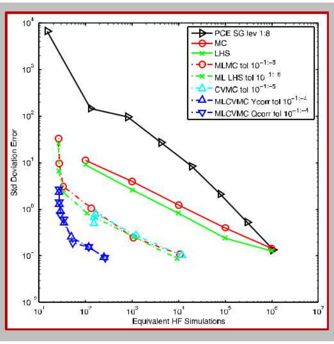
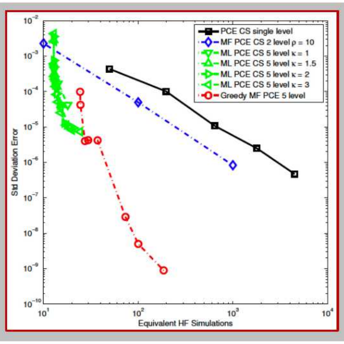
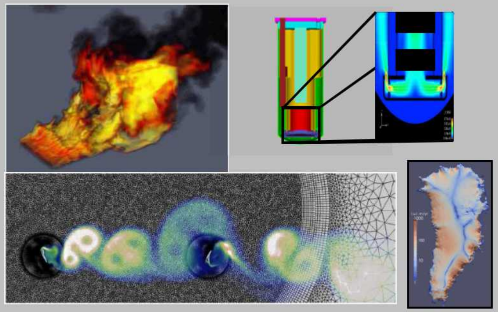
SAND2019-10986PE





DAKOTA

Explore and predict with confidence



Multilevel / Multifidelity Sampling and Emulation for Forward UQ

Michael S. Eldred, Gianluca Geraci, Alex Gorodetsky, and John Jakeman
Sandia National Laboratories, Albuquerque NM, and University of Michigan, Ann Arbor MI
Applied Math Visioning Workshop, Sept. 17—18, 2019

U.S. DEPARTMENT OF ENERGY

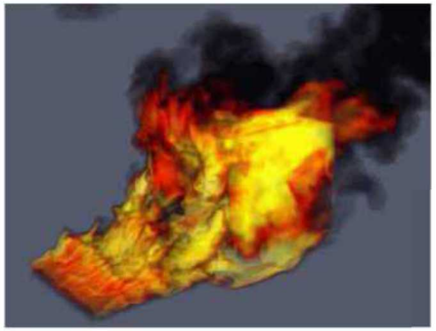
NNSA National Nuclear Security Administration

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.

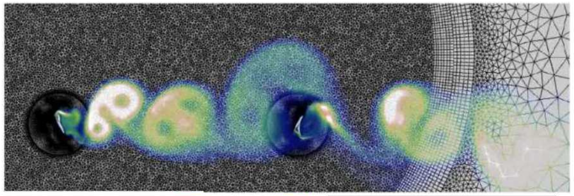
UQ & Optimization: DOE/DOD Mission Deployment

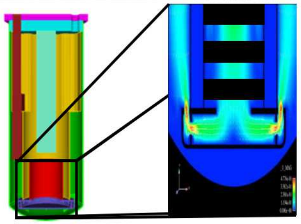
Sandia National Laboratories

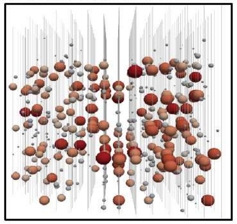
Stewardship (NNSA ASC)
Safety in abnormal environments



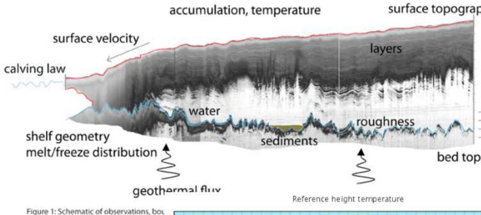
Energy (ASCR, EERE, NE)
Wind turbines, nuclear reactors

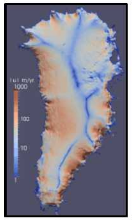


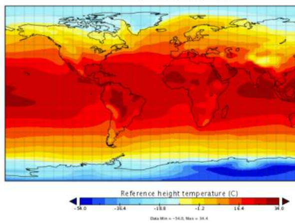




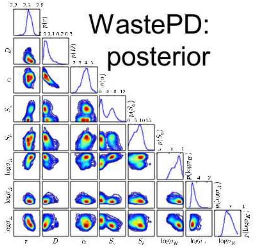
Climate (SciDAC, CSSEF, ACME)
Ice sheets, CISM, CESM, ISSM, CSDMS

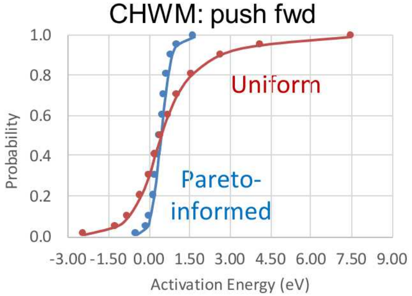


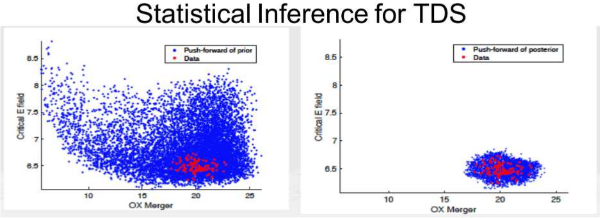




Addtnl. Office of Science:
(SciDAC, EFRC)
Comp. Matls: waste forms / hazardous matls (WastePD, CHWM)
MHD: Tokamak disruption (TDS)





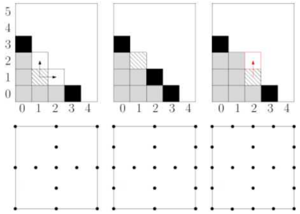


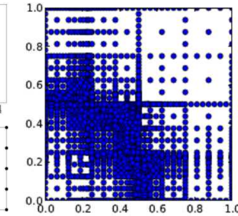
Common theme across these applications:

- 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)

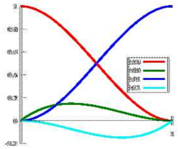
Research Thrusts for UQ

- *Focus:* Compute dominant uncertainty effects despite key challenges

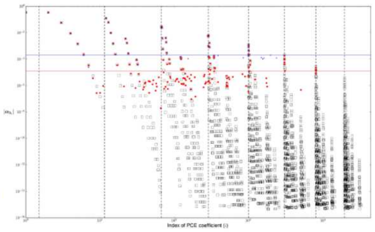




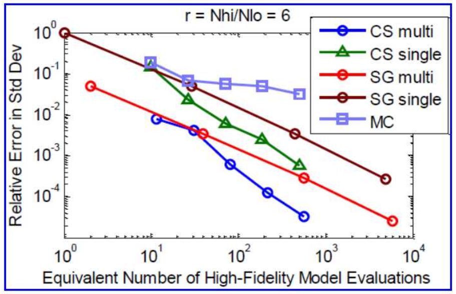
- Emphasize scalability through exploitation of special structure
  - *Adaptivity:* p- and h- refinement of stochastic expansions
  - *Adjoints:* gradient enhancement for PCE / SC / GP
  - *Sparsity:* compressed sensing
  - *Low Rank:* tensor / function train (w/ UMich)
  - *Dimension reduction:* active subspaces (w/ CU Boulder), adapted basis PCE (w/ USC)



$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ \pi_{0,j}(\vec{\xi}_i) & \pi_{1,j}(\vec{\xi}_i) & \cdots & \pi_{P,j}(\vec{\xi}_i) \\ \frac{\partial \pi_{0,j}}{\partial \xi_1}(\vec{\xi}_i) & \frac{\partial \pi_{1,j}}{\partial \xi_1}(\vec{\xi}_i) & \cdots & \frac{\partial \pi_{P,j}}{\partial \xi_1}(\vec{\xi}_i) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \pi_{0,j}}{\partial \xi_{n_\xi}}(\vec{\xi}_i) & \frac{\partial \pi_{1,j}}{\partial \xi_{n_\xi}}(\vec{\xi}_i) & \cdots & \frac{\partial \pi_{P,j}}{\partial \xi_{n_\xi}}(\vec{\xi}_i) \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{pmatrix} \vdots \\ \vec{u}^{(m,j)} \\ \vec{u}^{(m+1,j)} \\ \vdots \\ \vec{u}^{(m+n_\xi,j)} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \vec{u}_i \\ \frac{\partial \vec{u}_i}{\partial \xi_1} \\ \vdots \\ \frac{\partial \vec{u}_i}{\partial \xi_{n_\xi}} \\ \vdots \end{pmatrix}$$



- Compound efficiencies
  - Multilevel-Multifidelity with sampling & CS/FT surrogates (new: ROM, NN)
  - Active subspaces: subspace quadrature, enhance MF control variates



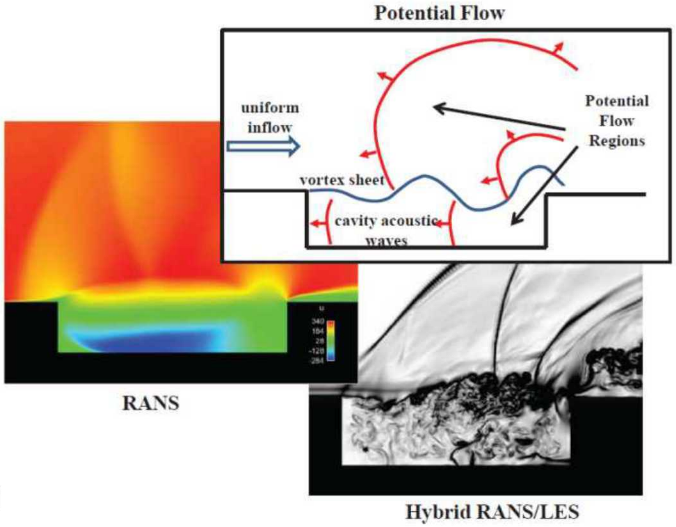
- Address complexity w/ component-based approach
  - Emulator-based Bayesian inference, Mixed aleatory-epistemic UQ, Optimization under uncertainty (new: Optimal experimental design)
- Position UQ for next generation architectures
  - *Current (imperative):* multilevel parallelism (MPI + local async)
  - *Future (declarative):* exploit DAG + AMT for ensemble workflows (w/ Stanford)

Multiple Model Forms in UQ & Opt

Discrete model choices for simulation of same physics

Sandia National Laboratories

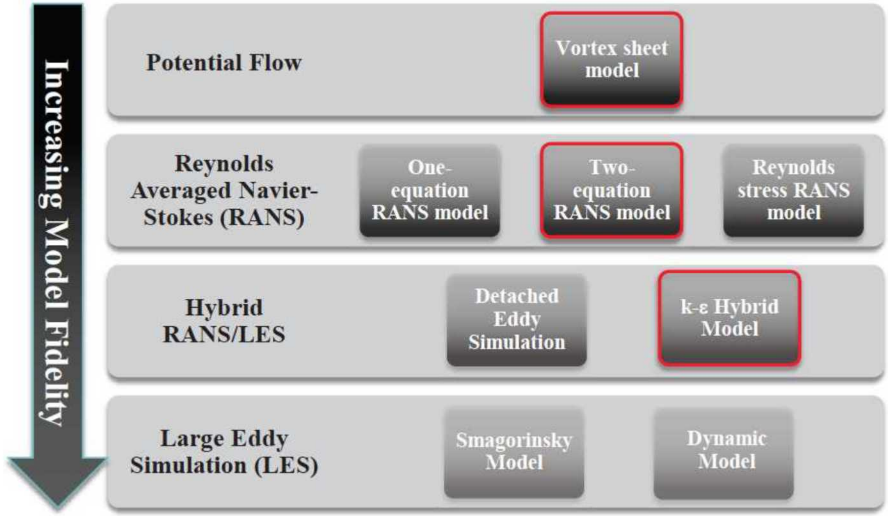
A clear hierarchy of fidelity (from low to high)
- Exploit less expensive models to render HF practical
  - *Multifidelity Opt, UQ, inference*
- Support general case of discrete model forms
  - Discrepancy does not go to 0 under refinement



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 → 0 at order of spatial/temporal convergence

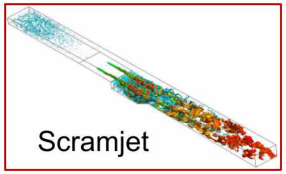
Combinations for multiphysics, multiscale

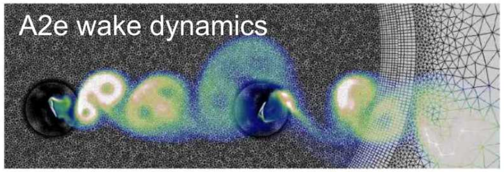


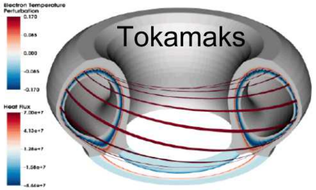
Research & Development in Multifidelity Methods

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



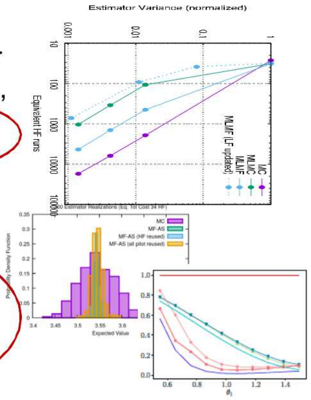




Emerging mission areas: abnormal thermal, Z-pinch MagLIF, quantum chemistry

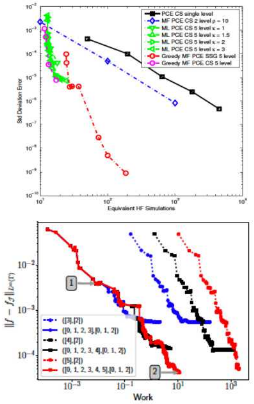
Monte Carlo UQ Methods

- *Production:* optimal resource allocation for multilevel, multifidelity, combined (DARPA x 2, Wind, Cardiovascular)
- *Emerging:* active dimensions ('18 EE LDRD), generalized fmwk for approx control variates (ASC V&V Methods)
- *On the horizon:* control of time avg; learning latent var relationships (CIS LDRD); model tuning / selection (CIS LDRD, DOE BES)



Surrogate UQ Methods (PCE, SC)

- *Production (v6.10):* ML PCE w/ projection & regression; ML SC w/ nodal/hierarchical interp; greedy ML adaptation (DARPA SEQUOIA)
- *Emerging:* multi-index stochastic collocation, multilevel function train (ASC V&V Methods)
- *On the horizon:* new surrogates (ROM, deep NN) with error mgmt ('19 EE LDRD, DOE BES); unification of surrogate + sampling approaches (CIS LDRD)



Optimization Under Uncertainty

- *Production:* manage simulation and/or stochastic fidelity
- *Emerging:*
  Derivative-based methods (DARPA SEQUOIA)
  - Multigrid optimization (MG/Opt)
  - Recursive trust-region model mgmt.: extend TRMM to deep hierarchies

  Derivative-free methods (DARPA ScramjetUQ)
  - SNOWPAC (w/ MIT, TUM) w/ MLMC error estimates
- *On the horizon:* Gaussian process-based approaches: multifidelity EGO (FASTMath OUU); Optimal experimental design (OED) (A2e)

Simple demonstration of key ML-MF concepts
Monte Carlo Sampling: MSE for mean estimator

Problem statement: We are interested in the expected value of $Q_M = \mathcal{G}(\mathbf{X}_M)$ 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:

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

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:

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

Accurate estimation $\Rightarrow$ **Large number** of samples at high (spatial) resolution

Simple demonstration of key ML-MF concepts
Multilevel MC: decomposition of estimator variance

Multilevel MC: Sampling from several approximations Q_M of Q (Multigrid...)

Ingredients:

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

$$Y_\ell \stackrel{\text{def}}{=} Q_{M_\ell} - Q_{M_{\ell-1}} \xrightarrow{\textit{linearity}} \mathbb{E}[Q_M] = \mathbb{E}[Q_{M_0}] + \sum_{\ell=1}^{L} \mathbb{E}\left[Q_{M_\ell} - Q_{M_{\ell-1}}\right] = \sum_{\ell=0}^{L} \mathbb{E}[Y_\ell]$$

- Multilevel Monte Carlo estimator

$$\hat{Q}_M^{\mathrm{ML}} \stackrel{\text{def}}{=} \sum_{\ell=0}^{L} \hat{Y}_{\ell,N_\ell}^{\mathrm{MC}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(Q_{M_\ell}^{(i)} - Q_{M_{\ell-1}}^{(i)} \right)$$

- The Mean Square Error is

$$\mathbb{E}\left[(\hat{Q}_M^{ML} - \mathbb{E}[Q])^2\right] = \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}ar(Y_\ell) + (\mathbb{E}[Q_M - Q])^2$$

Note If $Q_M \to Q$ (in a mean square sense), then $\mathbb{V}ar(Y_\ell) \xrightarrow{\ell \to \infty} 0$

Simple demonstration of key ML-MF concepts
Multilevel MC: optimal resource allocation

Let us consider the numerical cost of the estimator

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

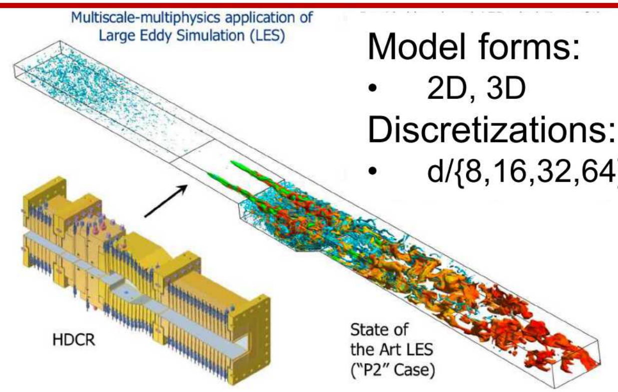
Determining the ideal number of samples per level (i.e. minimum cost at fixed variance)

$$\underbrace{\left.\begin{aligned} \mathcal{C}(\hat{Q}_M^{ML}) &= \sum_{\ell=0}^{L} N_\ell \mathcal{C}_\ell \\ \sum_{\ell=0}^{L} N_\ell^{-1} \mathbb{V}ar\left(Y_\ell\right) &= \varepsilon^2/2 \end{aligned}\right\}}_{\text{Balance ML estimator variance (stochastic error) and residual bias (deterministic error)}} \xrightarrow{\textit{Lagrange multiplier}} \underbrace{\boxed{N_\ell = \underbrace{\frac{2}{\varepsilon^2}\left[\sum_{k=0}^{L} \left(\mathbb{V}ar\left(Y_k\right)\mathcal{C}_k\right)^{1/2}\right]}_{\text{Level independent}} \underbrace{\sqrt{\frac{\mathbb{V}ar\left(Y_\ell\right)}{\mathcal{C}_\ell}}}_{\text{Level dependent}}}}_{\text{Optimal sample profile}}$$

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

Deployment Vignettes: ML, MF, MLMF Monte Carlo

Scramjet



Model forms:
- 2D, 3D

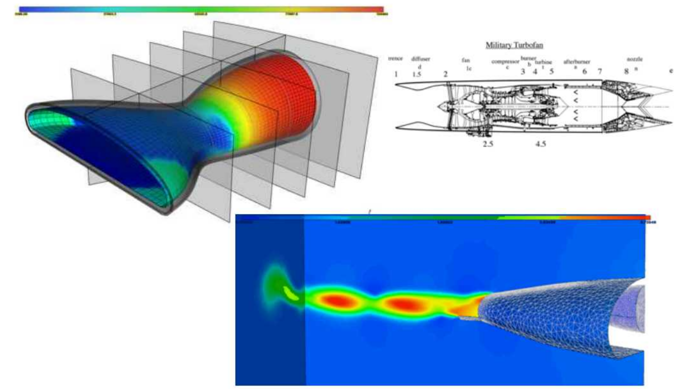
Discretizations:
- d/{8,16,32,64}

	$P_{0,mean}$	$P_{0,rms,mean}$	M_{mean}	TKE_{mean}	χ_{mean}
	P1				
$d/8$	4.02554e-03	1.90524e-06	1.99236e-02	3.34905e-07	4.24520e-03
$d/16$	4.03350e-07	7.77838e-08	6.68974e-05	1.74847e-08	4.40048e-05
	P1 updated				
$d/8$	4.05795e-03	1.90612e-06	1.60029e-02	7.53353e-07	9.41403e-04
$d/16$	2.85017e-04	7.36978e-07	2.07638e-03	2.99744e-07	2.57399e-02

Variance for the five QoIs of the P1 unit problem.

No variance decay for higher turbulence levels

UCAV Nozzle

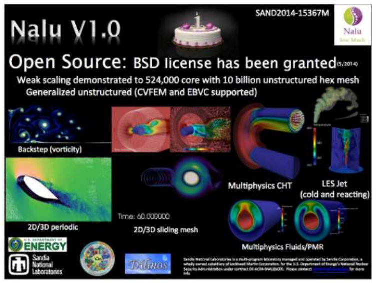


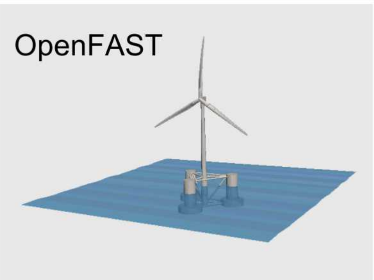
Non-predictive LF stress prior to reformulation

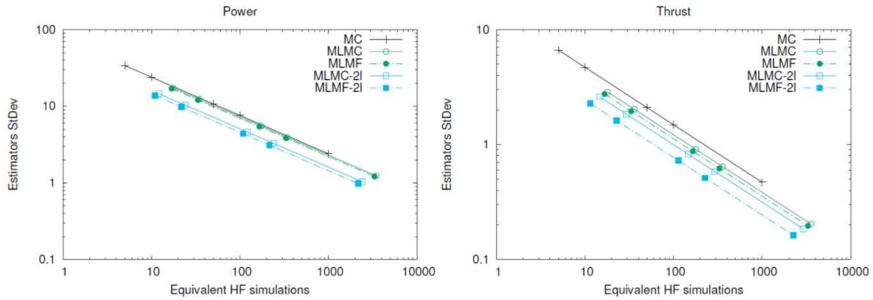
	LF correlation	LF Variance reduction [%]	LF (updated) correlation	LF (updated) Variance reduction [%]
Thrust	0.997	91.42	0.996	94.2
Mechanical Stress	2.31e-5	2.12e-3	0.944	89.2
Thermal Stress	0.391	12.81	0.987	93.4

Correlations and variance reduction for $\varepsilon^2/\varepsilon_0^2 = 0.001$.

Wind

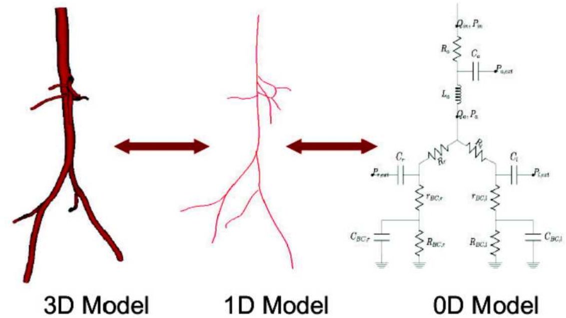


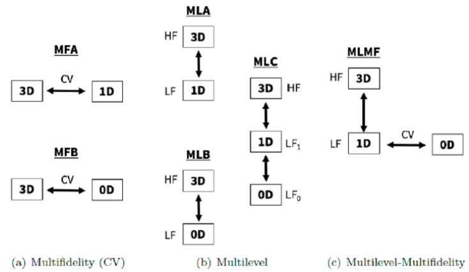




Nalu LES for Q0 is too coarse with limited predictive value

Cardiovascular





Method	Effective Cost (3D Simulations)	No. 3D Simulations	No. 1D Simulations	No. 0D Simulations
MC	9 885	9 885	–	–
MFA	56	21	15 681	–
MFB	39	36	–	154 880
MLA	305	212	41 990	–
MLB	156	150	–	342 060
MLC	165	156	1 324	351 940
MLMF	165	156	1 249	362 590

0D has greater predictive value, for which MF outperforms ML

Multilevel – Multifidelity Challenges

Key Challenge: existing ML/MF/MI performance is compelling on (elliptic) model problems, but significant generalization required for engineering apps. with non-trivial model relationships

- How well do we know the predictive value of each model a priori?
- Are the dependency relationships clear from the modeling source?
- Are they known to be highly correlated such that, e.g., control weights are not required?
- Conversely, can there be a penalty in greater generality with more weights to estimate, indicating the need for graph discovery?

Research directions:

- *Generalize:* start from a fully-connected, weighted structure
  - Compute correlations across full model ensemble
- *Optimize:* learn latent relationships for an optimized graph representation
  - Estimate reduced weight set from finite simulation instances

Multilevel – Multifidelity Sampling Methods

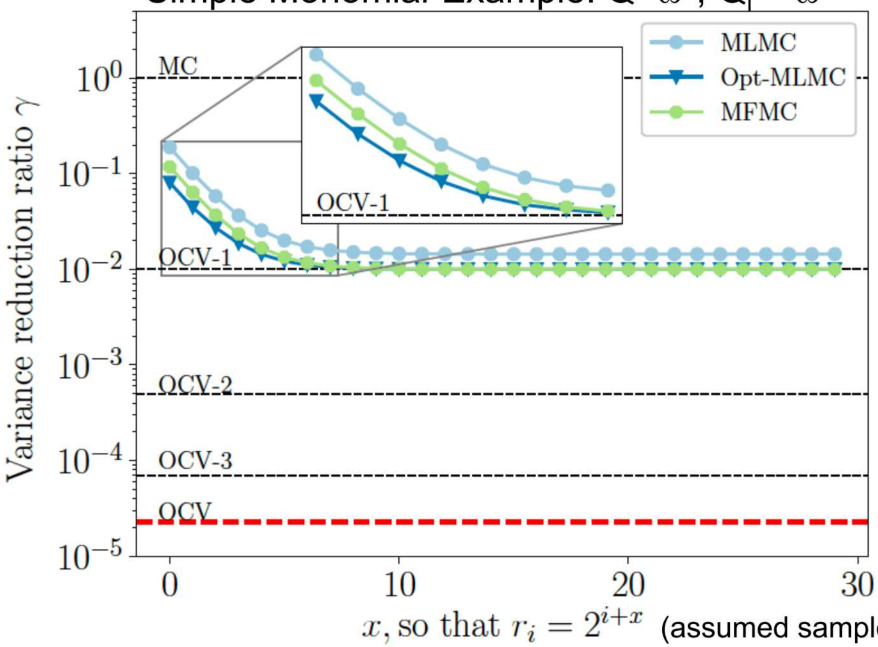
Generalized framework for approx. control variates

For M approximate models, look beyond (recursive) model pairings

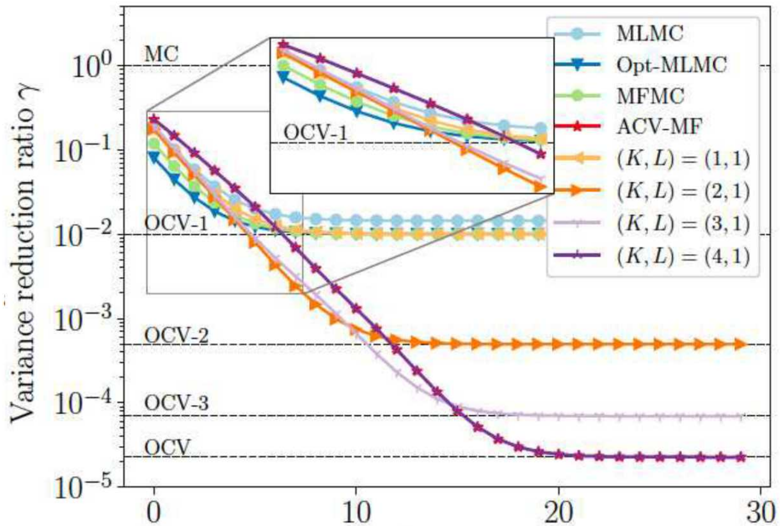
$$\hat{Q}^{\mathrm{CV}} = \hat{Q} + \sum_{i=1}^{M} \alpha_i \left(\hat{Q}_i - \mu_i \right)$$

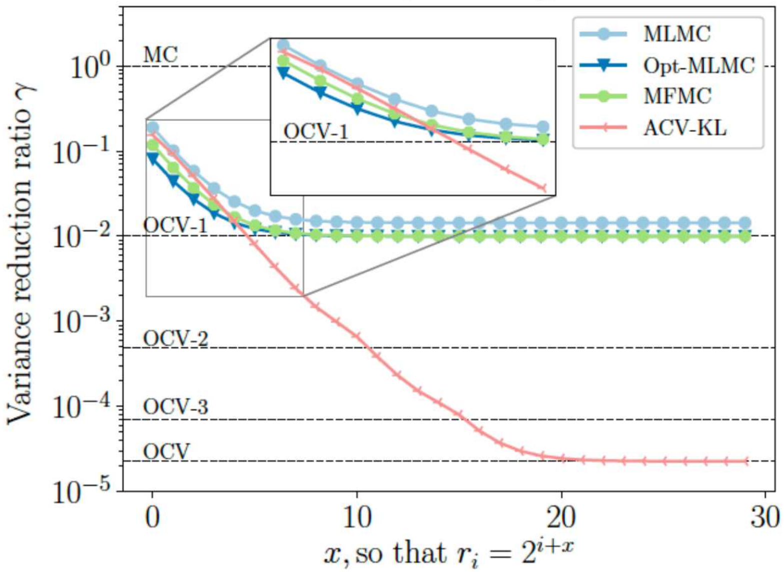
$$\underset{\underline{\alpha}}{\arg\min}\, \mathbb{V}\mathrm{ar}\left[\hat{Q}^{\mathrm{CV}}(\underline{\alpha}) \right] \Rightarrow \begin{cases} \boldsymbol{C} \in \mathbb{R}^{M \times M} & \text{covariance matrix among } Q_i \\ \boldsymbol{c} \in \mathbb{R}^{M} & \text{vector of covariances between } Q \\ \underline{\alpha}^* = \boldsymbol{C}^{-1}\boldsymbol{c} & \end{cases}$$

Simple Monomial Example: $Q=\omega^5$, $Q_i = \omega^{5-i}$



(assumed sample ratio)





A. Gorodetsky, G. Geraci, E., J. Jakeman "A generalized approximate control variate framework for multifidelity uncertainty quantification," arxiv.org/abs/1811.04988v3

Multilevel – Multifidelity Sampling Methods

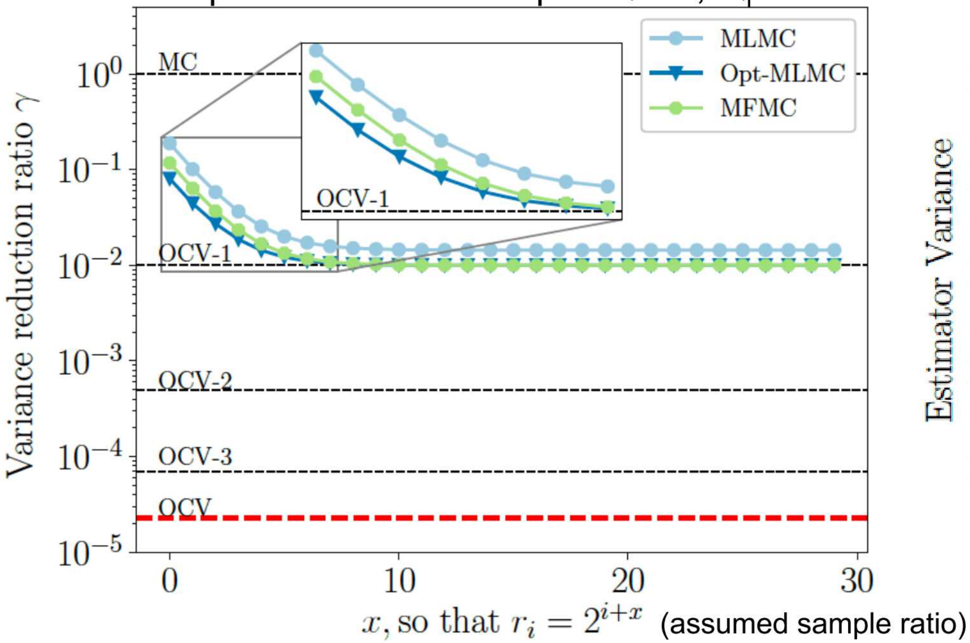
Generalized framework for approx. control variates

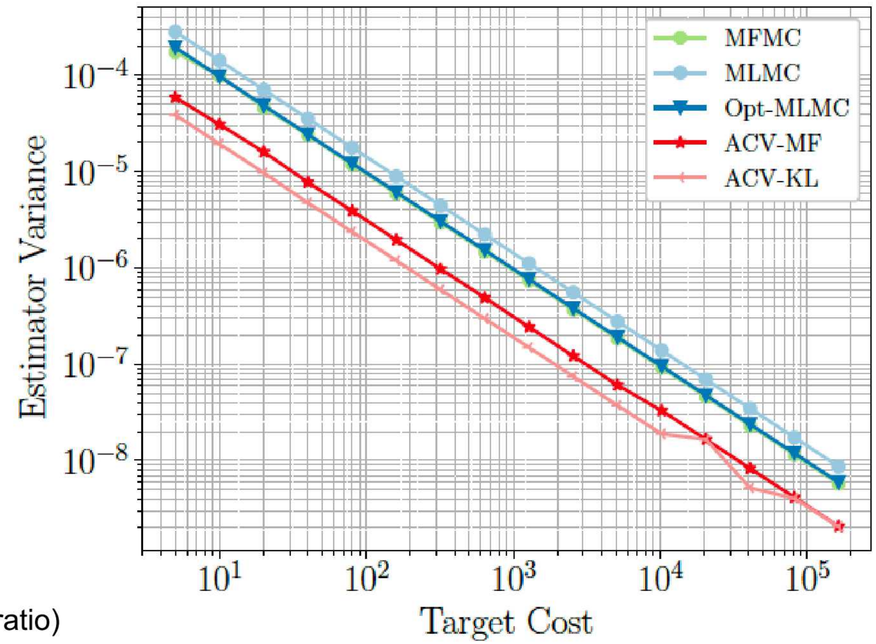
For M approximate models, look beyond (recursive) model pairings

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

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

Simple Monomial Example: $Q=\omega^5$, $Q_i = \omega^{5-i}$





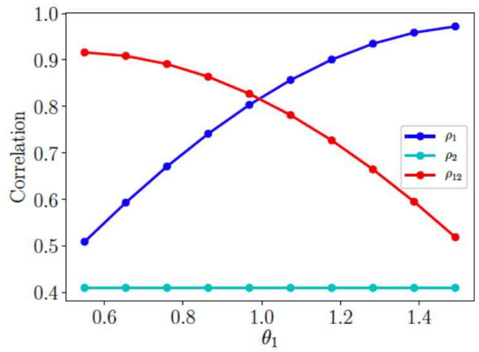
A. Gorodetsky, G. Geraci, E., J. Jakeman "A generalized approximate control variate framework for multifidelity uncertainty quantification," arxiv.org/abs/1811.04988v3

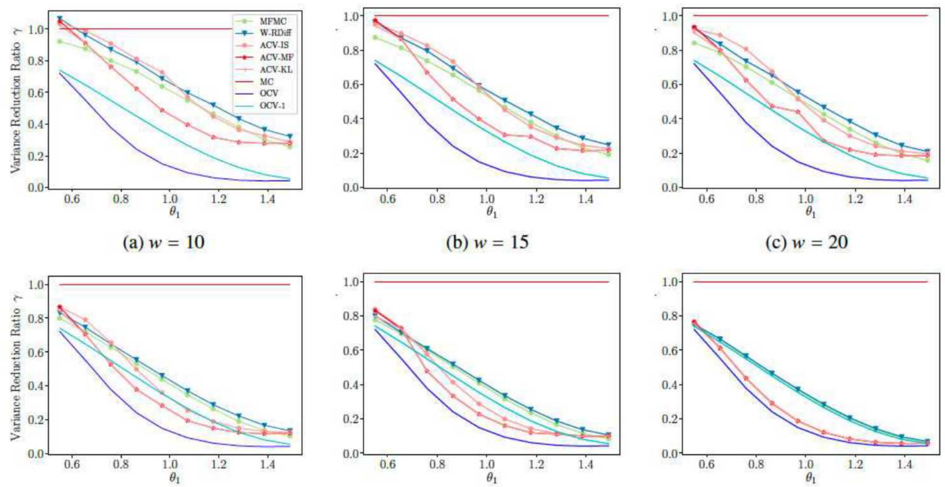
Generalized framework for approx. control variates

Tunable model problem

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

$A = \sqrt{11}$, $A_1 = \sqrt{7}$ and $A_2 = \sqrt{3}$ $\qquad \theta = \pi/2$, $\theta_2 = \pi/6$ and $\theta_2 < \theta_1 < \theta$.



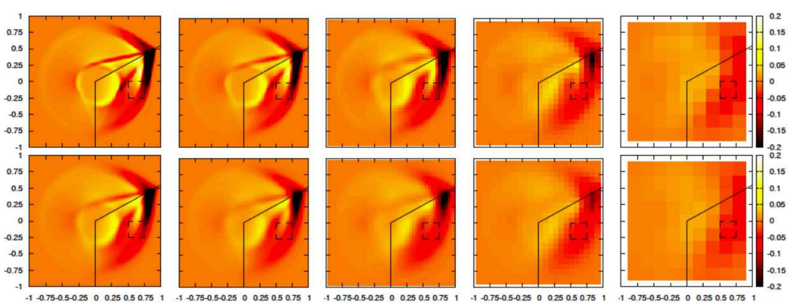


(a) $w = 10$ (b) $w = 15$ (c) $w = 20$

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

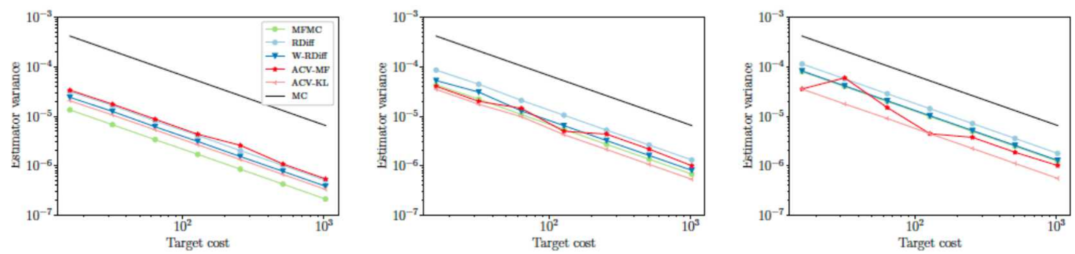
Exploiting the OCV vs. OCV-1 gap for different cost ratios

Elastic wave propagation in 2D



Shear stress contours for 2 fidelities x 5 discretizations

Well correlated discretization hierarchy → General multifidelity with model gaps



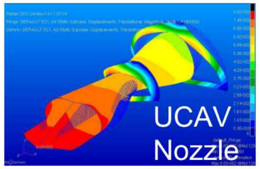
Perfomance of ACV-KL is robust to reduced model predictivity

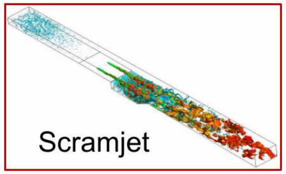
Research & Development in Multifidelity Methods

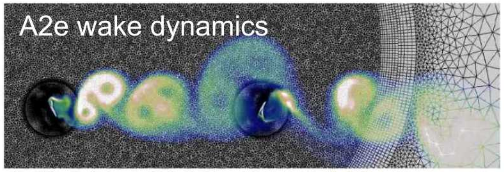
Sandia National Laboratories

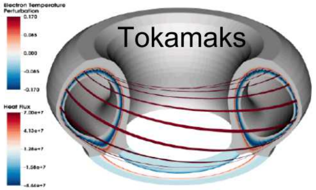
Recurring R&D theme: couple scalable algorithms with exploiting a (multi-dimensional) model hierarchy

- ➢ 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







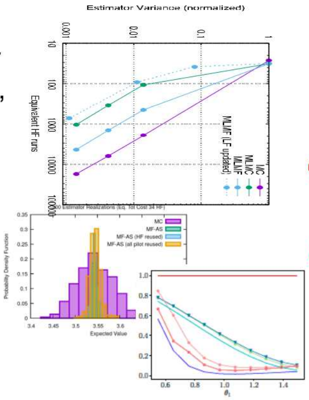


SECURE GC

Emerging mission areas: abnormal thermal, Z-pinch MagLIF, quantum chemistry

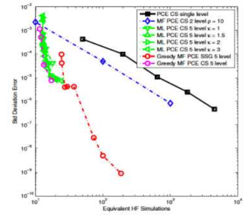
Monte Carlo UQ Methods

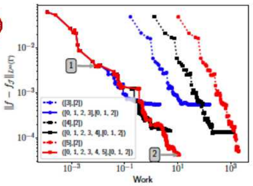
- *Production:* optimal resource allocation for multilevel, multifidelity, combined (DARPA SEQUOIA/ScramjetUQ)
- *Emerging:* active dimensions ('18 EE LDRD), generalized fmwk for approx control variates (ASC V&V Methods)
- *On the horizon:* control of time avg; learning latent var relationships (CIS LDRD); model tuning / selection (CIS LDRD, DOE BES)



Surrogate UQ Methods (PCE, SC)

- *Production (v6.10):* ML PCE w/ projection & regression; ML SC w/ nodal/hierarchical interp; greedy ML adaptation (DARPA SEQUOIA)
- *Emerging:* multi-index stochastic collocation, multilevel function train (ASC V&V Methods)
- *On the horizon:* new surrogates (ROM, deep NN) with error mgmt ('19 EE LDRD, DOE BES); unification of surrogate + sampling approaches (CIS LDRD)





Optimization Under Uncertainty

- *Production:* manage simulation and/or stochastic fidelity

- *Emerging:*
  Derivative-based methods (DARPA SEQUOIA)
  - Multigrid optimization (MG/Opt)
  - Recursive trust-region model mgmt.: extend TRMM to deep hierarchies

  Derivative-free methods (DARPA ScramjetUQ)
  - SNOWPAC (w/ MIT, TUM) w/ MLMC error estimates
- *On the horizon:* Gaussian process-based approaches: multifidelity EGO (FASTMath OUU)

Stochastic Expansions: Polynomial Chaos & Stochastic Collocation

Polynomial chaos: spectral projection using orthogonal polynomial basis fns

$$\boxed{R = \sum_{j=0}^{\infty} \alpha_j \Psi_j(\boldsymbol{\xi})}$$

using

$$
\begin{array}{lcccl}
\Psi_0(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_0(\xi_2) & = & 1 \\
\Psi_1(\boldsymbol{\xi}) & = & \psi_1(\xi_1)\,\psi_0(\xi_2) & = & \xi_1 \\
\Psi_2(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_1(\xi_2) & = & \xi_2 \\
\Psi_3(\boldsymbol{\xi}) & = & \psi_2(\xi_1)\,\psi_0(\xi_2) & = & \xi_1^2 - 1 \\
\Psi_4(\boldsymbol{\xi}) & = & \psi_1(\xi_1)\,\psi_1(\xi_2) & = & \xi_1 \xi_2 \\
\Psi_5(\boldsymbol{\xi}) & = & \psi_0(\xi_1)\,\psi_2(\xi_2) & = & \xi_2^2 - 1
\end{array}
$$

Distribution	Density function	Polynomial	Weight function	Support range
Normal	$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$	Hermite $He_n(x)$	$e^{\frac{-x^2}{2}}$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(x)$	1	$[-1, 1]$
Beta	$\frac{(1-x)^\alpha (1+x)^\beta}{2^{\alpha+\beta+1} B(\alpha+1,\beta+1)}$	Jacobi $P_n^{(\alpha,\beta)}(x)$	$(1-x)^\alpha (1+x)^\beta$	$[-1, 1]$
Exponential	e^{-x}	Laguerre $L_n(x)$	e^{-x}	$[0, \infty]$
Gamma	$\frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$	Generalized Laguerre $L_n^{(\alpha)}(x)$	$x^\alpha e^{-x}$	$[0, \infty]$

- Estimate α_j using regression or numerical integration: sampling, tensor quadrature, sparse grids, or cubature

$$\boxed{\alpha_j = \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} = \frac{1}{\langle \Psi_j^2 \rangle} \int_\Omega R\, \Psi_j\, \varrho(\boldsymbol{\xi})\, d\boldsymbol{\xi}}$$

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

$$\boxed{R(\boldsymbol{\xi}) \cong \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi})}$$

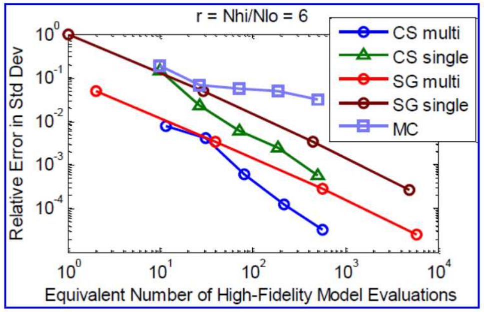
- ***Global:*** Lagrange (values) or Hermite (values+derivatives)
- ***Local:*** linear (values) or cubic (values+gradients) splines
- ***Nodal*** or ***Hierarchical*** interpolants

$$\boxed{L_j = \prod_{\substack{k=1 \\ k \neq j}}^{m} \frac{\xi - \xi_k}{\xi_j - \xi_k}} \Rightarrow \boxed{R(\boldsymbol{\xi}) \cong \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} r\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right) \left(L_{j_1}^{i_1} \otimes \cdots \otimes L_{j_n}^{i_n}\right)}$$

Sparse interpolants formed using Σ 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

Formulations for Multilevel PCE / SC



Starting point (2012): prescribed ML/MF resolutions w/ adaptivity

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

$$\boxed{N_{lo} >> N_{hi}}$$

Rate estimation

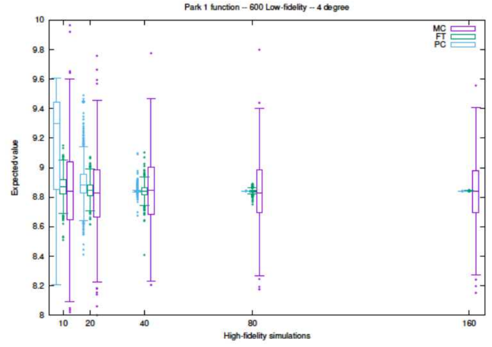
1. Optimal resource allocation: parameterize estimator variance → optimal N_l

Global κ and $\gamma > 0$

$$\boxed{Var[\hat{Y}_l] = \frac{Var[Y_l]}{\gamma N^{\kappa}} \quad \rightarrow \quad N_l = \sqrt[\kappa]{\frac{2}{\epsilon^2 \gamma} \sum_{q=0}^{L} \sqrt[\kappa+1]{Var[Y_q] C_q^{\kappa}}} \sqrt[\kappa+1]{\frac{Var[Y_l]}{C_l}}}$$

E., G. Geraci, J.D. Jakeman, "Multilevel Monte Carlo Hybrids Exploiting Multidelity Modeling and Sparse Polynomial Chaos Estimation," SIAM UQ 2016, Lausanne.

Main challenge: abrupt transitions in sparse / low rank recovery



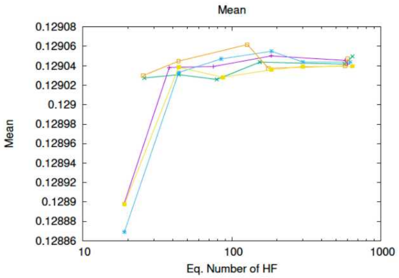
$N_{low} = 600$, **degree=4**

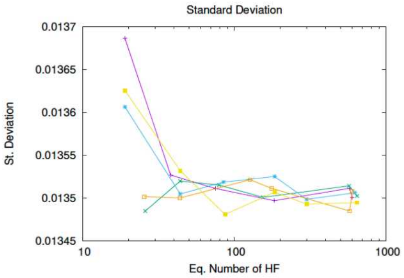
Recovery theory

2. Restricted Isometry Property (RIP) for sparse recovery (BLUE for OLS, FTT N_l scaling w/ rank)

$$\boxed{N_l \geq s_l \, log^3(s_l) \; L_l \, log(C_l)}$$

Jakeman, Narayan, and Zhou, 2016





Main challenge: compressible fns

→ increasing s

→ feedback not well controlled

Greedy

3. Greedy Multilevel refinement

ML competition with multiple level candidate generators

Main challenges: scalable refinement schemes, loss of precision

Surrogate approaches: Greedy multilevel refinement

$$\hat{Q}_L \approx \hat{Q}_0 + \sum_{l=1}^{L} \hat{\Delta}_L, \text{ for } \Delta_l \equiv Q_l - Q_{l-1}$$

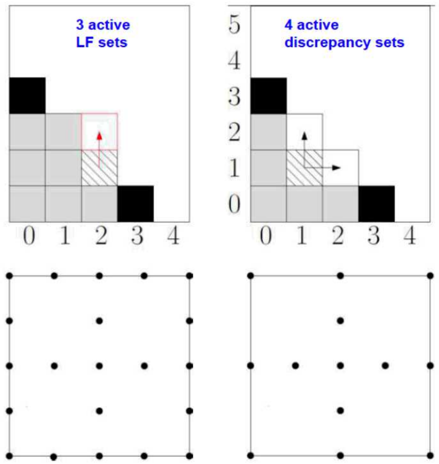
<u>Compete refinement candidates across model levels:</u> max induced change / cost

- 1 or more refinement candidates per model 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: $z/p/\beta/\beta^*$)

  normalized by relative cost of level increment (# new points * cost / point)
- Greedy selection of best candidate, which then generates new candidates for this model level

<u>Level candidate generators:</u>

- *Uniform refinement:* 1 exp order / grid level candidate per model level
  - Tensor / sparse grids: projection PCE, nodal/hierarchical SC
  - Regression PCE: least squares / compressed sensing
- *Anisotropic refinement:* 1 exp order / grid level candidate per model level
  - Tensor / sparse grids
- *Index-set refinement:* many candidates per level
  - Generalized sparse grids: projection PCE, nodal/hierarch SC
  - Regression PCE
- *Adapted candidate basis:* ~3 frontier advancements per model level
  - Regression PCE (Jakeman, E., Sargsyan, "Enhancing ℓ1-minimization estimates of polynomial chaos expansions using basis selection," *J. Comp. Phys.,* Vol. 289, May 2015.)



Multilevel / Multi-index PCE: greedy competition across models

Steady state diffusion

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

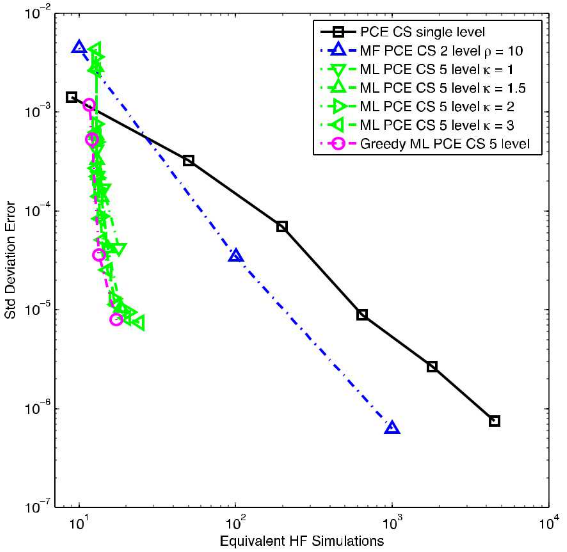
$$u(0,\boldsymbol{\xi}) = 0, \quad u(1,\boldsymbol{\xi}) = 0.$$

Advection diffusion

$$\frac{du}{dt}(x_1,t,Z) + a\frac{du}{dx}(x_1,t,Z) - \frac{d}{dx}\left[k(x_1,Z)\frac{du}{dx}(x_1,t,Z)\right] = g(x_1,t,Z) \qquad (x_1,Z)\in(0,1)\times\Gamma$$

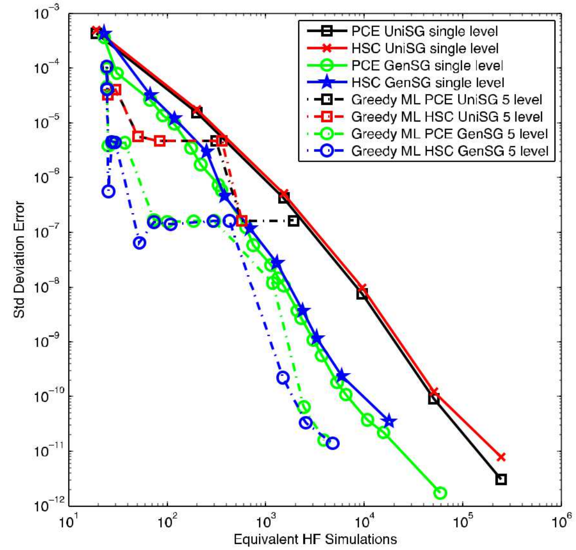
$$u(0,t,Z)=0 \quad u(1,t,Z)=0 \quad u(x_1,0,Z)=0,$$

Greedy ML PCE: compressed sensing with uniform candidate refinement



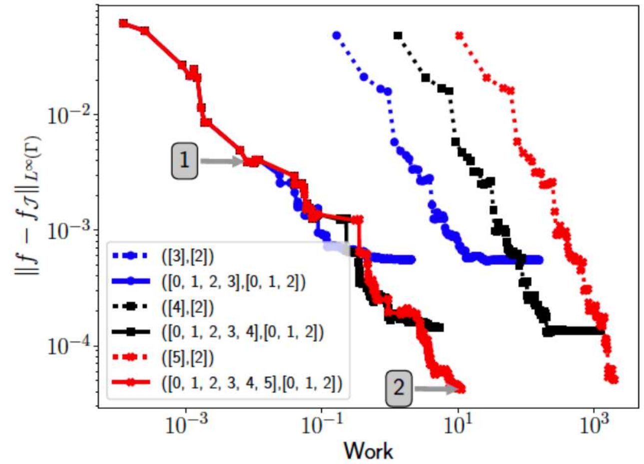
Conv Tol	N_1	N_2	N_3	N_4	N_5
1.e-1	198	9	9	9	9
1.e-2	644	198	9	9	9
1.e-3	1802	644	9	9	9
1.e-4	4505	1802	50	9	9

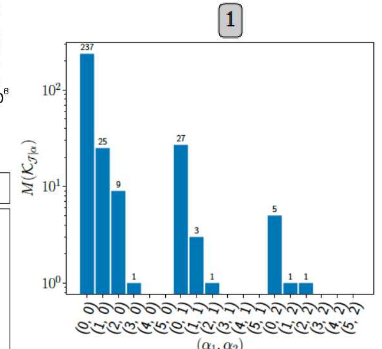
Greedy ML PCE: sparse grids with uniform / generalized refinement

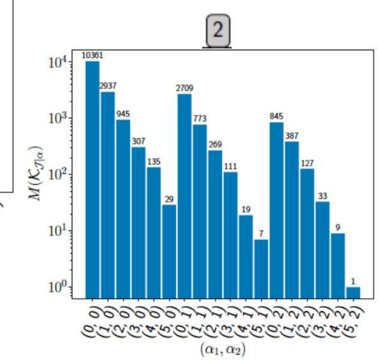


Conv Tol	N_1	N_2	N_3	N_4	N_5
1.e-2	43	23	19	19	19
1.e-4	211	83	19	19	19
1.e-6	391	271	156	19	19
1.e-8	1359	743	327	59	19
1.e-10	3535	2311	1039	391	19
1.e-12	10319	5783	2783	1343	43
1.e-14	26655	14991	8063	3703	1535

Greedy multi-index PCE: sparse grids with generalized refinement







Related ML-MF Topics (Time permitting)

Emulator-Based Bayesian inference

Trust-region Optimization / OUU

(ML-MF) Emulator-based Bayesian inference

MCMC sampling performed on emulator, leveraging differentiable emulator structure

- Pre-solve for MAP (maximum a posteriori probability) point: full Newton min of –log(posterior)
- Accurate MCMC proposal: emulator derivatives → Hessian of misfit → MVN proposal covariance
  - mitigates sample rejection in high D: for 10D Rosenbrock test, 98% rejection rate reduced to 30%

$$\boxed{p(\mathbf{d}|\xi) = \exp\left[-\frac{1}{2}(f(\xi)-\mathbf{d})^T\Gamma_{\mathbf{d}}^{-1}(f(\xi)-\mathbf{d})\right]}$$

Gaussian Likelihood

$$\boxed{-\log\left[p(\mathbf{d}|\xi)\right] = \frac{1}{2}(f(\xi)-\mathbf{d})^T\Gamma_{\mathbf{d}}^{-1}(f(\xi)-\mathbf{d}) = M(\xi)}$$

Negative Log Likelihood = Misfit

$$\boxed{\nabla_\xi^2 M(\xi) = \underbrace{\nabla_\xi f(\xi)^T\,\Gamma_{\mathbf{d}}^{-1}\,\nabla_\xi f(\xi)}_{\text{Gauss-Newton approx. Hessian (if only emulator grads)}} + \nabla_\xi^2 f(\xi)\cdot\left[\Gamma_{\mathbf{d}}^{-1}\,(f(\xi)-\mathbf{d})\right]}$$

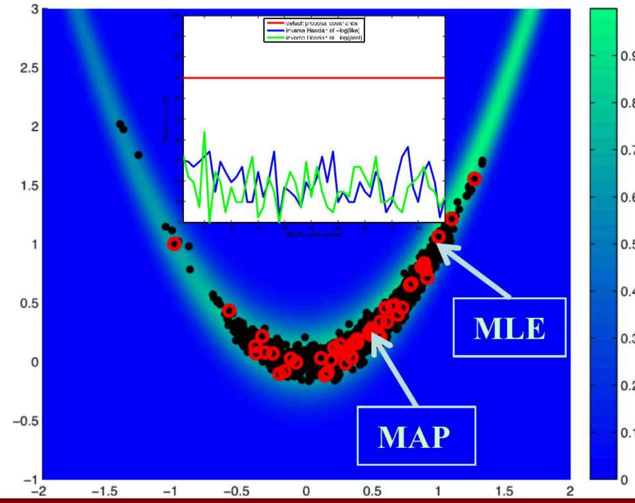
Hessian of Misfit

Laplace approx.: MVN proposal covariance defined by inverse Hessian of negative log posterior

$$\boxed{-\log\pi_d(\xi) = M(\xi) - \log\pi_0(\xi)}$$

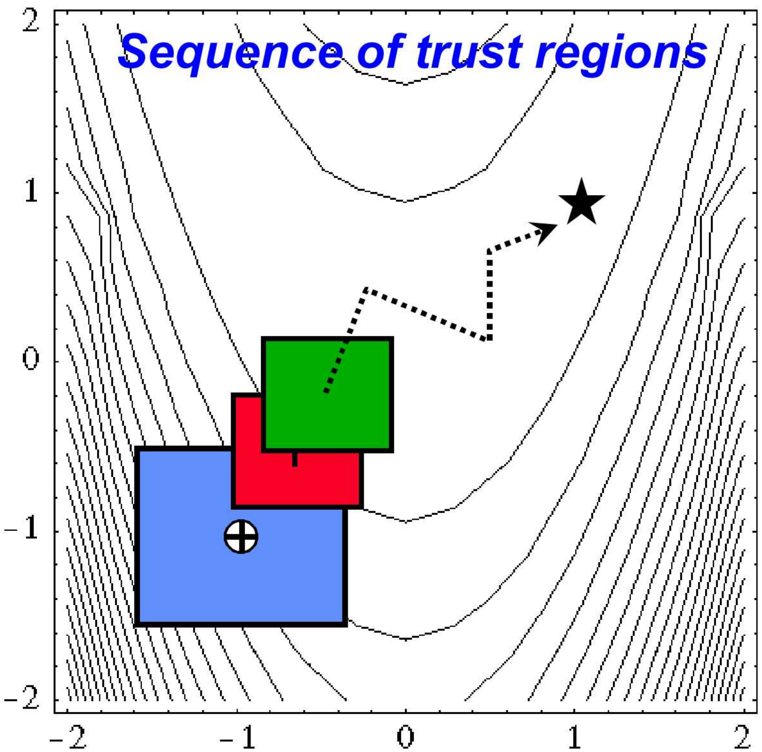
- augmenting misfit: Hessian of negative log prior provides regularization for priors w/ curvature (normal, beta, gamma)
- Posterior Hessian-based proposal balances likelihood and prior, performing better than either alone

Rosenbrock Problem; Prior ~ N(0,1)



Recursive Trust Region Model Management

Bi-fidelity opt. work in early 2000's



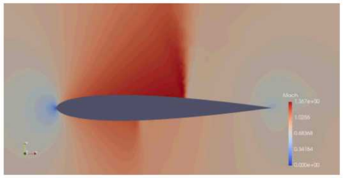
Extend to deep model hierarchies

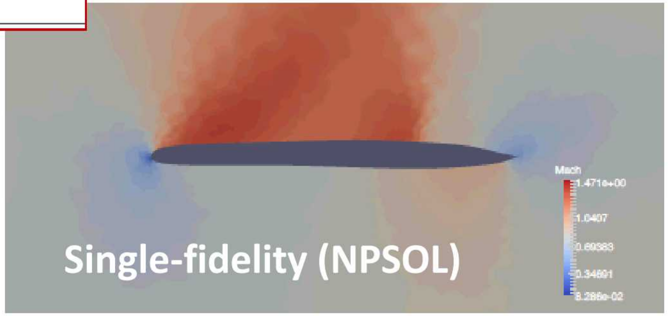
Algorithm 9 Recursive Trust Region Updating

procedure RECTR($r = 1$:LEN, x^r_c, x^r_*, $f^r_{\text{CORR}}(x)$)
  for r = len to 1 (bottom up: low to high || coarse to fine) **do**
    if State$_r$ = new candidate x^r_* **then**
      Test for new center: TR(x^r_c, x^r_*, $f^{r+1}_{\text{corr}}(x)$, $f^r_{\text{corr}}(x)$)
    end if
    if State$_r$ = new center x^r_c **then**
      Compute $f^{r-1}(x^r_c)$
      Compute CORRECTION($x^r_c, R, f^{r-1}(x^r_c), f^r(x^r_c)$)
      if Converged(x^r_c, $f^{r-1}_{\text{corr}}(x^r_c)$, L^{r-1}, U^{r-1}) **then**
        $x^{r-1}_* = x^r_c$ (new candidate)
      end if
    end if
  end for
  for r = 1 to len (top down: high to low || fine to coarse) **do**
    if State$_r$ = new center x^r_c **then**
      Recompute CORRECTION($x^r_c, R, f^{r-1}(x^r_c), f^r(x^r_c)$)
    end if
    if parent corrected **then**
      Recur updated corrections for $f^r_{\text{corr}}(x^r_c)$
    end if
    Reset State$_r$
  end for
end procedure

Transonic Airfoil Design

$$\underset{x}{\text{minimize}} \quad C_D(u(x))$$
$$\text{subject to} \quad C_L(u(x)) \geq \bar{C}_L$$
$$c_{\text{Euler}}(x, u(x)) = 0$$
$$-0.01 \leq x \leq 0.01$$





3-level Recursive TRMM for Euler

	Coarse evals	Medium evals	Fine evals	C_D	C_L
Reference NACA 0012	–	–	–	0.10345	0.80118
Single-fidelity SQP	–	–	806	0.064904	0.80118
Single-fidelity SLP	–	–	1190	0.065024	0.80199
Single-fidelity NIP BFGS	–	–	2294	0.066894	0.80119
Three-level 1^{st}-order TRMM	43630	4882	187	0.064968	0.80153
Three-level BFGS 2^{nd}-order TRMM	17020	248	93	0.068695	0.81966

- Order of magnitude fewer HF runs
- More aggressive profile shaping than MG/Opt

Summary Remarks

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, wind) → rich model ensemble, challenging QoI

Towards multilevel-multifidelity UQ tailored for smoothness and dimensionality

- Generalized MC framework for approximate control variates
  - Beyond assumed model relationships → CV wts for fully connected graphs → optimized LVN
  - Leverage dissimilar models (with dissimilar parameterizations) → discover shared physics
  - Well suited for high dimensionality and/or low regularity
- Multilevel PCE/SC:
  - Extend ML MC machinery with higher perf. estimators (sparsity, low rank, hierarchical stats), while extending previous heuristics (~2012) to include optimal allocations
    - *Rate estimation:* *Recovery theory* *Greedy refinement*
  - Achieve more rapid convergence (sufficient regularity, moderate dimensionality)
  - Enables emulator-based / derivative-enhanced approaches to inference, OUU, OED

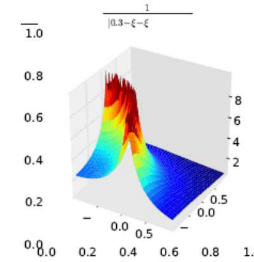
Vision:

- Robustly span range of application scenarios / mission needs
  - General framework with range of estimators, supporting different model features & analysis goals
  - Automatic discovery and exploitation of model relationships

Extra

Emphasis on Scalable Methods for High-fidelity UQ on HPC

Sandia National Laboratories

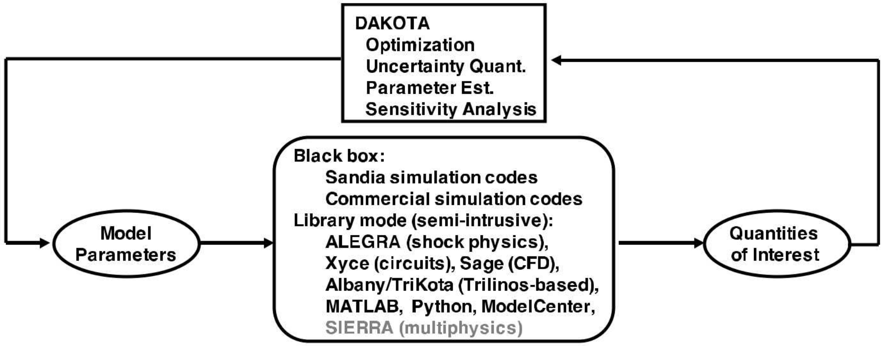


Compounding effects:

- Mixed aleatory-epistemic uncertainties (segregation → nested iteration)
- Requirement to evaluate probability of rare events (resolve PDF tails for QoI)
- Nonsmooth QoI (exp conv in spectral methods exploits smoothness)

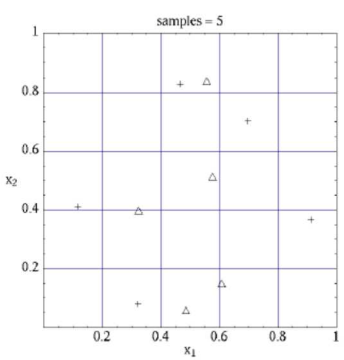
Steward Scalable Algorithms within DAKOTA

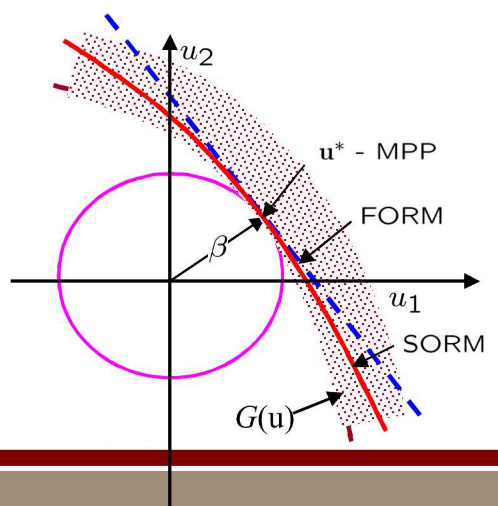
Explore and predict with confidence

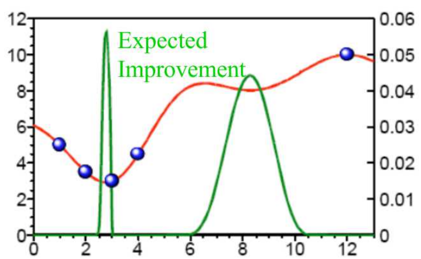


Core (Forward) UQ Capabilities:

- Sampling methods: MC, LHS, QMC, et al.
- Reliability methods: local (MV, AMV+, FORM, …), global (EGRA, GPAIS, POFDarts)
- Stochastic expansion methods: PCE, SC, fn train
- Epistemic methods: interval est., Dempster-Shafer evidence

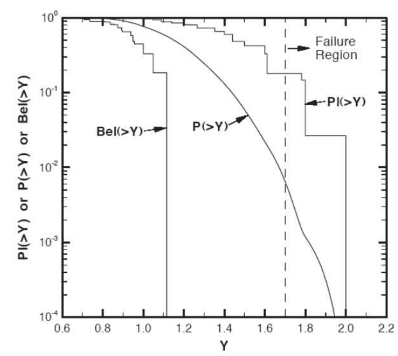






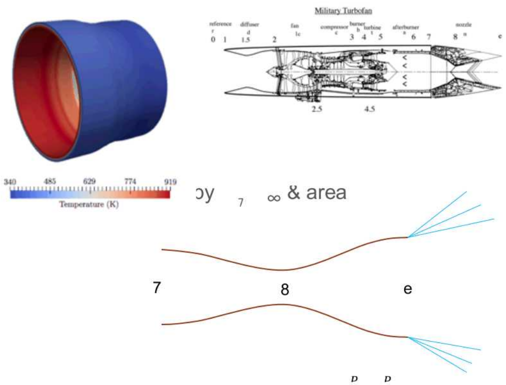
$$R = \sum_{j=0}^{\infty} \alpha_j \Psi_j(\boldsymbol{\xi})$$

$$R(\boldsymbol{\xi}) \cong \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi})$$

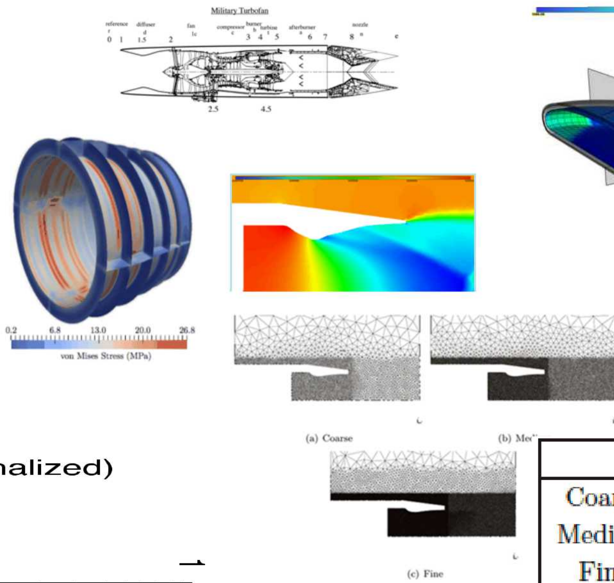


Initial Deployment of MLCV MC to UCAV Nozzle UQ

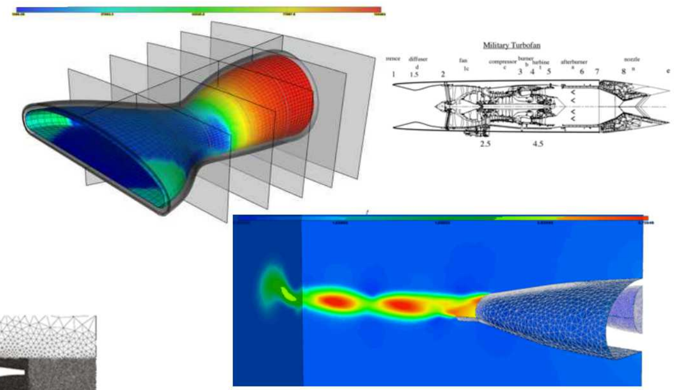
Low fidelity model: ~1D

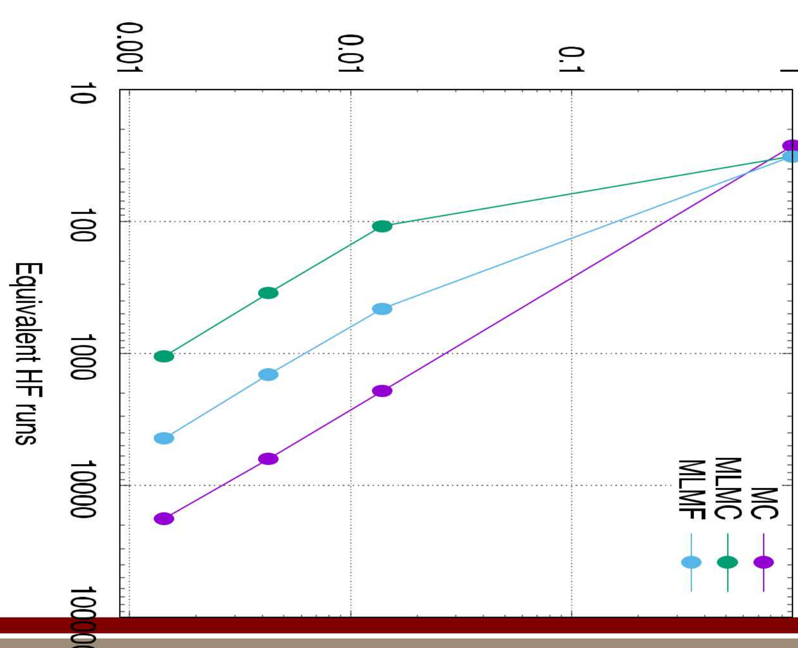


Medium fidelity model: ~2D



High fidelity model: ~3D





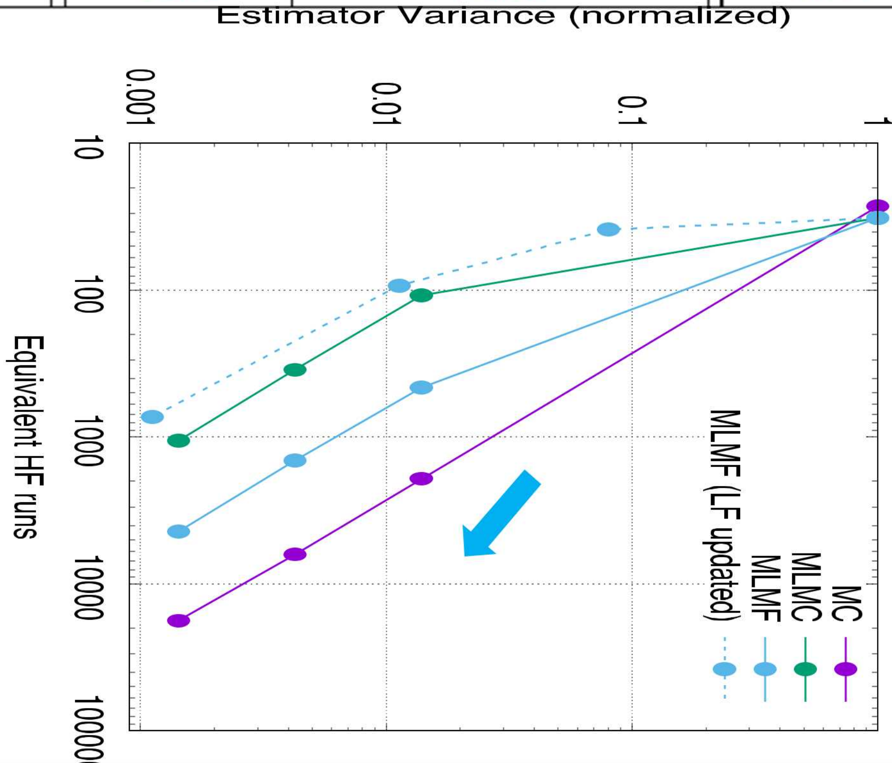
	LF	MF
Coarse	0.016	0.053
Medium	N/A	0.253
Fine	N/A	1.0

Optimal sample allocations for MLMF

Target accuracy	LF	MF		
	Coarse	Coarse	Medium	Fine
0.01	21143	1757	20	20
0.003	69580	5775	36	20
0.001	212828	17715	109	34

Updated Deployment of MLCV MC to UCAV Nozzle UQ

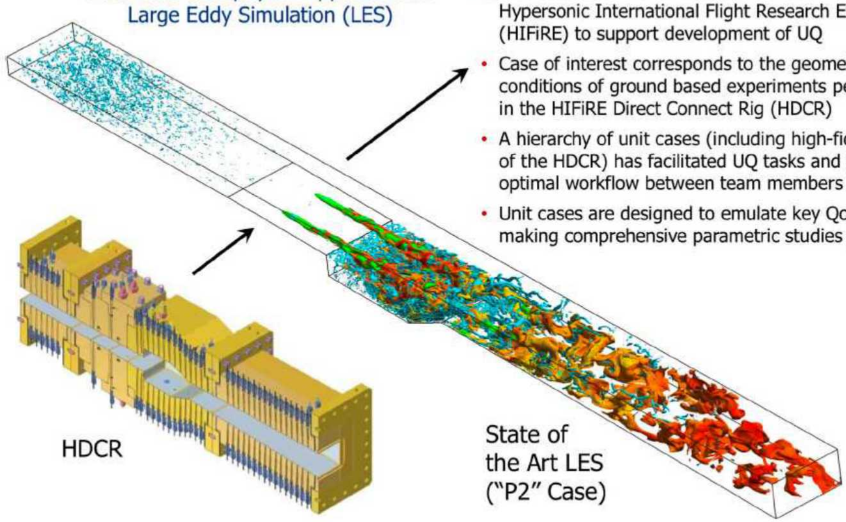
	LF		LF (updated)	
	correlation	Variance reduction [%]	correlation	Variance reduction [%]
Thrust	0.997	91.42	0.996	94.2
Mechanical Stress	2.31e-5	2.12e-3	0.944	89.2
Thermal Stress	0.391	12.81	0.987	93.4



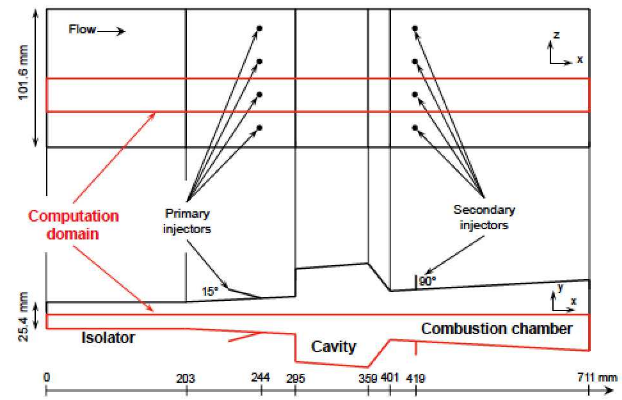
Accuracy ($\varepsilon^2/\varepsilon_0^2$)	LF	Medium Fidelity			LF (updated)	Medium Fidelity		
	Coarse	Coarse	Medium	Fine	Coarse	Coarse	Medium	Fine
0.1	N/A	N/A	N/A	N/A	404	20	20	20
0.01	21,143	1,757	20	20	3,091	177	31	20
0.003	69,580	5,775	36	20	N/A	N/A	N/A	N/A
0.001	212,828	17,715	109	34	32,433	1,773	314	20

Initial Deployment of MLCV MC for Scramjet UQ

Multiscale-multiphysics application of Large Eddy Simulation (LES)



- 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

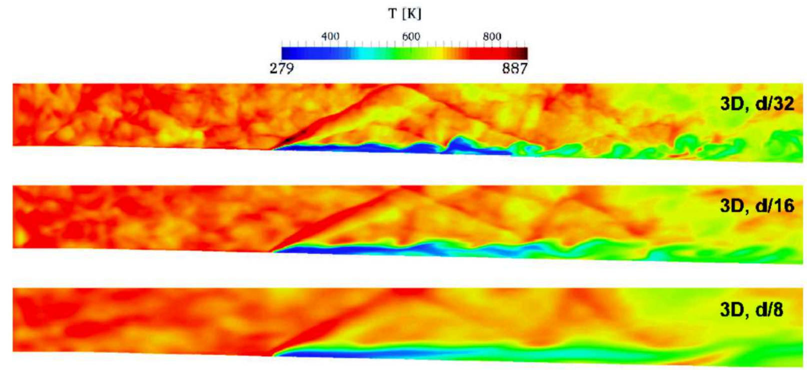


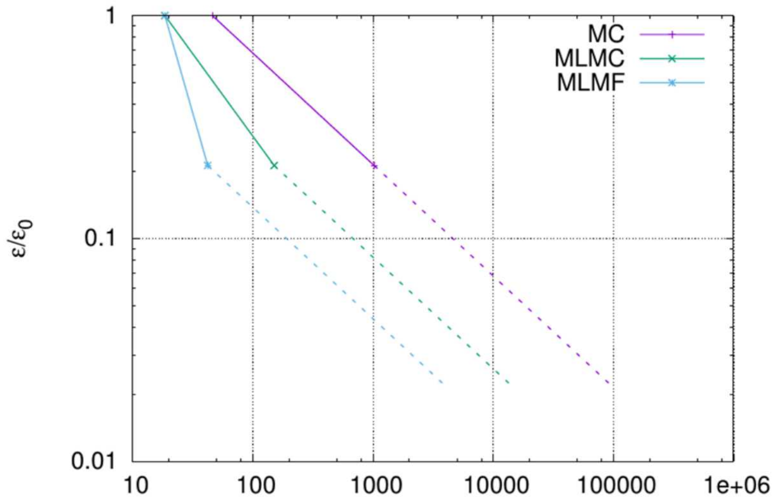
Model forms:
- 2D, 3D

Discretizations:
- d/{8,16,32,64}

	2D	3D
$d/8$	5E-4	0.11
$d/16$	0.014	1

TABLE: Computational cost.





	2D	3D
$d/8$	4,191	263
$d/16$	68	9

Optimal sample allocations for MLMF (MSE target = .045 of pilot MSE)

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

	$P_{0,mean}$	$P_{0,rms,mean}$	M_{mean}	TKE_{mean}	χ_{mean}
	P1				
$d/8$	4.02554e-03	1.90524e-06	1.99236e-02	3.34905e-07	4.24520e-03
$d/16$	4.03350e-07	7.77838e-08	6.68974e-05	1.74847e-08	4.40048e-05
	P1 updated				
$d/8$	4.05795e-03	1.90612e-06	1.60029e-02	7.53353e-07	9.41403e-04
$d/16$	2.85017e-04	7.36978e-07	2.07638e-03	2.99744e-07	2.57399e-02

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

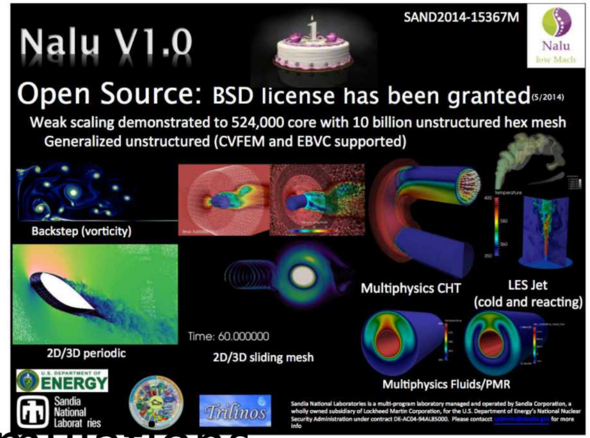
Observations from pilot sample: decay in variance across discretizations (LF d/8 and discrepancy d/16 – d/8) no longer observed for all QoI

- Need to engage additional refinement levels (i.e., d/32, d/64) to converge QoI statistics tied to resolution of turbulence
- Need robust MLMF strategies that *handle non-monotonicity while outperforming MC*

Computational Approach

- Low Fidelity: **OpenFAST-AeroDyn-Turbsim** (https://github.com/OpenFAST)
  - Turbsim generates turbulent atmospheric boundary layer flow field, semi-empirical
  - AereoDyn models the aerodynamic forces on the rotor
  - OpenFAST models the structural and controls response of the rotor (same for Nalu)

- High Fidelity: **Nalu** (https://github.com/NaluCFD)
  - LES, Solves the Navier-Stokes equations in the low-Mach number approximation with the one-equation, constant coefficient, TKE model for SGS, unstructured massively parallel.
  - Actuator Line model of the rotor
  - Single, uniform mesh (no nesting)



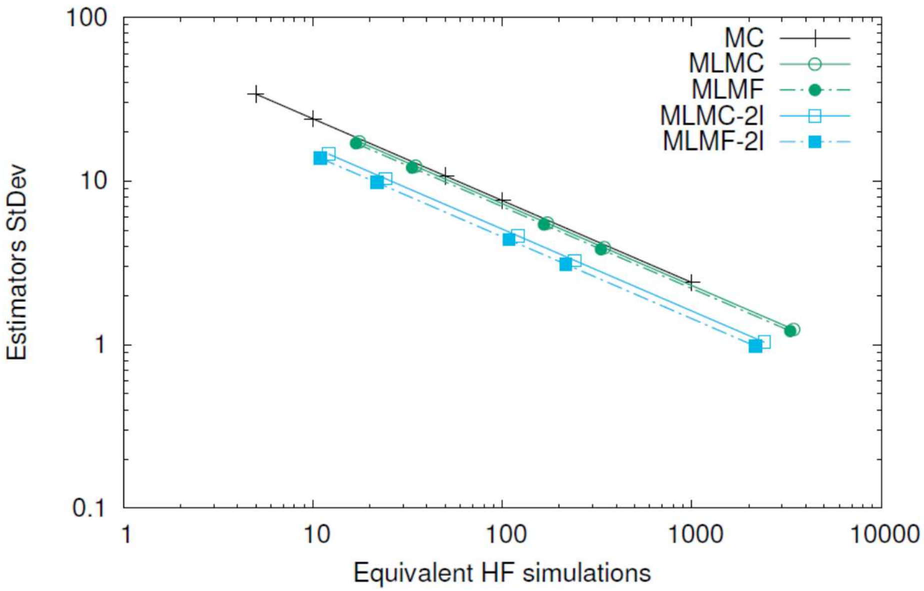
- Cost estimates for Nalu and OpenFAST simulations:

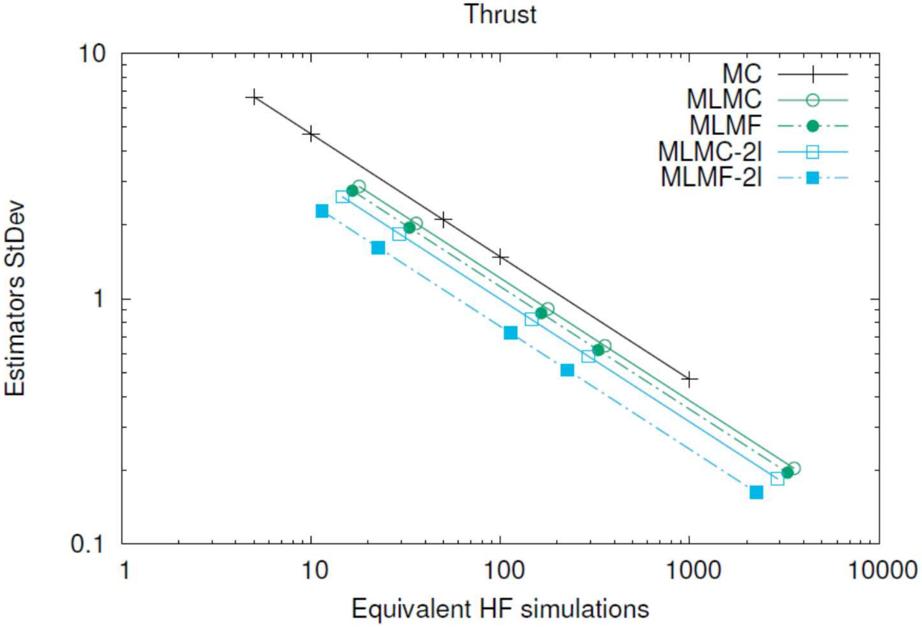
Case	Mesh size	Simulation time (seconds)	CPUs	Cost (CPU-hours)	Cost (relative)
OpenFAST		500	1	0.42	1
Coarse	100x50x50	2000	80	240	576
Medium	200x100x100	2000	160	960	2304
Fine	400x200x200	2000	400	6860	16500
Reference	800x200x200	2000	400	38400	91400

Domino, S. "Sierra Low Mach Module: Nalu Theory Manual 1.0", SAND2015-3107W, Sandia National Laboratories Unclassified Unlimited Release (UUR), 2015. https://github.com/NaluCFD/NaluDoc

Extrapolation

- Given the statistical properties estimated for power and thrust, we can extrapolate the behavior of several estimators:
  - Standard MC estimator
  - MLMC-3l: Multilevel $Q_0 + (Q_1 - Q_0) + (Q_2 - Q_1)$
  - MLMF-3l: MLMC-3l with CV for Q_0
  - MLMC-2l: Multilevel $Q_1 + (Q_2 - Q_1)$
  - MLMF-2l: MLMC-2l with CV for Q_1



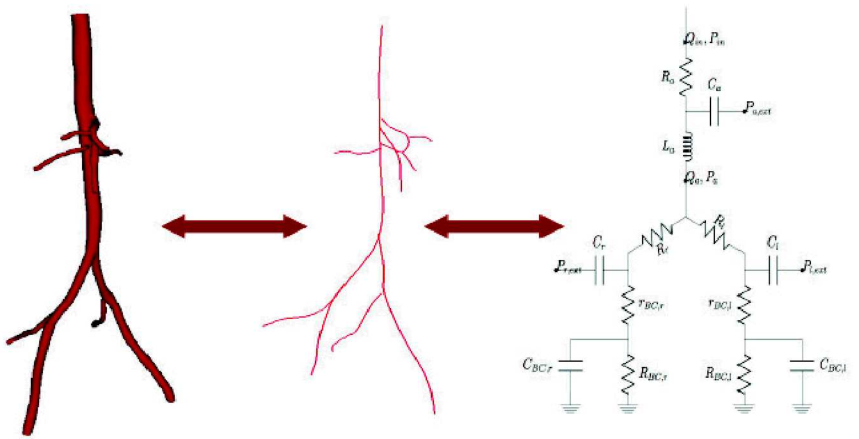


	Power			Thrust		
	MLMC	MLMF		MLMC	MLMF	
Level	Nalu	Nalu	OpenFAST	Nalu	Nalu	OpenFAST
0	161	137	2040	181	136	2887

Nalu LES for Q0 is too coarse will limited predictive value

Multilevel – Multifidelity Sampling Methods

Cardiovascular flow

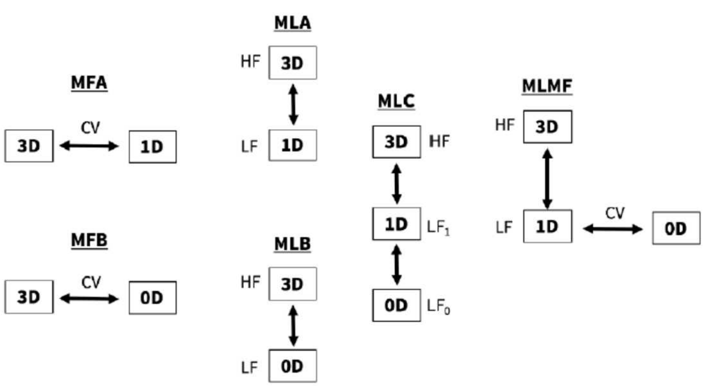


3D Model 1D Model 0D Model

Solver	Cost (1 simulation)	Effective Cost (No. 3D Simulations)
3D	96 hr	1
1D	11.67 min	2E-3
0D	5 sec	1.45E-5

Courtesy of C. Fleeter (Stanford), Prof. D. Schiavazzi (Notre Dame), Prof. A. Marden (Stanford)

Model relationships / graph topologies



(a) Multifidelity (CV) (b) Multilevel (c) Multilevel-Multifidelity

Costs to achieve prescribed error tolerance

Method	Effective Cost (3D Simulations)	No. 3D Simulations	No. 1D Simulations	No. 0D Simulations
MC	9 885	9 885	–	–
MFA	56	21	15 681	–
MFB	39	36	–	154 880
MLA	305	212	41 990	–
MLB	156	150	–	342 060
MLC	165	156	1 324	351 940
MLMF	165	156	1 249	362 590

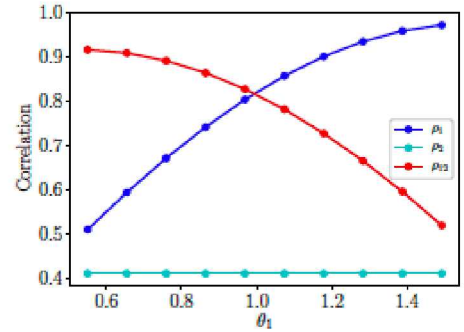
1D predictivity was insufficient and 0D contribution required control weighting

Generalized framework for approx. control variates

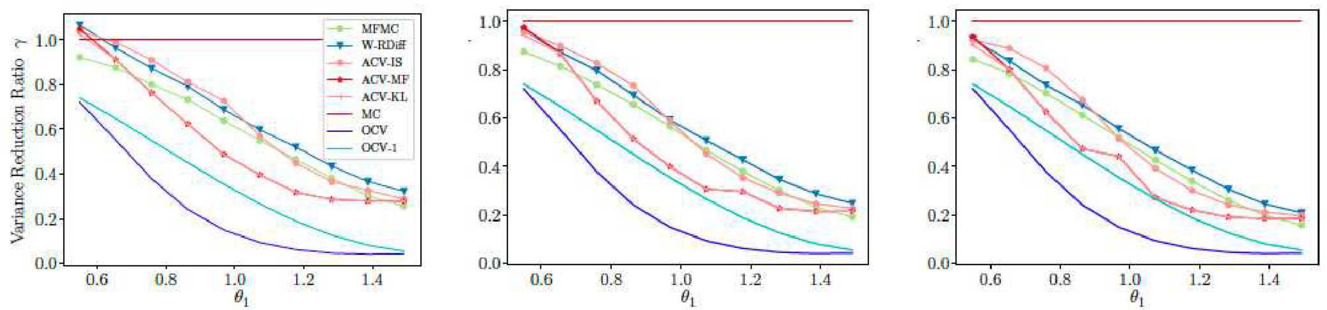
Tunable model problem

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

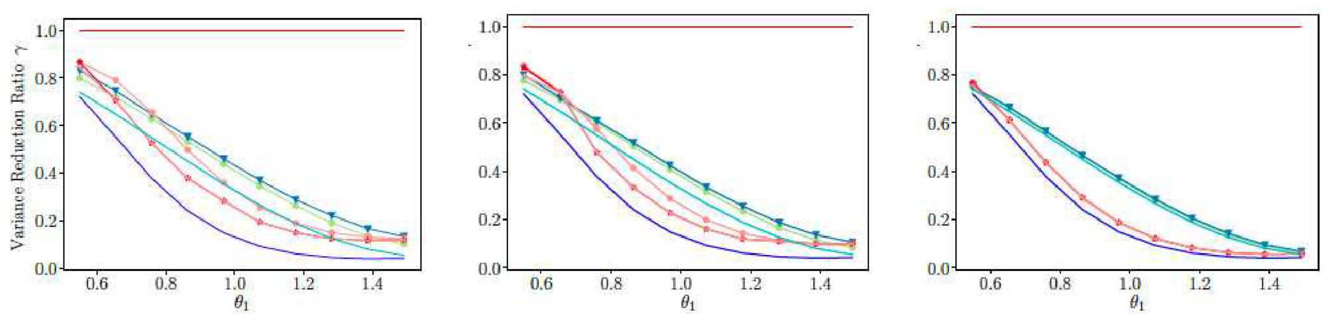
$A = \sqrt{11}$, $A_1 = \sqrt{7}$ and $A_2 = \sqrt{3}$ $\qquad$ $\theta = \pi/2$, $\theta_2 = \pi/6$ and $\theta_2 < \theta_1 < \theta$.



Exploiting the OCV vs. OCV-1 gap for different cost ratios



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



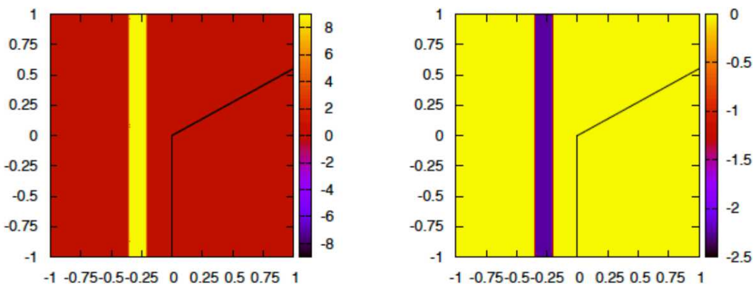
(d) $w = 50$ (e) $w = 100$ (f) $w = 1000$

A. Gorodetsky, G. Geraci, E., J. Jakeman “A Generalized Framework for Approximate Control Variates,” arxiv.org/abs/1811.04988v3

Generalized framework for approx. control variates

Two dimensional elasticity in heterogeneous media

Elastic wave propagation in 2D

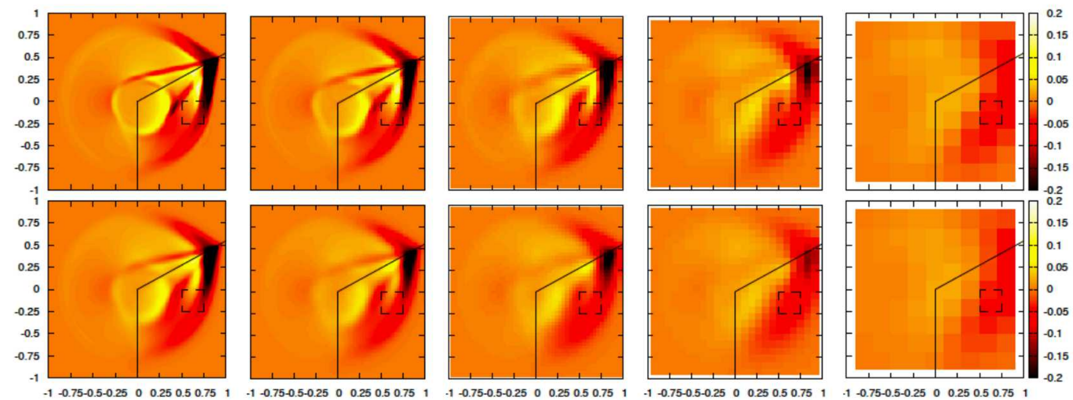


(a) Trace of the stress tensor.

(b) Velocity u component.

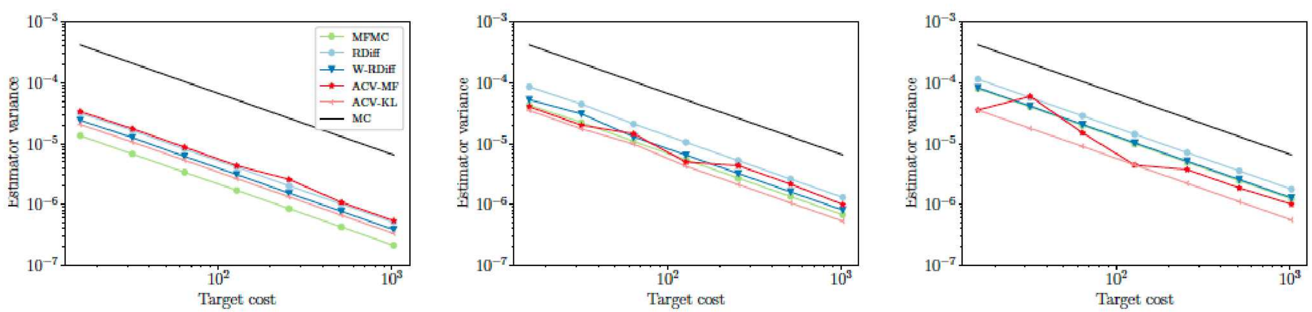
$$q_t + Aq_x + Bq_y = 0$$

Shear stress contours for 2 fidelities x 5 discretizations



Well correlated discretization hierarchy

General multifidelity with model gaps



Perfomance of ACV-KL is robust to reduced model predictivity

Stochastic Polynomial Expansion Methods

- Projection, Regression, Interpolation
- Multilevel | Multifidelity expansions (heuristic)
- Multilevel | Multifidelity expansions (optimized)

MF UQ with Spectral Stochastic Discrepancy Models

Sandia National Laboratories

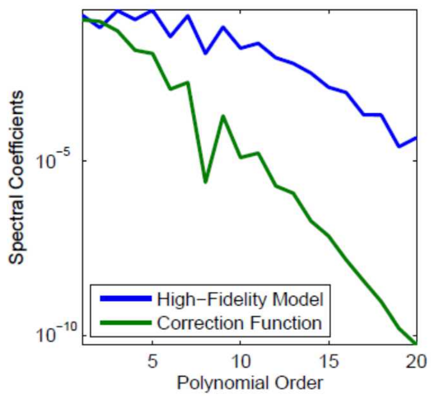
- High-fidelity simulations (e.g., RANS, LES) can be prohibitive for use in UQ
- Low fidelity "design" codes often exist that are predictive of basic trends
- Can we leverage LF codes w/i HF UQ in a rigorous manner? → global approxs. of model discrepancy

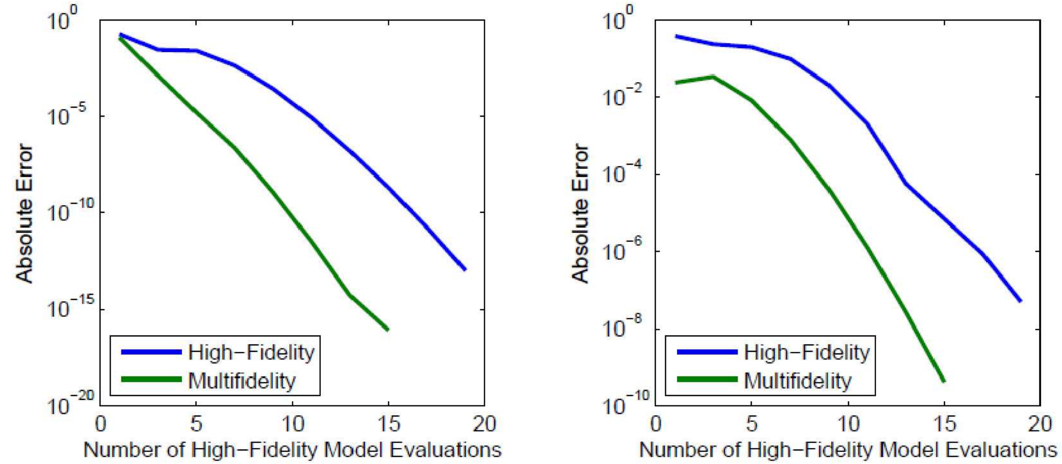
$$\boxed{\hat{f_{hi}}(\xi) = \sum_{j=1}^{N_{lo}} f_{lo}(\xi_j) L_j(\xi) + \sum_{j=1}^{N_{hi}} \Delta f(\xi_j) L_j(\xi)}$$

$$\boxed{N_{lo} >> N_{hi}}$$

$$R_{\text{high}}(\xi) = e^{-0.05\xi^2}\cos 0.5\xi \underbrace{- 0.5e^{-0.02(\xi-5)^2}}_{\text{discrepancy}}$$

$$R_{\text{low}}(\xi) = e^{-0.05\xi^2}\cos 0.5\xi,$$





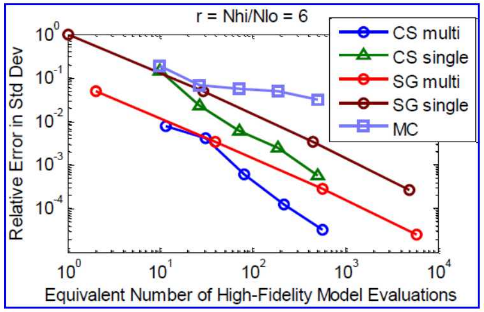
(a) Error in mean

(b) Error in standard deviation

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

Compressed sensing bi-fidelity: target sparsity

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



E., Ng, Barone, Domino, "Multifidelity Uncertainty Quantification Using Spectral Stochastic Discrepancy Models," Handbook of UQ, 2015.

ML PCE with rate estimation: Model Problem & UCAV Nozzle

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

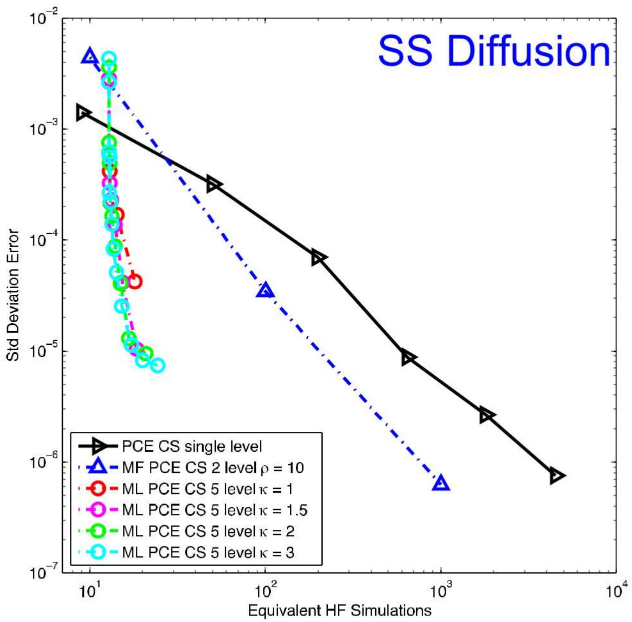
where x is the spatial coordinate, $\boldsymbol{\xi}$ a vector of independent random input parameters and $a(x,\boldsymbol{\xi})$ denotes the (random) diffusivity field. The following Dirichlet boundary conditions are also assumed

$$u(0,\boldsymbol{\xi}) = 0, \quad u(1,\boldsymbol{\xi}) = 0. \tag{23}$$

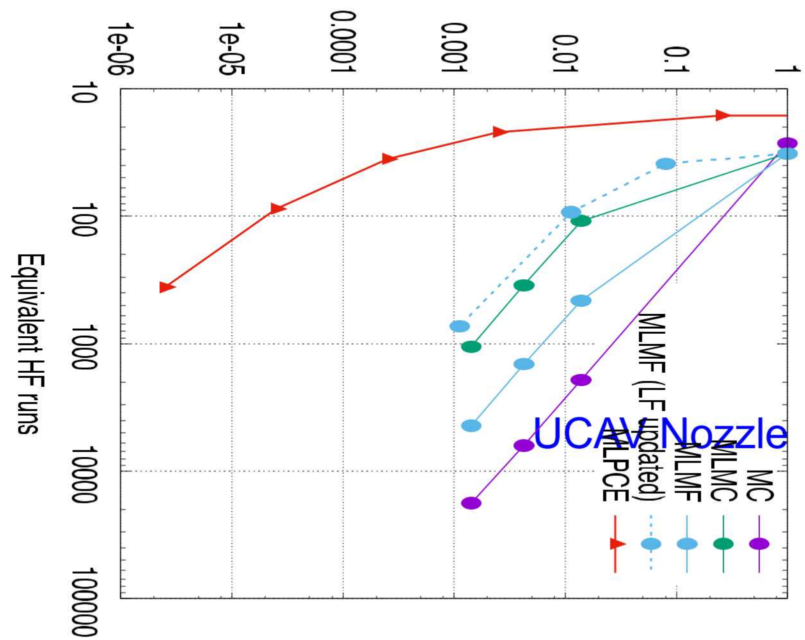
We are interested in quantifying the uncertainty in the solution u at specified spatial locations: $\bar{x} = 0.05, 0.5, 0.95$.

We represent the random diffusivity field a using the following expansion

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



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



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

ML PCE shows **more rapid convergence** using coarse/medium/fine discretizations:

- Exploits smoothness in moderate dim.

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

Algebraic Test Problems:
Numerical investigation of estimator variance relationships

- **Currin** *et al.* (1988) – Exponential function

$$f(\boldsymbol{\xi}) = \left[1 - \exp\left(-\frac{1}{2\xi_2}\right)\right] \frac{2300\xi_1^3 + 1900\xi_1^2 + 2092\xi_1 + 60}{100\xi^3 + 500\xi_1^2 + 4\xi_1 + 20}$$

$$\begin{aligned} f_{low}(\boldsymbol{\xi}) = &\frac{1}{4}\left[f(\xi_1 + 0.05, \xi_2 + 0.05) + f(\xi_1 + 0.05, \max(0, \xi_2 - 0.05)\right] \\ &+ \frac{1}{4}\left[f(\xi_1 - 0.05, \xi_2 + 0.05) + f(\xi_1 - 0.05, \max(0, \xi_2 - 0.05)\right] \end{aligned}$$

- **Short Column** (Eldred 2012 and Berchier 2016)

$$f(\boldsymbol{\xi}) = 1 - \frac{4M}{bh^2Y} - \left(\frac{P}{bhY}\right)^2$$

$$f_{low}(\boldsymbol{\xi}) = 1 - \frac{4P}{bh^2Y} - \left(\frac{P}{bhY}\right)^2$$

- **Park** (1991) – **F1**

$$f(\boldsymbol{\xi}) = \frac{\xi_1}{2}\left[\sqrt{1 + (\xi_2 + \xi_3^2)\frac{\xi_4}{\xi_1^2}}\right] + (\xi_1 + 3\xi_4)\exp(1 + \sin(\xi_3))$$

$$f_{low}(\boldsymbol{\xi}) = \left[1 + \frac{\sin(\xi_1)}{10}\right] f(\boldsymbol{\xi}) - 2\xi_1 + \xi_2^2 + \xi_3^2 + 0.5$$

- **Park** (1991) – **F2**

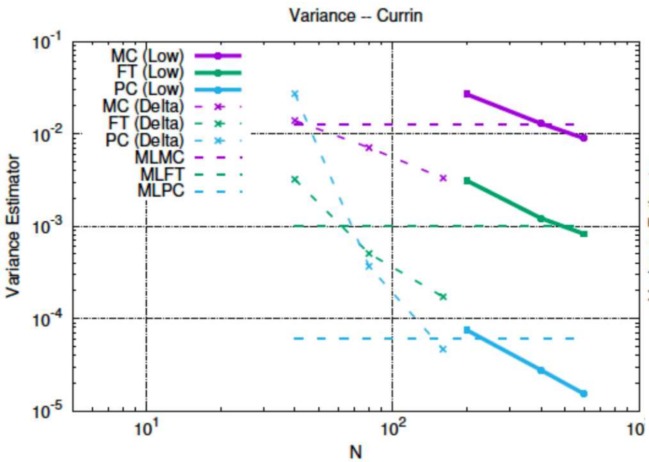
$$f(\boldsymbol{\xi}) = \frac{2}{3}\exp(\xi_1 + \xi_2) - \xi_4 \sin(\xi_3) + \xi_3$$

$$f_{low}(\boldsymbol{\xi}) = 1.2\, f(\boldsymbol{\xi}) - 1$$

Numerical Investigation of Estimator Variance:
ML MC, ML PCE CS, and ML FT

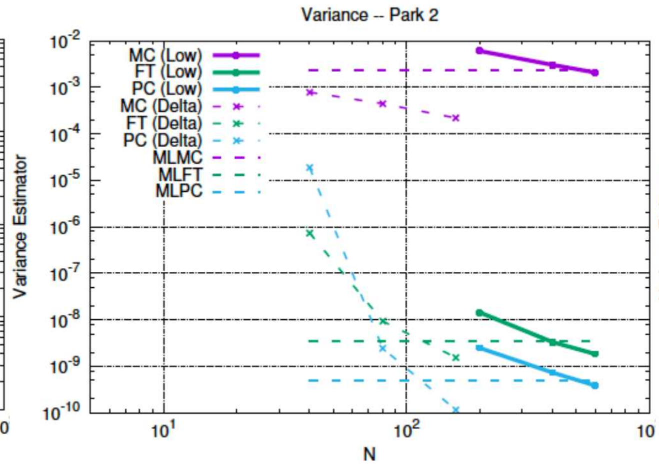
CURRIN

$\mathbb{V}ar\left(\hat{Y}_\ell\right)$



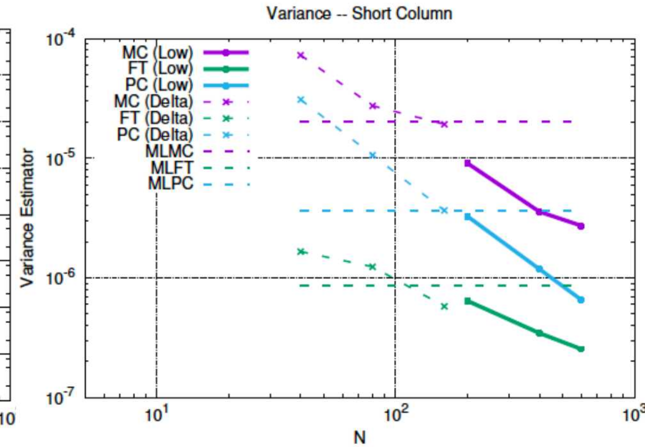
PARK 2

$\mathbb{V}ar\left(\hat{Y}_\ell\right)$



SHORT COLUMN

$\mathbb{V}ar\left(\hat{Y}_\ell\right)$



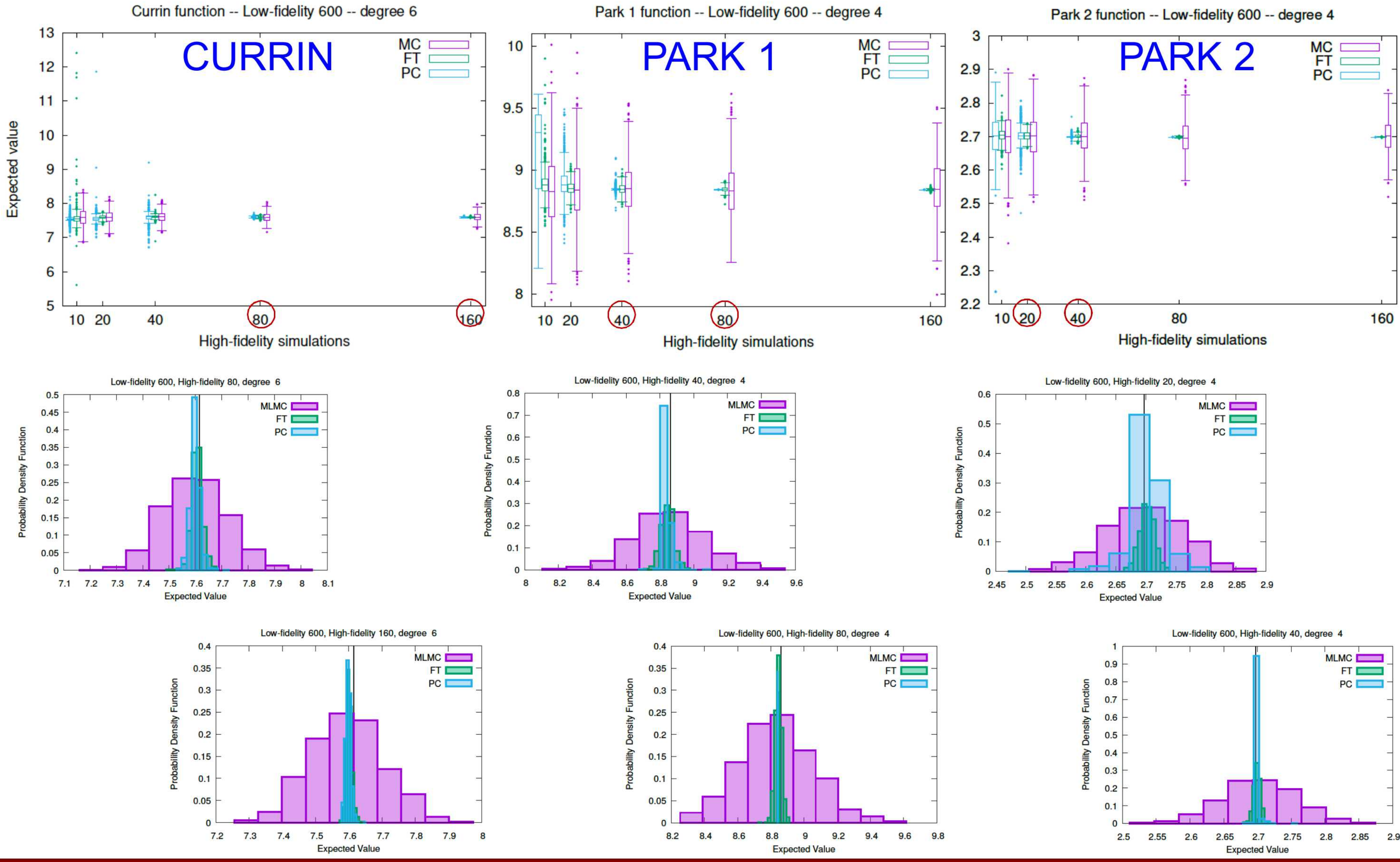
	γ		κ	
	Low	Δ	Low	Δ
MC	0.997	$8.47E-01$	1.005	$1.03E+00$
FT	2.737	$7.97E-02$	1.224	$2.11E+00$
PC	34.040	$1.26E-06$	1.446	$4.59E+00$

	γ		κ	
	Low	Δ	Low	Δ
MC	0.572	1.275	1.120	0.959
FT	33.024	97.752	0.848	0.759
PC	0.247	0.321	1.463	1.538

	γ		κ	
	Low	Δ	Low	Δ
MC	$1.13E+00$	$1.43E+00$	$9.78E-01$	$9.17E-01$
FT	$3.61E+03$	$5.33E-03$	$1.90E+00$	$4.44E+00$
PC	$5.82E+04$	$6.59E-11$	$1.70E+00$	$8.65E+00$

Kinks in estimated rate curves ?

Numerical Investigation of Estimator Variance:
ML MC, ML PCE, and ML FT



Abrupt threshold in CS / FFT: more important to ensure sufficient recovery than estimate smooth rate that follows

Multilevel-multifidelity expansion methods – beyond rate estimation
Algorithm 1: RIP sampling for multilevel sparse recovery

Address/avoid abrupt transition by ensuring sufficient sample density for accurate recovery

Restricted isometry property (RIP) based on sparsity *s*, cardinality *C*, mutual coherence *L*:

$$\boxed{N_l \geq s_l \, log^3(s_l) \; L_l \, log(C_l)}$$

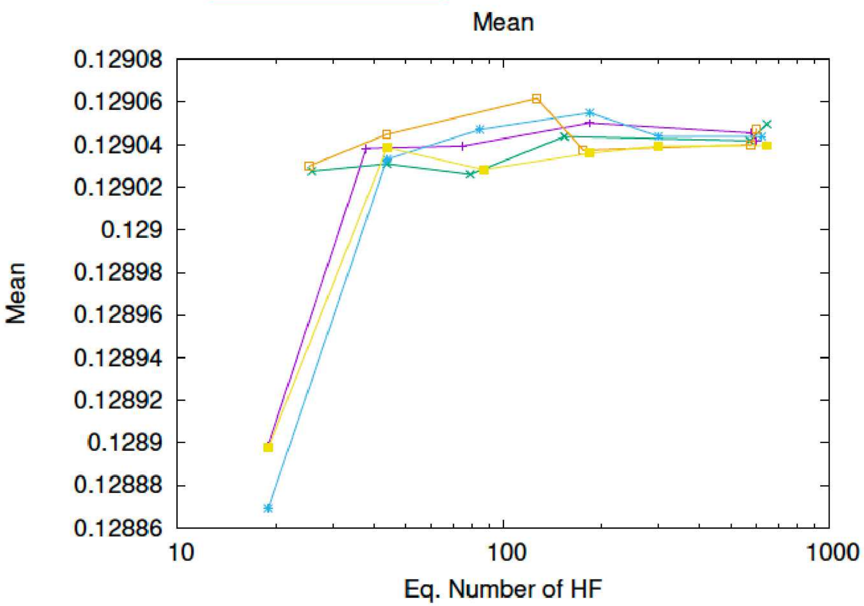
Jakeman, Narayan, and Zhou, 2016

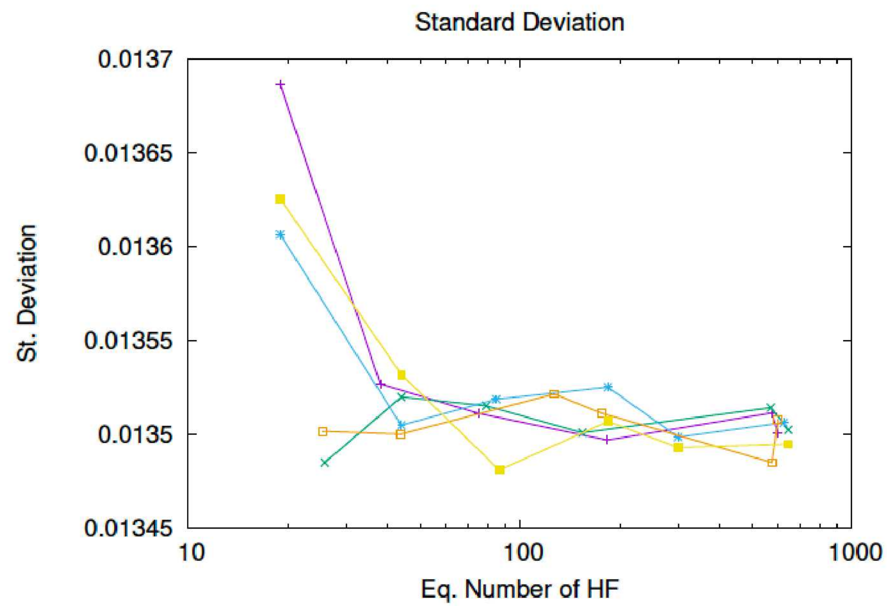
Basic Algorithm:

- Starting from a pilot sample, shape profile based on observed sparsity & iterate until convergence
- In practice, RIP sampling levels are quite conservative → enforce a constraint on the profile

$$\boxed{N_l \leq r C_l}$$

for some collocation ratio r





Challenge: feedback not well controlled for compressible fns → more samples allow accurate recovery of more terms

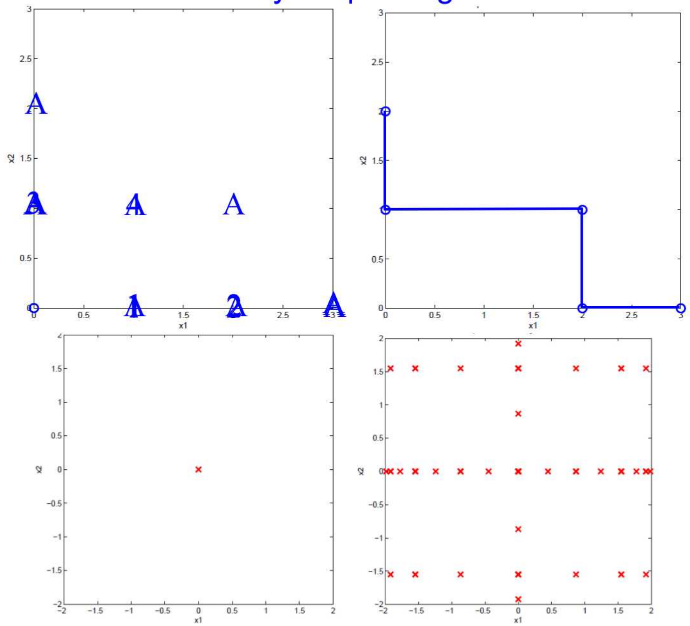
Multilevel-multifidelity expansion methods – beyond rate estimation

Algorithm 2. Greedy ML PCE with generalized sparse grids

GSG for a single level / fidelity:

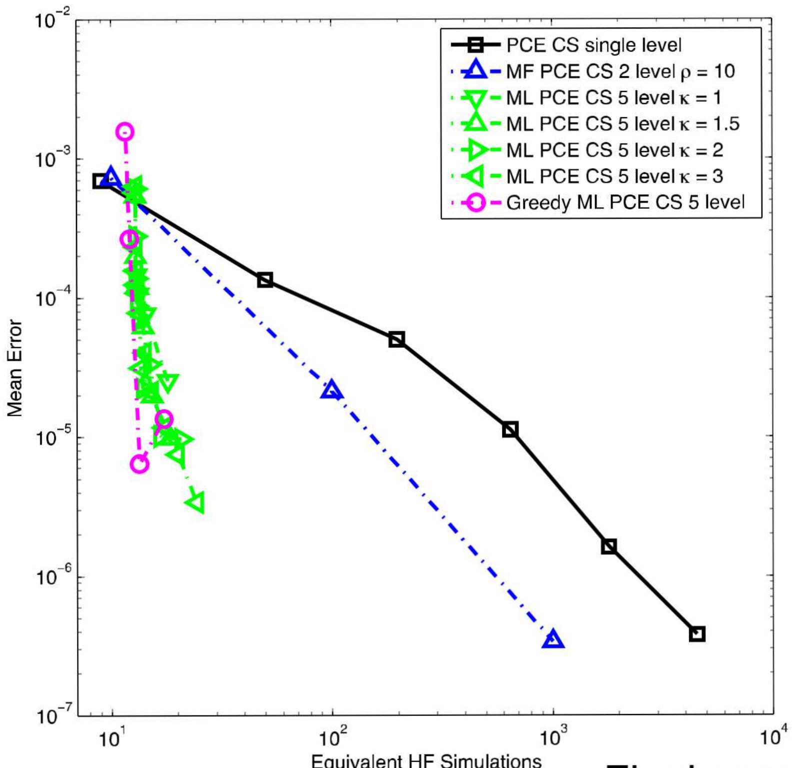
1. ***Initialization:*** Starting from reference grid (often w = 0 grid), define active index sets using admissible forward neighbors of all old index sets.
2. ***Trial set evaluation:*** For each trial index set, evaluate tensor grid, form tensor expansion, update combinatorial coefficients, and combine with reference expansion. Perform necessary bookkeeping to allow efficient restoration.
3. ***Trial set selection:*** Select trial index set that induces largest change in statistical QOI.
4. ***Update sets:*** If largest change > tolerance, then promote selected trial set from active to old and compute new admissible active sets; return to 2. If tolerance is satisfied, advance to step 5.
5. ***Finalization:*** Promote all remaining active sets to old set, update combinatorial coefficients, and perform final combination of tensor expansions to arrive at final result for statistical QOI.

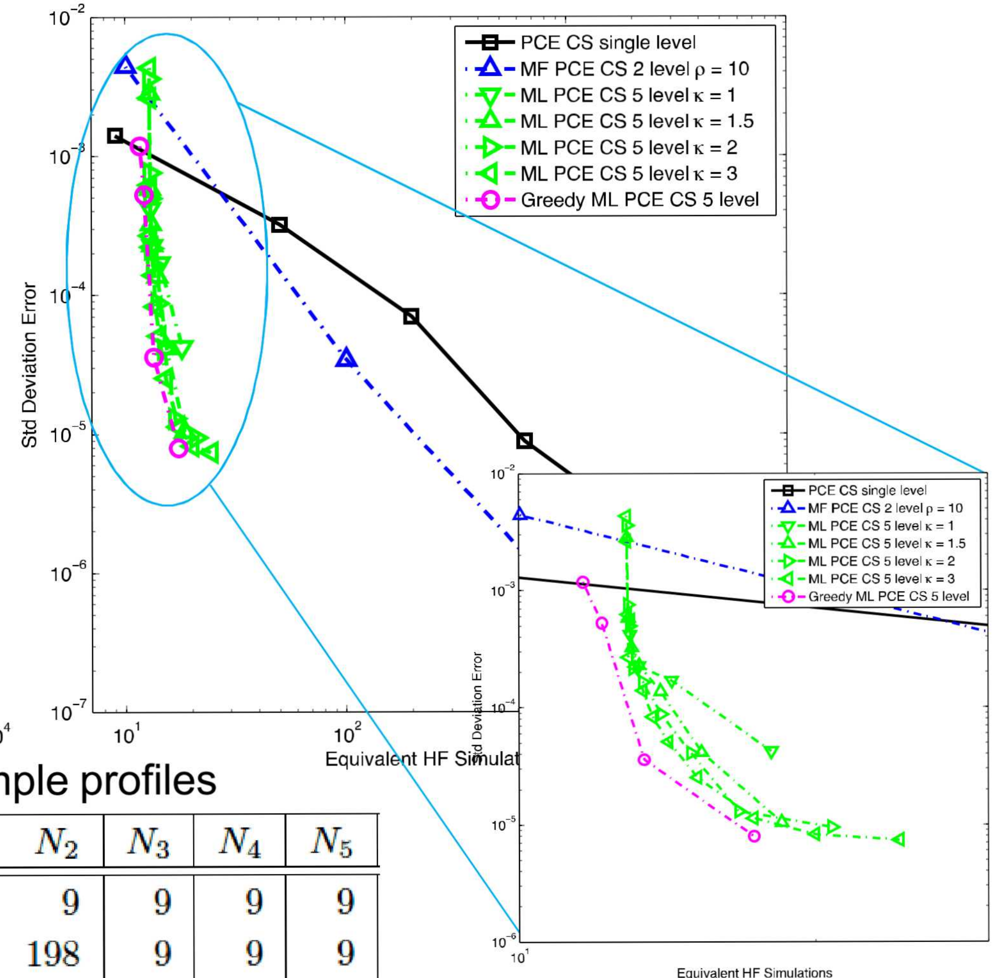
Smolyak sparse grid



Sandia National Laboratories

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

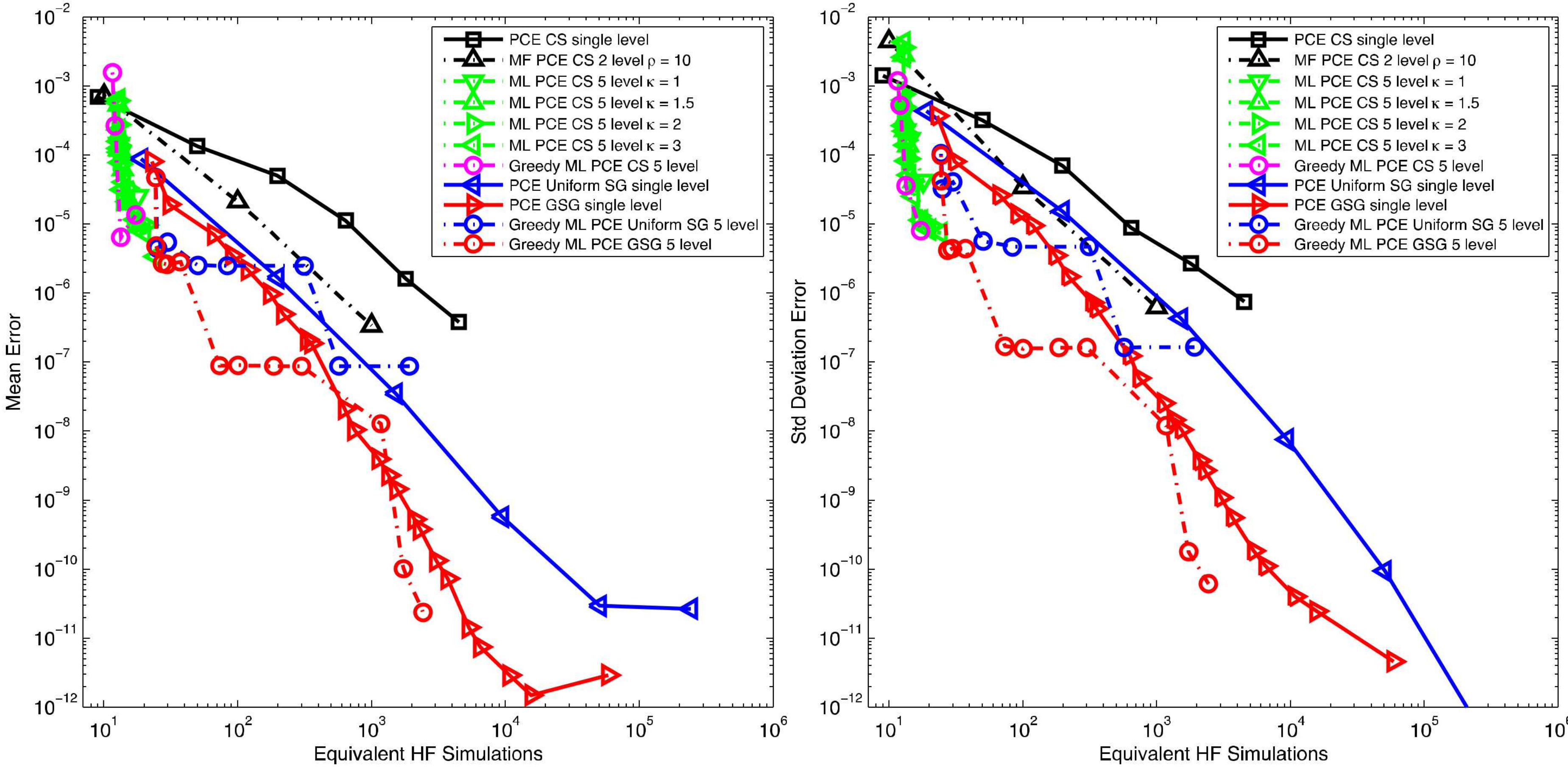




Final sample profiles

Conv Tol	N_1	N_2	N_3	N_4	N_5
1.e-1	198	9	9	9	9
1.e-2	644	198	9	9	9
1.e-3	1802	644	9	9	9
1.e-4	4505	1802	50	9	9

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

Emulator-based MCMC

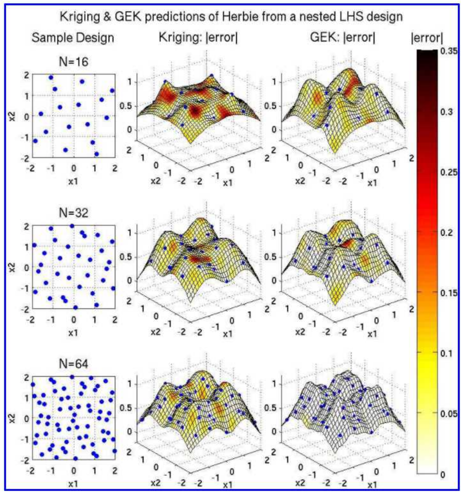
- Emulators eliminate direct interface of MCMC with expensive HF models
  - Polynomial chaos expansion (PCE), Stochastic collocation (SC), Gaussian process (GP)

Polynomial chaos: spectral projection w/ orthogonal poly basis

$$\boxed{R \cong \sum_{j=0}^{P} \alpha_j \Psi_j(\boldsymbol{\xi})}$$

$$\boxed{\alpha_j = \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} = \frac{1}{\langle \Psi_j^2 \rangle} \int_{\Omega} R \, \Psi_j \, \varrho(\boldsymbol{\xi}) \, d\boldsymbol{\xi}}$$

GPs/GEK: pivoted Cholesky



Stochastic collocation: form interpolants for known coeffs

- Global / local, nodal / hierarchical, value-based / deriv-enhanced

$$\boxed{R(\boldsymbol{\xi}) \cong \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi})}$$

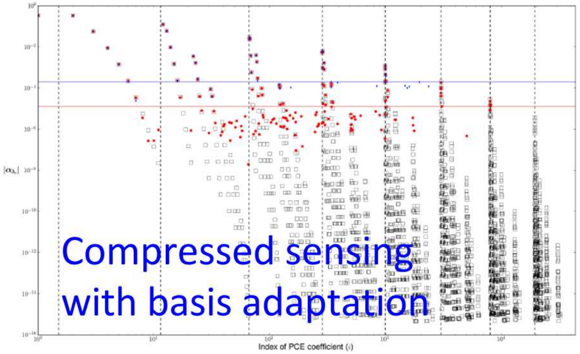
$$\boxed{L_j = \prod_{\substack{k=1 \\ k \neq j}}^{m} \frac{\xi - \xi_k}{\xi_j - \xi_k}}$$

Candidate basis terms (P)

Observations (N)

$$\begin{bmatrix} f(\mathbf{x}^{(1)}) \\ f(\mathbf{x}^{(2)}) \\ \vdots \\ f(\mathbf{x}^{(N)}) \end{bmatrix} = \begin{bmatrix} 1 & \phi_2(\mathbf{x}^{(1)}) & \phi_2(\mathbf{x}^{(1)}) & \dots & \phi_P(\mathbf{x}^{(1)}) \\ 1 & \phi_1(\mathbf{x}^{(2)}) & \phi_2(\mathbf{x}^{(2)}) & \dots & \phi_P(\mathbf{x}^{(2)}) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \phi_1(\mathbf{x}^{(N)}) & \phi_2(\mathbf{x}^{(N)}) & \dots & \phi_P(\mathbf{x}^{(N)}) \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_P \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

Over-determined (LS), under-determined (CS), Orthog. least interp.



Compressed sensing with basis adaptation

- **For this milestone, we employ the ML PCE emulator from the forward UQ study**

Exploit structure provided by emulator

Sandia National Laboratories

Form accurate proposal density using Hessian data

- Analytic derivatives of PCE/SC/GP → Hessian of misfit → proposal covariance

$$\boxed{p(\mathbf{d}|\xi) = \exp\left[-\frac{1}{2}(f(\xi)-\mathbf{d})^T\Gamma_{\mathbf{d}}^{-1}(f(\xi)-\mathbf{d})\right]}$$

Gaussian Likelihood

$$\boxed{-\log\left[p(\mathbf{d}|\xi)\right] = \frac{1}{2}(f(\xi)-\mathbf{d})^T\Gamma_{\mathbf{d}}^{-1}(f(\xi)-\mathbf{d}) = M(\xi)}$$

Negative Log Likelihood = Misfit

$$\boxed{\nabla_\xi M(\xi) = \nabla_\xi f(\xi)^T\,\Gamma_{\mathbf{d}}^{-1}\,(f(\xi)-\mathbf{d})}$$

Gradient of Misfit

$$\boxed{\nabla_\xi^2 M(\xi) = \underbrace{\nabla_\xi f(\xi)^T\,\Gamma_{\mathbf{d}}^{-1}\,\nabla_\xi f(\xi)}_{\text{Gauss-Newton approximate Hessian (if only emulator grads)}} + \nabla_\xi^2 f(\xi)\cdot\left[\Gamma_{\mathbf{d}}^{-1}\,(f(\xi)-\mathbf{d})\right]}$$

Hessian of Misfit

Use MVN proposal with covariance defined by inverse Hessian (Laplace approx) of:

- <u>*Negative log posterior*</u>, augmenting misfit with log prior

$$\boxed{-\log\pi_d(\xi) = M(\xi) - \log\pi_0(\xi)}$$

  - Hessian of negative log prior results in a regularization with many priors: normal, beta, gamma
  - No effect in other cases: uniform, exponential, triangular, histogram

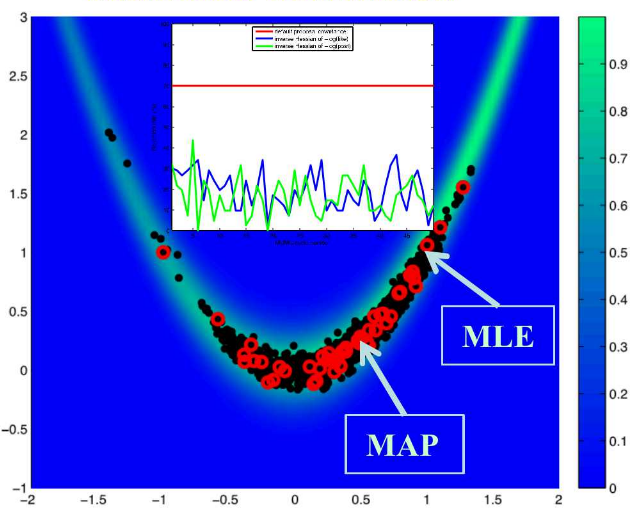
Multilevel emulator-based Bayesian inference: Model Problem (Rosenbrock)

Sandia National Laboratories

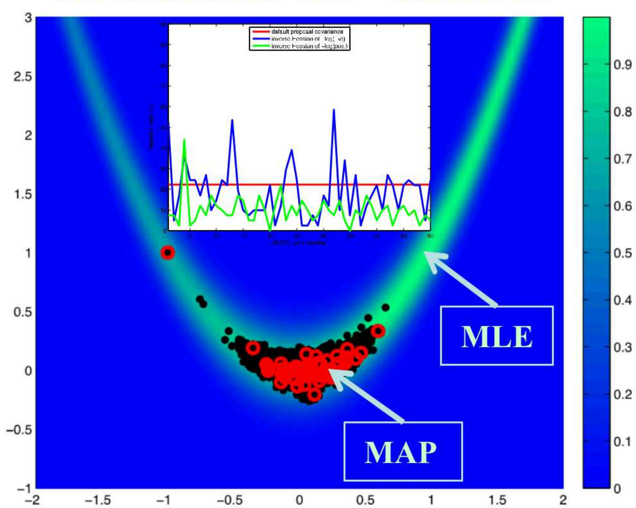
Greedy ML PCE expansion from forward UQ → efficient MCMC for data assimilation

- Define simulation-to-experiment or **simulation-to-reference-simulation** misfit function
- Markov Chain Monte Carlo (MCMC) performed on emulator, leveraging emulator structure
  - Pre-solve for MAP (maximum a posteriori probability) point: full Newton min of –log(posterior)
  - Accurate MCMC proposal mitigates sample rejection in high D
    - In 10D, 98% rejection rate reduced to 30%
  - Posterior Hessian-based proposal balances likelihood and prior, performing better than either alone

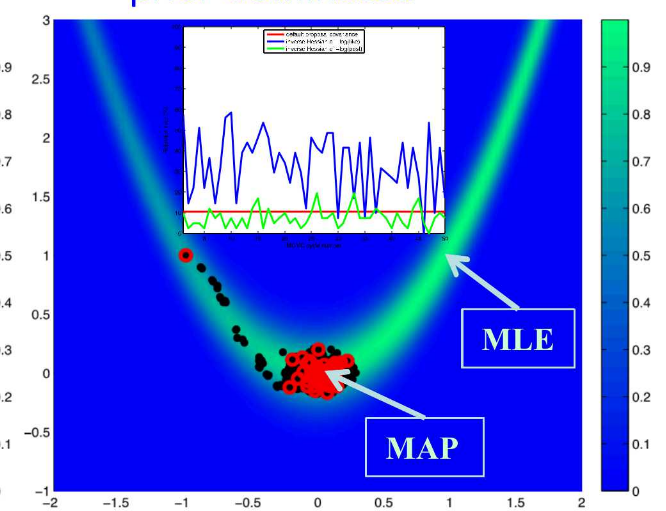
Weak prior ~ N(0,1)
Hessian of $-\log(\pi_d)$ is likelihood-dominated



Medium prior ~ N(0,.25)
Hessian of $-\log(\pi_d)$: L and π_0 contributions are ~ balanced



Strong prior ~ N(0,.1)
Hessian of $-\log(\pi_d)$ is prior-dominated



MCMC posterior samples in black, Hessian updates in red, insets plot rejection rate vs. Hessian-preconditioned MCMC cycle for 3 proposals

Multigrid optimization (MG/Opt)

SIAM J. SCI. COMPUT.
Vol. 26, No. 6, pp. 1811–1837

MODEL PROBLEMS FOR THE MULTIGRID OPTIMIZATION OF SYSTEMS GOVERNED BY DIFFERENTIAL EQUATIONS*

ROBERT MICHAEL LEWIS† AND STEPHEN G. NASH‡

SIAM REVIEW
Vol. 51, No. 2, pp. 361–395

Multigrid Methods for PDE Optimization*

Alfio Borzì†
Volker Schulz‡

As in multilevel Monte Carlo, exploit discretization hierarchy within optimization/OUU:

- Apply multigrid V cycle to hierarchy of *optimal* solns
  - Distinct from applying multigrid to KKT system
  - Distinct from successive refinement of optimal solns (employs bi-directional prolongation / restriction)

Recursively uses coarse resolution problems to generate search directions for finer-resolution solves

- Line search used to compute fine-resolution iterate from coarse-resolution search direction
- Globalization enables provable convergence

Special case of / component within generalized model management framework

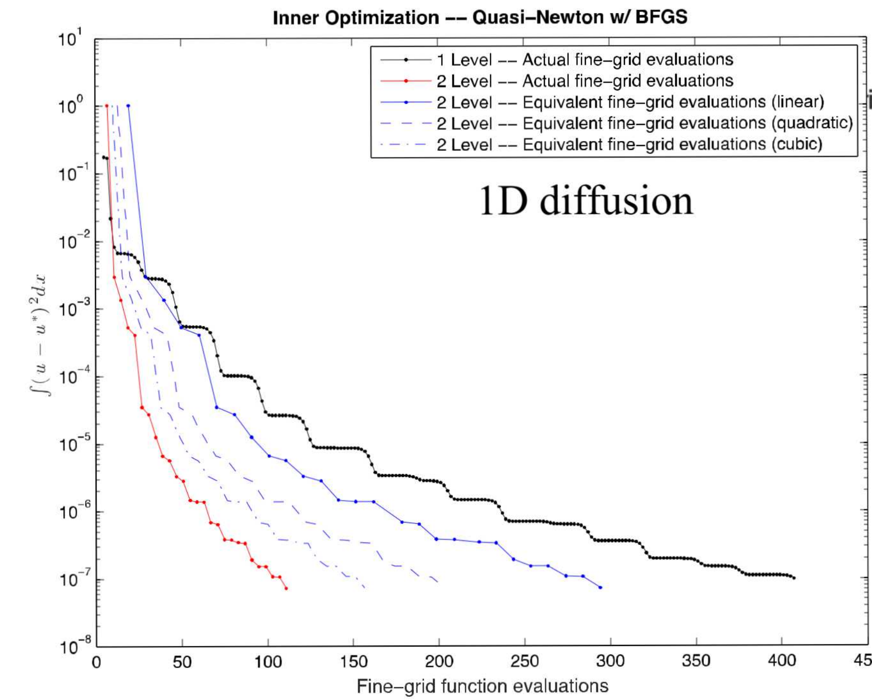
- Requires effective subproblem solver to generate a new iterate at a particular level
- Leverages 1st and (quasi, finite diff.) 2nd-order additive & combined corrections

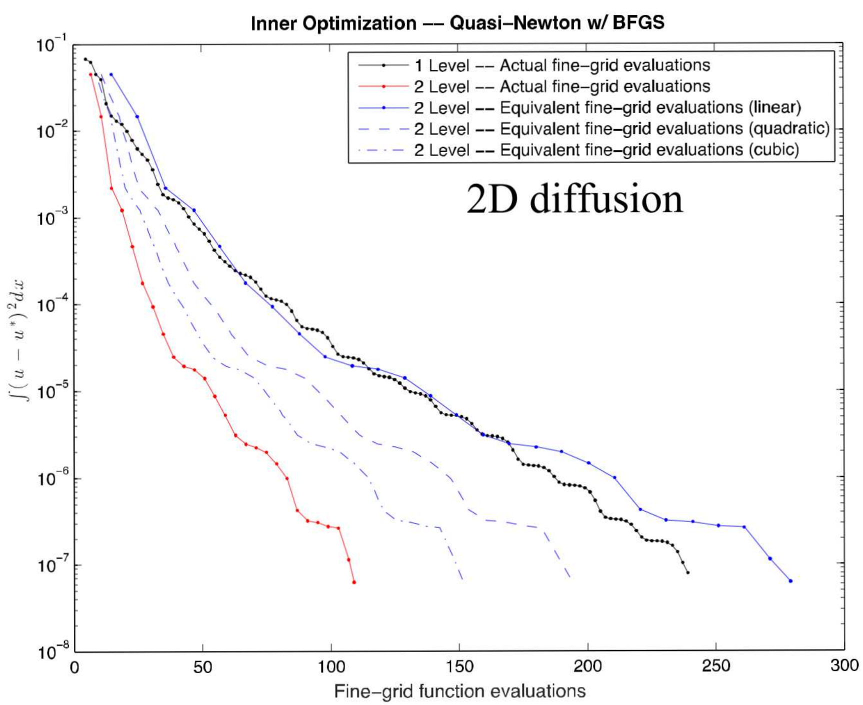
Algorithm 1 Multigrid Optimization

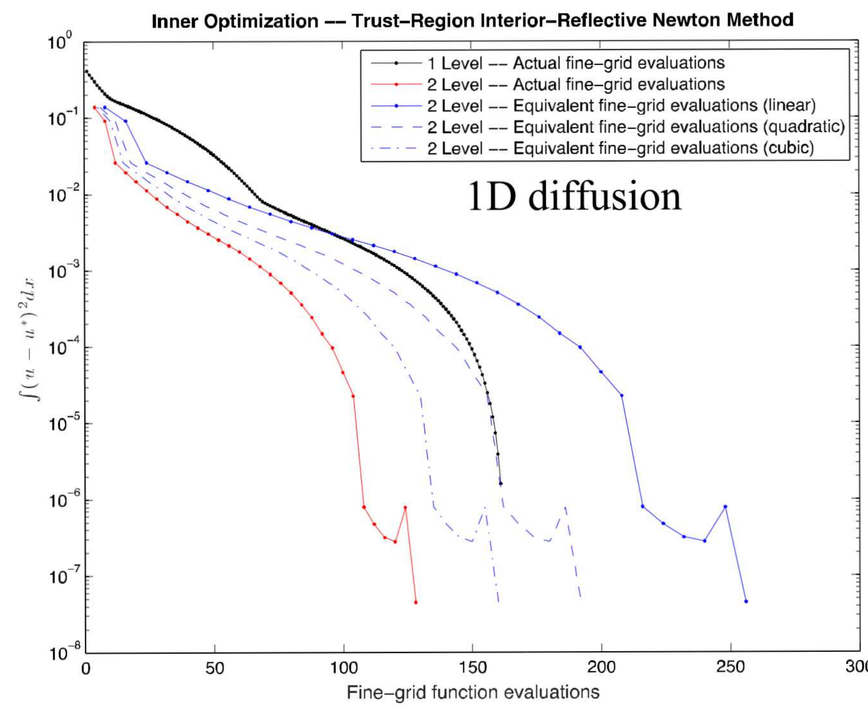
procedure MGOPT(k, $x_0^{(k)}$, $f^{(k)}(x)$, $v^{(k)}$)
  if $k = 0$ **then**
    $x_1^{(k)} = \arg\min_x \quad f^{(k)}(x) - \left[v^{(k)}\right]^T x$
    return $x_1^{(k)}$
  else
Dive
    Partially solve: $x_1^{(k)} = \arg\min_x \quad f^{(k)}(x) - \left[v^{(k)}\right]^T x$
    $x_1^{(k-1)} = R\left[x_1^{(k)}\right]$
    $v^{(k-1)} = \nabla f^{(k-1)}(x_1^{(k-1)}) - R\left[\nabla f^{(k)}(x_1^{(k)})\right]$
    $x_2^{(k-1)} = $ MGOPT($k-1$, $x_1^{(k-1)}$, $f^{(k-1)}(x)$, $v^{(k-1)}$)
Return
    $e = P\left[x_2^{(k-1)} - x_1^{(k-1)}\right]$
    $x_2^{(k)} = x_1^{(k)} + \alpha e$
    return $x_2^{(k)}$
  end if
end procedure

MG/Opt for Multilevel

- Coarse and fine discretizations (50, 100 pts) for LF model form
- Adjoint design gradients
- Prolongation/restriction via Lagrange interp.
- MATLAB Optimization Solvers (2)
- Grid scalings (1D and 2D diffusion)
- Solver cost scalings (linear, quadratic, cubic)





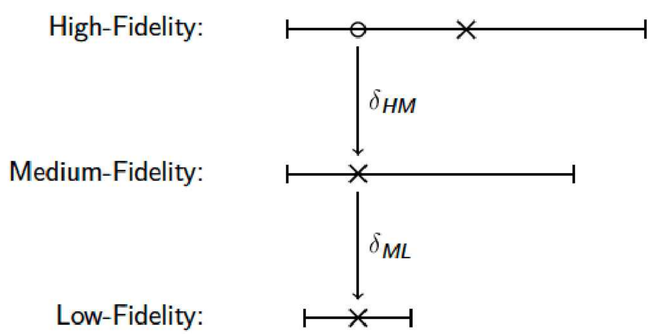


Recursive Trust Region Model Management

Algorithm 9 Recursive Trust Region Updating

procedure RECTR($r = 1$:LEN, x_c^r, x_*^r, $f_{\text{CORR}}^r(x)$)
  for $r =$ len to 1 (bottom up: low to high || coarse to fine) **do**
    if State$_r$ = new candidate x_*^r **then**
      Test for new center: TR(x_c^r, x_*^r, $f_{\text{corr}}^{r+1}(x)$, $f_{\text{corr}}^r(x)$)
    end if
    if State$_r$ = new center x_c^r **then**
      Compute $f^{r-1}(x_c^r)$
      Compute CORRECTION($x_c^r, R, f^{r-1}(x_c^r), f^r(x_c^r)$)
      if Converged(x_c^r, $f_{\text{corr}}^{r-1}(x_c^r)$, L^{r-1}, U^{r-1}) **then**
        $x_*^{r-1} = x_c^r$ (new candidate)
      end if
    end if
  end for
  for $r = 1$ to len (top down: high to low || fine to coarse) **do**
    if State$_r$ = new center x_c^r **then**
      Recompute CORRECTION($x_c^r, R, f^{r-1}(x_c^r), f^r(x_c^r)$)
    end if
    if parent corrected **then**
      Recur updated corrections for $f_{\text{corr}}^r(x_c^r)$
    end if
    Reset State$_r$
  end for
end procedure

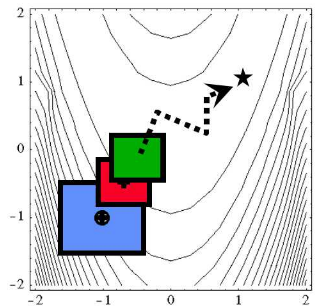
Repeat until convergence on high-fidelity model



SEQUOIA: DUU w/ Recursive Trust Region Model Management

Algorithm 9 Recursive Trust Region Updating

procedure RECTR($r = 1$:LEN, x_c^r, x_*^r, $f_{\text{CORR}}^r(x)$)
  for r = len to 1 (bottom up: low to high || coarse to fine) **do**
    if State$_r$ = new candidate x_*^r **then**
      Test for new center: TR(x_c^r, x_*^r, $f_{\text{corr}}^{r+1}(x)$, $f_{\text{corr}}^r(x)$)
    end if
    if State$_r$ = new center x_c^r **then**
      Compute $f^{r-1}(x_c^r)$
      Compute CORRECTION(x_c^r, R, $f^{r-1}(x_c^r)$, $f^r(x_c^r)$)
      if Converged(x_c^r, $f_{\text{corr}}^{r-1}(x_c^r)$, L^{r-1}, U^{r-1}) **then**
        $x_*^{r-1} = x_c^r$ (new candidate)
      end if
    end if
  end for
  for r = 1 to len (top down: high to low || fine to coarse) **do**
    if State$_r$ = new center x_c^r **then**
      Recompute CORRECTION(x_c^r, R, $f^{r-1}(x_c^r)$, $f^r(x_c^r)$)
    end if
    if parent corrected **then**
      Recur updated corrections for $f_{\text{corr}}^r(x_c^r)$
    end if
    Reset State$_r$
  end for
end procedure



Input Variables

Design Variables | Random Variables

- **Geometry**
  - interior wall shape (B-spline parameterization)
  - thermal layer thickness
  - baffle geometry
  - load layer thickness
  - stringer geometry
- **Material Properties**
  - density
  - elastic modulus
  - Poisson ratio
  - thermal conductivity
  - thermal expansion coef.
- **Inlet Conditions**
  - inlet stagnation temperature
  - inlet stagnation pressure
- **Environment**
  - atmospheric temperature
  - atmospheric pressure
  - heat transfer coef. to environment

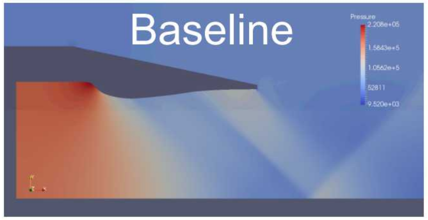
UCAV Nozzle Robust Design (Aero, Structural, Thermal)

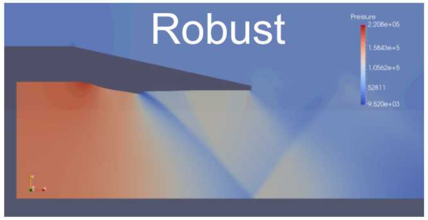
- TRMM
  - LF: L1 sparse grid w/ Euler COARSE
  - HF: L1 sparse grid w/ Euler MEDIUM
  - 1st-order consistent w/ FD gradients
  - Converged in 5 iterations using filter method
- Design variables
  - 21 B-spline & 8 thickness
- Random variables
  - 7 total

$$
\begin{aligned}
\min_{x} \quad & \mathrm{Var}(F(x,\xi)) \\
\text{subject to} \quad & \mathrm{E}[M(x)] \le M_{req} \\
& \mathrm{E}[F(x,\xi)] \ge F_{req} \\
& \mathrm{E}[\|T_0\|] \le T_{max,0} \\
& \mathrm{E}[\|S_0(\sigma_0(x,\xi))\|] \le 1 \\
& Ax \le b \\
& l \le x \le u
\end{aligned}
$$

single temperature constraint for entire geometry

von Mises stress used as single failure criterion





- Leverage multiple (deterministic) simulation fidelities / discretizations
- Leverage multiple (stochastic) UQ approaches / resolutions

DUU of Scramjet-Inspired Model Problem:
MLMC (UQ level) + noise-tolerant DFO informed by std error estimates

- **Objective:** Pressure loss (inlet outlet)

$$P_{loss} = \left(P_0^{in} - \frac{1}{L_y} \int_{out} P_0(y)\mathrm{d}y \right) / P_0^{in}$$

- **Constraint:** Average temperature along the cavity centerline

$$T_c = \frac{1}{L_x} \int_{L_c} T(x)\mathrm{d}x$$

- **OUU formulation:**

$$p^*_{\text{loss}}(\vec{x}^*, \vec{q}) = \min_{\vec{x}} \mathbb{E}[p_{\text{loss}}(\vec{x}, \vec{q})]$$

$$\text{s.t. } 593 \le \mathbb{E}[T_{\text{cav}}(\vec{x}, \vec{q})] - 3\sigma[T_{\text{cav}}(\vec{x}, \vec{q})]$$

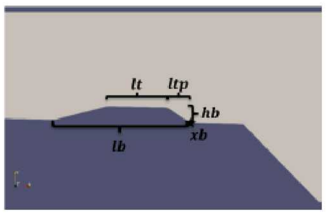


Figure: Design parameters

- **Design:** $\vec{x} = [hb, lt, ltp, lb, xb]^T$

$$0.5 \le hb \le 2.5$$
$$7.5 \le lt \le 11.5$$
$$2.5 \le ltp \le 4.5$$
$$17.5 \le lb \le 20.5$$
$$79.5 \le xb \le 85.5$$

- **Uncertain:** $\vec{q} = [p_{0,in}, T_{0,in}, M_{in}]^T$

$$p_{0,in} \sim \mathcal{N}(1.48e6, (7.4e3)^2)$$
$$T_{0,in} \sim \mathcal{N}(1550, 7.75^2)$$
$$M_{in} \sim \mathcal{N}(2.51, 0.01255^2)$$

Multilevel error estimators for $\mathbb{E}$, $\mathbb{V}$ and $\sqrt{\mathbb{V}}$.

Variance s^2_{ML}:

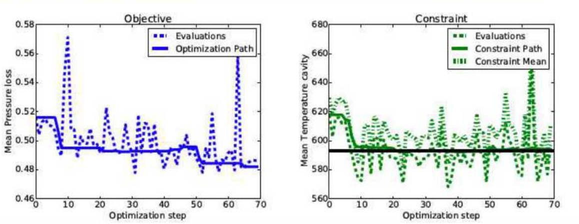
$$\mathbb{V}[s^2_{ML}] = \sum_{\ell=0}^{L} \mathbb{V}[\widehat{P^2_\ell}] + \mathbb{V}[\widehat{P^2_{\ell-1}}] - 2\mathrm{Cov}(\widehat{P^2_\ell}, \widehat{P^2_{\ell-1}})$$

where $\mathbb{V}[\widehat{P^2_\ell}] = \frac{1}{N_\ell}(\mu_{4,\ell} - \mathbb{V}[Q_\ell]^2) + \frac{2}{N_\ell(N_\ell - 1)}\mathbb{V}[Q_\ell]^2$

Standard deviation $\sqrt{s^2_{ML}}$:

$$SE(s^2_{ML}) \approx \frac{1}{2\sqrt{s^2_{ML}}}\sqrt{\mathbb{V}[s^2_{ML}]}$$

Stochastic MC OPTIMIZATION – Results



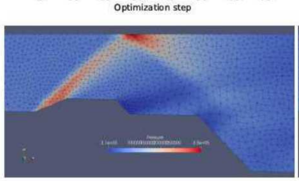


Figure: Initial design

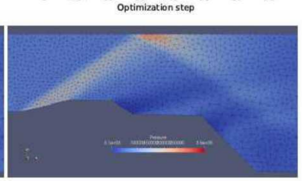
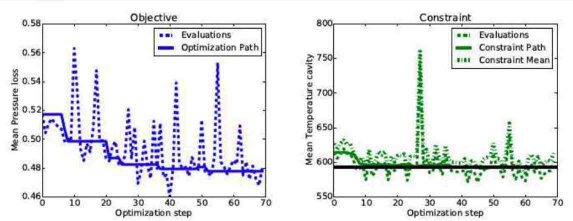


Figure: Final design

Stochastic MLMC OPTIMIZATION – Results



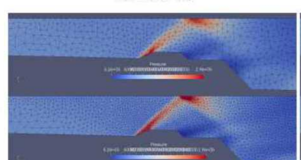


Figure: Initial design

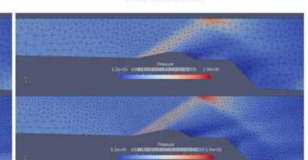
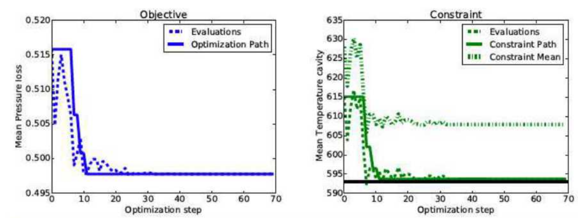


Figure: Final design

Stochastic MLPCE OPTIMIZATION – Results



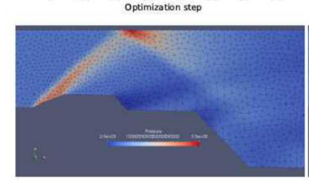


Figure: Initial design

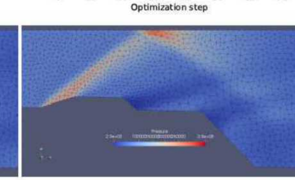


Figure: Final design

Each ML analysis 3x less expensive; MLPCE converges in ~10 iterations

Summary Remarks (ML-MF Opt/OUU)

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: higher performance estimators than MC; optimal resource allocations
  - CS / FT: exploit sparsity / low rank in δ; SC hierarchical interp; expand to model dimensions
  - *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 compr. sensing with exp order candidates; ML sparse grids with level / index set candidates
  - Achieve more rapid convergence (sufficient regularity, moderate dimensionality)

Multilevel-Multifidelity Opt / OUU: move beyond common bi-fidelity

- By exploiting richer model ensemble, computational effort can be pushed down the hierarchy, supporting case of only a handful of HF fine-grid evaluations
- Recursive TRMM looks promising relative to current multilevel approaches (MG/Opt)
- Multiple levels and multiple forms in simulation, UQ, or both