Approximation Errors for High-Dimensional Uncertainty Quantification

Sharif Rahman
The University of Iowa, Iowa City, IA 52242

2012 AMCS Seminar

Work supported by NSF (CMMI-1130147, CMMI-0969044)

Outline

Uncertainty Quantification

$$\text{Input } \boldsymbol{X} \in \mathbb{R}^N \rightarrow \boxed{\textbf{MATH MODEL}} \rightarrow \text{Output } y(\boldsymbol{X}) \in \mathbb{R}$$

$$\boldsymbol{X} \sim (\Omega, \mathcal{F}, P);\ y \in \mathcal{L}_2(\Omega, \mathcal{F}, P)$$

- **Objectives**
  - Statistical moments: $\mathbb{E}\left[y^l(\boldsymbol{X})\right] := \int_{\mathbb{R}^N} y^l(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x},\ l \in \mathbb{N}$
  - Rare-event probability: $P\left[y(\boldsymbol{X}) \in \Omega_F\right] = \int_{\mathbb{R}^N} I_{\Omega_F}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}$
  - Design in presence of uncertainties

Uncertainty Quantification

Input $\boldsymbol{X} \in \mathbb{R}^N \rightarrow$ **MATH MODEL** $\rightarrow$ Output $y(\boldsymbol{X}) \in \mathbb{R}$

$$\boldsymbol{X} \sim (\Omega, \mathcal{F}, P);\; y \in \mathcal{L}_2(\Omega, \mathcal{F}, P)$$

- **Objectives**
  - Statistical moments: $\mathbb{E}\left[y^l(\boldsymbol{X})\right] := \int_{\mathbb{R}^N} y^l(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x},\; l \in \mathbb{N}$
  - Rare-event probability: $P\left[y(\boldsymbol{X}) \in \Omega_F\right] = \int_{\mathbb{R}^N} I_{\Omega_F}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}$
  - Design in presence of uncertainties

Challenge/Motivation

High-dimensional input ($10 \leq N \leq 100$); exploit hidden structures for low-dimensional approximations

Dimensional Decomposition (Hoeffding, 1948)

$$y(\boldsymbol{X}) = y_0 + \sum_{i=1}^{N} y_i(X_i) + \sum_{i_1,i_2=1;i_1<i_2}^{N} y_{i_1 i_2}(X_{i_1}, X_{i_2}) + \cdots + \sum_{i_1,\cdots,i_s=1,i_1<\cdots<i_s}^{N} y_{i_1\cdots i_s}(X_{i_1}, \cdots, X_{i_s}) + \cdots + y_{12\cdots N}(X_1, \cdots, X_N)$$

$$\boldsymbol{X} = \{X_1, \cdots, X_N\}^T; \text{ indep.}; \; X_i \sim f_i(x_i) \text{ on } (\Omega_i, \mathcal{F}_i, P_i)$$

$$w(\boldsymbol{x}) = \prod_{i=1}^{N} w_i(x_i); \; w_{-u}(\boldsymbol{x}_{-u}) := \prod_{i=1, i\notin u}^{N} w_i(x_i)$$

$$\begin{aligned} y(\boldsymbol{X}) &= \sum_{u\subseteq\{1,\cdots,N\}} y_u(\boldsymbol{X}_u), \\ y_\emptyset &= \int_{\mathbb{R}^N} y(\boldsymbol{x}) w(\boldsymbol{x}) d\boldsymbol{x}, \\ y_u(\boldsymbol{X}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\boldsymbol{X}_u, \boldsymbol{x}_{-u}) w_{-u}(\boldsymbol{x}_{-u}) d\boldsymbol{x}_{-u} - \sum_{v\subset u} y_v(\boldsymbol{X}_v) \end{aligned}$$

ANOVA Dimensional Decomposition (ADD)

Select: $w(\boldsymbol{x})d\boldsymbol{x} = f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}$

$$
\begin{aligned}
y(\boldsymbol{X}) &= \sum_{u \subseteq \{1,\cdots,N\}} y_{u,A}(\boldsymbol{X}_u), \\
y_{\emptyset,A} &= \int_{\mathbb{R}^N} y(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}, \\
y_{u,A}(\boldsymbol{X}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\boldsymbol{X}_u, \boldsymbol{x}_{-u}) f_{\boldsymbol{X}_{-u}}(\boldsymbol{x}_{-u}) d\boldsymbol{x}_{-u} - \sum_{v \subset u} y_v(\boldsymbol{X}_v)
\end{aligned}
$$

- **Two Remarkable Properties**

$$
\left.
\begin{aligned}
\mathbb{E}\left[y_{u,A}(\boldsymbol{X}_u)\right] &= 0 \\
\mathbb{E}\left[y_{u,A}(\boldsymbol{X}_u) y_{v,A}(\boldsymbol{X}_v)\right] &= 0, \\
\emptyset \neq u, v \subseteq \{1,\cdots,N\},\ u &\neq v
\end{aligned}
\right\} \text{Independent } \boldsymbol{X}
$$

ADD component functions are orthogonal, but are difficult to obtain as they involve high-dimensional integrals

Referential Dimensional Decomposition (RDD)

Select: $w(\boldsymbol{x})d\boldsymbol{x}=\prod_{i=1}^{N}\delta(x_i - c_i)\,dx_i$

$$\boldsymbol{c} = (c_1, \cdots, c_N) \in \mathbb{R}^N$$

$$y(\boldsymbol{X}) = \sum_{u\subseteq\{1,\cdots,N\}} y_{u,R}(\boldsymbol{X}_u; \boldsymbol{c}),$$

$$y_{\emptyset,A} = y(\boldsymbol{c}),$$

$$y_{u,R}(\boldsymbol{X}_u; \boldsymbol{c}) = y(\boldsymbol{X}_u, \boldsymbol{c}_{-u}) - \sum_{v\subset u} y_{v,R}(\boldsymbol{X}_v; \boldsymbol{c})$$

RDD component functions lack orthogonal features, but are easy to obtain as they involve only function evaluations

Truncated ADD & Variances

- **S-variate ADD Approximation ($0 \le S < N$)**

$$\hat{y}_{S,A}(\boldsymbol{X}) = \sum_{\substack{u \subseteq \{1,\cdots,N\} \\ 0 \le |u| \le S}} y_{u,A}(\boldsymbol{X}_u)$$

- **Approximate Variance**

$$\hat{\sigma}^2_{S,A} := \mathbb{E}\left(\hat{y}_{S,A}(\boldsymbol{X}) - y_{\emptyset,A}\right)^2 = \sum_{s=1}^{S} \sum_{\substack{\emptyset \ne u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2; \; \sigma_u^2 := \mathbb{E}\left[y^2_{u,A}(\boldsymbol{X}_u)\right]$$

- **Exact Variance**

$$\sigma^2 := \mathbb{E}\left(y(\boldsymbol{X}) - y_\emptyset\right)^2 = \sum_{s=1}^{N} \sum_{\substack{\emptyset \ne u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2$$

When $S \to N$, $\hat{\sigma}^2_{S,A} \to \sigma^2$ ($\mathcal{L}_2$ convergence)

ADD Error

- **S-variate ADD Error**

$$e_{S,A} := \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X})\right)^2\right] := \int_{\mathbb{R}^N} \left[y(\boldsymbol{x}) - \hat{y}_{S,A}(\boldsymbol{x})\right]^2 f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

Using orthogonal properties of ADD,

$$e_{S,A} = \sum_{s=S+1}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2$$

- **Univariate ($S = 1$) & Bivariate ($S = 2$) ADD Errors**

$$e_{1,A} = \sum_{s=2}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2; \quad e_{2,A} = \sum_{s=3}^{N} \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2$$

ADD error completely eliminates the variance terms associated with S- and all lower-variate contributions

Optimality

- **Other Approximation Errors**

$$
\begin{aligned}
e_S &:= \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \\
&= \mathbb{E}\left[\left(\{y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X})\} + \{\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\}\right)^2\right] \\
&= \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X})\right)^2\right] + \mathbb{E}\left[\left(\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \\
&= e_{S,A} + \mathbb{E}\left[\left(\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \geq e_{S,A}
\end{aligned}
$$

$y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X}) \;\rightarrow\;$ higher than $S-$variate terms

$\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X}) \;\rightarrow\;$ at most $S-$variate terms

Optimality

- **Other Approximation Errors**

$$\begin{aligned} e_S &:= \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \\ &= \mathbb{E}\left[\left(\{y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X})\} + \{\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\}\right)^2\right] \\ &= \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X})\right)^2\right] + \mathbb{E}\left[\left(\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \\ &= e_{S,A} + \mathbb{E}\left[\left(\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X})\right)^2\right] \geq e_{S,A} \end{aligned}$$

$y(\boldsymbol{X}) - \hat{y}_{S,A}(\boldsymbol{X}) \;\rightarrow\;$ higher than $S-$variate terms

$\hat{y}_{S,A}(\boldsymbol{X}) - \hat{y}_S(\boldsymbol{X}) \;\rightarrow\;$ at most $S-$variate terms

ADD approximation is optimal in $\mathcal{L}_2$ sense

Truncated RDD

- **S-variate RDD Approximation ($0 \le S < N$)**

$$\hat{y}_{S,R}(\boldsymbol{X};\boldsymbol{c}) = \sum_{\substack{u\subseteq\{1,\cdots,N\}\\0\le|u|\le S}} y_{u,R}(\boldsymbol{X}_u;\boldsymbol{c})$$

- **Direct Form (Xu and Rahman, 2004)**

$$\hat{y}_{S,R}(\boldsymbol{X};\boldsymbol{c}) = \sum_{k=0}^{S}(-1)^k\binom{N-S+k-1}{k}\sum_{\substack{u\subseteq\{1,\cdots,N\}\\|u|=S-k}} y(\boldsymbol{X}_u,\boldsymbol{c}_{-u}),$$

$$\boldsymbol{c} = (c_1,\cdots,c_N)\in\mathbb{R}^N$$

Special Cases

- **Univariate RDD Approximation** $(S = 1)$

$$\hat{y}_{1,R}(\boldsymbol{X};\boldsymbol{c}) = \sum_{i=1}^{N} y(X_i, \boldsymbol{c}_{-\{i\}}) - (N-1)y(\boldsymbol{c})$$

- **Bivariate RDD Approximation** $(S = 2)$

$$\begin{aligned}\hat{y}_{2,R}(\boldsymbol{X};\boldsymbol{c}) &= \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} y(X_i, X_j, \boldsymbol{c}_{-\{i,j\}}) - (N-2)\sum_{i=1}^{N} y(X_i, \boldsymbol{c}_{-\{i\}}) \\ &\quad + \frac{1}{2}(N-1)(N-2)y(\boldsymbol{c})\end{aligned}$$

RDD Error

- **S-variate RDD Error**

$$\begin{aligned} e_{S,R}(\boldsymbol{c}) &:= \mathbb{E}\left[\left(y(\boldsymbol{X}) - \hat{y}_{S,R}(\boldsymbol{X};\boldsymbol{c})\right)^2\right] \\ &:= \int_{\mathbb{R}^N} \left[y(\boldsymbol{x}) - \hat{y}_{S,R}(\boldsymbol{x};\boldsymbol{c})\right]^2 f_{\boldsymbol{X}}(\boldsymbol{x})\,d\boldsymbol{x} \end{aligned}$$

- **Expected S-variate RDD Error**

If $\boldsymbol{c}$ is a randomly selected reference point with joint PDF $f_{\boldsymbol{X}}(\boldsymbol{c})$, then

$$\begin{aligned} \mathbb{E}\left[e_{S,R}(\boldsymbol{c})\right] &:= \int_{\mathbb{R}^N} e_{S,R}(\boldsymbol{c}) f_{\boldsymbol{X}}(\boldsymbol{c})\,d\boldsymbol{c} \\ &= \int_{\mathbb{R}^{2N}} \left[y(\boldsymbol{x}) - \hat{y}_{S,R}(\boldsymbol{x};\boldsymbol{c})\right]^2 f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{c})\,d\boldsymbol{x}\,d\boldsymbol{c} \end{aligned}$$

RDD vs. ADD Errors (Univariate)

Theorem

Let $\boldsymbol{c}$ be a random vector with joint PDF of the form $f_{\boldsymbol{X}}(\boldsymbol{c}) = \prod_{j=1}^{j=N} f_j(c_j)$, where f_j is the marginal PDF of its jth coordinate. Then the expected error committed by the univariate RDD approximation for $2 \leq N < \infty$ is

$$\mathbb{E}\left[e_{1,R}(\boldsymbol{c})\right] = \sum_{s=2}^{N} \left(s^2 - s + 2\right) \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2,$$

where $\sigma_u^2 = \mathbb{E}[y_{u,A}^2(\boldsymbol{X}_u)]$, $\emptyset \neq u \subseteq \{1, \cdots, N\}$.

RDD vs. ADD Errors (Univariate)

Theorem

Let $\boldsymbol{c}$ be a random vector with joint PDF of the form $f_{\boldsymbol{X}}(\boldsymbol{c}) = \prod_{j=1}^{j=N} f_j(c_j)$, where f_j is the marginal PDF of its jth coordinate. Then the expected error committed by the univariate RDD approximation for $2 \leq N < \infty$ is

$$\mathbb{E}\left[e_{1,R}(\boldsymbol{c})\right] = \sum_{s=2}^{N} \left(s^2 - s + 2\right) \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2,$$

where $\sigma_u^2 = \mathbb{E}[y_{u,A}^2(\boldsymbol{X}_u)]$, $\emptyset \neq u \subseteq \{1, \cdots, N\}$.

Corollary

$$4e_{1,A} \leq \mathbb{E}\left[e_{1,R}\right] \leq \left(N^2 - N + 2\right) e_{1,A}, \; 2 \leq N < \infty$$

RDD vs. ADD Errors (Bivariate)

Theorem

Let $\boldsymbol{c}$ *be a random vector with joint PDF of the form* $f_{\boldsymbol{X}}(\boldsymbol{c}) = \prod_{j=1}^{j=N} f_j(c_j)$, *where* f_j *is the marginal PDF of its* j*th coordinate. Then the expected error committed by the bivariate RDD approximation for* $3 \leq N < \infty$ *is*

$$\mathbb{E}\left[e_{2,R}(\boldsymbol{c})\right] = \sum_{s=3}^{N} \frac{1}{4}\left(s^4 - 2s^3 - s^2 + 2s + 8\right) \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2,$$

where $\sigma_u^2 = \mathbb{E}[y_{u,A}^2(\boldsymbol{X}_u)]$, $\emptyset \neq u \subseteq \{1, \cdots, N\}$.

RDD vs. ADD Errors (Bivariate)

Theorem

Let $\boldsymbol{c}$ *be a random vector with joint PDF of the form* $f_{\boldsymbol{X}}(\boldsymbol{c}) = \prod_{j=1}^{j=N} f_j(c_j)$, *where* f_j *is the marginal PDF of its* j*th coordinate. Then the expected error committed by the bivariate RDD approximation for* $3 \leq N < \infty$ *is*

$$\mathbb{E}\left[e_{2,R}(\boldsymbol{c})\right] = \sum_{s=3}^{N} \frac{1}{4}\left(s^4 - 2s^3 - s^2 + 2s + 8\right) \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2,$$

where $\sigma_u^2 = \mathbb{E}[y_{u,A}^2(\boldsymbol{X}_u)]$, $\emptyset \neq u \subseteq \{1, \cdots, N\}$.

Corollary

$$8e_{2,A} \leq \mathbb{E}\left[e_{2,R}\right] \leq \frac{1}{4}\left(N^4 - 2N^3 - N^2 + 2N + 8\right) e_{2,A},\ 3 \leq N < \infty$$

A Contrived Example

Consider a function of 100 variables with the following distribution of the variance terms: $\sum_{|u|=1} \sigma_u^2 = 0.999\sigma^2$, $\sum_{2\leq|u|\leq 99} \sigma_u^2 = 0$, $\sum_{|u|=100} \sigma_u^2 = 0.001\sigma^2$, $0 < \sigma^2 < \infty$

- **ADD Errors**

$$e_{1,A} = e_{2,A} = 0.001\sigma^2 \text{ (negligible)}$$

- **Expected RDD Errors**

$$\mathbb{E}\left[e_{1,R}(\boldsymbol{c})\right] \cong 9.9\sigma^2 \text{ (large)}$$

$$\mathbb{E}\left[e_{2,R}(\boldsymbol{c})\right] \cong 24,498\sigma^2 \text{ (enormous)}$$

A Contrived Example

Consider a function of 100 variables with the following distribution of the variance terms: $\sum_{|u|=1} \sigma_u^2 = 0.999\sigma^2$, $\sum_{2\leq|u|\leq 99} \sigma_u^2 = 0$, $\sum_{|u|=100} \sigma_u^2 = 0.001\sigma^2$, $0 < \sigma^2 < \infty$

- **ADD Errors**

$$e_{1,A} = e_{2,A} = 0.001\sigma^2 \text{ (negligible)}$$

- **Expected RDD Errors**

$$\mathbb{E}\left[e_{1,R}(\boldsymbol{c})\right] \cong 9.9\sigma^2 \text{ (large)}$$

$$\mathbb{E}\left[e_{2,R}(\boldsymbol{c})\right] \cong 24,498\sigma^2 \text{ (enormous)}$$

A higher-variate RDD approximation may commit a larger error than a lower-variate RDD approximation

RDD vs. ADD Errors (General)

Theorem

Let $\boldsymbol{c}$ *be a random vector with joint PDF of the form* $f_{\boldsymbol{X}}(\boldsymbol{c}) = \prod_{j=1}^{j=N} f_j(c_j)$, *where* f_j *is the marginal PDF of its* j*th coordinate. Then the expected error committed by the* S*-variate RDD approximation for* $0 \leq S < N$, $S+1 \leq N < \infty$ *is*

$$\mathbb{E}\left[e_{S,R}(\boldsymbol{c})\right] = \sum_{s=S+1}^{N} \left[1 + \sum_{k=0}^{S} \binom{s-S+k-1}{k}^2 \binom{s}{S-k}\right] \sum_{\substack{\emptyset \neq u \subseteq \{1,\cdots,N\} \\ |u|=s}} \sigma_u^2,$$

where $\sigma_u^2 = \mathbb{E}[y_{u,A}^2(\boldsymbol{X}_u)]$, $\emptyset \neq u \subseteq \{1,\cdots,N\}$.

RDD error eliminates S- and all lower-variate contributions, but with a stronger dependence on higher-variate terms

RDD vs. ADD Errors (General)

Corollary

The lower and upper bounds of the expected error $\mathbb{E}\left[e_{S,R}\right]$ from the S-variate RDD approximation, expressed in terms of the error $e_{S,A}$ from the S-variate ADD approximations, are

$$2^{S+1} e_{S,A} \leq \mathbb{E}\left[e_{S,R}\right] \leq \left[1 + \sum_{k=0}^{S} \binom{N-S+k-1}{k}^2 \binom{N}{S-k}\right] e_{S,A},$$

$0 \leq S < N < \infty$.

ADD approximations are exceedingly more precise than RDD approximations at higher-variate truncations

RDD vs. ADD Errors (General)

Corollary

The expected error $\mathbb{E}\left[e_{N-1,R}\right]$ *from the best RDD approximation, expressed in terms of the error* $e_{N-1,A}$ *from the best ADD approximation, where the best approximations are obtained by setting* $S = N-1$*, is*

$$\mathbb{E}\left[e_{N-1,R}\right] = 2^N e_{N-1,A},\ 1 \leq N < \infty.$$

The best RDD approximation error can be significantly larger than the best ADD approximation error

Conclusions

- New formulae for expected errors from various RDD approximations
- S-variate RDD error is at least 2^{S+1} times greater than the S-variate ADD error
- ADD approximation is optimal
- RDD approximation should be used with caution

Conclusions

- New formulae for expected errors from various RDD approximations
- S-variate RDD error is at least 2^{S+1} times greater than the S-variate ADD error
- ADD approximation is optimal
- RDD approximation should be used with caution

Future Works (Ph.D. topics)

- Dependent probability measures of random input (does ADD exist? with what properties?)
- Rare event probability (reliability, stochastic optimization)
- Adaptivity/sparsity (how to select S? how to pick y_u?)
- Multiplicative & hybrid dimensional decompositions