Bundle-Free Implicit Programming Approaches for the Optimal Control of Variational Inequalities of the First and Second Kind

Thomas M. Surowiec

Humboldt-Universität zu Berlin
Department of Mathematics
Joint Work with M. Hintermüller.

August 8, 2014

The "Lower-Level" Problem/Variational Inequality

Typical Variational Problems of Interest

- Contact problems in mechanics/free boundary problems
- Phase-field models with obstacle/nonsmooth potentials
- Volatility calibration in American options (Black-Scholes model)
- Parameter identification in image processing

The "Upper-Level" Problem/MPEC

How do we...

Bilevel Programming/Optimal Control/Parameter ID Problem

- Contact problems in mechanics/free boundary problems
  ...choose the applied force to achieve a desired state?
- Phase-field models with obstacle/nonsmooth potentials
  ...control the fluid to force a desired separation of phases?
- Volatility calibration in American options (Black-Scholes model)
  ...determine the true volatility based on market measurements?
- Parameter identification in image processing:
  ...obtain a robust (wrt stochasticity) or "distributed" regularization parameter?

General Modeling Framework

Consider VIs of the type:

$$\text{Find } y \in V : \varphi(y') \geq \varphi(y) + \langle u + f - Ay, y' - y \rangle, \ \forall y' \in V,$$

where (amongst other assumptions) $\varphi : V \to \mathbb{R}$ is convex.

V reflexive Banach space, $A : V \to V^*$ strongly monotone

$\Longrightarrow$

Solution mapping $V^* \ni u \mapsto y$ (denoted $S(u)$) is Lipschitz.

For parameter ID usually much less continuity (loc. Lipschitz, Hölder,...).
For today: We consider the Lipschitz case.

Implicit Programming vs. MPCC

General Modeling Framework: Implicit Programming

$$\min J(u, y) \text{ over } (u, y) \in H \times V,$$
$$\text{s.t. } y = S(Bu).$$

Other approaches:

- **"MPCC"** Replace $y = S(Bu)$ by introducing slack/KKT-multiplier consider MPCC (assuming complementarity conditions can be written!)
- **"Adapted Penalty"** Smooth and regularize the variational inequality, consider sequence of related control problems.

(Differential) Sensitivity of the Solution Map I

How smooth is S?

- In n-dimensions: S (loc.) Lipschitz $\Rightarrow$ S almost everywhere C^1 (Rademacher).
- In ∞-dimensions: S (loc.) Lipschitz $\Rightarrow$ S Gâteaux differentiable up to "small" sets (Aronszajn, Preiss, Zaijcek, et al.)

In general, we cannot rule out these "exceptional" set.

(Differential) Sensitivity of the Solution Map II

Case 1. $\varphi(y) := i_M(y)$ (Variational Inequalities of the First Kind)

- $M \neq \emptyset$ closed, convex subset of refl. Banach space V
- i_M is the usual indicator

Here, $S : V^* \to V$ is the solution mapping of

$$A(y) + N_M(y) \ni w$$

with $w \in V^*$. We let $B \in \mathcal{L}(H, V^*)$, e.g., an embedding. H refl. B. sp.

(Differential) Sensitivity of the Solution Map II

Theorem

If M is "polyhedric" in the sense of Mignot/Haraux and $A : V \to V^*$ is strongly monotone, Fréchet differentiable, and $A(0) = 0$, then

1. *The solution mapping S of the VI is Hadamard directionally differentiable.*
2. $d = S'(Bu, Bh)$ *is the unique solution of the VI:*

$$\text{Find } d \in \mathcal{K} : \langle A'(y)d - Bh, z - d\rangle \geq 0, \forall z \in \mathcal{K}.$$

$\mathcal{K} := T_M(y) \cap \{w - A(y)\}^{\perp}$ *("critical cone")*

Proof.

1. Use Mignot/Haraux (1976/1977), Levy & Rockafellar (1994). Allows one to "differentiate" the subdifferential $\partial\varphi$.
2. S Lipschitz $\Rightarrow$ generalized derivative $\equiv$ Hadamard directional derivative. □

$A(0) = 0 \Rightarrow A'(y)$ coercive (elliptic). E.g., Linear op., p-Laplacian ($p > 2$).

(Differential) Sensitivity of the Solution Map III

Case 2. $\varphi(y) := \int_\Omega |(Gy)(x)|_{n,m} dx$ (Variational Inequalities of the Second Kind)

- $\Omega \subset \mathbb{R}^n$ open and bounded, $n \in \mathbb{N}$
- $G : V \to L^2(\Omega)^{n,m}$ bounded and linear.
- $|\cdot|_{n,m}$: abs. val. $(n = m = 1)$, Euclid. $(n > 1,\ m = 1)$, Frob. $(n, m > 1)$

Here, $S : V^* \to V$ is the solution mapping of

$$A(y) + G^*\partial\|\cdot\|_{L^1}(Gy) \ni w$$

with $w \in V^*$. We let $B \in \mathcal{L}(H, V^*)$, e.g., an embedding. H refl. B. sp.

(Differential) Sensitivity of the Solution Map III

Examples

- Mechanics: 2D-(very!)-Simplified Friction

$$\varphi(\cdot) := ||\cdot||_{\mathbf{L^1(\Omega)}},\ B := E_{L^2\hookrightarrow H^{-1}},\ A = -\Delta,\ G = \beta Id.$$

- Petroleum Engineering: Steady-State Laminar Flow of Bingham Fluid

$$\varphi(\cdot) := ||\boldsymbol{\nabla}\cdot||_{\mathbb{L}^{\mathbf{1}}\mathbf{(\Omega)}},\ B := E_{L^2\hookrightarrow H^{-1}},\ A = -\Delta,\ G = \nabla.$$

- Digital Image Processing: Approximation of TV-Regularized Problem

$$\varphi(\cdot) := \beta||\boldsymbol{\nabla}\cdot||_{\mathbb{L}^{\mathbf{1}}\mathbf{(\Omega)}},\ B := K^*,\ A = -\alpha\Delta + K^*K,\ G = \nabla.$$

(Differential) Sensitivity of the Solution Map III

Theorem

If $n = m = 1$ and $A : V \to V^*$ is strongly monotone, Fréchet differentiable, and $A(0) = 0$, then

1. *The solution mapping S of the VI is Hadamard directionally differentiable.*
2. *$d = S'(Bu, Bh)$ is the unique solution of the VI:*

$$\text{Find } d \in \mathcal{K} : \langle A'(y)d - Bh, z - d\rangle \geq 0, \forall z \in \mathcal{K}.$$

$\mathcal{K}$ is a type of "generalized critical cone."

(Differential) Sensitivity of the Solution Map III

Generalized Critical Cone

Given u, $y = S(Bu)$, $q \in \partial||\cdot||_{L^1}(Gy)$. Define the **biactive** and **strongly active** sets by

$$\mathcal{A}^0 := \{x \in \Omega \,||(Gy)(x)| = 0,\ |q(x)| = 1\},$$
$$\mathcal{A}^+ := \{x \in \Omega \,||(Gy)(x)| = 0,\ |q(x)| < 1\}.$$

Then

$$\mathcal{K} := \left\{ w \in V \;\middle|\; \begin{array}{lcll} (Gw)(x) & = & 0, & \text{a.e. } x \in \mathcal{A}^+, \\ (Gw)(x) & \in & \mathrm{cone}(q(x)), & \text{a.e. } x \in \mathcal{A}^0. \end{array} \right\}$$

Here, $q(x) \in [-1, 1]$ we can split $\mathcal{A}^0$ into two further subsets:

$$\mathcal{A}^{0,1} := \left\{ x \in \mathcal{A}^0 \,|q(x) = 1 \right\}, \quad \mathcal{A}^{0,-1} := \left\{ x \in \mathcal{A}^0 \,|q(x) = -1 \right\}.$$

The cone constraints become:

$$(Gw)(x) \geq 0, \text{ a.e. } x \in \mathcal{A}^{0,1}, \quad (Gw)(x) \leq 0, \text{ a.e. } x \in \mathcal{A}^{0,-1}.$$

(Differential) Sensitivity of the Solution Map III

But what about $n > 1$?

(Differential) Sensitivity of the Solution Map III

But what about $n > 1$?

∞-dimensions:

Formulae for generalized derivatives available. Difficult to use in numerics.

N-dimensions

After discretization, much more possible if G and $V_h := \mathrm{span}\{\psi_1, \ldots, \psi_N\}$ "second-order compatible."

$d = S_h'(u; w)$ given as the (unique) solution of the following variational inequality of the first kind:

$$\text{Find } d \in \mathcal{K}_h : 0 \geq \langle B_h w - A_h'(y)d - \mathcal{Q}_h(y)d, d' - d \rangle, \; \forall d \in \mathcal{K}_h,$$

where $\mathcal{Q}_h(y)$ is the gradient associated with a positive semidefinite quadratic form.

Model MPEC

Assumptions

$$\min J(u,y) \text{ over } (u,y) \in H \times V,$$
$$\text{s.t. } y = S(Bu).$$

- V and H are Hilbert spaces
- $V \hookrightarrow H \equiv H^* \hookrightarrow V^*$ represents a Gelfand triple
- $J : H \times V \to \mathbb{R}$ is continuously Fréchet, bounded from below
- S is (Lipschitz, Hadamard dir. diff.) solution operator $S : V^* \to V$ for VI
- $B \in \mathcal{L}(H)$ with B compact from H to V^*
- $J(\cdot, S(B\cdot)) : H \to \mathbb{R}$ is coercive and weakly lower semi-continuous

B-Stationarity

Theorem

If $(u, y) \in H \times V$ is a (locally) optimal solution of the MPEC, then

$$\langle \nabla_y J(u, y), d \rangle_{V^*,V} + \langle \nabla_u J(u, y), w \rangle_{H^*,H} \geq 0, \ \forall (w, d) \in \mathrm{Gph}\, S'(Bu; B\cdot)$$

How can we use B-stationarity for a numerical method?

Towards a Conceptual Algorithm

Form Regularized Auxiliary Problem (RAP)

Let $y = S(Bu)$, define RAP:

$$\min F(h) := \frac{1}{2} b(h,h) + J_y(u,y)S'(Bu;Bh) + J_u(u,y)h \text{ over } h \in H. \quad \text{(RAP)}$$

$b(h,h) := (Qh,h)_H$ coercive (elliptic) and bounded quadratic form ($h \in H$).

RAP characterizes Solutions/B-stationarity

If (u,y) solves the MPEC, then $0 \in H$ solves the RAP

Descent Directions

If (u,y) not a solution, then solution h of RAP is a proper descent direction of reduced objective $\mathcal{J}(u) := J(u, S(Bu))$.

A Conceptual Algorithm

Algorithm 1 Conceptual Algorithm

Input: $u_0 \in H$; $\epsilon \geq 0$; $k := 0$

Set $y_0 = S(Bu_0)$.
Solve (RAP) with $(u, y) = (u_0, y_0)$ to obtain h_0.
while $||h_k||_H > \epsilon$ **do**
    Compute $u_{k+1} := u_k + t_k h_k$, $t_k > 0$, via a line search.
    Set $y_{k+1} = S(Bu_{k+1})$.
    Solve (RAP) with $(u, y) = (u_{k+1}, y_{k+1})$ to obtain h_{k+1}.
    Set $k := k + 1$.
end while

In general, this is an intractable method: (RAP) is an MPEC! But...

Obtaining Descent Directions

Exploiting the Sensitivity Analysis

Formulae for $S'(Bu; Bh) \Rightarrow S$ is Gâteaux differentiable if $\text{meas}(\mathcal{A}^0) = 0$.

Smooth case: $m(\mathcal{A}^0) = 0$ (no biactivity)

1. Explicit formula for $S'(Bu; Bh)$ allows us to calculate a descent direction of $\mathcal{J}$ (adjoint state exists!)
2. Obtain the gradient $\nabla_u \mathcal{J}(u)$ by solving adjoint equation.

Nonsmooth case: $m(\mathcal{A}^0) > 0$ (biactivity present)

1. Approximate the VI associated with $S'(Bu; Bh)$.
2. $\exists \gamma > 0$ (finite penalty parameter):

$$h_\gamma := Q^{-1}(B^* p_\gamma - \nabla_u J(u, y)),$$

   is a proper descent direction for $\mathcal{J}$.
3. p_γ solves linearization of the approximation of $S'(Bu; 0)$.

Applying the Idea

Optimal Control of a VI of Second Kind

$$\begin{aligned} &\min J(u,y) := \tfrac{1}{2}||y-y_d||^2_{L^2} + \tfrac{\alpha}{2}||u||^2_{L^2} \text{ over } (u,y) \in L^2(\Omega)\times H^1_0(\Omega),\\ &\text{s.t.}\quad y = \operatorname{argmin}\left\{\tfrac{1}{2}\int_\Omega |\nabla z|^2 dx - \int_\Omega (u+f)z dx + \int_\Omega |Gz| dx\right\}. \end{aligned} \quad (1)$$

Here, $\Omega \subset \mathbb{R}^n$, $n \in \{1,2,3\}$, is open and bounded; $\alpha > 0$; $f, y_d \in L^2(\Omega)$; and $G \in \mathcal{L}(H^1_0(\Omega), L^2(\Omega))$. B is the canonical embedding.

Same arguments for control of the obstacle problem
(need a few assumptions about the active sets).

The Directional Derivative of the Solution Map

For each $u \in L^2(\Omega)$ & $y = S(u)$ $S'(u;h) = d$; the unique solution of QP:

$$\begin{aligned} &\min \tfrac{1}{2}\int_\Omega |(\nabla w)(x)|^2 dx - \int_\Omega h(x)w(x)dx \text{ over } w \in H^1_0(\Omega)\\ &\text{s.t. } (Gw)(x) = 0, \text{ a.e. } x \in \mathcal{A}^+, \quad (Gw)(x) \geq 0, \text{ a.e. } x \in \mathcal{A}^{0,1}\\ &\quad\;\; (Gw)(x) \leq 0, \text{ a.e. } x \in \mathcal{A}^{0,-1} \end{aligned}$$

Obtaining Descent Directions

Smooth case: $m(\mathcal{A}^0) = 0$ (no biactivity)

1. $h = Q^{-1}(p - \alpha u)$ is a proper descent direction.
2. p solves the adjoint variational equation: $(Gp)(x) = 0,$ a.e. $x \in \mathcal{A}$ and

$$\int_\Omega \nabla p \cdot \nabla \psi dx = \int_\Omega (y_d - y)\psi dx, \ \forall \psi \in H_0^1(\Omega) : (G\psi)(x) = 0, \text{ a.e. } x \in \mathcal{A}.$$

Nonsmooth case: $m(\mathcal{A}^0) > 0$ (biactivity present)

1. Approximate the VI associated with $S'(u; h)$.
2. $\exists \gamma > 0$ (finite penalty parameter):

$$h_\gamma := Q^{-1}(p_\gamma - \alpha u),$$

   is a proper descent direction for $\mathcal{J}$.
3. p_γ solves linearization of the approximation of $S'(u; 0)$.

Obtaining Descent Directions in Nonsmooth Case

1. For some penalty map, e.g., $\beta(r) := \max(0, r)$, approximate $S'(u; h)$ by $d_\gamma(h)$, the solution of

$$-\Delta d + \gamma G^* \left[\chi_{\mathcal{A}^+} Gd + \chi_{\mathcal{A}^{0,1}} \beta(Gd) - \beta(-Gd)\right] = h.$$

Obtaining Descent Directions in Nonsmooth Case

1. For some penalty map, e.g., $\beta(r) := \max(0, r)$, approximate $S'(u; h)$ by $d_\gamma(h)$, the solution of

$$-\Delta d + \gamma G^* \left[\chi_{\mathcal{A}^+} Gd + \chi_{\mathcal{A}^{0,1}} \beta(Gd) - \beta(-Gd)\right] = h.$$

2. Consider smoothed RAP (assume (u, y) not B-stationary):

$$\min F_\gamma(h) := \frac{1}{2} b(h, h) + \alpha(u, h)_{L^2} + (y - y_d, d_\gamma(h))_{L^2}, \text{ over } h \in L^2(\Omega). \quad (2)$$

$h_\gamma := -\nabla_h F_\gamma(0) \neq 0$ is a proper descent direction of F_γ at zero. Here:

$$h_\gamma = Q^{-1}(p_\gamma - \alpha u).$$

where

$$-\Delta p_\gamma + \gamma G^* \left[\chi_{\mathcal{A}^+} G p_\gamma\right] = y_d - y.$$

Obtaining Descent Directions in Nonsmooth Case

1. For some penalty map, e.g., $\beta(r) := \max(0, r)$, approximate $S'(u; h)$ by $d_\gamma(h)$, the solution of
$$-\Delta d + \gamma G^* \left[\chi_{\mathcal{A}^+} Gd + \chi_{\mathcal{A}^{0,1}} \beta(Gd) - \beta(-Gd)\right] = h.$$
2. Consider smoothed RAP (assume (u, y) not B-stationary):
$$\min F_\gamma(h) := \frac{1}{2} b(h, h) + \alpha(u, h)_{L^2} + (y - y_d, d_\gamma(h))_{L^2}, \text{ over } h \in L^2(\Omega). \quad (2)$$
$h_\gamma := -\nabla_h F_\gamma(0) \neq 0$ is a proper descent direction of F_γ at zero. Here:
$$h_\gamma = Q^{-1}(p_\gamma - \alpha u).$$
where
$$-\Delta p_\gamma + \gamma G^* \left[\chi_{\mathcal{A}^+} Gp_\gamma\right] = y_d - y.$$
3. Moreover: $d_\gamma(\cdot) \overset{H_0^1}{\to} S'(u; \cdot)$ as $\gamma \to +\infty$.

Obtaining Descent Directions in Nonsmooth Case

1. For some penalty map, e.g., $\beta(r) := \max(0, r)$, approximate $S'(u; h)$ by $d_\gamma(h)$, the solution of

$$-\Delta d + \gamma G^* \left[\chi_{\mathcal{A}^+} Gd + \chi_{\mathcal{A}^{0,1}} \beta(Gd) - \beta(-Gd)\right] = h.$$

2. Consider smoothed RAP (assume (u, y) not B-stationary):

$$\min F_\gamma(h) := \frac{1}{2} b(h, h) + \alpha(u, h)_{L^2} + (y - y_d, d_\gamma(h))_{L^2}, \text{ over } h \in L^2(\Omega). \tag{2}$$

$h_\gamma := -\nabla_h F_\gamma(0) \neq 0$ is a proper descent direction of F_γ at zero. Here:

$$h_\gamma = Q^{-1}(p_\gamma - \alpha u).$$

where
$$-\Delta p_\gamma + \gamma G^* \left[\chi_{\mathcal{A}^+} Gp_\gamma\right] = y_d - y.$$

3. Moreover: $d_\gamma(\cdot) \overset{H_0^1}{\to} S'(u; \cdot)$ as $\gamma \to +\infty$.
4. Once γ fulfills
$$||d_\gamma(h_\gamma) - S'(u; h_\gamma)||_{L^2} < \frac{1}{||y - y_d||_{L^2}} \cdot \frac{c_1}{4} ||h_\gamma||^2_{L^2}, \tag{3}$$

then h_γ is a descent direction of $\mathcal{J}(u)$!

Algorithm I: A Descent Method for MPECs

Input: $u_0 \in L^2(\Omega)$; $\gamma_0 > 0$; $\varepsilon \geq 0$; $k := 0$; $\rho_1 > 1$, $\rho_2 > 1$;
Set $y_0 := S(u_0)$, $y_0^* := G^* q_0$ with $q_0 \in \partial || \cdot ||_{L^1}(Gy_0)$.
while stopping criterion not fulfilled **do**
  if no biactivity **then**
    Set $h_k = Q^{-1}(p_k - \alpha u_k)$, p_k solves adjoint eq.
  else
    Set $h_k = Q^{-1}(p_k - \alpha u_k)$, where p_k solves approx. adj. eq.
    while (3) fails **do**
      Choose $\tilde{\gamma}_k > \rho_1 \gamma_k$.
      Set $h_k = Q^{-1}(p_k - \alpha u_k)$, where p_k solves approx. adj. eq.
      Set $\gamma_k = \tilde{\gamma}_k$.
    end while
    Compute $u_{k+1} = u_k + t_k h_k$, $t_k > 0$, via a line search.
    Set $y_{k+1} := S(u_{k+1})$, $y_{k+1}^* := G^* q_{k+1}$ with $q_{k+1} \in \partial || \cdot ||_{L^1}(Gy_{k+1})$.
    Choose $\gamma_{k+1} > \rho_2 \gamma_k$.
  end if
  Set $k := k + 1$.

end while

Theoretical convergence proofs imply C-stationarity

$\Rightarrow$

Stop when C-stationarity holds (up to a small tolerance).

Details of Implementation I

Obtaining a feasible pair (u, y)

We solve the VI by rewriting as nonsmooth equation:

$$-\Delta y + G^* q = u + f,$$
$$Gy = \max(0, q + Gy - 1) + \max(0, -(1 + q + Gy)).$$

Use semismooth Newton (locally superlinearly convergent on each mesh).

Obtaining $d_\gamma(h)$

Use smoothed max-function $\beta_\varepsilon(r)$ and solve

$$-\Delta d + \gamma G^* \left[\chi_{\mathcal{A}^+} Gd + \chi_{\mathcal{A}^{0,1}} \beta_\varepsilon((Gd) - \beta_\varepsilon((-Gd)\right] = h$$

with standard Newton method.

The line search

Simple backtracking, Armijo-type...But what about Q?!

Details of Implementation II

1. $\Omega = [0,1] \times [0,1]$
2. $-\Delta$ discretized via finite differences, standard 5-point stencil
3. Overall method implemented within a nested-grid strategy: Solve on coarse grid, prolongate (9-point-star), solve on next finer grid.
4. Discrete L^2-norms used for residuals (OK considering regularity theory for the PDEs and VIs).

Examples

Example (Large Biactive Set, No Strongly Active Set, Discontinuous q)

Define

$$y^{\dagger}(\mathbf{x}_1,\mathbf{x}_2) = \beta_{\varepsilon}((-\Delta)^{-1}(\mu\sin((\mathbf{x}_1-0.5)(\mathbf{x}_2-0.5)))),$$
$$q^{\dagger}(\mathbf{x}_1,\mathbf{x}_2) = \chi_{\{y^{\dagger}>0\}}(\mathbf{x}_1,\mathbf{x}_2) - \chi_{\{y^{\dagger}\leq 0\}}(\mathbf{x}_1,\mathbf{x}_2) + \chi_{[0.5,1]\times[0,0.5]}(\mathbf{x}_1,\mathbf{x}_2),$$

where $\mu = 1\text{E}3$, $\varepsilon = 1\text{E-}2$,

$$\{y^{\dagger} > 0\} := \left\{(\mathbf{x}_1,\mathbf{x}_2)\in\Omega \,\middle|\, y^{\dagger}(\mathbf{x}_1,\mathbf{x}_2) > 0\right\},$$
$$\{y^{\dagger} \leq 0\} := \left\{(\mathbf{x}_1,\mathbf{x}_2)\in\Omega \,\middle|\, y^{\dagger}(\mathbf{x}_1,\mathbf{x}_2) \leq 0\right\}.$$

Moreover, we set

$$f = -\Delta y^{\dagger} - y^{\dagger} + q^{\dagger}, \quad y_d = y^{\dagger} - q^{\dagger} - \alpha\Delta y^{\dagger}.$$

In addition, $\alpha = 1$, $u_0 = 0$.

Examples

Example (Large Biactive Set, Large Strongly Active Set, Discontinuous q)

Define

$$y^{\dagger}(\mathrm{x}_1, \mathrm{x}_2) = \beta_{\varepsilon}((-\Delta)^{-1}(10\sin(5\mathrm{x}_1)cos(4\mathrm{x}_2))),$$
$$q^{\dagger}(\mathrm{x}_1, \mathrm{x}_2) = \chi_{\{y^{\dagger}>0\}}(\mathrm{x}_1, \mathrm{x}_2) - \chi_{\{y^{\dagger}<0\}}(\mathrm{x}_1, \mathrm{x}_2) + \chi_{\{y^{\dagger}==0\}}(\mathrm{x}_1, \mathrm{x}_2),$$

where $\varepsilon = 1\text{E-}2$, and $\{y^{\dagger} > 0\}, \{y^{\dagger} < 0\}$, and $\{y^{\dagger} = 0\}$ are defined as in Example 4. We again set

$$f = -\Delta y^{\dagger} - y^{\dagger} + q^{\dagger}, \quad y_d = y^{\dagger} - q^{\dagger} - \alpha\Delta y^{\dagger}.$$

and $\alpha = 1$, $u_0 = 0$.

Results $Q = Id$

Example 1							
DoF	k	Final $\|\|h_k\|\|_{L^2}$	Lin. Solves	ns	s	$\tau_{\min}$	ALSM
49	42	9.7193e-05	283	43	0	0.03125	6.7381
225	2	6.6522e-05	41	3	0	1	20.5
961	1	5.3036e-06	41	2	0	1	41^*
3969	1	2.9248e-05	31	2	0	1	31^*
16129	160	9.9635e-05	1063	161	0	0.015625	6.6438
65025	1	3.0284e-08	58	2	0	1	58^*
261121	1	1.9197e-06	99	2	0	1	99^*
Example 2							
DoF	k	Final $\|\|h_k\|\|_{L^2}$	Lin. Solves	ns	s	$\tau_{\min}$	ALSM
961	528	9.9895e-05	2703	529	0	0.0019531	5.1193
3969	69	9.95832e-05	431	70	0	0.0039062	6.2463
16129	4	9.3982e-05	68	5	0	0.0625	17
65025	223	9.9525e-05	1254	224	0	0.0019531	5.6233
261121	378	9.9820e-05	2037	379	0	0.0019531	5.3889

Adaptively Choosing $Q = \aleph_k Id$

A Two-Point/Barzilai-Borwein-Type Approach

Input: $(u_0, y_0, q_0), (u_1, y_1, q_1, h_1)$; $\gamma_1 > 0$; $\varepsilon \geq 0$; $k := 1$; $\rho_1 > 1$, $\rho_2 > 1$; $\aleph_0 = 1$.

while stopping criterion not fulfilled **do**
- **if** no biactivity **then**
  - Set $h_{k+1} = p_k - \alpha u_k$ p_k solves adj. eq.
  - Set $u_{k+1} = u_k + \aleph_k^{-1} h_{k+1}$.
- **else**
  - Set $h_{k+1} = p_k - \alpha u_k$, where p_k solves approx. adj. eq.
  - **while** (3) fails **do**
    - Choose $\tilde{\gamma}_k > \rho_1 > 1\gamma_k$.
    - Set $h_{k+1} = p_k - \alpha u_k$, where p_k solves approx. adj. eq.
    - Set $\gamma_k = \tilde{\gamma}_k$.
  - **end while**
  - Set $u_{k+1} = u_k + \aleph_k^{-1} h_{k+1}$
  - Set $y_{k+1} := S(u_{k+1})$, $y^*_{k+1} := G^* q_{k+1}$ with $q_{k+1} \in \partial || \cdot ||_{L^1}(Gy_{k+1})$.
  - Choose $\gamma_{k+1} > \rho_2 \gamma_k$.
- **end if**
- Set $\aleph_{k+1} = -\dfrac{(u_{k+1} - u_k, h_{k+1} - h_k)_{L^2}}{||u_{k+1} - u_k||^2_{L^2}}$
- Set $k := k + 1$.

end while

No theory yet, need to ensure $\{\aleph_k\}_k$ is bounded.

Results $Q = \aleph_k Id$

Example 1							
DoF	k	Final $\|\|h_k\|\|_{L^2}$	Lin. Solves	ns	s	$\aleph_{\min}$	$\aleph_{\max}$
49	3	8.464e-06	13	4	0	0.99993	1.0266
225	2	5.9638e-05	16	3	0	1	1.033
961	1	5.3777e-06	19	2	0	1	1.0001
3969	1	1.4151e-05	14	2	0	0.99999	1
16129	1	4.1134e-07	31	2	0	1	1.0002
65025	1	2.9916e-08	42	2	0	1	1.0001
261121	1	2.6744e-06	40	2	0	0.9874	1
Example 2							
DoF	k	Final $\|\|h_k\|\|_{L^2}$	Lin. Solves	ns	s	$\aleph_{\min}$	$\aleph_{\max}$
49	2	8.464e-06	12	3	0	1	1.0001
225	3	5.9638e-05	12	4	0	1	1.358
961	1	5.3777e-06	9	2	0	1	1
3969	16	1.4151e-05	51	17	0	0.96857	18.3984
16129	4	4.1134e-07	27	5	0	1	3.2726
65025	2	2.9916e-08	19	3	0	1	2.1232
261121	2	2.6744e-06	25	3	0	1	1.3604

Results $Q = \aleph_k Id$

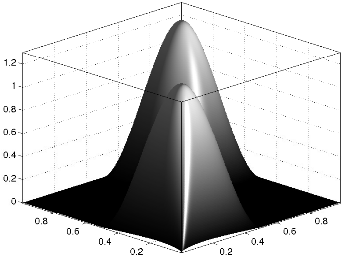

Figure: Optimal Controls u for Example 1 (l.) and 2 (r.)

Results $Q = \aleph_k Id$

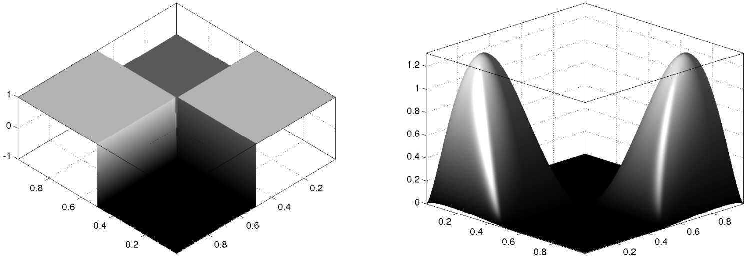

Figure: (l.) Subgradient q and (r.) State y for Example 1

Results $Q = \aleph_k Id$

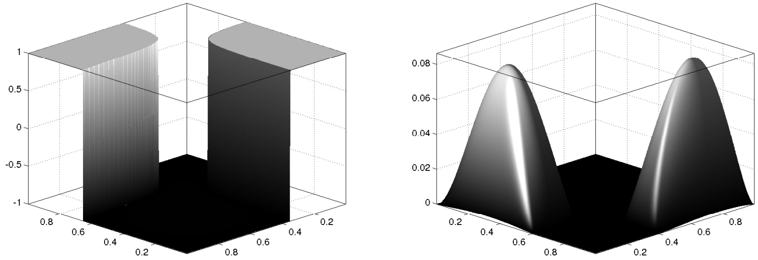

Figure: (l.) Subgradient q and (r.) State y for Example 2

Results $Q = \aleph_k Id$

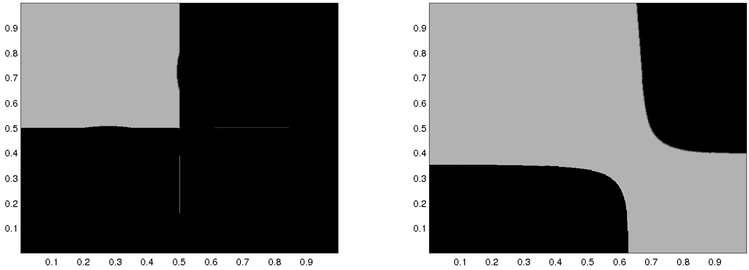

Figure: Biactive sets $\mathcal{A}^{0,-1}$ (lighter region) in Examples 1 (l.) and 2 (r.)