A surrogate method for density-based global sensitivity analysis

Sharif Rahman

College of Engineering, The University of Iowa, Iowa City, IA 52242, USA

ABSTRACT

This paper describes an accurate and computationally efficient surrogate method, known as the polynomial dimensional decomposition (PDD) method, for estimating a general class of density-based f-sensitivity indices. Unlike the variance-based Sobol index, the f-sensitivity index is applicable to random input following dependent as well as independent probability distributions. The proposed method involves PDD approximation of a high-dimensional stochastic response of interest, forming a surrogate input–output data set; kernel density estimations of output probability density functions from the surrogate data set; and subsequent Monte Carlo integration for estimating the f-sensitivity index. Developed for an arbitrary convex function f and an arbitrary probability distribution of input variables, the method is capable of calculating a wide variety of sensitivity or importance measures, including the mutual information, squared-loss mutual information, and $\mathcal{L}_1$-distance-based importance measure. Three numerical examples illustrate the accuracy, efficiency, and convergence properties of the proposed method in computing sensitivity indices derived from three prominent divergence or distance measures. A finite-element-based global sensitivity analysis of a leverarm was performed, demonstrating the ability of the method in solving industrial-scale engineering problems.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Global sensitivity analysis rooted in the output variance has been a longstanding staple for modeling and simulation of complex systems [1–4]. In many applications, however, the variance provides a restricted summary of output uncertainty. Therefore, sensitivity measures derived solely from the variance should be guardedly interpreted. A more rational sensitivity analysis should account for the entire probability distribution of an output variable, meaning that the probabilistic characteristics above and beyond the variance should be considered [5–9]. Addressing this fundamental concern has triggered the development of a generalized density-based sensitivity index, referred to as the f-sensitivity index, which is founded on the well-known f-divergence between conditional and unconditional output probability measures [6]. Unlike the variance-based Sobol index [1], the f-sensitivity is applicable to random input following dependent as well as independent probability distributions. Since the class of f-divergences supports numerous divergence or distance measures, a bevy of density-based sensitivity measures are possible, including the mutual information [10], squared-loss mutual information [11], and $\mathcal{L}_1$-distance-based importance measure [5], to name a few,

providing diverse choices to sensitivity analysis. A few researchers have applied existing divergence or distance measures for sensitivity analysis of engineering systems [12–15].

While the formulation of the f-sensitivity index is not overly complicated, its calculation in general is more intricate than the calculation of the variance-based sensitivity index. This is chiefly because the probability density functions required in defining the convex function f are harder to estimate than the variance. If the function f is already selected, resulting in a specific sensitivity index, then one can exploit the functional structure of f to devise accurate and efficient methods of calculation. This is exemplified in the works of Borgonovo [5], Liu and Homma [16], and Wei et al. [17], which present several estimation procedures for calculating the $\mathcal{L}_1$-distance-based importance measure. Here, the author delves into calculating the f-sensitivity index derived from a general convex function f, so that the method proposed is applicable to a host of density-based sensitivity indices. Nonetheless, if the sample size concomitant with a required accuracy in estimating the sensitivity index is very large, say, in the order of millions or more, then existing methods [5,16,17] will be limited to problems or functions that are inexpensive to evaluate. In a practical setting, however, the response function is often determined via time-consuming, costly finite-element analysis (FEA) or similar numerical calculations. In which case, an arbitrarily large sample size is no longer viable, and hence alternative routes to estimating the output probability densities, leading to the f-sensitivity index, should be explored.

This paper presents an accurate and computationally efficient surrogate method, known as the polynomial dimensional decomposition (PDD) method, for estimating a general class of density-based f-sensitivity indices. The method is based on (1) PDD approximation of a high-dimensional stochastic response, forming a surrogate input-output data set; (2) kernel density estimations (KDEs) of output probability density functions from the surrogate data set; and (3) subsequent Monte Carlo integration for estimating the f-sensitivity index. Section 2 formally defines the f-sensitivity index, summarizes its fundamental properties, and cites a few special cases. Section 3 briefly reviews the PDD approximation and explains how the integration of the PDD approximation with KDE leads to calculating the f-sensitivity index. Numerical results from two mathematical functions, as well as from a computationally intensive solid-mechanics problem, are reported in Section 4. Finally, conclusions are drawn in Section 5.

2. A general sensitivity measure

Let $\mathbb{N}:=\{1, 2, \ldots\}$, $\mathbb{N}_0:=\mathbb{N} \cup \{0\}$, $\mathbb{R}:=(-\infty, \infty)$, and $\mathbb{R}_0^+:=[0, \infty)$ represent the sets of positive integer (natural), non-negative integer, real, and non-negative real numbers, respectively. For $k \in \mathbb{N}$, denote by $\mathbb{R}^k$ the k-dimensional Euclidean space and by $\mathbb{N}_0^k$ the k-dimensional multi-index space. These standard notations will be used throughout the paper.

2.1. f-Divergence

Let $(\Psi, \mathcal{G})$ be a measurable space, where Ψ is a sample space and $\mathcal{G}$ is a σ-algebra of the subsets of Ψ, and μ be a σ-finite measure on $(\Psi, \mathcal{G})$. Let $\mathcal{P}$ be a set of all probability measures on $(\Psi, \mathcal{G})$, which are absolutely continuous with respect to μ. For two such probability measures $P_1, P_2 \in \mathcal{P}$, let $dP_1/d\mu$ and $dP_2/d\mu$ denote the Radon–Nikodym derivatives of P_1 and P_2 with respect to the dominating measure μ, that is, $P_1 << \mu$ and $P_2 << \mu$. For absolutely continuous measures in terms of probability theory, take Ψ to be the real line and μ to be the Lebesgue measure, that is, $d\mu = d\xi$, $\xi \in \mathbb{R}$, so that $dP_1/d\mu$ and $dP_2/d\mu$ are simply probability density functions, denoted by $f_1(\xi)$ and $f_2(\xi)$, respectively.

Let $f: [0, \infty) \to (-\infty, \infty]$ be an extended real-valued function, which is (1) continuous on $[0, \infty)$ and finite-valued on $(0, \infty)$; (2) convex on $[0, \infty)$; (3) strictly convex at $t=1$; and (4) equal to zero at $t=1$, that is, $f(1) = 0$. The f-divergence, describing the difference or discrimination between two probability measures P_1 and P_2, is defined by the integral [18–21]

$$D_f(P_1 \| P_2) := \int_{\mathbb{R}} f\left(\frac{f_1(\xi)}{f_2(\xi)}\right) f_2(\xi) d\xi. \tag{1}$$

The divergence measure in Eq. (1) was introduced in the 1960s by Csiszár [18,19], Ali and Silvey [20], and Morimoto [21].

2.2. f-sensitivity index

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is a sample space, $\mathcal{F}$ is a σ-field on Ω, and $P: \mathcal{F} \to [0, 1]$ is a probability measure. With $\mathcal{B}^N$ representing the Borel σ-field on $\mathbb{R}^N$, $N \in \mathbb{N}$, consider an $\mathbb{R}^N$-valued absolutely continuous random vector $\mathbf{X}:=(X_1, \ldots, X_N): (\Omega, \mathcal{F}) \to (\mathbb{R}^N, \mathcal{B}^N)$, describing the statistical uncertainties in all system and input parameters of a general stochastic problem. The probability law of $\mathbf{X}$, which may comprise independent or dependent random variables, is completely defined by its joint probability density function $f_{\mathbf{X}}: \mathbb{R}^N \to \mathbb{R}_0^+$. Let u be a non-empty subset of $\{1, \ldots, N\}$ with the complementary set $-u:=\{1, \ldots, N\} \setminus u$ and cardinality $1 \leq |u| \leq N$, and let $\mathbf{X}_u = (X_{i_1}, \ldots, X_{i_{|u|}})$, $1 \leq i_1 < \cdots < i_{|u|} \leq N$, be a subvector of $\mathbf{X}$ with

$\mathbf{X}_{-u} := \mathbf{X}_{\{1,\ldots,N\} \setminus u}$ defining its complementary subvector. Then, for a given $\varnothing \neq u \subseteq \{1, \ldots, N\}$, the marginal density function of $\mathbf{X}_u$ is $f_{\mathbf{X}_u}(\mathbf{x}_u) := \int_{\mathbb{R}^{N-|u|}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}_{-u}$.

Let $y(\mathbf{X}) := y(X_1, \ldots, X_N)$, a real-valued, continuous, measurable transformation on $(\Omega, \mathcal{F})$, define a general, square-integrable stochastic response function of interest. Define $Y := y(\mathbf{X})$ to be the associated output random variable. Denote by P_Y and $P_{Y|\mathbf{X}_u}$ the output probability measures and by $f_Y(\xi)$ and $f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)$ the probability density functions of random variables Y and $Y|\mathbf{X}_u$, respectively, where $Y|\mathbf{X}_u$ stands for Y conditional on $\mathbf{X}_u$, which is itself random. Setting $P_1 = P_Y$, $f_1 = f_Y$, $P_2 = P_{Y|\mathbf{X}_u}$, and $f_2 = f_{Y|\mathbf{X}_u}$ in Eq. (1), the f-divergence between the unconditional and conditional probability measures of Y becomes

$$D_f(P_Y \| P_{Y|\mathbf{X}_u}) := \int_{\mathbb{R}} f\left(\frac{f_Y(\xi)}{f_{Y|\mathbf{X}_u}(\xi|\mathbf{X}_u)}\right) f_{Y|\mathbf{X}_u}(\xi|\mathbf{X}_u) d\xi. \tag{2}$$

For density-based importance analysis, suppose that the sensitivity of Y with respect to a subset $\mathbf{X}_u$, $\varnothing \neq u \subseteq \{1, \ldots, N\}$, of input variables $\mathbf{X}$ is desired. As unveiled in a prequel [6], such a sensitivity measure can be linked to the f-divergence in Eq. (2). However, the f-divergence is random because $\mathbf{X}_u$ is random. Therefore, apply the expectation operator with respect to the probability measure of $\mathbf{X}_u$ on Eq. (2), thereby defining the f-divergence-rooted f-sensitivity index [6]

$$H_{u,f} := \mathbb{E}_{\mathbf{X}_u}\left[D_f(P_Y \| P_{Y|\mathbf{X}_u})\right] \tag{3}$$

of Y for $\mathbf{X}_u$. Applying the definition of the expectation operator in Eq. (3) and then substituting the expression of $D_f(P_Y \| P_{Y|\mathbf{X}_u})$ from Eq. (2) yields the f-sensitivity index

$$
\begin{aligned}
H_{u,f} &= \int_{\mathbb{R}^{|u|}} D_f(P_Y \| P_{Y|\mathbf{X}_u = \mathbf{x}_u}) f_{\mathbf{X}_u}(\mathbf{x}_u) d\mathbf{x}_u \\
&= \int_{\mathbb{R}^{|u|} \times \mathbb{R}} f\left(\frac{f_Y(\xi)}{f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)}\right) f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u) f_{\mathbf{X}_u}(\mathbf{x}_u) d\mathbf{x}_u d\xi \\
&= \int_{\mathbb{R}^{|u|} \times \mathbb{R}} f\left(\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u, Y}(\mathbf{x}_u, \xi)}\right) f_{\mathbf{X}_u, Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi,
\end{aligned}
\tag{4}
$$

where $P_{Y|\mathbf{X}_u = \mathbf{x}_u}$ and $f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u)$ are the probability measure and the probability density function, respectively, of Y conditional on $\mathbf{X}_u = \mathbf{x}_u$, and $f_{\mathbf{X}_u, Y}(\mathbf{x}_u, \xi)$ is the joint probability density function of $(\mathbf{X}_u, Y)$. The last equality in Eq. (4) is formed by recognizing $f_{\mathbf{X}_u, Y}(\mathbf{x}_u, \xi) = f_{Y|\mathbf{X}_u}(\xi|\mathbf{x}_u) f_{\mathbf{X}_u}(\mathbf{x}_u)$ and is useful for calculating the sensitivity index, to be discussed in Section 3.

2.3. General properties

It is important to emphasize a few general properties of the f-sensitivity index $H_{u,f}$ inherited from the f-divergence measure. The properties, originally derived in a prior work [6], are described in conjunction with six propositions as follows.

Proposition 1 (Non-negativity). The f-sensitivity index $H_{u,f}$ of Y for $\mathbf{X}_u$, $\varnothing \neq u \subseteq \{1, \ldots, N\}$, is non-negative and vanishes when Y and $\mathbf{X}_u$ are statistically independent.

Proposition 2 (Range of values). The range of values of $H_{u,f}$ is

$$0 \leq H_{u,f} \leq f(0) + f^*(0), \tag{5}$$

where $f(0) = \lim_{t \to 0_+} f(t)$ and $f^*(0) = \lim_{t \to 0_+} t f(1/t) = \lim_{t \to \infty} f(t)/t$.

Proposition 3 (Importance of all input variables). The f-sensitivity index $H_{\{1,\ldots,N\},f}$ of Y for all input variables $\mathbf{X} = (X_1, \ldots, X_N)$ is

$$H_{\{1,\ldots,N\},f} = f(0) + f^*(0), \tag{6}$$

where $f(0) = \lim_{t \to 0_+} f(t)$ and $f^*(0) = \lim_{t \to 0_+} t f(1/t) = \lim_{t \to \infty} f(t)/t$.

Proposition 4 (*Importance for an independent subset*). *Let* $\varnothing \neq u$, $v \subseteq \{1, ..., N\}$, $u \neq v$. *If* Y *and* $\mathbf{X}_v$ *are statistically independent, then*

$$H_{u \cup v, f} = H_{u \setminus v, f}. \tag{7}$$

In addition, if u *and* v *are disjoint subsets, that is,* $u \cap v = \varnothing$, *then*

$$H_{u \cup v, f} = H_{u, f}. \tag{8}$$

Proposition 5 (*Invariance*). *The* f-*sensitivity index* $H_{u, f}$ *of* Y *for* $\mathbf{X}_u$, $\varnothing \neq u \subseteq \{1, ..., N\}$ *is invariant under smooth and uniquely invertible transformations (diffeomorphisms) of* Y *and* $\mathbf{X}_u$.

Proposition 6 (*Bounds for metric* f-*divergences*). *Let* $\varnothing \neq u$, $v \subset \{1, ..., N\}$ *be two disjoint subsets such that* $u \cap v = \varnothing$. *For probability measures* P_Y, $P_{Y|\mathbf{X}_u}$, *and* $P_{Y|\mathbf{X}_{u \cup v}}$, *let* f *be a select convex generating function, which produces metric* f-*divergences from* P_Y *to* $P_{Y|\mathbf{X}_{u \cup v}}$, *from* P_Y *to* $P_{Y|\mathbf{X}_u}$, *and from* $P_{Y|\mathbf{X}_u}$ *to* $P_{Y|\mathbf{X}_{u \cup v}}$, *satisfying the triangle inequality*

$$D_f\left(P_Y \| P_{Y|\mathbf{X}_{u \cup v}}\right) \leq D_f\left(P_Y \| P_{Y|\mathbf{X}_u}\right) + D_f\left(P_{Y|\mathbf{X}_u} \| P_{Y|\mathbf{X}_{u \cup v}}\right). \tag{9}$$

Then

$$H_{u, f} \leq H_{u \cup v, f} \leq H_{u, f} + H_{v|u, f}, \tag{10}$$

where $H_{v|u, f} := \mathbb{E}_{\mathbf{X}_{u \cup v}}[D_f(P_{Y|\mathbf{X}_u} \| P_{Y|\mathbf{X}_{u \cup v}})]$ *is the conditional sensitivity index of* $Y|\mathbf{X}_u$ *for* $\mathbf{X}_{u \cup v}$. *Furthermore, if* $\mathbf{X}_u$ *and* $\mathbf{X}_v$ *are statistically independent, then*

$$H_{u, f} \leq H_{u \cup v, f} \leq H_{u, f} + H_{v, f}. \tag{11}$$

Readers interested in proofs of all six propositions are referred to the prior work [6].

Propositions 1 through 6 reveal a number of fundamental theoretical properties of the f-sensitivity index. First, the f-sensitivity index is non-negative. Second, the index is endowed with a range of values, where the smallest value is *zero*, but the largest value may be finite or infinite, depending on the generating function f chosen. The index vanishes or attains the largest value when the unconditional and conditional probability measures coincide or are mutually singular. Third, the index reaches the largest value for all input variables. Fourth, if the output variable and a subset of input variables are statistically independent, then there is no contribution from that subset of input variables to the sensitivity of the output variable. Fifth, unlike the variance-based Sobol index, which is invariant only under affine transformations, the f-sensitivity index is invariant under nonlinear but smooth and uniquely invertible transformations. Sixth, for a metric divergence, the resultant f-sensitivity index for a group of input variables increases from the unconditional sensitivity index for a subgroup of input variables, but is limited by the residual term emanating from the conditional sensitivity index.

2.4. Special cases

A multitude of f-sensitivity indices are possible by appropriately selecting the convex function f in Eqs. (2)–(4). Listed in Table 1 are 10 such convex functions and resultant sensitivity indices derived from the forward and reversed Kullback–Leibler divergences [22], total variational distance [23], Hellinger distance [24], Pearson χ^2 divergence [25], Neyman χ^2 divergence [25], α divergence [26], Vajda χ^α divergence [27], Jeffreys distance [28], and triangular discrimination [29]. Three prominent sensitivity indices, for example, the mutual information [10]

$$I_u := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \ln\left[\frac{f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi)}{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u)}\right] f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi) d\mathbf{X}_u d\xi =: H_{u, KL'} \tag{12}$$

between $\mathbf{X}_u$ and Y, the squared-loss mutual information [11]

$$I'_u := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left[\frac{f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi)}{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u)} - 1\right]^2 f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u) d\mathbf{X}_u d\xi$$

$$= \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \frac{f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi)}{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u)} \left[1 - \left\{\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u)}{f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi)}\right\}^2\right] f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi) d\mathbf{X}_u d\xi$$

$$=: H_{u, N} \tag{13}$$

between $\mathbf{X}_u$ and Y, and the $\mathcal{L}_1$-distance-based importance measure [5]

$$\delta_u := \frac{1}{2} \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left| f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u) - f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi) \right| d\mathbf{X}_u d\xi$$

$$= \frac{1}{2} \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left|\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{X}_u)}{f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi)} - 1\right| f_{\mathbf{X}_u, Y}(\mathbf{X}_u, \xi) d\mathbf{X}_u d\xi =: \frac{1}{2} H_{u, TV} \tag{14}$$

of $\mathbf{X}_u$ on Y, are rooted in reversed Kullback–Leibler, Neyman, and total variational divergences or distances, respectively. Indeed, many previously used sensitivity or importance measures are special cases of the f-sensitivity index derived from the f-divergence.

2.5. Univariate index

For practical applications, one may be interested in performing sensitivity analysis pertaining to individual random variables only. Setting $u = \{i\}$, where $i = 1, ..., N$, in Eqs. (3) and (4), the univariate f-sensitivity index of Y with respect to a variable X_i is

$$H_{\{i\}, f} := \mathbb{E}_{X_i}\left[D_f\left(P_Y \| P_{Y|X_i}\right)\right] := \int_{\mathbb{R} \times \mathbb{R}} f\left(\frac{f_Y(\xi) f_{X_i}(x_i)}{f_{X_i, Y}(x_i, \xi)}\right) f_{X_i, Y}(x_i, \xi) dx_i d\xi, \tag{15}$$

where $f_{X_i}(x_i)$ and $f_{X_i, Y}(x_i, \xi)$ are the marginal density of X_i and the joint density of (X_i, Y), respectively. Again, depending on the choice of the convex function f in Eq. (15), many univariate f-sensitivity indices from Table 1 with $u = \{i\}$ can be generated. For instance, the univariate f-sensitivity indices derived from the total variational distance, reversed Kullback–Leibler divergence, and Hellinger distance are

$$H_{\{i\}, TV} := \int_{\mathbb{R} \times \mathbb{R}} \left|\frac{f_Y(\xi) f_{X_i}(x_i)}{f_{X_i, Y}(x_i, \xi)} - 1\right| f_{X_i, Y}(x_i, \xi) dx_i d\xi, \tag{16a}$$

$$H_{\{i\}, KL'} := \int_{\mathbb{R} \times \mathbb{R}} \ln\left[\frac{f_{X_i, Y}(x_i, \xi)}{f_Y(\xi) f_{X_i}(x_i)}\right] f_{X_i, Y}(x_i, \xi) dx_i d\xi, \tag{16b}$$

and

$$H_{\{i\}, H} := \int_{\mathbb{R} \times \mathbb{R}} \left[\sqrt{\frac{f_Y(\xi) f_{X_i}(x_i)}{f_{X_i, Y}(x_i, \xi)}} - 1\right]^2 f_{X_i, Y}(x_i, \xi) dx_i d\xi, \tag{16c}$$

respectively. The univariate indices defined in Eqs. (16a), (16b) and (16c) will be referred to as total variational, Kullback–Leibler, and Hellinger sensitivity indices, respectively, in the paper.

3. The PDD method

According to the last line of Eq. (4), the f-sensitivity index $H_{u, f}$ is exactly calculated when both the integrand function f, which

Table 1
Ten special cases of the *f*-sensitivity index [6].

f-Divergence	$f(t)$[a]	Sensitivity index
Forward Kullback–Leibler divergence	$t \ln t$	$H_{u,KL} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \ln\left[\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Reversed Kullback–Leibler divergence	$-\ln t$	$H_{u,KL'} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \ln\left[\frac{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)}{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)} \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Total variational distance	$|t - 1|$	$H_{u,TV} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left| \frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} - 1 \right| f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Hellinger distance	$\left(\sqrt{t} - 1\right)^2$	$H_{u,H} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left[\sqrt{\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)}} - 1 \right]^2 f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Pearson χ^2 divergence	$t^2 - 1$	$H_{u,P} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left[\left\{ \frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \right\}^2 - 1 \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Neyman χ^2 divergence	$\dfrac{1 - t^2}{t}$	$H_{u,N} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \frac{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)}{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)} \left[1 - \left\{ \frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \right\}^2 \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
α-Divergence[b]	$\dfrac{4\left[1 - t^{(1-\alpha)/2} \right]}{1 - \alpha^2}$	$H_{u,\alpha} := \frac{4}{1 - \alpha^2} \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left[1 - \left[\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \right]^{(1-\alpha)/2} \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Vajda χ^α divergence[c]	$|t - 1|^\alpha$	$H_{u,V} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left| \frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} - 1 \right|^\alpha f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Jeffreys distance	$(t - 1) \ln t$	$H_{u,J} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \left[\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} - 1 \right] \ln\left[\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} \right] f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$
Triangular discrimination	$\dfrac{(t - 1)^2}{t + 1}$	$H_{u,\triangle} := \int_{\mathbb{R}^{|u|} \times \mathbb{R}} \frac{\left[\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} - 1 \right]^2}{\frac{f_Y(\xi) f_{\mathbf{X}_u}(\mathbf{x}_u)}{f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)} + 1} f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi) d\mathbf{x}_u d\xi$

[a] All generating functions have been normalized, that is, $f(1) = 0$.
[b] $\alpha \neq \pm 1$.
[c] $\alpha \geq 1$. If $\alpha = 1$, then $H_{u,V} = H_{u,TV}$.

depends on the probability densities $f_Y(\xi)$ and $f_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)$, and the subsequent $(|u| + 1)$-dimensional integral with respect to the probability measure of $(\mathbf{X}_u, Y)$ are exactly determined. For an arbitrary response function or probability distribution of random input, obtaining exact probability densities of responses is unrealistic. In which case, the probability densities should be assessed using approximate methods, say, the KDE methods, leading to their respective estimates $\bar{f}_Y(\xi)$ and $\bar{f}_{\mathbf{X}_u,Y}(\mathbf{x}_u, \xi)$. Furthermore, pursuing analytical integration is nearly impossible since the convex function f is arbitrary. Indeed, for a general function f, the integral must be approximated, for instance, by the Monte Carlo estimator

$$\bar{H}_{u,f}^{(L)} := \frac{1}{L} \sum_{l=1}^{L} f\left(\frac{\bar{f}_Y(\xi^{(l)}) \bar{f}_{\mathbf{X}_u}(\mathbf{x}_u^{(l)})}{\bar{f}_{\mathbf{X}_u,Y}(\mathbf{x}_u^{(l)}, \xi^{(l)})} \right) \tag{17}$$

of $H_{u,f}$, where $L \in \mathbb{N}$ is the sample size and $\{\mathbf{x}^{(l)}, \xi^{(l)}\}_{l=1,\ldots,L}$ is a data set comprising L input–output sample pairs of $(\mathbf{X}, Y)$. Unfortunately, generating such a data set for large L becomes computationally burdensome, if not prohibitive, when the output function y is determined via an expensive numerical analysis or algorithm. Therefore, Eq. (17) must be recast by exploiting mathematically rigorous yet computationally efficient surrogate approximations of the output function.

Modern surrogate approximations include polynomial chaos expansion (PCE) [30], stochastic collocation [31], and PDD [32,33], to name a few. All of these methods, commonly used for uncertainty quantification of complex systems, are known to offer significant computational advantages over crude Monte Carlo simulation. But there is one major limitation: most, if not all, surrogates are valid for independent input probability measures, although PDD has recently been extended to work also for

dependent variables [3]. For independent random input, the PDD and PCE are equivalent, but their respective approximations due to truncations are not the same. Indeed, numerical analyses of mathematical functions or simple dynamic systems reveal markedly higher convergence rates of the PDD approximation than the PCE approximation [34]. Furthermore, for truly high-dimensional problems, the PCE and collocation methods require astronomically large numbers of terms or coefficients, succumbing to the curse of dimensionality. The PDD reduces the computational effort, but more importantly, it deflates the curse of dimensionality to an extent determined by the degree of interaction among input variables [32,33]. This property, in addition to general applicability of PDD, was the principal motivation for coupling PDD, KDE, and a Monte Carlo integrator, to be described in the following subsections. The end product is referred to as the PDD method in this paper.

3.1. PDD approximation of stochastic response

For a given $\varnothing \neq u \subseteq \{1, \ldots, N\}$, $1 \leq |u| \leq N$, let $\mathbf{j}_{|u|} = (j_1, \ldots, j_{|u|}) \in \mathbb{N}_0^{|u|}$ represent a $|u|$-dimensional multi-index. Denote the associated space of the $|u|$-variate analysis-of-variance (ANOVA) [3] component functions of y by $\mathcal{L}_2(\Omega_u, \mathcal{F}_u, P_u)$, which is a Hilbert space. Let $\{\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\}$ be a set of multivariate orthonormal polynomials that is consistent with the probability measure $f_{\mathbf{X}_u}(\mathbf{x}_u) d\mathbf{x}_u$ of $\mathbf{X}_u$. Then $\{\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\}$, if its span is dense, constitutes a basis of $\mathcal{L}_2$. The standard Hilbert space theory states that every non-constant ANOVA component function of a square-integrable function y can be expanded in terms of $\{\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\}$. The result is the polynomial dimensional decomposition (PDD) [3,32,33],

$$y(\mathbf{X}) = y_\varnothing + \sum_{\varnothing \neq u \subseteq \{1,\ldots,N\}} \sum_{\mathbf{j}_{|u|} \in \mathbb{N}^{|u|}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \tag{18}$$

representing a finite, hierarchical expansion of y in terms of polynomials of input random variables with increasing dimensions. Here, the expansion coefficients of PDD are defined or determined from [3]

$$y_\varnothing := \int_{\mathbb{R}^N} y(\mathbf{x}) f_\mathbf{X}(\mathbf{x}) d\mathbf{x} \tag{19a}$$

and

$$C_{u\mathbf{j}_{|u|}} + \sum_{\substack{\varnothing \neq v \subseteq \{1,\ldots,N\} \\ v \cap u \neq \varnothing, v \not\subseteq u}} \sum_{\mathbf{k}_{|v|} \in \mathbb{N}^{|v|}} C_{v\mathbf{k}_{|v|}} J_{u\mathbf{j}_{|u|}, v\mathbf{k}_{|v|}} = I_{u\mathbf{j}_{|u|}}, \tag{19b}$$

where

$$I_{u\mathbf{j}_{|u|}} := \int_{\mathbb{R}^N} y(\mathbf{x}) \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) f_{\mathbf{X}_u}(\mathbf{x}_u) f_{\mathbf{X}_{-u}}(\mathbf{x}_{-u}) d\mathbf{x} \tag{19c}$$

and

$$J_{u\mathbf{j}_{|u|}, v\mathbf{k}_{|v|}} := \int_{\mathbb{R}^{|v \cup u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) \psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v) f_{\mathbf{X}_u}(\mathbf{x}_u) f_{\mathbf{X}_{v \cap -u}}(\mathbf{x}_{v \cap -u}) d\mathbf{x}_{v \cup u} \tag{19d}$$

are integrals that must be calculated. The definition of $y_\varnothing$ in Eq. (19a) is valid whether or not the random input $\mathbf{X}$ comprises independent or dependent variables. So is the linear system of equations, that is, Eq. (19b), from which the expansion coefficients $C_{u\mathbf{j}_{|u|}}$, $\varnothing \neq u \subseteq \{1,\ldots,N\}$, $\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}$, are calculated.

When the input variables are statistically independent with a product-type density function $f_\mathbf{X}(\mathbf{x}) = \prod_{i=1}^N f_{X_i}(x_i)$, then the definitions of the PDD expansion coefficients simplify such that [32,33]

$$y_\varnothing := \int_{\mathbb{R}^N} y(\mathbf{x}) \prod_{i=1}^N f_{X_i}(x_i) dx_i \tag{20a}$$

and

$$C_{u\mathbf{j}_{|u|}} := \int_{\mathbb{R}^N} y(\mathbf{x}) \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) \prod_{i=1}^N f_{X_i}(x_i) dx_i. \tag{20b}$$

Once the expansion coefficients are ascertained, whether obtained from Eqs. (19a) to (19d) or Eqs. (20a) and (20b), the PDD in Eq. (18) delivers a mean-square convergent expansion of any square-integrable function y [3,32,33].

Although Eq. (18) provides an exact mean-square representation, it contains an infinite number of coefficients, emanating from infinite numbers of orthonormal polynomials. In practice, the number of coefficients must be finite, say, by retaining at most mth-order polynomials. Furthermore, in many applications, the function y can be approximated by a sum of at most S-variate ANOVA or PDD component functions, where $1 \leq S \leq N$ is another truncation parameter, resulting in the S-variate, mth-order PDD approximation

$$\tilde{y}_{S,m}(\mathbf{X}) = y_\varnothing + \sum_{\substack{\varnothing \neq u \subseteq \{1,\ldots,N\} \\ 1 \leq |u| \leq S}} \sum_{k=|u|}^{m} \sum_{\substack{\|\mathbf{j}_{|u|}\|=k \\ \prod_{i=1}^{|u|} j_i \neq 0}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u), \tag{21}$$

which includes interactive effects of at most S input variables $X_{i_1}, \ldots, X_{i_S}$, $1 \leq i_1 \leq \cdots \leq i_S \leq N$, on y. Here, $\|\mathbf{j}_{|u|}\|$ is a suitable norm of $\mathbf{j}_{|u|}$, for instance, $\|\mathbf{j}_{|u|}\|_1 := j_1 + \cdots + j_{|u|}$ as the 1-norm or $\|\mathbf{j}_{|u|}\|_\infty := \max(j_1, \ldots, j_{|u|})$ as the ∞-norm. By selecting $S=1$ and $S=2$, the functions $\tilde{y}_{1,m}(\mathbf{X})$ and $\tilde{y}_{2,m}(\mathbf{X})$, respectively, provide univariate and bivariate mth-order PDD approximations, contain contributions from all input variables, and should not be viewed as first- and second-order approximations, nor do they limit the nonlinearity of $y(\mathbf{X})$. Depending on how the component functions

are constructed, arbitrarily high-order univariate and bivariate terms of $y(\mathbf{X})$ could be lurking inside $\tilde{y}_{1,m}(\mathbf{X})$ and $\tilde{y}_{2,m}(\mathbf{X})$. The fundamental conjecture underlying this decomposition is that the component functions arising in the function decomposition will exhibit insignificant S-variate effects cooperatively when $S \to N$, leading to useful lower-variate approximations of $y(\mathbf{X})$. When $S \to N$ and $m \to \infty$, $\tilde{y}_{S,m}(\mathbf{X})$ converges to $y(\mathbf{X})$ in the mean-square sense, that is, Eq. (21) generates a hierarchical and convergent sequence of approximations of $y(\mathbf{X})$. Further details of PDD are available elsewhere [3,32,33].

The determination of the PDD expansion coefficients $y_\varnothing$ and $C_{u\mathbf{j}_{|u|}}$ in Eqs. (19a)– (20b) requires calculating various N-dimensional integrals over $\mathbb{R}^N$. For large N, a full numerical integration employing an N-dimensional tensor product of a univariate quadrature formula is computationally prohibitive. The author proposes to apply the dimension-reduction Gaussian-integration scheme, originally developed by Xu and Rahman [35], to estimate the coefficients efficiently. The scheme entails approximating a high-dimensional integral of interest by a finite sum of lower-dimensional integrations and then using the Gauss quadrature formula. For instance, when the input variables are independent, this is accomplished by replacing the N-variate function y in Eqs. (20a) and (20b) with an R-variate truncation of the referential dimensional decomposition (RDD) at a chosen reference point [36], where $S \leq R \leq N$. The result is a reduced integration scheme, requiring evaluation of at most R-dimensional integrals [35]. When the reduced integrals for $R=S$ are numerically estimated, say, by a measure-consistent n-point Gauss quadrature rule ($n \in \mathbb{N}$), then the computational expense for an S-variate, mth-order PDD approximation involves a maximum of $\sum_{k=0}^{k=S} \binom{N}{k} n^k$ function evaluations [32]. In consequence, the computational complexity of the S-variate PDD approximation is an Sth-order polynomial with respect to the number of random variables or integration points. Therefore, PDD with dimension-reduction integration of the expansion coefficients alleviates the curse of dimensionality to an extent determined by S.

It is important to note that the dimension-reduction integration scheme is theoretically convergent [35]. In other words, when $R=N$, the resulting dimension-reduction reproduces the original integral, regardless of how the reference point is selected. When $R < N$, the error from the dimension-reduction integration depends not only on R, but also on the reference point. The integration error is related to the approximation error due to the truncation of RDD [36]. Furthermore, for a general function, one must use the Gauss quadrature formula to approximate all R- and lower-dimensional integrals, stemming from the dimension-reduction integration. The quadrature error can be regulated, given the continuity and/or smoothness of the integrand.

3.2. Kernel density estimation

Let $\tilde{Y}_{S,m} := \tilde{y}_{S,m}(\mathbf{X})$ define the S-variate, mth-order PDD approximation of the output random variable Y. Given the PDD truncation parameters S, m, and a sample size $L \in \mathbb{N}$, let $\left\{ \mathbf{x}^{(l)}, \tilde{\xi}_{S,m}^{(l)} \right\}_{l=1,\ldots,L}$ be the set of input–output sample pairs of $(\mathbf{X}, \tilde{Y}_{S,m})$, where the output samples are calculated from (21). Using this surrogate sample set, the KDEs of the probability density functions of $\tilde{Y}_{S,m}$ and $(\mathbf{X}_u, \tilde{Y}_{S,m})$ are obtained as [37,38]

$$\tilde{f}_{\tilde{Y}_{S,m}}(\xi) = \frac{1}{Lh_Y} \sum_{l=1}^L K_Y\left(\frac{\xi - \tilde{\xi}_{S,m}^{(l)}}{h_Y} \right) \tag{22a}$$

and

$$\bar{f}_{\mathbf{X}_u, \tilde{Y}_{S,m}}(\mathbf{x}_u, \xi)$$

$$= \frac{1}{L h_Y \prod_{i \in u} h_{X_i}} \sum_{l=1}^{L} K_Y\left(\frac{\xi - \tilde{\xi}_{S,m}^{(l)}}{h_Y}\right) \prod_{i \in u} K_{X_i}\left(\frac{x_i - x_i^{(l)}}{h_{X_i}}\right), \quad (22b)$$

respectively, where $K_{X_i}: \mathbb{R} \to \mathbb{R}$, $i \in u$, and $K_Y: \mathbb{R} \to \mathbb{R}$ are univariate kernel functions, and h_{X_i}, $i \in u$, and h_Y are smoothing parameters called the bandwidths. It is well known that the asymptotic mean-squared error committed by the KDE increases with the bandwidth size but decreases in the product of bandwidth and sample sizes. Therefore, for the KDE error to decline as $L \to \infty$, the bandwidth must decrease, but not at a rate faster than the sample size. This is sufficient to establish pointwise convergence of the KDE.

3.3. Calculation of sensitivity index

Substituting the probability densities in the last line of Eq. (4) with their KDEs in Eqs. (22a) and (22b) results in the PDD–KDE–Monte Carlo estimator

$$\tilde{H}_{u,f}^{(L,S,m)} := \frac{1}{L} \sum_{l=1}^{L} f\left(\frac{\bar{f}_Y(\tilde{\xi}_{S,m}^{(l)}) f_{\mathbf{X}_u}(\mathbf{x}_u^{(l)})}{\bar{f}_{\mathbf{X}_u, Y}(\mathbf{x}_u^{(l)}, \tilde{\xi}_{S,m}^{(l)})}\right) \quad (23)$$

of $H_{u,f}$. From the mean-square convergence properties [3,32,33], the sequence of PDD approximations $\{\tilde{y}_{S,m}; 1 \le S \le N, 1 \le m < \infty\}$ also converges to y in probability and in distribution as $S \to N$ and $m \to \infty$. Therefore, the estimator $\tilde{H}_{u,f}^{(L,S,m)}$ in Eq. (23) is anticipated to deliver a good approximation of $H_{u,f}$ when L is sufficiently large, provided that the PDD truncation parameters S and m are chosen wisely.

The input probability density $f_{\mathbf{X}}(\mathbf{x})$ of $\mathbf{X}$ is assumed to be known, so that the marginal density $f_{\mathbf{X}_u}(\mathbf{x}_u)$ of $\mathbf{X}_u$ for any $\varnothing \ne u \subseteq \{1, ..., N\}$ is exactly determined. This assumption is commonly invoked or fulfilled in uncertainty quantification, where the objective is propagating input uncertainties to determine the probabilistic characteristics of an output response. However, there are also data-driven stochastic problems, where raw input data, generated from either physical testing or field measurements, are supplied. In which case, the input density must also be estimated, say, by employing KDE to generate the approximate density $\bar{f}_{\mathbf{X}_u}(\mathbf{x}_u)$, imparting, therefore, an added layer of approximation to the proposed method.

Remark 1. Although the PDD approximations for dependent [3] or independent [32,33] random input were originally developed in past works, the current work describes how PDD can be effectively integrated with KDE and Monte Carlo simulation to estimate the density-based sensitivity indices. The integration proposed is novel in its own right.

Remark 2. Since the truncated PDD in Eq. (21) furnishes a surrogate approximation for an arbitrary random input, Eq. (23) is valid whether or not $\mathbf{X}$ comprises independent or dependent variables. In consequence, the proposed estimator can be viewed as a generalized version of the PDD estimator presented in an earlier work [6]. Moreover, it is important to emphasize that Eq. (23) constitutes single-loop (single simulation) samplings as opposed to double-loop computations required by several existing methods [5,16,17]. Therefore, the proposed estimator should be markedly more efficient than the existing methods even when computing old sensitivity indices.

Remark 3. It is important to emphasize that the output samples from the PDD approximation should not be confused with those generated from the original function. In a practical setting, the numerical calculations of y for generated input samples can be

computationally expensive or even prohibitive, particularly when the sample size needs to be very large. In contrast, the samples generated from the PDD approximation $\tilde{y}_{S,m}$ entail inexpensive evaluations of simple polynomial functions. Therefore, a relatively large sample size can be accommodated in the PDD method even when y is expensive to evaluate.

4. Numerical examples

Three numerical examples involving two elementary mathematical functions and an industrial-scale solid-mechanics problem are presented to illustrate the proposed PDD method for calculating three density-based sensitivity indices: the total variational sensitivity index $H_{\{i\},TV}$, the reversed Kullback–Leibler sensitivity index $H_{\{i\},KL}$, and the Hellinger sensitivity index $H_{\{i\},H}$. Two distinct norms of $\mathbf{j}_{|u|}$ in Eq. (21) were used: 1-norm for Example 1 and ∞-norm for Examples 2 and 3. In Examples 1 and 2, classical Hermite orthogonal polynomials, which are consistent with the dependent or independent Gaussian distributions of random input, were employed. In Example 3, classical Legendre polynomials were used for uniform distributions, whereas measure-consistent orthogonal polynomials generated from the Stieltjes procedure [32] were applied for lognormal distributions. The expansion coefficients involved in the first example were exactly calculated since $I_{u\mathbf{j}_{|u|}}$ and $J_{u\mathbf{j}_{|u|}, v\mathbf{k}_{|v|}}$ can be analytically determined [3]. In the remaining two examples, dimension-reduction integration and Gauss quadrature formulae consistent with orthogonal polynomials were employed to evaluate the PDD expansion coefficients. The dimension-reduction integration was performed with the mean input as the reference point and the truncation parameter $R = S$, so that an S-variate PDD approximation requires at most S-variate numerical integration [35]. Due to the large sample sizes chosen, only the means of Monte Carlo estimators are reported.

4.1. Example 1

The first example involves a quadratic transformation [3]

$$y(\mathbf{X}) = 12 + 4X_1 + 4X_2 + 4X_3 + X_1 X_2 + X_1 X_3 + X_2 X_3 \quad (24)$$

of a trivariate Gaussian random vector $\mathbf{X} = (X_1, X_2, X_3) \in \mathbb{R}^3$, which has mean $\mathbb{E}[\mathbf{X}] = \mathbf{0} \in \mathbb{R}^3$, positive-definite covariance matrix

$$\Sigma_{\mathbf{X}} = \mathbb{E}[\mathbf{X}\mathbf{X}^T] = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ (\text{sym.}) & & \sigma_3^2 \end{bmatrix} \in \mathbb{R}^{3\times3}, \quad (25)$$

comprising variances $\sigma_i^2 = 1$ of X_i for $i = 1, 2, 3$ and correlation coefficients ρ_{ij} between X_i and X_j, $i, j = 1, 2, 3$, $i \ne j$, and joint probability density function

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} (\det\Sigma_{\mathbf{X}})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\mathbf{x}^T\Sigma_{\mathbf{X}}^{-1}\mathbf{x}\right] =: \phi_{\mathbf{X}}(\mathbf{x}; \Sigma_{\mathbf{X}}). \quad (26)$$

Three distinct cases of correlation coefficients were examined: (1) $\rho_{12} = \rho_{13} = \rho_{23} = 1/5$ (equal correlation), (2) $\rho_{12} = 1/5$, $\rho_{13} = 2/5$, $\rho_{23} = 4/5$ (positive correlation), and (3) $\rho_{12} = -1/5$, $\rho_{13} = 2/5$, $\rho_{23} = -4/5$ (mixed correlation). The objective of this simple yet insightful example is to illustrate how the proposed method can be applied to determine f-sensitivity indices for dependent input, thereby ranking input variables concomitant with their correlation properties.

Given the Gaussian probability density function of $\mathbf{X}$, the marginal probability densities $\phi_{\mathbf{X}_u}(\mathbf{x}_u; \Sigma_{\mathbf{X}_u})$ of $\mathbf{X}_u$, $\varnothing \ne u \subseteq \{1, 2, 3\}$, are also Gaussian and are easily derived as

$$f_{\mathbf{X}_u}(\mathbf{x}_u) = (2\pi)^{-\frac{|u|}{2}}\big(\det\mathbf{\Sigma}_{\mathbf{X}_u}\big)^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\mathbf{x}_u^T\mathbf{\Sigma}_{\mathbf{X}_u}^{-1}\mathbf{x}_u\right]=:\phi_{\mathbf{X}_u}(\mathbf{x}_u;\mathbf{\Sigma}_{\mathbf{X}_u}), \quad (27)$$

where $\mathbf{\Sigma}_{\mathbf{X}_u}:=\mathbb{E}[\mathbf{X}_u\mathbf{X}_u^T]$ is the covariance matrix of $\mathbf{X}_u$. The probability density $\phi_{\mathbf{X}_u}(\mathbf{x}_u;\mathbf{\Sigma}_{\mathbf{X}_u})$ induces multivariate Hermite orthogonal polynomials

$$\tilde{\psi}_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) = \frac{(-1)^{|\mathbf{j}_{|u|}|}}{\phi_{\mathbf{X}_u}(\mathbf{x}_u;\mathbf{\Sigma}_{\mathbf{X}_u})}\left(\frac{\partial}{\partial\mathbf{x}_u}\right)^{\mathbf{j}_{|u|}}\phi_{\mathbf{X}_u}(\mathbf{x}_u;\mathbf{\Sigma}_{\mathbf{X}_u}),\ \mathbf{j}_{|u|}\in\mathbb{N}_0^{|u|}, \quad (28)$$

where $(\partial/\partial\mathbf{x}_u)^{\mathbf{j}_{|u|}}:=\partial^{j_1+\cdots+j_{|u|}}/\partial x_{i_1}^{j_1}\cdots\partial x_{i_{|u|}}^{j_{|u|}}$. Given the inner product $(y, z)_{\phi_{\mathbf{X}}}:=\int_{\mathbb{R}^N}yz\phi_{\mathbf{X}}(\mathbf{x};\mathbf{\Sigma}_{\mathbf{X}})d\mathbf{x}$, they eventually form a set of multivariate Hermite orthonormal polynomials

$$\left\{\psi_{u\mathbf{j}_{|u|}}:=\tilde{\psi}_{u\mathbf{j}_{|u|}}/\sqrt{(\tilde{\psi}_{u\mathbf{j}_{|u|}},\tilde{\psi}_{u\mathbf{j}_{|u|}})_{\phi_{\mathbf{X}}}},\ \mathbf{j}_{|u|}\in\mathbb{N}_0^{|u|}\right\} \quad (29)$$

that are consistent with the probability measure of $\mathbf{X}_u$. From these orthonormal polynomials and the function y, the integrals $I_{u\mathbf{j}_{|u|}}$ and $J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}}$ were exactly calculated from their definitions in Eqs. (19c) and (19d). Using $S=2$, $m=2$ in Eq. (21) and these two sets of integrals, a system of linear equations was generated and then solved to determine exactly the expansion coefficients $C_{u\mathbf{j}_{|u|}}$ for $\varnothing\neq u\subseteq\{1, 2, 3\}$ and $\|\mathbf{j}_{|u|}\|_1\leq m$. Since y is a sum of at most bivariate, second-order polynomials, the selection of $S=2$ and $m=2$ is adequate to reproduce y from the PDD approximation. Consequently, the sensitivity estimates generated from the original function y and its bivariate, second-order PDD approximation are identical, that is, $\bar{H}_{\{i\},f}^{(L)} = \tilde{H}_{\{i\},f}^{(L,2,2)}$.

For mixed correlation (Case 3), the bivariate, second-order ($S=2$, $m=2$) PDD approximation and the corresponding surrogate data set $\{\mathbf{x}^{(l)}, \tilde{\xi}_{S,m}^{(l)}\}_{l=1,\ldots,L}$ were employed in Eqs. (22a) and (22b) to obtain the KDE-generated densities $\bar{f}_{\tilde{Y}_{S,m}}(\xi)$ of $\tilde{Y}_{S,m}$ and $\bar{f}_{X_1,\tilde{Y}_{S,m}}(x_1, \xi)$ of $(X_1, \tilde{Y}_{S,m})$. Fig. 1(a) presents the marginal densities obtained for three progressively larger sample sizes: $L = 10^4$, $L = 10^6$, and $L = 10^8$. The joint densities, also produced for the same three sample sizes, are depicted in Figs. 1(b)–(d), respectively. The joint densities involving the two other input variables are similar and are, therefore, not shown or discussed. When the sample size increases, all KDE-generated densities, whether marginal or joint, converge as expected. In fact, the probability densities obtained for $L = 10^6$ and $L = 10^8$ are virtually indistinguishable to the naked eye.

Given a sample size L of a surrogate data set, define for the input variable X_i the $\mathcal{L}_1$ error

$$e_i^{(L,S,m)}:=\frac{\left|\tilde{H}_{\{i\},f}^{(L,S,m)}-\tilde{H}_{\{i\},f}^{(L',S,m)}\right|}{\tilde{H}_{\{i\},f}^{(L,S,m)}},\quad L'=L/10, \quad (30)$$

in estimating the univariate f-sensitivity index by the S-variate, mth-order PDD approximation and KDE, that is, by using Eq. (23). Since there exists no exact solution of $H_{\{i\},f}$, the estimated PDD solution obtained for the larger sample size was used as the reference solution. Again, for mixed correlation (Case 3), Figs. 2(a)–(c) display the $\mathcal{L}_1$ errors $e_i^{(L,S,m)}$ committed by the bivariate, second-order PDD approximation in estimating the univariate f-sensitivity indices $H_{\{i\},TV}$, $H_{\{i\},KL}$, and $H_{\{i\},H}$, respectively, for all three variables. The errors emanating from the PDD approximation drop with the sample size regardless of the variant of sensitivity indices examined.

Table 2 presents the PDD estimates $\tilde{H}_{\{i\},TV}^{(L,S,m)}$, $\tilde{H}_{\{i\},KL}^{(L,S,m)}$, and $\tilde{H}_{\{i\},H}^{(L,S,m)}$ of the univariate sensitivity indices of Y for three input variables,

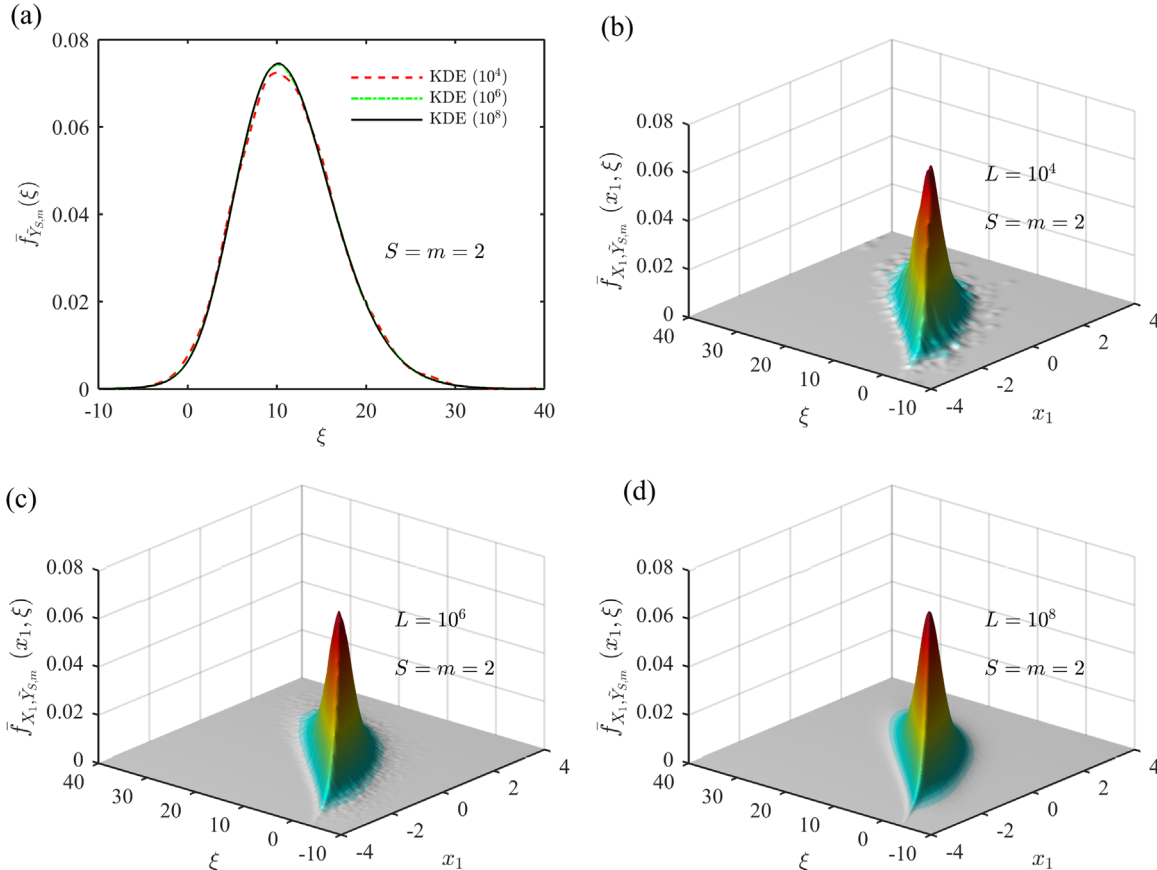Fig. 1. Various KDE-generated probability density functions in Example 1 using the bivariate, second-order $(S = m = 2)$ PDD method for mixed correlation; (a) marginal densities of $\tilde{Y}_{S,m}$ for $L = 10^4$, 10^6, and 10^8; (b) joint density of $(X_1, \tilde{Y}_{S,m})$ for $L = 10^4$; (c) joint density of $(X_1, \tilde{Y}_{S,m})$ for $L = 10^6$; (d) joint density of $(X_1, \tilde{Y}_{S,m})$ for $L = 10^8$.

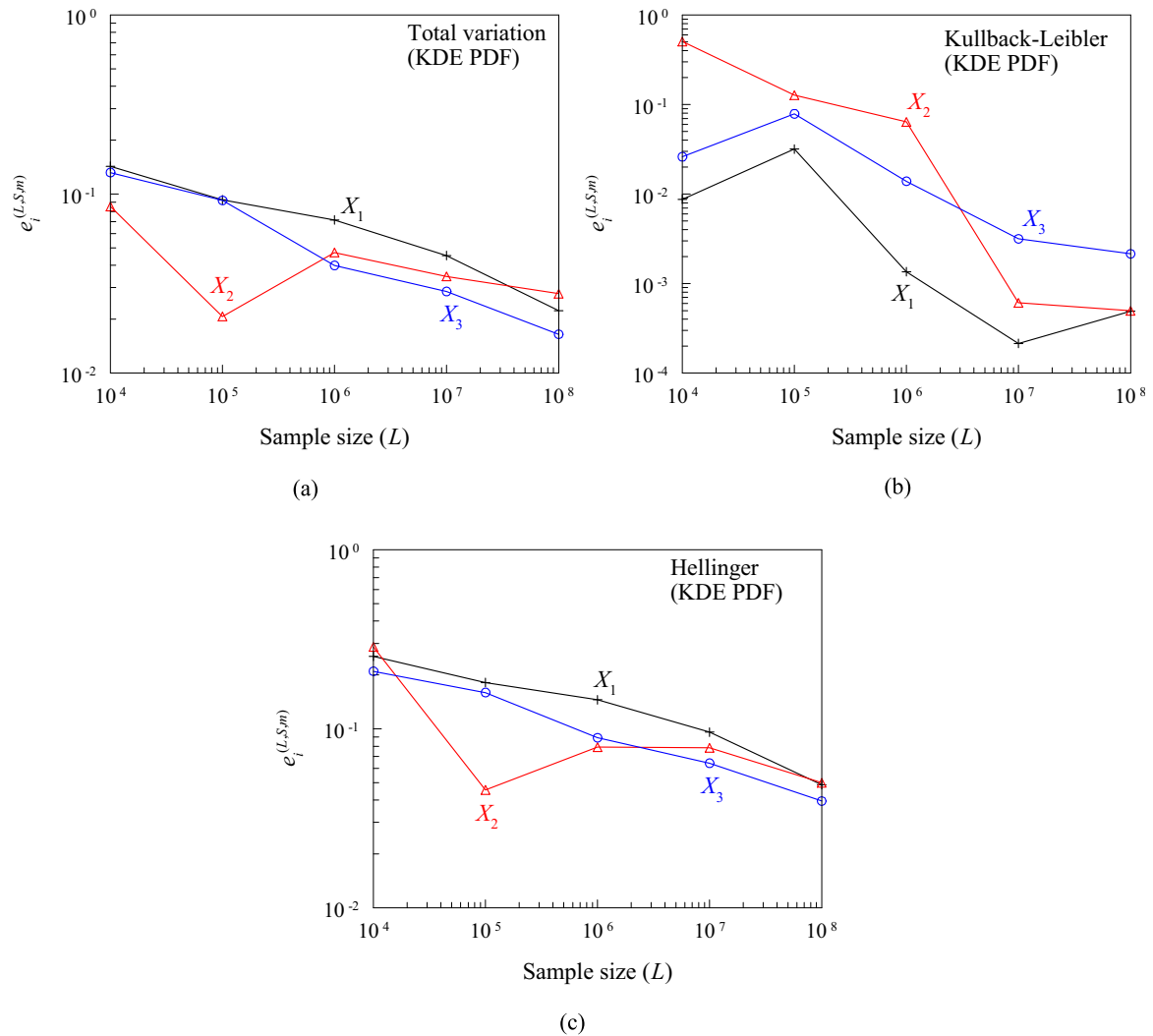**Fig. 2.** $\mathcal{L}_1$ errors committed by the bivariate, second-order ($S = m = 2$) PDD method in estimating three variants of sensitivity indices in Example 1 for mixed correlation; (a) total variational sensitivity index $H_{\{i\},TV}$; (b) Kullback–Leibler sensitivity index $H_{\{i\},KL}$; (c) Hellinger sensitivity index $H_{\{i\},H}$.

obtained using $S=2$, $m=2$, and $L = 10^8$. The results are tabulated for all three cases of correlation properties. The parenthetical numbers indicate relative rankings of all three variables, except when there is a tie. For identical correlation structures, such as the equally correlated case, all three variables are equally important, forming a three-way tie, as y is a symmetric function. For the positively correlated case, where the correlation coefficient increases monotonically from 1/5 to 4/5, X_1 and X_3 are the least and the most important variables, respectively, while the significance of X_2 is intermediary. The order of ranking should reverse if the correlation coefficient decreases monotonically. When the

correlation coefficients are both positive and negative, that is, for the mixedly correlated case, X_1 and X_2 become the most and the least important variables, respectively. Clearly, the correlation structure heavily influences global sensitivity analysis and, therefore, rankings of input variables.

4.2. Example 2

The next example originates from Bayesian sensitivity analysis involving a mixture of trigonometric and quadratic polynomial functions, described by [39]

Table 2

Bivariate, second-order ($S = m = 2$) PDD estimates of the total variational, Kullback–Leibler, and Hellinger sensitivity indices in Example 1 with $L = 10^8$.

Random variable	Case 1: equal correlation ($\rho_{12} = \rho_{13} = \rho_{23} = 1/5$)			Case 2: positive correlation ($\rho_{12} = 1/5, \rho_{13} = 2/5, \rho_{23} = 4/5$)			Case 3: mixed correlation ($\rho_{12} = -1/5, \rho_{13} = 2/5, \rho_{23} = -4/5$)		
	$\tilde{H}_{\{i\},TV}^{(L,S,m)}$	$\tilde{H}_{\{i\},KL}^{(L,S,m)}$	$\tilde{H}_{\{i\},H}^{(L,S,m)}$	$\tilde{H}_{\{i\},TV}^{(L,S,m)}$	$\tilde{H}_{\{i\},KL}^{(L,S,m)}$	$\tilde{H}_{\{i\},H}^{(L,S,m)}$	$\tilde{H}_{\{i\},TV}^{(L,S,m)}$	$\tilde{H}_{\{i\},KL}^{(L,S,m)}$	$\tilde{H}_{\{i\},H}^{(L,S,m)}$
X_1	0.569 (1)	0.328 (1)	0.168 (1)	0.550 (3)	0.301 (3)	0.158 (3)	0.869 (1)	0.737 (1)	0.353 (1)
X_2	0.569 (1)	0.328 (1)	0.168 (1)	0.789 (2)	0.604 (2)	0.306 (2)	0.158 (3)	0.043 (3)	0.018 (3)
X_3	0.569 (1)	0.328 (1)	0.168 (1)	0.944 (1)	0.899(1)	0.426 (1)	0.347 (2)	0.150 (2)	0.071 (2)

$$y(\mathbf{X}) = \mathbf{a}_1^T \mathbf{X} + \mathbf{a}_2^T \sin \mathbf{X} + \mathbf{a}_3^T \cos \mathbf{X} + \mathbf{X}^T \mathbf{M} \mathbf{X}, \tag{31}$$

where $\mathbf{X} = (X_1, \ldots, X_{15})^T \in \mathbb{R}^{15}$ is a 15-dimensional standard Gaussian input vector ($N=15$) with mean vector $\mathbb{E}[\mathbf{X}] = \{0, \ldots, 0\}^T =: \mathbf{0} \in \mathbb{R}^{15}$ and covariance matrix $\mathbb{E}[\mathbf{X}\mathbf{X}^T] = \mathrm{diag}[1, \ldots, 1] =: \mathbf{I} \in \mathbb{R}^{15 \times 15}$; $\sin \mathbf{X} := (\sin X_1, \ldots, \sin X_{15})^T \in \mathbb{R}^{15}$; $\cos \mathbf{X} := (\cos X_1, \ldots, \cos X_{15})^T \in \mathbb{R}^{15}$; and $\mathbf{a}_i \in \mathbb{R}^{15}$, $i = 1, 2, 3$, and $\mathbf{M} \in \mathbb{R}^{15 \times 15}$ are coefficient vectors and matrix, respectively, obtained from Oakley and O'Hagan's paper [39]. The coefficient vectors were selected so that the first five variables are unimportant, the next five variables are slightly important, and the last five variables are very important. The coefficient matrix reflects interactive effects from all pairs of random variables. The function y in Eq. (31) has been used for variance-based global sensitivity analysis [39,40].

Although $y: \mathbb{R}^{15} \to \mathbb{R}$, each term on the right side of Eq. (31) includes at most bivariate combinations of random variables. Therefore, a bivariate PDD approximation ($S=2$) is adequate and was employed to perform sensitivity analysis. The expansion coefficients were estimated using Hermite polynomials of largest orders $2 \le m \le 6$ and dimension-reduction numerical integration (Gauss–Hermite quadrature rule) with $R = S = 2$ and $n = m + 1$. Two bivariate PDD approximations with two distinct values of $m=2$ and $m=6$ and the associated surrogate data sets ($L = 10^6$) were employed to calculate by KDE the following PDD-generated densities: two marginal densities $\bar{f}_{\tilde{Y}_{S,m}}(\xi)$ of $\tilde{Y}_{S,m}$ and two joint densities $\bar{f}_{X_1, \tilde{Y}_{S,m}}(x_1, \xi)$ of $(X_1, \tilde{Y}_{S,m})$. For comparisons, both the marginal density $\bar{f}_Y(\xi)$ of Y and the joint density $\bar{f}_{X_1, Y}(x_1, \xi)$ of (X_1, Y) were also computed by KDE using the original function y and the corresponding data set ($L = 10^6$). Fig. 3(a) contrasts the two PDD-generated marginal densities with the density generated from the original function y. Similarly, Figs. 3(b) and (c) portray the two PDD-generated joint densities, whereas Fig. 3(d) depicts the joint density of (X_1, Y) obtained using the original function y. For larger m, that is, for $m=6$, all KDE-generated marginal and joint densities, whether obtained from the PDD approximation or the original function, are practically coincident.

Table 3 presents the estimates $\tilde{H}_{\{i\},TV}^{(L,S,m)}$, $\tilde{H}_{\{i\},KL}^{(L,S,m)}$, and $\tilde{H}_{\{i\},H}^{(L,S,m)}$ of the univariate sensitivity indices of Y for all 15 variables using the bivariate PDD method ($S=2$) for three polynomial orders, $m=2$, $m=4$, and $m=6$, including the requisite numbers of original function evaluations. For comparison, the estimates of respective sensitivity indices determined from the original function y, that is, $\tilde{H}_{\{i\},TV}^{(L)}$, $\tilde{H}_{\{i\},KL}^{(L)}$, and $\tilde{H}_{\{i\},H}^{(L)}$, are also enumerated in Table 3. In all cases, the sample size $L = 10^6$. When $m=2$, the lowest polynomial order selected, there exist some discrepancies between the indices estimated by the PDD and the original function, although the relative importance of input variables is clearly recognized already. When m increases, the accuracy of the estimated indices improves rapidly, however, with increased computational demand. When $m=6$, the largest polynomial order considered, the PDD-estimated indices are nearly identical to those generated from the original function. The computational efforts by the PDD method, measured in terms of the number of original function evaluations, vary from 451 to 3871, depending on the values of m chosen; nonetheless, they are markedly lower than the 10^6 function evaluations required in generating the benchmark results of Table 3. Therefore, for sensitivity analysis of complex systems, where y is expensive to evaluate, generating sensitivity indices solely from the original
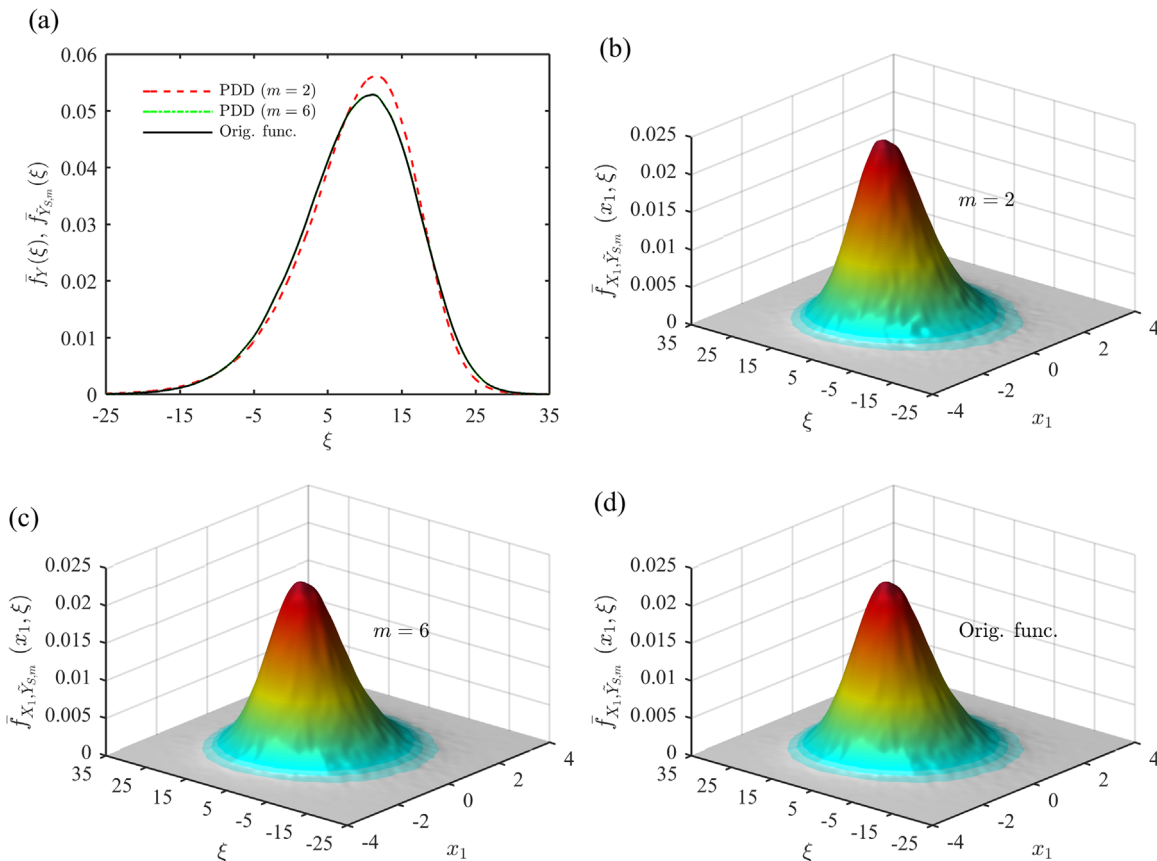
Fig. 3. Various KDE-generated probability density functions in Example 2 with $L = 10^6$; (a) marginal densities using bivariate, second-order ($S=2$, $m=2$) and bivariate, sixth-order ($S=2$, $m=6$) PDD approximations, and the original function; (b) joint density using bivariate, second-order ($S=2$, $m=2$) PDD approximation; (c) joint density using bivariate, sixth-order ($S=2$, $m=6$) PDD approximation; (d) joint density using the original function.

Table 3
Estimates of the total variational, Kullback–Leibler, and Hellinger sensitivity indices in Example 2 using various bivariate PDD ($S=2$) approximations and original function with $L=10^6$.

Random variable	$\tilde{H}_{\{i\},TV}^{(L,S,m)}, \tilde{H}_{\{i\},TV}^{(L)}$ (Total variational sensitivity index)				$\tilde{H}_{\{i\},KL'}^{(L,S,m)}, \tilde{H}_{\{i\},KL'}^{(L)}$ (Kullback–Leibler sensitivity index)				$\tilde{H}_{\{i\},H}^{(L,S,m)}, \tilde{H}_{\{i\},H}^{(L)}$ (Hellinger sensitivity index)			
	PDD ($m=2$)	PDD ($m=4$)	PDD ($m=6$)	Orig. func.	PDD ($m=2$)	PDD ($m=4$)	PDD ($m=6$)	Orig. func.	PDD ($m=2$)	PDD ($m=4$)	PDD ($m=6$)	Orig. func.
X_1	0.0643	0.0609	0.0607	0.0606	0.0103	0.0098	0.0097	0.0097	0.0028	0.0025	0.0025	0.0025
X_2	0.0442	0.0379	0.0377	0.0377	0.0078	0.0073	0.0073	0.0072	0.0017	0.0014	0.0014	0.0014
X_3	0.0681	0.0625	0.0621	0.0620	0.0132	0.0121	0.0120	0.0120	0.0039	0.0034	0.0034	0.0034
X_4	0.0608	0.0558	0.0556	0.0555	0.0086	0.0080	0.0080	0.0080	0.0021	0.0018	0.0018	0.0018
X_5	0.0550	0.0491	0.0489	0.0487	0.0070	0.0067	0.0067	0.0067	0.0015	0.0013	0.0013	0.0013
X_6	0.1131	0.1146	0.1151	0.1151	0.0190	0.0183	0.0182	0.0182	0.0071	0.0068	0.0067	0.0067
X_7	0.1142	0.1143	0.1147	0.1147	0.0201	0.0191	0.0191	0.0191	0.0075	0.0071	0.0071	0.0071
X_8	0.1236	0.1251	0.1256	0.1257	0.0222	0.0216	0.0215	0.0215	0.0083	0.0081	0.0080	0.0080
X_9	0.1371	0.1366	0.1371	0.1372	0.0289	0.0278	0.0278	0.0278	0.0114	0.0110	0.0110	0.0110
X_{10}	0.0870	0.0884	0.0889	0.0889	0.0127	0.0127	0.0127	0.0127	0.0042	0.0041	0.0041	0.0041
X_{11}	0.2249	0.2323	0.2335	0.2337	0.0639	0.0622	0.0621	0.0621	0.0280	0.0280	0.0279	0.0279
X_{12}	0.2431	0.2604	0.2630	0.2633	0.0741	0.0769	0.0772	0.0772	0.0329	0.0354	0.0356	0.0356
X_{13}	0.2154	0.2299	0.2318	0.2320	0.0608	0.0618	0.0618	0.0618	0.0265	0.0278	0.0279	0.0279
X_{14}	0.2127	0.2261	0.2283	0.2286	0.0571	0.0591	0.0593	0.0593	0.0248	0.0264	0.0265	0.0265
X_{15}	0.2382	0.2521	0.2543	0.2545	0.0703	0.0722	0.0722	0.0722	0.0311	0.0329	0.0330	0.0329
No. of func. eval.[a]	451	1741	3871	10^6	451	1741	3871	10^6	451	1741	3871	10^6

[a] The number of function evaluations for a bivariate, mth-order PDD approximation is $N(N-1)(n-1)^2/2 + N(n-1) + 1 = 105m^2 + 15m + 1$, where $N=15$, $n = m + 1$.

function will not be practical, and an approach invoking the PDD method becomes necessary; it will be demonstrated next.

4.3. Example 3

The final example illustrates the PDD method for sensitivity analysis of an industrial-scale, solid-mechanics problem. It involves FEA of a leverarm in a wheel loader, depicted in Fig. 4(a), commonly used in the heavy construction industry. The loading and boundary conditions of a single leverarm are shown in Fig. 4(b). Figs. 4(c)–(d) present a computer-aided design model and a finite-element mesh comprising 48,312 second-order tetrahedral elements of the leverarm, respectively. Two random loads P_H and P_V acting at pin E can be viewed as input loads due to other mechanical components of the wheel loader. The essential boundary conditions, sketched in Fig. 4(b), define random prescribed displacements u_{xF} and u_{yF} at pin F and u_{xG}, and u_{yG} at pin G. The leverarm is made of cast steel with random Young's modulus E and random Poisson's ratio ν. The input vector $\mathbf{X} = (P_H, P_V, E, \nu, u_{xF}, u_{yF}, u_{xG}, u_{yG}) \in \mathbb{R}^8$ comprises eight independent random variables with their statistical properties specified in Table 4. The von Mises stress and maximum principal strain distributions in Figs. 4(e) and (f), respectively, calculated for an arbitrarily selected sample input, are commonly used for examining material yielding or fatigue damage in mechanical systems. Several PDD approximations with measure-consistent orthogonal polynomials and Gauss quadrature rules were employed for density-based sensitivity analysis of a strain response.

Unlike the previous two examples, the PDD truncation parameters S and m for this complex problem cannot be ascertained *a priori*. Therefore, a limited convergence study was performed by progressively increasing the values of S (e.g., $S=1$, 2) and m (e.g., $m=2$, 3) and then watching what happens to the final result. A more rigorous approach entails an adaptive truncation strategy, where higher-variate and/or higher-order PDD component functions are automatically drawn in by fulfilling a judicious set of pruning criteria [41]. No adaptive strategy was employed in this work.

Table 5 presents the univariate ($S=1$) and bivariate ($S=2$) PDD estimates $\tilde{H}_{\{i\},TV}^{(L,S,m)}$, $\tilde{H}_{\{i\},KL'}^{(L,S,m)}$, and $\tilde{H}_{\{i\},H}^{(L,S,m)}$ of univariate sensitivity indices of the maximum principal strain for all eight variables. The PDD expansion coefficients were estimated by S-variate dimension-reduction integration, requiring one- ($S=1$) or at most two-dimensional ($S=2$) Gauss quadratures. The order m of orthogonal polynomials and the number n of Gauss quadrature points in the dimension-reduction numerical integration are $2 \leq m \leq 3$ and $n = m + 1$, respectively. The indices are broken down according to the choice of selecting $S=1$, 2 and $m=2$, 3. In all PDD approximations, the sample size $L=10^6$. The sensitivity indices by the PDD method in Table 5 for each variant quickly converge with respect to S and/or m. Since FEA is employed for response evaluations, the computational effort of the PDD method comes primarily from numerically determining the PDD expansion coefficients. The expenses involved in estimating the PDD coefficients vary from 25 to 33 FEA for the univariate PDD approximation and from 277 to 481 FEA for the bivariate PDD approximation, depending on the two values of m. Based on the sensitivity indices in Table 5, the horizontal boundary conditions (u_{xF} and u_{xG}) are highly important; the vertical load (P_V), elastic modulus (E), and vertical boundary conditions (u_{yF} and u_{yG}) are slightly important; and the horizontal load (P_H) and Poisson's ratio (ν) are unimportant in influencing the maximum principal strain. All three variants of the sensitivity indices lead to the same conclusion.

Note that the respective univariate and bivariate PDD solutions in this particular problem are practically the same. Therefore, the univariate PDD solutions are not only accurate, but also highly efficient. This is because of a realistic example chosen, where the individual main effects of input variables on the maximum principal strain are dominant over their interactive effects. Finally, this example also demonstrates the non-intrusive nature of the PDD method, which can be easily integrated with external commercial codes for analyzing large-scale engineering systems.

5. Conclusion

A surrogate method, referred to as the PDD method, is proposed to estimate a general class of density-based f-sensitivity
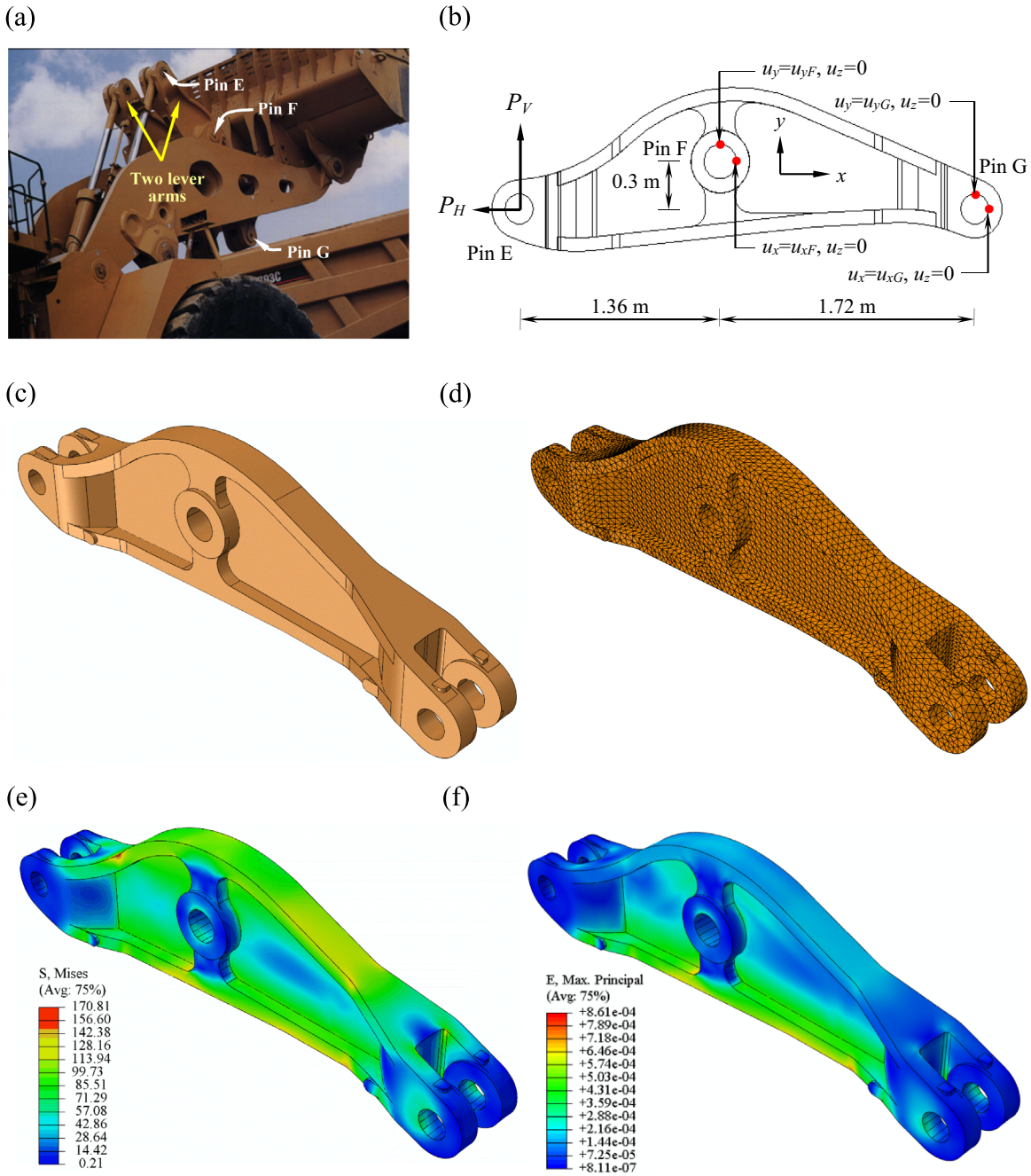
Fig. 4. FEA of a leverarm; (a) two leverarms in a wheel loader; (b) geometry and boundary conditions; (c) computer-aided design model; (d) finite-element mesh (48,312 elements); (e) von Mises stress contours for a sample input; (f) maximum principal strain contours for a sample input.

Table 4
Statistical properties of input variables in Example 3.

Random variable	Probability distribution	Mean	St. dev.
P_H[a], kN	Lognormal	507.69	76.15
P_V[a], kN	Lognormal	1517.32	227.60
E, GPa	Lognormal	203	10.15
ν	Lognormal	0.3	0.015
u_{xF}, mm	Uniform[b]	−5	$5/\sqrt{3}$
u_{yF}, mm	Uniform[c]	5	$5/\sqrt{3}$
u_{xG}, mm	Uniform[c]	5	$5/\sqrt{3}$
u_{yG}, mm	Uniform[b]	−5	$5/\sqrt{3}$

[a] To be distributed equally (halved) on the front and back sides of pin E.
[b] Uniformly distributed over [−10, 0] mm; to be applied on both sides.
[c] Uniformly distributed over [0, 10] mm; to be applied on both sides.

indices. The method is based on PDD approximation of a high-dimensional stochastic response, forming a surrogate input–output data set; kernel density estimations of output probability density functions from the surrogate data set; and subsequent Monte Carlo integration for estimating the f-sensitivity index. Developed for an arbitrary convex function f, the method is capable of calculating a wide variety of sensitivity or importance measures, including the mutual information, squared-loss mutual information, and $\mathcal{L}_1$-distance-based importance measure. In addition, the method can account for dependent as well as independent probability distributions of input random variables. The PDD method, predicated on an efficient surrogate approximation, becomes increasingly relevant when analyzing high-dimensional complex systems, demanding expensive function evaluations. Therefore, the computational burden of existing methods that

Table 5

Estimates of the total variational, Kullback–Leibler, and Hellinger sensitivity indices in Example 3 using various PDD approximations with $L = 10^6$.

Random variable	$\bar{H}_{\{i\},TV}^{(L,S,m)}$ (Total variational sensitivity index)				$\bar{H}_{\{i\},KL}^{(L,S,m)}$ (Kullback–Leibler sensitivity index)				$\bar{H}_{\{i\},H}^{(L,S,m)}$ (Hellinger sensitivity index)			
	Univariate PDD ($S=1$)[a]		Bivariate PDD ($S=2$)[b]		Univariate PDD ($S=1$)[a]		Bivariate PDD ($S=2$)[b]		Univariate PDD ($S=1$)[a]		Bivariate PDD ($S=2$)[b]	
	$m=2$	$m=3$	$m=2$	$m=3$	$m=2$	$m=3$	$m=2$	$m=3$	$m=2$	$m=3$	$m=2$	$m=3$
P_H	0.0204	0.0204	0.0204	0.0204	0.0046	0.0046	0.0046	0.0046	0.0003	0.0003	0.0003	0.0003
P_V	0.0856	0.0856	0.0843	0.0843	0.0173	0.0172	0.0166	0.0166	0.0056	0.0056	0.0053	0.0053
E	0.0333	0.0332	0.0328	0.0328	0.0060	0.0060	0.0059	0.0059	0.0009	0.0009	0.0009	0.0009
ν	0.0181	0.0181	0.0177	0.0178	0.0044	0.0044	0.0044	0.0044	0.0002	0.0002	0.0002	0.0002
u_{xF}	0.4251	0.4249	0.4231	0.4224	0.3604	0.3604	0.3622	0.3613	0.1120	0.1118	0.1108	0.1101
u_{yF}	0.0845	0.0845	0.0846	0.0847	0.0069	0.0069	0.0070	0.0070	0.0046	0.0046	0.0046	0.0046
u_{xG}	0.4249	0.4250	0.4233	0.4236	0.3597	0.3597	0.3614	0.3606	0.1117	0.1118	0.1110	0.1110
u_{yG}	0.0859	0.0859	0.0860	0.0862	0.0070	0.0070	0.0071	0.0071	0.0046	0.0046	0.0046	0.0046
No. of FEA	25	33	277	481	25	33	277	481	25	33	277	481

[a] The number of function evaluations for a univariate, mth-order PDD approximation is $Nn + 1 = 8(m + 1) + 1$, where $N=8$, $n = m + 1$.

[b] The number of function evaluations for a bivariate, mth-order PDD approximation is $N(N - 1)n^2/2 + Nn + 1 = 28(m + 1)^2 + 8(m + 1) + 1$, where $N=8$, $n = m + 1$.

mandate original function evaluations can be significantly deflated by the PDD method. Numerical results illuminate the accuracy, computational efficiency, and convergence properties of the proposed method in computing sensitivity indices derived from three prominent divergence or distance measures. Finally, a successful evaluation of an FEA-based sensitivity analysis of a leverarm demonstrates the ability of the proposed method in solving large-scale practical engineering problems.

References

[1] Sobol IM. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Math Comput Simul 2001;55:271–80.

[2] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis: the primer.West Sussex: Wiley; 2008.

[3] Rahman SA. Generalized ANOVA dimensional decomposition for dependent probability measures. SIAM/ASA J Uncertain Quantif 2014;2:670–97.

[4] Sudret B. Global sensitivity analysis using polynomial chaos expansions. Reliab Eng Syst Saf 2008;93:964–79.

[5] Borgonovo E. A new uncertainty importance measure. Reliab Eng Syst Saf 2007;92:771–84.

[6] Rahman S. The f-Sensitivity index. SIAM/ASA J Uncertain Quantif 2016;4:130–62.

[7] Borgonovo E, Castaings W. Tarantola, model emulation and moment-independent sensitivity analysis: an application to environmental modeling. Environ Model Softw 2012;34:105–15.

[8] Ratto M, Pagano A. Using recursive algorithms for the efficient identification of smoothing spline ANOVA models. Adv Stat Anal 2010;94:367–88.

[9] Borgonovo E, Hazen G, Plischke E. A common rationale for global sensitivity measures and their estimation. Risk Anal 2016:1–24. http://dx.doi.org/10.1111/risa.12555.

[10] Cover TM, Thomas JA. Elements of information theory.New York, NY: John Wiley and Sons, Inc; 1991.

[11] Suzuki T, Sugiyama M, Kanamori T, Sese, J. Mutual information estimation reveals global associations between stimuli and biological processes. BMC Bioinform 2009;10:S52:1–12.

[12] Liu H, Chen W, Sudjianto A. Relative entropy based method for probabilistic sensitivity analysis in engineering design. ASME J Mech Des 2006;128: p. 326–36.

[13] Greegar G, Manohar CS. Global response sensitivity analysis using probability distance measures and generalization of Sobol's analysis. Probab Eng Mech 2015;41:21–33.

[14] Greegar G, Manohar CS. Global response sensitivity analysis of uncertain structures. Struct Saf 2016;58:94–104.

[15] Abhinav S, Manohar CS. Global response sensitivity analysis of randomly excited dynamic structures. ASCE J Eng Mech 2015;142:04015094-1–10.

[16] Liu Q, Homma T. A new computational method of a moment-independent uncertainty importance measure. Reliab Eng Syst Saf 2009;94:1205–11.

[17] Wei P, Lu Z, Yuan X. Monte Carlo simulation for moment-independent sensitivity analysis. Reliab Eng Syst Saf 2013;110:60–7.

[18] Csiszár I. Eine informationstheoretische ungleichung und ihre anwendung auf den beweis der ergodizitat von markoffschen ketten. Magyar Tud Akad Mat Kutato Int Kozl 1963;8:85–108.

[19] Csiszár I. Information-type measures of difference of probability distributions and indirect observation. Stud Sci Math Hung 1967;2:299–318.

[20] Ali SM, Silvey SD. A general class of coefficients of divergence of one distribution from another. J R Stat Soc—Ser B: Methodol 1966;28:131–42.

[21] Morimoto T. Markov processes and the H-theorem. J Phys Soc Jpn 1963;18:328–31.

[22] Kullback S, Leibler AS. On information and sufficiency. Ann Math Stat 1951;22:79–86.

[23] Kolmogorov AN. A new invariant for transitive dynamical system. Dokl Akad Nauk USSR 1957;119:861–9.

[24] Hellinger E. Neue begründung der theorie quadratischen formen von unendlichen vielen veränderlichen. J Reine Angew Math 1909;136:210–71.

[25] Pearson K. On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonable supposed to have arisen from random sampling. Phil Mag 1900;50:157–72.

[26] Dragomir SS. Inequalities for Csiszár f-Divergence in information theory. Edited monograph, Victoria University, Melbourne, Australia; 2000. On line: ⟨http://rgmia.vu.edu.au⟩.

[27] Vajda I. On f-divergence and singularity of probability measures. Period Math Hung 1972;2:223–34.

[28] Jeffreys H. An invariant form for the prior probability in estimating problems. Proc R Soc Lond 1946;186A:453–61.

[29] Topsoe F. Some inequalities for information divergence and related measures of discrimination. Res Rep Coll RGMIA 1999;2:85–98.

[30] Wiener N. The homogeneous chaos. Am J Math 1938;60:897–936.

[31] Babuška I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J Numer Anal 2007;45:1005–34.

[32] Rahman S. Extended polynomial dimensional decomposition for arbitrary probability distributions. J Eng Mech 2009;135(12):1439–51.

[33] Rahman S. A polynomial dimensional decomposition for stochastic computing. Int J Numer Methods Eng 2008;76:2091–116.

[34] Rahman S, Yadav V. Orthogonal polynomial expansions for solving random eigenvalue problems. Int J Uncertain Quantif 2011;1:163–87.

[35] Xu H, Rahman S. A generalized dimension-reduction method for multi-dimensional integration in stochastic mechanics. Int J Numer Methods Eng 2004;61:1992–2019.

[36] Rahman S. Approximation errors in truncated dimensional decompositions. Math Comput 2014;83(290):2799–819.

[37] Rosenblatt M. Remarks on some non-parametric estimates of a density function. Ann Math Stat 1956;27:832–7.

[38] Parzen E. On estimation of a probability density function and mode. Ann Math Stat 1962;33:1065–76.

[39] Oakley J, O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. J R Stat Soc: Ser B 2004;66:751–69.

[40] Rahman S. Global sensitivity analysis by polynomial dimensional decomposition. Reliab Eng Syst Saf 2011;96(7):825–37.

[41] Yadav V, Rahman S. Adaptive-sparse polynomial dimensional decomposition for high-dimensional stochastic computing. Comput Methods Appl Mech Eng 2014;274:56–83.