SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. 2, pp. 670–697
© 2014 Sharif Rahman

A Generalized ANOVA Dimensional Decomposition for Dependent Probability Measures*

Sharif Rahman†

Abstract. This article explores the generalized analysis-of-variance, or ANOVA, dimensional decomposition (ADD) for multivariate functions of dependent random variables. Two notable properties, stemming from weakened annihilating conditions, reveal that the component functions of the generalized ADD have *zero* means and are hierarchically orthogonal. By exploiting these properties, a simple alternative approach is presented to derive a coupled system of equations that the generalized ADD component functions satisfy. The coupled equations, which subsume as a special case the classical ADD, reproduce the component functions for independent probability measures. To determine the component functions of the generalized ADD, a new constructive method is proposed by employing measure-consistent, multivariate orthogonal polynomials as bases and calculating the expansion coefficients involved from the solution of linear algebraic equations. New generalized formulae are presented for the second-moment characteristics, including triplets of global sensitivity indices, for dependent probability distributions. Furthermore, the generalized ADD leads to extended definitions of effective dimensions, reported in the current literature for the classical ADD. Numerical results demonstrate that the correlation structure of random variables can significantly alter the composition of component functions, producing widely varying global sensitivity indices and, therefore, distinct rankings of random variables. An application to random eigenvalue analysis demonstrates the usefulness of the proposed approximation.

Key words. ADD, ANOVA, dimensional decomposition, multivariate orthogonal polynomials, uncertainty quantification, global sensitivity analysis

AMS subject classifications. 26B49, 41A61, 49K30, 60H35, 65C60

DOI. 10.1137/120904378

1. Introduction.

Uncertainty quantification of complex systems, whether natural or engineered, is a crucial ingredient in numerous fields of engineering, science, and medicine. The remarkable growth of computing power, complemented by matching gains in algorithmic speed and accuracy, has led to near-ubiquity of computational methods for estimating the statistical moments, probability laws, and other relevant properties of such systems. However, most existing methods [4, 10, 32], while successful in tackling small to moderate numbers of random variables, begin to break down for truly high-dimensional problems. The root deterrence to practical computability is often related to high dimension of the multivariate integration or interpolation problem, known as the curse of dimensionality [1]. The dimensional decomposition of a multivariate function [9, 27, 18, 13] addresses the curse of dimensionality to some extent by developing an input-output behavior of complex systems with low effective

*Received by the editors January 2, 2013; accepted for publication (in revised form) August 4, 2014; published electronically November 13, 2014. This work was supported by the U.S. National Science Foundation under grants CMMI-0969044 and CMMI-1130147.

http://www.siam.org/journals/juq/2/90437.html

†Applied Mathematics & Computational Sciences, The University of Iowa, Iowa City, IA 52242 (sharif-rahman@uiowa.edu).

dimensions [2], wherein the degrees of interactions between input variables attenuate rapidly or vanish altogether.

A well-known prototype of dimensional decomposition is the analysis-of-variance, or ANOVA, dimensional decomposition (ADD), first presented by Hoeffding in the 1940s in relation to his seminal work on U-statistics [9]. Since then, ADD has been studied by numerous researchers in disparate fields of mathematics [24, 8], statistics [16, 6], finance [7], and basic and applied sciences [17], including engineering disciplines, where its polynomial versions have been successfully applied for uncertainty quantification of high-dimensional complex systems [21, 20, 35]. However, the existing ADD, referred to as the classical ADD in this paper, is strictly valid for independent, product-type probability measures of random input. In reality, there may exist significant correlation or dependence among input variables. The author rules out the Rosenblatt transformation [23] or others commonly used for mapping dependent to independent variables, as they may induce overly large nonlinearity to a stochastic response, potentially degrading the convergence properties of probabilistic solutions [20]. Therefore, the classical ADD must be generalized for an arbitrary, non-product-type probability measure. Doing so will require modifying the original annihilating conditions that will endow desirable orthogonal properties, insofar as is possible, to the generalization. Indeed, inspired by Stone [28] and employing a set of weakened annihilating conditions, Hooker [11] provided an existential proof of a unique ANOVA decomposition for dependent variables, referred to as the generalized ADD in this paper, subject to a mild restriction on the probability measure. Furthermore, he introduced a computational method for determining the component functions of the generalized ADD by minimizing a mean-squared error with weakened annihilating conditions as constraints. However, the method turns out to be computationally demanding and even potentially prohibitive when there exist a moderate number of variables. Chastaing, Gamboa, and Prieur [3] presented a boundedness assumption on the joint probability density function of random variables for the availability of a generalized ANOVA decomposition. Li and Rabitz [14] proposed combining hierarchically selected univariate orthogonal polynomials and regression, applied to input-output data, to approximate the component functions of the generalized ADD.

The purpose of this paper is threefold. First, a brief exposition of the classical ADD is given in section 3, setting the stage for the generalized ADD presented in section 4. Two propositions and a theorem, proven herein, reveal two special properties of the generalized ADD, leading to a coupled system of equations satisfied by the component functions. These theoretical results, which subsume the classical ADD as a special case, are shown to reproduce the component functions for independent probability measures. Second, section 5 introduces general multivariate orthogonal polynomials that are consistent with the probability measures of dependent input variables. A theorem and its proof, presented in this section, describe a new constructive method for finding the component functions of the generalized ADD in terms of measure-consistent, multivariate orthogonal polynomials. Third, the second-moment analysis of the generalized ADD is described in section 6. It entails global sensitivity analysis, including triplets of sensitivity indices, for dependent probability distributions. Using insights from the generalized ADD, extended definitions of two effective dimensions are proposed. Numerical results, including approximate solutions of a random eigenvalue problem, are reported in

section 7 to affirm the theoretical findings. Mathematical notation and conclusions are defined or drawn in sections 2 and 8, respectively.

2. Notation. Let $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{R}$, and $\mathbb{R}_0^+$ represent the sets of positive integer (natural), nonnegative integer, real, and nonnegative real numbers, respectively. For $k \in \mathbb{N}$, denote by $\mathbb{R}^k$ the k-dimensional Euclidean space and by $\mathbb{R}^{k\times k}$ the set of $k \times k$ real-valued matrices. These standard notations will be used throughout the paper.

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 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 high-dimensional stochastic problem. The probability law of $\mathbf{X}$, assumed to be continuous, is completely defined by its joint probability density function $f_{\mathbf{X}} : \mathbb{R}^N \to \mathbb{R}_0^+$. Let u be a subset of $\{1, \ldots, N\}$ with the complementary set $-u := \{1, \ldots, N\}\backslash u$ and cardinality $0 \le |u| \le N$, and let $\mathbf{X}_u = (X_{i_1}, \ldots, X_{i_{|u|}})$, $u \neq \emptyset$, $1 \le i_1 < \cdots < i_{|u|} \le N$, be a subvector of $\mathbf{X}$ with $\mathbf{X}_{-u} := \mathbf{X}_{\{1,\ldots,N\}\backslash u}$ defining its complementary subvector. Then, for a given $\emptyset \neq u \subseteq \{1, \ldots, N\}$, the marginal density function of $\mathbf{X}_u$ is $f_u(\mathbf{x}_u) := \int_{\mathbb{R}^{N-|u|}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}_{-u}$.

Let $y(\mathbf{X}) := y(X_1, \ldots, X_N)$, a real-valued, measurable transformation on $(\Omega, \mathcal{F})$, define a high-dimensional stochastic response of interest, and let $\mathcal{L}_2(\Omega, \mathcal{F}, P)$ represent a Hilbert space of square-integrable functions y with respect to the induced generic measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$ supported on $\mathbb{R}^N$. Although it is well known in the current literature, section 3 briefly describes the classical ADD, so that it can be contrasted with the generalized ADD presented in section 4, the main theme of this paper.

3. Classical ANOVA dimensional decomposition. The classical ADD is established by assuming independent coordinates of $\mathbf{X}$ and selecting a product-type probability measure, $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = \Pi_{i=1}^N f_{\{i\}}(x_i)dx_i$, of $\mathbf{X}$, where $f_{\{i\}} : \mathbb{R} \to \mathbb{R}_0^+$ is the marginal probability density function of X_i, defined on the probability triple $(\Omega_i, \mathcal{F}_i, P_i)$ with a bounded or an unbounded support on $\mathbb{R}$. The ADD, expressed by the compact form [24, 18, 13]

$$y(\mathbf{X}) = \sum_{u\subseteq\{1,\ldots,N\}} y_{u,C}(\mathbf{X}_u), \tag{3.1}$$

is a finite, hierarchical expansion in terms of its input variables with increasing dimensions, where $y_{u,C}$ is a $|u|$-variate component function describing a constant or the interactive effect of $\mathbf{X}_u$ on y when $|u| = 0$ or $|u| > 0$. The symbol C in the subscript of $y_{u,C}$ is a reminder that the component functions belong to the classical ADD. The summation in (3.1) comprises 2^N component functions, with each function depending on a group of variables indexed by a particular subset of $\{1, \ldots, N\}$, including the empty set $\emptyset$. Applying strong annihilating conditions, the component functions are endowed with desirable orthogonal properties, explained as follows.

3.1. Strong annihilating conditions. The strong annihilating conditions relevant to the classical ADD require all nonconstant component functions $y_{u,C}$ to integrate to *zero* with respect to the marginal density of each random variable in u, that is [26, 17, 18, 13],

$$\int_{\mathbb{R}} y_{u,C}(\mathbf{x}_u) f_{\{i\}}(x_i)dx_i = 0 \quad \text{for } i \in u \neq \emptyset, \tag{3.2}$$

resulting in two remarkable properties, described by Propositions 3.1 and 3.2.

Proposition 3.1. *The classical ADD component functions $y_{u,C}$, where $\emptyset \neq u \subseteq \{1,\dots,N\}$, have zero means, i.e.,*

$$\mathbb{E}\left[y_{u,C}(\mathbf{X}_u)\right] = 0.$$

Proposition 3.2. *Two distinct classical ADD component functions $y_{u,C}$ and $y_{v,C}$, where $\emptyset \neq u \subseteq \{1,\dots,N\}$, $\emptyset \neq v \subseteq \{1,\dots,N\}$, and $u \neq v$, are orthogonal; i.e., they satisfy the property*

$$\mathbb{E}\left[y_{u,C}(\mathbf{X}_u)y_{v,C}(\mathbf{X}_v)\right] = 0.$$

Integrating (3.1) with respect to the measure $f_{-u}(\mathbf{x}_{-u})d\mathbf{x}_{-u} = \prod_{i=1,i\notin u}^{N} f_{\{i\}}(x_i)dx_i$, that is, over all variables except $\mathbf{x}_u$, and using (3.2) yields the component functions [24, 18, 13]

$$y_{\emptyset,C} = \int_{\mathbb{R}^N} y(\mathbf{x}) \prod_{i=1}^{N} f_{\{i\}}(x_i)dx_i, \tag{3.3a}$$

$$y_{u,C}(\mathbf{X}_u) = \int_{\mathbb{R}^{N-|u|}} y(\mathbf{X}_u, \mathbf{x}_{-u}) \prod_{i=1,i\notin u}^{N} f_{\{i\}}(x_i)dx_i - \sum_{v\subset u} y_{v,C}(\mathbf{X}_v). \tag{3.3b}$$

In (3.3b), $(\mathbf{X}_u, \mathbf{x}_{-u})$ denotes an N-dimensional vector whose ith component is X_i if $i \in u$ and x_i if $i \notin u$. When $u = \emptyset$, the sum in (3.3b) vanishes, resulting in the expression of the constant function $y_{\emptyset,C}$ in (3.3a). When $u = \{1,\dots,N\}$, the integration in the last line of (3.3b) is on the empty set, reproducing identity (3.1) and hence finding the last function $y_{\{1,\dots,N\},C}$. Indeed, all component functions of y in (3.1) can be obtained by interpreting literally (3.3b).

Traditionally, (3.1), (3.3a), and (3.3b) with X_j, $j = 1,\dots,N$, following independent, standard uniform distributions, that is, $f_{\{i\}} = 1$, are identified as the classical ANOVA decomposition [24]. However, recent works reveal no fundamental requirement for a specific probability measure of $\mathbf{X}$, provided that the resultant integrals in (3.3a) and (3.3b) exist and are finite [18]. This generalization is trivial as long as $\mathbf{X}$ is endowed with a product-type probability measure.

3.2. Second-moment statistics. Applying the expectation operators on $y(\mathbf{X})$ in (3.1) and $(y(\mathbf{X}) - \mu)^2$ and recognizing Propositions 3.1 and 3.2, the mean of y is

$$\mu := \mathbb{E}\left[y(\mathbf{X})\right] = y_{\emptyset,C}, \tag{3.4}$$

whereas its variance

$$\sigma^2 := \mathbb{E}\left[(y(\mathbf{X}) - \mu)^2\right] = \sum_{\emptyset\neq u\subseteq\{1,\dots,N\}} \mathbb{E}\left[y_{u,C}^2(\mathbf{X}_u)\right] \tag{3.5}$$

splits into variances of all *zero*-mean, nonconstant component functions of y. According to (3.5), the variance decomposition follows the same structure of $y - y_{\emptyset,C}$ from (3.1), explaining why the acronym "ANOVA" was also coined for the function decomposition.

4. Generalized ANOVA dimensional decomposition. Consider a dependent random vector with an arbitrary non-product-type probability density function $f_{\mathbf{X}} : \mathbb{R}^N \to \mathbb{R}_0^+$ that has marginal probability density function f_u of $\mathbf{X}_u$, where $\emptyset \neq u \subseteq \{1, \ldots, N\}$. Assume that the support of $f_{\mathbf{X}}$ is grid-closed [11]. The grid closure implies that there exists a grid for every point $\mathbf{x}$ of $\text{supp}(f_{\mathbf{X}}) \subseteq \mathbb{R}^N$; that is, for any point $\mathbf{x} \in \text{supp}(f_{\mathbf{X}})$, one can traverse in each coordinate direction and find another point $\mathbf{x}' \in \text{supp}(f_{\mathbf{X}})$. Under this mild regularity requirement, fulfilled by common probability distributions, a square-integrable multivariate function y with respect to the marginal probability measure $f_u(\mathbf{x}_u)d\mathbf{x}_u$ supported on $\mathbb{R}^{|u|}$ also admits a unique, finite, hierarchical expansion [11]

$$y(\mathbf{X}) = \sum_{u \subseteq \{1,\ldots,N\}} y_{u,G}(\mathbf{X}_u), \tag{4.1}$$

referred to as the generalized ADD, in terms of component functions $y_{u,G}$, $u \subseteq \{1, \ldots, N\}$, of input variables with increasing dimensions. The existence and uniqueness of the decomposition in (4.1) have been proven under conditions (C.1) and (C.2) in [3], but (4.1) works well in practice under more general assumptions [11, 28]. The symbol G in the subscript of $y_{u,G}$ is meant to distinguish the component functions of the generalized ADD from those of the classical ADD. Similar to the classical ADD, the summation in (4.1) comprises 2^N component functions, with each function depending on a group of variables indexed by a particular subset of $\{1, \ldots, N\}$, including the empty set $\emptyset$. However, the component functions of the generalized ADD, different from those of the classical ADD, cannot be derived from the strong annihilating conditions expressed by (3.2). This is because some of the orthogonal properties that stem from (3.2) cannot be duplicated when the random variables are dependent. Having said so, the functions $y_{u,G}$, $u \subseteq \{1, \ldots, N\}$, can also be obtained from a similar perspective by adjusting classical annihilating conditions, described as follows.

4.1. Weak annihilating conditions. The weak annihilating conditions appropriate for the generalized ADD mandate all nonconstant component functions $y_{u,G}$ to integrate to *zero* with respect to the marginal density of $\mathbf{X}_u$ in each coordinate direction of u, that is [11],

$$\int_{\mathbb{R}} y_{u,G}(\mathbf{x}_u) f_u(\mathbf{x}_u) dx_i = 0 \quad \text{for } i \in u \neq \emptyset. \tag{4.2}$$

Compared with (3.2), (4.2) represents a milder version, but it still produces two remarkable properties of the generalized ADD, described by Propositions 4.1 and 4.2.

Proposition 4.1. *The generalized ADD component functions $y_{u,G}$, where $\emptyset \neq u \subseteq \{1, \ldots, N\}$, have zero means, i.e.,*

$$\mathbb{E}\left[y_{u,G}(\mathbf{X}_u)\right] = 0.$$

Proof. For any subset $\emptyset \neq u \subseteq \{1, \ldots, N\}$, let $i \in u$. Then

$$\begin{aligned}
\mathbb{E}\left[y_{u,G}(\mathbf{X}_u)\right] &:= \int_{\mathbb{R}^N} y_{u,G}(\mathbf{x}_u) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbb{R}^{|u|}} y_{u,G}(\mathbf{x}_u) f_u(\mathbf{x}_u) d\mathbf{x}_u \\
&= \int_{\mathbb{R}^{|u|-1}} \int_{\mathbb{R}} y_{u,G}(\mathbf{x}_u) f_u(\mathbf{x}_u) dx_i \prod_{j \in u, j \neq i} dx_j \\
&= 0,
\end{aligned}$$

where the last line follows from using (4.2). ■

Proposition 4.2. *Two distinct generalized ADD component functions $y_{u,G}$ and $y_{v,G}$, where $\emptyset \neq u \subseteq \{1, \ldots, N\}$, $\emptyset \neq v \subseteq \{1, \ldots, N\}$, and $v \subset u$, are orthogonal; i.e., they satisfy the property*

$$\mathbb{E}\left[y_{u,G}(\mathbf{X}_u) y_{v,G}(\mathbf{X}_v)\right] = 0.$$

Proof. For any two subsets $\emptyset \neq u \subseteq \{1, \ldots, N\}$ and $\emptyset \neq v \subseteq \{1, \ldots, N\}$, where $v \subset u$, the subset $u = v \cup (u \setminus v)$. Let $i \in (u \setminus v) \subseteq u$. Then

$$\begin{aligned}
\mathbb{E}\left[y_{u,G}(\mathbf{X}_u) y_{v,G}(\mathbf{X}_v)\right] &:= \int_{\mathbb{R}^N} y_{u,G}(\mathbf{x}_u) y_{v,G}(\mathbf{x}_v) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbb{R}^{|u|}} y_{u,G}(\mathbf{x}_u) y_{v,G}(\mathbf{x}_v) f_u(\mathbf{x}_u) d\mathbf{x}_u \\
&= \int_{\mathbb{R}^{|v|}} y_{v,G}(\mathbf{x}_v) \int_{\mathbb{R}^{|u \setminus v|}} y_{u,G}(\mathbf{x}_u) f_u(\mathbf{x}_u) d\mathbf{x}_{u \setminus v} d\mathbf{x}_v \\
&= \int_{\mathbb{R}^{|v|}} y_{v,G}(\mathbf{x}_v) \int_{\mathbb{R}^{|u \setminus v|-1}} \int_{\mathbb{R}} y_{u,G}(\mathbf{x}_u) f_u(\mathbf{x}_u) dx_i \prod_{\substack{j \in (u \setminus v) \\ j \neq i}} dx_j d\mathbf{x}_v \\
&= 0,
\end{aligned}$$

where the equality to *zero* results from using (4.2). ■

It is elementary to show that (4.2) shrinks to (3.2) for independent random variables, that is, when $\mathbf{X}_u$ has a product-type probability density $f_u(\mathbf{x}_u) = \Pi_{i \in u} f_{\{i\}}(x_i)$ for $\emptyset \neq u \subseteq \{1, \ldots, N\}$.

From Propositions 3.1 and 4.1, all nonconstant component functions of ADD, whether classical or generalized, have *zero* means. Therefore, a non-product-type probability measure, relevant to the generalized ADD, does not vitiate the first-moment properties of the classical ADD. However, Propositions 3.2 and 4.2, which describe the second-moment properties of ADD, tell a slightly different tale: any two distinct nonconstant component functions of the classical ADD are orthogonal, whereas two distinct nonconstant component functions of the generalized ADD are orthogonal only if the index set of one function is a proper subset of the index set of the other function. As an example, consider $N = 3$ with $2^3 - 1 = 7$ nonconstant component functions. Then the generalized ADD permits orthogonality between (1) $y_{\{i\},G}$ and

$y_{\{i_1,i_2\},G}$, where $i = 1,2,3$; $i_1 = i$, $i_2 = 1,2,3$, $i_1 < i_2$; $i_2 = i$, $i_1 = 1,2,3$, $i_1 < i_2$; (2) $y_{\{i\},G}$ and $y_{\{123\},G}$, where $i = 1,2,3$; and (3) $y_{\{i_1,i_2\},G}$ and $y_{\{1,2,3\},G}$, where $i_1, i_2 = 1,2,3$, $i_1 < i_2$. This nested orthogonality, originally presented and referred to as the hierarchical orthogonality by Hooker [11], is the result of imposing weakened annihilating conditions on the generalized ADD.

4.2. A coupled system of equations. Hooker [11] proposed finding component functions of the generalized ADD by minimizing a mean-squared error subject to the hierarchical orthogonality described by Propositions 4.1 and 4.2. A simpler alternative proposed here entails integrating (4.1) with respect to a judiciously chosen marginal probability measure and implementing the weak annihilating conditions when required. Lemma 4.3 and Theorem 4.4 illuminate this alternative approach, which sidesteps the need to solve the optimization problem altogether. The end result is a coupled system of equations satisfied by component functions.

Lemma 4.3. *The generalized ADD component functions $y_{v,G}$, $\emptyset \neq v \subseteq \{1,\dots,N\}$, of a square-integrable function $y : \mathbb{R}^N \to \mathbb{R}$, when integrated with respect to the probability measure $f_{-u}(\mathbf{x}_{-u})d\mathbf{x}_{-u}$, $u \subseteq \{1,\dots,N\}$, reduce to*

$$\int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} = \begin{cases} \int_{\mathbb{R}^{|v\cap -u|}} y_{v,G}(\mathbf{x}_v) f_{v\cap -u}(\mathbf{x}_{v\cap -u}) d\mathbf{x}_{v\cap -u} & \text{if } v\cap u \neq \emptyset \text{ and } v \nsubseteq u, \\ y_{v,G}(\mathbf{x}_v) & \text{if } v\cap u \neq \emptyset \text{ and } v \subseteq u, \\ 0 & \text{if } v \cap u = \emptyset. \end{cases} \tag{4.3}$$

Proof. For any two subsets $\emptyset \neq v \subseteq \{1,\dots,N\}$, $u \subseteq \{1,\dots,N\}$, one can write $(v \cap -u) \subseteq -u$ and $-u = (-u \setminus (v \cap -u)) \cup (v \cap -u)$. Let $v \cap u \neq \emptyset$, where $v \nsubseteq u$ in general. Then one of the two nontrivial results of (4.3) is obtained as

$$\begin{aligned} \int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} &= \int_{\mathbb{R}^{|v\cap -u|}} y_{v,G}(\mathbf{x}_v) \int_{\mathbb{R}^{N-|u|-|v\cap -u|}} f_{-u}(\mathbf{x}_{-u\setminus(v\cap -u)}, \mathbf{x}_{v\cap -u}) \\ &\quad \times d\mathbf{x}_{-u\setminus(v\cap -u)} d\mathbf{x}_{v\cap -u} \\ &= \int_{\mathbb{R}^{|v\cap -u|}} y_{v,G}(\mathbf{x}_v) f_{v\cap -u}(\mathbf{x}_{v\cap -u}) d\mathbf{x}_{v\cap -u}. \end{aligned} \tag{4.4}$$

If $v \subseteq u$, then $y_v(\mathbf{x}_v)$ does not depend on $\mathbf{x}_{-u}$, resulting in

$$\int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} = y_{v,G}(\mathbf{x}_v) \int_{\mathbb{R}^{N-|u|}} f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} = y_{v,G}(\mathbf{x}_v),$$

the other nontrivial result of (4.3). Finally, if $v \cap u = \emptyset$, then $v \cap -u = v$. Let $i \in v$. Therefore, the last line of (4.4), also valid for $v \cap u = \emptyset$, becomes

$$\begin{aligned} \int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} &= \int_{\mathbb{R}^{|v|}} y_{v,G}(\mathbf{x}_v) f_v(\mathbf{x}_v) d\mathbf{x}_v \\ &= \int_{\mathbb{R}^{|v|-1}} \left(\int_{\mathbb{R}} y_{v,G}(\mathbf{x}_v) f_v(\mathbf{x}_v) dx_i \right) \prod_{j\in v, j\neq i} dx_j \\ &= 0, \end{aligned}$$

where the equality to *zero* follows from using (4.2). ■

Theorem 4.4. *The generalized ADD component functions $y_{u,G}$, $u \subseteq \{1,\dots,N\}$, of a square-integrable function $y : \mathbb{R}^N \to \mathbb{R}$ for a given probability measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$ of $\mathbf{X} \in \mathbb{R}^N$ satisfy*

$$y_{\emptyset,G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \tag{4.5a}$$

$$\begin{aligned} y_{u,G}(\mathbf{X}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{X}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} - \sum_{v \subset u} y_{v,G}(\mathbf{X}_v) \\ &\quad - \sum_{\substack{\emptyset \neq v \subseteq \{1,\dots,N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \int_{\mathbb{R}^{|v \cap -u|}} y_{v,G}(\mathbf{X}_{v\cap u}, \mathbf{x}_{v \cap -u}) f_{v \cap -u}(\mathbf{x}_{v \cap -u}) d\mathbf{x}_{v \cap -u}. \end{aligned} \tag{4.5b}$$

Proof. Changing the dummy index from u to v, replacing $\mathbf{X}$ with $\mathbf{x}$, and integrating both sides of (4.1) with respect to the measure $f_{-u}(\mathbf{x}_{-u})d\mathbf{x}_{-u}$, that is, over all variables except $\mathbf{x}_u$, yields

$$\int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} = \sum_{v \subseteq \{1,\dots,N\}} \int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u}, \tag{4.6}$$

which is valid for any $u \subseteq \{1,\dots,N\}$, including the empty set $\emptyset$. To obtain the constant component function of y, let $u = \emptyset$. Then $-u = \{1,\dots,N\}$ and $f_{-u}(\mathbf{x}_{-u})d\mathbf{x}_{-u} = f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$, resulting in

$$\begin{aligned} \textstyle\int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} &= y_{\emptyset,G} + \sum_{\emptyset \neq v \subseteq \{1,\dots,N\}} \int_{\mathbb{R}^N} y_{v,G}(\mathbf{x}_v) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\ &= y_{\emptyset,G} + \sum_{\emptyset \neq v \subseteq \{1,\dots,N\}} \mathbb{E}\left[y_{v,G}(\mathbf{X}_v)\right]. \end{aligned} \tag{4.7}$$

Invoking Proposition 4.1, each expectation of the sum in (4.7) vanishes, yielding (4.5a). To derive the nonconstant component functions, apply Lemma 4.3, that is, (4.3), to simplify the right side of (4.6) into

$$\begin{aligned} &\sum_{v \subseteq \{1,\dots,N\}} \int_{\mathbb{R}^{N-|u|}} y_{v,G}(\mathbf{x}_v) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} \\ &= y_{u,G}(\mathbf{x}_u) + \sum_{v \subset u} y_{v,G}(\mathbf{x}_v) + \sum_{\substack{\emptyset \neq v \subseteq \{1,\dots,N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \int_{\mathbb{R}^{|v \cap -u|}} y_{v,G}(\mathbf{x}_v) f_{v \cap -u}(\mathbf{x}_{v \cap -u}) d\mathbf{x}_{v \cap -u} \end{aligned} \tag{4.8}$$

with $\subset$ representing the proper subset (strict inclusion). Substituting (4.8) into (4.6) and recognizing $v = (v \cap u) \cup (v \cap -u)$ produces (4.5b), completing the proof. ■

The constant component function of ADD, whether classical ($y_{\emptyset,C}$) or generalized ($y_{\emptyset,G}$), is the same as the expected value of $y(\mathbf{X})$. According to (3.3b), all nonconstant component functions of the classical ADD are hierarchically ordered in terms of the cardinality of subsets of $\{1,\dots,N\}$ and are determined sequentially. This is possible because for a given $\emptyset \neq u \subseteq$

$\{1,\ldots,N\}$, the component function $y_{u,C}$ depends only on the component functions $y_{v,C}$ such that $v \subset u$, including $\emptyset$. In contrast, the component functions of the generalized ADD, satisfying (4.5b), are coupled and must be solved simultaneously. In the latter case, for a given $\emptyset \neq u \subseteq \{1,\ldots,N\}$, the component function $y_{u,G}$ depends not only on the component functions $y_{v,G}$, where $v \subset u$, but also on the component functions $y_{v,G}$, where $v \cap u \neq \emptyset$, $v \nsubseteq u$. As an example, consider $u = \{1\}$ and $N = 3$. The classical and generalized component functions depending on x_1 are

$$y_{\{1\},C} = \int_{\mathbb{R}^2} y(x_1,x_2,x_3) f_{\{2\}}(x_2) f_{\{3\}}(x_3) dx_2 dx_3 - y_{\emptyset,C}$$

and

$$\begin{aligned} y_{\{1\},G} = &\int_{\mathbb{R}^2} y(x_1,x_2,x_3) f_{\{2,3\}}(x_2,x_3) dx_2 dx_3 - y_{\emptyset,G} \\ &- \int_{\mathbb{R}} y_{\{1,2\},G}(x_1,x_2) f_{\{2\}}(x_2) dx_2 - \int_{\mathbb{R}} y_{\{1,3\},G}(x_1,x_3) f_{\{3\}}(x_3) dx_3 \\ &- \int_{\mathbb{R}^2} y_{\{1,2,3\},G}(x_1,x_2,x_3) f_{\{2,3\}}(x_2,x_3) dx_2 dx_3, \end{aligned}$$

respectively. For the generalized ADD, there exist $2^N - 1$ such coupled equations, the right number of equations to determine uniquely all nonconstant component functions. A new computational method solving this system of equations will be formally presented in the following section.

Corollary 4.5. *The univariate, bivariate, and trivariate component functions of a square-integrable function $y : \mathbb{R}^N \to \mathbb{R}$, obtained by setting* (1) $u = \{i\}$, $i = 1,\ldots,N$, *and* $1 \leq N < \infty$; (2) $u = \{i_1, i_2\}$, $i_1 = 1,\ldots,N-1$, $i_2 = i_1+1,\ldots,N$, *and* $2 \leq N < \infty$; *and* (3) $u = \{i_1,i_2,i_3\}$, $i_1 = 1,\ldots,N-2$, $i_2 = i_1+1,\ldots,N-1$, $i_3 = i_2+1,\ldots,N$, *and* $3 \leq N < \infty$, *respectively, in* (4.5b) *are*

$$\begin{aligned} y_{\{i\},G}(X_i) = &\int_{\mathbb{R}^{N-1}} y(X_i, \mathbf{x}_{-\{i\}}) f_{-\{i\}}(\mathbf{x}_{-\{i\}}) d\mathbf{x}_{-\{i\}} - y_{\emptyset,G} \\ &- \sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \cap \{i\} \neq \emptyset, v \nsubseteq \{i\}}} \int_{\mathbb{R}^{|v \cap -\{i\}|}} y_{v,G}(\mathbf{X}_{v\cap\{i\}}, \mathbf{x}_{v\cap-\{i\}}) f_{v\cap-\{i\}}(\mathbf{x}_{v\cap-\{i\}}) d\mathbf{x}_{v\cap-\{i\}}, \end{aligned}$$

$$\begin{aligned} y_{\{i_1,i_2\},G}(X_{i_1},X_{i_2}) = &\int_{\mathbb{R}^{N-2}} y(X_{i_1},X_{i_2},\mathbf{x}_{-\{i_1,i_2\}}) f_{-\{i_1,i_2\}}(\mathbf{x}_{-\{i_1,i_2\}}) d\mathbf{x}_{-\{i_1,i_2\}} \\ &- y_{\emptyset,G} - y_{\{i_1\},G}(X_{i_1}) - y_{\{i_2\},G}(X_{i_2}) \\ &- \sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \cap \{i_1,i_2\} \neq \emptyset, v \nsubseteq \{i_1,i_2\}}} \int_{\mathbb{R}^{|v\cap-\{i_1,i_2\}|}} y_{v,G}(\mathbf{X}_{v\cap\{i_1,i_2\}}, \mathbf{x}_{v\cap-\{i_1,i_2\}}) \\ &\times f_{v\cap-\{i_1,i_2\}}(\mathbf{x}_{v\cap-\{i_1,i_2\}}) d\mathbf{x}_{v\cap-\{i_1,i_2\}}, \end{aligned}$$

$$
\begin{aligned}
y_{\{i_1,i_2,i_3\},G}(X_{i_1},X_{i_2},X_{i_3}) = &\int_{\mathbb{R}^{N-3}} y(X_{i_1},X_{i_2},X_{i_3},\mathbf{x}_{-\{i_1,i_2,i_3\}}) f_{-\{i_1,i_2,i_3\}}(\mathbf{x}_{-\{i_1,i_2,i_3\}}) d\mathbf{x}_{-\{i_1,i_2,i_3\}} \\
&- y_{\emptyset,G} - y_{\{i_1\},G}(X_{i_1}) - y_{\{i_2\},G}(X_{i_2}) - y_{\{i_3\},G}(X_{i_3}) \\
&- y_{\{i_1,i_2\},G}(X_{i_1},X_{i_2}) - y_{\{i_1,i_3\},G}(X_{i_1},X_{i_3}) - y_{\{i_2,i_3\},G}(X_{i_2},X_{i_3}) \\
&- \sum_{\substack{\emptyset\neq v\subseteq\{1,\ldots,N\} \\ v\cap\{i_1,i_2,i_3\}\neq\emptyset, v\not\subseteq\{i_1,i_2,i_3\}}} \int_{\mathbb{R}^{|v\cap-\{i_1,i_2,i_3\}|}} y_{v,G}(\mathbf{X}_{v\cap\{i_1,i_2,i_3\}},\mathbf{x}_{v\cap-\{i_1,i_2,i_3\}}) \\
&\times f_{v\cap-\{i_1,i_2,i_3\}}(\mathbf{x}_{v\cap-\{i_1,i_2,i_3\}}) d\mathbf{x}_{v\cap-\{i_1,i_2,i_3\}}.
\end{aligned}
$$

The specialized formulae for the component functions in Corollary 4.5 were previously derived by Li and Rabitz [14]. Theorem 4.4, in contrast, is general and provides a single master formula to concisely represent all component functions of the generalized ADD.

Corollary 4.6. *If $\mathbf{X} = (X_1,\ldots,X_N) \in \mathbb{R}^N$ comprises independent random variables, which follow arbitrary probability measures $f_{\{i\}}(x_i)dx_i$, $i = 1,\ldots,N$, then the generalized ADD degenerates to the classical ADD.*

Proof. For independent coordinates of $\mathbf{X}$, all joint probability density functions are products of their marginals, that is, $f_u(\mathbf{x}_u) = \Pi_{i\in u} f_{\{i\}}(x_i)$, where $\emptyset \neq u \subseteq \{1,\ldots,N\}$. Using this product structure of probability measures, which makes the strong and weak annihilating conditions coincide, it is elementary to show that $y_{u,G} = y_{u,C}$ for any $u \subseteq \{1,\ldots,N\}$, including $y_{\emptyset,G} = y_{\emptyset,C}$. Therefore, the generalized ADD reduces to the classical ADD. ■

5. A constructive method for determining component functions.

This section presents a new computational method, employing measure-consistent, multivariate orthonormal polynomials as basis functions, for solving the coupled system of equations satisfied by the component functions of the generalized ADD.

5.1. Multivariate orthonormal polynomials.

For the rest of the paper, the standard multi-index notation will be used in describing orthogonal polynomials in several variables. Accordingly, for a given $\emptyset \neq u \subseteq \{1,\ldots,N\}$, $1 \leq |u| \leq N$, let $\mathbf{j}_{|u|} = (j_1,\ldots,j_{|u|}) \in \mathbb{N}_0^{|u|}$ represent a $|u|$-dimensional multi-index with each component a nonnegative integer. For $\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}$ and $\mathbf{x}_u = (x_{i_1},\ldots,x_{i_{|u|}}) \in \mathbb{R}^{|u|}$, where $1 \leq i_1 < \cdots < i_{|u|} \leq N$, a monomial in $\mathbf{x}_u$ of index $\mathbf{j}_{|u|}$ is defined by $\mathbf{x}_u^{\mathbf{j}_{|u|}} := x_{i_1}^{j_1} \times \cdots \times x_{i_{|u|}}^{j_{|u|}}$. The nonnegative integer $|\mathbf{j}_{|u|}| := j_1 + \cdots + j_{|u|}$, which is equal to the 1-norm of $\mathbf{j}_{|u|}$, is called the total degree of $\mathbf{x}_u^{\mathbf{j}_{|u|}}$. A linear combination of $\mathbf{x}_u^{\mathbf{j}_{|u|}}$, where $|\mathbf{j}_{|u|}| = m_u$ and $m_u \in \mathbb{N}$, is a homogeneous polynomial of degree m_u. Denote by $\mathcal{P}_{m_u}^u := \operatorname{span}\{\mathbf{x}_u^{\mathbf{j}_{|u|}} : |\mathbf{j}_{|u|}| = m_u, \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\}$ the space of homogeneous polynomials of degree m_u, by $\Pi_{m_u}^u := \operatorname{span}\{\mathbf{x}_u^{\mathbf{j}_{|u|}} : |\mathbf{j}_{|u|}| \leq m_u, \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\}$ the space of polynomials of degree at most m_u, and by Π^u the space of all polynomials of $|u|$ variables. It is well known that [5]

$$
\dim \mathcal{P}_{m_u}^u = \binom{m_u + |u| - 1}{m_u} \text{ and } \dim \Pi_{m_u}^u = \binom{m_u + |u|}{m_u}.
$$

Assume that, for $\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}$, the moments $\int_{\mathbb{R}^{|u|}} \mathbf{x}_u^{\mathbf{j}_{|u|}} f_u(\mathbf{x}_u) d\mathbf{x}_u$ of $\mathbf{X}_u$ exist and are finite, and $\int_{\mathbb{R}^{|u|}} \psi_{u\mathbf{j}_{|u|}}^2(\mathbf{x}_u) f_u(\mathbf{x}_u) d\mathbf{x}_u > 0$ for every $\psi_{u\mathbf{j}_{|u|}} \in \Pi^u$, where $\psi_{u\mathbf{j}_{|u|}} \neq 0$ is a polynomial in $\mathbf{x}_u$ of

degree $\mathbf{j}_{|u|}$. Consistent with the probability measure $f_u(\mathbf{x}_u)d\mathbf{x}_u$, define an inner product

$$(g,h)_{f_u} := \int_{\mathbb{R}^{|u|}} g(\mathbf{x}_u)h(\mathbf{x}_u)f_u(\mathbf{x}_u)d\mathbf{x}_u =: \mathbb{E}\left[g(\mathbf{X}_u)h(\mathbf{X}_u)\right] \tag{5.1}$$

of two $|u|$-variate functions g and h. Then there exist orthogonal polynomials in $\mathbf{x}_u$ with respect to the inner product defined by (5.1). More precisely, a polynomial $\psi_{u\mathbf{j}_{|u|}} \in \Pi^u_{m_u}$ is called orthogonal with respect to $(\cdot,\cdot)_{f_u}$ if $(\psi_{u\mathbf{j}_{|u|}}, \psi_{u\mathbf{k}_{|u|}})_{f_u} = \mathbb{E}\big[\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)\psi_{u\mathbf{k}_{|u|}}(\mathbf{X}_u)\big] = 0$ for $|\mathbf{k}_{|u|}| < |\mathbf{j}_{|u|}|$, that is, for all $\psi_{u\mathbf{k}_{|u|}} \in \Pi^u_{m_u-1}$. This means that $\psi_{u\mathbf{j}_{|u|}}$ is orthogonal to all polynomials of lower degrees, but it may not be orthogonal to other orthogonal polynomials of the same degree. Define $\mathcal{V}^u_{m_u} := \{\psi_{u\mathbf{j}_{|u|}} \in \Pi^u_{m_u} : (\psi_{u\mathbf{j}_{|u|}}, \psi_{u\mathbf{k}_{|u|}})_{f_u} = 0,\ \psi_{u\mathbf{k}_{|u|}} \in \Pi^u_{m_u-1}\}$ as the space of orthogonal polynomials of degree of exactly m_u. It is elementary to show that the $\dim \mathcal{V}^u_{m_u} = \dim \mathcal{P}^u_{m_u}$. If, in addition, $(\psi_{u\mathbf{j}_{|u|}}, \psi_{u\mathbf{j}_{|u|}})_{f_u} = \mathbb{E}\big[\psi^2_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)\big] = 1$, then $\psi_{u\mathbf{j}_{|u|}}$ is called an orthonormal polynomial in $\mathbf{x}_u$ of degree $|\mathbf{j}_{|u|}|$, to be used in the remainder of this paper.

5.2. Fourier-Polynomial Expansion. 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 $f_u(\mathbf{x}_u)d\mathbf{x}_u$ of $\mathbf{X}_u$. For $\emptyset \neq u = \{i_1,\ldots,i_{|u|}\} \subseteq \{1,\ldots,N\}$, where $1 \le |u| \le N$ and $1 \le i_1 < \cdots < i_{|u|} \le N$, let $(\Omega_u, \mathcal{F}_u, P_u)$ be the probability triple of $\mathbf{X}_u = (X_{i_1},\ldots,X_{i_{|u|}})$. Denote the associated space of the $|u|$-variate component functions of y by

$$\mathcal{L}_2(\Omega_u, \mathcal{F}_u, P_u) := \left\{ y_{u,G} : \int_{\mathbb{R}^{|u|}} y^2_{u,G}(\mathbf{x}_u)f_u(\mathbf{x}_u)d\mathbf{x}_u < \infty \right\},$$

which is a Hilbert space. Then $\{\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u),\ \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, j_1,\ldots,j_{|u|} \neq 0\}$, if it is dense, constitutes a basis of $\mathcal{L}_2(\Omega_u, \mathcal{F}_u, P_u)$. The standard Hilbert space theory states that every nonconstant component function $y_{u,G} \in \mathcal{L}_2(\Omega_u, \mathcal{F}_u, P_u)$ of y can be expanded as [5]

$$y_{u,G}(\mathbf{X}_u) = \sum_{\substack{\mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|} \\ j_1,\ldots,j_{|u|} \neq 0}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u) \tag{5.2}$$

with

$$C_{u\mathbf{j}_{|u|}} := \int_{\mathbb{R}^{|u|}} y_{u,G}(\mathbf{x}_u)\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)d\mathbf{x}_u \tag{5.3}$$

defining associated expansion coefficients. Note that the summation in (5.2) precludes $j_1,\ldots,j_{|u|} = 0$; that is, the individual degree of each variable X_i in $\psi_{u\mathbf{j}_{|u|}}$, $i \in u$, cannot be *zero* since $y_{u,G}$ is a strictly $|u|$-variate function and has a *zero* mean following Proposition 4.1. For a more precise interpretation, the selection of multivariate Hermite polynomials as basis functions is described as follows.

Consider quadratic approximations of the univariate ($|u| = 1$) and bivariate ($|u| = 2$) component functions of $y(\mathbf{X})$, where $\mathbf{X} = (X_1,\ldots,X_N)$ is a *zero*-mean, N-dimensional Gaussian random vector with positive-definite covariance matrix $\mathbf{\Sigma_X} = \mathbb{E}[\mathbf{X}\mathbf{X}^T] = [\rho_{ij}\sigma_i\sigma_j]$, comprising

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

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} (\det \boldsymbol{\Sigma}_{\mathbf{X}})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}_{\mathbf{X}}^{-1}\mathbf{x}\right] =: \phi_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\Sigma}_{\mathbf{X}}). \tag{5.4}$$

The marginal probability densities of $\mathbf{X}_u$, $\emptyset \neq u \subseteq \{1, \ldots, N\}$, are also Gaussian, and are easily derived as

$$f_u(\mathbf{x}_u) = (2\pi)^{-\frac{|u|}{2}} (\det \boldsymbol{\Sigma}_u)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}{\mathbf{x}_u}^T \boldsymbol{\Sigma}_u^{-1}\mathbf{x}_u\right] =: \phi_u(\mathbf{x}_u; \boldsymbol{\Sigma}_u), \tag{5.5}$$

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

$$\tilde{\psi}_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) = \frac{(-1)^{|\mathbf{j}_{|u|}|}}{\phi_u(\mathbf{x}_u; \boldsymbol{\Sigma}_u)} \left(\frac{\partial}{\partial \mathbf{x}_u}\right)^{\mathbf{j}_{|u|}} \phi_u(\mathbf{x}_u; \boldsymbol{\Sigma}_u), \quad \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}, \tag{5.6}$$

where $(\partial/\partial \mathbf{x}_u)^{\mathbf{j}_{|u|}} := \partial^{j_1+\cdots+j_{|u|}}/\partial x_{i_1}^{j_1} \cdots \partial x_{i_{|u|}}^{j_{|u|}}$. They eventually form a set of multivariate Hermite orthonormal polynomials

$$\left\{\psi_{u\mathbf{j}_{|u|}} := \tilde{\psi}_{u\mathbf{j}_{|u|}}/(\tilde{\psi}_{u\mathbf{j}_{|u|}}, \tilde{\psi}_{u\mathbf{j}_{|u|}})_{\phi_u}, \mathbf{j}_{|u|} \in \mathbb{N}_0^{|u|}\right\} \tag{5.7}$$

that are consistent with the probability measure $\phi_u(\mathbf{x}_u; \boldsymbol{\Sigma}_u)d\mathbf{x}_u$ of $\mathbf{X}_u$. For example, when $u = \{i\}$, $i = 1, \ldots, N$, and $j_1 \leq 2$, (5.6) and (5.7) yield

$$\psi_{\{i\}0}(X_i) = 1, \quad \psi_{\{i\}1}(X_i) = X_i, \quad \psi_{\{i\}2}(X_i) = \frac{X_i^2 - 1}{\sqrt{2}},$$

a sequence of orthonormal polynomials for quadratic approximation of any square-integrable univariate function of X_i. Clearly, the complete basis set for a general function is $\{\psi_{\{i\}0}, \psi_{\{i\}1}, \psi_{\{i\}2}\}$. However, since $\mathbb{E}[y_{\{i\},G}(X_i)] = 0$ as per Proposition 4.1, only a linear combination of $\psi_{\{i\}1}$ and $\psi_{\{i\}2}$, without including $\psi_{\{i\}0}$—that is, the basis subset $\{\psi_{\{i\}1}, \psi_{\{i\}2}\}$—is sufficient to approximate $y_{\{i\},G}$. Similarly, when $u = \{i_1, i_2\}$, $i_1 = 1, \ldots, N-1$, $i_2 = i_1 + 1, \ldots, N$, and $|\mathbf{j}_2| \leq 2$, (5.6) and (5.7) result in

$$\psi_{\{i_1,i_2\}00}(X_{i_1}, X_{i_2}) = 1,$$

$$\psi_{\{i_1,i_2\}10}(X_{i_1}, X_{i_2}) = \frac{X_{i_1} - \rho_{i_1 i_2} X_{i_2}}{\sqrt{1 - \rho_{i_1 i_2}^2}}, \ \psi_{\{i_1,i_2\}01}(X_{i_1}, X_{i_2}) = \frac{X_{i_2} - \rho_{i_1 i_2} X_{i_1}}{\sqrt{1 - \rho_{i_1 i_2}^2}},$$

$$\psi_{\{i_1,i_2\}20}(X_{i_1}, X_{i_2}) = \frac{X_{i_1}^2 + \rho_{i_1 i_2}^2 \left(1 + X_{i_2}^2\right) - 2\rho_{i_1 i_2} X_{i_1} X_{i_2} - 1}{\sqrt{2}\left(1 - \rho_{i_1 i_2}^2\right)},$$

$$\psi_{\{i_1,i_2\}02}(X_{i_1}, X_{i_2}) = \frac{X_{i_2}^2 + \rho_{i_1 i_2}^2 \left(1 + X_{i_1}^2\right) - 2\rho_{i_1 i_2} X_{i_1} X_{i_2} - 1}{\sqrt{2}\left(1 - \rho_{i_1 i_2}^2\right)},$$

$$\psi_{\{i_1,i_2\}11}(X_{i_1}, X_{i_2}) = \frac{\sqrt{1 + \rho_{i_1 i_2}^2}}{\rho_{i_1 i_2}^2 - 1} \left[\frac{\rho_{i_1 i_2}\left(X_{i_1}^2 + X_{i_2}^2\right)}{1 + \rho_{i_1 i_2}^2} - X_{i_1} X_{i_2} + \frac{\rho_{i_1 i_2}\left(\rho_{i_1 i_2}^2 - 1\right)}{1 + \rho_{i_1 i_2}^2}\right],$$

a sequence of orthonormal polynomials for quadratic approximation of any square-integrable bivariate function of X_{i_1} and X_{i_2}. There are multiple ways to choose a set of basis functions for $y_{\{i_1,i_2\},G}$. The author proposes selecting a nested basis set $\{\psi_{\{i_1\}1}, \psi_{\{i_1\}2}, \psi_{\{i_2\}1}, \psi_{\{i_2\}2}, \psi_{\{i_1,i_2\}11}\}$, which subsumes the basis functions for $y_{\{i_1\},G}$ and $y_{\{i_2\},G}$. It is elementary to show that the members of such a nested basis set have *zero* means and are hierarchically orthogonal, that is, $\mathbb{E}[\psi_{\{i_1,i_2\}j_1j_2}(X_{i_1}, X_{i_2})\psi_{\{i_1\}j_1}(X_{i_1})] = 0$ and $\mathbb{E}[\psi_{\{i_1,i_2\}j_1j_2}(X_{i_1}, X_{i_2})\psi_{\{i_2\}j_2}(X_{i_2})] = 0$ for $|\mathbf{j}_2| \leq 2$, $j_1, j_2 \neq 0$. Again, since $\mathbb{E}[y_{\{i_1,i_2\},G}(X_{i_1}, X_{i_2})] = 0$ following Proposition 4.1, a constant multiplier of $\psi_{\{i_1,i_2\}11}$, excluding $\psi_{\{i_1\}1}$, $\psi_{\{i_1\}2}$, $\psi_{\{i_2\}1}$, $\psi_{\{i_2\}2}$—that is, the basis subset $\{\psi_{\{i_1,i_2\}11}\}$—is adequate to approximate $y_{\{i_1,i_2\},G}$. In both instances, the power j_k for each X_{i_k}, $k \in u$, of the monomial $\mathbf{X}_u^{\mathbf{j}_{|u|}} := X_{i_1}^{j_1} \times \cdots \times X_{i_{|u|}}^{j_{|u|}}$ in the basis subset, whether $|u| = 1$ or $|u| = 2$, is not *zero*. Therefore, the condition $j_1, \ldots, j_{|u|} \neq 0$ is required for selecting the basis functions of a general $|u|$-variate component function in (5.2) and (5.3).

The selection of nested basis functions, as explained in the preceding paragraph for $u = \{i_1, i_2\}$ and $v = \{i_1\}$ or $\{i_2\}$ and Gaussian probability measure, easily extends to a general $|u|$-variate function and general probability measure $f_u(\mathbf{x}_u)d\mathbf{x}_u$. Given $\emptyset \neq u \subseteq \{1, \ldots, N\}$, let $\{\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v), \emptyset \neq v \subseteq u, \mathbf{k}_{|v|} \in \mathbb{N}_0^{|v|}, k_1, \ldots, k_{|v|} \neq 0\}$ be a nested set of measure-consistent orthonormal polynomial basis functions for $y_{u,G}$, which comprises as a subset measure-consistent orthonormal polynomial basis functions for $y_{v,G}$, $\emptyset \neq v \subset u$. Then, from fundamental properties of multivariate orthogonal polynomials, (1) $\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)$ has a *zero* mean for any $\emptyset \neq u \subseteq \{1, \ldots, N\}$ and $j_1, \ldots, j_{|u|} \neq 0$; and (2) $\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)$ is orthogonal to $\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v)$ for any $\emptyset \neq v \subset u$, $j_1, \ldots, j_{|u|} \neq 0$, and $k_1, \ldots, k_{|v|} \neq 0$. Therefore, the *zero*-mean property of $y_{u,G}$ and hierarchical orthogonality between $y_{u,G}$ and $y_{v,G}$, $\emptyset \neq v \subset u$, as required by Propositions 4.1 and 4.2, are naturally satisfied.

The constant function $y_{\emptyset,G}$ defined in (4.5a) is an N-dimensional integral, which must be calculated or estimated by some means. The evaluation of nonconstant component functions $y_{u,G}(\mathbf{x}_u)$, $\emptyset \neq u \subseteq \{1, \ldots, N\}$, requires calculation of the expansion coefficients defined in (5.3), which are similar integrals on at most $\mathbb{R}^N$. But, since $y_{u,G}$ is unknown, the coefficients cannot be determined from their definitions alone. Two new results, Theorem 5.1 and Corollary 5.2, describe how these coefficients can be calculated from the solution of a linear system of algebraic equations.

Theorem 5.1. *Let y be a square-integrable function of $\mathbf{X}$, admitting a generalized ADD, where $\mathbf{X} = (X_1, \ldots, X_N)$ is an $\mathbb{R}^N$-valued dependent random vector with an arbitrary non-product-type joint probability density function $f_{\mathbf{X}} : \mathbb{R}^N \to \mathbb{R}_0^+$ and a marginal probability density function f_u of $\mathbf{X}_u$. Given $\emptyset \neq u \subseteq \{1, \ldots, N\}$, let $\{\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v), \emptyset \neq v \subseteq u, \mathbf{k}_{|v|} \in \mathbb{N}_0^{|v|}, k_1, \ldots, k_{|v|} \neq 0\}$ be a nested set of measure-consistent orthonormal polynomial basis functions such that $\mathbb{E}[\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)] = 0$ and $\mathbb{E}[\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v)] = 0$ for $\emptyset \neq v \subset u$, $j_1, \ldots, j_{|u|} \neq 0$, and $k_1, \ldots, k_{|v|} \neq 0$. Then the expansion coefficients of the polynomial representation of nonconstant component functions of y in (5.2) and (5.3) satisfy*

$$C_{u\mathbf{j}_{|u|}} + \sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \sum_{\substack{\mathbf{k}_{|v|} \in \mathbb{N}_0^{|v|} \\ k_1,\ldots,k_{|v|} \neq 0}} C_{v\mathbf{k}_{|v|}} J_{u\mathbf{j}_{|u|}, v\mathbf{k}_{|v|}} = I_{u\mathbf{j}_{|u|}}, \tag{5.8}$$

where the integrals

$$I_{u\mathbf{j}_{|u|}} := \int_{\mathbb{R}^N} y(\mathbf{x})\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)f_{-u}(\mathbf{x}_{-u})d\mathbf{x}, \tag{5.9a}$$

$$J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}} := \int_{\mathbb{R}^{|v\cup u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v)f_u(\mathbf{x}_u)f_{v\cap -u}(\mathbf{x}_{v\cap -u})d\mathbf{x}_{v\cup u}. \tag{5.9b}$$

Proof. Replace $y_{u,G}$ in (5.3) with the right side of (4.5b) to write

$$\begin{aligned} C_{u\mathbf{j}_{|u|}} = &\int_{\mathbb{R}^N} y(\mathbf{x})\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)f_{-u}(\mathbf{x}_{-u})d\mathbf{x} \\ &- \sum_{\emptyset\neq v\subset u} \textstyle\int_{\mathbb{R}^{|u|}} y_{v,G}(\mathbf{x}_v)\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)d\mathbf{x}_u \\ &- \sum_{\substack{\emptyset\neq v\subseteq\{1,\ldots,N\}\\ v\cap u\neq\emptyset, v\not\subseteq u}} \int_{\mathbb{R}^{|v\cup u|}} y_{v,G}(\mathbf{x}_v)\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)f_{v\cap -u}(\mathbf{x}_{v\cap -u})d\mathbf{x}_{v\cup u}, \end{aligned} \tag{5.10}$$

where in the second line the integral associated with $v = \emptyset$, that is,

$$\int_{\mathbb{R}^{|u|}} y_{\emptyset,G}\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)d\mathbf{x}_u = y_{\emptyset,G}\mathbb{E}\left[\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)\right]$$

drops out for all $\emptyset \neq u \subseteq \{1,\ldots,N\}$, consistent with the definition of orthonormal polynomials. Now substitute all component functions of y involved in (5.10) with their Fourier-polynomial expansions, as described by (5.2), which results in

$$\begin{aligned} C_{u\mathbf{j}_{|u|}} = &\int_{\mathbb{R}^N} y(\mathbf{x})\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)f_u(\mathbf{x}_u)f_{-u}(\mathbf{x}_{-u})d\mathbf{x} \\ &- \sum_{\emptyset\neq v\subset u} \sum_{\substack{\mathbf{k}_{|v|}\in\mathbb{N}_0^{|v|}\\ k_1,\ldots,k_{|v|}\neq 0}} C_{v\mathbf{k}_{|v|}} \textstyle\int_{\mathbb{R}^{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v)f_u(\mathbf{x}_u)d\mathbf{x}_u \\ &- \sum_{\substack{\emptyset\neq v\subseteq\{1,\ldots,N\}\\ v\cap u\neq\emptyset, v\not\subseteq u}} \sum_{\substack{\mathbf{k}_{|v|}\in\mathbb{N}_0^{|v|}\\ k_1,\ldots,k_{|v|}\neq 0}} C_{v\mathbf{k}_{|v|}} \\ &\times \int_{\mathbb{R}^{|v\cup u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v)f_u(\mathbf{x}_u)f_{v\cap -u}(\mathbf{x}_{v\cap -u})d\mathbf{x}_{v\cup u}. \end{aligned} \tag{5.11}$$

From fundamental properties of orthogonal polynomials and nested construction of basis functions, $\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)$ is orthogonal to $\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v)$, that is, the expectation or the integral

$$\mathbb{E}\left[\psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{X}_v)\right] = \int_{\mathbb{R}^{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v)f_u(\mathbf{x}_u)d\mathbf{x}_u = 0$$

for any $\emptyset \neq v \subset u$, $j_1, \ldots, j_{|u|} \neq 0$, and $k_1, \ldots, k_{|v|} \neq 0$. Therefore, (5.11) reduces to

$$
\begin{aligned}
C_{u\mathbf{j}_{|u|}} = &\int_{\mathbb{R}^N} y(\mathbf{x})\psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u) f_u(\mathbf{x}_u) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x} \\
&- \sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \sum_{\substack{\mathbf{k}_{|v|} \in \mathbb{N}_0^{|v|} \\ k_1,\ldots,k_{|v|} \neq 0}} C_{v\mathbf{k}_{|v|}} \\
&\times \int_{\mathbb{R}^{|v \cup u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{x}_u)\psi_{v\mathbf{k}_{|v|}}(\mathbf{x}_v) f_u(\mathbf{x}_u) f_{v\cap -u}(\mathbf{x}_{v\cap -u}) d\mathbf{x}_{v\cup u}.
\end{aligned} \tag{5.12}
$$

Defining integrals $I_{u\mathbf{j}_{|u|}}$ and $J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}}$, as in (5.9a) and (5.9b), (5.12) simplifies to (5.8), proving the theorem. ■

5.3. Finite-dimensional approximation. Equation (5.8) describes an infinite-dimensional system involving an infinite number of coefficients. Therefore, a finite-dimensional approximation, leading to approximate expansion coefficients and a truncated generalized ADD, must be used in practice. Corollary 5.2 provides such a solution.

Corollary 5.2. *When truncated at $|u| = S$ and $|\mathbf{j}_{|u|}| := j_1 + \cdots + j_{|u|} = m$, where $1 \leq S < N$ and $S \leq m < \infty$, the approximate expansion coefficients $\tilde{C}_{u\mathbf{j}_{|u|}}$ satisfy*

$$
\tilde{C}_{u\mathbf{j}_{|u|}} + \sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \sum_{k=1}^{m} \sum_{\substack{|\mathbf{k}_{|v|}| = k \\ k_1,\ldots,k_{|v|} \neq 0}} \tilde{C}_{v\mathbf{k}_{|v|}} J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}} = I_{u\mathbf{j}_{|u|}}, \ 1 \leq |u| \leq S, \ 1 \leq |\mathbf{j}_{|u|}| \leq m, \tag{5.13}
$$

and

$$
\tilde{C}_{u\mathbf{j}_{|u|}} = 0, \ 1 \leq |u| \leq S, \ m+1 \leq |\mathbf{j}_{|u|}| < \infty; \ S+1 \leq |u| \leq N, \ 1 \leq |\mathbf{j}_{|u|}| < \infty. \tag{5.14}
$$

Expressed compactly, the system of equations in (5.13) for nontrivial solutions of $\tilde{C}_{u\mathbf{j}_{|u|}}$ forms an $L_{S,m} \times L_{S,m}$ matrix equation: $\mathbf{Az} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{L_{S,m} \times L_{S,m}}$ contains integrals $J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}}$, $\mathbf{b} \in \mathbb{R}^{L_{S,m}}$ comprises integrals $I_{u\mathbf{j}_{|u|}}$, and $\mathbf{z} \in \mathbb{R}^{L_{S,m}}$ is the solution vector of the approximate expansion coefficients. The size of the matrix equation is

$$
L_{S,m} = \sum_{k=1}^{S} \binom{N}{k} \binom{m}{k},
$$

where $\binom{m}{k} = 0$ if $k > m$, when the measure-consistent orthonormal polynomials are constructed by satisfying the condition $j_1, \ldots, j_{|u|} \neq 0$, as explained previously. The matrix form of (5.13) is easy to implement and solve, and is scalable to higher dimensions in a straightforward way. It is elementary to show that $\tilde{C}_{u\mathbf{j}_{|u|}} \to C_{u\mathbf{j}_{|u|}}$ as $S \to N$ and $m \to \infty$.

The truncations introduced in Corollary 5.2 engender an S-variate, mth-order generalized ADD approximation

$$
\tilde{y}_{S,m}(\mathbf{X}) = y_{\emptyset,G} + \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ 1 \leq |u| \leq S}} \sum_{k=1}^{m} \sum_{\substack{|\mathbf{j}_{|u|}| = k \\ j_1,\ldots,j_{|u|} \neq 0}} \tilde{C}_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}} \tag{5.15}
$$

of $y(\mathbf{X})$ in (4.1), which is grounded on a fundamental conjecture known to be true in many real-world applications: given a high-dimensional function y, its $|u|$-variate component functions decay rapidly with respect to $|u|$, leading to accurate lower-variate approximations of y. For instance, by selecting $S = 1$ or 2, the functions $\tilde{y}_{1,m}$ and $\tilde{y}_{2,m}$, respectively, provide univariate and bivariate approximations of y. The higher the value of S and/or m, the higher the accuracy, but also the concomitant computational effort. When $S \to N$ and $m \to \infty$, $\tilde{y}_{S,m}$ converges to y in the mean-square sense, generating a hierarchical and convergent sequence of approximations.

The computational effort in determining the expansion coefficients $y_{\emptyset,G}$ and $\tilde{C}_{u\mathbf{j}_{|u|}}$ is rooted in efficient and accurate calculations of various N-dimensional integrals, including $I_{u\mathbf{j}_{|u|}}$, $u \subseteq \{1, \ldots, N\}$, $1 \le |u| \le S$, $|\mathbf{j}_{|u|}| \le m$. For large N, a full numerical integration employing an N-dimensional tensor product of a univariate quadrature rule is computationally prohibitive. Instead, a dimension-reduction integration scheme, developed by Xu and Rahman [33], can be applied to estimate the coefficients efficiently. The scheme entails approximating a high-dimensional integral of interest by a finite sum of lower-dimensional integrations. The computational complexity is Sth-order polynomial—for instance, linear or quadratic when $S = 1$ or 2—with respect to the number of variables or integration points, alleviating the curse of dimensionality to an extent determined by S. See the work of Xu and Rahman [33] for further details.

If the function y is a sum of at most S-variate, mth-order polynomials, then (5.13) and (5.14) yield the exact solution of the expansion coefficients $C_{u\mathbf{j}_{|u|}}$ and (5.15) exactly reproduces y, provided that the integrals $I_{u\mathbf{j}_{|u|}}$ and $J_{u\mathbf{j}_{|u|},v\mathbf{k}_{|v|}}$ are calculated exactly. Numerical results corroborating theoretical findings will be presented in section 7.

6. Second-moment analysis.

Once the component functions of the generalized ADD have been determined, subsequent evaluations of their second-moment characteristics, including global sensitivity analysis, are conducted as follows.

6.1. Mean and variance.

Applying the expectation operator on (4.1) and noting Proposition 4.1, the mean

$$\mu := \mathbb{E}[y(\mathbf{X})] = y_{\emptyset,G} \tag{6.1}$$

of $y(\mathbf{X})$ matches the constant component function of the generalized ADD. This is similar to (3.4), the result from the classical ADD, although the respective constants involved are not the same. Applying the expectation operator again, this time on $(y(\mathbf{X}) - \mu)^2$, and recognizing Proposition 4.2 results in the variance

$$\sigma^2 := \mathbb{E}\left[(y(\mathbf{X}) - \mu)^2\right] = \sum_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \mathbb{E}\left[y_{u,G}^2(\mathbf{X}_u)\right] + \sum_{\substack{\emptyset \neq u,v \subseteq \{1,\ldots,N\} \\ u \nsubseteq v \nsubseteq u}} \mathbb{E}\left[y_{u,G}(\mathbf{X}_u) y_{v,G}(\mathbf{X}_v)\right] \tag{6.2}$$

of $y(\mathbf{X})$, where the first sum represents variance contributions from all nonconstant component functions. In contrast, the second sum in (6.2) typifies covariance contributions from two distinct nonconstant component functions that are not orthogonal—a ramification of imposing the weak annihilating conditions appropriate for the generalized ADD. The latter sum

disappears altogether in the classical ADD because of the strong annihilating conditions that are possible to enforce for independent probability measures. Nonetheless, (6.1) and (6.2) furnish new generalized formulae for the second-moment statistics of $y(\mathbf{X})$ in terms of the moments of relevant component functions.

6.2. Global sensitivity indices. Mathematical modeling of complex systems often requires sensitivity analysis to determine how an output variable of interest is influenced by individual or subsets of input variables. A global sensitivity analysis constitutes the study of how the output variance from a mathematical model is divvied up, qualitatively or quantitatively, to distinct sources of input variation in the model [25]. There exist a multitudes of methods or techniques for calculating the global sensitivity indices of a function of independent variables: the random balance design method [31], state-dependent parameter metamodel [22], Sobol's method [26], polynomial chaos expansion [30], polynomial dimensional decomposition [19], nonparametric regression procedures [29], and many others. In contrast, only a few methods, such as those presented by Li and Rabitz [14], Kucherenko, Tarantola, and Annoni [12], and Chastaing, Gamboa, and Prieur [3], are available for models with dependent or correlated input. In this section, a triplet of global sensitivity indices is defined for problems involving dependent probability distributions of input variables.

Three $|u|$-variate global sensitivity indices of a stochastic response function $y(\mathbf{X})$ for a subset $\mathbf{X}_u$ of input variables $\mathbf{X}$, denoted by $S_{u,v}$, $S_{u,c}$, and S_u, are defined as the ratios

$$S_{u,v} := \frac{\mathbb{E}\left[y_{u,G}^2(\mathbf{X}_u)\right]}{\sigma^2}, \tag{6.3}$$

$$S_{u,c} := \frac{\displaystyle\sum_{\substack{\emptyset\neq v\subseteq\{1,\ldots,N\}\\ u\not\subseteq v\not\subseteq u}} \mathbb{E}\left[y_{u,G}(\mathbf{X}_u)y_{v,G}(\mathbf{X}_v)\right]}{\sigma^2}, \tag{6.4}$$

$$S_u := S_{u,v} + S_{u,c}, \tag{6.5}$$

provided that the variance $0 < \sigma^2 < \infty$. The first two indices, $S_{u,v}$ and $S_{u,c}$, represent the normalized versions of the variance contribution from $y_{u,G}$ to σ^2 and of the covariance contributions from $y_{u,G}$ and all $y_{v,G}$, such that $u \not\subseteq v \not\subseteq u$, to σ^2. They will be named the variance-driven global sensitivity index and the covariance-driven global sensitivity index, respectively, of $y(\mathbf{X})$ for $\mathbf{X}_u$. The third index, S_u, referred to as the total global sensitivity index of $y(\mathbf{X})$ for $\mathbf{X}_u$, is the sum of variance and covariance contributions from or associated with $y_{u,G}$ to σ^2. Since $\emptyset \neq u \subseteq \{1,\ldots,N\}$, there exist $2^N - 1$ such triplets of indices, adding up to

$$\sum_{\emptyset\neq u\subseteq\{1,\ldots,N\}} S_u = \sum_{\emptyset\neq u\subseteq\{1,\ldots,N\}} S_{u,v} + \sum_{\emptyset\neq u\subseteq\{1,\ldots,N\}} S_{u,c} = 1. \tag{6.6}$$

From the definitions, the variance-driven sensitivity index $S_{u,v}$ is a nonnegative, real-valued number. It reflects the contribution of $\mathbf{X}_u$ through $y_u(\mathbf{X}_u)$ in the system structure of $y(\mathbf{X})$. In contrast, the covariance-driven sensitivity index $S_{u,v}$ can be negative, positive, or zero, depending on the correlation between $\mathbf{X}_u$ and $\mathbf{X}_v$. It represents the contribution of $\mathbf{X}_u$ by the interaction of $y_u(\mathbf{X}_u)$ and $y_v(\mathbf{X}_v)$, when $u \nsubseteq v \nsubseteq u$, due to dependent probability distribution. Depending on whether $S_{u,c}$ is positive or negative, $S_{u,c}$ strengthens or weakens S_u, provided that $S_u > 0$. The individual sums of these two indices in (6.6) over all $\emptyset \neq u \subseteq \{1, \ldots, N\}$ may exceed unity or be negative, but the sum of these two individual sums is always equal to *one*. When the random variables are independent, the covariance-driven contribution to the total sensitivity index vanishes, leaving behind only one sensitivity index for the classical ADD. The global sensitivity indices, whether derived from the generalized or classical ADD, can be used to rank variables, fix unessential variables, and reduce dimensions of large-scale problems. They also facilitate a means to define effective dimensions of the function y, as follows.

Li et al. [15] have presented similar definitions of the three sensitivity indices under the names *structural*, *correlative*, and *total sensitivity* indices. However, their correlative sensitivity index, defined by

$$S'_{u,c} := \frac{\displaystyle\sum_{\substack{\emptyset \neq v \subseteq \{1,\ldots,N\} \\ v \neq u}} \mathbb{E}\left[y_{u,G}(\mathbf{X}_u) y_{v,G}(\mathbf{X}_v)\right]}{\sigma^2},$$

represents an unreduced version of the covariance-driven sensitivity index $S_{u,c}$ defined in (6.4). The difference between these two definitions stems from not recognizing the hierarchical orthogonality condition. Indeed, using Proposition 4.2, the condition $u \neq v$ reduces to $u \nsubseteq v \nsubseteq u$, resulting in $S'_{u,c} = S_{u,c}$.

6.3. Effective dimensions. For many practical applications, the multivariate function y of N variables, fortunately, can be effectively approximated by a sum of at most S-variate component functions $y_{u,G}$, $1 \leq |u| \leq S \leq N$, of the generalized ADD in (4.1). The truncation can be achieved by the notion of effective dimension, introduced by Caflisch, Morokoff, and Owen [2], who exploited the classical ADD-based low effective dimension to explain why the quasi–Monte Carlo method outperforms the crude Monte Carlo algorithm for evaluating a certain class of high-dimensional integrals. In this section, extended definitions of two generalized effective dimensions, stemming from the generalized ADD and global sensitivity indices, are presented.

Definition 6.1. *A square-integrable multivariate function y of $\mathbf{X} \in \mathbb{R}^N$ with finite variance $0 < \sigma^2 < \infty$ has a generalized effective dimension $1 \leq S_s \leq N$ in the superposition sense, henceforth denoted as the superposition dimension, if*

$$S_s := \min\left\{ S : 1 \leq S \leq N \text{ such that } \left| 1 - \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ 1 \leq |u| \leq S}} S_u \right| \leq 1 - p \right\}$$

and a generalized effective dimension $1 \leq S_t \leq N$ *in the truncation sense, henceforth denoted as the truncation dimension, if*

$$S_t := \min \left\{ S : 1 \leq S \leq N \text{ such that } \left| 1 - \sum_{u \subseteq \{1,\ldots,S\}} S_u \right| \leq 1 - p \right\},$$

where S_u *is the total global sensitivity index of* $y(\mathbf{X})$ *for* $\mathbf{X}_u$, $\emptyset \neq u \subseteq \{1, \ldots, N\}$, *and* $0 \leq p \leq 1$ *is a percentile threshold.*

Caflisch et al. used the 99th percentile for p, but it can be treated as a threshold parameter linked to the desired accuracy of a stochastic solution. Both definitions capture the notion in which the function y is almost S_s- or S_t-dimensional. The relevance of the truncation or superposition dimension depends on the nature of the function. The former signifies the number of important random variables and is appropriate when some variables are more important than others in an ordered set. In contrast, the latter determines whether the low-variate component functions of dimensional decomposition dominate the function and is appropriate when all variables are equally important. For truly high-dimensional problems, all variables contribute to a function value; therefore, the superposition dimension is more useful than the truncation dimension.

According to the definitions, evaluations of the generalized effective dimensions require calculating the variance σ^2 exactly, which is infeasible, if not impossible, for a general function y of an arbitrary number of variables. However, the S-variate, mth-order approximation $\tilde{y}_{S,m}(\mathbf{X})$, discussed in conjunction with Corollary 5.2, can be used to estimate the variance of $y(\mathbf{X})$, furnishing a practical means to calculate the effective dimensions. In this case, a convergence analysis with respect to S and m, or an adaptive version of (5.15), briefly described in section 6.4, will be required.

6.4. Adaptive-sparse approximation. The global sensitivity indices can be exploited to create an adaptive-sparse ADD approximation of a high-dimensional function. Let $\epsilon_1 \geq 0$ and $\epsilon_2 \geq 0$ denote two nonnegative error tolerances that specify the minimum values of $\tilde{S}_{u,m_u}$, which is the m_uth-order approximation of S_u, and $\Delta\tilde{S}_{u,m_u} := (\tilde{S}_{u,m_u} - \tilde{S}_{u,m_u-1})/\tilde{S}_{u,m_u-1}$, provided that $\tilde{S}_{u,m_u-1} \neq 0$. Then an adaptive-sparse ADD approximation

$$\bar{y}(\mathbf{X}) := y_{\emptyset,G} + \sum_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \sum_{m_u=1}^{\infty} \sum_{\substack{|\mathbf{j}_{|u|}|=m_u,\, j_1,\ldots,j_{|u|}\neq 0 \\ \tilde{S}_{u,m_u}>\epsilon_1, \Delta\tilde{S}_{u,m_u}>\epsilon_2}} C_{u\mathbf{j}_{|u|}} \psi_{u\mathbf{j}_{|u|}}(\mathbf{X}_u)$$

of $y(\mathbf{X})$ is formed by the subset of ANOVA component functions, satisfying two inclusion criteria: (1) $\tilde{S}_{u,m_u} > \epsilon_1$; and (2) $\Delta\tilde{S}_{u,m_u} > \epsilon_2$ for all $1 \leq |u| \leq N$ and $1 \leq m_u < \infty$. The first criterion requires the contribution of an m_uth-order polynomial approximation of $y_{u,G}(\mathbf{X}_u)$ toward the variance of $y(\mathbf{X})$ to exceed ϵ_1 in order to be accommodated within the resultant truncation. The second criterion identifies the augmentation in the variance contribution from $y_{u,G}(\mathbf{X}_u)$ evoked by a single increment in the polynomial order m_u and determines if it surpasses ϵ_2. In other words, these two criteria ascertain which interactive effects between two or more input random variables are retained and dictate the largest order of polynomials

in a component function, formulating a fully adaptive-sparse ANOVA approximation. No truncation parameters, that is, S, m of the truncated ADD, need to be selected a priori or arbitrarily. Although successfully developed for independent variables [34], additional efforts are needed for numerical implementation of the adaptive-sparse approximation for dependent variables.

7. Examples. Two illustrative examples, the one entailing an explicit mathematical function and the other involving an implicit function derived from finite-element analysis, are presented.

7.1. A mathematical function. Consider a quadratic polynomial function

$$y = (a_0 + a_1 X_1)(b_0 + b_1 X_2) + (a_0 + a_1 X_1)(c_0 + c_1 X_3) + (b_0 + b_1 X_2)(c_0 + c_1 X_3)$$

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

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

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, $i, j = 1, 2, 3$, $i \neq j$, and joint probability density function described by (5.4) for $N = 3$. Four sets of correlation coefficients with varied strengths and types of statistical dependence among random variables were examined: (1) $\rho_{12} = \rho_{13} = \rho_{23} = 0$ (uncorrelated); (2) $\rho_{12} = \rho_{13} = \rho_{23} = 1/5$ (equally correlated); (3) $\rho_{12} = 1/5$, $\rho_{13} = 2/5$, $\rho_{23} = 4/5$ (positively correlated); and (4) $\rho_{12} = -1/5$, $\rho_{13} = 2/5$, $\rho_{23} = -4/5$ (mixedly correlated). The deterministic parameters are $a_0 = b_0 = c_0 = 2$, $a_1 = b_1 = c_1 = 1$, rendering y a symmetric function. The objective of this simple yet insightful example is to explain how the proposed methods can be applied to determine the component functions of and global sensitivity indices from the generalized ADD.

Given the Gaussian probability density function of $\mathbf{X}$, the marginal probability densities of $\mathbf{X}_u$, $\emptyset \neq u \subseteq \{1, 2, 3\}$, are also Gaussian and are described by (5.5). The probability density $\phi_u(\mathbf{x}_u; \mathbf{\Sigma}_u)$ induces multivariate Hermite orthonormal polynomials $\{\psi_{u\mathbf{j}_{|u|}}\}$, as described by (5.6) and (5.7), 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 (5.9a) and (5.9b). Using Corollary 5.2, that is, (5.13), $S = 2$, $m = 2$, and these two sets of integrals, a system of linear equations was generated and then solved to determine exactly the expansion coefficients $\tilde{C}_{u\mathbf{j}_{|u|}}$ for $\emptyset \neq u \subseteq \{1, 2, 3\}$ and $|\mathbf{j}_{|u|}| \leq m$. Since y is a sum of at most bivariate, second-order polynomials, the selection of $S = 2$ and $m = 2$ is adequate to produce $\tilde{C}_{u\mathbf{j}_{|u|}} = C_{u\mathbf{j}_{|u|}}$, thereby exactly reproducing y from the generalized ADD.

Table 1 presents all eight component functions of y, obtained exactly using the proposed methods in sections 5 and 6, for four distinct cases of correlation properties of $\mathbf{X}$. It is elementary to verify that all component functions have zero means (Proposition 3.1 or 4.1) and are either fully orthogonal (Proposition 3.2) for Case 1 or hierarchically orthogonal (Proposition 4.2) for Cases 2 through 4. When there is no correlation between any two random

Table 1

The generalized ADD component functions and second-moment statistics of y.[(a)]

Case	Component functions	Moments
Case 1: Uncorrelated $(\rho_{12} = \rho_{13} = \rho_{23} = 0)$	$y_{\emptyset,G} = y_{\emptyset,C} = 12$ $y_{\{1\},G} = y_{\{1\},C} = 4X_1$ $y_{\{2\},G} = y_{\{2\},C} = 4X_2$ $y_{\{3\},G} = y_{\{3\},C} = 4X_3$ $y_{\{1,2\},G} = y_{\{1,2\},C} = X_1X_2$ $y_{\{1,3\},G} = y_{\{1,3\},C} = X_1X_3$ $y_{\{2,3\},G} = y_{\{2,3\},C} = X_2X_3$ $y_{\{1,2,3\},G} = y_{\{1,2,3\},C} = 0$	$\mu = 12$ $\sigma^2 = 51$
Case 2: Equally correlated $(\rho_{12} = \rho_{13} = \rho_{23} = 1/5)$	$y_{\emptyset,G} = \frac{63}{5}$ $y_{\{1\},G} = -\frac{5}{13} + 4X_1 + \frac{5}{13}X_1^2$ $y_{\{2\},G} = -\frac{5}{13} + 4X_2 + \frac{5}{13}X_2^2$ $y_{\{3\},G} = -\frac{5}{13} + 4X_3 + \frac{5}{13}X_3^2$ $y_{\{1,2\},G} = \frac{12}{65} - \frac{5}{26}X_1^2 + X_1X_2 - \frac{5}{26}X_2^2$ $y_{\{1,3\},G} = \frac{12}{65} - \frac{5}{26}X_1^2 + X_1X_3 - \frac{5}{26}X_3^2$ $y_{\{2,3\},G} = \frac{12}{65} - \frac{5}{26}X_2^2 + X_2X_3 - \frac{5}{26}X_3^2$ $y_{\{1,2,3\},G} = 0$	$\mu = \frac{63}{5}$ $\sigma^2 = \frac{1794}{25}$
Case 3: Positively correlated $(\rho_{12} = 1/5, \rho_{13} = 2/5, \rho_{23} = 4/5)$	$y_{\emptyset,G} = \frac{67}{5}$ $y_{\{1\},G} = -\frac{405}{754} + 4X_1 + \frac{405}{754}X_1^2$ $y_{\{2\},G} = -\frac{725}{1066} + 4X_2 + \frac{725}{1066}X_2^2$ $y_{\{3\},G} = -\frac{990}{1189} + 4X_3 + \frac{990}{1189}X_3^2$ $y_{\{1,2\},G} = \frac{12}{65} - \frac{5}{26}X_1^2 + X_1X_2 - \frac{5}{26}X_2^2$ $y_{\{1,3\},G} = \frac{42}{145} - \frac{10}{29}X_1^2 + X_1X_3 - \frac{10}{29}X_3^2$ $y_{\{2,3\},G} = \frac{36}{205} - \frac{20}{41}X_2^2 + X_2X_3 - \frac{20}{41}X_3^2$ $y_{\{1,2,3\},G} = 0$	$\mu = \frac{67}{5}$ $\sigma^2 = \frac{2514}{25}$
Case 4: Mixedly correlated $(\rho_{12} = -1/5, \rho_{13} = 2/5, \rho_{23} = -4/5)$	$y_{\emptyset,G} = \frac{57}{5}$ $y_{\{1\},G} = -\frac{115}{754} + 4X_1 + \frac{115}{754}X_1^2$ $y_{\{2\},G} = \frac{725}{1066} + 4X_2 - \frac{725}{1066}X_2^2$ $y_{\{3\},G} = \frac{170}{1189} + 4X_3 - \frac{170}{1189}X_3^2$ $y_{\{1,2\},G} = -\frac{12}{65} + \frac{5}{26}X_1^2 + X_1X_2 + \frac{5}{26}X_2^2$ $y_{\{1,3\},G} = \frac{42}{145} - \frac{10}{29}X_1^2 + X_1X_3 - \frac{10}{29}X_3^2$ $y_{\{2,3\},G} = -\frac{36}{205} + \frac{20}{41}X_2^2 + X_2X_3 + \frac{20}{41}X_3^2$ $y_{\{1,2,3\},G} = 0$	$\mu = \frac{57}{5}$ $\sigma^2 = \frac{774}{25}$

(a) $y = 12 + 4X_1 + 4X_2 + 4X_3 + X_1X_2 + X_1X_3 + X_2X_3$, where $a_0 = b_0 = c_0 = 2$, $a_1 = b_1 = c_1 = 1$.

variables, that is, $\rho_{12} = \rho_{13} = \rho_{23} = 0$, the proposed method replicates exactly the component functions of the classical ADD. Clearly, the component functions vary with the correlation structure, but when added together they reconstruct the function y regardless of whether or not the random variables are independent. It is important to note that the univariate parts of y, which are strictly linear functions of X_i, are exactly reproduced only when there is no correlation between any two random variables, that is, when invoking the classical ADD. In contrast, the univariate component functions derived from the generalized ADD with a non-trivial correlation structure contain second-order terms and are generally nonlinear. This is due to statistical dependence among random variables, inducing higher-order univariate terms that are not present in the original function to begin with. A similar behavior is observed when comparing the bivariate component functions in Table 1. The additional higher-order terms generated by dependent probability measures vanish when summing all component functions of a generalized ADD. The mean and variance of y for all four cases, calculated using (6.1) and (6.2), are also displayed in Table 1.

The component functions listed in Table 1 were employed for calculating the variance-driven, covariance-driven, and total global sensitivity indices defined in (6.3), (6.4), and (6.5). The expectations involved in (6.3) and (6.4) were exactly evaluated from their respective integral definitions. Table 2 enumerates the triplets of sensitivity indices, $S_{u,v}$, $S_{u,c}$, and S_u, of y for X_1, X_2, X_3, (X_1, X_2), (X_1, X_3), (X_2, X_3), and (X_1, X_2, X_3), calculated separately for the uncorrelated case and the three correlated cases. Three key findings jump out as follows. First, the total sensitivity indices from the generalized ADD for all three correlated cases comprise both variance- and covariance-driven contributions, whereas the total sensitivity indices from the generalized ADD for the uncorrelated case or from the classical ADD emanate solely from the variances of component functions. Second, for the mixedly correlated case, the sum of the variance-driven indices may exceed unity, while the sum of the covariance-driven indices may be negative, as specifically demonstrated when $\rho_{12} = -1/5$, $\rho_{13} = 2/5$, $\rho_{23} = -4/5$. Third, the stronger the correlations among random variables, the larger the covariance-driven contributions to the total sensitivity indices.

Using the total sensitivity indices in Table 2, the total effect sensitivity indices of y with respect to X_i, defined as $\bar{S}_i := \sum_{i \in u} S_u$, $i = 1, 2, 3$, were calculated to decipher the importance of each random variable. Table 3 displays the total effect sensitivity indices with respect to three random variables for the four 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 uncorrelated and equally correlated cases, all three variables are equally important, yielding 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 of random variables heavily influences the composition of component functions as well as global sensitivity analysis.

Table 2
Triplets of global sensitivity indices from the generalized ADD of y.[(a)]

$\mathbf{X}_u$	$S_{u,v}$	$S_{u,c}$	S_u	$S_{u,v}$	$S_{u,c}$	S_u
	Case 1: Uncorrelated ($\rho_{12} = \rho_{13} = \rho_{23} = 0$)			Case 2: Equally correlated ($\rho_{12} = \rho_{13} = \rho_{23} = 1/5$)		
X_1	0.313725	0	0.313725	0.227088	0.089780	0.316868
X_2	0.313725	0	0.313725	0.227088	0.089780	0.316868
X_3	0.313725	0	0.313725	0.227088	0.089780	0.316868
(X_1, X_2)	0.019608	0	0.019608	0.012349	0.004116	0.016465
(X_1, X_3)	0.019608	0	0.019608	0.012349	0.004116	0.016465
(X_2, X_3)	0.019608	0	0.019608	0.012349	0.004116	0.016465
(X_1, X_2, X_3)	0	0	0	0	0	0
$\sum$	1	0	1	0.718312	0.281688	1
	Case 3: Positively correlated ($\rho_{12} = 1/5, \rho_{13} = 2/5, \rho_{23} = 4/5$)			Case 4: Mixedly correlated ($\rho_{12} = -1/5, \rho_{13} = 2/5, \rho_{23} = -4/5$)		
X_1	0.164847	0.096992	0.261839	0.518299	0.103039	0.621337
X_2	0.168309	0.165600	0.333909	0.546677	−0.509771	0.036905
X_3	0.172897	0.202314	0.375211	0.518116	−0.201389	0.316728
(X_1, X_2)	0.008812	0.008812	0.017624	0.028623	−0.014311	0.014311
(X_1, X_3)	0.006049	0.004321	0.010370	0.019647	−0.014034	0.005614
(X_2, X_3)	0.000786	0.000262	0.001048	0.002553	0.002553	0.005105
(X_1, X_2, X_3)	0	0	0	0	0	0
$\sum$	0.5217	0.4783	1	1.63391	−0.63391	1

(a) $y = 12 + 4X_1 + 4X_2 + 4X_3 + X_1X_2 + X_1X_3 + X_2X_3$, where $a_0 = b_0 = c_0 = 2$, $a_1 = b_1 = c_1 = 1$.

Table 3
Total effects of random variables on the variance of y and relative rankings.[(a)]

Case	$\bar{S}_1$ (rank)	$\bar{S}_2$ (rank)	$\bar{S}_3$ (rank)
Case 1: Uncorrelated ($\rho_{12} = \rho_{13} = \rho_{23} = 0$)	0.352941 (Three-way tie)	0.352941 (Three-way tie)	0.352941 (Three-way tie)
Case 2: Equally correlated ($\rho_{12} = \rho_{13} = \rho_{23} = 1/5$)	0.349798 (Three-way tie)	0.349798 (Three-way tie)	0.349798 (Three-way tie)
Case 3: Positively correlated ($\rho_{12} = 1/5, \rho_{13} = 2/5, \rho_{23} = 4/5$)	0.289833 (3)	0.352581 (2)	0.386628 (1)
Case 4: Mixedly correlated ($\rho_{12} = -1/5, \rho_{13} = 2/5, \rho_{23} = -4/5$)	0.641262 (1)	0.056322 (3)	0.327446 (2)

(a) The total effect of random variable X_i is defined as $\bar{S}_i := \sum_{i \in u} S_u$, $i = 1, 2, 3$, $\emptyset \neq u \subseteq \{1, 2, 3\}$.

7.2. A random eigenvalue problem. The motivation for the next example lies in solving a practical problem involving uncertainty quantification of natural frequencies of a vibrating cantilever plate, as shown in Figure 1(a). The plate has the following deterministic geometric and material properties: length $L = 2$ in (50.8 mm), width $W = 1$ in (25.4 mm), Young's

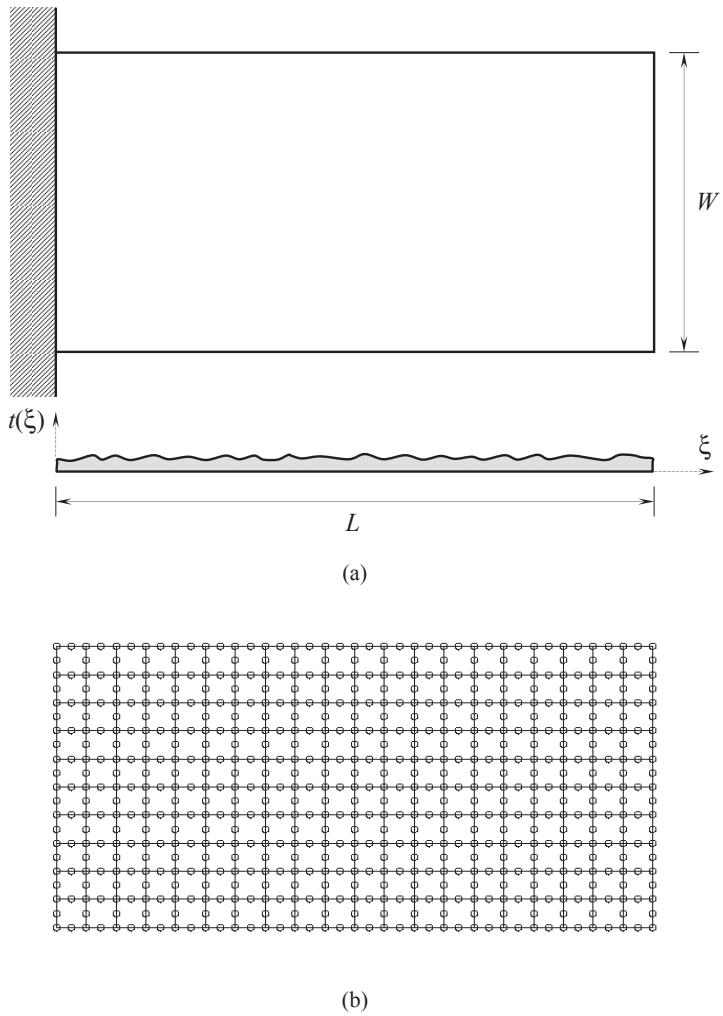

Figure 1. *A cantilever plate;* (a) *geometry;* (b) *finite-element discrete model.*

modulus, $E = 30 \times 10^6$ psi (206.8 GPa), Poisson's ratio $\nu = 0.3$, and mass density $\rho = 7.324 \times 10^{-4}$ lb-s^2/in^4 (7827 kg/mm^3). The randomness in natural frequencies arises due to random thickness $t(\xi)$, which is spatially varying in the longitudinal direction ξ only. The thickness is represented by a homogeneous, lognormal random field $t(\xi) = c\exp[\alpha(\xi)]$ with mean $\mu_t = 0.01$ in (0.254 mm), variance $\sigma_t^2 = v_t^2\mu_t^2$, and coefficient of variation $v_t = 0.2$, where $c = \mu_t/\sqrt{1+v_t^2}$ and $\alpha(\xi)$ is a zero-mean, homogeneous, Gaussian random field with variance $\sigma_\alpha^2 = \ln(1+v_t^2)$ and covariance function $\Gamma_\alpha(\tau) = \mathbb{E}[\alpha(\xi)\alpha(\xi+\tau) = \sigma_\alpha^2\exp[-|\tau|/(0.2L)]$. A 10×20 finite-element mesh of the plate, consisting of 200 eight-noded, second-order shell elements and 661 nodes, is shown in Figure 1(b). Using this mesh and the well-known midpoint method, the random field $\alpha(\xi)$ was discretized into a zero-mean, 20-dimensional, dependent Gaussian random vector $\mathbf{X}$ with covariance matrix $\mathbf{\Sigma_X} = [\Gamma_\alpha(\xi_i - \xi_j)]$, $i, j = 1, \ldots, 20$, where ξ_i is the coordinate of the center of the ith column of elements in Figure 1(b). The same mesh was used to calculate the natural frequencies, which are the square roots of eigenvalues.

Table 4
Means and standard deviations of the first four natural frequencies of the cantilever plate.

Mode	Generalized ADD ($S = 2,\ m = 2$)		Generalized ADD ($S = 2,\ m = 4$)		Crude MCS (5000 samples)	
	Mean (Hz)	St. dev. (Hz)	Mean (Hz)	St. dev. (Hz)	Mean (Hz)	St. dev. (Hz)
1	80.94	15.68	80.98	15.97	80.75	17.06
2	355.13	54.21	355.16	55.22	355.45	55.31
3	508.27	64.21	508.43	64.98	508.20	68.58
4	1169.63	137.45	1169.83	140.64	1170.55	142.29

The bivariate, second-order approximation and the bivariate, fourth-order approximation, that is, (5.15) truncated at $S = 2$, $m = 2$ and $S = 2$, $m = 4$, respectively, were employed to estimate various probabilistic characteristics of the first four natural frequencies of the plate. The construction of orthonormal polynomials is identical to that in the first example. However, the integrals involved in determining the coefficients of (5.15) were estimated from the dimension-reduction integration scheme [33], entailing at most two-dimensional integrations. For the three-point ($= m + 1$, $m = 2$) and five-point ($= m + 1$, $m = 4$) Gauss–Hermite quadrature rules selected, the two proposed approximations require $20 \times (20 - 1)(3 - 1)^2/2 + (20 \times (3 - 1) + 1 = 801$ and $20 \times (20 - 1)(5 - 1)^2/2 + (20 \times (5 - 1) + 1 = 3121$ finite-element analyses, respectively [33].

Table 4 presents the means and standard deviations of the first four natural frequencies, ω_i, $i = 1, \ldots, 6$, of the plate by three different methods: the two proposed bivariate approximations and crude Monte Carlo simulation (MCS). In all three methods, the solution of the matrix characteristic equation for a given input is equivalent to performing a finite-element analysis. Therefore, computational efficiency, even for this simple plate model, is a practical requirement in solving random eigenvalue problems. The statistics by the proposed methods were obtained using 5000 samples of (5.15), which consist of repeated yet inexpensive evaluations of elementary functions. Due to expensive finite-element analysis, however, crude MCS was conducted only up to 5000 realizations, which should be adequate for providing benchmark solutions of the second-moment characteristics. The agreement between the means and standard deviations by the proposed methods and crude MCS in Table 4 is very good even for the second-order approximation.

Figure 2 depicts the marginal probability densities of the four natural frequencies by the proposed approximations and crude MCS. Due to the computational expense inherent to finite-element analysis, the same 5000 samples generated for verifying the statistics in Table 4 were utilized to develop the histograms in Figure 2. However, since the proposed methods yield explicit eigenvalue approximations, an arbitrarily large sample size, e.g., 50,000 in this particular example, was selected to sample (5.15) for estimating the respective densities. Again, the results of the proposed methods and crude MCS match well, given the relatively small sample size of crude MCS. Nonetheless, there exist slight discrepancies in the tail regions of a few densities, suggesting a need for improvements by invoking higher-variate approximations.

The proposed methods, especially the bivariate, second-order approximation, are compu-

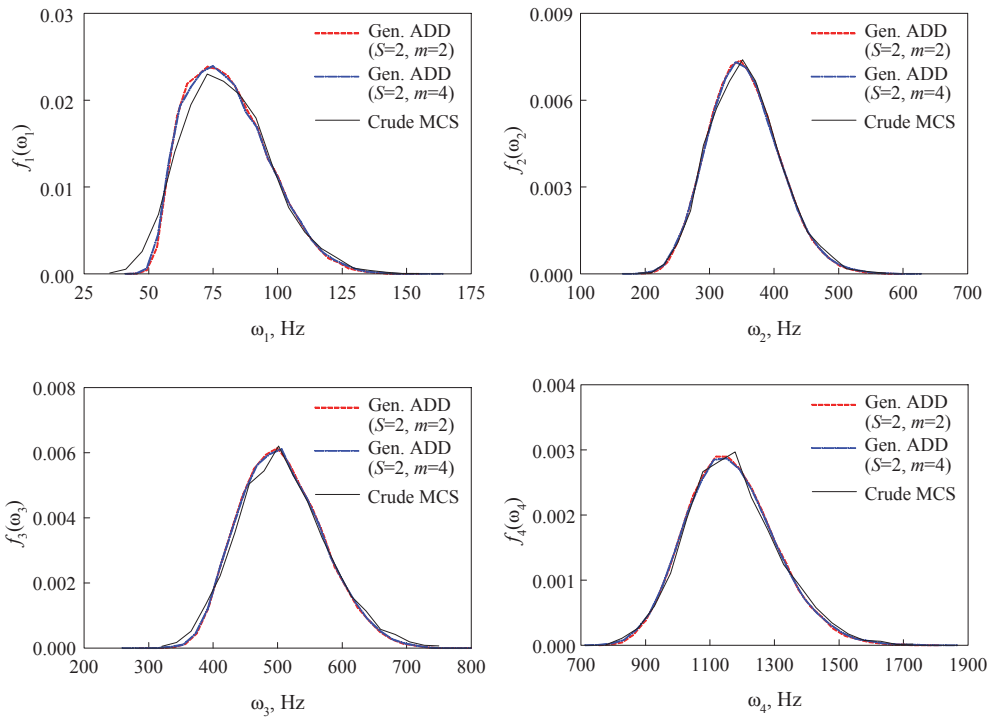

Figure 2. *Marginal probability density functions of the first four natural frequencies of the cantilever plate.*

tationally more efficient than crude MCS. Comparing the results of Table 4, the fourth-order approximation produces at most a modest improvement in the second-moment properties by the second-order approximation. Moreover, the respective marginal densities in Figure 2 obtained by both approximations are practically coincident. Therefore, the second-order approximation provides satisfactory results without incurring the significantly higher cost of the fourth-order approximation, at least for this example. Having said so, the cost scaling of a bivariate approximation, whether second-order or fourth-order, is still quadratic with respect to the number of random variables. Therefore, future efforts on developing adaptive-sparse approximations, where global sensitivity indices can be used to filter out unimportant component functions, should be explored.

8. Conclusion. A generalized ADD for dependent random variables, representing a finite sum of lower-dimensional component functions of a multivariate function, was studied. The classical annihilating conditions, when appropriately weakened, reveal two important properties of the generalized ADD: the component functions have *zero* means and are hierarchically orthogonal. A simple alternative approach is proposed for deriving the coupled system of equations satisfied by the component functions. The coupled equations, which subsume as a special case the classical ADD, reproduces the component functions for independent probability measures. By exploiting measure-consistent, multivariate orthogonal polynomials as bases, a new constructive method is proposed for determining the component functions of the generalized ADD. The method leads to a coupled system of linear algebraic equations

for the expansion coefficients of the component functions that is not only easy to implement and solve, but also supports scalability to higher dimensions. New generalized formulae are presented for the second-moment characteristics of a general stochastic function, including three distinct global sensitivity indices, relevant to dependent probability distributions. Analogous to the component functions, the generalized formulae shrink to the existing formulae from the classical ADD when the random variables are independent. Gaining insights from the generalized ADD, two generalized effective dimensions, one in the superposition sense and the other in the truncation sense, are defined. Numerical results from a simple yet insightful example indicate that the statistical dependence among random variables induces higher-order terms in the generalized ADD that may not be present in the original function or in the classical ADD. In addition, the component functions depend significantly on the correlation coefficients of random variables. Consequently, the global sensitivity indices may also vary widely, producing distinct rankings of random variables. Finally, an application to solving random eigenvalue problems demonstrates that the proposed approximation provides not only accurate but also computationally efficient estimates of the statistical moments and probability densities of natural frequencies.

Acknowledgments. The author wishes to thank the three anonymous reviewers and the associate editor for providing important comments that led to an improved, final version of the paper. The author benefited from discussions with Professors Giles Hooker (Cornell University), Palle Jorgensen (The University of Iowa), and Yuan Xu (University of Oregon).

REFERENCES

[1] R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.
[2] R. E. Caflisch, W. Morokoff, and A. Owen, *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*, J. Comput. Finance, 1 (1997), pp. 27–46.
[3] G. Chastaing, F. Gamboa, and C. Prieur, *Generalized Hoeffding-Sobol decomposition for dependent variables: Application to sensitivity analysis*, Electron. J. Stat., 6 (2012), pp. 2420–2448.
[4] M. Dev, I. Babuska, and J. Oden, *Solution of stochastic partial differential equations using Galerkin finite element techniques*, Comput. Methods Appl. Mech. Engrg., 48 (2001), pp. 6359–6372.
[5] C. F. Dunkl and Y. Xu, *Orthogonal Polynomials of Several Variables*, Encyclopedia Math. Appl. 81, Cambridge University Press, Cambridge, 2001.
[6] B. Efron and C. Stein, *The jackknife estimate of variance*, Ann. Statist., 9 (1981), pp. 586–596.
[7] M. Griebel and M. Holtz, *Dimension-wise integration of high-dimensional functions with applications to finance*, J. Complexity, 26 (2010), pp. 455–489.
[8] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal., 33 (1996), pp. 1995–2016. Corrected printing, 34 (1997), pp. 853–866.
[9] W. Hoeffding, *A class of statistics with asymptotically normal distributions*, Ann. Math. Statist., 20 (1948), pp. 293–325.
[10] M. Holtz, *Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance*, Ph.D. Dissertation, University of Bonn, Bonn, Germany, 2008.
[11] G. Hooker, *Generalized functional ANOVA diagnostics for high dimensional functions of dependent variables*, J. Comput. Graph. Statist., 16 (2007), pp. 709–732.
[12] S. Kucherenko, S. Tarantola, and P. Annoni, *Estimation of global sensitivity indices for models with dependent variables*, Comput. Phys. Comm., 183 (2012), pp. 937–946.
[13] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Wozniakowski, *On decompositions of multivariate functions*, Math. Comp., 79 (2010), pp. 953–966.
[14] G. Li and H. Rabitz, *General formulation of HDMR component functions with independent and corre-*

lated variables, J. Math. Chem., 50 (2012), pp. 99–130.

[15] G. Li, H. Rabitz, P. Elvington, O. Oluwole, F. Bacon, C. Kolb, and J. Schoendorf, *Global sensitivity analysis for systems with independent and/or correlated inputs*, J. Phys. Chem. A, 114 (2010), pp. 6022–6032.

[16] A. B. Owen, *The dimension distribution and quadrature test functions*, Statist. Sinica, 13 (2003), pp. 1–17.

[17] H. Rabitz and O. Alis, *General foundations of high-dimensional model representations*, J. Math. Chem., 25 (1999), pp. 197–233.

[18] S. Rahman, *Approximation errors in truncated dimensional decompositions*, Math. Comp., 83 (2014), pp. 2799–2819.

[19] S. Rahman, *Global sensitivity analysis by polynomial dimensional decomposition*, Reliab. Eng. Syst. Safe., 96 (2011), pp. 824–837.

[20] S. Rahman, *Extended polynomial dimensional decomposition for arbitrary probability distributions*, J. Eng. Mech-ASCE, 135 (2009), pp. 1439–1451.

[21] S. Rahman, *A polynomial dimensional decomposition for stochastic computing*, Internat. J. Numer. Methods Engrg., 76 (2008), pp. 2091–2116.

[22] M. Ratto, A. Pagano, and A. Young, *State dependent parameter metamodelling and sensitivity analysis*, Comput. Phys. Comm., 177 (2007), pp. 863–876.

[23] M. Rosenblatt, *Remarks on a multivariate transformation*, Ann. Math. Statist., 23 (1952), pp. 470–472.

[24] I. M. Sobol, *Theorems and examples on high dimensional model representations*, Reliab. Eng. Syst. Safe., 79 (2003), pp. 187–193.

[25] I. M. Sobol, *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, Math. Comput. Simulation, 55 (2001), pp. 271–280.

[26] I. M. Sobol, *Sensitivity estimates for nonlinear mathematical models*, Math. Model. Comput. Exp., 1 (1993), pp. 407–414.

[27] I. M. Sobol, *Multidimensional quadrature formulas and Haar functions*, Nauka, Moscow, 1969 (in Russian).

[28] C. J. Stone, *The use of polynomial splines and their tensor products in multivariate function estimation*, Ann. Statist., 22 (1994), pp. 118–171.

[29] C. Storlie, L. Swiller, J. Helton, and C. Sallaberry, *Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models*, Reliab. Eng. Syst. Safe., 94 (2009), pp. 1735–1763.

[30] B. Sudret, *Global sensitivity analysis using polynomial chaos expansions*, Reliab. Eng. Syst. Safe., 93 (2008), pp. 964–979.

[31] S. Tarantola, D. Gatelli, and T. Mara, *Random balance designs for the estimation of first-order global sensitivity indices*, Reliab. Eng. Syst. Safe., 91 (2006), pp. 717–727.

[32] N. Wiener, *The homogeneous chaos*, Amer. J. Math., 60 (1938), pp. 897–936.

[33] H. Xu and S. Rahman, *A generalized dimension-reduction method for multi-dimensional integration in stochastic mechanics*, Internat. J. Numer. Methods Engrg., 61 (2004), pp. 1992–2019.

[34] V. Yadav and S. Rahman, *Adaptive-sparse polynomial dimensional decomposition for high-dimensional stochastic computing*, Comput. Methods Appl. Math., 274 (2014), pp. 56–83.

[35] V. Yadav and S. Rahman, *Uncertainty quantification of high-dimensional complex systems by multiplicative polynomial dimensional decompositions*, Internat. J. Numer. Methods Engrg., 94 (2013), pp. 221–247.