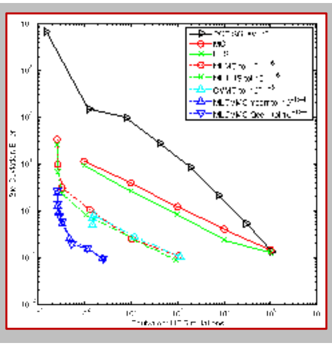
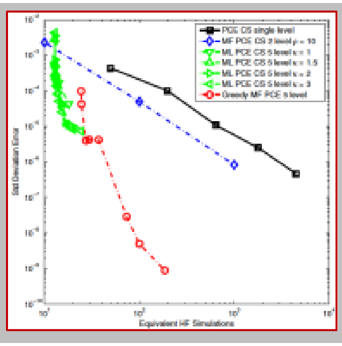
SAND2022-4390C

Exceptional service in the national interest

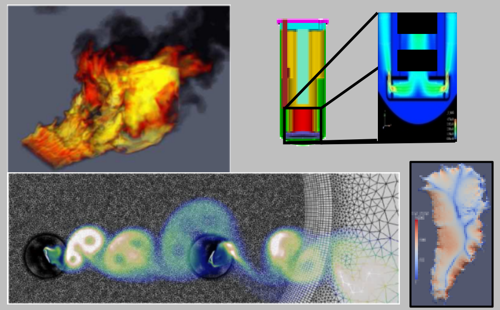
Sandia National Laboratories





DAKOTA

Explore and predict with confidence



All-at-Once (and Bi-Level) Model Tuning for Multifidelity Sampling

Michael S. Eldred[1], Gianluca Geraci[1] , Teresa Portone[1], Alex Gorodetsky[2], John Jakeman[1]

[1]*Optimization & Uncertainty Quantification Dept, Center for Computing Research, Sandia National Laboratories, Albuquerque NM*

[2]*Aerospace Engineering Department, University of Michigan, Ann Arbor MI*

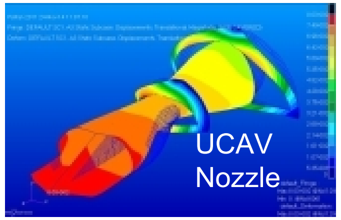
U.S. DEPARTMENT OF ENERGY

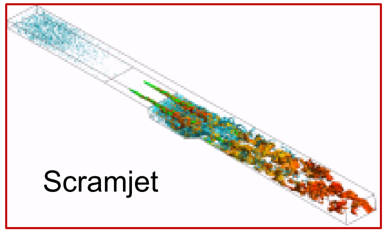
NNSA National Nuclear Security Administration

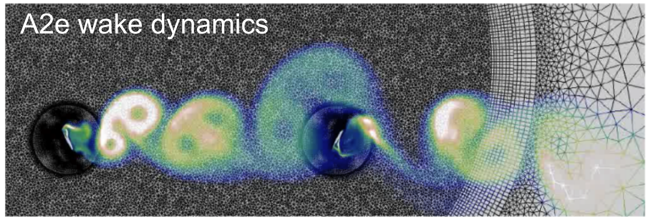
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

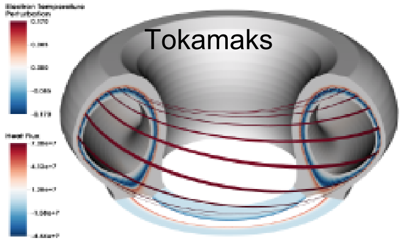
Multifidelity Methods: Sampling UQ, Surrogate UQ, OUU

2018/2019:

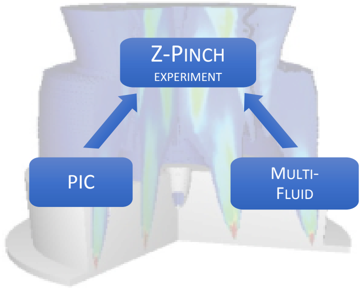


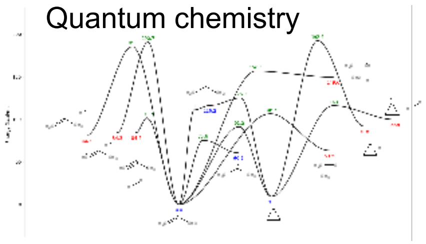


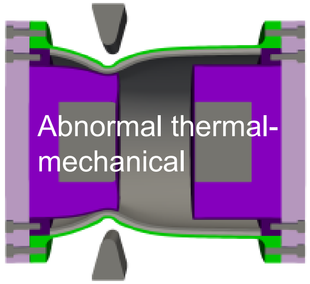


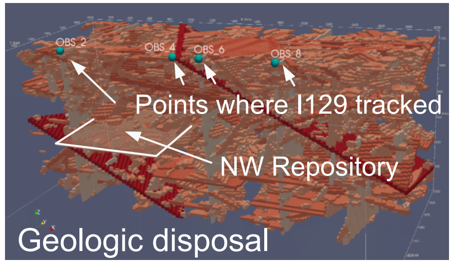


2020/2021:



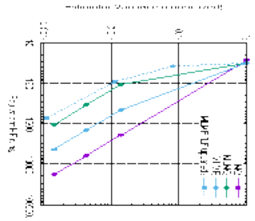


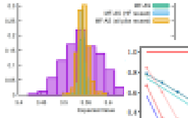


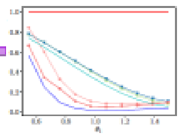


Monte Carlo UQ Methods

- *Production:* optimal resource allocation for multilevel, multifidelity, combined (DARPA EQUiPS, Wind, Cardiovascular)
- *Emerging:* active dimensions (LDRD, SciDAC), generalized fmwk for approx control variates (ASC V&V), goal orientation (rare events), hybrid methods for GSA
- *On the horizon:* control of time avg; model tuning / selection (LDRD)

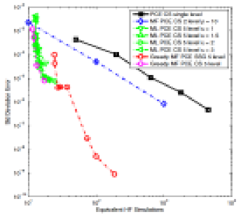


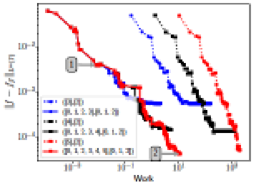




Surrogate UQ Methods (PCE, SC)

- *Production (v6.10+):* ML PCE w/ projection & regression; ML SC w/ nodal/hierarchical interp; greedy ML adaptation (DARPA SEQUOIA), multilevel fn train (ASC V&V)
- *Emerging:* multi-index stochastic collocation; multiphysics/multiscale integration (ASC V&V); new surrogates (GP, ROM, NN) w/ error mgmt. fmwk (LDRD, SciDAC); learning latent variable relationships (MFNets, LDRD)
- *On the horizon:* unification of surrogate + sampling approaches (LDRD)

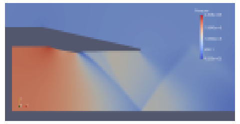




Optimization Under Uncertainty

- *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 Scramjet)
  - SNOWPAC (w/ MIT, TUM) with goal-oriented MLMC error estimates
- *On the horizon:* Gaussian process-based approaches: multifidelity EGO; Optimal experimental design (OED)



Non-hierarchical methods, alternative surrogates, goal-orientation, shared subspaces for dissimilar models

Key mission feedbacks

Multilevel performance on elliptic model PDEs is compelling, but does not accurately represent Sandia mission areas

- Extensions for complex multidimensional hierarchies → *multi-index collocation, multiphysics / multiscale*
- Investments in non-hierarchical MF methods → *ACV and MFNets*

Popular MF approaches neglect important practicalities

- "Oracle" correlations assumed → *iterated versions of MFMC, ACV*
- Imperfect data → *embedded cross validation*
- Dissimilar parameterizations → *shared subspaces*
- Free hyper-parameters → *model tuning*
- Stochastic simulation, simulation/surrogate error estimation → *extended error management framework*
- Ensemble management → *integration with HPC workflow managers, R&D in ensemble AMT*

MF methods most often utilize a fixed model ensemble determined by expert judgment

- Experts are often inaccurate in this context
  - SMEs from a physics discipline often have high predictivity standards and tend to over-estimate the LF accuracy required
- Leads to non-optimal correlation / cost trade-off and sub-optimal MF UQ

→ Initial explorations of hyper-parameter model tuning, within the context of particular estimators (ACV, MFMC, ...)

Background: paired ML/MF sampling methods of interest

Multilevel Monte Carlo

$$\mathbb{E}\left[Q_M^{\mathrm{HF}}\right] = \mathbb{E}\left[Q_{M_0}^{\mathrm{HF}}\right] + \sum_{\ell=1}^{L} \mathbb{E}\left[Q_{M_\ell}^{\mathrm{HF}} - Q_{M_{\ell-1}}^{\mathrm{HF}}\right]$$

$$\hat{Q}_M^{\mathrm{HF,ML}} = \sum_{\ell=0}^{L} \hat{Y}_{\ell,N_\ell}^{\mathrm{HF,MC}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} Y_\ell^{\mathrm{HF},(i)}$$

Minimize cost s.t. error balance:

$$N_\ell^{\mathrm{HF}} = \frac{2}{\varepsilon^2} \left[\sum_{k=0}^{L} \left(\mathbb{V}ar\left(Y_k^{\mathrm{HF}}\right) \mathcal{C}_k^{\mathrm{HF}} \right)^{1/2} \right] \sqrt{\frac{\mathbb{V}ar\left(Y_\ell^{\mathrm{HF}}\right)}{\mathcal{C}_\ell^{\mathrm{HF}}}}$$

M. Giles, "Multilevel Monte Carlo path simulation," 2008.

Control Variate Monte Carlo

$$Q_M^{\mathrm{HF,CV}} = Q_M^{\mathrm{HF}} + \alpha \left(Q_M^{\mathrm{LF}} - \mathbb{E}\left[Q_M^{\mathrm{LF}}\right] \right)$$

Classical control variate:

$$\alpha = -\rho \frac{\mathbb{V}ar^{1/2}\left(Q_M^{\mathrm{HF}}\right)}{\mathbb{V}ar^{1/2}\left(Q_M^{\mathrm{LF}}\right)}$$

LF oversample ratio:

$$r = \sqrt{\frac{\rho^2}{1-\rho^2} w}$$

Pasupathy et al, 2012; Ng and Willcox, 2014.

Multilevel-Control Variate Monte Carlo

$$\mathbb{E}\left[Q_M^{\mathrm{HF}}\right] = \mathbb{E}\left[Q_{M_0}^{\mathrm{HF}}\right] + \sum_{\ell=1}^{L} \mathbb{E}\left[Q_{M_\ell}^{\mathrm{HF}} - Q_{M_{\ell-1}}^{\mathrm{HF}}\right]$$

$$\simeq \sum_{\ell=0}^{L} \left(\hat{Y}_{M_\ell}^{\mathrm{HF,MC}} + \alpha_\ell \left(\hat{Y}_{M_\ell}^{\mathrm{LF,MC}} - \hat{\mathbb{E}}[Y_{M_\ell}^{\mathrm{LF}}] \right) \right)$$

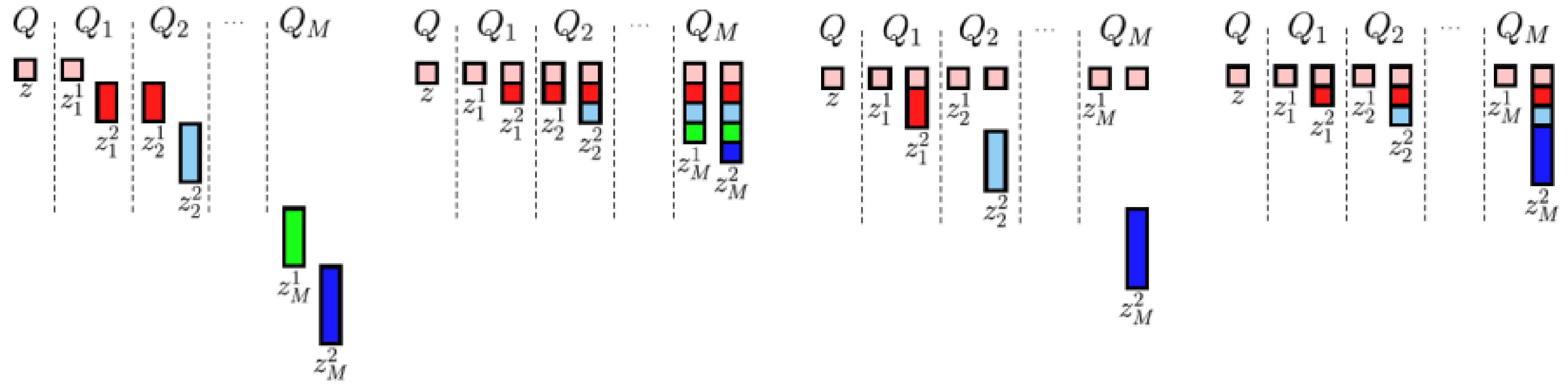
$$\begin{cases} r_\ell^\star = \sqrt{\dfrac{\rho_\ell^2}{1-\rho_\ell^2} w_\ell}, \\ N_\ell^{\mathrm{HF},\star} = \dfrac{2}{\varepsilon^2} \left[\displaystyle\sum_{k=0}^{L_{\mathrm{HF}}} \left(\dfrac{\mathbb{V}ar\left(Y_\ell^{\mathrm{HF}}\right) \mathcal{C}_\ell^{\mathrm{HF}}}{1-\rho_\ell^2} \right)^{1/2} \Lambda_\ell \right] \sqrt{\left(1-\rho_\ell^2\right) \dfrac{\mathbb{V}ar\left(Y_\ell^{\mathrm{HF}}\right)}{\mathcal{C}_\ell^{\mathrm{HF}}}} \end{cases}$$

G. Geraci, E., G. Iaccarino, "A multifidelity control variate approach for the multilevel Monte Carlo technique," CTR Res Briefs 2015.

Background: ensemble sampling methods of interest

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

Sample set definitions



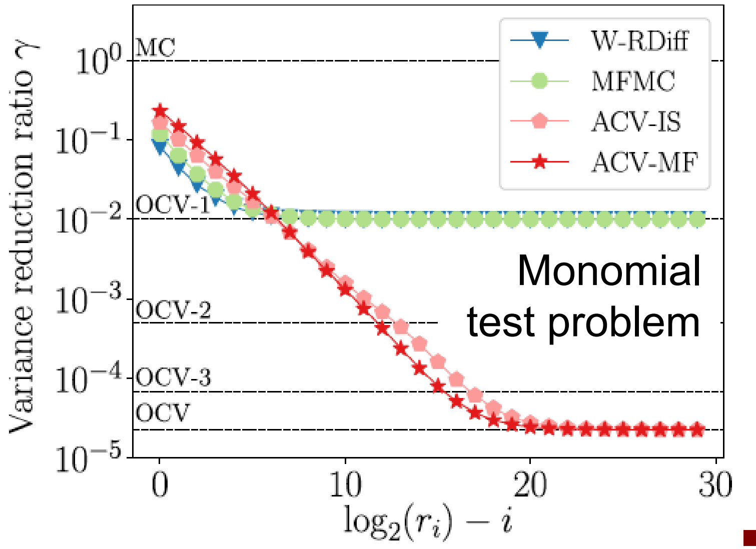
(a) W-RDiff sampling strat. (b) MFMC sampling strat. (c) ACV-IS sampling strat. (d) ACV-MF sampling strat.

Theoretical perf. bounds for recursive vs. non-recursive

- Recursive limited by variance reduction of perfect μ_1 (OCV-1)
- Non-recursive can exploit potential gap between OCV-1 and OCV

Methods minimize estimator variance over number of truth evals *N* and approximation oversample ratios ***r***

- MFMC has closed form for optimal ***r****,*N** (given ordered/reordered models)
- ACV solves numerically for ***r****,*N** (does not require ordering)



Gorodetsky, Geraci, E., Jakeman, "A generalized approximate control variate framework for multifidelity uncertainty quantification," JCP 408 (2020)

Iterated MFMC

Iterated ACV

Initialize: select a small shared pilot sample $N^{(0)}$ expected to under-shoot the optimal profile

1) Sample all models

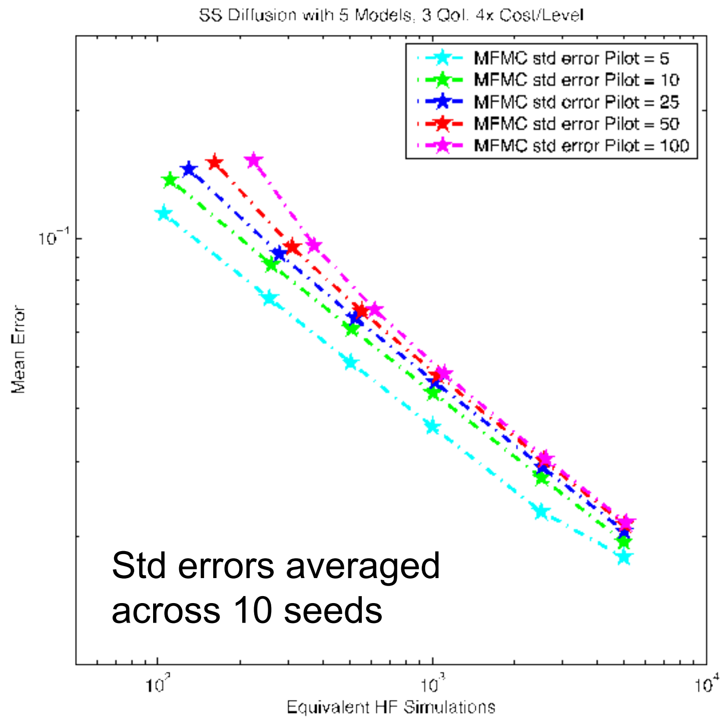
Iterated MFMC:

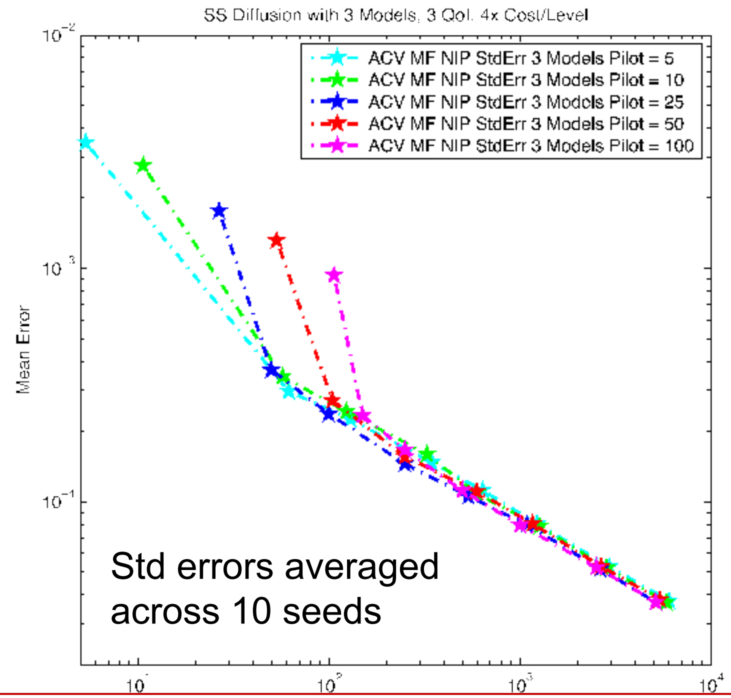
2) $N^{(i)}$ shared samples → Estimate $\rho^2_{LH}{}^{(i)}$ → Estimate $r^{(i)}$
3) Estimate $N^{(i+1)}$ using prescribed { budget C || tolerance ε }
4) Compute one-sided ΔN for shared samples from $N^{(i)}$ to $N^{(i+1)}$
   A. Optional: apply under-relaxation factor γ
   B. If non-zero increment, advance (i) and return to 1)

Iterated ACV:

1) $N^{(i)}$ shared samples → $Cov_{LL}{}^{(i)}$, $Cov_{LH}{}^{(i)}$ ("C", "c") → opt. solver → r^*, N^*
2) Compute one-sided ΔN for shared samples from $N^{(i)}$ to N^*
   A. Optional: apply under-relaxation factor γ
   B. If non-zero increment, advance (i) and return to 1)

Finalize: apply r* for LF eval increments, estimate α → apply controls to compute final expectation(s)





Performance degradation from pilot over-estimation is clearly evident
- Analytic r* reduces numerical burden but also limits flexibility

Performance degradation from pilot over-estimation is *not* significant
- ACV-MF demonstrates greater flexibility / resilience: locates near-optimal solutions that incorporate large pilots
- Starting pts on left are for budget = pilot (moves quickly from MC to ACV)

Model Tuning Approaches: All-At-Once and Bi-Level

Model tuning performed to maximize performance *of a particular estimator* (e.g., ACV-MF) using *tunable hyper-parameters* associated with one or more low-fidelity models (HF reference is immutable)

AAO optimization (in Python): hyper-parameters integrate as additional decision vars for minimizing EstVar

$$\arg\min_{\theta,\mathbf{r},N} \frac{Var[Q]}{N}(1 - R^2(\theta,\mathbf{r})) \quad s.t. \quad N\left(w + \sum_{i=1}^{M} w_i(\theta) r_i\right) \leq C$$

- Potential for greater efficiency: one integrated optimization solve
  - Need to emulate lower-level $\rho(\theta), w(\theta)$ to avoid expensive re-estimation at every change in θ

Bi-level optimization (in Dakota): inner loop optimization solve for each outer loop θ iterate

$$\arg\min_{\theta} \left[\arg\min_{\mathbf{r},N} \frac{Var[Q]}{N}(1 - R^2(\theta,\mathbf{r})) \quad s.t. \quad N\left(w + \sum_{i=1}^{M} w_i(\theta) r_i\right) \leq C\right]$$

- For nested numerical solution, outer loop must now contend with inner-loop solver noise
  - Noise and expense can be mitigated using pilot projections, with some loss of accuracy
- Can choose to emulate at a higher level, requiring fewer emulators (e.g. EGO, TRMM to min *EstVar*(θ)*)
  - Plug and play with surrogate-based methods (EGO, TRMM), MINLP, etc.
- Note: for analytic cases (e.g., MLMC, CVMC, standard MFMC), AAO collapses to single level argmin_{θ}

Exploration of model tuning for a parameterized model problem

Sandia National Laboratories

Tunable model problem (from JCP paper on ACV*)

- 1 parameter is tunable: θ_1
- 2 parameters are fixed: $\theta = \pi/2$, $\theta_2 = \pi/6$

Model Definitions

$$Q = \sqrt{11}y^5$$

$$Q_1 = \sqrt{7}\left(\cos\theta_1 x^3 + \sin\theta_1 y^3\right)$$

$$Q_2 = \sqrt{3}\left(\frac{\sqrt{3}}{2}x + \frac{1}{2}y\right), \quad \text{where} \quad x, y \sim \mathcal{U}(-1,1)$$

Correlations (analytic form available but not used in experiments)

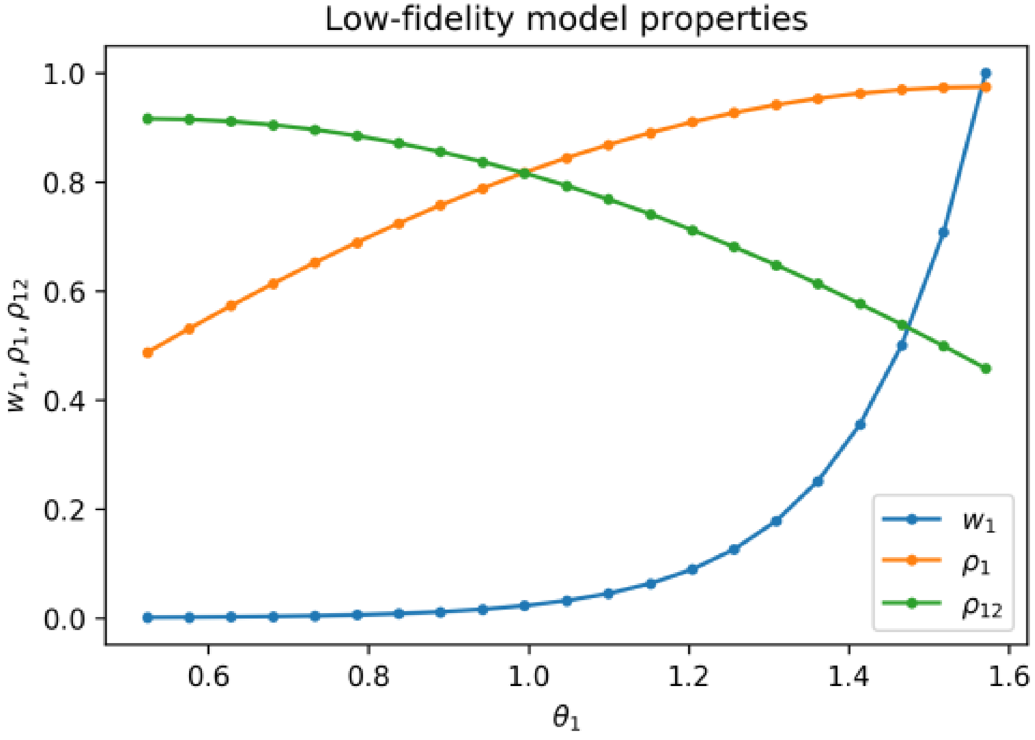
	Q	Q_1	Q_2
Q	1	$\frac{\sqrt{77}}{9}\sin\theta_1$	$\frac{\sqrt{33}}{14}$
Q_1	*sym*	1	$\frac{\sqrt{21}}{10}\left(\sin\theta_1 + \sqrt{3}\cos\theta_1\right)$
Q_2	*sym*	*sym*	1

θ_1 controls:

- Correlations among models ρ_1 and ρ_{12};
- Cost of evaluating Q_1 according to the cost law

$$\log w_1 = \log w_2 + \frac{\log w_2 - \log w}{\theta_2 - \theta}(\theta_1 - \theta_2)$$

$$\text{with} \quad w = 1 \quad \text{and} \quad w_2 = 10^{-3}$$



* Gorodetsky, Geraci, E., Jakeman, "A generalized approximate control variate framework for multifidelity uncertainty quantification," JCP 408 (2020)

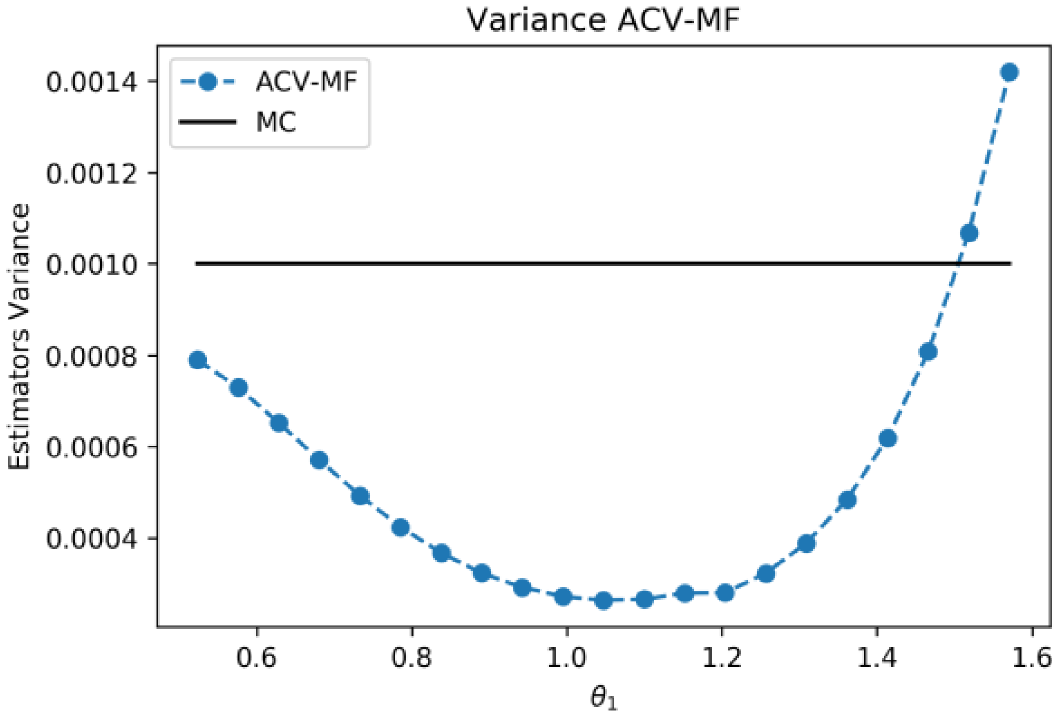
Tuning for parameterized model problem (Cont.)

Model tuning performed within the context of a particular estimator (here, ACV-MF)

$$\underset{\theta_1, N, r_1, r_2}{\text{argmin}} \ \frac{1}{N}\left(1 - R^2_{ACV-MF}(\theta_1, r_1, r_2)\right) \quad \text{s.t.} \quad \mathcal{C}^{tot} = N\left(w + \sum_{i=1}^{2} w_i r_i\right) \leq \mathcal{C}_{target} = 1000$$

AAO optimization (in Python):

- For ACV (and numerical MFMC), hyper-parameters integrate as additional decision vars for minimizing estimator variance



Mid-fidelity model (Q_1) is tuned for ACV at ~ midpoint $\theta_1^* = \pi/3$

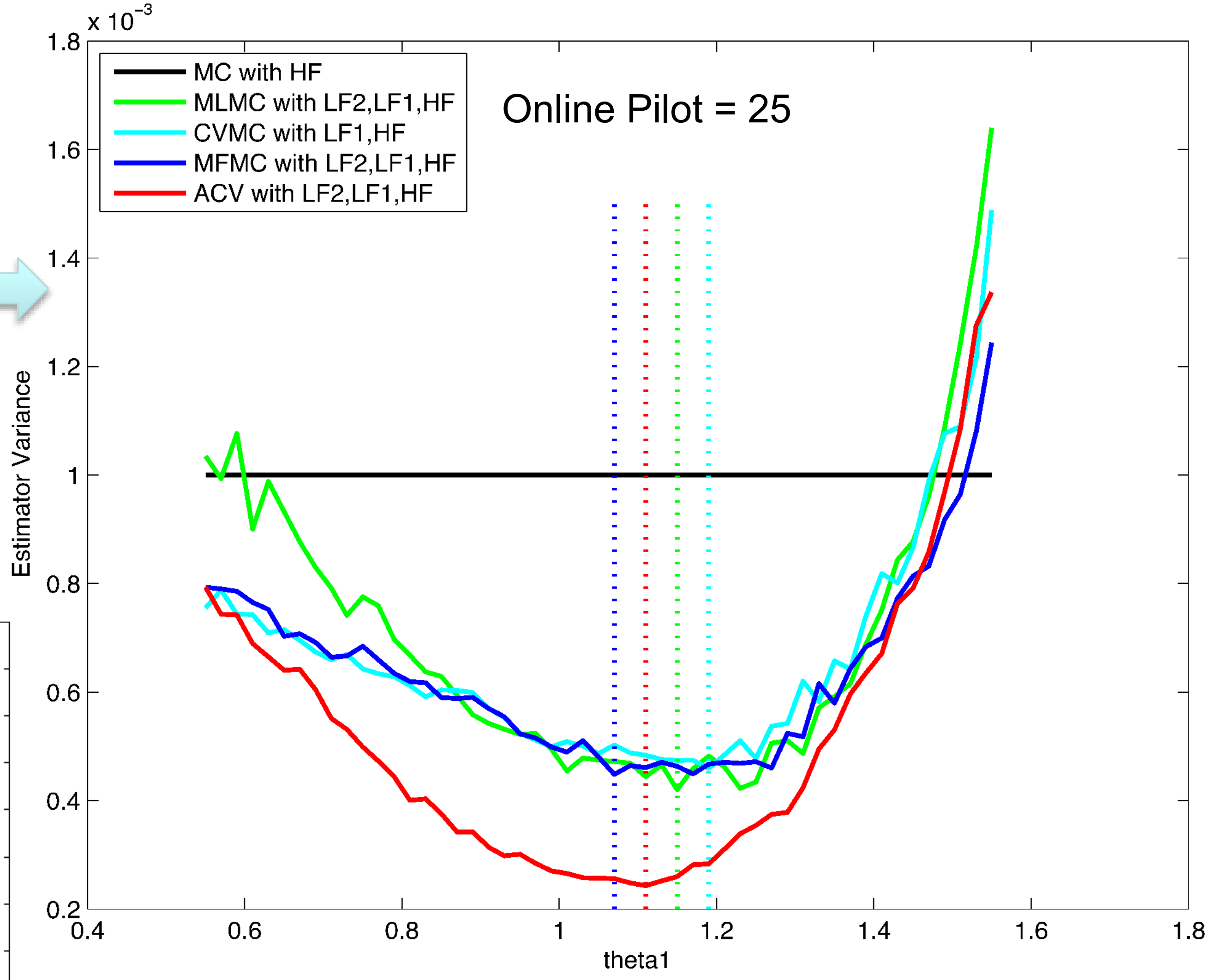
Tuning for parameterized model problem (Cont.)

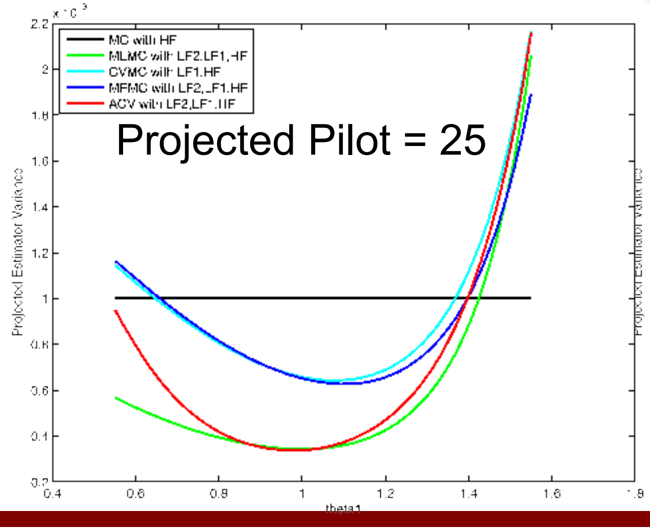
Sandia National Laboratories

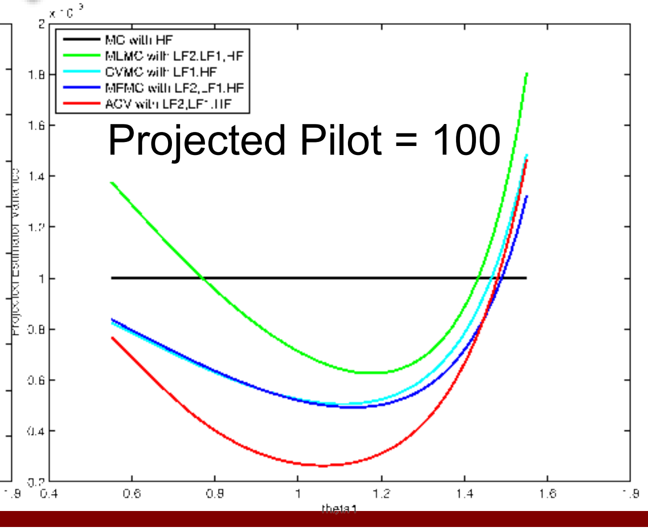
Bi-level optimization (in Dakota):

$$\arg\min_{\theta} \left[\arg\min_{\mathbf{r},N} \quad \frac{Var[Q]}{N}(1 - R^2(\theta, \mathbf{r})) \quad s.t. \quad N\left(w + \sum_{i=1}^{M} w_i(\theta) r_i\right) \leq C \right]$$

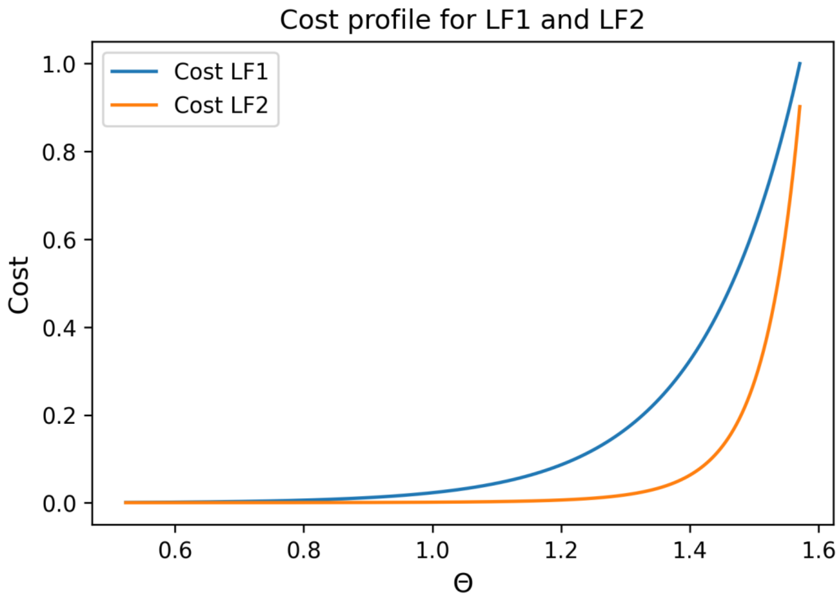
- For converged iteration (right), we observe some inner -loop solver noise
  - TO DO: explore additional solution modes offline / online max_iterations = 0
- For expensive problems, can tune based on pilot projection (bypassing iteration to convergence)
  - Eliminates some (but not all) sources of noise







Global Optimization of multiple hyper-parameters



Add cost model w_2 for LF2(θ_2): introduce δ, γ

$log(w) = log(w_{low}) - log(w_{low} / w_{high}) (\theta - \theta_{low})^{\delta} / \theta_{range}$

where $w_{low} = .001 * \gamma$, $w_{high} = 1. * \gamma$, $\theta_{low} = \pi / 6$, $\theta_{range} = \pi / 2 - \pi / 6$

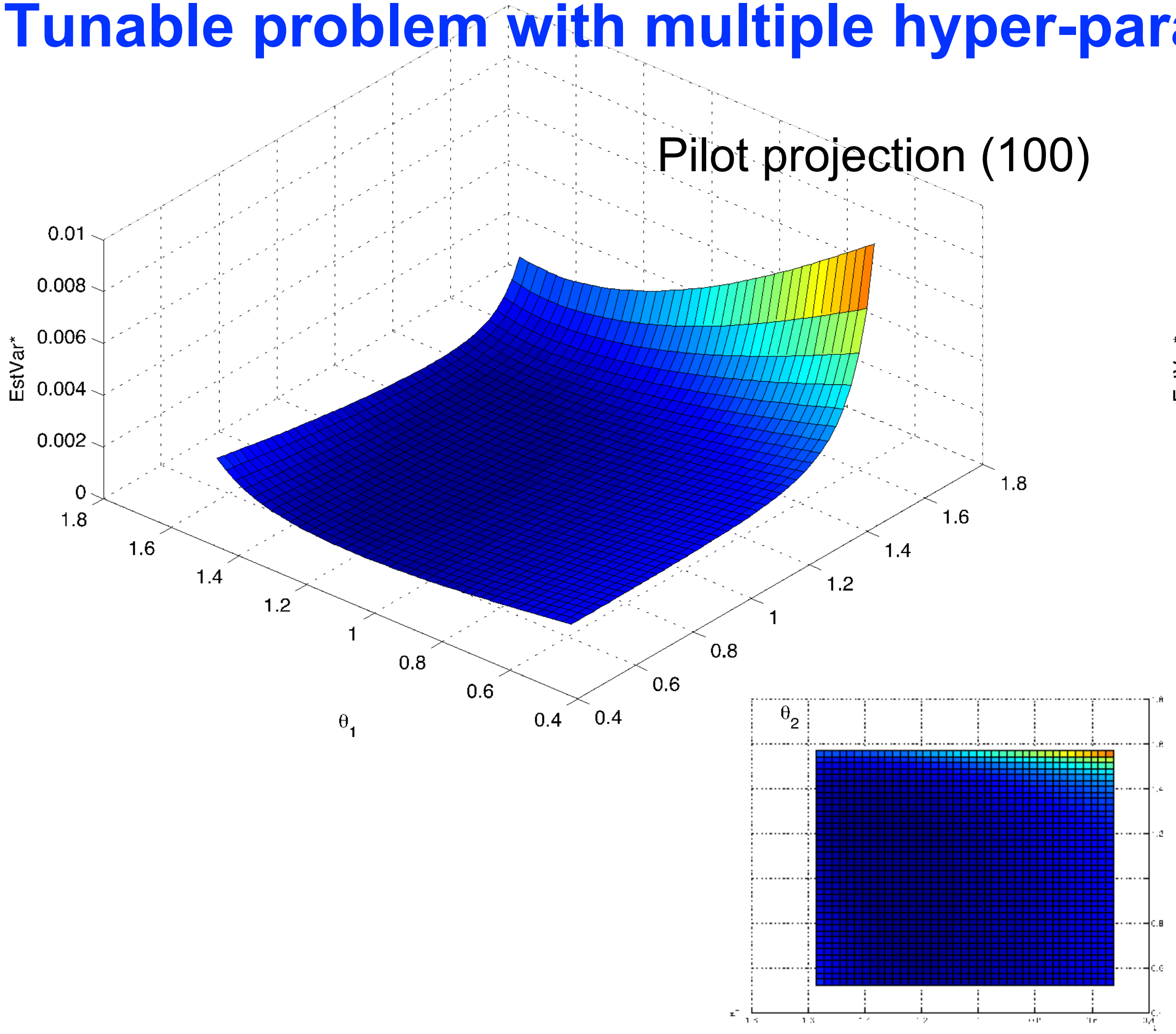
For w_1, $\delta = \gamma = 1$ (reproduces previous cost model)
For w_2, $\delta = 2.5$, $\gamma = 0.55$

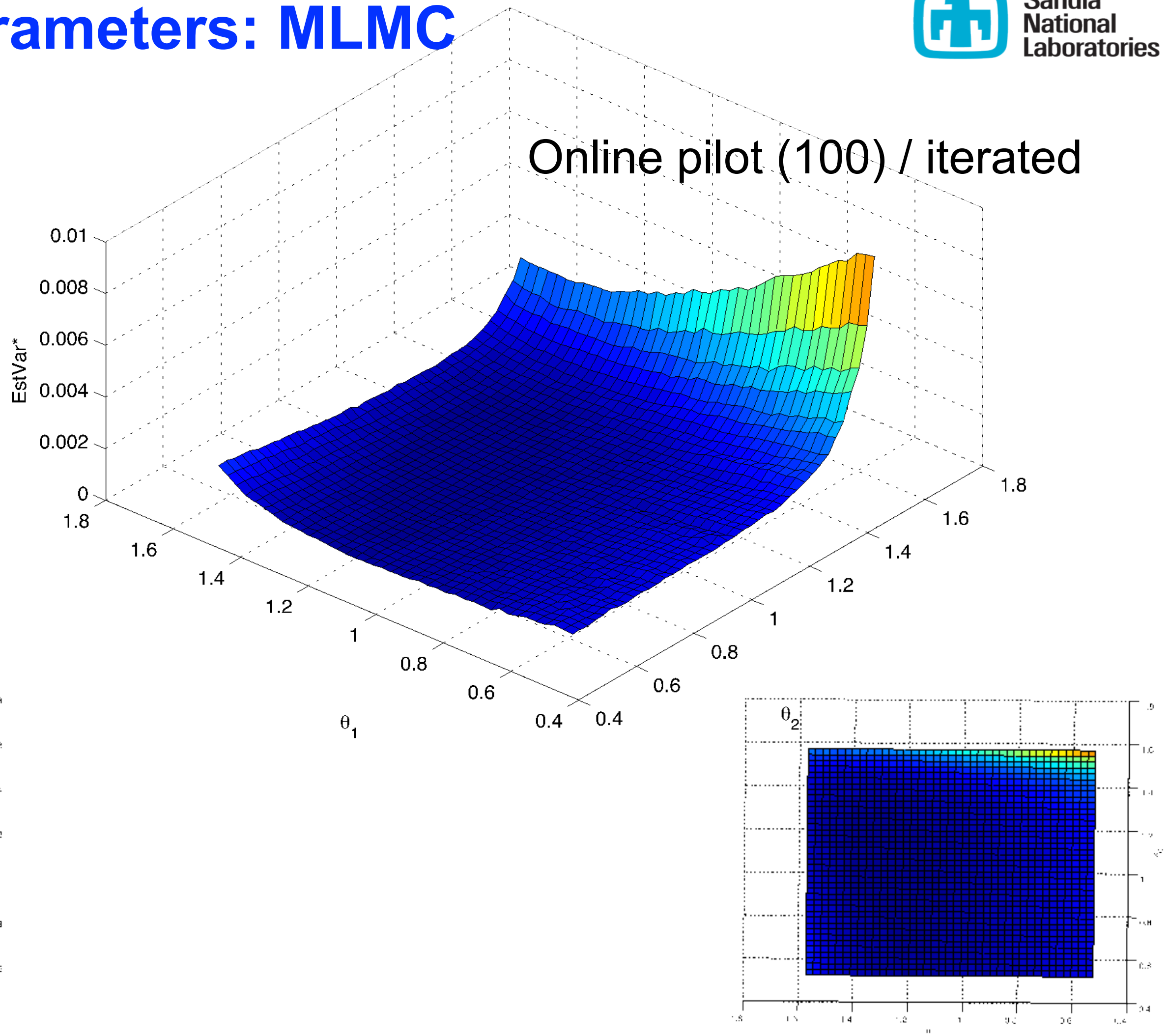
For a modest number of hyper-parameters, we have explored surrogate-based approaches

- Efficient Global Optimization (EGO)
- First-order trust region model management (TRMM, *aka* surrogate-based local optimization)

With care in declaring the relevant θ subset per model, we can leverage Dakota's evaluation cache and reuse pilot samples over θ (e.g., all HF pilots can be reused)

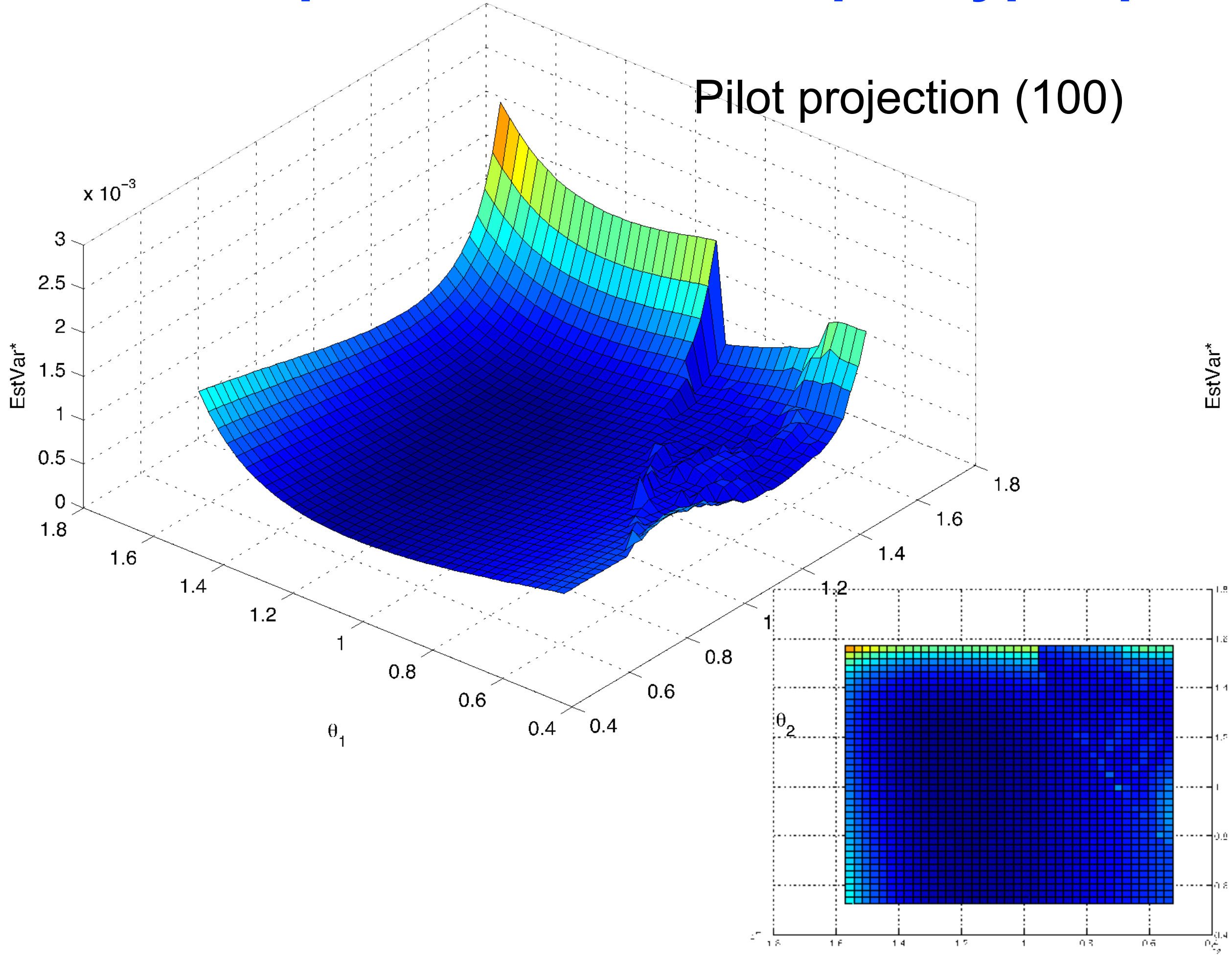
Tunable problem with multiple hyper-parameters: MLMC

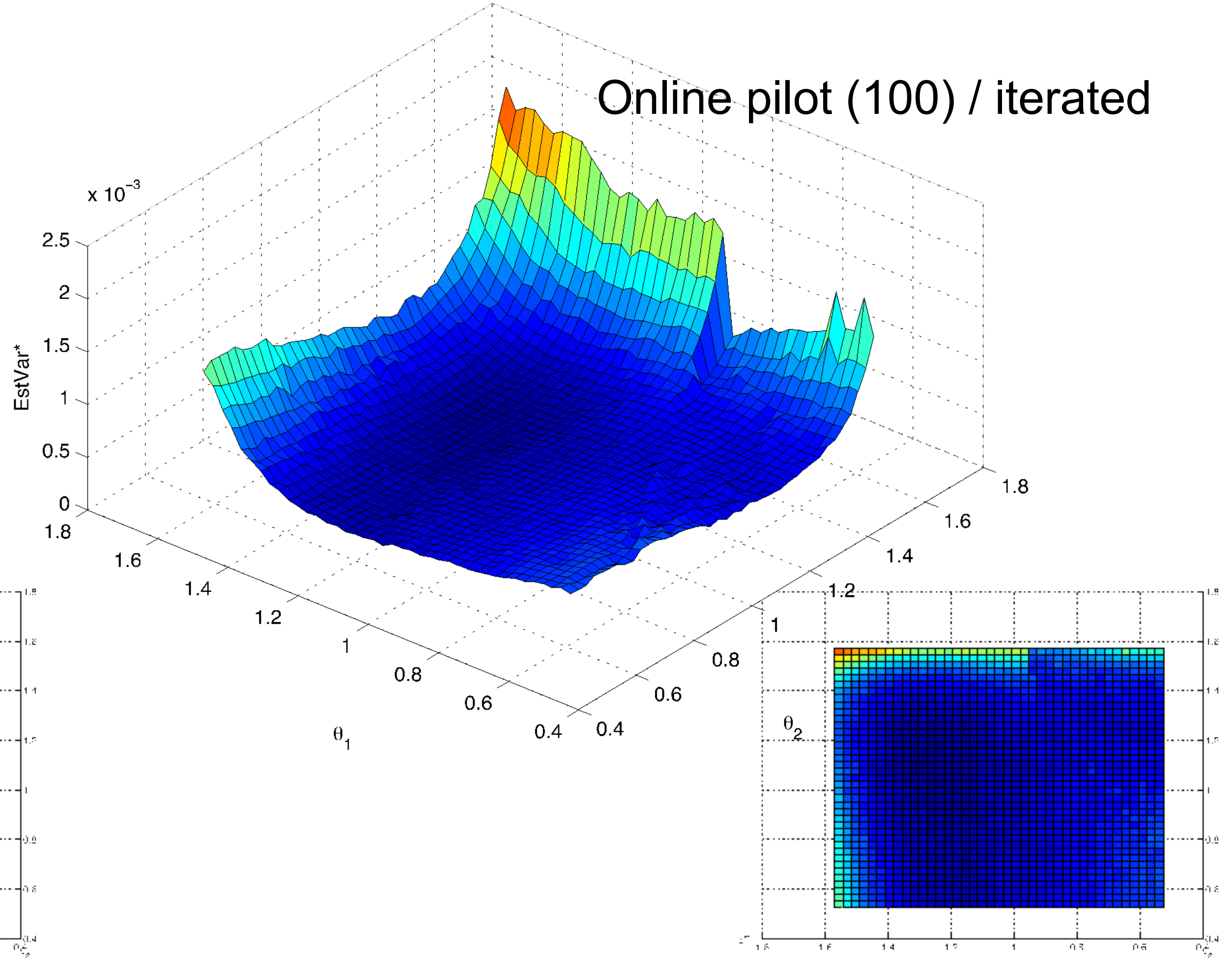




Less robust: significant performance loss for non-optimal theta (up to EstVar* = 0.01)

Tunable problem with multiple hyper-parameters: MFMC



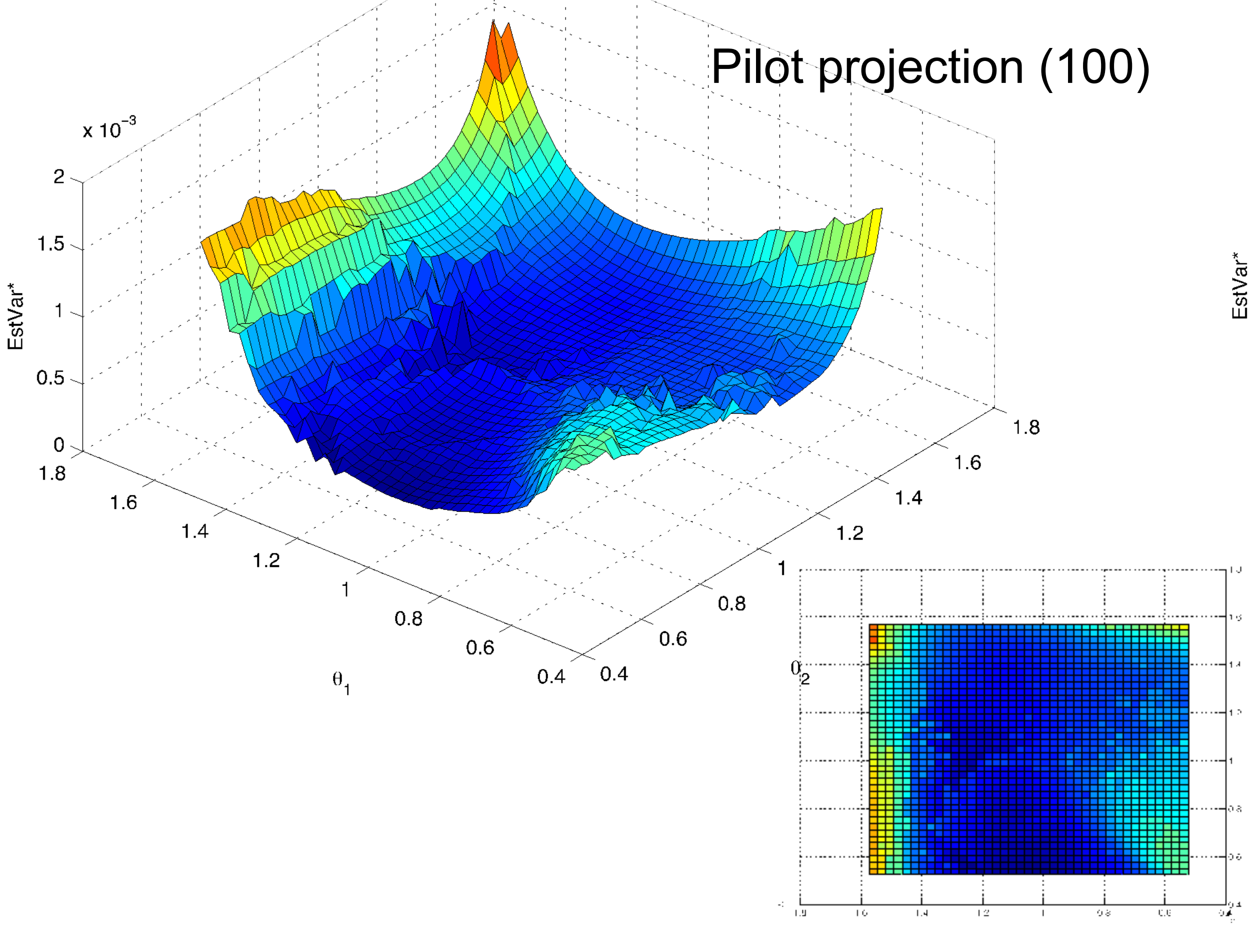


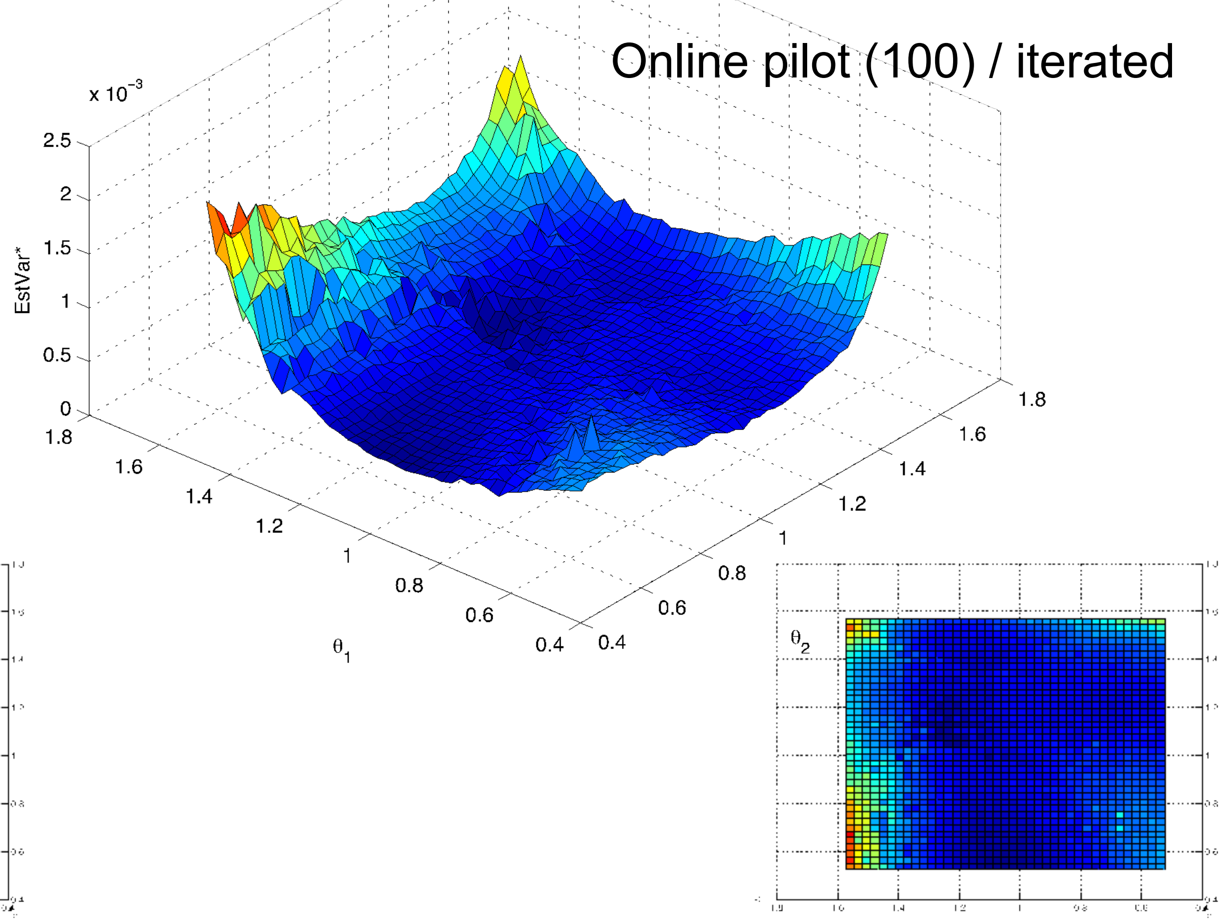
More consistent performance but susceptible to model mis-ordering:

- Dakota mitigates with switch to reordered numerical solve w/ pyramid constraint enforcement
- While noisier, performance relative to analytic looks promising
- Excepting discontinuity, generally unimodal

Tunable problem with multiple hyper-parameters: ACV

Sandia National Laboratories





Larger region of good performance and insensitive to model ordering:

- Multimodal: two LF1,LF2 configurations achieve best performance overall
  - Generally an algorithmic strength (as for adapting to over-estimated pilot), but a challenge for optimizers

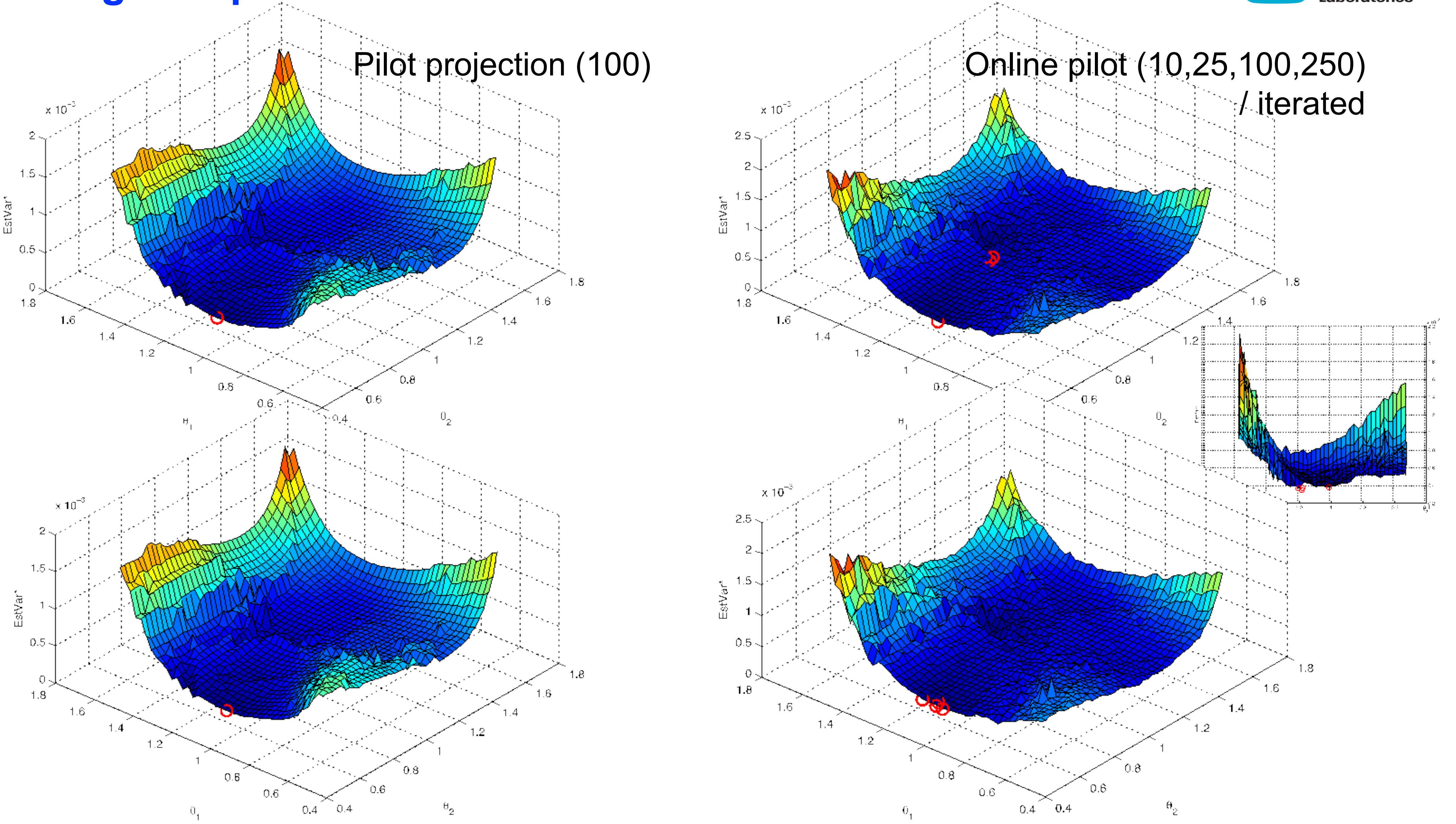
Bi-Level Surrogate Optimization results: ACV

Pilot projection (100)

Online pilot (10,25,100,250) / iterated

EGO

TRMM



TO DO: some run stats (iteration counts)

Current Directions: Multiple Deployments

- Stochastic simulations: (previous talk)
  - Turbulent flows/combustion with LES (H. Najm, C. Safta)
  - Subsurface transport for repositories (T. Portone, L. Swiler)
  - Radiation transport with PIC codes for HEDP (B. Reuter, G. Geraci, J. Jakeman)
- Spatial / temporal resolution
  - EDL (NASA: G. Bomarito, J. Warner, M. Thompson) (upcoming talk)
  - Thermal batteries (T. Portone, M. Eldred)
    - Two-dimensional model hierarchy: MLMC (1D slice), CVMC (1d slice), MLCV MC (2d), MFMC (flatten), ACV (flatten)
      - *LF = 2D reduced physics mode*; *HF = 2D full physics mode*
      - *Same coarse/medium/fine spatial resolutions* selected for both modes
      - Fine temporal resolution settings used for HF, *coarse temporal resolution settings for LF (tuning targets)*
- Data-driven surrogates with hyper-parameters: ROM, NN

Summary Observations

Realistic deployments of multifidelity methods encounter a variety of challenges

- Here we target the challenge of _optimally configuring multiple LF models_, given one or more DOF that trade accuracy vs. cost

Model Tuning Approaches

- AAO Optimization (in Python): hyper-parameters become additional decision vars in $\text{argmin}_{r,N,\theta}$ EstVar
  - Solve 1 integration optimization problem; emulate lower-level $\rho(\theta), w(\theta)$; avoids optimizing on top of solver noise
- Bi-level optimization (in Dakota): argmin_{θ} [$\text{argmin}_{r,N}$ EstVar]
  - Plug-and-play with surrogate-based optimizers to mitigate solver noise; either low-level or high-level emulators
  - AAO collapses to this in many cases (analytic allocations with ML, CV, MLCV, MF)
  - Implementation details: online cost recovery, solution modes, evaluation cache, bypass LF increments if only need EstVar
- Relative performance TBD

Numerical Experiments

- Tunable problem 1D (θ_1): ACV > MFMC > CVMC > MLMC
- Tunable problem 2D (θ_1, θ_2): ACV > MFMC > MLMC
  - Robustness obtained from numerical solves: can better adapt to pilot over-estimation, model sequencing
- Production thermal battery studies in flight, with feedbacks per below

Next steps

- Feedback from expensive deployments: streamline approaches, maximize data reuse, prune convenience synchronizations
- More thorough exploration of AAO benefits, when admissible

Extra

Background: multifidelity Monte Carlo (MFMC)

Sandia National Laboratories

Correlations
Costs → Optimal LF over-sample

$$r_i^* = \sqrt{\frac{w_1(\rho_{1,i}^2 - \rho_{1,i+1}^2)}{w_i(1-\rho_{1,2}^2)}}$$

→ HF samples from budget

$$m_1^* = \frac{p}{\boldsymbol{w}^T \boldsymbol{r}^*}$$

Following ρ estimation, budget p exhausted
→ No iteration

$$\alpha_i^* = \frac{\rho_{1,i}\sigma_1}{\sigma_i}$$

→ Expectations from shared, refined

Background: approximate control variate (ACV)

C = covariance matrix among Q_i
c = covariance vector among Q_i and Q

$$\underline{\alpha}^{\text{ACV-IS}} = -\left[\mathbf{C} \circ \mathbf{F}^{(IS)}\right]^{-1}\left[\operatorname{diag}\left(\mathbf{F}^{(IS)}\right) \circ \mathbf{c}\right]$$

$$\mathbb{V}\text{ar}\left[\hat{Q}^{\text{ACV-IS}}(\underline{\alpha}^{\text{ACV-IS}})\right] = \frac{\mathbb{V}\text{ar}[Q]}{N}\left(1 - R^2_{\text{ACV-IS}}\right), \text{ where } R^2_{\text{ACV-IS}} = \boldsymbol{a}^T\left[\mathbf{C} \circ \mathbf{F}^{(IS)}\right]^{-1}\boldsymbol{a}$$

$$\underline{\alpha}^{\text{ACV-MF}} = -\left[\mathbf{C} \circ \mathbf{F}^{(MF)}\right]^{-1}\left[\operatorname{diag}\left(\mathbf{F}^{(MF)}\right) \circ \mathbf{c}\right],$$

$$\mathbb{V}\text{ar}\left[\hat{Q}^{\text{ACV-MF}}(\underline{\alpha}^{\text{ACV-MF}})\right] = \frac{\mathbb{V}\text{ar}[Q]}{N}\left(1 - R^2_{\text{ACV-MF}}\right), \text{ where } R^2_{\text{ACV-MF}} = \boldsymbol{a}^T\left[\mathbf{C} \circ \mathbf{F}^{(MF)}\right]^{-1}\boldsymbol{a}$$

$\boldsymbol{a} = \left[\operatorname{diag}\left(\boldsymbol{F}^{(IS)}\right) \circ \bar{\mathbf{c}}\right]$ *and* $\boldsymbol{F}^{(IS)} \in \mathbb{R}^{M \times M}$ *has elements*

$$\boldsymbol{F}^{(IS)}{}_{ij} = \begin{cases} \frac{r_i-1}{r_i}\frac{r_j-1}{r_j} & \text{if } i \neq j \\ \frac{r_i-1}{r_i} & \text{otherwise} \end{cases}.$$

$\boldsymbol{a} = \left[\operatorname{diag}\left(\boldsymbol{F}^{(MF)}\right) \circ \bar{\mathbf{c}}\right]$ *and* $\boldsymbol{F}^{(MF)} \in \mathbb{R}^{M \times M}$ *has elements*

$$\boldsymbol{F}^{(MF)}{}_{ij} = \begin{cases} \frac{\min(r_i,r_j)-1}{\min(r_i,r_j)} & \text{if } i \neq j \\ \frac{r_i-1}{r_i} & \text{otherwise} \end{cases}.$$

← Differs only in off-diagonal terms + sample sets

$$\min_{N,\underline{r},K,L} \log(J_{ACV}(N,\underline{r},K,L)) \quad \text{subject to } N\left(w + \sum_{i=1}^{M} w_i r_i\right) \leq C, \quad N \geq 1, \quad r_1 \geq 1$$

Optimal r*,N* w/i budget from ***C,c*** estimates → No iteration

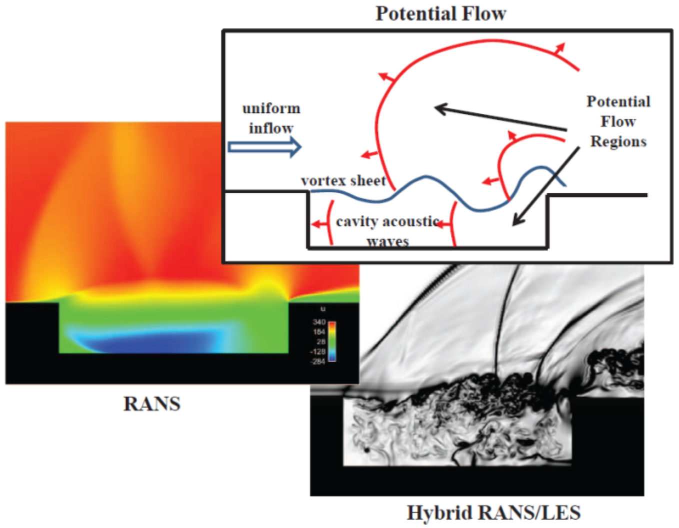
Peherstorfer, Willcox, Gunzburger, "Optimal Model Management For Multifidelity Monte Carlo Estimation", SISC, Vol. 38, No. 5.
Gorodetsky, Geraci, E., Jakeman, "A generalized approximate control variate framework for multifidelity uncertainty quantification," JCP 408 (2020)

Multiple Model Forms in UQ & Opt

Discrete model choices for simulation of same physics

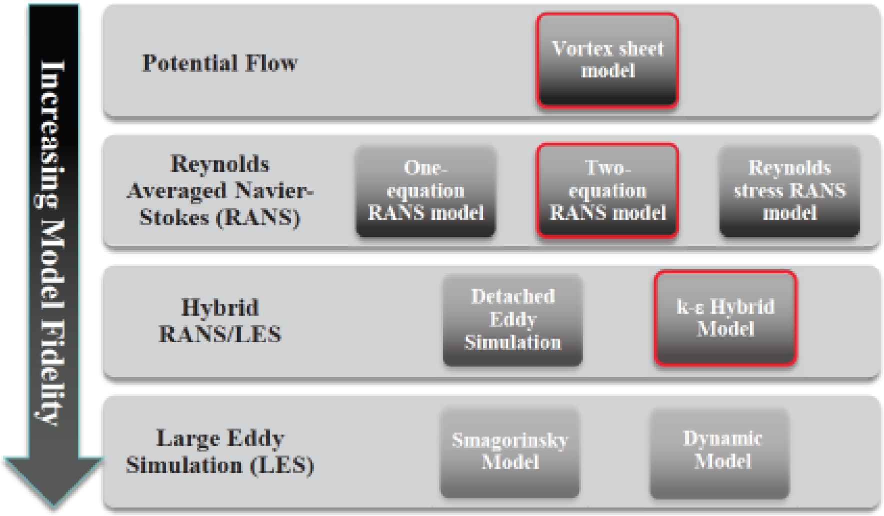
A clear hierarchy of fidelity (from low to high)

- 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 modeling options

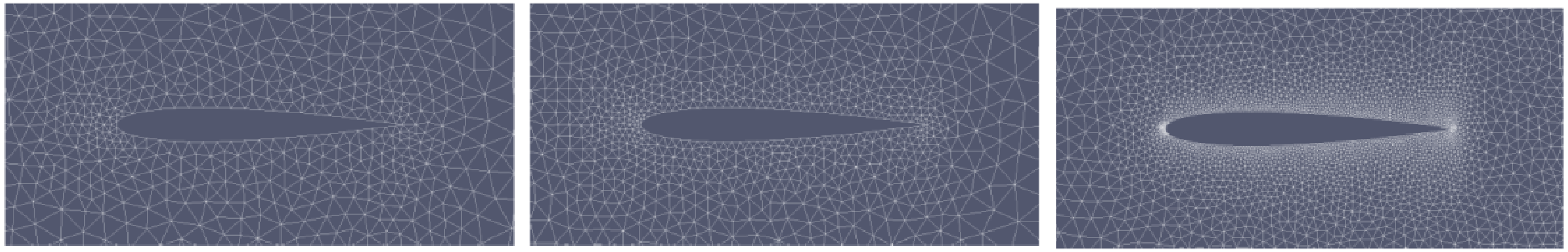
- *With data:* model selection, inadequacy characterization
    - Criteria: predictivity, discrepancy complexity
- *Without (adequate) data:* epistemic model form propagation
    - Intrusive, nonintrusive
- *In MF context:* correlation analysis, model tuning, ensemble selection



Discretization levels / resolution controls

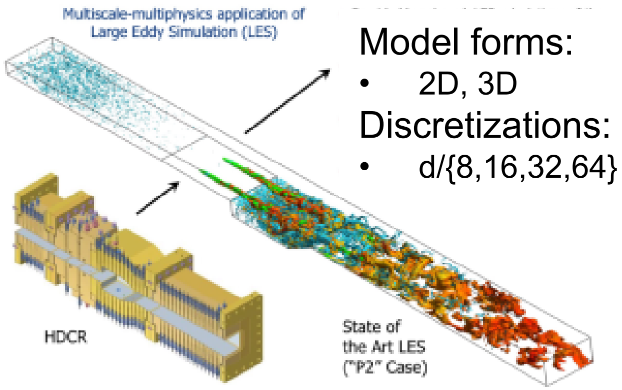
- Exploit special structure: discrepancy → 0 at order of spatial/temporal convergence

Combinations for multiphysics, multiscale



2018/2019 Deployments: ML, MF, MLMF Monte Carlo

Scramjet



Model forms:
- 2D, 3D

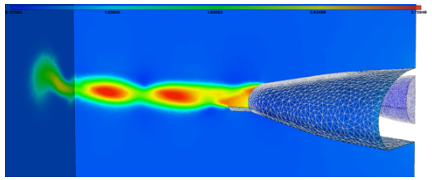
Discretizations:
- d/{8,16,32,64}

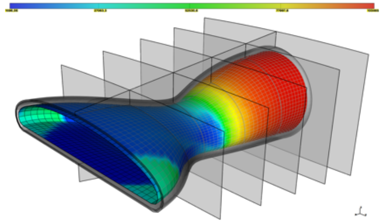
	$P_{0,mean}$	$P_{0,rms,mean}$	M_{mean}	TKE_{mean}	χ_{mean}
	P1				
$d/8$	4.02554e-03	1.90524e-06	1.99236e-02	3.34905e-07	4.24520e-03
$d/16$	4.03350e-07	7.77838e-08	6.68974e-05	1.74847e-08	4.40048e-05
	P1 updated				
$d/8$	4.05795e-03	1.90612e-06	1.60029e-02	7.53353e-07	9.41403e-04
$d/16$	2.85017e-04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

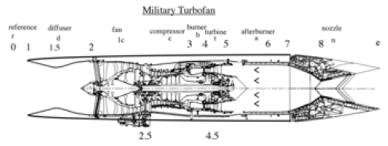
Table 2: Variance for the five QoIs of the P1 unit problem.

No variance decay for higher turbulence levels

UCAV Nozzle





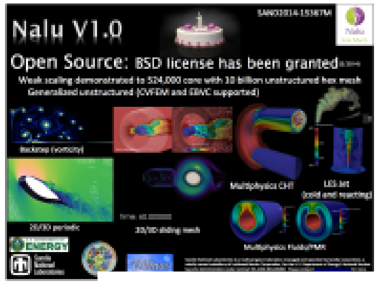


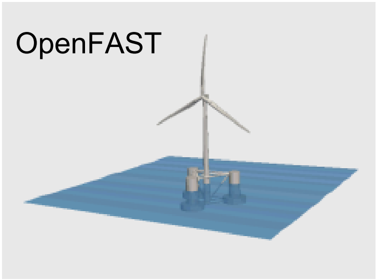
Non-predictive LF stress prior to reformulation

	LF		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

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

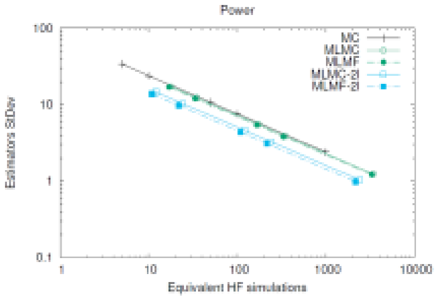
Wind

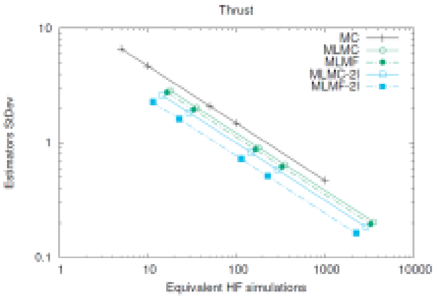




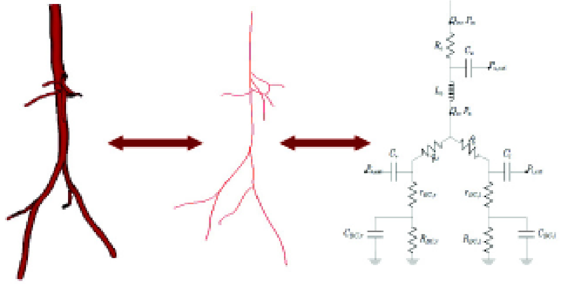
Nalu LES for Q_0 is too coarse with limited predictive value

Project basis for ML emulator-based inference to follow

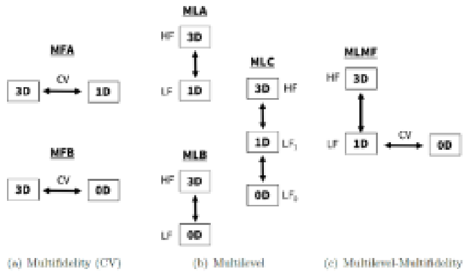




Cardiovascular



3D Model 1D Model 0D Model



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

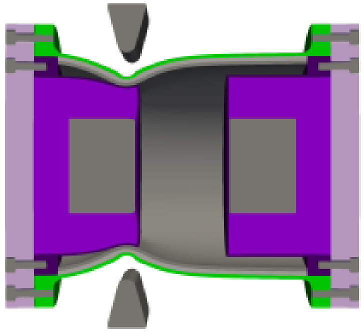
Recent Deployments: ML/MF Monte Carlo/Polynomial Chaos

Crash & Burn Multiphysics (ASC L2 Milestone)

Forward UQ w/ explicit (LF) + implicit (HF) SIERRA mechanics

- Multilevel MC across model resolutions for LF model
- Multifidelity MC with HF implicit + selection of most effective LF explicit

Successful demonstration of advanced UQ methods, integrated alongside emerging ASC workflows for multiphysics simulation

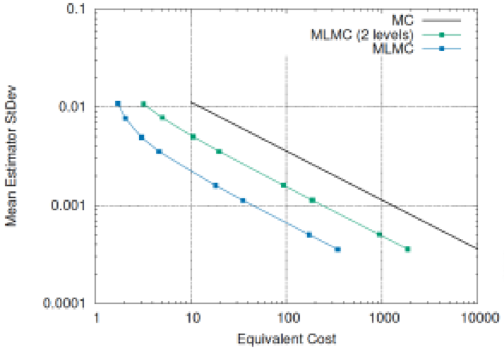


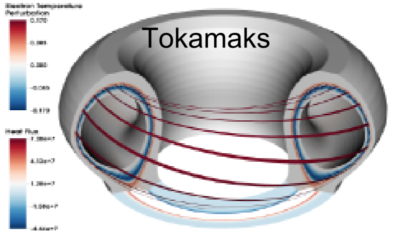
Mechanical loading of mock device

Prediction of Tokamak instability (SciDAC)

Magneto-hydrodynamics (Drekar)

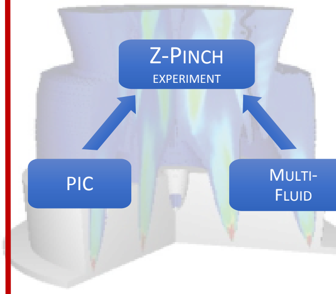
- Model resolutions are well correlated for demo problem
- MLMC is sufficient to obtain 30x reduction in cost for same accuracy



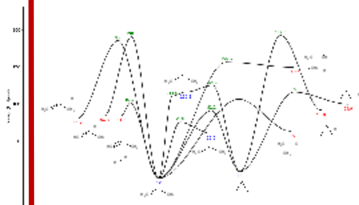


Estimator	N_{400}	N_{200}	N_{100}	Eq. Cost
MC	1273	-	-	1273
MLMC (2 levels)	1	1278	-	236.62
MLMC	1	8	1366	44.36

Emerging



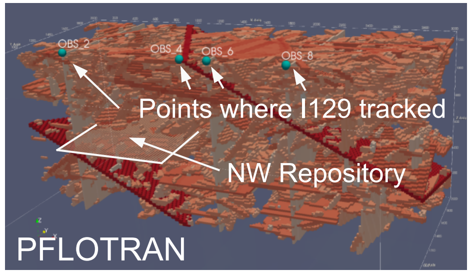
CIS LDRD: non-hierarchical ensemble (models + experiments)

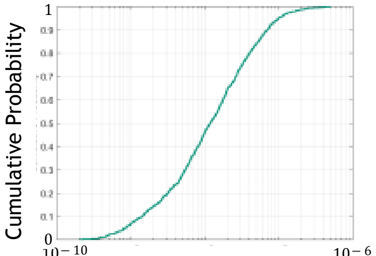


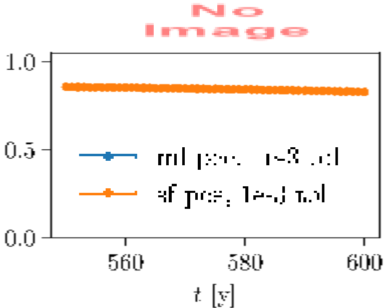
BES QC: exploration of the C_3H_6 PES with KinBot

Geologic Disposal

GDSA example simulation and QOI:



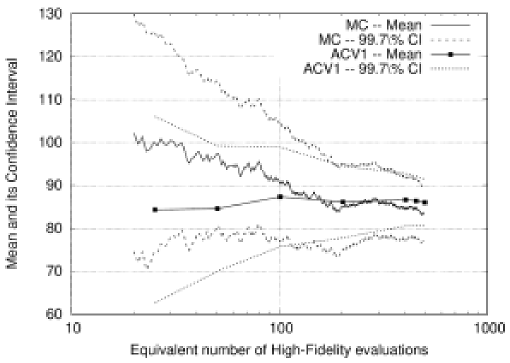




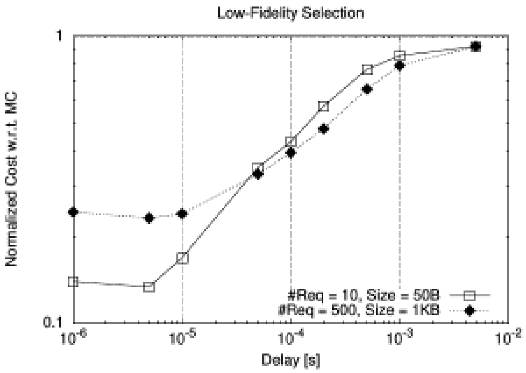
- Deployed MF PCE for GSA to a problem related to geologic disposal safety assessment (GDSA)
- Sobol' indices for model response as fn. of time
- Indices practically identical with ~80 equivalent HF evaluations for MF PCE compared to 713 evaluations for equivalent accuracy SF PCE.

Network Cybersecurity (SECURE GC LDRD)

- Deployed ACV for forward UQ with HF emulation (minimega) and LF discrete event simulation (ns-3)
- Investigated the efficiency of MF UQ by tuning ns-3 models
- Demonstrated increased efficiency for tail est. given a minimega dataset



Forward UQ: ACV1 vs MC



ns-3 tuning effect on ACV performance