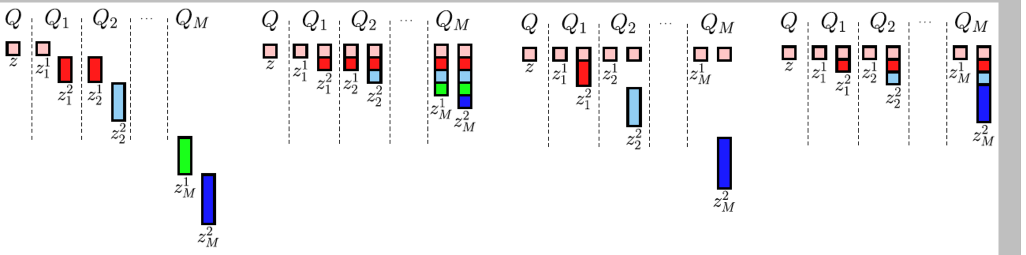
SAND2023-07098C

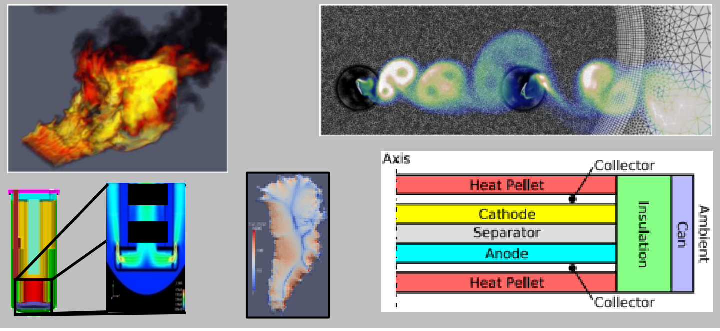
Exceptional service in the national interest



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

DAKOTA

Explore and predict with confidence



Model Ensemble Configuration for Multifidelity UQ

Michael S. Eldred[1], Gianluca Geraci[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*

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.

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.

Multilevel / Multifidelity Estimators based on Sampling

Sandia National Laboratories

Estimator	Type	Sample allocation
MLMC	1D: hierarchical, recursive	Analytic
MFMC	1D: hierarchical, recursive	Analytic, Numerical
MLMF MC	2D: HF,LF pair + resolutions	Analytic
ACV	Non-recursive / peer: all CV pairings target root	Numerical
Gen. ACV	Search over approx sets & DAGs (MFMC + ACV + intermediate)	Numerical
ML BLUE	Model groupings	Numerical

Motivation: production deployments of ML/MF methods encounter a variety of challenges that can impede performance

- Accurate a priori / offline estimations of Covar[Q] are often impractical, and should rather be integrated and optimized
  → *iterated pilot approaches*
- LF models often have parameters that trade accuracy vs. cost (set via SME judgment, but intuition often inaccurate in this context)
  → *hyper-parameter model tuning*
- Numerical solutions [ACV, GenACV] often suffer from multi-modality (and multiple solutions may exhibit similar performance)
  → augment local solutions: multi-start local from analytic initial guesses, global + local
- For general model ensembles, the best approximation selections and CV pairings/groupings are not known a priori
  → *ensemble selection and DAG enumeration*

Each of these concerns introduces additional sequenced or nested iteration, or expands the scale of an integrated optimization

Ensemble Configuration in Multifidelity Sampling

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

Covariance refinement: $\Delta N_{shared} \rightarrow$ Cover[Q]

Not swappable without N_{shared} re-eval.
Refinement is deferable:
tuning often projection-based

Ensemble selection: best subset (drop low value approx.)

CV selection (enumerate/optimize DAG pairings)

Separated for numerical reasons

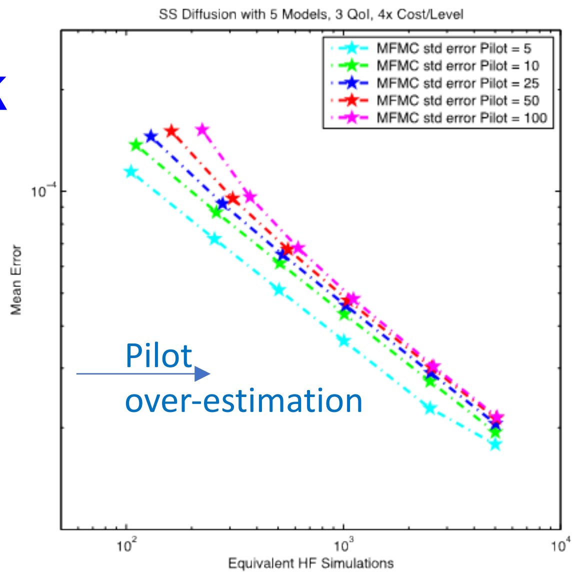
$$\arg\min_{\mathbf{r},N} \frac{Var[Q]}{N}(1 - R^2(\theta, \mathbf{r})) \quad s.t. \quad N\left(w + \sum_{i=1}^{M} w_i(\theta) r_i\right) \leq C$$

Resource Allocation: cost ⟷ accuracy

Fixed covariance (GenACV w/ MFMC, ACV as special cases)

Variable covariance: converge on N_{shared}, tune w.r.t. θ

Outer loops (modifying covariance): review of previous work



Iterated Pilots → integrate pilot sample as *online cost* and optimize total

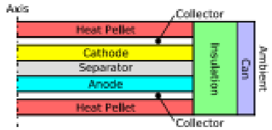
- {MFMC, ACV, GenACV} utilize a shared $N^{(i)}$ for estimation of Covar[Q] across models

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

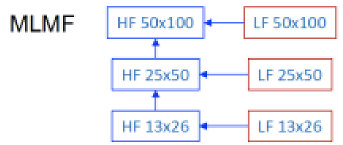
1) Sample all models
2) $N^{(i)}$ shared samples → $Cov_{LL}^{(i)}$, $Cov_{LH}^{(i)}$ → opt. solver → r^*, N^*
3) Compute one-sided ΔN for shared samples from $N^{(i)}$ to N^*
   A. Optional: apply under-relaxation factor γ
   B. If non-zero increment, advance (i) and return to 1)

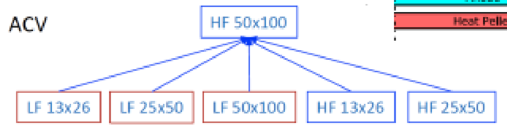
- Avoid inefficiency (over-est.) or inaccuracy (under-est.). Numerical solves provide resilience → find near-optimal solns. incorporating large pilots.

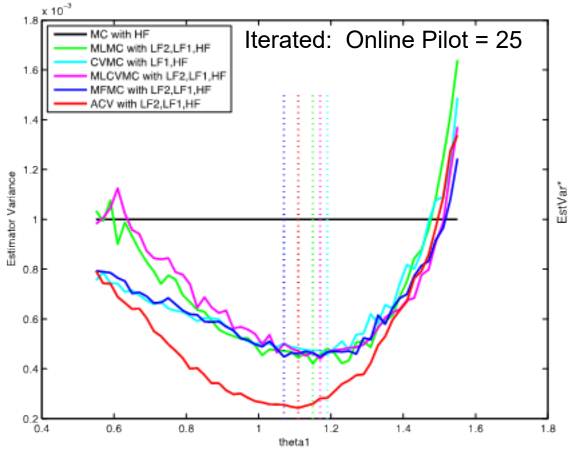
Hyper-Parameter Model Tuning → tune approximations to achieve best accuracy vs. cost trade-off

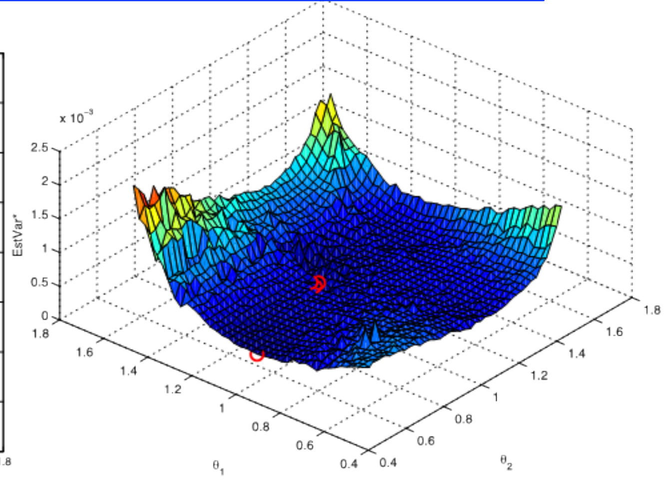


$$\arg\min_{\theta} \left[\arg\min_{\mathbf{r},N} \frac{Var[Q]}{N}(1 - R^2(\theta,\mathbf{r})) \quad s.t. \quad N\left(w + \sum_{i=1}^{M} w_i(\theta) r_i\right) \leq C \right]$$









Hand-tuned: refine across discrete combinations until ρ > 0.9 obtained for all QoI

Hand-tuned hyper-parameters:			
0.01	initial time step		27.3x
0.10	predictor-corrector tol		
0.10	nonlinear residual tol		
Projected MLCV Estimator Variance:	.050092		
Single fidelity accuracy for equiv cost:	1.3668	(969 HF)	
Single fidelity cost for equiv accuracy:	26,440 HF	(EstVar 0.050092)	

Hand-tuned hyper-parameters:			
0.01	initial time step		24.8x
0.10	predictor-corrector tol		
0.10	nonlinear residual tol		
Projected ACV Estimator Variance:	.053138		
Single fidelity accuracy for equiv cost:	1.3178	(1005 HF)	
Single fidelity cost for equiv accuracy:	24,925 HF	(EstVar .053138)	

EGO-tuned: global minimization of variance of selected estimator (max iter = 80)

Optimal hyper-parameters:			
0.0084097	initial time step		39.7x
0.0061138	predictor-corrector tol		
0.028493	nonlinear residual tol		
Projected MLCV Estimator Variance:	.034396		
Single fidelity accuracy for equiv cost:	1.3654	(970 HF)	
Single fidelity cost for equiv accuracy:	38,506 HF	(EstVar 0.034396)	

Optimal hyper-parameters:			
0.0067487	initial time step		143x
0.0010880	predictor-corrector tol		
0.046707	nonlinear residual tol		
Projected ACV Estimator Variance:	0.0092395		
Single fidelity accuracy for equiv cost:	1.3192	(1004 HF)	
Single fidelity cost for equiv accuracy:	143,340 HF	(EstVar 0.0092395)	

Greater tuning impact for more flexible estimators → GenACV

E., Geraci, Portone, Gorodetsky, Jakeman, "All-at-Once (and Bi-Level) Model Tuning for Multifidelity Sampling," SIAM UQ22.
Adams, E., Geraci, Portone, Ridgway, Stephens, Wildey, "Deployment of MF UQ for Thermal Battery Assessment, Part I," SAND2022-11856

Inner Loop (fixed covariance): Competed numerical solvers

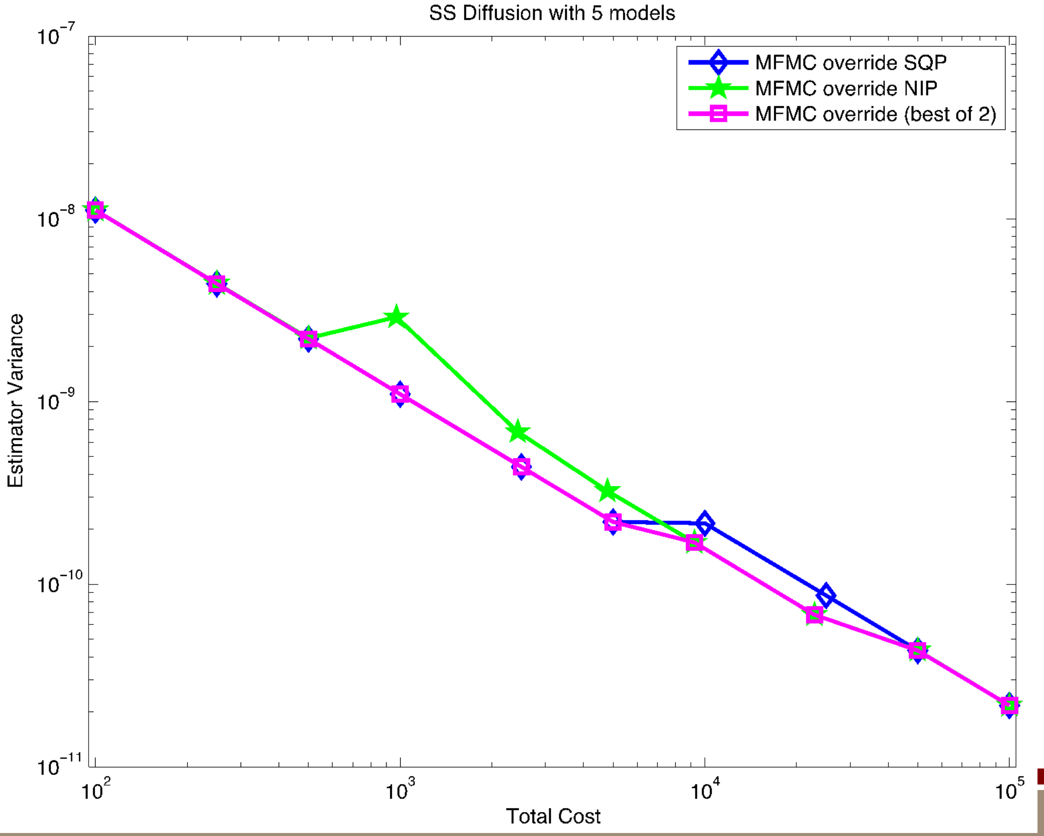
Sandia National Laboratories

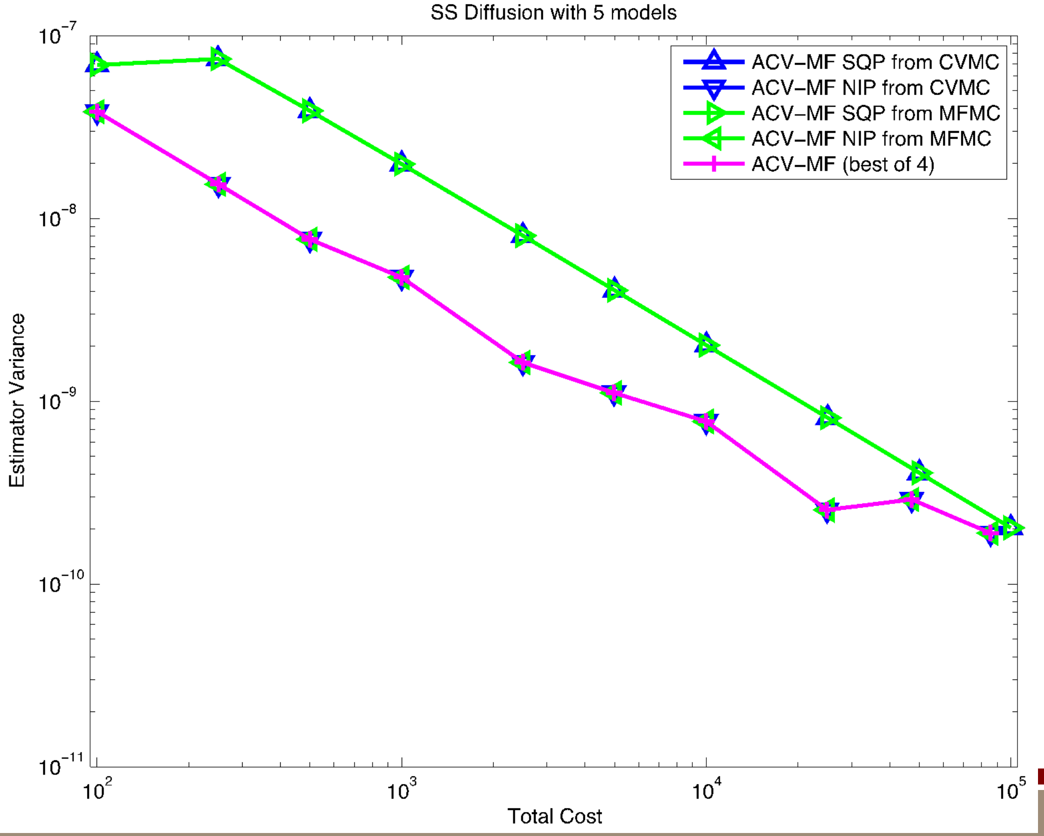
First iteration: multi-start / multi-solver

- Analytic solutions as initial guesses:
  - MFMC (ordered, reordered using average ρ^2_{LH})
  - Pairwise CVMC for given ensemble + DAG
- Solvers
  - SQP (via NPSOL)
  - NIP (via OPT++)

Subsequent iterations (if online/iterated)

- Warm start from previous best soln (1) keyed for active ensemble + DAG





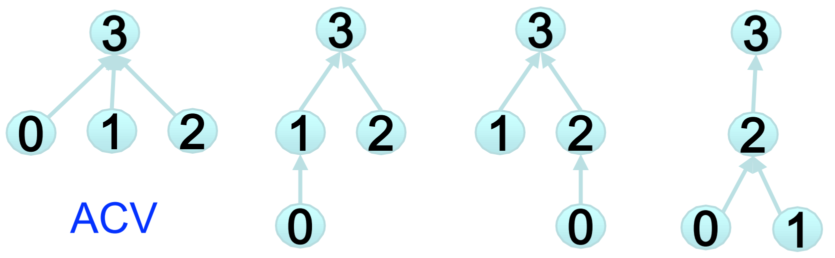
Inner Loop: graph enumeration + model selection (Generalized ACV)

Sandia National Laboratories

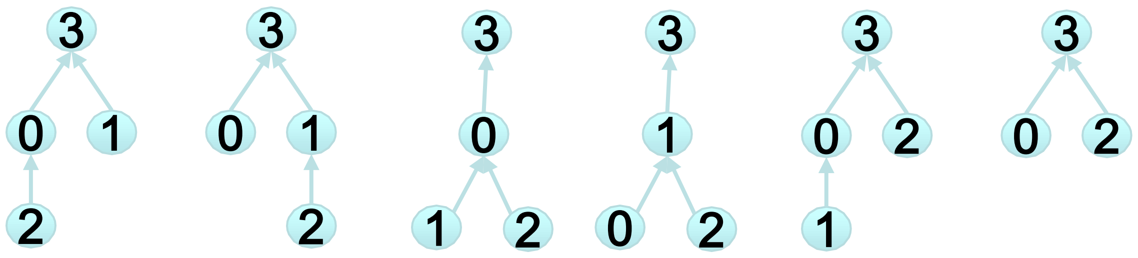
Explore possible model dependencies, as defined by DAGs that identify control variate pairings

- GenACV-MF is inclusive of current estimators MFMC and ACV-MF

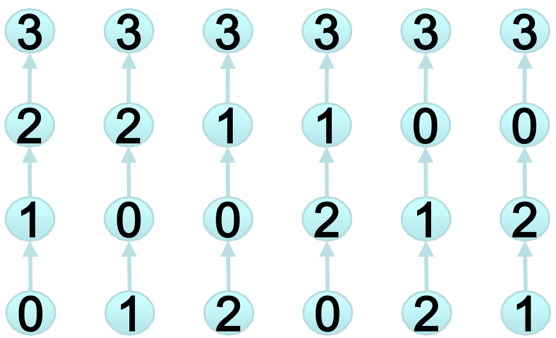
K-L (ordered depth = 2)



Partial (unordered depth = 2)

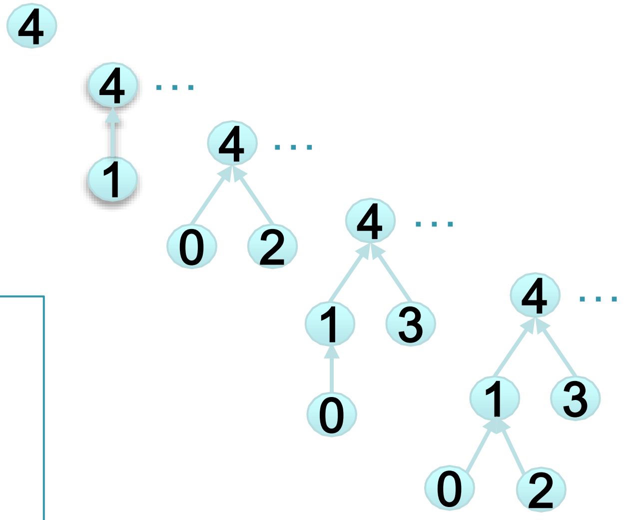


Full (unordered depth = all)



DAG enumeration

- For 3 models, total DAG = 3
- For 4 models, total DAG = 16 (depicted)
- For 5 models, total DAG = 125



DAG enumeration + model selection

Augments 1 set of 5 models (# DAG = 125)

- For 1 set of 1 model (MC), empty DAG = 1
- For each of 4 sets of 2 models, # DAG = 1
- For each of 6 sets of 3 models, # DAG = 3
- For each of 4 sets of 4 models, # DAG = 16

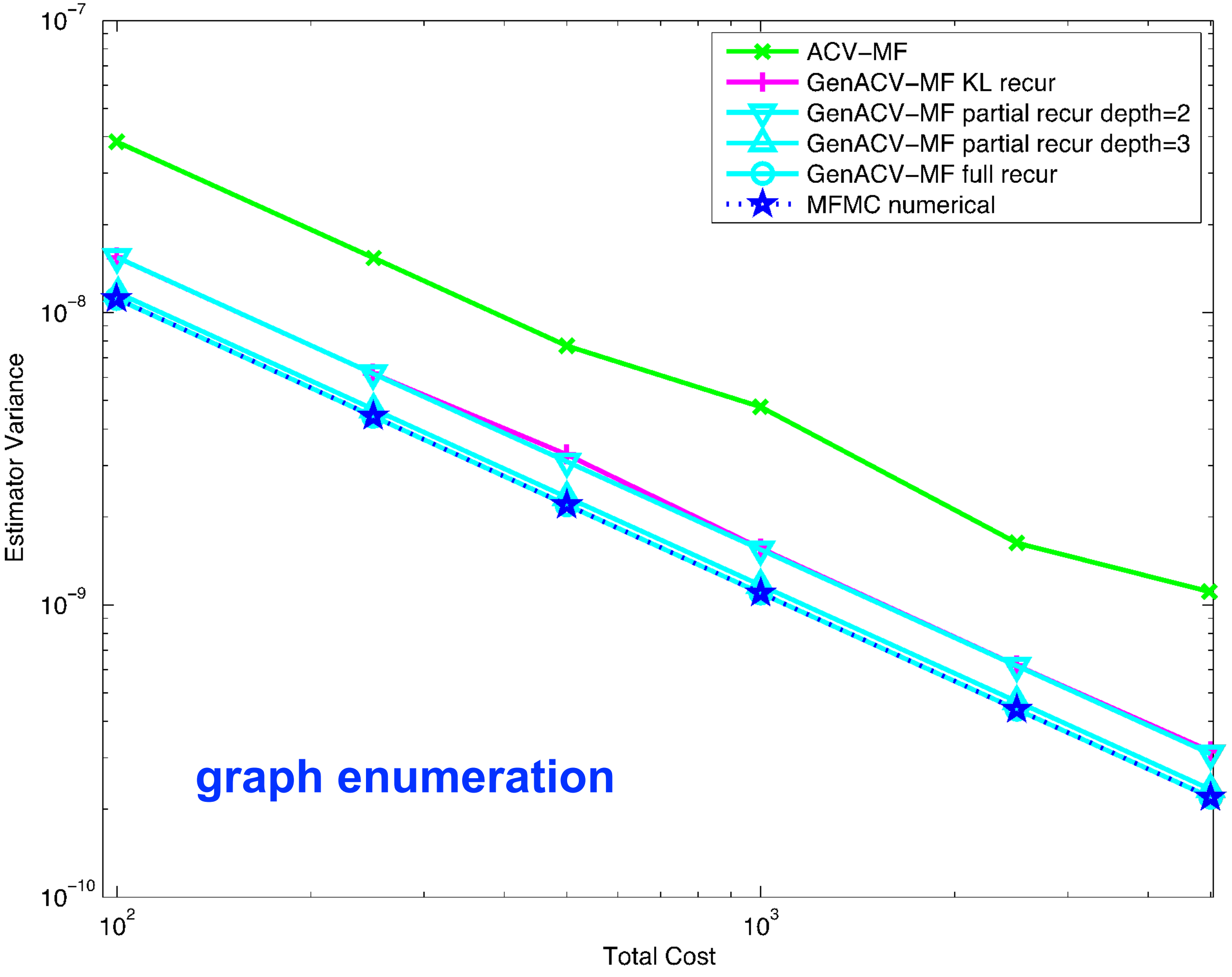
Total = 212 DAGs

Generalized ACV: test problem 1

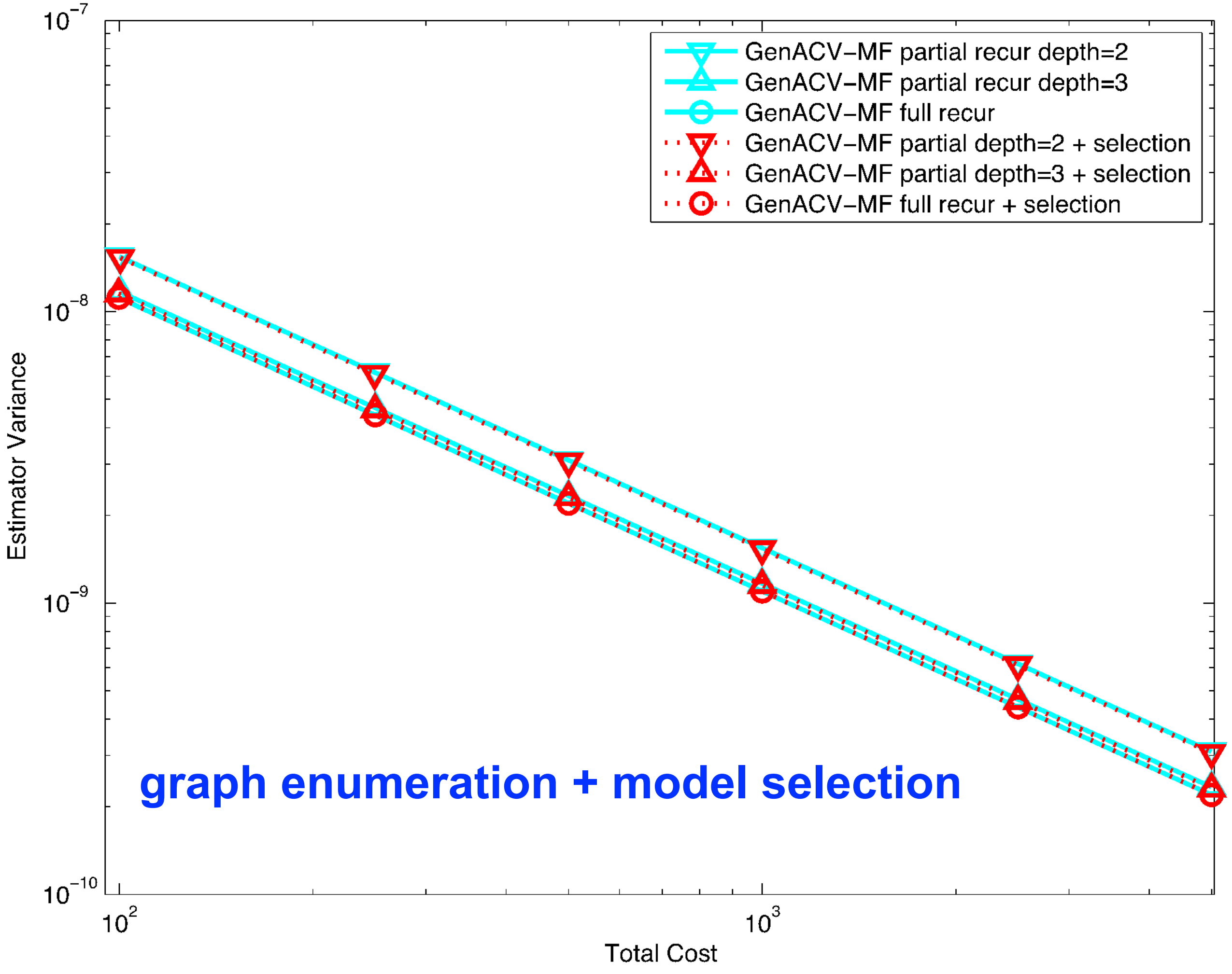
Steady state 1D diffusion: 5 well-ordered models
resolutions = {4,8,16,32,64}, relative cost = {1,4,16,64,256}

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

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



Peer DAG is not well suited → GenACV recovers MFMC at full depth



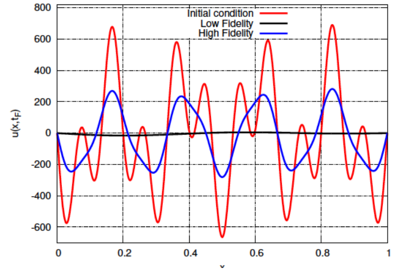
All approximations are useful (slight improvements for restricted depth)

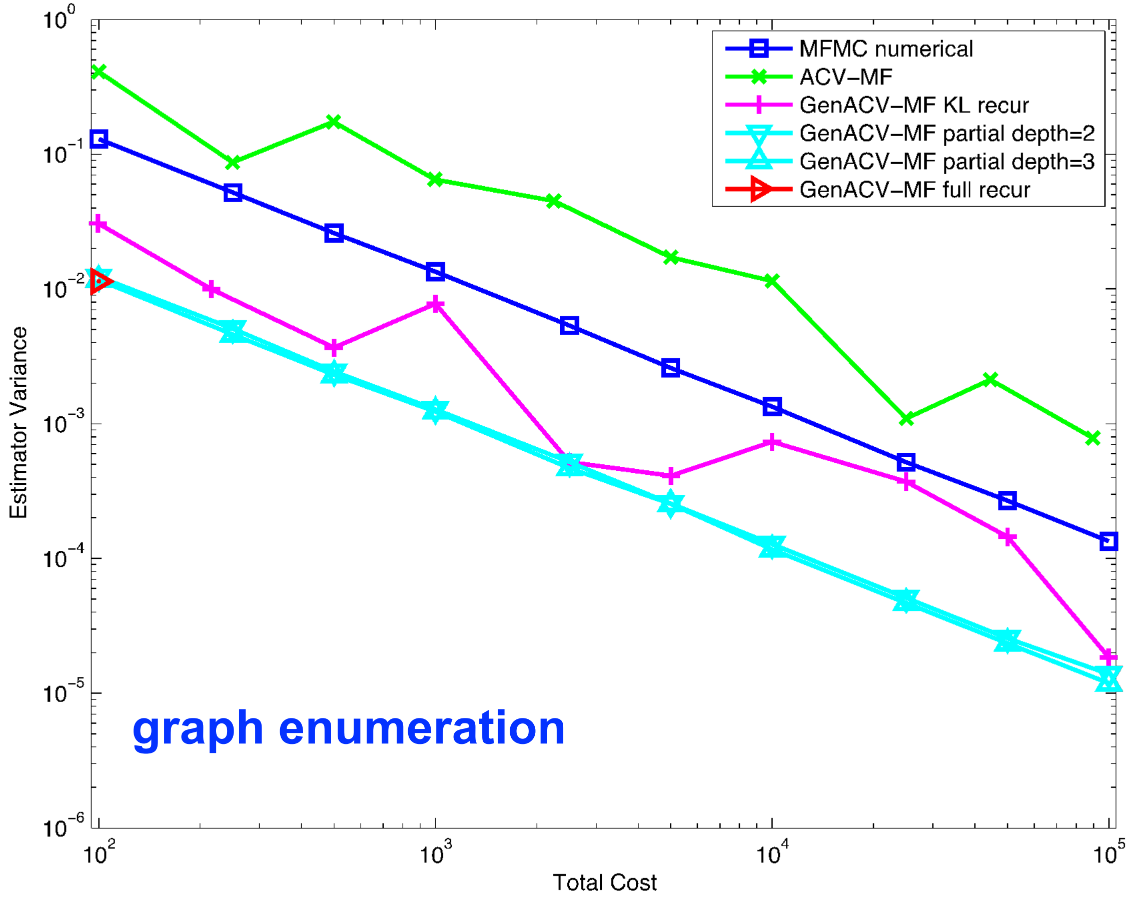
Generalized ACV: test problem 2

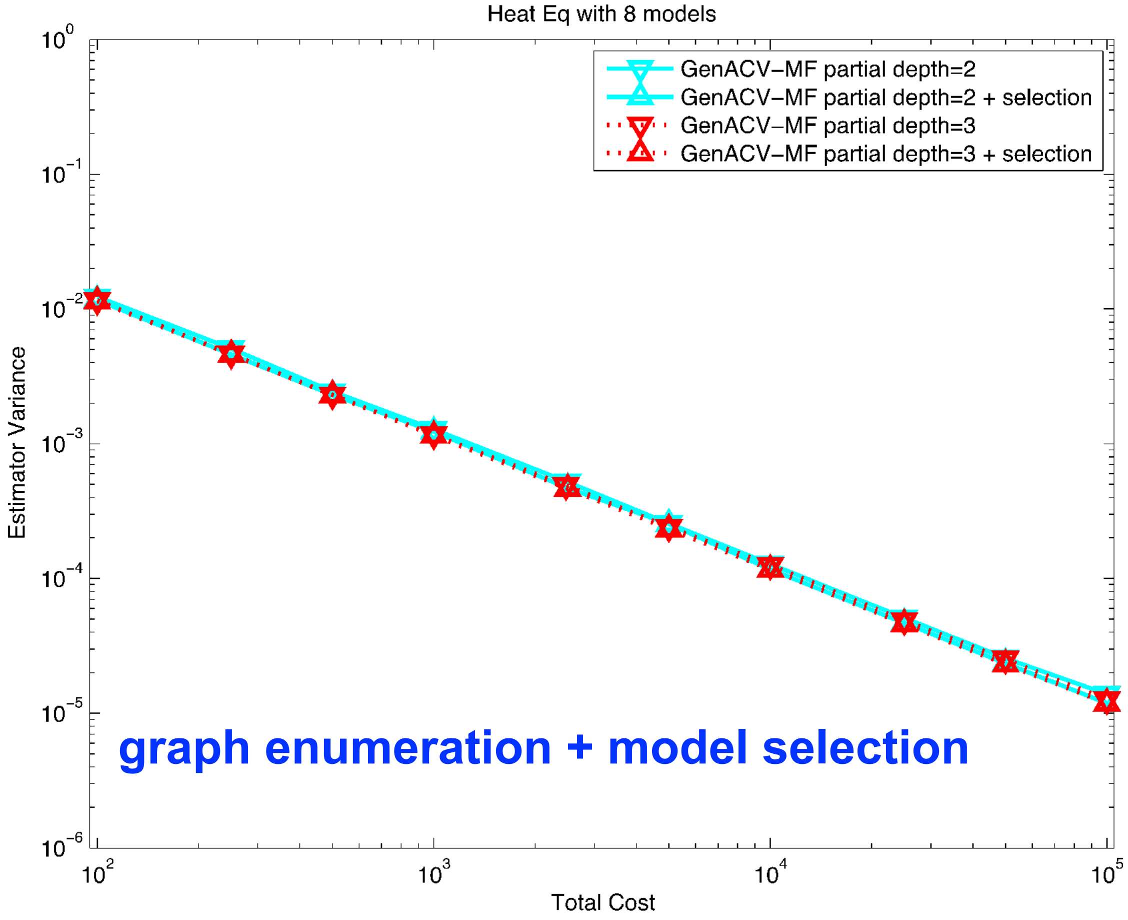
1D transient diffusion (“heat equation”)

- 8 models in 2D hierarchy: multifidelity + multilevel
- Fourier solution modes = 3 LF, 21 HF
- Spatial coordinates = {5 15 30 60} LF, {30 60 100 200} HF

$$\begin{cases} \dfrac{\partial u(x,\xi,t)}{\partial t} - \alpha(\xi)\dfrac{\partial^2 u(x,\xi,t)}{\partial x^2} = 0, & \alpha > 0,\ x \in [0,L] = \Omega \subset \mathbb{R} \\ u(x,\xi,0) = u_0(x,\xi), & t \in [0,t_F] \quad \text{and} \quad \xi \in \Xi \subset \mathbb{R}^d \\ u(x,\xi,t)|_{\partial\Omega} = 0 & \\ u_0(x,\xi) = \mathcal{G}(\xi)\mathcal{F}_1(x) + \mathcal{I}(\xi)\mathcal{F}_2(x) & \end{cases}$$







More complex hierarchy benefits significantly from DAG search
Solver noise could be smoothed with sample replicates

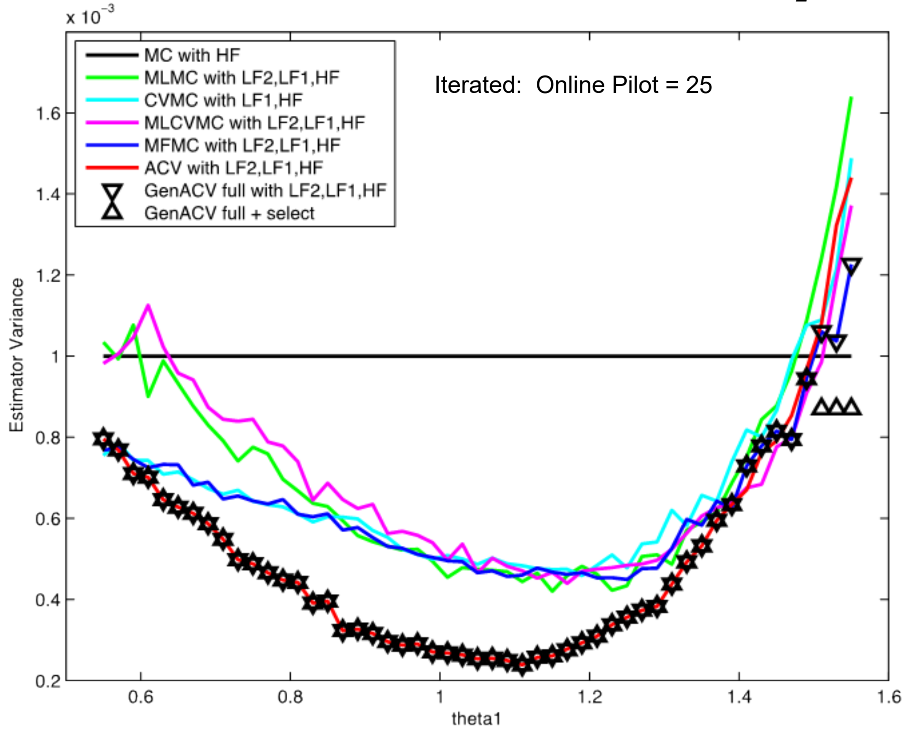
Model selection adds a small amount of additional performance, but all models generally providing utility

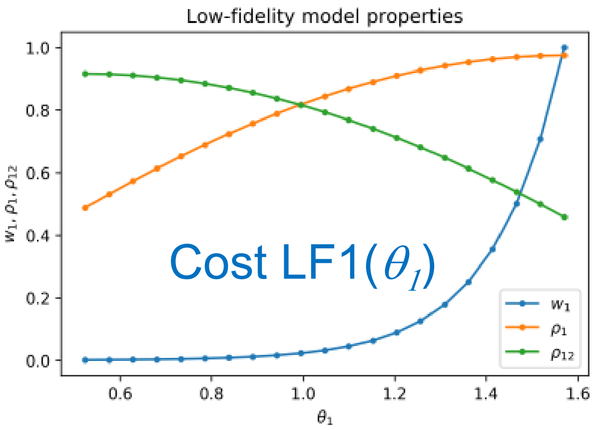
Putting it all together: tuning, iterated pilots, ensemble + DAG selection

“Tunable Model” Definitions (JCP 2020)

$$Q(\theta) = \sqrt{11}\ [\ \cos(\theta)\ x^5 + \sin(\theta)\ y^5\ ]$$
$$Q_1(\theta_1) = \sqrt{7}\ [\ \cos(\theta_1)\ x^3 + \sin(\theta_1)\ y^3\ ]$$
$$Q_2(\theta_2) = \sqrt{3}\ [\ \cos(\theta_2)\ x + \sin(\theta_2)\ y\ ]$$

Start with tuning 1 parameter (θ_1) for mid-fidelity high / low hyper-parameters fixed: $\theta = \pi/2$, $\theta_2 = \pi/6$

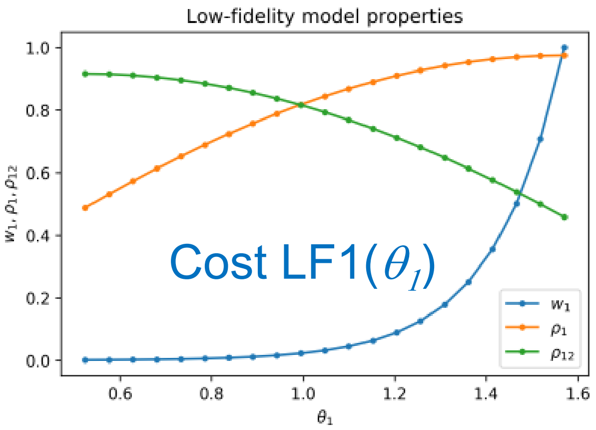


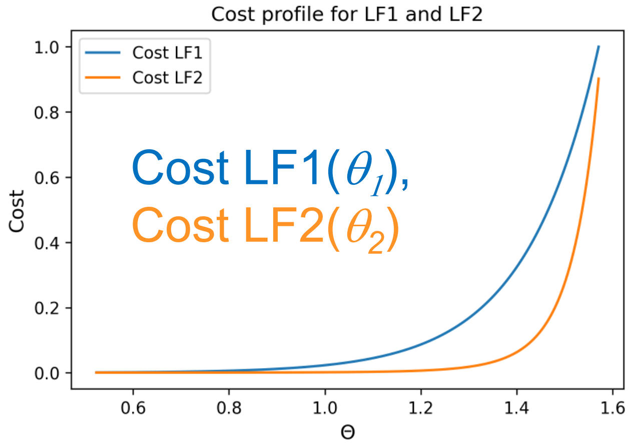


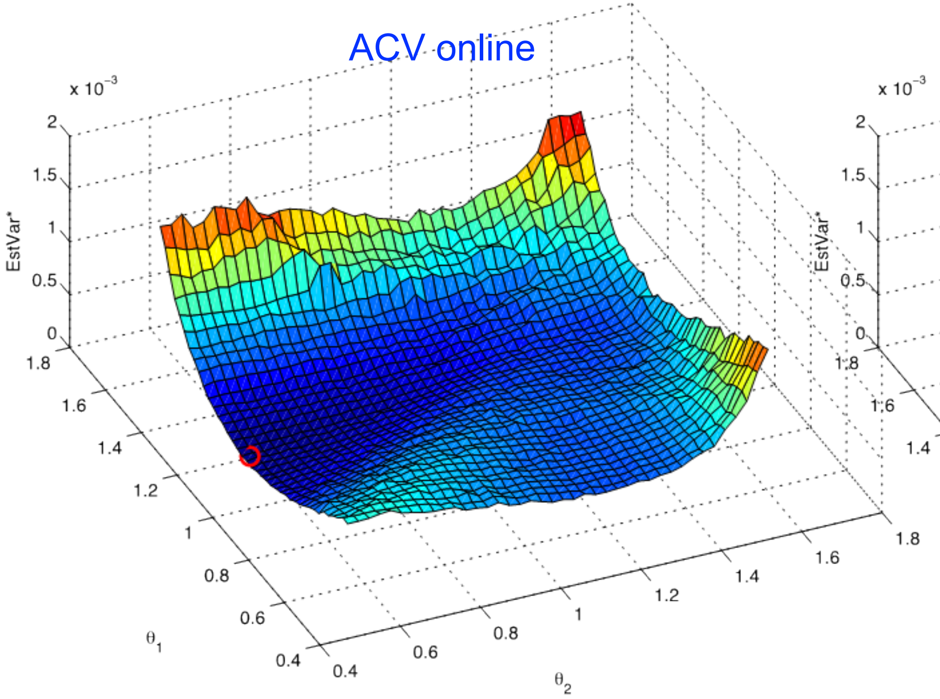
Putting it all together: tuning, iterated pilots, ensemble + DAG selection

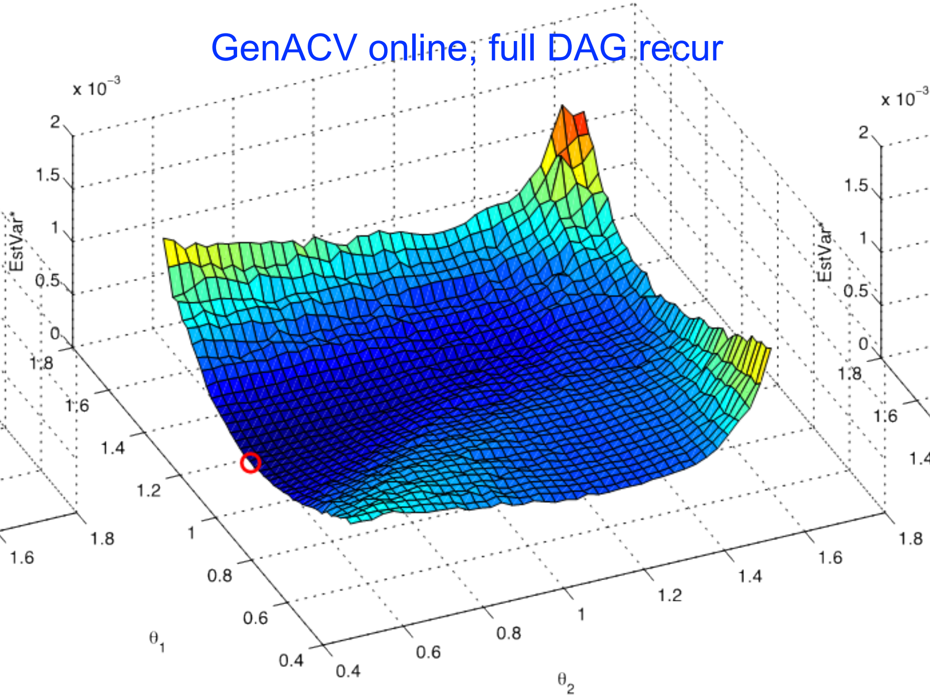
"Tunable Model" Definitions (JCP 2020)

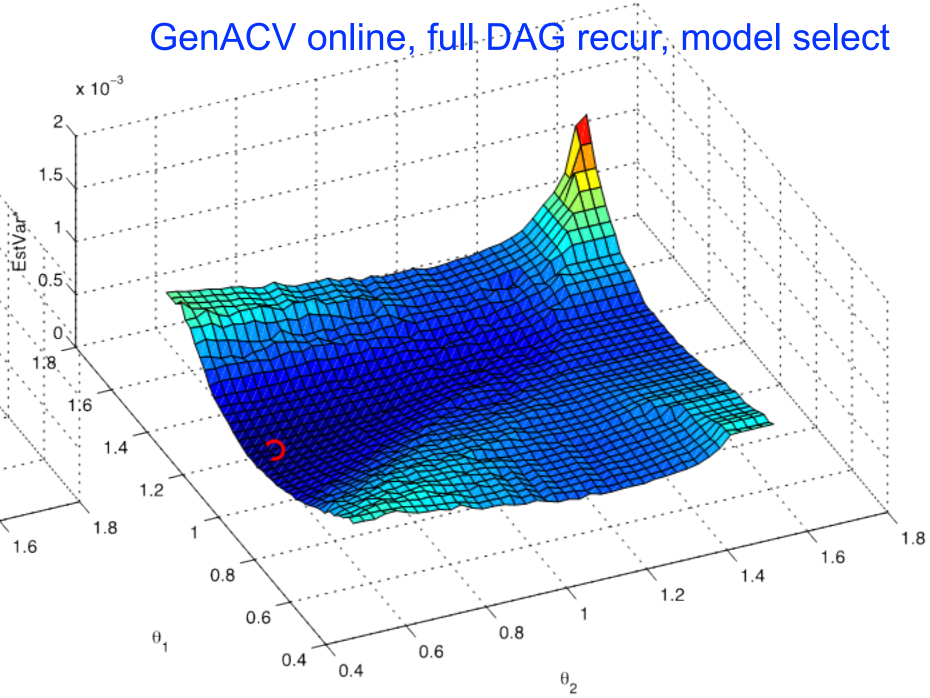
$$Q(\theta) = \sqrt{11}\ [\ \cos(\theta)\ x^5 + \sin(\theta)\ y^5\ ]$$
$$Q_1(\theta_1) = \sqrt{7}\ [\ \cos(\theta_1)\ x^3 + \sin(\theta_1)\ y^3\ ]$$
$$Q_2(\theta_2) = \sqrt{3}\ [\ \cos(\theta_2)\ x + \sin(\theta_2)\ y\ ]$$











In general, hyper-parameter model tuning amplifies the utility of non-hierarchical estimators and DAG flexibility

Summary Observations

Sandia National Laboratories

Production deployments of multifidelity methods encounter a variety of challenges

- Accurate offline estimations of Covar[Q] are expensive and should be integrated and optimized → ***Iterated online pilots***
- LF models often have parameters that trade accuracy vs. cost → ***Hyper-parameter model tuning***
- Numerical solutions are not always reliable w/ local solvers → ***Multistart/Multisolver, Global/Local***
- Best selections/pairings/groupings often unknown a priori → ***Model selection/DAG enumeration***

Outer loop (varying the inter-model covariance; previous work)

- Iterated pilots: avoid pilot under-/over-estimation or, with numerical solutions, mitigate the effects
- Hyper-parameter model tuning with bi-level & AAO approaches: especially effective with less structured estimators

Inner loop (fixing the inter-model covariance; recent work)

- Refine solver definitions with competition/sequencing *→ additional refinements in progress*
- Optimize CV pairings via DAG enumeration *→ clear benefit*
- Select most performant ensemble of approximations *→ marginal additional benefit for std test cases w/ good models; poor model cases discarded for model mis-tuning*

Next steps

- Streamline for efficiency (model tuning + large enumerations can become impractical even for simple test problems)
  - Solvers for MFMC/ACV/GenACV: multi-start/multi-solver → sequenced global-local search
  - ML-BLUE can unify ensemble configuration steps, SDP may aid in solver challenges?
  - MINLP / heuristic search strategies to short-circuit brute-force enumerations in GenACV

Extra

Inner Loop (fixed covariance): Sequenced numerical solvers

First iteration

- Global search as initial guess(es):
  - EGO (Bayesian opt; maximize EIF w/ DIRECT+local)
  - DIRECT (GP indirection may not add much in this case)
- Local refinement
  - SQP (via NPSOL)
  - NIP (via OPT++)

Subsequent iterations (if online/iterated)

- Warm start from previous best soln (1) keyed for active ensemble + DAG

TO DO LIST:

- Heat eq 8 models: identify best DAGs with/without model selection
- DIRECT + Local
- Model selection for KL
- ML BLUE
- Model selection for MFMC? (ACV fixed depth = all)
- Model selection without DAG enumeration (set ACV partial depth=1 and/or support “no recursion”)
- Include MC case in count and enumeration