DAKOTA: An Object-Oriented Framework for Simulation-Based Iterative Analysis

Michael S. Eldred

Optimization and Uncertainty Quantification Dept.

Sandia National Laboratories

Albuquerque, NM

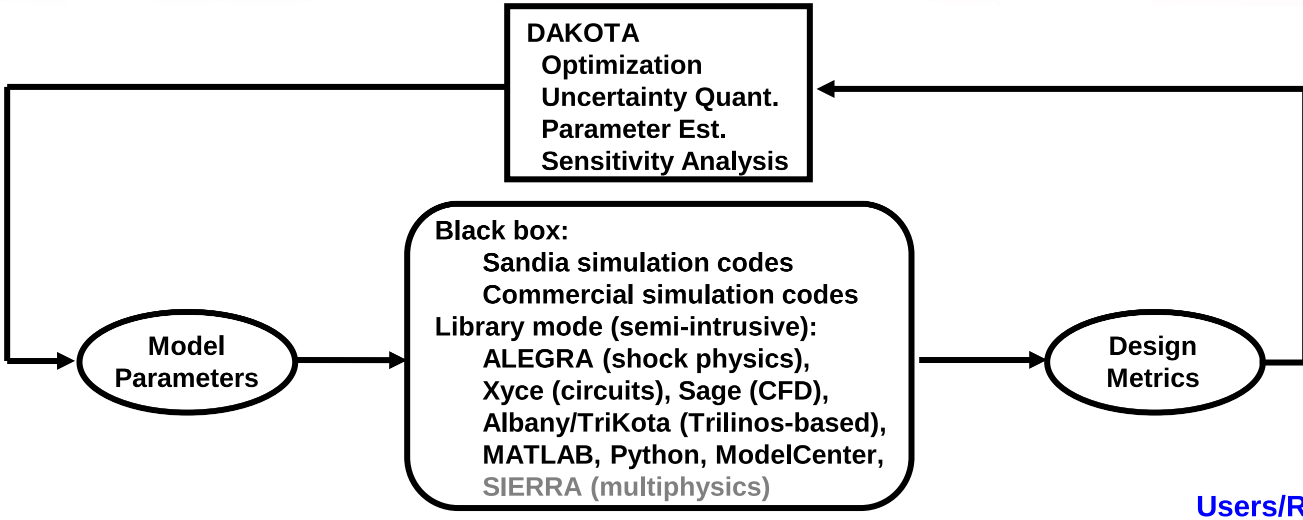
CSDMS 2013 Annual Meeting

Boulder, Colorado

March 23-25, 2013

Sandia is a multi program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

DAKOTA Project



Iterative systems analysis

Multilevel parallel computing

Simulation management

http://dakota.sandia.gov

Users/Ref/Dev/Theory Manuals + training matls. online

Began as LDRD in 1994 focused on optimization

Team: ~10 core personnel in NM/CA + TPL developers

Releases: Major/Interim, Stable/VOTD; ***5.3 released 1/31/13***

DAKOTA Training: ~35 sessions (~1000 students) since 2001

Outreach: Minitutorials at IMAC, SIAM CS&E, ASCR Exascale; SA/UQ short courses at NASA Langley, AFRL WPAFB.

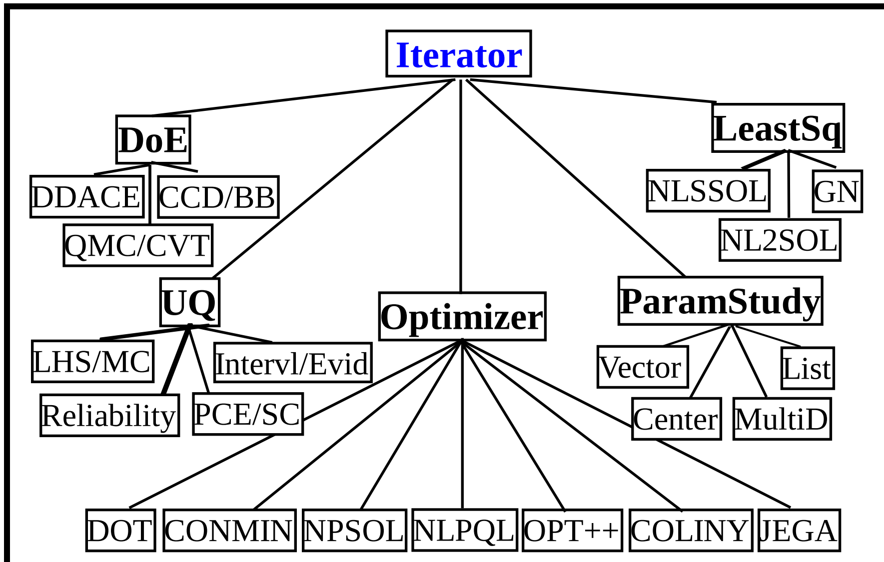
Modern SQE: Nightly builds/testing on Linux, Mac, Windows; subversion, TRAC, Cmake

GNU LGPL: free downloads worldwide (~10k unique ext. registrations, ~3500 distributions per yr.)

Community development: open checkouts now available

Community support: dakota-users, dakota-developers

C++ Framework



Model:

Parameters → **Interface** → Responses

Design
- continuous
- discrete range/set

Uncertain
- normal/lognorm.
- uniform/logunif.
- triang/exp/β/Γ
- EV: I, II, III
- histogram: bin/pt
- discrete: p/b/nb/g/hg
- interval

State
- continuous
- discrete range/set

Application
- system
- fork
- direct
- grid

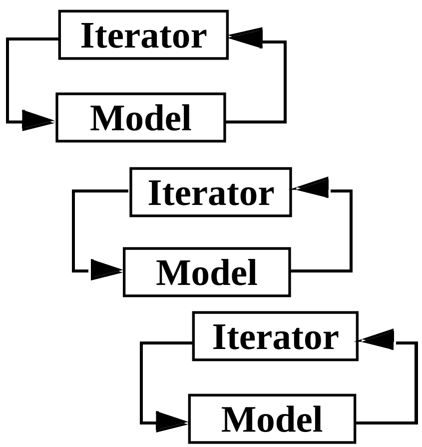
Approximation
- global
  - polynomial 1/2/3, NN, kriging, MARS, RBF
- multipoint – TANA3
- local – Taylor series
- multifidelity
- ROM

Functions
- objectives
- nonlin constraints
- least sq. terms
- generic

Gradients
- numerical
- analytic

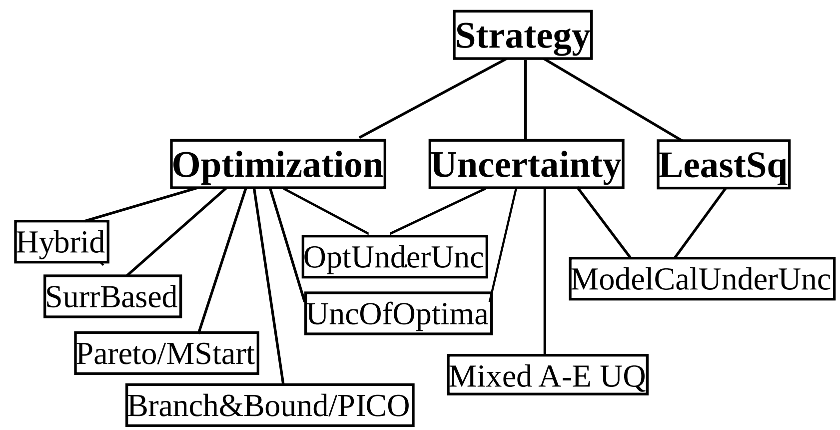
Hessians
- numerical
- analytic
- quasi

Strategy: control of multiple iterators and models



Coordination:
- Nested
- Surrogate
- Recast
- Sequential/Concurrent
- Adaptive/Interactive

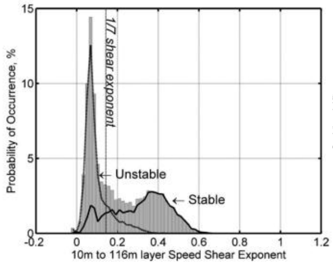
Parallelism:
- Asynchronous local
- Message passing
- Hybrid
- 4 nested levels with
  - Master-slave/dynamic
  - Peer/static

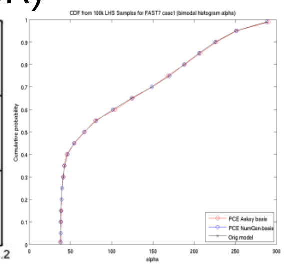


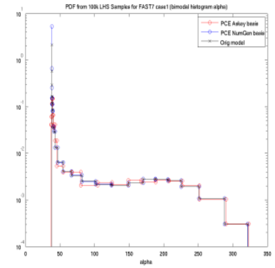
Six input components: Strategy, Method, Model, Variables, Interface, Responses

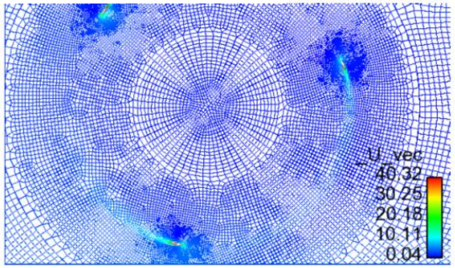
Office of Science Examples

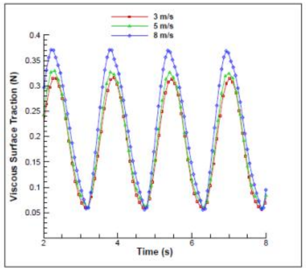
Wind energy (ASCR)



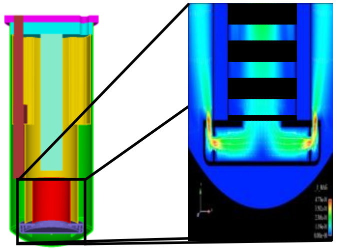


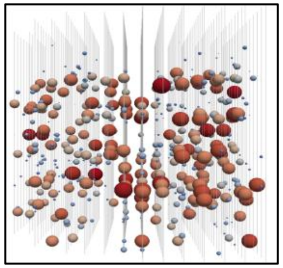




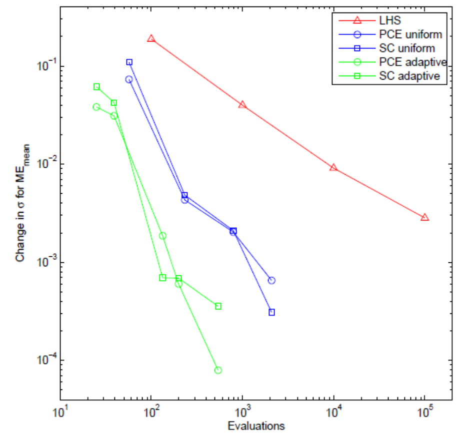


Nuclear reactors (CASL, NEAMS)



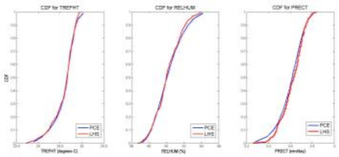


Adaptive refinement for CRUD

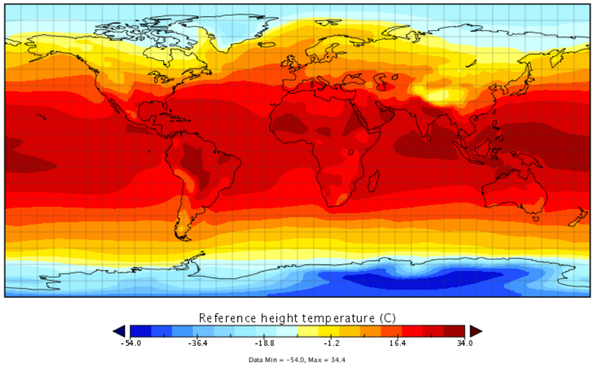


CSSEF: UQ w/ CAM4 (land, ocean, atmosphere)

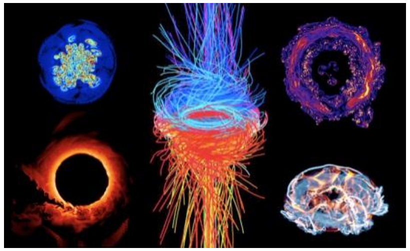
	RHMINL	RHMINH	ALFA	TAU	CZERO	KE
TREFHT	0.33	-0.06	-0.05	0.83	-0.04	0.04
T	0.58	-0.46	-0.35	-0.39	-0.05	0.19
U	-0.17	-0.37	0.07	0.82	-0.01	0.02
PS	0.29	-0.10	0.01	0.62	-0.04	0.03
RELHUM	0.05	0.58	-0.20	-0.74	-0.03	0.15
LHFLX	-0.30	0.31	0.10	0.82	0.01	-0.17
LWCF	-0.23	-0.72	-0.14	-0.59	-0.02	0.12
SWCF	0.92	0.31	0.04	0.21	-0.01	-0.03
PRECT	-0.40	0.38	0.05	0.74	0.03	-0.22
RADBAL	0.97	0.16	-0.03	-0.05	-0.02	0.01

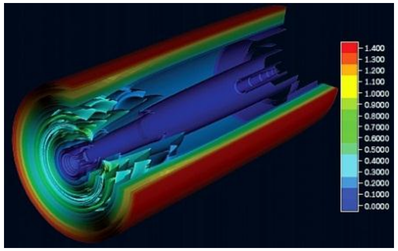


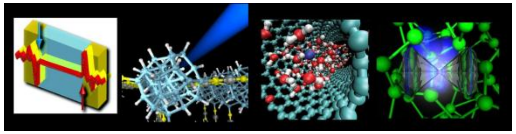
Reference height temperature

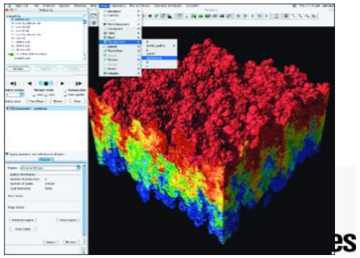


SciDAC-3: Fusion, BER, BES, nuclear, high energy



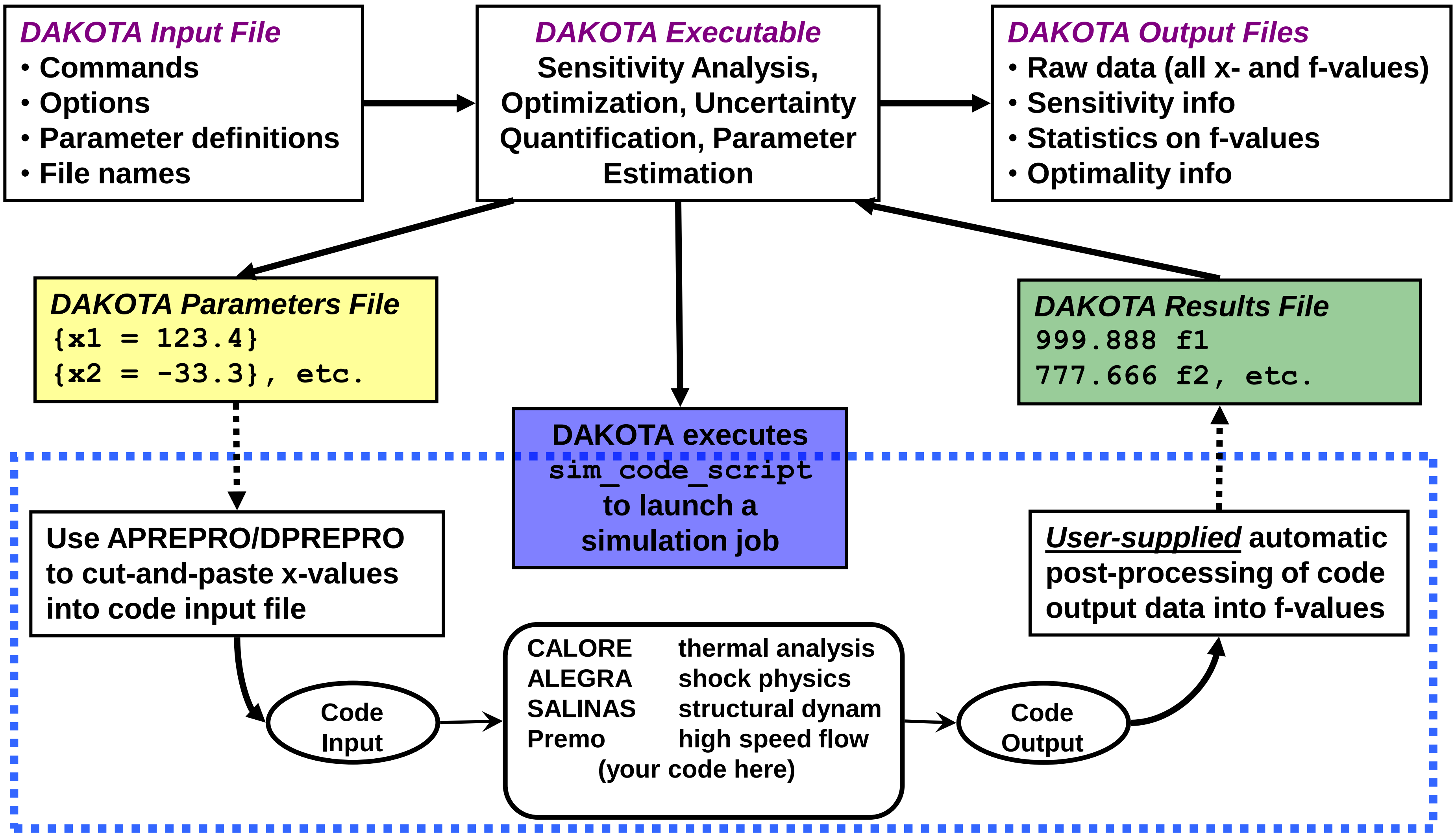






Executing DAKOTA

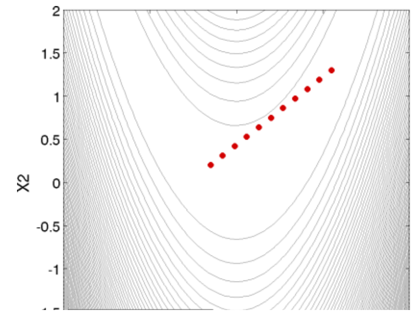
Simulation Management (Black Box case)

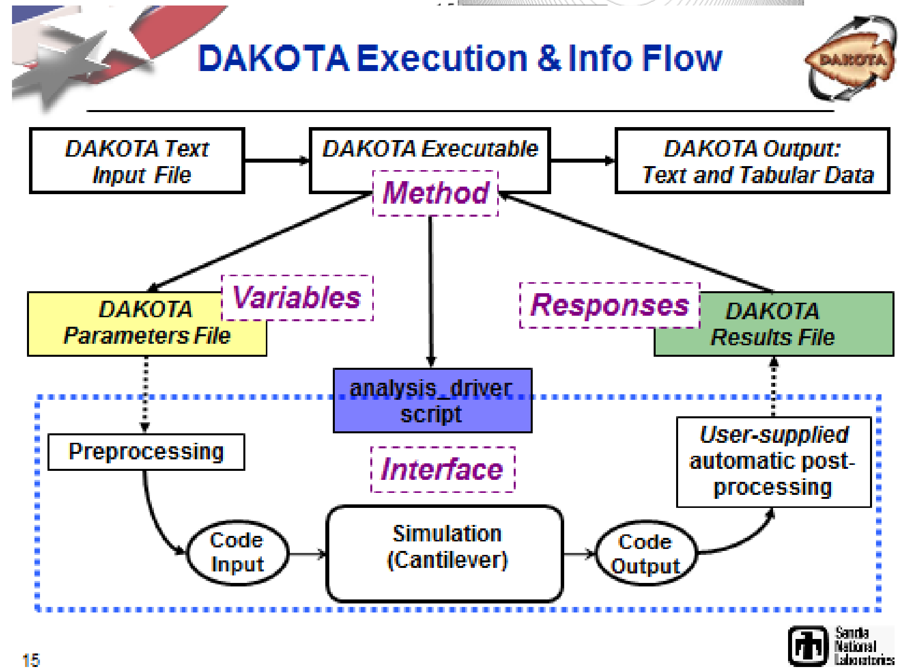


Configure DAKOTA Input File for a Vector Parameter Study

```
strategy
  single_method
    graphics

method
  vector_parameter_study
    num_steps = 10
    step_vector = 0.4 0.4 0. 0. 0. 0.

model
  single

variables
  continuous_design = 2
    initial_point  1.0 1.0
    descriptors    'w' 't'
  continuous_state = 4
    initial_state  40000. 29.E+6 500. 1000.
    descriptors    'R'    'E'    'X'   'Y'

interface
  direct # or system, fork
    analysis_driver = 'mod_cantilever'

responses
  response_functions = 3
  descriptors = 'area' 'stress' 'displacement'
  no_gradients
  no_hessians
```

Define Flow / Algorithm





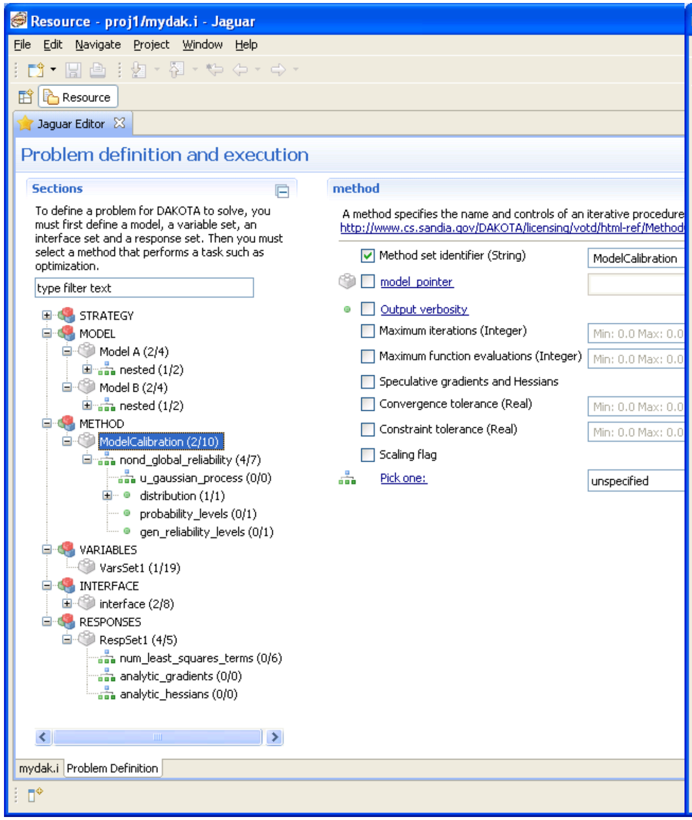
Define Problem / Mapping

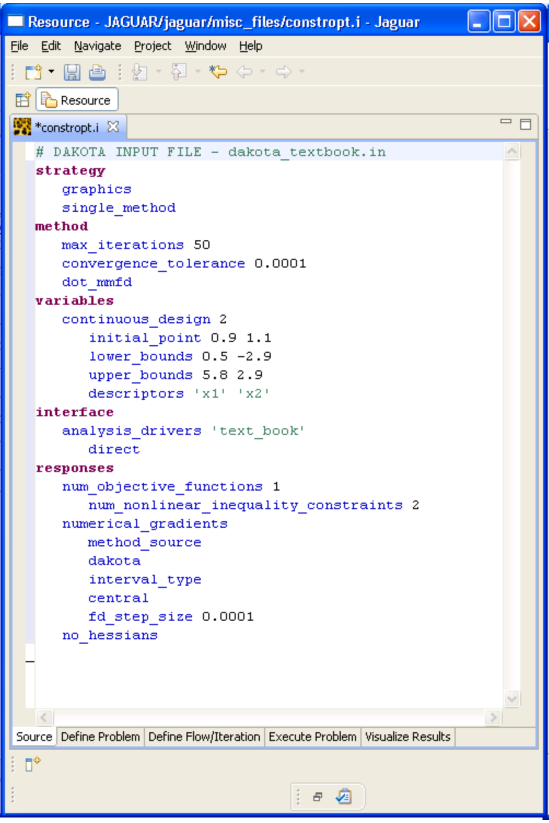
Deployment Initiative: JAGUAR User Interface

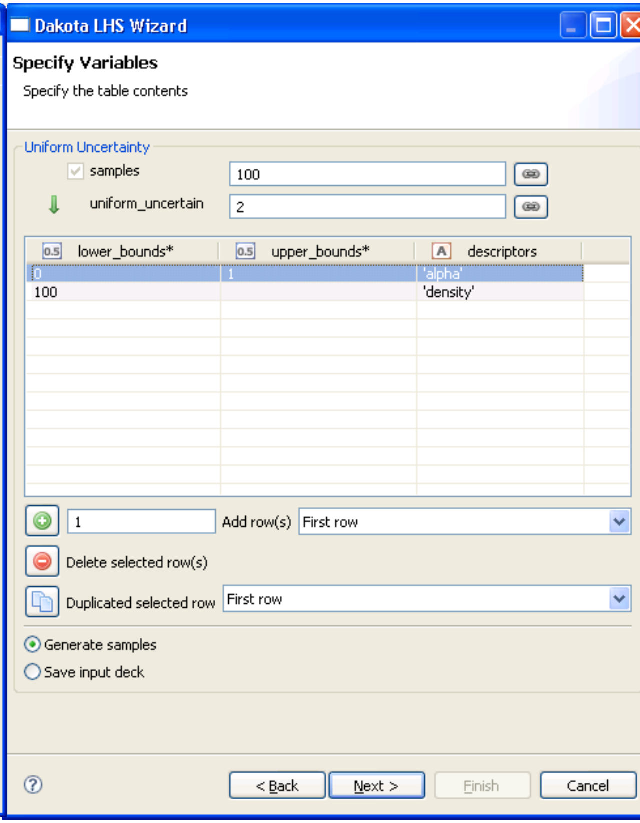
- Eclipse-based rendering of full DAKOTA input spec.
- Automatic syntax updates
- Tool tips, Web links, help
- Symbolics, sim. interfacing

- Flat text editor for experienced users
- Keyword completion
- Automatically synchronized with GUI widgets

- Simplified views for high-use applications (“Wizards”)







Deployment Initiative: Embedding

Make DAKOTA natively available within application codes

- Streamline problem set-up, reduce complexity, and lower barriers
  - A few additional commands within existing simulation input spec.
  - Eliminate analysis driver creation & streamline analysis (e.g., file I/O)
  - Simplify parallel execution
- Integrated options for algorithm intrusion ⇨

SNL Embedding

- Xyce, ALEGRA, Albany (Trikota/Trilinos)
- Planned: SIERRA Toolkit (STK)

External Embedding

- ModelCenter, university applications, R7 (INL), VERA (ORNL)
- Planned: QUESO (UT Austin)
- Expanding our external focus:
  - GPL → LGPL; svn restricted → open network
  - Tailored interfaces & algorithms

ModelEvaluator Levels

Non-intrusive

`ModelEvaluator: systems analysis`

- All residuals eliminated, coupling satisfied
- DAKOTA optimization & UQ

Intrusive to coupling

`ModelEvaluator: multiphysics`

- Individual physics residuals eliminated; coupling enforced by opt/UQ
- DAKOTA opt/UQ & MOOCHO opt.

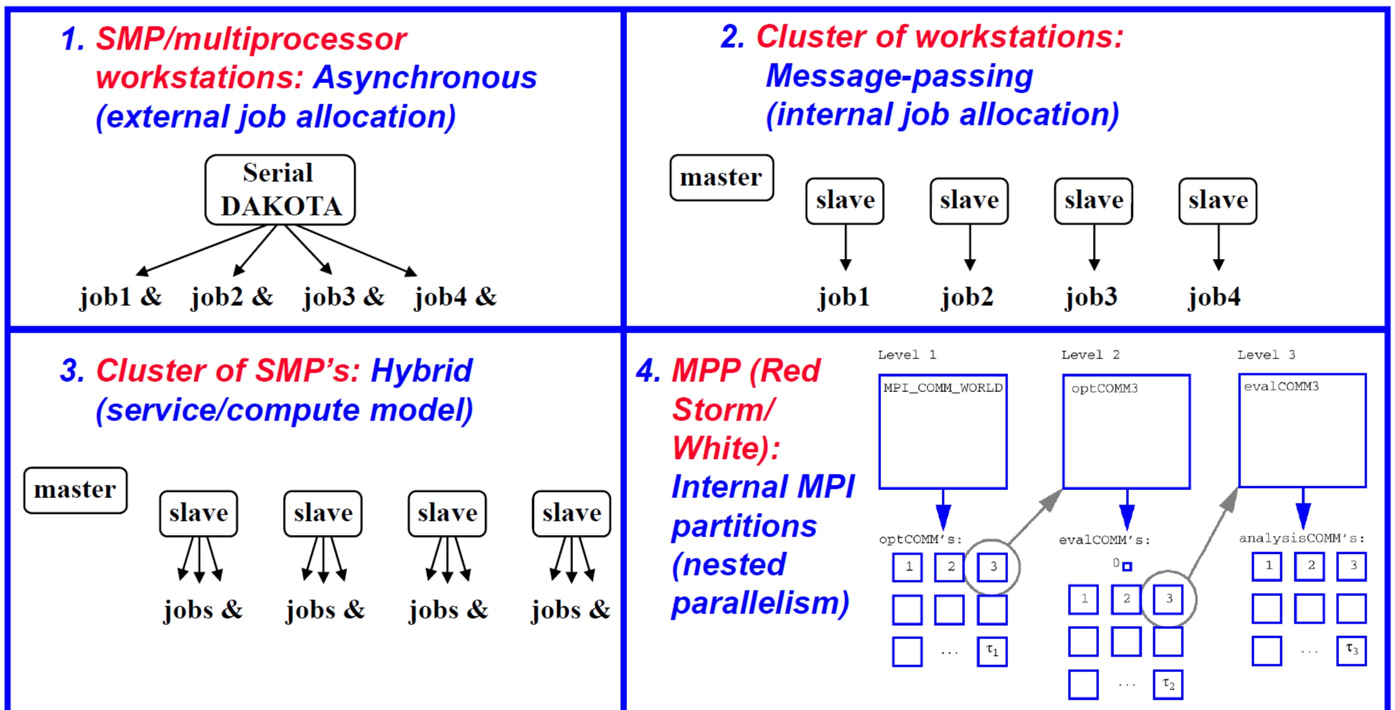
Intrusive to physics

`ModelEvaluator: single physics`

- No residuals eliminated
- MOOCHO opt., Stokhos UQ, NOX, LOCA

Parallelism Options: Multicore Desktops to MPP

1. *Algorithmic coarse-grained*: concurrency in data requests:
   - Iterators: Gradient-based, Nongradient-based, Surrogate-based
   - Strategies with concurrent Iterators: Multi-start, Pareto, Hybrid
   - Nested Models: OUU/MCUU, Mixed UQ
2. *Algorithmic fine-grained:* computing the internal linear algebra of an opt. algorithm in parallel
3. *Fn eval coarse-grained:* concurrent execution of separable simulations within each fn. eval.
4. *Fn eval fine-grained:* parallelization of the solution steps within a single analysis code



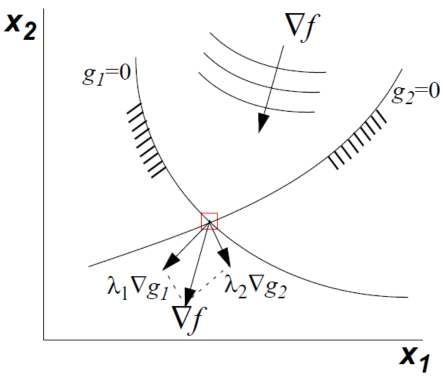
Sandia National Laboratories

Capability Overview
(with emphasis on UQ)

Sandia National Laboratories

Core Methods

Optimization: minimize/maximize objective(s) subject to constraints

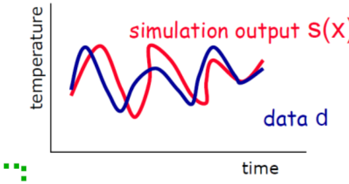


Karush-Kuhn-Tucker conditions:

$$\boxed{\nabla f - \sum_i \lambda_i \nabla g_i = 0}$$

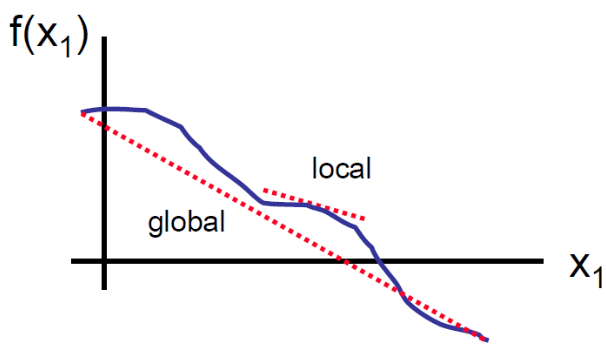
Achieve vector balance: objective fn grad contained within feasibility cone

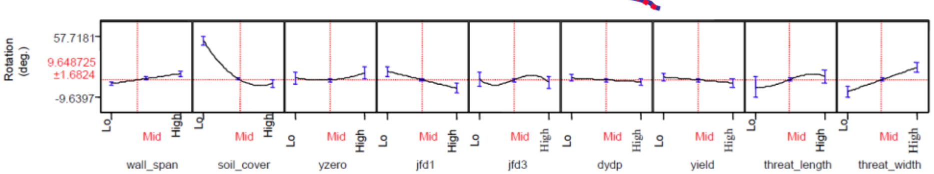
Model Calibration/Parameter Estimation: use nonlinear least squares to minimize errors between model and data



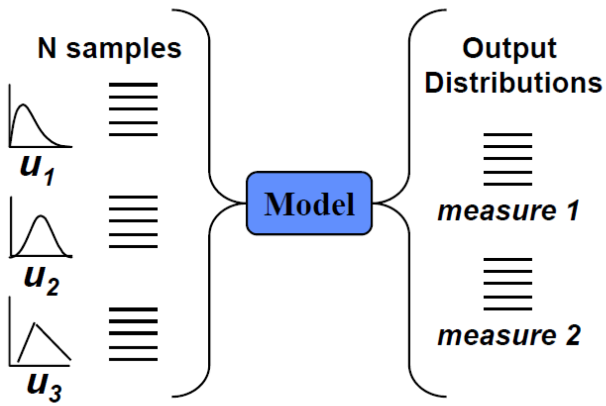
$$f(x) = \sum_{i=1}^{n} (\underbrace{s_i(x)}_{\text{Simulation output that depends on } x} - \underbrace{d_i}_{\text{Given data}})^2$$

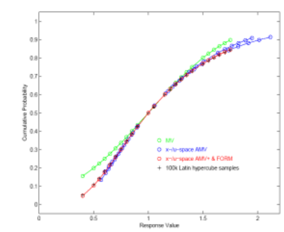
Sensitivity Analysis: identify most influential set of parameters for key response metrics

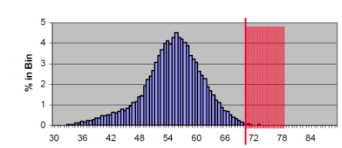




Uncertainty Quantification: quantify effect of random variables on key response metrics







Uncertainty Quantification Algorithms in DAKOTA: New methods bridge robustness/efficiency gap

	Production	Recently released	Under dev.	Future	Collabs.
Sampling	Latin Hypercube, Monte Carlo	Importance, Incremental		Bootstrap, Jackknife	FSU
Reliability	*Local:* Mean Value, 1st- & 2nd-order reliability methods (AMV+,FORM,SORM)	*Global:* adaptive GP methods (EGRA, GPAIS) with grad-enhancement	POF Darts	recursive emulation, TGP	*Local:* Notre Dame, *Global:* Vanderbilt
Stochastic expansion		PCE/SC with dimension-adaptive p-/h-refinement, grad-enhancement, compressed sensing	Local adapt refinement, orthog least interp, adjoint EE correction	Discrete random variables	Stanford, Purdue
Other probabilistic		Eff Subspace, Morris-Smale topology	Rand fields / stoch process	EDR	NCSU, Utah, Cornell, Maryland
Epistemic & Bayesian	Interval-valued/ Second-order prob. (nested sampling)	Opt-based interval est, Dempster-Shafer, model form, QUESO, GPMSA	Discrete GPs, model selection	Imprecise probability	LANL, UT Austin
Metrics & Global SA	Importance factors, rank/partial corr, main effects	Variance-based decomposition	Stepwise regression		LANL

Uncertainty Quantification Algorithms in DAKOTA: New methods bridge robustness/efficiency gap

	Production	Recently released	Under dev.	Future	Collabs.
Sampling	Latin Hypercube, Monte Carlo	Importance, Incremental		Bootstrap, Jackknife	FSU
Reliability	*Local:* Mean Value, 1st- & 2nd-order reliability methods (AMV+,FORM,SORM)	*Global:* adaptive GP methods (EGRA, G[illegible] er[illegible]	POF Darts	recursive emulation, [illegible]	*Local:* Notre Dame, *Global:* Vanderbilt
Stochastic expansion		PCE/SC with dimension-adaptive p-/h-refinement, grad-enhancement, compressed sensing	Local adapt refinement, orthog least interp, adjoint EE correction	Discrete random variables	Stanford, Purdue
Other probabilistic		Eff Subspace, Morris-Smale topology	Rand fields / stoch process	EDR	NCSU, Utah, Cornell, Maryland
Epistemic & Bayesian	Interval-valued/ Second-order prob. (nested sampling)	Opt-based interval est, Dempster-Shafer, model form, QUESO, GPMSA	Discrete GPs, model selection	Imprecise probability	LANL, UT Austin
Metrics & Global SA	Importance factors, rank/partial corr, main effects	Variance-based decomposition	Stepwise regression		LANL

Research: Scalability, Robustness, Goal-orientation

Uncertainty Quantification Algorithms in DAKOTA: New methods bridge robustness/efficiency gap

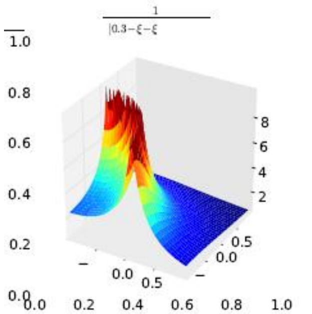
	Production	Recently released	Under dev.	Future	Collabs.
Sampling	Latin Hypercube, Monte Carlo	Importance, Incremental		Bootstrap, Jackknife	FSU
Reliability	*Local:* Mean Value, 1st- & 2nd-order reliability methods (AMV+,FORM,SORM)	*Global:* adaptive GP methods (EGRA, [illegible]	POF Darts	recursive emulation, [illegible]	*Local:* Notre Dame, *Global:* Vanderbilt
Stochastic expansion		PCE/SC with dimension-adaptive p-/h-refinement, grad-enhancement, compressed sensing	Local adapt refinement, orthog least interp, adjoint EE correction	Discrete random variables	Stanford, Purdue
Other probabilistic		Eff Subspace, Morris-Smale topology	Rand fields / stoch process	EDR	NCSU, Utah, Cornell, Maryland
Epistemic & Bayesian	Interval-valued/ Second-order prob. (nested sampling)	Opt-based interval est, Dempster-Shafer, model form, QUESO, GPMSA	Discrete GPs, model selection	Imprecise probability	LANL, UT Austin
Metrics & Global SA	Importance factors, rank/partial corr, main effects	Variance-based decomposition	Stepwise regression		LANL

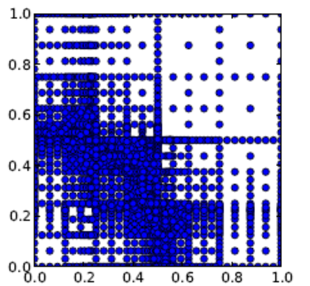
Research: Scalability, Robustness, Goal-orientation

Adv. Deployment

Fills Gaps

Scalable Methods for High-Dimensional UQ



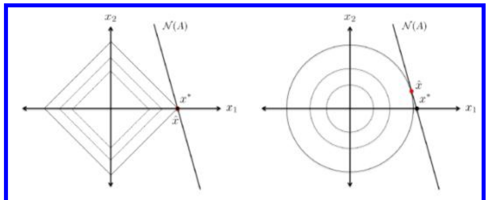


Key Challenges:

- Severe simulation budget constraints (e.g., a handful of HF runs)
- Moderate to high-dimensional in random variables: $O(10^1)$ to $O(10^2)$
- Compounding effects:
  - Mixed aleatory-epistemic uncertainties (→ nested iteration)
  - Requirement to evaluate probability of rare events (e.g., safety criteria)
  - Nonsmooth responses (→ difficulty with global basis spectral methods)

Algorithmic Capabilities:

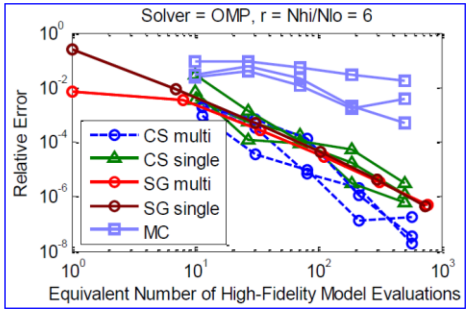
- Compute dominant uncertainty effects despite key challenges above
- Scalable UQ foundation
  - Goal-oriented adaptive refinement to reduce effective dimension
  - Adjoint techniques [given n (random dimension) > m (response QoI)]
  - Sparsity detection methods: compressive sensing, least interpolation
- Leverage foundation within higher-level studies
  - Multifidelity UQ
  - Mixed aleatory-epistemic UQ including model form
  - Bayesian inference, Optimization/calibration under uncertainty



(a) CS methodology (ℓ_1 objective) (b) Pseudo-inverse (ℓ_2 objective)

Mission Relevancy:

- NNSA: ASC
- Office of science: ASCR, SciDAC-3, CASL
- New UQ capabilities recently deployed in Dakota v5.3 (1/31/13)



Non-Intrusive Stochastic Expansions: Polynomial Chaos and Stochastic Collocation

Polynomial chaos: spectral projection using orthogonal polynomial basis fns

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

using

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

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

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

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

Stochastic collocation: instead of estimating coefficients for known basis functions, form <u>interpolants</u> 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

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

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: fault tolerance, decay, sparsity, error est.

Non-Intrusive Stochastic Expansions: Polynomial Chaos and Stochastic Collocation

Polynomial chaos: spectral projection using orthogonal polynomial basis fns

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

using

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

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

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

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

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

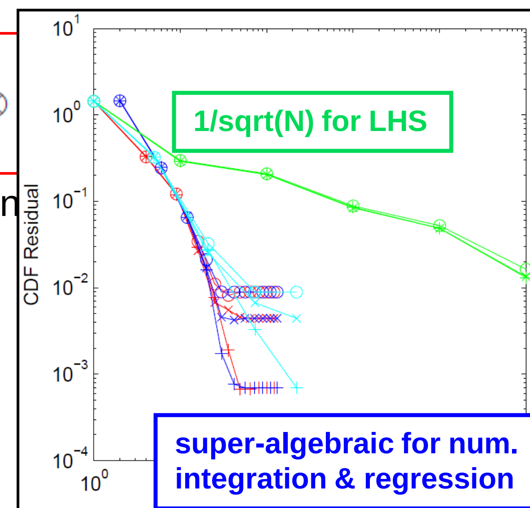
$$R = \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi})$$

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

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

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.



Approaches for forming PCE/SC Expansions

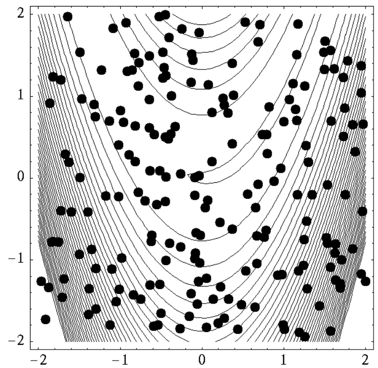
Random sampling: PCE

Expectation (sampling):

- Sample w/i distribution of ξ
- Compute expected value of product of R and each Ψ_j

Linear regression ("point collocation"):

- Sample w/i distribution of ξ
- Solves least squares data fit for all coefficients at once:



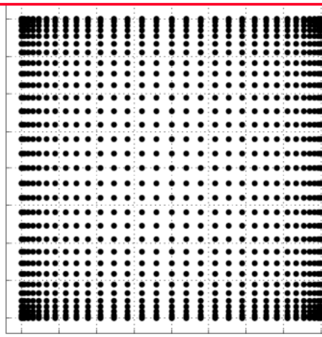
$$\boxed{\Psi \alpha = R}$$

Tensor-product quadrature: PCE/SC

$$\boxed{\mathscr{U}^i(f)(\xi) = \sum_{j=1}^{m_i} f(\xi_j^i)\, w_j^i}$$

$$\boxed{\mathcal{Q}_{\mathbf{i}}^n f(\xi) = \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_n}\right)(f)(\xi) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} f\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right)\left(w_{j_1}^{i_1} \otimes \cdots \otimes w_{j_n}^{i_n}\right)}$$

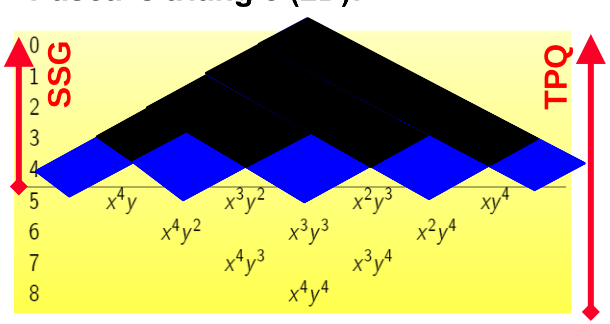
- **Every combination of 1-D rules**
- **Scales as m^n**
- **1-D Gaussian rule of order m → integrands to order $2m - 1$**
- **Assuming $R\,\Psi_j$ of order $2p$, select $m = p + 1$**

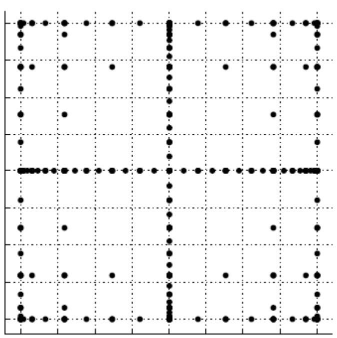


Smolyak Sparse Grid: PCE/SC

$$\boxed{\mathscr{A}(\mathrm{w}, n) = \sum_{\mathrm{w}+1 \le |\mathbf{i}| \le \mathrm{w}+n} (-1)^{\mathrm{w}+n-|\mathbf{i}|} \binom{n-1}{\mathrm{w}+n-|\mathbf{i}|} \cdot \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_n}\right)}$$

Pascal's triangle (2D):



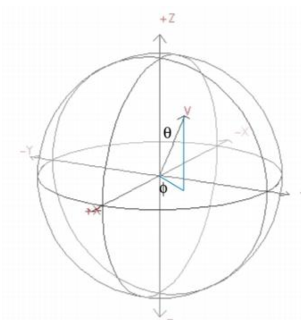


Cubature: PCE

Stroud and extensions (Xiu, Cools)

→ Low order PCE

→ global SA, anisotropy detection



Gaussian $i = 2$ → $p = 1$

$$\boxed{x_{k,2r-1} = \sqrt{2}\cos\frac{2rk\pi}{n+1}, \qquad x_{k,2r} = \sqrt{2}\sin\frac{2rk\pi}{n+1}}$$

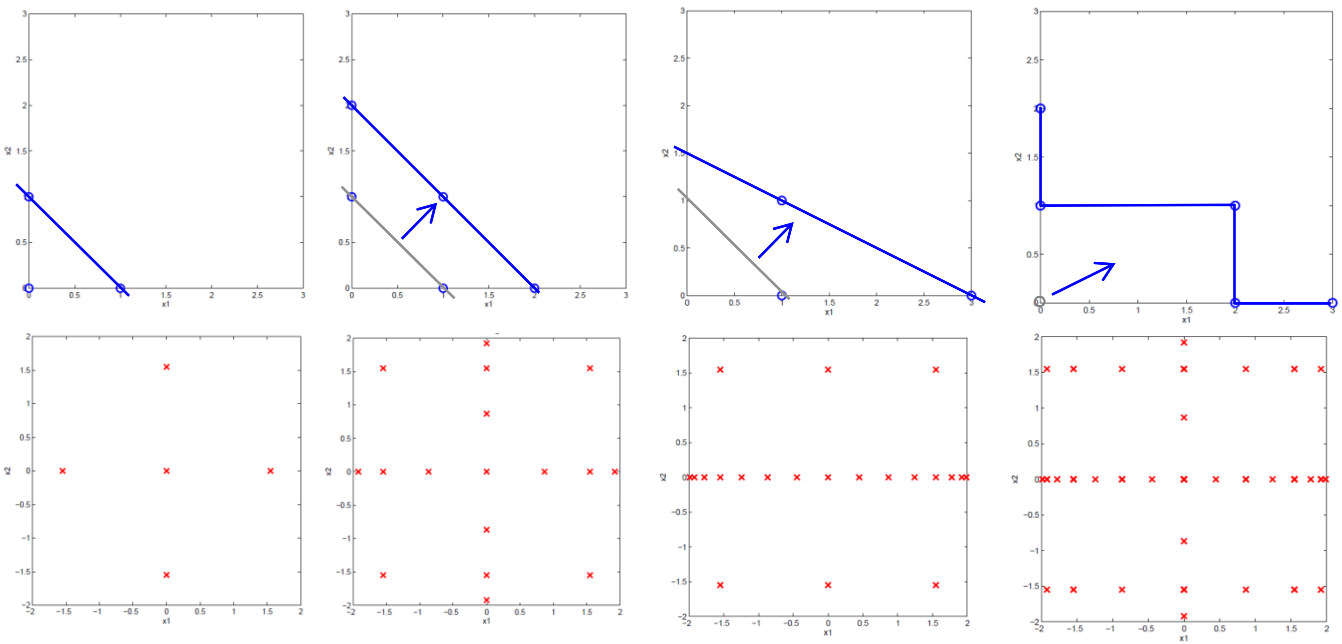
Arbitrary PDF $\boxed{t^{(k)} = \frac{1}{\gamma}\left[\sqrt{\gamma c_1}\, x^{(k)} - \delta\right]}$

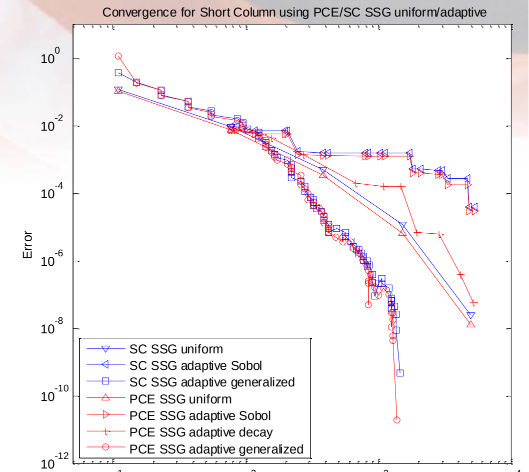
Sandia National Laboratories

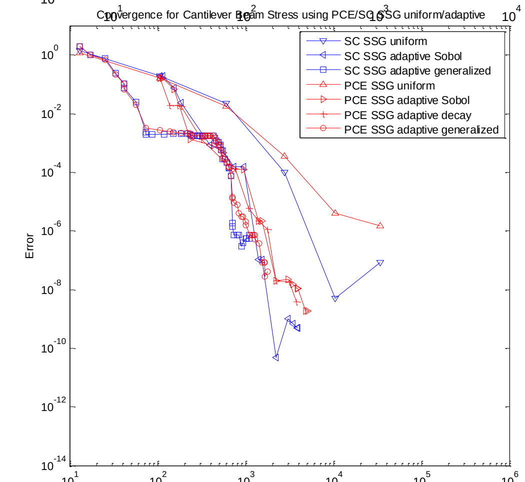
Stochastic Expansions on Structured Grids: Adaptive Collocation Methods

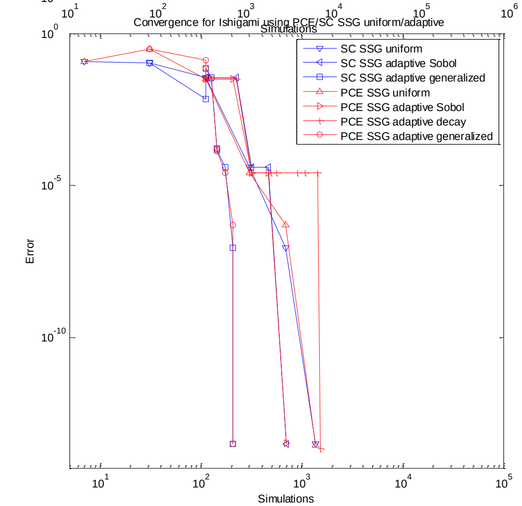
Polynomial order (p-) refinement approaches:

- **Uniform:** *isotropic* tensor/sparse grids
  - *Increment grid:* increase order/level, ensure change (restricted/nested)
  - *Assess convergence:* L^2 change in response covariance
- **Adaptive:** *anisotropic* tensor/sparse grids $\boxed{w\underline{\gamma} < \mathbf{i}\cdot\gamma \le w\underline{\gamma} + |\gamma|}$
  - **PCE/SC:** variance-based decomp. → total Sobol' indices → anisotropy
  - **PCE:** spectral coefficient decay rates → anisotropy
- **Goal-oriented adaptive:** *generalized* sparse grids
  - **PCE/SC:** change in QOI induced by trial index sets on active front
  - Fine-grained control: frontier not limited by index set constraint







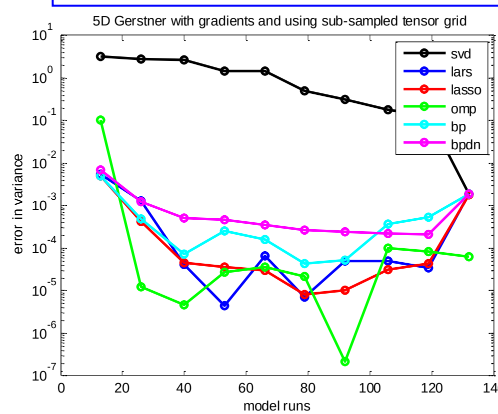


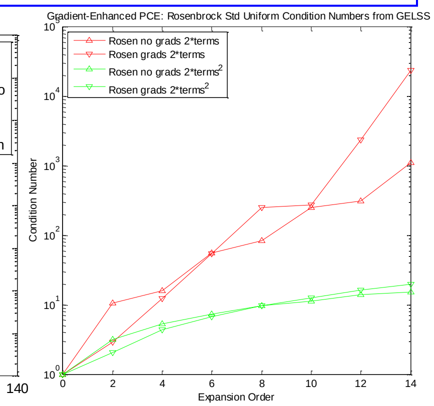
Extend Scalability: (Adjoint) Derivative-Enhancement

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

PCE:

- Linear regression including derivatives
  - Gradients/Hessians → addtnl. eqns.
  - Over-determined: SVD, eq-constrained LS
  - Under-determined: compressive sensing



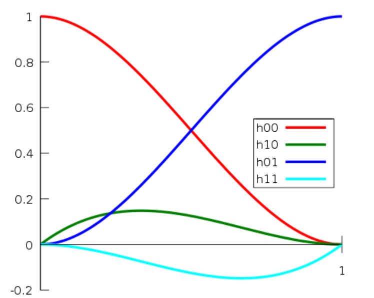


SC:

- Gradient-enhanced interpolants
  - Local: cubic Hermite splines
  - Global: Hermite interpolating polynomials

$$f = \sum_{i=1}^{N} f_i H_i^{(1)}(x_1) H_i^{(1)}(x_2) H_i^{(1)}(x_3) + \sum_{i=1}^{N} \frac{df_i}{dx_1} H_i^{(2)}(x_1) H_i^{(1)}(x_2) H_i^{(1)}(x_3) + \sum_{i=1}^{N} \frac{df_i}{dx_2} H_i^{(1)}(x_1) H_i^{(2)}(x_2) H_i^{(1)}(x_3) + \sum_{i=1}^{N} \frac{df_i}{dx_3} H_i^{(1)}(x_1) H_i^{(1)}(x_2) H_i^{(2)}(x_3)$$

Cubic shape fns: type 1 (value) & type 2 (gradient)

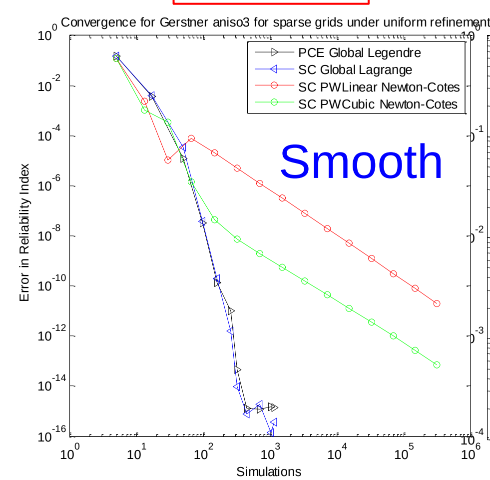


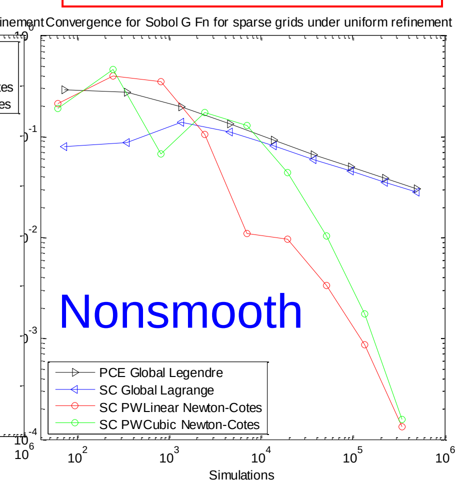
$$\mu = \sum_{i=1}^{N} f_i w_i^{(1)} w_i^{(1)} w_i^{(1)} + \sum_{i=1}^{N} \frac{df_i}{dx_1} w_i^{(2)} w_i^{(1)} w_i^{(1)} + \sum_{i=1}^{N} \frac{df_i}{dx_2} w_i^{(1)} w_i^{(2)} w_i^{(1)} + \sum_{i=1}^{N} \frac{df_i}{dx_3} w_i^{(1)} w_i^{(1)} w_i^{(2)}$$

and similar for higher-order moments

$$e^{-10x^2-5y^2}$$

$$f(\mathbf{x}) = 2\prod_{j=1}^{5} \frac{|4x_j - 2| + a_j}{1 + a_j}; \quad a = [0,1,2,4,8]$$

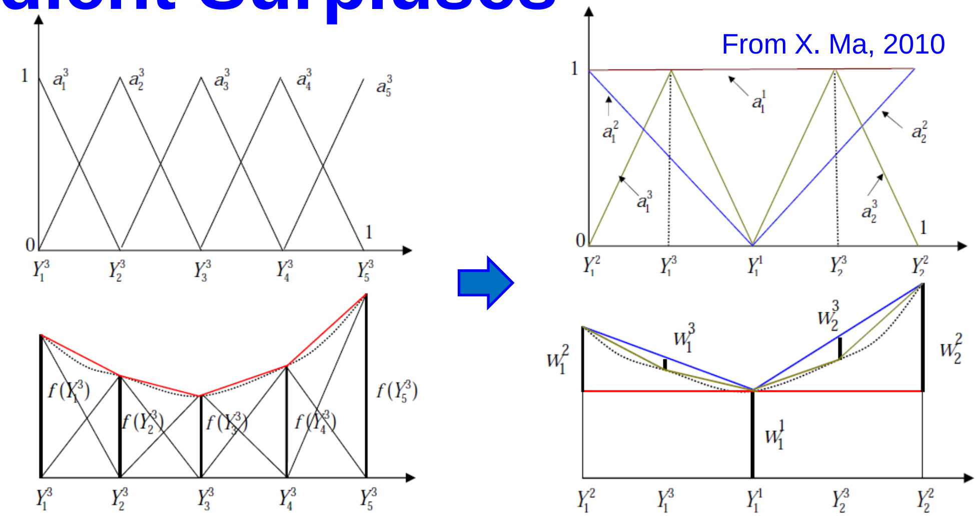




Local Error Estimation with Hierarchical Value/Gradient Surpluses

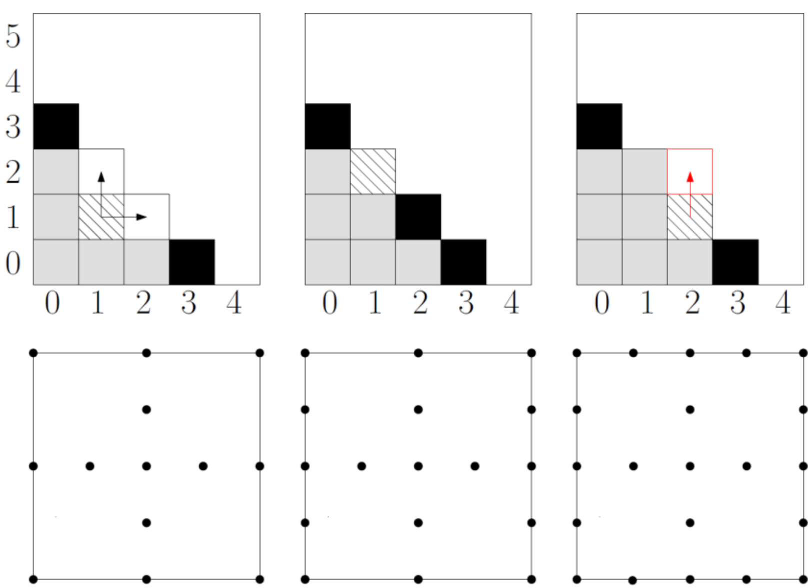
Hierarchical basis:

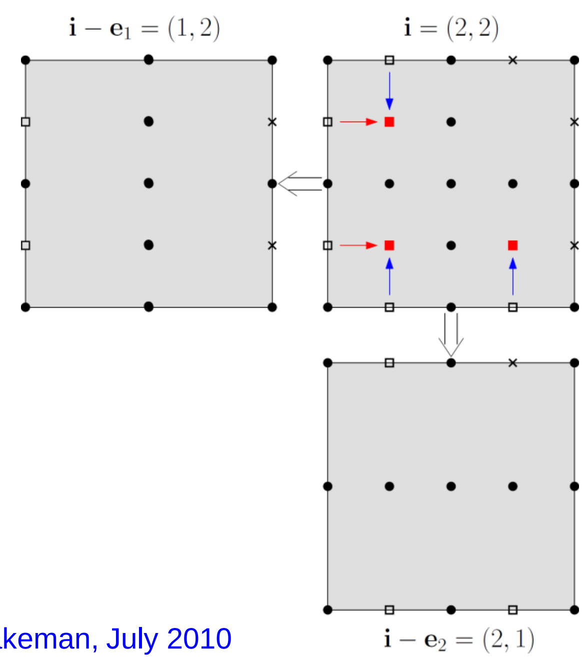
- *Improved precision* in QoI increments
- Surpluses provide error estimates for *local refinement* using local/global hierarchical interpolants
- New error indicators under development that leverage both *value and gradient surpluses*

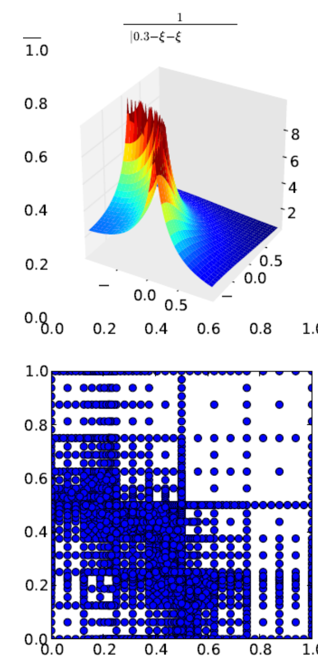


From X. Ma, 2010

$$\Delta\Sigma_{ij} = \Delta E[R_i R_j] - \mu_{R_i}\,\Delta E[R_j] - \mu_{R_j}\,\Delta E[R_i] - \Delta E[R_i]\Delta E[R_j] \rightarrow \Delta\sigma, \Delta\beta$$







From J. Jakeman, July 2010

Stochastic Expansions on Unstructured Grids: Compressive Sensing

$$\begin{bmatrix} f(\mathbf{x}^{(1)}) \\ f(\mathbf{x}^{(2)}) \\ \vdots \\ f(\mathbf{x}^{(N)}) \end{bmatrix} = \begin{bmatrix} 1 & \Phi_2(\mathbf{x}^{(1)}) & \Phi_2(\mathbf{x}^{(1)}) & \dots & \Phi_P(\mathbf{x}^{(1)}) \\ 1 & \Phi_1(\mathbf{x}^{(2)}) & \Phi_2(\mathbf{x}^{(2)}) & \dots & \Phi_P(\mathbf{x}^{(2)}) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \Phi_1(\mathbf{x}^{(N)}) & \Phi_2(\mathbf{x}^{(N)}) & \dots & \Phi_P(\mathbf{x}^{(N)}) \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_P \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

or in matrix notation

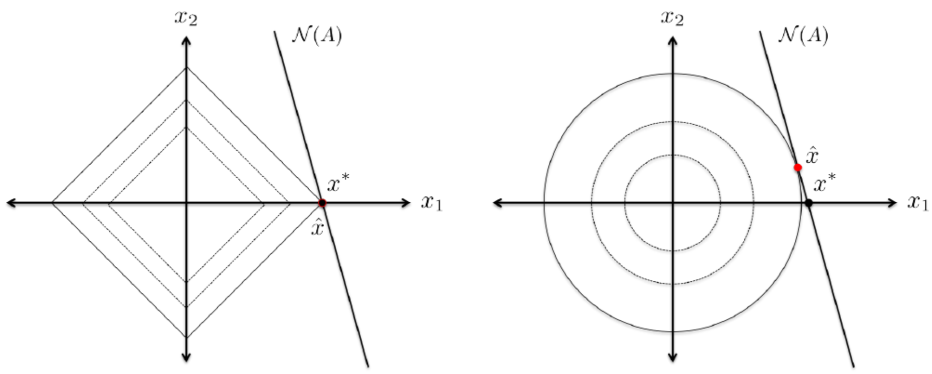
$$\mathbf{b} = \mathbf{A}\mathbf{x} + \varepsilon$$

and find the minimum norm solution

$$\min_{\mathbf{x}} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$$

or (more recently) find a sparse solution

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{such that} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 \le \varepsilon$$



(a) CS methodology (ℓ_1 objective) (b) Pseudo-inverse (ℓ_2 objective)

BP

$$\mathbf{c} = \arg\min \|\mathbf{c}\|_{\ell^1} \quad \text{such that} \quad \mathbf{\Phi c} = \mathbf{y}$$

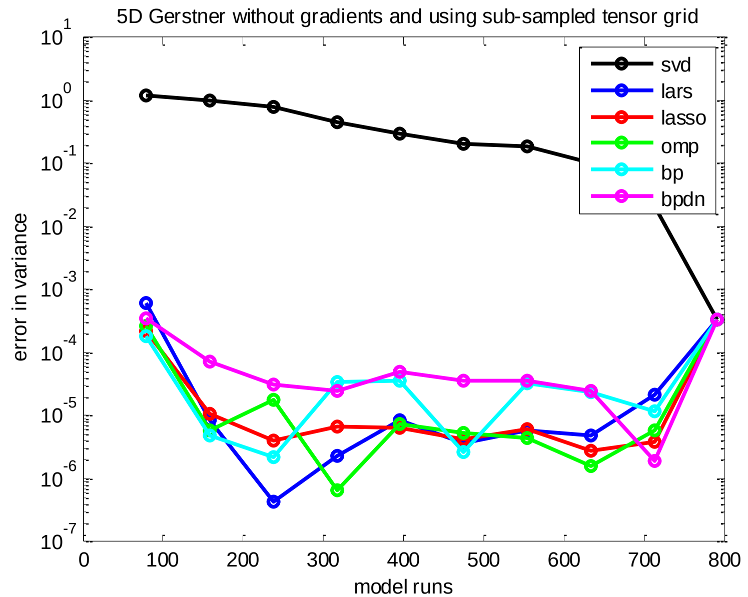
BPDN and OMP

$$\mathbf{c} = \arg\min \|\mathbf{c}\|_{\ell^1} \quad \text{such that} \quad \|\mathbf{\Phi c} - \mathbf{y}\|_{\ell^2} \le \varepsilon$$

LASSO and LARS

$$\mathbf{c} = \arg\min \|\mathbf{\Phi c} - \mathbf{y}\|_{\ell^2}^2 \quad \text{such that} \|\mathbf{x}\|_{\ell^1} \le \tau$$

Structured or unstructured grids
Value-based or gradient-enhanced



Multiple Model Forms in UQ

[Discrete model choices for same physics]

- A clear hierarchy of fidelity (low to high)

Low

Med

High

- An ensemble of models that are all credible (lacking a clear preference structure)

SA-RANS | KE-RANS-NBC | KE-RANS-DBC

  - With data: Bayesian model selection
  - *Without data:* epistemic model form uncertainty propagation

- Both

Potential flow | Vortex lattice

SA-RANS | KE-RANS-NBC | KE-RANS-DBC

Smagorinsky-LES | Germano-LES

DNS

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

Multifidelity UQ using Stochastic Expansions

Motivation:

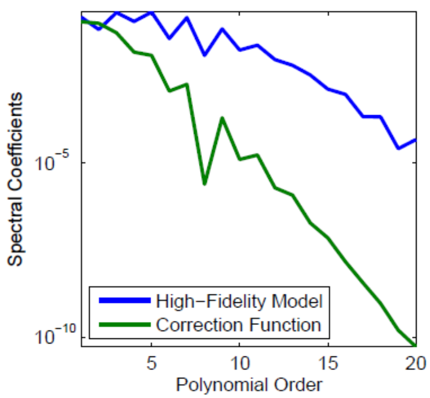
- High-fidelity simulations (e.g., RANS, LES) can be prohibitive for use in UQ
- Low fidelity “design” codes often exist that are predictive of basic trends
- Can we leverage LF codes w/i HF UQ in a rigorous manner? → global approxs. of model discrepancy

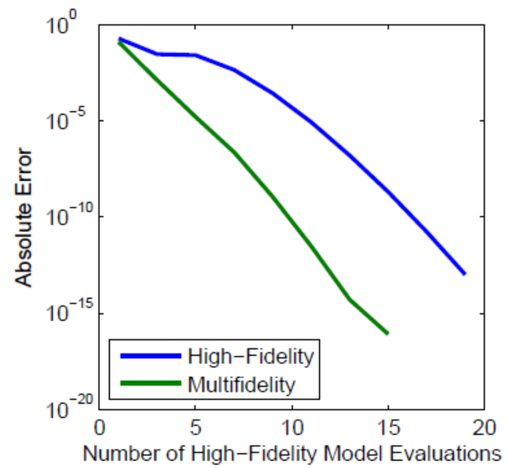
$$\boxed{\hat{f}_{hi}(\xi) = \sum_{j=1}^{N_{lo}} f_{lo}(\xi_j) L_j(\xi) + \sum_{j=1}^{N_{hi}} \Delta f(\xi_j) L_j(\xi)}$$

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

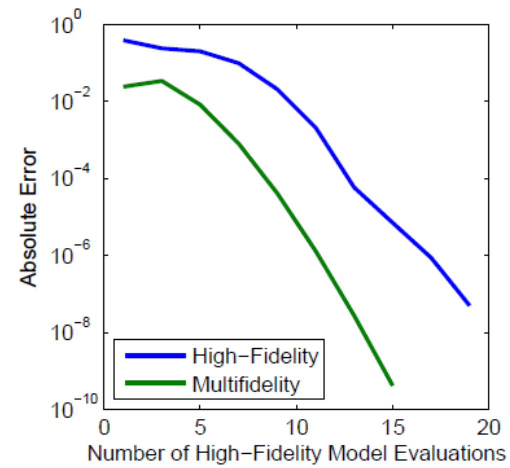
$$R_{\text{high}}(\xi) = e^{-0.05\xi^2} \cos 0.5\xi \; \underbrace{- \; 0.5e^{-0.02(\xi-5)^2}}_{\text{discrepancy}}$$

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





(a) Error in mean



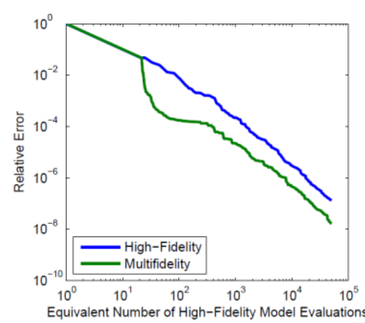
(b) Error in standard deviation

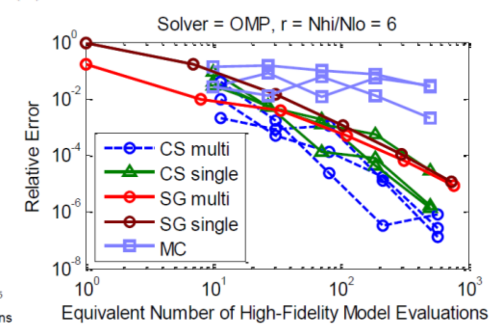
Adaptive sparse grid multifidelity algorithm:

- Gen. sparse grids for LF & discrepancy levels
- Greedy selection from grids: max ΔQoI/ΔCost
- Refine discrepancy where LF is less predictive

Compressive sensing multifidelity algorithm:

- Target sparsity within the model discrepancy





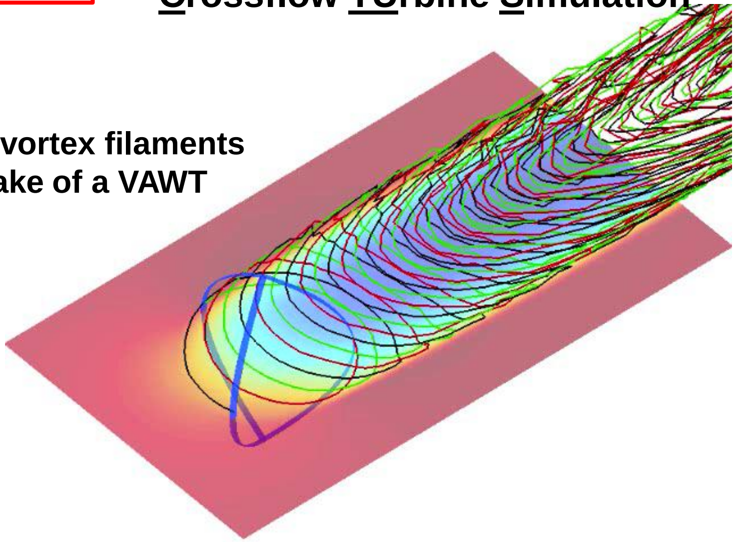
ASCR MF UQ example: VAWT CFD/FSI Modeling

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

Concluding Remarks

DAKOTA provides a variety of core algorithms for iterative analysis:

- *Optimization* -- *Sensitivity Analysis*
- *Calibration* -- *Uncertainty quantification*

as well as advanced capabilities for

- *Multilevel parallel computing*
- *Managing multiple iterative methods, models of varying fidelity, surrogates, nesting, recasting, etc.*

and advanced deployment initiatives that "lower the bar" for adoption

- *JAGUAR* -- *Library embedding*

UQ deployment faces a number of key challenges

- Severe simulation budget constraints and moderate to high random dimensionality
- Compounded by mixed uncertainties, nonsmoothness, rare events

Investments in scalable UQ R&D

- We are developing a broad suite of scalable and robust core UQ methods:
  - Goal-oriented adaptive refinement, (Adjoint) gradient-enhancement, Sparsity detection
- We are building on this foundation
  - Multifidelity UQ, Mixed aleatory-epistemic UQ, Bayesian inference

Impact and deployment

- Latest algorithm R&D deployed through DAKOTA (v5.3 released 1/31/13)
- Impact on NNSA (ASC) & Office of Science (ASCR, CASL, SciDAC)

Sandia National Laboratories

Extra Slides

List of Acronyms

ASCR: advanced scientific computing research
BP/BPDN: basis pursuit (denoising)
CASL: consortium for advanced simulation of light water reactors
CDF: cumulative distribution function
CFD: computational fluid dynamics
CS: compressive sensing
DNS: direct numerical simulation
EGRA: efficient global reliability analysis
FSI: fluid-structure interaction
GA/MOGA: (multiobjective) genetic algorithm
GP/GPAIS: gaussian process (adaptive importance sampling)
LARS: least angle regression
LASSO: least absolute shrinkage and selection operator
LES: large eddy simulation
LF/HF: low/high fidelity
LHS: latin hypercube sampling
LS/NLS: (nonlinear) least squares
MC: Monte Carlo
MVFOSM: mean value first-order second-moment
OMP: orthogonal matching pursuit
PCE: polynomial chaos expansion
PDF: probability density function
POD: proper orthogonal decomposition
POF: probability of failure
QoI: quantity of interest
RANS: Reynolds-averaged Navier-Stokes
SC: stochastic collocation
SciDAC: scientific discovery through advanced computing
SVD: singular value decomposition
UQ: uncertainty quantification
VAWT: vertical axis wind turbine
VBD: variance-based decomposition

Additional DAKOTA Resources

References

- Full list of research publications: http://dakota.sandia.gov/publications.html
- Selected application examples: http://dakota.sandia.gov/applications.html
- DAKOTA documentation: http://dakota.sandia.gov/documentation.html (see Theory, Users, and Reference Manuals)

Software Downloads

- DAKOTA: http://dakota.sandia.gov/download.html
- Version 5.3 release notes: http://dakota.sandia.gov/distributions/dakota/5.3/release-notes.html
- Related packages: http://dakota.sandia.gov/packages.html