A Surrogate Approach Towards Validation and Uncertainty Quantification of Multiphysics Reactor Simulation Codes

Thesis Prospectus

Artem Yankov

University of Michigan

April 17, 2014

Outline

Background

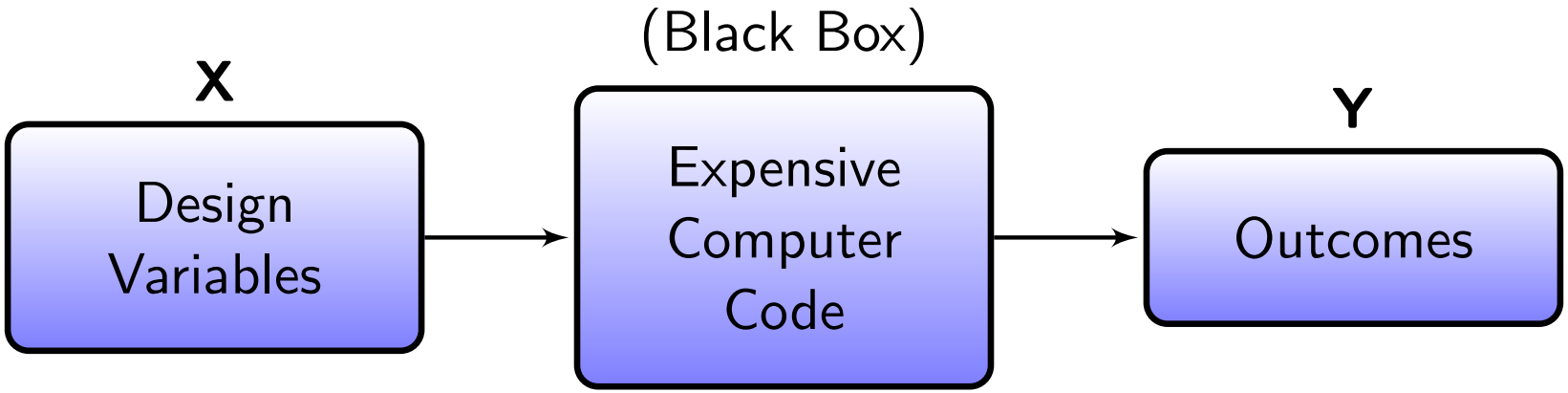
- Physical phenomena is studied by conducting experiments.
- Any data collected represents instances of underlying stochastic processes.
- We build predictive computer models in an attempt to reproduce such observed physical phenomena.
- To accurately capture stochastic element of experiments, computer models should be probabilistic.
- In other words, inputs to computer models have uncertainties associated with them that are propagated to any outputs of interest.
- Running computer simulations should be like conducting physical experiments. Computer experiments.

Why Surrogates?

- We run computer simulations to meet design objectives under certain constraints.
- Involves numerical optimization, calibration, and performing what-if analyses.
- Also, we're interested in determining which of our design variables have the greatest effects on simulation outcomes.
- Thousands of simulations required to make this possible but...
- Computer simulations that model real phenomenon like nuclear reactors often take $\mathcal{O}(\text{hours})$ or $\mathcal{O}(\text{days})$ to complete.
- Building a surrogate model for your expensive computer simulations can make everything listed above possible.

What is a Surrogate Model?

- A model for the outcomes of (likely) expensive computer simulations that can be rapidly evaluated while simultaneously preserving the predictive capabilities of the original simulations.
- Want to intelligently choose subspace $\{x_1, x_2, ..., x_N\} \subset \mathbf{X}$ to sample expensive computer code to get $\{y_1, y_2, ..., y_N\} \subset \mathbf{Y}$.
- Learn fast mapping that approximates $\mathbf{X} \rightarrow \mathbf{Y}$.



Proposed Application to Fuel Performance Modeling

- Fission Gas Release (FGR) refers to the phenomenon where Xenon and Krypton gases formed in UO_2 fuel rods are released into the rod filling gas.
- Causes pressure build-up and thermal conductivity degradation in the rod filling gas, potentially jeopardizing the safety of the reactor.
- Fission gas atoms generated in the fuel grains diffuse towards the grain boundaries.
- Majority of the gas diffuses into grain-face gas bubbles, giving rise to grain-face swelling.
- Bubble growth brings about bubble coalescence and interconnection, eventually leading to the formation of a tunnel network through which the fission gas is released.

SIFGRS FGR Model

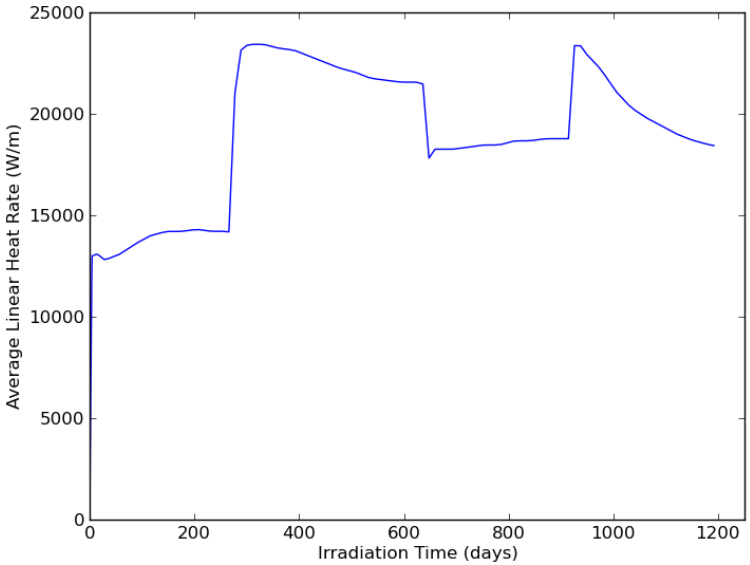
- Simple Integrated Fission Gas Release and Swelling (SIFGRS)
- Incorporates gas diffusion and precipitation in grains, growth and coalescence of gas bubbles at grain faces, thermal, athermal, steady-state, and transient gas release.
- Through a direct description of the grain face gas bubble development, the fission gas swelling and release are calculated as coupled processes.
- Parameterized by, among others, linear heat rate, gas diffusion coefficient, surface tension of grain face bubbles, hydrostatic pressure, fuel grain radius, fuel porosity, and grain boundary sweeping.

Risø AN3 Experiment

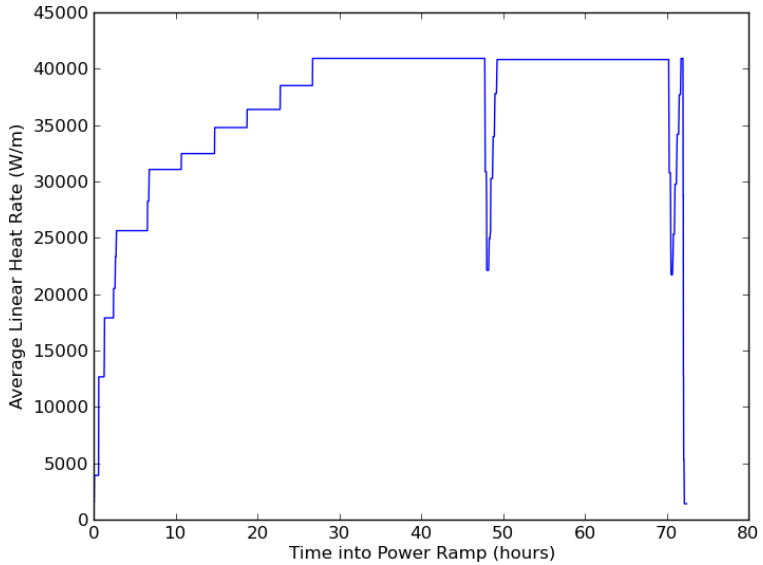
- Validation case for fuel performance modeling in the Fumex-II database.
- Experiment consists of a base irradiation of four reactor cycles in the Biblis A pressurized water reactor.
- After the base irradiation period, a fuel rod is extracted and refabricated to a shorter length before undergoing a power ramp.
- Refabricated fuel rod is outfitted with various instrumentation such that fuel centerline temperature, FGR and rod internal pressure measurements can be obtained.

Risø AN3 Experiment Irradiation Profiles

Base Irradiation History



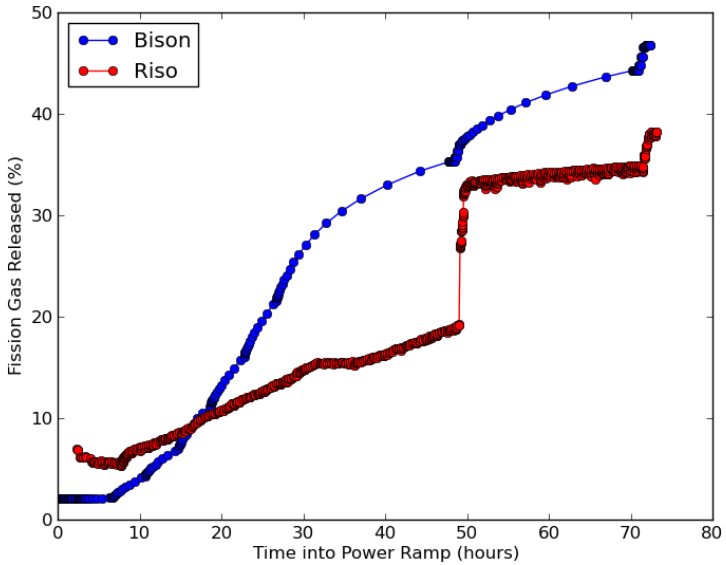
Power Ramp Experiment



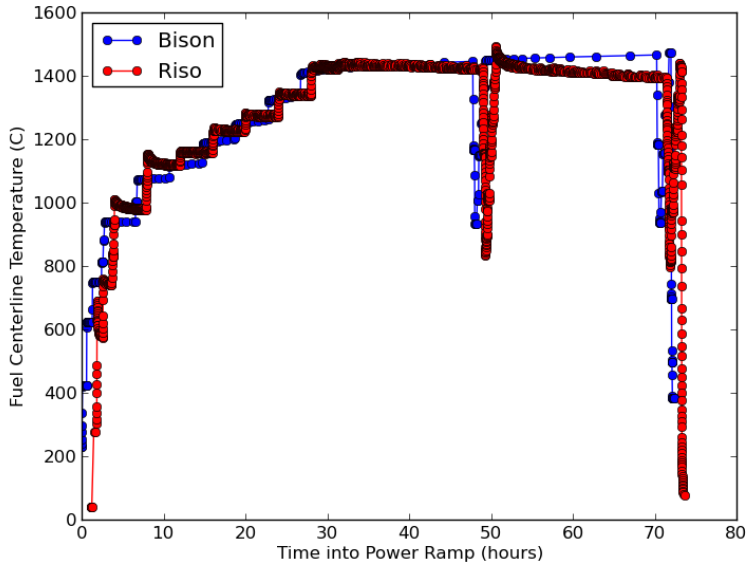
Modeling Risø AN3 Experiment with BISON

- BISON is a finite-element fuel performance modeling code that utilizes the SIFGRS model.
- SIFGRS parameters are quite generic and uncertain.

Fission Gas Release



Fuel Centerline Temperature



Modeling Risø AN3 Experiment with BISON/MPACT

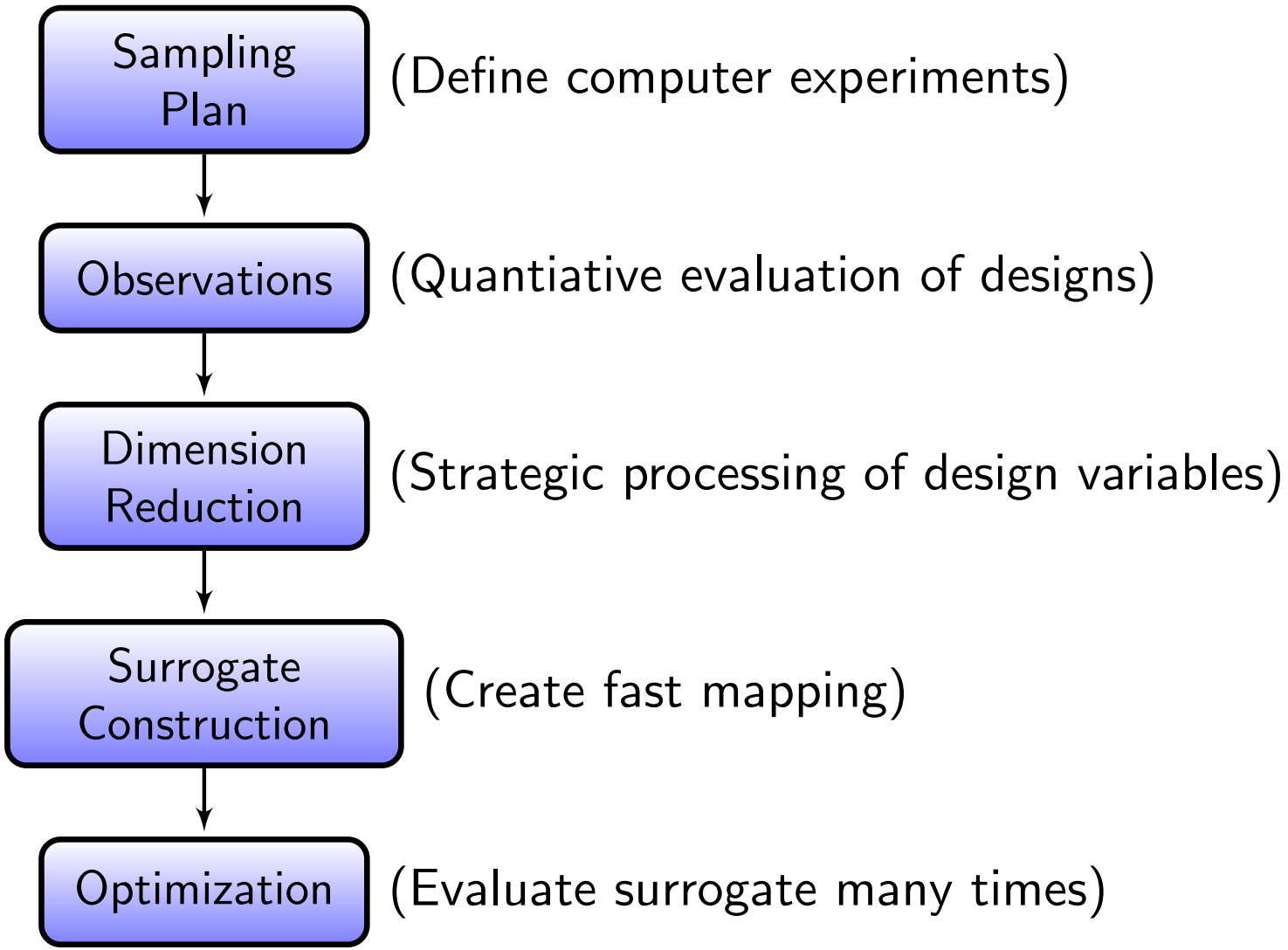
- No sense in comparing the output of a computer simulation to experimental data unless the computer simulation is of high fidelity and capable of reproducing the pertinent physics.
- MPACT is a neutronics code that provides detailed intrapin and azimuthally dependent neutronics data in the fuel elements.
- The two-way coupling scheme provided by BISON and MPACT provides the most accurate fuel performance modeling available for a nuclear reactor.
- Expensive!

Modeling Risø AN3 Experiment with BISON/MPACT

- BISON predictions of FGR and temperature fields stand to be improved by calibrating FGR parameters to experimental data.
- Calibration studies require $\mathcal{O}(10^3)$ function evaluations, which in this case are the coupled BISON/MPACT computer codes.
- Each simulation of the Risø AN3 experiment will take a few hours.
- It's necessary to construct a surrogate for the calibration study!

Outline

Classic Overview



Kriging vs. anchored-ANOVA Collocation

Kriging

- Dimension reduction processed separately
- Sampling points random
- User determines how many points to use for sampling plan
- Interpolation by covariance basis functions
- More statistical approach

anchored-ANOVA Collocation

- Dimension reduction inherent
- Sampling done on structured grid
- Sampling plan size dependent on number of design variables
- Polynomial interpolation
- More deterministic approach

Dimension Reduction for Kriging

- Kriging effective for $\mathcal{O}(10)$ design variables.
- For more design variables Kriging will defeat the purpose of having a surrogate in the first place.
- Fortunately, various engineering applications have shown that only a handful of design variables have non trivial impact on outputs of interest.
- How to identify the "important variables"?
- Morris' Algorithm.

Morris' Algorithm

- Premise: If the output parameter does not change with respect to a design variable then the variable can safely be ignored.
- Elementary effect $d_i(\mathbf{x})$ of design variable x_i:

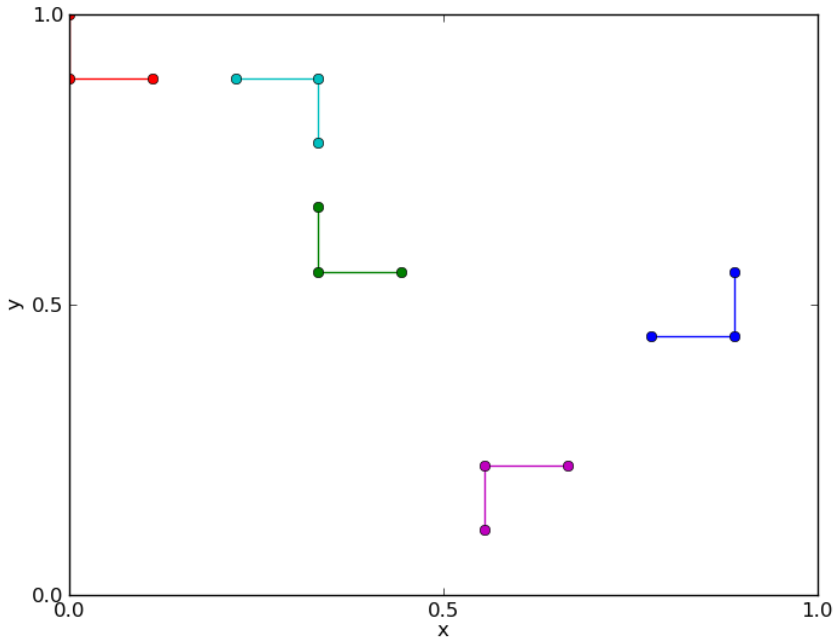
$$d_i(\mathbf{x}) = \frac{f(x_1, x_2, ..., x_{i-1}, x_i + \Delta, x_{i+1}, ..., x_k) - f(\mathbf{x})}{\Delta}$$

- Choosing a set of $\mathbf{x}$ carefully, it is possible to calculate an elementary effect for each of k design variables using only $k+1$ function evaluations using the random orientation matrix $\mathbf{B}^*$:

$$\mathbf{B}^* = \left(\mathbf{1}_{k+1,1}\mathbf{x}^* + \frac{\Delta}{2}\left[\left(2\mathbf{B} - \mathbf{1}_{k+1,k}\right)\mathbf{D}^* + \mathbf{1}_{k+1,k}\right]\right)\mathbf{P}^*.$$

Morris' Algorithm

- r random orientation matrices are created to obtain r elementary effects for each design variable.
- Plot mean and standard deviation of each variable's effects.
- Variables with negligible effect on function will cluster around origin.
- Large fluctuations in standard deviation indicative of nonlinear and interactive effects.

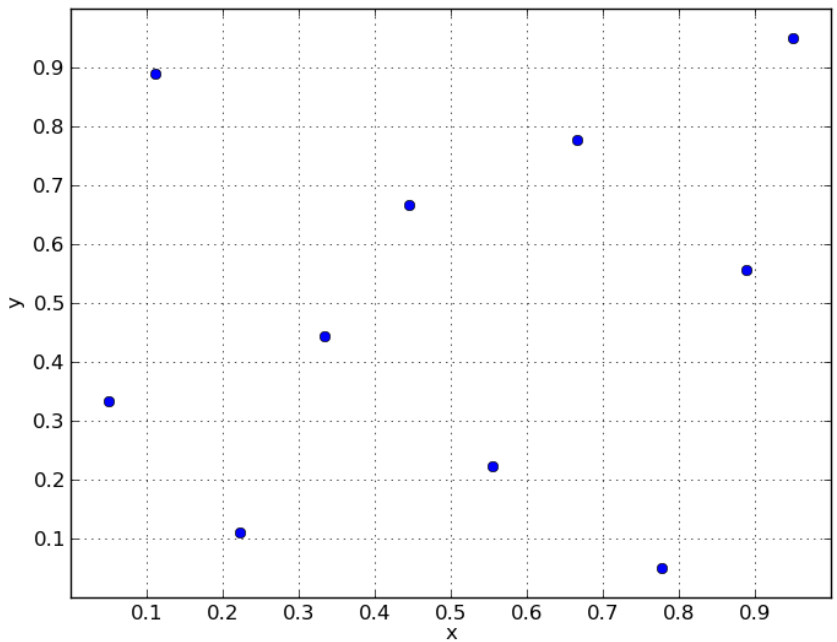


Designing a Kriging Sampling Plan

- All surrogate models are built around a set of points at which the objective computer code is actually evaluated.
- Intuitively, the surrogate accuracy is expected to decrease as one moves further away from such points.
- Important to spread N points as uniformly as possible across the design space.
- For Kriging, Latin Hypercube Sampling (LHS) is used to create a sampling plan.
- There is a notion of an optimized LHS sampling plan based on the maximin metric.

Latin Hypercube Sampling

- Basis of LHS rests upon dividing the normalized space of each design variable into n equally sized bins if n samples are required.
- As a result, when the n samples are taken it is guaranteed that the entire spectrum of each design variable's space has been visited.



Optimizing a LHS Plan

- The maximin metric describe by Morris and Mitchell makes use of two notions in an attempt to quantify the 'space-fillingness' of a sampling plan.
- Unique distances between all points in the plan sorted in ascending order $\{d_1, d_2, ..., d_m\}$.
- Corresponding number of occurrences of each distance $\{J_1, J_2, ..., J_m\}$.
- In words, the Morris and Mitchell criteria states that an optimized sampling plan will minimize all J_i while maximizing the corresponding d_i.
- The maximin sampling plan maximizes d_1, and among plans for which this is true, minimizes J_1, among plans for which this is true, maximizes d_2,....

Optimizing a LHS Plan

- The previous definition can be restated into a pseudo equivalent minimization problem.

$$\Phi_q(\mathbf{X}) = \left(\sum_{j=1}^{m} J_j d_j^{-q} \right)^{1/q}$$

- The minimization of this equation and the Morris and Mitchell definition of the maximin sampling plan are used in unison to obtain a locally optimal sampling plan.
- Generate initial sampling plan, optimize for set of q values using simulated annealing.
- Resulting set of plans are contested directly against each other by explicit application of Morris and Mitchell's maximin definition.

Kriging on a Sampling Plan

- Optimized sampling plan $\mathbf{X} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ...\mathbf{x}^{(n)}\}$.
- At each datum $\mathbf{x}^{(k)}$ a random process $Y(\mathbf{x}^{(k)})$ induces an observation $y^{(k)}$.
- Resulting random field can be described with a mean value of $\mathbf{1}\mu$ and a correlation matrix,

$$\boldsymbol{\Psi} = \begin{pmatrix} cor[Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(1)})] & \cdots & cor[Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(n)})] \\ \vdots & \ddots & \vdots \\ cor[Y(\mathbf{x}^{(n)}), Y(\mathbf{x}^{(1)})] & \cdots & cor[Y(\mathbf{x}^{(n)}), Y(\mathbf{x}^{(n)})] \end{pmatrix}$$

$$cor[Y(\mathbf{x}^{(i)}), Y(\mathbf{x}^{(l)})] = \exp\left(-\sum_{j=1}^{k} \theta_j |x_j^{(i)} - x_j^{(l)}|^{p_j}\right)$$

Kriging on a Sampling Plan

- Given the formulation of the observations occurring at $\mathbf{x}^{(k)}$ as instances of a stochastic process, the likelihood of seeing the observed data is,

$$L\left(\mathbf{Y}^{(1)},...,\mathbf{Y}^{(n)}|\mu,\sigma,\{\theta_1,...,\theta_k\},\{p_1,...,p_k\}\right) =$$

$$\frac{1}{(2\pi\sigma^2)^{n/2}\left|\boldsymbol{\Psi}\right|^{1/2}} \times \exp\left[\frac{(\mathbf{y}-\mathbf{1}\mu)^T\boldsymbol{\Psi}^{-1}(\mathbf{y}-\mathbf{1}\mu)}{2\sigma^2}\right].$$

- Maximizing the log likelihood,

$$\hat{\mu} = \frac{\mathbf{1}^T\boldsymbol{\Psi}^{-1}\mathbf{y}}{\mathbf{1}^T\boldsymbol{\Psi}^{-1}\mathbf{1}}$$

$$\hat{\sigma}^2 = \frac{(\mathbf{y}-\mathbf{1}\mu)^T\boldsymbol{\Psi}^{-1}(\mathbf{y}-\mathbf{1}\mu)}{n}.$$

Kriging on a Sampling Plan

- Substitute $\hat{\sigma}$ and $\hat{\mu}$ into log likelihood to get, concentrated ln-likelihood function.

$$\log(L) \approx -\frac{n}{2}\log\left(\hat{\sigma}^2\right) - \frac{1}{2}\log|\mathbf{\Psi}|$$

- Optimize with respect to the θ and p parameters using global search algorithm.
- Once all optimizing parameters are available the goal is to utilize the parameters to build a model that makes function predictions on new points $\mathbf{x}$.

Making Predictions with Kriging Surrogate

- Construct a vector of correlations with existing points and $\mathbf{x}$,

$$\boldsymbol{\psi} = \begin{pmatrix} cor[Y(\mathbf{x}^{(1)}), Y(\mathbf{x})] \\ \vdots \\ cor[Y(\mathbf{x}^{(n)}), Y(\mathbf{x})] \end{pmatrix}.$$

- New predictions can be made at $\mathbf{x}$ using the maximum likelihood estimator,

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}).$$

- Prediction using kriging works to estimate a function value at a certain point by computing a weighted average of known function values in the vicinity of the objective points.

Collocation and anchored-ANOVA Algorithmic Overview

- anchored-ANOVA: Decompose objective function into functions of one variable $f(x_i)$, two variables $f(x_i, x_j)$, three variables $f(x_i, x_j, x_k)$ as needed.
- Collocation: For each component in the decomposition construct a polynomial interpolant by sampling objective function at pre-defined points.
- The pre-defined points are determined by Smolyak Sparse Grids, and a selection of quadrature abscissas (e.g. Newton-Cotes).
- Combine interpolants in decomposition to get an effective surrogate.

Numerical Interpolation in 1D

- First, pick a set of m_i collocation points.
- Evaluate objective function at all collocation points.
- Interpolated function is a linear expansion of some basis a_j^i (e.g. Lagrange polynomials) with weights $f\left(x_j^i\right)$.

$$U^i = \sum_{j=1}^{m_i} f\left(x_j^i\right) a_j^i$$

Clenshaw-Curtis Collocation Points

- Clenshaw-Curtis points consist of the extrema of Chebyshev polynomials.
- $n+1$ abscissas can exactly integrate polynomials of degree n.
- Points have the advantage of being nested.

$$x_j^i = \begin{cases} \cos\frac{\pi(j-1)}{m_i-1} & j = 1, \ldots, m_i \text{ if } i > 1 \\ 0 & j = 1 \text{ if } i = 1 \end{cases}$$

$$m_i = \begin{cases} 2^{i-1}+1 & i > 1 \\ 1 & i = 1 \end{cases}$$

Basis Functions

- Lagrange characteristic polynomials are plagued by the fact that each evaluation requires $\mathcal{O}(m_i^2)$ operations and often the computation is numerically unstable.
- Instead the barycentric form of Lagrange characteristic polynomials is used to form a basis.

$$a_j^i = \begin{cases} 1 & \text{if } i = 1 \\ \dfrac{\frac{w_j^i}{x - x_j^i}}{\sum_{j=0}^{m_i} \frac{w_j^i}{x - x_j^i}} & j = 1, ..., m_i \text{ for } i > 1 \end{cases}$$

- For Clenshaw-Curtis collocation points the barycentric weights are given by,

$$w_j^i = (-1)^{j+1} \delta_j^i \qquad \delta_j^i = \begin{cases} .5 & j = 1 \text{ or } j = m_i \\ 1 & \text{else} \end{cases}.$$

Expanding to Multivariate Interpolation

- Combine 1D interpolation formulas using tensor products
- Suffers from "Curse of Dimensionality".
- Mitigate the curse using Smolyak Sparse Grids.

$$\left(U^{i_1} \otimes \cdots \otimes U^{i_d}\right)(f) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_d=1}^{m_{i_d}} f\left(x_{j_1}^{i_1}, \cdots, x_{j_d}^{i_d}\right)\left(a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_d}^{i_d}\right)$$

Smolyak Sparse Grids

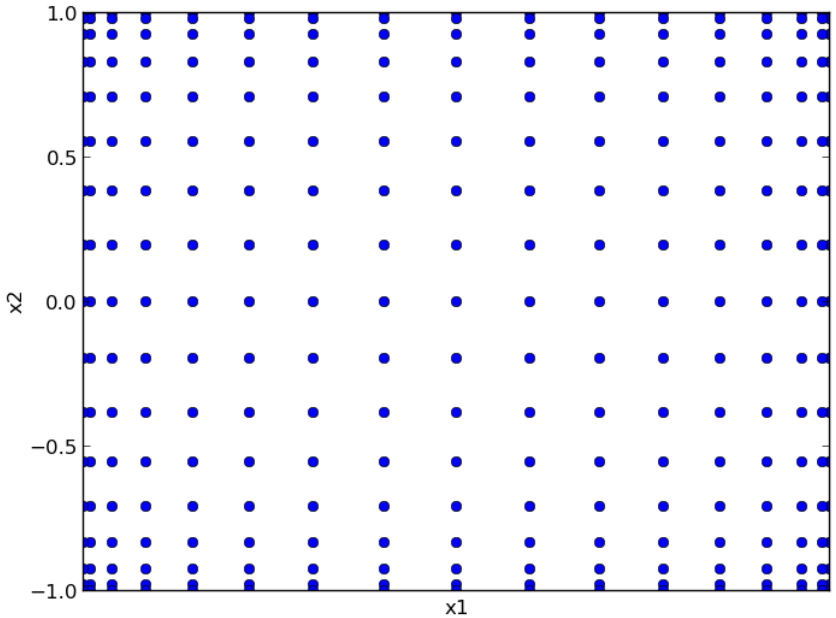
- Based on full tensor product formula the only difference being not all tensor products are used.
- In explicit form, the Smolyak formula is,

$$A_{q,d}(f) = \sum_{q-d+1 \leq |\mathbf{i}| \leq q} (-1)^{q-|\mathbf{i}|} \binom{d-1}{q-|\mathbf{i}|} \left(U^{i_1} \otimes \cdots \otimes U^{i_d}\right).$$

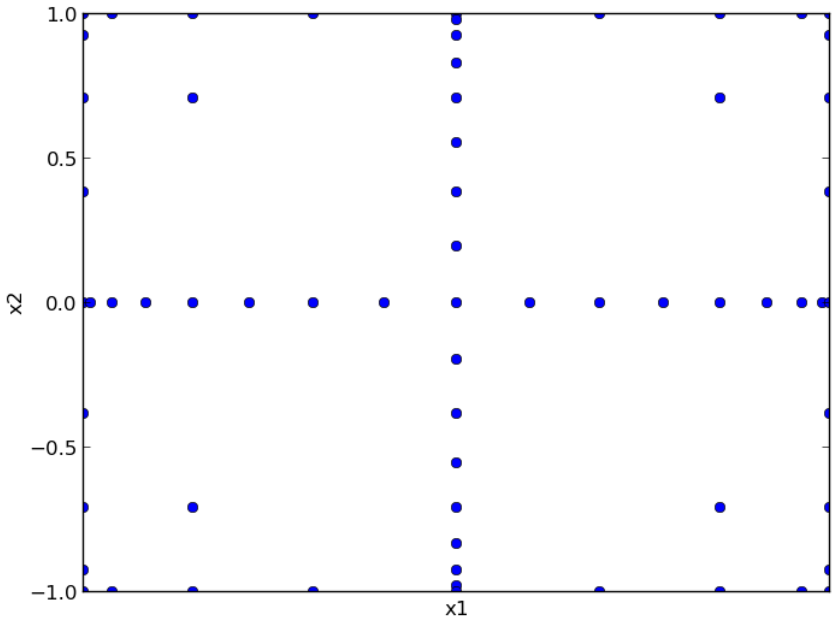
- i_k is the index corresponding to the level of interpolation in dimension k.
- The magnitude of $\mathbf{i}$ is $|\mathbf{i}| = |i_1 + \cdots + i_d|$.
- q keeps track of the level of interpolation of the Smolyak algorithm. More tensor product combinations are allowed as q increases.
- Smolyak algorithm reduces the total number of tensor product components by limiting the entries of $\mathbf{i}$.

Smolyak Sparse Grid Visualization

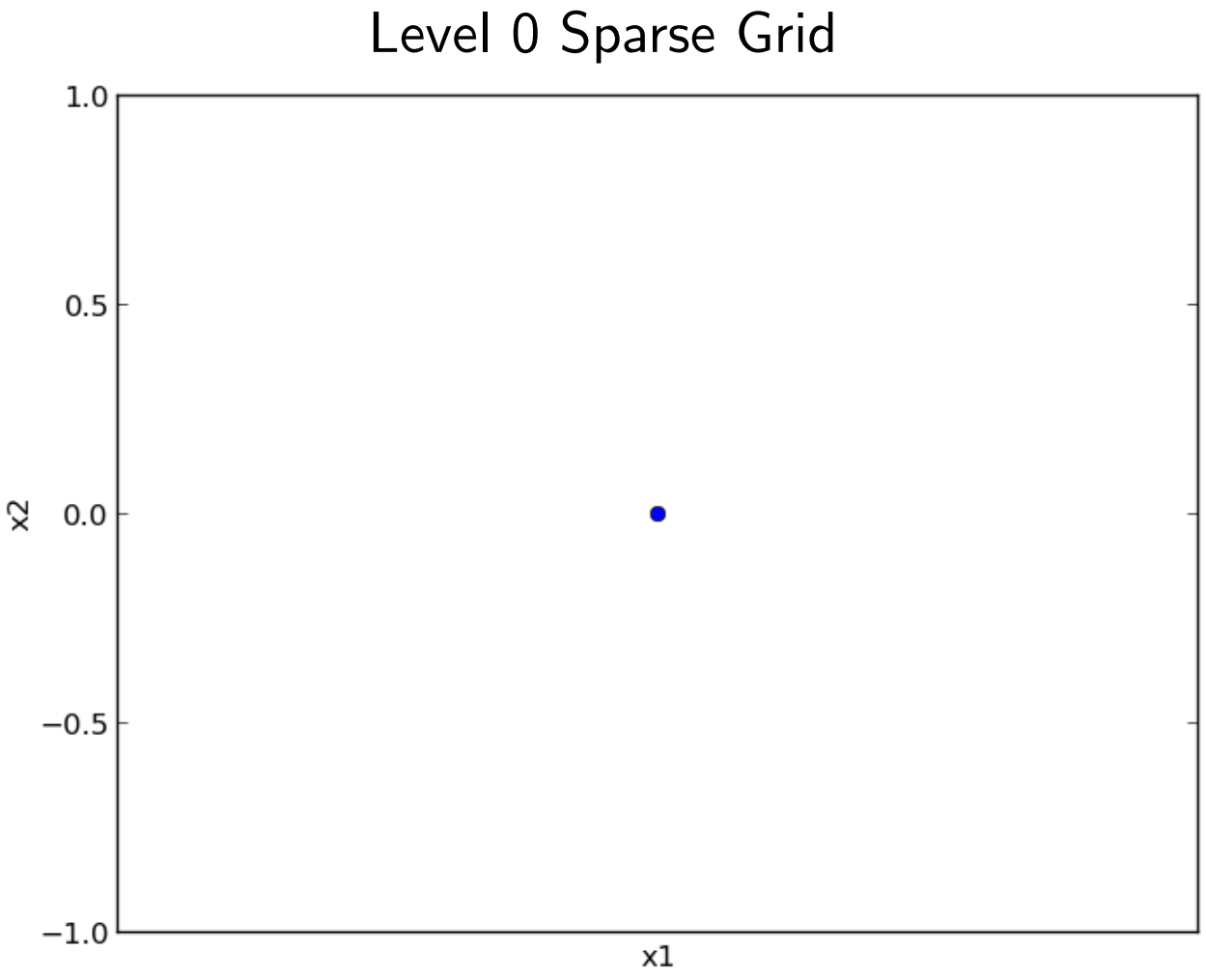
Level 4 Clenshaw-Curtis Full Tensor Product Grid



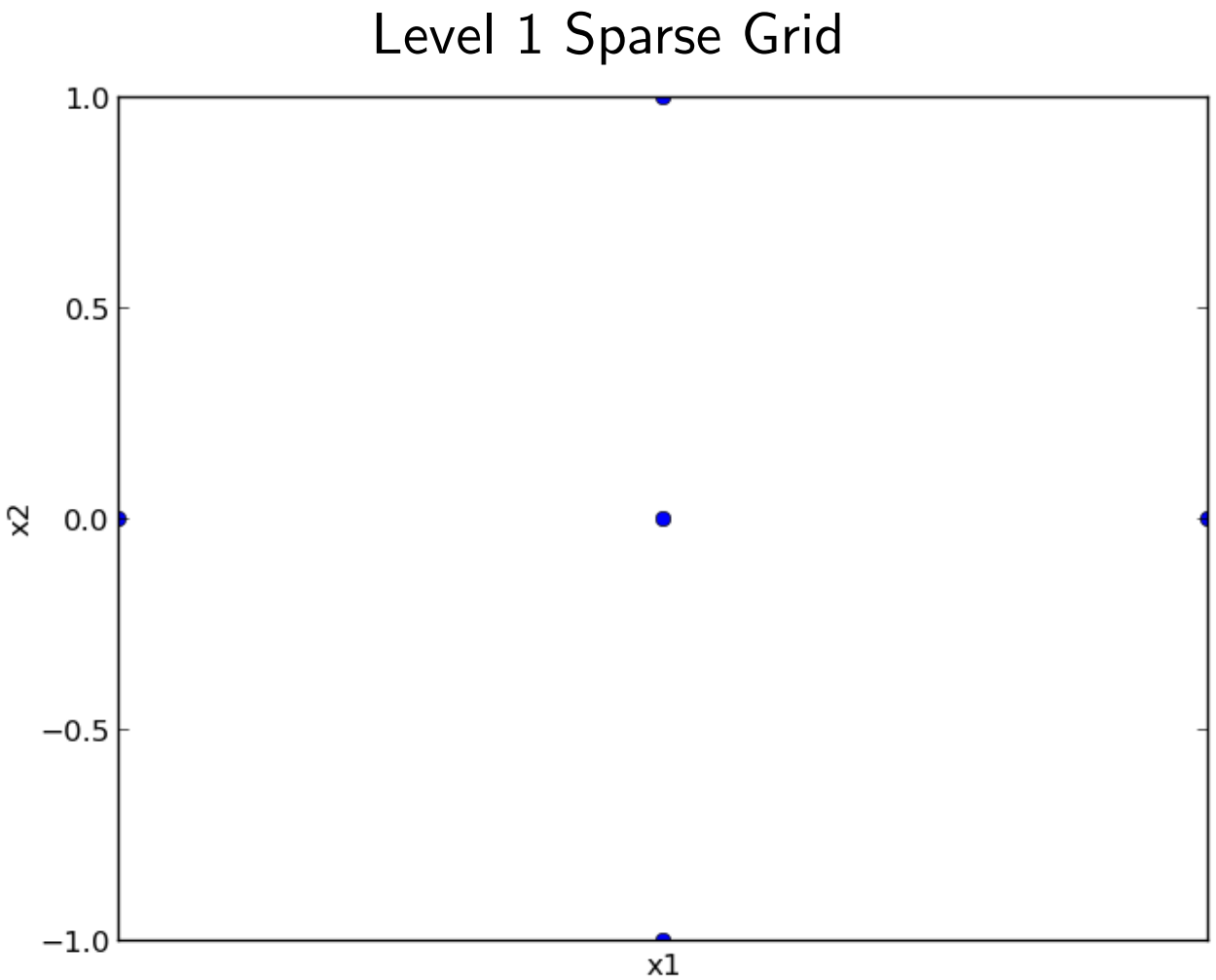
Level 4 Clenshaw-Curtis Smolyak Sparse Grid



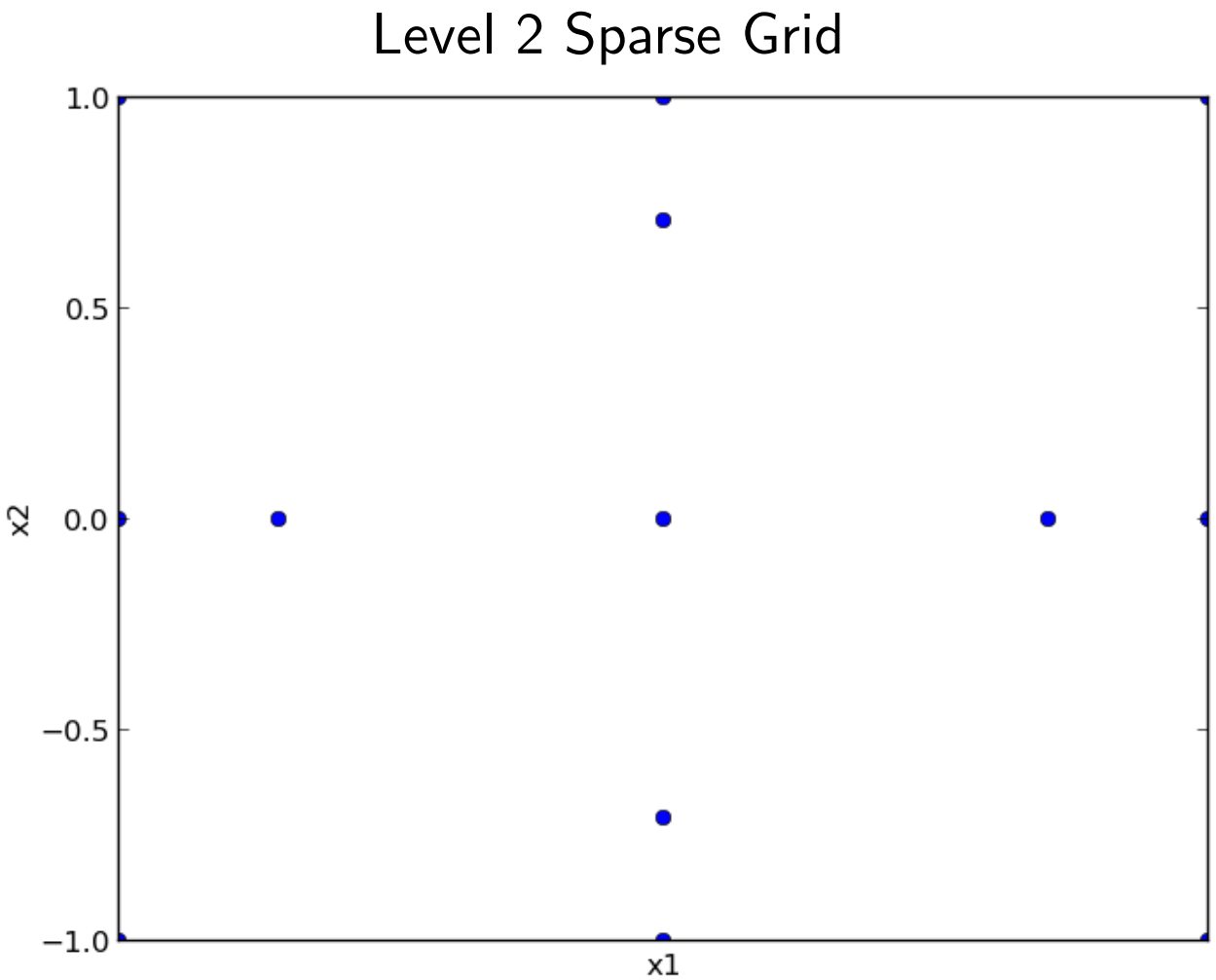
Smolyak Sparse Grid Visualization



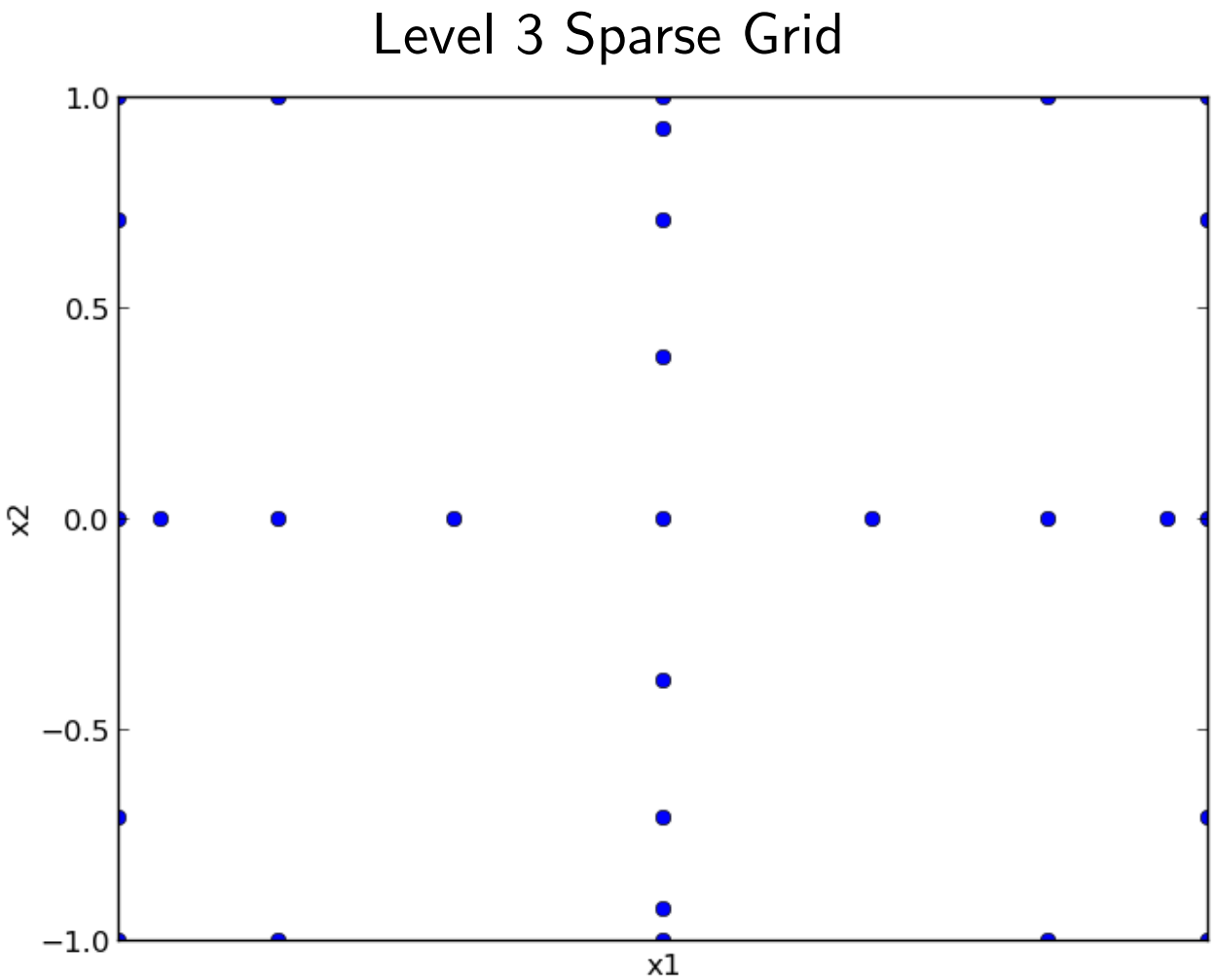
Smolyak Sparse Grid Visualization



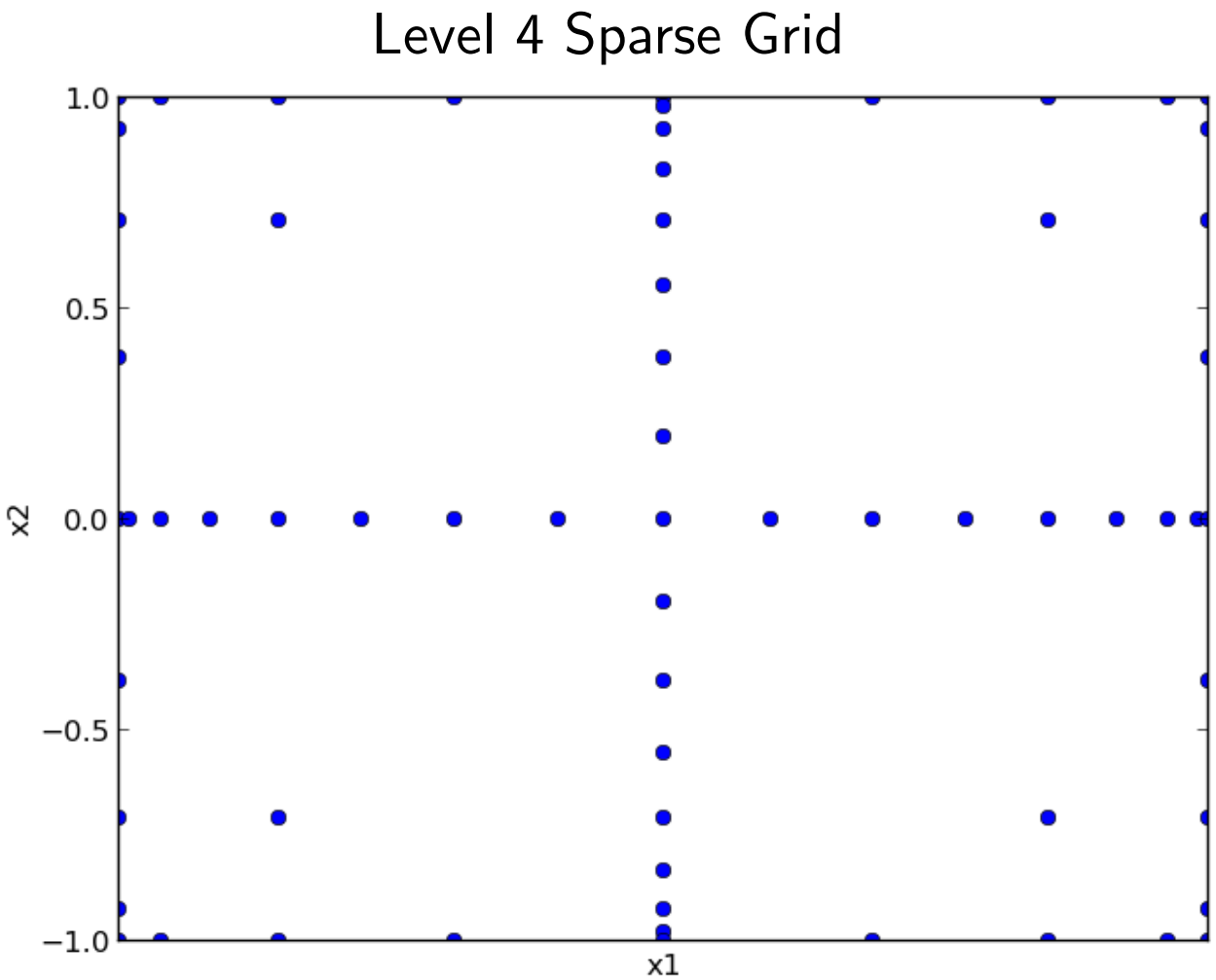
Smolyak Sparse Grid Visualization



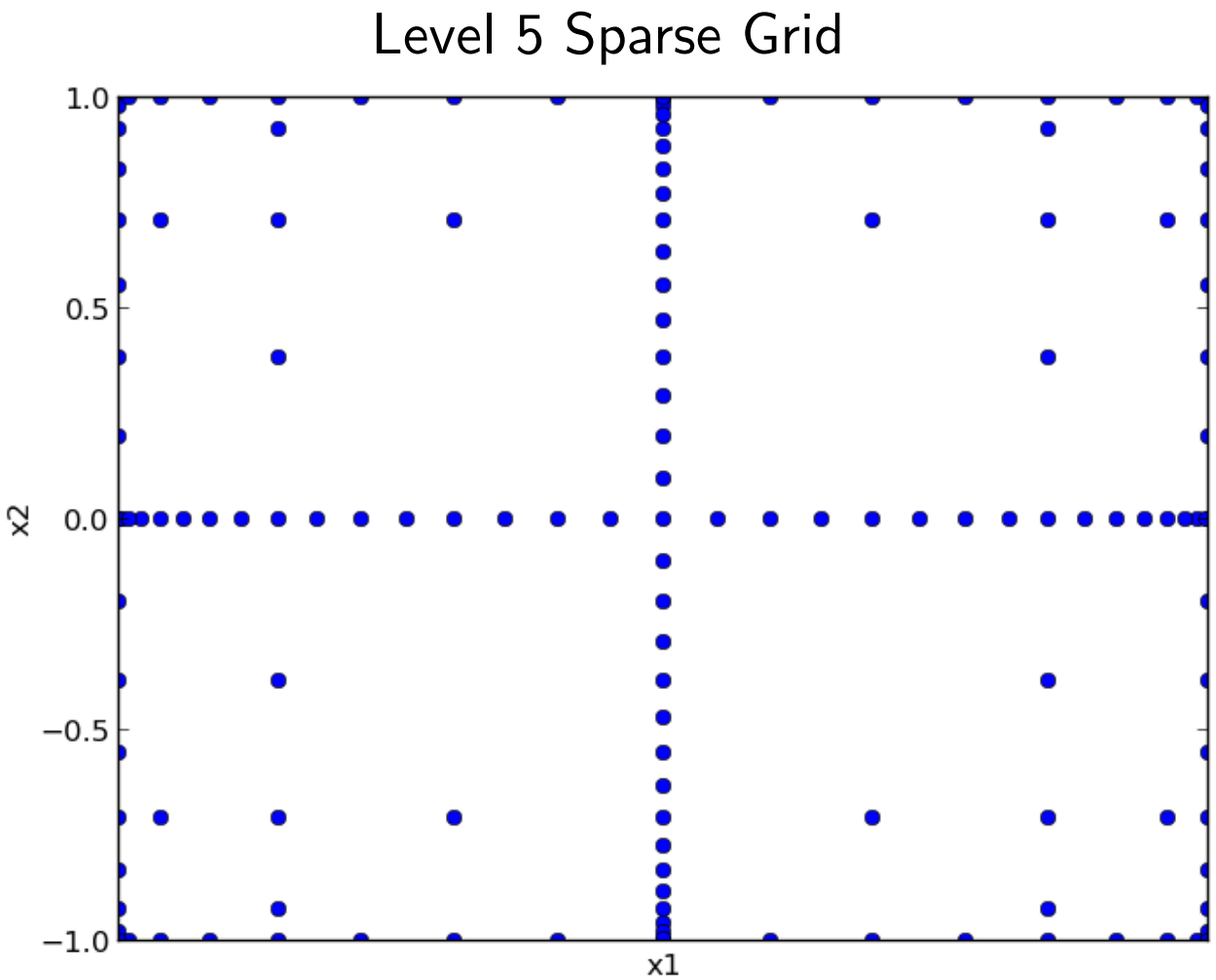
Smolyak Sparse Grid Visualization



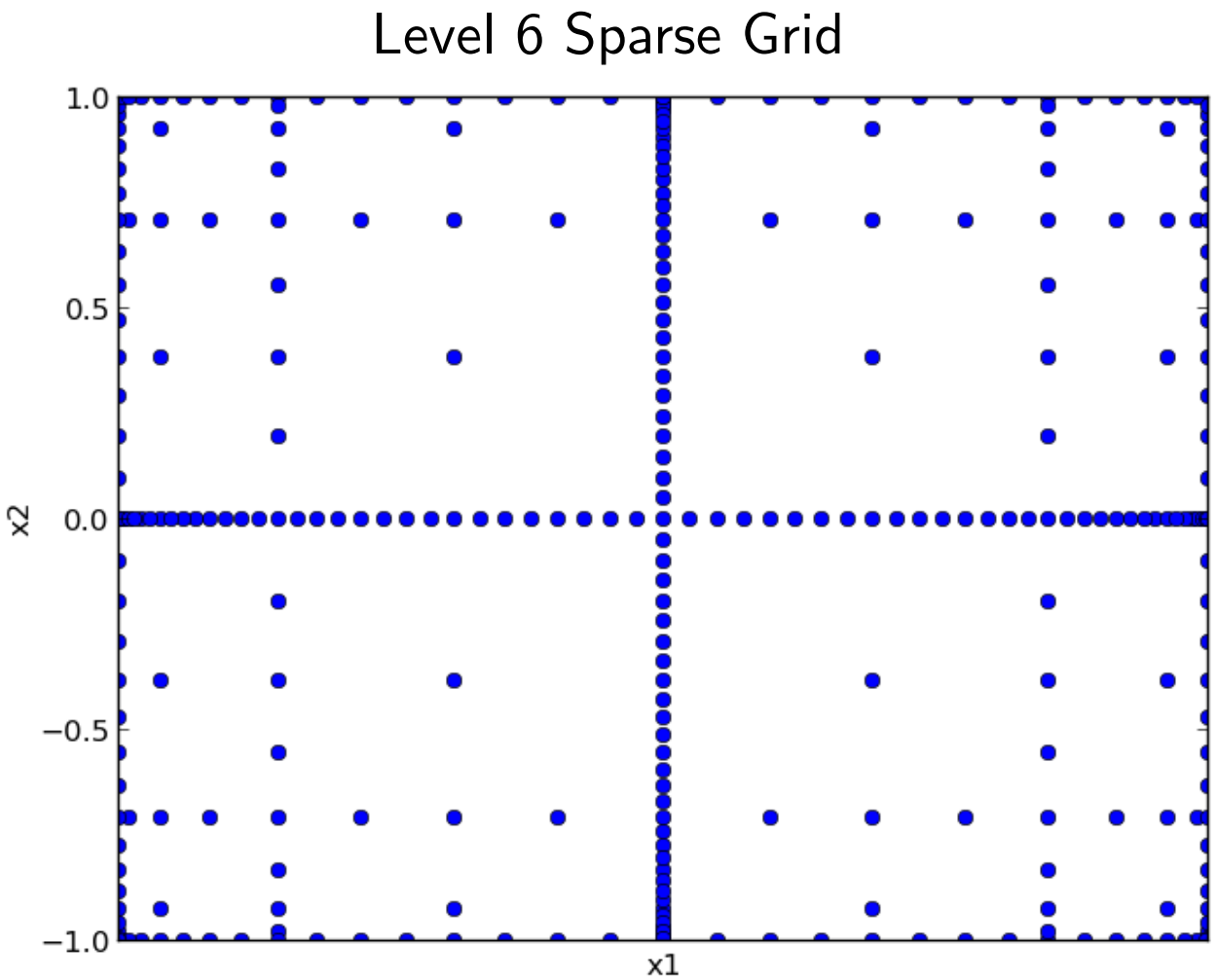
Smolyak Sparse Grid Visualization



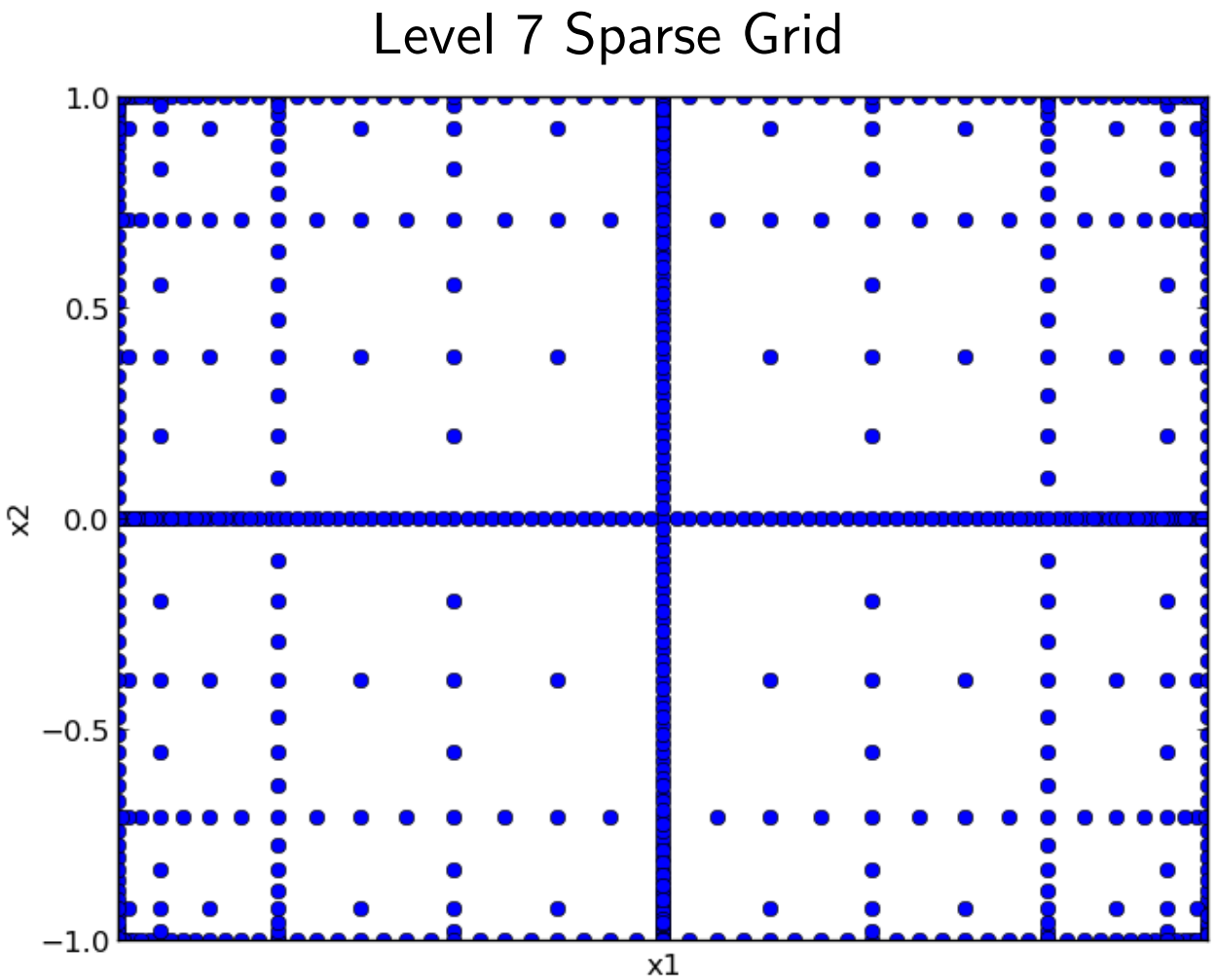
Smolyak Sparse Grid Visualization



Smolyak Sparse Grid Visualization



Smolyak Sparse Grid Visualization



Recursive Definition of Smolyak's Algorithm

- The explicit definition of Smolyak's algorithm can be rewritten in a recursive fashion.
- Able to increase interpolation level without having to start over each time.
- Hierarchical surplus serves as an error indicator. Lets you know how well the interpolation is going as the grid is refined.
- Use hierarchical surplus to determine convergence.

$$A_{q,d}(f) = A_{q-1,d}(f) + \Delta A_{q,d}$$

$$\Delta A_{q,d} = \sum_{|\mathbf{i}|=q} \left(\underbrace{f(x_{j_1}^{i_1}, ..., x_{j_d}^{i_d}) - A_{q-1,d}(x_{j_1}^{i_1}, ..., x_{j_d}^{i_d})}_{\text{Hierarchical Surplus}} \right) \cdot \left(a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_d}^{i_d} \right)$$

anchored-ANOVA Decomposition

- Consider a d-dimensional function $f(\mathbf{x}) = f(x_1, x_2, ..., x_d)$.
- The operator $P_{\mathbf{u}}$ projects from a d-dimensional space to a $|\mathbf{u}|$-dimensional space for some set $\mathbf{u} \subseteq \mathcal{D} = \{1, ..., d\}$.

$$f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) = P_{\mathbf{u}} f(\mathbf{x}) = \int_{\Omega^{d-|\mathbf{u}|}} f(\mathbf{x}) \, d\mu_{\mathcal{D}\setminus\mathbf{u}}(\mathbf{x})$$

- Use the Dirac measure $\delta(\mathbf{x} - \mathbf{a})d\mathbf{x}$ to evaluate the projection operator at an "anchor" point $\mathbf{a}$,

$$P_{\mathbf{u}} f(\mathbf{x}) = f(\mathbf{x})\,|_{\mathbf{x}=\mathbf{a}\setminus\mathbf{x}_{\mathbf{u}}}.$$

- For $\mathbf{u} \neq \mathbf{v}$ the following orthogonality relation holds:

$$(f_{\mathbf{u}}, f_{\mathbf{v}}) = 0$$

anchored-ANOVA Decomposition

- A function can be written in terms of its 2^d orthogonal components,

$$f(\mathbf{x}) = \sum_{\mathbf{u} \subseteq \mathcal{D}} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}).$$

- The component functions $f_{\mathbf{u}}$ are defined recursively,

$$f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) = P_{\mathbf{u}} f(\mathbf{x}) - \sum_{\mathbf{v} \subset \mathbf{u}} f_{\mathbf{v}}(\mathbf{x}_{\mathbf{v}}).$$

- Zeroth order component:

$$f_{\{\emptyset\}} = f(\overline{\mathbf{x}})$$

- First and second order components:

$$f_{\{x_i\}} = f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}\setminus x_i} - f_{\{\emptyset\}}$$
$$f_{\{x_i,x_j\}} = f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}\setminus\{x_i,x_j\}} - f_{\{\emptyset\}} - f_{\{x_i\}} - f_{\{x_j\}}$$

Note on anchored-ANOVA and Taylor Series

- The multivariate Taylor series of $f(\mathbf{x})$ about $\bar{\mathbf{x}}$,

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \sum_{i=1}^{d} \frac{\partial f(\mathbf{x})}{\partial x_i} (x_i - \bar{x}_i) + \frac{1}{2!} \sum_{i,j=1}^{d} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} (x_i - \bar{x}_i)(x_j - \bar{x}_j) + \ldots$$

- Evaluate at $\mathbf{x} = \mathbf{a} \setminus x_i$,

$$f(\mathbf{x})|_{\mathbf{x}=\mathbf{a}\setminus x_i} = f(\bar{\mathbf{x}}) + \frac{\partial f(\mathbf{x})}{\partial x_i} (x_i - \bar{x}_i) + \frac{1}{2!} \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2} (x_i - \bar{x}_i)^2 + \ldots$$

- Component functions are entire Taylor series expansions.
- Truncated anchored-ANOVA expansion will always be a better approximation than a truncated Taylor expansion.

Combining anchored-ANOVA and Smolyak Sparse Grids

- anchored-ANOVA decomposition forms foundation for surrogate model.
- Build each component function on a Smolyak sparse grid.
- Build all first-order components first.
- Compute sensitivity coefficient for each component.

$$\eta_i = \frac{\int_{\Omega_i} \left[f(\mathbf{x})|_{\mathbf{x}=\overline{\mathbf{x}}\setminus x_i} - f(\overline{\mathbf{x}}) \right] \rho(x_i) dx_i}{f_{\{\emptyset\}}}$$

- Build higher order components for only those variables whose sensitivity coefficients exceed a certain threshold.
- Continue to add higher order terms until convergence criteria met.

Outline

Infinite Lattice Problem Statement

- Reactor of infinite size is considered, effectively removing the effects of geometry in neutron transport
- System entirely characterized in terms of its material properties.
- Analytic solution available!

$$\begin{pmatrix} \Sigma_{a_1} + \Sigma_{1\to 2} & 0 \\ -\Sigma_{12} & \Sigma_{a_2} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \frac{1}{k_\infty} \begin{pmatrix} \nu\Sigma_{f_1} & \nu\Sigma_{f_2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

$$k_\infty = \frac{\Sigma_{a_2}\nu\Sigma_{f_1} + \Sigma_{1\to 2}\nu\Sigma_{f_2}}{\Sigma_{a_2}\left(\Sigma_{a_1} + \Sigma_{1\to 2}\right)}$$

Infinite Lattice Problem Statement

- Assume all variation in k_∞ can be attributed to its input cross sections, whose distributions follow a multivariate Gaussian.

			Correlation Coefficient Matrix				
	Mean	**Standard Dev.**	Σ_{a_1}	Σ_{a_2}	$\nu\Sigma_{f_1}$	$\nu\Sigma_{f_2}$	$\Sigma_{1\to2}$
Σ_{a_1}	1.04E-02	9.06E-05	1	0.07	-0.13	0.02	0.75
Σ_{a_2}	1.10E-01	2.31E-04	0.07	1	0.06	0.31	-0.07
$\nu\Sigma_{f_1}$	9.00E-03	4.85E-05	-0.13	0.06	1	0.33	-0.10
$\nu\Sigma_{f_2}$	1.91E-01	8.87E-04	0.02	0.31	0.33	1	0.01
$\Sigma_{1\to2}$	1.80E-02	2.18E-04	0.75	-0.07	-0.10	0.01	1

Infinite Lattice Problem Objective

- Propagate cross section uncertainties to find the mean and standard deviation of k_∞ using Monte Carlo sampling of true and surrogate functions, and the "Sandwich Equation".

$$\sigma^2(k_\infty) = S^T C S$$

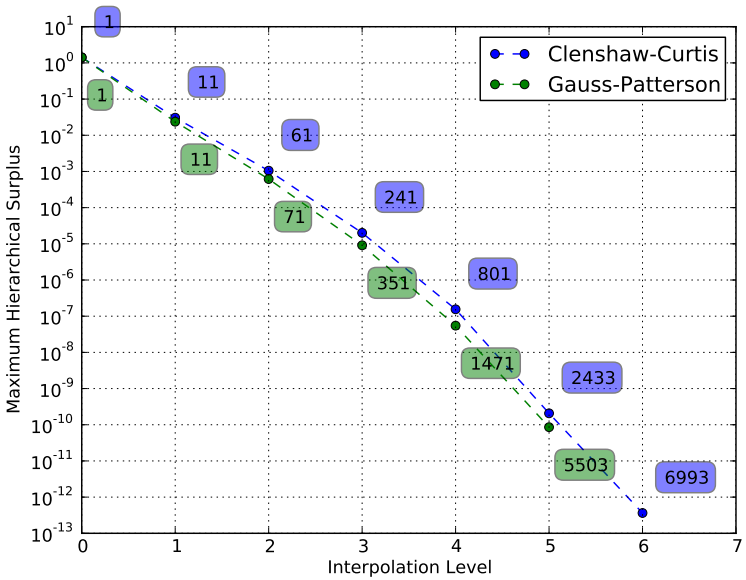
$$S^T = \left(\begin{array}{ccccc} \frac{\partial k_\infty}{\partial \Sigma_{a_1}} & \frac{\partial k_\infty}{\partial \Sigma_{a_2}} & \frac{\partial k_\infty}{\partial \nu\Sigma_{f_1}} & \frac{\partial k_\infty}{\partial \nu\Sigma_{f_2}} & \frac{\partial k_\infty}{\partial \Sigma_{1\to 2}} \end{array} \right)$$

- Also, find sensitivity coefficients analytically and numerically.

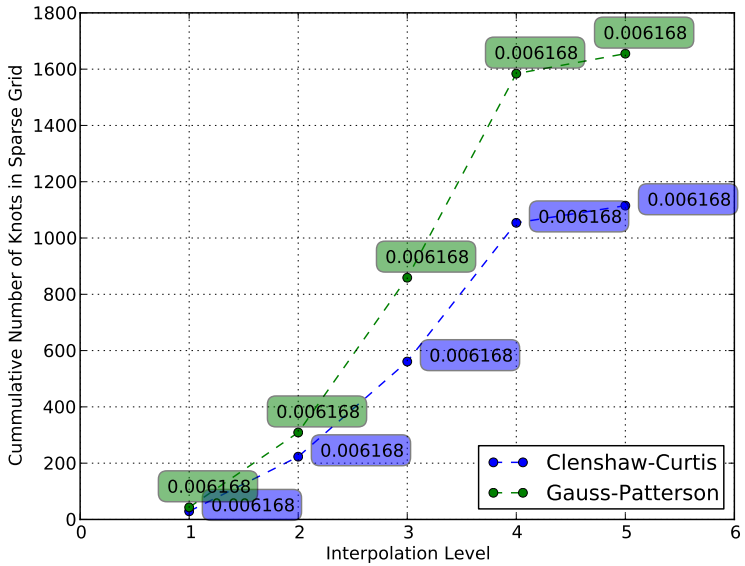
$$\left.\frac{\partial k_\infty}{\partial \Sigma_i}\right|_{\Sigma_{j\neq i}=\bar{\Sigma}_j} \approx \frac{k_\infty(\Sigma_i + \Delta\Sigma_i) - k_\infty(\Sigma_i - \Delta\Sigma_i)}{2\Delta\Sigma_i}$$

Convergence of 5-D Sparse Grid Interpolant

Hierarchical Surplus



Number of Abscissas



Mean and Variance of k_∞

- 1000 Monte Carlo samples, using the same random numbers, were gathered for each method.
- In "1D ANOVA" only the 5 first-order components are used.
- In "All ANOVA" all 32 components are used.
- Clenshaw-Curtis collocation points represented by "CC" while "GP" is Gauss-Patterson.

Method	**Mean**	**99% CI**	**Standard Dev.**	**99% CI**
5D Sparse Grid CC	1.41562	(1.41512, 1.41612)	0.006168	(0.005909, 0.006544)
5D Sparse Grid GP	1.41562	(1.41512, 1.41612)	0.006168	(0.005831, 0.006544)
1D ANOVA CC	1.41560	(1.41510, 1.41610)	0.006168	(0.005831, 0.006544)
All ANOVA CC	1.41562	(1.41512, 1.41612)	0.006168	(0.005831, 0.006544)
1D ANOVA GP	1.41560	(1.41510, 1.41610)	0.006168	(0.005831, 0.006544)
All ANOVA GP	1.41562	(1.41512, 1.41612)	0.006168	(0.005831, 0.006544)
True Function	1.41562	(1.41512, 1.41612)	0.006168	(0.005831, 0.006544)
Sandwich			0.006540	

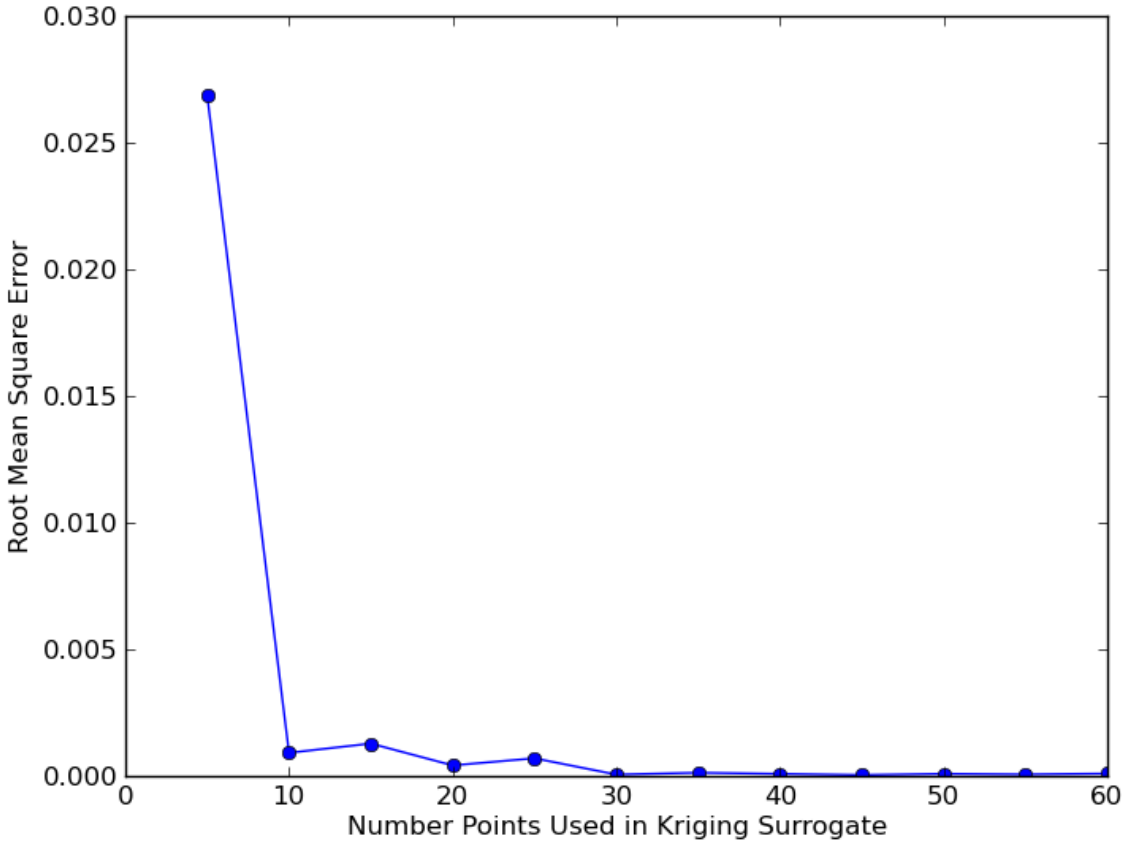
Sensitivity Coefficients for k_∞

- Central differencing... perturbations (1%) made to each cross section while holding the others at mean value.
- Models differ from analytic coefficients in fourth decimal place
- Expected due to $\mathcal{O}(\Delta\Sigma^2)$ convergence of central differencing.

	Normalized Sensitivity Coefficient of k_∞				
Method	Σ_{a_1}	Σ_{a_2}	$\nu\Sigma_{f_1}$	$\nu\Sigma_{f_2}$	$\Sigma_{1\to 2}$
5D Sparse Grid CC	-.367551	-.776087	.224060	.776010	.143491
5D Sparse Grid GP	-.367551	-.776087	.224060	.776010	.143491
1D ANOVA CC	-.367556	-.776098	.224063	.776020	.143493
All ANOVA CC	-.367551	-.776087	.224060	.776010	.143491
1D ANOVA GP	-.367556	-.776098	.224063	.776020	.143493
All ANOVA GP	-.367551	-.776087	.224060	.776010	.143491
Analytic	-.367520	-.775956	.224044	.775956	.143476
Central Difference	-.367551	-.776089	.224060	.776011	.143492

Kriging Surrogate for k_∞

- Observe that for ten evaluation points, twice the number of design variables in the objective function, error goes to zero.



Point Kinetics/Lumped Thermal Hydraulics Problem Statement

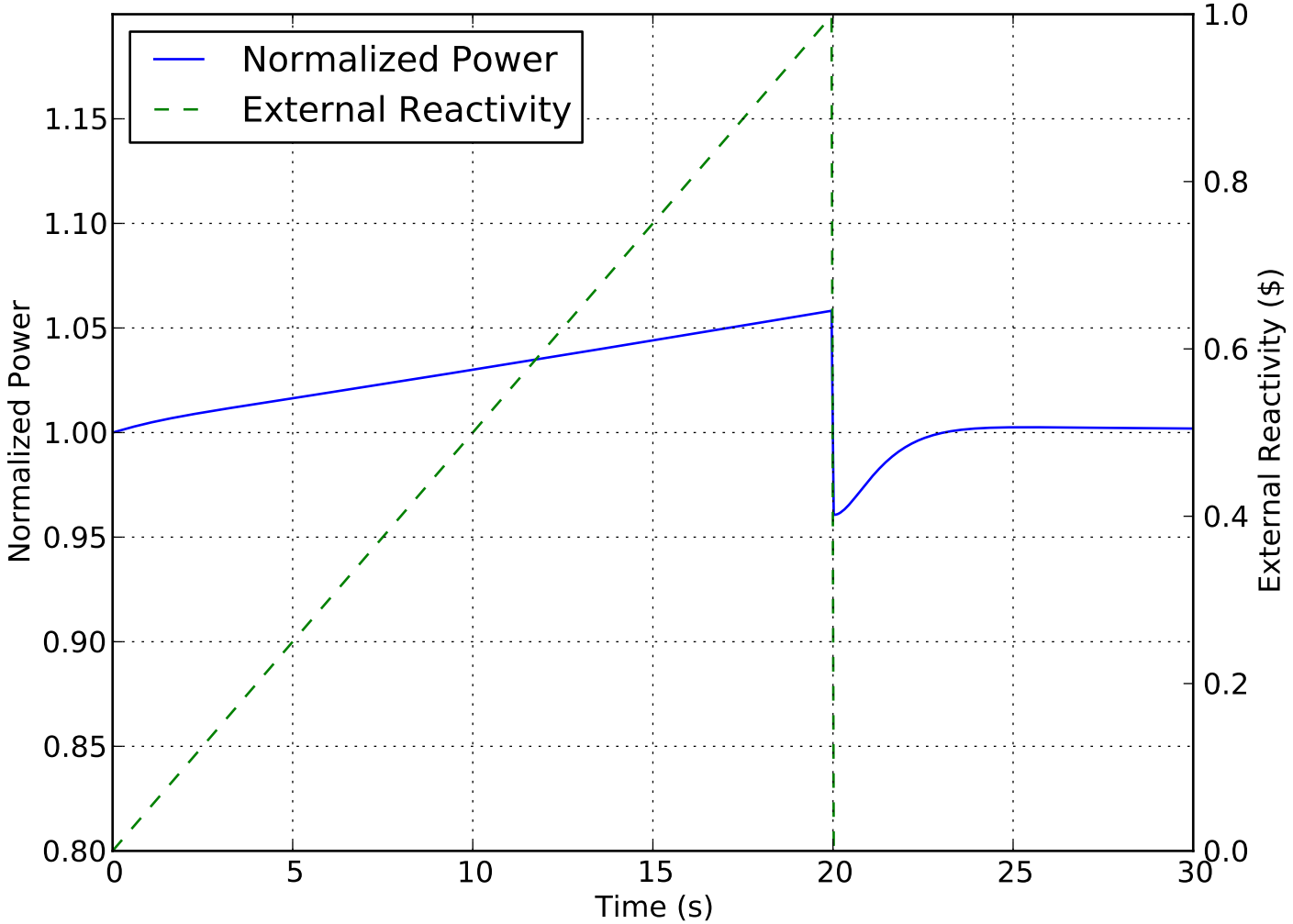
- Modeling a transient resulting from a half sawtooth external reactivity insertion in a BN800 sodium fast cooled reactor.

$$\rho_{ex}(t) = \begin{cases} t\rho_{max}/20 & t \leq 20 \\ 0 & t > 20 \end{cases}$$

- The physical model of the reactor consists of point kinetics to model the neutronics and lumped thermal hydraulics equations to describe temperature feedback.
- The coupled system is nonlinear and only has a time dependence.
- Kinetics coupled to thermal hydraulics through reactivity.

$$\rho(T_f, T_c, t) = \rho_{ex} + \alpha_d(T_f - T_f(0)) + \alpha_c(T_c - T_c(0))$$

Point Kinetics/Lumped Thermal Hydraulics Problem Statement



Point Kinetics/Lumped Thermal Hydraulics Problem Statement

- Reactor Power

$$\frac{dP}{dt} = \frac{\rho(T_f, T_c, t) - \beta}{\Lambda} P + \sum_{k=1}^{6} \lambda_k C_k$$

- Precursor Concentration

$$\frac{dC_k}{dt} = -\lambda_k C_k + \frac{\beta_k}{\Lambda} P$$

- Fuel Temperature

$$M_f c_{pf} \frac{dT_f}{dt} = P + Ah(T_c - T_f)$$

- Coolant Temperature

$$M_c c_{pc} \left(\frac{dT_c}{dt} + v \frac{T_c - T_{in}}{L} \right) = Ah(T_f - T_c)$$

Point Kinetics/Lumped Thermal Hydraulics Objective

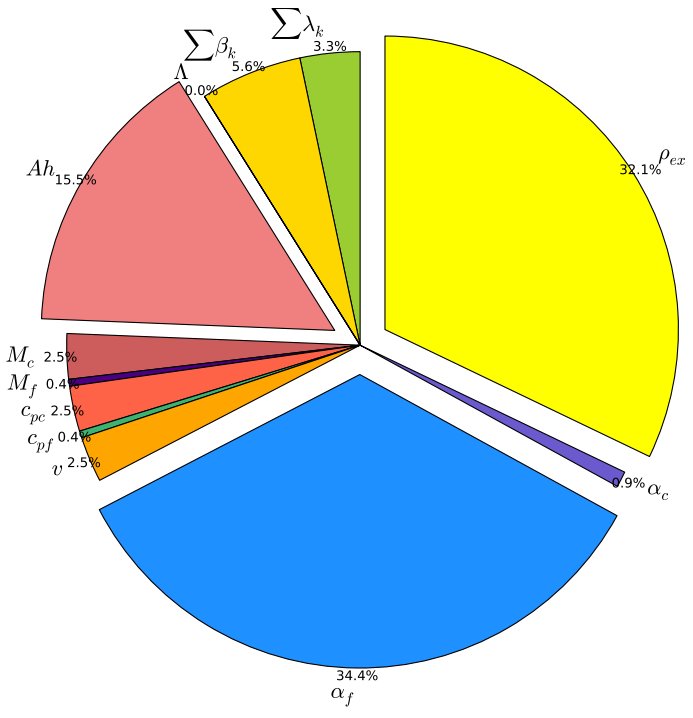
- A total of twenty two random variables will be investigated for their affect on the maximum fuel temperature attained during transient.

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$$
$$\Lambda, Ah, M_c, M_f, c_{pc}, c_{pf}, v, \alpha_d, \alpha_c, \rho_{max}$$

- All random variables are assumed to be independent of one another.
- Objective is to identify "important variables", perform sensitivity analysis, and calculate basic statistics using anchored-ANOVA collocation and Kriging.

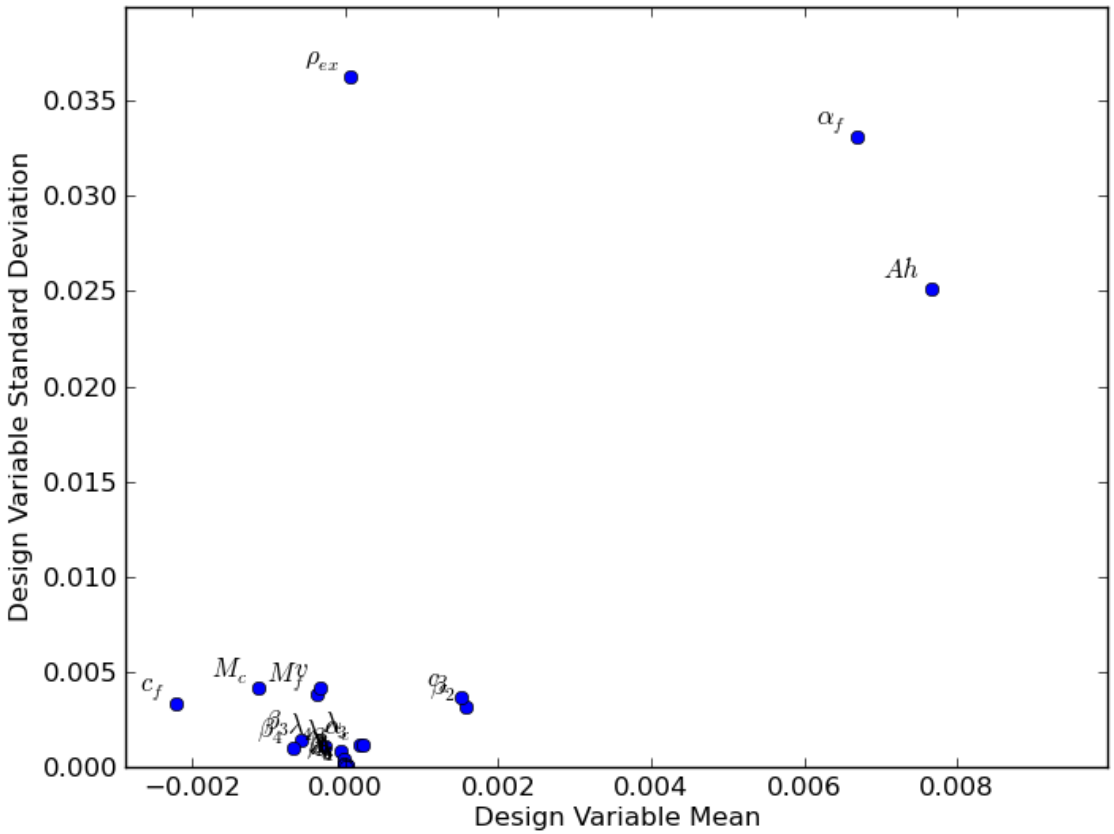
anchored-ANOVA Sensitivity Analysis

- For each first-order component in anchored-ANOVA decomposition, calculate a sensitivity coefficient.
- Variables Ah, α_d, and ρ_{max} comprise 82% of the total sensitivity.



Morris' Algorithm Sensitivity Analysis

- Algorithm visually identifies Ah, α_d, and ρ_{max} to be "important variables".

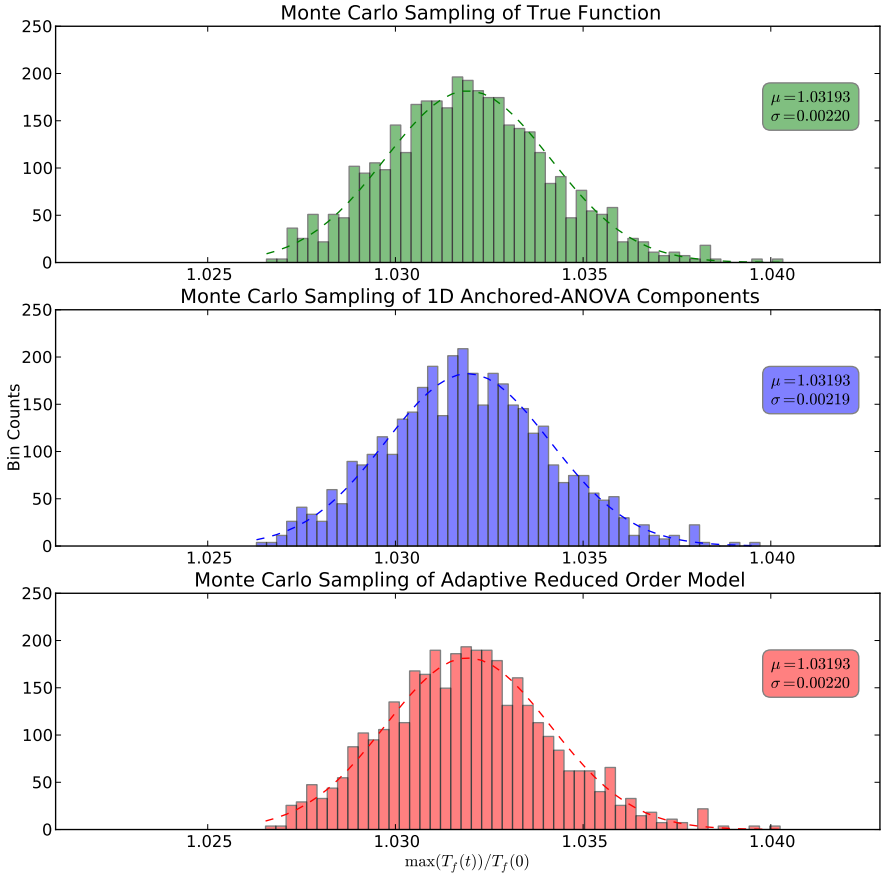


Statistics from anchored-ANOVA Collocation Surrogate

- Basic statistics for the maximum fuel temperature attained during transient.
- A total of 1000 Monte Carlo samples were used to get the statistics for each method (using the same random numbers).
- A Kriging surrogate based on 3 "important variables" was constructed and also sampled 1000 times for a mean normalized temperature of 1.03198 ± 0.002299.

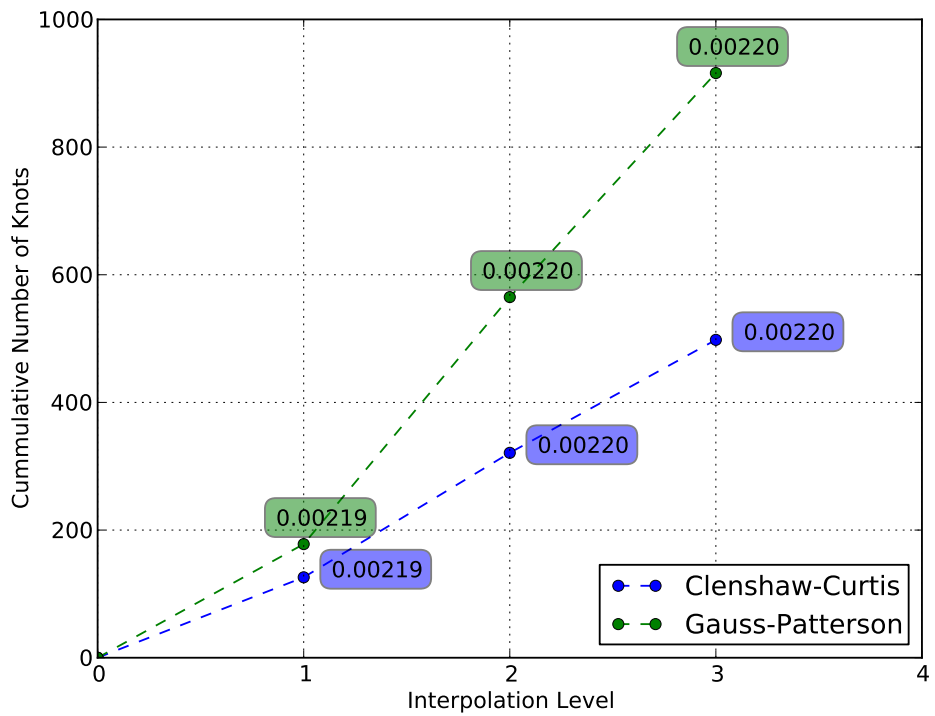
Method	**Mean**	**99% CI**	**Standard Dev.**	**99% CI**
1D ANOVA CC	1.03193	(1.03175, 1.03211)	0.002187	(0.002068, 0.002320)
All ANOVA CC	1.03193	(1.03175, 1.03211)	0.002196	(0.002076, 0.002330)
1D ANOVA GP	1.03193	(1.03175, 1.03211)	0.002187	(0.002068, 0.002320)
All ANOVA GP	1.03193	(1.03175, 1.03211)	0.002196	(0.002076, 0.002330)
True Function	1.03193	(1.03175, 1.03211)	0.002196	(0.002076, 0.002330)

Statistics from anchored-ANOVA Collocation Surrogate



Statistics from anchored-ANOVA Collocation Surrogate

- Higher-order ANOVA components built only to model interaction effects between Ah, α_d, and ρ_{max}.
- Surrogate consisting entirely of first-order components required 126 function evaluations.

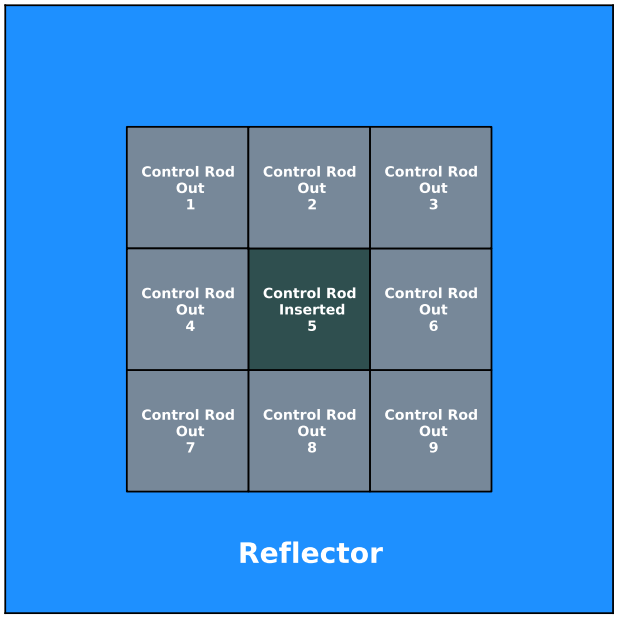


Calculating Sensitivity Coefficients on anchored-ANOVA Collocation Surrogates

Random Variable	1D ANOVA CC	All ANOVA CC	Central Diff.
λ_4	9.7308E-04	9.7309E-04	9.7379E-04
β_2	-2.6582E-03	-2.6582E-03	-2.6616E-03
Λ	-8.9294E-08	-8.9295E-08	-8.9364E-08
Ah	1.2553E-02	1.2553E-02	1.2584E-02
M_c	1.8753E-03	1.8753E-03	1.8716E-03
M_f	-3.6695E-04	-3.6695E-04	-3.6360E-04
c_{pc}	1.8753E-03	1.8753E-03	1.8903E-03
c_{pf}	-3.6695E-04	-3.6695E-04	-3.5976E-04
v	1.8838E-03	1.8839E-03	1.9177E-03
α_d	-2.6655E-02	-2.6656E-02	-2.6625E-02
α_c	8.4387E-04	8.4387E-04	8.7194E-04
ρ_{max}	3.1164E-02	3.1164E-02	3.1272E-02

Three Mile Island (TMI) Minicore Problem Statement

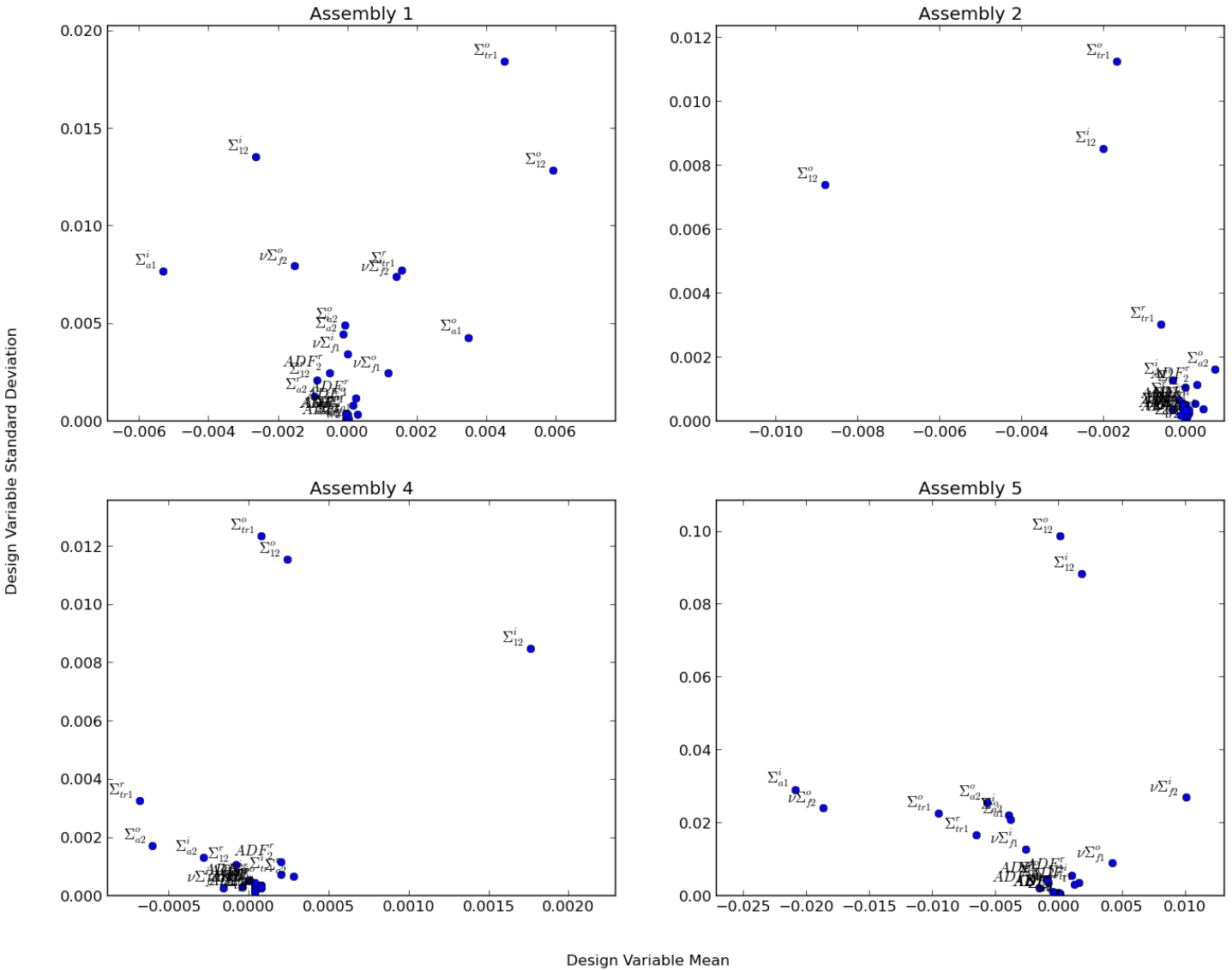
- Build surrogate for TMI Minicore power distribution using core simulator code PARCS.
- Total of 25 homogenized, two-group cross sections are used as design variables.



TMI Minicore Problem Data

- The few group, homogenized cross section description for each fuel assembly consists of transport, absorption, nu-fission, and scatter cross sections along with values for ADFs.
- For a two-group problem the total number of cross sections to describe an assembly is nine.
- Since the homogenized reflector region does not support fission only seven homogenized cross sections are required to describe it.
- Few-group covariance matrix is obtained using the 'Two-Step' method based on a total of 300 transport calculations with perturbed multigroup cross sections.

Morris' Algorithm Sensitivity Study



Kriging Surrogate

- Kriging surrogate is constructed for the core power as a function of Σ^i_{a1}, $\nu\Sigma^i_{f2}$, Σ^i_{12}, Σ^o_{tr1}, $\nu\Sigma^o_{f2}$ and Σ^o_{12}.
- Using these variables a sampling plan consisting of 20 true PARCS evaluations was designed.
- Based on the optimized sampling plan a Kriging surrogate was constructed and sampled 500 times in accordance with the six design variables' covariance matrix.

Sampling the Kriging Surrogate

Kriging Surrogate Statistics

Assembly	Mean	99% CI	Standard Dev.
1	0.8385	(0.8384, 0.8386)	0.0010
2	1.1499	(1.1497, 1.1500)	0.0013
3	0.8384	(0.8383, 0.8386)	0.0010
4	1.1499	(1.1497, 1.1500)	0.0013
5	1.0465	(1.0462, 1.0467)	0.0022
6	1.1499	(1.1497, 1.1500)	0.0013
7	0.8384	(0.8383, 0.8386)	0.0010
8	1.1499	(1.1497, 1.1500)	0.0013
9	0.8384	(0.8383, 0.8386)	0.0010

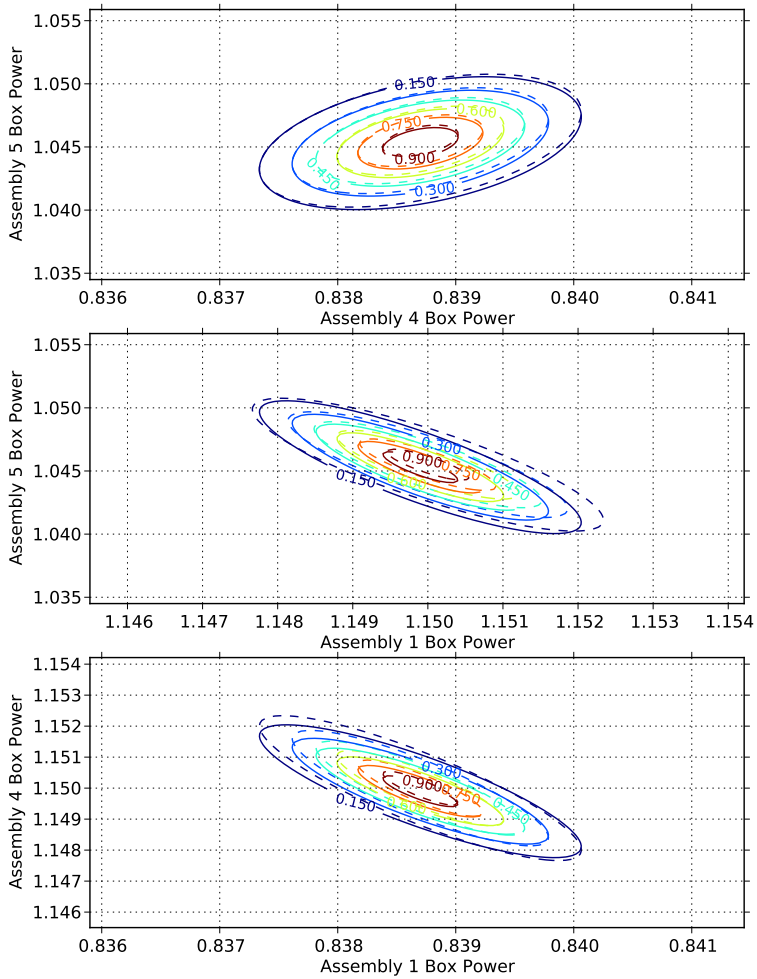
PARCS Two-Step Statistics

Assembly	Mean	99% CI	Standard Dev.
1	0.8387	(0.8386, 0.8388)	0.0007
2	1.1499	(1.1498, 1.1500)	0.0011
3	0.8387	(0.8386, 0.8388)	0.0007
4	1.1499	(1.1498, 1.1500)	0.0011
5	1.0453	(1.0450, 1.0456)	0.0027
6	1.1499	(1.1498, 1.1500)	0.0011
7	0.8387	(0.8386, 0.8388)	0.0007
8	1.1499	(1.1498, 1.1500)	0.0011
9	0.8387	(0.8386, 0.8388)	0.0007

anchored-ANOVA Collocation Surrogate

- Constructed surrogate using only first-order components.
- Required 123 evaluations of PARCS using Clenshaw-Curtis abscissas.
- Hierarchical surplus of 10^{-4} used to designate convergence.
- Surrogate sampled 500 times and multivariate distributions between assembly boxes calculated and compared to "true" solution.

Using anchored-ANOVA Collocation Surrogate to Reproduce Multivariate Distribution



Outline

Original Contributions of Research

- Surrogate use to replace expensive nuclear fuel performance simulations for calibration, optimization and inference.
- Folding in fuel performance experimental data with surrogate simulations to improve predictive capabilities of fuel performance codes.
- Use of the anchored-ANOVA decomposition coupled with collocation in nuclear engineering applications.
- Application of surrogate methods to multiphysics code systems.
- Use of dimension reduction techniques along with Kriging never before applied to nuclear fuel performance modeling.

Work Left to be Completed

- Apply methods described here to construct surrogate for Risø AN3 fission gas release model in BISON using DAKOTA framework.
- Use surrogate to infer "true" value of fission gas release parameters by folding in data from power ramp experiment.
- Replace BISON with coupled BISON/MPACT and repeat to study effects building a surrogate on a multiphysics code system.

The End

Thank you

Outline

Sparse Grid Mechanics Example

- Consider a function of two dimensions $f(x_1, x_2)$, $d = 2$.
- Let's calculate the Level 1 (q=3) Smolyak sparse grid interpolant for f, $A_{3,2}(f)$.

$$\begin{aligned} A_{3,2}(f) &= A_{2,2}(f) + \Delta A_{3,2} \\ &= \Delta A_{2,2} + \cancel{A_{1,2}(f)} + \Delta A_{3,2} \end{aligned}$$

- First let's calculate $\Delta A_{2,2}$.

$$\Delta A_{2,2} = \sum_{|\mathbf{i}|=2} \left(f(x_{j_1}^{i_1}, x_{j_2}^{i_2}) - \cancel{A_{1,2}(f)(x_{j_1}^{i_1}, x_{j_2}^{i_2})} \right) \cdot \left(a_{j_1}^{i_1} \otimes a_{j_2}^{i_2} \right)$$

Sparse Grid Mechanics Example

- $|\mathbf{i}| = 2 \rightarrow i_1 + i_2 = 2 \rightarrow$ sum over the set of i indices $\{(1,1)\}$.
- Using Clenshaw-Curtis for $i = 1$ we have $m_i = 1$, namely the origin.

$$\Delta A_{2,2} = f\left(x_1^1, x_1^1\right) a_1^1 a_1^1$$

- In this case, $a_1^1 = 1$ and $x_1^1 = 0$ so $\Delta A_{2,2}$ amounts to evaluating the objective function at the hypercube origin.

Sparse Grid Mechanics Example

- Now let's calculate $\Delta A_{3,2}$.
- $|\mathbf{i}| = 3 \rightarrow i_1 + i_2 = 3 \rightarrow$ sum over the set of i indices $\{(1,2),(2,1)\}$.
- Using Clenshaw-Curtis for $i = 2$ we have $m_i = 3$ so $j \in \{1,2,3\}$

$$\begin{aligned}\Delta A_{3,2} &= \sum_{|\mathbf{i}|=3} \left(f(x_{j_1}^{i_1}, x_{j_2}^{i_2}) - A_{2,2}(f)(x_{j_1}^{i_1}, x_{j_2}^{i_2}) \right) \cdot \left(a_{j_1}^{i_1} \otimes a_{j_2}^{i_2} \right) \\ &= \sum_{|\mathbf{i}|=3} \left(f(x_{j_1}^{i_1}, x_{j_2}^{i_2}) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1 \right) \cdot \left(a_{j_1}^{i_1} \otimes a_{j_2}^{i_2} \right)\end{aligned}$$

Sparse Grid Mechanics Example

$$
\begin{aligned}
\Delta A_{3,2} = & \left[f(x_1^2, x_1^1) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^2 a_1^1 \\
& + \left[f(x_2^2, x_1^1) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_2^2 a_1^1 \\
& + \left[f(x_3^2, x_1^1) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_3^2 a_1^1 \\
& + \left[f(x_1^1, x_1^2) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^1 a_1^2 \\
& + \left[f(x_1^1, x_2^2) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^1 a_2^2 \\
& + \left[f(x_1^1, x_3^2) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^1 a_3^2
\end{aligned}
$$

- Due to the nestedness of Clenshaw-Curtis, points can be reused.
- For example, $f(x_1^1, x_1^1) = f(x_1^1, x_2^2) = f(x_2^2, x_1^1)$.

Sparse Grid Mechanics Example

- Putting everything together, the level 1 Sparse grid interpolant is:

$$
\begin{aligned}
A_{3,2}(f) &= \Delta A_{2,2} + \Delta A_{3,2} \\
&= f\left(x_1^1, x_1^1\right) a_1^1 a_1^1 \\
&+ \left[f(x_1^2, x_1^1) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^2 a_1^1 \\
&+ \left[f(x_3^2, x_1^1) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_3^2 a_1^1 \\
&+ \left[f(x_1^1, x_1^2) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^1 a_1^2 \\
&+ \left[f(x_1^1, x_3^2) - f\left(x_1^1, x_1^1\right) a_1^1 a_1^1\right] a_1^1 a_3^2
\end{aligned}
$$