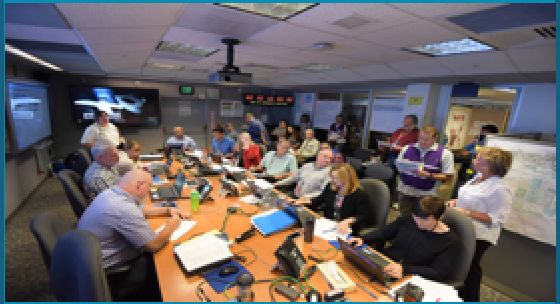
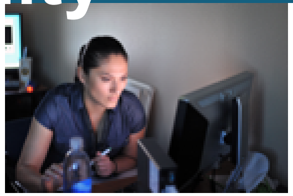
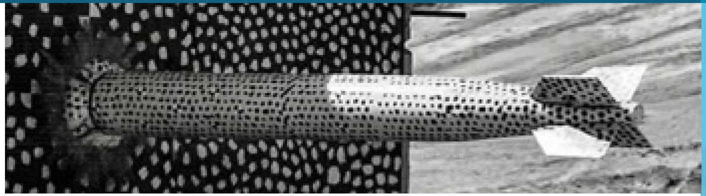
Sandia National Laboratories

SAND2021-6970C

Adaptive resource allocation for surrogate modeling of systems comprised of multiple disciplines with varying fidelity







John Jakeman Sandia National Laboratories

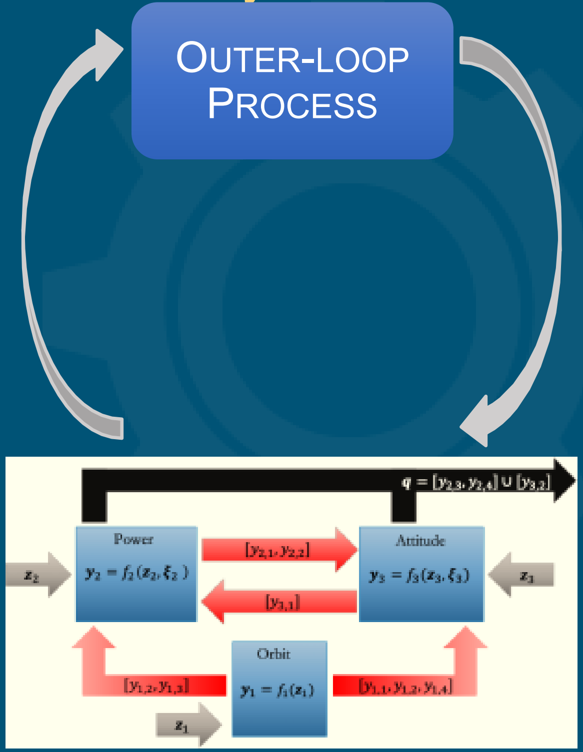
Samuel Friedman (TAMU), Michael Eldred (SNL), Lorenzo Tamellini (CNR-IMARTI), Alex Gorodetsky (UMich), Doug Allaire (TAMU)

IX International Conference on Coupled Problems in Science and Engineering

U.S. DEPARTMENT OF ENERGY NNSA

PROBLEM SUMMARY

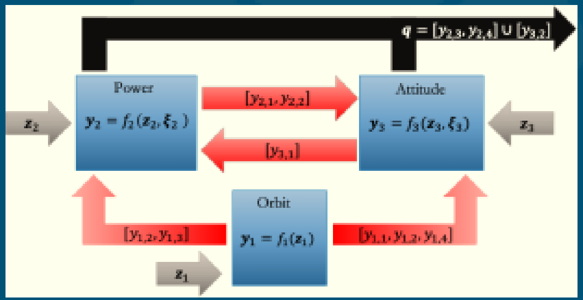
Goal:
UQ of coupled systems



Problem Summary

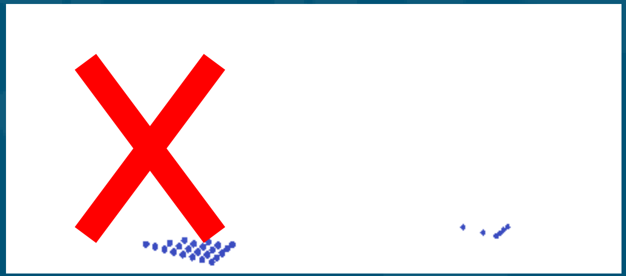
Goal:
UQ of coupled systems

Outer-loop Process



Challenge:
Expensive models intractable

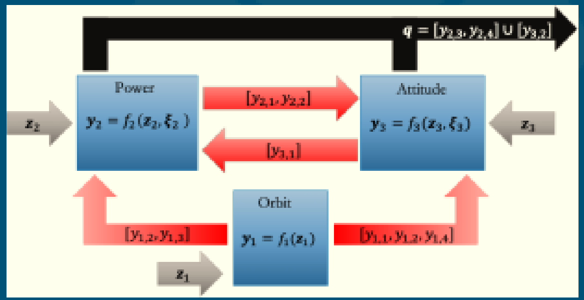
Outer-loop Process



Problem Summary

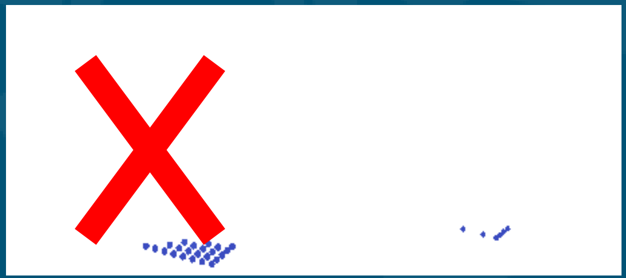
Goal:
UQ of coupled systems

Outer-loop Process



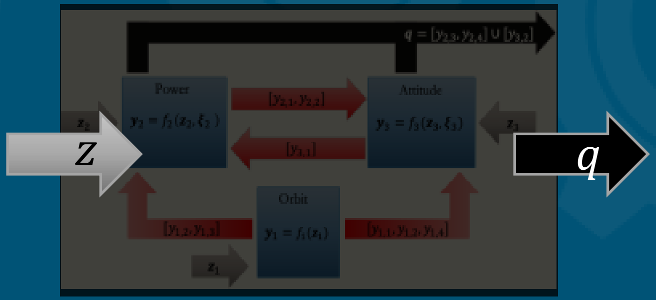
Challenge:
Expensive models intractable

Outer-loop Process



Existing approach:
Black-box surrogate

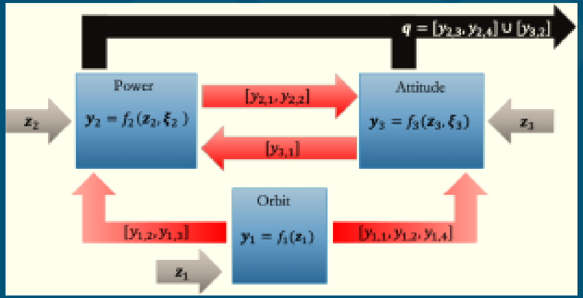
Outer-loop Process



Problem Summary

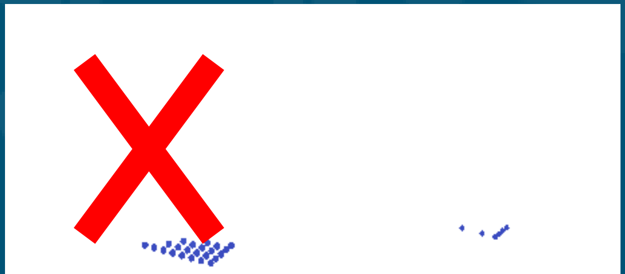
Goal:
UQ of coupled systems

Outer-loop Process



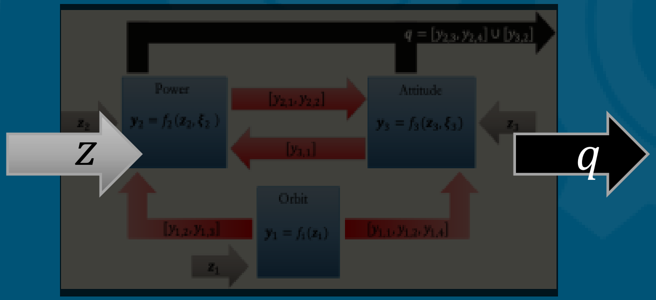
Challenge:
Expensive models intractable

Outer-loop Process



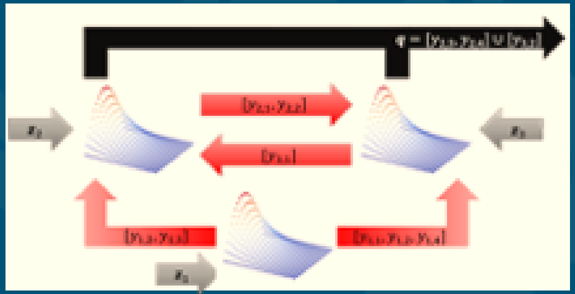
Existing approach:
Black-box surrogate

Outer-loop Process



Novel solution:
Build surrogates of components

Outer-loop Process



Outcome:
Reduce cost by 10-100X

Problem Setup

Map from system parameters to all outputs $\boldsymbol{y} = f(\boldsymbol{z}) : \Gamma \to \Upsilon$

Collection of all system parameters

$$\boldsymbol{z} = (z_1, \ldots, z_D)^\top \in \Gamma \subseteq \mathbb{R}^D$$

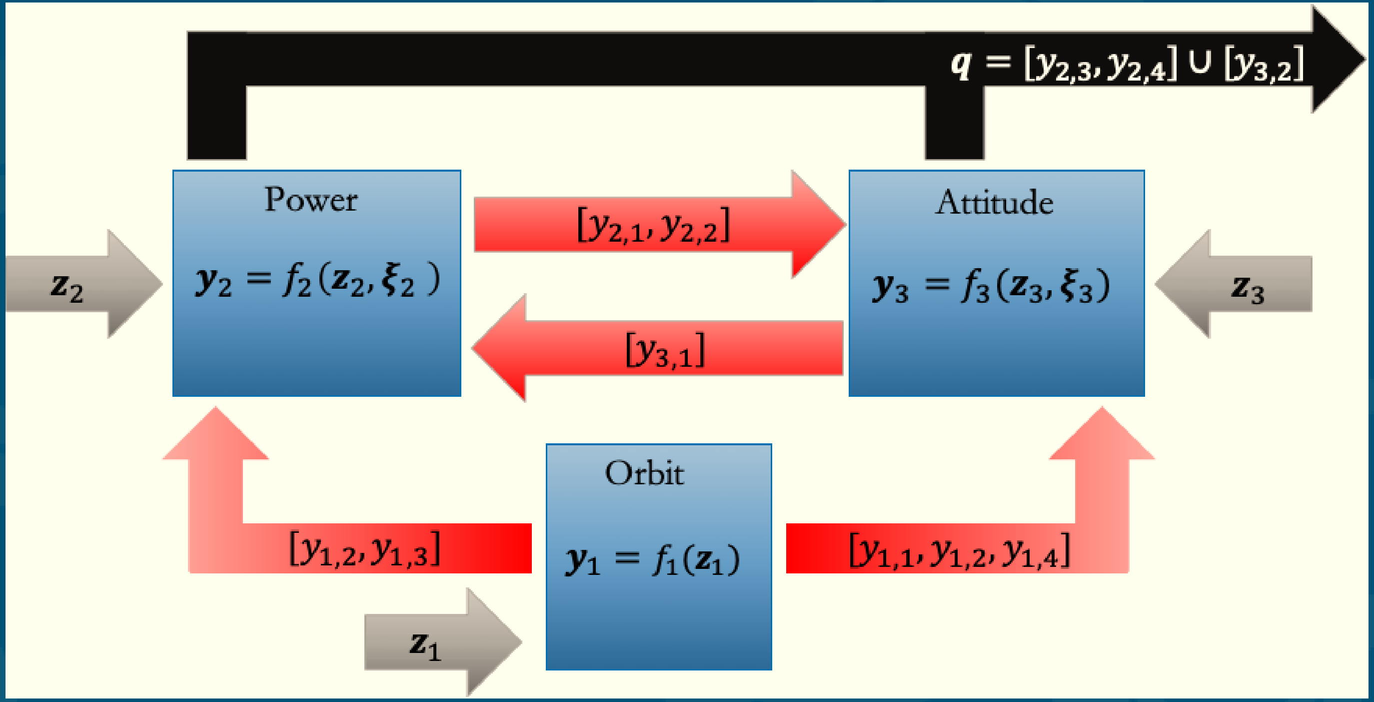
System comprised of K Components

$$\boldsymbol{y}_k = f_k(\boldsymbol{z}_k, \boldsymbol{\xi}_k) : \Gamma_k \times \Xi_k \to \Upsilon_k,$$

Components have subset of system variables

$$\boldsymbol{z}_k = \boldsymbol{R}_k^z \boldsymbol{z} \in \mathbb{R}^{D_k} \quad \boldsymbol{\xi}_k = \boldsymbol{R}_k^\xi \boldsymbol{y}, \in \mathbb{R}^{S_k}$$

Denote evaluation of system via the functional $G_z : (f_1, \ldots, f_K) \mapsto f(z)$.



Goal to predict subset of system outputs

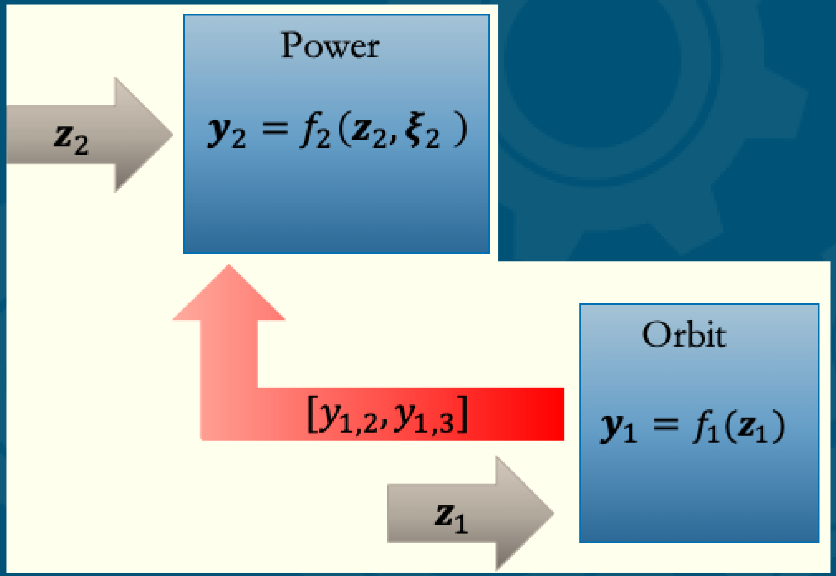
$$\boldsymbol{q} = \boldsymbol{R}^y \boldsymbol{y} \in \mathbb{R}^{Q^{\text{sys}}}$$

Evaluating Systems Of Components

Feed-forward coupling

$$y_k = f_k(z_k, \xi_k) \qquad \xi_k = y_{k-1} = f_{k-1}(z_{k-1}),$$

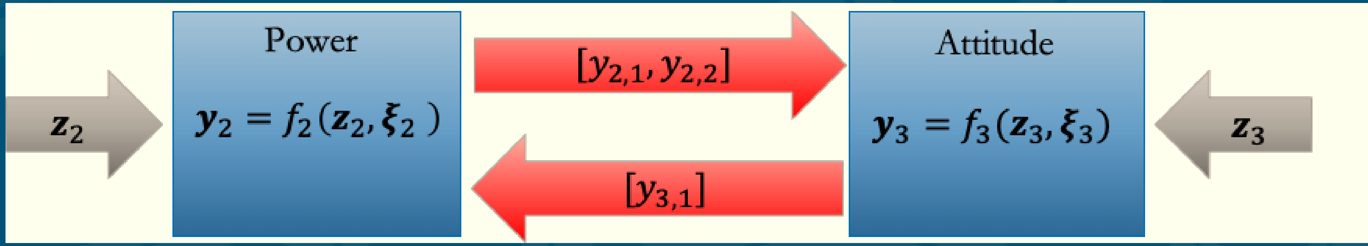
Simply pass output y_{k-1} as in put to f_k



Evaluating Systems Of Components

Feedback-coupling

$$\begin{cases} \boldsymbol{y}_j = f_j(\boldsymbol{z}_j, \boldsymbol{\xi}_j), & \boldsymbol{\xi}_j = \boldsymbol{y}_k \\ \boldsymbol{y}_k = f_k(\boldsymbol{z}_k, \boldsymbol{\xi}_k), & \boldsymbol{\xi}_j = \boldsymbol{y}_j \end{cases}$$



Use iteration function

$$F(\xi) = \begin{bmatrix} f_j(z_j, \xi_j) \\ f_k(z_k, \xi_k) \end{bmatrix} \qquad \xi = [\xi_j, \xi_k]^\top$$

Starting from ξ^0 we iterate until

$$\xi^t = F(\xi^{t-1}) \qquad \|\xi^t - \xi^{t-1}\| < \nu$$

Model Fidelity

Often models of varying fidelity (accuracy and cost) are available for a component

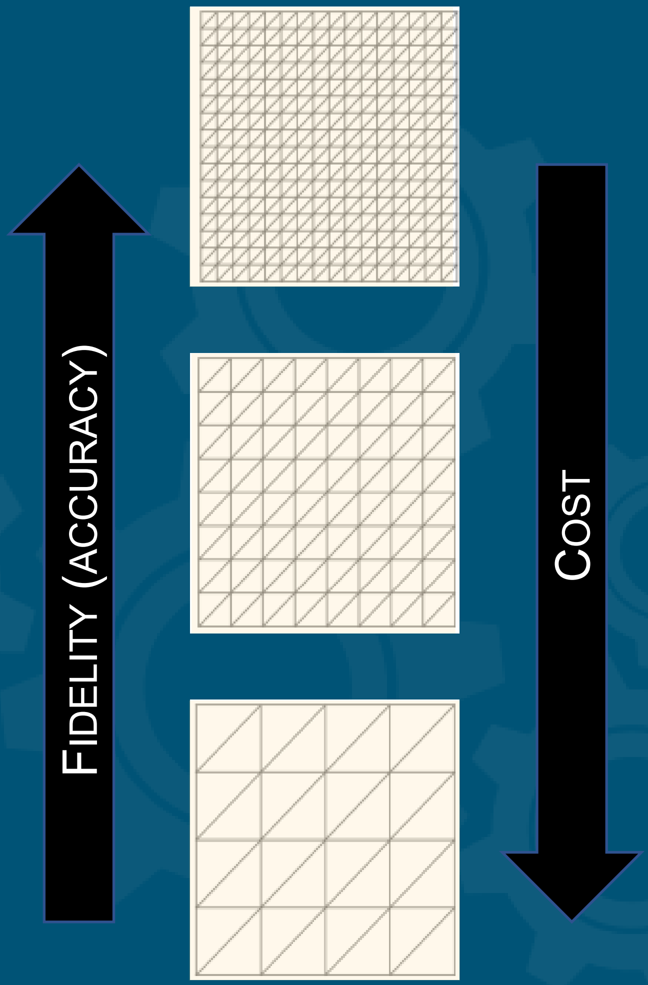
Assume models have hyper-parameters, e.g. mesh size, time-step, convergence tolerance, etc.

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{R_k}) \in \mathbb{N}^{R_k}$$

Assume error decreases with fidelity

$$\|f_{k,\boldsymbol{\alpha}^*} - f_k\| \le \|f_{k,\boldsymbol{\alpha}} - f_k\| \quad \boldsymbol{\alpha}^* \ge \boldsymbol{\alpha},$$

System prediction accuracy depends on component fidelity



Surrogate Modeling

Surrogates can reduce cost of using high-fidelity components

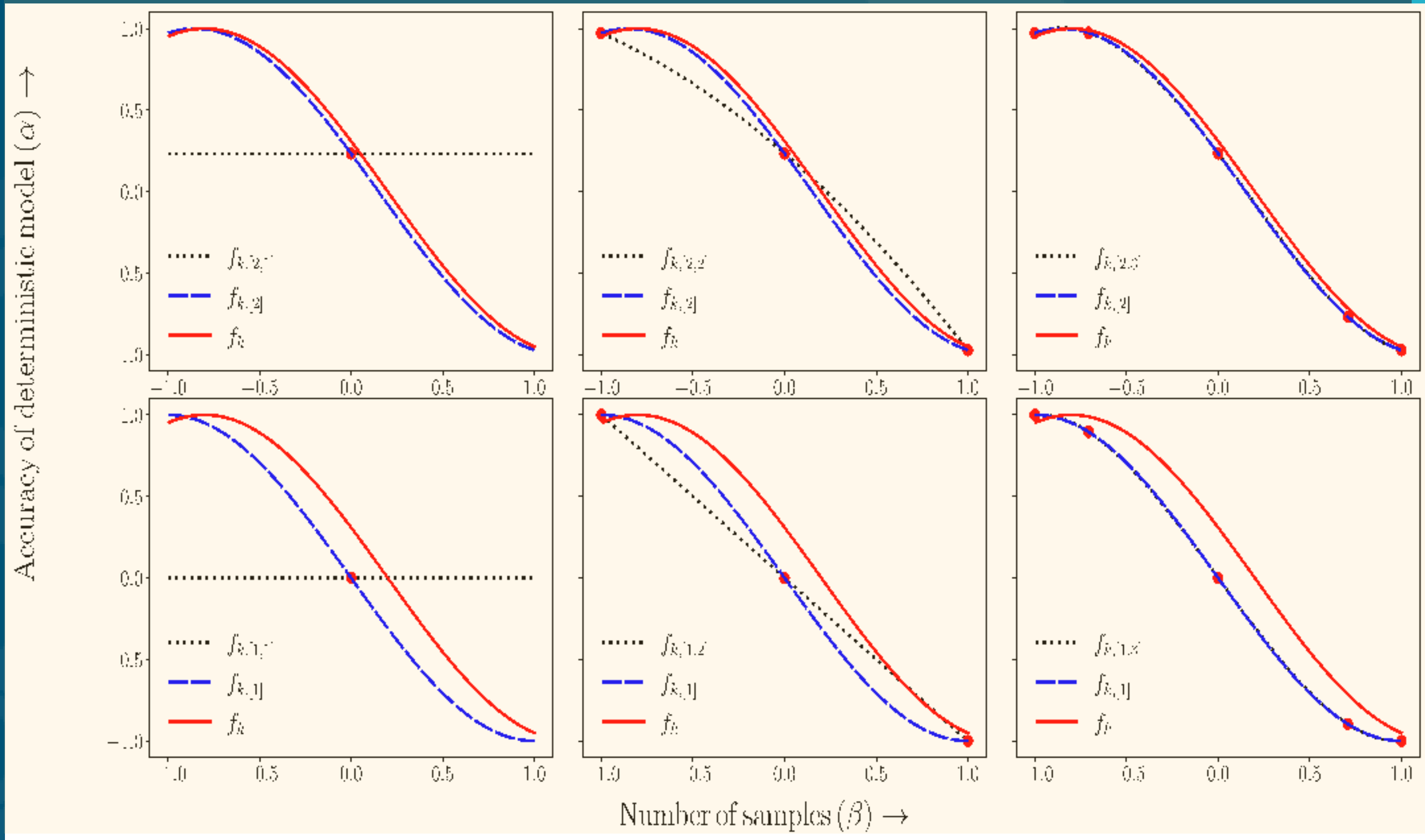
$$f_{k,[\boldsymbol{\alpha},\boldsymbol{\beta}]}(\boldsymbol{u}_k) \approx f_k(\boldsymbol{u}_k)$$

$$\boldsymbol{u}_k := [\boldsymbol{z}_k^\top, \boldsymbol{\xi}_k^\top]^\top$$

Use training data

$$\mathcal{U}_{k,\boldsymbol{\beta}} = \{\boldsymbol{u}_k^{(m)}\}_{m=1}^M$$

$$\{f_{k,\boldsymbol{\alpha}}(\boldsymbol{u}_k^{(m)})\}_{m=1}^M$$



Error depends on fidelity α and surrogate error β

$$\|f_k - f_{k,[\alpha,\beta]}\| \le \|f_k - f_{k,\alpha}\| + \|f_{k,\alpha} - f_{k,[\alpha,\beta]}\|$$

Characterizing The Coupling Variables

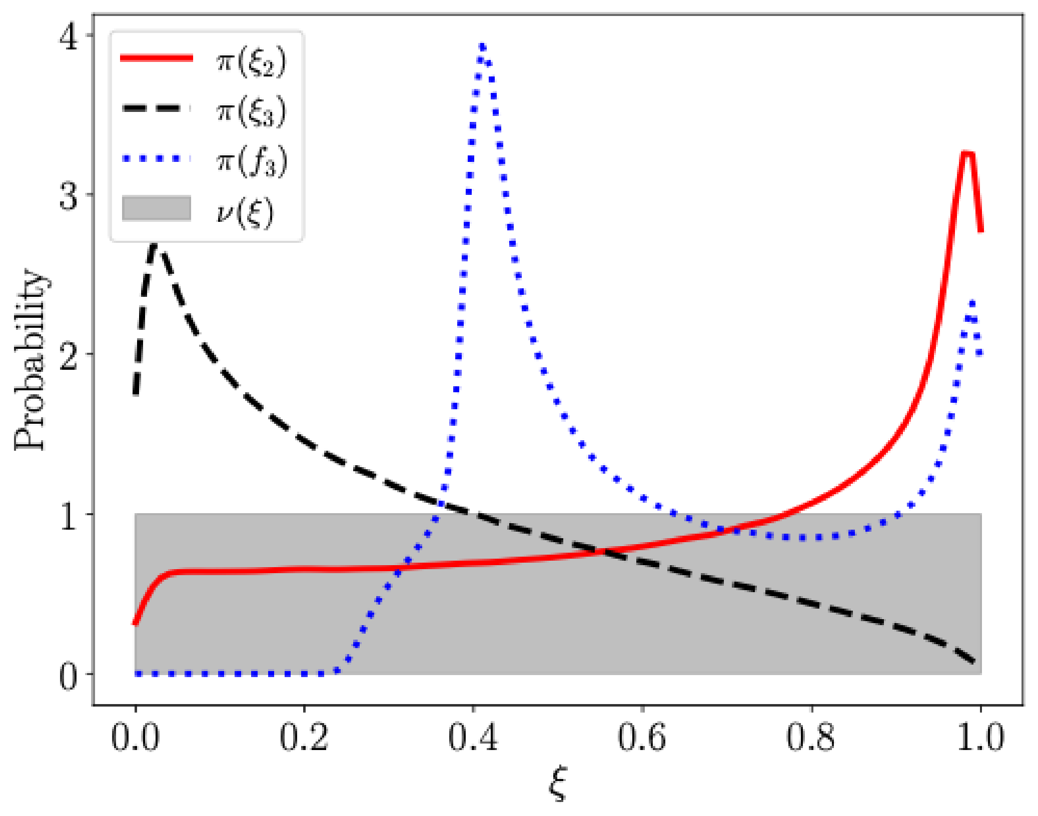
We must specify the ranges of the coupling variables.

$$\boldsymbol{y}_k = f_k(\boldsymbol{z}_k, \boldsymbol{\xi}_k)$$

The ranges of the coupling variables are unknown a priori

The estimated range of the coupling variables matters

The distribution of the coupling variables matters a lot less



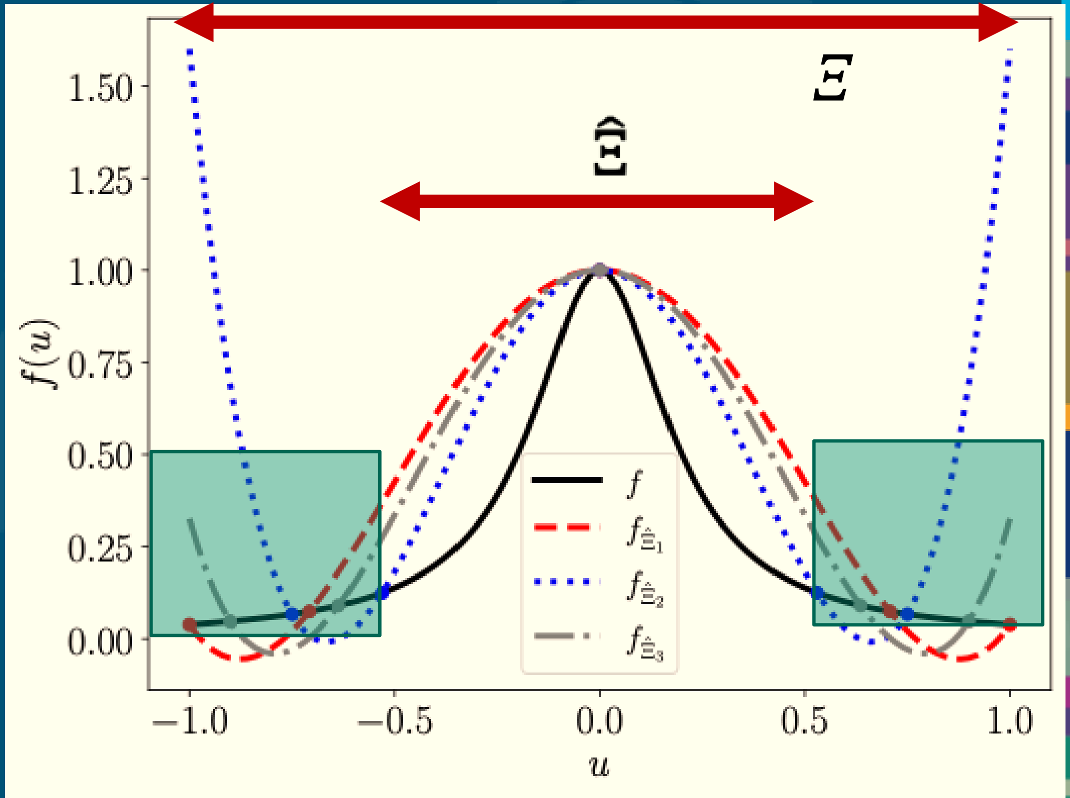
Characterizing The Coupling Variables

We must specify the ranges of the coupling variables.

Lemma 3.1 (Strong convergence [9]).

Let $\hat{\omega} : \hat{\Xi} \to \mathbb{R}$ *and* $\omega : \Xi \to \mathbb{R}$ *denote two densities which satisfy*

$$\delta = 1 - \int_{\Xi \setminus \hat{\Xi}} \omega(u)\,du$$



Characterizing The Coupling Variables

We must specify the ranges of the coupling variables.

Lemma 3.1 (Strong convergence [9]).

Let $\pi : \hat{\Xi} \to \mathbb{R}$ and $\omega : \Xi \to \mathbb{R}$ denote two densities which satisfy

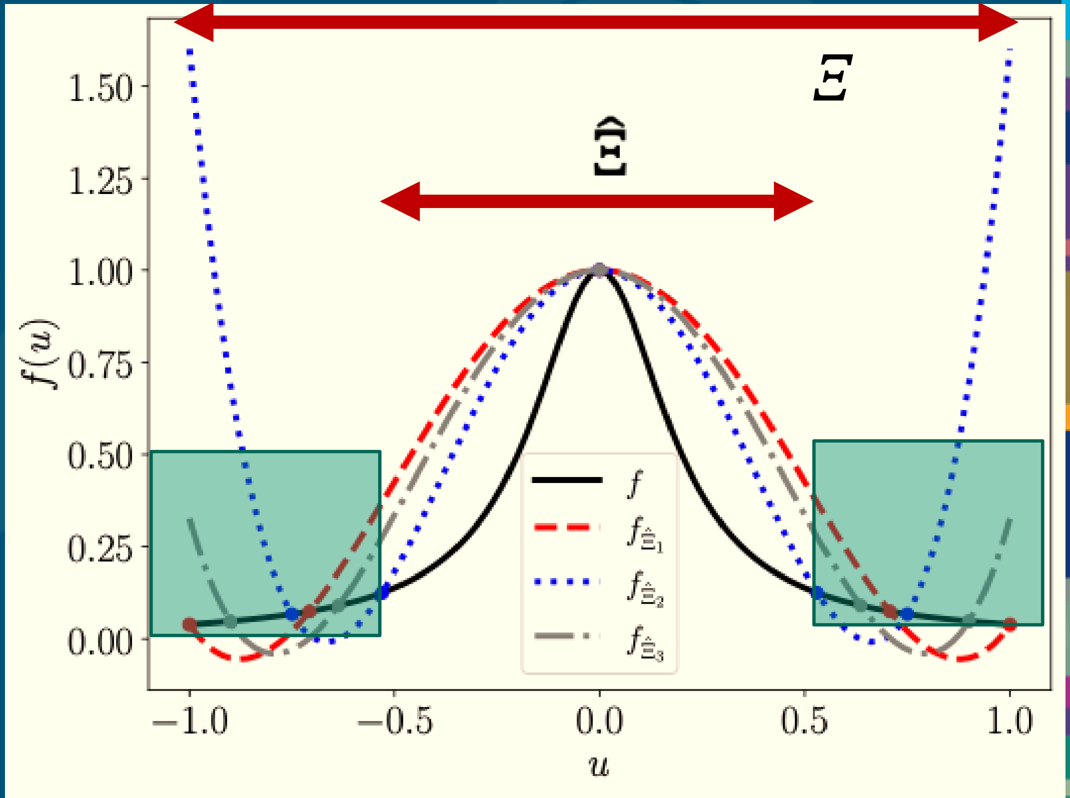
$$\delta = 1 - \int_{\Xi \cap \hat{\Xi}} \omega(u)\,du$$

Given an approximation f_π of f with approximation error ϵ, i.e.,

$$\epsilon := \|f - f_\pi\|_{L^p_\omega(\Xi)}, \quad p \geq 1,$$

then, if f is bounded with $C_f = \|f\|_{L^\infty(\Xi)}$, it holds that

$$\|f - f_\pi\|_{L^p_\omega(\Xi)} \leq C_r^{1/p}\epsilon + C_f\,\delta^{1/p}, \quad \text{provided} \quad C_r := \max_{u \in \Xi \cup \hat{\Xi}} \frac{\omega(u)}{\pi(u)} < \infty\,.$$



Evaluating Systems Using Component Surrogates

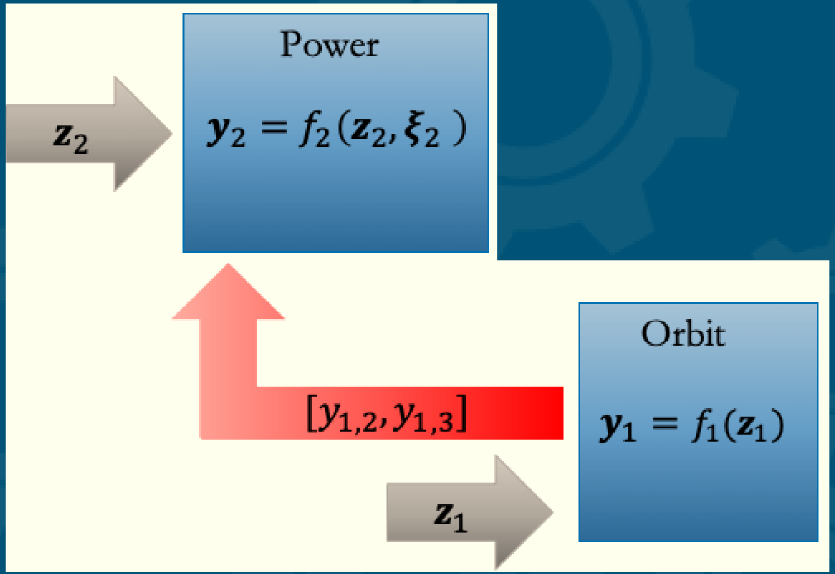
Proposition 3.1 (Feed-forward surrogate error)

Assume f_k is Lipschitz continuous. If

$$\|y_{k,q} - y_{k,\mathcal{I}_k,q}\|_{L^\infty(\Gamma)} \leq \epsilon_k \forall q = 1, \ldots, Q,$$

and

$$f = f_{1,\mathcal{I}_1} \circ f_{2,\mathcal{I}_2} \circ \cdots \circ f_{K,\mathcal{I}_K} \quad f = f_1 \circ f_2 \circ \cdots \circ f_K$$



Evaluating Systems Using Component Surrogates

Proposition 3.1 (Feed-forward surrogate error)

Assume f_k is Lipschitz continuous. If

$$\|y_{k,q} - y_{k,\mathcal{I}_k,q}\|_{L^\infty(\Gamma)} \leq \epsilon_k \forall q = 1,\ldots,Q,$$

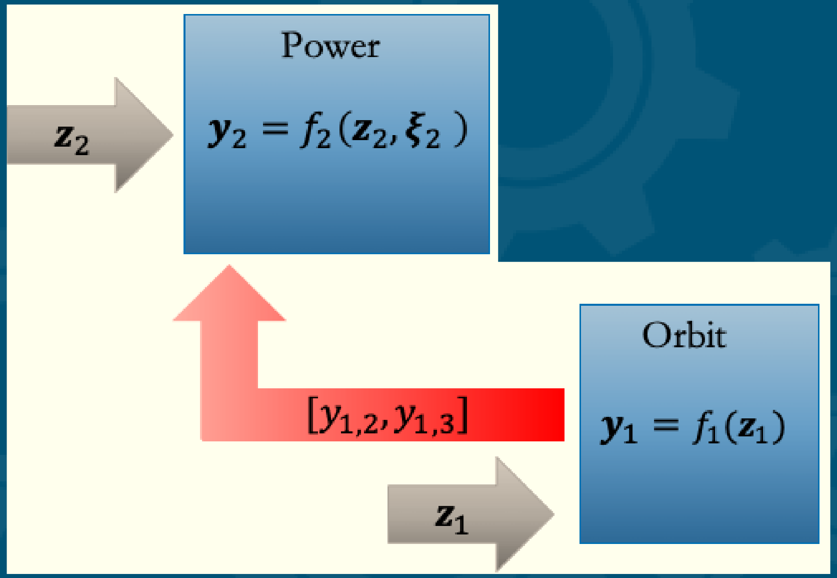
and

$$f = f_{1,\mathcal{I}_1} \circ f_{2,\mathcal{I}_2} \circ \cdots \circ f_{K,\mathcal{I}_K} \quad f = f_1 \circ f_2 \circ \cdots \circ f_K$$

then

$$\sup_{\boldsymbol{z}\in\Gamma} \max_{q=1,\ldots,Q} |f_q(\boldsymbol{z}) - \hat{f}_q(\boldsymbol{z})| \leq \epsilon \frac{1-L^K}{1-L}$$

where $L = \max_{k=1,\ldots,K} L_k$ *and* $\epsilon = \max_{k=1,\ldots,K} \epsilon_k$.



Evaluating Systems Using Component Surrogates

Proposition 3.1 (Feed-forward surrogate error)

Assume f_k is Lipschitz continuous. If

$$\|y_{k,q} - y_{k,\mathcal{I}_k,q}\|_{L^\infty(\Gamma)} \leq \epsilon_k \forall q = 1, \dots, Q,$$

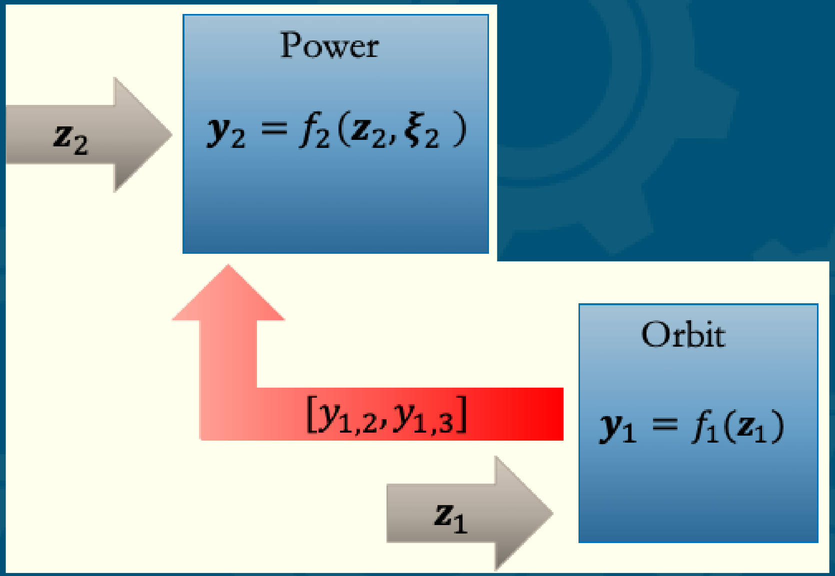
and

$$f = f_{1,\mathcal{I}_1} \circ f_{2,\mathcal{I}_2} \circ \cdots \circ f_{K,\mathcal{I}_K} \quad f = f_1 \circ f_2 \circ \cdots \circ f_K$$

then

$$\sup_{z \in \Gamma} \max_{q=1,\dots,Q} |f_q(z) - \hat{f}_q(z)| \leq \epsilon \frac{1 - L^K}{1 - L}$$

where $L = \max_{k=1,\dots,K} L_k$ *and* $\epsilon = \max_{k=1,\dots,K} \epsilon_k$.



Proof

$$\|f(z) - \hat{f}(z)\| = \|f_K \circ \cdots \circ f_1(z) - f_{K,\mathcal{I}_K} \circ \cdots \circ f_{1,\mathcal{I}_1}(z)\|$$

$$\leq \epsilon_K + L_K \|f_{K-1} \circ \cdots \circ f_1(z) - f_{K-1,\mathcal{I}_{K-1}} \circ \cdots \circ f_{1,\mathcal{I}_1}(z)\|$$

$$\leq \epsilon_K + \epsilon_{K-1} L_K + \epsilon_{K-2} L_K L_{K-1} + \cdots + \epsilon_1 \prod^{K} L_k$$

Multi-Index Stochastic Collocation

Approximate each component as a linear combination of surrogates built using different model fidelities.

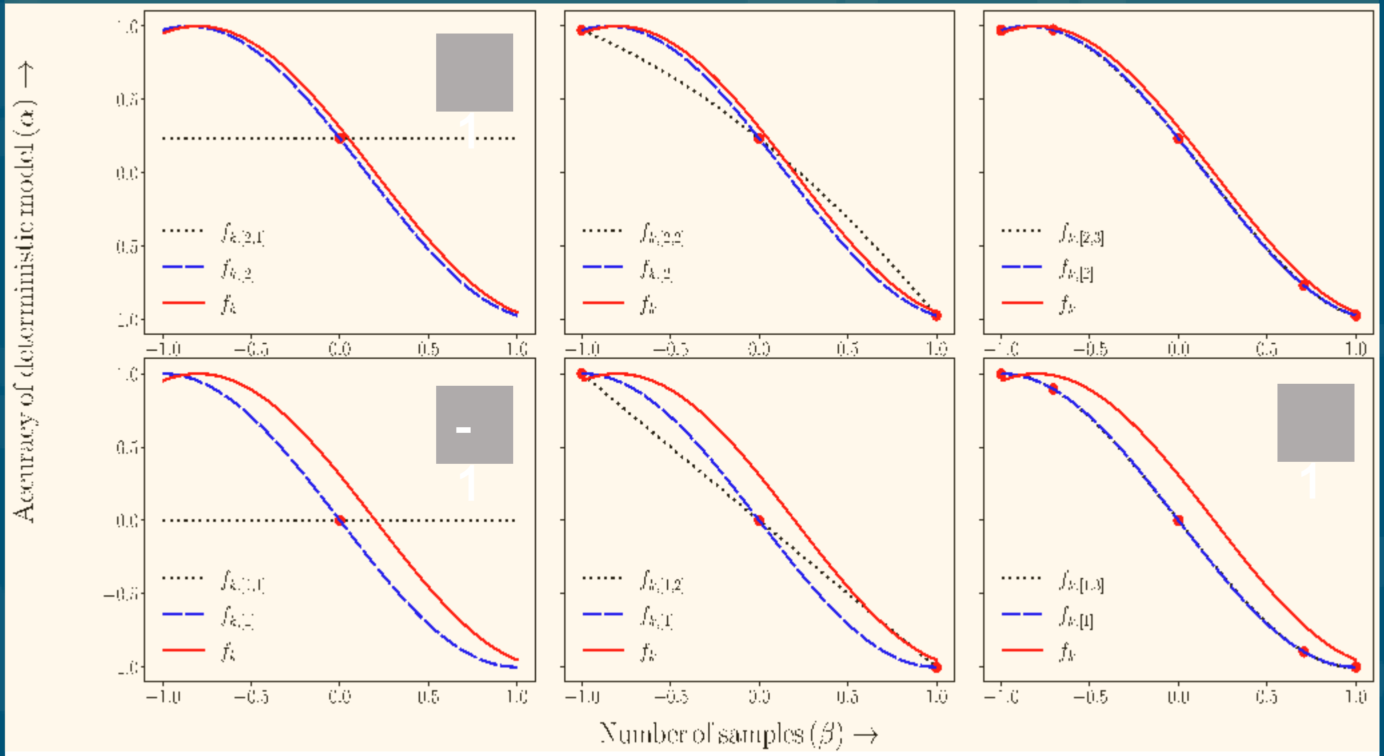
$$f(\boldsymbol{u}) \approx f_{\mathcal{I}}(\boldsymbol{u}) = \sum_{[\boldsymbol{\alpha},\boldsymbol{\beta}]\in\mathcal{I}} c_{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\, f_{[\boldsymbol{\alpha},\boldsymbol{\beta}]}(\boldsymbol{u})$$

Coefficients given by

$$c_{[\boldsymbol{\alpha},\boldsymbol{\beta}]} = \sum_{\substack{[\boldsymbol{i},\boldsymbol{j}]\in\{0,1\}^{R+N} \\ [\boldsymbol{\alpha}+\boldsymbol{i},\,\boldsymbol{\beta}+\boldsymbol{j}]\in\mathcal{I}_k}} (-1)^{\|[\boldsymbol{i},\boldsymbol{j}]\|_1}$$

We use tensor-product interpolants for each fidelity

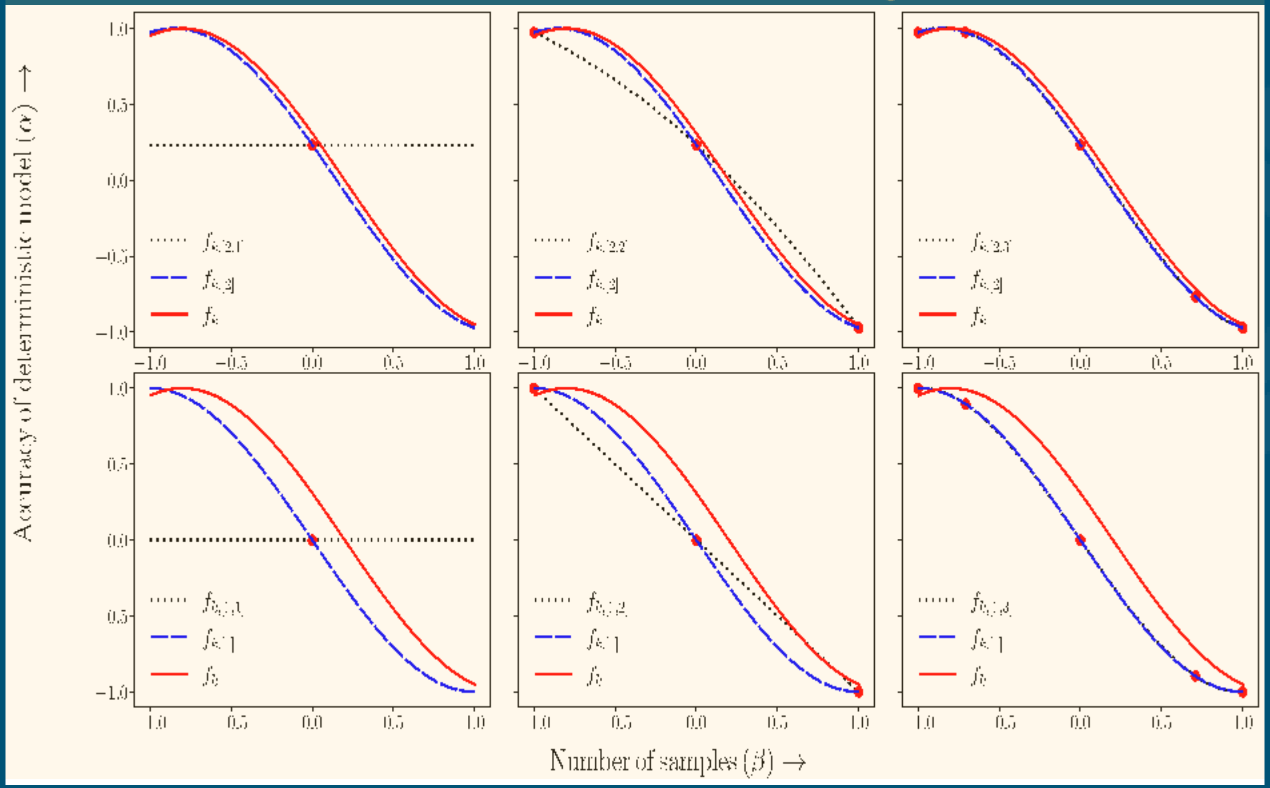
$$f_{[\boldsymbol{\alpha},\boldsymbol{\beta}]}(\boldsymbol{u}) = \sum_{\boldsymbol{j}\le m(\boldsymbol{\beta})} f_{\boldsymbol{\alpha}}(\boldsymbol{u}_{\boldsymbol{\beta}}^{(\boldsymbol{j})})\mathcal{L}_{\boldsymbol{\beta}}^{(\boldsymbol{j})}(\boldsymbol{u}).$$



Multi-Index Stochastic Collocation For A Component

Pose resource allocation as a binary knapsack problem

$$\max \sum_{[\alpha,\beta]\in\mathbb{N}_0^{n_\alpha+n_\beta}} \Delta E_{\alpha,\beta}\,\delta_{\alpha,\beta} \;\; such\ that \sum_{[\alpha,\beta]\in\mathbb{N}_0^{n_\alpha+n_\beta}} \Delta W_{\alpha,\beta}\,\delta_{\alpha,\beta} \le W_{max}$$



$\Delta E_{\alpha,\beta}$: Change in error (estimated) for each increment

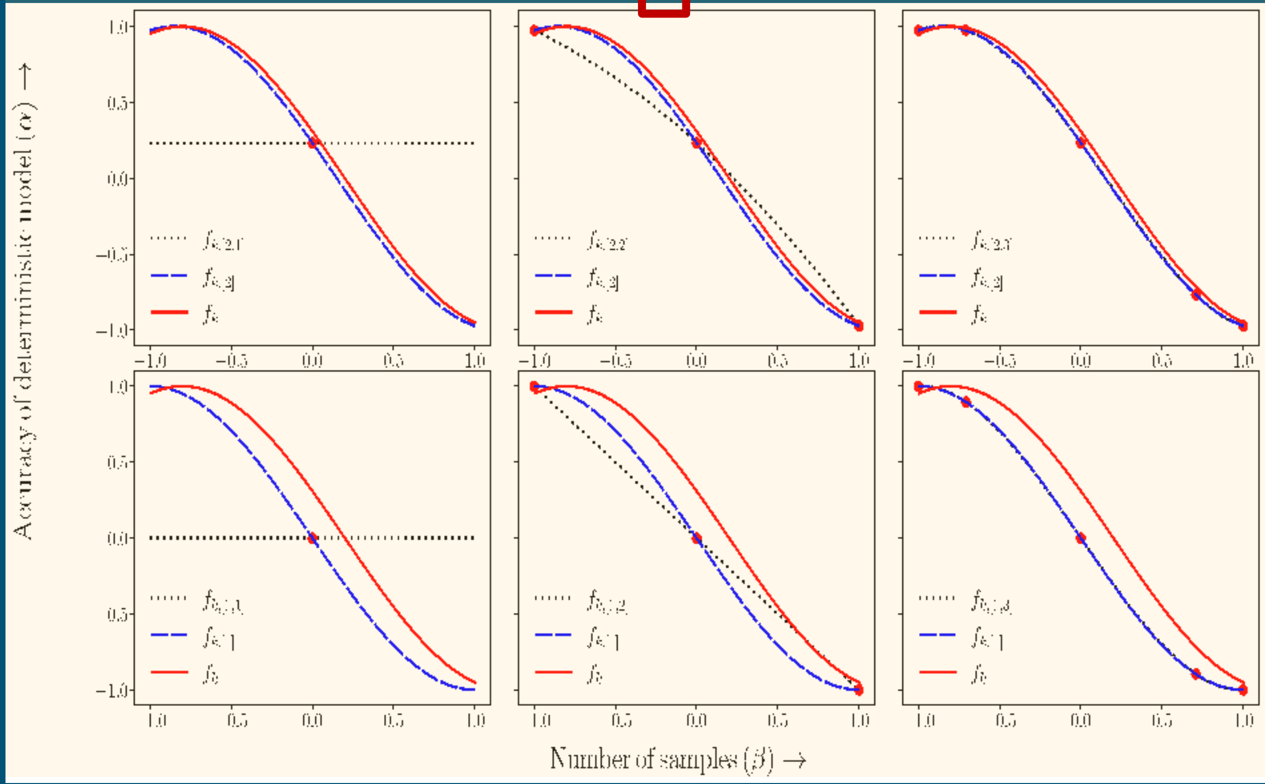
$\Delta W_{\alpha,\beta}$: Change in work for each increment

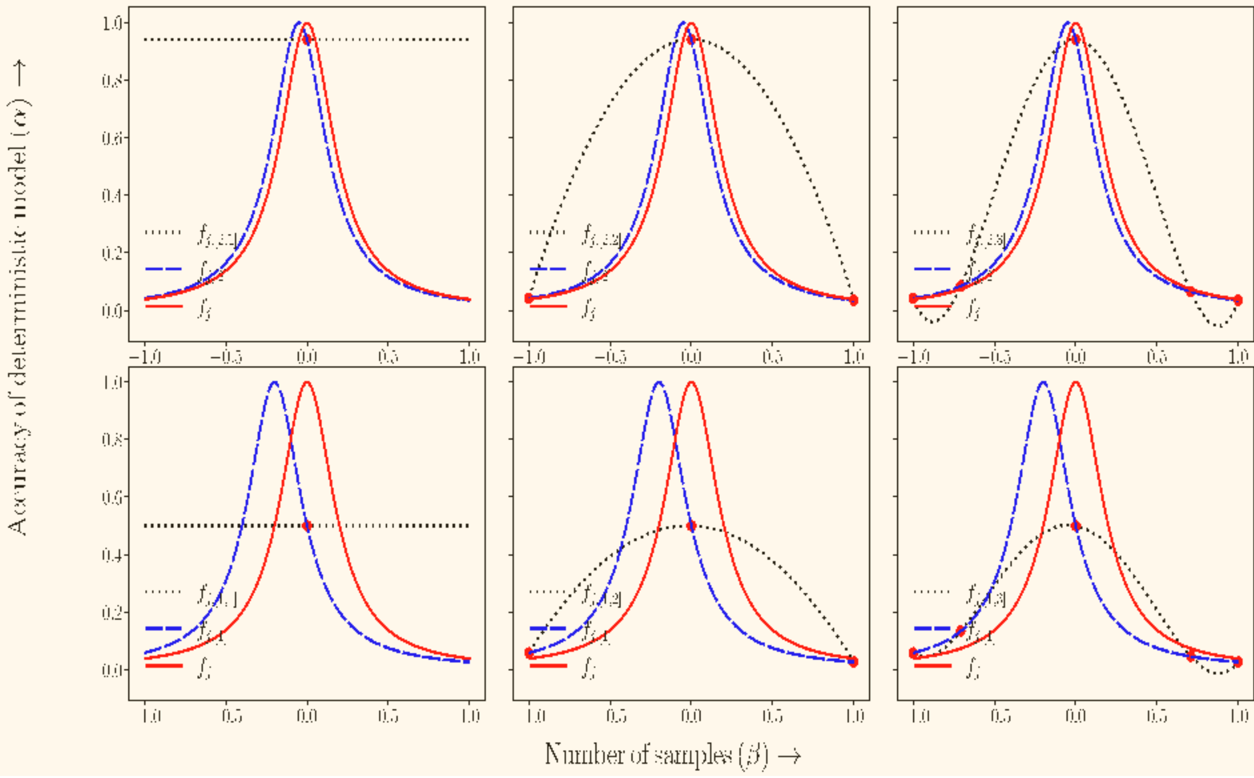
$$\gamma^k_{\alpha,\beta} = \frac{\Delta E^k_{[\alpha,\beta]}}{\Delta W^k_{[\alpha,\beta]}}$$

Multi-Index Stochastic Collocation For Systems

Must choose which component to refine based on error in system QoI

$$\max \sum_{[\alpha,\beta,k]} \Delta E_{\alpha,\beta,k}\delta_{\alpha,\beta,k} \text{ such that } \sum_{[\alpha,\beta,k]} \Delta W_{\alpha,\beta,k}\delta_{\alpha,\beta,k} \leq W_{max}$$





$\Delta E_{\alpha,\beta,k}$: Change in error (estimated) for each increment

Evaluate samples at a set of samples to compute change in system QoI

$$\gamma^k_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \frac{\Delta E^k_{[\boldsymbol{\alpha},\boldsymbol{\beta}]}}{\Delta W^k_{[\boldsymbol{\alpha},\boldsymbol{\beta}]}}$$

Estimating The Ranges Of The Coupling Variables

Interpolants built on tensor product of 1D Leja sequences

$$u^{(L+1)} = \operatorname*{argmax}_{u \in I} v(u) \prod_{l=1}^{L} |u - u^{(l)}|$$

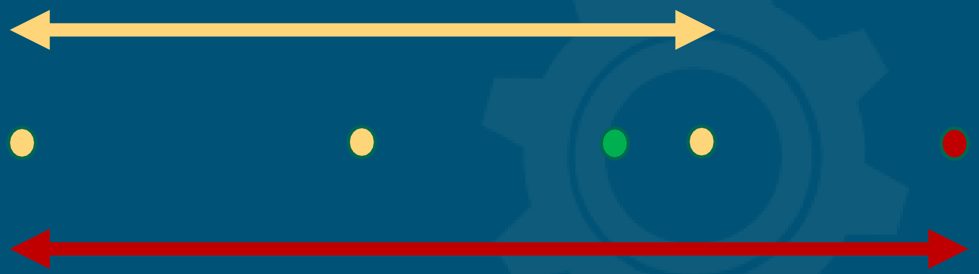
To compute error indicators we must evaluate integrated surrogates at set of random samples of $\boldsymbol{z}$

We update ranges according to

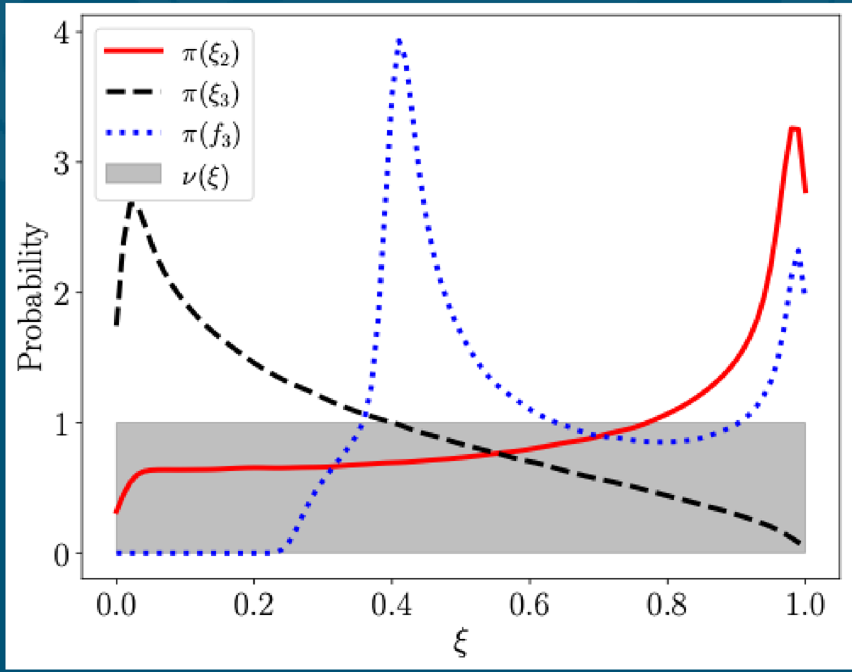
$$a_{k,i}^{t+1} = \min\left(a_{k,i}^{t}, \min_{\boldsymbol{z} \in \mathcal{U}^{\text{refine}}} f_{m,\mathcal{I}_m^{t+1},q}(\boldsymbol{z})\right)$$

$$b_{k,i}^{t+1} = \max\left(b_{k,i}^{t}, \max_{\boldsymbol{z} \in \mathcal{U}^{\text{refine}}} f_{m,\mathcal{I}_m^{t+1},q}(\boldsymbol{z})\right)$$

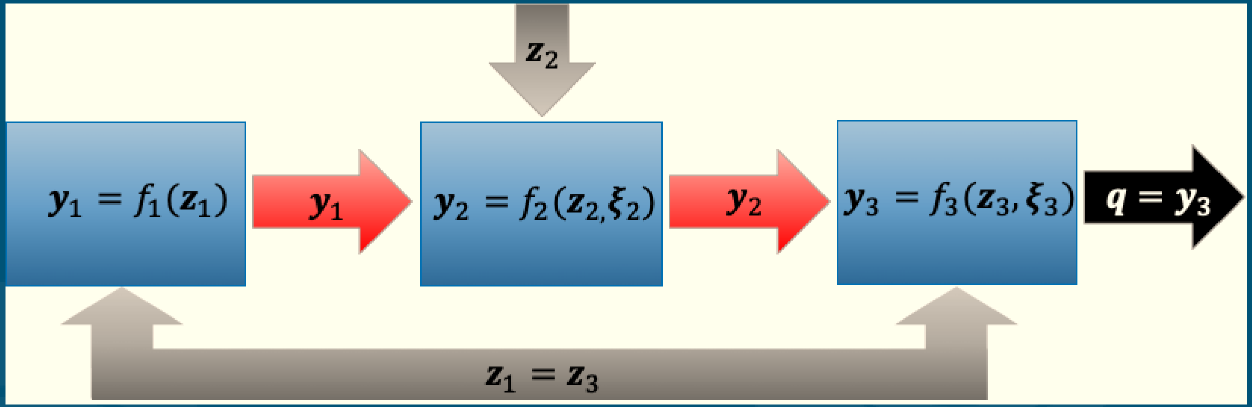
$$\Xi_{k,i}^{t} = [a_{k,i}^{t}, b_{k,i}^{t}]$$



$$\Xi_{k,i}^{t+1} = [a_{k,i}^{t+1}, b_{k,i}^{t+1}]$$



Estimating The Ranges Of The Coupling Variables

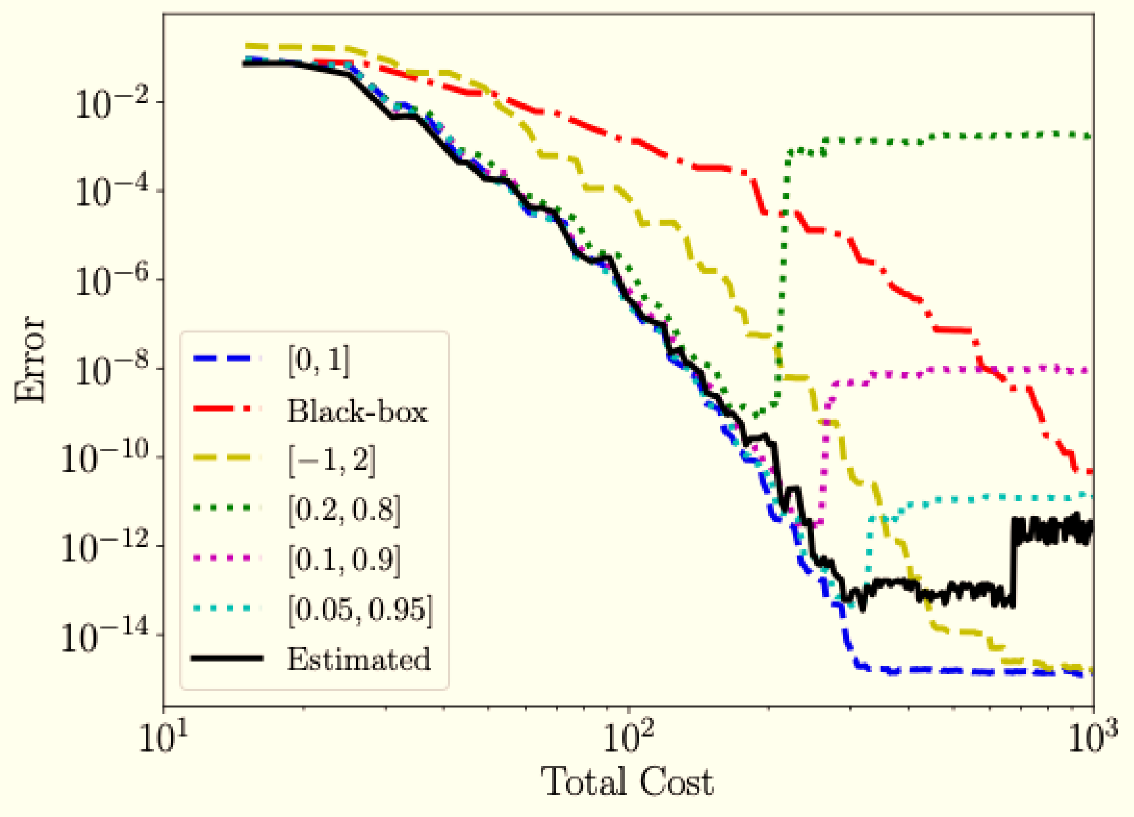


Adaptive component modeling is much more efficient than treating system as black-box

Under-estimating range introduces large errors

Over-estimation has little impact

Range can be estimated effectively

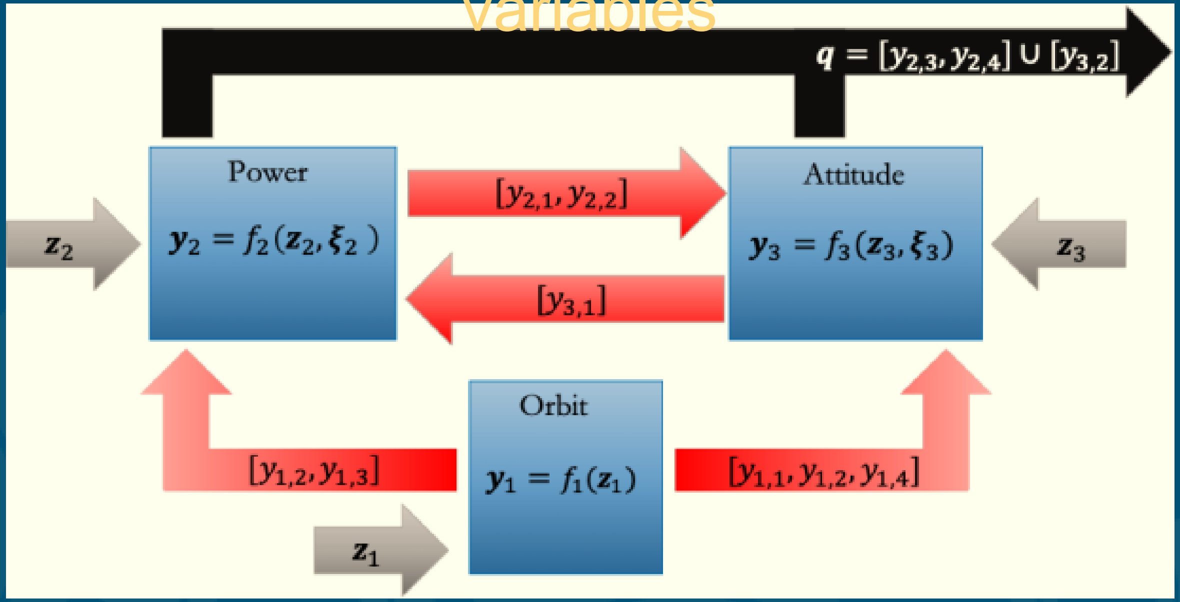


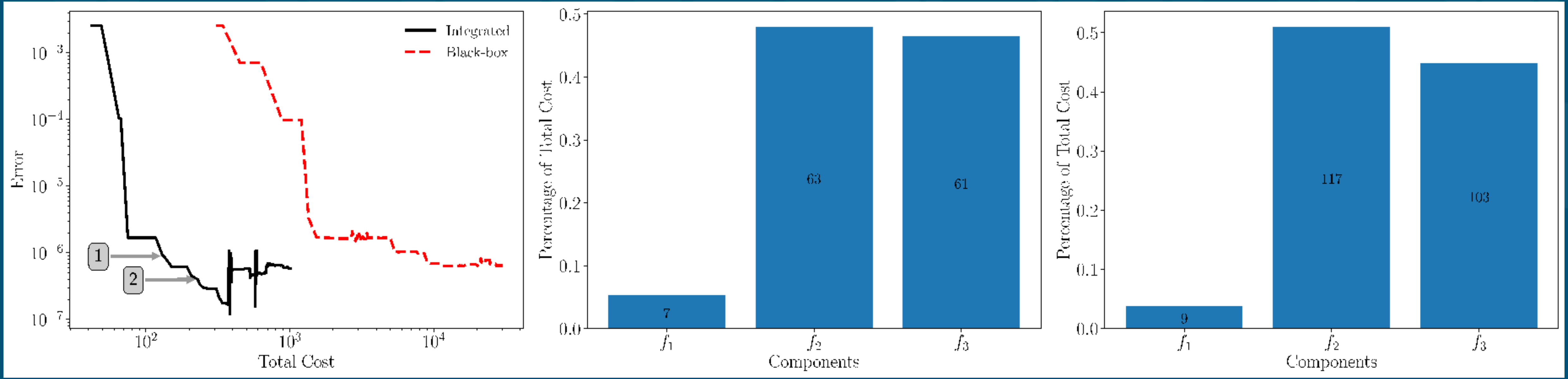
Fire-Detection Satellite

8 random variables

Approach can be used with feed-forward and feed-back coupling

Resources allocated based on sensitivity of system QoI to each component



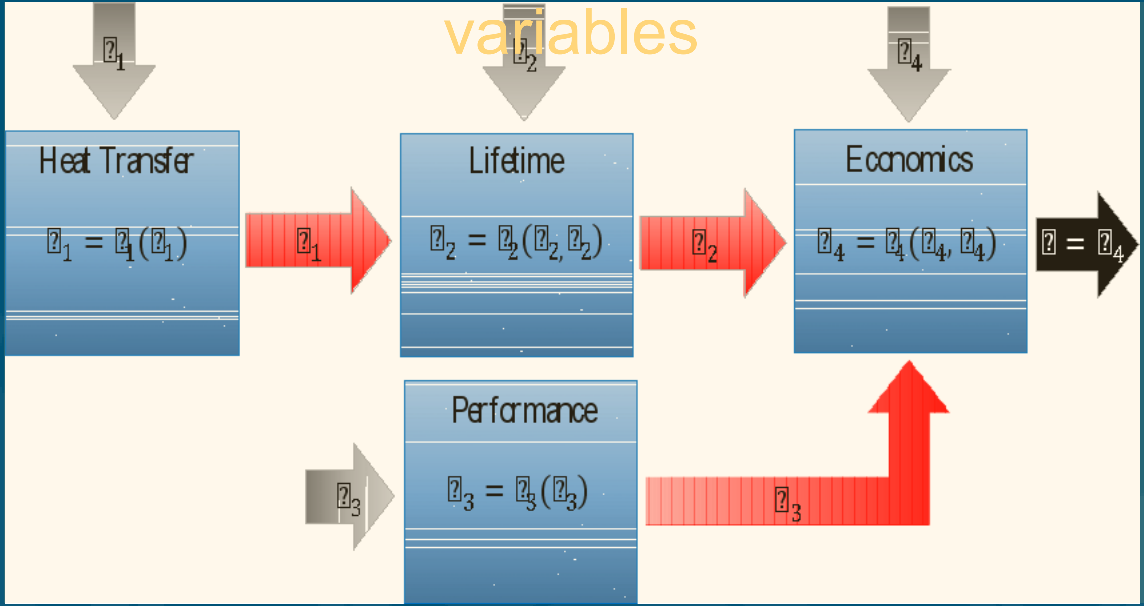


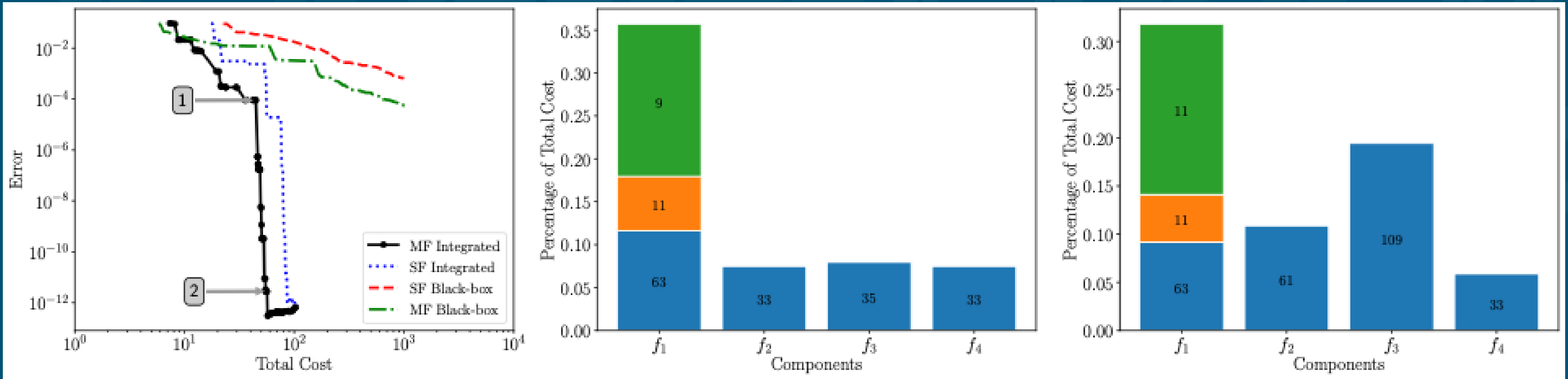
Gas-Turbine Operating Revenue

Approach can be used to when multiple fidelity models are available

Predictive accuracy of model fidelity is balanced with error in the surrogate

11 random variables





Conclusions

The structure of coupled systems should be exploited when constructing surrogates

We build surrogates of each component

The accuracy of each component surrogate depends on its predictive utility per unit cost

Our approach can reduce cost of UQ by 10-1000X