Structural and Multidisciplinary Optimization (2025) 68:242
https://doi.org/10.1007/s00158-025-04191-9

RESEARCH PAPER

Robust design optimization of structures with multimodal responses using Taylor expansions

Benedikt Kriegesmann[1] · **Jan Krüger**[1] · **Kai Steltner**[1] · **Timo Schmidt**[1] · **Robert Seifried**[1]

Received: 21 April 2025 / Revised: 22 October 2025 / Accepted: 24 October 2025

Abstract

This paper shows how a combination of multiple linear approximations can efficiently capture nonlinear and non-smooth response functions in the framework of robust design optimization. The basic idea is to split the domain of random parameters into multiple subdomains along discontinuities of the response. For each subdomain, a linear approximation is used to determine the stochastic moments of the response function due to the scatter in that subdomain. The total stochastic distribution of the structural response is considered to be a multimodal distribution, and its moments are composed of the approximations in the subdomains. The application to use cases shows that indeed the structural response is multimodal in these cases, and the stochastic moments are shown to be sufficiently accurate using the proposed method.

Keywords Robust design optimization · Probabilistic analysis · Taylor series approximation · Multimodal response

1 Introduction

Slender structures are known to be prone to buckling and, especially when being curved, imperfection sensitive [see, e.g., Arbocz et al. (1969)]. Imperfections, i.e., deviations from the ideal structure, are unknown in the design phase and hence uncertain. This uncertainty is typically tackled by the application of safety factors and/or knockdown factors. Alternatively, uncertain properties can be handled in a probabilistic sense (Elishakoff 2004). In the design optimization of structures, using knockdown factors is equivalent to performing deterministic optimization with a scale factor and can lead to sensitive optimum designs (Zang et al. 2002). Hence, taking uncertainties into account is essential in order to derive a design that is robust with respect to imperfections. This motivates the broad field of optimization under uncertainty or, more specifically, robust design optimization (RDO) and reliability-based design optimization (RBDO) when focusing on probabilistic approaches.

For uncertainty propagation (i.e., probabilistic analyses) Monte Carlo methods combined with surrogate models are frequently used today (Shi et al. 2023; Nan et al. 2024; Jing et al. 2024; Song et al. 2024). However, in optimization problems with many design variables, such as topology optimization or parameter-free shape optimization, there is still a need for much faster (albeit less accurate) approaches. Many of such fast approaches are based on a local approximation of some response function by a Taylor series. For RDO problems, the Taylor series is expanded at the mean vector of input parameters. RBDO approaches are based on approximations at the so-called limit state function, the function that divides the random space into a safe region and a failure region. RBDO approaches are computationally more expensive, as they require solving an optimization problem in each design optimization iteration, which is often circumvented by sequential approaches (Schuëller and Valdebenito 2010). The focus of the current paper lies on RDO problems.

Doltsinis and Kang (2004) used a Taylor expansion of the discretized state equation (static equilibrium), which is referred to as perturbation method. By using the adjoint method, Doltsinis and Kang were able to determine all derivatives that are required for a gradient-based design optimization. In Doltsinis et al. (2005), this approach was extended to nonlinear structural problems. Lazarov et al. (2012) applied the same approach to topology optimization, showing its advantages over sampling-based methods in the context of many design parameters. Instead of perturbing the state equation, one can use a Taylor series expansion of

---

Responsible Editor: Christian Gogu.

✉ Benedikt Kriegesmann
benedikt.kriegesmann@tuhh.de

[1] Hamburg University of Technology, Hamburg, Germany

the response function. This is referred to as (mean value) first-order second-moment method (FOSM), or second-order fourth-moment method (SOFM), depending on the order of the Taylor expansion. While most researchers use second-order approximations (see, e.g., Doltsinis and Kang 2004; Asadpoure et al. 2011), the authors showed that using a first-order approximation is also suitable for certain cases at much less computational cost (Kriegesmann and Lüdeker 2019). For second-order approaches, the computational cost scales at least linearly with the number of random parameters (Jansen et al. 2015; Krüger and Kriegesmann 2024), while the first-order approximation presented in Kriegesmann and Lüdeker (2019) requires solving only two systems of linear equations.

The local approximations, no matter of which polynomial order, are only applicable if the response function is sufficiently smooth. This condition is often violated for stability problems. For instance, the global buckling load of the stiffened panel optimized in Steltner et al. (2022) is a bimodal distribution, although all input distributions are unimodal, see Fig. 1. While here the Taylor expansion still provided a reasonable approximation, in extreme cases such as the von Mises-truss, local approximations completely fail (Kriegesmann 2021). The reason is that depending on the realization of random input parameters, the response may switch between different modes. A figurative example for such a response function is shown in Fig. 2.

In methods that are based on approximations at the limit state, the presence of multiple different failure modes is a well-known phenomenon (Ditlevsen and Madsen 1996; Der Kiureghian and Dakessian 1998). The solution to this is to search all design points and use local approximations there. Hence, for RBDO problems, solutions exist that handle multiple failure modes, no matter whether they originate from a nonlinear response or from multiple responses that are considered. The computational cost increases with each mode though. For RDO problems, however, the only option to handle responses that lead to a multimodal distribution is to use sampling-based methods. The current paper aims at filling this gap.

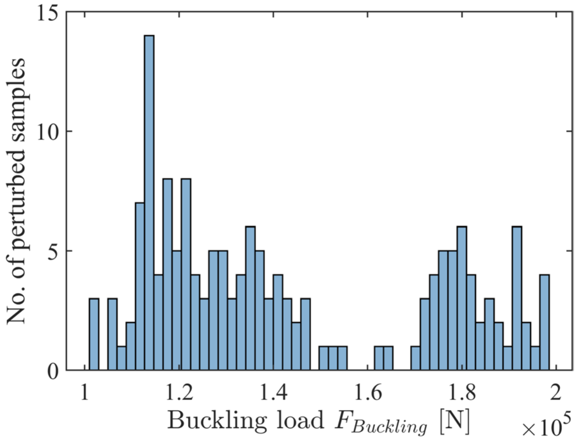


Fig. 1 Histogram of the global buckling load of the panel considered in Steltner et al. (2022) obtained by a Monte Carlo simulation

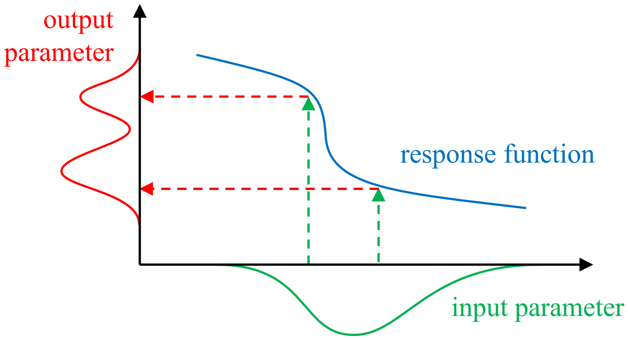


Fig. 2 Example of a function providing a bimodal output distribution for a unimodal input distribution

While sampling-based methods (i.e., Monte Carlo methods) differ fundamentally from those based on local Taylor expansions, there is a transition between both approaches. The Taylor series is used as a simple surrogate, which is used to determine stochastic moments analytically. Other surrogate-based approaches use more samples to set up the surrogate, which is then either used for analytic solutions or for a more efficient Monte Carlo sampling (Jing et al. 2024; Song et al. 2024). Some approaches for creating a surrogate also use gradient information (Nan et al. 2024; Kantarakias and Papadakis 2023), similar to Taylor expansion-based methods. In this spectrum, the current contribution can be considered as making a small step from the Taylor expansion toward gradient-enriched surrogate-based approaches.

For the sake of clarity, we want to emphasize that we do not seek an approach to handle multimodal input distribution as considered, for instance, in He et al. (2016), Hu and Du (2019), and Zhang et al. (2019). Furthermore, we do not address what is often referred to as multimodal optimization, which considers optimization strategies to find multiple minima (see for instance Das et al. 2011). The current paper presents an approach for the RDO of problems with multimodal responses using Taylor series approximations.

2 Robust design optimization of multimodal problems with Taylor expansion

The current section summarizes all theoretical ingredients that are required for the proposed approach. The approach is summarized at the end of this section.

We consider the following form of an optimization problem with the design vector $\boldsymbol{y}$.

$$\begin{aligned} \min_y \quad & \mu_g(y) + \kappa\ \sigma_g(y) \\ \text{s.t.} \quad & c_i(y) = 0] \\ & c_j(y) \leq 0 \end{aligned} \tag{1}$$

Here, μ_g and σ_g are the mean and standard deviation of a response function $g(\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{y}$ is the design vector and $\boldsymbol{x}$ are realizations of the random vector $\boldsymbol{X}$. For the time being, we consider the generic equality and inequality constraints c_i and c_j, respectively.

During the optimization, both stochastic moments (μ_g and σ_g) have to be determined for the current state of $\boldsymbol{y}$, i.e., for the function $g(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{y})$. In the objective, the standard deviation σ_g is weighted by the factor κ, which can be motivated by considering the optimization objective as an approximation of a quantile [see discussion in section 3.3 of Kriegesmann and Lüdeker (2019)]. Alternatively, this formulation can be considered as a multi objective optimization problem where κ is a weighting factor. The choice of κ influences the results, as increasing κ the standard deviation of the robustly optimized design decreases at the cost of an increasing mean value. The interpretation of the objective as a quantile helps for choosing the order of magnitude of κ. If the response function g is assumed to be Gauss distributed, choosing, for instance, $\kappa = 3$ provides the 99.87% quantile, and $\kappa = 4$ provides the 99.997% quantile. A different assumption on the distribution type leads to different values for κ, which, however, have the same order of magnitude.

This chapter presents a method to quickly determine the mean μ_g and the standard deviation σ_g and the gradients of a response function g, which is discontinuous. If the weighted sum of μ_g and σ_g is not the objective, but a constraint instead, all following derivations still apply.

2.1 Moment approximation based on Taylor expansion

Firstly, the well-known approach to determine mean and variance of an objective function using a Taylor series is recalled. For that, the Taylor series approximation of g is expanded in x at the mean vector $\boldsymbol{\mu}_X$ of $\boldsymbol{X}$.

$$\begin{aligned} g(\boldsymbol{x}, \boldsymbol{y}) = {} & g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right) + \sum_{i=1}^{n} \frac{\partial g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right)}{\partial x_i}\left(x_i - \mu_{X_i}\right) \\ & + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right)}{\partial x_i\, \partial x_j}\left(x_i - \mu_{X_i}\right)\left(x_j - \mu_{X_j}\right) + \ldots \end{aligned} \tag{2}$$

Inserting the first-order terms of (2) in the integrals that determine the mean μ_g and the variance σ_g^2 yields

$$\begin{aligned} \mu_g(y) &= \int_{-\infty}^{\infty} g(x, y) f_X(x) dx \approx g\left(\mu_X, y\right) \\ \sigma_g^2(y) &= \int_{-\infty}^{\infty} \left[g(x, y) - \mu_g(y)\right]^2 f_X(x) dx \\ &\approx \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial g\left(\mu_X, y\right)}{\partial x_i} \frac{\partial g\left(\mu_X, y\right)}{\partial x_j} \operatorname{Cov}\left(X_i, X_j\right) \end{aligned} \tag{3}$$

For a second-order approach, e.g., the second-order fourth-moment (SOFM) method, refer for instance to Kriegesmann et al. (2011).

The gradient-based solution of the optimization Problem (1) requires the derivatives of the stochastic moments with respect to each design variable y_k. Considering the FOSM approach, these are given by

$$\begin{aligned} \frac{\partial \mu_g(\boldsymbol{y})}{\partial y_k} &\approx \frac{\partial g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right)}{\partial y_k} \\ \frac{\partial \sigma_g^2(\boldsymbol{y})}{\partial y_k} &\approx 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right)}{\partial x_i \partial y_k} \frac{\partial g\left(\boldsymbol{\mu}_X, \boldsymbol{y}\right)}{\partial x_j} \operatorname{Cov}\left(X_i, X_j\right) \end{aligned} \tag{4}$$

The gradient of the mean equals the gradient of the deterministic function evaluated at the input mean, which is well known for many structural responses and available in commercial code. The gradient of the variance requires the derivative with respect to the random parameter x_j. Despite being less broadly available, this derivative has been given before for various cases (Lazarov et al. 2012; Kriegesmann and Lüdeker 2019). The second mixed partial derivatives are difficult to determine. In Kriegesmann and Lüdeker (2019) and Kranz et al. (2023) the authors avoid computing this derivative by using the adjoint method for the gradient of the variance. In Steltner et al. (2022), Krüger et al. (2023) and Krüger and Kriegesmann (2024), we present semi-intrusive approaches to approximate the derivatives. Note that here we assume that the stochastic distribution is independent of the design parameters, i.e., $\frac{\partial f_X}{\partial y_k} = 0$. For the case of a design-dependent input distribution, the reader is referred to Kriegesmann (2020).

2.2 Moment approximation with piecewise integration

The approximations given in the previous section are very inaccurate, if the objective function g behaves strongly nonlinear (Kriegesmann 2021). One reason for such strong nonlinearities can originate from a discontinuity of the objective function. Here, we present an approach which circumvents the inaccuracy due to discontinuities. For that, we consider the domain Ω_x of the random vector $\boldsymbol{X}$, which may or may not be bounded. As shown in Fig. 3, the domain is split by a function s into two complementary sets Ω_1 and Ω_2, i.e., $\Omega_x = \Omega_1 \cup \Omega_2$. Then, the governing equation for the mean value of the objective can be written as

$$\begin{aligned}\mu_g(y) &= \int_{\Omega_x} g(x,y)\, f_X(x)\, dx \\ &= \int_{\Omega_1} g(x,y)\, f_X(x)\, dx + \int_{\Omega_2} g(x,y)\, f_X(x)\, dx\end{aligned} \tag{5}$$

Now, we define the probability density functions (PDFs) f_1 and f_2 as

$$\begin{aligned} f_1(\boldsymbol{x}) &= \begin{cases} \xi_1 f_X(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \Omega_1 \\ 0 & \text{else} \end{cases} \\ f_2(\boldsymbol{x}) &= \begin{cases} \xi_2 f_X(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \Omega_2 \\ 0 & \text{else} \end{cases} \end{aligned} \tag{6}$$

where ξ_1 and ξ_2 are required to scale the PDF such that its integral over the whole domain of $\boldsymbol{x}$ equals 1. We now introduce the weight factors $\omega_i = 1/\xi_i$. By the definition of the PDFs in (6), we see that for each PDF

$$\begin{aligned}\int_{\Omega_1} f_1(x)\, dx &= \int_{\Omega_1} \xi_1 f_X\, dx \\ &= \frac{1}{\omega_1} \int_{\Omega_1} f_X\, dx = 1 \Leftrightarrow \omega_1 = \int_{\Omega_1} f_X\, dx\end{aligned} \tag{7}$$

Therefore, the sum of the weights must equal one.

$$1 = \int_{\Omega_x} f_X(x)\, dx = \int_{\Omega_1} f_X(\boldsymbol{x})\, dx + \int_{\Omega_2} f_X(x)\, dx = \omega_1 + \omega_2 \tag{8}$$

We can further define

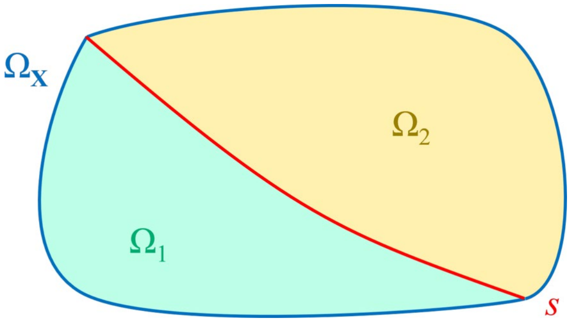


Fig. 3 Schematic illustration of a domain Ω_x split into two subdomains by a function s

$$\mu_1(y) = \int_{\Omega_1} g(x,y)\, f_1(x)\, dx \tag{9}$$

and

$$\mu_2(y) = \int_{\Omega_2} g(x,y)\, f_2(x)\, dx \tag{10}$$

Then, we can write Eq. (5) as

$$\mu_g = \frac{1}{\xi_1}\mu_1 + \frac{1}{\xi_2}\mu_2 = \omega_1\, \mu_1 + \omega_2\, \mu_2 \tag{11}$$

Here, μ_1 and μ_2 are the mean values of the objective g obtained when considering only a part of the domain that is split by the function s. The considerations are identical when the domain is split into more than two parts. Furthermore, the considerations can be extended to the variance. The resulting equations are the same as for mixture distributions, which are briefly described in "Appendix A1." There, the derivatives of the moments are given as well.

2.3 Summary of the proposed approach

With the aforementioned ingredients, it is possible to approximate the first stochastic moments of a discontinuous response function g with multiple Taylor series expansions. Thereby, even the moments of multimodal response distributions can be determined efficiently. The basic idea is to split the input distribution along the hyperplane s for which it is anticipated that the response function has a discontinuity. Then, for each part of the domain, a Taylor expansion-based approximation (for instance, FOSM) is used to determine "partial" stochastic moments. The obtained stochastic moments are combined with the weight factors that are derived from the decomposition of the input domain. The approach will be referred to as "multimodal FOSM" (MM-FOSM) in the following. The main challenge of this approach is that the shape and location of the discontinuity is generally unknown. While a hyperplane is used from Sect. 4 on, more complex shapes of a discontinuity are possible. Both the location and shape of discontinuities can be detected by sampling-based methods (Basudhar et al. 2008), which come with significant computational cost. However, there are engineering applications where the location of the discontinuity can be anticipated sufficiently well, as we show in Sect. 5.

The presented method also allows for determining the gradients of the stochastic moments and therefore, to embed the suggested moment approximations into robust design optimization. Hence, in each optimization iteration, the stochastic moments of g and their gradients are determined for the current design vector $\boldsymbol{y}$.

3 Analytic example of an objective function yielding a multimodal distribution

To demonstrate the general idea and challenges of the multimodal FOSM approach, firstly, an example is constructed that can be solved exactly. Let us consider a random variable X which is uniformly distributed in $[\underline{x}, \bar{x}]$, i.e.

$$f_X = \begin{cases} \frac{1}{\bar{x}-\underline{x}} & \text{for} \quad \underline{x} \leq x \leq \bar{x} \\ 0 & \text{else} \end{cases} \tag{12}$$

Further, we consider the piecewise linear function g with

$$g(x) = \begin{cases} m_L x + c_L & \text{for} \quad x \leq s \\ m_R x + c_R & \text{for} \quad x > s \end{cases} \tag{13}$$

The random distribution of g originating from the randomness of X is a bimodal distribution. The reason is the step of the function at $x = s$. The example is depicted in Fig. 4. For this plot, the following numerical values were chosen: $\underline{x} = 0.5, \bar{x} = 2, m_L = -2/3, m_R = -1/4, c_L = 2, c_R = 1$, and $s = 1$.

We now estimate the stochastic moments of g using the multimodal FOSM approach by splitting the input distribution. For the simple example, it is easy to see that realizations $x \leq s$ contribute to one mode of the output distribution, and realizations $x > s$ contribute to the other mode. Therefore, we split the input distribution at s. We hence get two new input distributions $f_{X,1}$ and $f_{X,2}$ given by

$$f_{X,1} = \begin{cases} \frac{1}{s-\underline{x}} & \underline{x} \leq x \leq s \\ 0 & \text{else} \end{cases} \tag{14}$$

and

$$f_{X,2} = \begin{cases} \frac{1}{\bar{x}-s} & s \leq x \leq \bar{x} \\ 0 & \text{else} \end{cases} \tag{15}$$

Now we perform a FOSM analysis of g for each input distribution using Eq. (3). That means we evaluate the function and its derivative at the mean of the first distribution $\mu_{X,1} = (s + \underline{x})/2$ and at the mean of the second distribution $\mu_{X,2} = (\bar{x} - s)/2$. In order to feed Eq. (3), the (co-)variances of the split distributions $f_{X,1}$ and $f_{X,2}$ are required, which are given by $\sigma^2_{X,1} = (s - \underline{x})^2/12$ and $\sigma^2_{X,2} = (\bar{x} - s)^2/12$.

Table 1 summarizes the results of the mean and standard deviation of the target function g when using the numerical values given before and used for Fig. 4. The FOSM solution considering the whole input distribution of course deviates significantly from the exact solution. The results of the FOSM analyses of the split input distributions are given in Table 1 as well. Weighting $f_{X,1}$ with a factor of $\omega_1 = 1/3$ and $f_{X,2}$ with $\omega_2 = 2/3$ using Eq. (37) yields the moments of the mixed distribution shown in the last row of Table 1.

The simple example shows that when performing FOSM analyses for parts of the input distribution, we get the exact solution of the target moments even if the objective function is discontinuous. However, we have to know the location of the discontinuity (in the example, the position s).

If for the same example, the discontinuity is removed from g, i.e., $m_L = m_R = -2/3$ and $c_L = c_R = 2$, then FOSM and MM-FOSM both provide the exact solution, which is given by $\mu_g = 1.167$ and $\sigma_g = 0.289$.

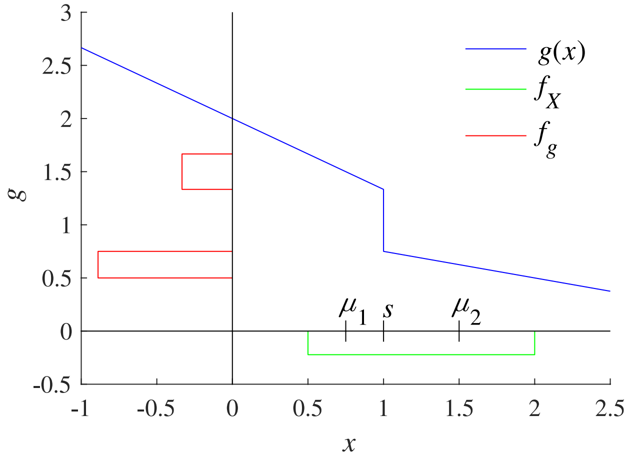


Fig. 4 Analytic example of a linear function (blue) with a step, uniform input distribution (green) and bimodal output distribution (red)

4 Splitting multivariate distributions

Given multiple correlated random parameters, splitting the domain of input parameters is not as straight forward as in the one-dimensional case. A split of the N-dimensional domain is given by a $(N-1)$-dimensional hypersurface. Performing the FOSM approach with each resulting distribution requires the determination of the mean vector and

Table 1 Mean value μ_g and standard deviation σ_g of the objective function of the analytic example

	μ_g	σ_g
Exact solution	0.917	0.420
FOSM solution	0.688	0.108
FOSM for $f_{X,1}$	1.500	0.096
FOSM for $f_{X,2}$	0.625	0.072
MM-FOSM	0.917	0.420

covariance matrix of the distributions obtained by the split. This is demonstrated in the following first for the multivariate standard Gauss distribution, before it is then generalized.

The vector of independent standard Gauss distributed parameters will be denoted $\boldsymbol{Z}$. The independence of $\boldsymbol{Z}$ allows to write its probability density function as $\varphi(z) = \varphi(z_1) \cdot \ldots \cdot \varphi(z_N)$.

4.1 Split along axes

First, we consider a two-dimensional distribution and a split at $z_1 = 0$ as shown in Figs. 5 and 6. The new random vectors are referred to as $\boldsymbol{Z}^+$ and $\boldsymbol{Z}^-$. The mean value of the first entry of $\boldsymbol{Z}^+$ is then given by

$$\begin{aligned} E(Z_1^+) = \mu_1^+ &= \int_0^\infty \int_{-\infty}^\infty z_1\, \xi^+\, \varphi(z_1) \cdot \varphi(z_2)\, dz_2 dz_1 \\ &= \int_0^\infty z_1\, 2\, \varphi(z_1)\, dz_1 \cdot \underbrace{\int_{-\infty}^\infty \varphi(z_2)\, dz_2}_{=1} = \sqrt{\frac{2}{\pi}} \end{aligned} \tag{16}$$

Here, the factor $\xi^+ = 2$ is required to scale φ such that its integral over the domain still equals one, since splitting the distribution at $z_1 = 0$ "halves" the distribution. The mean value μ_1^+ equals the mean of the (one-sided) truncated Gauss distributions. Hence, for a split at $z_1 = 0$ the mean vector of $\boldsymbol{Z}^+$ equals $\boldsymbol{\mu}^+ = (\sqrt{\frac{2}{\pi}}, 0)^T$. In the same manner, the (co-)variances and the moments of $\boldsymbol{Z}^-$ can be determined.

If the random parameters represent the amplitudes of geometric imperfections (as later in Sect. 5), one might want to split the domain not only for one parameter. To demonstrate the approach for multiple splits, we now consider two splits along the coordinate axes as shown in Fig. 7. The mean of the random vector $\boldsymbol{Z}^{++}$, which is defined in the first quadrant, is determined by

$$\begin{aligned} E(Z_1^{++}) &= \int_0^\infty \int_0^\infty z_1 2\varphi(z_1) \cdot 2\varphi(z_2)\, dz_2 dz_1 \\ &= \underbrace{\int_0^\infty z_1 2\varphi(z_1)\, dz_1}_{=\mu_{Z_1^+}} \cdot \underbrace{\int_0^\infty 2\varphi(z_2)\, dz_2}_{=1} \end{aligned} \tag{17}$$

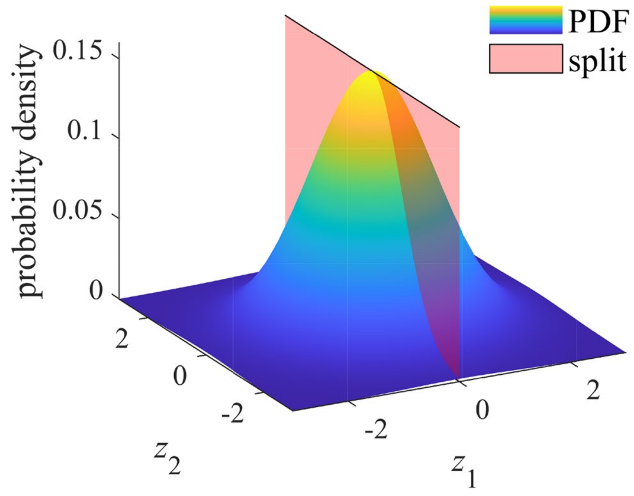


Fig. 5 Example of splitting a two-dimensional Gauss distribution at $z_1 = 0$

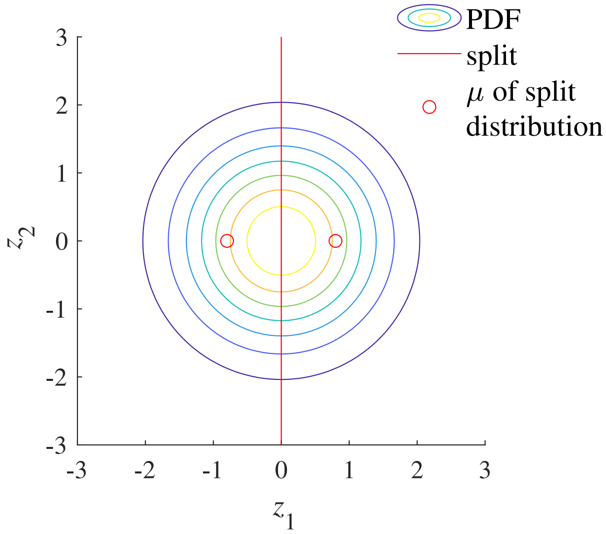


Fig. 6 Example of splitting a two-dimensional Gauss distribution at $z_1 = 0$ shown with iso lines

Again, only the moment of the one-dimensional, truncated Gauss distribution is required and the resulting mean vector

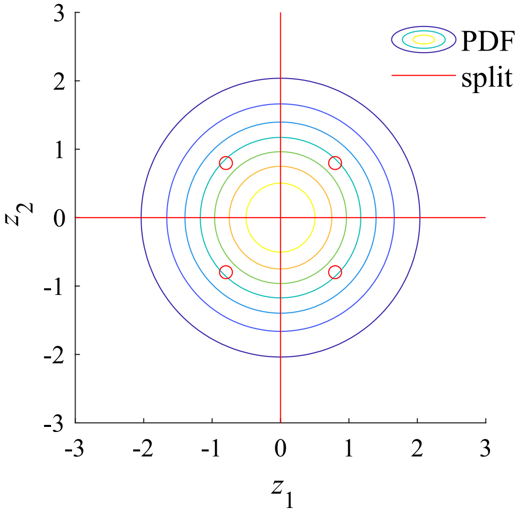


Fig. 7 Example of splitting a two-dimensional Gauss distribution along the coordinate axes

equals $\boldsymbol{\mu}^{++} = (\sqrt{\frac{2}{\pi}}, \sqrt{\frac{2}{\pi}})^T$. In the same manner, it can be shown that the variances in each direction equal $\sigma_1^{++2} = \sigma_2^{++2} = 1 - 2/\pi$ and that the covariance of Z_1^{++} and Z_2^{++} equals zero. Due to the symmetry of the split and following Eq. (7), the weights ω_i are all equal to one over the number of split, i.e., $\omega_i = 1/2^N$ for one split per dimension N.

This procedure can easily be transferred to higher dimensions. Figure 8 exemplarily shows a split of a three-dimensional input space along the axes, providing 8 subdomains. This illustrates that dividing the input space by (hyper-) planes along the axes doubles the number of subdomains with each additional dimension. The number of locations for which Taylor expansion are carried out hence equals 2^N. The computational cost increases exponentially and therefore makes the approach infeasible for high-dimensional problems. Therefore, an alternative approach for splitting the input domain is presented in the following subsection, for which the computational cost only scales linearly with the dimension.

4.2 Split along bisectors

Instead of splitting along the coordinate axes we now consider splits along the bisectors as shown in Fig. 9 for the two-dimensional space and in Fig. 10 for the three-dimensional space. Doing so, the number of subdomains equals $2N$ and hence only scales linearly with the dimension of the input space. Determining the mean values and covariances for the distributions in these subdomains is less straightforward than in the case of splits along axes.

Exemplarily, we consider the stochastic distribution in a subspace that is infinitely open in z_1-direction, which we denote by the superscript 1+. The mean value μ_1^{1+} of Z_1 in this subdomain is determined by

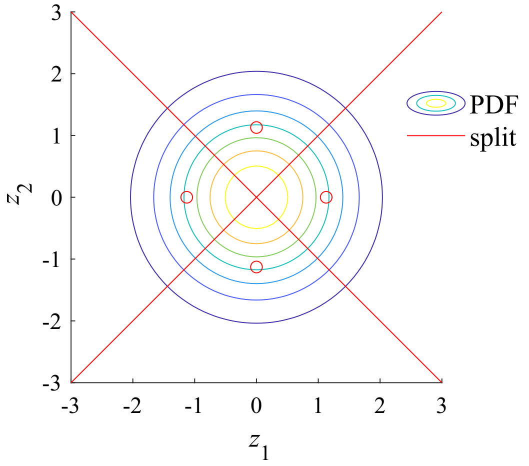


Fig. 9 Example of splitting a two-dimensional Gauss distribution along the bisectors

$$\mu_1^{1+} = \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} z_1\, \xi_1^+\, \varphi(z_1) \dots \varphi(z_N)\, dz_N \dots dz_1 \tag{18}$$

As given in "Appendix A3," this multi-dimensional integral can be reformulated as the following one-dimensional integral

$$\mu_1^{1+} = 2N \int_0^\infty z_1\, \varphi(z_1) \{2\,\Phi(z_1) - 1\}^{N-1} dz_1 \tag{19}$$

that can be solved quickly numerically. Here, Φ is the cumulative distribution function of the standard Gauss distribution. The mean values in all other directions equal zero,

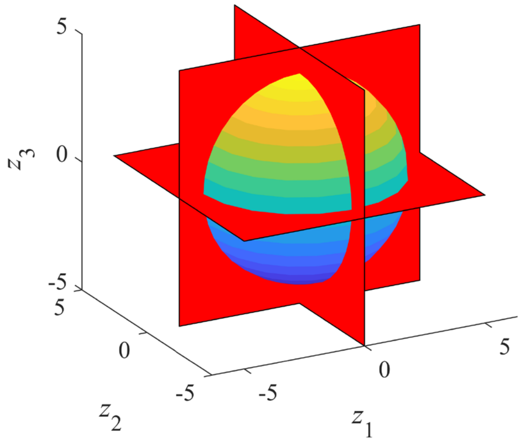


Fig. 8 Example of splitting a three-dimensional Gauss distribution along the coordinate axes(the sphere equals an iso-surface of equal probability density)

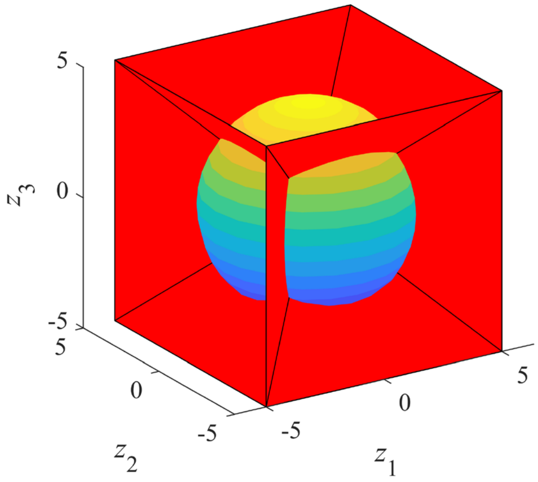


Fig. 10 Example of splitting a three-dimensional Gauss distribution along the bisectors (the sphere equals an iso-surface of equal probability density, the red planes are the spit locations)

which is obvious from symmetry and proven in "Appendix A3". Again, the factor ξ_1^+ scales the PDF such that its integral over the subdomain equals one. For the split along the axes, $\xi_i = 2N$ and hence, $\omega_i = 1/(2N)$ for all subdomains. When considering a different subdomain than “1+,” the absolute value of the mean value in the direction p of the domain still always equals $\mu_p^{p+} = \mu_1^{1+}$. This can be proven by replacing the index 1 in the previous equations by the arbitrary index p. For the subdomains pointing in opposite directions “$p-$,” the mean values equals $\mu_p^{p-} = -\mu_1^{1+}$, which is not only obvious from symmetry, but can also be determined by interchanging the integral bounds in Eq. (19). For different dimensions N, the mean values μ_1^{1+} are summarized in Table 11 in "Appendix A3".

For the (co)variance of the subdomain, four cases are distinguished:

- The variance of the variable pointing in the direction of the subdomain (here exemplarily $\mathrm{Var}(Z_1^{1+}) = \mathrm{Cov}(Z_1^{1+}, Z_1^{1+})$)
- The variance of any other direction (here exemplarily for $\mathrm{Var}(Z_2^{1+}) = \mathrm{Cov}(Z_2^{1+}, Z_2^{1+})$)
- The covariance of the variable pointing in the direction of the subdomain with some other variable (here exemplarily $\mathrm{Cov}(Z_1^{1+}, Z_2^{1+})$)
- The covariance of two variables not pointing the direction of the subdomain (here exemplarily $\mathrm{Cov}(Z_2^{1+}, Z_3^{1+})$)

The variance in z_1-direction is given by

$$\mathrm{Var}\left(Z_1^{1+}\right) = \int_0^{\infty} \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} \left(z_1 - \mu_1^+\right)^2 \xi_1^+ \, \varphi(\boldsymbol{x}) \, dz_N \dots dz_1 \tag{20}$$

and can be simplified to

$$\mathrm{Var}\left(Z_1^{1+}\right) = 2N \int_0^{\infty} \left(z_1 - \mu_1^+\right)^2 \varphi\left(z_1\right) \left\{2\,\Phi\left(z_1\right) - 1\right\}^{N-1} dz_1 \tag{21}$$

The variance in any other direction (here for instance z_2) in the “1+”-subdomain is given by

$$\mathrm{Var}\left(Z_2^{1+}\right) = \int_0^{\infty} \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} \left(z_2 - \mu_2\right)^2 \xi_1^+ \, \varphi(\boldsymbol{x}) \, dz_N \dots dz_1 \tag{22}$$

and can be simplified to

$$\mathrm{Var}\left(Z_2^{1+}\right) = 2N \int_0^{\infty} \varphi\left(z_1\right) \left\{2\Phi\left(z_1\right) - 2\,z_1\,\varphi\left(z_1\right) - 1\right\} \left\{2\Phi\left(z_1\right) - 1\right\}^{N-2} dz_1 \tag{23}$$

All covariances turn out to equal zero. All derivations are given in "Appendix A3" and the standard deviations of z_1 and z_2-direction obtained from solving the above integrals are given in Table 11. The above considerations, exemplarily shown for the “1+”-subdomain, apply similarly to all subdomains of the input domain.

Given a larger set of random parameters (e.g., originating from a discretized random field), one might want to investigate the influence of splitting the domain for each random parameter. For instance, one might start with the first term of a Karhunen–Loève expansion and then, in a second step, also split the direction of the second term and so on. For each increase of the number of splits, the position of the mean of each subdomain changes (see also table 11). In order not to rerun the analyses with steps into z_1-direction, the Taylor expansion has to be adjusted. Such an approach is given in "Appendix A2", where the FOSM approximations of stochastic moments are derived for the case that the objective function is not evaluated at the mean vector, but at some offset position.

So far, all derivations assume standard Gauss distribution and independence of random variables, which is typically not the case in practical application. Therefore, a generalization is discussed in the following subsection.

4.3 Generalization

Random variables of any type of distribution can be transformed to Gauss distributed variables through the Rosenblatt transformation (Rosenblatt 1952). This transformation can be highly nonlinear and may cause a huge loss of accuracy for the Taylor series approximations discussed here. The nonlinearity caused by the Rosenblatt transformation depends on the difference between the original distribution and the Gauss distribution (Meinen and Steenbergen 2025).

For the case of a random vector $\boldsymbol{X}$ with multivariate Gauss distribution (potentially obtained by Rosenblatt transformation) with mean vector $\boldsymbol{\mu}_{\boldsymbol{X}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{X}}$, any realization $\boldsymbol{x}$ is transformed to a standard Gauss distributed vector $\boldsymbol{z}$ with independent entries via

$$\boldsymbol{x} = \boldsymbol{\Sigma}_{\boldsymbol{X}}^{\frac{1}{2}} \boldsymbol{z} + \boldsymbol{\mu}_{\boldsymbol{X}} \Leftrightarrow \boldsymbol{z} = \boldsymbol{\Sigma}_{\boldsymbol{X}}^{-\frac{1}{2}} \left(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{X}}\right) \tag{24}$$

The root of the covariance matrix can be determined by its spectral decomposition, i.e.

$$\boldsymbol{\Sigma}_{\boldsymbol{X}}^{\frac{1}{2}} = \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \tag{25}$$

Here, $\boldsymbol{V}$ is a matrix with eigenvectors of $\boldsymbol{\Sigma}$ and $\boldsymbol{D}$ is a diagonal matrix with the eigenvalues of $\boldsymbol{\Sigma}$. Then, the entries of z represent the principal components of $\boldsymbol{x}$, which motivates creating subdomains in these directions. By employing Eq. (24), the split and also the Taylor series can then be carried out for the transformed vector z, which requires the derivative with respect to the transformed parameters, which is given by

$$\frac{\partial g}{\partial z} = \frac{\partial g}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial z} = \boldsymbol{\Sigma}_X^{\frac{1}{2}}\frac{\partial g}{\partial \boldsymbol{x}} \tag{26}$$

Then, the equations given in the previous subsection apply.

If the random parameters describe geometric imperfections, the mean vector of these imperfections typically does not equal the zero vector. In other words, the geometry does not scatter around the nominal, perfect geometry. However, the configuration of geometric imperfections at which the response function (e.g., a buckling load) varies discontinuously is typically the perfect geometry. Hence, the set of input parameters for which a discontinuity is expected is equal to the mean vector of input parameters. Another example is given in Sect. 5.1, where a beam is loaded eccentrically. The mean eccentricity is assumed to be different from zero, but the discontinuity in the response occurs for an eccentricity of zero (for the initial design). In such situations, it is beneficial to not split the domain at the mean vector, but at the zero vector in original coordinates.

After transforming the zero vector with the Rosenblatt transformation and/or Eq. (24), a non-zero vector is obtained in the space of independent standard Gauss parameters z_i. To be even more general, let us consider some vector $\boldsymbol{s}$, which represents the point of intersection of all splits in the z-space, as shown in Fig. 11. We still assume that the splits are along the bisectors, but now shifted such that they intersect in $\boldsymbol{s}$. Now, the mean vectors as well as the variances differ in all subdomains. Furthermore, the covariances are not equal to zero, i.e., the covariance matrix is not diagonal. Still, all required moments can be determined by solving one-dimensional integrals, which are given in "Appendix A4".

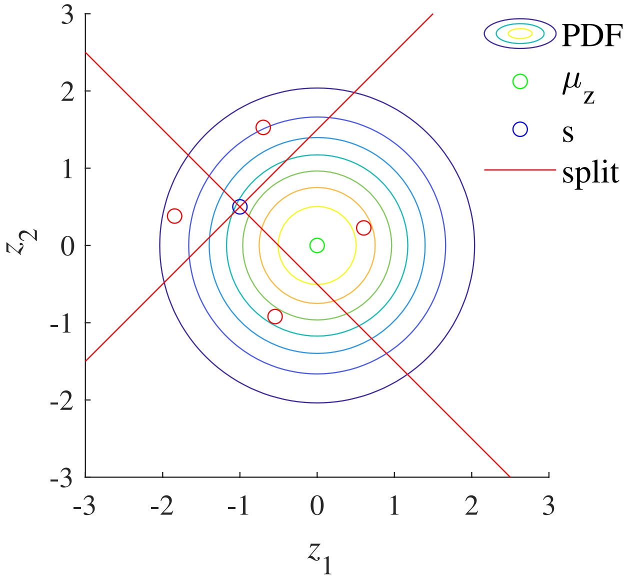


Fig. 11 Example of splitting a two-dimensional Gauss distribution with a split center different from the mean

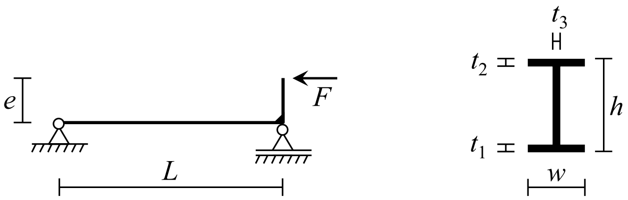


Fig. 12 Sketch of the simple example of a beam with eccentric load introduction (left) and the beam cross section (right)

5 Numerical examples of structural problems

5.1 Optimization of a plastically buckling beam

As a first simple example, we consider a beam that is axially loaded with an eccentricity e as shown in Fig. 12. The length equals $L = 10$ and the nominal value of the load eccentricity equals $e = 0.1$. For the cross section, an I-profile is chosen with a height of $h = 2$, a width of $w = 1$, and a thickness of $t_1 = t_2 = t_3 = 0.1$. An isotropic material is chosen with a Young's modulus of $E = 2.1 \cdot 10^5$, a Poisson's ratio of $\nu = 0.3$, and a yield stress of $\sigma_{pl} = 3 \cdot 10^4$. While the numbers of this academic example are chosen more or less randomly, the unit lengths can be chosen as [mm] and the unit of Young's modulus and yield stress as [MPa]. Load-driven geometrically nonlinear analyses are performed. The load–displacement curves in Fig. 13 of the initial design with two different eccentricities show that the load drops in the postbuckling region due to plastic yielding. The reaction force $\hat{F}$ at the end of the simulations is considered as objective function in the following, which is linked to the end compliance. The analyses are carried out using Abaqus [32], where no gradient information is available for I-beam profile parameters. Therefore, all derivatives required are determined using finite (central) differences for this example. A study on finite difference step size with the initial and the deterministically optimized design showed that a stable approximation of the first derivative is obtained for step sizes of $\Delta t = 0.0001$ and $\Delta e = 0.001$. Due to the size of the example, one finite element analysis took around 5 s, which, together with the small number of design and random parameters, makes it feasible to run with finite differences.

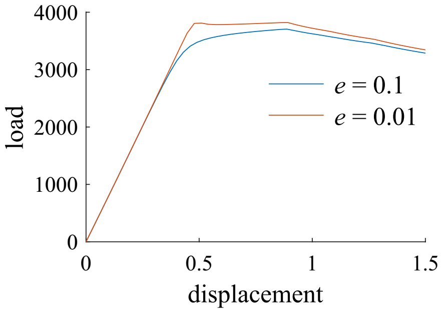


Fig. 13 Load over displacement of the initial design of the eccentrically loaded beam for different eccentricities

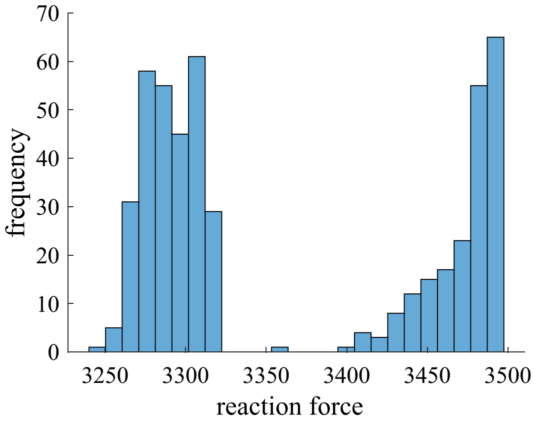


Fig. 14 Histogram of the Monte Carlo simulation of the deterministically optimized beam with eccentricity

For the subsequent optimization, the thicknesses of the profile $\boldsymbol{t} = (t_1, t_2, t_3)^T$ are considered as design parameters. In the robust design optimization, the eccentricity e is considered to be randomly distributed with a mean value of $\mu_e = 0.1$ and a standard deviation of $\sigma_e = 0.1$. Since only one random variable is considered in this simple example, the split is not a hyperplane, but a scalar denotes e_s. For the initial design, a C^1-discontinuity is present at $e = 0$, which is denoted e_0. The eccentricity is assumed to follow Gumbel distribution, for which the cumulative distribution function is given by.

$$F_X(x) = \exp(-\exp(-\frac{x-a}{b})) \tag{27}$$

For the chosen mean value and standard deviation, the Gumbel distribution parameters equal $a = 0.055$ and $b = 0.078$. Through Rosenblatt transformation, the parameters are transformed from the Gumbel distribution to the standard Gauss distribution. For instance, the value of the eccentricity $e_0 = 0$ transformed to standard Gauss space equals $z_0 = -1.1167$. While the stochastic distribution of the reaction force $\hat{F}$ of the initial design turns out to be unimodal, the distribution of the deterministically optimized design shown in Fig. 14 is bimodal.

Two different locations of the split of the input domain are considered for this problem. A split at the zero-eccentricity of $e_0 = 0$ and at the mean $\mu_e = 0.1$. The resulting stochastic moments in the transformed standard Gauss space as well as the weights ω are summarized in Table 2.

The design optimization problem reads

$$\begin{aligned} \min_{\boldsymbol{t}} \quad & \bar{g}(\boldsymbol{t}) \\ \text{s.t.} \quad & A(\boldsymbol{t}) = 0.38 \\ & \boldsymbol{R}(\boldsymbol{t}) = \boldsymbol{0} \\ & 0.01 \leq t_i \leq 0.3 \quad i = 1, 2, 3 \end{aligned} \tag{28}$$

deterministic opti.: $\bar{g}(\boldsymbol{t}) = -\hat{F}(\boldsymbol{t})$

RDO: $\bar{g}(\boldsymbol{t}) = -\mu_{\hat{F}}(\boldsymbol{t}) + 3\ \sigma_{\hat{F}}(\boldsymbol{t})$

where $\boldsymbol{R}$ is the residuum and A equals the cross-section area. While the residuum constraint is implicitly fulfilled by the finite element analysis, the upper and lower bounds of the design variables t_i and the equality constraint of the cross-section area are handled by the optimization algorithm. For the deterministic optimization, the eccentricity is fixed to $e = 0.1$ and the objective function simply reads $\bar{g} = -\hat{F}$, where for robust design optimization, $\bar{g} = -\mu_{\hat{F}}(\boldsymbol{t}) + 3\ \sigma_{\hat{F}}(\boldsymbol{t})$. The optimization problem is solved using the interior-point method implemented in Matlab [33].

The design parameters resulting from the deterministic optimization as well as from RDO using different Taylor series-based approaches are summarized in Table 3. Figure 15 shows the reaction forces $\hat{F}$ for eccentricities sampled in Monte Carlo simulations for the designs obtained by the different optimizations. These results reveal a C^1 discontinuity of the initial design at $e = 0$ and almost linear behavior anywhere else. The deterministic optimization provides a design with the highest observed reaction force of $\hat{F} = 6671$ for the nominal value of $e = 0.1$, which is, however, by far not reached in the Monte Carlo simulations (see Table 4). Moreover, the response is C^0-discontinuous at $e = 0.1$. The latter also holds for the designs obtained in RDO using FOSM and SOFM, where the FOSM design provides much higher reaction forces $\hat{F}$ for random eccentricities. Table 4 provides a cross validation of the stochastic moments obtained for all optimized designs analyzed by different probabilistic approaches. Each row provides the stochastic moments for one and the same design, estimated with different approximation methods. The Monte Carlo results in the last columns are considered as reference. The confidence intervals of the Monte Carlo estimators are given in "Appendix A5". The bold numbers represent the results obtained with the approach used within the optimization. For the

Table 2 Moments for subdomains for the two splits considered

Split at		Negative subsection			Positive subsection		
e-space	z-space	μ^-	σ^{-2}	ω^-	μ^-	σ^{-2}	ω^+
$e_0 = 0$	$z_0 = -1.1167$	−1.6194	0.1860	0.1321	0.2464	0.6641	0.8679
$\mu_e = 0.01$	$z_\mu = 0.1773$	−0.6885	0.4038	0.5704	0.9141	0.3265	0.4296

Table 3 Profile parameters of the optimized beam for different approaches

Optimization approach	Design parameters		
	t_1	t_2	t_3
Deterministic optimization	0.0952	0.1187	0.0930
RDO/FOSM	0.1704	0.1932	0.0100
RDO/SOFM	0.0911	0.1133	0.0978
RDO/MM-FOSM $e_s = 0$	0.1803	0.1833	0.0100
RDO/MM-FOSM $e_s = 0.1$	0.1786	0.1850	0.0100

MM-FOSM analyses, two different locations of the split e_s are considered: one at $e = 0$, which is correct for the initial design, and one for $e = 0.1$, which is correct for the deterministically optimized design and the design obtained with plain FOSM and SOFM. Consequently, MM-FOSM with $e_s = 0.1$ provides very good agreements with the Monte Carlo results for these cases. It is surprising that for the initial design, MM-FOSM with $e_s = 0.1$ is closer to the Monte Carlo results than MM-FOSM with $e_s = 0$. However, both are very close and despite the response function looks almost piecewise linear, there is still some nonlinearity causing deviations from Monte Carlo results. Overall, both MM-FOSM approximations show better comparison with Monte Carlo simulations. Only for the design optimized with the MM-FOSM approach, FOSM and SOFM provide reasonable estimations of the standard deviation. The partially poor approximation of the second-order approach is due to the sensitivity of the second-order derivative estimation to the finite difference step size.

Figure 15 reveals that the location of the discontinuity, which can be anticipated to be present for $e = 0$ for the initial design, is shifted during the optimizations toward the mean eccentricity. This raises the question, how to choose the location of the split e_s for the multimodal approach. We therefore investigate both assumptions one might make, a split at $e_s = e_0 = 0$ and a split at $e_s = \mu_e = 0.1$. For both

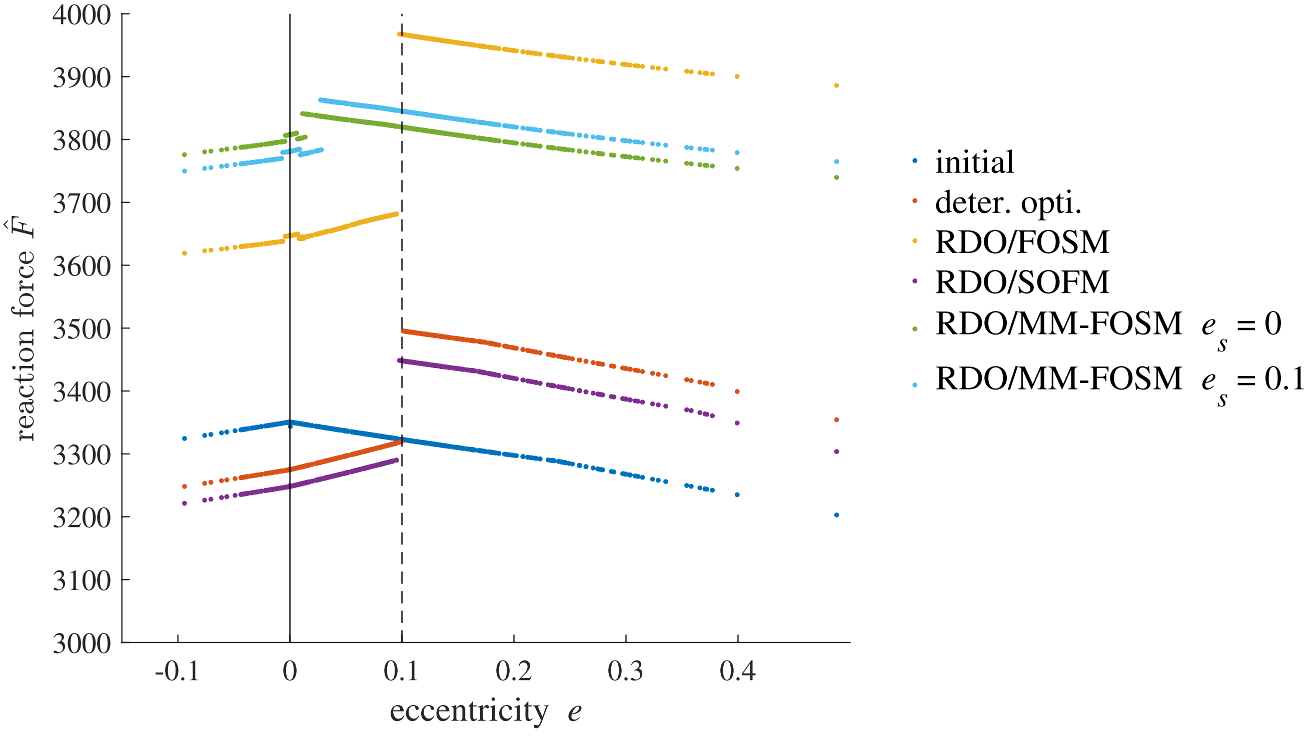


Fig. 15 Reaction force $\hat{F}$ for different eccentricities observed in Monte Carlo simulations of the different designs obtained in different optimizations

Table 4 Results of optimizations of the eccentrically loaded beam example

Optimization approach	Evaluated using									
					MM-FOSM					
	FOSM		SOFM		$e_s = 0$		$e_s = 0.1$		Monte Carlo	
	$\mu_{\hat{F}}$*	$\sigma_{\hat{F}}$	$\mu_{\hat{F}}$	$\sigma_{\hat{F}}$	$\mu_{\hat{F}}$	$\sigma_{\hat{F}}$	$\mu_{\hat{F}}$	$\sigma_{\hat{F}}$	$\mu_{\hat{F}}$	$\sigma_{\hat{F}}$
Initial design	3323	26.54	3322	27.12	3324	21.03	3325	24.97	3322	24.73
Deterministic optimization	**6671**	8798	$-33\cdot10^6$	$68\cdot10^6$	3464	79.43	3368	95.60	3366	91.93
RDO/FOSM	**3967**	**26.77**	3911	134.9	3921	114.4	3778	147.1	3781	144.9
RDO/SOFM	3448	25.82	**3452**	**22.68**	3419	72.31	3332	85.23	3332	82.6
RDO/MM-FOSM $e_s = 0$	3820	26.66	3823	23.83	**3814**	**21.82**	3821	23.57	3811	19.85
RDO/MM-FOSM $e_s = 0.1$	3845	26.77	3846	26.21	3833	33.26	**3846**	**23.57**	3822	32.82

*Equals the deterministic function value

split locations, we perform RDO and evaluate all optimized designs.

The designs obtained from RDO with different e_s in the multimodal approach are similar (see Table 3). The Monte Carlo results in Fig. 15 reveal that these designs lead to discontinuities which are neither at $e = 0$ nor at $e = 0.1$, but in between. Furthermore, the response shows slight discontinuities around $e = 0$, similar to the FOSM design.

Despite the fact, that the eccentricity in either case is assumed at the wrong location, the multimodal approaches outperform the classical FOSM and SOFM approaches. In the vast majority of cases, both multimodal approaches provide better approximations of the reference mean and variance (determined with Monte Carlo) than the classical approaches. This holds especially for the deterministically optimized design, where FOSM and SOFM are far off. Moreover, the RDOs using multimodal FOSM with either split location provide designs with higher mean values and less variance. This shows that the knowledge of the location of the discontinuity and the fact that this location changes during the optimization is not necessarily critical for the application of the approach.

The Monte Carlo simulations of this example are performed with 490 realization. The 90%-confidence intervals I of the estimated mean values and standard deviations are given in Table 5. The number in the I-columns has to be subtracted from the estimated moment to provide the 5% quantile, and adding them yields the 95% quantile of the estimator.

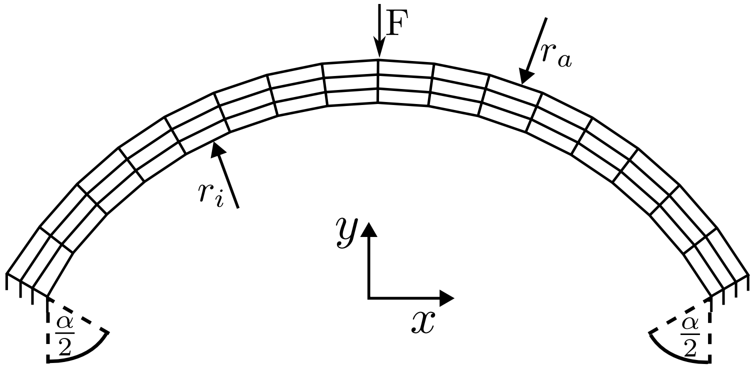
Table 5 Mean values and standard deviations obtained with Monte Carlo simulations together with the confidence intervals of the Monte Carlo estimator for the eccentrically loaded beam example in section 5.1

Design considered	$\mu_{\hat{F}}$	I_μ	$\sigma_{\hat{F}}$	I_σ
Initial design	3322	± 1.8	24.73	± 11.6
Deterministic optimization	3366	± 6.8	91.93	± 18.0
RDO / FOSM	3781	± 10.8	144.9	± 23.8
RDO / SOFM	3332	± 6.2	82.6	± 16.2
RDO / MM-FOSM $e_s = 0$	3811	± 1.5	19.85	± 6.7
RDO / MM-FOSM $e_s = 0.1$	3822	± 2.5	32.82	± 8.7

5.2 Shape optimization of a shallow arc

In a second example, the multimodal FOSM framework is applied to a 3D shape optimization of a shallow arch, which is known to be prone to buckling (Lindgaard and Lund 2011). The model is similar to the topology optimization example given in Krüger and Kriegesmann (2024). All relevant model parameters are summarized in Fig. 16. The shallow arch is discretized by a body-fitted mesh consisting of 320 × 12 × 12 hexahedral finite elements. It is clamped on both sides and loaded by a line load on the center line. The overall geometry is described by the inner radius $r_i = 180$ mm, the outer radius $r_a = 186$ mm, the thickness $t = 6$ mm, and the opening angle $\alpha = 50°$. Linear-elastic material behavior (Saint Venant-Kirchhoff constitutive law) with a Young's modulus of 2 GPa and a Poisson's ratio of 0.3 is assumed while the overall problem is solved using a displacement-driven geometric nonlinear analysis (Wriggers 2008). The displacement is applied in 80 equidistant steps where each step requires 5 to 30 inner iterations (FE solves). For this reason, this rather small example already requires huge computation time. All computations are performed using an in-house tool which is implemented in Matlab. We use fully integrated bi-linear elements. For the nonlinear analysis, a path following algorithm and a Green-Lagrange strain measure are used (Wriggers 2008). Parallelization is only used for computing the element stiffness matrices. The deterministic optimization is executed on a workstation with an AMD Ryzen 7950X with 16 cores while all other computations are performed on a high-performance cluster node with two AMD Epyc 9354 processors.

The shallow arch is a well-known snap-through problem with a load–displacement curve as shown in Fig. 17. Here, the overall reaction force is displayed for different



Inner radius r_i	183 mm
Outer radius r_a	186 mm
thickness t	6 mm
Angle α	50°
number of elements	46′080
Young's modulus	2 GPa
Poisson ratio	0.3
displacement steps u_i	[2 : 0.075 : 8] mm
objective κ	3
correlation length l_c	80 mm
Variance σ^2	0.5 mm

Fig. 16 Model sketch and parameters of the shallow arch

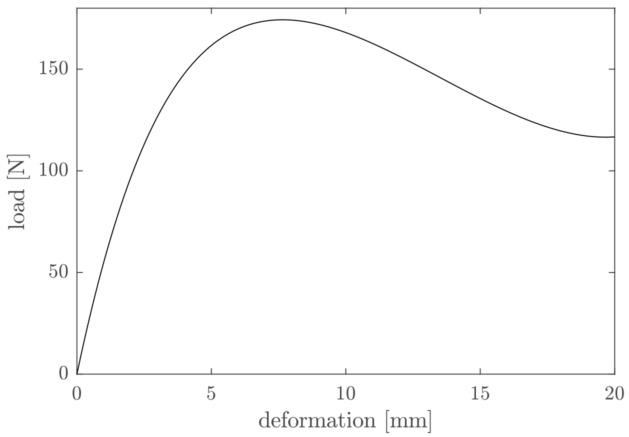


Fig. 17 Load–displacement curve of the initial design of the shallow arch

displacements of the loaded node at $z = 0$. One can clearly see that the load plot has a local maximum at a displacement of 7.5 mm, and afterward, it decreases. A load above the local maximum would lead to large deformations. Therefore, the maximum load is considered in the optimization. Since the maximum function is not differentiable, it is approximated by a P-mean-norm of loads F_i at different displacements u_i. The response function considered in this example is therefore given by

$$F_p = \frac{1}{n}\left(\sum_{i=1}^{n} F(u_i)^p\right)^{\frac{1}{p}} \tag{29}$$

with $p = 16$.

The shape optimization is parameterized using the nodal coordinates of the lower surface and the side surfaces ($z = -3$ mm and $z = +3$ mm) as design variables $\boldsymbol{y}$. Since node variations in tangential direction do not change the geometry, only variations perpendicular to the surface of the initial design are considered as design parameters. In order to maintain mesh quality and circumvent numerical instabilities during the optimization, the interior nodes are updated using Laplacian smoothing as shown in the study of Le et al. (2011), and the design nodes are filtered using a filter similar to the density filter in topology optimization (Sigmund 2007). Here, the nodes on the lower edge are filtered with a filter radius of 3 mm in z-direction and 9 mm in tangential direction. The nodes on the out-of-plane surfaces are filtered independently of the lower nodes with an filter radius of 3 mm in radial and 9 mm in tangential direction. The nodes on the upper edge of the model as well as the clamped nodes are not actively moved. However, a movement into z-direction occurs due to filtering effects and shared nodes with the other surfaces. A movement into the other directions is actively prevented by modifying the filter in the corresponding region. The parameter boundaries are chosen in a way that the edges are allowed to reduce or increase the length up to 60%.

In the context of robust design optimization all nodal coordinates (including interior nodes) are considered to be uncertain. While the design variables move the nodes orthogonal to the surface (i.e., x-, y-, and z- coordinates are not independently moved), it is assumed that imperfections on x-, y-, and z-coordinates are independent of each other. It is assumed that the coordinates are normal distributed with the variance $\sigma^2 = (0.5\ \text{mm})^2$, where the mean imperfections is assumed to be given by the current design configuration. Hence, the geometry scatters around the perfect structure. Correlation between coordinates of the same type is modeled using a square exponential random field with the correlation length $l_c = 80$ mm. To summarize, the design is defined by the outer nodes of the arc, where the top surface is not modified during the optimization. Inner nodes are updated by smoothing. The scatter of the geometry is described as a perturbation of all nodes of the perfect geometry. The position of the i-th node is hence given by

$$\boldsymbol{p}_i = \boldsymbol{y}_i + \boldsymbol{x}_i \tag{30}$$

where $\boldsymbol{p}_i$ is the physical position, $\boldsymbol{y}_i$ is the position of the perfect structure, and $\boldsymbol{x}_i$ is the random perturbation that describes the imperfection. Filter operators and Laplacian smoothing are not applied to the random variables. However, similar effects occur by using the random field also for interior nodes. A realistic interpretation of the parameterization could be a problem where the upper surface is already designed for some other purpose and should therefore not be changed. However, random effects might still lead to an imperfect structure of the upper edge. The overall optimization reads

$$\begin{aligned} \min_{\boldsymbol{y}} \quad & \bar{g}(\boldsymbol{x}, \boldsymbol{y}) \\ \text{s.t.} \quad & \tfrac{V_0}{V(\boldsymbol{x},\boldsymbol{y})} - 1 \leq 0 \\ & \boldsymbol{y}_{\min} \leq \boldsymbol{y} \leq \boldsymbol{y}_{\max} \\ & \boldsymbol{0} = \boldsymbol{R}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}) \end{aligned} \tag{31}$$

deterministic opti.: $\bar{g}(\boldsymbol{y}) = -F_p(\boldsymbol{y})$

RDO: $\bar{g}(\boldsymbol{y}) = -\mu_{F_p}(\boldsymbol{y}) + 3\,\sigma_{F_p}(\boldsymbol{y})$

with the objective function $\bar{g}$, the initial volume V_0, the parameter boundaries $\boldsymbol{y}_{\min}, \boldsymbol{y}_{\max}$, and the nonlinear residual $\boldsymbol{R}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u})$. In the deterministic case, the objective function equals the negative response function $\bar{g} = -F_p$. In the robust optimization the mean and standard deviation of the response function μ_{F_p} and σ_{F_p} are considered. The optimization is executed using the globally convergent method of moving asymptotes (GCMMA) of Svanberg (2007) with the initial move limit $s_{ini} = 0.05$. The stopping criterion is given by a relative change in the objective function of less than 0.01% per iteration. The stochastic quantities are computed using the classic first-order second-moment method as well as the multimodal approach.

For the robust design optimization using the multimodal approach, the distribution of the random variables has to be split. For a multivariate distribution, this split is most easily achieved when the random parameters are independent and standard Gauss distributed, as in the example in Sect. 4. Therefore, the random field is discretized using the EOLE method [similar to the Karhunen-Loève expansion (Sudret and Der Kiureghian 2000)]. By using so, the random parameters are the weights/amplitudes of the eigenvectors of the covariance matrix, i.e., of the eigenmodes of the random field. These eigenmodes are of similar shape as the buckling modes. Each random imperfection modes triggers a certain buckling response, which motivates splitting the reduced random space of amplitudes along their semi-axes. Storing the whole covariance matrix would require approximately 200 GB memory. In order to reduce the memory requirements, only the first 15 EOLE modes (equivalent to eigenvectors of covariance matrix) are considered. By doing this, the memory demand is reduced to 20 MB while the error of the variance is less than 0.4%. Afterward, the split along the bisectors is done for all considered modes.

The optimized designs are shown in Fig. 18, and the numeric results are shown in Table 6. All optimized designs are evaluated using the Monte Carlo method with a sample size of 1500. The relatively low sample size is chosen because the simulation time per sample is approximately 1 h. Comparing the full sample size with the half sample size reveals that the error is less than 1%.

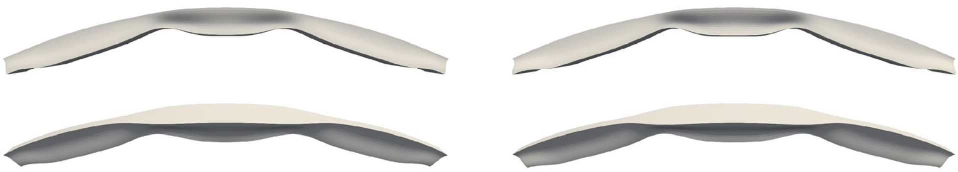
The deterministically optimized design can be interpreted as three bars which are connected by elastic joints. The joints only have a small bending stiffness in both bending directions and can therefore be interpreted as rotational bearings. Using this structure, the outer bars mainly carry compression loads while the middle bar is loaded by a bending moment. In consequence, the middle part is comparable to an I-shaped section. The side bars have a high thickness at the lower edge and a minimal thickness at the top. By doing this, the effective load angle is optimized although the clamped nodes are not allowed to move.

From engineering knowledge, it can be seen that this structure is not robust. As long as the problem is perfectly symmetric, the joints only have to transfer forces. In the presence of non-symmetric imperfections, the joints are exposed to moments and hence, the structure will deform severely (buckling). This can also be seen in the numeric results. The deterministically optimized design reaches good results in the ideal case, but in the presence of imperfections the mean value is 65% lower than the nominal function value.

A robust design optimization using the classic first-order second-moment method leads to a similar result as the deterministic optimization. A reason can be found in the large error of the FOSM approach at the deterministic optimum. Using a first-order Taylor series at the perfect geometry, buckling cannot be detected. Therefore, the mean is completely overestimated by 200%. In addition, the standard deviation is significantly underestimated.

In difference, the multimodal approach estimates much more pessimistic stochastic quantities. Compared to the Monte Carlo results, the mean is still significantly

Fig. 18 Optimized designs of the shallow arch. The results are shown from the front (top figure) and slightly tilted around the x-axis to show the 3D structure



(a) Deterministic optimized design.

(b) Optimized design using FOSM

(c) Optimized design using multimodal FOSM.

Table 6 Numeric results of the optimized shallow arch. The Monte Carlo simulations are done with a sample size of 1500

Optimization approach	Evaluated using								
	FOSM		MM-FOSM		Monte Carlo		No. iterations		Time (h)
	μ_g*	σ_g	μ_g	σ_g	μ_g	σ_g	Outer	All	
Determ. optimization	**395 N**	19 N	288 N	94 N	140 N	30 N	26	49	10
RDO/FOSM	**415 N**	**20 N**	290 N	98 N	141 N	30 N	34	73	48
RDO/MM-FOSM	169 N	6.8 N	**167 N**	**7.2 N**	164 N	6.2 N	32	32	20

*Equals the deterministic function value

overestimated by 100% while the standard deviation is overestimated by 200%. One reason for this deviation is that the relation of random parameters and response function is nonlinear. Therefore, the FOSM method only provides an approximation, no matter whether it is FOSM or MM-FOSM. Furthermore, the subdivision of the random space is chosen along the bisectors as shown in Figs. 9 and 10, but the true evolution of the discontinuity may differ. In fact, in-depth investigations implied that a split along the coordinate axes would have led to more accurate results. Despite this inaccuracy, using MM-FOSM provides a robust design. The reason is that the locations at which the difference Taylor series are expanded represent multiple different imperfection modes, which may trigger different buckling modes. The optimization therefore is able to counteract these different buckling modes. In terms of the resulting design, this is achieved by using joints that are able to transfer significant bending moments in combination with a bigger middle part. In difference to the buckling-sensitive designs, no I-shaped section is used. This might be caused by the fact that this section type is stiff against bending around the z-axis, but weak against bending with respect to the y-axis and torsion.

The superior performance of the multimodal optimized designs is further explained by the load–displacement curves shown in Fig. 19. Here, the left figure shows the result of the deterministically optimized design which is similar to the response of the design optimized using FOSM. The right figure shows the response of the structure which was optimized using multimodal FOSM. In absence of imperfections, the deterministically optimized structure has a perfect parabolic behavior which indicates snap-through behavior, but no bifurcation-type buckling. In difference to that, the design which was optimized using multimodal FOSM has a load–displacement curve which is typical for bifurcation buckling, even in the absence of imperfections. In consequence, imperfections do not change the overall response and the structure is much less imperfection sensitive.

From the previous sections, it is known that the multimodal approach needs more computational resources. Using the split along the bisectors, $2N$ FOSM evaluations at the cost of 2 deterministic evaluations have to be executed per iteration. Here, N represents the number of splits. Using 15 splits, the multimodal approach needs 60 deterministic evaluations per iteration. In Table 6, the number of iterations as well as the computation time are presented. Here, a distinction between outer iterations of the GCMMA and all iterations (inner and outer) is made. The deterministic optimization used 16 cores on a workstation and the optimization using FOSM used 4 cores on a high-performance cluster. The multimodal approach is parallelized to 15 workers with 4 cores each leading to overall 60 CPU cores. All optimization strategies needed approximately the same number of iterations. However, it turns out that the multimodal approach does not need any inner iterations because the

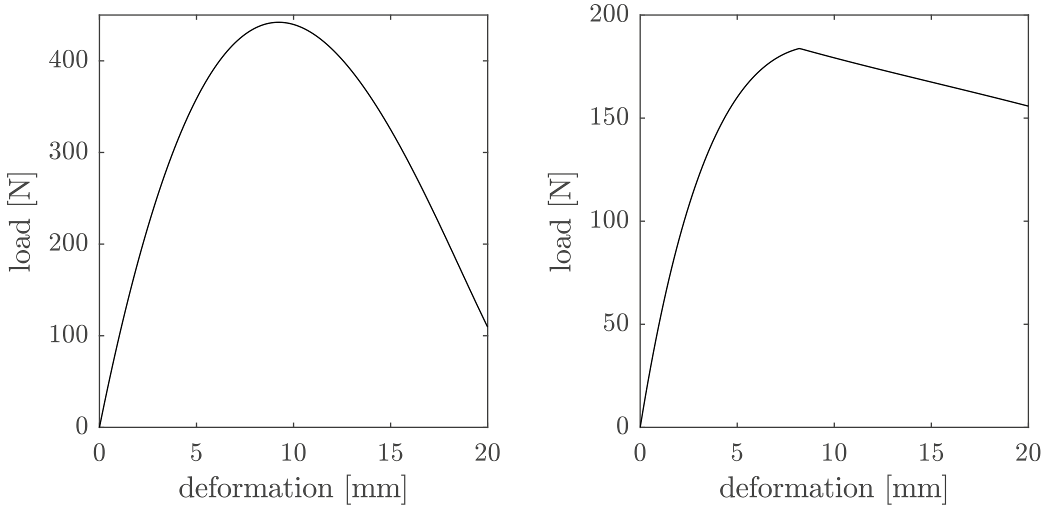


Fig. 19 Load–displacement curve of the final designs. Left: Deterministic design, which is similar to the design optimized with FOSM. Right: Design optimized with multimodal FOSM

multimodal optimization problem is smoother. This is due to the fact that no jumping between snap-through and bifurcation buckling occurs. Comparing the worker time (time × number workers) per iteration reveals two things: FOSM is three times more expensive than the deterministic optimization [where in theory a factor of 2 is expected (Krüger et al. (2023). The difference to theory is explained by the different hardware]. The multimodal approach is 50 times more expensive than a deterministic optimization, which is near the expected value of 60. The lower value is explained by the fact that the structural response is smoother in the multimodal optimization than in the deterministic optimization.

5.3 Topology optimization with contact

In the following, the multimodal FOSM framework is applied to a modified topology optimization benchmark for structures in unilateral contact of Strömberg and Klarbring (2010). Therein, the unilateral contact is modeled using Signorini's contact condition (Signorini 1959). In doing so, linear-elastic material behavior is assumed, while the unilateral contact constraints cause the underlying equilibrium to be both non-smooth and nonlinear. The equilibrium equations as well as the compliance optimization problem are derived in detail in Strömberg and Klarbring (2010). The latter also shows the necessity of using unilateral contact constraints in optimization, since the unilateral characteristic of contact has a significant influence on the optimized design. In fact, using the widely used bilateral penalty approaches lead to poor performing optimized designs, as the optimization will maximally exploit this non-physical model assumption.

In Schmidt et al. (2024), a robust topology optimization framework using FOSM is proposed to consider manufacturing tolerances at the surface of the contact support. It is shown that the non-smooth and nonlinear equilibrium does only depend linearly on the scattering surface of the contact support. Thus, higher-order methods such as SOFM cannot increase the accuracy, since all second-order derivatives with respect to the random variables are zero. This holds also true for the here discussed scattering of the radius of the contact structure. Additionally, FOSM becomes more and more inaccurate, if larger scatterings of the contact surface are considered, since larger scattering leads to an increasing number of switches of the contact conditions of the nodes at the contact support. More precisely, some nodes which are close to the contact support will come into contact, while others might lose contact. The on/off switching of the contact conditions cause discontinuous changes in the equilibrium equations. These changes cannot be captured by standard Taylor series-based methods approximating the mean and the standard deviation for robust optimization. Here, the proposed multimodal FOSM framework allows for considering discontinuous changes of the equilibrium equation, so that robust optimization can also be applied for larger scatterings.

The design space as well as the parameters of the benchmark example are shown in Fig. 20 and four different loads are considered. Each load case is analyzed separately, but considered as one response for the optimization through Eq. (32). Each load is applied at the continuous contact structure, moving the contact structure until it is in contact with the linear-elastic design domain. In this example, the radius of the contact structure r is considered as a normal distributed random variable with a mean of $\mu_r = 50$ mm and a standard deviation of $\sigma_r = 0.025$ mm. Thus, there is only one random variable, influencing all nodes of the linear-elastic domain at the contact surface. For the deterministic optimization, the objective is the average compliance $\tilde{c}$ of the four load cases

$$\bar{g} = \underbrace{\sum_{i=1}^{4} \frac{c(\boldsymbol{q}_{\mathrm{cj}}^{i})}{4}}_{\tilde{c}}, \tag{32}$$

while the robust optimization aims at the weighted sum of the average mean value $\mu_{\tilde{c}}$ and the average standard deviation $\sigma_{\tilde{c}}$ based on the four load cases

$$\bar{g} = \underbrace{\frac{1}{4}\sum_{i=1}^{4} \mu_{c(\boldsymbol{q}_{\mathrm{cj}}^{i})}}_{\mu_{\tilde{c}}} + 3 \underbrace{\frac{1}{4}\sum_{i=1}^{4} \sigma_{c(\boldsymbol{q}_{\mathrm{cj}}^{i})}}_{\sigma_{\tilde{c}}}. \tag{33}$$

In doing so, the resulting optimization problem reads

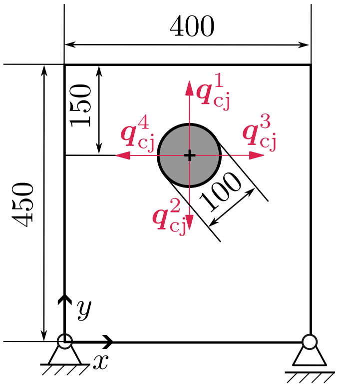
$$\begin{aligned} \min_{\boldsymbol{\rho}} \quad & \bar{g}(\boldsymbol{\rho}) \\ \text{s.t.} \quad & \boldsymbol{R}_{\mathrm{con}}(\boldsymbol{u}, \boldsymbol{y}, r) = \boldsymbol{0} \\ & \frac{V(\boldsymbol{y})}{V_0} - 1 \leq 0 \\ & \boldsymbol{\rho}_{\min} \leq \boldsymbol{\rho} \leq \boldsymbol{1} \end{aligned} \tag{34}$$

where $\boldsymbol{R}_{\mathrm{con}}$ is the nonlinear and non-smooth residual of the contact problem (Strömberg and Klarbring 2010). The design vector $\boldsymbol{\rho}$ represents the pseudodensities within a standard SIMP framework with density filter (Bendsøe and Sigmund 2004). The optimization problem is solved using the MMA (Svanberg 2007). The derivatives required to approximate the mean and the standard deviation using FOSM are derived in Schmidt et al. (2024) for a similar application. More precisely, the considered scattering of the radius is a subcase of the derived FOSM approach of Schmidt et al. (2024). One deterministic optimization step requires four nonlinear analyses of the four load cases, and four solutions of an adjoint system for determining the gradient. Performing RDO with FOSM requires two additional solutions of adjoint systems per load case. Since only one random parameter is considered and the random space is split into two domains, the cost of two FOSM analyses are required for one iteration using MM-FOSM. In order to apply MM-FOSM, the weightings w_i and the evaluation points of the mixture distribution of Eq. (37) must be set. The discontinuity is assumed to be at the nominal value of the radius, which equals the mean value. Consequently, the weighting factors equal $w_i = 0.5$ and the radius is set to $50 \pm 0.025\sqrt{\frac{2}{\pi}}$ for the evaluation points μ_i.

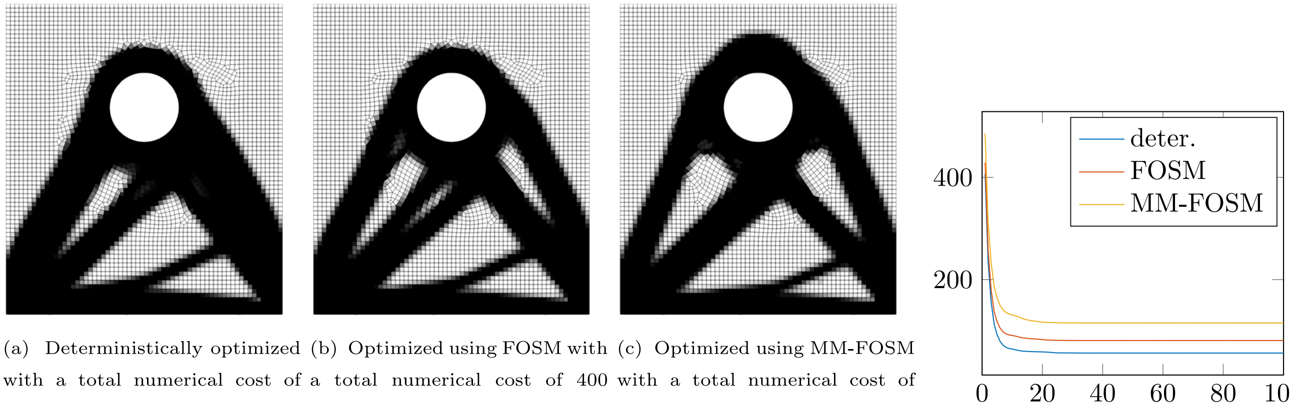
The optimized designs as well as the convergence history of the deterministic and the robust design optimizations are visualized in Fig.21. All optimizations converge as expected, and major differences in the final designs are observed. Especially, the differences between the final designs obtained using FOSM and MM-FOSM can be explained with the discontinuous characteristic of contact, since both methods will result in the same designs, if the considered problem has no discontinuity. This conclusion is confirmed by the results shown in Table 7.

In Table 7, the average mean and the average standard deviation of the initial design and the optimized designs are computed using FOSM, MM-FOSM and 1000 Monte Carlo samples. The FOSM approximations of both the average mean and the average standard deviation are inaccurate. In consequence, the optimized design performs poorly. In fact, the Monte Carlo simulations reveal that the optimized design based on FOSM has a higher average mean and a higher average standard deviation than the deterministically

Fig. 20 Design space of the contact problem



number of elements	4,444
Young's modulus	210,000 MPa
Poisson ratio	0.3
vol. fraction	0.5
filter radius	10 mm
$\lVert \boldsymbol{q}_{\mathrm{cj}}^{i} \rVert$	1,000 N
objective κ	3
mean radius μ_r	50 mm
variance radius σ_r^2	0.025^2 mm^2



(a) Deterministically optimized with a total numerical cost of 400 FEA evaluation and 400 adjoint FEA evaluations.

(b) Optimized using FOSM with a total numerical cost of 400 FEA evaluation and 800 adjoint FEA evaluations.

(c) Optimized using MM-FOSM with a total numerical cost of 800 FEA evaluation and 1,600 adjoint FEA evaluations.

(d) Convergence history.

Fig. 21 Design space and optimized designs for the contact example for $\kappa = 3$

optimized design. In contrast, the MM-FOSM approximations of the average mean are very accurate for all analyzed designs, while the average standard deviation is less accurate but still an improvement in comparison with FOSM. In consequence, the optimized design using MM-FOSM has a significant lower standard deviation than the deterministic optimization, proving that the robust optimization works as expected, if MM-FOSM is used. The results for each sample of the Monte Carlo simulations are visualized in Fig. 22, revealing that the contact leads to a C^1- discontinuity for the considered problem. Additionally, the mean values and the standard deviations of each load case are shown in Table 8.

In Table 9, the stochastic moments estimated with the Monte Carlo method are given per load case together with the confidence interval I of the estimator. Here, 1000 realizations were used. The confidence interval of the standard deviations is relatively high, which is due to the large kurtosis of obtained in the example. A convergence analysis of the estimators, however, showed a much higher accuracy than the confidence intervals might indicate.

5.4 Sizing optimization of a stiffened panel

In the following section, the multimodal FOSM framework is applied to a curved and stiffened panel commonly used in aircraft structures. The example was introduced in the context of robust design optimization in Steltner et al. (2022).

The design variables for the panel structure are the shell element thicknesses $\boldsymbol{t}$. The objective of the optimization is to maximize the sum of forces $F_S = \sum_{i=1}^{15} F_i(\boldsymbol{t})$ at predetermined displacements along the load–displacement curve without adding volume to the structure. The design optimization problem reads

$$\begin{aligned} \min_{\boldsymbol{t}} \quad & \bar{g}(\boldsymbol{t}) \\ \text{s.t.} \quad & V(\boldsymbol{t}) = V_0 \\ & \boldsymbol{R}(\boldsymbol{t}) = \boldsymbol{0} \\ & 0.2\, t_{i,0} \leq t_i \leq 2\, t_{i,0} \quad i = 1, .., n_e \end{aligned} \tag{35}$$

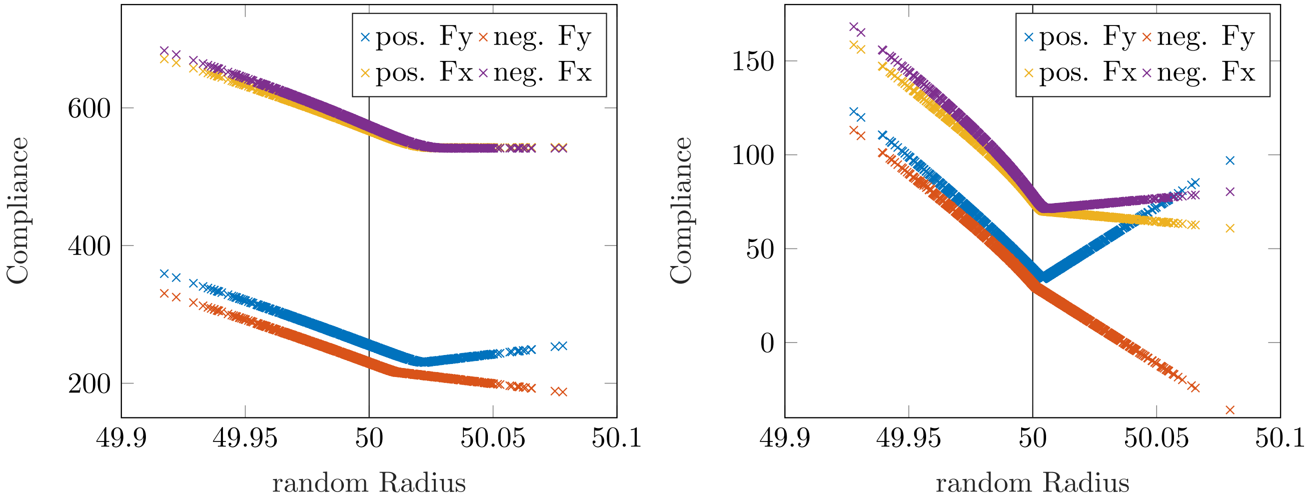
$$\text{deterministic opti.: } \bar{g}(\boldsymbol{t}) = -F_S(\boldsymbol{t})$$

$$\text{RDO: } \bar{g}(\boldsymbol{t}) = -\mu_{F_S}(\boldsymbol{t}) + \sigma_{F_S}(\boldsymbol{t})$$

The design variables are bound with respect to their initial values $t_{i,0}$. The optimization problem is solved using the

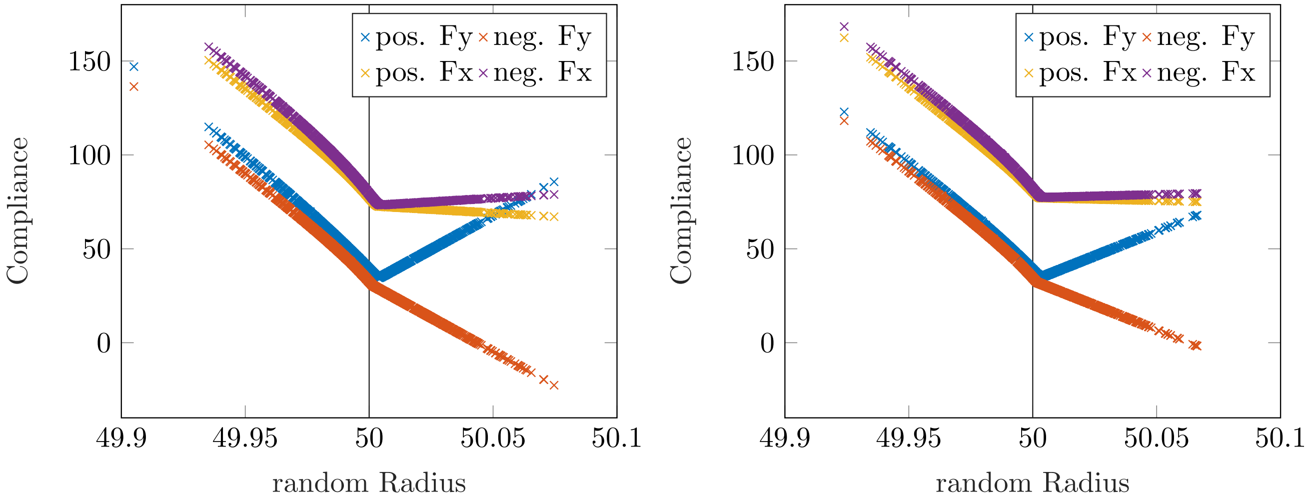
Table 7 Results of optimizations of the contact example

Optimization approach	Evaluated using						
	Deter.	FOSM		MM-FOSM		Monte Carlo	
	$\tilde{c}$	$\mu_{\tilde{c}}$	$\sigma_{\tilde{c}}$	$\mu_{\tilde{c}}$	$\sigma_{\tilde{c}}$	$\mu_{\tilde{c}}$	$\sigma_{\tilde{c}}$
Initial design	406.71	406.71	7.50	408.80	25.65	410.88	27.11
Deterministic optimization	56.13	56.13	8.38	66.94	17.32	66.42	20.72
RDO / FOSM	57.13	57.13	7.85	67.66	16.35	68.44	21.02
RDO / MM-FOSM	59.10	59.10	7.56	69.52	15.15	69.17	18.51



(a) Initial gray design.

(b) Final design using deter. Optimization.



(c) Final design using FOSM.

(d) Final design using MM-FOSM.

Fig. 22 Monte Carlo Sampling of the designs of Fig. 21

Table 8 Results of robust design optimization for the contact example of Fig. 21

Load case	Optimization approach	Deter. compl. $c(\boldsymbol{q}^i_{\mathrm{cj}})$	Opti. Appr.		Monte Carlo	
			$\mu_{c(\boldsymbol{q}^i_{\mathrm{cj}})}$	$\sigma_{c(\boldsymbol{q}^i_{\mathrm{cj}})}$	$\mu_{c(\boldsymbol{q}^i_{\mathrm{cj}})}$	$\sigma_{c(\boldsymbol{q}^i_{\mathrm{cj}})}$
Pos. Fy ($\boldsymbol{q}^1_{\mathrm{cj}}$)	Deter	38.92	-	-	55.15	16.50
	FOSM	39.28	39.28	7.30	56.56	17.34
	MM-FOSM	39.91	53.40	10.87	53.14	15.70
Neg. Fy ($\boldsymbol{q}^2_{\mathrm{cj}}$)	deter	31.28	-	-	35.28	25.09
	FOSM	32.55	32.55	7.37	37.63	25.59
	MM-FOSM	34.63	40.96	18.95	40.55	21.91
Pos. Fx ($\boldsymbol{q}^3_{\mathrm{cj}}$)	deter	75.85	-	-	85.13	20.31
	FOSM	77.37	77.37	8.27	87.79	20.10
	MM-FOSM	80.64	90.49	14.68	90.20	17.53
Neg. Fx ($\boldsymbol{q}^4_{\mathrm{cj}}$)	deter	78.45	-	-	90.12	20.99
	FOSM	79.32	79.32	8.47	91.78	21.05
	MM-FOSM	82.23	93.24	16.08	92.79	18.88

Table 9 Mean values and standard deviations obtained with Monte Carlo simulations together with the confidence intervals of the Monte Carlo estimator for the example with contact in Sect. 5.3

Design considered	LC1				LC2				LC3				LC4			
	μ_c	I_μ	σ_c	I_σ	μ_c	I_μ	σ_c	I_σ	μ_c	I_μ	σ_c	I_σ	μ_c	I_μ	σ_c	I_σ
Determ. optimization	55.1	± 0.86	16.5	± 4.7	35.3	± 1.31	25.1	± 6.4	85.1	± 1.06	20.3	± 5.4	90.1	± 1.09	21.0	± 5.8
RDO/FOSM	56.6	± 0.90	17.3	± 5.2	37.6	± 1.33	25.6	± 6.6	87.8	± 1.05	20.1	± 5.6	91.8	± 1.10	21.1	± 6.0
RDO/MM-FOSM	53.1	± 0.82	15.7	± 4.8	40.5	± 1.14	21.9	± 5.7	90.2	± 0.91	17.5	± 5.1	92.8	± 0.98	18.9	± 5.4

MMA as implemented in Tosca (https://www.3ds.com/products/simulia/tosca).

Geometric imperfections of the panel skin are considered as random parameters. For that, the geometric imperfections given in Zimmermann et al. (2006) are used as data basis. The random field of the imperfections is parameterized by a double Fourier series, which are mapped to the nodal position of the finite element mesh. The Fourier coefficients that parameterize the geometric imperfections are strongly correlated and therefore transformed to a set of uncorrelated parameters. For that, the coefficients are assembled in the random vector $\boldsymbol{X}$ with the mean vector $\boldsymbol{\mu}_{\boldsymbol{X}}$ and the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{X}}$, determined from the measurement data given in Zimmermann et al. (2006). The procedure is described in more detail in Kriegesmann et al. (2012).

As in the shallow arc example, a vector $\boldsymbol{Z}$ with uncorrelated parameters, zero mean and the identity matrix as covariance matrix is obtained from the transformation (24). The data from Zimmermann et al. (2006) consist of eight measurements of panels. Since the rank of the covariance matrix is always smaller than the number of measurements, only seven non-zero eigenvalues are found. Hence, Eq. (24) reduces the number of random parameters to seven. For the deterministic optimization, the mean imperfection, i.e., $\boldsymbol{z} = \boldsymbol{0}$, is imposed on the panel.

Figure 23 shows the deterministically optimized panel obtained by Steltner et al. (2022), where the thicknesses in the shell elements are considered as design parameters. This formulation turned out to implicitly increase the global buckling load and improve the postbuckling behavior. The discussion of the results shall not be repeated here and the interested reader is referred to Steltner et al. (2022). The example is chosen for this paper since the unimodal

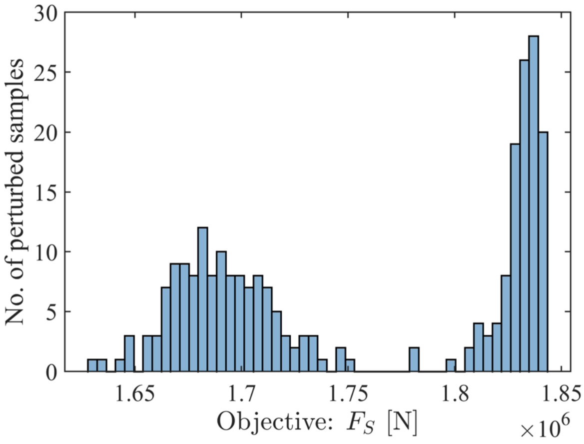


Fig. 24 Histogram of the optimization objective for the panel optimization conducted in Steltner et al. (2022)

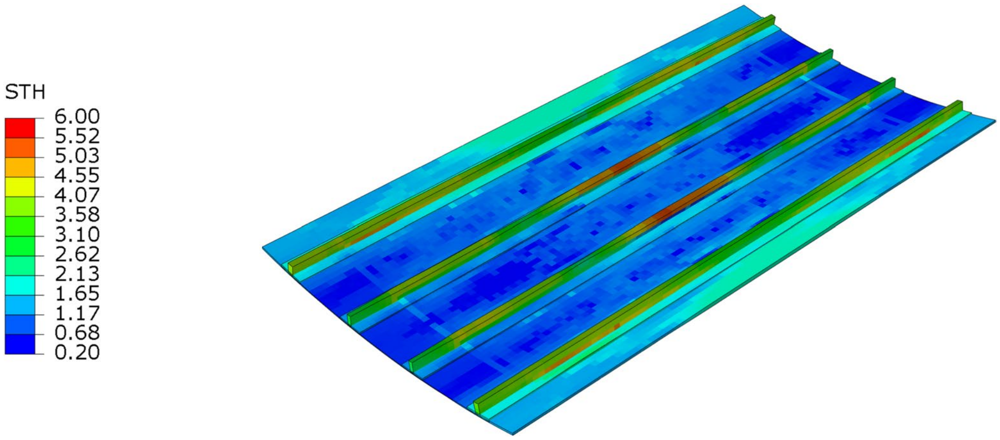


Fig. 23 Stiffened panel showing deterministically optimized shell thickness per element

Table 10 Numeric results of optimized stiffened panel

Optimization approach	Evaluated using					
	FOSM		MM-FOSM		Monte Carlo	
	μ_g*	σ_g	μ_g	σ_g	μ_g	σ_g
Determ. optimization	**1.81**	0.80	1.78	0.29	1.76	0.07
RDO / FOSM	**1.71**	**0.02**	1.72	0.12	1.72	0.05
RDO / MM-FOSM	1.71	0.01	**1.72**	**0.06**	1.71	0.01

*Equals the deterministic function value

multivariate distribution that describes the scatter of geometric imperfections leads to a bimodal distribution of the global buckling load (see Fig. 1) and the objective function considered in the optimization (sum of reaction forces) shown in Fig. 24. The robust design obtained in Steltner et al. (2022) does not show a multimodal distribution.

Following the reduction of the random parameters, all seven modes are considered for analyses and optimization using the multimodal FOSM method. The input distribution is split along the bisectors as described in Sect. 4. The mean and variance for each of the resulting 14 subdomains are given in Table 11.

Using the semi-intrusive approach published in Steltner et al. (2022), one FOSM analysis requires eight nonlinear finite element analyses for determining the mean and the variance of the response function and their gradients. Using the multimodal FOSM approach suggested in this paper, and splitting the random space into 14 subdomains, 14 FOSM analyses have to be run in each optimization iteration, totaling 112 analyses, which can be parallelized. The computational time required for one nonlinear finite element analysis using Abaqus (https://www.3ds.com/products/simulia/abaqus) varies along the optimization and is around 30 to 45 min.

Table 10 summarizes the results of deterministic and robust optimizations as well as probabilistic analyses conducted with FOSM, multimodal FOSM, and Monte Carlo simulations using 250 samples. Note that Steltner et al. (2022) used a very large finite difference step size of $\Delta z = 1.5$, which implicitly coarsely covered potential discontinuities. For a clean comparison, the same finite difference step size of $\Delta z = 0.1$ was used here for both FOSM and MM-FOSM. The stochastic moments, especially the standard deviation, of all designs are overestimated when using MM-FOSM. Similar to the shallow arc example, the relation of design response and random parameters is nonlinear, and the true evolution of the discontinuity is unknown. These effects lead to deviations of both FOSM methods and the Monte Carlo results, while being less severe in this example. Here it seems as if the plain FOSM method is more accurate than MM-FOSM, which, however, may occur coincidentally whenever a method is inaccurate. More importantly, the RDO using MM-FOSM provides a more robust design according to the Monte Carlo results. As for the shallow arc, this is due to the fact that for MM-FOSM, the response function is approximated at multiple locations in the random space, which represent multiple different imperfection modes, and therefore multiple buckling modes are captured during the optimization.

6 Conclusion

The multimodal first-order second-moment method introduced in this paper allows to predict the mean value and standard deviation of objective functions with discontinuities, which result in a multimodal stochastic distribution of the objective. The computational cost is higher than for approximations at the mean vectors, but still feasible to be embedded into a robust design optimization framework. The key challenge of the approach is that the location (hypersurface) of the discontinuity must be known, which is often not the case. There are, however, numerous cases such as buckling of structures or contact problems where the discontinuities can be anticipated, and hence, the method is applicable. An automated search of discontinuities would broaden the applicability of the new method. This comes most likely at the expense of significant computational cost and is subject of future research.

Appendix

A1 Mixture distribution

Another way to look at the proposed approach is to consider the distribution of the objective function as a mixture distribution. Therefore, a short introduction of mixture distributions is given based on Frühwirth-Schnatter (2006). A mixture distribution f_X, which is obtained by a weighted sum of N distributions $f_1, ..., f_N$, is given by

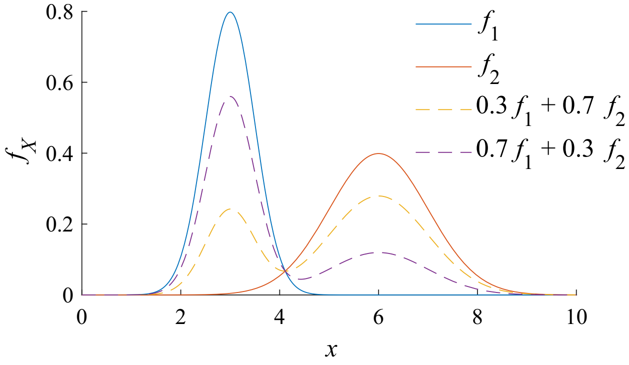


Fig. 25 Example of two probability density functions obtained from mixing two unimodal probability density functions with different weight factors

$$f_X(x) = \sum_{i=1}^{N} \omega_i f_i(x)$$
$$\text{with} \quad \omega_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{N} \omega_i = 1 \tag{36}$$

Examples of bimodal distributions obtained from mixing two unimodal distributions are shown in Fig. 25.

The moments of the mixture distribution can be determined from the moments of the mixture components as well. Here, we focus on the mean and variance, which are given by

$$\mu_X = \sum_{i=1}^{N} \omega_i \, \mu_i$$
$$\sigma_X^2 = \sum_{i=1}^{N} \omega_i \left(\mu_i^2 + \sigma_i^2\right) - \mu_X^2 \tag{37}$$

The derivatives with respect to a variable α are given by

$$\frac{d\mu_X}{d\alpha} = \sum_{i=1}^{N} \left(\frac{\partial \omega_i}{\partial \alpha} \mu_i + \omega_i \frac{\partial \mu}{\partial \alpha}{}_i \right) \tag{38}$$

and

$$\frac{d\sigma_X^2}{d\alpha} = \sum_{i=1}^{N} \frac{\partial \omega_i}{\partial \alpha} \left(\mu_i^2 + \sigma_i^2\right) + \sum_{i=1}^{N} \omega_i \left(2\mu_i \frac{\partial \mu_i}{\partial \alpha} + \frac{\partial \sigma_i^2}{\partial \alpha} \right) - 2\mu_X \frac{\partial \mu_X}{\partial \alpha} \tag{39}$$

A2 Taylor expansion with offset from mean

This section gives the governing equations of the FOSM method for the case that the Taylor series is not expanded at the mean, but at some offset location $\boldsymbol{\mu} + \boldsymbol{d}$. The approach is therefore referred to offset FOSM in this paper. Note that this section is independent of the consideration of multimodal distributions, but can be used in conjunction with a multimodal approach.

Inserting the Taylor expansion at $\boldsymbol{\mu} + \boldsymbol{d}$ given by

$$g(\boldsymbol{x}) = g(\boldsymbol{\mu} + \boldsymbol{d}) + \sum_{i=1}^{n} \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d})}{\partial x_i} \left(x_i - \mu_i - d_i\right) + \dots \tag{40}$$

into the governing equations for the mean and the variance (3) yields

$$\mu_g \approx g(\boldsymbol{\mu} + \boldsymbol{d}) - \sum_{i=1}^{n} \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d})}{\partial x_i} d_i \tag{41}$$

and

$$\sigma_g^2 \approx g^2(\boldsymbol{\mu} + \boldsymbol{d}) - \mu_g^2 - 2g(\boldsymbol{\mu} + \boldsymbol{d}) \sum_{i=1}^{n} \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d})}{\partial x_i} d_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d})}{\partial x_i} \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d})}{\partial x_j} \left[\operatorname{Cov}\left(X_i, X_j\right) + d_i d_j\right] \tag{42}$$

Inserting the mean allows to further simplify the variance to

$$\sigma_g^2 \approx \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial g\left(\boldsymbol{\mu}_X + \boldsymbol{d}\right)}{\partial x_i} \frac{\partial g\left(\boldsymbol{\mu}_X + \boldsymbol{d}\right)}{\partial x_j} \left[\operatorname{Cov}\left(X_i, X_j\right) + d_i d_j\right] - \left[\sum_{i=1}^{n} \frac{\partial g\left(\boldsymbol{\mu}_X + \boldsymbol{d}\right)}{\partial x_i} d_i \right]^2 \tag{43}$$

The derivatives of these approximations are given by

$$\frac{\partial \mu_g(\boldsymbol{y})}{\partial y_k} \approx \frac{\partial g(\boldsymbol{\mu} + \boldsymbol{d}, \boldsymbol{y})}{\partial y_k} - \sum_{i=1}^{n} \frac{\partial g^2(\boldsymbol{\mu} + \boldsymbol{d}, \boldsymbol{y})}{\partial x_i \partial y_k} d_i$$
$$\frac{\partial \sigma_g^2(\boldsymbol{y})}{\partial y_k} \approx 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 g\left(\boldsymbol{\mu}_X + \boldsymbol{d}\right)}{\partial x_i \partial y_k} \frac{\partial g\left(\boldsymbol{\mu}_X + \boldsymbol{d}\right)}{\partial x_j} \operatorname{Cov}\left(X_i, X_j\right) \tag{44}$$

It is worth noting that these approximations do not require more function evaluations or derivatives than the approach given in Sect. 2.1. The function is simply evaluated at a different location. Hence, the computational cost is the same.

A3 Stochastic moments of subdomains for splits along bisectors, centered at zero

In this subsection, the stochastic moments of subdomains given by bisectors are determined for a multivariate standard Gaussian distribution (zero mean vector, covariance matrix equals identity matrix). Due to independence of the random parameters, the probability density function of the random vector $\boldsymbol{Z}$ can be written as $\varphi(z) = \varphi\left(z_1\right) \cdot \ldots \cdot \varphi\left(z_N\right)$

We now consider the mean value of Z_1^{1+} for the case that the subdomain as illustrated in Fig. 9 is pointing in direction of z_1. The mean value is hence given by

$$\begin{aligned} E\left(Z_1^{1+}\right) = \mu_1^{1+} &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} z_1\, \xi_1^+\, \varphi\left(z_1\right) \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= 2N \int_0^\infty z_1\, \varphi\left(z_1\right) \left\{ \left[\Phi\left(z_2\right)\right]_{-z_1}^{z_1} \cdot \ldots \cdot \left[\Phi\left(z_N\right)\right]_{-z_1}^{z_1} \right\} dz_1 \\ &= 2N \int_0^\infty z_1\, \varphi\left(z_1\right) \left\{ \Phi\left(z_1\right) - \Phi\left(-z_1\right) \right\}^{N-1} dz_1 \end{aligned} \tag{45}$$

With $\Phi(-z) = 1 - \Phi(z)$, we get

$$\mu_1^{1+} = 2N \int_0^\infty z_1\, \varphi\left(z_1\right) \left\{ 2\, \Phi\left(z_1\right) - 1 \right\}^{N-1} dz_1 \tag{46}$$

This one-dimensional integral can quickly be solved with numeric integration. For various dimensions N, the mean values are summarized in Table 11. When considering the subdomain with points in negative z_1-direction, the mean value equals the negative value. This is not only obvious from symmetry consideration, but also follows from

$$\begin{aligned} E\left(Z_1^{1-}\right) = \mu_1^{1-} &= \int_{-\infty}^0 \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} z_1\, \xi_1^+\, \varphi\left(z_1\right) \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= -\int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} z_1\, \xi_1^+\, \varphi\left(z_1\right) \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= -E\left(Z_1^{1+}\right) = -\mu_1^{1+} \end{aligned} \tag{47}$$

Also, from symmetry, we can expect the mean value to equal zero for all other directions, which can be proven (exemplarily for Z_2) by

$$\begin{aligned} E\left(Z_2^{1+}\right) &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} z_2\, \xi_1^+\, \varphi\left(z_1\right) \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= 2N \int_0^\infty \varphi\left(z_1\right) \underbrace{\int_{-z_1}^{z_1} z_2\, \varphi\left(z_2\right) dz_2}_{=0} \prod_{k=3}^{N} \left[\Phi\left(z_k\right)\right]_{-z_1}^{z_1} dz_1 = 0 \end{aligned} \tag{48}$$

where the term $\int_{-z_1}^{z_1} z_2\, \varphi\left(z_2\right) dz_2$ equals zero due to symmetry of φ.

Table 11 Stochastic moments of the subdomain in p-direction for different dimensions N

N	μ_p^{p+}	σ_p^p	$\sigma_q^p (q \neq p)$
1	0.79788	0.60281	–
2	1.1284	0.60281	0.60281
3	1.3264	0.58597	0.66983
4	1.4647	0.5699	0.71409
5	1.5698	0.55621	0.74603
6	1.654	0.54466	0.77041
7	1.7239	0.53483	0.78976
8	1.7834	0.52634	0.80557
9	1.8351	0.51893	0.81878
10	1.8807	0.51238	0.83002
11	1.9215	0.50653	0.83972
12	1.9583	0.50127	0.8482
13	1.9918	0.4965	0.85568
14	2.0225	0.49214	0.86234
15	2.0509	0.48815	0.86831
16	2.0772	0.48446	0.87371
17	2.1017	0.48103	0.87861
18	2.1247	0.47785	0.88309
19	2.1462	0.47487	0.8872
20	2.1666	0.47208	0.89098

We now consider the covariance of the variable pointing in direction of the split, and some other variable, which here is exemplarily $\mathrm{Cov}(Z_1^{1+}, Z_2^{1+}))$ given by

$$\begin{aligned} \mathrm{Cov}\left(Z_1^{1+}, Z_2^{1+}\right) &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} \left(z_1 - \mu_1^{1+}\right)(z_2 - \underbrace{\mu_2^{1+}}_{=0})\, \xi_1^+\, \varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= 2N \int_0^\infty \left(z_1 - \mu_1^{1+}\right) \varphi\left(z_1\right) \{ \underbrace{\int_{-z_1}^{z_1} z_2\, \varphi\left(z_2\right) dz_2}_{=0} \\ &\quad \cdot \int_{-z_1}^{z_1} \varphi\left(z_3\right) dz_3 \cdot \ldots \cdot \int_{-z_1}^{z_1} \varphi\left(z_N\right) dz_N \}\, dz_1 = 0 \end{aligned} \tag{49}$$

Similarly,

$$\begin{aligned}\mathrm{Cov}\left(Z_2^{1+}, Z_3^{1+}\right) &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} (z_2 - \underbrace{\mu_2^{1+}}_{=0})(z_3 - \underbrace{\mu_3^{1+}}_{=0})\, \xi_1^+ \, \varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_n\right)\, dz_n \dots dz_1 \\ &= 2N \int_0^\infty \varphi\left(z_1\right) \{ \underbrace{\int_{-z_1}^{z_1} z_2\, \varphi\left(z_2\right) dz_2}_{=0} \cdot \underbrace{\int_{-z_1}^{z_1} z_3\, \varphi\left(z_3\right) dz_3}_{=0} \\ &\quad \cdot \int_{-z_1}^{z_1} \varphi\left(z_4\right) dz_4 \cdot \dots \cdot \int_{-z_1}^{z_1} \varphi\left(z_n\right) dz_n \}\, dz_1 = 0\end{aligned} \tag{50}$$

The variance of Z_1^{1+} is obtained by

$$\begin{aligned}\mathrm{Var}\left(Z_1^{1+}\right) &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} \left(z_1 - \mu_1^{1+}\right)^2 \xi_1^+ \, \varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= 2N \int_0^\infty \left(z_1 - \mu_1^{1+}\right)^2 \varphi\left(z_1\right) \\ &\quad \left\{ \int_{-z_1}^{z_1} \varphi\left(z_2\right) dz_2 \cdot \dots \cdot \int_{-z_1}^{z_1} \varphi\left(z_N\right) dz_N \right\} dz_1 \\ &= 2N \int_0^\infty \left(z_1 - \mu_1^{1+}\right)^2 \varphi\left(z_1\right) \left\{ \Phi\left(z_1\right) - \Phi\left(-z_1\right) \right\}^{N-1} dz_1\end{aligned} \tag{51}$$

With $\Phi(-z) = 1 - \Phi(z)$, we get

$$\mathrm{Var}\left(Z_1^{1+}\right) = 2N \int_0^\infty \left(z_1 - \mu_1^{1+}\right)^2 \varphi\left(z_1\right) \left\{ 2\,\Phi\left(z_1\right) - 1 \right\}^{N-1} dz_1 \tag{52}$$

We showed exemplarily for Z_2^{1+} how any other variance of the subdomain is determined by

$$\begin{aligned}\mathrm{Var}\left(Z_2^{1+}\right) &= \int_0^\infty \int_{-z_1}^{z_1} \dots \int_{-z_1}^{z_1} (z_2 - \underbrace{\mu_2}_{=0})^2\, \xi_1^+ \, \varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\, dz_N \dots dz_1 \\ &= 2N \int_0^\infty \varphi\left(z_1\right) \left\{ \int_{-z_1}^{z_1} z_2^2\, \varphi\left(z_2\right) dz_2 \cdot \int_{-z_1}^{z_1} \varphi\left(z_3\right) dz_3 \right. \\ &\quad \left. \cdot \dots \cdot \int_{-z_1}^{z_1} \varphi\left(z_N\right) dz_N \right\} dz_1\end{aligned} \tag{53}$$

The solution of $\int_{-z_1}^{z_1} z_2^2\, \varphi\left(z_2\right) dz_2$ is obtained by integration by parts. With

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad \text{and} \qquad \varphi'(z) = -\frac{z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \tag{54}$$

, we can write

$$\begin{aligned}\int_b^a z^2\, \varphi(z) dz &= -\int_b^a z\, \underbrace{\frac{-z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}_{\varphi'} dz = \int_b^a \varphi(z) dz - [z\, \varphi(z)]_b^a \\ &= \Phi(a) - \Phi(b) - a\, \varphi(a) + b\, \varphi(b)\end{aligned} \tag{55}$$

With $\varphi\left(z_1\right) = \varphi\left(-z_1\right)$ and $\Phi\left(-z_1\right) = 1 - \Phi\left(z_1\right)$, we get

$$\int_{-z_1}^{z_1} z_2^2\, \varphi\left(z_2\right) dz_2 = 2\Phi\left(z_1\right) - 1 - 2\, z_1\, \varphi\left(z_1\right) \tag{56}$$

Inserting this into Eq. (53) yields

$$\mathrm{Var}\left(Z_2^{1+}\right) = 2N \int_0^\infty \varphi\left(z_1\right) \left\{ 2\Phi\left(z_1\right) - 2\, z_1\, \varphi\left(z_1\right) - 1 \right\} \left\{ 2\Phi\left(z_1\right) - 1 \right\}^{N-2} dz_1 \tag{57}$$

A4 Stochastic moments of subdomains for splits along bisectors, centered at *s*

In difference to "Appendix A3", now the split is centered at some vector $\boldsymbol{s}$, as shown in Fig. 11. From now on, the index p indicates the direction in which the subdomain is pointing. All equations are derived for the case that the subdomain is open in positive p-direction, i.e., we consider the subdomain "$p+$." For the subdomain "$p-$," the integral bounds of p have to be replaced by $-$ inf as lower bound and s_p as upper bound, where s_p is the p-th entry of $\boldsymbol{s}$.

First, we simplify the integral that provides the weight factor ω_p^+.

$$\omega_p^+ = \int\limits_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int\limits_{s_p}^{\infty} \dots \int\limits_{s_p-z_p+s_N}^{z_p-s_p+s_N} \varphi(z_1) \dots \varphi(z_N)\, dz_N \dots dz_p \dots dz_1 = \int\limits_{s_p}^{\infty} \varphi(z_p) \left\{ \prod_{\substack{k=1\\k\neq p}}^{N} \Phi(z_p - s_p + s_k) - \Phi(s_p - z_p + s_k) \right\} dz_p \tag{58}$$

Since the term in curled brackets occurs more often, we define

$$\tilde{\Phi}_{k\neq p} = \prod_{\substack{k=1\\k\neq p}}^{N} \Phi(z_p - s_p + s_k) - \Phi(s_p - z_p + s_k) \tag{59}$$

where the subscript of $\tilde{\Phi}$ indicates the indices that are skipped in the product. With that, ω_p^+ reads

$$\omega_p^+ = \int\limits_{s_p}^{\infty} \varphi(z_p) \tilde{\Phi}_{k\neq p}\, dz_p \tag{60}$$

and the scale factor equals $\xi_p^+ = 1/\omega_p^+$.

The mean value in p-direction of the subdomain "$p+$" is given by

$$E\left(Z_p^{p+}\right) = \mu_p^{p+} = \int\limits_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int\limits_{s_p}^{\infty} \dots \int\limits_{s_p-z_p+s_N}^{z_p-s_p+s_N} z_p\, \xi_p^+\, \varphi(z_1) \dots \varphi(z_N)\, dz_N \dots dz_p \dots dz_1 = \xi_p^+ \int\limits_{s_p}^{\infty} z_p\, \varphi(z_p) \tilde{\Phi}_{k\neq p}\, dz_p \tag{61}$$

The mean value in any other direction $i \neq p$ of the subdomain "$p+$" is given by

$$E\left(Z_i^{p+}\right) = \mu_i^{p+} = \int\limits_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int\limits_{s_p}^{\infty} \dots \int\limits_{s_p-z_p+s_N}^{z_p-s_p+s_N} z_i\, \xi_p^+\, \varphi(z_1) \dots \varphi(z_N)\, dz_N \dots dz_1 = \xi_p^+ \int\limits_{s_p}^{\infty} \varphi(z_p) \left\{ \prod_{\substack{k=1\\k\neq p,i}}^{N} \int\limits_{s_p-z_p+s_k}^{z_p-s_p+s_k} \varphi(z_k)\, dz_k \right\} \left\{ \int\limits_{s_p-z_p+s_i}^{z_p-s_p+s_i} z_i\, \varphi(z_i)\, dz_i \right\} dz_p \tag{62}$$

To solve $\int\limits_{s_p-z_p+s_i}^{z_p-s_p+s_i} z_i\, \varphi(z_i)\, dz_i$ let us consider

$$\int\limits_b^a z\, \varphi(z) dz = \int\limits_b^a z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = -\int\limits_b^a \underbrace{\frac{-z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}_{\varphi'} dz = -\left[\varphi(z)\right]_b^a = \varphi(b) - \varphi(a) \tag{63}$$

Then,

$$\mu_i^{p+} = \xi_p^+ \int\limits_{s_p}^{\infty} \varphi(z_p) \tilde{\Phi}_{k\neq p,i} \left\{ \varphi(s_p - z_p + s_i) - \varphi(z_p - s_p + s_i) \right\} dz_p \tag{64}$$

Now we consider the covariance of two different parameters, which are not in p-direction, i.e., $i \neq j \neq p$

$$\operatorname{Cov}\left(Z_i^{p+}, Z_j^{p+}\right) = \int\limits_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int\limits_{s_p}^{\infty} \dots \int\limits_{s_p-z_p+s_N}^{z_p-s_p+s_N} \left(z_i - \mu_i^{p+}\right)\left(z_j - \mu_j^{p+}\right) \xi_p^+\, \varphi(z_1) \dots \varphi(z_N)\, dz_N \dots dz_1 = \xi_p^+ \int\limits_{s_p}^{\infty} \varphi(z_p) \left\{ \prod_{\substack{k=1\\k\neq p,i,j}}^{N} \int\limits_{s_p-z_p+s_k}^{z_p-s_p+s_k} \varphi(z_k)\, dz_k \right\} \left\{ \prod_{k=i,j} \int\limits_{s_p-z_p+s_k}^{z_p-s_p+s_k} \left(z_k - \mu_k^{p+}\right) \varphi(z_k)\, dz_k \right\} dz_p \tag{65}$$

To solve the integral in the last curled brackets, let us consider

$$\begin{aligned}\int_b^a (z-\mu)\,\varphi(z)dz &= \underbrace{\int_b^a z\,\varphi(z)dz}_{\text{see above}} - \mu\int_b^a \varphi(z)dz \\ &= \varphi(b) - \varphi(a) - \mu\,\Phi(a) + \mu\,\Phi(b)\end{aligned} \tag{66}$$

We now define

$$\begin{aligned}\tilde{\varphi}_k &= \int_{s_p-z_p+s_k}^{z_p-s_p+s_k} \left(z_k - \mu_k^{p+}\right)\varphi\left(z_k\right)\,dz_k \\ &= \varphi\left(s_p - z_p + s_k\right) - \varphi\left(z_p - s_p + s_k\right) \\ &\quad - \mu_k^{p+}\,\Phi\left(z_p - s_p + s_k\right) + \mu_k^{p+}\,\Phi\left(s_p - z_p + s_k\right)\end{aligned} \tag{67}$$

Then, the covariance equals

$$\mathrm{Cov}\left(Z_i^{p+}, Z_j^{p+}\right) = \xi_p^+ \int_{s_p}^{\infty} \varphi\left(z_p\right)\,\tilde{\Phi}_{k\neq p,i,j}\left\{\prod_{k=i,j}\tilde{\varphi}_k\right\}\,dz_p \tag{68}$$

Now, we consider the covariance for the case that one parameter involved is Z_p^{p+}.

$$\begin{aligned}\mathrm{Cov}\left(Z_p^{p+}, Z_i^{p+}\right) &= \int_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int_{s_p}^{\infty} \\ &\quad \dots \int_{s_p-z_p+s_N}^{z_p-s_p+s_N} \left(z_p - \mu_p^{p+}\right)\left(z_i - \mu_i^{p+}\right)\xi_p^+\,\varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\,dz_N \dots dz_p \dots dz_1 \\ &= \xi_p^+ \int_{s_p}^{\infty} \left(z_p - \mu_p^{p+}\right)\varphi\left(z_p\right) \\ &\quad \left\{\prod_{\substack{k=1\\ k\neq p,i}}^{N} \int_{s_p-z_p+s_k}^{z_p-s_p+s_k} \varphi\left(z_k\right)\,dz_k\right\} \\ &\quad \left\{\int_{s_p-z_p+s_i}^{z_p-s_p+s_i} \left(z_i - \mu_i^{p+}\right)\varphi\left(z_i\right)\,dz_i\right\} dz_p \\ &= \xi_p^+ \int_{s_p}^{\infty} \left(z_p - \mu_p^{p+}\right)\varphi\left(z_p\right)\;\tilde{\Phi}_{k\neq p,i}\;\tilde{\varphi}_i\,dz_p\end{aligned} \tag{69}$$

The variance of Z_p^{p+} is obtained by

$$\begin{aligned}\mathrm{Var}\left(Z_p^{p+}\right) &= \int_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int_{s_p}^{\infty} \dots \int_{s_p-z_p+s_N}^{z_p-s_p+s_N} \left(z_p - \mu_p^{p+}\right)^2 \xi_p^+\,\varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\,dz_N \dots dz_p \dots dz_1 \\ &= \xi_p^+ \int_{s_p}^{\infty} \left(z_p - \mu_p^{p+}\right)^2 \varphi\left(z_p\right) \\ &\quad \left\{\prod_{\substack{k=1\\ k\neq p}}^{N} \int_{s_1-z_1+s_k}^{z_1-s_1+s_k} \varphi\left(z_k\right)\,dz_k\right\} dz_p \\ &= \xi_p^+ \int_{s_p}^{\infty} \left(z_p - \mu_p^{p+}\right)^2 \varphi\left(z_p\right)\;\tilde{\Phi}_{k\neq p}\;dz_p\end{aligned} \tag{70}$$

Finally, we determined the variance in any direction $i \neq p$.

$$\begin{aligned}\mathrm{Var}\left(Z_i^{p+}\right) &= \int_{s_p-z_p+s_1}^{z_p-s_p+s_1} \dots \int_{s_p}^{\infty} \dots \int_{s_p-z_p+s_N}^{z_p-s_p+s_N} \left(z_i - \mu_i^{p+}\right)^2 \xi_p^+\,\varphi\left(z_1\right) \\ &\quad \dots \varphi\left(z_N\right)\,dz_N \dots dz_1 \\ &= \xi_p^+ \int_{s_p}^{\infty} \varphi\left(z_p\right)\left\{\prod_{\substack{k=1\\ k\neq p,i}}^{N} \int_{s_p-z_p+s_k}^{z_p-s_p+s_k} \varphi\left(z_k\right)\,dz_k\right\} \\ &\quad \left\{\int_{s_p-z_p+s_i}^{z_p-s_p+s_i} \left(z_i - \mu_i^{p+}\right)^2 \varphi\left(z_i\right)\,dz_i\right\} dz_p\end{aligned} \tag{71}$$

To solve $\int_{s_p-z_p+s_i}^{z_p-s_p+s_i} \left(z_i - \mu_i^{p+}\right)^2 \varphi\left(z_i\right)\,dz_i$ let us consider

$$\begin{aligned}&\int_b^a (z-\mu)^2\varphi(z)dz \\ &= \underbrace{\int_b^a z^2\varphi(z)dz}_{\text{see above}} - 2\mu\underbrace{\int_b^a z\,\varphi(z)dz}_{\text{see above}} + \mu^2\int_b^a \varphi(z)dz \\ &= \Phi(a)\left(1+\mu^2\right) - \Phi(b)\left(1+\mu^2\right) + \varphi(a)(2\mu - a) \\ &\quad + \varphi(b)(b - 2\mu)\end{aligned} \tag{72}$$

If we define

$$\begin{aligned}\Psi_i &= \int\limits_{s_p - z_p + s_i}^{z_p - s_p + s_i} \left(z_i - \mu_i^{p+}\right)^2 \varphi\left(z_i\right)\, dz_i \\ &= \Phi\left(z_p - s_p + s_i\right)\left(1 + \left(\mu_i^{p+}\right)^2\right) \\ &\quad - \Phi\left(s_p - z_p + s_i\right)\left(1 + \left(\mu_i^{p+}\right)^2\right) \\ &\quad + \varphi\left(z_p - s_p + s_i\right)\left(2\mu_i^{p+} - z_p - s_p + s_i\right) \\ &\quad + \varphi\left(s_p - z_p + s_i\right)\left(s_p - z_p + s_i - 2\mu_i^{p+}\right)\end{aligned} \tag{73}$$

, we can write the variance as

$$\mathrm{Var}\left(Z_i^{p+}\right) = \xi_p^{+} \int\limits_{s_p}^{\infty} \varphi\left(z_p\right)\, \tilde{\Phi}_{k \neq p,i}\, \Psi_i\, dz_p \tag{74}$$

Author contributions Benedikt Kriegesmann acquired funding, developed the methodology, produced results, and wrote the manuscript. Jan Krüger developed the methodology, produced the results, and wrote the manuscript. Kai Steltner produced the results and wrote the manuscript. Timo Schmidt produced the results and wrote the manuscript. Robert Seifried supervised and reviewed the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. Financial support of the German Research Foundation (Project Number 508865334) is acknowledged.

Data availability The paper contains no experimental data. The paper contains all information necessary to reproduce the results.

Declarations

Conflict of interest The authors state that there is no conflict of interest.

Replication of results The results for two out of the four examples (shallow arc and contact problem) were produced with unpublished finite element codes, using standard elements. Details on the considered nonlinearities (geometric and contact) are given in the example section. For two other examples (beam and panel), Abaqus was used for the finite element analysis. For these examples, the models will be provided upon reasonable request. Note that for the panel example, more detailed descriptions are also given in Steltner et al. (2022).

References

Abaqus. https://www.3ds.com/products/simulia/abaqus

Arbocz J, Babcock J, Charles D (1969) The effect of general imperfections on the buckling of cylindrical shells. J Appl Mech 36(1):28–38. https://doi.org/10.1115/1.3564582

Asadpoure A, Tootkaboni M, Guest JK (2011) Robust topology optimization of structures with uncertainties in stiffness—application to truss structures. Comput Struct 89(11):1131–1141. https://doi.org/10.1016/j.compstruc.2010.11.004. http://www.sciencedirect.com/science/article/pii/S004579491000266X

Basudhar A, Missoum S, Harrison Sanchez A (2008) Limit state function identification using Support Vector Machines for discontinuous responses and disjoint failure domains. Prob Eng Mech 23(1):1–11. https://doi.org/10.1016/j.probengmech.2007.08.004. https://www.sciencedirect.com/science/article/pii/S026689200700029X

Bendsøe MP, Sigmund O (2004) Topology optimization theory, methods, and applications, 2nd edn. Springer, Berlin. http://www.springer.com/us/book/9783540429920#aboutBook

Das S, Maity S, Qu B-Y, Suganthan PN (2011) Real-parameter evolutionary multimodal optimization—a survey of the state-of-the-art. Swarm Evol Comput 1(2):71–88. https://doi.org/10.1016/j.swevo.2011.05.005. https://www.sciencedirect.com/science/article/pii/S221065021100023X

Der Kiureghian A, Dakessian T (1998) Multiple design points in first and second-order reliability. Struct Saf 20(1):37–49. https://doi.org/10.1016/S0167-4730(97)00026-X. https://www.sciencedirect.com/science/article/pii/S016747309700026X

Ditlevsen O, Madsen HO (1996) Structural reliability methods. Wiley, Chichester

Doltsinis I, Kang Z (2004) Robust design of structures using optimization methods. Comput Methods Appl Mech Eng 193(23):2221–2237. https://doi.org/10.1016/j.cma.2003.12.055. http://www.sciencedirect.com/science/article/pii/S0045782504000787

Doltsinis I, Kang Z, Cheng G (2005) Robust design of non-linear structures using optimization methods. Comput Methods Appl Mech Eng 194(12):1779–1795. https://doi.org/10.1016/j.cma.2004.02.027. http://www.sciencedirect.com/science/article/pii/S0045782504004165

Elishakoff I (2004) Safety factors and reliability: Friends or foes? Springer, Berlin

Frühwirth-Schnatter S (2006) Finite mixture and Markov switching models. Springer. https://link.springer.com/book/10.1007/978-0-387-35768-3

He J, Guan X, Jha R (2016) Improve the accuracy of asymptotic approximation in reliability problems involving multimodal distributions. IEEE Trans Reliab 65(4):1724–1736. https://doi.org/10.1109/TR.2016.2604121

Hu Z, Du X (2019) Reliability methods for bimodal distribution with first-order approximation1. ASCE-ASME J Risk Uncert Eng Sys Part B Mech Eng 5(1):011005. https://doi.org/10.1115/1.4040000. https://asmedigitalcollection.asme.org/risk/article/doi/10.1115/1.4040000/366083/Reliability-Methods-for-Bimodal-Distribution-With

Jansen M, Lombaert G, Schevenels M (2015) Robust topology optimization of structures with imperfect geometry based on geometric nonlinear analysis. Comput Methods Appl Mech Eng 285:452–467. https://doi.org/10.1016/j.cma.2014.11.028. http://www.sciencedirect.com/science/article/pii/S004578251400454X

Jing S, Zhang Z, Meng X (2024) Surrogate-based robust design optimization by using Chebyshev-transformed orthogonal grid. Struct Multidisc Optim 67(7):127. https://doi.org/10.1007/s00158-024-03839-2

Kantarakias KD, Papadakis G (2023) Sensitivity-enhanced generalized polynomial chaos for efficient uncertainty quantification. J Comput Phys 491:112377. https://doi.org/10.1016/j.jcp.2023.112377. https://www.sciencedirect.com/science/article/pii/S0021999123004722

Kranz M, Lüdeker JK, Kriegesmann B (2023) A generalized approach for robust topology optimization using the first-order second-moment method for arbitrary response functions. Struct Multidisc Optim 66(5):98. https://doi.org/10.1007/s00158-023-03540-w

Kriegesmann B (2021) On the applicability of first-order approximations for design optimization under uncertainty. In: Challamel N, Kaplunov J, Takewaki I (eds) Modern trends in structural and solid mechanics 3 - non-deterministic mechanics. Wiley, pp. 39–60, section: 3 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119831839.ch3. https://doi.org/10.1002/9781119831839.ch3. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119831839.ch3

Kriegesmann B (2020) Robust design optimization with design-dependent random input variables. Struct Multidisc Optim 61(2):661–674. https://doi.org/10.1007/s00158-019-02388-3

Kriegesmann B, Lüdeker JK (2019) Robust compliance topology optimization using the first-order second-moment method. Struct Multidisc Optim 60(1):269–286. https://doi.org/10.1007/s00158-019-02216-8

Kriegesmann B, Rolfes R, Hühne C, Kling A (2011) Fast probabilistic design procedure for axially compressed composite cylinders. Compos Struct 93:3140–3149. https://doi.org/10.1016/j.compstruct.2011.06.017

Kriegesmann B, Jansen EL, Rolfes R (2012) Semi-analytic probabilistic analysis of axially compressed stiffened composite panels. Compos Struct 94(2):654–663. https://doi.org/10.1016/j.compstruct.2011.08.033

Krüger JC, Kranz M, Schmidt T, Seifried R, Kriegesmann B (2023) An efficient and non-intrusive approach for robust design optimization with the first-order second-moment method. Comput Methods Appl Mech Eng 414:116136. https://doi.org/10.1016/j.cma.2023.116136. https://linkinghub.elsevier.com/retrieve/pii/S0045782523002608

Krüger JC, Kriegesmann B, Robust design optimization using a non-intrusive second-order approximation of stochastic moments. Struct Multidisc Optim (2024) accepted for publication

Lazarov BS, Schevenels M, Sigmund O (2012) Topology optimization with geometric uncertainties by perturbation techniques. Int J Numer Meth Eng 90(11):1321–1336. https://doi.org/10.1002/nme.3361. http://onlinelibrary.wiley.com/doi/10.1002/nme.3361/abstract

Le C, Bruns T, Tortorelli D (2011) A gradient-based, parameter-free approach to shape optimization. Comput Methods Appl Mech Eng 200(9):985–996. https://doi.org/10.1016/j.cma.2010.10.004. http://www.sciencedirect.com/science/article/pii/S0045782510002823

Lindgaard E, Lund E (2011) A unified approach to nonlinear buckling optimization of composite structures. Comput Struct 89(3):357–370. https://doi.org/10.1016/j.compstruc.2010.11.008. https://www.sciencedirect.com/science/article/pii/S0045794910002749

MATLAB. https://de.mathworks.com/products/matlab.html

Meinen NE, Steenbergen RDJM (2025) Application of the Rosenblatt transformation in first-order system reliability approximations. Struct Saf 112:102521. https://doi.org/10.1016/j.strusafe.2024.102521. https://www.sciencedirect.com/science/article/pii/S0167473024000924

Nan H, Liang H, Di H, Li H (2024) A gradient-assisted learning strategy of Kriging model for robust design optimization. Reliab Eng Syst Saf 244:109944. https://doi.org/10.1016/j.ress.2024.109944. https://www.sciencedirect.com/science/article/pii/S095183202400019X

Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23(3):470–472. https://doi.org/10.1214/aoms/1177729394. https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-23/issue-3/Remarks-on-a-Multivariate-Transformation/10.1214/aoms/1177729394.full

Schmidt T, Kriegesmann B, Seifried R (2024) Robust contact-constrained topology optimization considering uncertainty at the contact support. Struct Multidisc Optim 67(4):46. https://doi.org/10.1007/s00158-024-03750-w

Schuëller GI, Valdebenito M (2010) Reliability-based optimization—an overview. Comput Technol Rev 1:121–155

Shi Y, Huang H-Z, Liu Y, Beer M (2023) Adaptive decoupled robust design optimization. Struct Saf 105:102378 https://doi.org/10.1016/j.strusafe.2023.102378. https://www.sciencedirect.com/science/article/pii/S0167473023000656

Sigmund O (2007) Morphology-based black and white filters for topology optimization. Struct Multidisc Optim 33(4–5):401–424. https://doi.org/10.1007/s00158-006-0087-x. https://link.springer.com/article/10.1007/s00158-006-0087-x

Signorini A (1959) Questioni di elasticità non linearizzata e semilinearizzata. Rendiconti di Matematica e delle sue Applicazioni 5:95–139

Song C, Shafieezadeh A, Xiao R, Sun B (2024) Analytical robust design optimization for hybrid design variables: an active-learning methodology based on polynomial chaos Kriging. Reliab Eng Syst Saf 250:110286. https://doi.org/10.1016/j.ress.2024.110286. https://www.sciencedirect.com/science/article/pii/S0951832024003582

Steltner K, Kriegesmann B, Pedersen CBW (2022) Robust sizing optimization of stiffened panels subject to geometric imperfections using fully nonlinear postbuckling analyses. Thin-Walled Struct 175:109195. https://doi.org/10.1016/j.tws.2022.109195. https://www.sciencedirect.com/science/article/pii/S0263823122001719

Strömberg N, Klarbring A (2010) Topology optimization of structures in unilateral contact. Struct Multidisc Optim 41:57–64. https://doi.org/10.1007/s00158-009-0407-z

Sudret B, Der Kiureghian A (2000) Stochastic finite element methods and reliability: a state-of-the-art report, Tech. Rep. UCB/SEMM-2000/08, Department of Civil and Environmental Engineering, University of California Berkeley, Berkeley

Svanberg K (2007) MMA and GCMMA—two methods for nonlinear optimization. Tech. rep, Optimization and Systems Theory, KTH, Stockholm, Sweden

Tosca. https://www.3ds.com/products/simulia/tosca

Wriggers P (2008) Nonlinear phenomena. In: Nonlinear finite element methods. Springer, Berlin, pp 7–18. https://doi.org/10.1007/978-3-540-71001-1_2. http://link.springer.com/chapter/10.1007/978-3-540-71001-1_2

Zang TA, Hemsch MJ, Hilburger MW, Kenny SP, Luckring JM, Maghami P, Padula SL, Stroud WJ (2002) Needs and opportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles, Tech. Rep. NASA/TM-2002-211462, National Aeronautics and Space Administration, Langley Research Center

Zhang Z, Wang J, Jiang C, Huang ZL (2019) A new uncertainty propagation method considering multimodal probability density functions. Struct Multidisc Optim 60(5):1983–1999. https://doi.org/10.1007/s00158-019-02301-y

Zimmermann R, Klein H, Kling A (2006) Buckling and postbuckling of stringer stiffened fibre composite curved panels-tests and computations. Compos Struct 73(2):150–161