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
**DAKOTA Project**

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

---

**DAKOTA Optimization**

**Uncertainty Quant.**

**Parameter Est.**

**Sensitivity Analysis**

**Black box:**
- Sandia simulation codes
- Commercial simulation codes
- Library mode (semi-intrusive):
  - ALEGRA (shock physics), Xyce (circuits), Sage (CFD), Albany/TriKota (Trilinos-based), MATLAB, Python, ModelCenter, SIERRA (multiphysics)

**Model Parameters**

**Design Metrics**

---

**Iterative systems analysis**

**Multilevel parallel computing**

**Simulation management**

**http://dakota.sandia.gov**

Users/Ref/Dev/Theory Manuals + training mats.
C++ Framework

Strategy: control of multiple iterators and models

- Iterator
- Model

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

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

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

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

Optimization
- Uncertainty
- LeastSq

Uncertainty
- OptUnderUnc
- UncOfOptima

LeastSq
- ModelCalUnderUnc

OptUnderUnc
- Mixed A-E UQ

UncOfOptima
- Branch&Bound/PICO

Hybrid
- SurrBased

Pareto/MStart

Strategy

Uncertainty

Mixed A-E UQ

Six input components: Strategy, Method, Model, Variables, Interface, Responses
Office of Science Examples

Wind energy (ASCR)

Nuclear reactors (CASL, NEAMS)

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

SciDAC-3: Fusion, BER, BES, nuclear, high energy
Executing DAKOTA
Simulation Management (Black Box case)

**DAKOTA Input File**
- Commands
- Options
- Parameter definitions
- File names

**DAKOTA Parameters File**
\{x1 = 123.4\}
\{x2 = -33.3\}, etc.

**Use APREPRO/DPREPRO to cut-and-paste x-values into code input file**

**Code Input**
- CALORE: thermal analysis
- ALEGRA: shock physics
- SALINAS: structural dynam
- Premo: high speed flow
  (your code here)

**Code Output**

**DAKOTA Executable**
- Sensitivity Analysis, Optimization, Uncertainty Quantification, Parameter Estimation

**DAKOTA Executes**
sim_code_script to launch a simulation job

**DAKOTA Results File**
999.888 f1
777.666 f2, etc.

**User-supplied automatic post-processing of code output data into f-values**

**DAKOTA Output Files**
- Raw data (all x- and f-values)
- Sensitivity info
- Statistics on f-values
- Optimality info
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 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
Capability Overview (with emphasis on UQ)
Core Methods

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

Karush-Kuhn-Tucker conditions:
\[ \nabla f - \sum_{i} \lambda_{i} \nabla g_{i} = 0 \]

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

\[ f(x) = \sum_{i=1}^{n} (s_{i}(x) - d_{i})^{2} \]

Simulation output that depends on x

Given data

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

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

N samples

Output Distributions

Model

measure 1

measure 2
<table>
<thead>
<tr>
<th>Category</th>
<th>Production</th>
<th>Recently released</th>
<th>Under dev.</th>
<th>Future</th>
<th>Collabs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling</strong></td>
<td>Latin Hypercube, Monte Carlo</td>
<td>Importance, Incremental</td>
<td></td>
<td>Bootstrap, Jackknife</td>
<td>FSU</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td><em>Local:</em> Mean Value, 1st- &amp; 2nd-order reliability methods (AMV+,FORM,SORM)</td>
<td><em>Global:</em> adaptive GP methods (EGRA, GPAIS) with grad-enhancement</td>
<td>POF Darts</td>
<td>recursive emulation, TGP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stochastic expansion</strong></td>
<td></td>
<td>PCE/SC with dimension-adaptive p-/h-refinement, grad-enhancement, compressed sensing</td>
<td>Local adapt refinement, orthog least interp, adjoint EE correction</td>
<td>Discrete random variables</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other probabilistic</strong></td>
<td></td>
<td>Eff Subspace, Morris-Smale topology</td>
<td>Rand fields / stoch process</td>
<td>EDR</td>
<td></td>
</tr>
<tr>
<td><strong>Epistemic &amp; Bayesian</strong></td>
<td>Interval-valued/Second-order prob. (nested sampling)</td>
<td>Opt-based interval est, Dempster-Shafer, model form, QUESO, GPMSA</td>
<td>Discrete GPs, model selection</td>
<td>Imprecise probability</td>
<td>LANL, UT Austin</td>
</tr>
<tr>
<td><strong>Metrics &amp; Global SA</strong></td>
<td>Importance factors, rank/partial corr, main effects</td>
<td>Variance-based decomposition</td>
<td>Stepwise regression</td>
<td></td>
<td>LANL</td>
</tr>
</tbody>
</table>
Uncertainty Quantification Algorithms in DAKOTA: New methods bridge robustness/efficiency gap

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

Research: Scalability, Robustness, Goal-orientation
## Uncertainty Quantification Algorithms in DAKOTA: New methods bridge robustness/efficiency gap

<table>
<thead>
<tr>
<th></th>
<th>Production</th>
<th>Recently released</th>
<th>Under dev.</th>
<th>Future</th>
<th>Collabs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling</strong></td>
<td>Latin Hypercube, Monte Carlo</td>
<td>Importance, Incremental</td>
<td></td>
<td></td>
<td>FSU</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td>Local: Mean Value, 1st- &amp; 2nd-order reliability methods (AMV+, FORM, SORM)</td>
<td>Global: adaptive GP methods (egra, GPAIS) with grad enhancement</td>
<td>POF Darts</td>
<td>recursive emulation, TOP</td>
<td>Local: Notre Dame, Global: Vanderbilt</td>
</tr>
<tr>
<td><strong>Stochastic expansion</strong></td>
<td></td>
<td>PCE/SC with dimension-adaptive p-/h-refinement, grad-enhancement, compressed sensing</td>
<td>Local adapt refinement, orthog least interp, adjoint EE correction</td>
<td>Discrete random variables</td>
<td>Stanford, Purdue</td>
</tr>
<tr>
<td><strong>Other probabilistic</strong></td>
<td></td>
<td>Eff Subspace, Morris-Smale topology</td>
<td>Rand fields / stoch process</td>
<td>EDR</td>
<td>NCSU, Utah, Cornell, Maryland</td>
</tr>
<tr>
<td><strong>Epistemic &amp; Bayesian</strong></td>
<td>Interval-valued/Second-order prob. (nested sampling)</td>
<td>Opt-based interval est, Dempster-Shafer, model form, QUESO, GPMSA</td>
<td>Discrete GPs, model selection</td>
<td>Imprecise probability</td>
<td>LANL, UT Austin</td>
</tr>
<tr>
<td><strong>Metrics &amp; Global SA</strong></td>
<td>Importance factors, rank/partial corr, main effects</td>
<td>Variance-based decomposition</td>
<td>Stepwise regression</td>
<td></td>
<td>LANL</td>
</tr>
</tbody>
</table>
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 ($\rightarrow$ nested iteration)
  - Requirement to evaluate probability of rare events (e.g., safety criteria)
  - Nonsmooth responses ($\rightarrow$ 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

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 Stochastic Expansions: Polynomial Chaos and Stochastic Collocation

**Polynomial chaos:** spectral projection using orthogonal polynomial basis functions

\[ R = \sum_{j=0}^{P} \alpha_j \Psi_j(\xi) \]

- Estimate \( \alpha_j \) using regression or numerical integration: sampling, tensor quadrature, sparse grids, or cubature

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

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

\[ L_i = \prod_{j=1 \atop j \neq i}^{m} \frac{x - x_j}{x_i - x_j} \]

Sparse interpolants formed using \( \sum \) 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 Stochastic Expansions: Polynomial Chaos and Stochastic Collocation

**Polynomial chaos:** spectral projection using orthogonal polynomial basis functions

\[ R = \sum_{j=0}^{P} \alpha_j \Psi_j(\xi) \]

- Estimate \( \alpha_j \) using regression or numerical integration: sampling, tensor quadrature, sparse grids, or cubature

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

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

\[ L_i = \prod_{j=1, j \neq i}^{m} \frac{x - x_j}{x_i - x_j} \]

\[ R(\xi) \approx \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} r_1(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}) \left( L_{j_1}^{i_1} \otimes \cdots \otimes L_{j_n}^{i_n} \right) \]

Sparse interpolants formed using \( \Sigma \) 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 $\xi$
- Compute expected value of product of $R$ and each $\Psi_j$

*Linear regression ("point collocation"):*
- Sample w/i distribution of $\xi$
- Solves least squares data fit for all coefficients at once:

$$\Psi \alpha = R$$

**Tensor-product quadrature: PCE/SC**

$$\mathcal{P}_i(f)(\xi) = \sum_{j=1}^{m_i} f(\xi_j) w_j$$

$$\mathcal{Q}_i^n f(\xi) = (\mathcal{P}_i^{1} \otimes \cdots \otimes \mathcal{P}_i^{n}) (f)(\xi) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} f(\xi_{j_1}, \ldots, \xi_{j_n}) (w_{j_1}^{1} \otimes \cdots \otimes w_{j_n}^{n})$$

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

**Smolyak Sparse Grid: PCE/SC**

$$\mathcal{S}(w, n) = \sum_{w+1 \leq |i| \leq w+n} (-1)^{w+n-|i|} \binom{n-1}{w+n-|i|} \cdot (\mathcal{P}_i^{1} \otimes \cdots \otimes \mathcal{P}_i^{n})$$

Pascal’s triangle (2D):

**Cubature: PCE**

Stroud and extensions (Xiu, Cools)
$\rightarrow$ Low order PCE
$\rightarrow$ global SA, anisotropy detection

Gaussian $i = 2 \rightarrow p = 1$

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

Arbitrary PDF

$$t^{(k)} = \frac{1}{\gamma} \left[ \sqrt{\gamma c_1 x^{(k)}} - \delta \right]$$
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
  - PCE/SC: variance-based decomp. \( \rightarrow \) total Sobol’ indices \( \rightarrow \) anisotropy
  - PCE: spectral coefficient decay rates \( \rightarrow \) 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

**PCE:**
- Linear regression including derivatives
  - Gradients/Hessians \( \rightarrow \) addtl. 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)

\[
 e^{-10x^2 - 5y^2}
\]

\[
 f(x) = 2 \prod_{j=1}^{3} \frac{|x_j - 2a|}{1 + a}; \quad a = [0, 1, 2, 4, 8]
\]

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

and similar for higher-order moments
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

\[
\Delta \Sigma_{ij} = \Delta E[R_i R_j] - \mu_{R_i} \Delta E[R_j] - \mu_{R_j} \Delta E[R_i] - \Delta E[R_i] \Delta E[R_j] \rightarrow \Delta \sigma, \Delta \beta
\]
Stochastic Expansions on Unstructured Grids: Compressive Sensing

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

or in matrix notation

\[
b = Ax + \varepsilon
\]

and find the minimum norm solution

\[
\min_x \|Ax - b\|_2
\]

or (more recently) find a sparse solution

\[
\min_x \|x\|_1 \text{ such that } \|Ax - b\|_2 \leq \varepsilon
\]

(a) CS methodology ($\ell_1$ objective)  (b) Pseudo-inverse ($\ell_2$ objective)

BP

\[
c = \arg\min_c \|c\|_{\ell^1} \text{ such that } \Phi c = y
\]

BPDN and OMP

\[
c = \arg\min_c \|c\|_{\ell^1} \text{ such that } \|\Phi c - y\|_{\ell^2} \leq \varepsilon
\]

LASSO and LARS

\[
c = \arg\min_c \|\Phi c - y\|^2_{\ell^2} \text{ such that } \|x\|_{\ell^1} \leq \tau
\]
Multiple Model Forms in UQ
[Discrete model choices for same physics]

- A clear hierarchy of fidelity (low to high)

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

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

- Both

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? \( \rightarrow \) global approxs. of model discrepancy

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

\( N_{lo} \gg N_{hi} \)

Adaptive sparse grid multifidelity algorithm:
- Gen. sparse grids for LF & discrepancy levels
- Greedy selection from grids: max \( \Delta \text{QoI}/\Delta \text{Cost} \)
- 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

Time = 0.0
Concluding Remarks

DAKOTA provides a variety of core algorithms for iterative analysis:
- Optimization
- Calibration

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

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)
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
- Related packages: http://dakota.sandia.gov/packages.html