Scalability of Global UQ Methods

For production UQ analyses, we prefer fast converging global methods:

- local approximate methods (reliability methods, moment-based methods) exhibit significant errors in presence of multimodal/nonsmooth/highly nonlinear responses
- MC/LHS are robust with dimension-independent convergence rates, but rates can be unacceptably slow

Spectral methods (e.g., PCE) provide a more effective balance of robustness and efficiency, especially when solution smoothness can be exploited

Exponential growth in expansion cardinality with *n* and *p*, and collocation requirements are >= the number of terms

To mitigate the curse of dimensionality:

- *A priori* model reduction methods (e.g., POD, Karhunen-Loeve) or other surrogate techniques (e.g., multifidelity)
- Goal-oriented adaptive refinement to reduce effective dimension
- Adjoint techniques [given *n* (random dimension) > *m* (response QoI)]
- Sparsity detection methods: compressive sensing, least interpolation

Primary focus is stochastic exp., but other adaptive sampling efforts are related
(and can be leveraged within an abstract refinement framework: EGRA, GPAIS, k-d darts, Morse-Smale)

Non-Intrusive Stochastic Expansions: Polynomial Chaos and Stochastic Collocation

Polynomial chaos: spectral projection using orthogonal polynomial basis fns

$$\boxed{R = \sum_{j=0}^{P} \alpha_j \Psi_j(\boldsymbol{\xi})} \quad \text{using} \quad \boxed{\begin{array}{lllll} \Psi_0(\xi) & = & \psi_0(\xi_1)\,\psi_0(\xi_2) & = & 1 \\ \Psi_1(\xi) & = & \psi_1(\xi_1)\,\psi_0(\xi_2) & = & \xi_1 \\ \Psi_2(\xi) & = & \psi_0(\xi_1)\,\psi_1(\xi_2) & = & \xi_2 \\ \Psi_3(\xi) & = & \psi_2(\xi_1)\,\psi_0(\xi_2) & = & \xi_1^2 - 1 \\ \Psi_4(\xi) & = & \psi_1(\xi_1)\,\psi_1(\xi_2) & = & \xi_1 \xi_2 \\ \Psi_5(\xi) & = & \psi_0(\xi_1)\,\psi_2(\xi_2) & = & \xi_2^2 - 1 \end{array}}$$

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

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

$$\boxed{\langle \Psi_j^2 \rangle \;=\; \prod_{i=1}^{n} \langle \psi_{m_i^j}^2 \rangle}$$

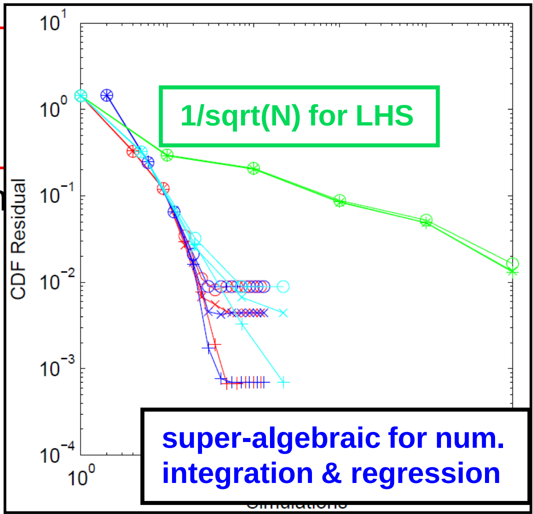
Stochastic collocation: instead of estimating coefficients for known basis functions, form interpolants for known coefficients

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

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

$$\boxed{L_i = \prod_{\substack{j=1 \\ j \neq i}}^{m} \frac{x - x_j}{x_i - x_j}} \Rightarrow \boxed{R(\boldsymbol{\xi}) \cong \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} r\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right) \left(L_{j_1}^{i_1} \otimes\right.}$$

Sparse interpolants formed using Σ of tensor in



- Tailor expansion form:
  - p-refinement: anisotropic tensor/sparse, generalized sparse
  - h-refinement: local bases with dimension & local refinement
- Method selection: fault tolerance, decay, sparsity, error est.

Core-Enabled UQ: Multiple Model Forms

Same physics:

- a clear hierarchy of fidelity (low to high) → multifidelity UQ

Low — Med — High

- an ensemble of models that could all be credible (lacking a clear preference structure) → model form uncertainty (inadequate data), model selection (rich data)

SA-RANS — KE-RANS-NBC — KE-RANS-DBC

Additional dimension(s) for multi-{physics,scale}

Sandia National Laboratories

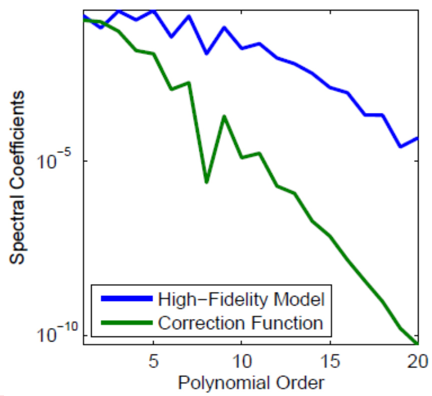
Multifidelity UQ using Stochastic Expansions

Multifidelity UQ through stochastic expansion of model discrepancy:

- Extension of multifidelity opt methods that converge to local HF optimum based on local corrections
- Converge to global HF statistics based on global corrections (0^{th}/1^{st} consistency @HF collocation pts)

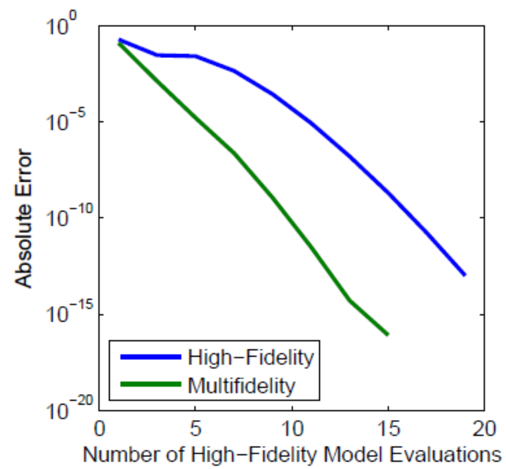
$$\hat{f_{hi}}(\xi) = \sum_{j=1}^{N_{lo}} f_{lo}(\xi_j) L_j(\xi) + \sum_{j=1}^{N_{hi}} \Delta f(\xi_j) L_j(\xi)$$

$$N_{lo} >> N_{hi}$$

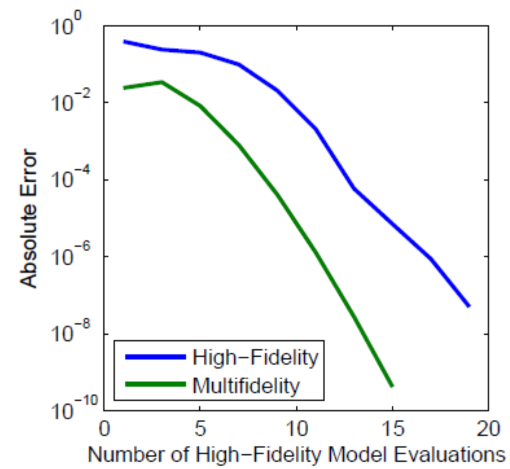


$$R_{\text{high}}(\xi) = e^{-0.05\xi^2} \cos 0.5\xi - 0.5 e^{-0.02(\xi-5)^2}$$

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



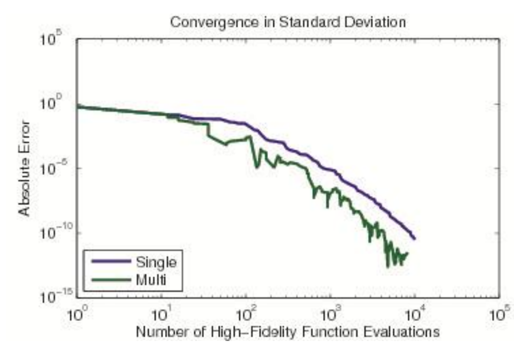
(a) Error in mean



(b) Error in standard deviation

Adaptive algorithm balances LF/HF cost and targets regions where LF predictive capabilities break down:

- Greedy selection of index sets for LF *or* model discrepancy based on ΔQOI/ΔCost



Multifidelity UQ with stochastic expansions

- A hierarchical approximation: resolve expansions for LF and ≥ 1 levels of model discrepancy
  - leverage information from less expensive low & medium-fidelity models
- Adaptive multifidelity algorithm → further generalization to generalized sparse grids
  - target regions where predictive capability of LF model breaks down
  - greedy selection of candidates that provide the greatest benefit to HF QoI per unit cost

Performance

- Ideal LF model for multifidelity UQ would result in discrepancy with the following properties:
  - discrepancy has lower complexity than HF model (spectrum of coefficients of discrepancy expansion decays more rapidly than HF expansion) → faster convergence rate (affects exponent)
  - discrepancy has lower variance than HF model → reduction in initial error (affects leading constant)
- Examples with good LF models → short column R_{low1}, elliptic PDE
  - ~ 80% reduction in HF evals for comparable statistical accuracy

$$|x_n - L| < Cn^{-p}$$

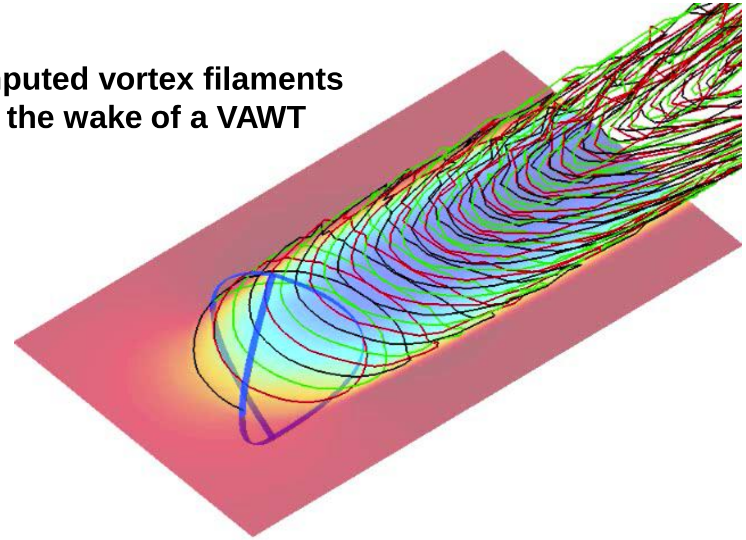
- In non-ideal cases, LF model is non-informative or omits/introduces high order information
  - horn acoustics: multifidelity performance did not degrade significantly from single-fidelity performance, as algorithm can fall back to reliance on resolving the original HF trends
- Additional directions:
  - adaptively discarding models from the hierarchy that are not providing value (low selection rate)
  - basis pursuit approaches (compressive sensing) that can directly target high-order discrepancy while benefiting from LF capture of low order trends

Current Focus: VAWT Performance Modeling

Vertical-axis Wind Turbine (VAWT)

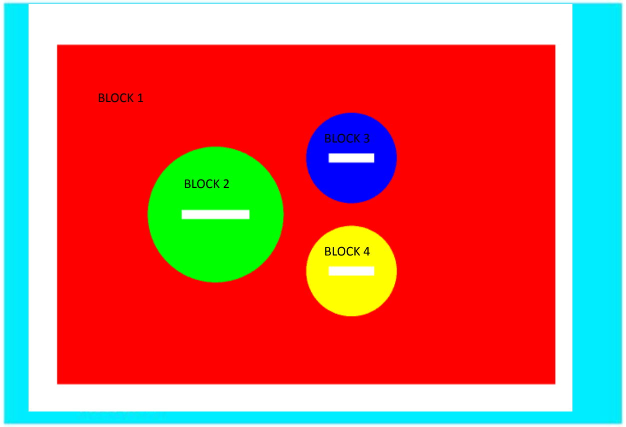
Low fidelity

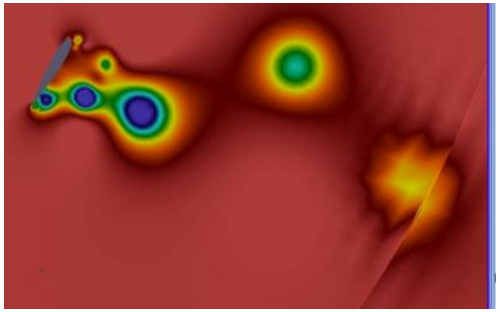
CACTUS: Code for Axial and Crossflow TUrbine Simulation

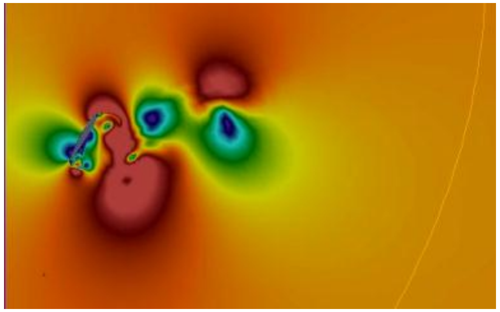


Computed vortex filaments in the wake of a VAWT

High fidelity: DG formulation for LES





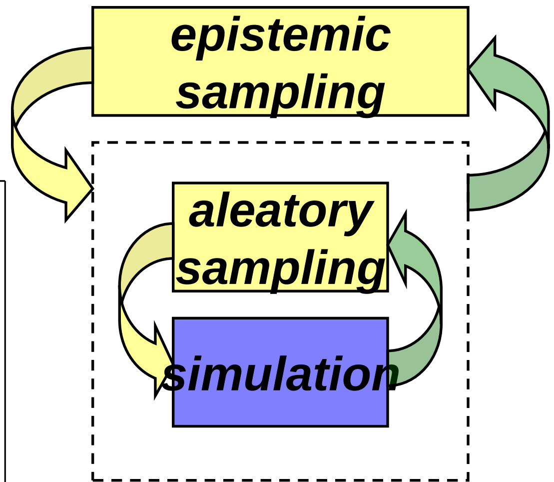


Sandia National Laboratories

Model Form Uncertainty Propagation

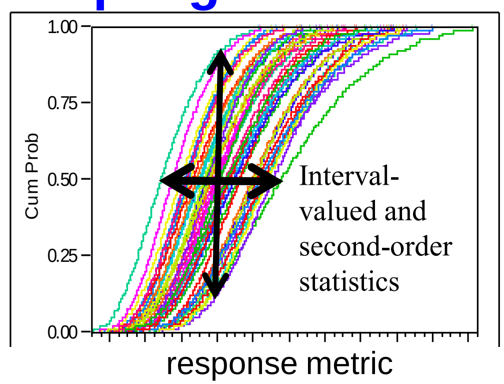
Mixed Aleatory-Epistemic UQ with Discrete Epistemic Model Forms

Epistemic uncertainty (aka: subjective, reducible, lack of knowledge uncertainty): insufficient info to specify objective probability distributions



Traditional approach: nested sampling

- Expensive sims → under-resolved sampling (especially @ outer loop)
- Under-prediction of credible outcomes



Algorithmic approaches

- Interval-valued probability (IVP), *aka* probability bounds analysis (PBA)
- Dempster-Shafer theory of evidence (DSTE)
- Second-order probability (SOP), *aka* probability of frequency

Increasing epistemic structure (stronger assumptions)

Address accuracy and efficiency

- Inner loop: stochastic exp. that are epistemic-aware (aleatory, combined)
- Outer loop:
  - IVP, DSTE: opt-based interval estimation, global (EGO) or local (NLP)
  - SOP: nested stochastic exp. (nested expectation is only post-processing in special cases)

$$\begin{array}{ll} \text{minimize} & M(s) \\ \text{subject to} & s_L \leq s \leq s_U \end{array}$$

$$\begin{array}{ll} \text{maximize} & M(s) \\ \text{subject to} & s_L \leq s \leq s_U \end{array}$$

Mixed Aleatory-Epistemic UQ:
IVP, SOP, and DSTE based on Stochastic Expansions

Interv Est Approach	UQ Approach	Expansion Variables	Evaluations (Fn, Grad)	Area	β

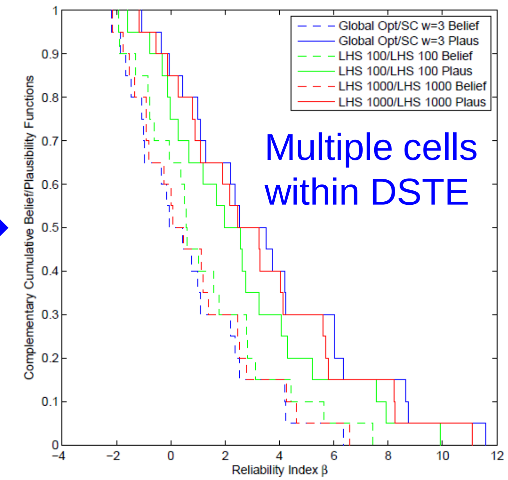
IVP SC SSG Aleatory: β interval converged to 5-6 digits by 300-400 evals

Interv Est Approach	UQ Approach	Expansion Variables	Evaluations (Fn, Grad)	Area	β
EGO	SC SSG w = 1	Aleatory	(84/91, 0/0)	[75.0002, 374.999]	[-2.26264, 11.8623]
EGO	SC SSG w = 2	Aleatory	(372/403, 0/0)	[75.0002, 374.999]	[-2.18735, 11.5900]
EGO	SC SSG w = 3	Aleatory	(1260/1365, 0/0)	[75.0002, 374.999]	[-2.18732, 11.5900]
EGO	SC SSG w = 4	Aleatory	(3564/3861, 0/0)	[75.0002, 374.999]	[-2.18732, 11.5900]
NPSOL	SC SSG w = 1	Aleatory	(21/77, 21/77)	[75.0000, 375.000]	[-2.26264, 11.8623]
NPSOL	SC SSG w = 2	Aleatory	(93/341, 93/341)	[75.0000, 375.000]	[-2.18735, 11.5901]
NPSOL	SC SSG w = 3	Aleatory	(315/1155, 315/1155)	[75.0000, 375.000]	[-2.18732, 11.5900]
NPSOL	SC SSG w = 4	Aleatory	(891/3267, 891/3267)	[75.0000, 375.000]	[-2.18732, 11.5900]

IVP nested LHS sampling: converged to 2-3 digits by 10^8 evals

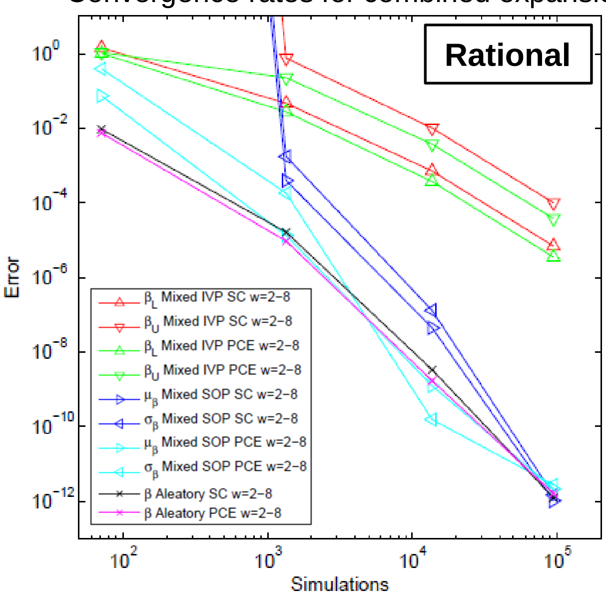
Interv Est Approach	UQ Approach	Expansion Variables	Evaluations (Fn, Grad)	Area	β
LHS 100	LHS 100	N/A	$(10^4/10^4, 0/0)$	[80.5075, 338.607]	[-2.14505, 8.64891]
LHS 1000	LHS 1000	N/A	$(10^6/10^6, 0/0)$	[76.5939, 368.225]	[-2.19883, 11.2353]
LHS 10^4	LHS 10^4	N/A	$(10^8/10^8, 0/0)$	[76.4755, 373.935]	[-2.16323, 11.5593]

Fully converged area interval = [75., 375.], β interval = [−2.18732, 11.5900]



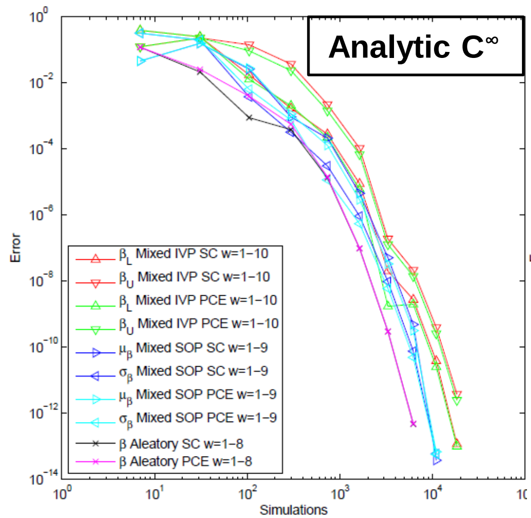
Multiple cells within DSTE

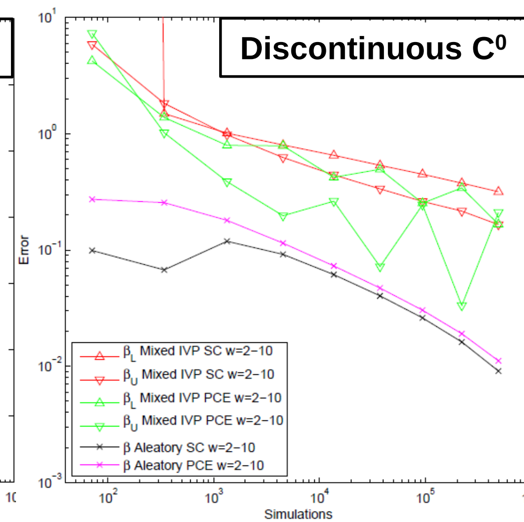
Convergence rates for combined expansions



L^∞ metrics: IVP mixed, DSTE mixed

L^2 metrics: Aleatory, SOP mixed



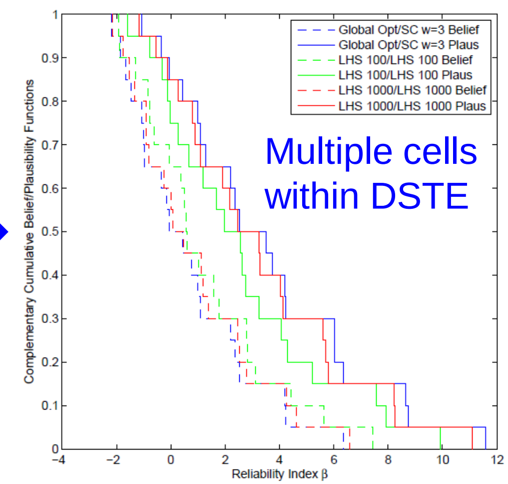


Sandia National Laboratories

Mixed Aleatory-Epistemic UQ:
IVP, SOP, and DSTE based on Stochastic Expansions

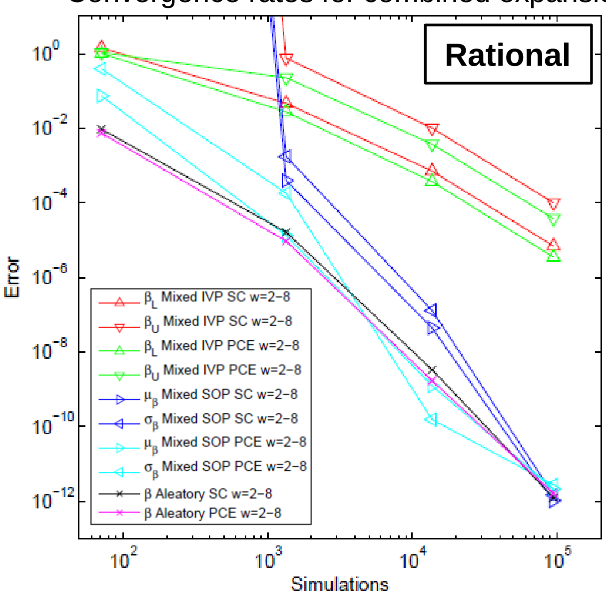
Interv Est Approach	UQ Approach	Expansion Variables	Evaluations (Fn, Grad)	Area	β
IVP SC SSG Aleatory: β interval converged to 5-6 digits by 300-400 evals					
EGO	SC SSG w = 1	Aleatory	(84/91, 0/0)	[75.0002, 374.999]	[-2.26264, 11.8623]
EGO	SC SSG w = 2	Aleatory	(372/403, 0/0)	[75.0002, 374.999]	[-2.18735, 11.5900]
EGO	SC SSG w = 3	Aleatory	(1260/1365, 0/0)	[75.0002, 374.999]	[-2.18732, 11.5900]
EGO	SC SSG w = 4	Aleatory	(3564/3861, 0/0)	[75.0002, 374.999]	[-2.18732, 11.5900]
NPSOL	SC SSG w = 1	Aleatory	(21/77, 21/77)	[75.0000, 375.000]	[-2.26264, 11.8623]
NPSOL	SC SSG w = 2	Aleatory	(93/341, 93/341)	[75.0000, 375.000]	[-2.18735, 11.5901]
NPSOL	SC SSG w = 3	Aleatory	(315/1155, 315/1155)	[75.0000, 375.000]	[-2.18732, 11.5900]
NPSOL	SC SSG w = 4	Aleatory	(891/3267, 891/3267)	[75.0000, 375.000]	[-2.18732, 11.5900]
IVP nested LHS sampling: converged to 2-3 digits by 10^8 evals					
LHS 100	LHS 100	N/A	($10^4/10^4$, 0/0)	[80.5075, 338.607]	[-2.14505, 8.64891]
LHS 1000	LHS 1000	N/A	($10^6/10^6$, 0/0)	[76.5939, 368.225]	[-2.19883, 11.2353]
LHS 10^4	LHS 10^4	N/A	($10^8/10^8$, 0/0)	[76.4755, 373.935]	[-2.16323, 11.5593]

Fully converged area interval = [75., 375.], β interval = [−2.18732, 11.5900]



Multiple cells within DSTE

Convergence rates for combined expansions



L^∞ metrics: IVP mixed, DSTE mixed

L^2 metrics: Aleatory, SOP mixed

Impact: render mixed UQ studies practical for large-scale applications

Current:

- **Global or local opt. for epistemic intervals**
  → accuracy *or* scaling w/ epistemic dimension
- **Global or local UQ for aleatory statistics**
  → accuracy *or* scaling w/ aleatory dimension

Future:

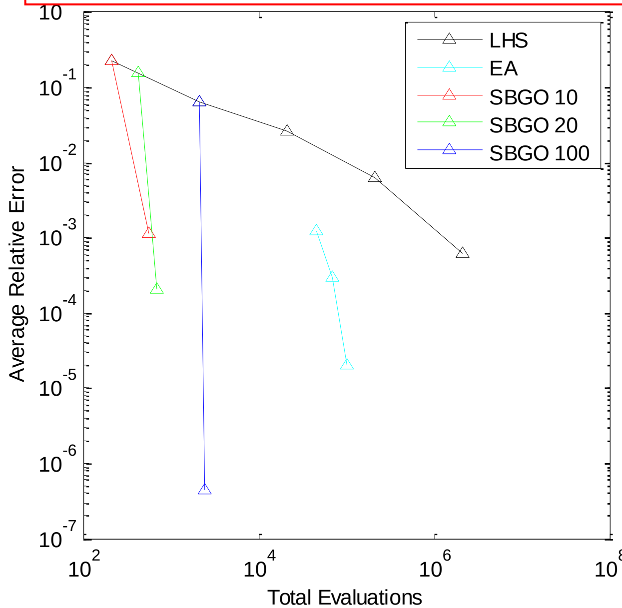
- ***adaptive* and *adjoint-enhanced* global methods**
  → accuracy *and* scaling

Addition of Discrete Epistemic Model Form

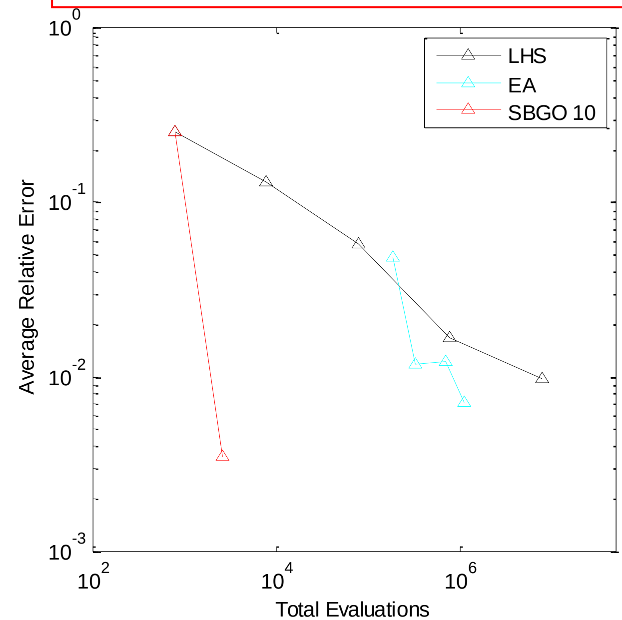
MINLP interval estimation approaches

- Latin hypercube sampling (LHS)
- Evolutionary algorithm (EA)
- Surrogate-based global optimization (SBGO)

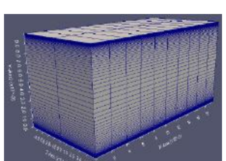
$$\text{Form 1}: \quad f_1 = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$
$$\text{Form 2}: \quad f_2 = 100(x_2 - x_1^2 + .2)^2 + (0.8 - x_1)^2$$

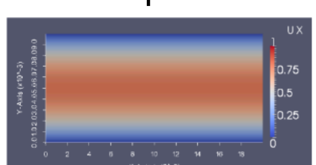


$$\text{Form 1}: \quad f_1 = 1 - \frac{4M}{bh^2Y} - \left(\frac{P}{bhY}\right)^2$$
$$\text{Form 2}: \quad f_2 = 1 - \frac{4P}{bh^2Y} - \left(\frac{P}{bhY}\right)^2$$
$$\text{Form 3}: \quad f_3 = 1 - \frac{4M}{bh^2Y} - \left(\frac{M}{bhY}\right)^2$$
$$\text{Form 4}: \quad f_4 = 1 - \frac{4M}{bh^2Y} - \left(\frac{P}{bhY}\right)^2 - \frac{4(P-M)}{bhY}$$



Drekar RANS turbulence: Spalart-Allmaras, k-ε





Method	Outer Evals	Total Evals	μ_{ux}	$\mu_{pressure}$
LHS	10	250	[0.727604, 2.78150]	[32.6109, 282.237]
SBGO	17	425	[0.622869, 4.44624]	[21.7321, 297.957]