Multilevel-Multifidelity Sampling and Emulation for Forward UQ

Michael S. Eldred, Gianluca Geraci, Alex A. Gorodetsky, John D. Jakeman
Sandia National Laboratories, Albuquerque NM / University of Michigan, Ann Arbor MI
IPAM Workshop: HPC and Data Science for Scientific Discovery, UCLA, October 15-19, 2018

**UQ & Optimization: DOE/DOD Mission Deployment**

**Stewardship** (NNSA ASC)
Safety in abnormal environments

**Energy** (ASCR, EERE, NE)
Wind turbines, nuclear reactors

**Climate** (SciDAC, CSSEF, ACME)
Ice sheets, CISM, CESM, ISSM, CSDMS

**Addtln. Office of Science:** (SciDAC, EFRC)
Comp. Matls: waste forms / hazardous matls (WastePD, CHWM)
MHD: Tokamak disruption (TDS)

**Common theme** across these applications:
- High-fidelity simulation models: push forward SOA in computational M&S w/ HPC
- Severe simulation budget **constraints** (e.g., a handful of runs)
- Significant dimensionality, driven by model complexity (multi-physics, multiscale)

Focus on scalable algorithms in combination with approaches that can exploit a modeling hierarchy.
Multiple Model Forms in UQ & Opt

Discrete model choices for simulation of same physics

A clear hierarchy of fidelity (from low to high)
• Exploit less expensive models to render HF practical
  • Multifidelity Opt, UQ, inference
• Support general case of discrete model forms
  • Discrepancy does not go to 0 under refinement

An ensemble of peer models lacking clear preference structure / cost separation: e.g., SGS models
• With data: model selection, inadequacy characterization
  • Criteria: predictivity, discrepancy complexity
• Without (adequate) data: epistemic model form uncertainty propagation
  • Intrusive, nonintrusive
• Within MF context: CV correlation

Discretization levels / resolution controls
• Exploit special structure: discrepancy $\rightarrow 0$ at order of spatial/temporal convergence

Combinations for multiphysics, multiscale
Multilevel-Multifidelity Concepts Have Broad Relevance

**Recurring R&D theme:** couple scalable algorithms with exploiting a (multi-dimensional) model hierarchy

- address scale and expense for high fidelity M&S applications in defense, energy, and climate
- render UQ / optimization / OUU tractable for cases where only a handful of HF runs are possible

---

**Monte Carlo Methods (UQ)**

Optimal resource allocation: multilevel (ML), multifidelity (MF), and combined (MLMF)

- More than order of magnitude speedup vs. MC
- Render HF UQ possible (e.g., only 9 LES in 24D)
- Integrate w/ active subspaces (enhance $p$, link $\xi$)
- Unification of ML, MF, MI approaches

**Polynomial Chaos Methods (UQ)**

- ML rate estimation, greedy ML adaptation
- Sparse grids, compressed sensing, fn train
- Exploit problem structure: sparsity, low rank
- Additional orders of magn. when regular
- Sparse grids in model space

**Recursive Trust Region Methods (OUU)**

- Extend trust-region model mgmt. to deep hierarchies
- Manage both simulation and stochastic fidelity
- Order of magnitude fewer HF runs
- More aggressive profile shaping than MG/Opt

---

![UCAV Nozzle](image1.png)

Scramjet

![A2e wake dynamics](image2.png)

![UCAV Nozzle OUU (Aero, Structural, Thermal)](image3.png)
Monte Carlo Methods
Monte Carlo Sampling Methods
MSE for mean estimator

**Problem statement:** We are interested in the expected value of $Q_M = g(X_M)$ where

- $M$ is (related to) the number of spatial degrees of freedom
- $\mathbb{E} [Q_M] \xrightarrow{M \to \infty} \mathbb{E} [Q]$ for some RV $Q : \Omega \to \mathbb{R}$

**Monte Carlo:**

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

two sources of error:

- **Sampling error:** replacing the expected value by a (finite) sample average
- **Spatial discretization:** finite resolution implies $Q_M \approx Q$

Looking at the Mean Square Error:

$$\mathbb{E} \left[ (\hat{Q}_{MC, N}^{M} - \mathbb{E} [Q])^2 \right] = N^{-1} \text{Var} (Q_M) + (\mathbb{E} [Q_M - Q])^2$$

Accurate estimation $\Rightarrow$ **Large number** of samples at high (spatial) resolution
Multilevel and Multifidelity Sampling Methods

Multilevel MC: decomposition of variance

Ingredients:

- \( \{M_\ell : \ell = 0, \ldots, L\} \) with \( M_0 < M_1 < \cdots < M_L \overset{\text{def}}{=} M \)
- Estimation of \( \mathbb{E}[Q_M] \) by means of correction w.r.t. the next lower level

\[
Y_\ell \overset{\text{def}}{=} Q_{M_\ell} - Q_{M_{\ell-1}} \quad \xrightarrow{\text{linearity}} \quad \mathbb{E}[Q_M] = \mathbb{E}[Q_{M_0}] + \sum_{\ell=1}^{L} \mathbb{E}[Q_{M_\ell} - Q_{M_{\ell-1}}] = \sum_{\ell=0}^{L} \mathbb{E}[Y_\ell]
\]

- Multilevel Monte Carlo estimator

\[
\hat{Q}_{ML}^{M_{\ell}} = \sum_{\ell=0}^{L} Y_{\ell,N_{\ell}} \quad \text{with} \quad Y_{\ell,N_{\ell}} = \sum_{\ell=0}^{L} \frac{1}{N_{\ell}} \sum_{i=1}^{N_{\ell}} (Q_{M_\ell}^{(i)} - Q_{M_{\ell-1}}^{(i)})
\]

- The Mean Square Error is

\[
\mathbb{E}[(\hat{Q}_{ML} - \mathbb{E}[Q])^2] = \sum_{\ell=0}^{L} N_{\ell}^{-1} \text{Var}(Y_\ell) + (\mathbb{E}[Q_M - Q])^2
\]

Note: If \( Q_M \rightarrow Q \) (in a mean square sense), then \( \text{Var}(Y_\ell) \xrightarrow{\ell \to \infty} 0 \)
Multilevel and Multifidelity Sampling Methods
Multilevel MC: optimal resource allocation

Let us consider the numerical cost of the estimator

$$C(\hat{Q}_M^{ML}) = \sum_{\ell=0}^{L} N_\ell C_\ell$$

Determining the ideal number of samples per level (i.e. minimum cost at fixed variance)

$$C(\hat{Q}_M^{ML}) = \sum_{\ell=0}^{L} N_\ell C_\ell \quad \sum_{\ell=0}^{L} N_\ell^{-1} \text{Var} (Y_\ell) = \varepsilon^2 / 2$$

Balance ML estimator variance (stochastic error) and residual bias (deterministic error) → don’t over-resolve one at the expense of the other

Lagrange multiplier

$$N_\ell = \frac{2}{\varepsilon^2} \left[ \sum_{k=0}^{L} (\text{Var} (Y_k) / C_k) \right]^{1/2} \frac{\sqrt{\text{Var} (Y_\ell) / C_\ell}}{C_\ell}$$

Level independent

Optimal sample profile

Level dependent
Multilevel and Multifidelity Sampling Methods
Classical Control Variate → Multifidelity MC

A Control Variate MC estimator (function $G$ with $\mathbb{E}[G]$ known)

$$\hat{Q}_N^{MCCV} = \hat{Q}_N^{MC} - \beta \left( \hat{G}_N^{MC} - \mathbb{E}[G] \right)$$

$$\underset{\beta}{\text{argmin}} \mathbb{V} \text{ar} \left( \hat{Q}_N^{MCCV} \right) \implies \beta = -\rho \frac{\mathbb{V} \text{ar}^{1/2}(Q)}{\mathbb{V} \text{ar}^{1/2}(G)}$$

$$\mathbb{V} \text{ar} \left( \hat{Q}_N^{MCCV} \right) = \mathbb{V} \text{ar} \left( \hat{Q}_N^{MC} \right) \left( 1 - \rho^2 \right)$$

In our context, $G$ is a low fidelity approximation of $Q$ and its expectation is not known a priori.

Let’s modify the high-fidelity QoI, $Q_{M}^{HF}$, to decrease its variance

$$\hat{Q}_{M,N}^{HF,\text{CV}} = \hat{Q}_{M,N}^{HF} + \alpha \left( \hat{Q}_{M,N}^{LF} - \mathbb{E}[Q_{M}^{LF}] \right)$$

additional and independent set $\Delta^{LF} = r^{NF}$

Minimize CV estimator variance → control param. as before:

$$\frac{d \mathbb{V} \text{ar}(\hat{Q}_M^{HF,\text{MF}})}{d \alpha} = 0 \implies \alpha = -\rho \frac{\mathbb{V} \text{ar}^{1/2}(Q_{M}^{HF})}{\mathbb{V} \text{ar}^{1/2}(Q_{M}^{LF})}$$

Minimize total cost → optimal sample ratio: $r^* = -1 + \sqrt{\frac{w \rho^2}{1 - \rho^2}}$

MFMC cost relative to MC

Pasupathy et al., 2012; Ng and Willcox, 2014; Peherstorfer, Willcox, & Gunzburger, 2016; et al.
Multilevel – Multifidelity Sampling Methods
Combining ML and CV for multidimensional model hierarchies

\[ \mathbb{E} \left[ Q_{M}^{HF} \right] = \sum_{l=0}^{L_{HF}} \mathbb{E} \left[ Y_{l}^{HF} \right] = \sum_{l=0}^{L_{HF}} \hat{Y}_{l}^{HF} \]

\( Y_{l}^{HF,*} = \hat{Y}_{l}^{HF} + \alpha_{l} \left( \hat{Y}_{l}^{LF} - \mathbb{E} \left[ Y_{l}^{LF} \right] \right) \)

\[ \L_{\ell}(r_{\ell}) \leq 1 - \rho_{\ell}^{2} \frac{r_{\ell}}{1 + r_{\ell}} \]

Cost per level is now \( C_{\ell}^{eq} = C_{\ell}^{HF} + C_{\ell}^{LF} (1 + r_{\ell}) \)

Optimal sample allocation across discretizations and model forms

\[ N_{l}^{HF,*} = 2 \frac{\varepsilon^{2}}{\var_{y_{l}^{HF}}(1 - \rho_{\ell}^{2})} \left[ \frac{\var_{y_{l}^{HF}}(1 - \rho_{\ell}^{2})}{C_{\ell}^{HF}} \right]^{1/2} \Lambda_{\ell} \]

Multilevel – Multifidelity Sampling Methods
Combining ML and CV for multidimensional model hierarchy

- Algorithmic-contained correlation improvement
- Optimality of the LF discrepancy

\[ \hat{Y}_{\ell}^{\text{LF}} = \gamma_{\ell} Q_{\ell}^{\text{LF}} - Q_{\ell-1}^{\text{LF}}, \]

where \( \gamma \) is chosen in order to maximize the correlation between \( Y_{\ell}^{\text{HF}} \) and \( \hat{Y}_{\ell}^{\text{LF}} \)

Cost min s.t.
error balance

\[ \theta_{\ell} = \frac{\text{Cov} \left( Y_{\ell}^{\text{LF}}, \hat{Y}_{\ell}^{\text{LF}} \right)}{\text{Cov} \left( Y_{\ell}^{\text{LF}}, Y_{\ell}^{\text{LF}} \right)} \]

\[ \tau_{\ell} = \frac{\text{Var} \left( \hat{Y}_{\ell}^{\text{LF}} \right)}{\text{Var} \left( Y_{\ell}^{\text{LF}} \right)} \]

\[ \Lambda_{\ell} = 1 - \rho_{\ell}^2 \frac{r_{\ell}^*}{1 + r_{\ell}^*} \]

\[ r_{\ell}^* = -1 + \frac{\rho_{\ell}^2 \sigma_{\ell}^2}{1 - \rho_{\ell}^2 \sigma_{\ell}^2} \omega_{\ell} \]

Can be rewritten as

\[ N_{\ell}^{\text{HF}} = \frac{2}{\varepsilon^2} \left[ \sum_{k=0}^{L_{\text{HF}}} \left( \frac{\text{Var} \left( Y_{k}^{\text{HF}} \right) C_{k}^{\text{HF}}}{1 - \rho_{HL}^2} \right)^{1/2} \Lambda_{k}(r_{k}) \right] \sqrt{\left( 1 - \rho_{HL}^2 \right) \frac{\text{Var} \left( Y_{\ell}^{\text{HF}} \right)}{C_{\ell}^{\text{HF}}}} \]

Compared to previous

\[ N_{\ell}^{\text{HF}} = \frac{2}{\varepsilon^2} \left[ \sum_{k=0}^{L_{\text{HF}}} \left( \frac{\text{Var} \left( Y_{k}^{\text{HF}} \right) C_{k}^{\text{HF}}}{1 - \rho_{HL}^2} \right)^{1/2} \Lambda_{k}(r_{k}) \right] \sqrt{\left( 1 - \rho_{HL}^2 \right) \frac{\text{Var} \left( Y_{\ell}^{\text{HF}} \right)}{C_{\ell}^{\text{HF}}}} \]

Multilevel – Multifidelity Sampling Methods

Results on model problem: wave propagation in composites

- Rod constituted by 50 layers, two alternated materials (A and B) with constitutive laws
  \[
  \begin{align*}
  \sigma_A &= K_1^A \epsilon + K_2^A \epsilon^2, \quad K_1^A = 1 \quad \text{and} \quad K_2^A = \xi_j \\
  \sigma_B &= K_1^B \epsilon + K_2^B \epsilon^2, \quad K_1^B = 1.5 \quad \text{and} \quad K_2^A = 0.8
  \end{align*}
  \]

- Uncertain initial static \((u(x, t = 0) = 0)\) pre-tension state:
  \[
  \sigma(x) = \begin{cases} 
  \xi_3 \exp \left( -\frac{(x - 0.35)(x - 0.25)}{2 \times 0.002} \right) & \text{if } 0 < x < 1/2 \quad \xi_3 \sim \mathcal{U}(0.5, 2) \\
  \xi_2 \exp \left( -\frac{(x - 0.65)(x - 0.75)}{2 \times 0.002} \right) & \text{if } 1/2 < x < 1 \quad \xi_2 \sim \mathcal{U}(0.5, 6.5)
  \end{cases}
  \]

- Spatially varying uncertain density: \(\rho(x) = \xi_1 + 0.5 \sin (2\pi x), \xi_1 \sim \mathcal{U}(1.5, 2)\)

- Clamped rod as B.C.

28 random variables
Two fidelities, each with 4 discretizations

<table>
<thead>
<tr>
<th>Level</th>
<th>MLMC (N_\ell)</th>
<th>MLMF-YI (N_{\ell}^{HF})</th>
<th>(N_{\ell}^{LF})</th>
<th>MLMF-QI (N_{\ell}^{HF})</th>
<th>(N_{\ell}^{LF})</th>
<th>(r_\ell)</th>
<th>(\rho_\ell^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>80029</td>
<td>5960</td>
<td>243178</td>
<td>40</td>
<td>0.97</td>
<td>4682</td>
<td>192090</td>
</tr>
<tr>
<td>1</td>
<td>6282</td>
<td>2434</td>
<td>12487</td>
<td>14</td>
<td>0.49</td>
<td>1049</td>
<td>13781</td>
</tr>
<tr>
<td>2</td>
<td>1271</td>
<td>262</td>
<td>3877</td>
<td>14</td>
<td>0.82</td>
<td>151</td>
<td>3657</td>
</tr>
<tr>
<td>3</td>
<td>212</td>
<td>47</td>
<td>966</td>
<td>19</td>
<td>0.84</td>
<td>34</td>
<td>754</td>
</tr>
</tbody>
</table>
DARPA SEQUOIA: Hierarchy of Fidelity Levels

Low fidelity model
- Quasi 1D ideal/non ideal nozzle aero
- 1D heat transfer
- Coarse axisymmetric FEM model
- 30 seconds on one core

Medium fidelity model
- 2D Euler/RANS axisymmetric CFD
- 1D heat transfer
- Coarse axisymmetric FEM model
- 5 minutes on one core (2D Euler)

High fidelity model
- 3D non-axisymmetric Euler/RANS CFD
- 1D heat transfer
- Full 3D FEM model
- 2 hours on 20 cores (3D RANS, coarse mesh)

Multiple mesh refinements available for Medium & High (ragged ML-MF)
**Context:** Analysis of performance of UCAV nozzles subject to environmental, material, and manufacturing uncertainties.

**Goal:** Explore utility of low fidelity model (potential flow, hoop stress) alongside discretizations for medium fidelity (Euler, FEM)

<table>
<thead>
<tr>
<th>Target accuracy</th>
<th>LF Coarse</th>
<th>LF Medium</th>
<th>LF Fine</th>
<th>MF Coarse</th>
<th>MF Medium</th>
<th>MF Fine</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>21143</td>
<td>1757</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>0.003</td>
<td>69580</td>
<td>5775</td>
<td>36</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>0.001</td>
<td>212828</td>
<td>17715</td>
<td>109</td>
<td>34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Optimal sample allocations based on relative cost, observed correlation between models, and observed variance distribution across levels.
Updated Deployment of MLCV MC to UCAV Nozzle UQ

<table>
<thead>
<tr>
<th></th>
<th>LF Variance reduction [%]</th>
<th>LF (updated) Variance reduction [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>correlation</td>
<td></td>
</tr>
<tr>
<td>Thrust</td>
<td>0.997</td>
<td>91.42</td>
</tr>
<tr>
<td>Mechanical Stress</td>
<td>2.31e-5</td>
<td>2.12e-3</td>
</tr>
<tr>
<td>Thermal Stress</td>
<td>0.391</td>
<td>12.81</td>
</tr>
</tbody>
</table>

Estimator Variance (normalized)

Equivalent HF runs

<table>
<thead>
<tr>
<th>Accuracy ($\varepsilon^2/\varepsilon_0^2$)</th>
<th>LF Coarse</th>
<th>Medium Fidelity</th>
<th>LF (updated)</th>
<th>Medium Fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coarse</td>
<td>Medium</td>
<td>Fine</td>
<td>Coarse</td>
</tr>
<tr>
<td>0.1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>404</td>
</tr>
<tr>
<td>0.01</td>
<td>21,143</td>
<td>1,757</td>
<td>20</td>
<td>3,091</td>
</tr>
<tr>
<td>0.003</td>
<td>69,580</td>
<td>5,775</td>
<td>36</td>
<td>N/A</td>
</tr>
<tr>
<td>0.001</td>
<td>212,828</td>
<td>17,715</td>
<td>109</td>
<td>32,433</td>
</tr>
</tbody>
</table>
DARPA EQUiPS (Scramjet UQ):
LES Models for Turbulent Reacting Flow in HIFiRE

- Provided benchmark LES calculations of the Hypersonic International Flight Research Experiment (HIFiRE) to support development of UQ
- Case of interest corresponds to the geometry and conditions of ground based experiments performed in the HIFiRE Direct Connect Rig (HDCR)
- A hierarchy of unit cases (including high-fidelity LES of the HDCR) has facilitated UQ tasks and provided optimal workflow between team members
- Unit cases are designed to emulate key QoIs while making comprehensive parametric studies possible

Model forms:
- 2D, 3D
Discretizations:
- d/{8,16,32,64}
3D LES simulation of scramjets is extremely expensive and a significant challenge for UQ; even more so for OUU.

**Goal:** Demonstrate UQ in moderately high D using only a “handful” of HF simulations, by leveraging lower fidelity 2D models and coarsened 2D/3D discretizations.

**UQ Approach:** MLCV algorithm described previously.

**Optimized allocation:** achieve MSE target for 3D LES in 24D using only 9 HF sims. (50 equiv HF)
Updated Deployment of MLCV MC for Scramjet UQ

**P1 updated:** re-formulate inputs in order to obtain an higher level of turbulence and, in turn, a more non-linear response of the system

<table>
<thead>
<tr>
<th></th>
<th>$P_0,\text{mean}$</th>
<th>$P_0,\text{rms,mean}$</th>
<th>$M_{\text{mean}}$</th>
<th>$\text{TKE}_{\text{mean}}$</th>
<th>$\chi_{\text{mean}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d/8$</td>
<td>4.02554e-03</td>
<td>1.90524e-06</td>
<td>1.99236e-02</td>
<td>3.34905e-07</td>
<td>4.24520e-03</td>
</tr>
<tr>
<td>$d/16$</td>
<td>4.03350e-07</td>
<td>7.77838e-08</td>
<td>6.68974e-05</td>
<td>1.74847e-08</td>
<td>4.40048e-05</td>
</tr>
</tbody>
</table>

Table 2: Variance for the five QoIs of the P1 unit problem.

**Observations from pilot sample:** decay in variance across discretizations (LF d/8 and discrepancy d/16 – d/8) no longer observed for all QoI

**Implications:** requires more focused analysis of deterministic convergence properties → Need to engage additional refinement levels (i.e., d/32, d/64) in order to converge QoI statistics that are closely tied to resolution of turbulence.
Multilevel – Multifidelity Sampling Methods

Cardiovascular flow

Model relationships / graph topologies

![Diagram showing model relationships and graph topologies](image)

Costs to achieve prescribed error tolerance

<table>
<thead>
<tr>
<th>Method</th>
<th>Effective Cost (3D Simulations)</th>
<th>No. 3D Simulations</th>
<th>No. 1D Simulations</th>
<th>No. 0D Simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>9.885</td>
<td>9.885</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MFA</td>
<td>56</td>
<td>21</td>
<td>15,681</td>
<td>–</td>
</tr>
<tr>
<td>MFB</td>
<td>39</td>
<td>36</td>
<td>–</td>
<td>154,880</td>
</tr>
<tr>
<td>MLA</td>
<td>305</td>
<td>212</td>
<td>41,990</td>
<td>–</td>
</tr>
<tr>
<td>MLB</td>
<td>156</td>
<td>150</td>
<td>–</td>
<td>342,060</td>
</tr>
<tr>
<td>MLC</td>
<td>165</td>
<td>156</td>
<td>1,324</td>
<td>351,940</td>
</tr>
<tr>
<td>MLMF</td>
<td>165</td>
<td>156</td>
<td>1,249</td>
<td>362,590</td>
</tr>
</tbody>
</table>

Courtesies of C. Fleeter (Stanford), Prof. D. Schiavazzi (Notre Dame), Prof. A. Marden (Stanford)
Multilevel – Multifidelity Sampling Methods
Research Direction: Leveraging active directions (ECCOMAS, WCCM)

• Active subspaces, ridge approximation, adapted basis, …

Let’s introduce the $m \times m$ matrix $C$

$$C = \int \left( \nabla f \right) \left( \nabla f \right)^T \rho(x) \, dx$$

Since $C$ is I) Positive semidefinite and II) Symmetric, it exists a real eigenvalue decomposition

$$C = W \Lambda W^T,$$ where

$W$ is the $m \times m$ orthogonal matrix whose columns are the normalized eigenvectors

$\Lambda = \text{diag} \{ \lambda_1, \ldots, \lambda_m \}$ and $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$

Let’s define two sets of variables

$$\begin{cases} y = W_A^T x \in \mathbb{R}^n & \text{(Active)} \\ z = W_I^T x \in \mathbb{R}^{(m-n)} & \text{(Inactive)} \end{cases}$$

$$\Rightarrow x = W_A y + W_I z \approx W_A y$$

• Main ideas:
  • For each model independently one can compute active directions
  • Sample along these shared active directions and map back to original model coords.
  • Principal directions for a shared QoI can bridge dissimilar parameterizations and demonstrate underlying shared processes

Multilevel – Multifidelity Sampling Methods
Research Direction: leveraging active directions (example 1)

Multilevel – Multifidelity Sampling Methods
Research Direction: leveraging active directions (example 1)

- Fixed computational budget of equivalent of 300 HF runs (LF cost ratio = 100)
- 1000 realizations for each estimator → pdf of estimated Expected Value
- Active subspace discovery for each realization during the pilot sample phase

Multilevel – Multifidelity Sampling Methods

Research Direction: leveraging active directions (example 2)

Wave propagation test problem

Enhances correlation (even if initially high) and links (dissimilar) model parameterizations

<table>
<thead>
<tr>
<th></th>
<th>$N_x$</th>
<th>$N_{t}$</th>
<th>$\Delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-fidelity</td>
<td>5</td>
<td>50</td>
<td>$36 \times 10^{-4}$</td>
</tr>
<tr>
<td>High-fidelity</td>
<td>801</td>
<td>600</td>
<td>$30 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

![Graph showing Active Direction Agnostic sampling with $\rho^2 = 0.89$ and Active Direction Aware sampling with $\rho^2 = 0.99$.]

![Histogram showing 250 Estimator Realizations with Expected Value ranging from 2.8 to 5.0.]

<table>
<thead>
<tr>
<th>Method</th>
<th>HF runs</th>
<th>LF runs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>40</td>
<td>5946</td>
</tr>
<tr>
<td>MC-MF</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>MC-MFAS</td>
<td>32</td>
<td>21185</td>
</tr>
</tbody>
</table>
Multilevel – Multifidelity Sampling Methods

Research Direction: Generalized control variates (in internal Sandia review)

- Unification of ML and CV approaches
- Look beyond (recursive) model pairings

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

\[ \operatorname{arg min}_\alpha \operatorname{Var} \left[ \hat{Q}^{CV}(\alpha) \right] \]

\[ C \in \mathbb{R}^{M \times M} \ \text{covariance matrix among } Q_i \]
\[ c \in \mathbb{R}^{M} \ \text{vector of covariances between } Q \]
\[ \alpha^* = C^{-1} c \]

A. Gorodetsky, G. Geraci, E., J. Jakeman “Approximate Control Variates,” (in internal SNL R&A; expected on arxiv next week).
Multilevel – Multifidelity Sampling Methods
Research Direction: Generalized control variates (in internal Sandia review)

- Unification of ML and CV approaches
- Look beyond (recursive) model pairings

\[
\hat{Q}^{\text{CV}} = \hat{Q} + \sum_{i=1}^{M} \alpha_i \left( \hat{q}_i - \mu_i \right)
\]

\[
\arg\min_{\alpha} \text{Var} \left[ \hat{Q}^{\text{CV}}(\alpha) \right] = \begin{cases} 
C \in \mathbb{R}^{M \times M} & \text{covariance matrix among } Q_i \\
 c \in \mathbb{R}^{M} & \text{vector of covariances between } Q \text{ and each } Q_i \\
 \alpha^* = C^{-1}c 
\end{cases}
\]

A. Gorodetsky, G. Geraci, E., J. Jakeman “Approximate Control Variates,” (in internal SNL R&A; expected on arxiv next week).
Summary: Monte Carlo Methods

The case for multilevel and multifidelity methods
- Push towards higher simulation fidelity can make propagation / inference / OUU untenable
- Multiple model fidelities / discretizations are often available that trade accuracy for cost
- Realistic deployments (nozzle, scramjet, cardio) → rich model ensemble, challenging QoI

Towards multilevel-multifidelity UQ tailored for smoothness and dimensionality
- Multilevel-multifidelity MC framework for cost-optimized variance reduction
  - ML MC targets variance decay within a discretization hierarchy
  - MF MC (control variates) targets correlation between HF and 1 or more LF models
  - ML-MF MC employs LF control variate at each HF discretization level → tailors approach to hierarchy type; leverages multiple variance reduction opportunities
  - Well suited for high dimensionality and/or low regularity

Research Directions
- Leverage active directions to enhance correlation and bridge dissimilar parameterizations
- Relax assumed model relationships / graph topologies to expose additional performance
Stochastic Polynomial Expansion Methods

• Projection, Regression, Interpolation
• Multilevel | Multifidelity expansions (heuristic)
**Stochastic Expansions: Polynomial Chaos & Stochastic Collocation**

**Polynomial chaos:** spectral projection using orthogonal polynomial basis fnns

\[ R = \sum_{j=0}^{\infty} \alpha_j \Psi_j(\xi) \]

- Estimate \( \alpha_j \) using regression or numerical integration: sampling, tensor quadrature, sparse grids, or cubature

**Stochastic collocation:** instead of estimating coefficients for known basis functions, form interpolants for known coefficients

- **Global:** Lagrange (values) or Hermite (values+derivatives)
- **Local:** linear (values) or cubic (values+gradients) splines

- **Nodal** or **Hierarchical** interpolants

\[ L_j = \prod_{k=1}^{m} \frac{\xi - \xi_k}{\xi_j - \xi_k} \]

Sparse interpolants formed using \( \sum \) of tensor interpolants

- **Tailor expansion form:**
  - p-refinement: anisotropic tensor/sparse, generalized sparse
  - h-refinement: local bases with dimension & local refinement

- **Method selection:** requirements for fault tolerance, decay, sparsity, error estimation
MF UQ with Spectral Stochastic Discrepancy Models

- High-fidelity simulations (e.g., RANS, LES) can be prohibitive for use in UQ
- Low fidelity “design” codes often exist that are predictive of basic trends
- Can we leverage LF codes w/i HF UQ in a rigorous manner? → global approxs. of model discrepancy

\[
\hat{f}_{hi}(\xi) = \sum_{j=1}^{N_{lo}} f_{lo}(\xi_j) L_j(\xi) + \sum_{j=1}^{N_{hi}} \Delta f(\xi_j) L_j(\xi)
\]

\(N_{lo} \gg N_{hi}\)

**Sparse grid bi-fidelity:** target reduced complexity in model discrepancy

**Compressed sensing bi-fidelity:** target sparsity

*(Functional) tensor train bi-fidelity:* target low rank

Stochastic Polynomial Expansion Methods
- Projection, Regression, Interpolation
- Multilevel | Multifidelity expansions (heuristic)
- Multilevel | Multifidelity expansions (optimized)
Formulations for Multilevel PCE / SC

1. **Optimal resource allocation:** parameterize estimator variance $\rightarrow$ optimal $N_l$
   
   Global $\kappa$ and $\gamma > 0$
   
   \[ \text{Var}[\hat{Y}_l] = \frac{\text{Var}[Y_l]}{\gamma N_l^\kappa} \rightarrow N_l = \sqrt[\kappa+1]{\frac{2}{\varepsilon^2 \gamma}} \sum_{q=0}^{L} \sqrt[\kappa+1]{\text{Var}[Y_q] C_q^\kappa} \sqrt[\kappa+1]{\frac{\text{Var}[Y_l]}{C_l}} \]


   **Main challenge:** abrupt transitions in sparse / low rank recovery

2. **Restricted Isometry Property (RIP) for sparse recovery**
   
   \[ N_l \geq s_l \log^3(s_l) L_l \log(C_l) \]

   Jakeman, Narayan, and Zhou, 2016

   Main challenge: compressible fns
   
   $\rightarrow$ increasing $s$
   
   $\rightarrow$ feedback not well controlled

3. **Greedy Multilevel refinement**
   
   ML competition with multiple level candidate generators

   Main challenges: scalable refinement schemes, loss of precision
ML PCE with rate estimation: Model Problem & UCAV Nozzle

\[ -\frac{d}{dx} \left[ a(x, \xi) \frac{du}{dx} (x, \xi) \right] = 10, \quad (x, \xi) \in (0, 1) \times I_\xi, \tag{22} \]

where \( x \) is the spatial coordinate, \( \xi \) a vector of independent random input parameters and \( a(x, \xi) \) denotes the (random) diffusivity field. The following Dirichlet boundary conditions are also assumed

\[ u(0, \xi) = 0, \quad u(1, \xi) = 0. \tag{23} \]

We are interested in quantifying the uncertainty in the solution \( u \) at specified spatial locations: \( x = 0.05, 0.5, 0.95 \).

We represent the random diffusivity field \( a \) using the following expansion

\[ a(x, \xi) = 1 + \sigma \sum_{k=1}^{d} \frac{1}{k^2} \cos(2\pi k x) \xi_k \tag{24} \]

Optimized resource allocation outperforms previous heuristics: \( \kappa > 1 \) is effective

Initial results were promising, but rate estimation impeded by abrupt transition in recovery

Optimal sample allocations based on relative cost, variance distribution across levels and \( \kappa = 2 \)

ML PCE shows more rapid convergence using coarse/medium/fine discretizations:
- Exploits smoothness in moderate dim.

ML PCE with rate estimation:
- Model Problem & UCAV Nozzle
- SS Diffusion
- Optimized resource allocation outperforms previous heuristics: \( \kappa > 1 \) is effective
- Initial results were promising, but rate estimation impeded by abrupt transition in recovery
- Optimal sample allocations based on relative cost, variance distribution across levels and \( \kappa = 2 \)
- ML PCE shows more rapid convergence using coarse/medium/fine discretizations:
  - Exploits smoothness in moderate dim.
Formulations for Multilevel PCE / SC

1. Optimal resource allocation: parameterize estimator variance → optimal $N_l$
   \[ \frac{\text{Var}[\hat{Y}_l]}{\gamma N_l^\kappa} \rightarrow N_l = \left( \frac{2}{c^2 \gamma} \sum_{q=0}^{L} \kappa+1 \text{Var}[Y_q] C_q \right)^{\frac{1}{\kappa+1}} \text{Var}[Y_l] C_l \]
   Global $\kappa$ and $\gamma > 0$
   
   Main challenge: abrupt transitions in sparse / low rank recovery


2. Restricted Isometry Property (RIP) for sparse recovery
   \[ N_l \geq s_l \log^3(s_l) L_l \log(C_l) \]
   Main challenge: compressible fns → increasing $s$ → feedback not well controlled

3. Greedy Multilevel refinement
   ML competition with multiple level candidate generators
   Main challenges: scalable refinement schemes, loss of precision
Multilevel-Multifidelity expansions – Greedy refinement

Compete level refinement candidates to maximize induced change per unit cost:
• 1 or more refinement candidates per level
• Measure impact on final QoI statistics (roll up multilevel estimates),
  • norm of change in response covariance (default)
  • norm of change in level mappings (goal-oriented: \( z/p/\beta/b^* \))
  normalized by relative cost of level increment (# new points * cost / point)
• Greedy selection of best candidate, which generates new candidate(s) for selected level

Level candidate generators:
• Uniform refinement of orders / levels (coarse-grained, 1 candidate per level)
  • Tensor / sparse grids: PCE and nodal/hierarchical SC
  • Regression PCE: least sq. / compressed sensing using fixed sample ratio
• Anisotropic refinement of orders / levels (coarse-grained, 1 candidate per level)
  • Tensor / sparse grids
• Index-set-based refinement (fine-grained, many candidates per level: exp growth w/ dim)
  • Generalized sparse grids: PCE and nodal/hierarchical SC
  • Regression PCE
• Adapted basis (coarse-grained, a few exp order frontier advancements per level)
  • Regression PCE

Multilevel-multifidelity expansion methods:
Greedy ML PCE: CS + uniform basis refinement

<table>
<thead>
<tr>
<th>Conv Tol</th>
<th>$N_1$</th>
<th>$N_2$</th>
<th>$N_3$</th>
<th>$N_4$</th>
<th>$N_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.e-1</td>
<td>198</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>1.e-2</td>
<td>644</td>
<td>198</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>1.e-3</td>
<td>1802</td>
<td>644</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>1.e-4</td>
<td>4505</td>
<td>1802</td>
<td>50</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>
Multilevel-multifidelity expansion methods
Greedy ML PCE: uniform/generalized sparse grids

Generalized sparse grid @ each level:
• Combinatorial growth in refinement candidates

![Graphs showing mean error and standard deviation error vs. equivalent HF simulations](image)

<table>
<thead>
<tr>
<th>Conv Tol</th>
<th>$N_1$</th>
<th>$N_2$</th>
<th>$N_3$</th>
<th>$N_4$</th>
<th>$N_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.e-2</td>
<td>43</td>
<td>23</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>1.e-4</td>
<td>211</td>
<td>83</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>1.e-6</td>
<td>391</td>
<td>271</td>
<td>156</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>1.e-8</td>
<td>1359</td>
<td>743</td>
<td>327</td>
<td>59</td>
<td>19</td>
</tr>
<tr>
<td>1.e-10</td>
<td>3535</td>
<td>2311</td>
<td>1039</td>
<td>391</td>
<td>19</td>
</tr>
<tr>
<td>1.e-12</td>
<td>10319</td>
<td>5783</td>
<td>2783</td>
<td>1343</td>
<td>43</td>
</tr>
<tr>
<td>1.e-14</td>
<td>26655</td>
<td>14991</td>
<td>8063</td>
<td>3703</td>
<td>1535</td>
</tr>
</tbody>
</table>

Greedy ML GSG sample profiles
Multilevel-multifidelity expansion methods
Greedy ML PCE: overlay all cases & references

CS approaches have greater flexibility at low sample levels (lower initialization cost), but accuracy currently limited by numerical issues for large systems allocated at coarse levels.
ML PCE / SC: Directions

Current developmental areas:

- Hierarchical interpolation (Δ precision for small grid increments)
- Functional tensor train (large systems: scalability of level solver, especially @ LF)
- Limiting number of level candidates (expanding front: MLMF scalability)
- Multidimensional model hierarchies → greedy sparse grids in model space

Interpolation via hierarchical surplus

From X. Ma, 2010

UCAV Nozzle for structural/thermal fidelity

J.D Jakeman, E., G. Geraci, A. Gorodetsky, “Adaptive Multi-index Collocation and Sensitivity Analysis” (in internal review)
Summary Remarks (ML PCE)

The case for multilevel and multifidelity methods
• Push towards higher simulation fidelity can make propagation / inference / OUU untenable
• Multiple model fidelities / discretizations are often available that trade accuracy for cost
• Deployments for CFD (nozzle, scramjet, wind) → rich model ensemble, challenging QoI

Towards multilevel-multifidelity UQ tailored for smoothness and dimensionality
• Multilevel-multifidelity MC framework for cost-optimized variance reduction
  • ML-MF MC employs LF control variate at each HF discretization level; tailor to hierarchy type
  • Well suited for high dimensionality and/or low regularity
• Multilevel PCE/SC: extend ML MC machinery with higher performance estimators
  extend heuristic multifidelity PCE/SC with optimal allocations
  • PCE CS / FT: exploit sparsity / low rank in δ; SC hierarchical interp: direct Δ calculation
  • Rate estimation of estimator variance: complicated by abrupt transitions in CS/FT recovery
  • RIP sampling: shape sample profile based on observed sparsity; issues w/ feedback
  • Greedy refinement: competition among multiple candidates per level, normalized by cost
    • ML compressed sensing with expansion order candidates
    • ML (generalized) sparse grids with level and index set candidates
  • Achieve more rapid convergence (sufficient regularity, moderate dimensionality)

Related Efforts:
• Multilevel Bayesian inference → exploit ML PCE/FT within emulator-based inference
• Multilevel Opt/OUU → move beyond common bi-fidelity to exploit deep / multi-D hierarchy