On the left: the black line shows the exact limit-state surface and the black cross denotes the exact MLFP. The red lines show the approximate limit-state surfaces given by the linear FE discretization for different choices of the mesh size h. The red crosses denote the approximate MLFPs. On the right: zoom-in of the left plot near the MLFPs.*

Figure 6 shows the exact limit-state surface and the limit-state surfaces given by the linear FE approximations. We observe that the exact as well as the approximate failure domains are convex. Indeed, the exact limit-state surface can be expressed as a function in terms of the first coordinate u_1 by

$$u_2(u_1) = 1/\widehat{x}\left(-1/3 - \left(\widehat{x}^3/6 - \widehat{x}^2/2 + \widehat{x}/3\right)\exp(-u_1/3 + 3)\right).$$

Thus, it holds $\partial A = \{(u_1, u_2(u_1)) : u_1 \in \mathbb{R}\}$. Since $-u_2''(u_1) > 0$, we conclude that $-u_2(u_1)$ is a convex function. Since the failure domain A has the same geometric properties as the *epigraph* of $-u_2(u_1)$, we conclude that A is convex. In a similar way, we can prove that A_h is convex. For more details on convex analysis we refer to [2]. Figure 6 also shows that the distances between the exact and approximate surfaces decrease for decreasing mesh size h. Since the limit-state surface is not a straight line, the FORM estimates of the probability of failure are not equal to the true ones, i.e., $P_f \neq P_f^{\text{FORM}}$ and $P_{f,h} \neq P_{f,h}^{\text{FORM}}$. The quantities P_f and $P_{f,h}$ are calculated by integrating numerically the standard normal PDF within the failure domain. We obtain the values $P_f \approx 1.71 \cdot 10^{-4}$ and $P_f^{\text{FORM}} \approx 2.08 \cdot 10^{-4}$.

Following the theoretical discussions, we expect that

$$|P_f - P_{f,h}| \le \widehat{C} h^s P_{f,h}^{\text{FORM}},$$

for $h > 0$ sufficiently small and $\widehat{C}$ given in the proof of Theorem 2.12. The order of convergence s is equal to the order of convergence of the FE discretization. The point evaluation of a linear FE approximation introduces an error of order two, since by [7, Theorem 1.1] it holds that the L^∞-error is bounded by

$$\|y(\cdot, \omega) - y_h(\cdot, \omega)\|_{L^\infty} \le C \|y(\cdot, \omega)\|_{W^{2,\infty}} h^2,$$

for a fixed $\omega \in \Omega$. Hence, we expect that the error bounds in Proposition 2.11 and Theorem 2.12 hold for $s = 2$.

As another discretization, we apply FEs with quadratic basis functions. In this case, the L^∞-error of the exact and FE solution converges with order $\mathcal{O}(h^3)$; see [7, Theorem 1.1]. Hence, we expect that the error bounds in Proposition 2.11 and Theorem 2.12 hold for $s = 3$.

Figure 7 shows the error of the probability of failure by linear and quadratic FEs for the mesh sizes $h_\ell = 1/2^\ell$, for $\ell = 1, \ldots, 9$. We observe that the approximate probability of failure $P_{f,h}$ converges to the exact probability of failure P_f for both discretizations. Similarly, $P_{f,h}^{\text{FORM}}$ converges to P_f^{FORM}. The true relative error $|P_f - P_{f,h}|/P_f$ as well as the relative error of the FORM estimates $|P_f^{\text{FORM}} - P_{f,h}^{\text{FORM}}|/P_f^{\text{FORM}}$ behaves as the discretization error of the FEM approximations. Moreover, the error bound in Theorem 2.12 behaves as the discretization error.
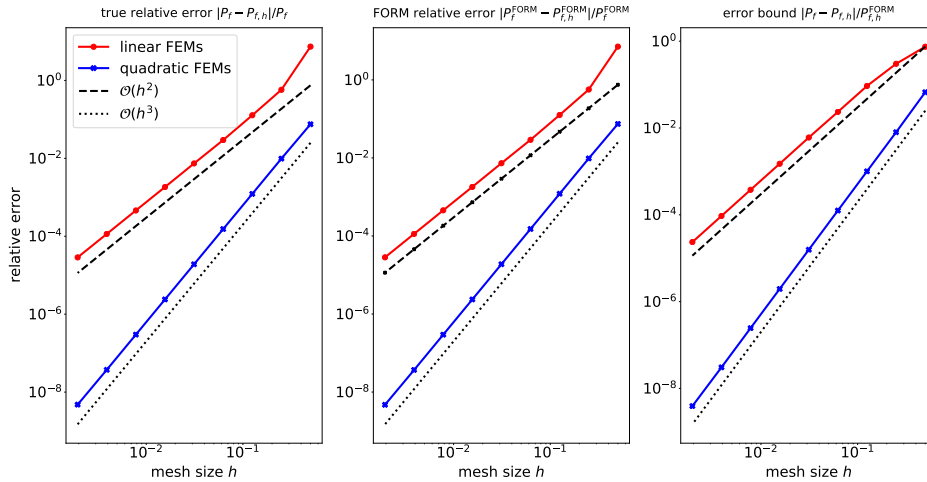
FIG. 7. *On the left: relative error of the approximate probability of failure $P_{f,h}$ with respect to the exact probability of failure P_f. In the middle: relative error of the approximate FORM estimate $P_{f,h}^{\text{FORM}}$ with respect to the exact FORM estimate P_f^{FORM}. On the right: bound of the error $|P_f - P_{f,h}|$. The black lines show the order of convergence.*

5.3. High-dimensional parameter space. In the following, we consider Example 2.8 of Section 2 for which it is not possible to calculate analytic expressions of the exact and approximate PDE solutions y and y_h, respectively. Moreover, the

limit-state surfaces A and A_h cannot be expressed explicitly. Therefore, we estimate P_f and $P_{f,h}$ with Sequential Importance Sampling (SIS) [24].

On the domain $D = (0,1)$, we seek a solution $y : \overline{D} \times \Omega \to \mathbb{R}$ which solves

(5.3)
$$-\frac{\partial}{\partial x}\left(a(x,\omega)\frac{\partial}{\partial x}y(x,\omega)\right) = 0, \quad \text{for } x \in (0,1),$$

$$\text{such that } y(0,\omega) = 1 \text{ and } y(1,\omega) = 0,$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$. The random field $a(x,\omega) = \exp(Z(x,\omega))$ is a log-normal random field and the underlying Gaussian field $Z(x,\omega)$ has constant mean $\mu_Z = 0.1$ and variance $\sigma_Z^2 = 0.04$. The covariance function of Z is $c(x_1, x_2) = \sigma_z^2 \exp\left(-\|x_1 - x_2\|_1/\lambda\right)$, with correlation length $\lambda = 0.3$. The random field Z is approximated via its truncated KLE with $n = 10$ leading terms, which captures around 93% of the variability of the random field. Failure is defined as the event that the flow rate $q(\cdot,\omega)$, given in (2.14), is larger than $q_{\max} = 1.7$ at $\widehat{x} = 1$. Hence, we express the LSF as $G(U(\omega)) = q_{\max} - q(1,\omega)$.

Linear FEs are applied with mesh size parameter $h > 0$ to obtain the approximate solution $y_h : D \times \Omega \to \mathbb{R}$ of (5.3). Accordingly, the approximate LSF is given by $G_h(U(\omega)) = q_{\max} - q_h(1,\omega)$. As discussed in Example 2.8, linear FEs yield a PDE discretization error of order one. Since the approximation error of the LSF is not uniformly bounded, Proposition 2.11 and Theorem 2.12 are not directly applicable. However, as noted in Remark 2.13, we expect that our error bounds also hold for $s = 1$.

The references P_f and $P_{f,h}$ are obtained by averaging over 100 SIS simulations with 10^4 samples, target coefficient of variation equal to 0.25 and using *Markov Chain Monte Carlo* (MCMC) with sampling from the von Mises–Fisher–Nakagami distribution. No burn-in is applied within the MCMC sampling and 10% of the samples are chosen as seeds of the simulated Markov chains via multinomial resampling. Details are given in [33]. We note that the coefficient of variation of the 100 probability of failure estimates is 10^{-2}. Hence, we expect that the sampling bias is negligible. The reference probability of failure is estimated as $P_f \approx 3.38 \cdot 10^{-4}$ on a mesh with discretization size $h = 2^{-12}$. Similar, the reference FORM estimate $P_f^{\mathrm{FORM}} \approx 4.66 \cdot 10^{-4}$ is obtained by FORM with mesh size $h = 2^{-12}$. The reference $P_{f,h}$ and $P_{f,h}^{\mathrm{FORM}}$ are obtained on a sequence of mesh sizes $h_\ell = 1/2^\ell$ for $\ell = 1, \ldots, 11$.

The upper left plot of Figure 8 shows the reference probability of failure P_f, approximations $P_{f,h}$ and FORM estimates $P_{f,h}^{\mathrm{FORM}}$. We observe that $P_{f,h}^{\mathrm{FORM}}$ is always larger than $P_{f,h}$ for a fixed mesh size h. This is a necessary condition for convex failure domains. However, we cannot show that the failure domains are indeed convex and unbounded. The upper right plot shows that the relative error $|P_f - P_{f,h}|/P_f$ behaves as the discretization error of the LSF. The same holds true for the relative error with respect to the FORM estimates, which is illustrated in the lower left plot. We expected this behaviour by Proposition 2.11. Moreover, the lower right plot shows the convergence of our error bound in Theorem 2.12. We also observe that the error bound gives an order one approximation which we have expected by Theorem 2.12.

5.3.1. 50-dimensional parameter space. We consider the problem setting of (5.3) with correlation length $\lambda = 0.1$. A smaller correlation length requires a larger number of leading KLE terms to acquire a similar resolution of the random field. Therefore, we consider $n = 50$ leading KLE terms, which captures around 96% of the variability of the random field. We adjust the threshold $q_{\max} = 1.5$ to achieve a

Fig. 8. *Upper left: reference probability of failure P_f, approximations $P_{f,h}$ and FORM estimates $P_{f,h}^{\mathrm{FORM}}$. Upper r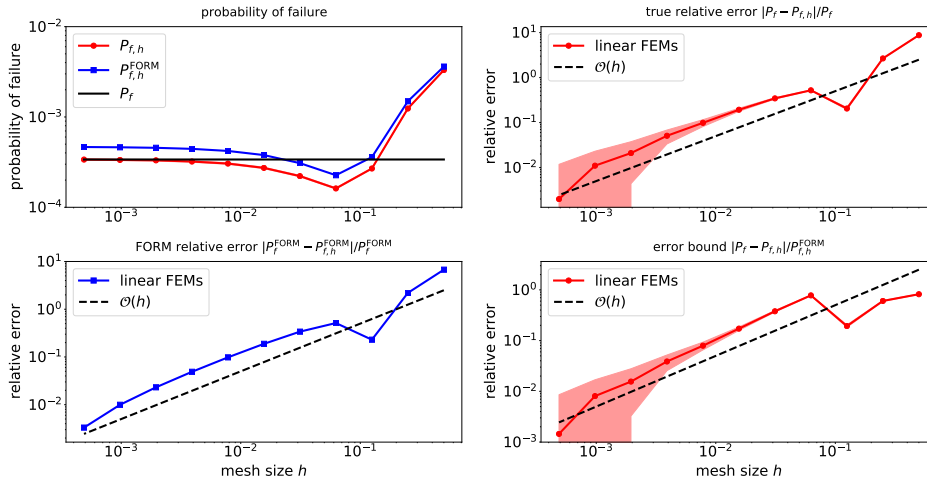ight: relative error of the approximations $P_{f,h}$ with respect to the reference probability P_f. Lower left: relative error of the approximate FORM estimate $P_{f,h}^{\mathrm{FORM}}$ with respect to the reference P_f^{FORM}. Lower right: derived bound of the error $|P_f - P_{f,h}|$. The dashed black lines show the order of convergence. The red areas show the standard deviations of the estimates.*

similar order of the probability of failure. As in the previous example, the references for the probability of failure are obtained by SIS and the settings as described above. The reference probability of failure is estimated as $P_f = 7.18 \cdot 10^{-5}$ on a mesh with discretization size $h = 2^{-12}$. The reference FORM estimate $P_f^{\mathrm{FORM}} = 1.52 \cdot 10^{-4}$ is obtained on the same discretization level. The reference $P_{f,h}$ and $P_{f,h}^{\mathrm{FORM}}$ are obtained on a sequence of mesh sizes $h_\ell = 1/2^\ell$ for $\ell = 1, \ldots, 11$.

The upper left plot of Figure 9 shows that $P_{f,h}^{\mathrm{FORM}}$ is always larger than $P_{f,h}$ for a fixed mesh size h. The upper right plot shows that the relative error $|P_f - P_{f,h}|/P_f$ has order $\mathcal{O}(h)$ of convergence for small discretization sizes h. For large h, we observe a plateau behaviour and then a fast decay until it converges with the expected order. The relative error with respect to the FORM estimates, which is illustrated in the lower left plot, has order $\mathcal{O}(h)$ of convergence and, hence, is the same as the convergence property of the LSF. Moreover, the lower right plot shows the convergence of our error bound. We also observe that the error bound gives an order one approximation for small h. This is exactly the order of convergence we expect from Proposition 2.11 and Theorem 2.12.

6. Conclusion and Outlook. In this manuscript, we have considered the approximation error of the probability of failure, which is induced through the approximation error of the LSF. We assume that the LSF depends on the evaluation of an elliptic PDE with stochastic diffusion parameter and Dirichlet boundary condition. We have shown in Theorem 2.12 under certain assumptions, that the approximation error of the probability of failure behaves as the PDE discretization error multiplied by the FORM estimate of the probability of failure. Moreover, we have shown in Proposition 2.11 that the relative error of the FORM estimates behaves as the PDE discretization error. If the LSF is affine linear with respect to the stochastic param-
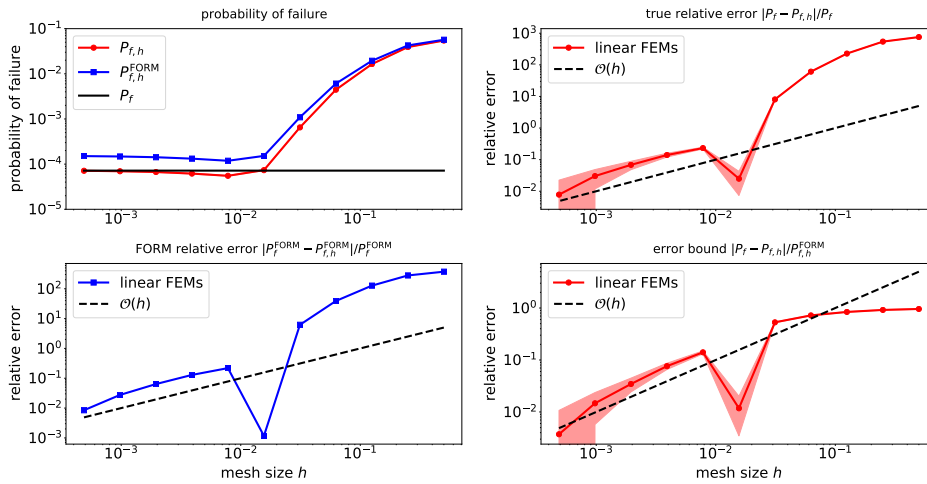
FIG. 9. *Upper left: reference probability of failure P_f, approximations $P_{f,h}$ and FORM estimates $P_{f,h}^{\text{FORM}}$. Upper right: relative error of the approximations $P_{f,h}$ with respect to the reference probability P_f. Lower left: relative error of the approximate FORM estimate $P_{f,h}^{\text{FORM}}$ with respect to the reference P_f^{FORM}. Lower right: derived bound of the error $|P_f - P_{f,h}|$. The dashed black lines show the order of convergence. The red areas show the standard deviations of the estimates.*

eter, the derived error bound gives an upper bound for the relative approximation error of the probability of failure.

Our provided error bounds are only applicable for uniformly elliptic and bounded diffusion coefficients. We outline an idea to treat pathwise elliptic and bounded diffusion coefficients. However, we have not provided a complete proof. In several numerical experiments, we observe that our provided error bounds also hold true for pathwise elliptic and bounded diffusion coefficients. In these experiments, we have shown that the approximation error of the probability of failure indeed behaves as the derived error bound given in Theorem 2.12. The same holds true for the bound of the relative error of the FORM estimates given in Proposition 2.11.

The manuscript can be used as a starting point to derive an error bound, which is applicable for a broader range of LSFs. The derivation of an error bound for the relative error, which does not consist of the FORM estimate, is still of high interest.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Department of Commerce, National Bureau of Standards, 1964.

[2] J. M. BORWEIN AND A. S. LEWIS, *Convex Analysis and Nonlinear Optimization, Theory and Examples*, Springer, New York, NY, 2 ed., 2006, https://doi.org/10.1007/978-0-387-31256-9.

[3] D. BRAESS, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 3 ed., 2007, https://doi.org/10.1017/CBO9780511618635.

[4] J. CHARRIER, R. SCHEICHL, AND A. L. TECKENTRUP, *Finite element error analysis of elliptic*

PDEs with random coefficients and its application to multilevel Monte Carlo methods, SIAM Journal on Numerical Analysis, 51 (2013), pp. 322–352, https://doi.org/10.1137/110853054.

[5] F. J. CORNATON, Y. PARK, S. D. NORMANI, E. A. SUDICKY, AND J. F. SYKES, *Use of groundwater lifetime expectancy for the performance assessment of a deep geologic waste repository*, Water Resources Research, 44 (2008), https://doi.org/10.1029/2007WR006208.

[6] A. DER KIUREGHIAN AND P.-L. LIU, *Structural reliability under incomplete probability information*, Journal of Engineering Mechanics, 112 (1986), pp. 85–104, https://doi.org/10.1061/(ASCE)0733-9399(1986)112:1(85).

[7] J. DOUGLAS, T. DUPONT, AND L. WAHLBIN, *Optimal L_∞ error estimates for Galerkin approximations to solutions of two-point boundary value problems*, Mathematics of Computation, 29 (1975), pp. 475–483, https://doi.org/10.1090/S0025-5718-1975-0371077-0.

[8] D. ELFVERSON, F. HELLMAN, AND A. MÅLQVIST, *A multilevel Monte Carlo method for computing failure probabilities*, SIAM/ASA Journal on Uncertainty Quantification, 4 (2016), pp. 312–330, https://doi.org/10.1137/140984294.

[9] O. G. ERNST, B. SPRUNGK, AND H.-J. STARKLOFF, *Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 823–851, https://doi.org/10.1137/140981319.

[10] L. C. EVANS, *Partial Differential Equations*, American Mathematical Society, 2 ed., 2010.

[11] R. EYMARD, T. GALLOUËT, AND R. HERBIN, *Finite volume methods*, in Solution of Equation in $\mathbb{R}^n$ (Part 3), Techniques of Scientific Computing (Part 3), vol. 7 of Handbook of Numerical Analysis, Elsevier, 2000, pp. 713–1018, https://doi.org/10.1016/S1570-8659(00)07005-8.

[12] G. A. FENTON AND D. V. GRIFFITHS, *Bearing-capacity prediction of spatially random $c-\phi$ soils*, Canadian Geotechnical Journal, 40 (2003), pp. 54–65, https://doi.org/10.1139/t02-086.

[13] A. GARBUNO-INIGO, F. HOFFMANN, W. LI, AND A. M. STUART, *Interacting langevin diffusions: Gradient structure and ensemble Kalman sampler*, SIAM Journal on Applied Dynamical Systems, 19 (2020), pp. 412–441, https://doi.org/10.1137/19M1251655.

[14] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer–Verlag, New York, 1991, https://doi.org/10.1007/978-1-4612-3094-6.

[15] R. D. GORDON, *Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument*, Ann. Math. Statist., 12 (1941), pp. 364–366, https://doi.org/10.1214/aoms/1177731721.

[16] W. HACKBUSCH, *Elliptic Differential Equations*, Springer Series in Computational Mathematics, Springer, Berlin, Heidelberg, 2 ed., 2017, https://doi.org/10.1007/978-3-662-54961-2.

[17] A. HASOFER AND N. LIND, *An exact and invariant first order reliability format*, Journal of Engineering Mechanics, 100 (1974), pp. 111–121.

[18] M. HOHENBICHLER AND R. RACKWITZ, *Non-normal dependent vectors in structural safety*, Journal of the Engineering Mechanics Division, 107 (1981), pp. 1227–1238.

[19] A. KIUREGHIAN, *First- and second-order reliability methods*, in Engineering Design Reliability Handbook, E. Nikolaidis, D. M. Ghiocel, and S. Singhal, eds., CRC Press, 2004, ch. 14.

[20] R. J. LEVEQUE, *Finite Difference Methods for Ordinary and Partial Differential Equations*, Society for Industrial and Applied Mathematics, 2007, https://doi.org/10.1137/1.9780898717839.

[21] G. J. LORD, C. E. POWELL, AND T. SHARDLOW, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2014, https://doi.org/10.1017/CBO9781139017329.

[22] U. NOSECK, D. BECKER, C. FAHRENHOLZ, E. FEIN, J. FLÜGGE, K.-P. KRÖHN, J. MÖNIG, I. MÜLLER-LYDA, T. ROTHFUCHS, A. RÜBEL, AND J. WOLF, *Assessment of the long-term safety of repositories*, Gesellschaft für Anlage und Reaktorsicherheit (GRS) mbH, 2008, https://www.grs.de/sites/default/files/pdf/GRS-237.pdf.

[23] F. W. OLVER, D. W. LOZIER, R. F. BOISVERT, AND C. W. CLARK, *NIST Handbook of Mathematical Functions*, Cambridge University Press, USA, 1 ed., 2010.

[24] I. PAPAIOANNOU, C. PAPADIMITRIOU, AND D. STRAUB, *Sequential importance sampling for structural reliability analysis*, Structural Safety, 62 (2016), pp. 66–75, https://doi.org/10.1016/j.strusafe.2016.06.002.

[25] B. PEHERSTORFER, B. KRAMER, AND K. WILLCOX, *Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation*, SIAM/ASA Journal on Uncertainty Quantification, 6 (2018), pp. 737–761, https://doi.org/10.1137/17M1122992.

[26] R. RANNACHER AND B. VEXLER, *A priori error estimates for the finite element discretization of elliptic parameter identification problems with pointwise measurements*, SIAM Journal on Control and Optimization, 44 (2005), pp. 1844–1863, https://doi.org/10.1137/040611100.

[27] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning, MIT Press, 1 ed., 2006.

[28] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Wellesley-Cambridge Press, 2 ed., 1997.

[29] D. Straub, I. Papaioannou, and W. Betz, *Bayesian analysis of rare events*, Journal of Computational Physics, 314 (2016), pp. 538–556, https://doi.org/10.1016/j.jcp.2016.03.018.

[30] A. L. Teckentrup, R. Scheichl, M. B. Giles, and E. Ullmann, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numerische Mathematik, 125 (2013), pp. 569–600, https://doi.org/10.1007/s00211-013-0546-4.

[31] E. Ullmann and I. Papaioannou, *Multilevel estimation of rare events*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 922–953, https://doi.org/10.1137/140992953.

[32] B. Vexler, *Adaptive Finite Element Methods for Parameter Identification Problems*, PhD thesis, University of Heidelberg, 2004, https://doi.org/10.11588/heidok.00004603.

[33] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann, *Multilevel sequential importance sampling for rare event estimation*, SIAM Journal on Scientific Computing, 42 (2020), pp. A2062–A2087, https://doi.org/10.1137/19M1289601.

C. Article: The ensemble Kalman filter for rare event estimation

Summary

In this article, we study the ensemble Kalman filter (EnKF) as a novel sampling-based method to estimate the probability of rare events. The starting point of the EnKF algorithm is the work of [57, 106], where the EnKF is applied to inverse problems. To apply the EnKF to rare event estimation, we reformulate the rare event problem as an inverse problem and apply the EnKF to generate failure samples. With the generated samples, we fit a distribution model and estimate the probability of the rare event with importance sampling. We show that the resulting EnKF densities yield a novel way to approximate the optimal importance sampling density. We suggest a novel adaptive algorithm to apply the EnKF to multi-modal failure domains. Therefore, we combine the approach of [102] with a clustering approach to determine adaptively the localization parameter.

Another contribution of this article is the analysis of the continuous-time limit of the EnKF update tailored to rare event estimation. Under the assumptions that the EnKF is applied without noise and the limit-state function is affine linear, we derive the continuous-time limit of the EnKF update. We show that the mean of the EnKF particles converges to a convex combination of the most likely failure point and the mean of the optimal importance sampling density. Moreover, we show that the safe particles move to the limit-state surface and remain there. Thus, the covariance of the safe particles in direction perpendicular to the limit-state surface shrinks to zero, which is also known as ensemble collapse. Contrary, the covariance of the safe particles in direction parallel to the limit-state surface remains constant.

In numerical experiments, we compare the performance of the EnKF with Sequential Importance Sampling (SIS). We observe that for unimodal failure domains, the EnKF requires less computational costs than SIS for a fixed level of accuracy. For multi-modal failure domains, the performance of the EnKF is similar with SIS. However, the application of SIS to multi-modal failure domains is more straightforward, while the application of the EnKF should be considered carefully.

Statement of individual contribution

I came up with the idea of implementing the EnKF for rare event estimation. In my master's thesis, I have studied the EnKF for deep learning and I have observed similarities between rare event estimation and Bayesian inverse problems.
I was fully responsible for the implementation and execution of numerical experiments. Also, I was in charge to work out the technical steps in the proofs of our main statements. I was supported by insightful discussions with Iason Papaionnou and Elisabeth Ullmann concerning technical details. Moreover, I was fully responsible for writing and preparing this manuscript. Iason Papaioannou, Elisabeth Ullmann, and I proofread and polished the article together.

Permission to include

Fabian Wagner, Iason Papaioannou, Elisabeth Ullmann.
The ensemble Kalman filter for rare event estimation.
Submitted for publication in SIAM/ASA Journal on Uncertainty Quantification (2021).
arXiv preprint: `https://arxiv.org/abs/2106.10062`
(See also article [120] in the bibliography)

On the following page, a copy of the E-Mail by SIAM may be found with the confirmation that the manuscript "The ensemble Kalman filter for rare event estimation" may be included in my dissertation. Moreover, a digital version of the consent to publish may be found at

`https://www.siam.org/publications/journals/about-siam-journals/`
`information-for-authors`

(Accessed on 21 June 2021)

Von:	Kelly Thomas <Thomas@siam.org>
Gesendet:	Dienstag, 27. April 2021 22:41
An:	fabian.wagner@ma.tum.de
Betreff:	RE: Contact Request From SIAM

Dear Mr. Wagner:

SIAM is happy to give permission to reuse material from your article mentioned below in your thesis.

Sincerely,

Kelly Thomas
Managing Editor
Society for Industrial and Applied Mathematics
3600 Market Street - 6th Floor
Philadelphia, PA 19104
thomas@siam.org / (267) 350-6387

-----Original Message-----
From: James Haines
Sent: Tuesday, April 27, 2021 4:30 PM
To: Kelly Thomas <Thomas@siam.org>
Subject: FW: Contact Request From SIAM

-----Original Message-----
From: No Reply <noreply@siam.org>
Sent: Tuesday, April 27, 2021 3:30 PM
To: juq@siam.org
Subject: Contact Request From SIAM

Dear juq@siam.org,

This message was sent to you from the SIAM contact request system on behalf of Fabian Wagner.

Below is the information provided by Fabian Wagner:

Email address: fabian.wagner@ma.tum.de
Phone number:
This request was made from the following page on the SIAM website:
 https://www.siam.org/publications/journals/siam-asa-journal-on-uncertainty-quantification-juq

The following message was provided by Fabian Wagner:

Dear Sirs and Madams,

My name is Fabian Wagner, I am one of the authors of the article "The ensemble Kalman filter for rare event estimation" (M140411) that is submitted to the SIAM/ASA Journal on Uncertainty Quantification.
Since I am pursuing a cumulative doctoral thesis, it is necessary by the rules of my university (TU Munich, Germany) that I provide a "letter of approval" by the publisher to all my articles that I want to include in my thesis. I hereby kindly ask for such a confirmation (by Email or weblink to your terms and condition page stating such a regulation), allowing me to use the article mentioned above in my doctoral thesis.

1

If you have any questions beforehand, do not hesitate to ask.

Best wishes,
Fabian Wagner

_____
Thank you.

_-*
If you would like to respond to the requester, please contact them using the information they provided above.
Please do not reply to this message as it is sent from an unmonitored email address.
This contact request was generated from SIAM'S website using SIAM's Contact Request System (CRS). This system seeks to facilitate communication while mitigating email abuse and protecting the privacy of your contact information.
If you believe this message was sent to you in error or you have any other concerns please forward this entire email to SIAM's Customer Service Department at service@siam.org along with your concerns.

Request ID: 2499

The first two books in the SIAM Data Science Book Series have now published! Learn more<https://my.siam.org/Store/Home/BookSeries/22>.

Notice of publication and copyright

The manuscript "The ensemble Klaman filter for rare event estimation" is submitted for publication in SIAM/ASA Journal on Uncertainty Quantification (2021), published by the Society for Industrial and Applied Mathematics (SIAM).

On the following page, a copy of the E-Mail by SIAM may be found with the confirmation that the manuscript "The ensemble Kalman filter for rare event estimation" is submitted to SIAM/ASA Journal on Uncertainty Quantification.

Thereafter, the arXiv preprint version of the manuscript "The ensemble Kalman filter for rare event estimation" is shown. The arXiv preprint can be accessed at `https://arxiv.org/abs/2106.10062`.

Fabian Wagner

Dear Mr. Wagner,

Thank you for submitting "The ensemble Kalman filter for rare event estimation."

Your paper has been assigned MS#M140411.

You can check on the status of the paper by logging in at https://juq.siam.org and following the live manuscripts link.

You will receive notice after the review process has been completed. You can send inquiries to juq@siam.org at any time. It's helpful if you refer to your paper by title, by the manuscript number, and by your name.

The authors listed for your submission are Fabian Wagner, Iason Papaioannou, and Elisabeth Ullmann, and as Corresponding Author you represent that any and all co-authors are aware of and agree with this submission. The Editors-in-Chief have been notified of your submission and all further correspondence will be sent only to you as the Corresponding Author.

We are including additional information below that we encourage you keep on hand as reference. Please do not hesitate to write if you have any questions.

Regards,

Heather Blythe
Senior Publications Coordinator
SIAM
Email: blythe@siam.org
Phone: 215-382-9800, ext. 352
Fax: 215-386-7999

Information for the author:

In order to facilitate electronic searches of our articles and to help readers who may be scanning our journals or Tables of Contents for articles of interest, we ask that you give special consideration to the following items, should you be asked to revise your paper, or should it be accepted for publication.

Title: The title should accurately and specifically reflect the subject matter addressed in the paper. Do not use mathematical symbols or equations in the title unless absolutely necessary.

Abstract: Abstracts are often reproduced, separate from the body of the article, in review journals and in electronic search facilities. Many readers will decide whether to read your article based on the information contained in the abstract. Thus, the abstract should be brief (250 words or less), but should accurately summarize the contents of the paper. Specifically: indicate the conclusions reached in the paper and the possible applications of the research performed; avoid using mathematical equations; do not cite references by number.

The ensemble Kalman filter for rare event estimation

F. Wagner, I. Papaioannou, E. Ullmann

21st June 2021

Abstract

We present a novel sampling-based method for estimating probabilities of rare or failure events. Our approach is founded on the Ensemble Kalman filter (EnKF) for inverse problems. Therefore, we reformulate the rare event problem as an inverse problem and apply the EnKF to generate failure samples. To estimate the probability of failure, we use the final EnKF samples to fit a distribution model and apply Importance Sampling with respect to the fitted distribution. This leads to an unbiased estimator if the density of the fitted distribution admits positive values within the whole failure domain. To handle multi-modal failure domains, we localise the covariance matrices in the EnKF update step around each particle and fit a mixture distribution model in the Importance Sampling step. For affine linear limit-state functions, we investigate the continuous-time limit and large time properties of the EnKF update. We prove that the mean of the particles converges to a convex combination of the most likely failure point and the mean of the optimal Importance Sampling density if the EnKF is applied without noise. We provide numerical experiments to compare the performance of the EnKF with Sequential Importance Sampling.

Keywords: Reliability analysis, importance sampling, ensemble Kalman filter, inverse problems

1 Introduction

Estimating the probability of failure is a frequent and crucial task in reliability analysis and risk management [2, 36]. Failure of a system is determined by the outcome of a *limit-state function* (LSF). By convention, if the outcome is larger than zero, the given state keeps the system in a safe mode. For a negative outcome, the state leads to failure. The *probability of failure* is defined as the probability mass of all failure states. Since failure probabilities are likely to be small, estimation of the failure probability requires the simulation of *rare events*.

Often, the evaluation of the LSF requires the evaluation of a computational expensive model, a partial differential equation, which makes crude Monte Carlo sampling [16, 45] prohibitive. Variance reduction techniques like *Subset Simulation* (SuS) [3, 4], *Sequential Importance Sampling* (SIS) [42, 54] or the *cross-entropy* based Importance Sampling (IS) method [29, 41, 56] have been developed to reduce computational costs while preserving an accurate estimate. In *line sampling* [8, 28, 43], sampling is performed on a hyperplane perpendicular to an important direction. Alternative to sampling methods, approximation methods, like the *first* and *second order reliability method* [35] (FORM/SORM), determine the *most likely failure point* (MLFP) and approximate the surface of the failure

domain. In our novel approach, we investigate the *Ensemble Kalman filter* (EnKF) for inverse problems proposed by [24, 46] and apply it to rare event estimation. The EnKF is a sampling-based method.

The Kalman filter [26] has been originally proposed for data assimilation problems. If the observation operator and dynamic are linear, and the initial and noise distribution are Gaussian, the Kalman filter is exact. The Extended Kalman filter [32, Section 4.2.2] is applied for nonlinear dynamics. In this case, the dynamic is approximated via its linearization and the Kalman update is applied. However, this requires derivative information, which might be costly to obtain. The EnKF [15] approximates the derivative via an ensemble of particles. Thus, the EnKF does not require derivative information. However, the EnKF is only asymptotically exact if the dynamic is linear. To apply the EnKF to systems with multimodal distributions, the authors of [13, 33, 48] propose to fit a GM distribution in each EnKF update step and to update the particles belonging to each mixture term separately.

Compared to SIS and SuS, the EnKF has several advantages which are our motivation to implement the EnKF for rare event estimation. The EnKF is easier to implement since a *Markov chain Monte Carlo* (MCMC) [20] algorithm is not required. Thus, the EnKF contains fewer hyperparameters which have to be tuned. Moreover, it requires no burn-in and no tuning for the optimal acceptance rate has to be performed. Since the EnKF particles are equally weighted, no computational costs are wasted for degenerated samples with very small weights or for rejected samples. Moreover, the EnKF is applicable for high-dimensional problems.

In recent years, several theoretical properties of the EnKF have been studied. In [24], it is shown that the EnKF satisfies the *subspace property*, i.e., the particles stay for all iterations within the subspace spanned by the initial ensemble. The authors of [6, 46, 47] investigate the continuous time limit of the EnKF update step, which results in a coupled system of *stochastic differential equations* (SDEs). For a fixed ensemble size, linear system, and considering the limit $t \to \infty$, the ensemble collapses to its mean value. The authors of [18, 21] study the mean and covariance of the EnKF particles for an infinite ensemble in the linear Gaussian setting. The distribution of the EnKF particles is equal to the posterior distribution for $t = 1$ in the continuous time limit [18]. The limit $t \to \infty$ yields again ensemble collapse. In the nonlinear case, the work of [14] shows that the distribution of the ensemble particles does not converge to the posterior distribution even for an infinite ensemble. Instead, the particles approximate the distribution of a so-called *analysis variable.*

To apply the EnKF algorithm for rare event simulation, we formulate the rare event problem as an inverse problem via an auxiliary LSF. The auxiliary LSF is the concatenation of the rectified linear unit (ReLU) and the original LSF. We apply the EnKF for inverse problems to this reformulation. Since the distribution of the analysis variable differs from the posterior distribution in general, we fit a distribution model with the final EnKF particles and apply IS with respect to the fitted distribution to estimate the probability of failure. This procedure is similar to [29, Algorithm 3.1]. In particular, we apply the *Gaussian mixture* (GM) and the *von Mises–Fisher–Nakagami mixture* (vMFNM) distribution model, which have also been used in the context of the cross-entropy method [41]. Due to the properties of IS, we achieve an unbiased estimator for the probability of failure if the density of the fitted distribution admits positive values within the whole failure domain. Additionally, we apply an adaptive approach to determine the sequence of discretization steps. This approach is similar to adaptive *Sequential Monte Carlo* [5] and is already applied to the EnKF in [23]. We demonstrate that the sequence of the resulting EnKF

densities is a piecewise smooth approximation of the discontinuous optimal IS density. Since the EnKF particles always approximate the distribution of one single mode, we apply a multi-modal strategy to handle multi-modal failure domains. In our work, we consider the approach of [44], where the means and covariance matrices of the EnKF update are localised around each particle. Therefore, the particles evolve individually (within their nearest neighbourhood) to distinct failure modes. The approach of [44] requires the choice of a localization parameter, which we avoid through employing a clustering algorithm. Moreover, we translate some of the already observed theoretical properties for the EnKF to the rare event setting. We derive the continuous time limit of the particle dynamic for affine linear LSFs. If the EnKF is applied without noise in the data space, we prove that the ensemble mean converges to a convex combination of the MLFP and of the mean of the optimal IS density.

The manuscript is structured as follows. Section 2 reviews the general setting of rare event estimation and introduces an equivalent formulation as an inverse problem. Thereafter, *Bayesian inverse problems* (BIPs) and the EnKF for inverse problems are discussed. Section 3 contains the formulation of the EnKF for estimating the probability of failure. Section 4 shows theoretical properties of the EnKF for affine linear LSFs. The proofs are given in the Appendix A. In Section 5, the EnKF is applied to numerical experiments and its performance is compared with SIS. We end this manuscript with a conclusion in Section 6.

2 Problem Setting

We start by defining failure events and the probability of failure. We discuss the standard formulation and introduce an alternative formulation that draws an analogy to BIPs. We introduce BIPs and the well-known Bayes' theorem. Moreover, we discuss the EnKF for inverse problems and its theoretical properties.

2.1 Rare event estimation

The following notation is based on [40, 42]. It is common to define failure events via an LSF that distinguishes safe and failure states. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Given is an LSF $G : \mathbb{R}^d \to \mathbb{R}$, which models the performance of a system. Importantly, the LSF often depends on a computationally intensive numerical model of the system. The state $u \in \mathbb{R}^d$ leads to failure if $G(u) \leq 0$; otherwise, u is a safe state.

The input state $u \in \mathbb{R}^d$ is a realization of an $\mathbb{R}^d$-valued random variable $U : \Omega \to \mathbb{R}^d$. We will sometimes use the equivalent definition of a failure event by $\{\omega \in \Omega : G(U(\omega)) \leq 0\}$ to emphasize that the outcome of the LSF depends on an event $\omega \in \Omega$. The random variable U is distributed according to the *probability density function* (pdf) $\mu_0 : \mathbb{R}^d \to [0, \infty[$.

The goal is to estimate the probability mass of the states $u \in \mathbb{R}^d$, or equivalently, of events $\omega \in \Omega$, which lead to failure. This probability is called *probability of failure* and is denoted by P_f. Indeed, the probability of failure is given by

$$P_f := \mathbb{P}\left(\{\omega \in \Omega : G(U(\omega)) \leq 0\}\right) = \int_{u \in \mathbb{R}^d} I(G(u) \leq 0)\mu_0(u)\mathrm{d}u, \qquad (2.1)$$

where I denotes the indicator function. Since the *failure domain* $\{u \in \mathbb{R}^d : G(u) \leq 0\}$ is a priori unknown, estimating P_f accurately is a nontrivial task.

In IS [1, 45], the integral in (2.1) is expressed as an expectation with respect to an IS

density $p : \mathbb{R}^d \to \mathbb{R}$

$$P_f = \int_{u \in \mathbb{R}^d} I(G(u) \leq 0) w(u) p(u) \mathrm{d}u = \mathbb{E}_p[I(G(u) \leq 0) w(u)],$$

where $w(u) := \mu_0(u)/p(u)$ is the *importance weight*. Using J samples $\{u^{(j)}\}_{j=1}^J$ which are distributed according to p, the IS estimator for P_f is given by

$$\hat{P}_f^{\mathrm{IS}} := \frac{1}{J} \sum_{j=1}^J I(G(u^{(j)}) \leq 0) w(u^{(j)}). \tag{2.2}$$

If the support of p contains the intersection of the support of μ_0 and the failure domain, (2.2) gives an unbiased estimator for P_f. The optimal IS density is

$$p_{\mathrm{opt}}(u) := \frac{1}{P_f} I(G(u) \leq 0) \mu_0(u), \tag{2.3}$$

which leads to a zero-variance estimator. Sampling-based methods like SuS [3, 4] or SIS [42, 54] aim to generate samples from the optimal IS density. This viewpoint can be interpreted as seeking states $u \in \mathbb{R}^d$ which result in $G(u) \leq 0$. It is possible to define this task as an inverse problem. In inverse problems, the goal is to identify the inputs of a model, whose outcome produces a set of given data $y^\dagger$. To define an equivalent inverse problem, we apply the ReLU function to the outcome of G and define the auxiliary LSF $\widetilde{G} : \mathbb{R}^d \to \mathbb{R}$ as

$$\widetilde{G}(u) := \max\{0, G(u)\}. \tag{2.4}$$

With $\widetilde{G}$ and the data $y^\dagger = 0$, we reformulate the rare event problem as an inverse problem. We seek all $u \in \mathbb{R}^d$ such that

$$\widetilde{G}(u) = 0. \tag{2.5}$$

The overall goal of this manuscript is to apply the EnKF to generate samples from the failure domain and estimate the probability of failure with IS. Since inverse problems of the form (2.5) are ill-posed, we discuss BIPs in the following section.

Remark 2.1. *An alternative Bayesian interpretation of the rare event simulation problem is obtained through observation of equation* (2.3). *The optimal IS density p_{opt} can be interpreted as a posterior density, where the indicator function $I(G(u) \leq 0)$ is the likelihood function, μ_0 is the prior density, and P_f is the evidence. This observation is also discussed in [52]. However, this interpretation does not allow application of the EnKF algorithm, which is developed for standard BIPs with equality-type information. The latter implies that the model outcome is compared to the data with an equality sign. The likelihood $I(G(u) \leq 0)$ provides inequality-type information.*

2.2 Bayesian inverse problems

We consider a forward model $\mathcal{G} : \mathbb{R}^d \to \mathbb{R}^m$, which maps the input to the output space. Given is the data $y^\dagger \in \mathbb{R}^m$. The goal in an inverse problem with noisy observations is to find a state $u^\dagger \in \mathbb{R}^d$ such that

$$y^\dagger = \mathcal{G}(u^\dagger) + \eta, \tag{2.6}$$

where η is the observational noise and is assumed to be distributed as $N(0, \Gamma)$, where Γ is a symmetric positive definite covariance matrix. Since the classical inverse problem (2.6) is ill-posed, (u, y) are modelled as realisations from a jointly varying random variable (U, Y) [7, 49]. With this Bayesian viewpoint and under mild assumptions on the forward model, noise distribution and prior distribution [30, 49], the inverse problem is well-posed and the solution is the posterior density $\mu_{y^\dagger}$. By virtue of Bayes' theorem [7, Theorem 14], the posterior density is given by

$$\mu_{y^\dagger}(u) = \frac{1}{Z(y^\dagger)} \exp\left(-\Psi\left(u; y^\dagger\right)\right) \mu_0(u),$$

where μ_0 is the prior density and Ψ is a potential, defined as $\Psi\left(u; y^\dagger\right) := 1/2\|y^\dagger - \mathcal{G}(u)\|_\Gamma^2$, where $\|\cdot\|_\Gamma := \|\Gamma^{-1/2}\cdot\|_2$. The term $\exp\left(-\Psi\left(u; y^\dagger\right)\right)$ is the likelihood function and returns the density of the data given a parameter state. The normalizing constant is given by

$$Z(y^\dagger) := \int_{\mathbb{R}^d} \exp\left(-\Psi\left(u; y^\dagger\right)\right) \mu_0(u)\mathrm{d}u > 0.$$

For linear models $\mathcal{G}(u) = Au$ with $A \in \mathbb{R}^{m \times d}$ and a Gaussian prior $\mu_0 \sim N(0, \Gamma_0)$, [49, Theorem 2.4] shows that the posterior mean $m \in \mathbb{R}^d$ and posterior covariance $C \in \mathbb{R}^{d \times d}$ are given by

$$m = \left(A^T \Gamma^{-1} A + \Gamma_0^{-1}\right)^{-1} A^T \Gamma^{-1} y, \qquad C = \left(A^T \Gamma^{-1} A + \Gamma_0^{-1}\right)^{-1}. \tag{2.7}$$

Similar to rare event estimation, sampling-based methods like Sequential Monte Carlo [9, 12] have been developed to shift samples from the prior density μ_0 to the posterior density $\mu_{y^\dagger}$. The EnKF for inverse problems [46] is another sampling-based method.

2.3 EnKF for inverse problems

In this section, we use the notation and derivation of [46], which is motivated by Sequential Monte Carlo. In the EnKF for inverse problems, the posterior is reached in a sequential manner. Starting from the prior density μ_0, the sequence of densities is defined by

$$\mu_n(u) \propto \exp\left(-\frac{nh}{2}\|y^\dagger - \mathcal{G}(u)\|_\Gamma^2\right) \mu_0(u), \quad \text{for } n = 1, \dots, N,$$

where $h := 1/N$ and N is the user-defined number of steps to reach the posterior density. It immediately follows that $\mu_N = \mu_{y^\dagger}$ is the posterior density. By the sequential definition, it holds that

$$\mu_{n+1}(u) \propto \exp\left(-\frac{h}{2}\|y^\dagger - \mathcal{G}(u)\|_\Gamma^2\right) \mu_n(u), \quad \text{for } n = 0, \dots, N-1.$$

Remark 2.2. *The n-th step is equivalent to the solution of the inverse problem (2.6) where the noise η is distributed according to $N(0, (nh)^{-1}\Gamma)$. If $nh > 1$, i.e. $n > N$, the noise covariance is down-scaled.*

The sequence of densities μ_n is approximated via an ensemble of equally weighted particles. The initial ensemble $\mathbf{u_0} = \{u_0^{(j)}\}_{j=1}^J$ is distributed according to the prior density μ_0. In one step of the EnKF, an ensemble of $J \in \mathbb{N}$ samples $\mathbf{u_n} = \{u_n^{(j)}\}_{j=1}^J$, which is distributed

(approximately) according to the density μ_n, is transformed into the ensemble $\mathbf{u_{n+1}} = \{u_{n+1}^{(j)}\}_{j=1}^J$, which is distributed (approximately) as μ_{n+1}. Formally, the ensemble $\mathbf{u_n}$ is updated via

$$u_{n+1}^{(j)} = u_n^{(j)} + C_{\text{up}}(\mathbf{u_n}) \left(C_{\text{pp}}(\mathbf{u_n}) + \frac{1}{h}\Gamma\right)^{-1} \left(y_{n+1}^{(j)} - \mathcal{G}(u_n^{(j)})\right), \quad (2.8)$$

for $j = 1, \ldots, J$, where $y_{n+1}^{(j)}$ is the data $y^\dagger$ perturbed by an additive Gaussian noise

$$y_{n+1}^{(j)} = y^\dagger + \xi_{n+1}^{(j)}, \quad (2.9)$$

and $\xi_{n+1}^{(j)} \sim \mathrm{N}(0, h^{-1}\Gamma)$ represents the observational noise scaled by the step size h^{-1}. The matrices $C_{\text{pp}} \in \mathbb{R}^{m \times m}$ and $C_{\text{up}} \in \mathbb{R}^{d \times m}$ are the empirical covariance and cross-covariance matrices and are given by

$$C_{\text{pp}}(\mathbf{u_n}) = \frac{1}{J} \sum_{j=1}^J \left(\mathcal{G}(u_n^{(j)}) - \overline{\mathcal{G}}\right) \otimes \left(\mathcal{G}(u_n^{(j)}) - \overline{\mathcal{G}}\right), \quad (2.10)$$

$$C_{\text{up}}(\mathbf{u_n}) = \frac{1}{J} \sum_{j=1}^J \left(u_n^{(j)} - \overline{u}\right) \otimes \left(\mathcal{G}(u_n^{(j)}) - \overline{\mathcal{G}}\right), \quad (2.11)$$

where $\otimes$ denotes the outer product of two vectors and $\overline{u}, \overline{\mathcal{G}}$ are the empirical means

$$\overline{u} = \frac{1}{J} \sum_{j=1}^J u_n^{(j)}, \qquad \overline{\mathcal{G}} = \frac{1}{J} \sum_{j=1}^J \mathcal{G}(u_n^{(j)}).$$

The authors of [46] show that in the *continuous time limit* $h \to 0$, the update (2.8) leads to the following SDE

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = C_{\text{up}}(\mathbf{u})\Gamma^{-1}\left(y^\dagger - \mathcal{G}(u^{(j)})\right) + C_{\text{up}}(\mathbf{u})\Gamma^{-1/2}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t} \quad (2.12)$$

$$= \frac{1}{J} \sum_{k=1}^J \left\langle \mathcal{G}(u^{(k)}) - \overline{\mathcal{G}}, y^\dagger - \mathcal{G}(u^{(j)}) + \Gamma^{1/2}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t} \right\rangle_\Gamma \left(u^{(k)} - \overline{u}\right),$$

where $W^{(j)}$ is a Brownian motion [37], which represents the observational noise. If we consider the noise free case, i.e. $y_{n+1}^{(j)} = y^\dagger$, and assume that the forward model $\mathcal{G}$ is linear, i.e. $\mathcal{G}(u) = Au$ for $A \in \mathbb{R}^{m \times d}$, it holds that

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C_{uu}(\mathbf{u})D_u\Psi(u^{(j)}; y^\dagger) = -C_{uu}(\mathbf{u})\left(A^T\Gamma^{-1}Au^{(j)} - A^T\Gamma^{-1}y^\dagger\right), \quad (2.13)$$

where $C_{uu}(\mathbf{u})$ is the empirical covariance matrix of the ensemble $\mathbf{u}$. Equation (2.13) implies that the samples move in the direction of the negative gradient of the data misfit. The multiplication with $C_{uu}(\mathbf{u})$ can be interpreted as a pre-conditioner of the particle dynamic [18]. In Section 4, we will show that a similar expression holds true for the setting of rare event estimation if we consider $\mathcal{G} = \widetilde{G}$ as defined in (2.4).

The authors in [18] show for the linear case and the mean field limit $J \to \infty$ that the mean $m(t)$ and covariance $C(t)$ of the EnKF particles satisfy the following differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}m(t) = -C(t)\left(A^T\Gamma^{-1}Am(t) - A^T\Gamma^{-1}y^\dagger\right),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}C(t) = -C(t)A^T\Gamma^{-1}AC(t),$$

which lead to the solutions

$$m(t) = \left(A^T \Gamma^{-1} A t + \Gamma_0^{-1}\right)^{-1} A^T \Gamma^{-1} y^\dagger, \qquad C(t) = \left(A^T \Gamma^{-1} A t + \Gamma_0^{-1}\right)^{-1}.$$

Thus for $t = 1$, it holds that the mean and covariance of the EnKF ensemble is equal to the posterior mean and posterior covariance (2.7). However, for $t \to \infty$, the covariance $C(t)$ converges to zero, which gives ensemble collapse. Therefore, the EnKF for inverse problems gives samples of the posterior for $t = 1$. For $t \to \infty$, the EnKF can be seen as an optimization method instead as a sampling method [18, 46]. If $\mathcal{G}$ is nonlinear, the work of [14] shows that the EnKF gives approximate samples from the analysis variable

$$U^a := U + \mathrm{Cov}(U, \mathcal{G}(U) + \eta)\mathrm{Cov}(\mathcal{G}(U) + \eta)^{-1}(y^\dagger - \mathcal{G}(U) - \eta). \tag{2.14}$$

We note that the distribution of U^a is in general different from the posterior distribution. Moreover, U^a represents a single EnKF update. Thus, for the continuous time limit and the limit $t \to \infty$, the analysis variable cannot be used to derive the limit of the mean $m(t)$, since the EnKF applies infinitely many update steps.

Remark 2.3. *The authors of [18] show that the EnKF can be adjusted in a way that the covariance of the ensemble does not collapse for $t \to \infty$. Instead, the posterior is reached for $t \to \infty$. In their approach, the noise is not added in the data space as in equation (2.9). Instead, it is added in the input space. However, for rare event estimation, the noise η and noise covariance Γ are artificial. Therefore, it is useful to consider $t > 1$ for the EnKF approach in [46], which constitutes a downscaling of the noise as noted in Remark 2.2. We will use a stopping criterion to determine when the downscaling of the noise is sufficient.*

3 EnKF for rare event estimation

In the following, we apply the EnKF approach of [46] to the forward model $\mathcal{G} = \widetilde{G}$. With $\widetilde{G}$, the EnKF generates failure samples. To reduce the number of iterations, we apply the adaptive approach of [23], which is similar to the adaptive approach of SIS given in [42]. Since the EnKF does not provide samples from the posterior distribution but gives approximate samples from the analysis variable U^a, we apply an IS strategy to estimate the probability of failure. Additionally, we discuss an approach to handle multi-modal failure domains.

3.1 EnKF with adaptive step size

We consider the case that the output space of the LSF G is one-dimensional. This is the usual setting in rare event estimation and simplifies the following considerations. However, the EnKF for rare event estimation can be also applied to a general output space $\mathbb{R}^m$. By [11, 22], we assume, without loss of generality, that the underlying random variable $U : \Omega \to \mathbb{R}^d$ is a d-variate standard normally distributed random variable. Thus, we assume that the prior $\mu_0(u) = \varphi_d(u)$ is the d-variate standard Gaussian density. Translating the rare event setting into the Bayesian framework results in the data $y^\dagger = 0$ and the auxiliary LSF $\widetilde{G}(u) = \max\{0, G(u)\}$, as mentioned in Section 2.1. Now, we can define the rare event problem as an inverse problem. The goal is to find $u^\dagger \in \mathbb{R}^d$ such that

$$0 = \widetilde{G}(u^\dagger) + \eta, \tag{3.1}$$

where $\eta \sim \mathrm{N}(0, \Gamma)$ is observational noise and $\Gamma > 0$. We note that the noise is artificial and we do not know Γ. However, it turns out that Γ is not relevant for the EnKF with adaptive step size and the desired noise level is specified by the user. The posterior density of the inverse problem (3.1) is given by

$$\mu_{y^\dagger}(u) \propto \exp\left(-\frac{1}{2\Gamma}\widetilde{G}(u)^2\right)\varphi_d(u). \tag{3.2}$$

Remark 3.1. *Since the auxiliary LSF $\widetilde{G}$ is always nonlinear, even if the original LSF is linear, the approximation of the derivative by the particles is not exact and the posterior is non-Gaussian. Thus by [14], the EnKF generates approximate samples from the analysis variable U^a (2.14) which differs from the posterior density $\mu_{y^\dagger}$ (3.2).*

Compared with the SIS approach of [42], we can see that (3.2) is an alternative way to define a piecewise smooth approximation of the optimal IS density p_{opt} in (2.3). In the same manner as in [23, 42], we determine the sequence of the EnKF densities adaptively. We define a sequence of temperatures σ_n with $\infty = \sigma_0 > \sigma_1 > \sigma_2 > \cdots > \sigma_N > 0$, where N is the number of a priori unknown EnKF iterations. The sequence of densities in the EnKF is given by

$$\mu_n(u) \propto \exp\left(-\frac{1}{2\sigma_n\Gamma}\widetilde{G}(u)^2\right)\varphi_d(u). \tag{3.3}$$

By the sequential definition of the EnKF it holds that

$$\mu_{n+1}(u) \propto \exp\left(-\frac{1}{2\Gamma}\left(\frac{1}{\sigma_{n+1}} - \frac{1}{\sigma_n}\right)\widetilde{G}(u)^2\right)\mu_n(u).$$

By defining $t_n = 1/\sigma_n$, the temperatures σ_n can be viewed as a time discretization of the SDE (2.12). Considering the limit $\sigma_n \to 0$ in (3.3), or $t_n \to \infty$, leads to pointwise convergence of the EnKF densities to the optimal IS density since

$$\lim_{\sigma_n \to 0} \exp\left(-\frac{1}{2\sigma_n\Gamma}\widetilde{G}(u)^2\right) = I(G(u) \leq 0).$$

This behaviour is illustrated in Figure 1. We note that the limit $\sigma_n \to 0$ implies a complete downscaling of the noise Γ, which results in a classical ill-posed inverse problem. The approximation of the indicator function by SIS [42] is given by

$$\lim_{\sigma_n \to 0} \Phi\left(-\frac{G(u)}{\sigma_n}\right) = I(G(u) \leq 0), \quad \text{for } G(u) \neq 0,$$

where $\Phi(\cdot)$ is the cumulative distribution function of the one-dimensional standard normal distribution. Figure 1 shows that the approximation by the EnKF is equal to the indicator function for $G(u) \leq 0$, while the SIS approximation is symmetric around $G(u) = 0$. Similar to [42], we use the coefficient of variation of the likelihood weights of two subsequent densities to determine the sequence σ_n. Therefore, we use a user specific target coefficient of variation δ_{target}. Given the temperature σ_n, we determine σ_{n+1} by

$$\sigma_{n+1} = \operatorname*{argmin}_{\sigma \in (0, \sigma_n)} \frac{1}{2}\left(\delta_{w_{n+1}} - \delta_{\mathrm{target}}\right)^2, \tag{3.4}$$
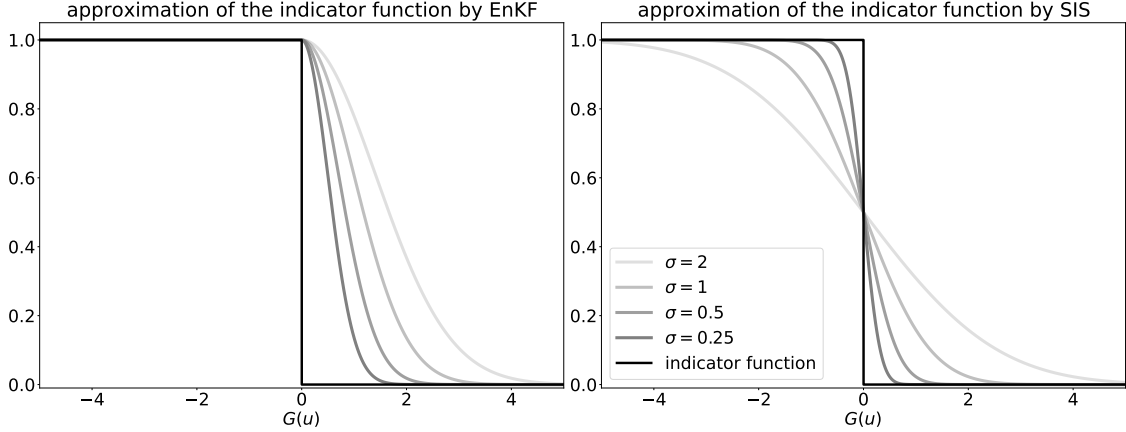
Figure 1: Approximation of the indicator function $I(G(u) \leq 0)$ by the EnKF and SIS.

where $\delta_{w_{n+1}}$ is the coefficient of variation of the likelihood weights $\mathbf{w_{n+1}} = \{w_{n+1}^{(j)}\}_{j=1}^{J}$ of the consecutive EnKF densities μ_n and μ_{n+1}. These weights are given by

$$w_{n+1}^{(j)} = \exp\left(-\frac{1}{2\Gamma}\left(\frac{1}{\sigma_{n+1}} - \frac{1}{\sigma_n}\right)\widetilde{G}(u_n^{(j)})^2\right), \quad \text{for } j = 1, \ldots, J.$$

Through this adaptive procedure, it is not necessary to specify the variance Γ of the observational noise. Therefore, we assume, without loss of generality, that $\Gamma = 1$.

With the new temperature σ_{n+1}, we define the step size of the EnKF as $h_{n+1} = 1/\sigma_{n+1} - 1/\sigma_n$. Hence in one step of the EnKF, the ensemble $\mathbf{u_n}$ is updated by

$$u_{n+1}^{(j)} = u_n^{(j)} + C_{\text{up}}(\mathbf{u_n})\left(C_{\text{pp}}(\mathbf{u_n}) + \frac{1}{h_{n+1}}\right)^{-1}\left(\xi_{n+1}^{(j)} - \widetilde{G}(u_n^{(j)})\right), \quad (3.5)$$

for $j = 1, \ldots, J$, where $\xi_{n+1}^{(j)} \sim \mathrm{N}(0, h_{n+1}^{-1})$ is an additive scaled noise in the output space. The matrices $C_{\text{pp}}(\mathbf{u_n})$ and $C_{\text{up}}(\mathbf{u_n})$ are determined by (2.10) and (2.11), respectively, using $\widetilde{G}$ instead of G. Since the output space is one-dimensional, it holds that $C_{\text{pp}}(\mathbf{u_n}) \in \mathbb{R}$ and $C_{\text{up}}(\mathbf{u_n}) \in \mathbb{R}^d$. Hence, (3.5) does not require the solution of a linear system and the computational costs of one EnKF step behaves as $\mathcal{O}(d \cdot J)$.

As in SIS [42], we use the target coefficient of variation δ_{target} as the stopping criterion of the EnKF iteration. In particular, the EnKF stops if the coefficient of variation of the likelihood weights with respect to the optimal IS density and the current EnKF density is less than δ_{target}. These likelihood weights are given by

$$w_{\text{opt}}^{(j)} = I\left(\widetilde{G}(u_n^{(j)}) = 0\right)\exp\left(\frac{1}{2\sigma_n}\widetilde{G}(u_n^{(j)})^2\right), \quad \text{for } j = 1, \ldots, J. \quad (3.6)$$

The EnKF iteration stops if $\delta_{\mathbf{w}_{\text{opt}}} \leq \delta_{\text{target}}$. By the indicator function and $\widetilde{G}$, the weights $w_{\text{opt}}^{(j)}$ are either zero or one and, therefore, the stopping criterion can be easily interpreted by the *effective sample size* [31, Section 3.4]. This implies that a certain percentage of the particles have to belong to the failure domain to stop the EnKF iteration. This percentage depends on δ_{target}. If δ_{target} is small, then a high percentage of the particles has to belong to the failure domain. For instance, if $\delta_{\text{target}} = 0.25$, then around 94% of the particles has to be in the failure domain. For $\delta_{\text{target}} = 3$, around 10% of the particles has to be in the failure domain. We investigate different values for δ_{target} in the numerical experiments.

Remark 3.2. *We note that SIS applies an MCMC algorithm to shift samples between one density and its consecutive density. Therefore, the generated samples are distributed according to the target densities and the probability of failure can be directly estimated using these samples. However, the EnKF generates approximate samples from the analysis variable [14]. Therefore, we do not use these samples to estimate the probability of failure directly. Instead, we use a single IS step after the EnKF iteration is finished, which yields an unbiased estimator for the probability of failure.*

3.2 Estimation of the probability of failure

Once the EnKF iteration is finished, the final ensemble is denoted by $\mathbf{u}_N$ and the final EnKF density is given by

$$\mu_N(u) \propto \exp\left(-\frac{1}{2\sigma_N}\tilde{G}(u)^2\right)\mu_0(u).$$

We follow [29, Algorithm 3.1] to construct an accurate and unbiased estimator for the probability of failure. Their approach is based on IS. Therefore, we fit a certain distribution model with the final ensemble $\mathbf{u}_N$. Consequently, we generate the ensemble $\hat{\mathbf{u}} = \{\hat{u}^{(j)}\}_{j=1}^J$ from the fitted distribution and estimate the probability of failure by

$$\hat{P}_f = \frac{1}{J}\sum_{j=1}^J I(G(\hat{u}^{(j)}) \leq 0)\frac{\varphi_d(\hat{u}^{(j)})}{p(\hat{u}^{(j)})}, \tag{3.7}$$

where $p(\cdot)$ is the pdf of the fitted distribution. In particular, we consider the GM [17, Section 1] and the vMFNM distribution model [41]. Indeed, the vMFNM performs well in high-dimension while the performance of the GM deteriorates with increase of the dimension.

The GM is defined by the sum of K Gaussian distributions, where each of them is defined by a mean vector m_k and a covariance matrix C_k yielding the pdf $\varphi(\cdot \mid m_k, C_k)$. Thus, the density of the GM distribution is given by

$$p_{\mathrm{GM}}(u) = \sum_{k=1}^K \pi_k\varphi(u \mid m_k, C_k), \tag{3.8}$$

where $\pi_k \geq 0$ is the weight of the kth mixture component with $\sum_{k=1}^K \pi_k = 1$. The parameters of the GM are determined by maximum likelihood estimation using the final EnKF ensemble $\mathbf{u_N}$. For $K > 1$ mixture components, the parameters of the GM are not analytically given and the *expectation-maximization* (EM) algorithm is applied. We refer to [34] for a detailed explanation of the EM algorithm.

The vMFNM distribution is based on the polar decomposition $u = r \cdot a$, where $r = \|u\|_2$ and $a = u/r$. For a single mixture component, the pdf of the vMFN distribution is given by

$$p_{\mathrm{vMFN}}(r, a \mid \nu, \kappa, s, \gamma) = p_{\mathrm{N}}(r \mid s, \gamma) \cdot p_{\mathrm{vMF}}(a \mid \nu, \kappa),$$

where p_{N} is the density of the Nakagami distribution and p_{vMF} is the pdf of the von Mises–Fisher distribution. The Nakagami [38] distribution determines the distribution of the radius r and depends on a shape parameter s and spread parameter γ. The vMF distribution [56] determines the distribution of the direction a and depends on the mean

direction ν and concentration parameter κ. Similar to (3.8), the vMFNM distribution is defined by

$$p_{\text{vMFNM}}(r, a) = \sum_{k=1}^{K} \pi_k p_{\text{vMFN}}(r, a \mid \nu_k, \kappa_k, s_k, \gamma_k).$$

The parameters of the vMFNM distribution are determined by maximum likelihood estimation and the EM algorithm. Algorithm 1 shows the complete algorithm to estimate the probability of failure with the EnKF.

Algorithm 1: EnKF for rare event estimation

1: generate the initial ensemble $\mathbf{u_0}$ by sampling J independent samples from $\varphi_d(u)$
2: evaluate the auxiliary LSF $\widetilde{G}$ for the current ensemble $\mathbf{u_0}$
3: $\sigma_0 \leftarrow \infty$, $n \leftarrow 0$
4: **while** EnKF is not finished **do**
5:     determine σ_{n+1} from the optimization problem (3.4)
6:     define the current step size $h_{n+1} = 1/\sigma_{n+1} - 1/\sigma_n$
7:     calculate $C_{\text{pp}}(\mathbf{u_n})$ and $C_{\text{up}}(\mathbf{u_n})$ based on (2.10),(2.11)
8:     generate J independent samples $\xi_{n+1}^{(j)} \sim \text{N}\left(0, h_{n+1}^{-1}\right)$
9:     update the ensemble $\mathbf{u_n}$ to $\mathbf{u_{n+1}}$ based on (3.5)
10:     evaluate the auxiliary LSF $\widetilde{G}$ for the current ensemble $\mathbf{u_{n+1}}$
11:     determine the coefficient of variation $\delta_{\mathbf{w}_{\text{opt}}}$ based on (3.6)
12:     **if** $\delta_{\mathbf{w}_{\text{opt}}} < \delta_{\text{target}}$ **then**
13:         EnKF is finished
14:     **end if**
15:     $n \leftarrow n + 1$
16: **end while**
17: fit a distribution model based on the final ensemble $\mathbf{u_n}$
18: generate J samples $\hat{\mathbf{u}}$ from the fitted distribution
19: estimate the probability of failure $\hat{P}_f$ based on (3.7)
20: **return** $\hat{P}_f$

We have now defined the procedure of estimating the probability of failure via the EnKF. As we have seen, the GM and vMFNM distributions are able to capture mixture distributions. However, the EnKF in its current form is not able to generate samples from a mixture, since the samples are always concentrated around a single mean value.

3.3 Multi-modal EnKF

In this section, we investigate the approach of [44] such that the EnKF is able to generate samples from a multi-modal failure domain. This property is necessary for rare event estimation since it is possible that the failure domain consists of various distinct modes. To achieve the multi-modal property, the empirical covariance and cross-covariance matrices $C_{\text{pp}}, C_{\text{up}}$ are localised for each particle of the ensemble. Therefore, we calculate a weight matrix $W \in \mathbb{R}^{J \times J}$, which represents the distances of the particles. The entries of the weight matrix are given by

$$W_{i,j} = \exp\left(-\frac{1}{2\alpha}\|u_n^{(i)} - u_n^{(j)}\|_2^2\right), \quad \text{for } i, j = 1, \ldots, J, \tag{3.9}$$

where $\alpha > 0$ is a parameter chosen by the user. In addition, the weight matrix is normalised such that each column sums up to one. A large value of α leads to higher weights for the neighbouring particles. For small α, the weights are only large for the nearest neighbours, which yields more localised covariance matrices and the particles move slowly since the particles do not interact with their neighbours [44].

Remark 3.3. *We note that (3.9) is the value of the pdf $\varphi_d(u_n^{(i)} \mid u_n^{(j)}, \alpha \cdot \mathrm{Id}_d)$, i.e., the Gaussian density with mean vector $u_n^{(j)}$ and covariance matrix $\alpha \cdot \mathrm{Id}_d$, where $\mathrm{Id}_d \in \mathrm{R}^{d \times d}$ denotes the identity matrix.*

To avoid the selection of the parameter α, we propose the following adaptive approach. We first apply a distribution-based clustering through fitting a mixture distribution model. Either of the two distribution models discussed in Section 3.2 can be used for this purpose. The ensemble $\mathbf{u_n}$ is thus split into K clusters. For each cluster $k = 1, \ldots, K$, we determine the empirical covariance matrix C_k by using all particles which belong to the cluster k. Finally, for the particle $u_n^{(j)}$ which belongs to the cluster k, the jth column of the weight matrix is determined by

$$W_{i,j} = \exp\left(-\frac{1}{2} \| C_k^{-1/2} (u_n^{(i)} - u_n^{(j)}) \|_2^2 \right), \quad \text{for } i = 1, \ldots, J. \tag{3.10}$$

Remark 3.4. *Alternative to this adaptive procedure, the parameter α in (3.9) can be chosen as $\alpha \propto d$. This approach could be applied in high-dimensional problems to address the curse of dimensionality of the Euclidean norm. This is due to the fact that the Euclidean distance of the points $u_1 = (0, \ldots, 0) \in \mathbb{R}^d$ and $u_2 = (\epsilon, \ldots, \epsilon)^T \in \mathbb{R}^d$ satisfies*

$$\| u_1 - u_2 \|_2^2 = d\epsilon^2.$$

With the weight matrix W, we calculate for each particle localised means and covariance matrices. The localised means for the particle $u^{(j)}$ are given by

$$\bar{u}_{\mathrm{loc}}^{(j)} = \sum_{i=1}^J W_{i,j} u_n^{(i)}, \qquad \overline{\widetilde{G}}_{\mathrm{loc}}^{(j)} = \sum_{i=1}^J W_{i,j} \widetilde{G}(u_n^{(i)}).$$

With these localised means, we determine the localised covariance matrices by

$$C_{\mathrm{loc,pp}}(u_n^{(j)}) = \sum_{i=1}^J W_{i,j} \left(\widetilde{G}(u_n^{(i)}) - \overline{\widetilde{G}}_{\mathrm{loc}}^{(j)} \right) \otimes \left(\widetilde{G}(u_n^{(i)}) - \overline{\widetilde{G}}_{\mathrm{loc}}^{(j)} \right), \tag{3.11}$$

$$C_{\mathrm{loc,up}}(u_n^{(j)}) = \sum_{i=1}^J W_{i,j} \left(u_n^{(i)} - \bar{u}_{\mathrm{loc}}^{(j)} \right) \otimes \left(\widetilde{G}(u_n^{(i)}) - \overline{\widetilde{G}}_{\mathrm{loc}}^{(j)} \right). \tag{3.12}$$

Finally, one iteration of the EnKF with localised covariance matrices is given by

$$u_{n+1}^{(j)} = u_n^{(j)} + C_{\mathrm{loc,up}}(u_n^{(j)}) \left(C_{\mathrm{loc,pp}}(u_n^{(j)}) + \frac{1}{h_{n+1}} \right)^{-1} \left(\xi_{n+1}^{(j)} - G(u_n^{(j)}) \right), \tag{3.13}$$

for $j = 1, \ldots, J$. After the EnKF iteration is finished, we fit a mixture distribution model and estimate the probability of failure as in Section 3.2. We summarise the EnKF update with localised covariances in Algorithm 2.

Algorithm 2: EnKF update with localised covariances
1: **if** α is given **then**
2:     **for** all particles $u_n^{(j)}$ **do**
3:         determine the jth column of the normalized weight matrix by (3.9)
4:     **end for**
5: **else**
6:     split the ensemble $\mathbf{u_n}$ into K clusters
7:     **for** all clusters k **do**
8:         determine all particles which belong to the cluster k
9:         determine the empirical covariance matrix C_k
10:         **for** all particles $u_n^{(j)}$ belonging to the cluster k **do**
11:             determine the jth column of the normalized weight matrix by (3.10)
12:         **end for**
13:     **end for**
14: **end if**
15: **for** all particles $u_n^{(j)}$ **do**
16:     determine $C_{\mathrm{loc,pp}}(u_n^{(j)})$, $C_{\mathrm{loc,up}}(u_n^{(j)})$ by (3.11), (3.12)
17:     update $u_n^{(j)}$ to $u_n^{(j+1)}$ by (3.13)
18: **end for**

4   EnKF for affine linear LSFs: theoretical properties

In this section, we derive theoretical properties of the EnKF for rare event estimation. We consider the case where the LSF G is affine linear, i.e. $G(u) = a^T u - b$, where $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$. The linearity of G implies that the auxiliary LSF $\tilde{G} = \max\{0, G(u)\}$ is piecewise linear.

In particular, we derive the continuous time limit $h \to 0$ of the EnKF update (3.5) and investigate the large particle limit $J \to \infty$ of the resulting SDE. Thereafter, we study the large time properties $t \to \infty$ to determine the limit of the ensemble mean. We will always consider the noise free case, i.e., $y_{n+1}^{(j)} = y^\dagger$. We do not analyse the analysis variable in (2.14), since we are particularly interested in the continuous time limit and the large time properties.

In the next theorem, we derive the continuous time limit $h \to 0$ of the EnKF update (3.5). We note that we add the failure surface $\{G = 0\}$ to the safe domain for the remaining section. This has no influence on the probability of failure estimate since the failure surface is a set with Lebesgue measure equal to zero. However, adding $\{G = 0\}$ to the safe domain simplifies the following considerations since safe particles remain safe states and failure particles remain failure states for all $t \geq 0$.

Theorem 4.1. *Denote by $S = \{k \in [J] : G(u^{(k)}) \geq 0\}$ and $F = \{k \in [J] : G(u^{(k)}) < 0\}$ the index sets of the safe and failure particles. For the LSF $G(u) = a^T u - b$ and for the noise free case $y_{n+1}^{(j)} = y^\dagger$, the safe particles satisfy the flow*

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C_{\mathrm{uu}}(\mathbf{u})D_u\left(\frac{1}{2}G(u^{(j)})^2\right) + \frac{G(u^{(j)})}{J}\sum_{k \in F}G(u^{(k)})(u^{(k)} - \overline{u}), \qquad (4.1)$$

while failure particles do not move.

Remark 4.2. *The first summand of (4.1) implies that the safe particles move in the direction where the LSF G decreases. If the mean $\overline{u}$ is in the safe domain, it follows that*

13

$G(u^{(k)})(u^{(k)} - \overline{u})$ *points away from the failure surface since* $G(u^{(k)}) < 0$ *for* $k \in F$. *Thus, the second term slows down the movement to the failure surface. Indeed, we prove that the safe particles converge to the surface of the failure domain.*

From Theorem 4.1, we observe that initial failure particles do not move in the EnKF iterations. If initial safe particles reach the failure surface, they will stay at the failure surface for the remaining time. Thus, initial safe particles follow the dynamic (4.1) and do not contribute to the second term in (4.1) for all $t \geq 0$. If the initial ensemble does not contain failure particles, the second term in (4.1) is zero for all $t \geq 0$ and the dynamic simplifies. Therefore, we distinguish the cases if initial failure particles are present or not. In the following, we consider the continuous time limit (4.1) for an infinite ensemble $J \to \infty$. We denote by $U(t)$ the random variable which is distributed according to the particle density of the EnKF particles $\{u^{(j)}\}_{j=1}^{\infty}$ at the time point $t \geq 0$. Moreover, we denote by $m(t) := \mathbb{E}[U(t)]$ the mean of the ensemble and by $C(t) := \mathbb{E}[(U(t) - m(t)) \otimes (U(t) - m(t))]$ the covariance matrix of the ensemble. We derive the limit of the ensemble mean under the following assumptions.

Assumption 4.3. *We assume that*

  (i) $G(u) = a^T u - b$ *with* $a = (1, 0, \dots, 0)^T \in \mathbb{R}^d$ *and* $b < 0$,

  (ii) *the distribution of the input random variable* U *is* d-*variate independent standard normal,*

  (iii) *the EnKF is applied without noise, i.e.,* $y_{n+1}^{(j)} = y^{\dagger}$.

We note that Assumption 4.3 (i) can be satisfied without loss of generality by rotation invariance of the standard normal density.

4.1   Mean-field limit: no failure particles

We begin with the case that the initial ensemble does not contain failure particles. Thus, the second term in (4.1) is zero and the distinction of safe and failure states is not necessary. We note that this assumption is not valid, if we consider the large particle limit $J \to \infty$ and a Gaussian initial ensemble, since the initial ensemble will contain failure states. However, this assumption simplifies the proofs of the statements and it enables us to obtain insights in the particle dynamic and mean-field limit. We consider the case that the initial ensemble contains failure samples in Section 4.2.

In the next lemma, we show that the ensemble mean $m(t)$ converges to the MLFP $u^{\mathrm{MLFP}} \in \mathbb{R}^d$ under Assumption 4.3 with the restriction that the initial ensemble contains no failure particles. The MLFP is the solution of the minimization problem

$$\min_{u \in \mathbb{R}^d} \quad \frac{1}{2} \|u\|_2^2, \quad \text{such that } G(u) = 0.$$

For general nonlinear problems, the FORM approximation of P_f is given by $P_f^{\mathrm{FORM}} = \Phi\left(-\|u^{\mathrm{MLFP}}\|_2\right)$ if $G(0) > 0$. If $G(0) < 0$, $P_f^{\mathrm{FORM}} = 1 - \Phi\left(-\|u^{\mathrm{MLFP}}\|_2\right)$. Indeed, P_f^{FORM} is equal to P_f if the LSF G is affine linear. Assumption 4.3 (i) implies that $u^{\mathrm{MLFP}} = (b, 0, \dots, 0)^T$, $G(0) > 0$, and $\|u^{\mathrm{MLFP}}\|_2 = |b|$.

Theorem 4.4. *Let Assumption 4.3 hold under the restriction that the initial ensemble contains no failure particles even for an infinite ensemble. In the large particle limit $J \to \infty$, the ensemble mean satisfies*

$$m_1(t) = b\left(1 - \frac{1}{\sqrt{2t+1}}\right), \qquad m_i(t) = 0, \quad for \ i = 2, \ldots, d,$$

while the covariance of the ensemble satisfies

$$C(t) = \begin{pmatrix} 1/(1+2t) & \\ & \mathrm{Id}_{d-1} \end{pmatrix}. \tag{4.2}$$

Thus, for the limit $t \to \infty$, it holds

$$\lim_{t\to\infty} m(t) = u^{\mathrm{MLFP}}.$$

Corollary 4.5. *Let Assumption 4.3 hold under the restriction that the initial ensemble contains no failure particles even for an infinite ensemble. The relative distance between u^{MLFP} and the ensemble mean $m(t)$ satisfies*

$$\frac{\|m(t) - u^{\mathrm{MLFP}}\|_2}{\|u^{\mathrm{MLFP}}\|_2} = \frac{1}{\sqrt{2t+1}}.$$

Remark 4.6. *From equation (4.2), we see that the covariance $C(t)$ implies ensemble collapse in the first component for $t \to \infty$. However, the ensemble does not collapse to a single point but it collapses to the surface of the failure domain. Thus, we see that the particles move only in direction perpendicular to the failure surface, or equivalently, in direction of a, until all particles are on the failure surface.*

4.2 Mean-field limit: with failure particles

In this section, we consider the case that the initial ensemble contains failure particles. To derive the mean-field equation, we split the ensemble into the safe and failure particles. We define $U_S(t)$ as the random variable which is distributed according to the particle density of the safe particles $\{u^{(j)} : G(u^{(j)}) \geq 0\}_{j=1}^{\infty}$ at $t \geq 0$. Similar, $U_F(t)$ is the random variable which is distributed according to the particle density of the failure particles $\{u^{(j)} : G(u^{(j)}) < 0\}_{j=1}^{\infty}$ at $t \geq 0$. We note that the distribution of $U_F(t)$ stays constant with respect to t since failure particles do not move.

Since the portion of failure particles in the initial ensemble is equal to P_f for $J \to \infty$, it holds that $\mathbb{P}(U(t) = U_S(t)) = (1 - P_f)$ and $\mathbb{P}(U(t) = U_F(t)) = P_f$. Thus, the mean of the ensemble is given by

$$m(t) = \mathbb{E}[U(t)] = (1 - P_f)\mathbb{E}[U_S(t)] + P_f\mathbb{E}[U_F(t)]$$
$$=: (1 - P_f)m_S(t) + P_f m_F(t). \tag{4.3}$$

The mean $m_F(t)$ of the failure particles is constant and is equal to the mean of the optimal IS density, which is given by

$$m_F = \mathbb{E}[U \mid G(U) < 0] = \mathbb{E}[U \mid U_1 < b] = \left(-\frac{\varphi_1(b)}{\Phi(b)}, 0, \ldots, 0\right)^T =: u^{\mathrm{opt}}, \tag{4.4}$$

where we apply the formula of the mean of a truncated Gaussian [25, Section 10.1]. Indeed, u^{opt} is the mean of a standard Gaussian random variable truncated at $U_1 < b$. The following theorem states that the ensemble mean $m(t)$ converges to a convex combination of the MLFP u^{MLFP} and the mean of the optimal IS density u^{opt} under Assumption 4.3.

Theorem 4.7. *Let Assumption 4.3 hold. For the large particle limit $J \to \infty$, the ensemble mean satisfies*

$$\lim_{t \to \infty} m(t) = (1 - P_f)u^{\mathrm{MLFP}} + P_f u^{\mathrm{opt}}. \tag{4.5}$$

In the following, we point out the structure of the proof of Theorem 4.7, since this gives important insights to the dynamic of the EnKF particles. The formal proofs are given in the Appendix A. From (4.3) and (4.4), it is sufficient to show that the mean of the safe particles $m_S(t)$ converges to u^{MLFP}. To prove this statement, we derive the mean field equation of the particle dynamic (4.1).

Lemma 4.8. *Let Assumption 4.3 hold. For the large particle limit $J \to \infty$, the mean field equation of the safe particles is given by*

$$\frac{\mathrm{d}u_S(t)}{\mathrm{d}t} = -C(t)D_u \left(\frac{1}{2}G(u_S(t))^2 \right) + G(u_S(t))P_f \left((1, 0, \ldots, 0)^T - m(t)(u_1^{\mathrm{opt}} - b) \right). \tag{4.6}$$

Remark 4.9. *Since $D_u \left(\frac{1}{2}G(u)^2 \right) = (G(u), 0, \ldots, 0)^T$, $C(0) = \mathrm{Id}_d$ and $m(0) = 0$, we see from (4.6) that the dynamic acts only on the first component $u_{S,1}(t)$ of the particles for $t = 0$. Moreover, this movement is independent of the other components $u_{S,i}(t)$ for $i = 2, \ldots, d$. Hence, the covariance matrix $C(t)$ and the mean $m(t)$ will only change in their fist components $C_{1,1}(t)$ and $m_1(t)$. Inductively, we conclude that the particle dynamic (4.6) acts only on the first component $u_{S,1}(t)$ for all $t \geq 0$ while all other components $u_{S,i}(t)$ for $i = 2, \ldots, d$ remain constant. We note that this observation is also shown in (4.2) for the case that no failure particles are present.*

The above remark implies that it is sufficient to consider the case $d = 1$. In this case, the mean field equation of the safe particles reads as

$$\frac{\mathrm{d}u_S(t)}{\mathrm{d}t} = -C(t)G(u_S(t)) + G(u_S(t))P_f \left(1 - m(t)(u^{\mathrm{opt}} - b) \right),$$

where $C(t) \in \mathbb{R}$ is the variance of the ensemble. Thus, the mean $m_S(t)$ of the safe particles satisfies the flow

$$\frac{\mathrm{d}m_S(t)}{\mathrm{d}t} = -G(m_S(t)) \left(C(t) - P_f \left(1 - m(t)(u^{\mathrm{opt}} - b) \right) \right). \tag{4.7}$$

To prove Theorem 4.7, we show that the mean of the safe particles converges to u^{MLFP}. Since $m_S(t) \geq u^{\mathrm{MLFP}}$ and $G(m_S(t)) \geq 0$, it is sufficient to show that

$$C(t) - P_f \left(1 - m(t)(u^{\mathrm{opt}} - b) \right) > 0, \quad \text{for all } t \geq 0, \tag{4.8}$$

as long as $m_S(t) \neq u^{\mathrm{MLFP}}$. This guarantees that the dynamic (4.7) is negative and, thus, m_S converges to u^{MLFP}. Since the ensemble contains always the initial failure particles, we can bound the variance $C(t)$ from below.

Lemma 4.10. *We consider $d = 1$. If $m_S(t) \neq u^{\mathrm{MLFP}}$ and the safe particles are not collapsed to a single point, the variance $C(t)$ is bounded from below by*

$$C(t) > (1 - P_f)m_S(t)^2 + P_f(1 + bu^{\mathrm{opt}}) - ((1 - P_f)m_S(t) + P_f u^{\mathrm{opt}})^2.$$

In the proof of Theorem 4.7, see Appendix A.2, we show with Lemma 4.10 that (4.8) is valid for all $t \geq 0$ as long as $m_S(t) \neq u^{\mathrm{MLFP}}$. Together with the fact that $\mathrm{d}m_S(t)/\mathrm{d}t = 0$ for $m_S(t) = u^{\mathrm{MLFP}}$, we conclude that

$$\lim_{t \to \infty} m_S(t) = u^{\mathrm{MLFP}},$$

which implies that (4.5) holds true.

Remark 4.11. *We note that the derived theoretical properties only hold true if the EnKF is applied without noise. In this case, the safe particles converge to the surface of the failure domain and remain there. However, the EnKF is more robust if noise is added to the observations. Therefore, we propose to employ the EnKF with noise for practical applications.*

Remark 4.12. *Theorem 4.4 and 4.7 give a justification to consider the limit $t \to \infty$ or $\sigma \to 0$ even for nonlinear LSFs since we expect that the particles converge to the failure domain and a high number of particles is in proximity of the MLFP.*

Remark 4.13. *Since P_f is typically small, we see from Theorem 4.7 that the EnKF can be applied to estimate u^{MLFP} and to approximate P_f by P_f^{FORM}.*

5   Numerical experiments

We consider four numerical experiments to test the performance of the EnKF for rare event estimation. In all experiments, we compare the results of the EnKF with SIS. We consider the SIS algorithm given in [54] which applies the vMFNM distribution as proposal density in the MCMC algorithm. In all experiments, we apply SIS without burn-in and 10% of the samples are chosen as seeds for the simulated Markov chains via multinomial resampling. Hence, the simulated Markov chains have length equal to 10. The EnKF is always applied with noise $\xi_{n+1}^{(j)} \sim \mathrm{N}(0, h_{n+1}^{-1})$, where h_{n+1} is the step size of the EnKF update. We refer to the standard EnKF iteration (3.5) as the *EnKF with global covariances* and the multi-modal approach of Section 3.3 as the *EnKF with local covariances.*

The performance is measured via the required computational costs and the achieved *relative root mean square error* (relRMSE), which is defined as

$$\mathrm{relRMSE} := \frac{\left(\mathbb{E}\left[(\hat{P}_f - P_f)^2 \right] \right)^{\frac{1}{2}}}{P_f},$$

where $\hat{P}_f$ is the estimated probability of failure by the EnKF or SIS. The computational costs are equal to the number of required LSF evaluations. In all plots, we remove the probability of failure estimates which are larger than the 99th percentile of the estimates. This removes outliers[1] which might occur during the iterations. We are aware that these occur due to the choice of the parametric model of the IS density p which leads to large likelihood weights in (3.7) in very infrequent simulation runs. This observation in also discussed in [39, Section 9.3]. The percentage of outliers in the EnKF estimates is between

---

[1] An estimate $\hat{x}$ is an outlier if $\hat{x} \geq x_{0.75} + 3(x_{0.75} - x_{0.25})$, where $x_{0.25}$ and $x_{0.75}$ are the 25th and 75th percentiles. This criterion is based on Tukey's fences [50, Section 2D] and indicates that an estimate is *far out.*

$0.6\% - 2.0\%$ for the first three examples and 3.2% for the fourth example. Hence, by removing particles which are larger than the 99th percentile, not all outliers are removed. The percentage of outliers in the SIS estimates is smaller than 1%.

We begin with three examples which are also considered in [42]. These examples do not require expensive forward model evaluations and their inputs are two-dimensional independent standard normal random variables. Therefore, we can visualize how the ensemble of particles is evolving during the EnKF iterations and compare them with the samples generated by SIS. In the fourth example, we consider the diffusion equation in one-dimensional space with stochastic diffusion coefficient. This example is also considered in [51, 54] and has a high-dimensional parameter space. The relRMSE and the average computational cost are estimated with 500 independent simulation runs for the first three examples and with 100 runs for the fourth example.

5.1 Convex limit-state function

We consider the following convex LSF, which is given in [27],

$$G(u) = 0.1(u_1 - u_2)^2 - \frac{1}{\sqrt{2}}(u_1 + u_2) + 2.5.$$

The corresponding probability of failure is $4.21 \cdot 10^{-3}$ and the failure domain has a single mode [42]. We consider $J \in \{250, 500, 1000, 2000, 5000, 10000\}$ as the ensemble sizes and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$ as the target coefficients of variation. Since the failure domain is unimodal, we apply the EnKF with global covariances. We apply either the GM or vMFNM distribution model with one mixture component to fit the distribution of the final EnKF particles. We apply SIS with the vMFNM distribution model with one mixture component as the proposal density in the MCMC step.

Figure 2 shows the samples of the final iteration of the EnKF and SIS for the convex LSF. Note that these are not the samples of the fitted distribution after the final step of the EnKF. We see that a small target coefficient of variation leads to more samples within the failure domain. This holds true for both, the EnKF and SIS. We observe that the EnKF samples are more spread along the surface of the failure domain. By Theorem 4.4 this observation is expected since the covariance of the ensemble stays constant in the direction parallel to the failure surface. In addition, the mean of the EnKF particles is in proximity to the MLFP and the mean of the optimal IS density as expected by Theorem 4.7. In contrast, the SIS samples are centered around a certain mean value and are more similar to the optimal IS density. For $\delta_{\text{target}} = 1.00$, the EnKF moves more samples into the failure domain than SIS. Therefore, we expect that for larger δ_{target} the EnKF performs better than SIS.

Figure 3 shows the relRMSE on the horizontal axis and the computational costs on the vertical axis for the convex LSF. We observe that the EnKF reaches a slightly higher level of accuracy than SIS. For a larger number of samples per level, both methods have a smaller error. However, δ_{target} has a smaller influence on the error for the EnKF than for SIS. For $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00\}$ the error stays constant for the EnKF with the vMFNM distribution. The error increases only for the two largest target coefficients of variations. In contrast, the error for SIS increases as δ_{target} increases. For the EnKF with GM, the behaviour of the error is similar to SIS. In summary, the EnKF requires less computational costs than SIS for a fixed level of accuracy, which fits our expectations from Figure 2.
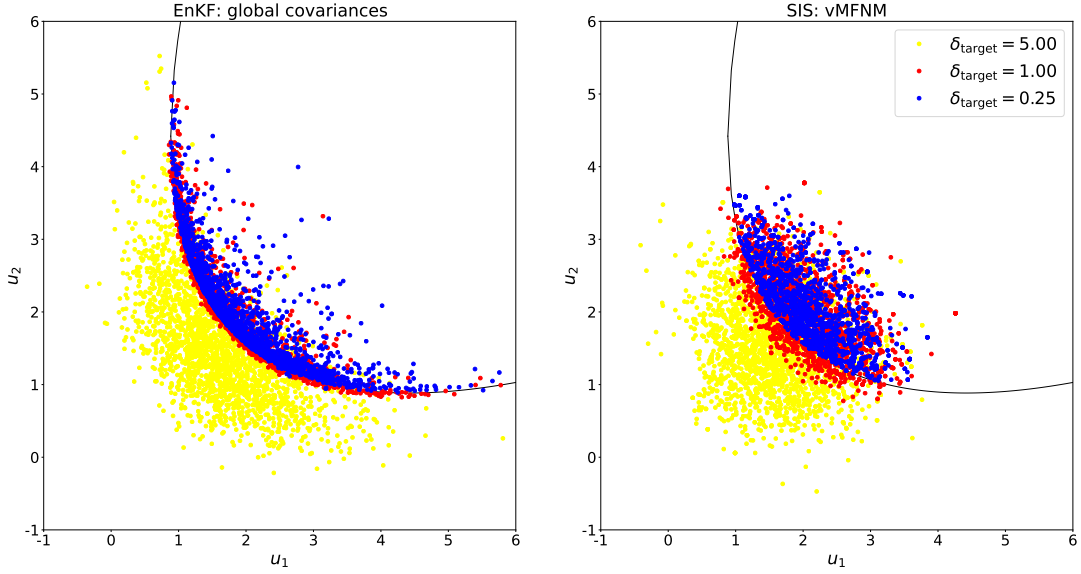
Figure 2: Convex LSF: Samples at the end of the iterations of the EnKF and SIS for $\delta_{\text{target}} \in \{0.25, 1.00, 5.00\}$ and for 2000 samples per level. The black lines show the boundary of the failure domain. Left: Samples of the EnKF with global covariances. Right: Samples of SIS with vMFNM and one mixture component.
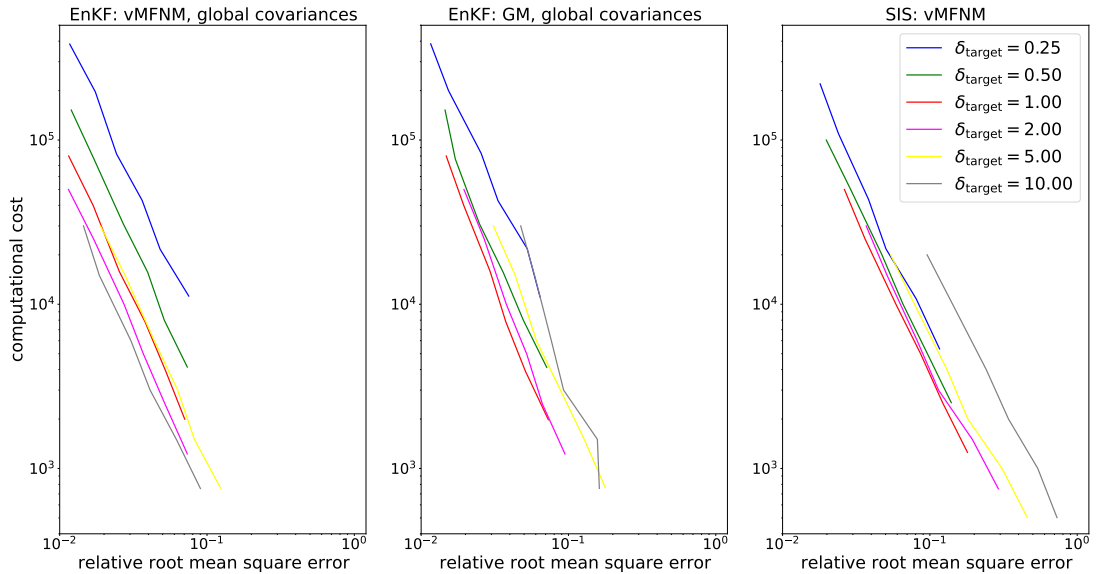

Figure 3: Convex LSF: Computational costs and relRMSE of the EnKF and SIS averaged over 500 runs for $J \in \{250, 500, 1000, 2000, 5000, 10000\}$ samples per level and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. Left: EnKF with vMFNM and global covariances; Middle: EnKF with GM and global covariances; Right: SIS with vMFNM. One mixture component is applied in the EnKF and SIS.

5.2 Parabolic limit-state function

In [10], the following parabolic LSF is proposed

$$G(u) = 5 - u_2 - \frac{1}{2}(u_1 - 0.1)^2.$$

The exact probability of failure is $3.01 \cdot 10^{-3}$ [42]. In this case, the failure domain consists of two distinct areas with high probability mass. Therefore, we apply the EnKF with local covariances. We set $\alpha = 2$ and apply two mixture components to fit the GM and vMFNM distribution model in the final IS step of the EnKF. SIS is applied with the vMFNM distribution model with two mixture components. Moreover, we do not consider $J = 10000$ for the ensemble size. Apart from this, we consider the same settings as for the convex LSF in the previous section.
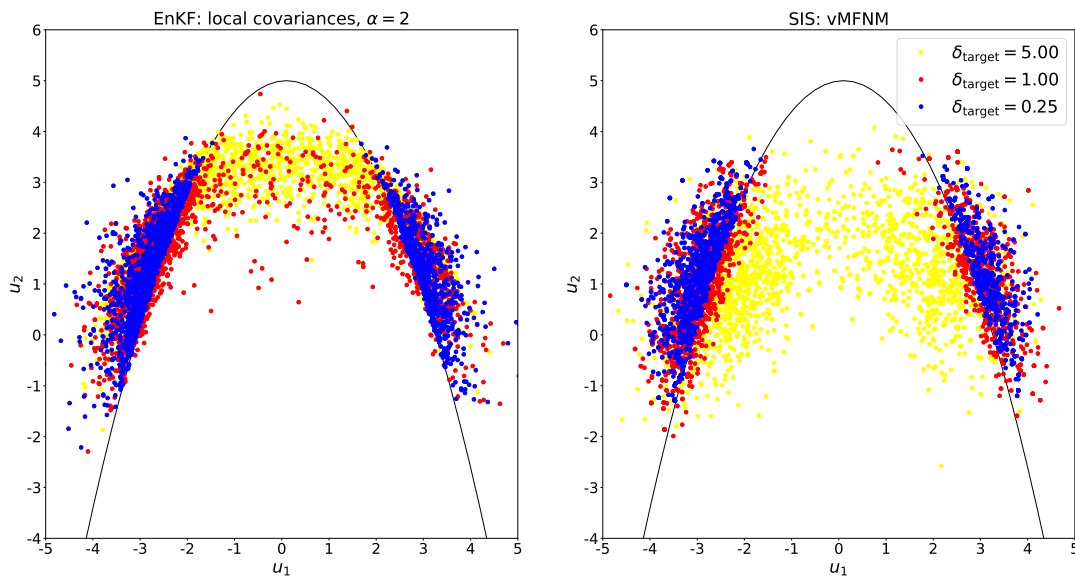

Figure 4: Parabolic LSF: Samples at the end of the iterations of the EnKF with local covariances and SIS for $\delta_{\text{target}} \in \{0.25, 1.00, 5.00\}$ and for 2000 samples per level. Two mixtures are considered for the distribution models. The black lines show the boundary of the failure domain. Left: Samples of the EnKF with local covariances and $\alpha = 2$. Right: Samples of SIS with vMFNM and two mixtures.

Figure 4 shows the evolved samples of the final iteration of the EnKF and SIS for the parabolic LSF. We see for both methods that the generated samples concentrate near the two separated failure modes. However, for the EnKF with $\delta_{\text{target}} = 5$, many samples are not contained in the failure domain but are in between of the two failure modes. For SIS with $\delta_{\text{target}} = 5$, the samples are more concentrated around the failure modes. As in the previous example, we observe that the samples of the EnKF are more spread along the surface of the failure domain.

Figure 5 shows the relRMSE and the computational costs for the parabolic LSF. Again, we observe that the EnKF reaches the same level of accuracy as SIS. For a larger number of samples per level, both methods have a smaller error. For $\delta_{\text{target}} \in \{0.25, 0.50, 1.00\}$ the error of the EnKF is similar while for larger target coefficient of variations the error is larger for both distribution models. This behaviour is also observed in Figure 4. For $\delta_{\text{target}} = 5$, the samples do not separate clearly in the two failure modes. For SIS, we observe that the error decrease for decreasing δ_{target}, which implies larger computational
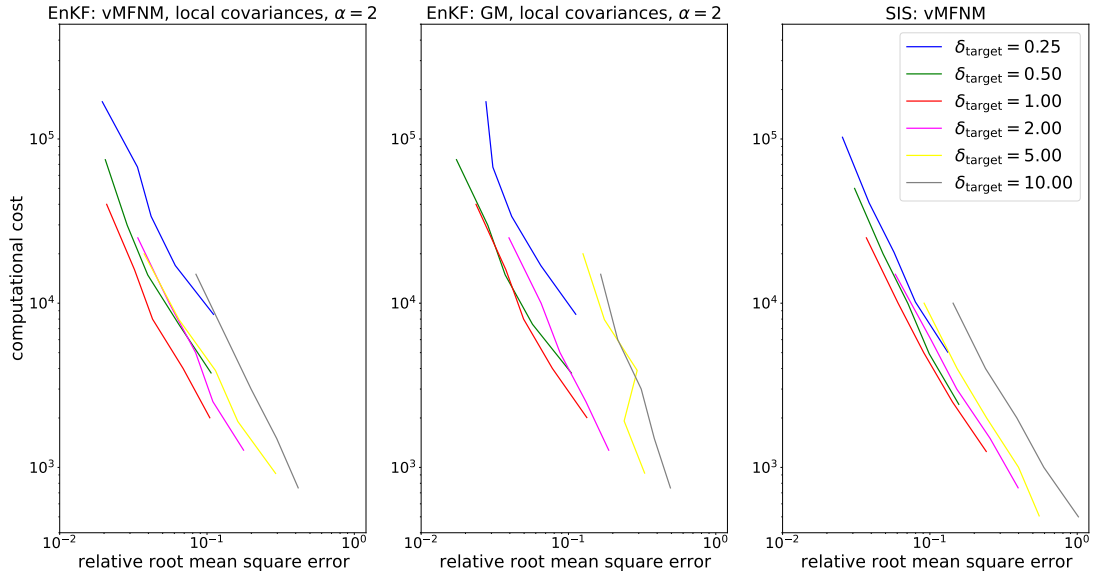
Figure 5: Parabolic LSF: Computational costs and relRMSE of the EnKF and SIS averaged over 500 runs for $J \in \{250, 500, 1000, 2000, 5000\}$ samples per level and $\delta_{\mathrm{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. Left: EnKF with vMFNM and local covariances, $\alpha = 2$; Middle: EnKF with GM and local covariances, $\alpha = 2$; Right: SIS with vMFNM. Two mixture components are applied in the EnKF and SIS.

costs. As for the convex LSF, the EnKF requires less computational costs than SIS for a fixed level of accuracy. Since $\alpha = 2$ yields promising results, we do not consider a parameter study for α, nor do we consider the adaptive approach given in Section 3.3.

5.3 Series system reliability problem

In the third example, we consider a series system reliability problem given in [53], which is defined by the LSF

$$G(u) = \min \left\{ \begin{array}{c} 0.1(u_1 - u_2)^2 - (u_1 + u_2)/\sqrt{2} + 3 \\ 0.1(u_1 - u_2)^2 + (u_1 + u_2)/\sqrt{2} + 3 \\ u_1 - u_2 + 7/\sqrt{2} \\ u_2 - u_1 + 7/\sqrt{2} \end{array} \right\}.$$

The corresponding probability of failure is $2.2 \cdot 10^{-3}$ and the failure domain consists of four distinct modes [42]. We apply SIS with the vMFNM distribution model with four mixtures components. The EnKF is applied with local covariances and four mixture components in the final fitting step. For this example, we consider the two approaches given in Section 3.3 for determining the weight matrix W in (3.9). We start by performing a parameter study for the parameter α. Thereafter, we consider the adaptive approach given in Section 3.3, which we reference as the *EnKF with adaptive local covariances*.

We consider the parameter values $\alpha \in \{0.10, 0.25, 0.50, 0.75, 1.00, 2.00\}$ and apply the EnKF with local covariances and 2000 samples per level. Figure 6 shows the generated samples for varying α and $\delta_{\mathrm{target}} \in \{1.00, 5.00\}$. We observe that the samples are more concentrated for small α. Indeed, for smaller values for α, the samples separate into four failure modes. For $\alpha \geq 1$, nearly all samples are contained in two failure modes. In summary, for $\alpha \in \{0.25, 0.50, 0.75\}$ we expect good results since the samples capture well the four modes. For $\alpha = 0.10$, the samples are too concentrated. Thus, we choose
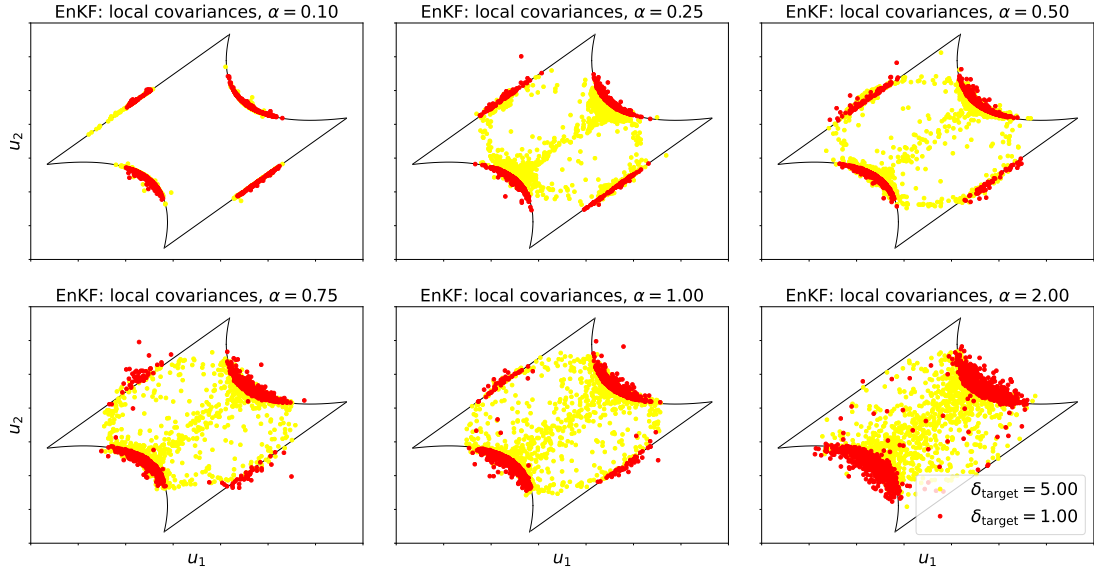
Figure 6: Series system LSF: Samples at the end of the iterations of the EnKF with local covariances for $\delta_{\text{target}} \in \{1.00, 5.00\}$, $\alpha \in \{0.10, 0.25, 0.50, 0.75, 1.00, 2.00\}$ and 2000 samples per level. The black lines show the boundary of the failure domain.
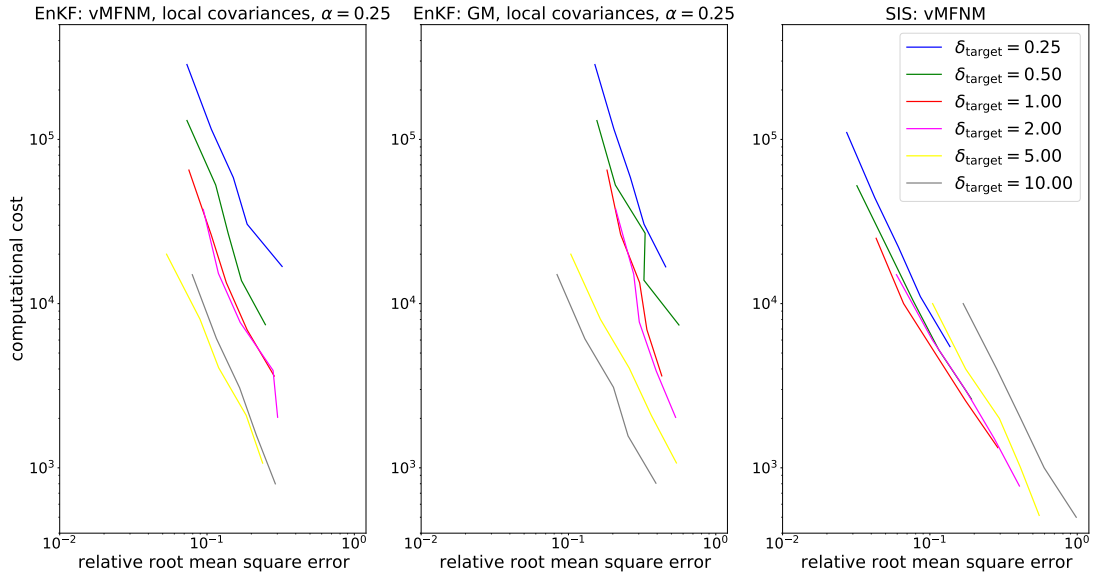

Figure 7: Series system LSF: Computational costs and relRMSE of the EnKF and SIS averaged over 500 runs for $J \in \{250, 500, 1000, 2000, 5000\}$ samples per level and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. Left: EnKF with vMFNM and local covariances, $\alpha = 0.25$; Middle: EnKF with GM and local covariances, $\alpha = 0.25$; Right: SIS with vMFNM. Four mixture components are applied in the EnKF and SIS.

$\alpha = 0.25$ and investigate the respective performance in more detail.

Figure 7 shows the performance of the EnKF with $\alpha = 0.25$. The ensemble size is $J \in \{250, 500, 1000, 2000, 5000\}$ and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$ is the target coefficient of variation. We observe that a larger value for δ_{target} leads to a smaller error. This is due to the fact that for small δ_{target} the EnKF particles are more concentrated and do not always split into the four failure modes. Therefore, the estimates contain a bias. In particular for $\delta_{\text{target}} \in \{5.00, 10.00\}$, the EnKF yields a good performance. However, SIS requires less computational costs for a fixed level of accuracy.

In the following, we consider the adaptive approach of Section 3.3 to determine the weight matrix W. In particular, we fit the GM or vMFNM distribution model with four mixture components in each EnKF step to split the particles in a cluster with four components. Consequently, we calculate the empirical covariance matrix for each cluster and determine the weight matrix W by (3.10).
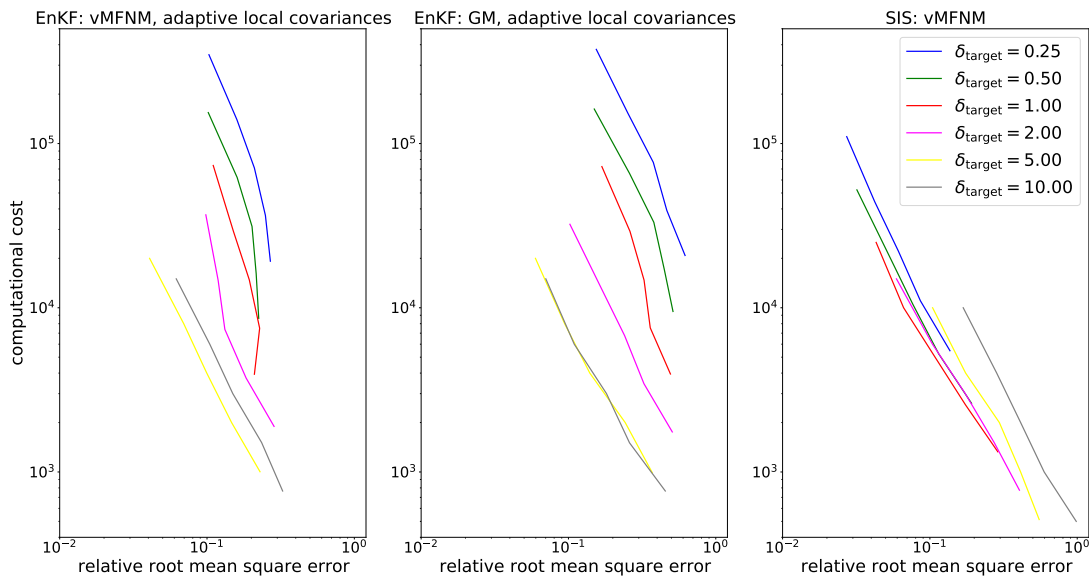

Figure 8: Series system LSF: Computational costs and relRMSE of the EnKF and SIS averaged over 500 runs for $J \in \{250, 500, 1000, 2000, 5000\}$ samples per level and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. Left: EnKF with vMFNM and adaptive local covariances; Middle: EnKF with GM and adaptive local covariances; Right: SIS with vMFNM. Four mixture components are applied in the EnKF and SIS.

Figure 8 shows the relRMSE and the computational costs for the EnKF with adaptive local covariances. We observe a similar behaviour of the error as in Figure 7. A larger value for δ_{target} yields a smaller error. However, the EnKF with the vMFNM and $\delta_{\text{target}} = 5.00$ requires less computational costs than SIS for a fixed level of accuracy.

We note that both local covariance approaches require more tempering steps. This is due to the fact that the particles interact less with their neighbours and, thus, move slower. This observation is also made in [44].

5.4 1D diffusion equation

The final example considers the diffusion equation in the one-dimensional domain $D = (0, 1)$. For $\mathbb{P}$-almost every (a.e.) $\omega \in \Omega$, we seek the weak solution $y(\cdot, \omega) \in H_0^1(D)$ such

that for all $v \in H_0^1(D)$ it holds

$$\int_D a(x,\omega)\frac{\partial}{\partial x}y(x,\omega) \cdot \frac{\partial}{\partial x}v(x)\mathrm{d}x = \int_D v(x)\mathrm{d}x.$$

The diffusion coefficient $a(x,\omega) = \exp(Z(x,\omega))$ is a log-normal random field. It is specified by its mean function $\mathrm{E}[a(x,\cdot)] = 1$ and standard deviation $\mathrm{Std}[a(x,\cdot)] = 0.1$. Thus, the mean function of Z is $\mu_Z = \log(\mathrm{E}[a(x,\cdot)]) - \sigma_Z^2/2$ and the variance is given by $\sigma_Z^2 = \log\left((\mathrm{Std}[a(x,\cdot)]^2 + \mathrm{E}[a(x,\cdot)]^2)/\mathrm{E}[a(x,\cdot)]^2\right)$. Moreover, we assume that Z has an exponential type covariance function $c(x,y) = \sigma_Z^2 \exp\left(-\|x - y\|_1/\lambda\right)$ with correlation length $\lambda = 0.01$. The truncated Karhunen–Loève (KL) expansion

$$Z_d(x,\omega) = \mu_Z + \sigma_Z \sum_{m=1}^{d} \sqrt{\nu_m}\theta_m(x)U_m(\omega)$$

gives an approximation to the infinite-dimensional random field Z. Thus, we define $a_d = \exp(Z_d)$ as an approximation for a. The associated solution of the weak form is denoted by y_d. The eigenpairs (ν_m, θ_m) can be analytically calculated as explained in [19, Section 2.3.3]. Moreover, $U := \{U_m\}_{m=1}^d$ are independent standard normal Gaussian random variables. We set $d = 150$ which captures 87% of the variability of $\log(a)$.

In addition, we approximate the solution y_d by piecewise linear, continuous finite elements on a uniform grid with mesh size $h = 1/512$. The finite element approximation is denoted by y_h. Finally, we call the event $\omega \in \Omega$ a failure event if the solution $y_h(\cdot,\omega)$ is larger than 0.535 at $x = 1$. This gives the LSF

$$G(U(\omega)) := 0.535 - y_h(x = 1,\omega).$$

By crude Monte Carlo sampling with $2 \cdot 10^8$ samples, the probability of failure is estimated as $P_f = 1.682 \cdot 10^{-4}$. In the following, this value is referred to as the reference solution. We note that the truncation of the KL expansion and the discretization parameter h induces an error in G. Thus, the probability of failure P_f is an approximation to the exact one which requires the exact solution y. Since we always consider a fixed discretization level and fix the number of KL terms, the error is not present in the estimates. For an error analysis with respect to the discretization size h, we refer to [55].

The probability of failure is estimated by the EnKF and SIS. The estimation is performed for $J \in \{250, 500, 1000, 2000\}$ samples per level and target coefficient of variation equal to $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. For the EnKF, we use the vMFNM as distribution model with one mixture. Moreover, we apply global covariances since we expect one single failure mode. The GM distribution is not considered, since it does not perform well in high dimensions. Similar we apply SIS with sampling from the vMFNM distribution with one mixture.

Figure 9 shows the relRMSE and the computational costs for the diffusion equation problem. The EnKF yields the same level of accuracy as SIS. Indeed, the EnKF yields the smallest error with the largest target coefficient of variation. We note that $\delta_{\text{target}} = 10$ requires only one step to reach the stopping criterion. The surface of the failure domain might be highly nonlinear. If δ_{target} is large, the final EnKF particles are more spread, which yields a better fitting distribution in this particular case and a smaller error. In contrast, a smaller value for δ_{target} leads to a smaller error for SIS. This is the complete opposite observation as for the EnKF. This is due to the fact that for small δ_{target}, two consecutive densities in SIS are more similar and the estimation of the probability of failure is more robust. In summary, we conclude that the EnKF requires less computational costs for a fixed level of accuracy.
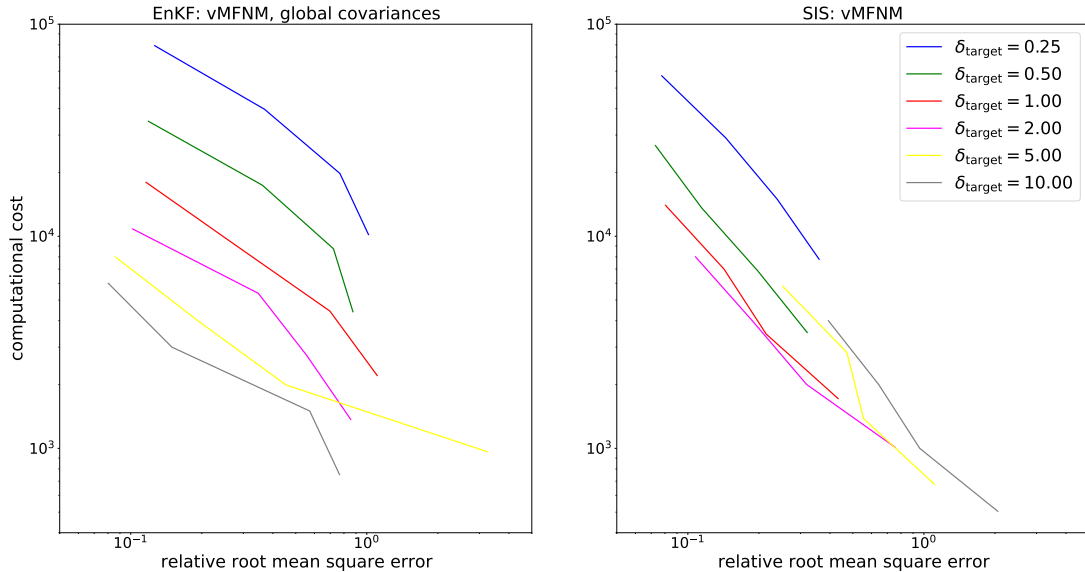
Figure 9: Diffusion equation: Computational costs and relRMSE of the EnKF and SIS averaged over 100 runs for $J \in \{250, 500, 1000, 2000\}$ samples per level and $\delta_{\text{target}} \in \{0.25, 0.50, 1.00, 2.00, 5.00, 10.00\}$. Left: EnKF with vMFNM and global covariances; Right: SIS with vMFNM. One mixture component is applied in the EnKF and SIS.

6 Conclusion and Outlook

We introduce a novel sampling method for estimating small probabilities of failure that employs the EnKF sampler. The proposed method reformulates the rare event problem as an inverse problem by concatenating the ReLU function with the LSF. For this reformulation, the EnKF is applied to generate failure samples in an adaptive manner. Consequently, a distribution model is fitted with the generated samples and the probability of failure is estimated with IS. We have shown that the EnKF densities define an alternative sequence of piecewise smooth approximations of the optimal IS density as compared to the sequence employed in SIS.

For affine linear LSFs, we have derived the particle dynamic for the continuous time limit of the EnKF update. Under the assumption that the EnKF is applied without noise, we have proven that the ensemble mean converges to a convex combination of the most likely failure point and the mean of the optimal IS density in the large particle and large time limit.

To handle multi-modal failure domains, we localise the covariance matrices in the EnKF update around each particle. This localisation can be made adaptively using a clustering approach.

In numerical experiments, we compare the EnKF with SIS in terms of the relative root mean square error and the required computational costs. For single modal failure domains, the EnKF requires less computational costs than SIS for a fixed level of accuracy. However, for multi-modal failure domains, the application of the EnKF is not straightforward and SIS yields a better performance in some cases.

For future work, our manuscript can be used as a starting point to analyse the EnKF for rare events for more general settings. In particular, the analysis for nonlinear LSFs and the analysis of the mean field limit of the EnKF with noise is still an open question. Moreover, the combination of the EnKF with a multilevel strategy would be beneficial for LSFs which can be approximated by a hierarchy of discretization levels. Finally, the

EnKF algorithm could be potentially extended for estimating the probabilities of failure associated with multivariate outputs. This approach could be beneficial in cases where the multiple outputs are based on a single model solve.

A Proofs of Section 4

Proof of Theorem 4.1. The authors of [46] show for the continuous time limit $h \to 0$ of (3.5) and the noise free case that the particles satisfy the flow

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = \frac{1}{J}\sum_{k=1}^{J}\langle \widetilde{G}(u^{(k)}) - \overline{\overline{\widetilde{G}}}, -\widetilde{G}(u^{(j)})\rangle(u^{(k)} - \overline{u}), \tag{A.1}$$

where we used that $y^\dagger = 0$ and $\Gamma = 1$. If $\widetilde{G}(u^{(j)}) = 0$, (A.1) implies that

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = 0.$$

Thus, failure particles do not move. Now we consider the case that $\widetilde{G}(u^{(j)}) = G(u^{(j)}) > 0$. The ensemble mean $\overline{\overline{\widetilde{G}}}$ can be expressed as

$$\overline{\overline{\widetilde{G}}} = \frac{1}{J}\sum_{k=1}^{J}\widetilde{G}(u^{(k)}) = \frac{1}{J}\sum_{k\in S}G(u^{(k)}) = \frac{1}{J}\sum_{k=1}^{J}G(u^{(k)}) - \frac{1}{J}\sum_{k\in F}G(u^{(k)}) = G(\overline{u}) - C,$$

since G is affine linear for all u and we define $C := \frac{1}{J}\sum_{k\in F}G(u^{(k)}) < 0$. Splitting up (A.1) into the safe and failure states and using that $\widetilde{G}(u^{(k)}) = G(u^{(k)})$ for $k \in S$ and $\widetilde{G}(u^{(k)}) = 0$ for $k \in F$, we get

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = \frac{1}{J}\left(\sum_{k\in S}\langle G(u^{(k)}) - G(\overline{u}) + C, -G(u^{(j)})\rangle(u^{(k)} - \overline{u})\right. \tag{A.2}$$

$$\left. + \sum_{k\in F}\langle -G(\overline{u}) + C, -G(u^{(j)})\rangle(u^{(k)} - \overline{u})\right). \tag{A.3}$$

The sum in (A.2) is equal to

$$\sum_{k\in S}\langle G(u^{(k)}) - G(\overline{u}), -G(u^{(j)})\rangle(u^{(k)} - \overline{u}) + \sum_{k\in S}\langle C, -G(u^{(j)})\rangle(u^{(k)} - \overline{u}) := S_1 + S_2,$$

while the sum in (A.3) is equal to

$$\sum_{k\in F}\langle G(u^{(k)}) - G(\overline{u}), -G(u^{(j)})\rangle(u^{(k)} - \overline{u}) + \sum_{k\in F}\langle C - G(u^{(k)}), -G(u^{(j)})\rangle(u^{(k)} - \overline{u})$$
$$:= F_1 + F_2,$$

where we have used the linearity of the scalar product. Adding S_1 and F_1 and multiplying with $1/J$ gives

$$\frac{1}{J}(S_1 + F_1) = \frac{1}{J}\sum_{k=1}^{J}\langle G(u^{(k)}) - G(\overline{u}), -G(u^{(j)})\rangle(u^{(k)} - \overline{u})$$

$$= -C_{\mathrm{uu}}(\mathbf{u})D_u\left(\frac{1}{2}G(u^{(j)})^2\right), \tag{A.4}$$

where we have applied (2.13) since G is affine linear. Adding the remaining parts gives

$$\frac{1}{J}(S_2 + F_2) = \frac{\langle C, -G(u^{(j)}) \rangle}{J} \sum_{k=1}^{J} (u^{(k)} - \overline{u}) + \frac{1}{J} \sum_{k \in F} \langle G(u^{(k)}), G(u^{(j)}) \rangle (u^{(k)} - \overline{u})$$

$$= \frac{G(u^{(j)})}{J} \sum_{k \in F} G(u^{(k)})(u^{(k)} - \overline{u}), \tag{A.5}$$

since $\sum_{k=1}^{J} \left(u^{(k)} - \overline{u} \right) = 0$. Adding (A.4) and (A.5) gives the desired result. $\qquad\square$

A.1 Proof of Section 4.1

Proof of Lemma 4.4. We follow the proof of [18, Lemma 3.2] and adjust it to the rare event setting. The second part of (4.1) is zero for all $t \geq 0$ since all initial particles are in the safe domain and, at the time point a particle reaches the failure surface, it does not move anymore. Therefore, the dynamic of all particles satisfies for all $t \geq 0$

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C_{\mathrm{uu}}(\mathbf{u})D_u \left(\frac{1}{2} G(u^{(j)})^2 \right) = -C_{\mathrm{uu}}(\mathbf{u}) \left(aa^T u^{(j)} - ab \right).$$

The large particle limit $J \to \infty$ leads to the mean field equation at $t \geq 0$

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -C(t) \left(aa^T u(t) - ab \right), \tag{A.6}$$

where $u(t)$ is a realisation of $U(t)$. With (A.6), we can derive the dynamic of the mean and covariance matrix. The ensemble mean satisfies

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = -C(t) \left(aa^T m(t) - ab \right). \tag{A.7}$$

By defining $e(t) = U(t) - m(t)$, we get

$$\frac{\mathrm{d}e(t)}{\mathrm{d}t} = -C(t)aa^T e(t). \tag{A.8}$$

For the covariance it holds that $C(t) = \mathbb{E}[e(t) \otimes e(t)]$. Differentiating $C(t)$ with respect to t and plugging in (A.8), it follows

$$\frac{\mathrm{d}C(t)}{\mathrm{d}t} = \mathbb{E}\left[\frac{\mathrm{d}e(t)}{\mathrm{d}t} \otimes e + e \otimes \frac{\mathrm{d}e(t)}{\mathrm{d}t} \right] = -2C(t)aa^T C(t). \tag{A.9}$$

From (A.9), it follows for the inverse of the covariance matrix

$$\frac{\mathrm{d}C^{-1}(t)}{\mathrm{d}t} = -C^{-1}(t) \left(\frac{\mathrm{d}C(t)}{\mathrm{d}t} \right) C^{-1}(t) = 2aa^T. \tag{A.10}$$

With the initial condition $C(0) = \mathrm{Id}_d$ and from (A.10), it follows

$$C(t) = \left(\mathrm{Id}_d + 2aa^T t \right)^{-1} = \begin{pmatrix} 1/(1 + 2t) & \\ & \mathrm{Id}_{d-1} \end{pmatrix},$$

where $\mathrm{Id}_{d-1} \in \mathbb{R}^{(d-1) \times (d-1)}$ is the identity matrix. Inserting the expression of the covariance matrix $C(t)$ in (A.7) gives

$$\frac{\mathrm{d}m_1(t)}{\mathrm{d}t} = \frac{b - m_1(t)}{1 + 2t}, \qquad \frac{\mathrm{d}m_i(t)}{\mathrm{d}t} = 0, \quad \text{for } i = 2, \ldots, d.$$

With the initial condition $m(0) = 0$, the entries of the mean are given by

$$m_1(t) = b\left(1 - \frac{1}{\sqrt{2t+1}}\right), \qquad m_i(t) = 0, \quad \text{for } i = 2, \ldots, d.$$

Since $\lim_{t \to \infty} m_1(t) = b$, we conclude that $\lim_{t \to \infty} m(t) = u^{\mathrm{MLFP}}$, which is the desired result. $\qquad \square$

A.2 Proofs of Section 4.2

Proof of Lemma 4.8. Theorem 4.1 gives the continuous time limit of the particle dynamic. We consider the mean field limit $J \to \infty$ for the two parts in (4.1) separately. For the first part, it holds that

$$\lim_{J \to \infty} -C_{\mathrm{uu}}(\mathbf{u}) D_u \left(\frac{1}{2} G(u^{(j)})^2\right) = -C(t) D_u \left(\frac{1}{2} G(u^{(j)})^2\right).$$

Now, we consider the second part in (4.1). We split the sum into two parts as

$$\frac{G(u^{(j)})}{J} \sum_{k \in F} G(u^{(k)})(u^{(k)} - \overline{u})$$

$$= G(u^{(j)}) \frac{|F|}{J} \left(\frac{1}{|F|} \sum_{k \in F} G(u^{(k)}) u^{(k)} - \frac{\overline{u}}{|F|} \sum_{k \in F} G(u^{(k)})\right). \qquad (A.11)$$

At first, we see that $\lim_{J \to \infty} |F|/J = P_f$ since $|F|$ is the number of failure particles in the initial ensemble. The limit of the first sum in (A.11) is

$$\lim_{J \to \infty} \frac{1}{|F|} \sum_{k \in F} G(u^{(k)}) u^{(k)} = \mathbb{E}[G(U) U \mid G(U) < 0]$$

$$= \mathbb{E}[U_1 U \mid U_1 < b] - b \mathbb{E}[U \mid U_1 < b]. \qquad (A.12)$$

The second tern in (A.12) is equal to $-bu^{\mathrm{opt}}$. Since the components of U are independent, it holds that $\mathbb{E}[U_1 U \mid U_1 < b] = (\mathbb{E}[U_1^2 \mid U_1 < b], 0, \ldots, 0)^T$.
Since $\mathrm{Var}[U_1 \mid U_1 < b] = \mathbb{E}[U_1^2 \mid U_1 < b] - \mathbb{E}[U_1 \mid U < b]^2$ we conclude that

$$\mathbb{E}[U_1^2 \mid U_1 < b] = \mathrm{Var}[U_1 \mid U_1 < b] + \mathbb{E}[U_1 \mid U_1 \leq b]^2$$

$$= 1 - b\frac{\varphi(b)}{\Phi(b)} - \frac{\varphi^2(b)}{\Phi^2(b)} + \left(-\frac{\varphi(b)}{\Phi(b)}\right)^2 = 1 + bu_1^{\mathrm{opt}},$$

where we used the formula of the mean and variance of a truncated Gaussian [25, Section 10.1]. In summary we get

$$\lim_{J \to \infty} \frac{1}{|F|} \sum_{k \in F} G(u^{(k)}) u^{(k)} = (1, 0, \ldots, 0)^T.$$

For the second sum in (A.11), we get

$$\lim_{J \to \infty} -\frac{\overline{u}}{|F|} \sum_{k \in F} G(u^{(k)}) = -m(t) \mathbb{E}[G(U) \mid G(U) < 0] = -m(t)(u_1^{\mathrm{opt}} - b)$$

by linearity of G. In summary we conclude that

$$\lim_{J \to \infty} \frac{G(u^{(j)})}{J} \sum_{k \in F} G(u^{(k)})(u^{(k)} - \bar{u}) = G(u^{(j)}) P_f \left((1, 0, \ldots, 0)^T - m(t)(u_1^{\text{opt}} - b) \right).$$

Together with the first limit, we get for $J \to \infty$ that

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = -C(t) D_u \left(\frac{1}{2} G(u^{(j)})^2 \right) + G(u^{(j)}) P_f \left((1, 0, \ldots, 0)^T - m(t)(u_1^{\text{opt}} - b) \right),$$

which is the desired result. □

Proof of Lemma 4.10. We consider $d = 1$. For the variance it holds that

$$\begin{aligned}
C(t) = \mathrm{Var}[U(t)] &= \mathbb{E}[U(t)^2] - \mathbb{E}[U(t)]^2 \\
&= (1 - P_f)\mathbb{E}[U(t)^2 \mid U(t) \geq b] + P_f \mathbb{E}[U(t)^2 \mid U(t) < b] \\
&\quad - ((1 - P_f)m_S(t) + P_f m_F(t))^2.
\end{aligned}$$

Again, we use

$$\mathbb{E}[U(t)^2 \mid U(t) \geq b] = \mathrm{Var}[U(t) \mid U(t) \geq b] + \mathbb{E}[U(t) \mid U(t) \geq b]^2 > m_S(t)^2,$$

since the variance is always positive as long as $U_S(t)$ is not collapsed to a single point. Using that $\mathbb{E}[U(t)^2 \mid U(t) < b] = 1 + bu^{\text{opt}}$ and $m_F = u^{\text{opt}}$, we get

$$C(t) > (1 - P_f)m_S(t)^2 + P_f(1 + bu^{\text{opt}}) - ((1 - P_f)m_S(t) + P_f u^{\text{opt}})^2.$$

□

Proof of Theorem 4.7. It remains to show that

$$C(t) - P_f \left(1 - m(t)(u^{\text{opt}} - b) \right) > 0. \tag{A.13}$$

We check that this is true if $m_S(t) > u^{\text{MLFP}}$. We start by using Lemma 4.10 and splitting up $m(t)$. Thus,

$$\begin{aligned}
&C(t) - P_f \left(1 - m(t)(u^{\text{opt}} - b) \right) \\
&> (1 - P_f)m_S(t)^2 + P_f(1 + bu^{\text{opt}}) - ((1 - P_f)m_S(t) + P_f u^{\text{opt}})^2 \\
&\quad - P_f \left(1 - \left((1 - P_f)m_S(t) + P_f u^{\text{opt}} \right) (u^{\text{opt}} - b) \right) \\
&= P_f(1 - P_f) \left(m_S(t)^2 + m_S(t)(-u^{\text{opt}} - b) + bu^{\text{opt}} \right) := P_f(1 - P_f)f(m_S(t)).
\end{aligned}$$

The quadratic function f has the roots $m_{S,1} = b$ and $m_{S,2} = u^{\text{opt}}$. Moreover, it holds that $f'(b) > 0$ since $b > u^{\text{opt}}$. It follows that $f(m_S(t)) > 0$ for $m_S(t) > b$. Together with $P_f(1 - P_f) > 0$, (A.13) holds for all $t \geq 0$ and, thus,

$$\lim_{t \to \infty} m_S(t) = u^{\text{MLFP}}.$$

□

29

References

[1] S. AGAPIOU, O. PAPASPILIOPOULOS, D. SANZ-ALONSO, AND A. M. STUART, *Importance sampling: intrinsic dimension and computational cost*, Statistical Science, 32 (2017), pp. 405–431, https://doi.org/10.1214/17-STS611.

[2] A. AGARWAL, S. DE MARCO, E. GOBET, AND G. LIU, *Rare event simulation related to financial risks: efficient estimation and sensitivity analysis*, HAL, (2017), https://hal-polytechnique.archives-ouvertes.fr/hal-01219616. Working paper.

[3] S.-K. AU AND J. L. BECK, *Estimation of small failure probabilities in high dimensions by subset simulation*, Probabilistic Engineering Mechanics, 16 (2001), pp. 263–277, https://doi.org/10.1016/S0266-8920(01)00019-4.

[4] S.-K. AU AND Y. WANG, *Engineering Risk Assessment with Subset Simulation*, John Wiley & Sons, Ltd, 2014, https://doi.org/10.1002/9781118398050.

[5] A. BESKOS, A. JASRA, N. KANTAS, AND A. THIERY, *On the convergence of adaptive sequential Monte Carlo methods*, The Annals of Applied Probability, 26 (2016), pp. 1111–1146, https://doi.org/10.1214/15-AAP1113.

[6] D. BLÖMKER, C. SCHILLINGS, P. WACKER, AND S. WEISSMANN, *Well posedness and convergence analysis of the ensemble Kalman inversion*, Inverse Problems, 35 (2019), p. 085007, https://doi.org/10.1088/1361-6420/ab149c.

[7] M. DASHTI AND A. M. STUART, *The Bayesian approach to inverse problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer, Cham, 2017, pp. 311–428, https://doi.org/10.1007/978-3-319-12385-1_7.

[8] M. DE ANGELIS, E. PATELLI, AND M. BEER, *Advanced line sampling for efficient robust reliability analysis*, Structural Safety, 52 (2015), pp. 170–182, https://doi.org/10.1016/j.strusafe.2014.10.002.

[9] P. DEL MORAL, A. DOUCET, AND A. JASRA, *Sequential Monte Carlo samplers*, Journal of the Royal Statistical Society. Series B (Statistical Methodology), 68 (2006), pp. 411–436, https://doi.org/10.1111/j.1467-9868.2006.00553.x.

[10] A. DER KIUREGHIAN AND T. DAKESSIAN, *Multiple design points in first and second-order reliability*, Structural Safety, 20 (1998), pp. 37–49, https://doi.org/10.1016/S0167-4730(97)00026-X.

[11] A. DER KIUREGHIAN AND P.-L. LIU, *Structural reliability under incomplete probability information*, Journal of Engineering Mechanics, 112 (1986), pp. 85–104, https://doi.org/10.1061/(ASCE)0733-9399(1986)112:1(85).

[12] A. DOUCET AND A. M. JOHANSEN, *A tutorial on particle filtering and smoothing: fifteen years later*, in The Oxford Handbook of Nonlinear Filtering, D. Crisan and B. Rozovskii, eds., Oxford University Press, Oxford, 2011, pp. 656–704.

[13] L. DOVERA AND E. DELLA ROSSA, *Multimodal ensemble Kalman filtering using Gaussian mixture models*, Computational Geosciences, 15 (2011), pp. 307–323, https://doi.org/10.1007/s10596-010-9205-3.

[14] O. G. Ernst, B. Sprungk, and H.-J. Starkloff, *Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 823–851, `https://doi.org/10.1137/140981319`.

[15] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*, Springer, Berlin, Heidelberg, 2 ed., 2006, `https://doi.org/10.1007/978-3-642-03711-5`.

[16] G. S. Fishman, *Monte Carlo: Concepts, Algorithms and Applications*, Springer Series in Operations Research, Springer, New York, NY, 1 ed., 1996, `https://doi.org/10.1007/978-1-4757-2553-7`.

[17] S. Frühwirth-Schnatter, *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, Springer, New York, NY, 1 ed., 2006, `https://doi.org/10.1007/978-0-387-35768-3`.

[18] A. Garbuno-Inigo, F. Hoffmann, W. Li, and A. M. Stuart, *Interacting Langevin diffusions: gradient structure and ensemble Kalman sampler*, SIAM Journal on Applied Dynamical Systems, 19 (2020), pp. 412–441, `https://doi.org/10.1137/19M1251655`.

[19] R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, NY, 1 ed., 1991, `https://doi.org/10.1007/978-1-4612-3094-6`.

[20] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57 (1970), pp. 97–109, `https://doi.org/10.2307/2334940`.

[21] M. Herty and G. Visconti, *Kinetic methods for inverse problems*, Kinetic & Related Models, 12 (2019), pp. 1109–1130, `https://doi.org/10.3934/krm.2019042`.

[22] M. Hohenbichler and R. Rackwitz, *Non-normal dependent vectors in structural safety*, Journal of the Engineering Mechanics Division, 107 (1981), pp. 1227–1238, `https://doi.org/10.1061/JMCEA3.0002777`.

[23] M. Iglesias, M. Park, and M. V. Tretyakov, *Bayesian inversion in resin transfer molding*, Inverse Problems, 34 (2018), p. 105002, `https://doi.org/10.1088/1361-6420/aad1cc`.

[24] M. A. Iglesias, K. Law, and A. M. Stuart, *Ensemble Kalman methods for inverse problems*, Inverse Problems, 29 (2013), p. 045001, `https://doi.org/10.1088/0266-5611/29/4/045001`.

[25] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1 of Wiley Series in Probability and Statistics, Wiley, 2 ed., 1994.

[26] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering, 82 (1960), pp. 35–45, `https://doi.org/10.1115/1.3662552`.

[27] S. Katsuki and D. M. Frangopol, *Hyperspace division method for structural reliability*, Journal of Engineering Mechanics, 120 (1994), pp. 2405–2427, `https://doi.org/10.1061/(ASCE)0733-9399(1994)120:11(2405)`.

[28] P. S. KOUTSOURELAKIS, H. J. PRADLWARTER, AND G. I. SCHUËLLER, *Reliability of structures in high dimensions, part I: algorithms and applications*, Probabilistic Engineering Mechanics, 19 (2004), pp. 409 – 417, https://doi.org/10.1016/j.probengmech.2004.05.001.

[29] D. P. KROESE, R. Y. RUBINSTEIN, AND P. W. GLYNN, *Chapter 2 - The cross-entropy method for estimation*, in Handbook of Statistics, C. R. Rao and V. Govindaraj, eds., vol. 31 of Handbook of Statistics, Elsevier, 2013, pp. 19–34, https://doi.org/10.1016/B978-0-444-53859-8.00002-3.

[30] J. LATZ, *On the well-posedness of Bayesian inverse problems*, SIAM/ASA Journal on Uncertainty Quantification, 8 (2020), pp. 451–482, https://doi.org/10.1137/19M1247176.

[31] J. LATZ, I. PAPAIOANNOU, AND E. ULLMANN, *Multilevel sequential Monte Carlo for Bayesian inverse problems*, Journal of Computational Physics, 368 (2018), pp. 154–178, https://doi.org/10.1016/j.jcp.2018.04.014.

[32] K. LAW, A. M. STUART, AND K. ZYGALAKIS, *Data Assimilation: A Mathematical Introduction*, Texts in Applied Mathematics, Springer, Cham, 1 ed., 2015, https://doi.org/10.1007/978-3-319-20325-6.

[33] R. LI, V. PRASAD, AND B. HUANG, *Gaussian mixture model-based ensemble Kalman filtering for state and parameter estimation for a PMMA process*, Processes, 4 (2016), https://doi.org/10.3390/pr4020009.

[34] G. MCLACHLAN AND D. PEEL, *Ml fitting of mixture models*, in Finite Mixture Models, John Wiley & Sons, Ltd, 2000, ch. 2, pp. 40–80, https://doi.org/10.1002/0471721182.ch2.

[35] R. E. MELCHERS AND A. T. BECK, *Structural Reliability Analysis and Prediction*, John Wiley & Sons, Ltd, 3 ed., 2017, https://doi.org/10.1002/9781119266105.

[36] J. MORIO AND M. BALESDENT, *Estimation of Rare Event Probabilities in Complex Aerospace and other Systems*, Woodhead Publishing, 1 ed., 2015, https://doi.org/10.1016/C2014-0-02344-1.

[37] P. MÖRTERS AND Y. PERES, *Brownian Motion*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010, https://doi.org/10.1017/CBO9780511750489.

[38] M. NAKAGAMI, *The m-distribution, a general formula of intensity distribution of rapid fading*, in Statistical Methods in Radio Wave Propagation, W. Hoffman, ed., Pergamon, 1960, pp. 3–36, https://doi.org/10.1016/B978-0-08-009306-2.50005-4.

[39] A. B. OWEN, *Monte Carlo theory, methods and examples*, 2013, https://statweb.stanford.edu/~owen/mc/.

[40] I. PAPAIOANNOU, W. BETZ, K. ZWIRGLMAIER, AND D. STRAUB, *MCMC algorithms for subset simulation*, Probabilistic Engineering Mechanics, 41 (2015), pp. 89–103, https://doi.org/10.1016/j.probengmech.2015.06.006.

[41] I. Papaioannou, S. Geyer, and D. Straub, *Improved cross entropy-based importance sampling with a flexible mixture model*, Reliability Engineering & System Safety, 191 (2019), p. 106564, `https://doi.org/10.1016/j.ress.2019.106564`.

[42] I. Papaioannou, C. Papadimitriou, and D. Straub, *Sequential importance sampling for structural reliability analysis*, Structural Safety, 62 (2016), pp. 66–75, `https://doi.org/10.1016/j.strusafe.2016.06.002`.

[43] R. Rackwitz, *Reliability analysis—a review and some perspectives*, Structural Safety, 23 (2001), pp. 365–395, `https://doi.org/10.1016/S0167-4730(02)00009-7`.

[44] S. Reich and S. Weissmann, *Fokker–Planck particle systems for Bayesian inference: computational approaches*, SIAM/ASA Journal on Uncertainty Quantification, 9 (2021), pp. 446–482, `https://doi.org/10.1137/19M1303162`.

[45] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd, 3 ed., 2016, `https://doi.org/10.1002/9781118631980`.

[46] C. Schillings and A. M. Stuart, *Analysis of the ensemble Kalman filter for inverse problems*, SIAM Journal on Numerical Analysis, 55 (2017), pp. 1264–1290, `https://doi.org/10.1137/16M105959X`.

[47] C. Schillings and A. M. Stuart, *Convergence analysis of ensemble Kalman inversion: the linear, noisy case*, Applicable Analysis, 97 (2018), pp. 107–123, `https://doi.org/10.1080/00036811.2017.1386784`.

[48] K. W. Smith, *Cluster ensemble Kalman filter*, Tellus A: Dynamic Meteorology and Oceanography, 59 (2007), pp. 749–757, `https://doi.org/10.1111/j.1600-0870.2007.00246.x`.

[49] A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numerica, 19 (2010), p. 451–559, `https://doi.org/10.1017/S0962492910000061`.

[50] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley Publishing Company, 1977.

[51] E. Ullmann and I. Papaioannou, *Multilevel estimation of rare events*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 922–953, `https://doi.org/10.1137/140992953`.

[52] F. Uribe, I. Papaioannou, Y. M. Marzouk, and D. Straub, *Cross-entropy-based importance sampling with failure informed dimension reduction for rare event estimation*, arXiv, (2020), `https://arxiv.org/abs/2006.05496`.

[53] P. H. Waarts, *Structural reliability using finite element methods. An appraisal for DARS: Directional Adaptive Response surface Sampling*, Delft University Press, 2000.

[54] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann, *Multilevel sequential importance sampling for rare event estimation*, SIAM Journal on Scientific Computing, 42 (2020), pp. A2062–A2087, `https://doi.org/10.1137/19M1289601`.

[55] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann, *Error analysis for probabilities of rare events with approximate models*, arXiv, (2021), `https://arxiv.org/abs/2008.06368`.

[56] Z. Wang and J. Song, *Cross-entropy-based adaptive importance sampling using von Mises–Fisher mixture for high dimensional reliability analysis*, Structural Safety, 59 (2016), pp. 42–52, `https://doi.org/10.1016/j.strusafe.2015.11.002`.