Computationally Efficient Para SAND2019-1694C
Analysis for PDE-Constrained Optimization

Joseph Hart†
with Bart van Bloemen Waanders†

† Sandia National Laboratories[1]
Center for Computing Research
Optimization and UQ Group

SIAM CS&E
Spokane, Wa
February 26, 2019

[1] Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

1. Sensitivity Analysis for PDE-Constrained Optimization

2. Method and Computation
   - Overview of the Method
   - Computational Considerations

3. Numerical Results

Problem Formulation

Consider the PDE-constrained optimization problem

$$\min_{u,z} J(u, z) \tag{1}$$
$$s.t. \quad c(u, z, \boldsymbol{\theta}) = 0$$

where

- J is an objective function
- c is a partial differential equation (PDE)
- u is a (infinite dimensional) state variable
- z is a (possibly infinite dimensional) design or control variable
- $\boldsymbol{\theta}$ corresponding to (possibly infinite dimensional) uncertain parameters in c

Goal: Determine the sensitivity of the **solution** of (1) to changes in $\boldsymbol{\theta}$.
This is the sensitivity of (1), not simply the sensitivity of J or c.

Motivation

$$\min_{u,z} J(u,z) \tag{1}$$
$$s.t. \quad c(u,z,\boldsymbol{\theta}) = 0$$

Our ultimate goal is to produce a robust or risk-averse optimal design or control. This requires

- engineering design to mitigate uncertainties
- data acquisition and inverse problems to characterize or mitigate uncertainties
- a robust or risk-averse formulation of (1)

This is challenging when the model

- is large scale
- is multi-physics
- is multi-scale
- contains many different uncertain parameters

Goal: Use sensitivity analysis to screen and prioritize parameters.

1. Sensitivity Analysis for PDE-Constrained Optimization

2. Method and Computation
   - Overview of the Method
   - Computational Considerations

3. Numerical Results

Optimization Problem

$$\min_{u,z} J(u,z) \qquad (1)$$
$$s.t. \quad c(u,z,\boldsymbol{\theta}) = 0$$

Let

$$\mathcal{L}(u,z,\lambda,\boldsymbol{\theta}) = J(u,z) + \langle c(u,z,\boldsymbol{\theta}), \lambda \rangle$$

denote the Lagrangian for (1) and assume that $(u^\star, z^\star, \lambda^\star)$ is a local minimum of (1) with $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Under mind assumptions [2], there exists a function

$$f : \mathcal{N}(\boldsymbol{\theta}_0) \to \mathcal{N}(u^\star, z^\star, \lambda^\star)$$

such that

$$\nabla \mathcal{L}(f(\boldsymbol{\theta}), \boldsymbol{\theta}) = 0$$

for any $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}_0$.

---

[2] K. Brandes and R. Griesse, Quantitative stability analysis of optimal solutions in PDE-constrained optimization.

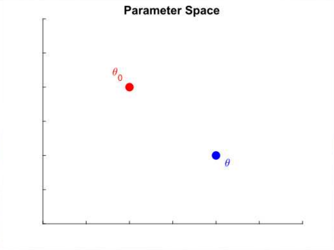
Parameter to Optimal Solution Mapping

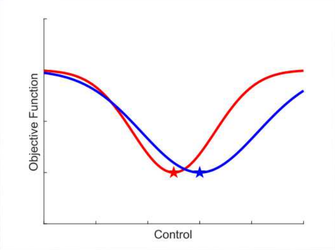
$$f : \mathcal{N}(\boldsymbol{\theta}_0) \to \mathcal{N}(u^\star, z^\star, \lambda^\star)$$

such that

$$\nabla \mathcal{L}(f(\boldsymbol{\theta}), \boldsymbol{\theta}) = 0$$

$$\boldsymbol{\theta} \qquad \mapsto \qquad f(\boldsymbol{\theta})$$





Parameters $\mapsto$ Optimal Control Solution

Local Sensitivity

$$\min_{u,z} J(u,z) \tag{1}$$
$$s.t. \quad c(u,z,\boldsymbol{\theta}) = 0$$

$$f : \mathcal{N}(\boldsymbol{\theta}_0) \to \mathcal{N}(u^\star, z^\star, \lambda^\star)$$

- $f(\boldsymbol{\theta})$ is the solution of (1) with parameters $\boldsymbol{\theta}$
  - it depends on the particular local minimum $(u^\star, z^\star, \lambda^\star)$
  - evaluating $f(\boldsymbol{\theta})$ requires the solving (1) (expensive)
- following the work of Roland Griesse (now Hertzog), the Fréchet derivative of f with respect to $\boldsymbol{\theta}$ is given by

$$\mathcal{D}f(\boldsymbol{\theta}) = -\mathcal{K}^{-1}\mathcal{B}$$

where $\mathcal{K}$ is the hessian of $\mathcal{L}$ and $\mathcal{B}$ is the Fréchet derivative of $\nabla\mathcal{L}$ with respect to $\boldsymbol{\theta}$ (both evaluated at $(u^\star, z^\star, \lambda^\star)$)

Local Sensitivity

$$\mathcal{D}f(\boldsymbol{\theta}) = -\mathcal{K}^{-1}\mathcal{B}$$

- $||\mathcal{D}f(\boldsymbol{\theta})\psi||$ measures the local sensitivity of the optimal solution $(u^\star, z^\star, \lambda^\star)$ to perturbations of $\boldsymbol{\theta}$ in the direction ψ
- evaluating $\mathcal{D}f(\boldsymbol{\theta})\psi$ requires solving a large linear system (whose coefficient matrix is the hessian of $\mathcal{L}$)
- define the projection operator Π by $\Pi(u, z, \lambda) = z$
- discretize $\boldsymbol{\theta}$ in a finite dimensional subspace $V = \text{span}\{\psi_1, \psi_2, \dots, \psi_p\}$
- the local sensitivity of the design/control solution $z^\star$ with respect to $\boldsymbol{\theta}$ in the direction ψ_i is given by

$$L_i = ||\Pi\mathcal{K}^{-1}\mathcal{B}\psi_i|| \qquad i = 1, 2, \dots, p$$

Local Sensitivity: Computation

$$L_i = ||\Pi \mathcal{K}^{-1} \mathcal{B} \psi_i|| \qquad i = 1, 2, \dots, p$$

- requires solving the PDE-constrained optimization problem once, followed by p linear system solves with coefficient matrix $\mathcal{K}$
- p is large when $V = \text{span}\{\psi_1, \psi_2, \dots, \psi_p\}$ discretizes a spatially dependent $\boldsymbol{\theta}$
- in many applications $\mathcal{B}$ possesses a low rank structure (because it encodes parametric sensitivities), leveraging the singular value decomposition yields

$$L_i = ||\Pi \mathcal{K}^{-1} \mathcal{B} \psi_i|| = \left|\left| \sum_{k=1}^{\infty} \sigma_k (v_k, \psi_i) u_k \right|\right| \approx \sqrt{\sum_{k=1}^{m} \sigma_k^2 (v_k, \psi_i)^2}$$

  where (σ_k, u_k, v_k), $k = 1, 2, \dots$ are the singular triplets of $\Pi \mathcal{K}^{-1} \mathcal{B}$
- the singular triplets (σ_k, u_k, v_k), $k = 1, 2, \dots, m$, may frequently be estimated with less than p applications of $\Pi \mathcal{K}^{-1} \mathcal{B}$
- estimating the truncated SVD of $\Pi \mathcal{K}^{-1} \mathcal{B}$ allows efficient estimation of L_i, $i = 1, 2, \dots, p$

Extension to Global Sensitivity Indices

$$\min_{u,z} J(u,z) \quad s.t. \quad c(u,z,\boldsymbol{\theta}) = 0 \tag{1}$$

Local sensitivity index

$$L_i = ||\Pi \mathcal{K}^{-1} \mathcal{B} \psi_i|| \qquad i = 1, 2, \ldots, p$$

- L_i depends on $\boldsymbol{\theta}_0$ and $u^\star, z^\star, \lambda^\star$
- sample different $\boldsymbol{\theta}_0$'s to explore the parameter space
- sample different initial optimization iterates to explore the various local minima
- estimate L_i, $i = 1, 2, \ldots, p$, for each parameter and initial iterate sample
- need to compute L_i's efficiently

1. Sensitivity Analysis for PDE-Constrained Optimization

2. Method and Computation
   - Overview of the Method
   - **Computational Considerations**

3. Numerical Results

Singular Value Decomposition (SVD)

- The computational challenge is computing the truncated SVD (with appropriate inner products) of

$$\Pi \mathcal{K}^{-1} \mathcal{B} \tag{2}$$

- Let M_θ and M_z denote the mass matrices (defining discretized functional space inner products using coordinate representations)
- The SVD of the discretized operator (2) may be extracted from the generalized eigenvalue problem

$$Ax = \lambda Mx$$

where

$$A = \begin{pmatrix} 0 & M_z \Pi \mathcal{K}^{-1} \mathcal{B} \\ \mathcal{B}^T \mathcal{K}^{-1} \Pi M_z & 0 \end{pmatrix}$$

and

$$M = \begin{pmatrix} M_z & 0 \\ 0 & M_\theta \end{pmatrix}$$

Generalized Eigenvalue Problem

$$\begin{pmatrix} 0 & M_z \Pi \mathcal{K}^{-1} \mathcal{B} \\ \mathcal{B}^T \mathcal{K}^{-1} \Pi M_z & 0 \end{pmatrix} \begin{pmatrix} \tilde{z} \\ \tilde{\boldsymbol{\theta}} \end{pmatrix} = \lambda \begin{pmatrix} M_z & 0 \\ 0 & M_{\boldsymbol{\theta}} \end{pmatrix} \begin{pmatrix} \tilde{z} \\ \tilde{\boldsymbol{\theta}} \end{pmatrix}$$

- computational cost dominated by applying $\mathcal{K}^{-1}$ (a large linear system solve)
- use conjugate gradient since $\mathcal{K}$ is symmetric positive definite
- use randomized generalized eigenvalue solver from [3] to facilitate parallelism
  - $\mathcal{K}^{-1}$ must be applied sequentially if an iterative eigenvalue solver is used
  - the randomized solver allows them to be executed in parallel since the matrix vector products are independent
- postprocess dominant eigenvalues/vectors to estimate the truncated SVD (and hence local sensitivities)

[3] A.K. Saibaba, J. Lee,and P.K. Kitanidis, Randomized algorithms for generalized Hermitian eigenvalue problems with application to computing Karhunen-Loève expansion

Computational Cost Outline

$$\min_{u,z} J(u,z) \tag{1}$$
$$s.t. \quad c(u,z,\boldsymbol{\theta}) = 0$$

Cost to compute one local sensitivity (assuming all eigenvalue solver matrix vector products are parallelized)

1. solve (1) once
2. solve 4 large linear systems (whose coefficient matrix is the hessian of the Lagrangian of (1) evaluated at local minimum) for each matrix vector product required by the eigenvalue solver
3. many other inexpensive matrix vector products

Chemical Vapor Deposition Reactor

$$\min_{u,z} \frac{1}{2}\int_\Omega (\nabla \times v)dx + \frac{\gamma}{2}\int_{\Gamma_c} z^2 dx$$

$s.t.$

$$u = (v_1, v_2, p, T)$$

$$-\epsilon(p)\nabla^2 v + (v\cdot\nabla)v + \nabla p + \eta(\boldsymbol{\theta})Tg = 0 \qquad \text{in } \Omega$$

$$\nabla \cdot v = 0 \qquad \text{in } \Omega$$

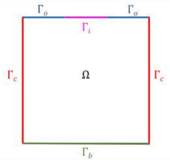
$$-\kappa(\boldsymbol{\theta})\Delta T + v\cdot\nabla T = 0 \qquad \text{in } \Omega$$

$$T = 0 \quad \text{and} \quad v = v_i \qquad \text{on } \Gamma_i$$

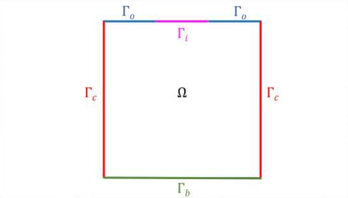
$$\kappa(\boldsymbol{\theta})\frac{\partial T}{\partial n} = 0 \quad \text{and} \quad v = v_o \qquad \text{on } \Gamma_o$$

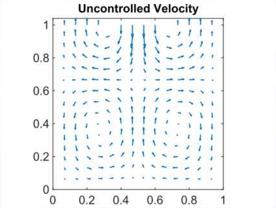
$$T = T_b(\boldsymbol{\theta}) \quad \text{and} \quad v = 0 \qquad \text{on } \Gamma_b$$

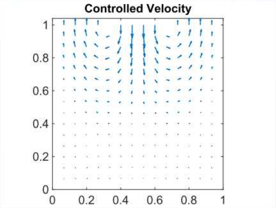
$$\kappa(\boldsymbol{\theta})\frac{\partial T}{\partial n} + \nu(\boldsymbol{\theta})(z - T) = 0 \quad \text{and} \quad v = 0 \qquad \text{on } \Gamma_c$$



Chemical Vapor Deposition Reactor







- particles are injected into the top of a container
- the temperature on side walls is controlled to minimize vorticities in the fluid
- uncertainties enter through properties of the fluid and spatially distributed thermal boundary conditions

Chemical Vapor Deposition Reactor: Continued

- total of 153 uncertain parameters
- total of 40,000 degrees of freedom in the discretization of the state
- local sensitivities are computed for 20 different parameter samples
- the randomized eigenvalue solver uses 16 random vector sample
- 1280 processors are used to parallelize the computation with minimal communication
- implemented in the Rapid Optimization Library [4] in Trilinos

[4] D. P. Kouri, G. von Winckel, and D. Ridzal, ROL: Rapid Optimization Library.

Control Solutions

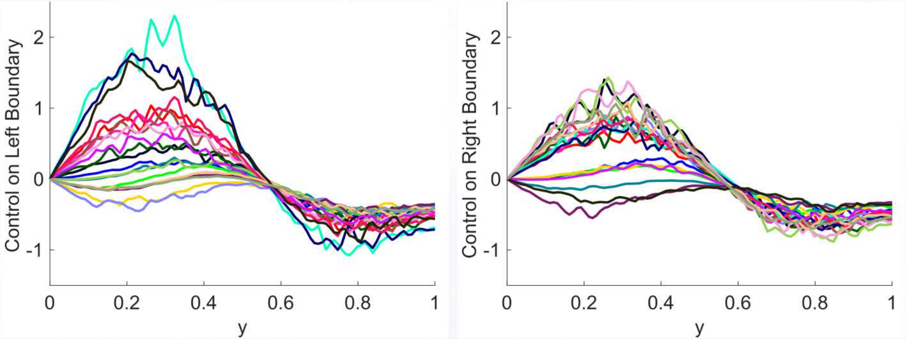


Figure: Control solutions corresponding to 20 different parameter samples. The left and right panels are the controllers on the left and right boundaries, respectively. Each curve is a control solution for a given parameter sample.

Singular Values

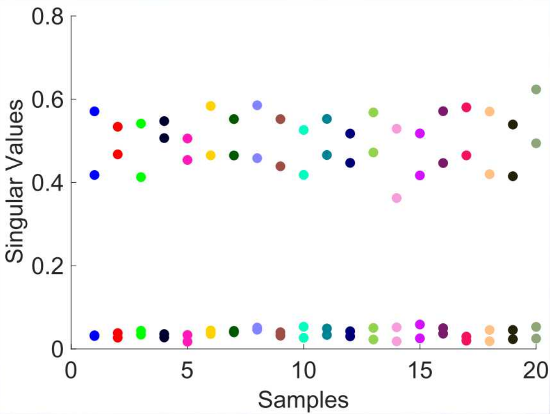


Figure: Leading 4 singular values at 20 different parameter samples. Each vertical slice corresponds to the leading 4 singular values for a fixed sample.

Local Sensitivities

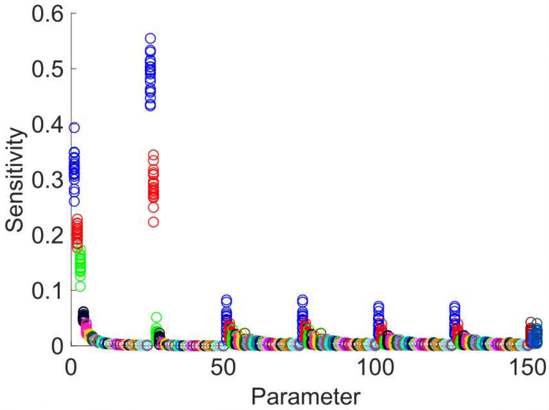


Figure: Local sensitivities for the 153 uncertain parameters. The 20 circles in each vertical slice indicates the sensitivity index for a fixed parameter as it varies over the 20 parameter samples.

Controller Singular Vectors

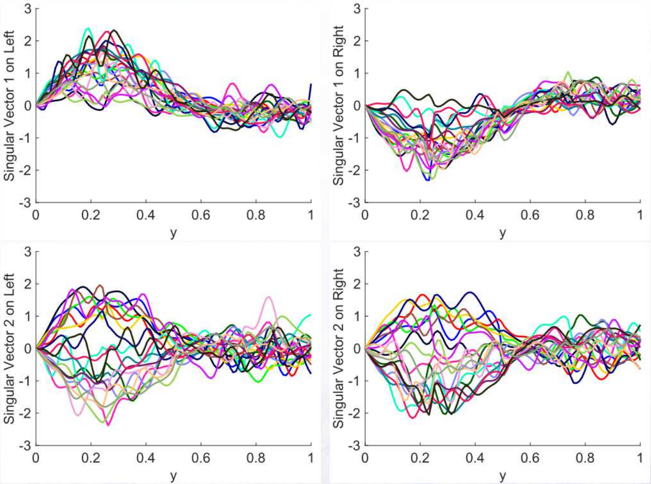


Figure: The top (bottom) row shows the first (second) singular vector on the left and right boundaries, respectively. Each curve corresponds to a different parameter sample.

Observations

- The local sensitivity analysis yields similar results for each parameter sample.
- Only around 10% of the uncertain parameters exhibit significant influence on the control strategy.
- The bottom boundary condition, T_b, has the greatest influence on the control strategy.

Summary

- interested in the sensitivity of the solution of an optimization problem to uncertain parameters
- following the work of Roland Griesse (now Hertzog), a derivative-based approach is proposed
- the Singular Value Decomposition and a randomized solver are used to enable efficient computation of local sensitivities
- the algorithm is implemented in the Rapid Optimization Library (ROL) of Trilinos which enables parallel linear algebra constructs and scalable performance for nonlinear multi-physics problems
- the method is illustrated on a nonlinear multi-physics model for a chemical vapor deposition reactor

Questions?

Joseph Hart

Sandia National Laboratories

joshart@sandia.gov