Exploration of Multifidelity Approaches for Unc SAND2019-6849C Quantification in Network Applications

Gianluca Geraci[1], Laura P. Swiler[1], Jonathan Crussell[2]
and Bert J. Debusschere[3]

3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP2019)
Crete, Greece
June 26th 2019

Sandia National Laboratories

U.S. DEPARTMENT OF ENERGY NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Plan of the talk

- Network modeling
- Multifidelity UQ
- Numerical Experiments
- Concluding remarks

Network modeling

Network models
Simulation vs Emulation

Why are we interested in network models?

- Network **operators**: understand the potential impacts of changes **before implementing** them
- Network **designers**: understand trade-offs **before** network **creation**

Network modeling refers to:

- **Simulation:** similar to their physics-modeling counterparts and they are **based on a deep understanding of the underlying processes** to simulate network components and interactions in software
- **Emulation:** run the real software on virtualized hardware thus it is able to **capture unknown or not well-understood behaviors**

A crude fluid dynamics analogy: Let's consider a **straight wing** flying...

- ...at nearly 0° angle-of-attach with laminar attached flow $\rightarrow$ **physics well-known** $\rightarrow$ thin airfoil theory (<u>simulation</u>)
- ...at high angle-of-attach with turbulent detached flow $\rightarrow$ **physics poorly understood** $\rightarrow$ Large Eddy Simulations (<u>emulation</u>)

Emulation has a higher-degree of realism, but this comes at an higher computational cost

Network models

Simulation vs Emulation: strengths and differences

Simulation

- ✔ Fast to develop
- ✔ Run faster than real-time since they control the clock
- ✔ Easy to run in parallel because they are neither time-dependent or reliant on virtualized hardware (which might be limited)
- ✗ Unable to capture emergent behaviors

Emulation

- ✔ Run the real software therefore closely resembles a physical testbed
- ✗ Requires more hardware and therefore the number of concurrent evaluations are limited

Emulytics

- Combine Emulation and analytics
- Includes hardware in the loop, simulation, and emulation (and human in the loop)

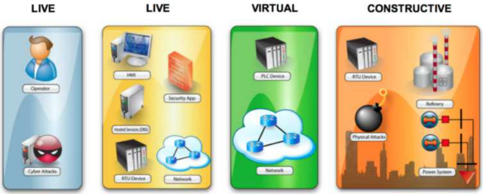


Figure courtesy of David Fritz, SAND2018-3927

We focus here only on the **virtual layer**

Network models

Why network modeling at Sandia? (Courtesy of David Fritz, SAND2018-3927[1])

- **DevOps:** Ensure operation of new hardware, software, services in high-consequence environments. Predictive analysis to detect malfunctions, misconfigurations and malicious consequences
- **Malware:** Understanding of malware through pseudo-in situ execution
- **ICS/SCADA:** Under uncertain threats, what are the best countermeasures for my IT-connected ICS systems? Can we detect attacks? Can we assess resiliency of the IT-controls over the entire power grids?
- **Nuclear Weapons:** Can we assure Communication, Command and Control regardless of network state and threats?

For all these applications we operate in an extreme uncertain environment/scenario and we need to quantify the probability of obtain certain desired responses by our systems

$\Downarrow$

Uncertainty Quantification

[1] http://minimega.org/presentations/gt_2018.slide#7

Multifidelity Uncertainty Quantification

Uncertainty Quantification
Forward Propagation – Why Sampling Methods?

UQ context at a glance:

- High-dimensionality, non-linearity and bifurcations/discontinuities
- Large set of modeling choices available (network topology, operative conditions, etc.)

Natural candidate:

- **Sampling**-based (MC-like) approaches because they are **non-intrusive**, **robust** and **flexible**...
- **Drawback:** Slow convergence $\mathcal{O}(N^{-1/2}) \rightarrow$ many realizations to build reliable statistics

Goal of MF UQ:

Reducing the computational cost of obtaining MC reliable statistics by combining several models

Pivotal idea:

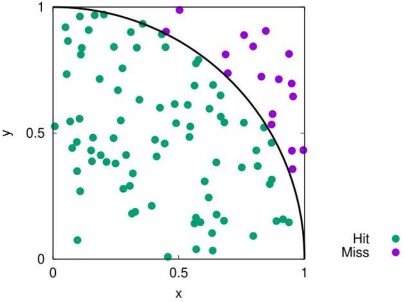
- Simplified (**low-fidelity**) models are **inaccurate** but **computationally inexpensive** $\Rightarrow$ **low-variance** estimates
- **High-fidelity** models are **costly**, but **accurate** $\Rightarrow$ **low-bias** estimates

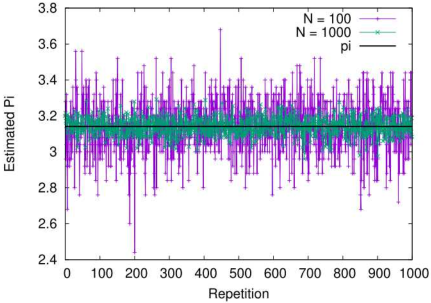
Monte Carlo
Generalities

Let consider a **random variable Q**, we want to compute **its expected value $\mathbb{E}[Q]$** (or some high-order moment):

$$\hat{Q}_N^{\mathrm{MC}} = \frac{1}{N} \sum_{i=1}^{N} Q^{(i)}$$

Let's use MC to compute the value $\pi = \dfrac{\#\mathrm{Hit}}{N}$





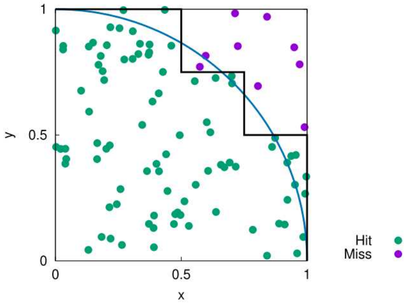
Monte Carlo

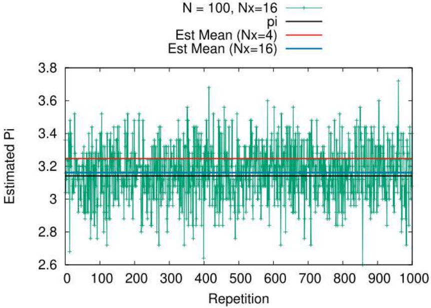
Introducing the notion of fidelity

Whenever a numerical problems **cannot be resolved with infinite accuracy** (discretization error), the MC estimator for a specific **M**th level

$$\hat{Q}^{MC}_{\mathbf{M},N} \overset{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} Q_{\mathbf{M}}{}^{(i)}$$

Let's use MC to compute the value $\pi = \dfrac{\#\text{Hit}}{N}$





Monte Carlo
Overall estimator error

Two sources of error in the **Mean Square Error**:

$$\mathbb{E}\left[(\hat{Q}^{MC}_{M,N} - \mathbb{E}[Q])^2\right] = \mathbb{V}\mathbf{ar}\left(\hat{\mathbf{Q}}^{\mathbf{MC}}_{\mathbf{M,N}}\right) + (\mathbb{E}[\mathbf{Q_M} - \mathbf{Q}])^{\mathbf{2}}$$

- **Sampling error:** replacing the expected value by a (finite) sample average, *i.e.*

$$\mathbb{V}\mathbf{ar}\left(\hat{\mathbf{Q}}^{\mathbf{MC}}_{\mathbf{M,N}}\right) = \frac{\mathbb{V}\mathbf{ar}(\mathbf{Q})}{\mathbf{N}}$$

- **Model fidelity (e.g. discretization):** finite accuracy

Accurate estimation $\Rightarrow$ **Large number** of samples evaluated for the **high fidelity** model

$$\mathbb{E}[Q_M] - \hat{Q}^{MC}_{M,N} \sim \sqrt{\frac{\mathbb{V}\mathbf{ar}(\mathbf{Q_M})}{\mathbf{N}}}\mathcal{N}(0,1)$$

In our network application we operate under the assumptions that

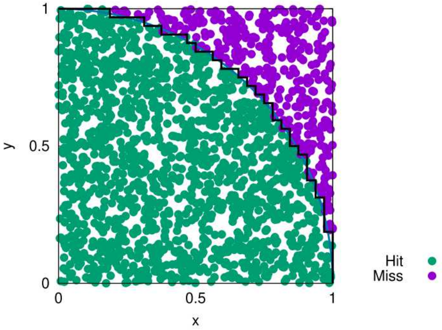
- The emulytics is the highest **unbiased** fidelity model, *i.e.* $(\mathbb{E}[\mathbf{Q_M} - \mathbf{Q}])^{\mathbf{2}} = 0$
- Our goal is to solely **reduce the variance** of the estimator by introducing low-fidelity evaluations

Accelerating Monte Carlo
Bringing multiple fidelity models into the picture

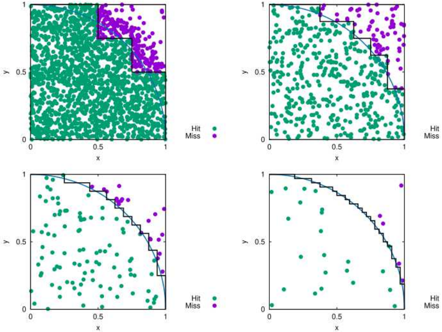
Pivotal idea:

- **High-fidelity** models are **costly**, but **accurate**
  - **low-bias** estimates
- Simplified (**low-fidelity**) models are **inaccurate** but **cheap**
  - **low-variance** estimates

Single Fidelity



Multi Fidelity



Control variate

Several ways of accelerating MC convergence

In this talk we focus on reducing the variance of the estimator

$$\mathbb{V}ar\left(\hat{Q}\right) = \frac{\mathbb{V}ar(Q)}{N}$$

What can we do to drive down the variance of the estimator?

#0 **Increasing the number of samples** $\rightarrow$ this is going to cost us more

#1 Changing the problem with another one under the assumption that the mean is the same, but the new variance is smaller

#2 **Change the problem with a computational cheapest one** (that preserves the mean in this case)

#1: Variance reduction techniques

- Act on the sampling (Stratification, Important Sampling *etc.*)
- **Act on the function (control variate)**

Approximate Control Variates

Optimal Control Variate

M low-fidelity models with known expected value

Let's consider M **low-fidelity models with known mean**. The Optimal Control Variate (OCV) is generated by adding M unbiased terms to the MC estimator

$$\hat{Q}^{\mathrm{CV}} = \hat{Q} + \sum_{i=1}^{M} \alpha_i \left(\hat{Q}_i - \mu_i \right)$$

- $\hat{Q}_i$ MC estimator for the ith low-fidelity model
- μ_i known expected value for the ith low-fidelity model
- $\underline{\alpha} = [\alpha_i, \ldots, \alpha_M]^{\mathrm{T}}$ set of weights (to be determined)

Let's define

- The covariance matrix among all the low-fidelity models: $\mathbf{C} \in \mathbb{R}^{M \times M}$
- The vector of covariances between the high-fidelity Q and each low-fidelity Q_i: $\mathbf{c} \in \mathbb{R}^M$
- $\bar{\mathbf{c}} = \mathbf{c}/\mathbb{V}ar(Q) = [\rho_1 \mathbb{V}ar(Q_1), \ldots, \rho_M \mathbb{V}ar(Q_M)]^{\mathrm{T}}$, where ρ_i is the correlation coefficient (Q, Q_i)

The optimal weights are obtained as $\underline{\alpha}^\star = -\mathbf{C}^{-1}\mathbf{c}$ and the variance of the OCV estimator

$$\begin{aligned} \mathbb{V}ar\left(\hat{Q}^{\mathrm{CV}}\right) &= \mathbb{V}ar\left(\hat{Q}\right) \left(1 - \bar{\mathbf{c}}^{\mathrm{T}} \mathbf{C}^{-1} \bar{\mathbf{c}}\right) \\ &= \mathbb{V}ar\left(\hat{Q}\right) (1 - R^2_{OCV}), \quad 0 \leq R^2_{OCV} \leq 1. \end{aligned}$$

For a single low-fidelity model: $R^2_{OCV-1} = \rho_1^2$

Approximate Control Variate

M low-fidelity models with unknown expected value

For complex engineering models the **expected values of the M low-fidelity models are unknown** *a priori*

- Let's define the set of sample used for the high-fidelity model: $\mathbf{z}$
- Let's consider N_i ordered evaluations for Q_i: $\mathbf{z_i}$ (we assume $N_i = \lceil r_i N \rceil$)
- Let's partition $\mathbf{z_i}$ in two ordered subsets $\mathbf{z_i^1} \bigcup \mathbf{z_i^2} = \mathbf{z_i}$ (note that in general $\mathbf{z}_i^1 \bigcap \mathbf{z}_i^2 \neq \emptyset$)

The **generic Approximate Control Variate** is defined as

$$\tilde{Q}\left(\underline{\alpha}, \underline{\mathbf{z}}\right) = \hat{Q}(\mathbf{z}) + \sum_{i=1}^{M} \alpha_i \left(\hat{Q}_i(\mathbf{z_i^1}) - \hat{\mu}_i(\mathbf{z_i^2})\right) = \hat{Q}(\mathbf{z}) + \sum_{i=1}^{M} \alpha_i \Delta_i(\mathbf{z_i}) = \hat{Q} + \underline{\alpha}^{\mathrm{T}} \underline{\Delta},$$

The **optimal weights** and **variance** can be obtained as

$$\underline{\alpha}^{\mathrm{ACV}} = -\mathbb{C}ov\left[\underline{\Delta}, \underline{\Delta}\right]^{-1} \mathbb{C}ov\left[\underline{\Delta}, \hat{Q}\right]$$

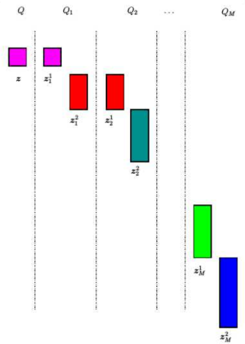
$$\mathbb{V}ar\left(\tilde{Q}\left(\underline{\alpha}^{ACV}\right)\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - \mathbb{C}ov\left[\underline{\Delta}, \hat{Q}\right]^{\mathrm{T}} \frac{\mathbb{C}ov\left[\underline{\Delta}, \underline{\Delta}\right]^{-1}}{\mathbb{V}ar\left(\hat{Q}\right)} \mathbb{C}ov\left[\underline{\Delta}, \hat{Q}\right]\right)$$

$$= \mathbb{V}ar\left(\hat{Q}\right)\left(1 - R^2_{\mathrm{ACV}}\right).$$

For a single low-fidelity model: $R^2_{ACV-1} = \frac{r_1 - 1}{r_1} \rho_1^2$ (this result does not depend on the partitioning of $\mathbf{z_1}$)

Multilevel Monte Carlo

A recursive partitioning with independent estimators (Given a prescribed bias)



MLMC can be obtained from ACV with

- $\mathbf{z_i^1} = \mathbf{z}$
- $\mathbf{z_i^2} = \mathbf{z_{i+1}^1}$ for $i = 1, \dots, M-1$
- $\alpha_\mathbf{i} = -\mathbf{1}$ for all i

$$\hat{Q}^{\text{MLMC}}(\mathbf{z}) = \hat{Q} + \sum_{i=1}^{M} (-\mathbf{1}) \left(\hat{Q}_i(\mathbf{z_i^1}) - \hat{\mu}_i(\mathbf{z_i^2}) \right)$$

$$\mathbb{V}ar\left(\hat{Q}^{\text{MLMC}}\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - R^2_{\text{MLMC}}\right)$$

$$R^2_{\text{MLMC}} = -\sum_{i=1}^{M} \frac{\bar{r}_i + \bar{r}_{i-1}}{\bar{r}_i \bar{r}_{i-1}} \tau_i^2 + 2 \sum_{i=1}^{M-1} \frac{\rho_{i,i+1}\tau_i\tau_{i+1}}{\bar{r}_i} - 2\rho_1\tau_1, \quad \tau_i = \sqrt{\mathbb{V}ar(Q_i)/\mathbb{V}ar(Q)}$$

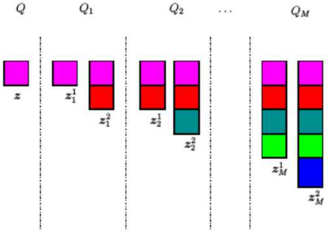
where

- $\mathbf{z}_i^1$ and $\mathbf{z}_i^2$ is $\bar{r}_{i-1}N$ and $\bar{r}_i N$ and $\bar{r}_0 = 1$, it holds that $r_i = \bar{r}_i + \bar{r}_{i-1}$

Given the recursive nature of MLMC, we can show that $R^2_{\text{MLMC}} < \rho_1^2$

Multifidelity Monte Carlo

An approximated control variate with a recursive partitioning



MFMC can be obtained from ACV with

- $\mathbf{z}_i^1 = \mathbf{z}_{i-1}$ and $\mathbf{z}_i^2 = \mathbf{z}_i$ for $i = 2, \ldots, M$
- $\mathbf{z}_1^1 = \mathbf{z}$ and $\mathbf{z}_1^2 = \mathbf{z}_1$

$$\alpha_i^{\text{MFMC}} = -\frac{\mathbb{C}ov\left[Q, Q_i\right]}{\mathbb{V}ar(Q_i)}, \quad \text{for} \quad i = 1, \ldots, M,$$

and the variance of the estimator is

$$\mathbb{V}ar\left(\underline{\alpha}^{\text{MFMC}}\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - R_{\text{MFMC}}^2\right)$$

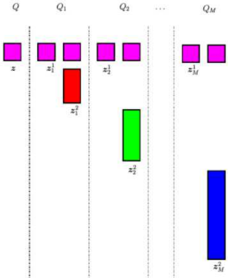
$$R_{\text{MFMC}}^2 = \sum_{i=1}^{M} \frac{r_i - r_{i-1}}{r_i r_{i-1}} \rho_i^2 = \rho_1^2 \left(\frac{r_1 - 1}{r_1} + \sum_{i=2}^{M} \frac{r_i - r_{i-1}}{r_i r_{i-1}} \frac{\rho_i^2}{\rho_1^2}\right).$$

Given the recursive nature of MFMC, we can show that $R_{\text{MFMC}}^2 < \rho_1^2$

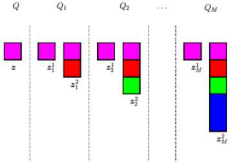
Examples of Convergent Estimators

Is it possible to overcome the limitation of the recursive sampling schemes?

We proposed two sampling strategies that overcome the limitation of the recursive schemes



(a) ACV-IS sampling strategy.



(b) ACV-MF sampling strategy.

As an example, let's consider the **ACV-MF estimator**

$$R^2_{\mathrm{ACV-MF}} = \left[\mathrm{diag}\left(\mathbf{F}^{(\mathrm{MF})}\right) \circ \bar{\mathbf{c}}\right]^{\mathrm{T}} \left[\mathbf{C} \circ \mathrm{diag}\left(\mathbf{F}^{(\mathrm{MF})}\right)\right]^{-1} \left[\mathrm{diag}\left(\mathbf{F}^{(\mathrm{MF})}\right) \circ \bar{\mathbf{c}}\right].$$

The matrix $\mathbf{F}^{(\mathrm{MF})} \in \mathbb{R}^{M \times M}$ encodes the particular sampling strategy and is defined as

$$\mathbf{F}^{(\mathrm{MF})}_{ij} = \begin{cases} \frac{\min(r_i, r_j) - 1}{\min(r_i, r_j)} & \text{if } i \neq j \\ \frac{r_i - 1}{r_i} & \text{otherwise} \end{cases}, \quad \text{for} \quad \mathbf{r_i} \to \infty, \quad \mathbf{F}^{(\mathrm{MF})} \to \mathbf{1_M} \quad \text{and} \quad \mathbf{R^2_{ACV-MF}} \to \mathbf{R^2_{OCV}}$$

A parametric model problem

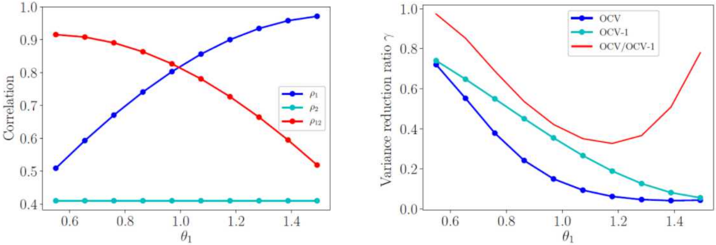
What happens for a limited number of low-fidelity simulations?

We designed a parametric test problem to explore different cost and correlation scenarios ($x, y \sim \mathcal{U}(-1, 1)$)

$$Q = A\left(\cos\theta\, x^5 + \sin\theta\, y^5\right)$$
$$Q_1 = A_1\left(\cos\theta_1\, x^3 + \sin\theta_1\, y^3\right)$$
$$Q_2 = A_2\left(\cos\theta_2\, x + \sin\theta_2\, y\right)$$

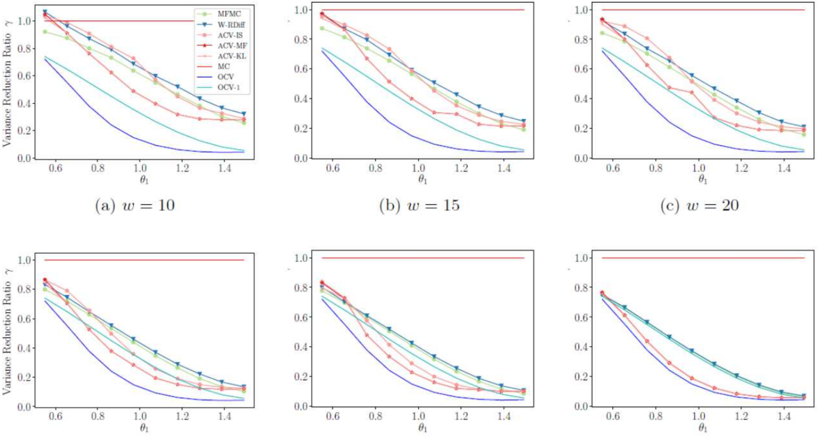
We use the following definitions

- $A = \sqrt{11}$, $A_1 = \sqrt{7}$, and $A_2 = \sqrt{3}$ (give unitary variance for each model)
- $\theta = \pi/2$ and $\theta_2 = \pi/6$ and θ_1 varies uniformly in the bounds $\theta_2 < \theta_1 < \theta$
- We consider a fixed cost ratio between models, *i.e.* a relative cost of 1 for Q, $1/w$ for Q_1 and $1/w^2$ for Q_2



A PARAMETRIC MODEL PROBLEM

COMPARISON OF DIFFERENT ESTIMATORS (EQ. COST 100 HF)



(a) $w = 10$ (b) $w = 15$ (c) $w = 20$

(d) $w = 50$ (e) $w = 100$ (f) $w = 1000$

FIGURE: Variance reduction for cost ratios of $[1, 1/w, 1/w^2]$ for Q, Q_1, and Q_2

Simulation/Emulation tools

Simulation tool

ns-3

ns-3

- ns-3 is a **discrete event** simulator for IP and non-IP addresses
- Software written in C++ with bindings available for Python
- GNU GPLv2-licensed
- Possible to construct simulations from **reusable components** to configure nodes, topologies and applications

Discrete-event simulation

- Virtual time evolves from event to event
- A single-threaded event list is executed
- Events are scheduled to occur at specific virtual/simulation time
- Events can generate additional events
- Simulation ends when a specific time is reached or there are no more events

Emulation tool

minimega

minimega

- Tool for launching and managing virtual machines
- It can run on your laptop or distributed across a cluster
- Open source GNU GPLv3-licensed, publicly available and active project
- Integrate real hardware with virtual experiments

Numerical Experiments

NS3 TEST PROBLEM
1 CLIENT - 1 SERVER NETWORK CONFIGURATION

Network Configuration
- 1 client - 1 server (possible to extend to multiple clients)
- 100 Requests

Uncertain Parameters
- `DataRate` $\sim \mathcal{U}(5, 500) Mbps$
- `Delay` $\sim \mathcal{U}(1, 3) ms$

Fidelity definition
- HF: `ResponseSize` $16MB$ – runtime $20min$
- LF: `ResponseSize` $1MB$ – runtime $50s$
- LF*: `ResponseSize` $500B$ and 10 Requests – runtime $0.15s$

	$\mathcal{C}$
HF	1
LF	0.0417
LF*	0.000125

TABLE: Normalized Cost

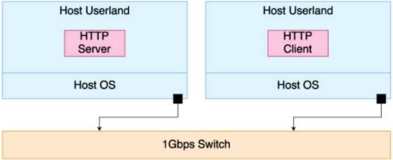


FIGURE: Network Configuration

Uncertainty Quantification
MC versus Multifidelity estimator

Requests/second (Expected Value)

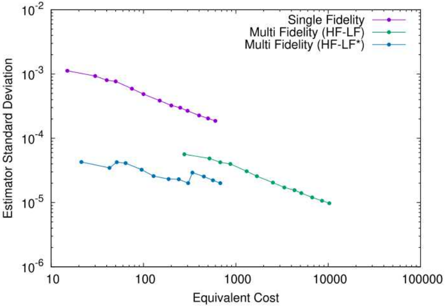


Figure: Estimators Standard Deviation.

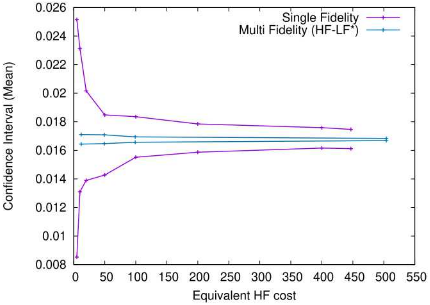


Figure: Confidence Interval convergence

First minimega-NS3 demonstration
Network Configuration: 1 Client - 1 Server

Network Configuration

- 1 client - 1 server (possible to extend to multiple clients)
- 100 Requests

Uncertain Parameters

- `DataRate` $\sim \mathcal{U}(5, 500)Mbps$
- `ResponseSize` $\sim \ln \mathcal{U}(500, 16 \times 10^6)B$

Fidelity definition

- `minimega` – HF: 100 Requests (average over 10 repetitions)
- `ns3` – LF: 10 Requests (`Delay` $50ms$)
- `ns3` – LF*: 1 Requests (`Delay` $5ms$)

	$\mathcal{C}$
HF	1
LF	0.016
LF*	0.002

Table: Normalized Cost

We assume **serial execution for the low-fidelity model**, however we might easily increase the efficiency of LF (`ns3`) by running multiple concurrent evaluations

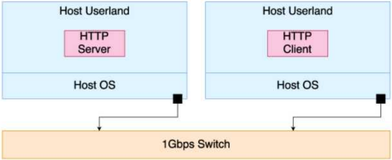


Figure: Network Configuration

FIRST minimega-NS3 DEMONSTRATION
ESTIMATOR STANDARD DEVIATION

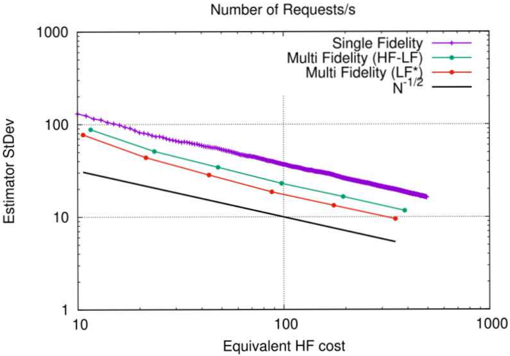


FIGURE: Exp. Value StDev

- The **variance reduction** we obtain w.r.t. MC is
$$\mathbb{V}ar\left(\tilde{Q}\left(\underline{\alpha}^{ACV}\right)\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - \frac{\mathbf{r_1 - 1}}{\mathbf{r_1}}\rho_{\mathbf{1}}^{\mathbf{2}}\right)$$
- The **number of low-fidelity simulations** is $N_{LF} = N \times r_1$ where
$$r_1 = \sqrt{\frac{\mathcal{C}_{HF}}{\mathcal{C}_{LF}} \frac{\rho_1^2}{1 - \rho_1^2}}$$
- For each HF simulation we need to spend an **extra cost** in LF simulations
$$\text{Eq.Cost}: \quad \mathcal{C}_{tot} = N\left(1 + \mathbf{r_1}\frac{\mathcal{C}_{\mathbf{LF}}}{\mathcal{C}_{\mathbf{HF}}}\right)$$
- For this case

	ρ_1	r_1	$r_1 \mathcal{C}_{LF}/\mathcal{C}_{HF}$
LF	0.86	4.69	0.075
LF*	0.90	10.83	0.022

FIRST minimega-NS3 DEMONSTRATION
EXPECTED VALUE ESTIMATION

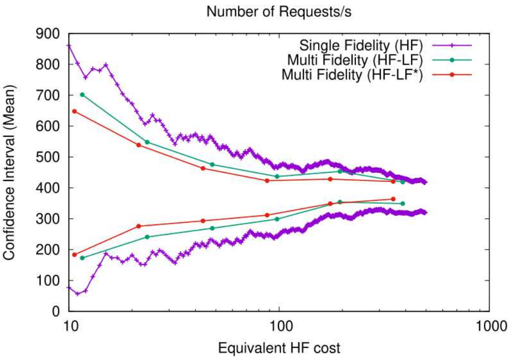


FIGURE: Exp. Value Confidence Interval

- The **variance reduction** we obtain w.r.t. MC is

$$\mathbb{V}ar\left(\tilde{Q}\left(\underline{\alpha}^{ACV}\right)\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - \frac{\mathbf{r_1 - 1}}{\mathbf{r_1}}\rho_1^2\right)$$

- The **number of low-fidelity simulations** is $N_{LF} = N \times r_1$ where

$$r_1 = \sqrt{\frac{\mathcal{C}_{HF}}{\mathcal{C}_{LF}} \frac{\rho_1^2}{1 - \rho_1^2}}$$

- For each HF simulation we need to spend an **extra cost** in LF simulations

$$\text{Eq.Cost}: \quad \mathcal{C}_{tot} = N\left(1 + \mathbf{r_1} \frac{\mathcal{C}_{\mathbf{LF}}}{\mathcal{C}_{\mathbf{HF}}}\right)$$

- For this case

	ρ_1	r_1	$r_1 \mathcal{C}_{LF}/\mathcal{C}_{HF}$
LF	0.86	4.69	0.075
LF*	0.90	10.83	0.022

FIRST minimega-NS3 DEMONSTRATION
ESTIMATOR STANDARD DEVIATION

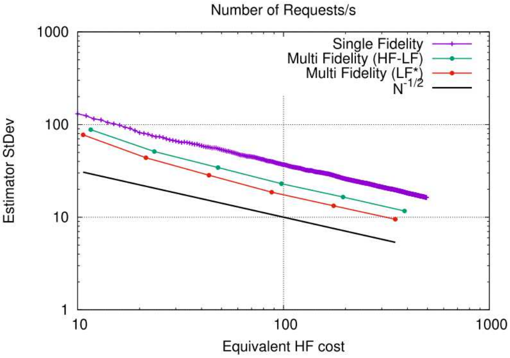


FIGURE: Exp. Value StDev

	ρ_1	r_1	$r_1 \mathcal{C}_{LF} / \mathcal{C}_{HF}$
LF	0.86	4.69	0.075
LF*	0.90	10.83	0.022

Example (for LF*)

- Number of **HF runs**: $N = 500$
- Number of **LF* runs**: $r_1 \times N = 5415$
- Equivalent **LF cost**: $r_1 \times N \times \frac{\mathcal{C}_{LF}}{\mathcal{C}_{HF}} = 11$
- **Total** estimator **cost** (HF + LF*): $\mathcal{C}_{tot} = 500 + 11 = 511$
- **Variance reduction**: $\left(1 - \frac{r_1 - 1}{r_1} \rho_1^2\right) = 0.23$

More than **70% of variance reduction** is obtained by adding **only an equivalent cost of 11 HF runs**

FIRST minimega-NS3 DEMONSTRATION
BEYOND THE SINGLE MODEL MULTIFIDELITY

Is it efficient to leverage multiple low-fidelity models at the same time?

	HF	**LF**	**LF***
HF	1	0.86	0.90
LF	0.86	1	0.99
LF*	0.90	0.99	1

TABLE: Correlation matrix

	OCV	**ACV**
HF+LF	0.26	0.39
HF+LF*	0.19	0.23
HF+LF+LF*	0.08	N/A

TABLE: Variance Reduction, $1 - R^2$

$$\mathbb{V}ar\left(\tilde{Q}\right) = \mathbb{V}ar\left(\hat{Q}\right)\left(1 - R^2\right)$$

NOTE:

- OCV assumes that the LF expected values are known, *i.e.* maximum attainable variance reduction

Concluding Remarks

Conclusions

Preliminary results: Multifidelity UQ for network applications

State-of-the-art

- Multifidelity Uncertainty Quantification proved to be effective for many different applications
- Encouraging preliminary results have been obtained for simple network configurations

Future Directions

- Extension to additional statistics (Tails, risk measures, *etc.*)
- Multifidelity Sensitivity Analysis
- Extension to discrete variables
- Extension to more complex network configurations/topologies
- Exploration of data-driven approaches for LF modelling (model reduction, active directions, *etc.*)
- Exploration of surrogate-based approaches

THANKS!

Acknowledgements

- Alex Gorodetsky, Mike Eldred and John Jakeman contributed to the ACV algorithm development
- Laboratory Directed Research & Development Funds @ Sandia – LDRD Grand Challenge SECURE

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA-0003525.