UNCERTAINTY QUANTIFICATION BY ALTERNATIVE DECOMPOSITIONS OF MULTIVARIATE FUNCTIONS[*]

SHARIF RAHMAN[†]

Abstract. This article advocates factorized and hybrid dimensional decompositions (FDD/HDD), as alternatives to analysis-of-variance dimensional decomposition (ADD), for second-moment statistical analysis of multivariate functions. New formulas revealing the relationships between component functions of FDD and ADD are proposed. While ADD or FDD is relevant when a function is strongly additive or strongly multiplicative, HDD, whether formed linearly or nonlinearly, requires no specific dimensional hierarchies. Furthermore, FDD and HDD lead to alternative definitions of effective dimension, reported in the current literature only for ADD. New closed-form or analytical expressions are derived for univariate truncations of all three decompositions, followed by mean-squared error analysis of univariate ADD, FDD, and HDD approximations. The analysis finds appropriate conditions when one approximation is better than the other. Numerical results affirm the theoretical finding that HDD is ideally suited to a general function approximation that may otherwise require higher-variate ADD or FDD truncations for rendering acceptable accuracy in stochastic solutions.

Key words. ADD, ANOVA, dimensional decomposition, FDD, HDD, stochastic analysis

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

DOI. 10.1137/12089168X

1. Introduction. Uncertainty quantification of complex systems entails stochastic computing for a large number of random variables. Although the sampling-based methods can solve any stochastic problem, they generally require numerous deterministic trials and are, therefore, cost prohibitive when each analysis demands expensive finite-element or similar numerical calculations. Existing analytical or approximate methods require additional assumptions, mostly for computational expediency, that begin to deteriorate when the input-output mapping is highly nonlinear and the input variance is arbitrarily large. Furthermore, truly high-dimensional problems are all but impossible to solve using most existing methods, including numerical integration. The root deterrence to practical computability is often related to the high dimension of the multivariate integration or interpolation problem, known as the *curse of dimensionality* [1]. The dimensional decomposition of a multivariate function [8, 18, 12, 9] addresses the curse of dimensionality to some extent by developing an input-output behavior of complex systems with low *effective dimension* [2], wherein the degrees of interactions between input variables attenuate rapidly or vanish altogether.

A prominent variant of dimensional decomposition is the well-known analysis-of-variance or ANOVA dimensional decomposition (ADD), first presented by Hoeffding in the 1940s in relation to his seminal work on *U*-statistics [8]. Since then, ADD has been studied by numerous researchers in disparate fields of mathematics [16, 6, 7], statistics [10, 3], finance [5], and basic and applied sciences [11], including engineering disciplines, mostly for uncertainty quantification [21, 15, 13]. However, ADD constitutes a finite sum of lower-dimensional component functions of a multivariate

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

function, and is, therefore, predicated on the additive nature of a function decomposition. In contrast, when a response function is dominantly of a multiplicative nature, suitable multiplicative-type decompositions, such as factorized dimensional decomposition (FDD) [20], should be explored. But existing truncations of FDD are limited to only univariate or bivariate approximations, because FDD component functions of three or more variables have yet to be determined. No error analyses exist comparing ADD and FDD, even for respective univariate approximations. Nonetheless, ADD or FDD is relevant as long as the dimensional hierarchy of a stochastic response is also additive or multiplicative. Unfortunately, the dimensional structure of a response function, in general, is not known a priori. Therefore, indiscriminately using ADD or FDD for general stochastic analysis is not desirable. Further complications may arise when a complex system exhibits a response that is dominantly neither additive nor multiplicative. In the latter case, hybrid approaches coupling both additive and multiplicative decompositions, preferably selected optimally, are needed. For such decompositions, it is unknown which truncation parameter should be selected when compared with that for ADD or FDD. Is it possible to solve a stochastic problem by selecting a lower truncation parameter for hybrid decompositions than for ADD or FDD? If the answer is yes, then a significant, positive impact on high-dimensional uncertainty quantification is anticipated. These enhancements, some of which are indispensable, should be pursued without sustaining significant additional cost.

The purpose of this paper is threefold. First, a brief exposition of ADD and FDD is given in section 3. A theorem, proven herein, reveals the relationship between all component functions of FDD and ADD, so far available only for univariate and bivariate component functions. Three function classes, comprising purely additive functions, purely multiplicative functions, and their mixtures, are examined to illustrate when and how one decomposition or approximation is better than the other. Second, a new hybrid approach optimally blending ADD and FDD approximations, referred to as hybrid dimensional decomposition (HDD), is presented in section 4 for second-moment analysis. Both linear and nonlinear mixtures of ADD and FDD approximations are supported. Gaining insights from FDD and HDD, alternative definitions of effective dimension are proposed. Third, section 5 reports new explicit formulas for respective univariate approximations derived from ADD, FDD, and HDD. The mean-squared error analyses pertaining to univariate ADD, FDD, and HDD approximations are also described. Numerical results from four elementary yet illuminating examples and a practical engineering problem are reported in sections 3 through 5 as relevant. There are nine new theoretical results stated or proved in this paper: Theorems 3.4, 4.1, and 5.4, Corollaries 3.5 and 4.2, Propositions 5.1 and 5.2, and Lemmas 3.3 and 5.3. Mathematical notations and conclusions are defined or drawn in sections 2 and 6, 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 $\mathbf{X} := (X_1, \ldots, X_N) : (\Omega, \mathcal{F}) \to (\mathbb{R}^N, \mathcal{B}^N)$, which describes the statistical uncertainties in all system and input parameters of a high-dimensional stochastic problem. The probability law of $\mathbf{X}$ is completely defined by its joint probability density function

$f_{\mathbf{X}} : \mathbb{R}^N \to \mathbb{R}_0^+$. Assuming independent coordinates of $\mathbf{X}$, its joint probability density $f_{\mathbf{X}}(\mathbf{x}) = \Pi_{i=1}^{i=N} f_i(x_i)$ is expressed by a product of marginal probability density functions f_i of X_i, $i = 1, \ldots, N$, defined on the probability triple $(\Omega_i, \mathcal{F}_i, P_i)$ with a bounded or an unbounded support on $\mathbb{R}$. For a given $u \subseteq \{1, \ldots, N\}$, $f_{-u}(\mathbf{x}_{-u}) := \prod_{i=1, i \notin u}^{N} f_i(x_i)$ defines the marginal density function of $\mathbf{X}_{-u} := \mathbf{X}_{\{1,\ldots,N\}\backslash u}$.

3. ANOVA and factorized dimensional decompositions. 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 $\mathcal{L}_2(\Omega, \mathcal{F}, P)$ 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 $\mathbb{R}^N$. Two useful dimensional decompositions, namely, ADD and FDD of y, are briefly described as follows.

3.1. ADD. The ADD of y, expressed by the recursive form [12, 9, 16],

$$(3.1\text{a}) \qquad y(\mathbf{X}) = \sum_{u \subseteq \{1,\ldots,N\}} y_u(\mathbf{X}_u),$$

$$(3.1\text{b}) \qquad y_\emptyset = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x},$$

$$(3.1\text{c}) \qquad y_u(\mathbf{X}_u) = \int_{\mathbb{R}^{N-|u|}} y(\mathbf{X}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\mathbf{x}_{-u} - \sum_{v \subset u} y_v(\mathbf{X}_v),$$

is a finite, hierarchical expansion in terms of its input variables with increasing dimensions, where $u \subseteq \{1, \ldots, N\}$ is a subset with the complementary set $-u = \{1, \ldots, N\}\backslash u$ and cardinality $0 \le |u| \le N$, and y_u is a $|u|$-variate component function describing a constant or the interactive effect of $\mathbf{X}_u = (X_{i_1}, \ldots, X_{i_{|u|}})$, $1 \le i_1 < \cdots < i_{|u|} \le N$, a subvector of $\mathbf{X}$, on y when $|u| = 0$ or $|u| > 0$. The summation in (3.1a) comprises 2^N terms, with each term depending on a group of variables indexed by a particular subset of $\{1, \ldots, N\}$, including the empty set $\emptyset$. In (3.1c), $(\mathbf{X}_u, \mathbf{x}_{-u})$ 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 (3.1c) vanishes, resulting in the expression of the constant function $y_\emptyset$ in (3.1b). When $u = \{1, \ldots, N\}$, the integration in (3.1c) is on the empty set, reproducing (3.1a) and hence finding the last function $y_{\{1,\ldots,N\}}$. Indeed, all component functions of y can be obtained by interpreting literally (3.1c). The nonconstant component functions satisfy the annihilating conditions [12, 9, 16]

$$\int_{\mathbb{R}} y_u(\mathbf{x}_u) f_i(x_i) dx_i = 0 \text{ for } i \in u,$$

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

PROPOSITION 3.1. *The ADD component functions y_u, $\emptyset \ne u \subseteq \{1, \ldots, N\}$, have zero means, i.e.,*

$$\mathbb{E}\left[y_u(\mathbf{X}_u)\right] = \int_{\mathbb{R}^{|u|}} y_u(\mathbf{x}_u) f_u(\mathbf{x}_u) d\mathbf{x}_u = 0.$$

PROPOSITION 3.2. *Two distinct ADD component functions y_u and y_v, where $\emptyset \ne u \subseteq \{1, \ldots, N\}$, $\emptyset \ne v \subseteq \{1, \ldots, N\}$, and $u \ne v$, are orthogonal, i.e., they satisfy the property*

$$\mathbb{E}\left[y_u(\mathbf{X}_u) y_v(\mathbf{X}_v)\right] = \int_{\mathbb{R}^{|u \cup v|}} y_u(\mathbf{x}_u) y_v(\mathbf{x}_v) f_{u \cup v}(\mathbf{x}_{u \cup v}) d\mathbf{x}_{u \cup v} = 0.$$

Remark 1. Traditionally, (3.1a)–(3.1c) with X_j, $j = 1, \ldots, N$, following independent, standard uniform distributions, has been identified as the ANOVA decomposition [16]; however, more recent works [15, 13] reveal no fundamental requirement for a specific probability measure of $\mathbf{X}$, provided that the resultant integrals in (3.1a)–(3.1c) exist and are finite. In this work, the ADD should be interpreted with respect to an arbitrary but product-type probability measure, instilling desirable orthogonal properties.

3.1.1. ADD approximation. The S-variate ADD approximation $\tilde{y}_S(\mathbf{X})$, say, of $y(\mathbf{X})$, where $0 \leq S < N$, is obtained by truncating the right side of (3.1a) at $0 \leq |u| \leq S$, yielding

$$(3.2) \qquad \tilde{y}_S(\mathbf{X}) = \sum_{\substack{u \subseteq \{1, \ldots, N\} \\ 0 \leq |u| \leq S}} y_u(\mathbf{X}_u).$$

Applying the expectation operator on $y(\mathbf{X})$ and $\tilde{y}_S(\mathbf{X})$ from (3.1a) and (3.2), respectively, and noting Proposition 3.1, the mean of the S-variate ADD approximation

$$(3.3) \qquad \mathbb{E}\left[\tilde{y}_S(\mathbf{X})\right] = y_\emptyset$$

matches the exact mean $\mathbb{E}[y(\mathbf{X})] := \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = y_\emptyset$, regardless of S. Applying the expectation operator again, this time on $(\tilde{y}_S(\mathbf{X}) - y_\emptyset)^2$, and recognizing Proposition 3.2 results in splitting the variance [12, 17]

$$(3.4) \qquad \tilde{\sigma}_S^2 := \mathbb{E}\left[(\tilde{y}_S(\mathbf{X}) - y_\emptyset)^2\right] = \sum_{\substack{\emptyset \neq u \subseteq \{1, \ldots, N\} \\ 1 \leq |u| \leq S}} \sigma_u^2 = \sum_{s=1}^{S} \sum_{\substack{\emptyset \neq u \subseteq \{1, \ldots, N\} \\ |u| = s}} \sigma_u^2$$

of the S-variate ADD approximation into variances $\sigma_u^2 := \mathbb{E}[y_u^2(\mathbf{X}_u)]$, $\emptyset \neq u \subseteq \{1, \ldots, N\}$, of *zero*-mean ADD component functions y_u. Clearly, the approximate variance in (3.4) approaches the exact variance

$$\sigma^2 := \mathbb{E}\left[(y(\mathbf{X}) - y_\emptyset)^2\right] = \sum_{\emptyset \neq u \subseteq \{1, \ldots, N\}} \sigma_u^2 = \sum_{s=1}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1, \ldots, N\} \\ |u| = s}} \sigma_u^2,$$

the sum of all variance terms, when $S \to N$. A normalized version σ_u^2/σ^2 is often called the global sensitivity index of y for $\mathbf{X}_u$ [17].

3.2. FDD. Consider a multiplicative form,

$$(3.5) \qquad y(\mathbf{X}) = \prod_{u \subseteq \{1, \ldots, N\}} [1 + z_u(\mathbf{X}_u)],$$

of the dimensional decomposition of y, where z_u, $u \subseteq \{1, \ldots, N\}$, are various component functions of input variables with increasing dimensions. Like the sum in (3.1a), the product in (3.5) comprises 2^N terms, with each term depending on a group of variables indexed by a particular subset of $\{1, \ldots, N\}$, including the empty set $\emptyset$. This multiplicative decomposition exists and is unique for any square-integrable function $y \in \mathcal{L}_2(\Omega, \mathcal{F}, P)$ with a nonzero mean.

The FDD, originally proposed by Tunga and Demiralp [20] under the name of factorized high-dimensional model representation, has yet to receive due attention for

uncertainty quantification of complex systems. A prime reason why FDD is not on a par with ADD is the lack of explicit relationships between their component functions. Lemma 3.3 and Theorem 3.4 reveal the desired relationships.

LEMMA 3.3. *The FDD component functions, z_v, $\emptyset \neq v \subseteq \{1, \ldots, N\}$, of a square-integrable function $y : \mathbb{R}^N \to \mathbb{R}$ with a nonzero mean $\mathbb{E}[y(\mathbf{X})] = y_\emptyset \neq 0$ satisfy*

$$1 + z_v(\mathbf{X}_v) \neq 0.$$

Proof. Suppose to the contrary that $1 + z_v(\mathbf{X}_v) = 0$ for any $\emptyset \neq v \subseteq \{1, \ldots, N\}$. Then it follows from (3.5) that $y(\mathbf{X}) = 0$ and hence $\mathbb{E}[y(\mathbf{X})] = 0$. This contradicts the assumption that $\mathbb{E}[y(\mathbf{X})] = y_\emptyset \neq 0$, completing the proof. $\square$

THEOREM 3.4. *The recursive relationships between component functions of the ADD and FDD of a nonzero mean, square-integrable function $y : \mathbb{R}^N \to \mathbb{R}$, represented by* (3.1a) *and* (3.5)*, respectively, are*

$$(3.6) \qquad 1 + z_u(\mathbf{X}_u) = \frac{\displaystyle\sum_{v \subseteq u} y_v(\mathbf{X}_v)}{\displaystyle\prod_{v \subset u} [1 + z_v(\mathbf{X}_v)]}, \ u \subseteq \{1, \ldots, N\}.$$

Proof. Since (3.1a) and (3.5) represent the same function y,

$$(3.7) \qquad \sum_{u \subseteq \{1,\ldots,N\}} y_u(\mathbf{X}_u) = \prod_{u \subseteq \{1,\ldots,N\}} [1 + z_u(\mathbf{X}_u)],$$

which, as is, is unwieldy to solve for z_u. Instead, expand the right side of (3.5) to form

$$(3.8)$$
$$y(\mathbf{X}) = 1 + z_\emptyset + \sum_{\substack{u \subseteq \{1,\ldots,N\} \\ |u|=1}} r_u(z_v(\mathbf{X}_v); v \subseteq u) + \sum_{\substack{u \subseteq \{1,\ldots,N\} \\ |u|=2}} r_u(z_v(\mathbf{X}_v); v \subseteq u) + \cdots$$
$$+ r_{\{1,\ldots,N\}}(z_v(\mathbf{X}_v); v \subseteq \{1, \ldots, N\})$$
$$= \sum_{u \subseteq \{1,\ldots,N\}} r_u(z_v(\mathbf{X}_v); v \subseteq u),$$

where $r_u(z_v(\mathbf{X}_v); v \subseteq u)$ is a function of at most $|u|$-variate multiplicative component functions of y. For instance, when $u = \emptyset$, $u = \{i\}$, and $u = \{i_1, i_2\}$, $i, i_1, i_2 = 1, \ldots, N$, $i_2 > i_1$, the corresponding r_u-functions are $r_\emptyset(z_\emptyset) = 1 + z_\emptyset$, $r_{\{i\}}(z_\emptyset, z_{\{i\}}(X_i))$, and $r_{\{i_1,i_2\}}(z_\emptyset, z_{\{i_1\}}(X_{i_1}), z_{\{i_2\}}(X_{i_2}), z_{\{i_1,i_2\}}(X_{i_1}, X_{i_2}))$, respectively. Comparing (3.1a) and (3.8) yields the recursive relationship,

$$(3.9) \qquad r_u(z_v(\mathbf{X}_v); v \subseteq u) = y_u(\mathbf{X}_u),$$

which, on inversion, expresses z_u, $u \subseteq \{1, \ldots, N\}$, in terms of the additive ANOVA component functions y_v, $v \subseteq u$. Therefore, all remaining additive or multiplicative component functions of (3.7) not involved can be ignored. Indeed, setting $y_v = z_v = 0$ for all $v \not\subseteq u$ in (3.7) results in

$$(3.10) \qquad \sum_{v \subseteq u} y_v(\mathbf{X}_v) = \prod_{v \subseteq u} [1 + z_v(\mathbf{X}_v)] = [1 + z_u(\mathbf{X}_u)] \prod_{v \subset u} [1 + z_v(\mathbf{X}_v)],$$

where Lemma 3.3 assures that $1 + z_v(\mathbf{X}_v) \neq 0$ for any $\emptyset \neq v \subseteq u$. On inversion, (3.10) yields (3.6), completing the proof. $\square$

COROLLARY 3.5. *Recursive evaluations of* (3.6) *eliminate* $1 + z_v$, $v \subset u$, *leading to an explicit form of*

$$(3.11) \qquad 1 + z_u(\mathbf{X}_u) = \cfrac{\sum_{w_{|u|} \subseteq u} y_{w_{|u|}}\left(\mathbf{X}_{w_{|u|}}\right)}{\displaystyle\prod_{w_{|u|} \subset u} \cfrac{\sum_{w_{|u|-1} \subseteq w_{|u|}} y_{w_{|u|-1}}\left(\mathbf{X}_{w_{|u|-1}}\right)}{\displaystyle\prod_{w_{|u|-1} \subset w_{|u|}} \cfrac{\ddots}{\ddots \displaystyle\prod_{w_2 \subset w_3} \cfrac{\sum_{w_1 \subseteq w_2} y_{w_1}\left(\mathbf{X}_{w_1}\right)}{\displaystyle\prod_{w_1 \subset w_2} \cfrac{\sum_{w_0 \subseteq w_1} y_{w_0}\left(\mathbf{X}_{w_0}\right)}{1}}}}$$

for any $u \subseteq \{1, \ldots, N\}$, *solely in terms of the ANOVA component functions.*

COROLLARY 3.6. *The multiplicative constant, univariate, and bivariate component functions of a square-integrable function* $y : \mathbb{R}^N \to \mathbb{R}$, *obtained by setting* $u = \emptyset$, $u = \{i\}$; $i = 1, \ldots, N$, *and* $u = \{i_1, i_2\}$; $i_1 < i_2 = 1, \ldots, N$, *respectively, in* (3.6) *or* (3.11) *are*

$$(3.12) \qquad 1 + z_\emptyset = y_\emptyset,$$

$$(3.13) \qquad 1 + z_{\{i\}}(X_i) = \frac{y_\emptyset + y_{\{i\}}(X_i)}{y_\emptyset},$$

and

$$(3.14) \quad 1 + z_{\{i_1, i_2\}}(X_{i_1}, X_{i_2}) = \frac{y_\emptyset + y_{\{i_1\}}(X_{i_1}) + y_{\{i_2\}}(X_{i_2}) + y_{\{i_1, i_2\}}(X_{i_1}, X_{i_2})}{y_\emptyset \left[\dfrac{y_\emptyset + y_{\{i_1\}}(X_{i_1})}{y_\emptyset}\right]\left[\dfrac{y_\emptyset + y_{\{i_2\}}(X_{i_2})}{y_\emptyset}\right]}.$$

Remark 2. Equations (3.12), (3.13), and (3.14) can also be obtained employing the identity and first- and second-degree idempotent operators [20]. However, to obtain similar expressions for trivariate and higher-variate multiplicative component functions, an extensive amount of algebra associated with third- and higher-degree idempotent operators will be required. This is a primary reason why component functions with three or more variables have yet to be reported in the current literature. Theorem 3.4, in contrast, is simpler and, more important, provides a general expression—(3.6) or (3.11)—that is valid for a multiplicative component function of an arbitrary number of variables.

Remark 3. When $y_\emptyset$ is *zero* or is close to *zero*, (3.5) through (3.14) may fail or become ill-conditioned, raising questions about the suitability of an FDD approximation in such conditions. However, they do not necessarily imply that the FDD cannot be used. Indeed, all of these problems can be remedied by appropriately conditioning the function y. For instance, by adding a nonzero constant to y or multiplying y with a nonzero constant, (3.5) through (3.14) for the preconditioned y remain valid and well-behaved.

3.2.1. FDD approximation. The S-variate FDD approximation $\hat{y}_S(\mathbf{X})$, say, of $y(\mathbf{X})$, where $0 \leq S < N$, is obtained by truncating the right side of (3.5) at $0 \leq |u| \leq S$, yielding

$$(3.15) \qquad \hat{y}_S(\mathbf{X}) = \prod_{\substack{u \subseteq \{1,\ldots,N\} \\ 0 \leq |u| \leq S}} [1 + z_u(\mathbf{X}_u)].$$

It is important to note that the right side of (3.15) contains products of at most S-dimensional ANOVA component functions of y. Therefore, the term "S-variate" used for the FDD approximation should be interpreted in the context of including at most S-degree interaction of input variables, even though $\hat{y}_S(\mathbf{X})$ is not strictly an S-variate function.

Unlike (3.3) and (3.4), the mean and variance of $\hat{y}_S(\mathbf{X})$, respectively defined as

$$(3.16) \qquad \mathbb{E}\left[\hat{y}_S(\mathbf{X})\right] := \int_{\mathbb{R}^N} \hat{y}_S(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

and

$$(3.17) \qquad \hat{\sigma}_S^2 := \mathbb{E}\left[\left(\hat{y}_S(\mathbf{X}) - \mathbb{E}\left[\hat{y}_S(\mathbf{X})\right]\right)^2\right] := \int_{\mathbb{R}^N} \left(\hat{y}_S(\mathbf{x}) - \mathbb{E}\left[\hat{y}_S(\mathbf{X})\right]\right)^2 f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x},$$

do not produce closed-form or analytic expressions in terms of $y_\emptyset$ and σ_u^2 if S is selected arbitrarily. This is one drawback of FDD when compared with ADD. Having said so, they are easily estimated by sampling methods, such as quasi- and crude Monte Carlo simulations, or even numerical integration if N is not overly large. Not wishing to preempt subsequent discussion, these two moments can be obtained in closed form when $S = 1$, that is, for the univariate FDD approximation, to be explained in section 5.

Given a truncation S, which approximation stemming from ADD and FDD is more precise? The answer depends on the class of multivariate function being approximated. Indeed, the relative advantage or disadvantage of ADD and FDD should be judged by examining three special classes of functions: purely multiplicative functions, purely additive functions, and their mixtures, described as follows.

3.3. Purely multiplicative functions. Consider a function of pure multiplicative form

$$(3.18) \qquad y(\mathbf{X}) = \nu_0 \prod_{i=1}^N h_i(X_i),$$

where $\nu_0 \in \mathbb{R}$ is a constant and $h_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, N$, are square-integrable univariate functions with

$$\nu_i := \mathbb{E}\left[h_i(X_i)\right] := \int_{\mathbb{R}} h_i(x_i) f_i(x_i) dx_i,$$

$$\delta_i^2 := \mathbb{E}\left[\left(h_i(X_i) - \nu_i\right)^2\right] := \int_{\mathbb{R}} \left(h_i(x_i) - \nu_i\right)^2 f_i(x_i) dx_i$$

denoting their first two moments. For a function y satisfying (3.18), both ADD and FDD component functions and their respective variances can be determined with little

effort. Indeed, when (3.1b) and (3.1c) are used, the outcome is

$$(3.19) \qquad y_\emptyset = 1 + z_\emptyset = \nu_0 \prod_{i=1}^{N} \nu_i,$$

$$(3.20) \qquad y_u(\mathbf{X}_u) = \nu_0 \left(\prod_{i \notin u} \nu_i \right) \prod_{i \in u} [h_i(X_i) - \nu_i], \; \emptyset \neq u \subseteq \{1, \dots, N\},$$

whereas applying (3.19) and (3.20) to (3.6) or (3.11) results in

$$(3.21) \qquad 1 + z_u(\mathbf{X}_u) = \begin{cases} \dfrac{h_i(X_i)}{\nu_i} & \text{if } u = \{i\}, \; i = 1, \dots, N, \\ 1 & \text{if } |u| \geq 2. \end{cases}$$

The ADD component functions lead to the S-variate variance approximation

$$(3.22) \qquad \tilde{\sigma}_S^2 = \sum_{s=1}^{S} \sum_{\substack{\emptyset \neq u \subseteq \{1, \dots, N\} \\ |u| = s}} \nu_0^2 \left(\prod_{i \notin u} \nu_i^2 \right) \prod_{i \in u} \delta_i^2$$

that approaches the exact variance

$$(3.23) \qquad \sigma^2 = \nu_0^2 \left[\prod_{i=1}^{N} \left(\delta_i^2 + \nu_i^2 \right) - \prod_{i=1}^{N} \nu_i^2 \right]$$

of y as $S \to N$. In contrast, the FDD component functions, due to their multiplicative structure, yield the exact variance, that is, $\hat{\sigma}_S^2 = \sigma^2$ for any $1 \leq S \leq N$. The implication is that the univariate truncation of FDD is adequate for calculating the variance or any other probabilistic characteristics of a purely multiplicative function y. This is because such a function is exactly reproduced by its univariate FDD approximation. The same does not hold true for an ADD approximation, requiring, therefore, higher-variate truncations to approximate y progressively accurately. Clearly, FDD is more relevant than ADD for the multiplicative class of functions.

3.4. Purely additive functions. Now consider a function of purely additive form

$$(3.24) \qquad y(\mathbf{X}) = \mu_0 + \sum_{i=1}^{N} g_i(X_i).$$

Here $\mu_0 \in \mathbb{R}$ is another constant and $g_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \dots, N$, are square-integrable univariate functions with

$$\mu_i := \mathbb{E}[g_i(X_i)] := \int_{\mathbb{R}} g_i(x_i) f_i(x_i) dx_i,$$

$$\lambda_i^2 := \mathbb{E}\left[(g_i(X_i) - \mu_i)^2 \right] := \int_{\mathbb{R}} (g_i(x_i) - \mu_i)^2 f_i(x_i) dx_i$$

designating their first two moments. Employing (3.1a)–(3.1c) and (3.6), the ADD and FDD component functions of y in (3.24) are

$$(3.25) \qquad y_\emptyset = 1 + z_\emptyset = \mu_0 + \sum_{i=1}^{N} \mu_i,$$

$$(3.26) \qquad y_u(\mathbf{X}_u) = \begin{cases} g_i(X_i) - \mu_i & \text{if } u = \{i\},\ i = 1, \ldots, N, \\ 0 & \text{if } |u| \geq 2, \end{cases}$$

and

$$(3.27) \qquad 1 + z_u(\mathbf{X}_u) = \frac{\mu_0 + \left(\displaystyle\sum_{i=1}^{N} \mu_i\right) + \displaystyle\sum_{i \in u} [g_i(X_i) - \mu_i]}{\displaystyle\prod_{v \subset u} [1 + z_v(\mathbf{X}_v)]},$$

with the nonconstant FDD component functions expressed recursively. In this case, the variance due to an S-variate ADD approximation yields the exact variance

$$(3.28) \qquad \tilde{\sigma}_S^2 = \sigma^2 = \sum_{i=1}^{N} \lambda_i^2$$

for any $1 \leq S \leq N$. However, due to the complicated form of (3.27), an explicit formula for the variance of a generic S-variate FDD approximation is not possible. But given a truncation S, the corresponding variance can be estimated by sampling methods or numerical integration. Nonetheless, if a function has a purely additive structure, the univariate ADD approximation exactly reproduces that function, thereby needing at most the univariate truncation of ADD. In contrast, FDD will now require higher-variate truncations for rendering gradually accurate approximations. For purely additive functions, ADD is, therefore, more appropriate than FDD, turning the table on the latter decomposition.

3.5. Mixtures of multiplicative and additive functions. Finally, consider a blended function of the form

$$(3.29) \qquad y(\mathbf{X}) = \nu_0 \left(\prod_{i=1}^{N} h_i(X_i)\right) + \mu_0 + \sum_{i=1}^{N} g_i(X_i),$$

built on adding the purely multiplicative and purely additive pieces from (3.18) and (3.24). Depending on relative orders of these constituents, y can be dominantly multiplicative or dominantly additive or neither. Regardless, (3.1a)–(3.1c) result in its ADD component functions

$$(3.30) \quad y_\emptyset = \nu_0 \left(\prod_{i=1}^{N} \nu_i\right) + \mu_0 + \sum_{i=1}^{N} \mu_i,$$

(3.31)

$$y_u(\mathbf{X}_u) = \begin{cases} \nu_0 \left(\displaystyle\prod_{j=1, j\neq i}^{N} \nu_j\right) [h_i(X_i) - \nu_i] + g_i(X_i) - \mu_i & \text{if } u = \{i\}, i = 1, \ldots, N, \\ \nu_0 \left(\displaystyle\prod_{i \notin u} \nu_i\right) \displaystyle\prod_{i \in u} [h_i(X_i) - \nu_i] & \text{if } |u| \geq 2, \end{cases}$$

lending themselves to calculate the variance

$$
(3.32) \quad \tilde{\sigma}_S^2 =
\begin{cases}
\displaystyle\sum_{i=1}^{N} \left[\nu_0^2 \left(\prod_{j=1, j\neq i}^{N} \nu_j^2 \right) \delta_i^2 + \lambda_i^2 + 2\nu_0 \left(\prod_{j=1, j\neq i}^{N} \nu_j \right) \eta_i^2 \right] & \text{if } S = 1, \\
\displaystyle\tilde{\sigma}_1^2 + \sum_{s=2}^{S} \sum_{\substack{\emptyset\neq u\subseteq\{1,\dots,N\} \\ |u|=s}} \nu_0^2 \left(\prod_{i\notin u} \nu_i^2 \right) \prod_{i\in u} \delta_i^2 & \text{if } S \geq 2,
\end{cases}
$$

of an S-variate ADD approximation, where

$$
\eta_i^2 := \mathbb{E}\left[(h_i(X_i) - \nu_i)(g_i(X_i) - \mu_i)\right] := \int_{\mathbb{R}} (h_i(x_i) - \nu_i)(g_i(x_i) - \mu_i) f_i(x_i) dx_i
$$

is the covariance between $h_i(X_i)$ and $g_i(X_i)$. The approximate variance approaches the exact variance

$$
(3.33) \quad \sigma^2 = \nu_0^2 \left[\prod_{i=1}^{N} (\delta_i^2 + \nu_i^2) - \prod_{i=1}^{N} \nu_i^2 \right] + \sum_{i=1}^{N} \lambda_i^2 + 2\nu_0 \left(\prod_{i=1}^{N} \nu_i \right) \sum_{i=1}^{N} \frac{\eta_i^2}{\nu_i},
$$

where the first two terms are the variances of the purely multiplicative and purely additive pieces, while the third term represents the covariance between these two pieces. The formulas for $y_u(\mathbf{X}_u)$, $\tilde{\sigma}_S^2$, and σ^2 given in (3.30) through (3.33) are slightly general, shrinking to (3.19) through (3.23) for purely multiplicative functions and to (3.25) through (3.28) for purely additive functions, as expected. Although the FDD component functions can be easily derived as per (3.6), again, an explicit formula for the variance $\hat{\sigma}_S^2$ of FDD approximation remains impalpable for a generic truncation S. Unlike the two pure function classes discussed in preceding subsections, it is not obvious which approximation between ADD and FDD is better for this mixed class of functions.

3.6. Example 1. Consider two functions,

$$
(3.34) \quad y_1 = \nu_0 \prod_{i=1}^{N} X_i \text{ and } y_2 = \mu_0 + \sum_{i=1}^{N} X_i,
$$

endowed with a purely multiplicative and a purely additive structure, respectively, where X_i, $i = 1, \dots, N$, are N independent and identically distributed random variables, each following uniform distribution over [0,1], $\nu_0 = 100$, and $\mu_0 = 0$. The mean and variance of $y_1(\mathbf{X})$ are $\nu_0/2^N$ and $\nu_0^2/3^N - \nu_0^2/2^{2N}$, respectively, and of $y_2(\mathbf{X})$ are $\mu_0 + N/2$ and $N/12$, respectively. The ADD and FDD component functions of y_1 or y_2 were obtained using (3.19) through (3.21) or (3.25) through (3.27). The means and variances of ADD approximations were calculated exactly using (3.19) or (3.25) and (3.22) or (3.28), whereas the means and variances of FDD approximations [(3.16), (3.17)] were estimated by a fully symmetric multidimensional integration rule with nine generators [4]. The purpose of this example is to compare the variances of both functions obtained using various ADD and FDD approximations.

Tables 3.1 and 3.2, respectively, present the mean-squared errors by univariate ($S = 1$) to decavariate ($S = 10$) truncations of ADD for y_1 and by univariate ($S = 1$) to pentavariate ($S = 5$) truncations of FDD for y_2 when $N = 6, 7, 8, 9, 10$. The error is defined as the absolute difference between the exact (σ^2) and approximate

TABLE 3.1
ADD approximation errors[a] in variances of y_1 in Example 1.

S	N				
	6	7	8	9	10
1	0.566974	0.640558	0.703332	0.75646	0.801087
2	0.206118	0.281116	0.357219	0.43174	0.502717
3	4.5738×10^{-2}	8.1426×10^{-2}	0.126477	0.179179	0.237499
4	5.6430×10^{-3}	1.4862×10^{-2}	3.0335×10^{-2}	5.2899×10^{-2}	8.2789×10^{-2}
5	2.9700×10^{-4}	1.5496×10^{-3}	4.6969×10^{-3}	1.0806×10^{-2}	2.0905×10^{-2}
6	0	7.0437×10^{-5}	4.2391×10^{-4}	1.4518×10^{-3}	3.7149×10^{-3}
7	$-$[b]	0	1.6956×10^{-5}	1.1548×10^{-4}	4.4062×10^{-4}
8	$-$[b]	$-$[b]	0	4.1244×10^{-6}	3.1328×10^{-5}
9	$-$[b]	$-$[b]	$-$[b]	0	1.0106×10^{-6}
10	$-$[b]	$-$[b]	$-$[b]	$-$[b]	0

(a) The error is defined as the absolute difference between exact and approximate variances, divided by the exact variance.
(b) Not applicable.

TABLE 3.2
FDD approximation errors[a] in variances of y_2 in Example 1.

S	N				
	6	7	8	9	10
1	2.3436×10^{-2}	2.0641×10^{-2}	1.8420×10^{-2}	1.6620×10^{-2}	1.5134×10^{-2}
2	1.1620×10^{-3}	9.2133×10^{-4}	7.4728×10^{-4}	6.1788×10^{-4}	5.1923×10^{-4}
3	1.0918×10^{-4}	7.8170×10^{-5}	5.5198×10^{-5}	4.2094×10^{-5}	3.2728×10^{-5}
4	1.7728×10^{-5}	1.3099×10^{-5}	9.3640×10^{-6}	6.7126×10^{-6}	4.8768×10^{-6}
5	2.5480×10^{-6}	2.5579×10^{-6}	1.9704×10^{-6}	1.4152×10^{-6}	9.9971×10^{-7}

(a) The error is defined as the absolute difference between exact and approximate variances, divided by the exact variance.

($\tilde{\sigma}_S^2$ or $\hat{\sigma}_S^2$) variances, divided by the exact variance. Given a problem size (N), the errors decay with increasing S for both functions as expected, although the ADD approximations for a fixed S worsen when N grows larger. This implies that an ADD approximation becomes less precise for high-dimensional multiplicative functions. No such degradations are observed for FDD approximations of high-dimensional additive functions. Both errors vanish altogether when $S = N = 10$, as the decavariate ADD and FDD approximations coincide with y_1 and y_2, respectively. Due to the purely multiplicative and purely additive functional forms examined, their univariate FDD and univariate ADD approximations are the same as y_1 and y_2, producing respective variances exactly.

4. Hybrid dimensional decomposition. When a desired stochastic response exhibits neither a dominantly additive nor a dominantly multiplicative nature, then a mixed approach that optimally combines ADD and FDD approximations is needed. Three hybrid approximations, employing a nonlinear and two linear mixtures of ADD and FDD, are proposed.

4.1. Hybrid approximations. Given S-variate ADD and FDD approximations $\tilde{y}_S(\mathbf{X})$ and $\hat{y}_S(\mathbf{X})$, let

$$\bar{y}_S(\mathbf{X}; \alpha_S, \beta_S, \gamma_S, \ldots) := \begin{cases} y_\emptyset & \text{if } S = 0, \\ h(\tilde{y}_S(\mathbf{X}), \hat{y}_S(\mathbf{X}); \alpha_S, \beta_S, \gamma_S, \ldots) & \text{if } 1 \leq S < N, \\ y(\mathbf{X}) & \text{if } S = N, \end{cases}$$

define a general, S-variate hybrid approximation of $y(\mathbf{X})$, where h is a chosen linear or

nonlinear model function such that $\mathbb{E}[\bar{y}_S(\mathbf{X}; \alpha_S, \beta_S, \gamma_S, \ldots)] = y_\emptyset$ and $\alpha_S, \beta_S, \gamma_S, \ldots$ are the associated model parameters. With no loss of generality, the *zero*-mean functions, defined by $w(\mathbf{X}) := y(\mathbf{X}) - y_\emptyset$, $\tilde{w}_S(\mathbf{X}) := \tilde{y}_S(\mathbf{X}) - y_\emptyset$, $\hat{w}_S(\mathbf{X}) := \hat{y}_S(\mathbf{X}) - \mathbb{E}[\hat{y}_S(\mathbf{X})]$, and $\bar{w}_S(\mathbf{X}; \alpha_S, \beta_S, \gamma_S, \ldots) := \bar{y}_S(\mathbf{X}; \alpha_S, \beta_S, \gamma_S, \ldots) - y_\emptyset$, will be used for the remainder of this section. It is important to note that the expectation of the univariate FDD approximation $\hat{y}_1(\mathbf{X})$ is $y_\emptyset$, which is also the mean of $y(\mathbf{X})$. Since a higher-variate approximation cannot be worse than a lower-variate approximation, one may conjecture, in the absence of a rigorous proof, that the bivariate and higher-variate FDD approximations also yield the exact mean. However, such conjecture is neither required nor used for hybrid approximations developed here. Theorem 4.1 and Corollaries 4.2 and 4.3 describe three optimal hybrid approximations $\bar{y}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n})$, $\bar{y}_{S,l}(\mathbf{X}; \alpha_{S,l}, \beta_{S,l})$, and $\bar{y}_{S,l'}(\mathbf{X}; \alpha_{S,l'})$ for $1 \leq S < N$, each producing the exact mean $y_\emptyset$, where the subscripts n, l, and l' refer to a nonlinear and two linear models examined in this work. Their zero-mean counterparts are defined as $\bar{w}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) := \bar{y}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) - y_\emptyset$, $\bar{w}_{S,l}(\mathbf{X}; \alpha_{S,l}, \beta_{S,l}) := \bar{y}_{S,l}(\mathbf{X}; \alpha_{S,l}, \beta_{S,l}) - y_\emptyset$, and $\bar{w}_{S,l'}(\mathbf{X}; \alpha_{S,l'}) := \bar{y}_{S,l'}(\mathbf{X}; \alpha_{S,l'}) - y_\emptyset$, respectively.

THEOREM 4.1. *Given an integer $1 \leq S < N < \infty$, let $\tilde{w}_S(\mathbf{X})$ and $\hat{w}_S(\mathbf{X})$ represent zero-mean, S-variate ADD and FDD approximations with variances $\tilde{\sigma}_S^2 := \mathbb{E}[\tilde{y}_S(\mathbf{X}) - y_\emptyset]^2 = \mathbb{E}[\tilde{w}_S^2(\mathbf{X})]$ and $\hat{\sigma}_S^2 := \mathbb{E}[\hat{y}_S(\mathbf{X}) - \mathbb{E}[\hat{y}_S(\mathbf{X})]]^2 = \mathbb{E}[\hat{w}_S^2(\mathbf{X})]$, respectively, of a real-valued, zero-mean, square-integrable function $w(\mathbf{X})$. Then an optimal, nonlinear, S-variate hybrid approximation of $w(\mathbf{X})$, if it exists, is*

$$(4.1) \quad \begin{aligned} \bar{w}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) = &\alpha_{S,n}\tilde{w}_S(\mathbf{X}) + \beta_{S,n}\hat{w}_S(\mathbf{X}) \\ &+ \gamma_{S,n}[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X}) - \mathbb{E}\{\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\}], \end{aligned}$$

where $(\alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) \in \mathbb{R}^3$ is the solution of

$$(4.2)$$
$$\begin{bmatrix} \tilde{\sigma}_S^2 & \mathbb{E}[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})] & \mathbb{E}[\tilde{w}_S^2(\mathbf{X})\hat{w}_S(\mathbf{X})] \\ & \hat{\sigma}_S^2 & \mathbb{E}[\tilde{w}_S(\mathbf{X})\hat{w}_S^2(\mathbf{X})] \\ (sym.) & & \mathbb{E}[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X}) - \mathbb{E}\{\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\}]^2 \end{bmatrix} \begin{Bmatrix} \alpha_{S,n} \\ \beta_{S,n} \\ \gamma_{S,n} \end{Bmatrix}$$

$$= \begin{Bmatrix} \tilde{\sigma}_S^2 \\ \mathbb{E}[w(\mathbf{X})\hat{w}_S(\mathbf{X})] \\ \mathbb{E}[w(\mathbf{X})\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})] \end{Bmatrix}.$$

Proof. Define

$$\bar{e}_{S,n} := [w(\mathbf{X}) - \bar{w}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n})]^2$$

as the square of the difference between $w(\mathbf{X})$ and its S-variate hybrid approximation $\bar{w}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n})$. For the mean-squared error of $\bar{w}_{S,n}$ to be minimum, set

$$(4.3) \quad \frac{\partial \mathbb{E}[\bar{e}_{S,n}]}{\partial \alpha_{S,n}} = \frac{\partial \mathbb{E}[\bar{e}_{S,n}]}{\partial \beta_{S,n}} = \frac{\partial \mathbb{E}[\bar{e}_{S,n}]}{\partial \gamma_{S,n}} = 0.$$

Let $\mathcal{A}$ be an open subset of $\mathbb{R}^3$. Suppose that $\bar{e}_{S,n} : \mathbb{R}^N \times \mathcal{A} \to \mathbb{R}$ satisfies the following regularity conditions: (1) $\bar{e}_{S,n}$ is a Lebesgue-integrable function of $\mathbf{X}$ for each $(\alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) \in \mathcal{A}$; (2) for almost all $\mathbf{x} \in \mathbb{R}^N$, the partial derivatives $\partial \bar{e}_{S,n}/\partial \alpha_{S,n}$, $\partial \bar{e}_{S,n}/\partial \beta_{S,n}$, and $\partial \bar{e}_{S,n}/\partial \gamma_{S,n}$ exist for all $(\alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) \in \mathcal{A}$; and (3) there exists an integrable function $\theta : \mathbb{R}^N \to \mathbb{R}$ such that $|\partial \bar{e}_{S,n}/\partial \alpha_{S,n}| \leq \theta(\mathbf{x})$, $|\partial \bar{e}_{S,n}/\partial \beta_{S,n}| \leq \theta(\mathbf{x})$, and $|\partial \bar{e}_{S,n}/\partial \gamma_{S,n}| \leq \theta(\mathbf{x})$ for all $(\alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) \in \mathcal{A}$. Then

for all $(\alpha_{S,n}, \beta_{S,n}, \gamma_{S,n}) \in \mathcal{A}$, the differential and expectation (integral) operators in (4.3) can be interchanged, that is,

$$(4.4) \qquad \mathbb{E}\left[\frac{\partial \bar{e}_{S,n}}{\partial \alpha_{S,n}}\right] = \mathbb{E}\left[\frac{\partial \bar{e}_{S,n}}{\partial \beta_{S,n}}\right] = \mathbb{E}\left[\frac{\partial \bar{e}_{S,n}}{\partial \gamma_{S,n}}\right] = 0.$$

Substituting the expression of $\bar{w}_{S,n}(\mathbf{X}; \alpha_{S,n}, \beta_{S,n}, \gamma_{S,n})$ from (4.1) into (4.4) produces (4.2), proving the theorem. □

COROLLARY 4.2. *Neglecting the nonlinear term in* (4.1) *through* (4.4) *creates an optimal, linear, S-variate hybrid approximation*

$$(4.5) \qquad \bar{w}_{S,l}(\mathbf{X}; \alpha_{S,l}, \beta_{S,l}) = \alpha_{S,l}\tilde{w}_S(\mathbf{X}) + \beta_{S,l}\hat{w}_S(\mathbf{X})$$

of $w(\mathbf{X})$, $1 \leq S < N < \infty$, *where the optimal model parameters*

$$(4.6) \qquad \alpha_{S,l} = \frac{\hat{\sigma}_S^2\tilde{\sigma}_S^2 - \mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\mathbb{E}\left[w(\mathbf{X})\hat{w}_S(\mathbf{X})\right]}{\tilde{\sigma}_S^2\hat{\sigma}_S^2 - \left(\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)^2}$$

and

$$(4.7) \qquad \beta_{S,l} = \frac{\tilde{\sigma}_S^2\mathbb{E}\left[w(\mathbf{X})\hat{w}_S(\mathbf{X})\right] - \tilde{\sigma}_S^2\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]}{\tilde{\sigma}_S^2\hat{\sigma}_S^2 - \left(\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)^2}.$$

COROLLARY 4.3. *Neglecting the nonlinear term and constraining the sum of two remaining model parameters to be unity in* (4.1) *through* (4.4) *creates yet another optimal, linear, S-variate hybrid approximation*

$$(4.8) \qquad \bar{w}_{S,l'}(\mathbf{X}; \alpha_{S,l'}) = \alpha_{S,l'}\tilde{w}_S(\mathbf{X}) + (1 - \alpha_{S,l'})\hat{w}_S(\mathbf{X})$$

of $w(\mathbf{X})$, $1 \leq S < N < \infty$, *where the optimal model parameter*

$$(4.9) \qquad \alpha_{S,l'} = \frac{\tilde{\sigma}_S^2 + \hat{\sigma}_S^2 - \mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right] - \mathbb{E}\left[w(\mathbf{X})\hat{w}_S(\mathbf{X})\right]}{\tilde{\sigma}_S^2 + \hat{\sigma}_S^2 - 2\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]}.$$

The second linear hybrid approximation $\bar{w}_{S,l'}$ for $S = 1$ or 2 presented in Corollary 4.3 coincides with that proposed by Tunga and Demiralp [19]. However, the results of the first linear hybrid approximation $\bar{w}_{S,l}$ and nonlinear hybrid approximation $\bar{w}_{S,n}$—that is, Theorem 4.1 and Corollary 4.2—are new. Furthermore, the two linear approximations, $\bar{w}_{S,l}$ and $\bar{w}_{S,l'}$, are not the same for a general truncation $2 \leq S < N$. Indeed,

$$\alpha_{S,l} - \alpha_{S,l'} =$$

$$\frac{\left(\hat{\sigma}_S^2 - \mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)\left(\tilde{\sigma}_S^2 - \mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)\left(\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right] - \mathbb{E}\left[w(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)}{\left(\tilde{\sigma}_S^2 + \hat{\sigma}_S^2 - 2\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)\left(\tilde{\sigma}_S^2\hat{\sigma}_S^2 - \left\{\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right\}^2\right)}$$

does not vanish for arbitrary S. Nor do the model parameters of the first linear hybrid approximation, which satisfy

$$\alpha_{S,l} + \beta_{S,l} = 1 - \frac{\left(\tilde{\sigma}_S^2 - \mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)\left(\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right] - \mathbb{E}\left[w(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right)}{\tilde{\sigma}_S^2\hat{\sigma}_S^2 - \left\{\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right]\right\}^2},$$

sum to unity, as required in the second linear hybrid approximation. Equation (4.5) is endowed with two independent parameters and hence more flexibility in forming

an optimal approximation. Therefore, the first approximation is expected to produce more accurate results than the second approximation, to be illustrated in Example 2.

Remark 4. The nonlinear hybrid approximation $\bar{w}_{S,n}$ presented in Theorem 4.1 is motivated on furnishing more precise stochastic solutions than those by either version of the linear hybrid approximation. This is possible when all higher-order moments involved in (4.2) exist and the system matrix in (4.2) is invertible. The need for a nonlinear approximation becomes significant when only a univariate truncation is feasible, but not necessarily yielding sufficiently accurate statistical solutions by linear approximations. The numerical results from linear and nonlinear approximations will be contrasted in section 5.

Remark 5. Since $\mathbb{E}[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})] \leq \tilde{\sigma}_S^2\hat{\sigma}_S^2$ (Cauchy–Schwarz inequality), the submatrix formed by the first two rows and columns of the system matrix in (4.2) is invertible. Therefore, the linear hybrid approximations exist for any function y or w with a finite variance. Under appropriate conditions, the linear approximations, $\bar{w}_{S,l}$ and $\bar{w}_{S,l'}$, for a given $1 \leq S < N < \infty$, are capable of reproducing the exact solution: (1) If $w = \tilde{w}_S$, then (4.6), (4.7), and (4.9) yield $\alpha_{S,l} = \alpha_{S,l'} = 1$ and $\beta_{S,l} = 0$, and hence $\bar{w}_{S,l} = \bar{w}_{S,l'} = \tilde{w}_S = w$; (2) if $w = \hat{w}_S$, then (4.6), (4.7), and (4.9) produce $\alpha_{S,l} = \alpha_{S,l'} = 0$ and $\beta_{S,l} = 1$, and therefore $\bar{w}_{S,l} = \bar{w}_{S,l'} = \hat{w}_S = w$. In Example 1, for instance, both functions y_1 and y_2 and their respective linear, univariate HDD approximations are coincident, leading to exact variances by the HDD approximation.

4.2. Second-moment properties. Applying the expectation operator on (4.1), (4.5), and (4.8) yields the exact mean

$$(4.10) \quad \mathbb{E}\left[\bar{y}_{S,n}(\mathbf{X};\alpha_{S,n},\beta_{S,n},\gamma_{S,n})\right] = \mathbb{E}\left[\bar{y}_{S,l}(\mathbf{X};\alpha_{S,l},\beta_{S,l})\right] = \mathbb{E}\left[\bar{y}_{S,l'}(\mathbf{X};\alpha_{S,l'})\right] = y_\emptyset,$$

by all three hybrid approximations. However, their respective variances,

$$(4.11)$$
$$\begin{aligned}
\bar{\sigma}_{S,n}^2 &:= \mathbb{E}\left[\bar{w}_{S,n}^2(\mathbf{X};\alpha_{S,n},\beta_{S,n},\gamma_{S,n})\right] = \alpha_{S,n}^2\tilde{\sigma}_S^2 + \beta_{S,n}^2\hat{\sigma}_S^2 \\
&\quad + \gamma_{S,n}^2\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X}) - \mathbb{E}\left\{\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right\}\right]^2 + 2\alpha_{S,n}\beta_{S,n}\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right] \\
&\quad + 2\alpha_{S,n}\gamma_{S,n}\mathbb{E}\left[\tilde{w}_S^2(\mathbf{X})\hat{w}_S(\mathbf{X})\right] + 2\beta_{S,n}\gamma_{S,n}\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S^2(\mathbf{X})\right],
\end{aligned}$$

$$(4.12) \quad \bar{\sigma}_{S,l}^2 := \mathbb{E}\left[\bar{w}_{S,l}^2(\mathbf{X};\alpha_{S,l},\beta_{S,l})\right] = \alpha_{S,l}^2\tilde{\sigma}_S^2 + \beta_{S,l}^2\hat{\sigma}_S^2 + 2\alpha_{S,l}\beta_{S,l}\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right],$$

and

$$(4.13) \quad \begin{aligned}
\bar{\sigma}_{S,l'}^2 &:= \mathbb{E}\left[\bar{w}_{S,l'}^2(\mathbf{X};\alpha_{S,l'},\beta_{S,l'})\right] \\
&= \alpha_{S,l'}^2\tilde{\sigma}_S^2 + (1 - \alpha_{S,l'})^2\hat{\sigma}_S^2 + 2\alpha_{S,l'}(1 - \alpha_{S,l'})\mathbb{E}\left[\tilde{w}_S(\mathbf{X})\hat{w}_S(\mathbf{X})\right],
\end{aligned}$$

while approximate, emerge steadily more accurate as $S \to N$. Compared with ADD and FDD approximations, the HDD approximations proposed require expectations of various products of $\tilde{w}_S(\mathbf{X})$ and $\hat{w}_S(\mathbf{X})$ to calculate their variances. They simplify greatly for the univariate HDD approximation, to be discussed in the following section.

4.3. Example 2. Consider the blended function

$$(4.14) \quad y = y_\emptyset + \frac{\sum_{i=1}^{N} X_i - \mathbb{E}\left[\sum_{i=1}^{N} X_i\right]}{\sqrt{\mathbb{E}\left(\sum_{i=1}^{N} X_i - \mathbb{E}\left[\sum_{i=1}^{N} X_i\right]\right)^2}} + \frac{\prod_{i=1}^{N} X_i - \mathbb{E}\left[\prod_{i=1}^{N} X_i\right]}{\sqrt{\mathbb{E}\left(\prod_{i=1}^{N} X_i - \mathbb{E}\left[\prod_{i=1}^{N} X_i\right]\right)^2}}$$

TABLE 4.1
ADD, FDD, and HDD approximation errors[a] in variances of y in Example 2.

S	ADD	FDD	HDD (1st linear model)	HDD (2nd linear model)
			Approximation	
1	0.139942	9.8251×10^{-2}	2.2734×10^{-2}	2.2734×10^{-2}
2	3.9452×10^{-2}	2.8834×10^{-2}	1.5503×10^{-3}	5.5528×10^{-3}
3	5.9550×10^{-3}	3.1717×10^{-3}	1.8656×10^{-4}	5.7923×10^{-3}
4	3.7219×10^{-4}	3.7008×10^{-4}	1.4099×10^{-5}	5.0992×10^{-4}

(a) The error is defined as the absolute difference between exact and approximate
variances, divided by the exact variance.

of N random variables, X_i, $i = 1, \ldots, N$, which are independent, identical, and uniformly distributed over [0,1]. From elementary calculations, the mean and variance of $y(\mathbf{X})$ are $y_\emptyset$ and $2 + 2\sqrt{N \times 3^{N-1}/(2^{2N} - 3^N)}$, respectively. Since y in (4.14) follows the general structure of (3.29), all ADD component functions [(3.30), (3.31)] and the resultant second-moment statistics [(3.30), (3.32)] of an S-variate ADD approximation were obtained exactly. However, the mean and variance of an S-variate FDD approximation were estimated by a fully symmetric multidimensional integration rule with nine generators [4]. The variance of an S-variate HDD approximation was obtained using (4.12), where the expectation in the last term and those involved in the model parameters [(4.6), (4.7)] were calculated by the same integration rule. The objective of this example is to evaluate the accuracy of various ADD, FDD, and HDD (both linear models) approximations in calculating several probabilistic characteristics of y for $N = 5$ and $y_\emptyset = 5$.

The means of FDD approximations for $S = 1$, 2, 3, and 4, obtained using the aforementioned multidimensional rule and retaining five digits after the decimal point, are 5.00000, 5.00014, 5.00021, and 5.00002. They are practically equal to 5—the value of $y_\emptyset$—demonstrating that all four FDD approximations also produce the exact mean, at least in this example. However, a formal proof of this conjecture, applicable to FDD approximations of arbitrary functions, remains elusive.

Table 4.1 presents the errors in variances of y calculated by univariate ($S = 1$) to quadrivariate ($S = 4$) truncations of ADD, FDD, and HDD. The definition of the error is the same as in Example 1. Both ADD and FDD commit smaller errors when the truncation S increases, as expected. Since the purely additive and purely multiplicative pieces of y have been standardized, that is, they have zero means and unit variances, the ADD and FDD errors have similar orders and trends. In contrast, the HDD errors from the first linear approximation, given $1 \leq S \leq 4$, are consistently lower than the corresponding ADD or FDD errors by almost an order of magnitude. However, the same observation does not hold true for the second linear hybrid approximation, which provides more precise results than ADD or FDD only for $S = 1$ or 2, but the accuracy does not improve or even degrades for $S = 3$ and 4. The superior accuracy of the first linear approximation is attributed to two independent model parameters, compared with a single model parameter of the second linear approximation. Nonetheless, the hybrid approximation proposed is more accurate than the ADD or FDD approximation, at least, for second-moment analysis.

Does the improvement of HDD approximation extend to other probabilistic characteristics of y? Figure 4.1 depicts the probability density functions (PDFs) of y and its univariate to quadrivariate ADD, FDD, and HDD (first linear model) approximations. The PDFs, essentially the normalized histograms, were estimated from 10^6 Monte Carlo simulations of y and their various approximations. At univariate truncation ($S = 1$), neither ADD nor FDD provide acceptable results when compared with
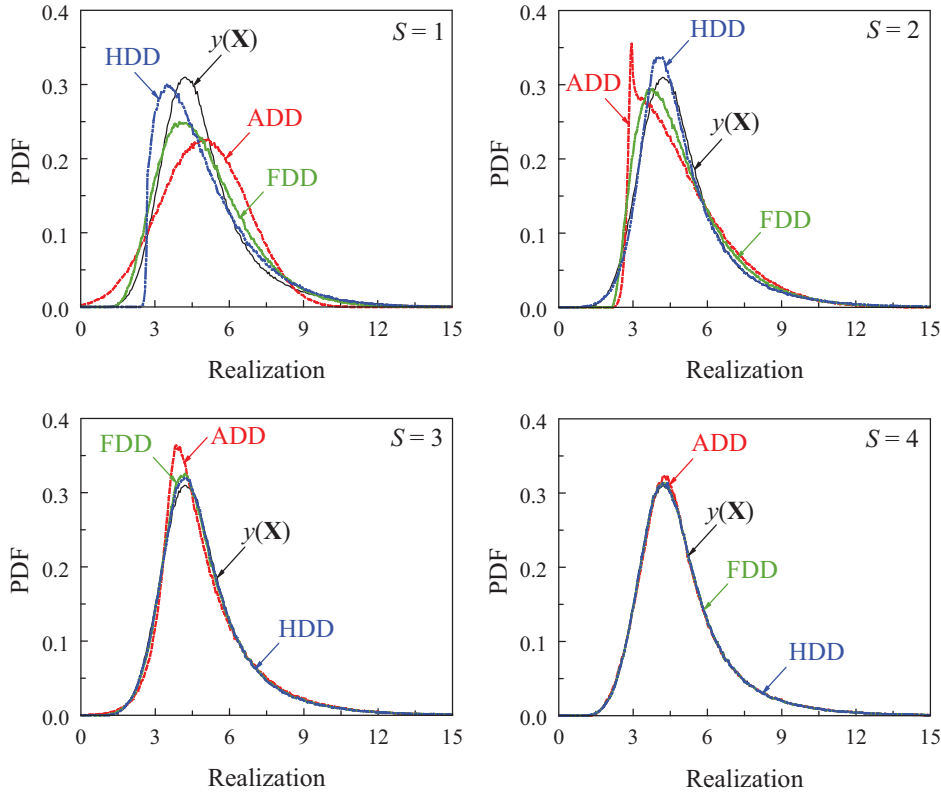
FIG. 4.1. *PDFs of y and its ADD, FDD, and HDD approximations in Example 2; note: univariate (top, left); bivariate (top, right); trivariate (bottom, left); quadrivariate (bottom, right).*

the PDF of y. In contrast, the PDF derived from the univariate or bivariate HDD approximation matches with the PDF of y better than the corresponding ADD or FDD approximation. The discrepancy between the PDFs of y and its approximation, whether ADD, FDD, or HDD, diminishes as S increases. At quadrivariate truncation ($S = 4$), all four PDFs are practically the same, even at the tail regions. Any distinction, especially between the results of y and its FDD or HDD approximation, is hardly visible to the naked eye. Judging from the rates at which the PDFs of approximations are converging to the PDF of y, the HDD approximation may also be beneficial for calculating the probability distribution of a stochastic response. This in large measure depends on the smoothness properties of the function chosen. However, strictly speaking, there does not exist a formal proof yet that an HDD approximation will always lead to an accurate calculation of probabilities. The topic merits further study.

4.4. Effective dimensions. When employing dimensional decomposition of a square-integrable function y, an important decision revolves around selecting the truncation parameter S. The truncation can be achieved by the notion of effective dimension, introduced by Caflisch, Morokoff, and Owen [2], who exploited 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 subsection, a complete set of definitions, including two new alternative ones, of effective dimension, stemming from ADD, FDD, and HDD are presented.

DEFINITION 4.4. *A square-integrable multivariate function y with finite variance $0 < \sigma^2 < \infty$ has the ADD-based effective dimension $1 \le S_{ADD} \le N$ if*

$$(4.15) \qquad S_{ADD} := \min\left\{S : 1 \le S \le N \text{ such that } |\sigma^2 - \tilde{\sigma}_S^2| \le (1-p)\sigma^2\right\},$$

the FDD-based effective dimension $1 \le S_{FDD} \le N$ if

$$(4.16) \qquad S_{FDD} := \min\left\{S : 1 \le S \le N \text{ such that } |\sigma^2 - \hat{\sigma}_S^2| \le (1-p)\sigma^2\right\},$$

and the HDD-based effective dimension $1 \le S_{HDD} \le N$ if

$$(4.17) \qquad S_{HDD} := \min\left\{S : 1 \le S \le N \text{ such that } |\sigma^2 - \bar{\sigma}_S^2| \le (1-p)\sigma^2\right\},$$

where $\bar{\sigma}_S^2 := \mathbb{E}[\bar{y}_S(\mathbf{X}; \alpha_S, \beta_S, \gamma_S, \ldots) - y_\emptyset]^2 = \mathbb{E}[\bar{w}_S^2]$ is the variance due to an S-variate (linear or nonlinear) HDD approximation of $w(\mathbf{X})$ and $0 \le p \le 1$ is a percentile threshold close to one.

The definitions of effective dimension capture the notion in which y, represented by its ADD, FDD, or HDD, is almost S_{ADD}-, S_{FDD}-, or S_{HDD}-dimensional. The knowledge of effective dimension, whether rooted in ADD, FDD, or HDD, is valuable, because ANOVA terms up to the effective dimension contribute most to the function. Among these three definitions, the ADD approximation results in the effective superposition dimension (S_{ADD}) of Caflisch, Morokoff, and Owen—a concept ideally suited to a function comprising strongly additive dimensional structure. In contrast, the effective dimension S_{FDD} proposed is appropriate when the dimensional structure is strongly multiplicative. However, the dimensional hierarchy of a multivariate function in general is neither dominantly additive nor dominantly multiplicative. Therefore, the effective dimension S_{HDD} should be viewed as a generalized effective dimension that is applicable to a broader class of functions than the ADD- and FDD-based effective dimensions. Furthermore, given a truncation S, an HDD approximation can be more precise than both ADD and FDD approximations. In which case, the HDD-based effective dimension is expected to be lower than or equal to the ADD- or FDD-based effective dimensions. It is, therefore, possible to optimally blend lower-variate ANOVA terms in HDD to contribute more to the function than by the same ANOVA terms in ADD or FDD alone. See Example 3 for numerical results.

Remark 6. Caflisch, Morokoff, and Owen used the 99th percentile for p, but it can be treated as a threshold parameter linked to the desired accuracy of a stochastic solution. They also introduced effective dimension in the truncation sense, referred to as effective truncation dimension, which is appropriate when some variables are more important than others in an ordered set of the additive decomposition. It is possible to define similar effective dimensions grounded on FDD and HDD approximations as well. In contrast, the effective superposition dimensions in (4.15), (4.16), and (4.17) determine whether the low-variate component functions of dimensional decomposition dominate the function and are appropriate when all variables are equally important. For truly high-dimensional problems, all variables contribute to a function value; therefore, the effective superposition dimensions examined are more useful than the effective truncation dimension.

4.5. Example 3. Consider the functions y_1 and y_2 described by (3.34) in Example 1 and y described by (4.14) in Example 2. The purpose of this example is to compare the effective dimensions of these three functions, stemming from various dimensional decompositions and demonstrate the benefit of the HDD-based effective dimension.

TABLE 4.2
Effective dimensions of functions in Examples 1 and 2 for $p = 0.99$.

| | Function y_1 in (3.34) | | | Function y_2 in (3.34) | | | Function y in (4.14) | | |
N	S_{ADD}	S_{FDD}	S_{HDD}	S_{ADD}	S_{FDD}	S_{HDD}	S_{ADD}	S_{FDD}	S_{HDD}
5	—[a]	—[a]	—[a]	—[a]	—[a]	—[a]	3	3	2
6	4	1	1	1	2	1	—[a]	—[a]	—[a]
7	5	1	1	1	2	1	—[a]	—[a]	—[a]
8	5	1	1	1	2	1	—[a]	—[a]	—[a]
9	6	1	1	1	2	1	—[a]	—[a]	—[a]
10	6	1	1	1	2	1	—[a]	—[a]	—[a]

(a) Not applicable as $6 \leq N \leq 10$ in Example 1 and $N = 5$ in Example 2.

Table 4.2 enumerates S_{ADD}, S_{FDD}, and S_{HDD} of y_1 and y_2 for $N = 6, 7, 8, 9, 10$ and of y for $N = 5$ when $p = 0.99$. They are easily obtained from respective definitions of effective dimensions given by (4.15), (4.16), and (4.17), and the second-moment errors listed in Tables 3.1 and 3.2. For all functions or problem sizes examined, HDD leads to least effective dimensions. This is because the variance of a multivariate function, given the largest dimension of ANOVA component functions retained, is better estimated by HDD than ADD or FDD approximations.

5. Univariate approximations. The univariate truncation ($S = 1$) of each decomposition, whether ADD, FDD, or HDD, constitutes the simplest possible approximation of y other than its mean. In which case, the approximate solutions of statistical moments and error analysis simplify and can be obtained or conducted in closed-form or analytically. They are of huge practical interest, because modeling and simulation of large-scale complex systems carry a heavy computational price tag, for which only a univariate truncation is feasible.

5.1. ADD and FDD error analyses. Consider the univariate ADD approximation

$$(5.1) \qquad \tilde{y}_1(\mathbf{X}) = y_\emptyset + \sum_{i=1}^{N} y_{\{i\}}(X_i)$$

and the univariate FDD approximation

$$(5.2) \qquad \hat{y}_1(\mathbf{X}) = (1 + z_\emptyset)\prod_{i=1}^{N} \left[1 + z_{\{i\}}(X_i)\right] = y_\emptyset\prod_{i=1}^{N}\frac{y_\emptyset + y_{\{i\}}(X_i)}{y_\emptyset}$$

of $y(\mathbf{X})$, obtained by fixing $S = 1$ in (3.2) and (3.15), respectively. Applying the expectation operator on $\tilde{y}_1(\mathbf{X})$, $\hat{y}_1(\mathbf{X})$, and their squares, the means

$$\mathbb{E}[\tilde{y}_1(\mathbf{X})] = \mathbb{E}[\hat{y}_1(\mathbf{X})] = \mathbb{E}[y(\mathbf{X})] := y_\emptyset$$

are identical and exact, although the respective variances,

$$(5.3) \qquad \tilde{\sigma}_1^2 := \mathbb{E}\left[(\tilde{y}_1(\mathbf{X}) - y_\emptyset)^2\right] = \sum_{i=1}^{N} \sigma_{\{i\}}^2 \text{ and}$$

$$(5.4) \qquad \hat{\sigma}_1^2 := \mathbb{E}\left[(\hat{y}_1(\mathbf{X}) - y_\emptyset)^2\right] = y_\emptyset^2\left[\prod_{i=1}^{N}\left(1 + \frac{\sigma_{\{i\}}^2}{y_\emptyset^2}\right) - 1\right],$$

are approximate, where $\sigma_{\{i\}}^2 := \mathbb{E}[y_{\{i\}}^2(X_i)]$ is the variance of the *zero*-mean, univariate ADD component function $y_{\{i\}}$, $i = 1, \ldots, N$. The last expression of (5.4) is obtained from the realization that $\mathbb{E}[\{y_\emptyset + y_{\{i\}}(X_i)\}/y_\emptyset] = 1$ and $\mathbb{E}[\{y_\emptyset + y_{\{i\}}(X_i)\}^2/y_\emptyset^2] = 1 + \sigma_{\{i\}}^2/y_\emptyset^2$. Comparing (5.3) and (5.4), $\hat{\sigma}_1^2 \to \tilde{\sigma}_1^2$ when $\sigma_{\{i\}}^2/y_\emptyset^2 \to 0$ or when $\sigma_{\{i\}}^2/y_\emptyset^2 < 1$ and $N \to \infty$. See the original work of Hoeffding [8] for the existence of ADD for $N \to \infty$.

Given a general square-integrable function y, are these univariate approximations adequate? Which approximation between ADD and FDD is more precise? The resolution of these questions depends on error analysis, where

$$(5.5) \qquad \tilde{e}_1 := \mathbb{E}\left[y(\mathbf{X}) - \tilde{y}_1(\mathbf{X})\right]^2 \text{ and } \hat{e}_1 := \mathbb{E}\left[y(\mathbf{X}) - \hat{y}_1(\mathbf{X})\right]^2$$

define two second-moment errors, stemming from univariate ADD and FDD approximations of y. Applying (3.1a), (5.1), (5.2) into (5.5) and invoking Propositions 3.1 and 3.2,

$$(5.6) \qquad \tilde{e}_1 = \sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \sigma_u^2$$

is a sum of variance terms contributed by bivariate and higher-variate component functions of y, whereas
(5.7)

$$\hat{e}_1 = \sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \sigma_u^2 + y_\emptyset^2 \, \mathbb{E}\left[\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \prod_{i \in u} \frac{y_{\{i\}}(X_i)}{y_\emptyset}\right]^2$$

$$- 2y_\emptyset \sum_{s=2}^{N} \mathbb{E}\left[\left\{\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \prod_{i \in u} \frac{y_{\{i\}}(X_i)}{y_\emptyset}\right\}\left\{\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} y_u(\mathbf{X}_u)\right\}\right]$$

includes two additional terms, comprising expectations of products of three or more component functions of y. Since the first term of (5.7) is the same as $\tilde{e}_1$, the univariate FDD approximation incurs less (more) error than the univariate ADD approximation if the second term is smaller (larger) than the third term. Therefore, the condition

$$\mathbb{E}\left[\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \prod_{i \in u} \frac{y_{\{i\}}(X_i)}{y_\emptyset}\right]^2$$

$$< \frac{2}{y_\emptyset} \sum_{s=2}^{N} \mathbb{E}\left[\left\{\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \prod_{i \in u} \frac{y_{\{i\}}(X_i)}{y_\emptyset}\right\}\left\{\sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} y_u(\mathbf{X}_u)\right\}\right]$$

determines when the univariate FDD approximation is more precise than the uni-

variate ADD approximation and vice versa. When $N = 2$, for instance, $\hat{e}_1 < \tilde{e}_1$ if $\mathbb{E}[y_{\{1\}}^2(X_1)y_{\{2\}}^2(X_2)] < 2y_\emptyset \mathbb{E}[y_{\{1\}}(X_1)y_{\{2\}}(X_2)y_{\{12\}}(X_1, X_2)]$ and vice versa.

Remark 7. The ADD approximation $\tilde{y}_1(\mathbf{X})$ is called univariate because (5.1) comprises a sum of at most univariate component functions, describing only the main effect of $\mathbf{X}$. In contrast, the FDD approximation $\hat{y}_1(\mathbf{X})$ in (5.2) contains products of various univariate functions. Therefore, some effects of interactions between two input variables X_i and X_j, $i \neq j$, subsist in $\hat{y}_1(\mathbf{X})$. For example, consider a function $y = y_\emptyset + y_{\{1\}}(X_1) + y_{\{2\}}(X_2) + y_{\{1\}}(X_1)y_{\{2\}}(X_2)/y_\emptyset$ of two variables, containing a sum and a product of its univariate ANOVA component functions. The univariate ADD approximation, $\tilde{y}_{1,m} = y_\emptyset + y_{\{1\}}(X_1) + y_{\{2\}}(X_2)$, captures only the main effects of X_1 and X_2, and may produce nonnegligible errors if the product term of y is significant. On the other hand, the univariate FDD approximation, $\hat{y}_1 = (1 + z_\emptyset)[1 + z_{\{1\}}(X_1)][1 + z_{\{2\}}(X_2)] = y_\emptyset + y_{\{1\}}(X_1) + y_{\{2\}}(X_2) + y_{\{1\}}(X_1)y_{\{2\}}(X_2)/y_\emptyset$, obtained using the relationships in (3.12) and (3.13), exactly reproduces y, thereby capturing not only the main effects, but also the interactive effect of input variables. Similar conditions prevail for a function of an arbitrary, but finite, number of variables, provided that the higher-variate ANOVA component functions are products of univariate ANOVA component functions. However, the ANOVA component functions with distinct dimensions are unrelated in general. Therefore, the error committed by an FDD approximation may or may not be lower than that by an ADD approximation. Nonetheless, the term "univariate" used in this paper for the FDD approximation should be interpreted in the context of including at most univariate component functions, not necessarily preserving only the main effects.

There exist two special cases where one of the two univariate approximations does not perpetrate any error. First, consider a purely additive function y, where its zero-variate and univariate component functions are arbitrary, but its bivariate and higher-variate ANOVA component functions vanish, that is, $y_u(\mathbf{X}_u) = 0$ for $2 \leq |u| \leq N$. In this case, (5.6) and (5.7) yield $\tilde{e}_1 = 0$, while

$$\hat{e}_1 = y_\emptyset^2 \, \mathbb{E}\left[\sum_{\substack{s=2}}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \prod_{i \in u} \frac{y_{\{i\}}(X_i)}{y_\emptyset}\right]^2 > 0,$$

indicating superiority of univariate ADD over univariate FDD approximations. Second, consider a function y, where bivariate and higher-variate ANOVA component functions are products of zero-variate and univariate component functions, distributed as follows: $y_u(\mathbf{X}_u) = y_\emptyset \prod_{i \in u \subseteq \{1,\ldots,N\}} y_{\{i\}}(X_i)/y_\emptyset]$ for $2 \leq |u| \leq N$. In the latter case, (5.6) and (5.7) deliver

$$\tilde{e}_1 = y_\emptyset^2 \sum_{\substack{s=2}}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ |u|=s}} \mathbb{E}\left[\prod_{i \in u} \frac{y_{\{i\}}^2(X_i)}{y_\emptyset^2}\right] > 0,$$

but $\hat{e}_1 = 0$, reversing the trend of the previous case.

5.2. HDD. The S-variate hybrid approximations and their statistics proposed in the preceding section also simplify for univariate truncations of ADD and FDD.

Both linear and nonlinear hybrid approximations are described by the following two propositions.

PROPOSITION 5.1. *A linear, univariate HDD approximation of $w(\mathbf{X})$, obtained by setting $S = 1$ in (4.5) through (4.7), is*

$$(5.8) \qquad \bar{w}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l}) = \alpha_{S,l}\tilde{w}_1(\mathbf{X}) + \beta_{1,l}\hat{w}_1(\mathbf{X}),$$

where the model parameters

$$(5.9) \qquad \alpha_{1,l} = \frac{\hat{\sigma}_1^2 - \mathbb{E}\left[w(\mathbf{X})\hat{w}_1(\mathbf{X})\right]}{\hat{\sigma}_1^2 - \tilde{\sigma}_1^2} \ and \ \beta_{1,l} = \frac{\mathbb{E}\left[w(\mathbf{X})\hat{w}_1(\mathbf{X})\right] - \tilde{\sigma}_1^2}{\hat{\sigma}_1^2 - \tilde{\sigma}_1^2}.$$

The result in Proposition 5.1 is obtained using the relationships, $\mathbb{E}[w(\mathbf{X})\tilde{w}_1(\mathbf{X})] = \mathbb{E}[\tilde{w}_1(\mathbf{X})\hat{w}_1(\mathbf{X})] = \mathbb{E}[\tilde{w}_1^2(\mathbf{X})] =: \tilde{\sigma}_1^2$, that stem from Propositions 3.1 and 3.2. The remaining expectation

$$(5.10) \qquad \mathbb{E}\left[w(\mathbf{X})\hat{w}_1(\mathbf{X})\right] = \mathbb{E}\left[y(\mathbf{X})\hat{y}_1(\mathbf{X})\right] - y_\emptyset^2$$

cannot be reduced further as it involves the original function y or w, but it can be estimated by sampling methods or numerical integration.

Remark 8. The two parameters of the first linear model, described by (5.8) and (5.9), add up to *one*. This is due to special properties of $\tilde{y}_1(\mathbf{X})$ and $\hat{y}_1(\mathbf{X})$ discussed in subsection 5.1. Therefore, the second linear model, described by (4.8) and (4.9) at univariate truncation ($S = 1$), is redundant, as it leads to the same solution of the first linear model. However, due to complicated forms of $\hat{y}_S(\mathbf{X})$, where $2 \leq S < N$, the same relationship does not hold for two generic, linear, S-variate HDD approximations. See Corollaries 4.2 and 4.3 for further insights.

The mean of $\bar{w}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})$ is *zero* and, therefore, $\mathbb{E}[\bar{y}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})] = y_\emptyset$, matching the exact mean. The variance of $\bar{w}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})$ or $\bar{y}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})$ is

$$
\begin{aligned}
(5.11) \qquad \bar{\sigma}_{1,l}^2 &:= \mathbb{E}\left[\bar{w}_{1,l}^2(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})\right] \\
&= \left(\alpha_{1,l}^2 + 2\alpha_{1,l}\beta_{1,l}\right)\tilde{\sigma}_1^2 + \beta_{1,l}^2\hat{\sigma}_1^2 \\
&= \left(2\alpha_{1,l} - \alpha_{1,l}^2\right)\tilde{\sigma}_1^2 + \left(1 - \alpha_{1,l}\right)^2\hat{\sigma}_1^2,
\end{aligned}
$$

a linear combination of variances from univariate ADD and FDD approximations.

PROPOSITION 5.2. *A nonlinear, univariate HDD approximation of $w(\mathbf{X})$, obtained by setting $S = 1$ in (4.1) through (4.4), is*

$$\bar{w}_{1,n}(\mathbf{X}; \alpha_{1,n}, \beta_{1,n}, \gamma_{1,n}) = \alpha_{1,n}\tilde{w}_1(\mathbf{X}) + \beta_{1,n}\hat{w}_1(\mathbf{X}) + \gamma_{1,n}\left[\tilde{w}_1(\mathbf{X})\hat{w}_1(\mathbf{X}) - \tilde{\sigma}_1^2\right],$$

where the model parameters are the solution of

$$
(5.12)
\begin{bmatrix}
\tilde{\sigma}_1^2 & \tilde{\sigma}_1^2 & \mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1(\mathbf{X})\right] \\
 & \hat{\sigma}_1^2 & \mathbb{E}\left[\tilde{w}_1(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right] \\
(sym.) & & \mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right] - \tilde{\sigma}_1^4
\end{bmatrix}
\begin{Bmatrix}
\alpha_{1,n} \\
\beta_{1,n} \\
\gamma_{1,n}
\end{Bmatrix}
=
\begin{Bmatrix}
\tilde{\sigma}_1^2 \\
\mathbb{E}\left[w(\mathbf{X})\hat{w}_1(\mathbf{X})\right] \\
\mathbb{E}\left[w(\mathbf{X})\tilde{w}_1(\mathbf{X})\hat{w}_1(\mathbf{X})\right]
\end{Bmatrix}.
$$

Starting from (5.1) and (5.2) and applying Propositions 3.1 and 3.2, the additional expectations involved in (5.12) are

$$\mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1(\mathbf{X})\right] = \frac{2}{y_\emptyset}\sum_{i_1=1}^{N-1}\sum_{i_2=i_1+1}^{N}\sigma_{\{i_1\}}^2\sigma_{\{i_2\}}^2 + \sum_{i=1}^{N}\mathbb{E}\left[y_{\{i\}}^3(X_i)\right],$$

$$\mathbb{E}\left[\tilde{w}_1(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right] = \left(\hat{\sigma}_1^2 + y_\emptyset^2\right)\sum_{i=1}^{N}\frac{2y_\emptyset\sigma_{\{i\}}^2 + \mathbb{E}\left[y_{\{i\}}^3(X_i)\right]}{\sigma_{\{i\}}^2 + y_\emptyset^2} - 2y_\emptyset\tilde{\sigma}_1^2,$$

$$\mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right] = \left(\hat{\sigma}_1^2 + y_\emptyset^2\right)\left(\sum_{i=1}^{N}\frac{y_\emptyset^2\sigma_{\{i\}}^2 + 2y_\emptyset\mathbb{E}\left[y_{\{i\}}^3(X_i)\right] + \mathbb{E}\left[y_{\{i\}}^4(X_i)\right]}{\sigma_{\{i\}}^2 + y_\emptyset^2}\right.$$
$$\left. + 2\sum_{i_1=1}^{N-1}\sum_{i_2=i_1+1}^{N}\frac{(2y_\emptyset\sigma_{\{i_1\}}^2 + \mathbb{E}[y_{\{i_1\}}^3(X_{i_1})])(2y_\emptyset\sigma_{\{i_2\}}^2 + \mathbb{E}[y_{\{i_2\}}^3(X_{i_2})])}{(\sigma_{\{i_1\}}^2 + y_\emptyset^2)(\sigma_{\{i_2\}}^2 + y_\emptyset^2)}\right)$$
$$+ y_\emptyset\tilde{\sigma}_1^2 - 2y_\emptyset\sum_{i=1}^{N}\mathbb{E}\left[y_{\{i\}}^3(X_i)\right] - 4\sum_{i_1=1}^{N-1}\sum_{i_2=i_1+1}^{N}\sigma_{\{i_1\}}^2\sigma_{\{i_2\}}^2,$$

and

(5.13) $\mathbb{E}\left[w(\mathbf{X})\tilde{w}_1(\mathbf{X})\hat{w}_1(\mathbf{X})\right] = \mathbb{E}\left[y(\mathbf{X})\tilde{y}_1(\mathbf{X})\hat{y}_1(\mathbf{X})\right] - y_\emptyset\mathbb{E}\left[y(\mathbf{X})\hat{y}_1(\mathbf{X})\right] - 2y_\emptyset\tilde{\sigma}_1^2,$

expressed in terms of $y_\emptyset$ and various moments of univariate ANOVA component functions. Compared with the linear, hybrid model, however, they require third- and fourth-order moments that must be furnished. Similar to (5.10), (5.13) also involves w or y and cannot be reduced further.

Since the mean of $\bar{w}_{1,n}(\mathbf{X};\alpha_{1,n},\beta_{1,n},\gamma_{1,n})$ is *zero*, $\mathbb{E}[\bar{y}_{1,n}(\mathbf{X};\alpha_{1,n},\beta_{1,n},\gamma_{1,n})] = y_\emptyset$, matching the exact mean as well. However, the variance of $\bar{w}_{1,n}(\mathbf{X};\alpha_{1,n},\beta_{1,n},\gamma_{1,n})$ or $\bar{y}_{1,n}(\mathbf{X};\alpha_{1,n},\beta_{1,n},\gamma_{1,n})$ is

$$\bar{\sigma}_{1,n}^2 := \mathbb{E}\left[\bar{w}_{1,n}^2(\mathbf{X};\alpha_{1,n},\beta_{1,n},\gamma_{1,n})\right]$$

(5.14)
$$= \alpha_{1,n}^2\tilde{\sigma}_1^2 + \beta_{1,n}^2\hat{\sigma}_1^2 + \gamma_{1,n}^2\left(\mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right] - \tilde{\sigma}_1^4\right) + 2\alpha_{1,n}\beta_{1,n}\tilde{\sigma}_1^2$$
$$+ 2\alpha_{1,n}\gamma_{1,n}\mathbb{E}\left[\tilde{w}_1^2(\mathbf{X})\hat{w}_1(\mathbf{X})\right] + 2\beta_{1,n}\gamma_{1,n}\mathbb{E}\left[\tilde{w}_1(\mathbf{X})\hat{w}_1^2(\mathbf{X})\right],$$

a nonlinear combination of variances and higher-order moments from univariate ADD and FDD approximations. It is trivial to show that (5.14) reduces to (5.11) if $\alpha_{1,n} = \alpha_{1,l}$, $\beta_{1,n} = \beta_{1,l}$, and $\gamma_{1,n} = 0$.

5.3. HDD error analysis. For the univariate truncation, which approximation stemming from ADD, FDD, and HDD is most accurate? Lemma 5.3 and Theorem 5.4 demonstrate that the HDD approximation commits the lowest error.

LEMMA 5.3. *The variance of the univariate FDD approximation is greater than or equal to the variance of the univariate ADD approximation, that is, $\hat{\sigma}_1^2 \geq \tilde{\sigma}_1^2$.*

Proof. From (5.4),

$$\hat{\sigma}_1^2 = y_\emptyset^2\left[\prod_{i=1}^{N}\left(1 + \frac{\sigma_{\{i\}}^2}{y_\emptyset^2}\right) - 1\right]$$
$$= \sum_{i=1}^{N}\sigma_{\{i\}}^2 + y_\emptyset^2\sum_{s=2}^{N}\sum_{\substack{\emptyset\neq u\subseteq\{1,\ldots,N\}\\|u|=s}}\prod_{i\in u}\frac{\sigma_{\{i\}}^2}{y_\emptyset^2}$$
$$\geq \tilde{\sigma}_1^2,$$

where the last line follows from (5.3) and the recognition that the second term of the second line is nonnegative. □

THEOREM 5.4. *Let* $\tilde{e}_1 := \mathbb{E}\left[y(\mathbf{X}) - \tilde{y}_1(\mathbf{X})\right]^2$, $\hat{e}_1 := \mathbb{E}\left[y(\mathbf{X}) - \hat{y}_1(\mathbf{X})\right]^2$, *and* $\bar{e}_{1,l} := \mathbb{E}\left[y(\mathbf{X}) - \bar{y}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})\right]^2$ *define the mean-squared errors committed by the univariate ADD, univariate FDD, and univariate HDD (linear) approximations, respectively, of a real-valued, square-integrable function y. Then*

$$\bar{e}_{1,l} \leq \tilde{e}_1, \ \bar{e}_{1,l} \leq \hat{e}_1.$$

Proof. From (5.3) and (5.6),

$$(5.15) \qquad \tilde{e}_1 = \sigma^2 - \sum_{i=1}^{N} \sigma_{\{i\}}^2 = \sigma^2 - \tilde{\sigma}_1^2.$$

Since y, $\hat{y}_1$, and $\bar{y}_{1,l}$ have the same mean,

$$(5.16) \qquad \begin{aligned} \hat{e}_1 &= \mathbb{E}\left[w(\mathbf{X}) - \hat{w}_1(\mathbf{X})\right]^2 \\ &= \sigma^2 + \hat{\sigma}_1^2 - 2\mathbb{E}\left[w(\mathbf{X})\hat{w}_1(\mathbf{X})\right] \\ &= \sigma^2 - 2\alpha_{1,l}\tilde{\sigma}_1^2 - \left(1 - 2\alpha_{1,l}\right)\hat{\sigma}_1^2, \end{aligned}$$

$$(5.17) \qquad \begin{aligned} \bar{e}_{1,l} &= \mathbb{E}\left[w(\mathbf{X}) - \bar{w}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})\right]^2 \\ &= \sigma^2 + \bar{\sigma}_{1,l}^2 - 2\mathbb{E}\left[w(\mathbf{X})\bar{w}_{1,l}(\mathbf{X}; ; \alpha_{1,l}, \beta_{1,l})\right] \\ &= \sigma^2 + \left(2\alpha_{1,l} - \alpha_{1,l}^2\right)\tilde{\sigma}_1^2 + \left(1 - \alpha_{1,l}\right)^2 \hat{\sigma}_1^2 \\ &\quad - 2\mathbb{E}\left[w(\mathbf{X})\left\{\alpha_{1,l}\tilde{w}_1(\mathbf{X}) + (1 - \alpha_{1,l})\hat{w}_1(\mathbf{X})\right\}\right] \\ &= \sigma^2 - \left(2\alpha_{1,l} - \alpha_{1,l}^2\right)\tilde{\sigma}_1^2 - \left(1 - \alpha_{1,l}\right)^2 \hat{\sigma}_1^2. \end{aligned}$$

The third line of (5.16) is obtained using (5.9). In (5.17), the third line is derived using (5.8), (5.9), and (5.11), whereas the fourth line is attained using (5.9) and the understanding that $\mathbb{E}[w(\mathbf{X})\tilde{w}_1(\mathbf{X})] = \tilde{\sigma}_1^2$. Subtracting each of (5.15) and (5.16) from (5.17) yields

$$\begin{aligned} \bar{e}_{1,l} - \tilde{e}_1 &= -\left(1 - \alpha_{1,l}\right)^2 \left(\hat{\sigma}_1^2 - \tilde{\sigma}_1^2\right) \leq 0, \\ \bar{e}_{1,l} - \hat{e}_1 &= -\alpha_{1,l}^2 \left(\hat{\sigma}_1^2 - \tilde{\sigma}_1^2\right) \leq 0, \end{aligned}$$

where the inequalities follow from Lemma 5.3, completing the proof. □

Although the result of Theorem 5.4 is expected, presenting the theorem and a formal proof is appropriate, given that such a result has yet to appear in the literature. It is less simple to follow suit for the univariate nonlinear hybrid approximation or for a general S-variate hybrid approximation.

5.4. Example 4. Consider the function

$$(5.18) \qquad y = \left[\frac{2}{N}\sum_{i=1}^{N} X_i\right]^m$$

of N independent, identical, and uniformly distributed random variables X_i, $i = 1, \dots, N$ over [0,1], where $N = 10$ and $m \in \mathbb{N}$ is an exponent. The function y in (5.18) is purely additive when $m = 1$, but it transitions from strongly additive to strongly multiplicative as m grows larger. The objective of this example is to compare univariate ADD, univariate FDD, and univariate HDD (linear and nonlinear) approximations, that is, $\tilde{y}_1(\mathbf{X})$, $\hat{y}_1(\mathbf{X})$, $\bar{y}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l}) = \bar{y}_{1,l'}(\mathbf{X}; \alpha_{1,l'})$, and
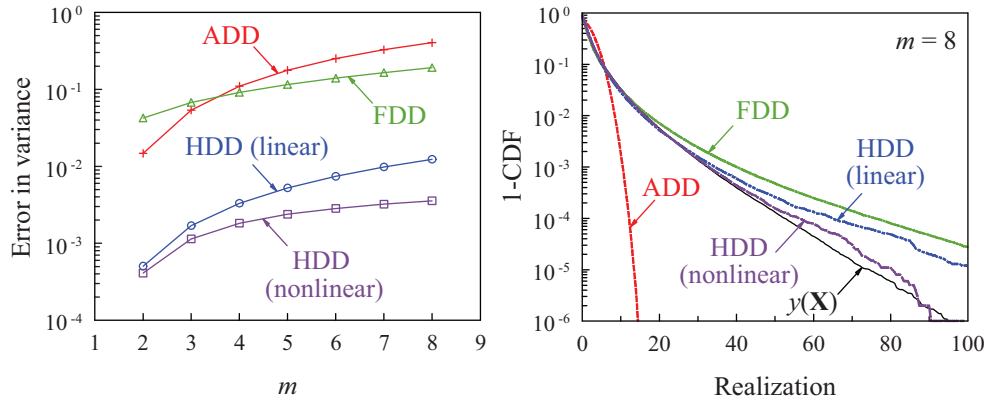
FIG. 5.1. *Results of univariate ADD, FDD, and HDD approximations of y in Example 4; note: mean-squared errors (left); complementary CDF (right).*

$\bar{y}_{1,n}(\mathbf{X}; \alpha_{1,n}, \beta_{1,n}, \gamma_{1,n})$ in calculating the variance and rare-event probabilities of $y(\mathbf{X})$ for $m = 2, 3, 4, 5, 6, 7, 8$. The second-moment properties of $y(\mathbf{X})$, given m, were calculated exactly. The variances of $\tilde{y}_1(\mathbf{X})$ [(5.3)], $\hat{y}_1(\mathbf{X})$ [(5.4)], and $\bar{y}_{1,l}(\mathbf{X}; \alpha_{1,l}, \beta_{1,l})$ [(5.11)], including all univariate ADD and FDD components functions, were calculated analytically. The expectations $\mathbb{E}[w(\mathbf{X})\hat{w}(\mathbf{X})]$ and $\mathbb{E}[w(\mathbf{X})\tilde{w}(\mathbf{X})\hat{w}(\mathbf{X})]$, involved in determining the optimal model parameters [(5.9), (5.12)] and the variance of $\bar{y}_{1,n}(\mathbf{X}; \alpha_{1,n}, \beta_{1,n}, \gamma_{1,n})$ [(5.14)], were estimated by a fully symmetric multidimensional integration rule with nine generators [4].

Figure 5.1 (left) plots the respective errors in calculating the variance of y by four univariate ($S = 1$) truncations of ADD, FDD, HDD (linear), and HDD (nonlinear) against m. The definition of the error is the same as in Example 1 or 2. When $m = 2$ or 3, the function is still strongly additive and, therefore, the univariate ADD approximation is better than the univariate FDD approximation. But the trend reverses when $4 \leq m \leq 8$, the range of higher values examined. This is because the function switches from dominantly additive ($m \leq 3$) to dominantly multiplicative ($m > 3$) as m increases. Nonetheless, for all values of m considered, the univariate HDD approximation, whether linear or nonlinear, commits lower errors than either univariate ADD or univariate FDD approximation. The nonlinear version of the univariate HDD approximation is even more precise than its linear counterpart—a trend that becomes pronounced when m is larger. However, the improvement of the nonlinear model comes with a price, as the model requires calculations of higher-order moments, alluded to in the preceding subsection. Nonetheless, the HDD models proposed provide a means to calculate the second-moment properties more accurately than ADD and FDD approximations.

Figure 5.1 (right) displays the complementary cumulative distribution functions (CDFs) of y for $m = 8$ and its four univariate approximations, each obtained by 10 million Monte Carlo samples. The chosen scale of the vertical axis is logarithmic to delineate rare-event probabilities that are commonly used for reliability analysis of complex systems. Compared with univariate ADD and FDD approximations, both variants of the univariate HDD approximation developed provide better estimates of the tail probabilistic characteristics of y. The ADD approximation significantly underestimates the tail behavior, whereas the FDD approximation overestimates the complementary CDF by a moderate amount. The nonlinear HDD approximation is
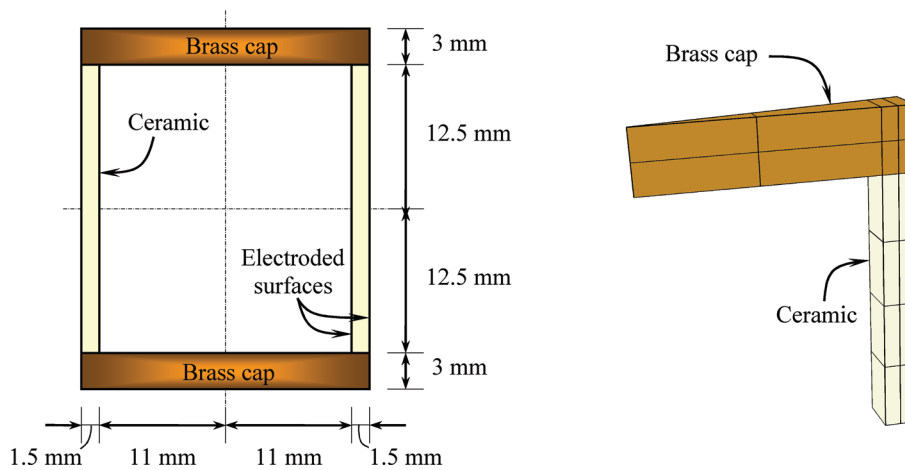
FIG. 5.2. *A piezoelectric transducer; note: geometry (left); finite-element discrete model (right).*

more precise than the linear version, especially when the probabilities are very low. Indeed, an HDD approximation is desirable for uncertainty quantification of high-dimensional, complex systems, where only univariate truncations are feasible, but not necessarily producing adequate accuracy by either ADD or FDD approximation alone.

5.5. Example 5. The final example is motivated on solving a practical engineering problem, which entails eigenspectrum analysis of a piezoelectric transducer commonly used for converting electrical pulses to mechanical vibrations, and vice versa. Figure 5.2 (left) shows a 25-mm-diameter cylinder made of a piezoelectric ceramic PZT4 (lead zirconate titanate) with brass end caps. The thicknesses of the transducer and end caps are 1.5 mm and 3 mm, respectively. The cylinder, 25 mm long, was electroded on both the inner and outer surfaces. The random variables include (1) ten nonzero constants defining elasticity, piezoelectric stress coefficients, and dielectric properties of PZT4; (2) elastic modulus and Poisson's ratio of brass; and (3) mass densities of brass and PZT4 [14]. The statistical properties of all 14 random variables are listed in Table 5.1. The random variables are independent and follow lognormal distributions. Due to axisymmetry, a 20-noded finite-element discrete model of a slice of the transducer, shown in Table 5.2 (right), was created. The objective is to evaluate various univariate approximations in calculating the second-moment properties of the natural frequencies of the transducer. For this problem, the ANOVA component functions of each natural frequency response were approximated by third-order polynomial (Hermite) expansions in terms of orthogonal polynomials [15], where the expansion coefficients were estimated by dimension-reduction integration [21].

Tables 5.2 and 5.3 present the means and standard deviations, respectively, of the first six natural frequencies, Ω_i, $i = 1, \ldots, 6$, of the transducer by four dimensional decomposition methods: univariate ADD, univariate FDD, univariate HDD (linear), and bivariate ADD approximations; and crude Monte Carlo simulation. The second-moment properties by decomposition methods, obtained by four-point Gauss–Hermite quadrature rules for estimating the expansion coefficients, are judged to be converged responses, as their changes due to further increases in polynomial order or number of quadrature nodes are negligibly small. Therefore, the univariate and bivariate approx-

TABLE 5.1
Statistical properties of the random input for the piezoelectric transducer.

Random variable	Property[a]	Mean	Coefficient of variation
X_1, GPa	D_{1111}	115.4	0.15
X_2, GPa	D_{1122}, D_{1133}	74.28	0.15
X_3, GPa	D_{2222}, D_{3333}	139	0.15
X_4, GPa	D_{2233}	77.84	0.15
X_5, GPa	$D_{1212}, D_{2323}, D_{1313}$	25.64	0.15
X_6, Coulomb/m^2	e_{111}	15.08	0.1
X_7, Coulomb/m^2	e_{122}, e_{133}	-5.207	0.1
X_8, Coulomb/m^2	e_{212}, e_{313}	12.71	0.1
X_9, nF/m	D_{11}	5.872	0.1
X_{10}, nF/m	D_{22}, D_{33}	6.752	0.1
X_{11}, GPa	E_b	104	0.15
X_{12}	ν_b	0.37	0.05
X_{13}, g/m^3	ρ_b	8500	0.15
X_{14}, g/m^3	ρ_c	7500	0.15

(a) D_{ijkl} are elastic moduli of ceramic; e_{ijk} are piezoelectric stress coefficients of ceramic; D_{ij} are dielectric constants of ceramic; E_b, ν_b, ρ_b are elastic modulus, Poisson's ratio, and mass density of brass; ρ_c is mass density of ceramic.

TABLE 5.2
Means of natural frequencies of the piezoelectric transducer.

Frequency	Approximation				Crude Monte Carlo
	Univariate ADD	Univariate FDD	Univariate HDD	Bivariate ADD	
Ω_1, kHz	19.45	19.45	19.45	19.38	19.35
Ω_2, kHz	42.31	42.31	42.31	42.27	42.25
Ω_3, kHz	59.23	59.23	59.23	59.42	59.33
Ω_4, kHz	67.45	67.45	67.45	67.21	67.23
Ω_5, kHz	90.57	90.57	90.57	90.60	90.62
Ω_6, kHz	101.69	101.69	101.69	101.66	101.57

TABLE 5.3
Standard deviations of natural frequencies of the piezoelectric transducer.

Frequency	Approximation				Crude Monte Carlo
	Univariate ADD	Univariate FDD	Univariate HDD	Bivariate ADD	
Ω_1, kHz	2.30	2.30	2.68	2.54	2.66
Ω_2, kHz	7.03	7.06	7.08	7.09	7.11
Ω_3, kHz	6.65	6.66	6.74	6.82	6.83
Ω_4, kHz	6.94	6.95	7.00	7.09	7.00
Ω_5, kHz	7.37	7.38	7.48	7.58	7.51
Ω_6, kHz	9.42	9.43	9.35	9.29	9.29

imations require $14 \times (14-1)+1 = 43$ and $14 \times (14-1)(4-1)^2/2+(14 \times (4-1)+1 = 862$ finite-element analyses, respectively [15]. Due to expensive finite-element analysis, crude Monte Carlo simulation was conducted only up to 50,000 realizations, which should be adequate for providing benchmark solutions of the second-moment characteristics. The agreement between the means by approximate and Monte Carlo methods in Table 5.2 is excellent. A comparison of standard deviations in Table 5.3 reveals superiority of the bivariate ADD approximation over univariate ADD and FDD approximations, as expected, but at a computational cost markedly higher than the

univariate approximations. More important, the univariate HDD approximation proposed is better than either univariate ADD or univariate FDD approximation and produces standard deviations very close to those by the bivariate ADD and crude Monte Carlo methods without carrying the computational burden of the latter methods.

6. Conclusions and outlook. It is time to take stock and recap what has been accomplished so far and what remains to be done.

Two dimensional decompositions, namely, FDD and HDD, of a multivariate function, representing finite products and sum-product mixtures of lower-dimensional component functions, were developed. A theorem, proven herein, reveals the relationship between all component functions of FDD and ADD, so far available only for univariate and bivariate component functions. Three function classes, comprising purely additive functions, purely multiplicative functions, and their mixtures, were examined to illustrate when and how an FDD approximation is more precise or relevant than an ADD approximation and vice versa. However, when a function is not endowed with a specific dimensional hierarchy, an HDD approximation, optimally blending ADD and FDD approximations, is more appropriate than either ADD or FDD approximation. Furthermore, the FDD and HDD lead to alternative definitions of effective dimension, reported in the literature associated with ADD only. New closed-form or analytical expressions were derived for calculating the variances stemming from univariate truncations of all three decompositions. The subsequent mean-squared error analysis pertaining to univariate ADD, FDD, and HDD approximations finds appropriate conditions when one approximation is better than the other. Numerical results from four simple yet insightful examples and a practical engineering problem indicate that an HDD approximation, when called for, commits lower errors than does ADD or FDD approximation. Therefore, HDD, whether formed linearly or nonlinearly, is ideally suited to a general function approximation that may otherwise require higher-variate ADD or FDD truncations for rendering acceptable accuracy in stochastic solutions.

Future work involves developing an adaptive strategy for FDD and HDD approximations to solve industrial-scale, stochastic problems encountered in engineering and applied sciences.

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] B. Efron and C. Stein, *The Jackknife estimate of variance*, Ann. Statist., 9 (1981), pp. 586–596.
[4] A. C. Genz and A. A. Malik, *An imbedded family of fully symmetric numerical integration rules*, SIAM J. Numer. Anal., 20 (1983), pp. 580–588.
[5] M. Griebel and M. Holtz, *Dimension-wise integration of high-dimensional functions with applications to finance*, J. Complexity, 26 (2010), pp. 455–489.
[6] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal., 33 (1996), pp. 1995–2016.
[7] F. J. Hickernell, *Erratum: Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal., 34 (1997), pp. 853–866.
[8] W. Hoeffding, *A class of statistics with asymptotically normal distributions*, Ann. Math. Statist., 20 (1948), pp. 293–325.
[9] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Wozniakowski, *On decompositions of multivariate functions*, Math. Comp., 79 (2010), pp. 953–966.
[10] A. B. Owen, *The dimension distribution and quadrature test functions*, Statist. Sinica, 13 (2003), pp. 1–17.

[11] H. RABITZ AND O. ALIS, *General foundations of high-dimensional model representations*, J. Math. Chem., 25 (1999), pp. 197–233.

[12] S. RAHMAN, *Approximation errors in truncated dimensional decompositions*, Math. Comp., to appear.

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

[14] S. RAHMAN, *Probability distributions of natural frequencies of uncertain dynamic systems*, AIAA J., 47 (2009), pp. 1579–1589.

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

[16] I. M. SOBOL, *Theorems and examples on high dimensional model representations*, Reliab. Engrg. System Safety, 79 (2003), pp. 187–193.

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

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

[19] M. A. TUNGA AND M. DEMIRALP, *Hybrid high dimensional model representation (HHDMR) on the partitioned data*, J. Comput. Appl. Math., 185 (2006), pp. 107–132.

[20] M. A. TUNGA AND M. DEMIRALP, *A factorized high-dimensional model representation on the nodes of a finite hyperprismatic regular grid*, Appl. Math. Comput., 164 (2005), pp. 865–883.

[21] 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.