Sparse controls for elliptic and parabolic partial differential equations

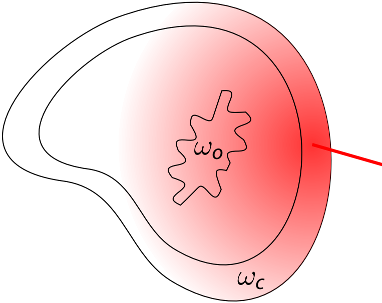
Eduardo Casas **Christian Clason** Karl Kunisch

Institute for Mathematics and Scientific Computing
Karl-Franzens-Universität Graz

1st IFAC Workshop on Control of Systems Modeled by PDEs
Paris, September 25, 2013

Motivation

- Optimization of light source locations in optical tomography
- Standard approach (discrete): choose combinations from initial set of candidate locations
  $\rightsquigarrow$ combinatorial in DOFs
- Here: Consider fictitious distributed "control field", apply sparse control techniques [Stadler 2009]
  $\rightsquigarrow$ localization of sources
- Goal: Homogeneous illumination (application in photochemotherapy)



Motivation: sparse control

$$\min_u \frac{1}{2}\|y - z\|_{L^2}^2 + \alpha\|u\|_{L^1} \quad \text{with} \quad Ay = u$$

- z desired state, A partial differential operator
- L^1 norm promotes sparsity in optimization
- But: L^1 lacks necessary compactness for existence of minimizers (Dunford–Pettis requires additional equi-integrability)
- $\rightsquigarrow$ Consider minimizers in measure space $\mathcal{M} \supset L^1$

  (alternative: control constraints)

  [Stadler 2009, D./G. Wachsmuth 2011, Herzog/Casas/G. Wachsmuth 2012]

Numerical solution

Difficulty: $\mathcal{M}$ not reflexive, hence

- choice of topology important
  (weak topology $\neq$ weak-$*$ topology)
- norm not differentiable (but convex)

Approach:

1. Fenchel duality in appropriate topology
2. conforming discretization of measure space
3. semismooth Newton method

Control problem

$$\begin{cases} \min\limits_{u\in\mathcal{M}(\Omega)} \frac{1}{2}\|y-z\|^2_{L^2(\Omega)} + \alpha\|u\|_{\mathcal{M}(\Omega)} \\ \text{with} \quad Ay = u, \quad y|_{\partial\Omega} = 0 \end{cases}$$

- $\mathcal{M}(\Omega) = C_0(\Omega)^*$ space of Radon measures, norm
$$\|u\|_{\mathcal{M}(\Omega)} = \sup_{\|\varphi\|_{C_0(\Omega)}\leq 1} \int \varphi\, du \left(= \|u\|_{L^1(\Omega)} \text{ for } u \in L^1(\Omega)\right)$$
- Unique solution $y \in W_0^{1,q}(\Omega)$, $1 < q < \frac{n}{n-1}$, for all $u \in \mathcal{M}(\Omega)$
- (Non-negativity, partial control and observation possible)

Optimality system

- Existence of minimizer $\bar{u} \in \mathcal{M}(\Omega)$ by standard arguments (weak-$*$ topology on $\mathcal{M}(\Omega)$), unique if A injective
- Fenchel duality: transformation to state constraint problem; minimizer $\bar{u} \in \mathcal{M}(\Omega)$, adjoint $\bar{p} \in C_0(\Omega)$ satisfies for any $\gamma > 0$

Optimality conditions

$$\begin{cases} A\bar{y} = \bar{u} \\ A^*\bar{p} = \bar{y} - \bar{z} \\ -\bar{u} = \max(0, -\bar{u} + \gamma(\bar{p} - \alpha)) + \min(0, -u + \gamma(\bar{p} + \alpha)) \end{cases}$$

Numerical solution: discretization

Goal:

- conforming discretization $U_h \subset \mathcal{M}(\Omega)$
- discrete optimality conditions for coefficients
  $\rightsquigarrow$ numerical solution of measure space problem
- (discretize-then-optimize = optimize-then-discretize)

Approach:

1. choose (adjoint-consistent) Galerkin discretization $Y_h \subset C_0(\Omega)$
2. construct discrete dual $U_h = Y_h^*$ with respect to appropriate consistent topology

Discretization: (adjoint) state space

- $\mathcal{T}_h$ shape regular triangulation of $\Omega \subset \mathbb{R}^n$, mesh size h
- $\{x_j\}_{j=1}^{N_h}$ interior nodes of $\mathcal{T}_h$
- $\{e_j\}_{j=1}^{N_h}$ nodal basis of continuous piecewise linear functions

$$Y_h = \left\{ y_h \in C_0(\Omega) : y_h = \sum_{j=1}^{N_h} y_j e_j, \text{ with } \{y_j\}_{j=1}^{N_h} \subset \mathbb{R} \right\}$$

with norm

$$\|y_h\|_C = \max_{1 \le j \le N_h} |y_j| =: |\vec{y}_h|_\infty$$

($\vec{y}_h \in \mathbb{R}^{N_h}$ coefficient vector)

Spatial discretization: control space

$$U_h = \left\{ u_h \in \mathcal{M}(\Omega) : u_h = \sum_{j=1}^{N_h} u_j \delta_{x_j}, \text{ with } \{u_j\}_{j=1}^{N_h} \subset \mathbb{R} \right\}$$

with norm

$$\|u_h\|_{\mathcal{M}} = \sup_{\|v\|_C = 1} \sum_{j=1}^{N_h} u_j \langle \delta_{x_j}, v \rangle = \sum_{j=1}^{N_h} |u_j| =: |\vec{u}_h|_1$$

$\rightsquigarrow$ U_h topological dual of Y_h with respect to duality pairing

$$\langle u_h, y_h \rangle_{\mathcal{M}, C_0} = \sum_{j=1}^{N_h} u_j y_j = \vec{u}_h^T \vec{y}_h$$

Discretization

Semidiscrete control problem

$$\min_{u \in \mathcal{M}(\Omega)} \frac{1}{2} \|y_h(u) - z\|^2_{L^2(\Omega)} + \alpha \|u\|_{\mathcal{M}(\Omega)}$$

- $y_h(u) \in Y_h$ solves Galerkin approximation of state equation
- Control not discretized
  (variational discretization [Hinze 2005])
- Solution in $\mathcal{M}(\Omega)$, unique solution $\bar{u}_h \in U_h$
- Convergence for $h \to 0$

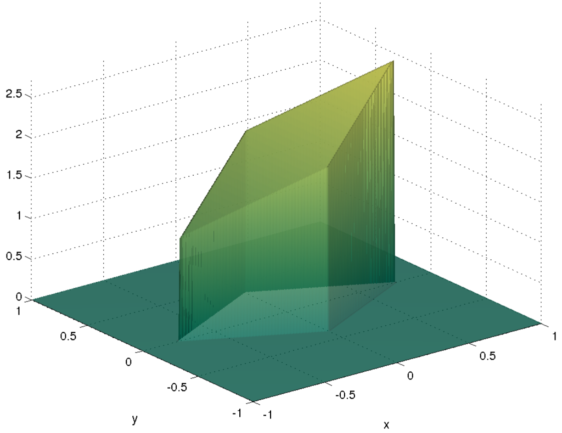
Numerical solution

Discrete optimality system

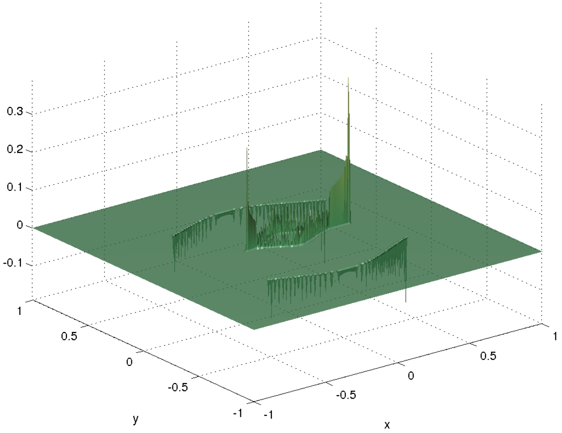
$$\begin{cases} A_h\vec{y}_h = \vec{u}_h \\ A_h^T\vec{p}_h = M_h(\vec{y}_h - \vec{z}) \\ -\vec{u}_h = \max(0, -\vec{u}_h + \gamma(\vec{p}_h - \alpha)) + \min(0, -\vec{u}_h + \gamma(\vec{p}_h + \alpha)) \end{cases}$$

- A_h stiffness matrix, M_h mass matrix, max / min componentwise
- Semismooth in $\mathbb{R}^{N_h}$ $\rightsquigarrow$ semismooth Newton method
- Combine with (generalized) Moreau–Yosida regularization, homotopy $\gamma \to \infty$ (globalization)

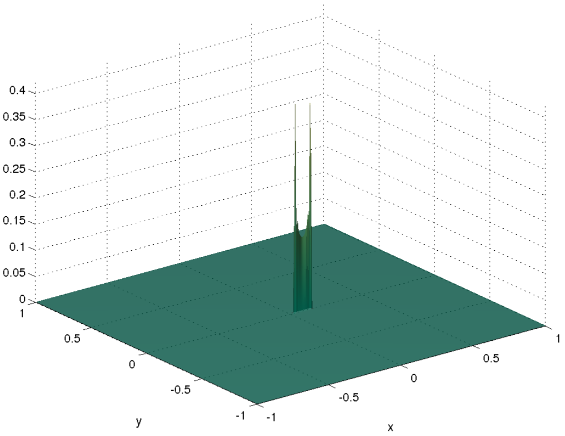
Numerical example: target



Numerical example: control ($\alpha = 10^{-2}$)



Numerical example: control ($\alpha = 10^{-1}$)



Control problem

$$\begin{cases} \min_{u \in \mathrm{L}^2(\mathcal{M})} \frac{1}{2}\|y - z\|^2_{L^2(\Omega_T)} + \alpha\|u\|_{\mathrm{L}^2(\mathcal{M})} \\ \text{with} \quad y_t + Ay = u, \quad y|_{\partial\Omega} = 0, \quad y(0) = 0 \end{cases}$$

- $y \in L^2(\Omega_T)$ very weak solution of state equation
- Existence, uniqueness, optimality conditions for control $\bar{u} \in \mathrm{L}^2(0, T; \mathcal{M}(\Omega)) = \mathrm{L}^2(0, T; C_0(\Omega))^*$, duality pairing

$$\langle u, z \rangle_{\mathrm{L}^2(\mathcal{M}), \mathrm{L}^2(C_0)} = \int_0^T \langle u(t), z(t) \rangle_{\mathcal{M}, C_0} \, \mathrm{d}t$$

- $\rightsquigarrow$ sparsity in space, not time

Time discretization: (adjoint) state space

- Temporal grid $0 = t_0 < t_1 < \ldots < t_{N_\tau} = T$
- Time steps $\tau_k = t_k - t_{k-1}$,
- $\{\chi_k e_j\}_{j,k}$ basis, piecewise constant in time, linear in space

$$Y_\sigma = \{y_\sigma \in \mathrm{L}^2(0,T;Y_h) : y_\sigma|_{I_k} \in Y_h,\ 1 \leq k \leq N_\tau\}$$

with norm

$$\|y_\sigma\|^2_{\mathrm{L}^2(C_0)} = \sum_{k=1}^{N_\tau} \tau_k \left(\max_{1\leq j\leq N_h} |y_{kj}|\right)^2 =: \sum_{k=1}^{N_\tau} \tau_k |\vec{y}_k|^2_\infty$$

Time discretization: control space

$$U_\sigma = \{u_\sigma \in L^2(0,T;U_h) : u_\sigma|_{I_k} \in U_h,\ 1 \le k \le N_\tau\}$$

with norm

$$\|u_\sigma\|^2_{L^2(\mathcal{M})} = \int_0^T \Big\| \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj}\chi_k \delta_{x_j} \Big\|^2_{\mathcal{M}} dt = \sum_{k=1}^{N_\tau} \tau_k |\vec{u}_k|_1^2$$

$\rightsquigarrow$ U_σ topological dual of Y_σ with respect to duality pairing

$$\langle u_\sigma, y_\sigma \rangle_{L^2(\mathcal{M}),L^2(C_0)} = \sum_{k=1}^{N_\tau} \tau_k \sum_{j=1}^{N_h} u_{kj} y_{kj} = \sum_{k=1}^{N_\tau} \tau_k (\vec{u}_k^T \vec{y}_k)$$

Semidiscrete control problem

$$\min_{u\in \mathrm{L}^2(\mathcal{M})} \frac{1}{2}\|y_\sigma(u) - z\|^2_{L^2(\Omega_T)} + \alpha\|u\|_{\mathrm{L}^2(\mathcal{M})}$$

- $y_\sigma(u) \in Y_\sigma$ solves dG(0)cG(1) discretization of state equation
- Unique minimizer $\bar{u}_\sigma \in U_\sigma \subset \mathrm{L}^2(0,T;\mathcal{M}(\Omega))$ with

$$\bar{u}_\sigma = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h} \bar{u}_{kj}\chi_k\delta_{x_j}$$

$\rightsquigarrow$ compute coefficients $\bar{u}_{kj}, 1 \leq k \leq N_\tau,\ 1 \leq j \leq N_h$

- Problem: characterization of subdifferential of $\mathrm{L}^2(\mathcal{M})$ norm
  $\rightsquigarrow$ use discrete Fenchel duality, reformulation

Discrete dual problem

$$\min_{p_\sigma \in \mathbb{R}^{N_\sigma}} \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k ([L_\sigma^T p_\sigma]_k - M_h z_k)^T M_h^{-1} ([L_\sigma^T p_\sigma]_k - M_h z_k) + \delta_\alpha(p_\sigma)$$

- $L_\sigma y_\sigma = u_\sigma$ dG(0)cG(1) discretization of state equation,

$$\delta_\alpha(q) := \begin{cases} 0 & \text{if } \left(\sum_{k=1}^{N_\tau} \tau_k |q_k|_\infty^2\right)^{1/2} \leq \alpha \\ \infty & \text{otherwise} \end{cases}$$

- Spatiotemporal coupling of $\delta_\alpha \rightsquigarrow$ define

$$c_k := |p_k|_\infty, \quad 1 \leq k \leq N_\tau$$

Reformulation

$$\begin{cases} \min\limits_{p_\sigma \in \mathbb{R}^{N_\sigma}, c_\sigma \in \mathbb{R}^{N_\tau}} \dfrac{1}{2} \displaystyle\sum_{k=1}^{N_\tau} \tau_k ([L_\sigma^T p_\sigma]_k - M_h z_k)^T M_h^{-1} ([L_\sigma^T p_\sigma]_k - M_h z_k) \\ \text{with} \quad |p_k|_\infty \leq c_k \quad 1 \leq k \leq N_\tau \quad \text{and} \quad \displaystyle\sum_{k=1}^{N_\tau} \tau_k c_k^2 = \alpha^2 \end{cases}$$

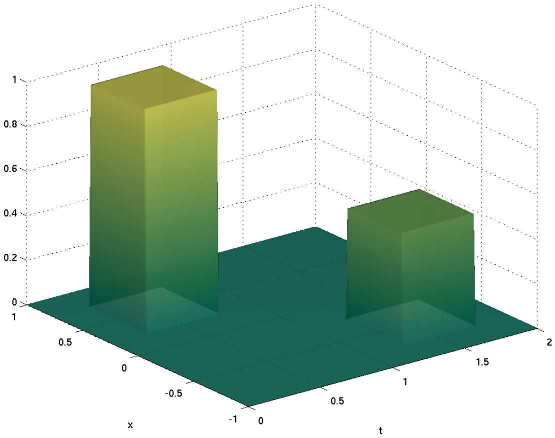
- $c_\sigma = (c_1, \ldots, c_{N_\tau})^T \in \mathbb{R}^{N_\tau}$
- Equivalent reformulation $\rightsquigarrow$ unique solution $(\bar{p}_\sigma, \bar{c}_\sigma)$
- Optimality conditions (Maurer–Zowe condition)
- Complementarity formulation

Optimality conditions

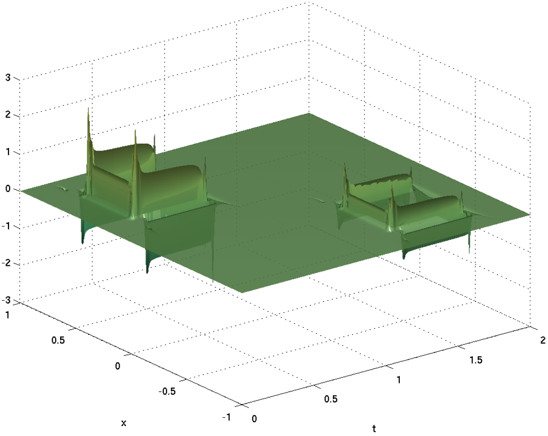
$$\begin{cases} L_\sigma \bar{y}_\sigma - \bar{u}_\sigma = 0, \qquad L_\sigma^T \bar{p}_\sigma - M_\sigma(\bar{y}_\sigma - z_\sigma) = 0 \\ \bar{u}_k + \max(0, -\bar{u}_k + \gamma(\bar{p}_k - \bar{c}_k)) + \min(0, -\bar{u}_k + \gamma(\bar{p}_k + \bar{c}_k)) = 0 \\ \sum_{j=1}^{N_h} [-\max(0, -\bar{u}_k + \gamma(\bar{p}_k - \bar{c}_k)) + \min(0, -\bar{u}_k + \gamma(\bar{p}_k + \bar{c}_k))]_j \\ \qquad + 2\lambda \bar{c}_k = 0 \\ \sum_{k=1}^{N_\tau} \tau_k \bar{c}_k^2 - \alpha^2 = 0 \end{cases}$$

Semismooth in $\mathbb{R}^{N_\sigma}$ $\rightsquigarrow$ solve using semismooth Newton method

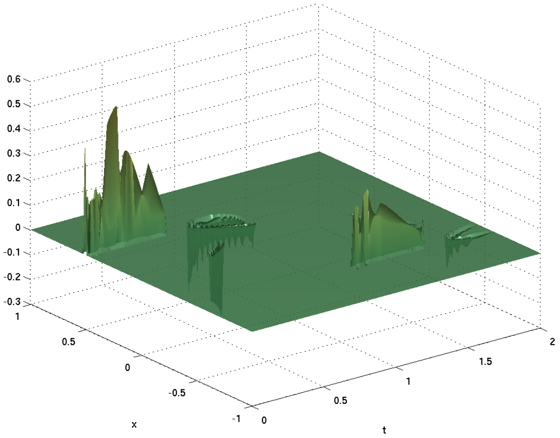
Numerical example (1D): target



Numerical example (1D): control ($\alpha = 10^{-3}$)



Numerical example (1D): control ($\alpha = 10^{-1}$)



Conclusion

- Measure space appropriate framework for sparse control
- Conforming discretization retains structural properties
  $\rightsquigarrow$ numerical solution of measure space problem
- Efficient solution by convex analysis, semismooth Newton method

Outlook:

- optimal sensor placement via sparse control
- directional sparsity (control in $\mathcal{M}(\Omega; L^2(0,T))$)
- optimal placement/identification of dipoles (control in $C^1(\Omega)^*$)

Preprints, codes:

`http://www.uni-graz.at/~clason/publications.html`