Scalable algorithms for PDE-constrained optimization under uncertainty

Omar Ghattas (ICES/Geosciences/Mech Eng, UT Austin)

Joint work with:
Peng Chen (ICES, UT Autin) and Umberto Villa (ICES, UT Austin)

Earlier work with:
Alen Alexanderian (NCSU), Noémi Petra (UC Merced), Georg Stadler (NYU)

16 April 2018
SIAM Conference on Uncertainty Quantification
Orange County, California, USA

Supported by DARPA EQUiPS, DOE MMICCS, AFOSR Comp Math

PDE-constrained optimization under uncertainty

Decision-making under uncertainty is often the ultimate goal of UQ:

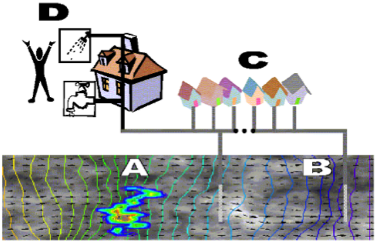
- Inverse problem: Infer uncertain model parameters given data
- Optimal experimental design problem: How should we acquire data to reduce uncertainty in inferred parameters?
- Forward problem: Propagate uncertain parameters through forward model
- Optimal design/control problem: Find design/control variables that optimize a desired uncertain objective

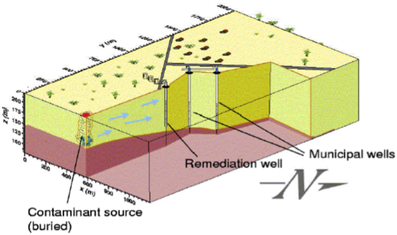
Fundamental difficulty: OUU amounts to many forward UQ problems

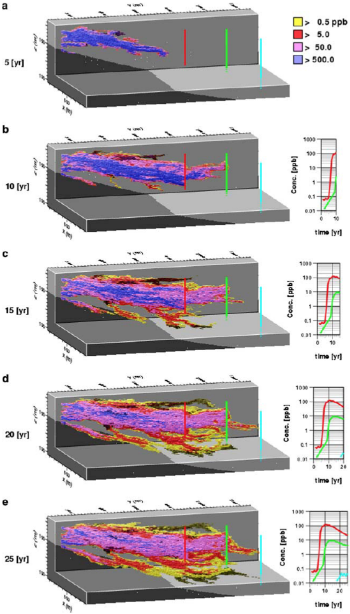
Sessions here at UQ18:

- IP8: Johannes Royset: *Good and bad uncertainty: Consequences in UQ and design*
- MT8: Drew Kouri: *Optimization and control under uncertainty*
- OED sessions: MS2/15, MS31/37, MS51/64/77
- OUU sessions: MS11/24, CP9, MS46/60, MS73, MS97/110

Example: Groundwater contaminant remediation







Source: Reed Maxwell, CSM

Classes of PDE-constrained optimization under uncertainty

- Inverse problem
  - Infer initial contaminant field from measurements of pressure at wells and from a model of subsurface flow and transport with random log permeability field
- Optimal experimental design problem
  - Where should new observation wells be placed so that initial condition is “best” inferred? (alphabetic optimality criteria, Bayes risk, expected information gain)
- Optimal design problem
  - Where should remediation wells be placed so that (uncertain) contaminant concentrations at municipal wells are minimized?
- Optimal control problem
  - What should the rates of pumping/injection at remediation wells be so that (uncertain) contaminant concentrations at municipal wells are minimized?

PDE-constrained optimal control under uncertainty

Find control/design variables z (pumping rates $q_j^z(t)$ and locations $\boldsymbol{x}_j^z$ of remediation wells) so that the expected value of the contaminant concentration c at the drinking wells is minimized:

$$\min_{z \in Z} \mathcal{J}(z) := \sum_i \int_X \int_0^T \int_\Omega w\, c(m, z)\, \delta_\varepsilon(\boldsymbol{x}_i^w)\, d\boldsymbol{x}\, dt\, d\mu(m) + \sum_j \beta_j \int_0^T (q_j^z)^2\, dt$$

where concentration c depends on random log permeability field m, control variables $q_j^z(t)$, and design variables $\boldsymbol{x}_j^z$ through the coupled groundwater flow and contaminant transport equations

$$\phi \rho c_t \frac{\partial p}{\partial t} - \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = -\sum_i q_i^w\, \delta_\varepsilon(\boldsymbol{x}_i^w) - \sum_j q_j^z\, \delta_\varepsilon(\boldsymbol{x}_j^z)$$

$$\frac{\mu \phi}{\exp(m)} \boldsymbol{v} + \boldsymbol{\nabla} p = \boldsymbol{0}$$

$$\frac{\partial (R \phi c)}{\partial t} + \boldsymbol{\nabla} \cdot (\phi c \boldsymbol{v}) - \boldsymbol{\nabla} \cdot (\phi \boldsymbol{D} \boldsymbol{\nabla} c) = -\sum_i c(\boldsymbol{x}_i^w)\, q_i^w \delta_\varepsilon(\boldsymbol{x}_i^w) - \sum_j c(\boldsymbol{x}_j^z)\, q_j^z\, \delta_\varepsilon(\boldsymbol{x}_j^z)$$

$$\boldsymbol{D} \approx (\alpha_T |\boldsymbol{v}| + \hat{D}) \boldsymbol{I} + (\alpha_L - \alpha_T) \frac{\boldsymbol{v} \otimes \boldsymbol{v}}{|\boldsymbol{v}|}$$

Mean-variance PDE-constrained optimal control

- Weak form of forward PDE model with random and control variables:

$$\text{find } u \in \mathcal{U} \text{ such that } r(u, v, m, z) = 0 \quad \forall v \in \mathcal{V}$$

  where $u \in \mathcal{U}$ is state, $v \in \mathcal{V}$ adjoint, $m \in \mathcal{M}$ random field, $z \in \mathcal{Z}$ control
- Objective function: Consider mean–variance of control functional $Q(\cdot, \cdot) : \mathcal{U} \times \mathcal{M} \to \mathbb{R}$:

$$\mathcal{J}(z) = \mathbb{E}[Q] + \beta \mathrm{Var}[Q] + \mathcal{P}(z)$$

  where $\mathcal{P}(z)$ is cost of controls (or regularization)
- Optimal control problem: find $z^* \in \mathcal{Z}$, s.t.

$$z^* = \arg\min_{z \in \mathcal{Z}} \mathcal{J}(z), \text{ subject to } r(u, v, m, z) = 0$$

- Sample average approximation (SAA) is prohibitive: entails as many (nonlinear) PDE constraints as required for accurate estimation of $\mathbb{E}[Q]$

$$z^* = \arg\min_{z \in \mathcal{Z}} \mathcal{J}^{\mathrm{MC}}(z), \text{ subject to } r(u, v, m_i, z) = 0 \quad i = 1, \ldots, M$$

$\Longrightarrow$ "Many-PDE-constrained optimization"

Mean-variance PDE-constrained optimal control

- Weak form of forward PDE model with random and control variables:

$$\text{find } u \in \mathcal{U} \text{ such that } r(u, v, m, z) = 0 \quad \forall v \in \mathcal{V}$$

  where $u \in \mathcal{U}$ is state, $v \in \mathcal{V}$ adjoint, $m \in \mathcal{M}$ random field, $z \in \mathcal{Z}$ control
- Objective function: Consider mean–variance of control functional $Q(\cdot, \cdot) : \mathcal{U} \times \mathcal{M} \to \mathbb{R}$:

$$\mathcal{J}(z) = \mathbb{E}[Q] + \beta \text{Var}[Q] + \mathcal{P}(z)$$

  where $\mathcal{P}(z)$ is cost of controls (or regularization)
- Optimal control problem: find $z^* \in \mathcal{Z}$, s.t.

$$z^* = \arg\min_{z \in \mathcal{Z}} \mathcal{J}(z), \text{ subject to } r(u, v, m, z) = 0$$

- Sample average approximation (SAA) is prohibitive: entails as many (nonlinear) PDE constraints as required for accurate estimation of $\mathbb{E}[Q]$

$$z^* = \arg\min_{z \in \mathcal{Z}} \mathcal{J}^{\text{MC}}(z), \text{ subject to } r(u, v, m_i, z) = 0 \quad i = 1, \dots, M$$

$\Longrightarrow$ "Many-PDE-constrained optimization"

Some existing approaches for PDE-constrained OUU

- Schulz & Schillings, *Problem formulations and treatment of uncertainties in aerodynamic design*, AIAA J, 2009.
- Borzì & von Winckel, *Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients*, SISC, 2009.
- Borzì, Schillings, & von Winckel, *On the treatment of distributed uncertainties in PDE-constrained optimization*, GAMM-Mitt. 2010.
- Borzì & von Winckel, *A POD framework to determine robust controls in PDE optimization*, Computing and Visualization in Science, 2011.
- Gunzburger & Ming, *Optimal control of stochastic flow over a backward-facing step using ROM*, SISC 2011.
- Hou, Lee, & Manouzi, *Finite element approximations of stochastic optimal control problems constrained by stochastic elliptic PDEs*, J Math Anal Appl, 2011.
- Gunzburger, Lee, & Lee, *Error estimates of stochastic optimal Neumann boundary control problems*, SINUM, 2011.
- Rosseel & Wells, *Optimal control with stochastic PDE constraints and uncertain controls*, CMAME, 2012.
- Tiesler, Kirby, Xiu, & Preusser, *Stochastic collocation for optimal control problems with stochastic PDE constraints*, SICON, 2012.
- Kouri, Heinkenschloss, Ridzal, & Van Bloemen Waanders, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SISC, 2013.
- Chen, Quarteroni, & Rozza, *Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations*, SINUM, 2013.
- Kunoth & Schwab, *Analytic regularity and gPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs*, SICON, 2013.
- Kouri, *A multilevel stochastic collocation algorithm for optimization of PDEs with uncertain coefficients*, JUQ, 2014.
- Chen & Quarteroni, *Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint*, JUQ, 2014.
- Ng & Willcox, *Multifidelity approaches for optimization under uncertainty*, IJNME, 2014.
- Kouri, Heinkenschloss, Ridzal, & van Bloemen Waanders, *Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty*, SISC, 2014.
- Chen, Quarteroni, & Rozza, *Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations*, Num. Math. 2015.
- Ng & Willcox, *Monte Carlo information-reuse approach to aircraft conceptual design optimization under uncertainty*, J Aircraft, 2015.
- P. Benner, A. Onwunta, and M. Stoll. Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs. SIMAX, 2016.
- A.A. Ali, E. Ullmann, & M. Hinze, Multilevel Monte Carlo analysis for optimal control of elliptic PDEs with random coefficients, SIAM/ASA JUQ, 2017.

Quadratic approximation in infinite dimensions

- We approximate Q by a quadratically-truncated Taylor expansion

$$Q(m) \approx Q_{\mathsf{quad}}(m) = Q(\bar{m}) + \langle g_m(\bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \mathcal{H}_m(\bar{m})(m - \bar{m}), m - \bar{m} \rangle$$

- For a Gaussian random field m with $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$, Q_{quad} is non-Gaussian, but we can still express[1]

$$\mathbb{E}[Q_{\mathsf{quad}}] = Q(\bar{m}) + \frac{1}{2}\mathrm{tr}(\tilde{\mathcal{H}})$$

$$\mathrm{Var}[Q_{\mathsf{quad}}] = \langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m}) \rangle + \frac{1}{2}\mathrm{tr}(\tilde{\mathcal{H}}^2)$$

  where $\tilde{\mathcal{H}} = \mathcal{C}^{1/2}\mathcal{H}_m(\bar{m})\mathcal{C}^{1/2}$ is the *covariance-preconditioned Hessian*
- Q_{quad} is corrected by using it as a control variate (cf. multifidelity methods[2]
- Need to efficiently evaluate $\mathrm{tr}(\tilde{\mathcal{H}})$ and $\mathrm{tr}(\tilde{\mathcal{H}}^2)$ and their gradients w.r.t. z

[1] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations, *SIAM/ASA JUQ*, 2017.

[2] L. Ng & K. Willcox, *Multifidelity approaches for optimization under uncertainty*, IJNME, 2014.

Quadratic approximation in infinite dimensions

- We approximate Q by a quadratically-truncated Taylor expansion

$$Q(m) \approx Q_{\text{quad}}(m) = Q(\bar{m}) + \langle g_m(\bar{m}), m - \bar{m} \rangle + \frac{1}{2}\langle \mathcal{H}_m(\bar{m})(m - \bar{m}), m - \bar{m} \rangle$$

- For a Gaussian random field m with $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$, Q_{quad} is non-Gaussian, but we can still express[1]

$$\mathbb{E}[Q_{\text{quad}}] = Q(\bar{m}) + \frac{1}{2}\text{tr}(\tilde{\mathcal{H}})$$

$$\text{Var}[Q_{\text{quad}}] = \langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m}) \rangle + \frac{1}{2}\text{tr}(\tilde{\mathcal{H}}^2)$$

  where $\tilde{\mathcal{H}} = \mathcal{C}^{1/2}\mathcal{H}_m(\bar{m})\mathcal{C}^{1/2}$ is the *covariance-preconditioned Hessian*

- Q_{quad} is corrected by using it as a control variate (cf. multifidelity methods[2]
- Need to efficiently evaluate $\text{tr}(\tilde{\mathcal{H}})$ and $\text{tr}(\tilde{\mathcal{H}}^2)$ and their gradients w.r.t. z

---

[1] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations, *SIAM/ASA JUQ*, 2017.

[2] L. Ng & K. Willcox, *Multifidelity approaches for optimization under uncertainty*, IJNME, 2014.

Quadratic approximation in infinite dimensions

- We approximate Q by a quadratically-truncated Taylor expansion

$$Q(m) \approx Q_{\mathsf{quad}}(m) = Q(\bar{m}) + \langle g_m(\bar{m}), m - \bar{m} \rangle + \frac{1}{2} \langle \mathcal{H}_m(\bar{m})(m - \bar{m}), m - \bar{m} \rangle$$

- For a Gaussian random field m with $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$, Q_{quad} is non-Gaussian, but we can still express[1]

$$\mathbb{E}[Q_{\mathsf{quad}}] = Q(\bar{m}) + \frac{1}{2}\mathsf{tr}(\tilde{\mathcal{H}})$$

$$\mathsf{Var}[Q_{\mathsf{quad}}] = \langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m}) \rangle + \frac{1}{2}\mathsf{tr}(\tilde{\mathcal{H}}^2)$$

  where $\tilde{\mathcal{H}} = \mathcal{C}^{1/2}\mathcal{H}_m(\bar{m})\mathcal{C}^{1/2}$ is the *covariance-preconditioned Hessian*

- Q_{quad} is corrected by using it as a control variate (cf. multifidelity methods[2]
- Need to efficiently evaluate $\mathsf{tr}(\tilde{\mathcal{H}})$ and $\mathsf{tr}(\tilde{\mathcal{H}}^2)$ and their gradients w.r.t. z

[1] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations, *SIAM/ASA JUQ*, 2017.

[2] L. Ng & K. Willcox, *Multifidelity approaches for optimization under uncertainty*, IJNME, 2014.

How to compute $\mathrm{tr}(\tilde{\mathcal{H}})$ efficiently?

- When the eigenvalues decay rapidly (as is common for Hessians), the trace can be approximated efficiently with small N by

$$\mathrm{tr}(\tilde{\mathcal{H}}) \approx \sum_{j=1}^{N} \lambda_j(\tilde{\mathcal{H}}) \;\text{ and }\; \mathrm{tr}(\tilde{\mathcal{H}}^2) \approx \sum_{j=1}^{N} \lambda_j^2(\tilde{\mathcal{H}})$$

  where λ_j, $j = 1, \ldots, N$, are the dominant eigenvalues of $\tilde{\mathcal{H}}$, or the dominant generalized eigenvalues of $(\mathcal{H}_m(\bar{m}), \mathcal{C}^{-1})$, i.e.,

$$\mathcal{H}_m(\bar{m})\psi_j = \lambda_j \mathcal{C}^{-1}\psi_j$$

  where ψ_j are the $\mathcal{C}^{-1}$-orthonormal eigenfunctions, i.e., $\langle \psi_i, \mathcal{C}^{-1}\psi_j \rangle = \delta_{ij}$
- Prohibitive to compute $\mathcal{H}_m$ by itself; instead can form action in a given direction at cost of pair of linearized forward/adjoint PDE solves
- $\Longrightarrow$ Need operator-free eigensolver that can capture dominant spectrum in number of operator applications that scales with effective rank

Computing the trace of $\tilde{\boldsymbol{H}}$ via randomized SVD

- Double-pass randomized SVD algorithm estimates trace at cost of $2r$ products of $\tilde{\boldsymbol{H}}$ with random vectors ($r = N + p$, N is rank of $\tilde{\boldsymbol{H}}$, p is oversampling #)
- Resulting cost is $2r$ pairs of incremental forward/adjoint solves w/same PDE operator and $4r$ Poisson solves
- Covariance operator and Hessian are often compact (Q is sensitive to limited number of modes) so composition is low-rank

Randomized SVD (double pass algorithm)

1. Generate i.i.d. Gaussian matrix $\boldsymbol{R} \in \mathbb{R}^{n\times r}$ with r = numerical rank of $\tilde{\boldsymbol{H}}$ ($r \ll n$)
2. Form $\boldsymbol{Y} = \tilde{\boldsymbol{H}}\boldsymbol{R}$
3. Compute $\boldsymbol{Q}$ = orthonormal basis for $\boldsymbol{Y}$
4. Define $\boldsymbol{B} \in \mathbb{R}^{r\times r} := \boldsymbol{Q}^T\tilde{\boldsymbol{H}}\boldsymbol{Q}$
5. Decompose $\boldsymbol{B} = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^T$
6. Low-rank approximation: $\tilde{\boldsymbol{H}} \approx \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T$, where $\boldsymbol{V} \in \mathbb{R}^{n\times r} := \boldsymbol{Q}\boldsymbol{Z}$
7. Trace estimation: $\mathrm{tr}(\tilde{\boldsymbol{H}}) \approx \mathrm{tr}(\boldsymbol{B})$

- Thus often $r \ll n$, independent of parameter dimension n, and with high probability

$$|\mathrm{tr}(\tilde{\boldsymbol{H}}) - \mathrm{tr}(\boldsymbol{B})| \leq c(p) \sum_{r<i\leq n} |\lambda_i(\tilde{\boldsymbol{H}})|$$

- Quadratic-based approximations of $\mathbb{E}[Q]$ and $\mathrm{Var}[Q]$ require $4r$ linearized PDE solves with same PDE operator (small multiple of highly nonlinear forward solve)
- See Saibaba, Alexanderian, Ipsen, *Numerische Mathematik* 2017 for analysis of trace estimate by more general randomized subspace iterations

Eigenproblem-constrained optimization

With the trace computed via randomized SVD, we obtain

$$\mathcal{J}_{\mathsf{quad}}(z) = \underbrace{Q(\bar{m}) + \frac{1}{2}\sum_{j=1}^{N}\lambda_j(\tilde{\mathcal{H}})}_{\mathbb{E}[Q]} + \underbrace{\beta\langle g_m(\bar{m}), \mathcal{C}g_m(\bar{m})\rangle + \frac{\beta}{2}\sum_{j=1}^{N}\lambda_j^2(\tilde{\mathcal{H}})}_{\beta\mathsf{Var}[Q]} + \mathcal{P}(z)$$

where $Q(\bar{m}) := \bar{Q}$ is obtained by solving the forward problem for $u \in \mathcal{U}$

$$\langle \tilde{v}, \partial_v \bar{r}(u, \tilde{v}, z)\rangle = 0, \quad \forall \tilde{v} \in \mathcal{V}$$

with $\bar{r}(u, \tilde{v}, z) = r(u, \tilde{v}, \bar{m}, z)$ for short. By defining the Lagrangian

$$\mathcal{L}(u, v, \bar{m}, z) = Q(u) + \bar{r}(u, v, z)$$

the gradient $g_m(\bar{m})$ is found from

$$\langle \tilde{m}, g_m(\bar{m})\rangle = \langle \tilde{m}, \partial_m \mathcal{L}\rangle = \langle \tilde{m}, \partial_m \bar{r}(u, v, z)\rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

for which we need to compute $v \in \mathcal{V}$ by solving the adjoint problem

$$\langle \tilde{u}, \partial_u \bar{r}(u, v, z)\rangle = -\langle \tilde{u}, \partial_u \bar{Q}\rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

Eigenproblem-constrained optimization

With the trace computed via randomized SVD, we obtain

$$\mathcal{J}_{\text{quad}}(z) = \underbrace{Q(\bar{m}) + \frac{1}{2}\sum_{j=1}^{N} \lambda_j(\tilde{\mathcal{H}})}_{\mathbb{E}[Q]} + \underbrace{\beta\langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m})\rangle + \frac{\beta}{2}\sum_{j=1}^{N} \lambda_j^2(\tilde{\mathcal{H}})}_{\beta \text{Var}[Q]} + \mathcal{P}(z)$$

where $Q(\bar{m}) := \bar{Q}$ is obtained by solving the forward problem for $u \in \mathcal{U}$

$$\langle \tilde{v}, \partial_v \bar{r}(u, \tilde{v}, z)\rangle = 0, \quad \forall \tilde{v} \in \mathcal{V}$$

with $\bar{r}(u, \tilde{v}, z) = r(u, \tilde{v}, \bar{m}, z)$ for short. By defining the Lagrangian

$$\mathcal{L}(u, v, \bar{m}, z) = Q(u) + \bar{r}(u, v, z)$$

the gradient $g_m(\bar{m})$ is found from

$$\langle \tilde{m}, g_m(\bar{m})\rangle = \langle \tilde{m}, \partial_m \mathcal{L}\rangle = \langle \tilde{m}, \partial_m \bar{r}(u, v, z)\rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

for which we need to compute $v \in \mathcal{V}$ by solving the adjoint problem

$$\langle \tilde{u}, \partial_u \bar{r}(u, v, z)\rangle = -\langle \tilde{u}, \partial_u \bar{Q}\rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

Eigenproblem constrained optimization

To compute λ_j, which satisfies for $j = 1, \ldots, N$

$$\mathcal{H}_m(\bar{m})\psi_j = \lambda_j \mathcal{C}^{-1}\psi_j, \text{ and } \langle \psi_j, \mathcal{C}^{-1}\psi_j \rangle = \delta_{ij}$$

we need Hessian action in a direction $\hat{m}$, for which we form the Lagrangian

$$\mathcal{L}^H(u, v, m, z; \hat{u}, \hat{v}, \hat{m}) = \underbrace{\langle \hat{m}, \partial_m \bar{r} \rangle}_{\text{gradient}} + \underbrace{\langle \hat{v}, \partial_v \bar{r} \rangle}_{\text{forward}} + \underbrace{\langle \hat{u}, \partial_u \bar{r} + \partial_u \bar{Q} \rangle}_{\text{adjoint}}$$

which involves the gradient, the forward problem, and the adjoint problem. The Hessian action is given by the variation of $\mathcal{L}^H$ with respect to m:

$$\langle \tilde{m}, \mathcal{H}_m(\bar{m})\,\hat{m} \rangle = \langle \tilde{m}, \partial_m \mathcal{L}^H\,\hat{m} \rangle = \langle \tilde{m}, \partial_{mv}\bar{r}\,\hat{v} + \partial_{mu}\bar{r}\,\hat{u} + \partial_{mm}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

where $\hat{u} \in \mathcal{U}$ is the solution of the incremental forward problem, $\partial_u \mathcal{L}^H = 0$

$$\langle \tilde{v}, \partial_{vu}\bar{r}\,\hat{u} \rangle = -\langle \tilde{v}, \partial_{vm}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{v} \in \mathcal{V}$$

and $\hat{v} \in \mathcal{V}$ is the solution of the incremental adjoint problem, $\partial_v \mathcal{L}^H = 0$

$$\langle \tilde{u}, \partial_{uv}\bar{r}\,\hat{v} \rangle = -\langle \tilde{u}, \partial_{uu}\bar{r}\,\hat{u} + \partial_{uu}\bar{Q}\,\hat{u} + \partial_{um}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

Eigenproblem constrained optimization

To compute λ_j, which satisfies for $j = 1, \ldots, N$

$$\mathcal{H}_m(\bar{m})\psi_j = \lambda_j \mathcal{C}^{-1}\psi_j, \text{ and } \langle \psi_j, \mathcal{C}^{-1}\psi_j \rangle = \delta_{ij}$$

we need Hessian action in a direction $\hat{m}$, for which we form the Lagrangian

$$\mathcal{L}^H(u, v, m, z; \hat{u}, \hat{v}, \hat{m}) = \underbrace{\langle \hat{m}, \partial_m \bar{r} \rangle}_{\text{gradient}} + \underbrace{\langle \hat{v}, \partial_v \bar{r} \rangle}_{\text{forward}} + \underbrace{\langle \hat{u}, \partial_u \bar{r} + \partial_u \bar{Q} \rangle}_{\text{adjoint}}$$

which involves the gradient, the forward problem, and the adjoint problem. The Hessian action is given by the variation of $\mathcal{L}^H$ with respect to m:

$$\langle \tilde{m}, \mathcal{H}_m(\bar{m})\,\hat{m} \rangle = \langle \tilde{m}, \partial_m \mathcal{L}^H\,\hat{m} \rangle = \langle \tilde{m}, \partial_{mv}\bar{r}\,\hat{v} + \partial_{mu}\bar{r}\,\hat{u} + \partial_{mm}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

where $\hat{u} \in \mathcal{U}$ is the solution of the incremental forward problem, $\partial_u \mathcal{L}^H = 0$

$$\langle \tilde{v}, \partial_{vu}\bar{r}\,\hat{u} \rangle = -\langle \tilde{v}, \partial_{vm}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{v} \in \mathcal{V}$$

and $\hat{v} \in \mathcal{V}$ is the solution of the incremental adjoint problem, $\partial_v \mathcal{L}^H = 0$

$$\langle \tilde{u}, \partial_{uv}\bar{r}\,\hat{v} \rangle = -\langle \tilde{u}, \partial_{uu}\bar{r}\,\hat{u} + \partial_{uu}\bar{Q}\,\hat{u} + \partial_{um}\bar{r}\,\hat{m} \rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

Eigenproblem constrained optimization

To compute λ_j, which satisfies for $j = 1, \dots, N$

$$\mathcal{H}_m(\bar{m})\psi_j = \lambda_j \mathcal{C}^{-1}\psi_j, \text{ and } \langle \psi_j, \mathcal{C}^{-1}\psi_j \rangle = \delta_{ij}$$

we need Hessian action in a direction $\hat{m}$, for which we form the Lagrangian

$$\mathcal{L}^H(u, v, m, z; \hat{u}, \hat{v}, \hat{m}) = \underbrace{\langle \hat{m}, \partial_m \bar{r} \rangle}_{\text{gradient}} + \underbrace{\langle \hat{v}, \partial_v \bar{r} \rangle}_{\text{forward}} + \underbrace{\langle \hat{u}, \partial_u \bar{r} + \partial_u \bar{Q} \rangle}_{\text{adjoint}}$$

which involves the gradient, the forward problem, and the adjoint problem. The Hessian action is given by the variation of $\mathcal{L}^H$ with respect to m:

$$\langle \tilde{m}, \mathcal{H}_m(\bar{m})\, \hat{m} \rangle = \langle \tilde{m}, \partial_m \mathcal{L}^H\, \hat{m} \rangle = \langle \tilde{m}, \partial_{mv}\bar{r}\, \hat{v} + \partial_{mu}\bar{r}\, \hat{u} + \partial_{mm}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

where $\hat{u} \in \mathcal{U}$ is the solution of the incremental forward problem, $\partial_u \mathcal{L}^H = 0$

$$\langle \tilde{v}, \partial_{vu}\bar{r}\, \hat{u} \rangle = -\langle \tilde{v}, \partial_{vm}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{v} \in \mathcal{V}$$

and $\hat{v} \in \mathcal{V}$ is the solution of the incremental adjoint problem, $\partial_v \mathcal{L}^H = 0$

$$\langle \tilde{u}, \partial_{uv}\bar{r}\, \hat{v} \rangle = -\langle \tilde{u}, \partial_{uu}\bar{r}\, \hat{u} + \partial_{uu}\bar{Q}\, \hat{u} + \partial_{um}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

Eigenproblem constrained optimization

To compute λ_j, which satisfies for $j = 1, \ldots, N$

$$\mathcal{H}_m(\bar{m})\psi_j = \lambda_j \mathcal{C}^{-1}\psi_j, \text{ and } \langle \psi_j, \mathcal{C}^{-1}\psi_j \rangle = \delta_{ij}$$

we need Hessian action in a direction $\hat{m}$, for which we form the Lagrangian

$$\mathcal{L}^H(u, v, m, z; \hat{u}, \hat{v}, \hat{m}) = \underbrace{\langle \hat{m}, \partial_m \bar{r} \rangle}_{\text{gradient}} + \underbrace{\langle \hat{v}, \partial_v \bar{r} \rangle}_{\text{forward}} + \underbrace{\langle \hat{u}, \partial_u \bar{r} + \partial_u \bar{Q} \rangle}_{\text{adjoint}}$$

which involves the gradient, the forward problem, and the adjoint problem. The Hessian action is given by the variation of $\mathcal{L}^H$ with respect to m:

$$\langle \tilde{m}, \mathcal{H}_m(\bar{m})\, \hat{m} \rangle = \langle \tilde{m}, \partial_m \mathcal{L}^H\, \hat{m} \rangle = \langle \tilde{m}, \partial_{mv}\bar{r}\, \hat{v} + \partial_{mu}\bar{r}\, \hat{u} + \partial_{mm}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{m} \in \mathcal{M}$$

where $\hat{u} \in \mathcal{U}$ is the solution of the incremental forward problem, $\partial_u \mathcal{L}^H = 0$

$$\langle \tilde{v}, \partial_{vu}\bar{r}\, \hat{u} \rangle = -\langle \tilde{v}, \partial_{vm}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{v} \in \mathcal{V}$$

and $\hat{v} \in \mathcal{V}$ is the solution of the incremental adjoint problem, $\partial_v \mathcal{L}^H = 0$

$$\langle \tilde{u}, \partial_{uv}\bar{r}\, \hat{v} \rangle = -\langle \tilde{u}, \partial_{uu}\bar{r}\, \hat{u} + \partial_{uu}\bar{Q}\, \hat{u} + \partial_{um}\bar{r}\, \hat{m} \rangle, \quad \forall \tilde{u} \in \mathcal{U}$$

OUU problem with quadratic approximation $\mathcal{J}_{\text{quad}}$

$$\min_{z \in \mathcal{Z}} \mathcal{J}_{\text{quad}}(z) := Q(\bar{m}) + \frac{1}{2}\sum_{j=1}^{N} \lambda_j(\tilde{\mathcal{H}}) + \beta \left(\langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m}) \rangle + \frac{1}{2}\sum_{j=1}^{N} \lambda_j^2(\tilde{\mathcal{H}}) \right) + \mathcal{P}(z)$$

where:

forward	$\langle v^*, \partial_v \bar{r} \rangle = 0 \quad \forall v^* \in \mathcal{V}$
adjoint	$\langle u^*, \partial_u \bar{r} + \partial_u \bar{Q} \rangle = 0 \quad \forall u^* \in \mathcal{U}$
eigenvalue	$\langle \psi_j^*, (\mathcal{H}_m(\bar{m}) - \lambda_j \mathcal{C}^{-1}) \psi_j \rangle = 0 \quad \forall \psi_j^* \in \mathcal{M} \quad j = 1, \dots, N$
orthonormality	$\lambda_j^* (\langle \psi_j, \mathcal{C}^{-1} \psi_j \rangle - 1) = 0 \quad \forall \lambda_j^* \in \mathbb{R} \quad j = 1, \dots, N$
incremental forw	$\langle \hat{v}_j^*, \partial_{vu} \bar{r}\, \hat{u}_j + \partial_{vm} \bar{r}\, \psi_j \rangle = 0 \quad \forall \hat{v}_j^* \in \mathcal{V} \quad j = 1, \dots, N$
incremental adj	$\langle \hat{u}_j^*, \partial_{uv} \bar{r}\, \hat{v}_j + \partial_{uu} \bar{r}\, \hat{u}_j + \partial_{uu} \bar{Q}\, \hat{u}_j + \partial_{um} \bar{r}\, \psi_j \rangle = 0 \;\; \forall \hat{u}_j^* \in \mathcal{U} \;\; j = 1, \dots$

Lagrangian of the OUU problem

$$\mathcal{L}_{\text{quad}}\left(u, v, \{\lambda_j\}, \{\psi_j\}, \{\hat{u}_j\}, \{\hat{v}_j\}, u^*, v^*, \{\lambda_j^*\}, \{\psi_j^*\}, \{\hat{u}_j^*\}, \{\hat{v}_j^*\}, z\right) :=$$

quad obj	$= Q(\bar{m}) + \frac{1}{2}\sum_{j=1}^{N} \lambda_j(\tilde{\mathcal{H}}) + \beta\left(\langle g_m(\bar{m}), \mathcal{C} g_m(\bar{m})\rangle + \frac{1}{2}\sum_{j=1}^{N} \lambda_j^2(\tilde{\mathcal{H}})\right) + \mathcal{P}(z)$
forward	$+\ \langle v^*, \partial_v \bar{r}\rangle$
adjoint	$+\ \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q}\rangle$
eigen. prob.	$+\ \sum_{j=1}^{N} \langle \psi_j^*, (\mathcal{H}_m(\bar{m}) - \lambda_j \mathcal{C}^{-1})\psi_j\rangle$
orth. cond.	$+\ \sum_{j=1}^{N} \lambda_j^* (\langle \psi_j, \mathcal{C}^{-1}\psi_j\rangle - 1)$
inc. fwd.	$+\ \sum_{j=1}^{N} \langle \hat{v}_j^*, \partial_{vu}\bar{r}\, \hat{u}_j + \partial_{vm}\bar{r}\, \psi_j\rangle$
inc. adj.	$+\ \sum_{j=1}^{N} \langle \hat{u}_j^*, \partial_{uv}\bar{r}\, \hat{v}_j + \partial_{uu}\bar{r}\, \hat{u}_j + \partial_{uu}\bar{Q}\, \hat{u}_j + \partial_{um}\bar{r}\, \psi_j\rangle$

Gradient of $\mathcal{J}_{\text{quad}}$ (assuming λ_j distinct)

- Variation of $\mathcal{L}_{\text{quad}}$ wrt λ_j vanishes:

$$\psi_j^* = \frac{1+2\beta\lambda_j}{2}\psi_j, \quad j = 1,\ldots,N$$

- Variation of $\mathcal{L}_{\text{quad}}$ wrt $\hat{v}_j$ vanishes:

$$\hat{u}_j^* = \frac{1+2\beta\lambda_j}{2}\hat{u}_j, \quad j = 1,\ldots,N$$

- Variation of $\mathcal{L}_{\text{quad}}$ wrt $\hat{u}_j$ vanishes:

$$\hat{v}_j^* = \frac{1+2\beta\lambda_j}{2}\hat{v}_j, \quad j = 1,\ldots,N$$

- Variation of $\mathcal{L}_{\text{quad}}$ wrt v vanishes: find $u^* \in \mathcal{U}$ s.t. (incr forward operator)

$$\begin{aligned}
\langle \tilde{v}, \partial_{vu}\bar{r}\, u^* \rangle &= -2\beta \langle \tilde{v}, \partial_{vm}\bar{r}\,(\mathcal{C}\partial_m \bar{r})\rangle \\
&\quad - \sum_{j=1}^{N} \langle \tilde{v}, \partial_{vmu}\bar{r}\,\hat{u}_j\,\psi_j^* + \partial_{vmm}\bar{r}\,\psi_j\,\psi_j^* \rangle \\
&\quad - \sum_{j=1}^{N} \langle \tilde{v}, \partial_{vuu}\bar{r}\,\hat{u}_j\,\hat{u}_j^* + \partial_{vum}\bar{r}\,\psi_j\,\hat{u}_j^* \rangle, \quad \forall \tilde{v} \in \mathcal{V}
\end{aligned}$$

Computing the gradient of the OUU problem

- Variation of $\mathcal{L}_{\text{quad}}$ wrt u vanishes: find $v^* \in \mathcal{V}$ s.t. (incr adjoint operator)

$$
\begin{aligned}
&\langle \tilde{u}, \partial_{uv}\bar{r}\, v^* \rangle = \\
&- \langle \tilde{u}, \partial_u \bar{Q} \rangle - 2\beta \langle \tilde{u}, \partial_{um}\bar{r}\, (\mathcal{C}\partial_m \bar{r}) \rangle \\
&- \langle \tilde{u}, \partial_{uu}\bar{r}\, u^* + \partial_{uu}\bar{Q}\, u^* \rangle \\
&- \sum_{j=1}^{N} \langle \tilde{u}, \partial_{umv}\bar{r}\, \hat{v}_j\, \psi_j^* + \partial_{umu}\bar{r}\, \hat{u}_j\, \psi_j^* + \partial_{uum}\bar{r}\, \psi_j\, \psi_j^* \rangle \\
&- \sum_{j=1}^{N} \langle \tilde{u}, \partial_{uvu}\bar{r}\, \hat{u}_j\, \hat{v}_j^* + \partial_{uvm}\bar{r}\, \psi_j\, \hat{v}_j^* \rangle \\
&- \sum_{j=1}^{N} \langle \tilde{u}, \partial_{uuv}\bar{r}\, \hat{v}_j\, \hat{u}_j^* + \partial_{uuu}\bar{r}\, \hat{u}_j\, \hat{u}_j^* + \partial_{uuu}\bar{Q}\, \hat{u}_j\, \hat{u}_j^* + \partial_{uum}\bar{r}\, \psi_j\, \hat{u}_j^* \rangle, \forall \tilde{u} \in \mathcal{U},
\end{aligned}
$$

- Finally the gradient of the cost functional can be computed as

$$
D_z \mathcal{J}_{\text{quad}}(z) = \partial_z \mathcal{L}_{\text{quad}}(\text{primal}, \text{dual}, z)
$$

- Total cost: 1 forward PDE solve, $1 + 4(N + p) + 2N + 2$ linearized PDE solves (independent of uncertain parameter or control dimensions!)

Quadratic approximation as a variance reduction

- Statistics computed by quadratic approximation may be **biased**
- Use Monte Carlo quadrature to correct quadratic approximation

$$\mathbb{E}[Q] = \mathbb{E}[Q_{\text{quad}}] + \mathbb{E}[\underbrace{Q - Q_{\text{quad}}}_{Y}] \approx \mathbb{E}[Q_{\text{quad}}] + \underbrace{\hat{Y}}_{\text{MC estimator}}$$

- Mean squared error (MSE) of MC estimate of $\mathbb{E}[Q]$ and $\mathbb{E}[Y]$

$$\text{MSE}(Q) \asymp \frac{1}{M}\text{Var}[Q] \quad \text{vs.} \quad \text{MSE}(Y) \asymp \frac{1}{M}\text{Var}[Y]$$

- A much smaller number of MC samples is required for $\mathbb{E}[Y]$ as

$$\text{Var}[Y] \ll \text{Var}[Q]$$

  provided Q_{quad} is a good approximation of (highly correlated to) Q
- Similar variance reduction can be applied for the variance $\text{Var}[Q]$.

The unbiased cost functional with variance reduction

We obtain an unbiased evaluation of the cost functional as

$$\mathcal{J}^{\mathsf{MC}}_{\mathsf{quad}}(z) = \widehat{Q}_{\mathsf{quad}} + \beta \widehat{V}^{\mathsf{quad}}_{Q} + \mathcal{P}(z)$$

where

$$\begin{aligned}\widehat{Q}_{\mathrm{quad}} &= Q(\bar{m}) + \frac{1}{2}\mathsf{tr}(\mathcal{H}) \\ &\quad + \frac{1}{M}\sum_{i=1}^{M}\Big(Q(m_i) - Q(\bar{m}) - \langle m_i - \bar{m}, g_m(\bar{m})\rangle \\ &\qquad - \frac{1}{2}\langle m_i - \bar{m}, \mathcal{H}_m(\bar{m})\,(m_i - \bar{m})\rangle\Big)\end{aligned}$$

and

$$\begin{aligned}\widehat{V}^{\mathsf{quad}}_{Q} &:= \langle \mathcal{C} g_m(\bar{m}), g_m(\bar{m})\rangle + \frac{1}{4}(\mathsf{tr}(\mathcal{H}))^2 + \frac{1}{2}\mathsf{tr}(\mathcal{H}^2) \\ &\quad + \frac{1}{M}\sum_{i=1}^{M}\Big((Q(m_i) - Q(\bar{m}))^2 \\ &\quad - \Big(\langle m_i - \bar{m}, g_m(\bar{m})\rangle + \frac{1}{2}\langle m_i - \bar{m}, \mathcal{H}_m(\bar{m})\,(m_i - \bar{m})\rangle\Big)^2\Big) \\ &\quad - \Big(\frac{1}{2}\mathsf{tr}(\mathcal{H}) + \frac{1}{M_2}\sum_{i=1}^{M_2}\Big(Q(m_i) - \langle m_i - \bar{m}, g_m(\bar{m})\rangle \\ &\qquad - \frac{1}{2}\langle m_i - \bar{m}, \mathcal{H}_m(\bar{m})\,(m_i - \bar{m})\rangle\Big)\Big)^2\end{aligned}$$

OUU Lagrangian w/variance reduction using quad approx

$$
\begin{aligned}
&\mathcal{L}^{\mathsf{MC}}_{\mathsf{quad}}\Big(u, v, \{u_i\}, \{\lambda_j\}, \{\psi_j\}, \{\hat{u}_j\}, \{\hat{v}_j\}, \{\hat{u}_i\}, \{\hat{v}_i\},\\
&\qquad u^*, v^*, \{v_i\}, \{\lambda^*_j\}, \{\psi^*_j\}, \{\hat{u}^*_j\}, \{\hat{v}^*_j\}, \{\hat{u}^*_i\}, \{\hat{v}^*_i\}, z\Big)\\
&= \mathcal{J}^{\mathsf{MC}}_{\mathsf{quad}} + \langle v^*, \partial_v \bar{r}\rangle + \langle u^*, \partial_u \bar{r} + \partial_u \bar{Q}\rangle + \sum_{i=1}^{M} r(u_i, v_i, m_i, z).\\
&+ \sum_{j=1}^{N} \langle \psi^*_j, (\mathcal{H}_m(\bar{m}) - \lambda_j \mathcal{C}^{-1})\psi_j\rangle\\
&+ \sum_{j=1}^{N} \lambda^*_j (\langle \psi_j, \mathcal{C}^{-1}\psi_j\rangle - 1)\\
&+ \sum_{j=1}^{N} \langle \hat{v}^*_j, \partial_{vu}\bar{r}\,\hat{u}_j + \partial_{vm}\bar{r}\,\psi_j\rangle\\
&+ \sum_{j=1}^{N} \langle \hat{u}^*_j, \partial_{uv}\bar{r}\,\hat{v}_j + \partial_{uu}\bar{r}\,\hat{u}_j + \partial_{uu}\bar{Q}\,\hat{u}_j + \partial_{um}\bar{r}\,\psi_j\rangle\\
&+ \sum_{i=1}^{M} \langle \hat{v}^*_i, \partial_{vu}\bar{r}\,\hat{u}_i + \partial_{vm}\bar{r}\,m_i\rangle\\
&+ \sum_{i=1}^{M} \langle \hat{u}^*_i, \partial_{uv}\bar{r}\,\hat{v}_i + \partial_{uu}\bar{r}\,\hat{u}_i + \partial_{uu}\bar{Q}\,\hat{u}_i + \partial_{um}\bar{r}\,m_i\rangle.
\end{aligned}
$$

Total: $1 + M$ forward PDE solves and $3 + 4(N + p) + 4N + 5M$ linearized PDE solves

Optimal design of acoustic metamaterial cloak: Setup

Helmholtz equation:

$$\Delta u + k^2 u = 0 \quad \text{in } D$$
$$u = u_{\text{in}} \quad \text{on } \Gamma_{\text{in}}$$
$$\lim_{r\to\infty} r(\partial_r u^s - iku^s) = 0$$

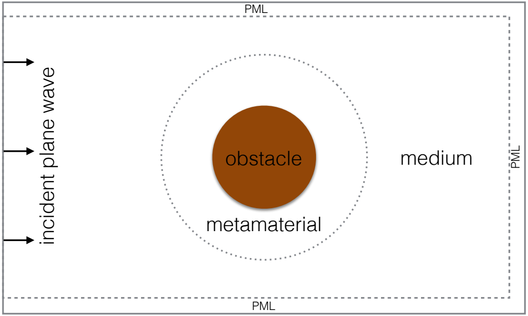
Absorbing BC on Γ_{out} via PML

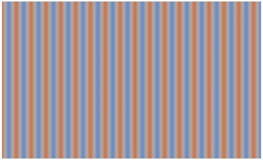
u: (complex) total field $=$ incident field $+$ scattered field

$$u = u^i + u^s$$

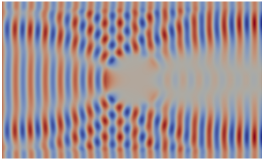
k: wavenumber ω^2/c^2, given by

$$k = \begin{cases} k_m & \text{in medium} \\ k_o & \text{in obstacle} \\ k_m e^{m+z} & \text{in metamaterial} \end{cases}$$





incident field



total field

Optimal design of acoustic cloak: Setup

- The complex field $u = u_r + iu_i$ and adjoint $v = v_r + iv_i$, are defined in the Hilbert space $(u_r, u_i), (v_r, v_i) \in V = H^1_{\Gamma_{\text{in}}} \times H^1_{\Gamma_{\text{in}}}$, where

$$H^1_{\Gamma_{\text{in}}} = \{w \in L^2(D), |\nabla w| \in L^2(D), w|_{\Gamma_{\text{in}}} = 0\}.$$

- The weak form is given by: find $(u_r, u_i) \in V$ such that

$$r(u, v, m, z) = 0 \quad \forall (v_r, v_i) \in V,$$

with $r(u, v, m, z) = r_1(u, v_r, m, z) + ir_2(u, v_i, m, z)$, where

$$r_1(u, v_r, m, z) = \int_D A_r \nabla u_r \cdot \nabla v_r + A_i \nabla u_i \cdot \nabla v_r dx - \int_D K_r u_r v_r + K_i u_i v_r dx,$$

$$r_2(u, v_i, m, z) = \int_D -A_r \nabla u_i \cdot \nabla v_i + A_i \nabla u_r \cdot \nabla v_i dx - \int_D K_r u_i v_i - K_i u_r v_i dx,$$

where A_r, A_i, K_r, K_i depend on (m, z) through the wavenumber k.

- The objective function is given by the misfit in the background medium D_{back}

$$Q(u(m, z)) = \int_{D_{\text{back}}} |u(m, z) - u_{\text{back}}|^2 dx.$$

The regularization term uses L_1-norm to promote design sparsity in metamaterial

$$\mathcal{P}(z) = \alpha \int_{D_{\text{meta}}} |z| dx.$$

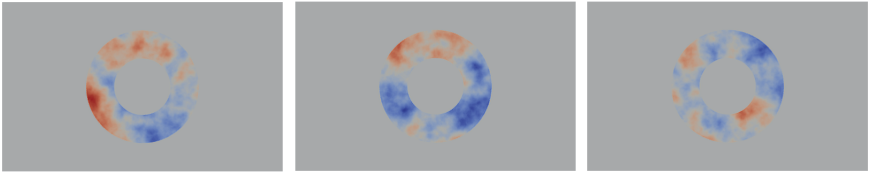
Optimal design of acoustic cloak: Samples

DOF for FE discretization of state, random, and design variables (FEniCS)

DOF	mesh1	mesh2	mesh3	mesh4	mesh5
u (P2)	40,194	159,746	636,930	2,543,618	10,166,274
m, z (P1)	940	3,336	12,487	48,288	189,736

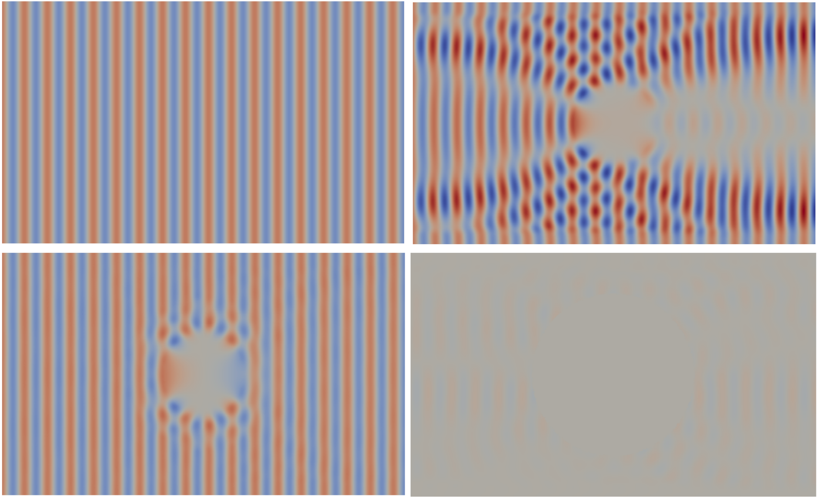
The random field $m \sim \mathcal{N}(\bar{m}, \mathcal{C})$ with mean $\bar{m} = 0$ and covariance

$$\mathcal{C} = (-\gamma\Delta + \delta I)^{-2} \quad \text{with correlation length} \sim \sqrt{\frac{\gamma}{\delta}}$$



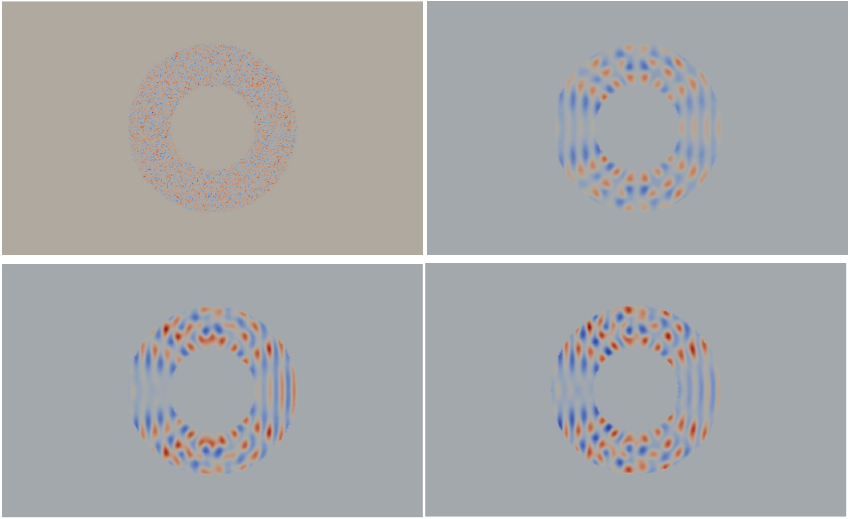
Samples of the random field m ($\gamma = \delta = 50$, corresponding to manufacturing error of $10\% \sim 15\%$ of material property)

Optimal design of acoustic cloak: Fields



Top: No cloak: Incident field and total field with obstacle
Bottom: Optimal cloak: Total field and scattered field

Optimal design of acoustic cloak: Optimal design



Optimal design (∞-dim design field z) with different approximations

Top: Random design, deterministic; Bottom: quadratic, SAA

Optimal design of cloak: Deterministic vs stochastic

Table: Estimates of $q = (Q - \bar{Q})^2$ and mean square errors with 100 samples

design	$\hat{q}$	MSE($\hat{q}$)	MSE($q - q_{\text{lin}}$)	MSE($q - q_{\text{quad}}$)
z_{random}	1.01e+01	2.97e+00	1.90e+00	1.50e-03
z_{deter}	1.13e+01	4.89e+00	4.89e+00	7.32e-02
z_{quad}	1.30e+00	4.06e-02	3.81e-02	1.07e-02
z_{saa}	1.41e+00	2.54e-02	1.54e-02	2.89e-04

Variance reduction of 100X–1000X by quadratic approximation

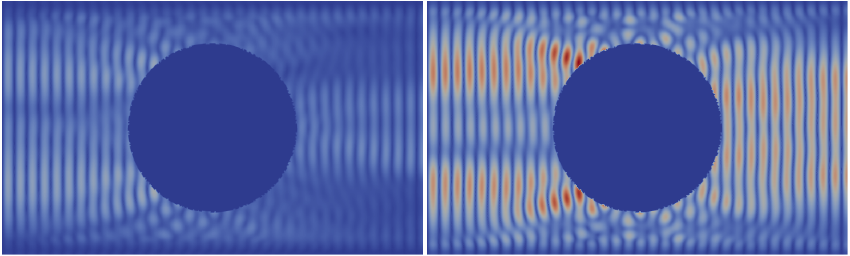


Figure: Std of the scattered field at optimal design z_{quad} (left) and z_{deter} (right)

Optimal design of acoustic cloak: Quad vs SAA

Table: Estimates of misfit Q and mean square errors with 100 samples

design	$\hat{Q}$	MSE($\hat{Q}$)	MSE($Q - Q_{\text{lin}}$)	MSE($Q - Q_{\text{quad}}$)
z_{random}	6.56e+01	9.67e-02	9.80e-03	1.63e-05
z_{deter}	2.55e+00	4.75e-02	4.75e-02	7.30e-04
z_{quad}	1.17e+00	4.85e-03	4.31e-03	6.74e-04
z_{saa}	6.46e+00	1.01e-02	1.29e-03	3.37e-05

Variance reduction of 10X–1000X by quadratic approximation

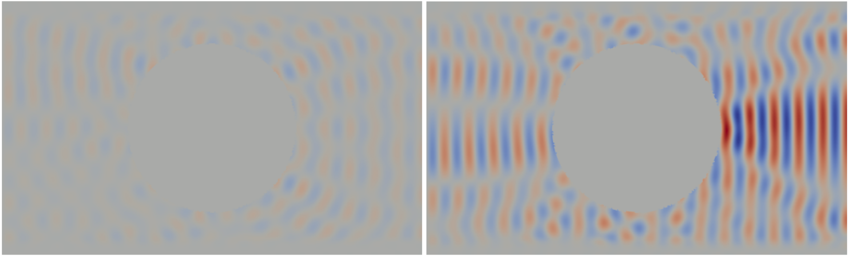
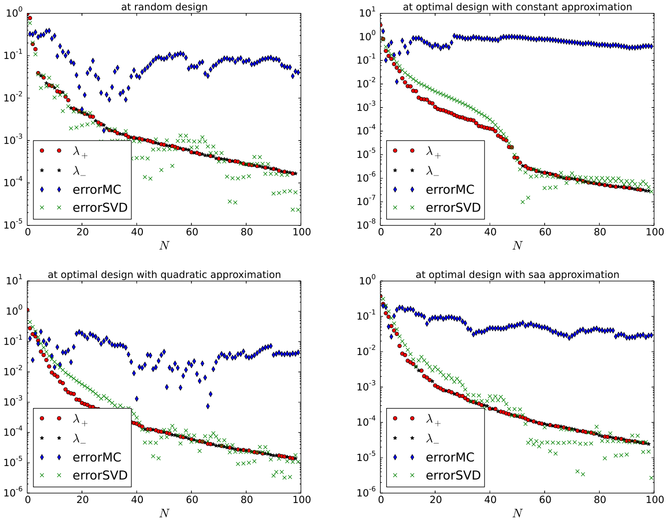


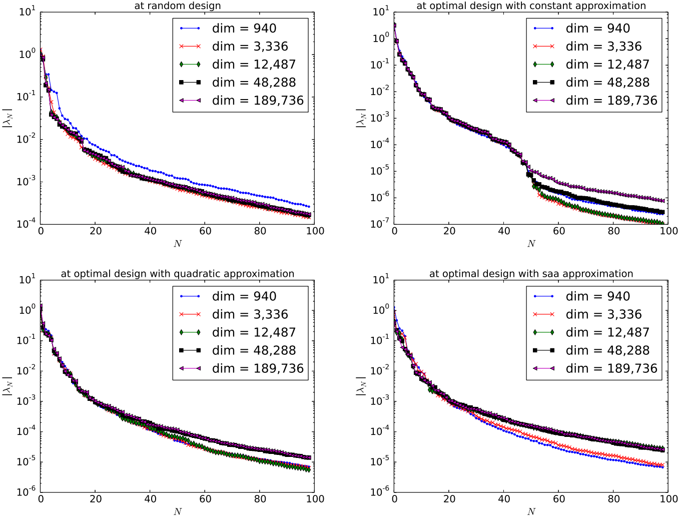
Figure: Mean of the scattered field at optimal design z_{quad} (left) and z_{saa} (right)

Optimal design of acoustic cloak: Trace estimate



Eigenvalues $\lambda_N(\mathcal{C}^{1/2}\mathcal{H}_m(\bar{m})\mathcal{C}^{1/2})$ (first 100 out of 189,736) and trace estimation errors by MC and randomized SVD

Optimal design of acoustic cloak: Scalability I



Spectrum decay of the covariance-preconditioned Hessian is scalable

Optimal design of acoustic cloak: Scalability II

Table: Estimates of misfit Q and mean square errors with 100 samples

dimension	$\hat{Q}$	MSE($\hat{Q}$)	MSE($Q - Q_{\text{lin}}$)	MSE($Q - Q_{\text{quad}}$)
940	7.33e+01	1.25e-01	7.01e-03	7.16e-05
3,336	6.87e+01	1.56e-01	9.29e-03	7.51e-05
12,487	6.56e+01	9.67e-02	9.80e-03	1.63e-05
48,288	6.94e+01	1.00e-01	1.04e-02	1.13e-04

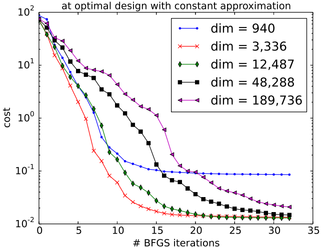
Variance reduction of 1000X (at random design) is scalable

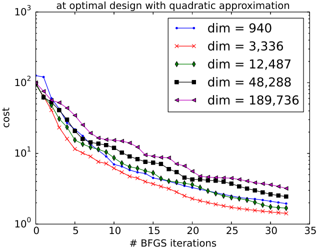
Table: Estimates of $q = (Q - Q_0)^2$ and mean square errors with 100 samples

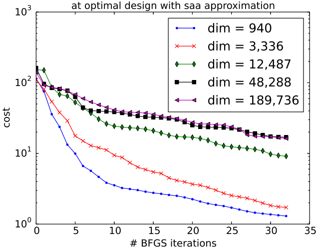
dimension	$\hat{q}$	MSE($\hat{q}$)	MSE($q - q_{\text{lin}}$)	MSE($q - q_{\text{quad}}$)
940	1.44e+01	3.19e+00	1.42e+00	7.53e-03
3,336	2.06e+01	1.13e+01	3.10e+00	1.99e-02
12,487	1.01e+01	2.97e+00	1.90e+00	1.50e-03
48,288	1.21e+01	4.82e+00	2.52e+00	4.92e-03

Variance reduction of 1000X (at random design) is scalable

Optimal design of acoustic cloak: Scalability III



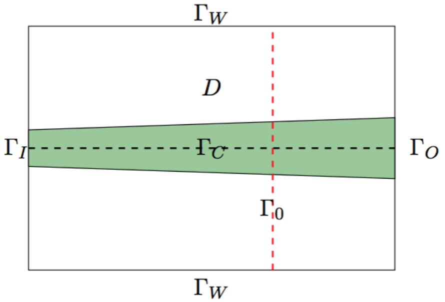




Optimization (# BFGS inter) is scalable by quadratic approximation

Optimal control for turbulent jet flow: setup

- **Control** is horizontal velocity profile at inlet boundary Γ_I
- **Objective** is to maximize jet width at Γ_0
- **Random input** is an inadequacy field for turbulent viscosity, upto 10^6 dimensions.
- **Constraint** on inlet momentum: $\int_{\Gamma_I} (\mathbf{u} \cdot \mathbf{n})^2 ds = M_I$



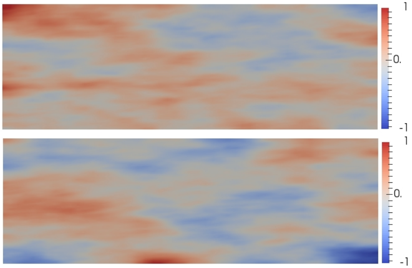


Figure: Left: sketch of the physical domain of the turbulence jet flow, with inlet boundary Γ_I, outlet boundary Γ_O, top and bottom wall Γ_W, the center axis Γ_C, and the cross-section Γ_0. The computational domain D is the top part of the physical domain. Right: two random samples drawn from the Gaussian measure $\mathcal{N}(0, \mathcal{C})$ with $\mathcal{C} = (-\nabla \cdot (\Theta \nabla) + \alpha I)^{-2}$.

Optimal control for turbulent jet flow: model

$$
\begin{aligned}
-\nabla\cdot\left((\nu+\gamma\nu_{t,0})\left(\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^\top\right)\right)+(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}+\nabla p &= 0, &&\text{in } D,\\
\nabla\cdot\boldsymbol{u} &= 0, &&\text{in } D,\\
-\nabla\cdot\left(\left(\nu+(\gamma+e^m)\,\nu_{t,0}\right)\nabla\gamma\right)+\boldsymbol{u}\cdot\nabla\gamma-\frac{1}{2}\frac{\boldsymbol{u}\cdot\boldsymbol{e}_1}{x_1+b}\gamma &= 0, &&\text{in } D,\\
\sigma_n(\boldsymbol{u})\cdot\boldsymbol{\tau}=0,\quad \boldsymbol{u}\cdot\boldsymbol{n}+\chi_W\,\phi(z) &= 0, &&\text{on } \Gamma_I,\\
\sigma_n(\boldsymbol{u})\cdot\boldsymbol{n}=0,\quad \boldsymbol{u}\cdot\boldsymbol{\tau} &= 0, &&\text{on } \Gamma_O\cup\Gamma_W,\\
\sigma_n(\boldsymbol{u})\cdot\boldsymbol{\tau}=0,\quad \boldsymbol{u}\cdot\boldsymbol{n} &= 0, &&\text{on } \Gamma_C,\\
\gamma-\gamma_0 &= 0, &&\text{on } \Gamma_I\cup\Gamma_W,\\
\sigma_n^\gamma(\gamma)\cdot\boldsymbol{n} &= 0, &&\text{on } \Gamma_O\cup\Gamma_C.
\end{aligned}
$$

$$
\begin{aligned}
\sigma_n(\boldsymbol{u}) &= (\nu+\gamma\nu_{t,0})\left(\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^\top\right)\cdot\boldsymbol{n}\\
\sigma_n^\gamma(\gamma) &= (\nu+(\gamma+e^m)\nu_{t,0})\nabla\gamma\cdot\boldsymbol{n}\\
\nu_{t,0}=C\sqrt{M}(x_1+aW)^{1/2}\ \text{with}\ M &= \int_{\Gamma_I}||\boldsymbol{u}_{\mathsf{dns}}||^2 ds\\
\gamma_0=0.5-0.5\tanh\left(5\left(\frac{30-x_1}{30}\right)(x_2-1-0.5x_1)\right)&.
\end{aligned}
$$

Optimal control for turbulent jet flow: model – weak form

The weak form is defined for $u = (\boldsymbol{u}, p, \gamma), v = (\boldsymbol{v}, q, \eta) \in \mathcal{V} := (H^1(D))^2 \times L^2(D) \times H^1(D)$

$$r(u, v, m, z) = \text{Model}(u, v, m) + \text{Stablization}(u, v) + \text{Nitsche}(u, v, m, z),$$

where the first term represents the model in the weak form given by

$$\begin{aligned}
&\text{Model}(u, v, m) \\
&= \int_D (\nu + \nu_t) 2S(\boldsymbol{u}) \cdot S(\boldsymbol{v}) dx + \int_D [(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}] \cdot \boldsymbol{v} dx - \int_D p \nabla \cdot \boldsymbol{v} dx \\
&+ \int_D q \nabla \boldsymbol{u} dx \\
&+ \int_D \left(\nu + (\gamma + e^m)\, \nu_{t,0}\right) \nabla\gamma \cdot \nabla\eta dx + \int_D [\boldsymbol{u} \cdot \nabla\gamma]\eta dx - \int_D \frac{1}{2} \frac{\boldsymbol{u} \cdot \boldsymbol{e}_1}{x_1 + b} \gamma\eta dx.
\end{aligned}$$

The second term represents the stabilization by Galerkin Least-Squares (GLS) method

$$\begin{aligned}
\text{Stablization}(u, v) = &\int_D \tau_1 L_1(u) \cdot D_u L_1(u)(v) dx \\
&+ \int_D \tau_2 (\nabla \cdot \boldsymbol{u})(\nabla \cdot \boldsymbol{v}) dx + \int_D \tau_3 (\boldsymbol{u} \cdot \nabla\gamma)(\boldsymbol{u} \cdot \nabla\eta) dx
\end{aligned}$$

where $L_1(u)$ represents the strong form of the momentum equation of (32), τ_1, τ_2 and τ_3 are properly chosen stabilization constants associated with the local Péclet number.

Optimal control for turbulent jet flow: model – weak form

The weak form is defined for $u = (\boldsymbol{u}, p, \gamma), v = (\boldsymbol{v}, q, \eta) \in \mathcal{V} := (H^1(D))^2 \times L^2(D) \times H^1(D)$

$$r(u, v, m, z) = \text{Model}(u, v, m) + \text{Stablization}(u, v) + \text{Nitsche}(u, v, m, z),$$

The third term represents the weak imposition of the boundary condition and the control function by Nitsche's method [Bazilevs et al., 2007], which reads

$$\begin{aligned}
\text{Nitsche}(u, v, m, z) &= C_d \int_{\Gamma_O \cup \Gamma_W} h^{-1}(\nu + \nu_t)(\boldsymbol{u} \cdot \boldsymbol{\tau})(\boldsymbol{v} \cdot \boldsymbol{\tau}) ds \\
&- \int_{\Gamma_O \cup \Gamma_W} (\sigma_n(\boldsymbol{u}) \cdot \boldsymbol{\tau})(\boldsymbol{v} \cdot \boldsymbol{\tau}) + (\sigma_n(\boldsymbol{v}) \cdot \boldsymbol{\tau})(\boldsymbol{u} \cdot \boldsymbol{\tau}) ds \\
&+ C_d \int_{\Gamma_I} h^{-1}(\nu + \nu_t)(\boldsymbol{u} \cdot \boldsymbol{n} + \chi_W \phi(z))(\boldsymbol{v} \cdot \boldsymbol{n}) ds \\
&- \int_{\Gamma_I} (\sigma_n(\boldsymbol{u}) \cdot \boldsymbol{n})(\boldsymbol{v} \cdot \boldsymbol{n}) + (\sigma_n(\boldsymbol{v}) \cdot \boldsymbol{n})(\boldsymbol{u} \cdot \boldsymbol{n} + \chi_W \phi(z)) ds,
\end{aligned}$$

where the first and the third terms enforce the boundary conditions while the second and the fourth terms represent the compatibility conditions. C_d is a constant, set as $C_d = 10^5$. h is a local length of the boundary edge along the Dirichlet boundaries.

Optimal control for turbulent jet flow: trace estimate

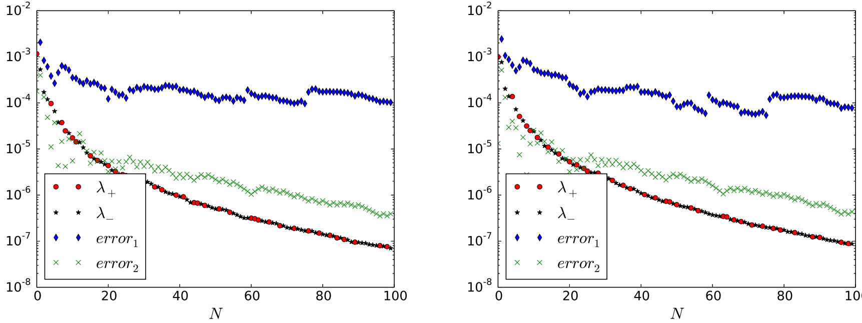


Figure: The decay of the generalized eigenvalues, λ_+ for positive eigenvalues and λ_- for negative eigenvalues, and the errors, denoted error_1 for the randomized trace estimator $\widehat{T}_1$ and error_2 for the randomized SVD-based trace estimator $\widehat{T}_2$, with respect to the number of estimate terms N. Left: $z = z_0$; right: $z = z^{\text{MC}}_{\text{quad}}$.

Optimal control for turbulent jet flow: optimal control

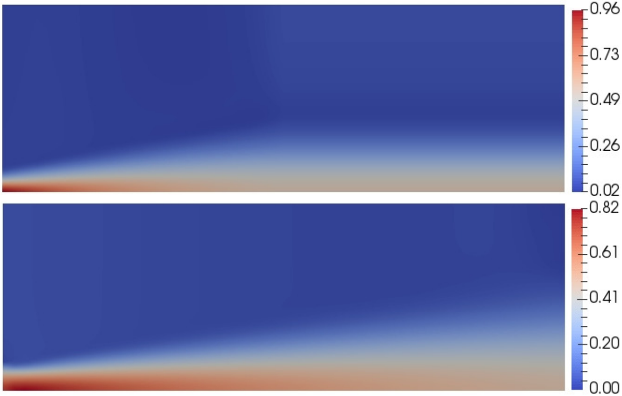


Figure: The velocity field corresponding to the initial control (top) and the optimal control with quadratic approximation and variance reduction (bottom).

Optimal control for turbulent jet flow: scalability I

- total cost = # PDE solves/iteration × # optimization iterations
- # PDE solves/iteration depends on decay of generalized eigenvalues

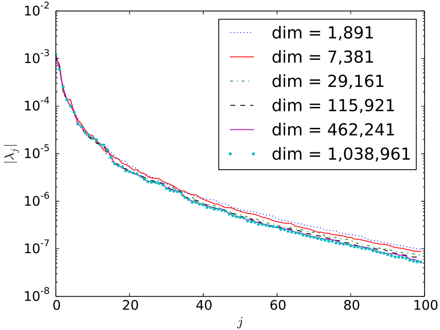


Figure: Decay of the generalized eigenvalues (in absolute value) with different parameter dimensions (dim) at optimal control. Results indicate dimension independence of per-optimization-iteration cost.

Optimal control for turbulent jet flow: scalability II

- total cost = # PDE solves/iteration × # optimization iterations
- # iterations depends how well the BFGS Hessian approximates true Hessian

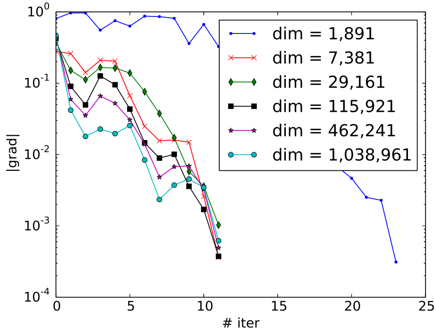


Figure: Gradient reduction with number of BFGS iterations. Simple control variable continuation employed. Results indicate dimension independence of BFGS iterations. (modulo initialization of 1M case)

Optimal control for turbulent jet flow: scalability III

How does variance reduction behave as parameter dimension increases?

Table: Mean-square-error (MSE) for $\mathbb{E}[Q]$ at optimal control.

parameter dimension	$\mathbb{E}^{\text{MC}}[-Q]$	MSE(Q)	MSE($Q - Q_{\text{lin}}$)	MSE($Q - Q_{\text{quad}}$)
1,891	-1.71e+00	7.40e-06	2.68e-08	1.81e-09
7,381	-1.59e+00	7.94e-06	1.57e-07	1.46e-08
29,161	-1.44e+00	3.82e-06	7.23e-08	1.66e-08

Table: MSE for Var$[Q]$, where $q = (Q - Q_0)^2$, $q_{\text{lin}} = (Q_{\text{lin}} - Q_0)^2$, $q_{\text{quad}} = (Q_{\text{quad}} - Q_0)^2$.

parameter dimension	$\mathbb{E}^{\text{MC}}[q]$	MSE(q)	MSE($q - q_{\text{lin}}$)	MSE($q - q_{\text{quad}}$)
1,891	8.05e-05	9.37e-10	1.76e-11	8.77e-13
7,381	8.13e-05	1.15e-09	8.87e-12	1.48e-12
29,161	5.60e-05	6.59e-10	3.39e-11	4.04e-12

100×–1000× speedup with quadratic approximation as control variate

Low-rank Hessian-based variance reduction for OUU

- Construct 2nd order Taylor approximation (wrt random parameters) of control objective, and use as a variance reduction tool for mean-variance OUU
- Hessian of parameter-to-objective map is compact, with fast decaying eigenvalues.
- Randomized SVD used to accurately and efficiently capture the low-rank
- Leads to an optimization problem constrained by a Hessian eigenvalue problem, with state and adjoint PDE constraints to define the gradient entering the objective approximation, and incremental state and adjoint PDE constraints to define the Hessian action
- Solved for sequence of OUU problems with up to 1 million random parameters, demonstrated scalability (i.e., # of PDE solves constant with increasing random parameter and control dimensions)
  - Trace estimation by randomized SVD is scalable
  - Quasi-Newton optimization iterations are scalable
  - Variance reduction is scalable
  - $\Longrightarrow$ Overall method is scalable
- Taylor approximation is local; variance reduction can deteriorate for large 3rd derivatives or large variances
- Current work: higher order Taylor and other approximations, chance constraints, alternatives to low-rank approximation, applications to deep learning of complex PDE maps

References

- P. Chen, U. Villa, and O. Ghattas. *Taylor approximation and variance reduction for PDE-constrained optimal control problems under uncertainty.* http://arxiv.org/abs/1804.04301
- A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations*, **SIAM/ASA Journal on Uncertainty Quantification**, 5(1):1166–1192, 2017. https://doi.org/10.1137/16M106306X
- A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems*, **SIAM Journal on Scientific Computing**, 38(1):A243–A272, 2016.
- T. Bui-Thanh and O. Ghattas, *Analysis of the Hessian for inverse scattering problems. Parts I, II, III*, **Inverse Problems** 2012a, 2012b; **Inverse Problems and Imaging** 2013.
- N. Alger, V. Rao, A. Myers, T. Bui-Thanh, and O. Ghattas, *Scalable matrix-free adaptive convolution-product approximation for locally translation-invariant operators*, working paper.