Hierarchical Sparse Adaptive Sampling in High Dimensional Design and Parametric Spaces

Justin Winokur, Olivier Le Maître, Omar Knio
Duke University, Durham, NC, USA

Abstract

We investigate two approaches to building sparse, adaptive representations of quantities of interest depending on uncertain parameters and deterministic design variables. We explore a nested approach, wherein we perform adaptive pseudo-spectral projections (aPSP) in the space of design variables and conduct, independently at a design point, the adaptation in the space of uncertain variables. We also develop an alternative approach, in which aPSP is conducted in the design-random parameter product space, and introduce a decomposition methodology to guide the refinement of sparse grids. Specifically, we use the decomposed projection surplus estimates to tune both the grid adaptation and selective termination criteria. We compare the performance for the two approaches in a simple test problem, and then examine their performance for a shock-tube design experiment involving a high-dimensional system of stiff ODEs. Computed results indicate that, whereas both methods provide effective means for tuning the quality of the representation in the deterministic and stochastic spaces, adaptive refinement in the product space is generally more efficient than the nested approach.

Background

Polynomial Chaos Expansions: Let $F(\boldsymbol{\xi})$ be the model output, parameterized by the $d-$dimensional random vector $\boldsymbol{\xi}$. Expand $F(\boldsymbol{\xi})$ in *truncated* series of an orthonormal polynomial basis

$$F(\boldsymbol{\xi}) \approx \sum_{\boldsymbol{k}\in\mathcal{K}} F_{\boldsymbol{k}}\Psi_{\boldsymbol{k}}(\boldsymbol{\xi}) \qquad \Rightarrow F_{\boldsymbol{k}} = \langle F\Psi_{\boldsymbol{k}}\rangle$$

Estimate $\langle F\Psi_{\boldsymbol{k}}\rangle$ with appropriate multi-dimensional quadrature
$\langle F\Psi_{\boldsymbol{k}}\rangle \approx \langle F\Psi_{\boldsymbol{k}}\rangle_Q = \sum_{i=1}^{N_Q} F(\boldsymbol{\xi}_i)\Psi_{\boldsymbol{k}}(\boldsymbol{\xi}_i)w_i$

Sparse Pseudo-Spectral Projection (PSP): Build projection with sparse construction to avoid exponential scaling (of fully tensorized construction) and internal aliasing (from insufficient exactness of sparse quadrature). Let $\mathrm{P}^1_\ell F$ be a 1D, level ℓ projection operator of F. Express $\mathrm{P}^1_\ell F$ in terms of a telescoping sum:

$$\mathrm{P}^1_\ell F = \sum_{l=1}^{\ell}\left(\mathrm{P}^1_l - \mathrm{P}^1_{l-1}\right)F = \sum_{l=1}^{\ell}\Delta^{\mathrm{P}}_l F, \qquad \Delta^{\mathrm{P}}_1 F \doteq \mathrm{P}^1_1 F$$

Then, the multi-dimensional, tensor-product projection is:

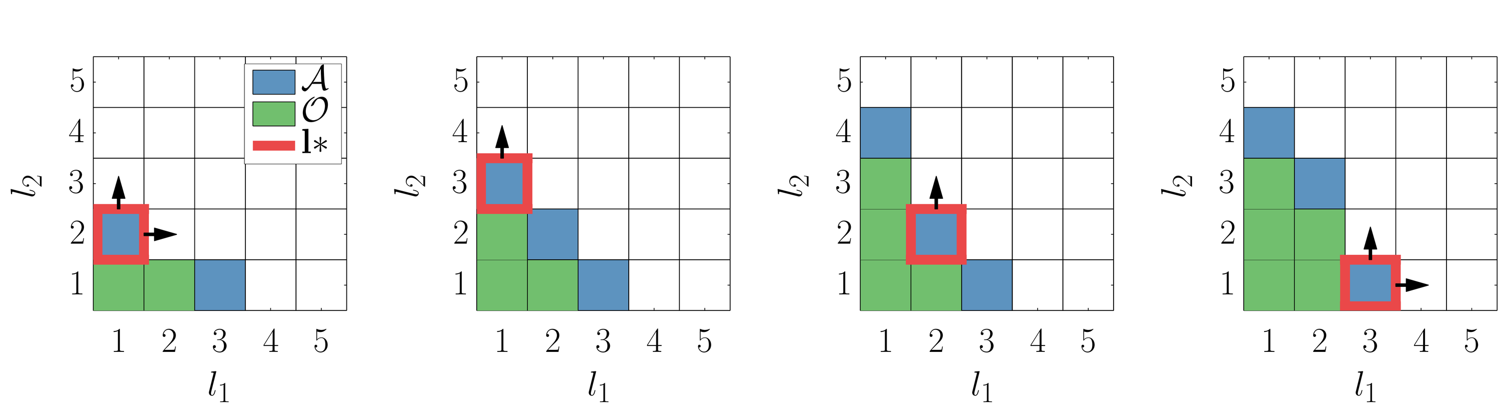
$$\mathbf{P}^d_{\boldsymbol{\ell}}F = \left(\mathrm{P}^1_{\ell_1}\otimes\ldots\otimes\mathrm{P}^1_{\ell_d}\right)F = \left(\left(\sum_{l_1=1}^{\ell_1}\Delta^{\mathrm{P}}_{l_1}\right)\otimes\ldots\otimes\left(\sum_{l_d=1}^{\ell_d}\Delta^{\mathrm{P}}_{l_d}\right)\right)F$$

$$= \sum_{l_1=1}^{\ell_1}\ldots\sum_{l_d=1}^{\ell_d}\left(\Delta^{\mathrm{P}}_{l_1}\otimes\ldots\otimes\Delta^{\mathrm{P}}_{l_d}\right)F = \sum_{\boldsymbol{l}\in\mathscr{L}}\left(\Delta^{\mathrm{P}}_{l_1}\otimes\ldots\otimes\Delta^{\mathrm{P}}_{l_d}\right)F.$$

Adaptive PSP (aPSP): $\mathscr{L}$ can be any admissible multi-index and can be built adaptively. Associated with each $\boldsymbol{l}\in\mathscr{L}$ is a projection surplus:

$$\epsilon(\boldsymbol{l}) = \frac{1}{c_0}\left|\left|\left(\Delta^{\mathrm{P}}_{l_1}\otimes\ldots\otimes\Delta^{\mathrm{P}}_{l_d}\right)F\right|\right| \qquad c_0 = \text{normalizing constant}$$

- Define two sets, Active, $\mathcal{A}$ and Old, $\mathcal{O}$, where $\mathscr{L} = \mathcal{A}\cup\mathcal{O}$
- while $\boxed{\eta^2 \doteq \sum_{\boldsymbol{l}\in\mathcal{A}}\epsilon(\boldsymbol{l})^2} \leq \texttt{Tol}$, Choose $\boldsymbol{l}^*\in\mathcal{A}$ with highest $\epsilon(\boldsymbol{l})$
  - $\mathcal{O} := \mathcal{O}\cup\boldsymbol{l}^*$
  - $\mathcal{A} := \mathcal{A}\setminus\boldsymbol{l}^*$
  - Add all $\mathcal{O}$-admissible forward neighbors of $\boldsymbol{l}^*$ to $\mathcal{A}$



Approach 1: Nested Hierarchical Adaptation

Nested Adaptation performs local "inner-space" $(\boldsymbol{u})$ adaptation at each realization of the "outer-space" $(\boldsymbol{p})$ adaptation on $F(\boldsymbol{p},\boldsymbol{u})$

- Allows for independent tolerances in both $\boldsymbol{p}$ and $\boldsymbol{u}$ spaces
  - Good control of $\boldsymbol{p}$ space realizations
- Enforces full tensorization of design variable and uncertain parameter spaces
  - Incurs a non-trivial cost
- Can be easily extended to hybrid approaches combining different methods for inner or outer representations

Nested adaption:

- Independently adapt in $\boldsymbol{p}$ with aPSP
- At each p_i realization of the $\boldsymbol{p}$ adaption
  - Adapt in $\boldsymbol{u}$ on $g^{(i)}(\boldsymbol{u}) = F(\boldsymbol{p},\boldsymbol{u}|\boldsymbol{p}=p_i)$
  - Independent, local $\mathcal{K}^{(i)}$.

$$g^{(i)}(\boldsymbol{u}) \approx \sum_{\boldsymbol{k}\in\mathcal{K}^{(i)}} g^{(i)}_{\boldsymbol{k}}\Psi_{\boldsymbol{k}}(\boldsymbol{u}).$$

  - Compute $\left|\left|g^{(i)}(\boldsymbol{u})\right|\right|$ as QoI for $\boldsymbol{p}$-adaptation
- After $\boldsymbol{p}$-adaptation converges, determine global $\tilde{g}_k$, Note: $\mathcal{K}\doteq\bigcup_{i=1}^{N_p}\mathcal{K}^{(i)}$

$$g^{(i)}(\boldsymbol{u}) \approx \sum_{\boldsymbol{k}\in\mathcal{K}}\tilde{g}^{(i)}_{\boldsymbol{k}}\Psi_{\boldsymbol{k}}(\boldsymbol{u}) \quad \tilde{g}^{(i)}_{\boldsymbol{k}} = \begin{cases} g^{(i)}_{\boldsymbol{k}} & \boldsymbol{k}\in\mathcal{K}^{(i)},\\ 0 & \text{otherwise.}\end{cases}$$

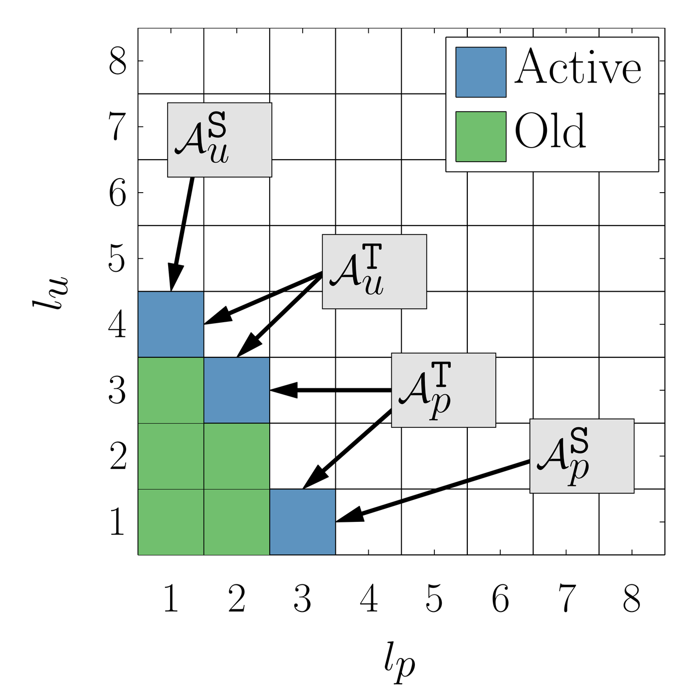
- Project the $\boldsymbol{p}$ dependent $g^{(i)}(\boldsymbol{u})$ in $\boldsymbol{p}$ space for a global representation of $F(\boldsymbol{p},\boldsymbol{u})$

Approach 2: Sensitivity Tuned Adaptation in Product Space

Consider product space of $\boldsymbol{\xi} = (\boldsymbol{p},\boldsymbol{u})$ and perform aPSP in that product space.
- Exploits sparsity between the two spaces
- Desirable to be able to asses and tune adaptation along different directions

Tuning adaptation:
Apply a decomposition to $\mathcal{A}$ where $\mathcal{A}^{\mathrm{S}}_{\boldsymbol{p}}$ is indices only along the $\boldsymbol{p}$ axis and $\mathcal{A}^{\mathrm{T}}_{\boldsymbol{p}}$ includes mixed indices.
Define: ($\bullet = \boldsymbol{p}$ or $\boldsymbol{u}$, direction of interest)

$$\left(\eta^{\mathrm{S}}_\bullet\right)^2 \doteq \sum_{\boldsymbol{l}\in\mathcal{A}^S_\bullet}\epsilon(\boldsymbol{l})^2 \qquad \left(\eta^{\mathrm{T}}_\bullet\right)^2 \doteq \sum_{\boldsymbol{l}\in\mathcal{A}^T_\bullet}\epsilon(\boldsymbol{l})^2$$

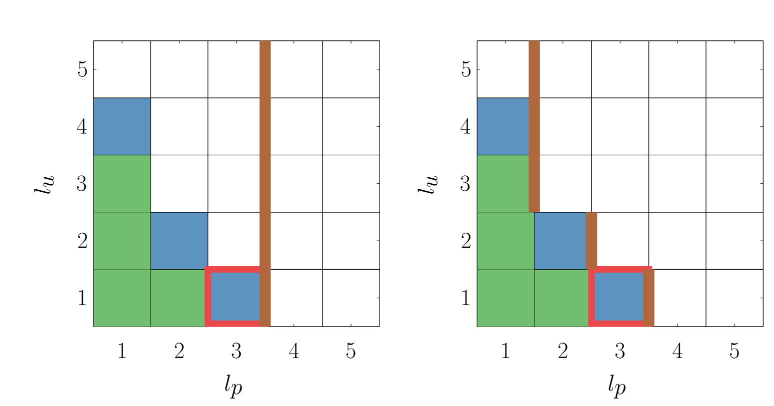


For simplicity, set $\eta_\bullet$ to be a combination of $\eta^{\mathrm{S}}_\bullet$ and $\eta^{\mathrm{T}}_\bullet$.
Adapt until *either* of the following is met

1: $\eta_{\boldsymbol{p}} \leq \texttt{Tol}_{\boldsymbol{p}}$ *and* $\eta_{\boldsymbol{u}} \leq \texttt{Tol}_{\boldsymbol{u}}$
2: $\eta \leq \texttt{Tol}$

If the $\eta_{\boldsymbol{p}}$ or $\eta_{\boldsymbol{u}}$ fall below the respective tolerance, halt adaptation in one of two ways:

T1: Do not allow further refinement beyond the highest level reached in the converged direction but allow for additional mixed terms.
T2: Do not admit any forward neighbors in the converged direction, and consequently restrict the inclusion of many mixed terms.



T1 T2

Monte Carlo Error Estimate

For model F and surrogate (PCE) $\widetilde{F}$, set $\chi(\boldsymbol{\xi}) \doteq F(\boldsymbol{\xi}) - \widetilde{F}(\boldsymbol{\xi})$ and define *a posteriori* error $\zeta^2 = \frac{\mathbb{E}\{\chi^2\}}{\mathbb{E}\{F^2\}}$ and for each space

$$\left(\zeta^{\mathrm{S}}_\bullet\right)^2 = \frac{\mathbb{E}\left\{\mathbb{E}\left\{\chi|\boldsymbol{\xi}_\bullet\right\}^2\right\}}{\mathbb{E}\left\{F^2\right\}} \Rightarrow \frac{\mathbb{V}\left\{\chi\right\}\mathrm{S}^\chi_\bullet + \mathbb{E}\left\{\chi\right\}^2}{\mathbb{E}\left\{F^2\right\}}, \qquad \left(\zeta^{\mathrm{T}}_\bullet\right)^2 = \frac{\mathbb{V}\left\{\chi\right\}\mathrm{T}^\chi_\bullet + \mathbb{E}\left\{\chi\right\}^2}{\mathbb{E}\left\{F^2\right\}}$$

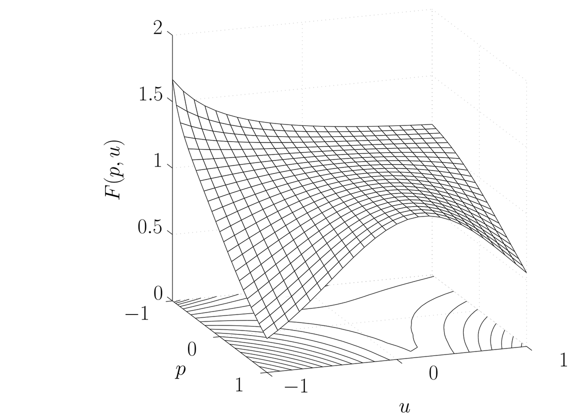
where $\mathrm{S}^\chi_\bullet$ and $\mathrm{T}^\chi_\bullet$ are the first and total sensitivity indices of χ, $\mathbb{E}$ and $\mathbb{V}$ denote the expectation and variance, respectively
Let $\zeta_\bullet$ be a combination of $\zeta^{\mathrm{S}}_\bullet$ and $\zeta^{\mathrm{T}}_\bullet$. Estimate quantities with Monte-Carlo sampling.

Work supported by the US Department of Energy, Office of Advanced Scientific Computing Research, Award DE-SC0008789

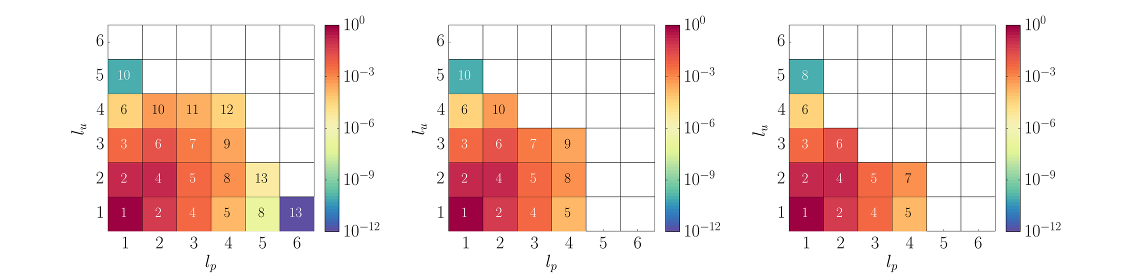
Low-Dimensional Test Problem

Consider a simple test problem with a single design variable and stochastic parameter:

$$F(\boldsymbol{p},\boldsymbol{u}) = \left(1 + \frac{1/3}{2p+u+7/2}\right) \times \exp\left[-\left(\frac{1}{2}\left(u-\frac{1}{5}\right)(p+1)\right)^2\right]$$

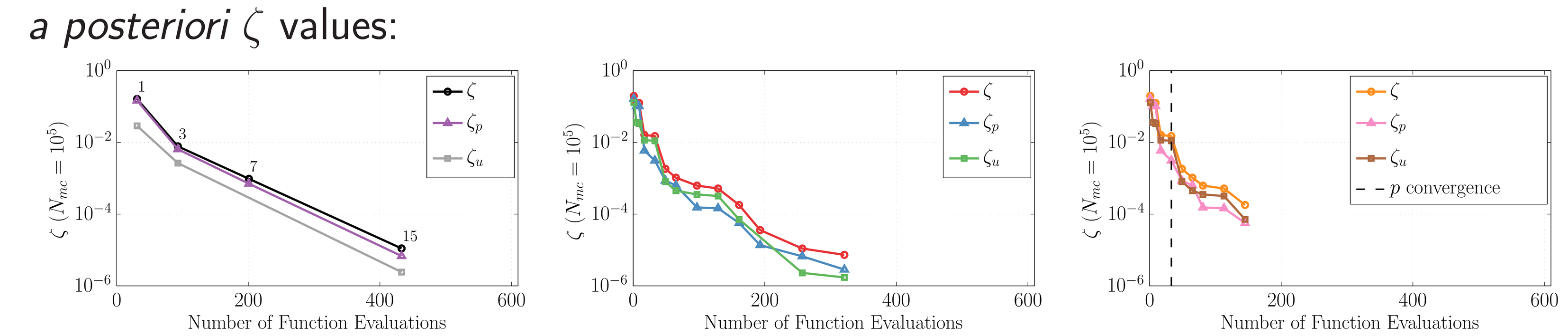


Final adaptive path and $\log_{10}|\epsilon(\boldsymbol{l})| \quad \forall \quad \boldsymbol{l}\in\mathscr{L}$ for the different methods:



No termination T1 termination T2 termination

a posteriori ζ values:



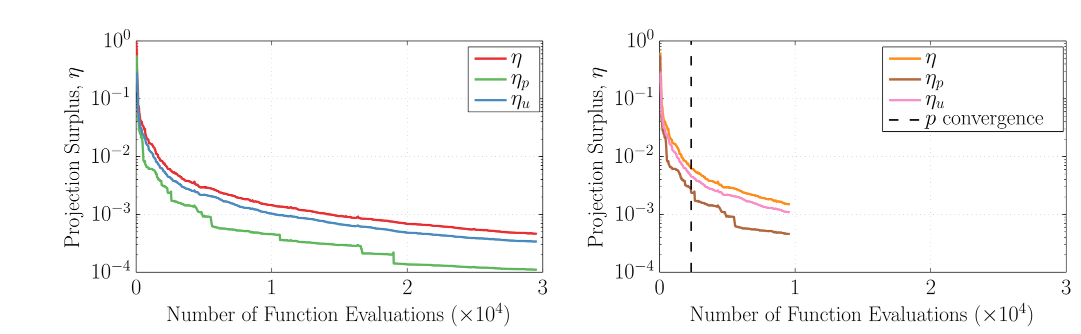
Nested Product Space (no termination) Product Space (T1 termination)

Nested consistently reduced error in inner space. Both product-space methods out-performed the nested on realization count efficiency. Untuned product-space adaptation reduced error but required more total realizations and had the *same* monomials in $\boldsymbol{u}$.

Large Scale Demonstration: Methane Combustion

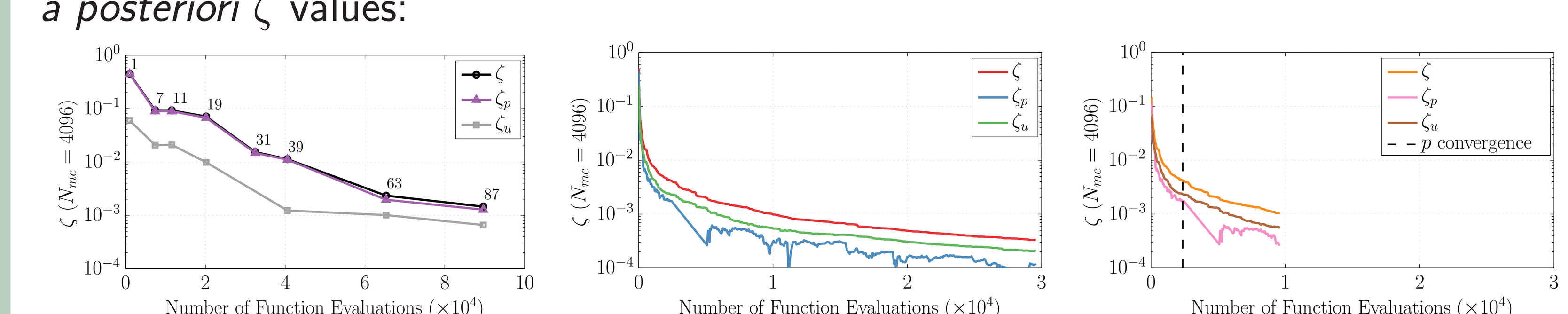
Consider a shock-tube methane combustion experiment with 22-stochastic parameters and 3 design variables. Measure peak electron concentration

Termination indicator, η and directional $\eta_{\boldsymbol{p}}$ and $\eta_{\boldsymbol{u}}$ for the product-space adaptation:



No Termination T1 termination

a posteriori ζ values:



Nested (note change of axis limit) Product Space (no termination) Product Space (T1 termination)

Nested adaption had good error properties but *extremely* expensive. In this example, the guided adaptation had little affect on the adaptivity but provided improved termination control. Note: small MC sample sized caused artifacts for some measures of ζ_p and ζ_u

Conclusions

- Nested adaptation provides good error control but the tensor-product nature of $\boldsymbol{p}-\boldsymbol{u}$ coupling makes it expensive
- Tuned product-space adaptation is a simple modification to aPSP providing enhanced adaptivity and termination criteria
  - Prevents adaptivity from ignoring important but low-variance directions
  - Allows for a sparse coupling of $\boldsymbol{p}$ and $\boldsymbol{u}$ dependence
  - Easily adaptable to more than two groupings of directions
  - Different directional termination methods provide greater control of mixed terms
- Directional η values are also a useful *diagnostic* quantity

Department of Mechanical Engineering & Materials Science, Duke University, Durham, NC, USA
Mail: Justin.Winokur@duke.edu
http://www.mems.duke.edu/