Multigrid methods for optimal control problems with PDEs

Alfio Borzì

Institute for Mathematics and Scientific Computing
Karl-Franzens-University, Graz, Austria

Thirty-second Conference of the Dutch and Flemish Numerical Analysis Communities, Woudschoten Conference, Zeist, The Netherlands, 3 - 5 October, 2007

Research framework

Aim at modeling and simulating application problems is to achieve better understanding of real world systems possibly with the purpose of controlling these systems in a desired way

We discuss the following systems

- Equilibrium systems with constraits
- Biological/chemical/physiological reaction-diffusion models

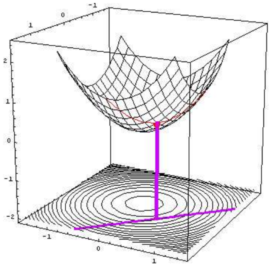
Applications:

Control of equilibrium and bio-chemical models: Optimal configuration in equilibrium systems, open dissipative systems, prey-predator systems, chemical turbulence, electrical fields in human tissues.

The formulation of optimal control problems

- A model of the dynamical system
- A description of the control mechanism
- A criterion that models the purpose of the control and the cost of its action

We have a constrained minimization problem



$$\begin{cases} \text{minimize} & J(y,u) \\ \text{under the constraint} & c(y,u) = 0 \end{cases}$$

Optimization with PDE constraints

$$\min_{u \in U_{ad}} J(y,u) \qquad J : Y \times U \to \mathbb{R}$$
$$\text{s.t.} \quad c(y,u) = 0$$

The existence of c_y^{-1} enables a distinction between y, the state variable, and $u \in U_{ad} \subset U$, the optimization variable in the admissible set.
So we have the mapping $u \mapsto J(y(u),u)$ in the form

$$u \overset{\text{IFT}}{\mapsto} y(u) \mapsto J(y(u),u) =: \hat{J}(u)$$

The solution of this optimization problem is characterized by the following optimality system

$$\begin{aligned} c(y,u) &= 0 \\ c_y(y,u)^* \, p &= -h'(y) \\ (\nu \, g'(u) + c_u^* \, p \, , v - u) &\geq 0 \quad \text{for all } v \in U_{ad} \end{aligned}$$

assuming $J(y,u) = h(y) + \nu \, g(u)$, $\nu > 0$. We have
$\nabla \hat{J}(u) = \nu \, g'(u) + c_u^* \, p(u)$

Outline of the talk I: Equilibrium systems

- Control-constrained nonlinear elliptic optimal control problems
- Optimality systems
- Discretization of the optimality system
- Smoothing and multigrid methods
- A state-constrained elliptic optimal control problem

Control-constrained elliptic optimal control problems

Consider a two-dimensional material plate Ω whose state is described by the temperature distribution y.

Assume thermal radiation ($G(y) < 0$) or positive temperature feedback ($G(y) > 0$) due to chemical reactions.

We may control y to come close to a given target profile $z \in L^2(\Omega)$, by acting with a (boundary or) distributed source term u, the control function.

$$\left\{ \begin{array}{rcll} \min_{u \in U_{ad}} J(y,u) & := & \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} & \\ \Delta y + G(y) & = & u + f & \text{in } \Omega \\ y & = & 0 & \text{on } \partial\Omega \end{array} \right.$$

$$U_{ad} = \{u \in L^2(\Omega) \,|\, u_L(\mathbf{x}) \leq u(\mathbf{x}) \leq u_H(\mathbf{x}) \text{ a.e. in } \Omega\}$$

Optimality system

Optimal solutions are characterized by the following optimality system

$$\begin{aligned} \Delta y + G(y) - u = & \quad f \quad \text{in } \Omega, \\ y = & \quad 0 \quad \text{on } \partial\Omega, \\ \Delta p + G'(y)\, p + y = & \quad z \quad \text{in } \Omega, \\ p = & \quad 0 \quad \text{on } \partial\Omega, \\ (\nu u - p,\ v - u) \geq & \quad 0 \quad \text{for all } v \in U_{ad}. \end{aligned}$$

The last equation gives the optimality condition. It is equivalent to

$$u = \max\{u_L, \min\{u_H, \frac{1}{\nu} p(u)\}\} \text{ in } \Omega, \quad \nu > 0$$

Nondifferentiability !

Optimality system (continue)

The case with $\nu = 0$ is characterized by the following system

$$
\begin{aligned}
\Delta y + G(y) - u &= f && \text{in } \Omega, \\
y &= 0 && \text{on } \partial\Omega, \\
\Delta p + G'(y)\,p + y &= z && \text{in } \Omega, \\
p &= 0 && \text{on } \partial\Omega, \\
p &= \min\{0, p + u - u_L\} + \max\{0, p + u - u_H\} && \text{in } \Omega.
\end{aligned}
$$

Nondifferentiability prevents the use of classical Newton or gradient techniques, requiring more sophisticated methods based on generalized differentiability concepts.

Alternative: MG approach

FDM Discretization

Consider the finite-difference framework [Hackbusch,Süli].
Let Ω be rectangular domain. Introduce the discrete L_h^2-scalar product $(v_h, w_h)_{L_h^2} = h^2 \sum_{\mathbf{x}\in\Omega_h} v_h(\mathbf{x})\, w_h(\mathbf{x})$, with norm $|v_h|_0 = (v_h, v_h)_{L_h^2}^{1/2}$.

First-order backward and forward partial derivatives of v_h in the x_i direction are denoted by ∂_i^- and ∂_i^+, respectively.
Assume sufficiently smooth functions $v \in C^k(\bar{\Omega})$, $k = 0, 1, \ldots$, and denote with $(R_h v)(x) = v(x)$ the restriction operator on $\bar{\Omega}$. We have
The second-order five-point Laplacian

$$\tilde{\Delta}_h = \partial_1^+ \partial_1^- + \partial_2^+ \partial_2^-$$

The fourth-order nine-point Laplacian

$$\Delta_h = (1 - \frac{h^2}{12} \partial_1^+ \partial_1^-)\, \partial_1^+ \partial_1^- + (1 - \frac{h^2}{12} \partial_2^+ \partial_2^-)\, \partial_2^+ \partial_2^- .$$

A priori accuracy estimate

In the linear case, $G(y) = g\, y$ with $g \leq 0$ and $U_{ad} = L^2(\Omega)$. We have the following discrete optimality system

$$\begin{aligned} \Delta_h y_h + g\, y_h - p_h/\nu &= f_h \\ \Delta_h p_h + g\, p_h + y_h &= z_h \end{aligned}$$

Theorem

Let $y \in C^{k+2}(\bar{\Omega})$, $k = 2, 4$, and $p \in C^{l+2}(\bar{\Omega})$, $l = 2, 4$, be solutions to the optimality system, and let y_h and p_h be solutions to the discrete optimality system. Then there exists a constant c, depending on Ω, and independent of h, such that

$$|y_h - R_h y|_0^2 + \frac{1}{\nu}\, |p_h - R_h p|_0^2 \leq c\, (h^{2k}\, \|y\|^2_{C^{k+2}(\bar{\Omega})} + h^{2l}\, \frac{1}{\nu}\, \|p\|^2_{C^{l+2}(\bar{\Omega})}).$$

Results of numerical experiments give evidence that it appears to hold also in the presence of nonlinearity and of constraints.

Discretization of the optimality system

The one-dimensional expanded form of $\Delta_h v(x)$ is

$$\frac{1}{12h^2}(-v(x-2h)+16\,v(x-h)-30\,v(x)+16\,v(x+h)-v(x+2h)).$$

This scheme results in a system which is neither diagonally dominant nor of non-negative type [Bramble Hubbard]. Nevertheless it satisfies a max principle.

We can express the action of Δ (resp. $\tilde{\Delta}$) on the function v_h in the following compact form

$$\Delta_h v_h|_{ij} = \frac{1}{h^2}\Big(\sum_{s,t\in\omega_{ij},\, s,t\neq i,j} c_{st}\, v_{st} - c_{ij} v_{ij}\Big).$$

and for convenience set

$$A_{ij} = \sum_{s,t\in\omega_{ij},\, s,t\neq i,j} c^y_{st}\, y_{st} - h^2 f_{ij} \quad \text{and} \quad B_{ij} = \sum_{s,t\in\omega_{ij},\, s,t\neq i,j} c^p_{st}\, p_{st} - h^2 z_{ij}$$

Discretization of the optimality system (continue)

We have the following set of equations at i, j for the three scalar variables y_{ij}, p_{ij}, and u_{ij}:

$$
\begin{aligned}
A_{ij} - c_{ij}^{y}\, y_{ij} + h^2\, G(y_{ij}) - h^2\, u_{ij} &= 0 \\
B_{ij} - c_{ij}^{p}\, p_{ij} + h^2\, G'(y_{ij})\, p_{ij} + h^2\, y_{ij} &= 0 \\
(\nu u_{ij} - p_{ij}) \cdot (v_{ij} - u_{ij}) &\geq 0 \quad \text{for all } v_h \in U_{ad\,h}
\end{aligned}
$$

Solving these equations at each grid point in a given order results in a robust smoother.

Smoothing

- Compute the inverse of the Jacobian for the y, p system

$$J_{ij}^{-1} = \frac{1}{\det J_{ij}} \begin{pmatrix} -c_{ij}^{p} + h^2 G'(y_{ij}) & 0 \\ -h^2(1 + G''(y_{ij})\, p_{ij}) & -c_{ij}^{y} + h^2 G'(y_{ij}) \end{pmatrix}$$

- Define a local Newton update for y_{ij} and p_{ij} at i, j

$$\begin{pmatrix} y_{ij}(u_{ij}) \\ p_{ij}(u_{ij}) \end{pmatrix} = \begin{pmatrix} y_{ij} \\ p_{ij} \end{pmatrix} + J_{ij}^{-1} \begin{pmatrix} r_{ij}^{y}(u_{ij}) \\ r_{ij}^{p} \end{pmatrix},$$

Where

$$r_{ij}^{y} = -(A_{ij} - c_{ij}^{y}\, y_{ij} + h^2\, G(y_{ij}) - h^2\, u_{ij})$$

$$r_{ij}^{p} = -(B_{ij} - c_{ij}^{p}\, p_{ij} + h^2\, G'(y_{ij})\, p_{ij} + h^2\, y_{ij})$$

Smoothing (continue)

- Find u^*_{ij} such that $J'(y(u),u) = \nu u^*_{ij} - p_{ij}(u^*_{ij}) = 0$.

$$u^*_{ij} = \left(\nu + \frac{(1 + G''(y_{ij})\, p_{ij})\, h^4}{\det J_{ij}}\right)^{-1} \times$$
$$[p_{ij} + \frac{1}{\det J_{ij}} \left(h^2(1 + G''(y_{ij})\, p_{ij})\, r^y_{ij} + (c^y_{ij} - h^2 G'(y_{ij}))\, r^p_{ij}\right)].$$

- Set (projection)

$$u_{ij} = \begin{cases} u_{H\,ij} & \text{if} \quad u^*_{ij} \geq u_{H\,ij} \\ u^*_{ij} & \text{if} \quad u_{L\,ij} < u^*_{ij} < u_{H\,ij} \\ u_{L\,ij} & \text{if} \quad u^*_{ij} \leq u_{L\,ij}. \end{cases}$$

- Update state and adjoint variables $y_{ij} = y_{ij}(u_{ij})$ and $p_{ij} = p_{ij}(u_{ij})$.

Multigrid FAS-V(m_1, m_2)-Cycle [Brandt]

Set $B_1(w_1^{(0)}) \approx A_1^{-1}$ (e.g., iterating with S_1 starting with $w_1^{(0)}$). For $k = 2, \ldots, L$ define B_k in terms of B_{k-1} as follows.

1. Set the starting approximation $w_k^{(0)}$.
2. Pre-smoothing. Define $w_k^{(l)}$ for $l = 1, \ldots, m_1$, by
$$w_k^{(l)} = S_k(w_k^{(l-1)}, f_k).$$
3. Coarse grid correction.
Set $w_k^{(m_1+1)} = w_k^{(m_1)} + I_{k-1}^{k}(w_{k-1} - \hat{I}_k^{k-1} w_k^{(m_1)})$ where
$$w_{k-1} = B_{k-1}(\hat{I}_k^{k-1} w_k^{(m_1)}) \;\; I_k^{k-1}(f_k - A_k(w_k^{(m_1)})) + A_{k-1}(\hat{I}_k^{k-1} w_k^{(m_1)}) \; .$$
4. Post-smoothing. Define $w_k^{(l)}$ for $l = m_1 + 2, \cdots, m_1 + m_2 + 1$, by
$$w_k^{(l)} = S_k(w_k^{(l-1)}, f_k).$$
5. Set $B_k(w_k^{(0)})\, f_k = w_k^{(m_1+m_2+1)}$.

Local Fourier analysis

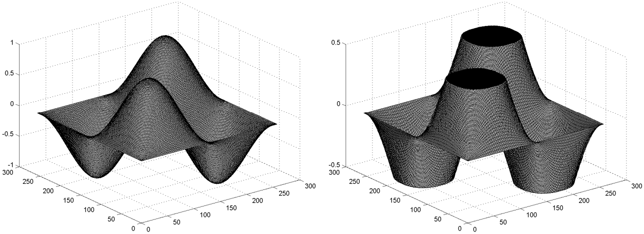
Estimates for convergence factors and smoothing factors (linear case)

		$\Delta_h y,\ \Delta_h p$		$\tilde{\Delta}_h y,\ \tilde{\Delta}_h p$	
(m_1, m_2)	ν	μ	$\rho(TG_k^{k-1})$	μ	$\rho(TG_k^{k-1})$
(1,1)	10^{-4}	0.5362	0.2429	0.5020	0.1939
(2,2)	10^{-4}	0.5362	0.1233	0.5020	0.0851
(1,1)	10^{-8}	0.6089	0.3457	0.5491	0.2772
(2,2)	10^{-8}	0.6089	0.1933	0.5491	0.1255
		$\Delta_h y,\ \tilde{\Delta}_h p$		$\tilde{\Delta}_h y,\ \Delta_h p$	
(m_1, m_2)	ν	μ	$\rho(TG_k^{k-1})$	μ	$\rho(TG_k^{k-1})$
(1,1)	10^{-4}	0.5346	0.2413	0.5346	0.2413
(2,2)	10^{-4}	0.5346	0.1215	0.5346	0.1215
(1,1)	10^{-8}	0.5787	0.3094	0.5787	0.3094
(2,2)	10^{-8}	0.5787	0.1566	0.5787	0.1566

A control-constrained nonlinear optimal control problem

$$\begin{aligned} \Delta y + y^4 &= u + f, \\ \Delta p + 4y^3 p + y &= z \\ (\nu u - p,\, v - u) &\geq 0 \quad \text{for all } v \in U_{ad} \end{aligned}$$

$$U_{ad} = \{u \in L^2(\Omega) \,|\, -1/2 \leq u(\mathbf{x}) \leq 1/2 \ \text{a.e. in } \Omega\}$$



Solution for $\nu = 10^{-6}$. The state (left) and the control (right). $z(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2) + O(\nu)$

Results of experiments

Fourth-order scheme versus second-order scheme for both equations

$\Delta_h y,\ \Delta_h p$						
ν	Mesh	$\lvert y_h - R_h y\rvert_0$	$\lvert u_h - R_h u\rvert_0$	$\lvert y_h - R_h z\rvert_0$	ρ	CPU s
10^{-3}	256^2	0.19(-8)	0.11(-7)	0.39(-1)	0.100	1.23
10^{-3}	1024^2	0.76(-11)	0.43(-10)	0.39(-1)	0.115	13.59
10^{-6}	256^2	0.55(-9)	0.39(-6)	0.39(-4)	0.116	1.28
10^{-6}	1024^2	0.21(-11)	0.15(-8)	0.39(-4)	0.117	13.50

$\tilde{\Delta}_h y,\ \tilde{\Delta}_h p$						
ν	Mesh	$\lvert y_h - R_h y\rvert_0$	$\lvert u_h - R_h u\rvert_0$	$\lvert y_h - R_h z\rvert_0$	ρ	CPU s
10^{-3}	256^2	0.24(-4)	0.13(-3)	0.39(-1)	0.04	0.81
10^{-3}	1024^2	0.15(-5)	0.86(-5)	0.39(-1)	0.03	10.79
10^{-6}	256^2	0.71(-5)	0.48(-2)	0.42(-4)	0.03	0.82
10^{-6}	1024^2	0.43(-6)	0.30(-3)	0.39(-4)	0.04	13.29

Results of experiments

Mixed fourth-order/second-order schemes

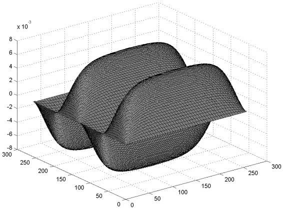
		$\Delta_h y$, $\tilde{\Delta}_h p$				
ν	Mesh	$\lvert y_h - R_h y\rvert_0$	$\lvert u_h - R_h u\rvert_0$	$\lvert y_h - R_h z\rvert_0$	ρ	CPU s
10^{-3}	256^2	0.54(-7)	0.99(-5)	0.39(-1)	0.115	1.26
10^{-3}	1024^2	0.35(-8)	0.62(-6)	0.39(-1)	0.114	12.68
10^{-6}	256^2	0.18(-8)	0.77(-6)	0.39(-4)	0.116	1.26
10^{-6}	1024^2	0.11(-9)	0.25(-7)	0.39(-4)	0.117	12.81

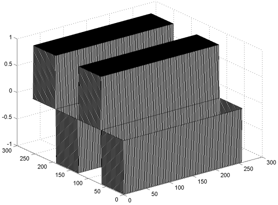
		$\tilde{\Delta}_h y$, $\Delta_h p$				
ν	Mesh	$\lvert y_h - R_h y\rvert_0$	$\lvert u_h - R_h u\rvert_0$	$\lvert y_h - R_h z\rvert_0$	ρ	CPU s
10^{-3}	256^2	0.24(-4)	0.12(-3)	0.39(-1)	0.04	0.82
10^{-3}	1024^2	0.15(-5)	0.80(-5)	0.39(-1)	0.03	12.34
10^{-6}	256^2	0.71(-5)	0.48(-2)	0.42(-4)	0.03	1.00
10^{-6}	1024^2	0.43(-6)	0.30(-3)	0.39(-4)	0.04	15.82

Bang-bang control

The case $\nu = 0$ and box constraints $u_L = -1$ and $u_H = 1$, and a non attainable target function given by $z(x_1, x_2) = \sin(4\pi x_1)$

	$\Delta_h y$ $\Delta_h p$		$\tilde{\Delta}_h y$, $\tilde{\Delta}_h p$	
Mesh	ρ	CPU s	ρ	CPU s
128×128	0.45	0.35	0.40	0.25
256×256	0.45	1.39	0.52	1.01
512×512	0.45	5.10	0.50	3.95
1024×1024	0.45	21.29	0.45	15.98
2048×2048	0.45	87.28	0.45	64.07

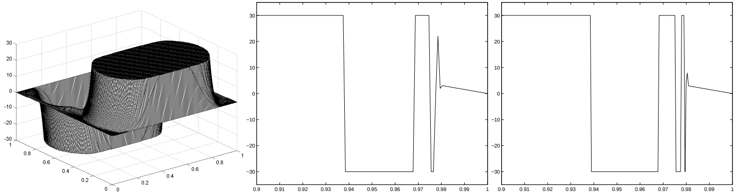




Bang-bang and chattering phenomena

Consider the objective function $z(x_1, x_2) = \sin(2\pi x_1)\sin(\pi x_2)$ and box constraints $u_L = -30$ and $u_H = 30$.

Results for $\nu = 10^{-6}$ and $\nu = 0$.



Bang-bang and switching of the control function for $x_1 = 3/4$ and $x_2 \in [0.9, 1]$ obtained with $\nu = 0$ on increasingly finer meshes: 1025×1025 and 8193×8193.

A state-constrained elliptic optimal control problem

$$\begin{cases} \min_{u \in L^2(Q)} J(y,u) & := & \frac{1}{2}||y-z||^2_{L^2(\Omega)} + \frac{\nu}{2}||u||^2_{L^2(\Omega)} & \\ \Delta y & = & u + f & \text{in } \Omega \\ y & = & 0 & \text{on } \partial\Omega \\ y_L \leq y & \leq & y_H & \text{in } \Omega \end{cases}$$

The solution approach through Lagrange multipliers associated with the state constraints leads to difficulties:

- The Lagrange multipliers associated with the state constraints are regular Borel measures.
- Methods relying on Lagrange multipliers must be adapted.

Remedy: Lavrentiev-type or Moreau-Yosida regularizations.

Regularized state-constrained optimal control problem

$$\begin{cases} \min_{u \in L^2(Q)} J(y,u) &:= \frac{1}{2}||y-z||^2_{L^2(\Omega)} + \frac{\nu}{2}||u||^2_{L^2(\Omega)} \\ \Delta y = u + f & \text{in } \Omega \\ y = 0 & \text{on } \partial\Omega \\ y_L \leq y - \lambda u \leq y_H & \text{in } \Omega, \quad \lambda > 0 \end{cases}$$

Set $v = y - \lambda u$. It becomes a 'control-constrained' optimal control problem

$$\begin{cases} \min_{v \in L^2(Q)} J(y,v) &:= \frac{1}{2}||y-z||^2_{L^2(\Omega)} + \frac{\nu}{2\lambda^2}||y-v||^2_{L^2(\Omega)} \\ \Delta y - y/\lambda + v/\lambda = f & \text{in } \Omega \\ y = 0 & \text{on } \partial\Omega \\ y_L \leq v \leq y_H & \text{in } \Omega \end{cases}$$

- The associated Lagrange multipliers are $L^2(\Omega)$.

Optimality system

The objective functional $J(y, v)$ is strictly convex and lower semicontinuous. One can prove existence and uniqueness of the optimal solution.

This solution is characterized by the following optimality system

$$
\begin{aligned}
\Delta y - y/\lambda + v/\lambda &= f \\
\Delta p - p/\lambda + (y - z) + k\,(y - v) &= 0 \\
(p/\lambda - k\,(y - v), t - v) &\geq 0 \quad \text{for all } t \in V_{ad}
\end{aligned}
$$

where $k = \nu/\lambda^2$ and

$$
V_{ad} = \{v \in L^2(\Omega) \,|\, y_L(x) \leq v(x) \leq y_H(x) \ \text{ a.e. in } \Omega\},
$$

Smoothing

- Compute the Jacobian for the y, p system

$$J_{ij}^{-1} = \frac{1}{\det J_{ij}} \begin{pmatrix} -(c_{ij}^p + \frac{h^2}{\lambda}) & 0 \\ -h^2(1+k) & -(c_{ij}^y + \frac{h^2}{\lambda}) \end{pmatrix},$$

- Define a local Newton update for y_{ij} and p_{ij} at i, j

$$\begin{pmatrix} y_{ij}(v_{ij}) \\ p_{ij}(v_{ij}) \end{pmatrix} = \begin{pmatrix} y_{ij} \\ p_{ij} \end{pmatrix} + J_{ij}^{-1} \begin{pmatrix} r_{ij}^y(v_{ij}) \\ r_{ij}^p(v_{ij}) \end{pmatrix},$$

- Find v_{ij}^* such that $\frac{p_{ij}(v_{ij}^*)}{\lambda} - k\,(y_{ij}(v_{ij}^*) - v_{ij}^*) = 0$. Set

$$v_{ij} = \begin{cases} y_{H\,ij} & \text{if} \quad v_{ij}^* \geq y_{H\,ij} \\ v_{ij}^* & \text{if} \quad y_{L\,ij} < v_{ij}^* < y_{H\,ij} \\ y_{L\,ij} & \text{if} \quad v_{ij}^* \leq y_{L\,ij}. \end{cases}$$

- Set $p_{ij} = p_{ij}(v_{ij})$ and $y_{ij} = y_{ij}(v_{ij})$

Numerical results

Consider $z(x_1, x_2) = \sin(2\pi x_1)\sin(\pi x_2)$ and $y_L(x) = -1/2$ and $y_L(x) = 1/2$. Mesh 513×513

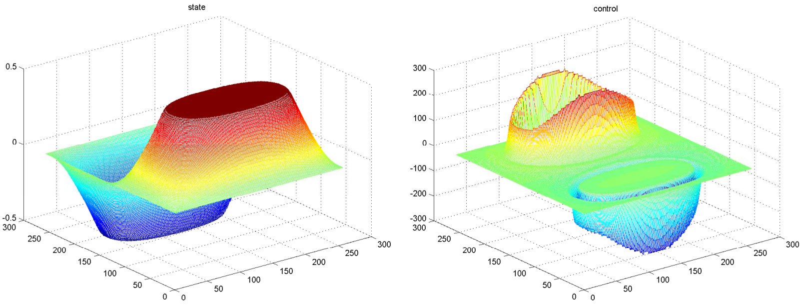
Convergence factors choosing $\nu = \lambda^2$

	λ	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
2nd-order	ρ	0.03	0.04	0.10	0.10	0.09
4th-order	ρ	0.09	0.07	0.12	0.08	0.08

Further results for ρ with $\nu = 10^{-6}$ and $\lambda = 10^{-3}$

Mesh	257×257	513×513	1025×1025
2nd-order	0.12	0.10	0.10
4th-order	0.15	0.12	0.09

Solution for a state-constrained optimal control problem



Multigrid convergence theory

1. Multigrid convergence theory for scalar elliptic equation

$$-\Delta y = f \text{ in } \Omega \text{ and } y = 0 \text{ on } \partial\Omega.$$

The matrix form of this problem is $\hat{A}_k\, y_k = f_k$.

Convergence results are given in terms of the error operator $\hat{E}_k := I_k - \hat{B}_k \hat{A}_k$. We have (for $m_1 = 1$, $m_2 = 0$)

$$\hat{E}_k\, y \quad = \quad [(I_k - I_{k-1}^k \hat{P}_{k-1}) + I_{k-1}^k\, \hat{E}_{k-1}\, \hat{P}_{k-1}]\hat{S}_k\, y.$$

Theorem 1: There exists a positive constant $\hat{\delta} < 1$ such that

$$(\hat{A}_k \hat{E}_k y, \hat{E}_k y)_k \leq \hat{\delta}^2\, (\hat{A}_k y, y)_k \quad \text{for all } y \in M_k, , \qquad k = L$$

2. Consider the decoupled symmetric system

$$\begin{aligned} -\nu\Delta y &= \nu g && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \\ -\Delta p &= z && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega. \end{aligned}$$

This system is exactly two copies of Poisson problem. Hence the multigrid convergence theory for this system inherits the properties of the scalar case.

Multigrid convergence theory (continue)

3. To analyze the optimality system define

$$\hat{\mathcal{A}}_k = \begin{pmatrix} \nu\, \hat{A}_k & 0 \\ 0 & \hat{A}_k \end{pmatrix}$$

and analogously $\hat{\mathcal{B}}_k$, $\hat{\mathcal{E}}_k$, etc., as counterparts of $\hat{B}_k$, $\hat{E}_k$, etc..

Theorem 2: There exists a positive constant $\hat{\delta} < 1$ such that

$$(\hat{\mathcal{A}}_L \hat{\mathcal{E}}_L \mathbf{w}, \hat{\mathcal{E}}_L \mathbf{w})_L \leq \hat{\delta}^2 (\hat{\mathcal{A}}_L \mathbf{w}, \mathbf{w})_L \quad \text{for all } \mathbf{w} = (y, p) \in \mathcal{M}_L,$$

Consider

$$\mathcal{A}_k = \hat{\mathcal{A}}_k + \mathcal{D}_k,$$

where

$$\mathcal{D}_k = \begin{pmatrix} 0 & -I_k \\ I_k & 0 \end{pmatrix}.$$

Note that $|(\mathcal{D}_k(u, v), (y, p))| \leq C\,|(u, v)|\,|(y, p)|$.
With $\mathcal{B}_k$, $\mathcal{A}_k$, etc., replacing $\hat{B}_k$, $\hat{A}_k$, etc..

$$\mathcal{E}_k = [\mathcal{I}_k - \mathcal{I}_{k-1}^k \mathcal{P}_{k-1} + \mathcal{I}_{k-1}^k\, \mathcal{E}_{k-1}\, \mathcal{P}_{k-1}]\, \mathcal{S}_k$$

Theorem 3: There exist positive constants h_0 and $\delta < 1$ such that for all $h_1 < h_0$ we have

$$(\hat{\mathcal{A}}_k \mathcal{E}_k \mathbf{w}, \mathcal{E}_k \mathbf{w})_k \leq \delta^2 (\hat{\mathcal{A}}_k \mathbf{w}, \mathbf{w})_k \quad \text{for all } \mathbf{w} \in \mathcal{M}_k, \qquad k = L$$

where $\delta = \hat{\delta} + C h_1$ and $\hat{\delta} = Ck/(Ck+1)$.

Outline of the talk II: Dynamical systems

- Reaction diffusion process controlled through source terms or boundary terms
- Optimality systems
- Discretization of the optimality system
- Smoothing and multigrid methods
- Applications
- Local Fourier analysis

Reaction diffusion process controlled through source terms

$$\begin{cases} \min_{u \in L^2(Q)} J(y,u) & & & \\ -\partial_t y + G(y) + \sigma\,\Delta y & = & u & \text{in } Q = \Omega \times (0,T) \\ y & = & y_0 & \text{in } \Omega \times \{t=0\} \\ y & = & 0 & \text{on } \Sigma = \partial\Omega \times (0,T) \end{cases}$$

Control required to

track a desired trajectory $y_d(\mathbf{x}, t)$

reach a desired terminal state $y_T(\mathbf{x})$

For this purpose, the following cost functional can be considered

$$J(y,u) = \frac{\alpha}{2} ||y - y_d||^2_{L^2(Q)} + \frac{\beta}{2} ||y(\cdot, T) - y_T||^2_{L^2(\Omega)} + \frac{\nu}{2} ||u||^2_{L^2(Q)}$$

Reaction diffusion process controlled through boundary terms

$$\begin{cases} \min_{u \in L^2(\Sigma)} J(y,u) & & & \\ -\partial_t y + G(y) + \sigma\, \Delta y & = & 0 & \text{in } Q = \Omega \times (0,T) \\ y & = & y_0 & \text{in } \Omega \times \{t=0\} \\ -\frac{\partial y}{\partial n} & = & u & \text{on } \Sigma = \partial\Omega \times (0,T) \end{cases}$$

Control required to
track a desired trajectory $y_d(\mathbf{x},t)$
reach a desired terminal state $y_T(\mathbf{x})$

For this purpose, the following cost functional can be considered

$$J(y,u) = \frac{\alpha}{2} ||y - y_d||^2_{L^2(Q)} + \frac{\beta}{2} ||y(\cdot,T) - y_T||^2_{L^2(\Omega)} + \frac{\nu}{2} ||u||^2_{L^2(\Sigma)}$$

Optimality systems

Distributed source control: The solution is characterized by

$$
\begin{aligned}
-\partial_t y + G(y) + \sigma \Delta y &= u && \text{in } Q \\
\partial_t p + G'(y)p + \sigma \Delta p + \alpha(y - y_d) &= 0 && \text{in } Q \\
\nu u - p &= 0 && \text{in } Q \\
y = 0,\ p &= 0 && \text{on } \Sigma
\end{aligned}
$$

Neumann boundary control: The optimal solution satisfies

$$
\begin{aligned}
-\partial_t y + G(y) + \sigma \Delta y &= 0 && \text{in } Q \\
\partial_t p + G'(y)p + \sigma \Delta p + \alpha(y - y_d) &= 0 && \text{in } Q \\
\nu u - p &= 0 && \text{on } \Sigma \\
-\frac{\partial y}{\partial n} = u,\ -\frac{\partial p}{\partial n} &= 0 && \text{on } \Sigma
\end{aligned}
$$

With initial condition $y(\mathbf{x}, 0) = y_0(\mathbf{x})$ for the state variable (evolving forward in time). And the terminal condition for the adjoint variable (evolving backward in time) $p(\mathbf{x}, T) = \beta(y(\mathbf{x}, T) - y_T(\mathbf{x}))$.

FDM Discretization

The optimality systems are discretized by, e.g., finite differences and backward Euler scheme.

Ω_h defines the set of interior mesh-points, (x_i, y_j), $2 \le i, j \le N_x$.

The space-time grid is defined by

$$Q_{h,\delta t} = \{(\mathbf{x}, t_m) : \mathbf{x} \in \Omega_h,\ t_m = (m-1)\delta t,\ 1 \le m \le N_t + 1,\ \delta t = T/N_t\}$$

Time difference operators

$$\partial_t^+ y_h^m = \frac{y_h^m - y_h^{m-1}}{\delta t} \quad \text{and} \quad \partial_t^- p_h^m = \frac{p_h^{m+1} - p_h^m}{\delta t}$$

Example distributed control:

$$\begin{aligned} -\partial_t^+ y_h^m + G(y_h^m) + \sigma \Delta_h y_h^m &= u_h^m \\ \partial_t^- p_h^m + G'(y_h^m) p_h^m + \sigma \Delta_h p_h^m + \alpha(y_h^m - y_{dh}^m) &= 0 \\ \nu u_h^m - p_h^m &= 0 \end{aligned}$$

Boundary control: Consider the optimality system on the boundary and discretize the boundary derivative using a second-order centered scheme.

Multigrid FAS-V(m_1, m_2)-Cycle – On space-time cylinder

Set $B_1(w_1^{(0)}) \approx A_1^{-1}$ (e.g., iterating with S_1 starting with $w_1^{(0)}$). For $k = 2, \dots, L$ define B_k in terms of B_{k-1} as follows.

1. Set the starting approximation $w_k^{(0)}$.
2. Pre-smoothing. Define $w_k^{(l)}$ for $l = 1, \dots, m_1$, by
$$w_k^{(l)} = S_k(w_k^{(l-1)}, f_k)$$
3. Coarse grid correction.
Set $w_k^{(m_1+1)} = w_k^{(m_1)} + I_{k-1}^{k}(w_{k-1} - \hat{I}_k^{k-1} w_k^{(m_1)})$ where
$$w_{k-1} = B_{k-1}(\hat{I}_k^{k-1} w_k^{(m_1)}) \;\; I_k^{k-1}(f_k - A_k(w_k^{(m_1)})) + A_{k-1}(\hat{I}_k^{k-1} w_k^{(m_1)}) \quad .$$
4. Post-smoothing. Define $w_k^{(l)}$ for $l = m_1 + 2, \cdots, m_1 + m_2 + 1$, by
$$w_k^{(l)} = S_k(w_k^{(l-1)}, f_k)$$
5. Set $B_k(w_k^{(0)})\, f_k = w_k^{(m_1+m_2+1)}$.

Multigrid components: Coarsening

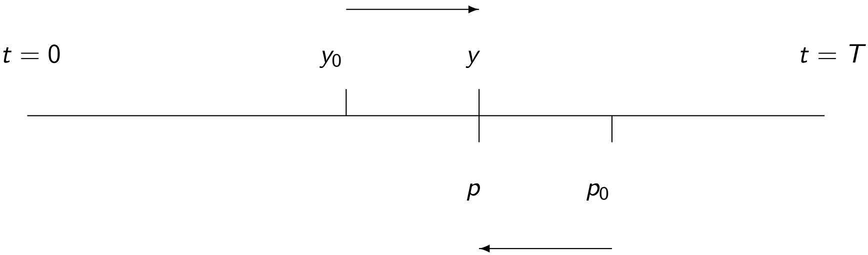
Coarsening strategy Consider L levels, $k = 1, \dots, L$.
Coarsening in the space directions: $h = h_k = h_1/2^{k-1}$.
Coarsening in time direction: $\delta t = \delta t_k = \delta t_1/s^{k-1}$,
$s = 1$ semicoarsening in space;
$s = 2$ standard time coarsening;
$s = 4$ double time coarsening.
Transfer operators
I_k^{k-1} denotes the injection operator for restriction.
I_{k-1}^k denotes bilinear interpolation in space and
If $s = 1$ no interpolation in time is needed,
if $s \in \{2, 4\}$ then I_{k-1}^k corresponds to bilinear interpolation in space and in time.

Multigrid components: Smoothing

Smoothing iteration: Two requirements

- Coupling between state and control variables.
- Preserve opposite orientation of state and adjoint equations.

Backward Euler discretization



Pointwise smoothing: Time-Splitted Collective Gauss-Seidel Iteration (TS-CGS)

$$\begin{pmatrix} y \\ p \end{pmatrix}^{(1)}_{i\,j\,m} = \begin{pmatrix} y \\ p \end{pmatrix}^{(0)}_{i\,j\,m} + \begin{pmatrix} -(1+4\sigma\gamma)+\delta t\, G' & -\delta t/\nu \\ \delta t(\alpha + G''\, p) & -(1+4\sigma\gamma)+\delta t\, G' \end{pmatrix}^{-1} \begin{pmatrix} r_y \\ r_p \end{pmatrix}_{i\,j\,m}$$

1. Set the starting approximation.
2. For $m = 2, \ldots, N_t$ do
3. For ij in, e.g., lexicographic order do

$$y^{(1)}_{i\,j\,m} = y^{(0)}_{i\,j\,m} + \frac{[-(1+4\sigma\gamma)+\delta t G']\, r_y(w) + \frac{\delta t}{\nu}\, r_p(w)}{[-(1+4\sigma\gamma)+\delta t G']^2 + \frac{\delta t^2}{\nu}(\alpha + G'' p)}\Big|^{(0)}_{i\,j\,m}$$

$$p^{(1)}_{i\,j\,N_t-m+2} = p^{(0)}_{i\,j\,N_t-m+2} + \frac{[-(1+4\sigma\gamma)+\delta t G']\, r_p(w) - \delta t(\alpha + G'' p)\, r_y(w)}{[-(1+4\sigma\gamma)+\delta t G']^2 + \frac{\delta t^2}{\nu}(\alpha + G'' p)}\Big|^{(0)}_{i\,j\,N_t-m+}$$

4. end

Blockwise smoothing: Time-Line Collective Gauss-Seidel Iteration (TL-CGS)

Consider the discrete optimality system at i,j and for all time steps and solve the resulting block-tridiagonal system.

$$\begin{pmatrix} y \\ p \end{pmatrix}_{ij}^{(1)} = \begin{pmatrix} y \\ p \end{pmatrix}_{ij}^{(0)} + M^{-1} \begin{pmatrix} r_y \\ r_p \end{pmatrix}_{ij}$$

The block-tridiagonal system has the following form

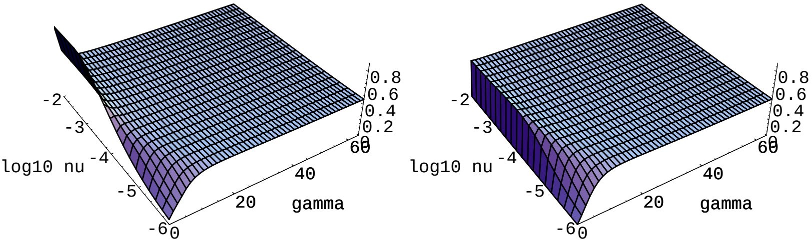
$$M = \begin{bmatrix} A_2 & C_2 & & & \\ B_3 & A_3 & C_3 & & \\ & B_4 & A_4 & C_4 & \\ & & & & C_{N_t} \\ & & & B_{N_t+1} & A_{N_t+1} \end{bmatrix}, A_m = \begin{bmatrix} -(1+4\sigma\gamma) + \delta t G' & -\frac{\delta t}{\nu} \\ \delta t(\alpha + G''\, p) & -(1+4\sigma\gamma) + \delta t G' \end{bmatrix}$$

$$B_m = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } C_m = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Centered at t_m, the entries B_m, A_m, C_m refer to the variables (y,p) at t_{m-1}, t_m, and t_{m+1}, respectively.

Smoothing factors

Pointwise smoother: Time-Splitted Collective Gauss-Seidel Iteration (TS-CGS)
Blockwise smoother: Time-Line Collective Gauss-Seidel Iteration (TL-CGS)

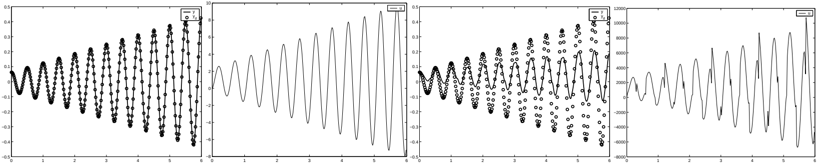


Smoothing factors of TS-CGS scheme (left) and TL-CGS scheme as functions of ν and $\gamma = \delta t/h^2$

Long time intervals: MG - Receding horizon techniques

Consider the optimal control problem of tracking y_d for $t \geq 0$.
Define time windows of size Δt. In each time window, an optimal control problem with tracking ($\alpha = 1$) and terminal observation ($\beta = 1$) is solved.

Multigrid Receding Horizon Scheme (MG-RH)
Solid fuel ignition model: $-\partial_t y + \sigma\, \Delta y + \delta e^y = f$

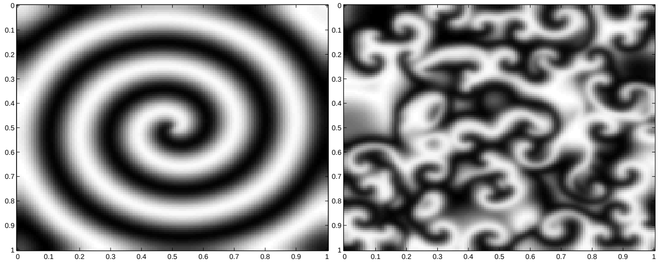


Time evolution of the state y and the desired trajectory y_d (left) and the optimal control u (right) at $(x_1, x_2) = (0.5, 0.5)$. Distributed control (left) and boundary control (right).

Prey/predator and chemical turbulence models: The Lambda-Omega system

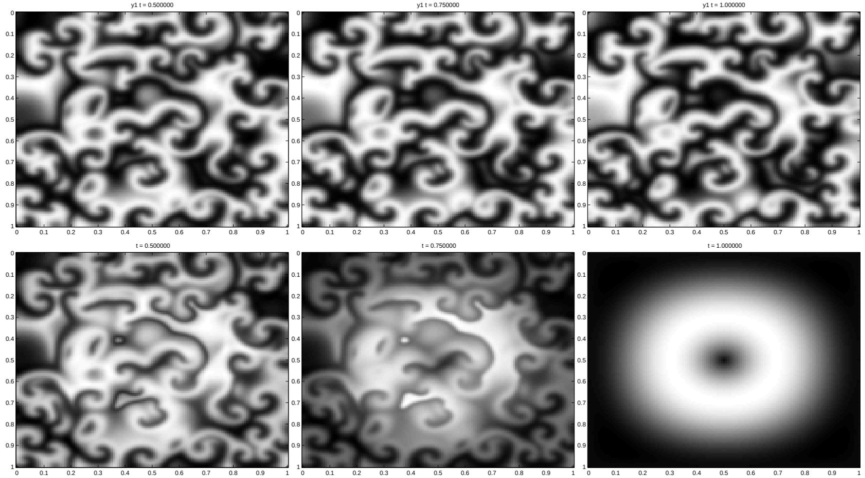
$$\frac{\partial}{\partial t}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} u(y_1,y_2) & -\omega(y_1,y_2) \\ \omega(y_1,y_2) & u(y_1,y_2) \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sigma\,\Delta\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

where (Kuramoto & Koga '81)
$u(y_1,y_2) = 1-(y_1^2+y_2^2)$ and $\omega(y_1,y_2) = -\beta(y_1^2+y_2^2)$.



The emergence of spatial patterns requires σ sufficiently small, e.g., $\sigma = 10^{-4}$. This Patterns are unstable for β large, e.g., $\beta = 2$.

Controlled and uncontrolled evolution



Convergence and tracking properties

Neumann b.c.. Mesh $N_x \times N_y \times N_t = 128 \times 128 \times 100L$; $\beta = 2$, $\sigma = 10^{-4}$. Use FAS-$V(2,2)$-cycle.

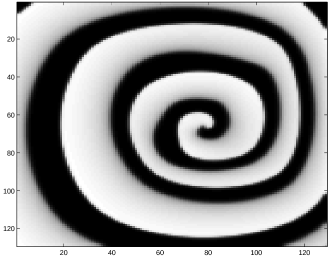
Convergence and tracking properties depending on $\nu_1 = \nu_2 = \nu$; $\beta_i = 0$ and $\alpha_i = 1$, $i = 1, 2$ (tracking type).

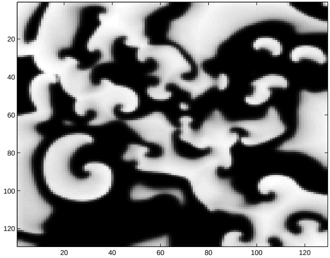
γ	ρ	$\|y_1 - y_{1d}\|_{L^2(Q)}$	$\|y_2 - y_{2d}\|_{L^2(Q)}$
10^{-3}	0.001	$8.58\,10^{-2}$	$7.00\,10^{-2}$
10^{-5}	0.001	$5.81\,10^{-3}$	$4.95\,10^{-3}$

Convergence and tracking properties depending on $\nu_1 = \nu_2 = \gamma$; $\beta_i = 1$ and $\alpha_i = 0$, $i = 1, 2$ (terminal observation).

γ	ρ	$\|y_1 - y_{1T}\|_{L^2(\Omega)}$	$\|y_2 - y_{2T}\|_{L^2(\Omega)}$
10^{-3}	0.66	$6.48\,10^{-4}$	$5.82\,10^{-4}$
*10^{-5}	0.62	$6.49\,10^{-6}$	$5.82\,10^{-6}$
10^{-7}	0.44	$6.49\,10^{-8}$	$5.82\,10^{-8}$

Cardiac arrhythmia





Aliev-Panfilov's model of cardiac excitation

$$\frac{\partial y_1}{\partial t} = -ky_1(y_1 - a)(y_1 - 1) - y_1\, y_2 + \sigma\, \Delta y_1 + u$$

$$\frac{\partial y_2}{\partial t} = [\epsilon_0 + \frac{\mu_1 y_2}{\mu_2 + y_1}][-y_2 - ky_1(y_1 - b - 1)]$$

y_1 - transmembrane potential and y_2 - membrane conductance

The parameters can be adjusted to reproduce the characteristics of the given cardiac tissue.

$k = 8,\ a = 0.1,\ b = 0.1,\ \epsilon_0 = 0.01,\ \mu_1 = 0.07,\ \mu_2 = 0.3,\ \sigma = 2.5\, 10^{-5}$.

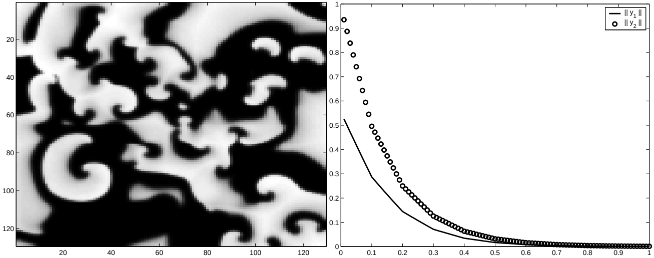
The model evolves from an initial planar (half) wave to a turbulent electrical pattern.

Determine a control response in the form of an electrical field able to drive the system from a turbulent pattern to a uniform pattern as in the case of no stimulus.

Control can be realized by a defibrillator

Towards the control of cardiac arrhythmia

Control u applies only to the transmembrane potential eq.



Initial state (left). Result of the optimal control mechanism: the desired fast decay of $|y_1|_0$ and $|y_2|_0$ in time can be observed.
Computed by TL-CGS multigrid receding horizon scheme with no tracking and zero terminal state; $\alpha = 0$, $\beta = 1$, $\nu = 0.1$. Ten time windows of size $\Delta t = 0.1$ are considered. On each window, the optimal control problem is solved efficiently on a $128 \times 128 \times 10$ grid by 3 FAS-$V(2,2)$-cycles.

Local Fourier analysis I: Introduction

Assume infinite grids and the Fourier components

$$\phi(\mathbf{j}, \boldsymbol{\theta}) = e^{i\mathbf{j}\cdot\boldsymbol{\theta}}$$

where $i = \sqrt{-1}$, $\mathbf{j} = (j_x, j_t) \in \mathbb{Z} \times \mathbb{Z}$, $\boldsymbol{\theta} = (\theta_x, \theta_t) \in [-\pi, \pi)^2$, and $\mathbf{j} \cdot \boldsymbol{\theta} = j_x\theta_x + j_t\theta_t$.

The frequency domain is spanned by the following two sets of frequencies (harmonics)

$$\boldsymbol{\theta}^{(0,0)} := (\theta_x, \theta_t) \text{ and } \boldsymbol{\theta}^{(1,0)} := (\overline{\theta}_x, \theta_t),$$

where $(\theta_x, \theta_t) \in ([-\pi/2, \pi/2) \times [-\pi, \pi))$ and $\overline{\theta}_x = \theta_x - sign(\theta_x)\pi$.
$\phi(\cdot, \boldsymbol{\theta}^{(0,0)})$: low frequencies components in space.
$\phi(\cdot, \boldsymbol{\theta}^{(1,0)})$: high frequencies components in space direction.
Both have all frequencies components in time direction.
Using semicoarsening, we have that $\phi(\mathbf{j}, \boldsymbol{\theta}^{(0,0)}) = \phi(\mathbf{j}, \boldsymbol{\theta}^{(1,0)})$ on the coarse grid.
The action of multigrid:
1) reduce the high frequency error components by smoothing
2) reduce the low frequency error components by coarse-grid correction

$$\mathbf{TG}_k^{k-1} = \mathcal{S}_k^{m_2} \mathbf{CG}_k^{k-1} \mathcal{S}_k^{m_1} = \mathcal{S}_k^{m_2} \left[\mathcal{I}_k - \mathcal{I}_{k-1}^k (\mathcal{A}_{k-1})^{-1} \mathcal{I}_k^{k-1} \mathcal{A}_k\right] \mathcal{S}_k^{m_1}$$

Denote with $E_k^{\theta} = span[\phi_k(\cdot, \boldsymbol{\theta}^{\boldsymbol{\alpha}}) : \boldsymbol{\alpha} \in \{(0,0), (1,0)\}]$ and assume that $(\mathcal{A}_{k-1})^{-1}$ exists
The twogrid operator $\mathbf{TG}_k^{k-1}$ on the space $E_k^{\theta} \times E_k^{\theta}$ is represented by a 4×4 matrix (Fourier symbol)

$$\mathbf{TG}_k^{k-1}(\boldsymbol{\theta}) = \hat{\mathcal{S}}_k(\boldsymbol{\theta})^{m_2}\ \mathbf{CG}_k^{k-1}(\boldsymbol{\theta})\ \hat{\mathcal{S}}_k(\boldsymbol{\theta})^{m_1}$$

Local Fourier analysis I: Steps 1.-3.

Consider the action of the operators on the couple $(y,p) \in E_k^\theta \times E_k^\theta$, where:

$$y = \sum_{p=0,1} Y^{(p,0)}\, \phi_k(\mathbf{j}, \boldsymbol{\theta}^{(p,0)}) \text{ and } p = \sum_{p=0,1} P^{(p,0)}\, \phi_k(\mathbf{j}, \boldsymbol{\theta}^{(p,0)}).$$

Here, $Y^{\boldsymbol{\alpha}}$ and $P^{\boldsymbol{\alpha}}$ are the Fourier amplitudes.

1. Compute the symbol $\hat{\mathcal{S}}_k(\boldsymbol{\theta})$ (e.g. TS-CGS)

$$\hat{\mathcal{S}}_k(\boldsymbol{\theta}) = diag\{\sigma(\boldsymbol{\theta}^{(0,0)}), \sigma(\boldsymbol{\theta}^{(1,0)}), \sigma(\boldsymbol{\theta}^{(0,0)}), \sigma(\boldsymbol{\theta}^{(1,0)})\},$$

where

$$\sigma(\boldsymbol{\theta}^{(p,q)}) = \frac{\nu\gamma(2\gamma+1)e^{i\theta_x^p}}{\delta t^2 + \nu[(2\gamma+1)^2 - \gamma(2\gamma+1)e^{-i\theta_x^p} - (2\gamma+1)e^{-i\theta_t^q}]}\,.$$

2. The symbol of full-weighting restriction operator (in space)

$$\hat{\mathcal{I}}_k^{k-1}(\boldsymbol{\theta}) = \frac{1}{2} \begin{matrix} (1+\cos(\theta_x)) & (1-\cos(\theta_x)) & 0 & 0 \\ 0 & 0 & (1+\cos(\theta_x)) & (1-\cos(\theta_x)) \end{matrix}\,,$$

Bilinear prolongation $\hat{\mathcal{I}}_{k-1}^k(\boldsymbol{\theta}) = \hat{\mathcal{I}}_k^{k-1}(\boldsymbol{\theta})^T$.

3. The symbol of the fine-grid operator is

$$\mathcal{A}_k(\boldsymbol{\theta}) = \begin{matrix} a_y(\boldsymbol{\theta}^{(0,0)}) & 0 & -\delta t/\nu & 0 \\ 0 & a_y(\boldsymbol{\theta}^{(1,0)}) & 0 & -\delta t/\nu \\ \delta t & 0 & a_p(\boldsymbol{\theta}^{(0,0)}) & 0 \\ 0 & \delta t & 0 & a_p(\boldsymbol{\theta}^{(1,0)}) \end{matrix}\,,$$

where

$$a_y(\boldsymbol{\theta}^{(p,q)}) = 2\gamma\cos(\theta_x^p) - e^{-i\theta_t^q} - 2\gamma - 1 \quad \text{and} \quad a_p(\boldsymbol{\theta}^{(p,q)}) = 2\gamma\cos(\theta_x^p) - e^{i\theta_t^q} - 2\gamma - 1$$

Local Fourier analysis III: Step 4.

4. The symbol of the coarse-grid operator follows

$$\mathcal{A}_{k-1}(\boldsymbol{\theta}) = \begin{pmatrix} \gamma \cos(2\theta_x)/2 - e^{-i\theta_t} - \gamma/2 - 1 & -\delta t/\nu \\ \delta t & \gamma \cos(2\theta_x)/2 - e^{i\theta_t} - \gamma/2 - 1 \end{pmatrix}.$$

Note: δt remains unchanged while $\gamma \longrightarrow \gamma/4$ by coarsening.
Characterization of the smoothing property of $\mathcal{S}_k$: assume an ideal coarse grid correction.
Let Q_k^{k-1} be a projection operator which annihilates the low frequency error components and leaves the high frequency error components unchanged

$$Q_k^{k-1}\,\phi(\boldsymbol{\theta},\cdot) = \begin{cases} 0 & \text{if } \boldsymbol{\theta} = \boldsymbol{\theta}^{(0,0)}, \\ \phi(\cdot,\boldsymbol{\theta}) & \text{if } \boldsymbol{\theta} = \boldsymbol{\theta}^{(1,0)}. \end{cases}$$

On the space $E_k^{\theta} \times E_k^{\theta}$ we then have

$$\mathbf{Q}_k^{k-1}(\boldsymbol{\theta}) = \begin{pmatrix} Q_k^{k-1} & 0 \\ 0 & Q_k^{k-1} \end{pmatrix} \quad \text{for } \boldsymbol{\theta} \in ([-\pi/2,\pi/2) \times [-\pi,\pi)).$$

The smoothing property of $\mathcal{S}_k$ is defined as follows

$$\mu = \max\{ r(\mathbf{Q}_k^{k-1}(\boldsymbol{\theta})\,\hat{\mathcal{S}}_k(\boldsymbol{\theta})) : \ \boldsymbol{\theta} \in ([-\pi/2,\pi/2) \times [-\pi,\pi)) \},$$

where r is the spectral radius.

Local Fourier analysis IV: Estimates

The twogrid convergence factor is given by

$$\rho(\mathbf{TG}_k^{k-1}) = \sup\{r(\mathbf{TG}_k^{k-1}(\boldsymbol{\theta})) : \boldsymbol{\theta} \in ([-\pi/2, \pi/2) \times [-\pi, \pi))\}.$$

Semicoarsening: (left) smoothing factor as a function of ν and γ; (right) twogrid convergence factor as a function of ν and γ ($m_1 = m_2 = 1$).

Assuming an ideal coarse grid correction, the convergence factor of the twogrid scheme is given by

$$\rho \approx \mu^{m_1+m_2}$$

Fourier analysis: ρ for semicoarsening; $\delta t = 1/64$.

	ν		
γ	10^{-4}	10^{-6}	10^{-8}
16	0.126	0.102	$4.28\,10^{-4}$
32	0.130	0.120	$5.13\,10^{-3}$
48	0.131	0.127	$1.72\,10^{-2}$
64	0.132	0.130	$3.38\,10^{-2}$

Results of experiments with semicoarsening

$\nu = 10^{-4}$				
$N_x \times N_y \times N_t$	γ	ρ	$\|\|y-z\|\|_0$	$\|\|r_y\|\|_0,\ \|\|r_p\|\|_0$
$32 \times 32 \times 64$	16	0.146	$1.55\,10^{-3}$	$4.5\,10^{-10},\ 7.6\,10^{-12}$
$64 \times 64 \times 64$	64	0.164	$1.55\,10^{-3}$	$9.1\,10^{-10},\ 1.0\,10^{-11}$
$128 \times 128 \times 64$	256	0.159	$1.55\,10^{-3}$	$1.1\,10^{-9},\ 8.1\,10^{-12}$
$\nu = 10^{-6}$				
$N_x \times N_y \times N_t$	γ	ρ	$\|\|y-z\|\|_0$	$\|\|r_y\|\|_0,\ \|\|r_p\|\|_0$
$32 \times 32 \times 64$	16	0.147	$4.03\,10^{-5}$	$1.4\,10^{-10},\ 1.9\,10^{-13}$
$64 \times 64 \times 64$	64	0.140	$4.23\,10^{-5}$	$2.6\,10^{-10},\ 2.1\,10^{-13}$
$128 \times 128 \times 64$	256	0.165	$4.27\,10^{-5}$	$3.3\,10^{-10},\ 5.8\,10^{-13}$
$\nu = 10^{-8}$				
$N_x \times N_y \times N_t$	γ	ρ	$\|\|y-z\|\|_0$	$\|\|r_y\|\|_0,\ \|\|r_p\|\|_0$
$32 \times 32 \times 64$	16	0.008	$9.09\,10^{-7}$	$4.7\,10^{-15},\ 1.1\,10^{-18}$
$64 \times 64 \times 64$	64	0.06	$1.73\,10^{-6}$	$9.1\,10^{-12},\ 7.6\,10^{-16}$
$128 \times 128 \times 64$	256	0.134	$2.06\,10^{-6}$	$9.1\,10^{-11},\ 8.1\,10^{-15}$

All results with $m_1 = m_2 = 1$

References

- E. Arian and S. Ta'asan, *Multigrid one shot methods for optimal control problems: Infinite dimensional control*, ICASE Report No. 94-52, 1994.
- A. Borzì, *Smoothers for control- and state-constrained optimal control problems*, Computing and Visualization in Science.
- A. Borzì, *High-order discretization and multigrid solution of elliptic nonlinear constrained optimal control problems*, Journal of Computational and Applied Mathematics.
- A. Borzì, *Multigrid methods for parabolic distributed optimal control problems*, J. Comp. Appl. Math., 157 (2003), pp. 365–382.
- A. Borzì, *Space-time multigrid methods for solving unsteady optimal control problems*, (Chapter 5) in L.T. Biegler, O. Ghattas, M. Heinkenschloss, D. Keyes and B. van Bloemen Waanders (Eds.), Real-Time PDE-Constrained Optimization, Computational Science and Engineering, Vol. 3, SIAM, Philadelphia, 2007.
- A. Borzì and G. Borzì, *An algebraic multigrid method for a class of elliptic differential systems*, SIAM J. Sci. Comp., 25(1) (2003), pp. 302–323.
- A. Borzì and R. Griesse, *Distributed optimal control of lambda-omega systems*, J. Num. Math., 14 (2006), pp. 17–40.
- A. Borzì, K. Kunisch, and D.Y. Kwak, *Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system*, SIAM J. Control Opt., 41(5) (2003), pp. 1477-1497.
- Th. Dreyer, B. Maar, and V. Schulz, *Multigrid optimization in applications*, J. Comput. Appl. Math., 120 (2000), pp. 67–84.