PDE-Constrained Optimization in Physics, Chemistry & Biology: Modelling and Numerical Methods

John Pearson

School of Mathematics, University of Edinburgh

Multiscale Computational Modelling, ICMS, Edinburgh
16 May 2018

Based on work with: Sergey Dolgov, Philip Maini, Martin Stoll, Andy Wathen

PDE-Constrained Optimization

- Many PDE-constrained optimization problems may be written as

$$\begin{aligned} \min_{y,u} \quad & \frac{1}{2}\|y-\widehat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \quad & \mathcal{D}y = u \quad \textbf{+BCs}. \end{aligned}$$

- Applications in far-reaching areas such as flow control, semiconductor design, electromagnetic inverse problems, weather forecasting, medical imaging, and finance.

PDE-Constrained Optimization

- Many PDE-constrained optimization problems may be written as

$$\min_{y,u} \ \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)}$$
$$\text{s.t.} \ \ \mathcal{D}y = u \quad \textbf{+BCs}.$$

- Applications in far-reaching areas such as flow control, semiconductor design, electromagnetic inverse problems, weather forecasting, medical imaging, and finance.

- For instance, suppose a company wishes to store a foodstuff or chemical as close as possible to some "ideal" atmospheric conditions, and sets up a controlled atmosphere to do this.
- This is likely to be expensive, so one may also wish to minimize the cost of doing this.

PDE-Constrained Optimization

- Many PDE-constrained optimization problems may be written as

$$\min_{y,u} \ \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)}$$
$$\text{s.t.} \ \ \mathcal{D}y = u \quad \textbf{+BCs}.$$

- Applications in far-reaching areas such as flow control, semiconductor design, electromagnetic inverse problems, weather forecasting, medical imaging, and finance.

- For instance, suppose a company wishes to store a foodstuff or chemical as close as possible to some "ideal" atmospheric conditions, and sets up a controlled atmosphere to do this.
- This is likely to be expensive, so one may also wish to minimize the cost of doing this.
- Then one could solve this using a PDE-constrained optimization problem, with $\mathcal{D}$ some system of PDEs to describe the physics behind the atmospheric conditions, and

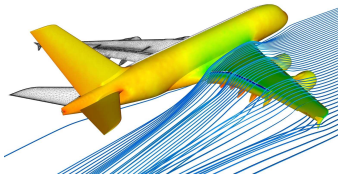
  $y =$ **atmospheric conditions**,

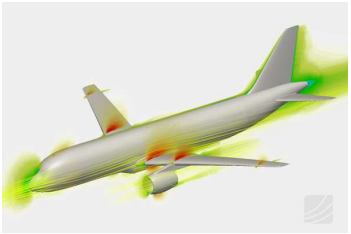
  $u =$ **energy expended**, or financial cost,

  $\hat{y} =$ **"ideal" atmospheric conditions** for chemical,

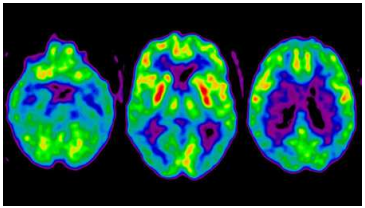
  $\beta =$ **parameter determining at what ratio** atmospheric conditions and cost is prioritized.

Can model problems in fluid flow control, ...





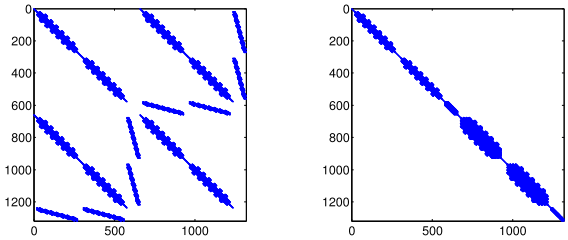
... imaging, including medical imaging, ...



... and chemical processes:

PDE-Constrained Optimization

- We must accurately model the processes, and design efficient numerical solvers.
- Using a finite element method → matrix system of very high dimension.
- Most effective approach for solving these systems is to construct iterative methods which are accelerated by powerful preconditioners.
- When solving matrix system $\mathcal{A}\mathbf{x} = \mathbf{b}$, a good preconditioner $\mathcal{P}$ will be such that $\mathcal{P}^{-1}$ is cheap to apply & $\mathcal{P}^{-1}\mathcal{A}$ has desirable properties.
- Many advantages: can exploit sparsity and structure of matrices, don't have to store the entire system, can solve large problems rapidly & in parallel.



Matrix $\mathcal{A}$ & Preconditioner $\mathcal{P}$

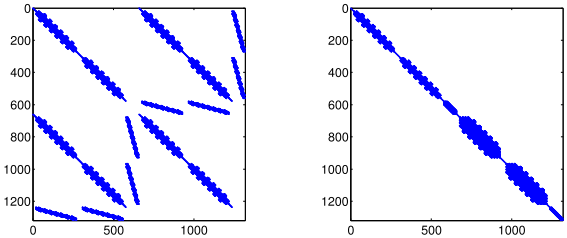
$$\mathcal{A}\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad \mathcal{P}^{-1}\mathcal{A}\mathbf{x} = \mathcal{P}^{-1}\mathbf{b}$$

Preconditioning

- For certain iterative methods, convergence is controlled by

$$\kappa(\mathcal{A}) = \frac{\lambda_{\max}(\mathcal{A})}{\lambda_{\min}(\mathcal{A})} \quad \text{or} \quad \kappa(\mathcal{P}^{-1}\mathcal{A}) = \frac{\lambda_{\max}(\mathcal{P}^{-1}\mathcal{A})}{\lambda_{\min}(\mathcal{P}^{-1}\mathcal{A})}.$$

- So if $\kappa(\mathcal{P}^{-1}\mathcal{A}) \ll \kappa(\mathcal{A})$, convergence is achieved in many fewer iterations.
- But $\mathcal{P}^{-1}$ must be cheap to apply, otherwise 'cost per iteration' is prohibitive.



Matrix $\mathcal{A}$ & Preconditioner $\mathcal{P}$

Saddle Point Systems

- In the problems we consider, the matrices are of *saddle point* structure:

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}.$$

- Two preconditioners for $\mathcal{A}$ are [Kuznetsov, 1995], [Murphy, Golub & Wathen, 2000]

$$\mathcal{P}_D = \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}, \qquad \mathcal{P}_T = \begin{bmatrix} A & 0 \\ B & -S \end{bmatrix}.$$

- Here, $S = BA^{-1}B^T$ is the (negative) *Schur complement*.
- Excellent spectral properties: if $\mathcal{P}_D^{-1}\mathcal{A}$ and $\mathcal{P}_T^{-1}\mathcal{A}$ are nonsingular [Ipsen, 2001]:

$$\lambda(\mathcal{P}_D^{-1}\mathcal{A}) \in \left\{1, \frac{1}{2}(1 \pm \sqrt{5})\right\},$$
$$\lambda(\mathcal{P}_T^{-1}\mathcal{A}) \in \{1\}.$$

- In general A, S are not practical preconditioners, so we devise approximations $\widehat{A}$, $\widehat{S}$.

$$\begin{aligned} \min_{y,u} \quad & \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \quad & -\nabla^2 y = u \quad \text{in } \Omega \\ & y = f \quad \text{on } \partial\Omega \end{aligned}$$

$$\begin{aligned} \min_{y,u} \quad & \frac{1}{2}\|y-\widehat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \quad & -\nabla^2 y = u \quad \text{in } \Omega \\ & y = f \quad \text{on } \partial\Omega \end{aligned}$$

Distributed Poisson Control

- Differentiating (with respect to $\mathbf{y}$, $\mathbf{u}$, $\mathbf{p}$) the cost functional:

$$\mathcal{L}(\mathbf{y},\mathbf{u},\mathbf{p}) = \frac{1}{2}(\mathbf{y}-\widehat{\mathbf{y}})^T M(\mathbf{y}-\widehat{\mathbf{y}}) + \frac{\beta}{2}\mathbf{u}^T M\mathbf{u} + \mathbf{p}^T(K\mathbf{y} - M\mathbf{u} - \mathbf{d}),$$

where M is a finite element *mass matrix*, and K a *stiffness matrix*, gives

$$\left[\begin{array}{cc|c} M & 0 & K \\ 0 & \beta M & -M \\ \hline K & -M & 0 \end{array}\right] \left[\begin{array}{c} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{array}\right] = \left[\begin{array}{c} M\widehat{\mathbf{y}} \\ \mathbf{0} \\ \mathbf{d} \end{array}\right].$$

$$\begin{aligned} \min_{y,u} \quad & \frac{1}{2}\|y-\hat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L_2(\Omega)} \\ \text{s.t.} \quad & -\nabla^2 y = u \quad \text{in } \Omega \\ & y = f \quad \text{on } \partial\Omega \end{aligned}$$

Distributed Poisson Control

- Differentiating (with respect to **y**, **u**, **p**) the cost functional:

$$\mathcal{L}(\mathbf{y},\mathbf{u},\mathbf{p}) = \frac{1}{2}(\mathbf{y}-\hat{\mathbf{y}})^T M(\mathbf{y}-\hat{\mathbf{y}}) + \frac{\beta}{2}\mathbf{u}^T M\mathbf{u} + \mathbf{p}^T(K\mathbf{y} - M\mathbf{u} - \mathbf{d}),$$

  where M is a finite element *mass matrix*, and K a *stiffness matrix*, gives

$$\left[\begin{array}{cc|c} M & 0 & K \\ 0 & \beta M & -M \\ \hline K & -M & 0 \end{array}\right] \left[\begin{array}{c} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{array}\right] = \left[\begin{array}{c} M\hat{\mathbf{y}} \\ \mathbf{0} \\ \mathbf{d} \end{array}\right].$$

- This is a saddle point system with

$$A = \begin{bmatrix} M & 0 \\ 0 & \beta M \end{bmatrix}, \qquad S = KM^{-1}K + \frac{1}{\beta}M.$$

- We may precondition A using Chebyshev semi-iteration to approximate M^{-1}.

Approximating the Schur Complement – Matching Strategy

- We aim to capture both terms of the Schur complement by writing

$$S = KM^{-1}K + \frac{1}{\beta}M, \qquad \widehat{S} = \left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right).$$

- This ensures that [Pearson & Wathen, 2012]:

$$\boxed{\lambda(\widehat{S}^{-1}S) \in \left[\frac{1}{2}, 1\right]}$$

Approximating the Schur Complement – Matching Strategy

- We aim to capture both terms of the Schur complement by writing

$$S = KM^{-1}K + \frac{1}{\beta}M, \qquad \widehat{S} = \left(K + \frac{1}{\sqrt{\beta}}M\right) M^{-1} \left(K + \frac{1}{\sqrt{\beta}}M\right).$$

- This ensures that [Pearson & Wathen, 2012]:

$$\boxed{\lambda(\widehat{S}^{-1}S) \in \left[\frac{1}{2}, 1\right]}$$

- Our preconditioner requires four ingredients:
  1. Saddle point approximation,
  2. Approximation of mass matrix by Chebyshev semi-iteration,
  3. Matching strategy for Schur complement,
  4. Effective multigrid method for $K + \frac{1}{\sqrt{\beta}}M$ to apply $\widehat{S}$.
- Only $\sim$ 15 iterations required for 6 digits of accuracy using MINRES.

J. W. Pearson and A. J. Wathen, *A New Approximation of the Schur Complement in Preconditioners for PDE-Constrained Optimization*, Numerical Linear Algebra with Applications.

Time-Dependent Problems

$$\min_{y,u} \quad \frac{1}{2}\int_0^T \int_\Omega \left(y(\mathbf{x},t) - \hat{y}(\mathbf{x},t)\right)^2 \, \mathrm{d}\Omega \mathrm{d}t + \frac{\beta}{2}\int_0^T \int_\Omega \left(u(\mathbf{x},t)\right)^2 \, \mathrm{d}\Omega \mathrm{d}t \qquad \text{s.t.} \quad \frac{\partial y}{\partial t} - \nabla^2 y = u \quad + \textbf{ICs} \quad + \textbf{BCs}$$

Time-Dependent Problems

$$\min_{y,u} \ \frac{1}{2}\int_0^T \int_\Omega (y(\mathbf{x},t)-\hat{y}(\mathbf{x},t))^2 \ \mathrm{d}\Omega \mathrm{d}t + \frac{\beta}{2}\int_0^T \int_\Omega (u(\mathbf{x},t))^2 \ \mathrm{d}\Omega \mathrm{d}t \qquad \text{s.t.} \quad \frac{\partial y}{\partial t} - \nabla^2 y = u \quad + \ \textbf{ICs} \ + \ \textbf{BCs}$$

The matrix system is of the form $\mathcal{A}\mathbf{x} = \mathbf{b}$, where $\mathcal{A}$ is given by

$$\begin{bmatrix}
\frac{\tau}{2}M & & & & 0 & & & & M+\tau K & -M & & \\
& \tau M & & & & 0 & & & & M+\tau K & -M & \\
& & \ddots & & & & \ddots & & & & \ddots & \ddots \\
& & & \frac{\tau}{2}M & & & & 0 & & & & M+\tau K \\
\hline
0 & & & & \frac{\beta\tau}{2}M & & & & -\tau M & & & \\
& 0 & & & & \beta\tau M & & & & -\tau M & & \\
& & \ddots & & & & \ddots & & & & \ddots & \\
& & & 0 & & & & \frac{\beta\tau}{2}M & & & & -\tau M \\
\hline
M+\tau K & & & & -\tau M & & & & 0 & & & \\
-M & M+\tau K & & & & -\tau M & & & & 0 & & \\
& \ddots & \ddots & & & & \ddots & & & & \ddots & \\
& & -M & M+\tau K & & & & -\tau M & & & & 0
\end{bmatrix}.$$

Table of Contents

1. Chemistry: Chemical Reactions
2. Biology: Pattern Formation & Chemotaxis
3. Physics: Fluid Flow Control & Image Metamorphosis

Table of Contents

1 Chemistry: Chemical Reactions

2 Biology: Pattern Formation & Chemotaxis

3 Physics: Fluid Flow Control & Image Metamorphosis

Reaction–Diffusion Control Problems for Chemical Reactions

- A problem which we are now keen to consider is the following optimal control problem involving reaction–diffusion equations. We wish to minimize

$$\mathcal{J} = \frac{\alpha_y}{2}\int_0^T\int_\Omega (y-\widehat{y})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_z}{2}\int_0^T\int_\Omega (z-\widehat{z})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_c}{2}\int_0^T\int_{\partial\Omega} c^2 \ \mathrm{d}s\mathrm{d}t,$$

subject to the following PDE constraints:

$$\begin{aligned}
y_t - D_1\nabla^2 y + k_1 y &= -\gamma_1 yz && \text{in } \Omega\times(0,T),\\
z_t - D_2\nabla^2 z + k_2 z &= -\gamma_2 yz && \text{in } \Omega\times(0,T),\\
D_1\frac{\partial y}{\partial n} &= c && \text{on } \partial\Omega\times(0,T),\\
D_2\frac{\partial z}{\partial n} + \epsilon z &= 0 && \text{on } \partial\Omega\times(0,T),\\
y(\mathbf{x},0) &= y_0(\mathbf{x}) && \text{in } \Omega,\\
z(\mathbf{x},0) &= z_0(\mathbf{x}) && \text{in } \Omega.
\end{aligned}$$

- We may also incorporate the control constraints $c_- \leqslant c \leqslant c_+$ a.e. on $\partial\Omega\times(0,T)$.

Optimality Conditions

- On the continuous level, we consider the Lagrangian

$$\begin{aligned}\mathcal{L}(y,z,c,p,q) = \frac{\alpha_y}{2}\int_0^T\int_\Omega (y-\hat{y})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_z}{2}\int_0^T\int_\Omega (z-\hat{z})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_c}{2}\int_0^T\int_{\partial\Omega} c^2 \ \mathrm{d}s\mathrm{d}t \\ + \int_0^T\int_\Omega p_\Omega\left(y_t - D_1\nabla^2 y + k_1 y + \gamma_1 yz\right) \mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_\Omega q_\Omega\left(z_t - D_2\nabla^2 z + k_2 z + \gamma_2 yz\right) \mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_{\partial\Omega} p_{\partial\Omega}\left(D_1\frac{\partial y}{\partial n} - c\right) \mathrm{d}s\mathrm{d}t + \int_0^T\int_{\partial\Omega} q_{\partial\Omega}\left(D_2\frac{\partial z}{\partial n} + \epsilon z\right) \mathrm{d}s\mathrm{d}t.\end{aligned}$$

Optimality Conditions

- On the continuous level, we consider the Lagrangian

$$\begin{aligned}\mathcal{L}(y,z,c,p,q) = \frac{\alpha_y}{2}\int_0^T\int_\Omega (y-\widehat{y})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_z}{2}\int_0^T\int_\Omega (z-\widehat{z})^2 \ \mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_c}{2}\int_0^T\int_{\partial\Omega} c^2 \ \mathrm{d}s\mathrm{d}t \\ + \int_0^T\int_\Omega p_\Omega\left(y_t - D_1\nabla^2 y + k_1 y + \gamma_1 yz\right)\ \mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_\Omega q_\Omega\left(z_t - D_2\nabla^2 z + k_2 z + \gamma_2 yz\right)\ \mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_{\partial\Omega} p_{\partial\Omega}\left(D_1\frac{\partial y}{\partial n} - c\right)\ \mathrm{d}s\mathrm{d}t + \int_0^T\int_{\partial\Omega} q_{\partial\Omega}\left(D_2\frac{\partial z}{\partial n} + \epsilon z\right)\ \mathrm{d}s\mathrm{d}t.\end{aligned}$$

- Differentiating with respect to p, q gives the *state equations*:

$$y_t - D_1\nabla^2 y + k_1 y = -\gamma_1 yz, \qquad z_t - D_2\nabla^2 z + k_2 z = -\gamma_2 yz.$$

Optimality Conditions

- On the continuous level, we consider the Lagrangian

$$\begin{aligned}\mathcal{L}(y,z,c,p,q) = &\frac{\alpha_y}{2}\int_0^T\int_\Omega (y-\hat{y})^2\,\mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_z}{2}\int_0^T\int_\Omega (z-\hat{z})^2\,\mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_c}{2}\int_0^T\int_{\partial\Omega} c^2\,\mathrm{d}s\mathrm{d}t\\ &+\int_0^T\int_\Omega p_\Omega\left(y_t - D_1\nabla^2 y + k_1 y + \gamma_1 yz\right)\mathrm{d}\Omega\mathrm{d}t\\ &+\int_0^T\int_\Omega q_\Omega\left(z_t - D_2\nabla^2 z + k_2 z + \gamma_2 yz\right)\mathrm{d}\Omega\mathrm{d}t\\ &+\int_0^T\int_{\partial\Omega} p_{\partial\Omega}\left(D_1\frac{\partial y}{\partial n} - c\right)\mathrm{d}s\mathrm{d}t + \int_0^T\int_{\partial\Omega} q_{\partial\Omega}\left(D_2\frac{\partial z}{\partial n} + \epsilon z\right)\mathrm{d}s\mathrm{d}t.\end{aligned}$$

- Differentiating with respect to p, q gives the *state equations*:

$$y_t - D_1\nabla^2 y + k_1 y = -\gamma_1 yz, \qquad z_t - D_2\nabla^2 z + k_2 z = -\gamma_2 yz.$$

- Differentiating with respect to y, z gives the *adjoint equations*:

$$\begin{aligned}-p_t - D_1\nabla^2 p + k_1 p + \gamma_1 pz + \gamma_2 qz + \alpha_y y &= \alpha_y\hat{y},\\ -q_t - D_2\nabla^2 q + k_2 q + \gamma_2 qy + \gamma_1 py + \alpha_z z &= \alpha_z\hat{z}.\end{aligned}$$

Optimality Conditions

- On the continuous level, we consider the Lagrangian

$$\begin{aligned}\mathcal{L}(y,z,c,p,q) = \frac{\alpha_y}{2}\int_0^T\int_\Omega (y-\widehat{y})^2 \,\mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_z}{2}\int_0^T\int_\Omega (z-\widehat{z})^2 \,\mathrm{d}\Omega\mathrm{d}t + \frac{\alpha_c}{2}\int_0^T\int_{\partial\Omega} c^2 \,\mathrm{d}s\mathrm{d}t \\ + \int_0^T\int_\Omega p_\Omega\left(y_t - D_1\nabla^2 y + k_1 y + \gamma_1 yz\right)\,\mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_\Omega q_\Omega\left(z_t - D_2\nabla^2 z + k_2 z + \gamma_2 yz\right)\,\mathrm{d}\Omega\mathrm{d}t \\ + \int_0^T\int_{\partial\Omega} p_{\partial\Omega}\left(D_1\frac{\partial y}{\partial n} - c\right)\,\mathrm{d}s\mathrm{d}t + \int_0^T\int_{\partial\Omega} q_{\partial\Omega}\left(D_2\frac{\partial z}{\partial n} + \epsilon z\right)\,\mathrm{d}s\mathrm{d}t.\end{aligned}$$

- Differentiating with respect to p, q gives the *state equations*:

$$y_t - D_1\nabla^2 y + k_1 y = -\gamma_1 yz, \qquad z_t - D_2\nabla^2 z + k_2 z = -\gamma_2 yz.$$

- Differentiating with respect to y, z gives the *adjoint equations*:

$$\begin{aligned}-p_t - D_1\nabla^2 p + k_1 p + \gamma_1 pz + \gamma_2 qz + \alpha_y y = \alpha_y\widehat{y}, \\ -q_t - D_2\nabla^2 q + k_2 q + \gamma_2 qy + \gamma_1 py + \alpha_z z = \alpha_z\widehat{z}.\end{aligned}$$

- Differentiating with respect to c gives the *gradient equation* $\alpha_c c - p = 0$ on $\partial\Omega \times (0,T)$.

Newton Iteration – Matrix System

- In matrix form, the Newton system is written as

$$\begin{bmatrix} \alpha_y \mathrm{Id} & \gamma_1\bar{p}+\gamma_2\bar{q} & 0 & \mathcal{D}_y' & \gamma_2\bar{z} \\ \gamma_1\bar{p}+\gamma_2\bar{q} & \alpha_z \mathrm{Id} & 0 & \gamma_1\bar{y} & \mathcal{D}_z' \\ 0 & 0 & \alpha_c D_1^{-1}\mathrm{Id} & -D_1^{-1}\mathrm{Id} & 0 \\ \mathcal{D}_y & \gamma_1\bar{y} & -D_1^{-1}\mathrm{Id} & 0 & 0 \\ \gamma_2\bar{z} & \mathcal{D}_z & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_y \\ s_z \\ s_c \\ s_p \\ s_q \end{bmatrix} = b,$$

where

$$\mathcal{D}_y = \frac{\partial}{\partial t} - D_1\nabla^2 + k_1\mathrm{Id} + \gamma_1\bar{z}, \qquad \mathcal{D}_y' = -\frac{\partial}{\partial t} - D_1\nabla^2 + k_1\mathrm{Id} + \gamma_1\bar{z},$$
$$\mathcal{D}_z = \frac{\partial}{\partial t} - D_2\nabla^2 + k_2\mathrm{Id} + \gamma_2\bar{y}, \qquad \mathcal{D}_z' = -\frac{\partial}{\partial t} - D_2\nabla^2 + k_2\mathrm{Id} + \gamma_2\bar{y}.$$

- The vector b represents the terms from the previous Newton iteration.

Newton Iteration – Matrix System

- Applying a finite element method at the Newton step, we obtain the matrix:

$$\begin{bmatrix} \tau\alpha_y\mathbf{M} & \tau(\gamma_1\mathbf{M}_p + \gamma_2\mathbf{M}_q) & 0 & \mathbf{L}_{y,C}^T & \tau\gamma_2\mathbf{M}_z \\ \tau(\gamma_1\mathbf{M}_p + \gamma_2\mathbf{M}_q) & \tau\alpha_z\mathbf{M} & 0 & \tau\gamma_1\mathbf{M}_y & \mathbf{L}_{z,C}^T \\ 0 & 0 & \tau\alpha_c D_1^{-1}\mathbf{M}_\Gamma & -\tau D_1^{-1}\mathbf{N}^T & 0 \\ \mathbf{L}_{y,C} & \tau\gamma_1\mathbf{M}_y & -\tau D_1^{-1}\mathbf{N} & 0 & 0 \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} & 0 & 0 & 0 \end{bmatrix},$$

  where

$$\mathbf{L}_{y,C} = \mathbf{M}_E + \tau D_1\mathbf{K} + \tau k_1\mathbf{M} + \tau\gamma_1\mathbf{M}_z, \quad \mathbf{L}_{z,C} = \mathbf{M}_E + \tau D_2\mathbf{K} + \tau k_2\mathbf{M} + \tau\gamma_2\mathbf{M}_y.$$

- Here $\mathbf{M}$ and $\mathbf{K}$ are block diagonal matrices with mass and stiffness matrices for each time-step, $\mathbf{M}_\Gamma$ is associated boundary mass matrix, $\mathbf{N}$ the trace operator mapping onto the boundary, and $\mathbf{M}_E$ mass matrices from time-stepping.
- All other $\mathbf{M}_\psi = \mathrm{blkdiag}(M_\psi, \ldots, M_\psi)$ are obtained from evaluating integrals of the form $[M_\psi]_{ij} = \int \psi\phi_i\phi_j$ for each matrix entry.

$$\left[\begin{array}{ccc|cc} \tau\alpha_y\mathbf{M} & \tau(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q) & 0 & \mathbf{L}_{y,C}^T & \tau\gamma_2\mathbf{M}_z \\ \tau(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q) & \tau\alpha_z\mathbf{M} & 0 & \tau\gamma_1\mathbf{M}_y & \mathbf{L}_{z,C}^T \\ 0 & 0 & \tau\alpha_c D_1^{-1}\mathbf{M}_\Gamma & -\tau D_1^{-1}\mathbf{N}^T & 0 \\ \hline \mathbf{L}_{y,C} & \tau\gamma_1\mathbf{M}_u & -\tau D_1^{-1}\mathbf{N} & 0 & 0 \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} & 0 & 0 & 0 \end{array}\right]$$

Preconditioning the Matrix System – $(1,1)$-block

- Let us apply saddle point theory when approximating the $(1,1)$-block, and take

$$\widehat{A} = \tau \left[\begin{array}{cc|c} \alpha_y\mathbf{M} - \alpha_z^{-1}(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q)\mathbf{M}^{-1}(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q) & 0 & 0 \\ \gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q & \alpha_z\mathbf{M} & 0 \\ \hline 0 & 0 & \alpha_c D_1^{-1}\mathbf{M}_\Gamma \end{array}\right].$$

- We apply Chebyshev semi-iteration to approximate $\mathbf{M}_\Gamma^{-1}$.
- Preconditioner will be non-symmetric → apply BiCG or GMRES.

JWP and M. Stoll, *Fast Iterative Solution of Reaction–Diffusion Control Problems Arising from Chemical Reactions*, SIAM Journal on Scientific Computing.

$$\left[\begin{array}{ccc|cc} \tau\alpha_y\mathbf{M} & \tau(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q) & 0 & \mathbf{L}_{y,C}^T & \tau\gamma_2\mathbf{M}_z \\ \tau(\gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q) & \tau\alpha_z\mathbf{M} & 0 & \tau\gamma_1\mathbf{M}_y & \mathbf{L}_{z,C}^T \\ 0 & 0 & \tau\alpha_c D_1^{-1}\mathbf{M}_\Gamma & -\tau D_1^{-1}\mathbf{N}^T & 0 \\ \hline \mathbf{L}_{y,C} & \tau\gamma_1\mathbf{M}_y & -\tau D_1^{-1}\mathbf{N} & 0 & 0 \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} & 0 & 0 & 0 \end{array}\right]$$

Preconditioning the Matrix System – Schur Complement

- We now approximate

$$S = \frac{1}{\tau}\begin{bmatrix} \mathbf{L}_{y,C} & \tau\gamma_1\mathbf{M}_y \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} \end{bmatrix}\mathbf{A}_{(1,2)}^{-1}\begin{bmatrix} \mathbf{L}_{y,C}^T & \tau\gamma_2\mathbf{M}_z \\ \tau\gamma_1\mathbf{M}_y & \mathbf{L}_{z,C}^T \end{bmatrix} + \frac{\tau}{\alpha_c D_1}\begin{bmatrix} \mathbf{N}\mathbf{M}_\Gamma^{-1}\mathbf{N}^T & 0 \\ 0 & 0 \end{bmatrix},$$

  where

$$\mathbf{A}_{(1,2)} = \begin{bmatrix} \alpha_y\mathbf{M} & \gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q \\ \gamma_1\mathbf{M}_p+\gamma_2\mathbf{M}_q & \alpha_z\mathbf{M} \end{bmatrix}.$$

- We make use of our matching strategy derived earlier to write:

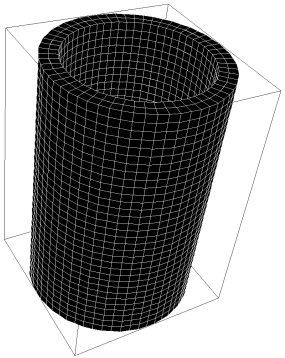
$$\widehat{S} = \frac{1}{\tau}\begin{bmatrix} \mathbf{L}_{y,C}+\widehat{\mathbf{M}} & \tau\gamma_1\mathbf{M}_y \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} \end{bmatrix}\mathbf{A}_{(1,2)}^{-1}\begin{bmatrix} \mathbf{L}_{y,C}^T+\widehat{\mathbf{M}} & \tau\gamma_2\mathbf{M}_z \\ \tau\gamma_1\mathbf{M}_y & \mathbf{L}_{z,C}^T \end{bmatrix}.$$

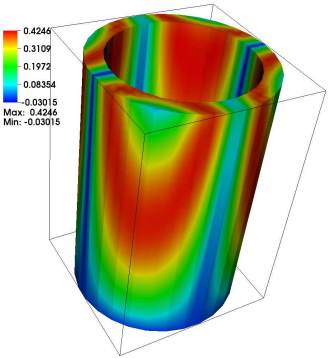
$$\widehat{S} = \frac{1}{\tau}\begin{bmatrix} \mathbf{L}_{y,C} + \widehat{\mathbf{M}} & \tau\gamma_1\mathbf{M}_y \\ \tau\gamma_2\mathbf{M}_z & \mathbf{L}_{z,C} \end{bmatrix} \mathbf{A}^{-1}_{(1,2)} \begin{bmatrix} \mathbf{L}^T_{y,C} + \widehat{\mathbf{M}} & \tau\gamma_2\mathbf{M}_z \\ \tau\gamma_1\mathbf{M}_y & \mathbf{L}^T_{z,C} \end{bmatrix}$$

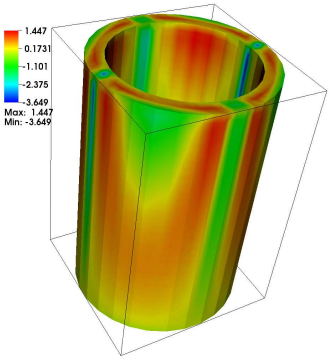
Some Observations

- To apply $\widehat{S}^{-1}$ in practice, use fixed number of iterations of an Uzawa scheme, coupled with algebraic multigrid routine to approximate the diagonal blocks.
- Good lower bound of $\lambda(\widehat{S}^{-1}S)$.
- Greater variation in upper bound due to range of parameters: mesh size h, τ, α_y, α_z, α_c, D_1, D_2, k_1, k_2, γ_1, γ_2, ϵ, c_-, c_+.
- Best case scenario: when one term in S strongly dominates.
- Worst case scenario: when first term of S is (close to) indefinite.

$$\hat{y} = t\,|\sin(2x_1x_2x_3)| + 0.3, \quad \hat{z} = 0, \quad k_1 = k_2 = D_1 = D_2 = 1, \quad \gamma_1 = \gamma_2 = 0.15$$







DoF	$\alpha_c = 10^{-3}$			$\alpha_c = 10^{-5}$		
	Time	Newton	Iterations	Time	Newton	Iterations
538,240	1,995	step 1	17	1,726	step 1	16
		step 2	20		step 2	16
		step 3	20		step 3	16
3,331,520	14,757	step 1	28	14,904	step 1	28
		step 2	31		step 2	27
		step 3	29		step 3	34

Table of Contents

1 Chemistry: Chemical Reactions

2 Biology: Pattern Formation & Chemotaxis

3 Physics: Fluid Flow Control & Image Metamorphosis

Parameter Identification in Pattern Formation Processes

- A useful reaction–diffusion control problem arises in pattern formation processes. Here we wish to minimize

$$\mathcal{J}(y,z,a,b) = \frac{\beta_1}{2}\left\|y-\widehat{y}(\mathbf{x},t)\right\|^2_{L_2(\Omega)} + \frac{\beta_2}{2}\left\|z-\widehat{z}(\mathbf{x},t)\right\|^2_{L_2(\Omega)} + \frac{\nu_1}{2}\left\|a(\mathbf{x},t)\right\|^2_{L_2(\Omega)} + \frac{\nu_2}{2}\left\|b(\mathbf{x},t)\right\|^2_{L_2(\Omega)},$$

  subject to PDE constraints given by the Schnakenberg equations

$$\begin{aligned} y_t - D_y\nabla^2 y + \gamma(y-y^2z) - \gamma a &= 0 \quad \text{in } \Omega, \\ z_t - D_z\nabla^2 z + \gamma y^2 z - \gamma b &= 0 \quad \text{in } \Omega, \\ y(\mathbf{x},0) = y_0(\mathbf{x}), &\qquad z(\mathbf{x},0) = z_0(\mathbf{x}), \\ \frac{\partial y}{\partial n} = \frac{\partial z}{\partial n} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

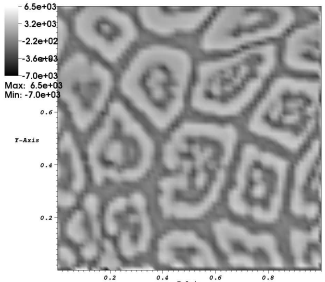
- The state variables here are y and z, with the control variables a and b.
- We may again include control constraints.

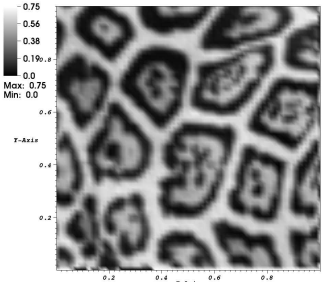
Parameter Identification in Pattern Formation

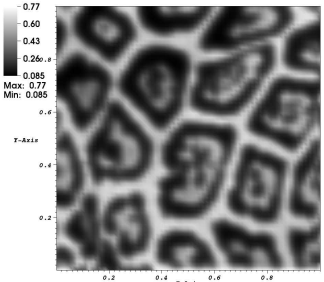
$$\min_{y,z,a,b} \quad \frac{\beta_1}{2}\|y-\widehat{y}(\mathbf{x},t)\|^2_{L_2(\Omega)} + \frac{\beta_2}{2}\|z-\widehat{z}(\mathbf{x},t)\|^2_{L_2(\Omega)} + \frac{\nu_1}{2}\|a(\mathbf{x},t)\|^2_{L_2(\Omega)} + \frac{\nu_2}{2}\|b(\mathbf{x},t)\|^2_{L_2(\Omega)}$$

s.t.

$$\begin{aligned} y_t - D_y\nabla^2 y + \gamma(y - y^2 z) - \gamma a = 0 \quad &\text{in } \Omega, \\ z_t - D_z\nabla^2 z + \gamma y^2 z - \gamma b = 0 \quad &\text{in } \Omega, \\ y(\mathbf{x},0) = y_0(\mathbf{x}), \qquad & z(\mathbf{x},0) = z_0(\mathbf{x}), \\ \frac{\partial y}{\partial n} = \frac{\partial z}{\partial n} = 0 \quad &\text{on } \partial\Omega. \end{aligned}$$

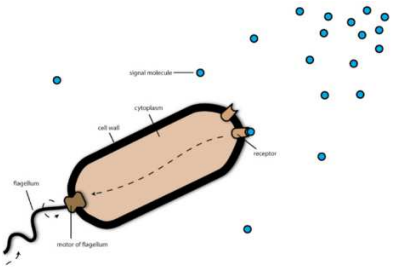


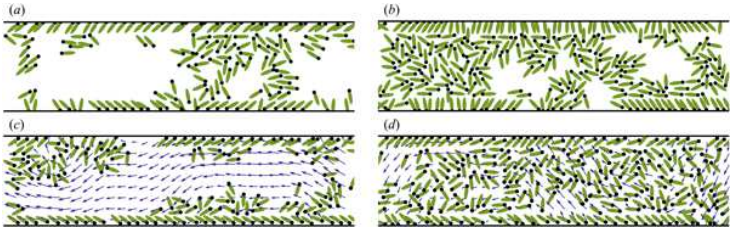




Optimal Control Problems in Chemotaxis

- Chemotaxis is the movement of cells/organisms in a directed fashion as a response to external chemical signals.
- Keller and Segel presented a mathematical model for bacterial chemotaxis in 1971.
- In essence, for large numbers of bacteria, the bacteria will on average move up gradients of the chemoattractant concentration.
- Inverse problem: Given an observed cell concentration profile, what can be said about chemoattractant at boundaries of the domain?
- Shown numerically by Lebiedz and Brandt-Pollmann that "it is possible to systematically control spatiotemporal dynamical behavior".





Optimal Control of Bacterial Chemotaxis System

- We wish to examine a boundary control problem arising from a bacterial chemotaxis system:

$$
\begin{aligned}
\min_{z,c,u} \quad & \frac{1}{2}\int_\Omega (z(\mathbf{x},T)-\widehat{z})^2 + \frac{\gamma_c}{2}\int_\Omega (c(\mathbf{x},T)-\widehat{c})^2 + \frac{\gamma_u}{2}\int_0^T\int_{\partial\Omega} u^2 \\
\text{s.t.} \quad & \frac{\partial z}{\partial t} - D_z\nabla^2 z - \alpha\nabla\cdot\left(\frac{\nabla c}{(1+c)^2}z\right) = 0 && \text{on } \Omega\times(0,T), \\
& \frac{\partial c}{\partial t} - \nabla^2 c + \rho c - w\frac{z^2}{1+z^2} = 0 && \text{on } \Omega\times(0,T), \\
& \frac{\partial z}{\partial n} = 0 && \text{on } \partial\Omega\times(0,T), \\
& \frac{\partial c}{\partial n} + \zeta c = \zeta u && \text{on } \partial\Omega\times(0,T), \\
& z(\mathbf{x},0) = z_0(\mathbf{x}) && \text{on } \Omega, \\
& c(\mathbf{x},0) = c_0(\mathbf{x}) && \text{on } \Omega
\end{aligned}
$$

- z denotes cell density, c is concentration of chemoattractant.
- z_0, c_0 are given initial conditions, and γ_c, γ_u, D_z, α, ρ, w, ζ given (positive) parameters.
- Can also consider additional control constraints $u_-(\mathbf{x},t) \leqslant u \leqslant u_+(\mathbf{x},t)$.

Optimal Control of Bacterial Chemotaxis System

- We wish to examine a boundary control problem arising from a bacterial chemotaxis system:

$$
\begin{aligned}
\min_{z,c,u} \quad & \frac{1}{2}\int_\Omega (z(\mathbf{x},T)-\widehat{z})^2 + \frac{\gamma_c}{2}\int_\Omega (c(\mathbf{x},T)-\widehat{c})^2 + \frac{\gamma_u}{2}\int_0^T\int_{\partial\Omega} u^2 \\
\text{s.t.} \quad & \frac{\partial z}{\partial t} - D_z\nabla^2 z - \alpha\nabla\cdot\left(\frac{\nabla c}{(1+c)^2}z\right) = 0 && \text{on } \Omega\times(0,T), \\
& \frac{\partial c}{\partial t} - \nabla^2 c + \rho c - w\frac{z^2}{1+z^2} = 0 && \text{on } \Omega\times(0,T), \\
& \frac{\partial z}{\partial n} = 0 && \text{on } \partial\Omega\times(0,T), \\
& \frac{\partial c}{\partial n} + \zeta c = \zeta u && \text{on } \partial\Omega\times(0,T), \\
& z(\mathbf{x},0) = z_0(\mathbf{x}) && \text{on } \Omega, \\
& c(\mathbf{x},0) = c_0(\mathbf{x}) && \text{on } \Omega
\end{aligned}
$$

- z denotes cell density, c is concentration of chemoattractant.
- z_0, c_0 are given initial conditions, and γ_c, γ_u, D_z, α, ρ, w, ζ given (positive) parameters.
- Can also consider additional control constraints $u_-(\mathbf{x},t) \leqslant u \leqslant u_+(\mathbf{x},t)$.

Table of Contents

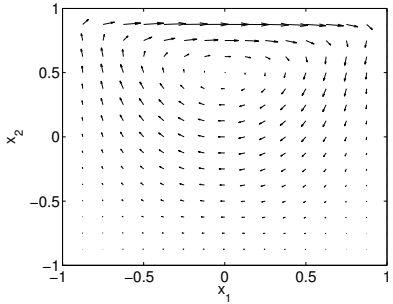
1 Chemistry: Chemical Reactions

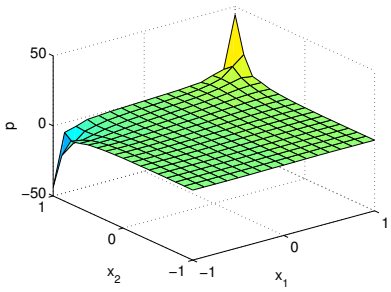
2 Biology: Pattern Formation & Chemotaxis

3 Physics: Fluid Flow Control & Image Metamorphosis

Optimal Control Problems in Fluid Dynamics

$$\min_{\underline{\mathbf{y}},\underline{\mathbf{u}}} \quad \frac{1}{2}\left\|\underline{\mathbf{y}}-\widehat{\underline{\mathbf{y}}}\right\|^2_{L_2(\Omega)}+\frac{\beta}{2}\left\|\underline{\mathbf{u}}\right\|^2_{L_2(\Omega)}, \qquad \mathcal{D}=\begin{bmatrix} \frac{\partial}{\partial t}-\nu\nabla^2+\underline{\mathbf{y}}\cdot\nabla & \nabla \\ -\nabla\cdot & 0 \end{bmatrix}$$





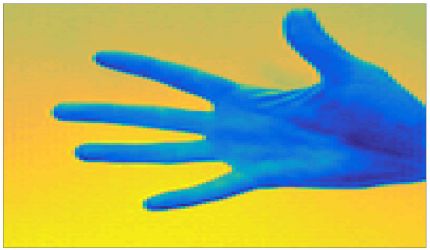
- Crucial application area is that of *flow control* problems.
- Can tackle a range of such problems, with preconditioners explicitly based on physical features of fluid flow.

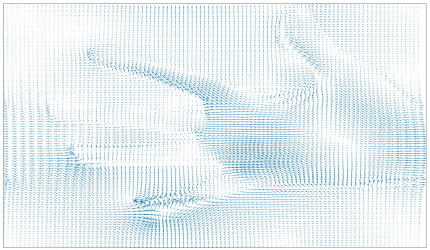
JWP, *On the Development of Parameter-Robust Preconditioners and Commutator Arguments for Solving Stokes Control Problems*, Electronic Transactions on Numerical Analysis.

JWP, *Preconditioned Iterative Methods for Navier-Stokes Control Problems*, Journal of Computational Physics.

Optimal Transport for Image Metamorphosis

$$\mathcal{E}(y, \mathbf{m}) = \frac{1}{2\gamma} \int_\Omega (y(\mathbf{x}, 1) - \hat{y}_1(\mathbf{x}))^2 \, \mathrm{d}\Omega + \frac{\delta}{2} \int_0^1 \int_\Omega (y(\mathbf{x}, t) - \hat{y}(\mathbf{x}, t))^2 \, \mathrm{d}\Omega \mathrm{d}t + \frac{\beta}{2} \int_0^1 \int_\Omega (Q\mathbf{m}(\mathbf{x}, t))^2 \, \mathrm{d}\Omega \mathrm{d}t$$





- We may also investigate the solution of an optimization problem subject to a transport equation arising from the modelling of image metamorphosis.
- Models the apparent 'motion' of an image, in a movie for example.
- Good numerical results, using both finite difference approach, and radial basis functions.

R. Herzog, JWP, and M. Stoll, *Fast Iterative Solvers for an Optimal Transport Problem*.

Concluding Remarks

- PDE-constrained optimization provides a valuable tool for examining scientific processes, provided suitable mathematical models are developed.
- If this can be done, the main challenge is then devising fast and effective numerical methods for solving the models.
- We considered preconditioned iterative methods for solving the matrix systems arising from these problems, using saddle point approximations, coupled with strategies for approximating the $(1,1)$-block and Schur complement.
- We were only required to store matrices which were much smaller than the matrix system as a whole.

Thank you for your attention!