Preconditioning of a full-space trust-region SQP algorithm for PDE optimization

Denis Ridzal

Sandia National Laboratories

Workshop on PDE-constrained Optimization with Random Coefficients

Oberwolfach, Feb-1-2013

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
Research sponsored by the NNSA Advanced Scientific Computing Program under CSAR.

Sandia applications in large-scale and PDE-constrained optimization

Optimal design, control and inverse problems (PDECO):

- heat conduction, acoustics, nonlinear elasticity
- coupled thermal/fluid flow, coupled elastic/acoustic
- magnetohydrodynamics, semiconductors, shock-hydrodynamics

Optimization-based modeling (NLP):

- constrained mesh interpolation, feature-preserving interpolation
- optimization-based mesh remap of scalars, with box constrains, and mesh remap of vectors, with determinant constraints

And many, many more.

— Computational regime: Often $\mathcal{O}(10,000)$ compute nodes ...

— Would be nice to have a scalable, generic optimization solver ...

— Start from a generic formulation, derive a *matrix-free* algorithm, specialize linear algebra to the problems of interest.

Outline

- Matrix-free trust-region SQP for equality-constrained optimization
  - Condensed summary (R. 2006, Heinkenschloss/R. 2008/9/11)
- Preconditioning of optimality systems
  - 'Optimal' solvers of Rees, Dollar, Wathen (2010)
  - The idea in our context
- Numerical examples
  - Optimal design, control and inverse problems
- Hedging against uncertainty
  - Algorithmic infrastructure must support the UQ questions

Matrix-free trust-region SQP for equality-constrained optimization

Sequential quadratic programming

Solve equality-constrained optimization problem:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned}$$

where $f : \mathcal{X} \to \mathbb{R}$ and $c : \mathcal{X} \to \mathcal{C}$, for some Hilbert spaces $\mathcal{X}$ and $\mathcal{C}$, and f and c are twice continuously Fréchet differentiable.

Define Lagrangian functional $\mathcal{L} : \mathcal{X} \times \mathcal{C} \to \mathbb{R}$:

$$\mathcal{L}(x, \lambda) = f(x) + \langle \lambda, c(x) \rangle_{\mathcal{C}}$$

If *regular* point x_* is a local solution of the NLP, then there exists a $\lambda_* \in \mathcal{C}$ satisfying the *1st order necessary optimality conditions*:

$$\begin{aligned} \nabla_x f(x_*) + c_x(x_*)^* \lambda_* &= 0 \\ c(x_*) &= 0 \end{aligned}$$

Sequential quadratic programming

Newton's method applied to optimality conditions:

$$\begin{pmatrix} \nabla_{xx}\mathcal{L}(x_k,\lambda_k) & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} s \\ z \end{pmatrix} = - \begin{pmatrix} \nabla_x f(x_k) + c_x(x_k)^*\lambda_k \\ c(x_k) \end{pmatrix}$$

If $\nabla_{xx}\mathcal{L}(x_k,\lambda_k)$ is positive definite on the null space of $c_x(x_k)$, the above KKT system is necessary and sufficient for solving the QP:

$$\begin{aligned} \min_{s\in\mathcal{X}} \quad & \frac{1}{2}\langle \nabla_{xx}\mathcal{L}(x_k,\lambda_k)s, s\rangle_{\mathcal{X}} + \langle \nabla_x\mathcal{L}(x_k,\lambda_k), s\rangle_{\mathcal{X}} + \mathcal{L}(x_k,\lambda_k) \\ \text{s.t.} \quad & c_x(x_k)s + c(x_k) = 0 \end{aligned}$$

Sequential quadratic programming

Newton's method applied to optimality conditions:

$$\begin{pmatrix} \nabla_{xx}\mathcal{L}(x_k,\lambda_k) & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} s \\ z \end{pmatrix} = - \begin{pmatrix} \nabla_x f(x_k) + c_x(x_k)^*\lambda_k \\ c(x_k) \end{pmatrix}$$

If $\nabla_{xx}\mathcal{L}(x_k,\lambda_k)$ is positive definite on the null space of $c_x(x_k)$, the above KKT system is necessary and sufficient for solving the QP:

$$\begin{aligned} \min_{s\in\mathcal{X}} \quad & \frac{1}{2}\langle \nabla_{xx}\mathcal{L}(x_k,\lambda_k)s, s\rangle_{\mathcal{X}} + \langle \nabla_x\mathcal{L}(x_k,\lambda_k), s\rangle_{\mathcal{X}} + \mathcal{L}(x_k,\lambda_k) \\ \text{s.t.} \quad & c_x(x_k)s + c(x_k) = 0 \end{aligned}$$

Solve a sequence of *nonconvex* quadratic trust-region subproblems:

$$\begin{aligned} \min_{s\in\mathcal{X}} \quad & \frac{1}{2}\langle \nabla_{xx}\mathcal{L}(x_k,\lambda_k)s, s\rangle_{\mathcal{X}} + \langle \nabla_x\mathcal{L}(x_k,\lambda_k), s\rangle_{\mathcal{X}} + \mathcal{L}(x_k,\lambda_k) \\ \text{s.t.} \quad & c_x(x_k)s + c(x_k) = 0\,, \quad \|s\|_{\mathcal{X}} \le \Delta_k \end{aligned}$$

Possible incompatibility of constraints: composite-step approach.

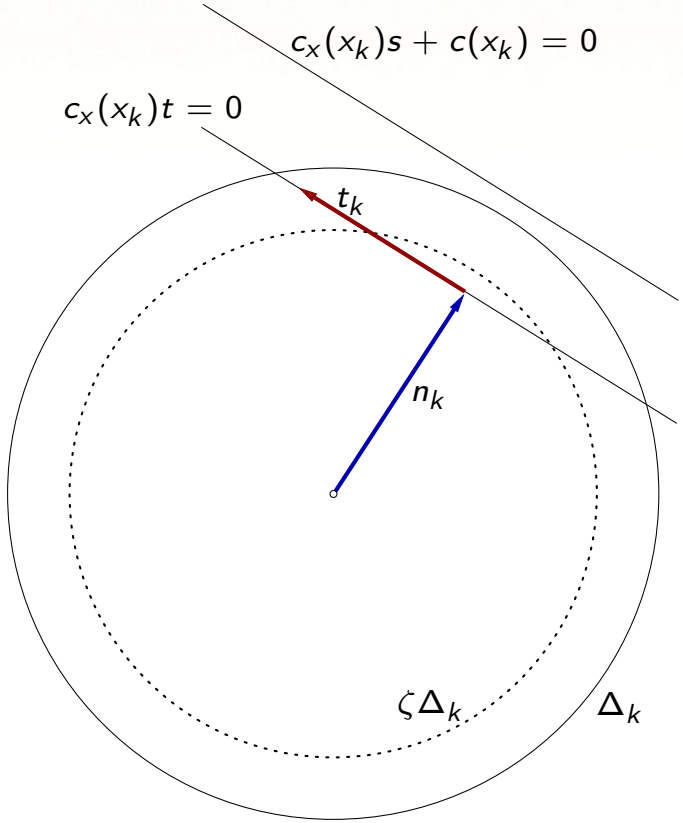
Composite-step approach for the solution of the trust-region subproblem

- **Trust-region step**:
  $s_k = n_k + t_k$

- **Quasi-normal step** n_k:
  reduces linear infeasibility

$$\begin{aligned} \min_{n \in \mathcal{X}} \quad & \|c_x(x_k)n + c(x_k)\|_{\mathcal{C}}^2 \\ \text{s.t.} \quad & \|n\|_{\mathcal{X}} \leq \zeta \Delta_k \end{aligned}$$

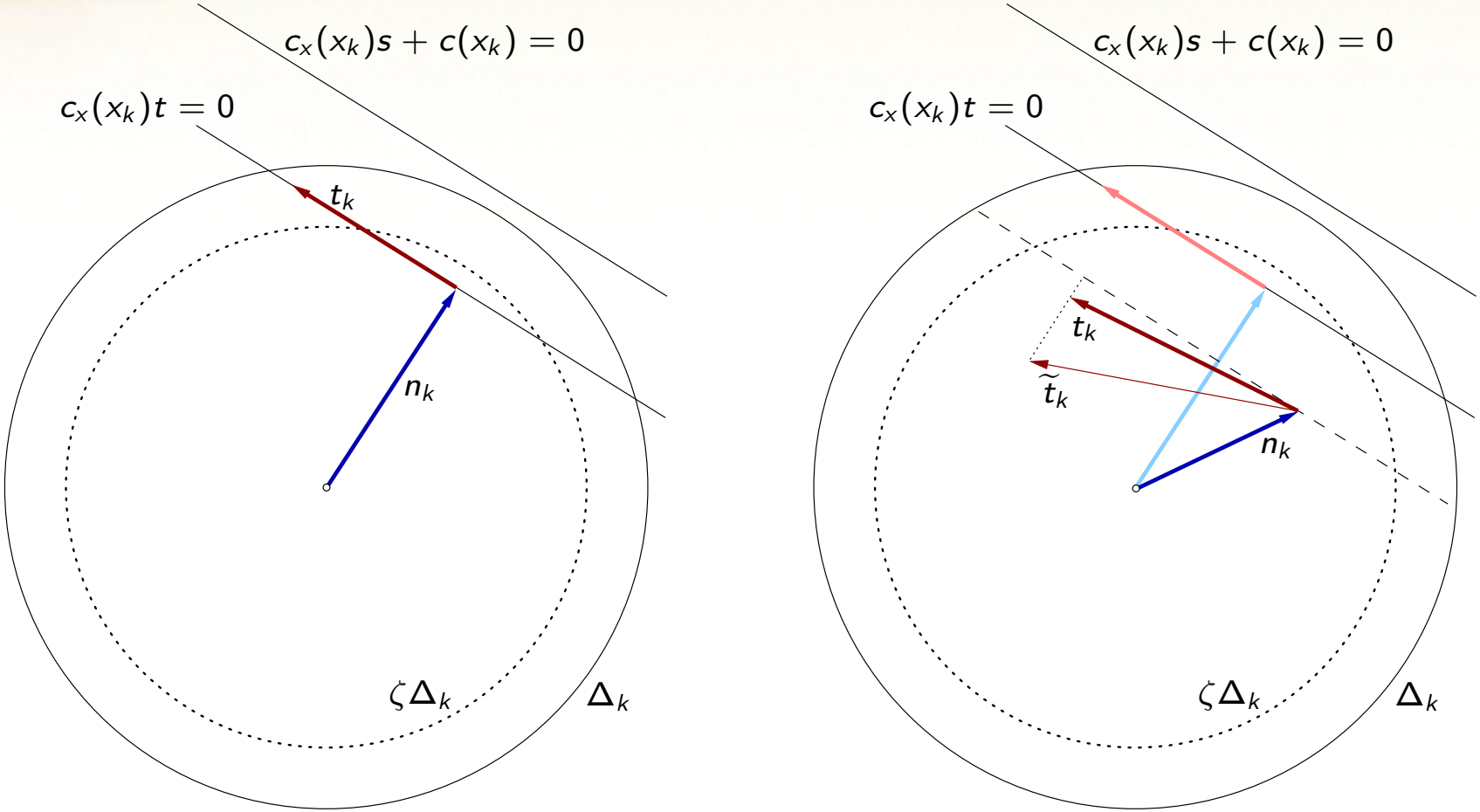
- **Tangential step** t_k:
  improves optimality while staying in the null space of the linearized constraints



$$\begin{aligned} \min_{t \in \mathcal{X}} \quad & \frac{1}{2}\langle \nabla_{xx}\mathcal{L}(x_k,\lambda_k)(t+n_k), t+n_k\rangle_{\mathcal{X}} + \langle \nabla_x \mathcal{L}(x_k,\lambda_k), t+n_k\rangle_{\mathcal{X}} + \mathcal{L}(x_k,\lambda_k) \\ \text{s.t.} \quad & c_x(x_k)t = 0\,, \quad \|t+n_k\|_{\mathcal{X}} \leq \Delta_k \end{aligned}$$

Omojokun (1989), Byrd, Hribar, Nocedal (1997), Dennis, El-Alem, Maciel (1997)

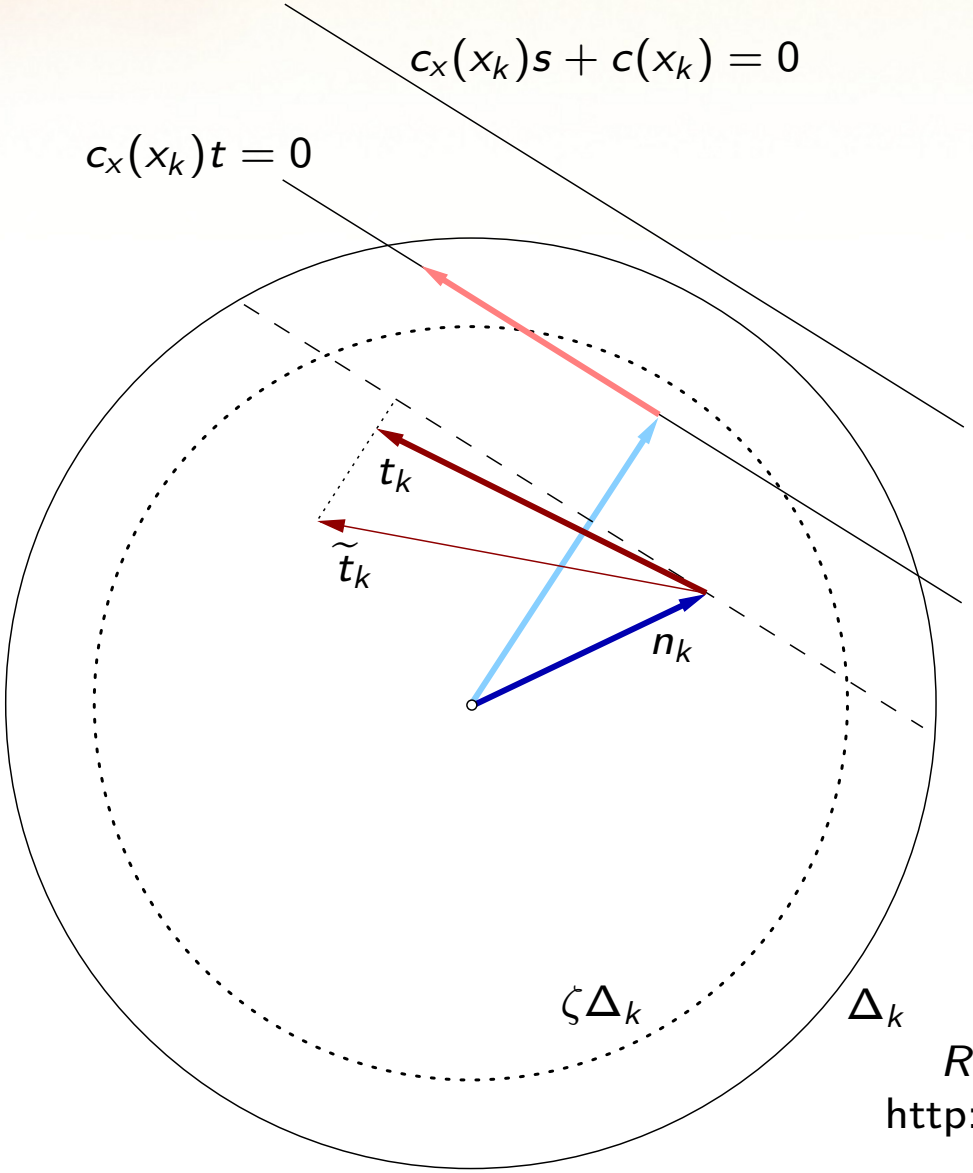
Composite step for matrix-free TR-SQP



With matrix-free linear algebra — such as iterative linear solvers — the quasi-normal and tangential steps are computed **inexactly**.

Special care must be taken to guarantee convergence!

Sketch of the matrix-free TR-SQP algorithm



Composite step:

$s_k = n_k + t_k$

1. Compute quasi-normal step n_k using **inexact Powell dogleg**.
2. Solve tangential subproblem for $\widetilde{t}_k$ using **inexact projected ST-CG**.
3. Restore linearized feasibility, yielding tangential step t_k.
4. Update Lagrange multipliers λ_{k+1}.
5. Evaluate progress.

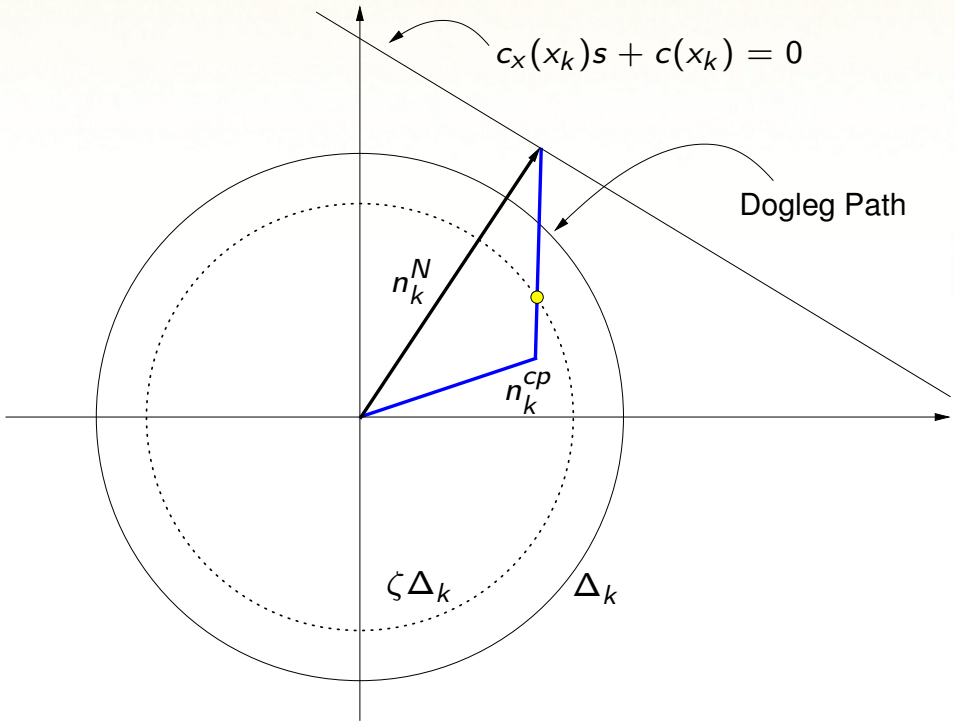
Ridzal, Aguiló, Heinkenschloss, SAND2011-9346
http://prod.sandia.gov/sand_doc/2011/119346.pdf

Inexact quasi-normal step

Cauchy step:

$$n_k^{cp} = \min_{\alpha \geq 0, \|n\|_{\mathcal{X}} \leq \zeta \Delta_k} \|c_x(x_k)n + c(x_k)\|_{\mathcal{C}}^2$$

$$\text{s.t.} \qquad n = -\alpha c_x(x_k)^* c(x_k)$$



Newton step:

n_k^N = minimum norm solution of

$$\min_n \quad \|c_x(x_k)n + c(x_k)\|_{\mathcal{C}}^2$$

The minimum norm solution n_k^N can be computed by solving:

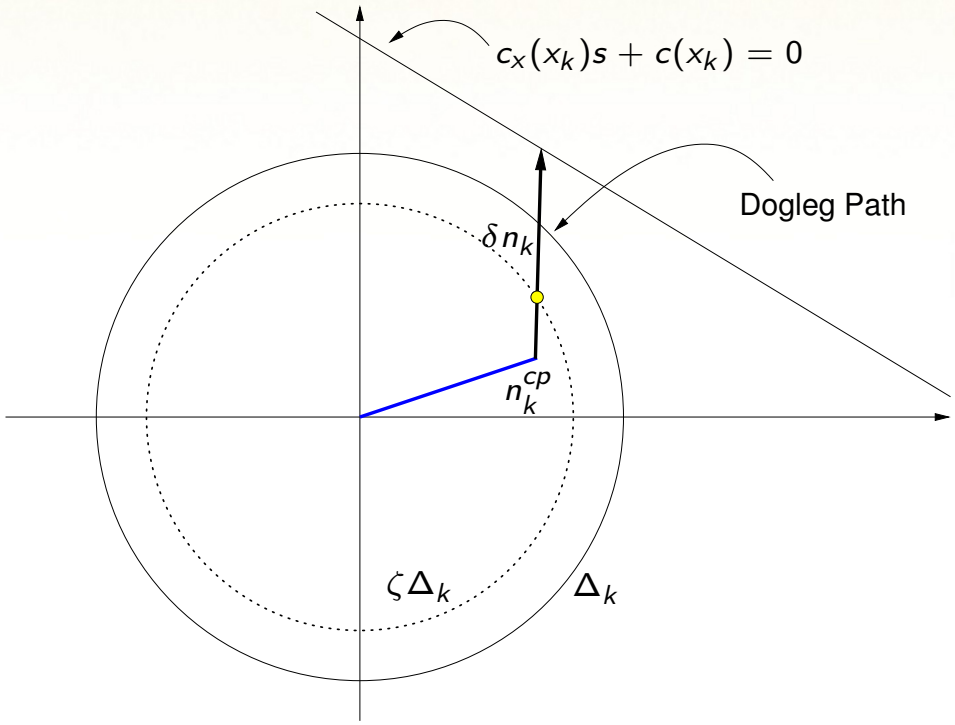
$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} n_k^N \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ -c(x_k) \end{pmatrix}.$$

Inexact quasi-normal step

Cauchy step:

$$n_k^{cp} = \min_{\alpha \geq 0, \|n\|_{\mathcal{X}} \leq \zeta \Delta_k} \|c_x(x_k)n + c(x_k)\|_{\mathcal{C}}^2$$

$$\text{s.t.} \qquad n = -\alpha c_x(x_k)^* c(x_k)$$



Newton step:

n_k^N = minimum norm solution of

$$\min_n \quad \|c_x(x_k)n + c(x_k)\|_{\mathcal{C}}^2$$

With inexactness, we solve for $\delta n_k = n_k^N - n_k^{cp}$:

$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} \delta n_k \\ y \end{pmatrix} = \begin{pmatrix} -n_k^{cp} + e^1 \\ -c_x(x_k)n_k^{cp} - c(x_k) + e^2 \end{pmatrix}.$$

For a fixed $0 \leq C^{qn} \leq 1$, the linear system residual must satisfy:

$$\|e^1\|_{\mathcal{X}}^2 + \|e^2\|_{\mathcal{C}}^2 \leq (C^{qn})^2 \|c_x(x_k)n_k^{cp} + c(x_k)\|_{\mathcal{C}}^2 .$$

Inexact tangential step

First, we solve a slightly reformulated tangential subproblem using an inexact projected Steihaug-Toint CG algorithm:

- by construction handles nonconvex trust-region subproblems,
- null-space projections are computed inexactly,
- the 3-term CG recurrence is lost, uses full orthogonalization,
- Steihaug-Toint stopping conditions are modified substantially.

Second, we perform another null-space projection, accurate enough to:

- restore linearized feasibility lost in the previous step,
- maintain progress in the tangential direction.

Throughout, we carefully control projection residuals.

Linear system 1
Quasi-normal step

Given a quasi-normal Cauchy point n_k^{cp}, we solve for $\delta n_k = n_k^N - n_k^{cp}$, where n_k^N is the desired Newton step:

$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} \delta n_k \\ y \end{pmatrix} = \begin{pmatrix} -n_k^{cp} \\ -c_x(x_k) n_k^{cp} - c(x_k) \end{pmatrix} + \begin{pmatrix} e^1 \\ e^2 \end{pmatrix}$$

Stopping condition for the residual $\begin{pmatrix} e^1 & e^2 \end{pmatrix} \in \mathcal{X} \times \mathcal{C}$:

$$\|e^1\|_{\mathcal{X}}^2 + \|e^2\|_{\mathcal{C}}^2 \leq (\xi^{qn})^2 \, \|c_x(x_k) n_k^{cp} + c(x_k)\|_{\mathcal{C}}^2 \,,$$

where $0 < \xi^{qn} \leq 1$.

Linear system 2
Tangential subproblem

At every CG iteration i, we compute an inexact projection $\widetilde{z}_i = \mathcal{W}_k(\widetilde{r}_i)$:

$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} \widetilde{z}_i \\ y \end{pmatrix} = \begin{pmatrix} \widetilde{r}_i \\ 0 \end{pmatrix} + \begin{pmatrix} e_i^1 \\ e_i^2 \end{pmatrix}$$

Stopping condition for the residual $\begin{pmatrix} e_i^1 & e_i^2 \end{pmatrix} \in \mathcal{X} \times \mathcal{C}$:

$$\|e_i^1\|_{\mathcal{X}} + \|e_i^2\|_{\mathcal{C}} \leq \xi^{proj} \min\left\{\|\widetilde{z}_i\|_{\mathcal{X}}, \|\widetilde{r}_i\|_{\mathcal{X}}\right\},$$

with $0 < \xi^{proj} \leq 1$.

Linear system 3
Restoring linearized feasibility

We perform another inexact null space projection,

$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} t_k \\ y \end{pmatrix} = \begin{pmatrix} \widetilde{t}_k \\ 0 \end{pmatrix} + \begin{pmatrix} e^1 \\ e^2 \end{pmatrix}$$

Stopping condition for the residual $\begin{pmatrix} e^1 & e^2 \end{pmatrix} \in \mathcal{X} \times \mathcal{C}$:

$$\|e^1\|_{\mathcal{X}} + \|e^2\|_{\mathcal{C}} \leq \Delta_k \min\left\{\Delta_k, \|n_k + t_k\|_{\mathcal{X}}, \xi^{tang}\|\widetilde{t}_k\|_{\mathcal{X}}/\Delta_k\right\},$$

for $0 < \xi^{tang} \leq 1$.

Linear system 4
Lagrange multipliers

Let $\widehat{x}_k = x_k + n_k + t_k$. We solve for $\Delta\lambda = \lambda_{k+1} - \lambda_k$:

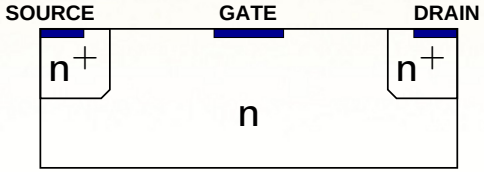
$$\begin{pmatrix} I & c_x(\widehat{x}_k)^* \\ c_x(\widehat{x}_k) & 0 \end{pmatrix} \begin{pmatrix} z \\ \Delta\lambda \end{pmatrix} = \begin{pmatrix} -\nabla_x f(\widehat{x}_k) - c_x(\widehat{x}_k)^*\lambda_k + e^1 \\ e^2 \end{pmatrix}$$

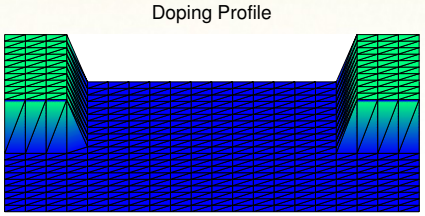
Stopping condition for the residual $\begin{pmatrix} e^1 & e^2 \end{pmatrix} \in \mathcal{X} \times \mathcal{C}$:

$$\|e^1\|_{\mathcal{X}} + \|e^2\|_{\mathcal{C}} \leq \min\left\{\xi^{lmg}, \xi^{lmh}\|\nabla_x f(\widehat{x}_k) + c_x(\widehat{x}_k)^*\lambda_k\|_{\mathcal{X}}\right\},$$

for $0 < \xi^{lmh} \leq 1$ and a fixed $\xi^{lmg} > 0$ independent of k.

Distributed control of Poisson-Boltzmann





Match the current measured at an Ohmic contact Γ_o of a semiconductor device to a prescribed value, by modifying a reference doping profile $\widehat{g}$. We consider a simplified model problem:

$$\min_{u,g} \quad \frac{\alpha}{2}\int_{\Gamma_o}(\nabla u\cdot\nu - \nabla\widehat{u}\cdot\nu)^2\,ds + \frac{\beta}{2}\int_{\Omega}(g-\widehat{g})^2\,dx + \frac{\gamma}{2}\int_{\Omega}|\nabla g|^2\,dx$$

subject to

$$\begin{aligned} -\Delta u &= \exp(u) - \exp(-u) - g && \text{in } \Omega, \\ u &= 0 && \text{on } \Gamma_D, \\ \nabla u\cdot\nu &= 0 && \text{on } \Gamma_N, \end{aligned}$$

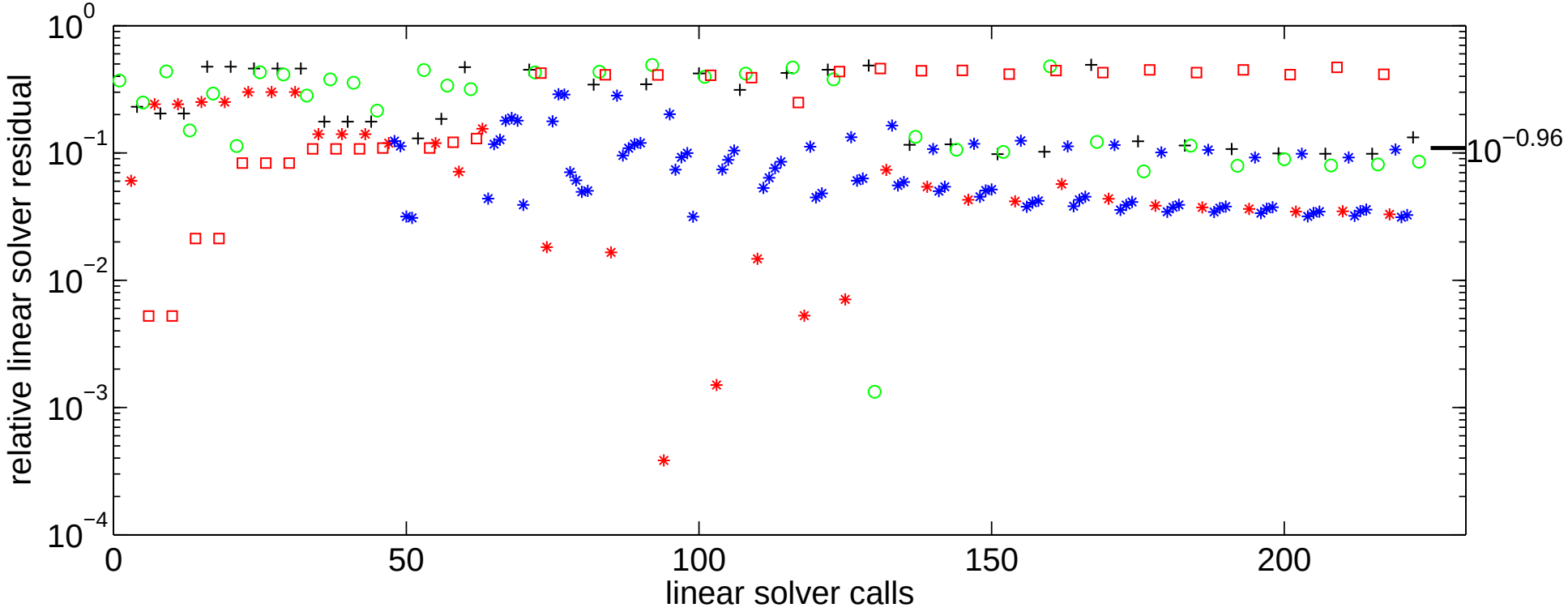
where $\alpha, \beta, \gamma > 0$, u is the potential, and g is the doping. We use a finite element model with 4096 triangles and an ILU KKT preconditioner.

INEXACT ALGORITHM

Nominal Tol tol^{LS}	**0.5**	**10^{-1}**	**10^{-2}**	**10^{-3}**	**10^{-4}**	**10^{-5}**	**10^{-6}**	**10^{-7}**	**10^{-8}**
(S)uccess/(F)ailure	S	S	S	S	S	S	S	S	S
GMRES iters	381	380	514	671	844	1032	1218	1408	1601
GMRES calls	223	136	120	120	120	120	120	120	120
GMRES iters/call	1.7	2.8	4.3	5.6	7.0	8.6	10.2	11.7	13.3
GMRES prec builds	42	29	24	24	24	24	24	24	24
STCG iters	117	68	59	59	59	59	59	59	59
SQP iters	34	21	18	18	18	18	18	18	18

'CONVENTIONAL' ALGORITHM

Nominal Tol tol^{LS}	**0.5**	**10^{-1}**	**10^{-2}**	**10^{-3}**	**10^{-4}**	**10^{-5}**	**10^{-6}**	**10^{-7}**	**10^{-8}**
(S)uccess/(F)ailure	F	F	F	F	F	F	F	F	F
GMRES iters	—	—	—	—	—	—	—	—	—
SQP iters	—	—	—	—	—	—	—	—	—



- □ red squares denote quasi-normal step computations
- * red asterisks correspond to the first STCG iterations
- * blue asterisks denote the subsequent STCG iterations
- + black crosses correspond to tangential step computations
- ○ green circles denote Lagrange multiplier computations

Preconditioning of optimality systems

The linear system to solve

$$\begin{pmatrix} I & c_x(x_k)^* \\ c_x(x_k) & 0 \end{pmatrix} \begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \end{pmatrix}$$

'AUGMENTED SYSTEM'

- These are KKT systems for the **convex quadratic programs**

$$\begin{aligned} \min \quad & \frac{1}{2}\langle z, z\rangle - \langle b^1, z\rangle \\ \text{s.t.} \quad & c_x(x_k)z = b^2. \end{aligned}$$

- True even if the trust-region subproblems

$$\begin{aligned} \min \quad & \frac{1}{2}\langle \nabla_{xx}\mathcal{L}(x_k,\lambda_k)s, s\rangle_{\mathcal{X}} + \langle \nabla_x \mathcal{L}(x_k,\lambda_k), s\rangle_{\mathcal{X}} + \mathcal{L}(x_k,\lambda_k) \\ \text{s.t.} \quad & c_x(x_k)s + c(x_k) = 0, \quad \|s\|_{\mathcal{X}} \le \Delta_k \end{aligned}$$

are *not convex* !

The augmented system in PDE optimization

- Split variables x into **state variables** u and **control variables** g, i.e.:

$$x_k = (u_k, g_k)$$

- Write augmented system matrices as 3×3 block matrices

$$\begin{pmatrix} I & 0 & c_u(x_k)^* \\ 0 & I & c_g(x_k)^* \\ c_u(x_k) & c_g(x_k) & 0 \end{pmatrix}$$

- Compress notation:

$$\begin{pmatrix} I & 0 & C_u^T \\ 0 & I & C_g^T \\ C_u & C_g & 0 \end{pmatrix}$$

Optimal solvers for PDE optimization

- Rees, Dollar, Wathen: SIAM J. Sci. Comp., 32 (2010), pp. 271-298.
- Motivated by the **3-iteration preconditioner** for saddle-point systems (Murphy, Golub, Wathen).

$$\begin{aligned} \min \quad & \frac{1}{2} u^T M u + \frac{\beta}{2} g^T M g - u^T b_u^1 \\ \text{s.t.} \quad & C_u u + C_g g = b^2 \end{aligned}$$

$$\mathcal{K} = \begin{pmatrix} M & 0 & C_u^T \\ 0 & \beta M & C_g^T \\ C_u & C_g & 0 \end{pmatrix} \quad \mathcal{P}_{\mathcal{K}}^* = \begin{pmatrix} M^{-1} & 0 & 0 \\ 0 & \frac{1}{\beta} M^{-1} & 0 \\ 0 & 0 & (C_u M^{-1} C_u^T + \frac{1}{\beta} C_g M^{-1} C_g^T)^{-1} \end{pmatrix}$$

Optimal solvers for PDE optimization

- Rees, Dollar, Wathen: SIAM J. Sci. Comp., 32 (2010), pp. 271-298.
- Motivated by the **3-iteration preconditioner** for saddle-point systems (Murphy, Golub, Wathen).

$$\begin{aligned} \min \quad & \frac{1}{2} u^T M u + \frac{\beta}{2} g^T M g - u^T b_u^1 \\ \text{s.t.} \quad & C_u u + C_g g = b^2 \end{aligned}$$

$$\mathcal{K} = \begin{pmatrix} M & 0 & C_u^T \\ 0 & \beta M & C_g^T \\ C_u & C_g & 0 \end{pmatrix} \qquad \mathcal{P}_{\mathcal{K}} = \begin{pmatrix} M^{-1} & 0 & 0 \\ 0 & \frac{1}{\beta} M^{-1} & 0 \\ 0 & 0 & C_u^{-T} M C_u^{-1} \end{pmatrix}$$

Eigenvalue estimates

Three clusters:

$$\lambda = 1$$

$$\frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+\frac{2\alpha_2}{\beta}}\right)$$

$$\frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_2}{\beta}}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+\frac{2\alpha_1 h^4}{\beta}}\right)$$

where $\alpha_1, \alpha_2 \ll 1$ are independent of h and β.

Setup:

- distributed control of the Poisson equation
- quadrilateral grid with linear elements (Q1)
- distributed measurements

The approximate Schur preconditioner for augmented systems

$$\mathcal{A} = \begin{pmatrix} I & 0 & C_u^T \\ 0 & I & C_g^T \\ C_u & C_g & 0 \end{pmatrix} \qquad \mathcal{P}_{\mathcal{A}}^* = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & (C_u C_u^T + C_g C_g^T)^{-1} \end{pmatrix}$$

- Independent of the objective function.
- Linearized 'forward' solve followed by an 'adjoint' solve.
- No second-order derivative information needed.
- Setup cost can be amortized within each optimization iteration.
- Of course, no free lunch: The objective function is handled through the inexact projected Steihaug-Toint CG iteration.

The approximate Schur preconditioner for augmented systems

$$\mathcal{A} = \begin{pmatrix} I & 0 & C_u^T \\ 0 & I & C_g^T \\ C_u & C_g & 0 \end{pmatrix} \qquad \mathcal{P}_{\mathcal{A}} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & C_u^{-T} C_u^{-1} \end{pmatrix}$$

- Independent of the objective function.
- Linearized 'forward' solve followed by an 'adjoint' solve.
- No second-order derivative information needed.
- Setup cost can be amortized within each optimization iteration.
- Of course, no free lunch: The objective function is handled through the inexact projected Steihaug-Toint CG iteration.

Eigenvalue estimates

Three clusters:

$$\lambda = 1$$

$$\frac{1}{2}\left(1+\sqrt{5+2\alpha_1 h^4}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+2\alpha_2}\right)$$

$$\frac{1}{2}\left(1-\sqrt{5+2\alpha_2}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+2\alpha_1 h^4}\right)$$

where $\alpha_1, \alpha_2 \ll 1$ are independent of h and β.

Setup:

- distributed control of the Poisson equation
- quadrilateral grid with linear elements (Q1)
- distributed measurements

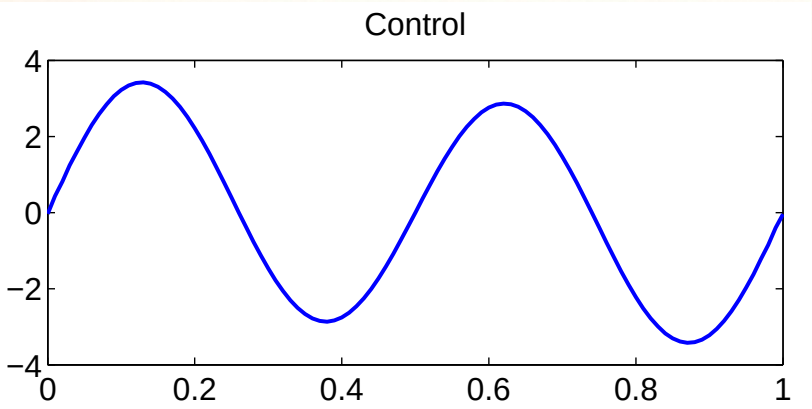
Numerical Results

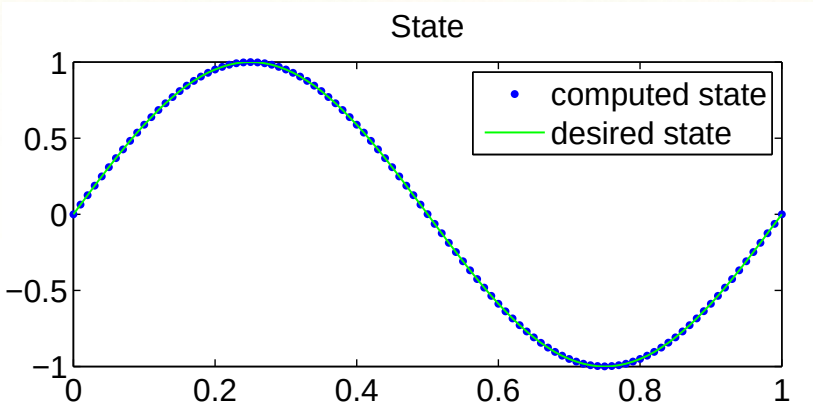
- Distributed control of the Burgers equation.
- Boundary control of the Navier-Stokes equations.
- Source inversion subject to advection-dominated diffusion.
- Wave focusing: Distributed control of the Helmholtz equation.
- Plane wave matching: Partial distributed control of Helmholtz.
- Parameter identification (inversion) in the Poisson equation.
- Parameter identification (inversion) in nonlinear elasticity.

Notes:

- In all experiments, we set a *nominal tolerance* of $\mathbf{10^{-2}}$ for all linear system solves. Our SQP algorithm dynamically adjusts this tolerance, as needed for convergence.
- We use sparse direct solvers when applying $\mathcal{P}_{\mathcal{A}}$, however, where applicable, a single V cycle of multigrid suffices.

Burgers: Distributed Control





$$\underset{u,g}{\text{Minimize}} \quad \frac{1}{2}\int_\Omega (u-\widehat{u})^2\,dx + \frac{\alpha}{2}\int_\Omega g^2\,dx$$

subject to

$$\begin{aligned} -\nu u_{xx} + u\,u_x &= g && \text{in} \quad \Omega \\ u &= 0 && \text{on} \quad \partial\Omega\,. \end{aligned}$$

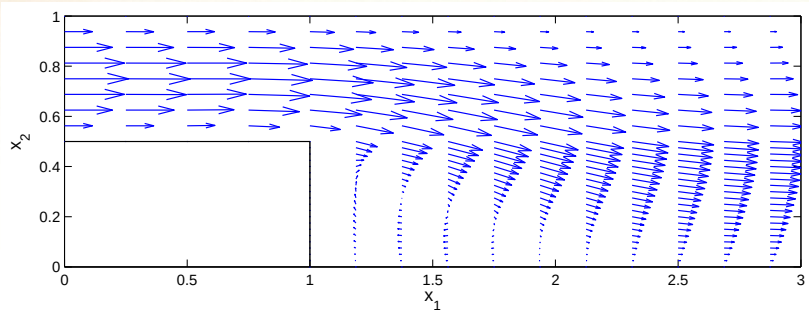
Here u and g are the *state* and *control* variables, respectively, $\alpha = 10^{-5}$ is a control penalty parameter, $\nu = 10^{-2}$ is the viscosity parameter, and $\widehat{u} = \widehat{u}(x) = \sin(2\pi x)$ is the desired state.

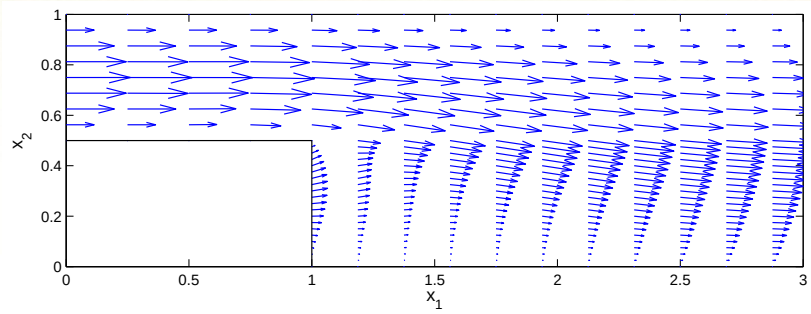
Mesh Size	Avg. GMRES Iters
100	7.9
200	8.2
400	8.1

```
 k    f(x_k)      || c(x_k) || ||L_x(x_k,lam_k)||  Delta_k   || n_k || || t_k || itcg
 ...
 1  1.80267e-01  1.044878e+00    5.137360e-02     6.25e+00  1.00e+01  1.25e+01     2
 2  5.81544e-02  5.013597e-01    1.833405e-02     6.25e+00  5.00e+00  6.25e+00     1
 3  2.80816e-02  2.171727e-01    7.333203e-03     5.60e+01  5.00e+00  6.25e+00     2
 3  2.80816e-02  2.171727e-01    7.333203e-03     2.80e+01  3.40e+00  5.60e+01     8
 4  1.21109e-02  1.445566e-01    3.595189e-03     2.80e+01  3.40e+00  2.80e+01     8
 5  2.47416e-03  1.077326e-01    1.041845e-03     2.80e+01  5.67e-01  2.80e+01     6
 6  3.57513e-04  1.720253e-02    1.533407e-04     1.36e+02  1.10e+00  1.94e+01    12
 7  2.80732e-05  5.013467e-03    2.019718e-05     1.36e+02  1.28e-01  8.08e+00    12
 8  2.51331e-05  5.029674e-05    6.713842e-07     1.36e+02  3.77e-02  8.55e-01    12
 9  2.50114e-05  1.476124e-06    6.052325e-09     1.36e+02  2.41e-04  1.30e+00    34
10  2.50113e-05  2.947095e-10    1.252652e-10     1.36e+02  4.90e-06  4.91e-02    35

                                CG iters:  137
```

Navier-Stokes: Boundary Control





$$\underset{\mathbf{u},p,\mathbf{g}}{\text{Minimize}} \quad \frac{1}{2}\int_D (\partial_{x_1}\mathbf{u}_2 - \partial_{x_2}\mathbf{u}_1)^2\,dx + \frac{\alpha}{2}\int_{\Gamma_c} |\mathbf{g}|^2\,ds,$$

subject to

$$\begin{aligned}
-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p &= \mathbf{0} && \text{in} && \Omega,\\
\nabla\cdot\mathbf{u} &= 0 && \text{in} && \Omega,\\
(\nu\nabla\mathbf{u} - pI)\mathbf{n} &= \mathbf{0} && \text{on} && \Gamma_{out},\\
(\nu\nabla\mathbf{u} - pI)\mathbf{n} + \frac{1}{\delta}\mathbf{u} &= \frac{1}{\delta}\mathbf{g} && \text{on} && \Gamma_c,\\
\mathbf{u} &= \mathbf{b} && \text{on} && \Gamma_{in},\\
\mathbf{u} &= \mathbf{0} && \text{on} && \partial\Omega \setminus (\Gamma_{in}\cup\Gamma_c\cup\Gamma_{out}),
\end{aligned}$$

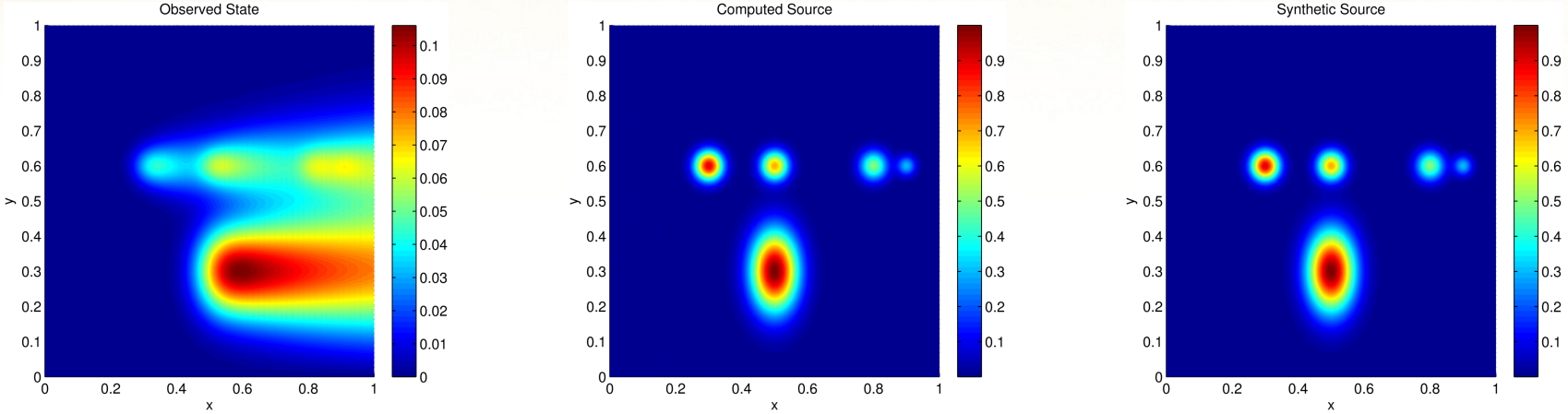
where $\nu = 1/200$ and $\alpha = 10^{-1}$.

Mesh Size	Avg. GMRES Iters
352	14.5
1408	17.4
5632	21.4

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	0.00000e+00	7.531966e-02	0.000000e+00	1.00e+02			
1	1.12909e-01	2.389013e-02	7.184773e-02	1.00e+02	3.27e+00	2.57e+00	17
2	3.46935e-01	7.848287e-03	1.090567e-01	1.00e+02	1.47e+00	1.73e+00	19
3	5.20101e-01	1.187366e-02	1.170088e-01	1.00e+02	6.87e-01	3.06e+00	19
4	2.57485e-01	3.965730e-03	6.613416e-02	1.00e+02	1.29e+00	9.33e-01	16
5	1.85397e-01	8.340092e-04	3.119255e-02	1.00e+02	5.10e-01	4.60e-01	17
6	1.81074e-01	1.225354e-04	2.118968e-03	1.00e+02	8.07e-02	2.29e-01	16
7	1.80438e-01	3.670772e-07	1.638772e-05	1.00e+02	1.17e-02	1.67e-02	19
8	1.80439e-01	7.554073e-12	2.883976e-10	1.00e+02	4.91e-05	5.92e-03	20

CG iters: 143

Advection-Dominated Diffusion: Source Inversion



$$\underset{u,\, g}{\text{Minimize}} \quad \frac{1}{2}\int_\Omega (u-\widehat{u})^2\, dx + \frac{\alpha}{2}\int_\Omega g^2\, dx$$

subject to

$$\begin{aligned} -\kappa\, \Delta u + \mathbf{V}\cdot\nabla u &= g && x\in\Omega \\ u &= d && x\in\partial\Omega_D \\ \nabla u\cdot\mathbf{n} &= 0 && x\in\partial\Omega_N, \end{aligned}$$

where u is the simulated concentration, $\widehat{u}$ is the measured concentration, g is the source and $\alpha = 10^{-6}$.

$\kappa = 10^{-2}$	
Mesh Size	**Avg. GMRES Iters**
64×64	10.6
128×128	12.0
256×256	12.5

128×128	
Diffusion	**Avg. GMRES Iters**
$\kappa = 10^{-2}$	12.0
$\kappa = 10^{-3}$	12.6
$\kappa = 10^{-4}$	13.0

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	1.63768e-02	4.448427e-02	1.194189e-03	1.00e+02			
1	1.21839e-06	7.301545e-13	1.679715e-06	1.00e+02	1.13e+01	7.74e+00	9
2	8.96338e-09	1.212529e-17	1.664197e-08	1.00e+02	1.73e-11	2.87e+00	30
3	8.37210e-09	1.183593e-17	1.322777e-10	1.00e+02	6.78e-16	2.06e-01	61
4	8.37188e-09	1.216858e-17	1.261303e-12	1.00e+02	9.55e-16	1.27e-02	114

CG iters: 214 / 179 / 171

$\kappa = 10^{-2}$

Mesh Size	Avg. GMRES Iters
64×64	10.6
128×128	12.0
256×256	12.5

128×128

Diffusion	Avg. GMRES Iters
$\kappa = 10^{-2}$	12.0
$\kappa = 10^{-3}$	12.6
$\kappa = 10^{-4}$	13.0

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	5.49479e-04	0.000000e+00	1.218077e-04	1.00e+02			
1	1.24653e-06	1.889448e-17	8.893031e-07	1.08e+02	0.00e+00	1.54e+01	9
2	8.97088e-09	1.864353e-17	8.181410e-09	1.08e+02	1.98e-15	5.62e+00	29
3	8.49700e-09	1.887460e-17	7.424629e-11	1.08e+02	2.10e-15	3.17e-01	53
4	8.49682e-09	1.957668e-17	6.415283e-13	1.08e+02	2.03e-15	1.57e-02	88

CG iters: 214 / 179 / 171

$\kappa = 10^{-2}$

Mesh Size	Avg. GMRES Iters
64×64	10.6
128×128	12.0
256×256	12.5

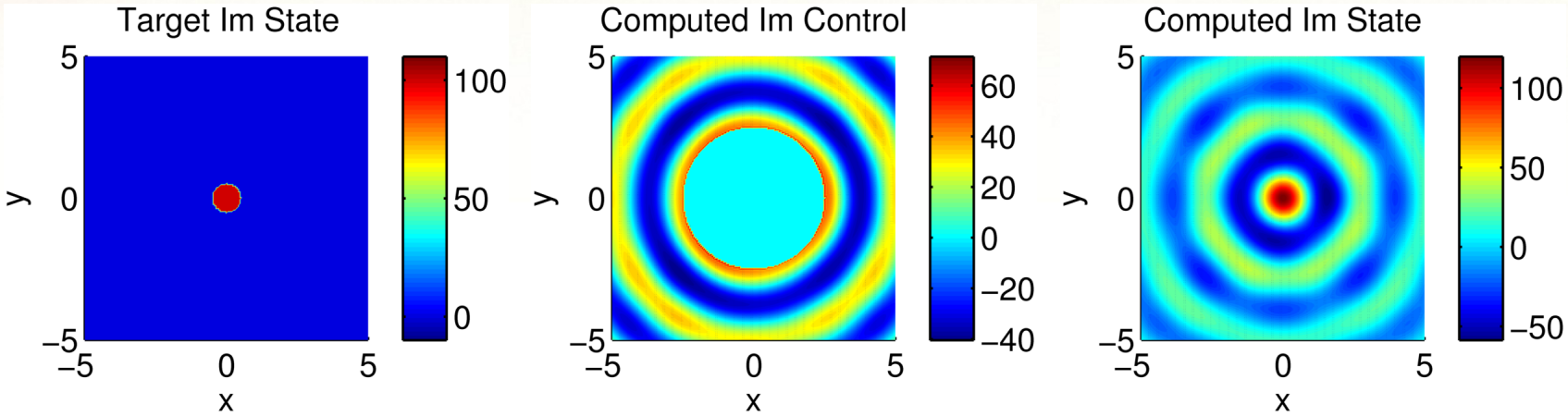
128×128

Diffusion	Avg. GMRES Iters
$\kappa = 10^{-2}$	12.0
$\kappa = 10^{-3}$	12.6
$\kappa = 10^{-4}$	13.0

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	5.47943e-04	0.000000e+00	6.072364e-05	1.00e+02			
1	1.25268e-06	3.480946e-17	4.469691e-07	2.16e+02	0.00e+00	3.08e+01	9
2	8.92941e-09	3.474085e-17	3.788849e-09	2.16e+02	8.90e-15	1.12e+01	29
3	8.52932e-09	3.484492e-17	3.288983e-11	2.16e+02	9.19e-15	5.62e-01	51
4	8.52918e-09	3.479454e-17	3.116473e-13	2.16e+02	9.47e-15	2.52e-02	82

CG iters: 214 / 179 / 171

Helmholtz: Wave Focusing



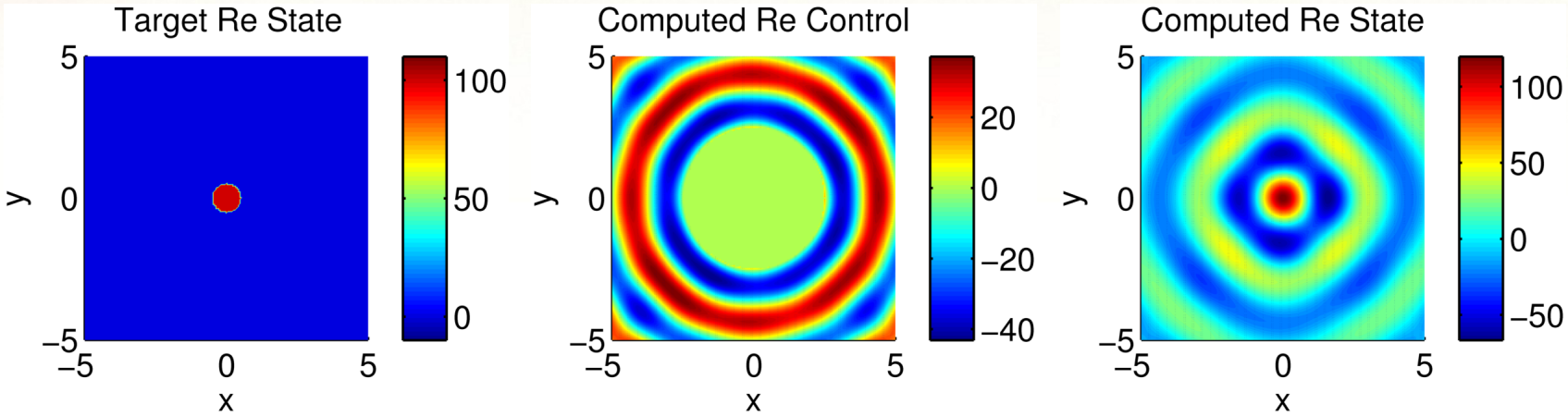
$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-4}$.

Helmholtz: Wave Focusing



$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

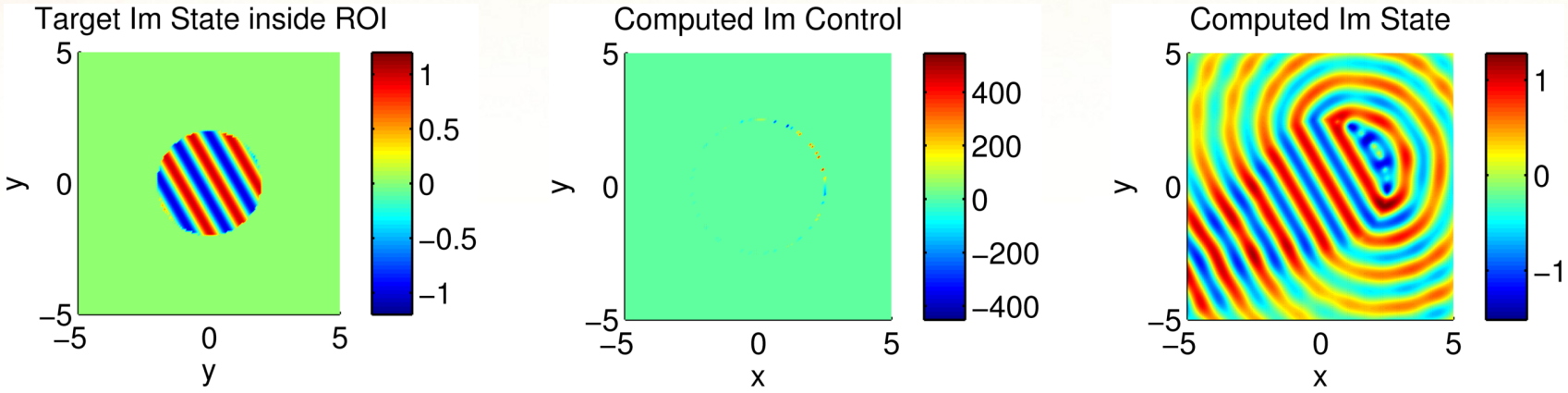
where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-4}$.

Mesh Size	Freq / Wave #	Avg. GMRES Iters
80×80	225 Hz / 8	7.9
160×160	450 Hz / 16	5.8
320×320	900 Hz / 32	6.5

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	7.50834e+03	8.926706e-01	1.648215e+00	1.00e+02			
1	7.35619e+03	3.536454e-06	1.631174e+00	7.92e+02	5.30e+01	1.00e+02	1
2	6.12139e+03	1.786289e-08	1.485670e+00	5.55e+03	1.10e-04	7.92e+02	1
3	7.05854e+02	1.348830e-08	4.671433e-01	3.88e+04	7.97e-08	5.55e+03	1
4	1.11686e+02	2.188792e-07	7.493318e-04	3.88e+04	1.74e-08	2.54e+03	1
5	1.11260e+02	5.464315e-08	7.121193e-06	3.88e+04	2.91e-07	2.16e+03	10
6	1.11260e+02	1.131356e-09	9.927443e-08	3.88e+04	4.28e-08	7.47e+01	17
7	1.11260e+02	4.597865e-11	1.933821e-09	3.88e+04	1.02e-09	7.63e-01	18

CG iters: 49

Helmholtz: Plane Wave Matching



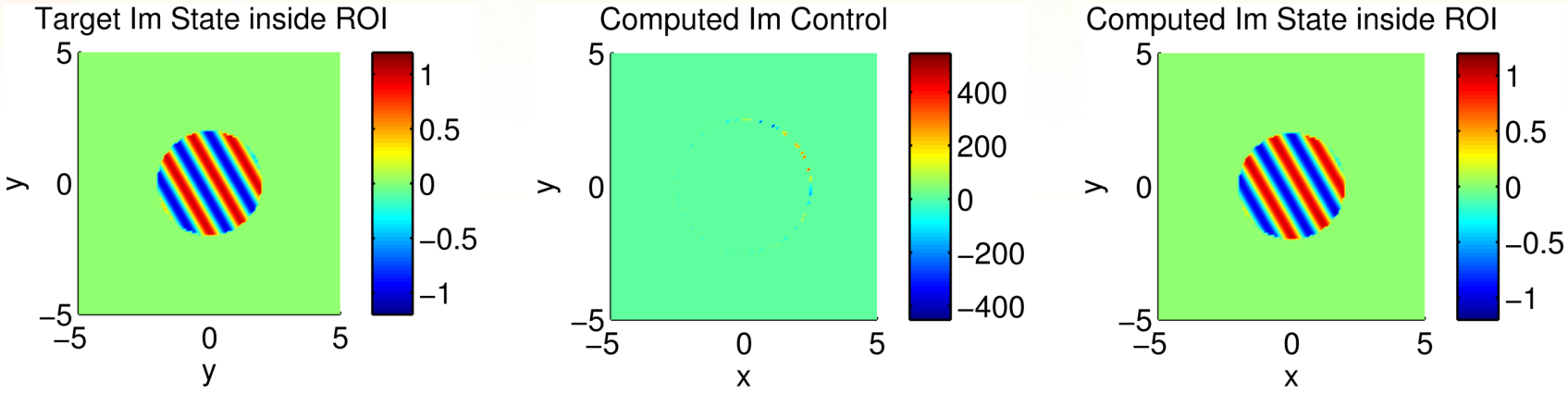
$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-5}$.

Helmholtz: Plane Wave Matching



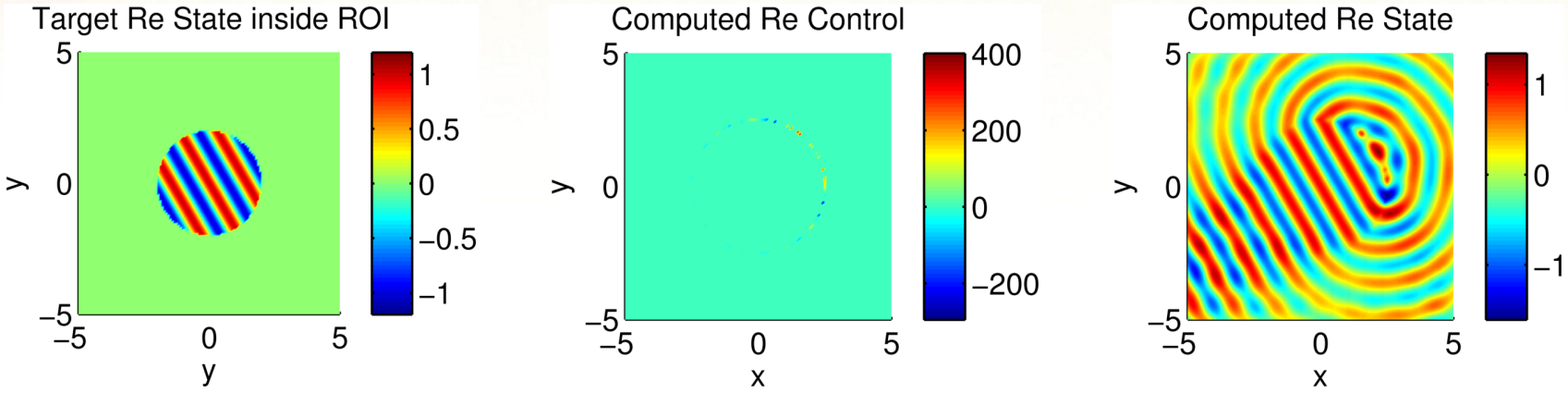
$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-5}$.

Helmholtz: Plane Wave Matching



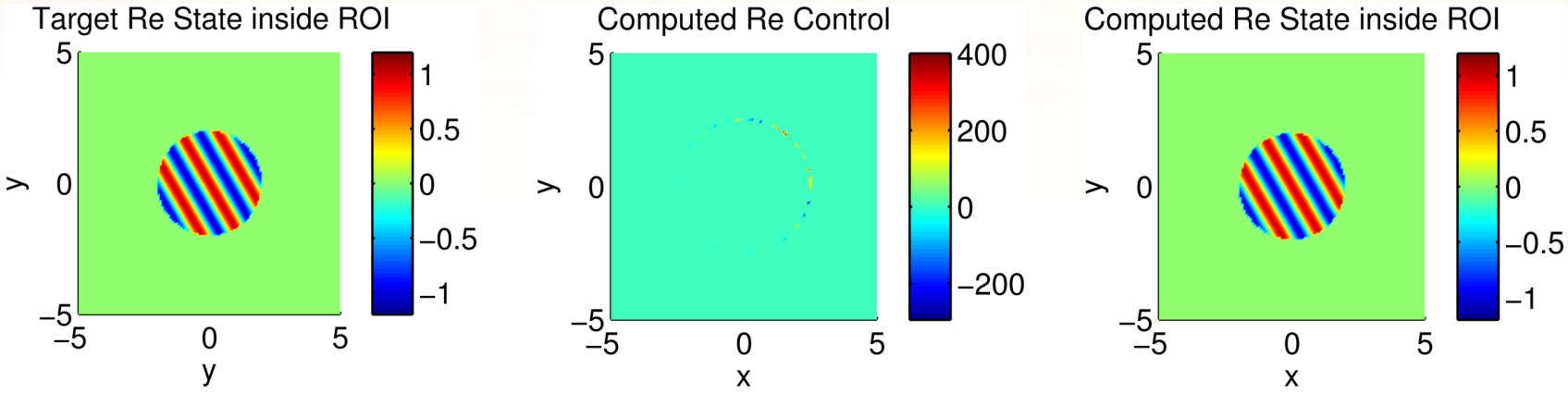
$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-5}$.

Helmholtz: Plane Wave Matching



$$\underset{p,g}{\text{Minimize}} \quad \frac{1}{2}\int_{\widehat{\Omega}} (p-\widehat{p})\overline{(p-\widehat{p})}\,dx + \frac{\alpha}{2}\int_{\Omega_c} g\overline{g}\,dx$$

subject to

$$\begin{aligned} -\Delta p - k^2 p &= g && \text{in} \quad \Omega, \\ \frac{\partial p}{\partial n} &= -ikp && \text{on} \quad \partial\Omega, \end{aligned}$$

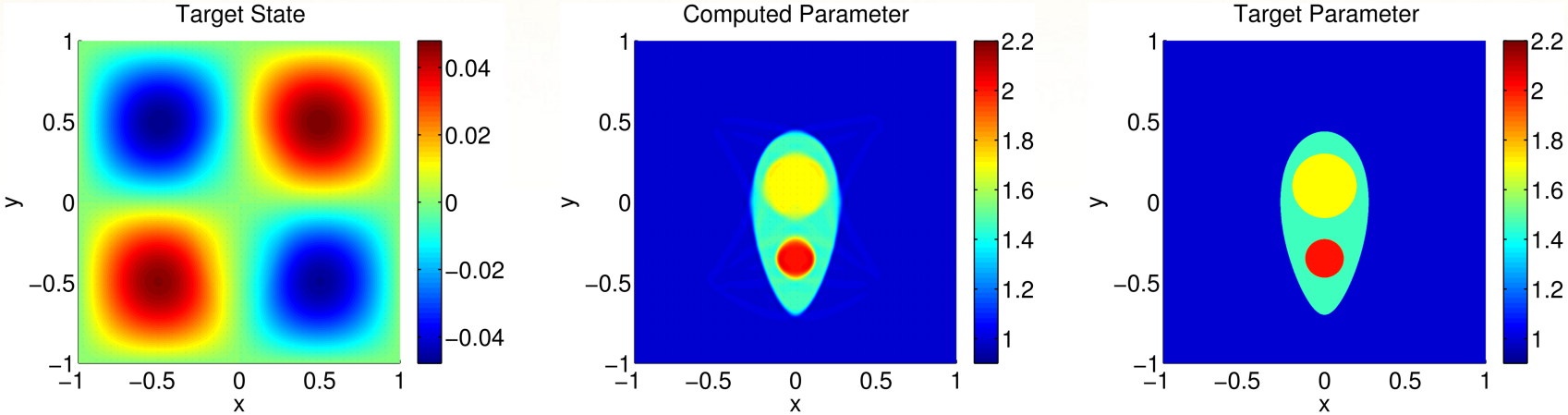
where p is the pressure, g is the volume force, k is the wave number, $\widehat{p}$ is the desired pressure and $\alpha = 10^{-5}$.

Mesh Size	Freq / Wave #	Avg. GMRES Iters
64×64	112 Hz / 4	9.1
128×128	225 Hz / 8	7.7
256×256	450 Hz / 16	6.4

```
k    f(x_k)      || c(x_k) || ||L_x(x_k,lam_k)||  Delta_k   || n_k || || t_k || itcg
0  6.02194e+00  2.989234e-02    1.810687e-03     1.00e+05
1  3.14125e-01  1.293450e-08    1.462836e-05     1.00e+05  5.28e-01  6.61e+03     7
2  3.12119e-01  5.829430e-11    1.453873e-07     1.00e+05  9.15e-08  4.78e+02    18
3  3.12118e-01  2.771248e-12    1.071073e-09     1.00e+05  1.56e-10  8.53e+00    20
4  3.12118e-01  9.716601e-14    8.239398e-12     1.00e+05  2.60e-12  5.99e-02    21

CG iters:  66
```

Poisson: Parameter Identification



$$\underset{\{u,\kappa\}\in\,\mathcal{U}\times\mathcal{K}}{\text{Minimize}} \quad \frac{1}{2}\int_\Omega (u-\widehat{u})^2\,dx + \frac{\alpha}{2}\sqrt{|\nabla\kappa|^2+\beta}$$

subject to

$$-\nabla\cdot(\kappa\nabla u) = f \text{ in } \Omega,$$

where $\mathcal{U} = \{u\colon u\in H^1(\Omega),\ u=0 \text{ on } \partial\Omega\}$, $\mathcal{K} = \{\kappa\colon \kappa\in H^1(\Omega),\ \kappa>0\}$ and $\alpha = 10^{-8}$, $\beta = 10^{-10}$.

Mesh Size	Avg. GMRES Iters
64×64	4.7
128×128	4.6
256×256	4.7

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	1.30240e-03	2.109265e-15	8.549061e-05	1.00e+02			
1	3.40849e-04	1.048282e-03	2.583983e-05	1.14e+02	4.03e-13	1.62e+01	8
2	7.39239e-05	8.268766e-04	7.580455e-06	1.45e+02	3.27e-01	2.07e+01	7
3	9.27937e-06	4.704437e-04	1.822981e-06	1.45e+02	1.89e-01	1.86e+01	6
4	7.53843e-07	7.155762e-04	2.777201e-07	1.45e+02	8.29e-02	1.78e+01	13
5	1.71122e-07	2.617447e-04	3.876637e-08	1.45e+02	2.22e-02	9.81e+00	22
6	1.24507e-07	5.506526e-05	3.219040e-09	1.45e+02	5.39e-03	4.31e+00	33
7	1.24282e-07	2.291248e-06	1.023034e-10	1.45e+02	1.06e-03	8.40e-01	35
8	1.24309e-07	3.396201e-09	1.979246e-12	1.45e+02	3.99e-05	2.93e-02	34

CG iters: 158

Material inversion in nonlinear elasticity

$$\underset{\{u,\mu,\kappa\}\in\,\mathcal{U}\times\mathcal{G}\times\mathcal{K}}{\text{Minimize}} \quad \frac{1}{2}\int_{\Omega^0}(u_i-\widehat{u}_i)^2\,dx + R(\mu)+R(\kappa)$$

subject to

$$-(F_{ik}S_{kj})_{,j}=0 \qquad \text{in} \quad \Omega^0,$$

$$(F_{ik}S_{kj})\,n_j=\tau_i \qquad \text{in} \quad \partial\Omega^0_\tau$$

where $\mathcal{U}=\left\{u_i\colon u_i\in H^1\left(\Omega^0\right),\ u_i=0 \text{ on } \partial\Omega^0_u\right\}$, $\mathcal{G}=\{\mu\colon \mu\in L^2(\Omega^0),\ \mu>0\}$, and $\mathcal{B}=\{\kappa\colon \kappa\in L^2(\Omega^0),\ \kappa>0\}$.

— Here $\{u_i\}_{i=1,\ldots,d}$ is the displacement in the i-th direction and τ_i is the surface traction in the i-th direction.

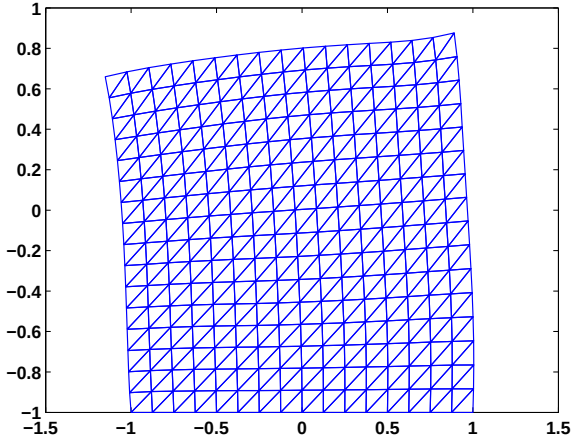
— The second Piola-Kirchhoff stress tensor for a Saint-Venant Kirchhoff material is given by $S_{ij}=\mathbf{C}_{ijkl}\,E_{kl}$, where the fourth-order tensor of elastic moduli is

$$\mathbf{C}_{ijkl}=(\kappa-\frac{2}{3})\,\delta_{ij}\,\delta_{kl}+\mu\left(\delta_{ik}\,\delta_{jl}+\delta_{il}\,\delta_{jk}\right).$$

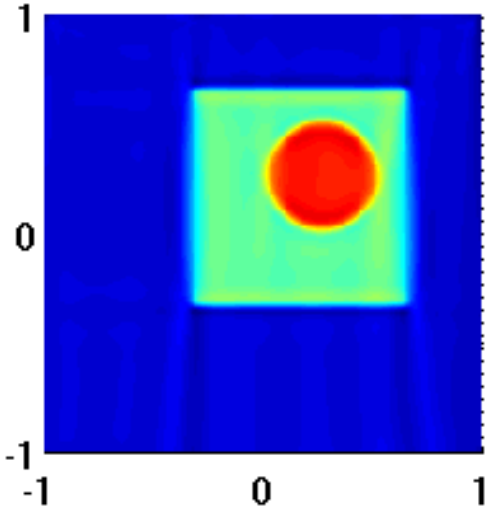
The Green strain tensor is given by $E_{ij}=\frac{1}{2}\left(F_{ki}F_{kj}-\delta_{ij}\right)$, with the deformation gradient $F_{ij}=\frac{\partial u_i}{\partial x^0_j}+\delta_{ij}$, Kronecker delta δ_{ij}, and a reference material point x^0.

Elastic inversion – Solution

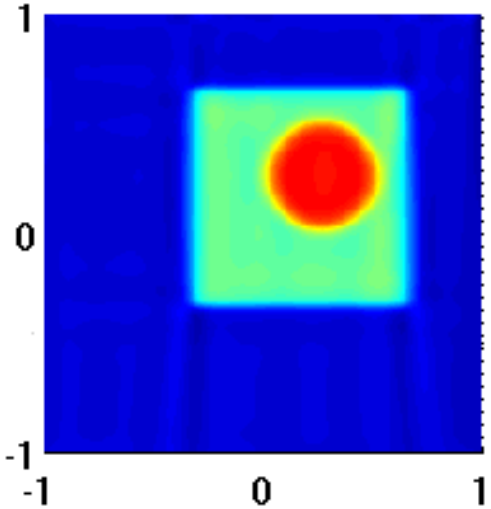
Nonlinear elastic deformation



Computed shear modulus μ



Computed bulk modulus κ



Mesh Size	Avg. GMRES Iters
32×32	9.1
64×64	9.0
128×128	9.7

k	f(x_k)	|| c(x_k) ||	||L_x(x_k,lam_k)||	Delta_k	|| n_k ||	|| t_k ||	itcg
0	1.74452e-03	1.588574e-15	2.846439e-03	1.00e+02			
1	1.37411e-04	1.997871e-03	5.556572e-05	1.00e+02	3.16e-14	3.83e+00	56
2	4.73910e-06	4.318096e-03	2.443777e-06	1.00e+02	2.52e-01	6.38e+00	7
3	3.10439e-06	5.943678e-04	1.150126e-06	1.00e+02	1.13e-01	3.17e+00	7
4	2.14699e-06	6.900686e-04	9.066461e-08	1.00e+02	8.58e-03	3.52e+00	15
5	2.14079e-06	8.124125e-06	7.329380e-08	1.00e+02	6.31e-03	2.41e-01	8
6	2.13387e-06	1.199189e-05	4.643000e-09	1.00e+02	3.78e-05	4.29e-01	23
7	2.13393e-06	3.457890e-08	1.033532e-09	1.00e+02	8.60e-05	1.35e-02	11
8	2.13393e-06	2.451197e-09	4.289974e-11	1.00e+02	5.27e-07	6.11e-03	24

CG iters: 151

Summary

- A simple approximate Schur preconditioner works very well for augmented systems arising in our matrix-free SQP algorithm.
- Augmented systems enable a *separation of concerns* – optimality vs. PDE constraints – at the level of the preconditioner.
- Numerical examples demonstrate **mesh independence** for various physics models.
- **Only mildly affected by the governing PDEs**.
- Enables the scalable application of **full-space matrix-free SQP** to very large PDE-constrained optimization problems.
- Not shown: Inequality constraints affect the scalability only mildly.

- **Advances in 'all-at-once' KKT preconditioning will improve the overall performance by reducing the number of CG iterations.**

Hedging against uncertainty

After collocation in stochastic space, with deterministic control

$$
\begin{pmatrix}
I & 0 & \dots & 0 & 0 & w^1(C_u^1)^T & 0 & \dots & 0 \\
0 & I & \dots & 0 & 0 & 0 & w^2(C_u^2)^T & \dots & 0 \\
\vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & \dots & I & 0 & 0 & 0 & \dots & w^n(C_u^n)^T \\
0 & 0 & \dots & 0 & I & w^1(C_g^1)^T & w^2(C_g^2)^T & \dots & w^n(C_g^n)^T \\
w^1 C_u^1 & 0 & \dots & 0 & w^1 C_g^1 & 0 & 0 & \dots & 0 \\
0 & w^2 C_u^2 & \dots & 0 & w^2 C_g^2 & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \ddots & 0 & \vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & \dots & w^n C_u^n & w^n C_g^n & 0 & 0 & \dots & 0
\end{pmatrix}
$$