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

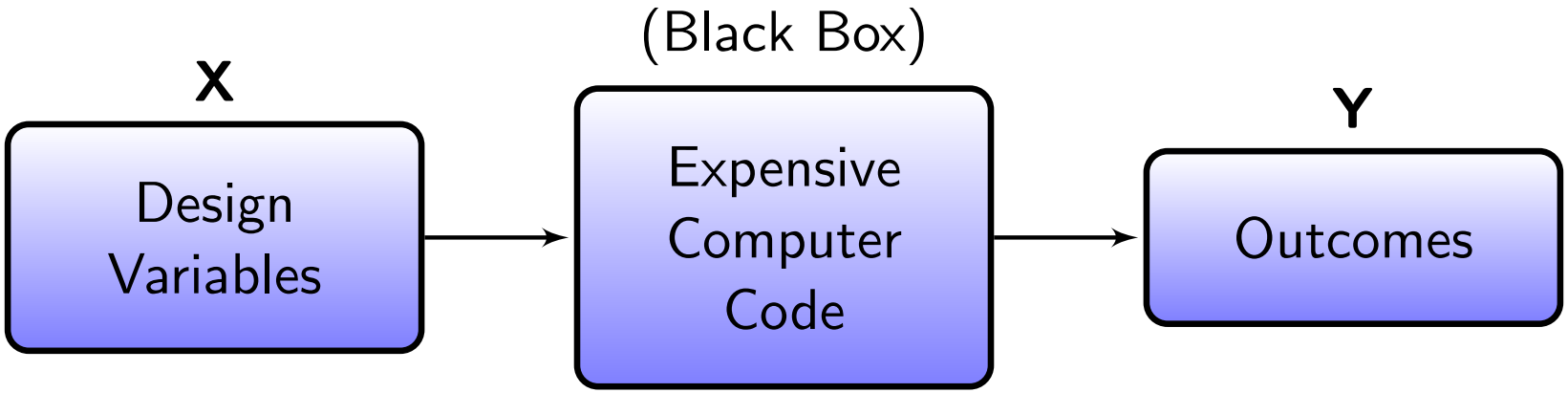
- Physical phenomena is studied by conducting experiments.
- Any data collected represents instances of underlying stochastic processes.
- We build predictive computer models in an attempt to reproduce such observed physical phenomena.
- To accurately capture stochastic element of experiments, computer 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

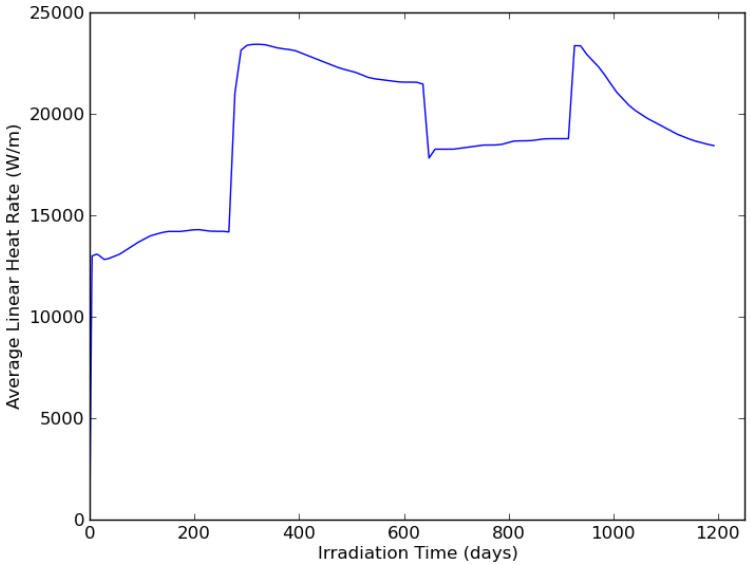
- Simple Integrated Fission Gas Release and Swelling (SIFGRS)
- Incorporates gas diffusion and precipitation in grains, growth and coalescence of gas bubbles at grain faces, thermal, athermal, steady-state, and transient gas release.
- Through a direct description of the grain face gas bubble development, 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

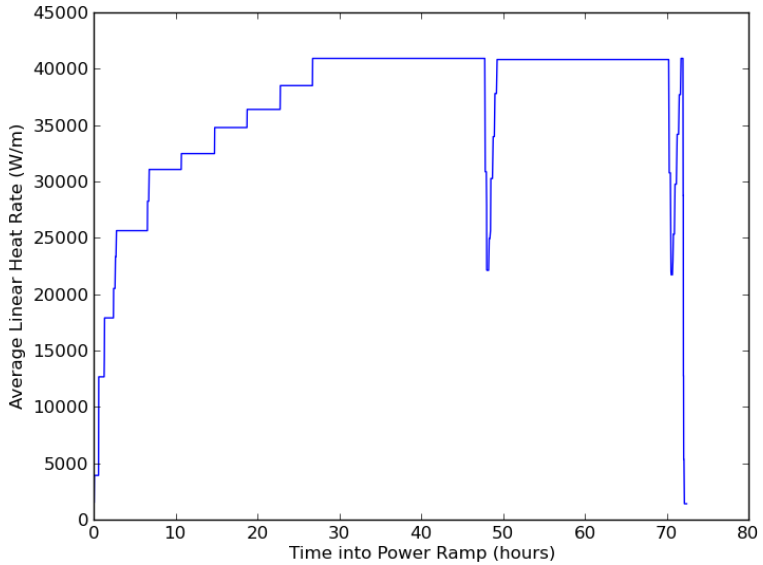
- Validation case for fuel performance modeling in the Fumex-II database.
- Experiment consists of a base irradiation of four reactor cycles in the Biblis A pressurized water reactor.
- After the base irradiation period, a fuel rod is extracted and refabricated to a shorter length before undergoing a 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



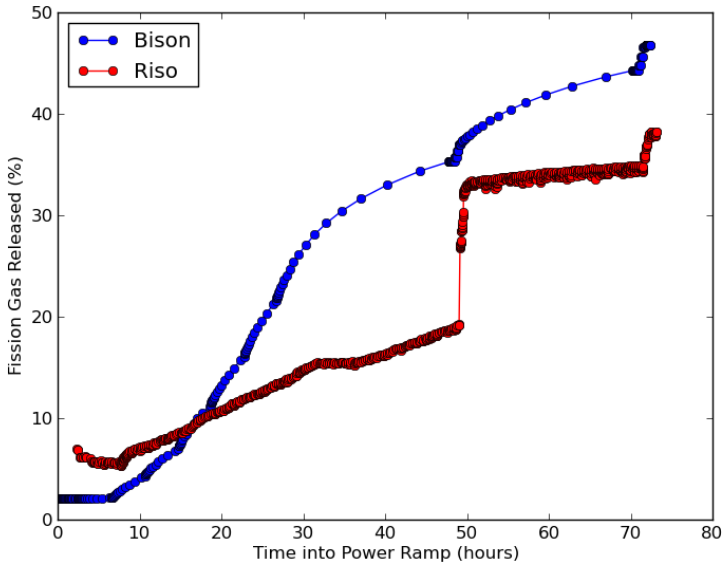
Power Ramp Experiment



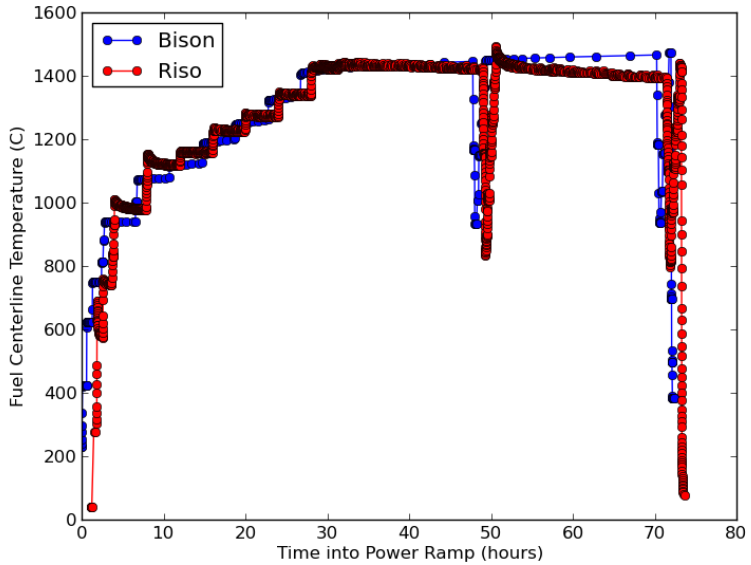
Modeling Risø AN3 Experiment with BISON

- 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

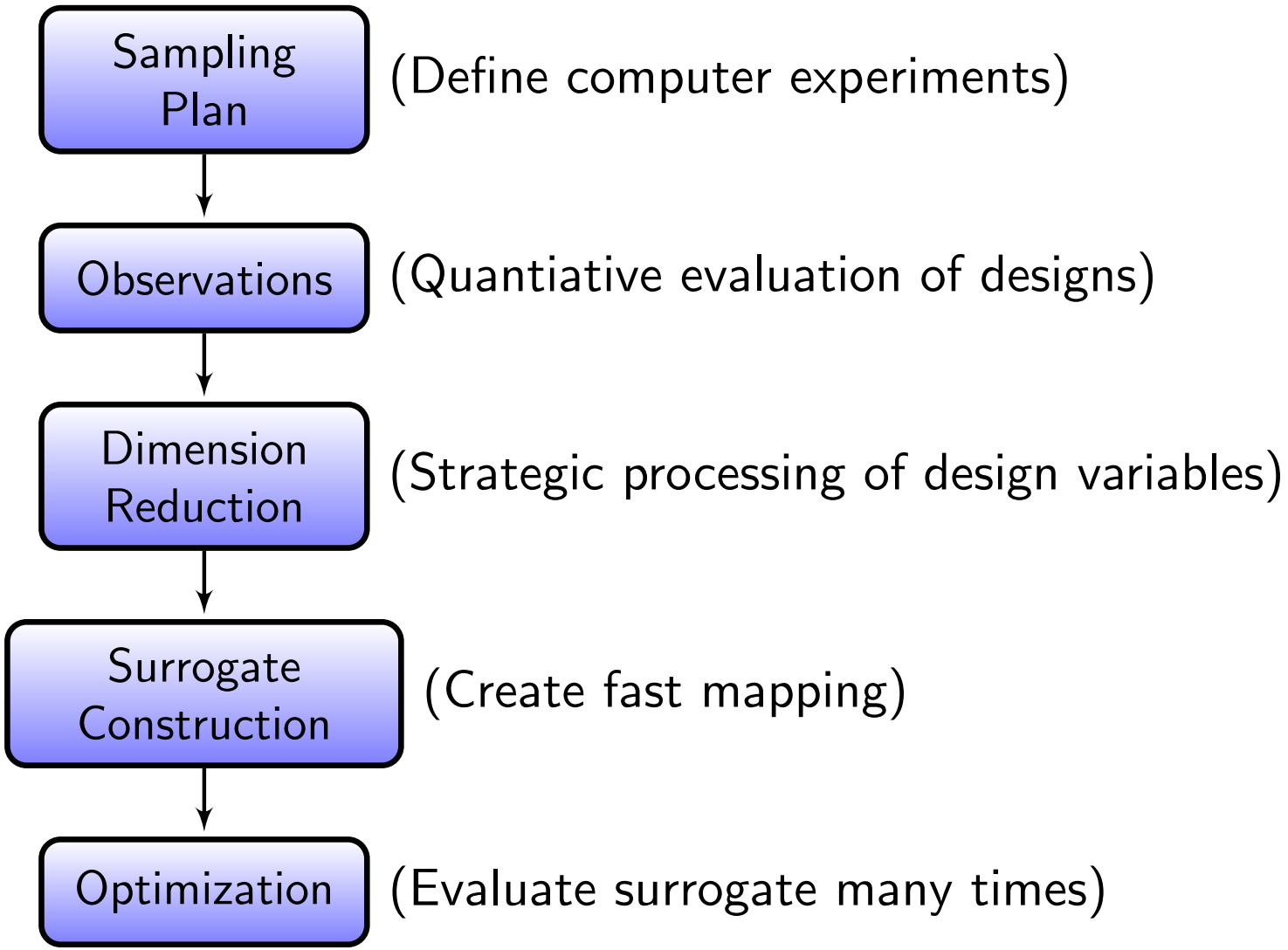
- No sense in comparing the output of a computer simulation to experimental data unless the computer simulation is of high fidelity and capable of reproducing the pertinent physics.
- MPACT is a neutronics code that provides detailed intrapin and azimuthally dependent neutronics data in the fuel elements.
- 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

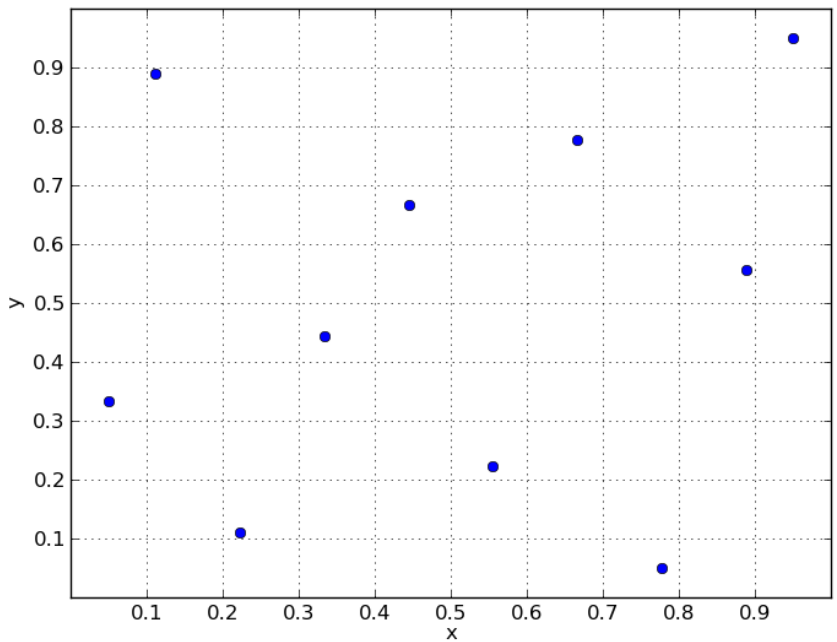
- Dimension reduction inherent
- Sampling done on structured grid
- Sampling plan size dependent on number of design variables
- Polynomial interpolation
- More deterministic approach

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,

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

Outline

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(q=3, d=2).

$$\begin{aligned} A(3,2) &= A(2,2) + \Delta A(3,2) \\ &= \Delta A(2,2) + \Delta A(3,2) \end{aligned}$$