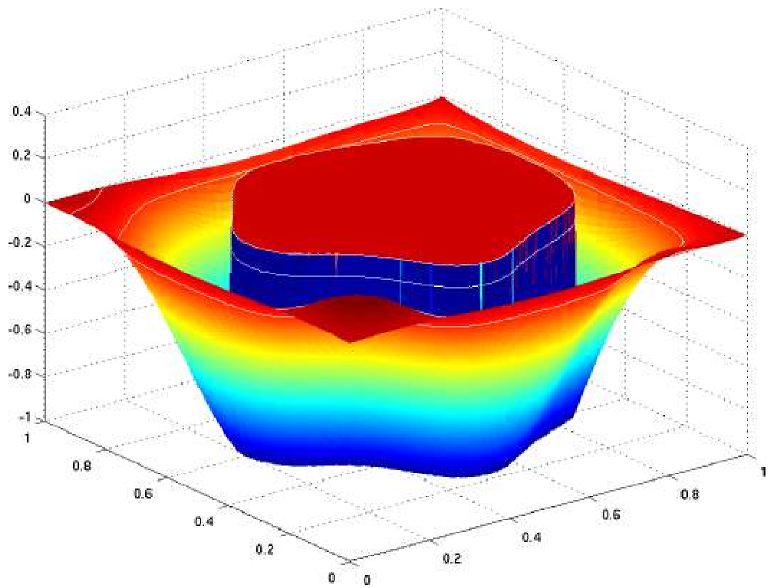
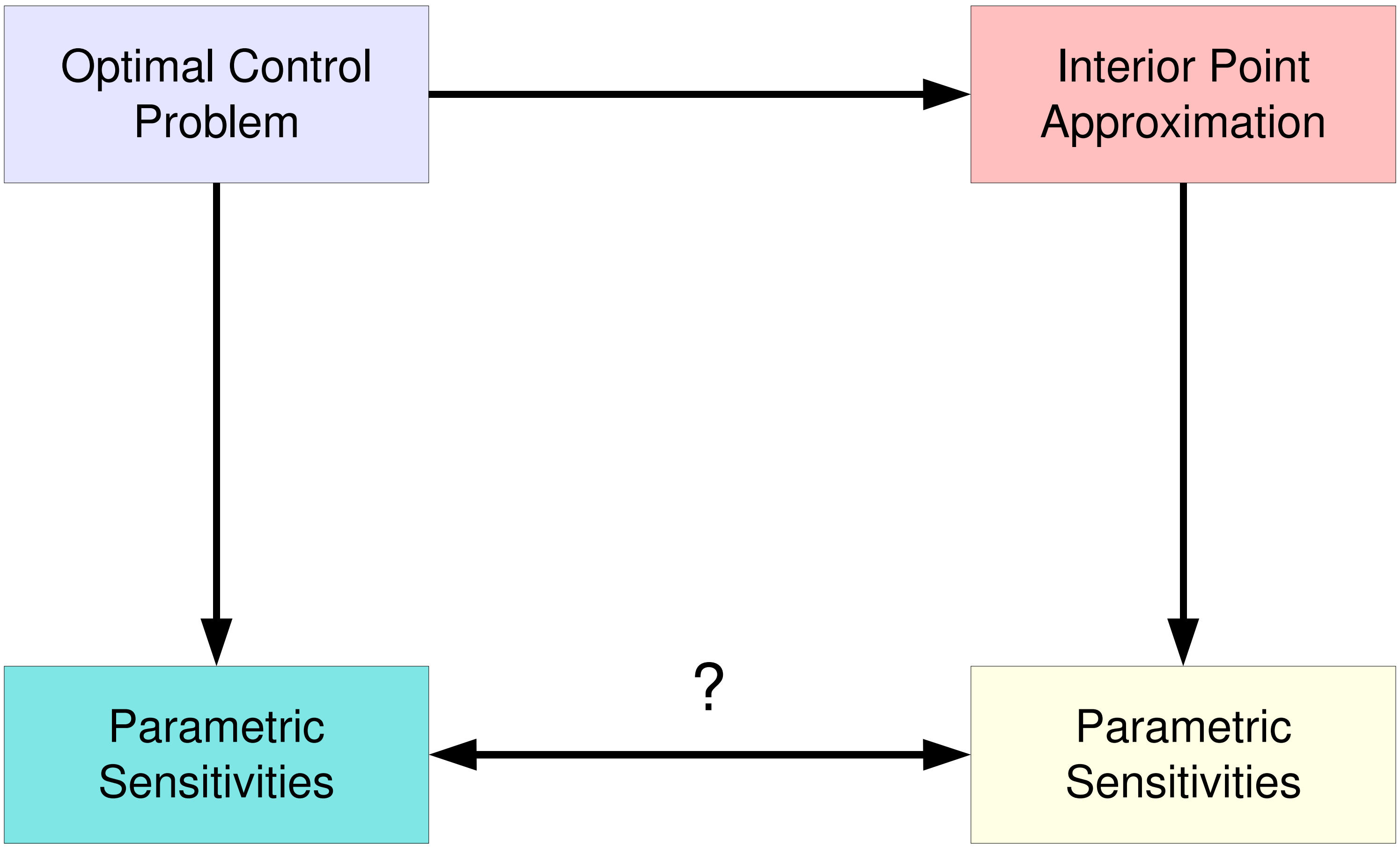


Martin Weiser
Roland Griesse

Interior Point Methods and Parametric Sensitivity in Optimal Control


Optimal Control Problem
Interior Point Approximation
Parametric Sensitivities
?
Parametric Sensitivities

$$\min_u J(u;p) \quad \text{subject to} \quad \begin{array}{l} Gu+g\geq 0 \\ u\in L_\infty(\Omega) \end{array}$$

Optimal Control Problem

$$J(u;p)=\langle u,(\alpha(p)+K(p))u\rangle+\langle u,l(p)\rangle$$

$\alpha(p)\geq\underline{\alpha}>0$

$K(p): L_2\to L_\infty$

continuously differentiable w.r.t. $p\in\mathbb{R}$

$(Gu)(x)=\overline{G}u(x)$ for almost all $x\in\Omega$ — linear Nemyckii operator

$\overline{G}:\mathbb{R}\to\mathbb{R}^m$

Solution $u(p)$

Multiplier $\eta(p)$

Optimal Control Problem

Distributed control problems

$$\min_{u,y} \|y - y_d(p)\|_{L_2}^2 + \alpha(p)\|u\|_{L_2}^2$$

s.t. $u \geq 0, \quad -\Delta y = u, \quad y = 0$ on $\partial\Omega$

Regularized obstacle problems

$$\min_{u} \|\nabla u\|_{L_2}^2 + \langle u, l \rangle$$

s.t. $u \geq 0, \quad u = u_d$ on $\partial\Omega$

Parametric Sensitivities

Theorem

The mapping $p \to (u(p), \eta(p))$ is Lipschitz continuous and directionally differentiable at p_0.

The directional derivative $(u_p(p), \eta_p(p))$ is the unique solution of the auxiliary problem

$$\min_u J_{uu}(p_0)[u_p, u_p] + J_{up}(u(p_0); p_0) u_p$$

s.t. $u_p = 0$ a.e. on Ω_a

$u_p \geq 0$ a.e. on Ω_0

$\Omega_a = \{x \in \Omega : \eta > 0\}$	active set
$\Omega_i = \{x \in \Omega : Gu + g > 0\}$	inactive set
$\Omega_0 = \{x \in \Omega : Gu + g + \eta = 0\}$	weakly active set

Parametric Sensitivities

- Stability of optimal solutions
- Fast prediction of optimal solutions for varied parameters

$$u(p+\delta p)\approx u(p)+u_p(p)\delta p$$

caveat:

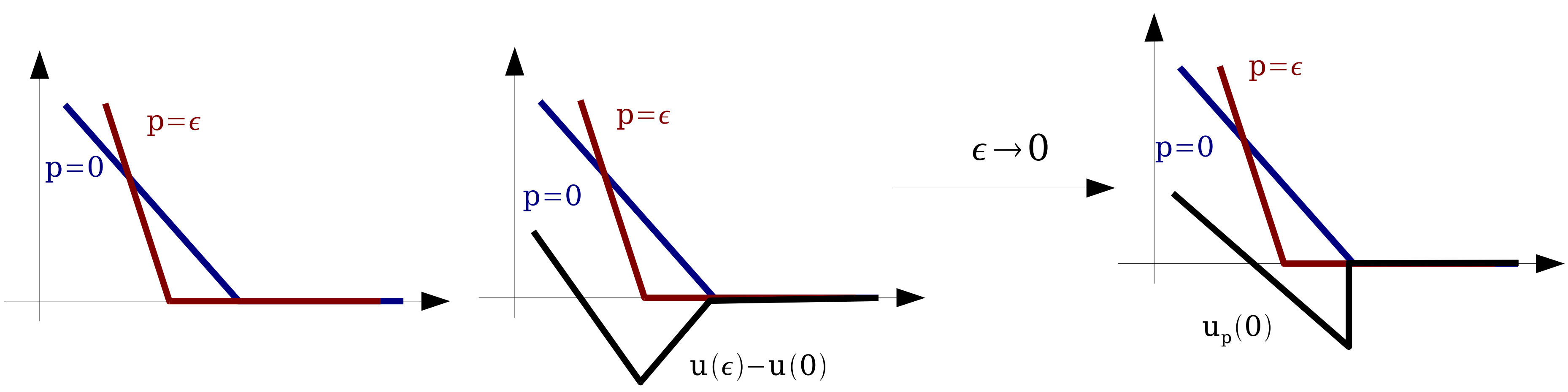
- active set usually does not change on update!

$$\min_u J_{uu}(p_0)[u_p,u_p]+J_{up}(u(p_0);p_0)u_p$$

s.t. $u_p=0$ a.e. on Ω_a

$u_p\geq 0$ a.e. on Ω_0

- update will in general violate constraints!



Primal barrier formulation

Interior Point Approximation

$$\min_u J(u;p) \quad \text{s.t.} \quad Gu+g \geq 0 \quad \longrightarrow \quad \min_u \; J(u;p) - \mu \int_\Omega \ln(Gu+g)\, dx$$

$$J_u(u;p) - G^* \frac{\mu}{Gu+g} = 0$$

Primal-dual formulation $\eta = \frac{\mu}{Gu+g}$

$$v = (u,\eta) \qquad F(v) = \begin{bmatrix} J_u(u;p) - G^* \eta \\ \eta * (Gu+g) - \mu \end{bmatrix} = 0$$

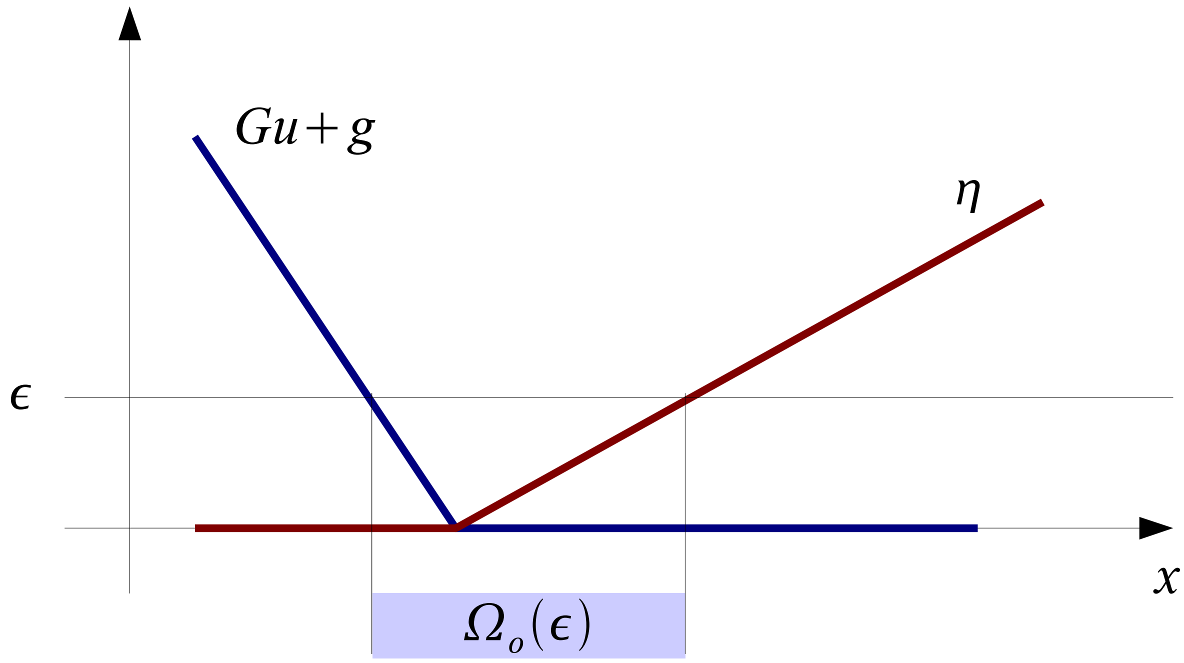
Homotopy $\mu \to 0$: central path solutions $u(p,\mu)$

parametric sensitivities $u_p(p,\mu)$

Definition

For a feasible point $v=(u,\eta)$, the indecisive set is $\Omega_o(\epsilon)=\{x\in\Omega:\eta+(Gu+g)\leq\epsilon\}$

v is strictly complementary of order r, if $|\Omega_o|=O(\epsilon^r)$



Interior Point Approximation

Theorem

Assume that $v(p)$ is strictly complementary of order $r \leq 1$

Then, $\|v(p,\mu)-v(p)\|_{L_q} \leq c\,\mu^{\frac{r+q}{2q}}$ holds for $2 \leq q \leq \infty$

Proof

(i) central path derivative $F_v(v;p,\mu)\,v_\mu(p,\mu)+F_\mu(v;p,\mu)=0$

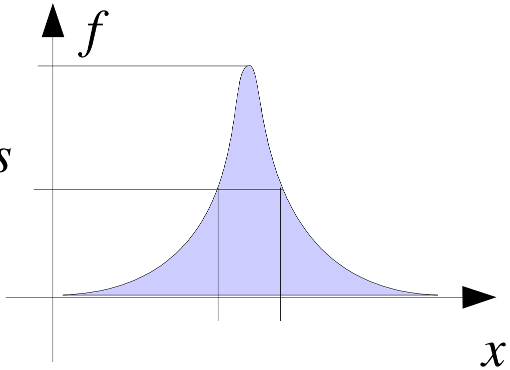
(ii) show that $\|F_v(v;p,\mu)^{-1}[a,b]\|_{L_q} \leq c\left(\|a\|_{L_q}+\|b/(Gu+g+\eta)\|_{L_q}\right)$

(iii) show that $\|v_\mu(p,\mu)\|_{L_q} \leq c\,\|(Gu+g+\eta)^{-1}\|_{L_q} \leq c\,\mu^{\frac{r-q}{2q}}$

use $|\{x\in\Omega : |f(x)|>s\}| \leq \psi(s) \Rightarrow \|f\|_{L_q}^q \leq q\int_0^\infty s^{q-1}\,\psi(s)\,ds$

$|\{x\in\Omega : (Gu+g+\eta)^{-1}>s\}| \leq c\,s^{-r}$

$Gu+g+\eta \geq 2\sqrt{\mu}$



(iv) integrate $\|v(p,\mu)-v(p)\|_{L_q} \leq \int_0^\mu \|v_\mu(p,\tau)\|_{L_q}\,d\tau$

(ii) show that $\|F_v(v;p,\mu)^{-1}[a,b]\|_{L_q} \le c\left(\|a\|_{L_q} + \|b/(Gu+g+\eta)\|_{L_q}\right)$

Interior Point Approximation

(a) $q=2$
symmetrize

$$\begin{bmatrix} J_{uu} & -G^* \\ -G & -(Gu+g)/\eta \end{bmatrix}\begin{bmatrix}\bar u \\ \bar\eta\end{bmatrix} = \begin{bmatrix} a \\ -b/\eta \end{bmatrix}$$

(b) almost active/inactive sets $\chi_A = 1,\ \eta > Gu+g,\ \ 0$ otherwise

$$\chi_I = 1-\chi_A$$

eliminate nearly inactive

$$\chi_I\bar\eta = \chi_I \frac{\eta}{Gu+g}(b/\eta - G\bar u)$$

$$\begin{bmatrix} J_{uu} + G^*\chi_I\eta/(Gu+g)G & -G^* \\ -G & -\chi_A(Gu+g)/\eta \end{bmatrix}\begin{bmatrix}\bar u \\ \chi_A\bar\eta\end{bmatrix} = \begin{bmatrix} a \\ -\chi_A b/\eta \end{bmatrix}$$

saddle point lemma

$$\|\bar u\|_{L_2} + \|\chi_A\bar\eta\|_{L_2} \le c(\|a\|_{L_2} + \|\chi_A b/\eta\|_{L_2})$$

(c) $q>2$ $\quad J_{uu} = \alpha + K$

$$\begin{bmatrix} \alpha + G^*\chi_I\eta/(Gu+g)G & -G^* \\ -G & -\chi_A(Gu+g)/\eta \end{bmatrix}\begin{bmatrix}\bar u \\ \chi_A\bar\eta\end{bmatrix} = \begin{bmatrix} a - K\bar u \\ -\chi_A b/\eta \end{bmatrix}$$

apply saddle point lemma pointwisely

$$|\bar u| + |\chi_A\bar\eta| \le c(|a| + |\chi_A b/\eta| + \|K\|_{L_2\to L_\infty}\|\bar u\|_{L_2}) \quad \text{a.e.}$$

Theorem

Parametric Sensitivities

Assume that $v(p)$ is strictly complementary of order $r \le 1$

Then, $\|v_p(p,\mu) - v_p(p)\|_{L_q} \le c\,\mu^{\frac{r}{2q}}$ holds for $2 \le q \le \infty$

Proof

(i) differentiate $F_v v_p = -F_p$: $\quad F_{vv}[v_p, v_\mu] + \underbrace{F_{v\mu}}_{=0} v_p + F_v v_{p\mu} = -F_{pv} v_\mu - \underbrace{F_{p\mu}}_{=0}$

(ii) $F_v v_p = -F_p = -\begin{bmatrix} J_{up} \\ 0 \end{bmatrix} \Rightarrow v_p$ bounded

(iii) $F_v v_{p\mu} = \begin{bmatrix} a \\ b \end{bmatrix}$ with $\|a\|_{L_q} + \|b\|_{L_q} \le c\,\mu^{\frac{r-q}{2q}} \Rightarrow \|v_{p\mu}\|_{L_q} \le c\,\mu^{\frac{r-2q}{2q}}$

since $\|(Gu+g+\eta)^{-1}\|_{L_\infty} \le \mu^{-1/2}$

(iv) integrate $\|v_p(p,\mu) - v_p(p)\|_{L_q} \le \int_0^\mu \|v_{p\mu}(p,\tau)\|_{L_q}\, d\tau$

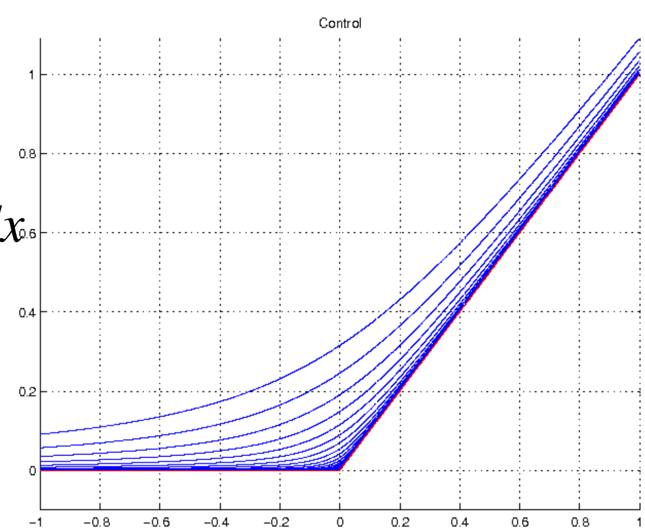
$r=1$	IP approximation $\|v(p,\mu)-v(p)\|_{L_q} \le \mu^{\frac{r+q}{2q}}$	IP sensitivities $\|v_p(p,\mu)-v_p(p)\|_{L_q} \le \mu^{\frac{r}{2q}}$
L_2	$\mu^{\frac{3}{4}}$	$\mu^{\frac{1}{4}}$
L_∞	$\mu^{\frac{1}{2}}$	no convergence

Parametric Sensitivities

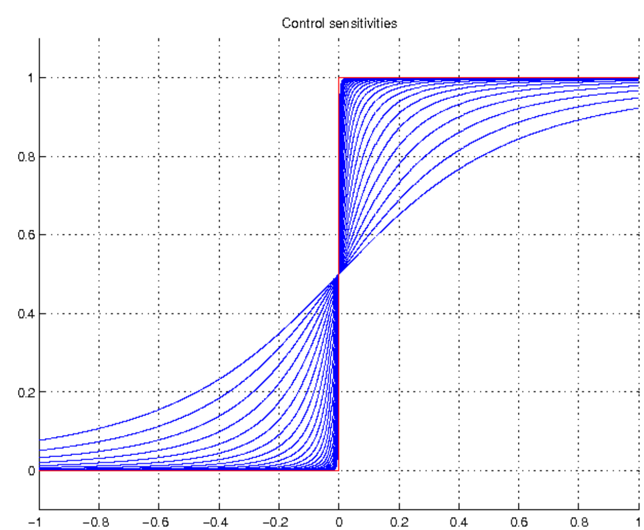
Example

$$\min_{u} \int_{-1}^{1} (u(x)-x-p)^2\,dx$$

s.t. $u \ge 0$



$u(x)$



$u_p(x)$

$$\min_u \; \|y - y_d + p\|_{L_2}^2 + \alpha \|u\|_{L_2}^2$$

s.t. $-40 \le u \le 40$

$-\Delta y = u$ on $(0,1)$

$y(0) = y(1) = 0$

$\alpha = 10^{-4}$

$y_d = \sin(3\pi x)$

$p = 0$

Framework:

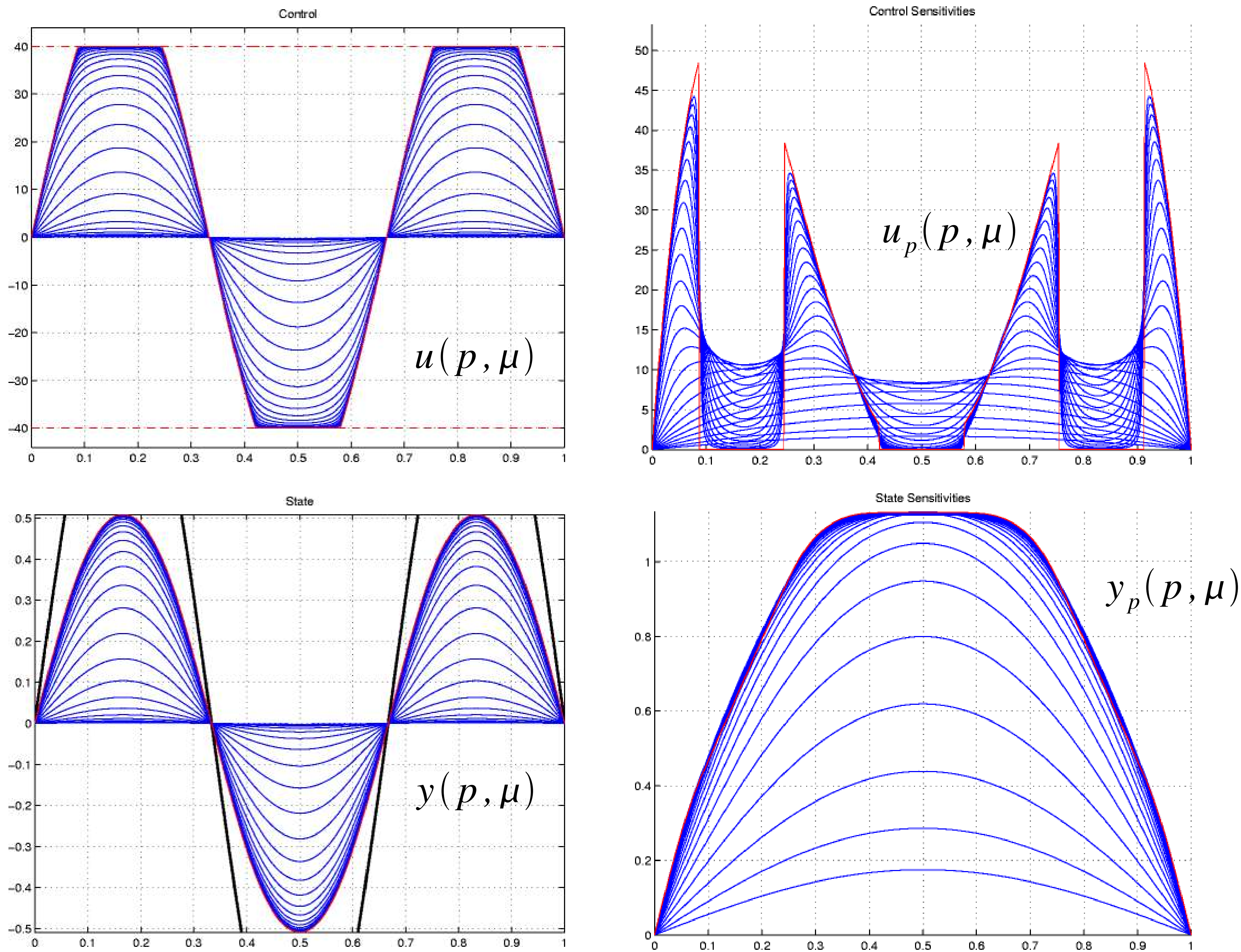
$$y = Su$$

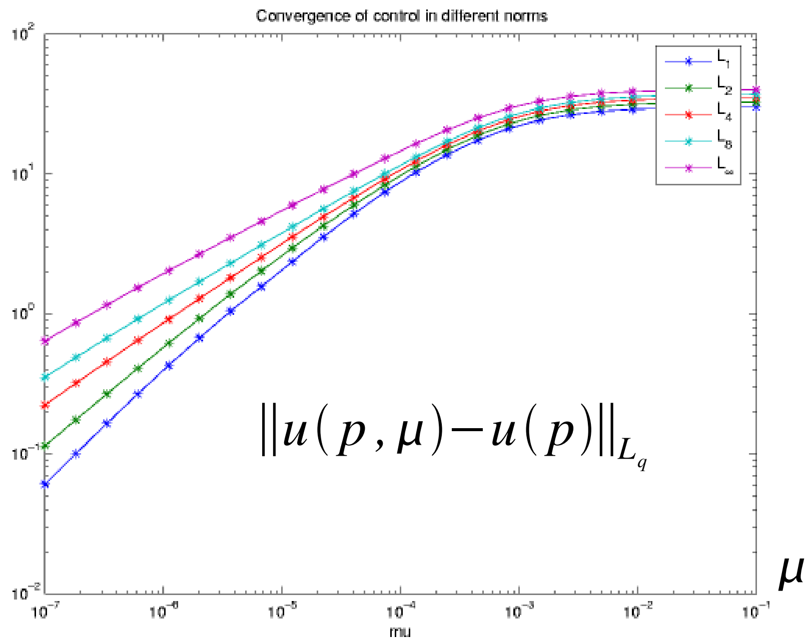
$$K = S^* S : L_2 \to L_\infty$$

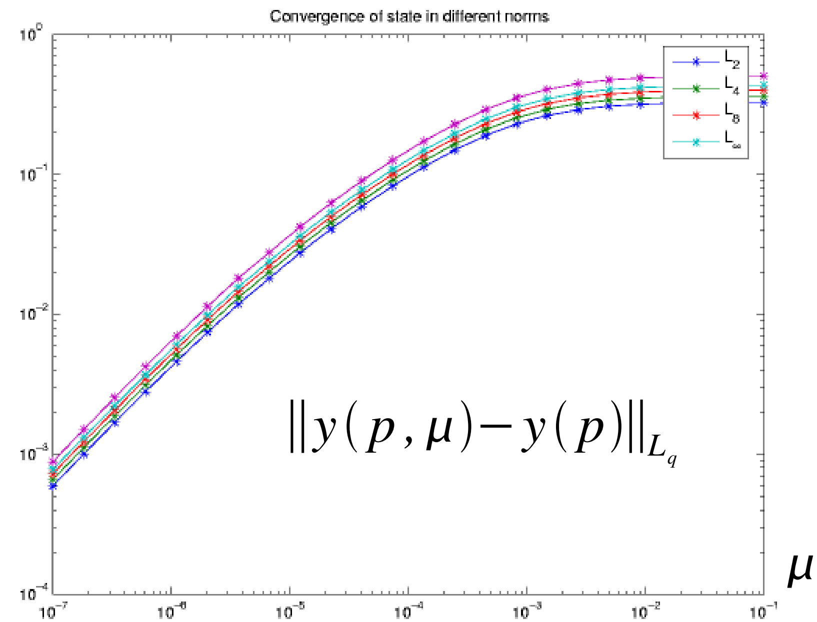
$$l = 2 S^* (p - y_d)$$

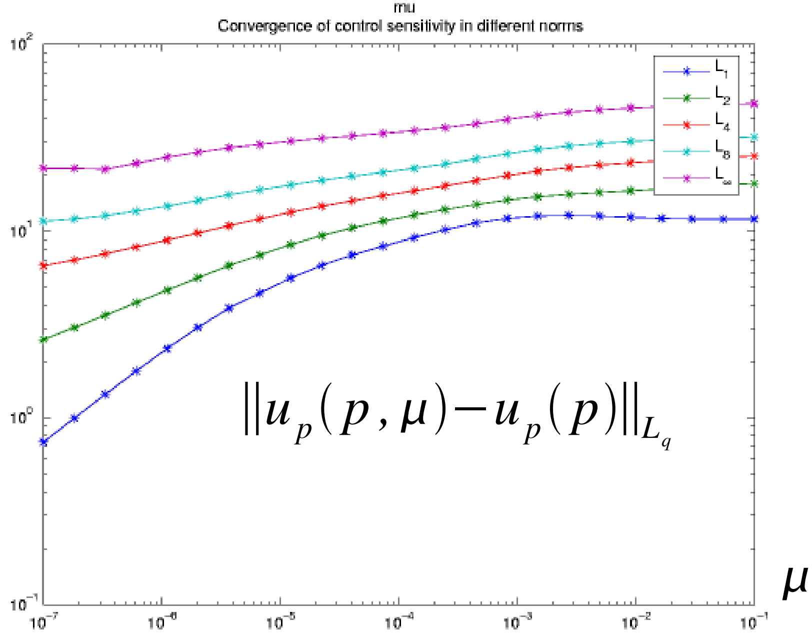
Numerics:

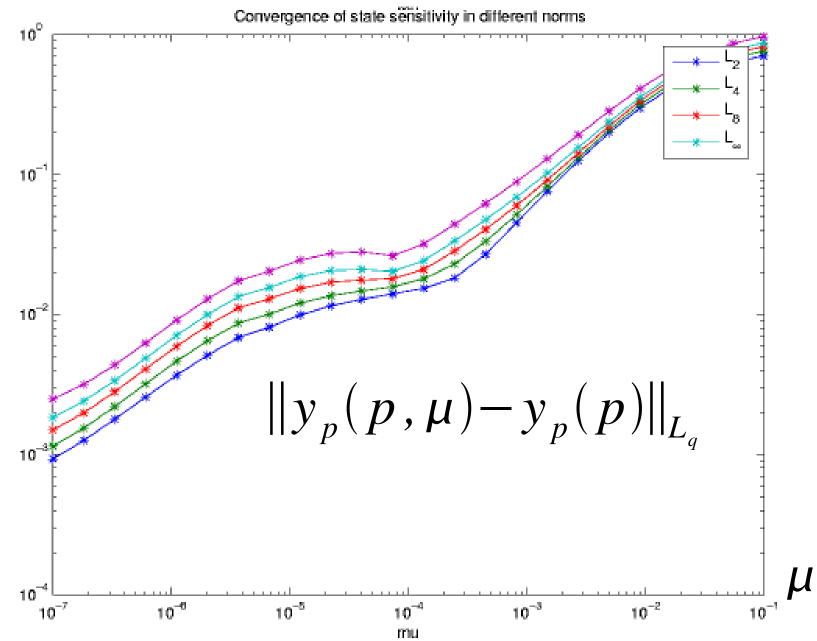
- finite differences
- equidistant grid (512 nodes)
- $\mu \in [10^{-7}, 10^{-1}]$


Control
$u(p,\mu)$
Control Sensitivities
$u_p(p,\mu)$
State
$y(p,\mu)$
State Sensitivities
$y_p(p,\mu)$


Convergence of control in different norms
L_1
L_2
L_4
L_8
L_∞
$\|u(p,\mu)-u(p)\|_{L_q}$
mu
μ


Convergence of state in different norms
L_2
L_4
L_8
L_∞
$\|y(p,\mu)-y(p)\|_{L_q}$
μ


Convergence of control sensitivity in different norms
L_1
L_2
L_4
L_8
L_∞
$\|u_p(p,\mu)-u_p(p)\|_{L_q}$
mu
μ


Convergence of state sensitivity in different norms
L_2
L_4
L_8
L_∞
$\|y_p(p,\mu)-y_p(p)\|_{L_q}$
mu
μ

$$\min_u \|\nabla u\|_{L_2}^2 + p\langle u, l\rangle$$

s.t. $u \geq -1$

$u = 0$ on $\partial\Omega$

$$\Omega = (0,1)^2$$

Framework:

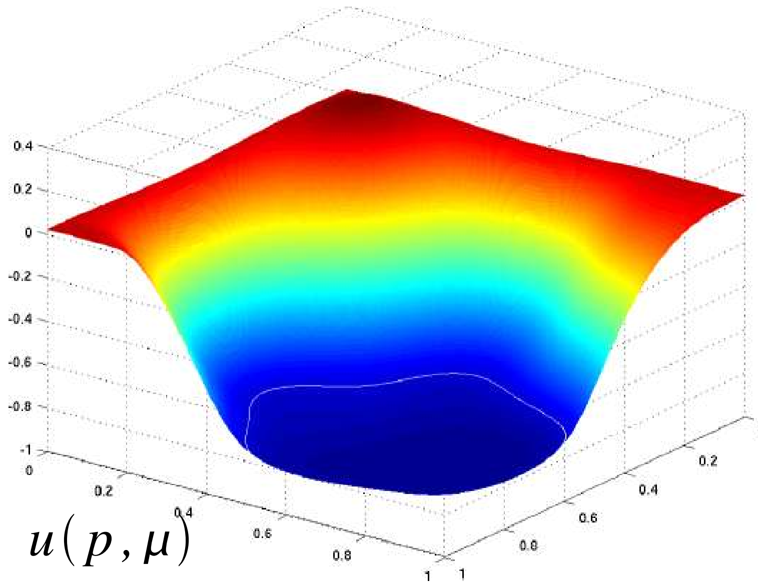
dual problem $\min_\eta \langle \eta, -\Delta^{-1}\eta\rangle + p\langle \eta, \Delta^{-1} l\rangle + \alpha\|\eta\|_{L_2}^2$

s.t. $\eta \geq 0$

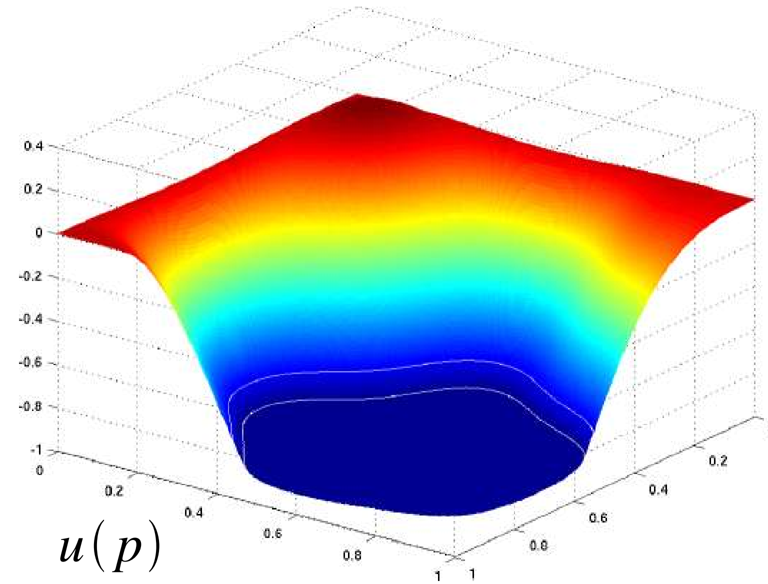
Numerics:

- finite differences
- equidistant grid 300x300
- $\mu \in [10^{-7}, 1/2]$

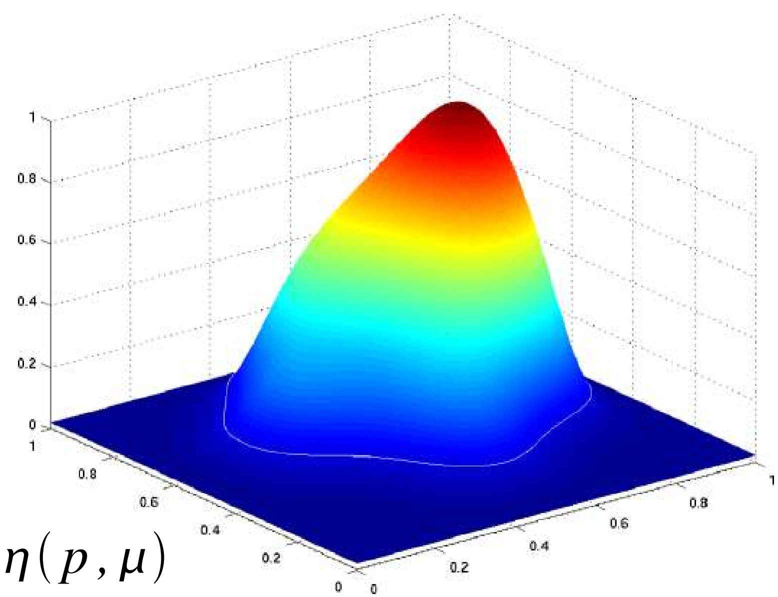
$\mu = 2 \cdot 10^{-2}$ Solutions



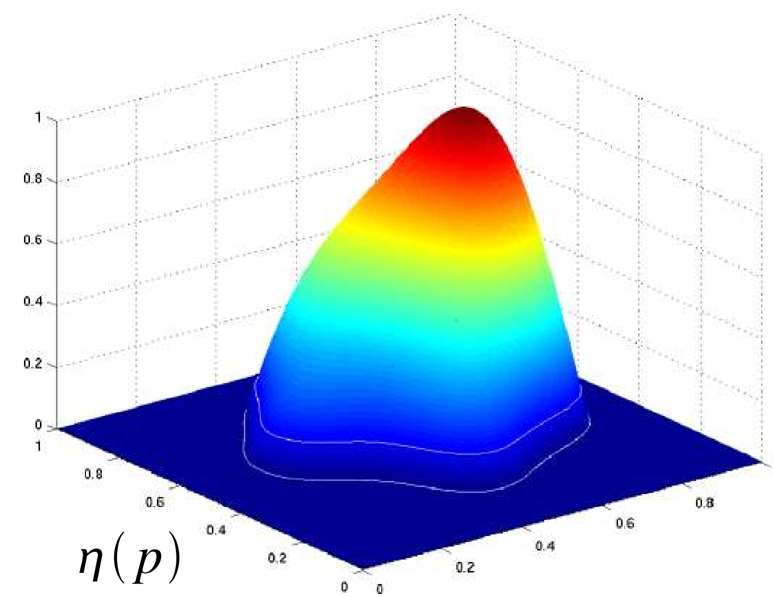
$u(p,\mu)$



$u(p)$

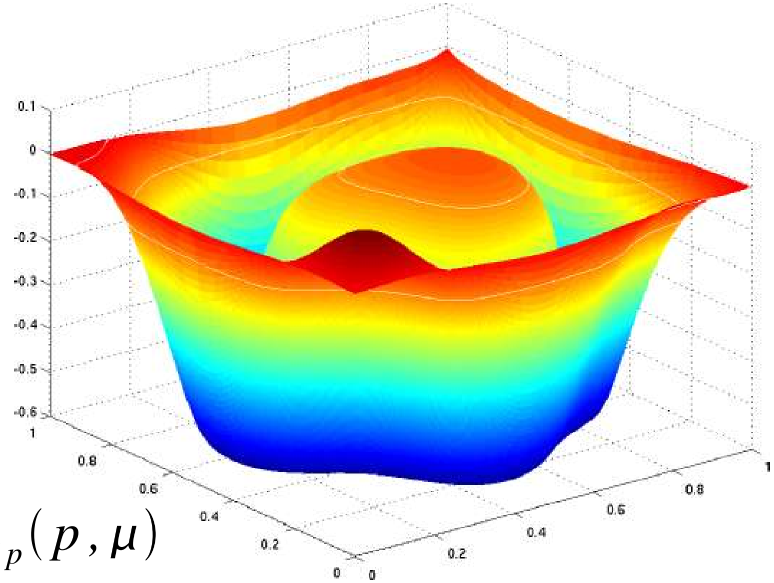


$\eta(p,\mu)$

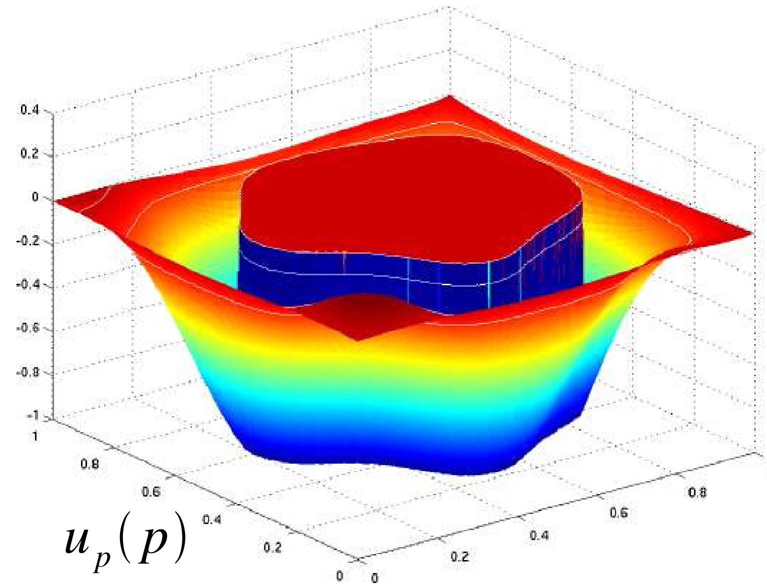


$\eta(p)$

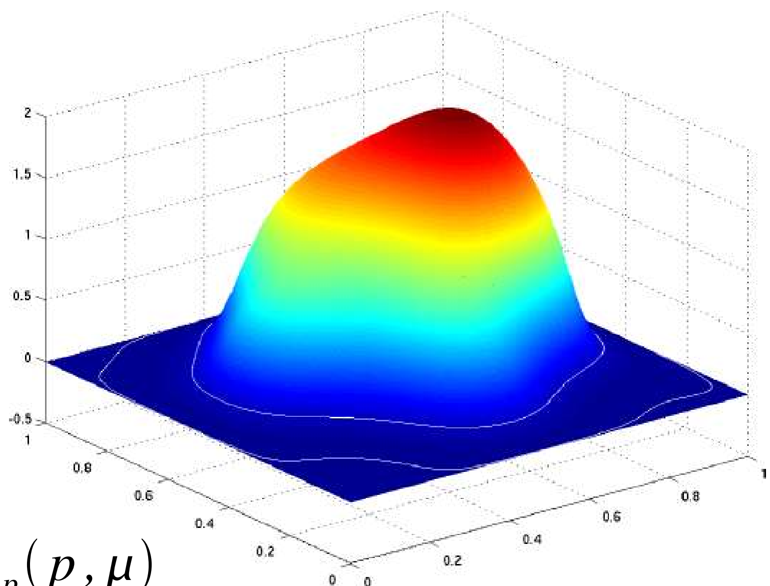
$\mu = 2 \cdot 10^{-2}$ Sensitivities



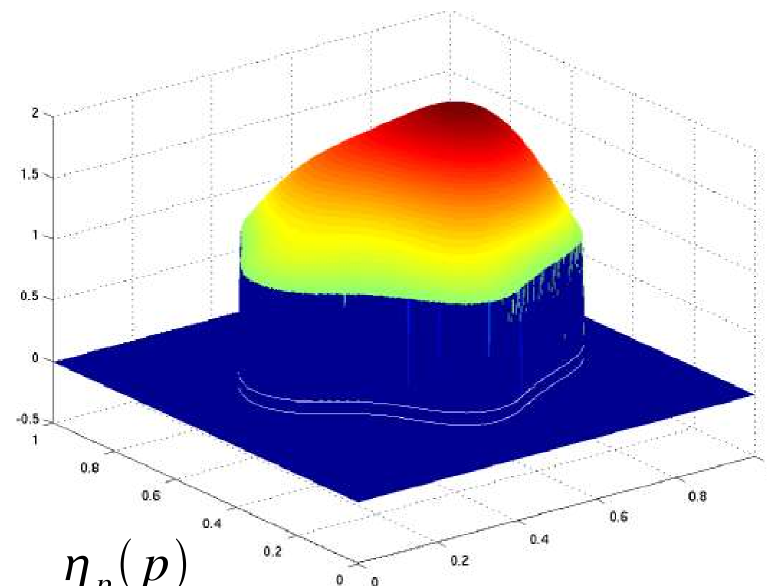
$u_p(p, \mu)$



$u_p(p)$

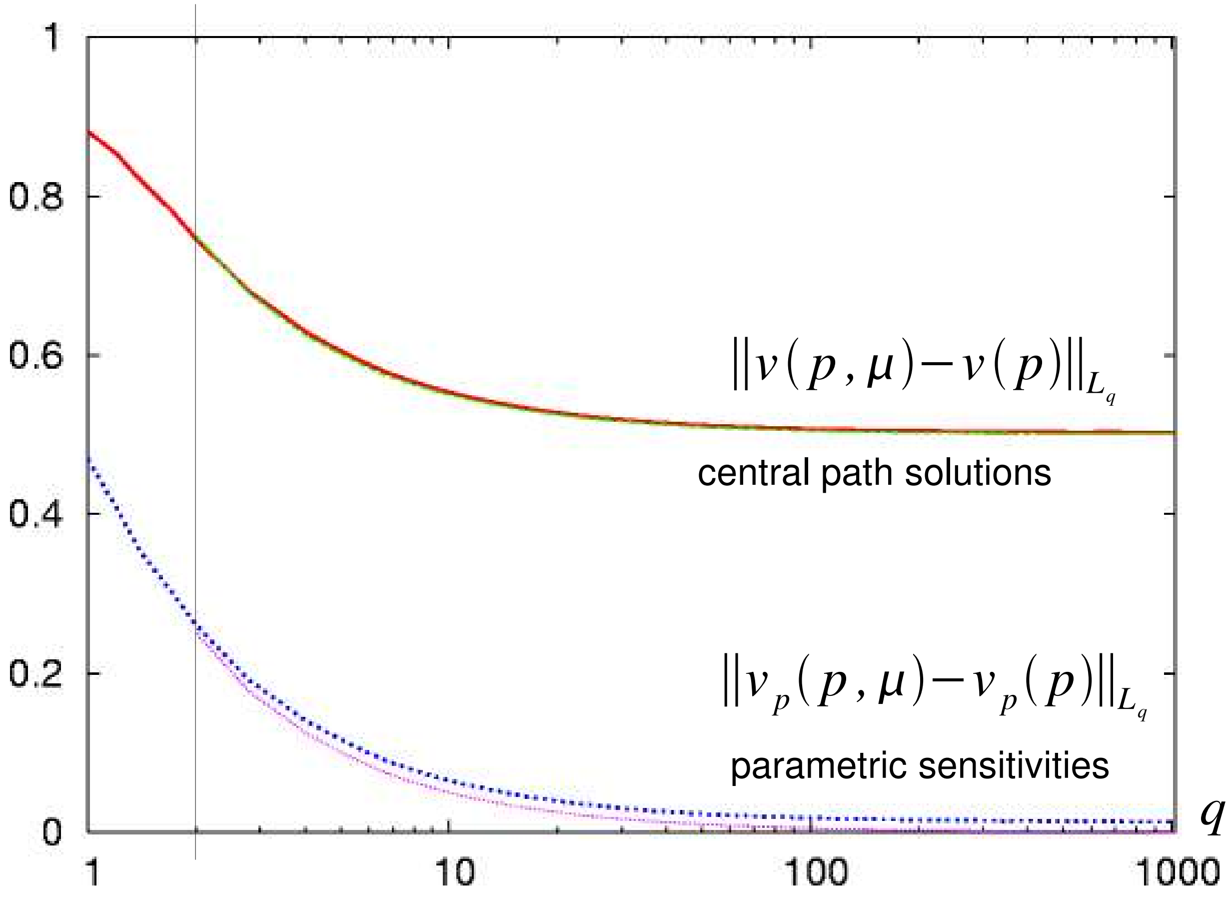


$\eta_p(p, \mu)$



$\eta_p(p)$

Numerically observed convergence rates



Thank you for your attention!