An introduction to domain decomposition methods for optimal control problems

Gabriele Ciaramella

MOX, Politecnico di Milano

Outline

1 Part I: Introduction

Optimal control problems

In these lectures, we consider two stationary/elliptic optimal control problems.

- Parts I, II, III: linear-quadratic control problem:

$$\min_{y,u} \; J(y,u) := \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\nu}{2}\|u\|_{L^2}^2,$$

s.t.

$$\begin{aligned} -\Delta y &= f + u \text{ in } \Omega, \\ y &= 0 \text{ on } \partial\Omega. \end{aligned}$$

- Part IV: nonlinear and nonsmooth control problem:

$$\min_{y,u} \; J(y,u) := \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1},$$

s.t.

$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + u \text{ in } \Omega, \\ y &= 0 \text{ on } \partial\Omega, \end{aligned}$$

and

$$u \in U_{\text{ad}} := \{v \in L^2(\Omega) \;:\; u_\ell(x) \le u(x) \le u_u(x) \text{ in } \Omega\}.$$

First-order optimality systems

- First part: linear-quadratic control problem:

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p &= y_d - y && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ \nu u &= p && \text{in } \Omega. \end{aligned}$$

- Second part: nonlinear and nonsmooth control problem:

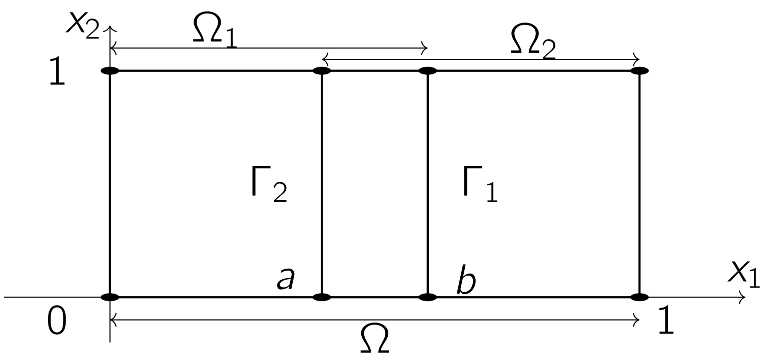
$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p + cp + b\varphi'(y)p &= y_d - y && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ \langle \nu u - p + \beta\lambda, v - u \rangle_{L^2} &\geq 0 && \text{for all } v \in U_{\text{ad}}, \\ \lambda &\in \partial\|u\|_{L^1} && \partial\|u\|_{L^1} \text{ subdifferential of } \|\cdot\|_{L^1} \text{ at } u. \end{aligned}$$

Parallel Schwarz method (Dirichlet)

We eliminate the equation $\nu u = p$:

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p &= y_d - y && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ \nu u &= p && \text{in } \Omega. \end{aligned} \implies \begin{aligned} -\Delta y &= f + \frac{1}{\nu} p && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p &= y_d - y && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega. \end{aligned}$$

Let $\Omega = (0,1)^2$. Consider an overlapping decomposition $\Omega = \Omega_1 \cup \Omega_2$:



$$0 < a < b < 1$$

$L = b - a$ is the overlap

Γ_1 and Γ_2 are the interfaces

The parallel Schwarz method (PSM) is

$$\begin{aligned} -\Delta y_1^n &= f + p_1^n/\nu && \text{in } \Omega_1, \\ y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_1, \\ y_1^n &= y_2^{n-1} && \text{on } \Gamma_1, \\ -\Delta p_1^n &= y_d - y_1^n && \text{in } \Omega_1, \\ p_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_1, \\ p_1^n &= p_2^{n-1} && \text{on } \Gamma_1, \end{aligned} \qquad \begin{aligned} -\Delta y_2^n &= f + p_2^n/\nu && \text{in } \Omega_2, \\ y_2^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_2, \\ y_2^n &= y_1^{n-1} && \text{on } \Gamma_2, \\ -\Delta p_2^n &= y_d - y_2^n && \text{in } \Omega_2, \\ p_2^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_2, \\ p_2^n &= p_1^{n-1} && \text{on } \Gamma_2. \end{aligned}$$

Optimized Parallel Schwarz method (Robin)

The optimized Schwarz method (OSM) is obtained by Robin transmission conditions:

$$
\begin{aligned}
-\Delta y_1^n &= f + p_1^n/\nu && \text{in } \Omega_1, & -\Delta y_2^n &= f + p_2^n/\nu && \text{in } \Omega_2,\\
y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_1, & y_2^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_2,\\
\mathcal{B}_1(y_1^n) &= \mathcal{B}_1(y_2^{n-1}) && \text{on } \Gamma_1, & \mathcal{B}_2(y_2^n) &= \mathcal{B}_2(y_1^{n-1}) && \text{on } \Gamma_2,\\
-\Delta p_1^n &= y_d - y_1^n && \text{in } \Omega_1, & -\Delta p_2^n &= y_d - y_2^n && \text{in } \Omega_2,\\
p_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_1, & p_2^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_2,\\
\mathcal{B}_1(p_1^n) &= \mathcal{B}_1(p_2^{n-1}) && \text{on } \Gamma_1, & \mathcal{B}_2(p_2^n) &= \mathcal{B}_2(p_1^{n-1}) && \text{on } \Gamma_2,
\end{aligned}
$$

where $\mathcal{B}_j(v) = \partial_{n_j} v + qv$, with $q > 0$ (original from Lions 1990).

Different possibilities (Benamou 1996):

- Coupled transmission conditions ($j = 1, 2$ and $\mu < 0$):

$$
\begin{aligned}
\partial_{n_j} y_j^n - \frac{\mu}{\sqrt{\nu}} p_j^n &= \partial_{n_j} y_{3-j}^{n-1} - \frac{\mu}{\sqrt{\nu}} p_{3-j}^{n-1} \quad \text{on } \Gamma_j,\\
\partial_{n_j} p_j^n + \frac{\mu}{\sqrt{\nu}} y_j^n &= \partial_{n_j} p_{3-j}^{n-1} + \frac{\mu}{\sqrt{\nu}} y_{3-j}^{n-1} \quad \text{on } \Gamma_j.
\end{aligned}
$$

- General coupled transmission conditions ($j = 1, 2$, $\mu < 0$ and $q > 0$):

$$
\begin{aligned}
\partial_{n_j} y_j^n + q y_j^n - \frac{\mu}{\sqrt{\nu}} p_j^n &= \partial_{n_j} y_{3-j}^{n-1} + q y_{3-j}^{n-1} - \frac{\mu}{\sqrt{\nu}} p_{3-j}^{n-1} \quad \text{on } \Gamma_j,\\
\partial_{n_j} p_j^n + q p_j^n + \frac{\mu}{\sqrt{\nu}} y_j^n &= \partial_{n_j} p_{3-j}^{n-1} + q p_{3-j}^{n-1} + \frac{\mu}{\sqrt{\nu}} y_{3-j}^{n-1} \quad \text{on } \Gamma_j.
\end{aligned}
$$

Schwarz preserves optimization

In general, Schwarz subdomain problems are optimality systems.

- PSM:

$$\begin{aligned} -\Delta y_j^n &= f + p_j^n/\nu && \text{in } \Omega_j, \\ y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ y_j^n &= y_{3-j}^{n-1} && \text{on } \Gamma_j, \\ -\Delta p_j^n &= y_d - y_j^n && \text{in } \Omega_j, \\ p_j^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ p_j^n &= p_{3-j}^{n-1} && \text{on } \Gamma_j. \end{aligned}$$

$$\begin{aligned} \min_{y_j^n, u_j^n} J_j(y_j^n, u_j^n) &:= \frac{1}{2}\int_{\Omega_j} |y_j^n - y_d|^2 + \frac{\nu}{2}\int_{\Omega_j} |u_j^n|^2 + \int_{\Gamma_j} \partial_{n_j} y_j^n p_{3-j}^{n-1}, \\ \text{s.t.} - \Delta y_j^n &= f + u_j^n \text{ in } \Omega_j, \\ y_j^n &= y_{3-j}^{n-1} \quad \text{on } \Gamma_j, \\ y_j^n &= 0 \quad \text{on } \partial\Omega \cap \partial\Omega_j. \end{aligned}$$

- OSM (Lions, $\mathcal{B}_j(v) = \partial_{n_j} v + qv$):

$$\begin{aligned} -\Delta y_j^n &= f + p_j^n/\nu && \text{in } \Omega_j, \\ y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ \mathcal{B}_j(y_j^n) &= \mathcal{B}_j(y_{3-j}^{n-1}) && \text{on } \Gamma_j, \\ -\Delta p_j^n &= y_d - y_j^n && \text{in } \Omega_j, \\ p_j^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ \mathcal{B}_j(p_j^n) &= \mathcal{B}_j(p_{3-j}^{n-1}) && \text{on } \Gamma_j. \end{aligned}$$

$$\begin{aligned} \min_{y_j^n, u_j^n} J_j(y_j^n, u_j^n) &:= \frac{1}{2}\int_{\Omega_j} |y_j^n - y_d|^2 + \frac{\nu}{2}\int_{\Omega_j} |u_j^n|^2 - \int_{\Gamma_j} y_j^n \mathcal{B}_j(p_{3-j}^{n-1}), \\ \text{s.t.} - \Delta y_j^n &= f + u_j^n \text{ in } \Omega_j, \\ \mathcal{B}_j(y_j^n) &= \mathcal{B}_j(y_{3-j}^{n-1}) \quad \text{on } \Gamma_j, \\ y_j^n &= 0 \quad \text{on } \partial\Omega \cap \partial\Omega_j. \end{aligned}$$

Remark: use your favorite solver (linear system/root finder or optimization algorithm).

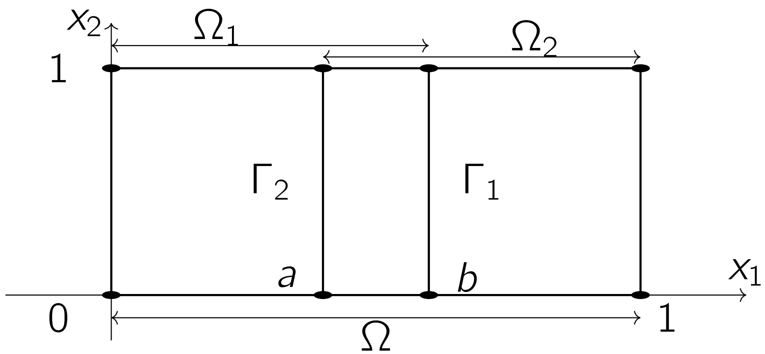
References

- *Iterative Methods and Preconditioners for Systems of Linear Equations*,
  Ciaramella and Gander, SIAM Fundamentals of Algorithms (2022)
- *On the Schwarz method for unconstrained elliptic optimal control problems*,
  Ciaramella and Kwok, in preparation (2022)
- *Nonlinear optimized Schwarz preconditioner for elliptic optimal control problems*,
  Ciaramella, Kwok, Müller, DD26 Proceedings (2022)
- *Schwarz methods over the course of time*,
  Gander, ETNA (2008)
- *Domain Decomposition Methods - Algorithms and Theory*,
  Toselli and Widlund, Springer (2005)
- *A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations*,
  Benamou, SIAM J. Numer. Anal. (1996)
- *On the Schwarz alternating method. III*,
  Lions, Third DD Conference Papers (1990)
- *On the Schwarz alternating method. I*,
  Lions, First DD Conference Papers (1988)

2 Part II: Parallel Schwarz methods

Convergence: a two-subdomain overview

$$\begin{aligned}
-\Delta y_j^n &= f + p_j^n/\nu && \text{in } \Omega_j,\\
y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j,\\
y_j^n &= y_{3-j}^{n-1} && \text{on } \Gamma_j,\\
-\Delta p_j^n &= y_d - y_j^n && \text{in } \Omega_j,\\
p_j^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j,\\
p_j^n &= p_{3-j}^{n-1} && \text{on } \Gamma_j.
\end{aligned}$$



$a = \frac{1}{2} - \frac{L}{2}$
$b = \frac{1}{2} + \frac{L}{2}$

Expand the errors $\mathrm{y}_j^n := y - y_j^n$ and $\mathrm{p}_j^n := p - p_j^n$ in Fourier sine series ($k_m = \pi m$):

$$\mathrm{y}_j^n(x_1, x_2) = \sum_{m=1}^{\infty} \widehat{\mathrm{y}}_j^n(x_1, k_m)\sin(k_m x_2) \quad \text{and} \quad \mathrm{p}_j^n(x_1, x_2) = \sum_{m=1}^{\infty} \widehat{\mathrm{p}}_j^n(x_1, k_m)\sin(k_m x_2).$$

Then, one can show that (study $\widehat{\mathrm{y}}_j^n$ and $\widehat{\mathrm{p}}_j^n$ and use Parseval) there is a $C > 0$ such that

$$\max(\|\mathrm{y}_j^n\|_{L^2}, \|\mathrm{p}_j^n\|_{L^2}) \le C \max_{m=1,2,\dots} \rho(k_m, \nu, L)^n \max(\|\mathrm{y}_j^0\|_{L^2}, \|\mathrm{p}_j^0\|_{L^2}),$$

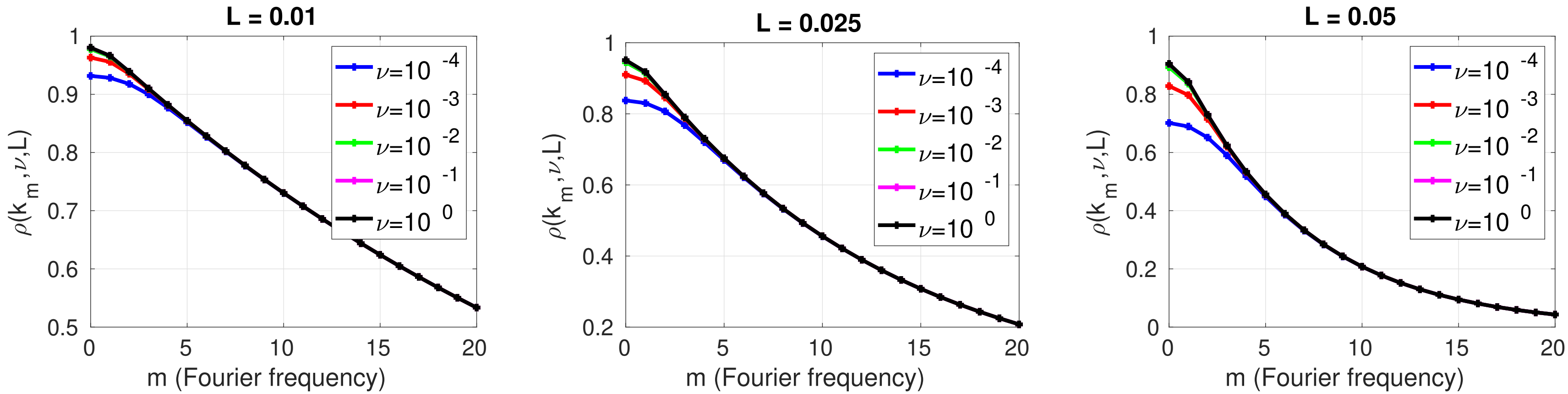
where

$$\rho(k_m, \nu, L) = \left| \frac{\sinh(\lambda(1/2 - L))}{\sinh(\lambda(1/2 + L))} \right|,$$

with $\lambda = g_1 + i g_2$, $g_1 = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{\nu} + k_m^4} + k_m^2}$ and $g_2 = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{\nu} + k_m^4} - k_m^2}$.

Convergence: a two-subdomain overview

$$\rho(k_m, \nu, L) = \left| \frac{\sinh(\lambda(1/2 - L))}{\sinh(\lambda(1/2 + L))} \right|$$

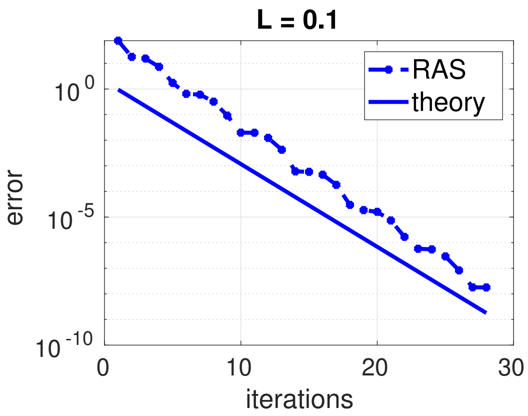


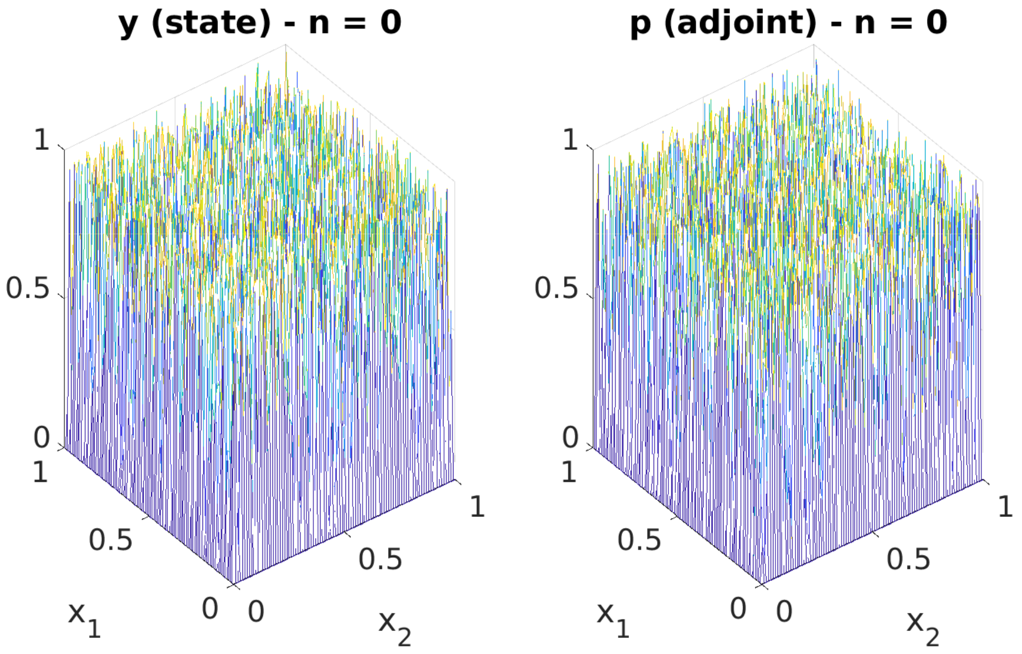
Remark:

- In 2D: $m = 1, 2, \ldots$
- In 1D: $m = 0$.

Convergence: a two-subdomain overview

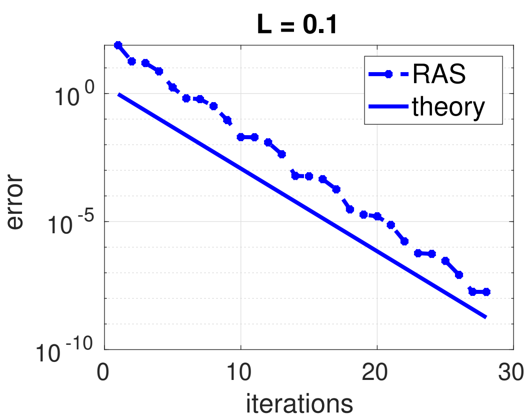
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

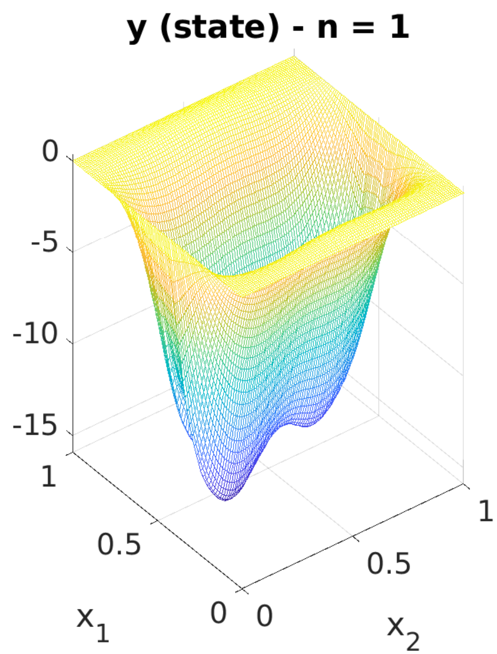


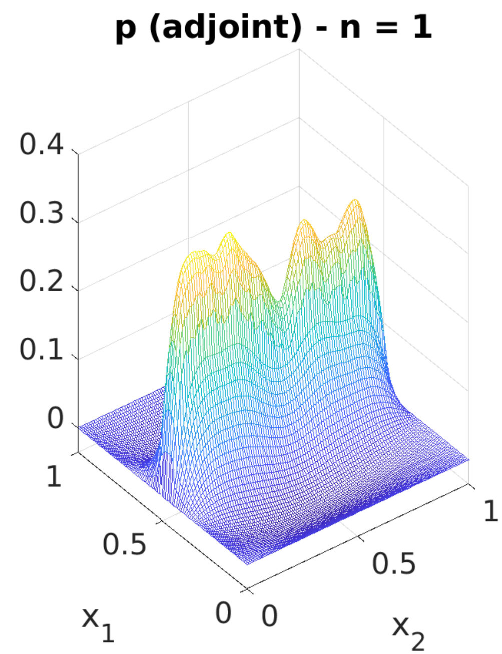


Convergence: a two-subdomain overview

Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

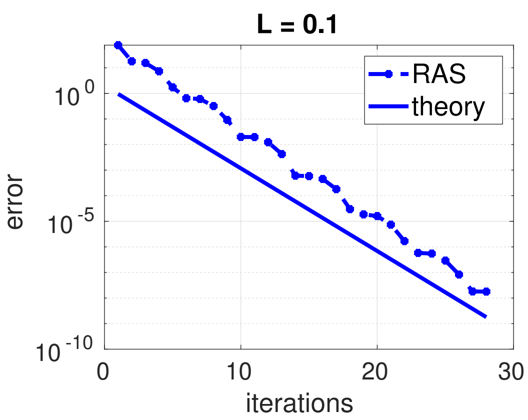


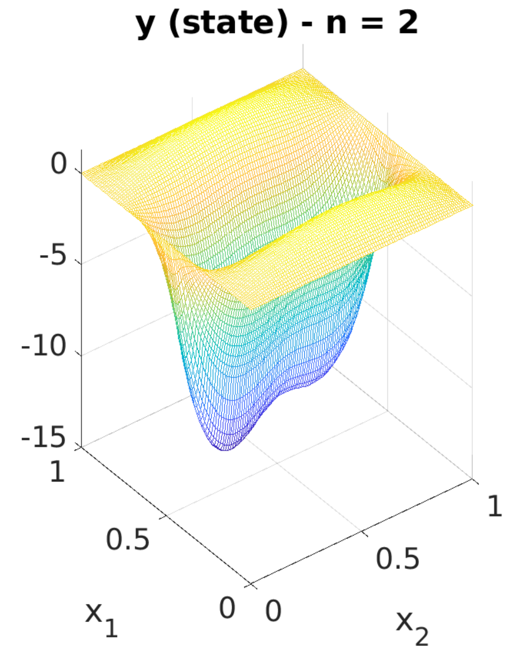


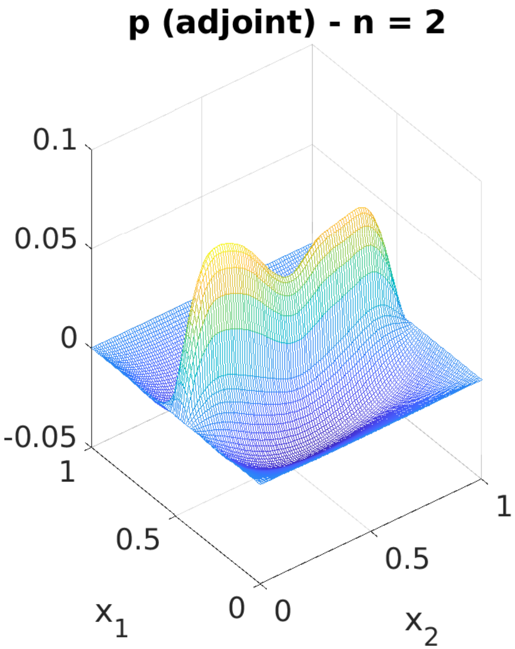


Convergence: a two-subdomain overview

Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

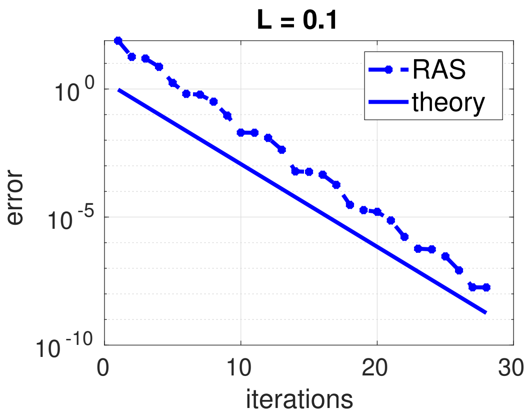


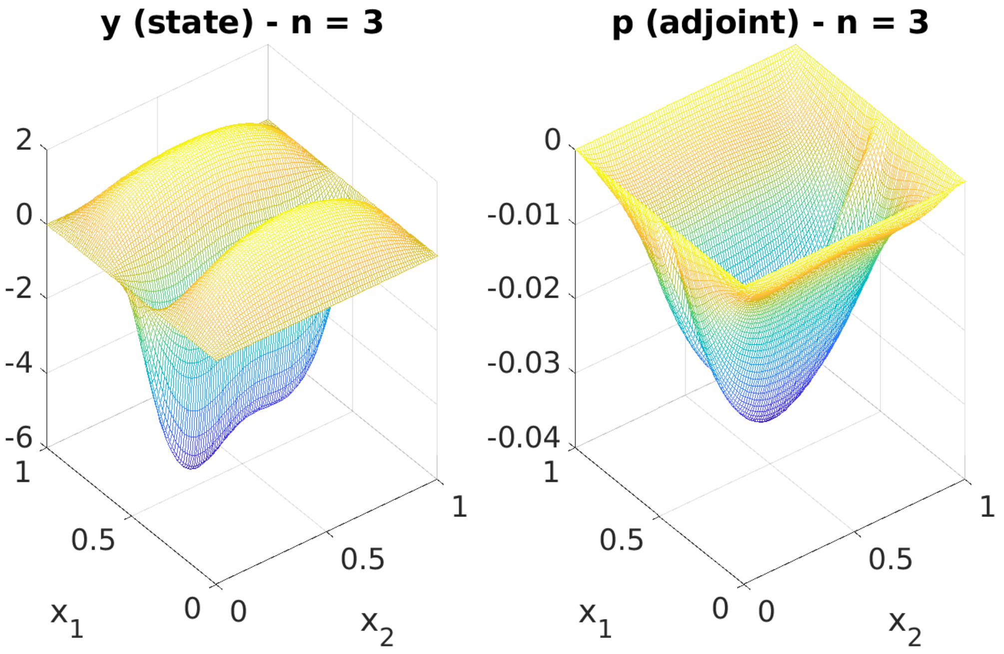




Convergence: a two-subdomain overview

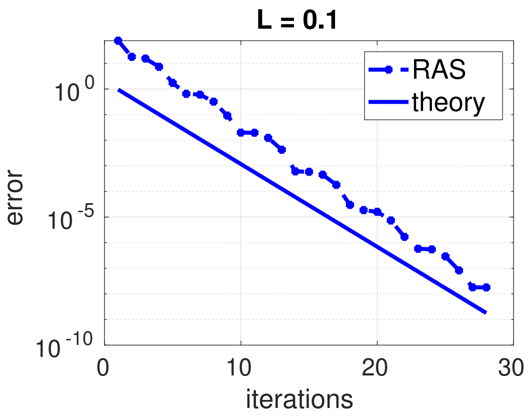
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

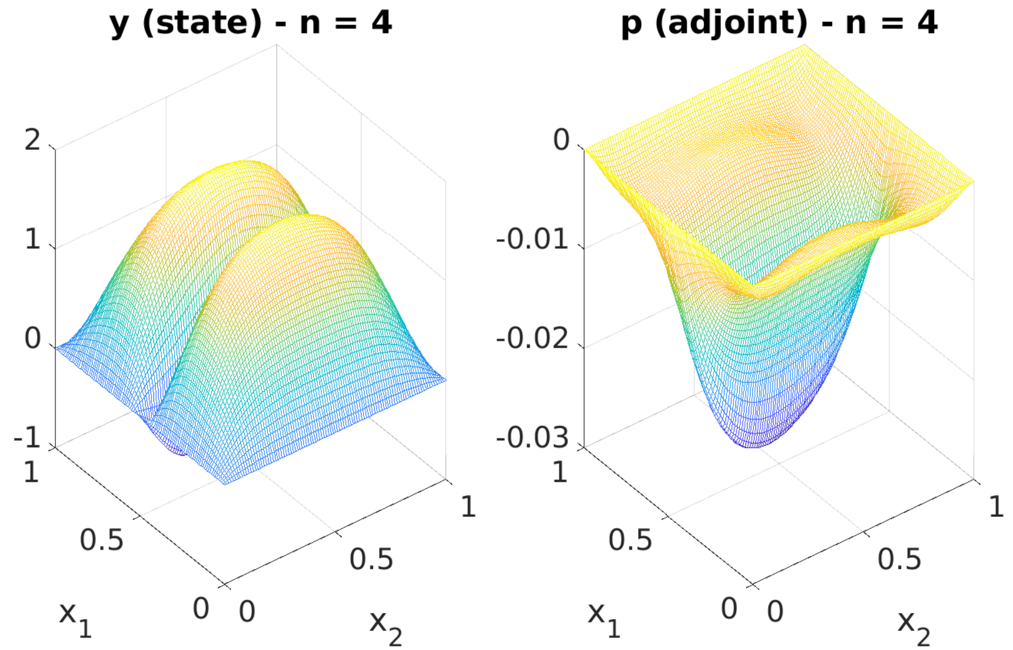




Convergence: a two-subdomain overview

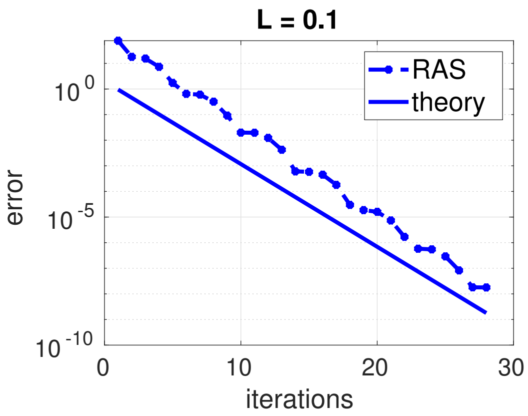
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

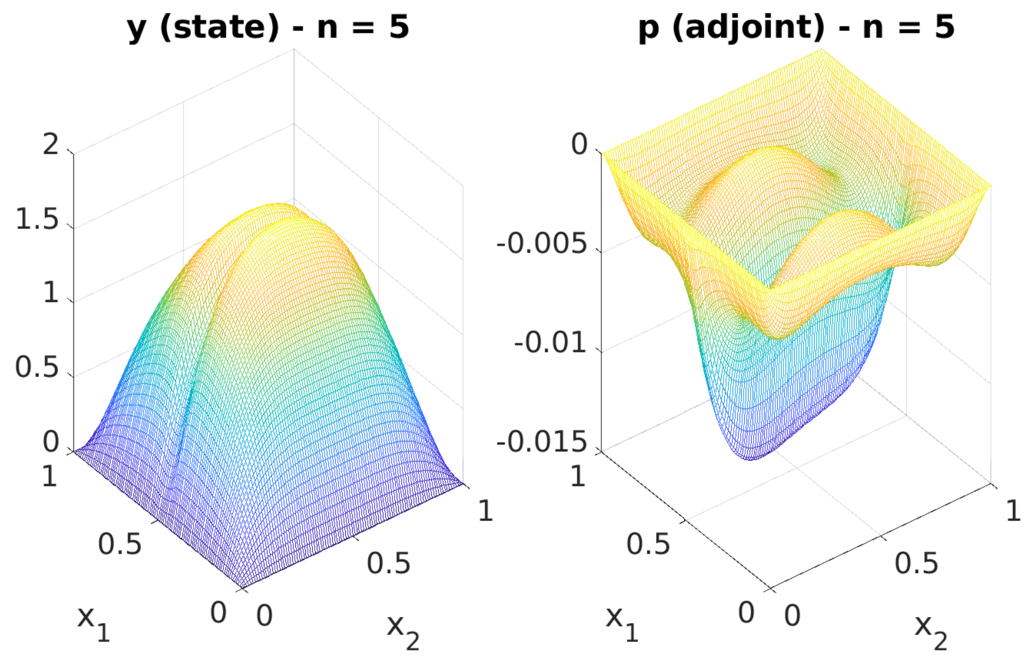




Convergence: a two-subdomain overview

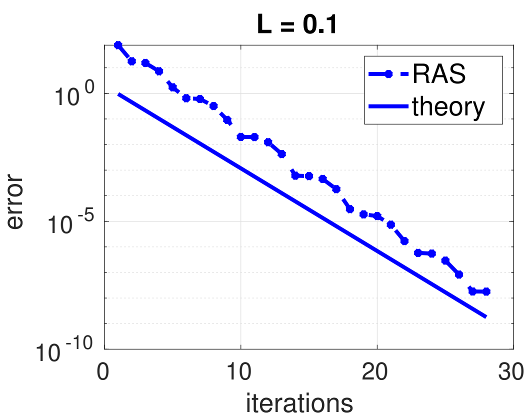
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

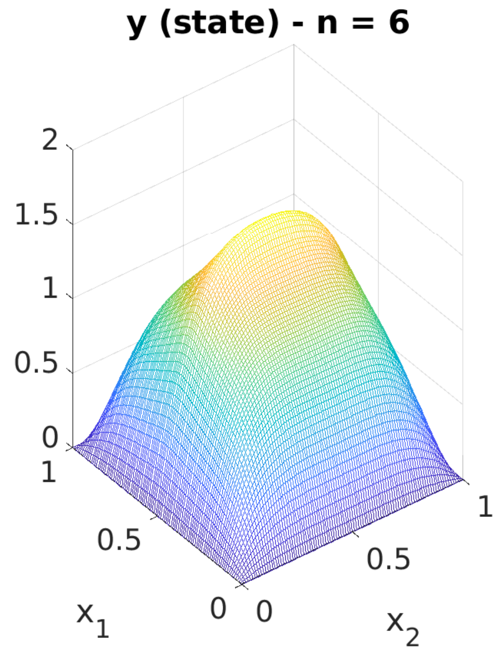


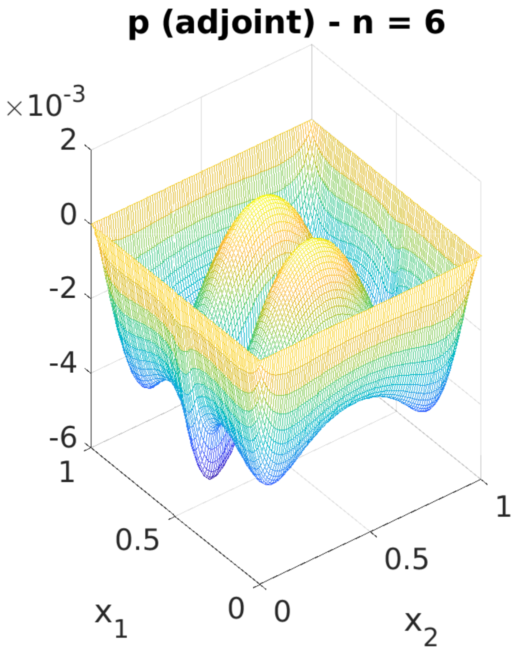


Convergence: a two-subdomain overview

Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

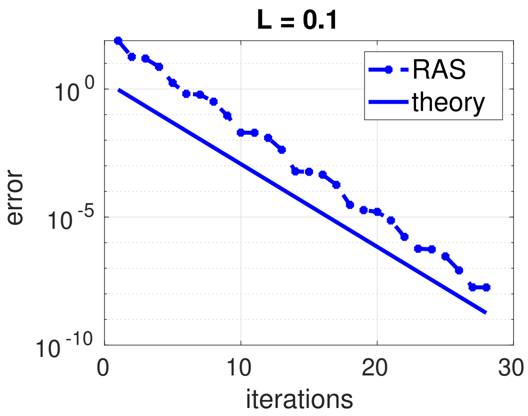


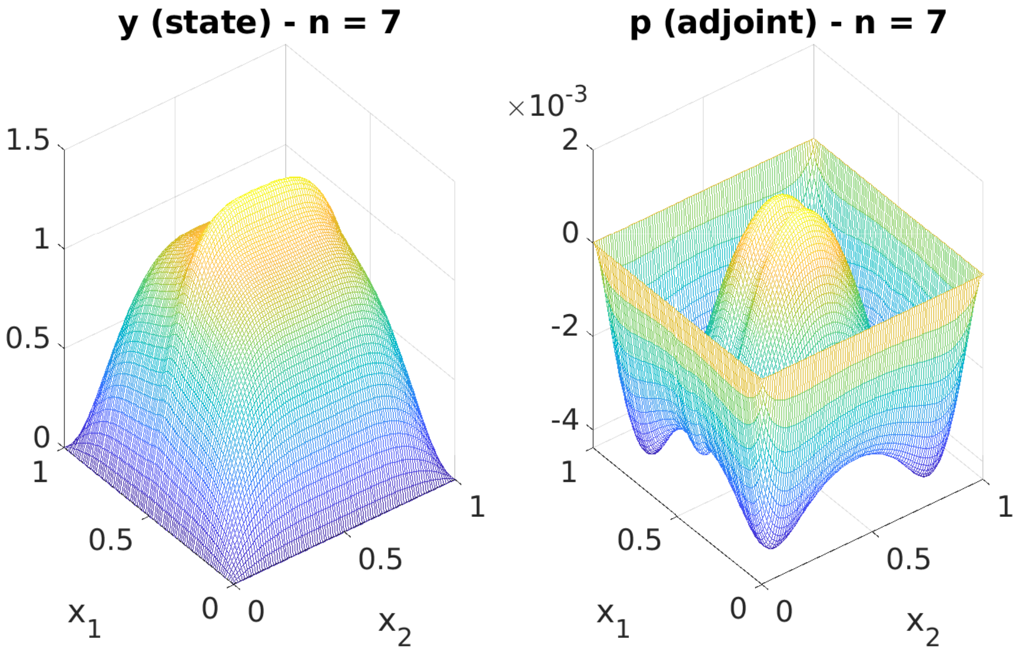




Convergence: a two-subdomain overview

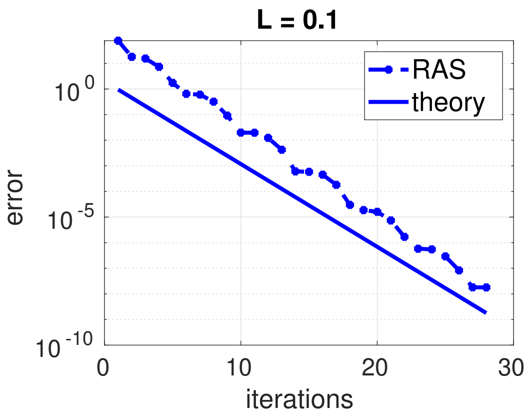
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

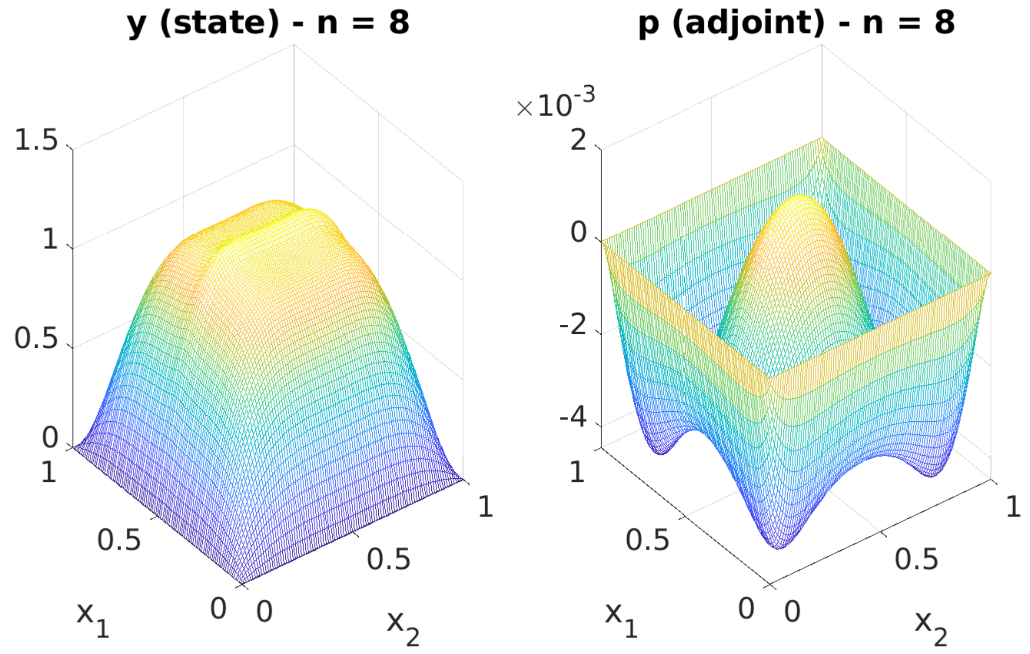




Convergence: a two-subdomain overview

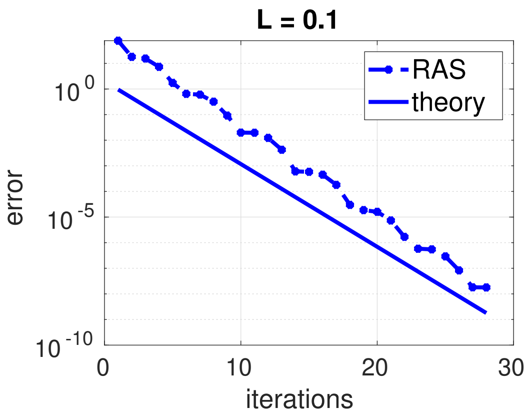
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

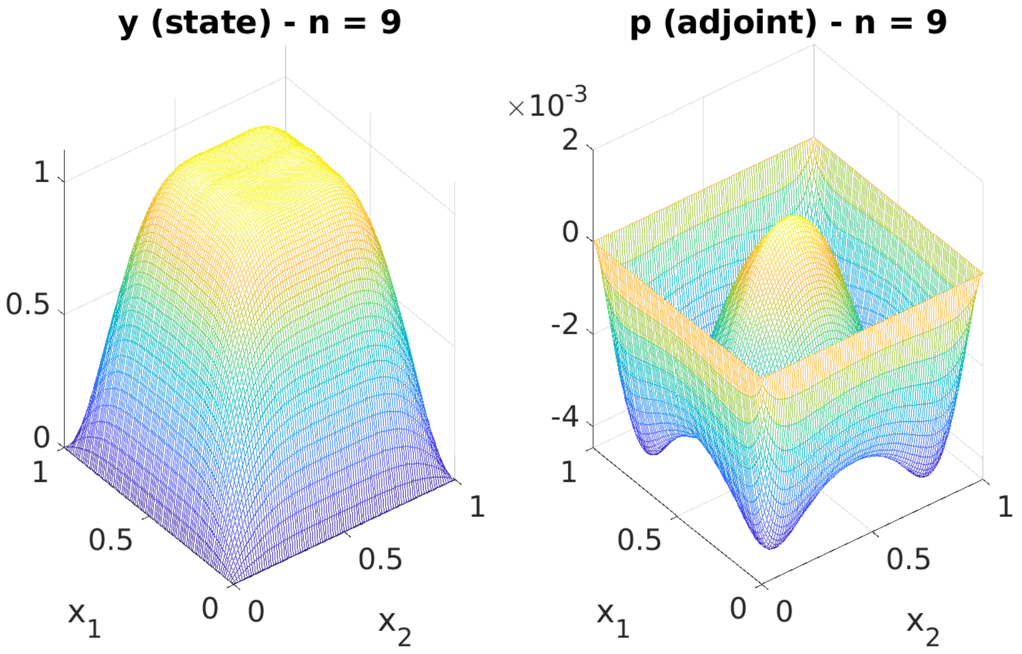




Convergence: a two-subdomain overview

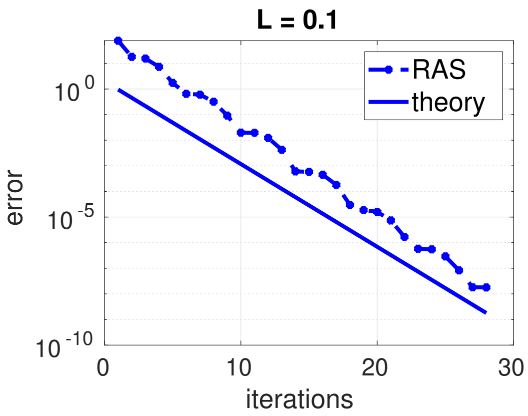
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

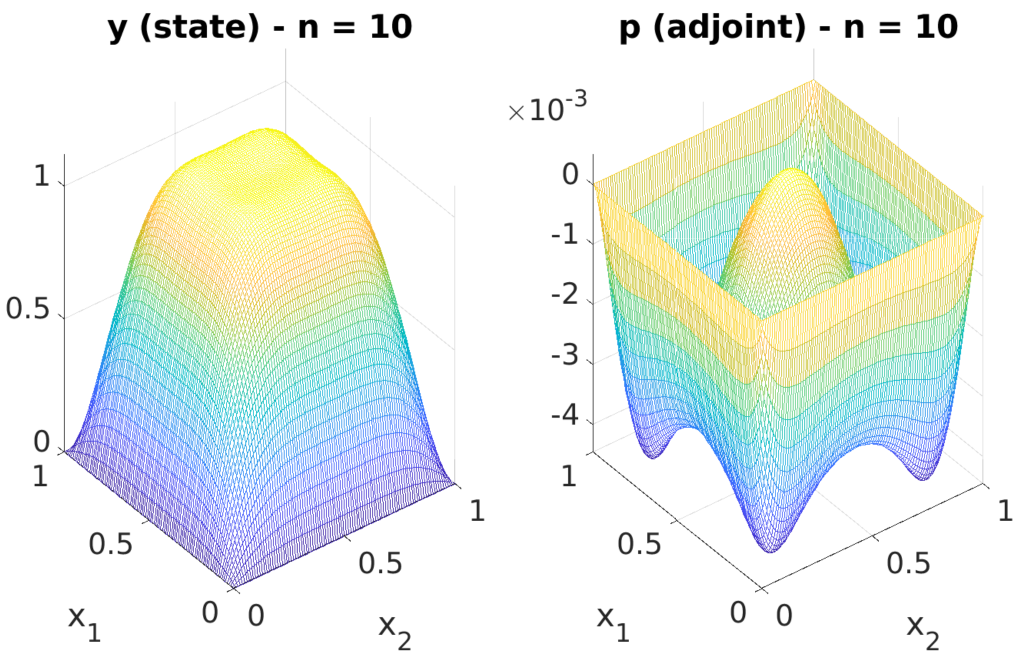




Convergence: a two-subdomain overview

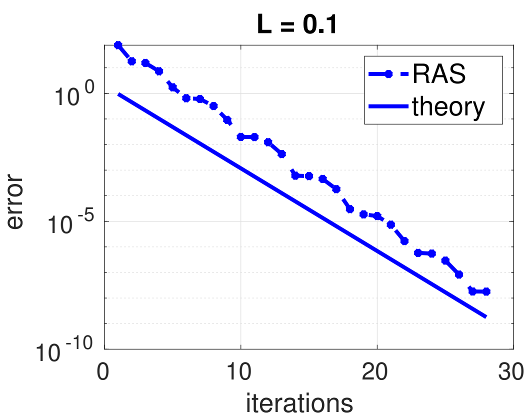
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

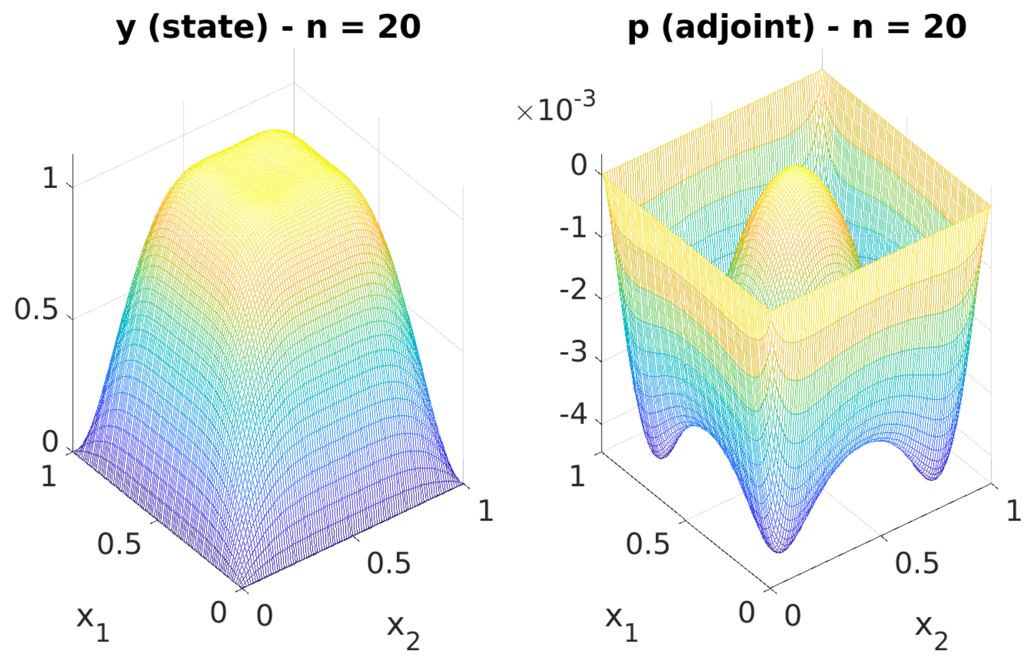




Convergence: a two-subdomain overview

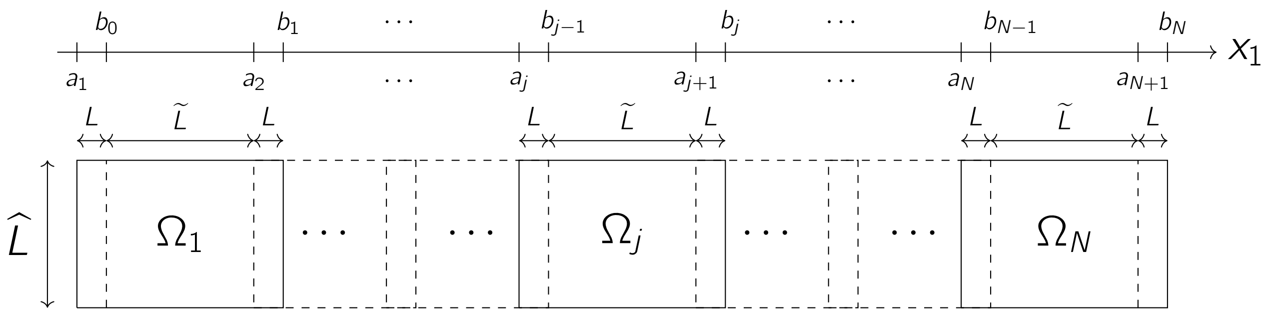
Example: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.





Convergence analysis by Fourier series

Consider the following multiple subdomain decomposition $\Omega = \cup_{j=1}^{N} \Omega_j$.



The PSM for the solution of this problem is given by

$$\begin{aligned} -\Delta y_j^n &= f + p_j^n/\nu \ \text{ in } \Omega_j, \\ y_j^n(\cdot,0) &= 0,\ y_j^n(\cdot,\widehat{L}) = 0, \\ y_j^n(a_j,\cdot) &= y_{j-1}^{n-1}(a_j,\cdot), \\ y_j^n(b_j,\cdot) &= y_{j+1}^{n-1}(b_j,\cdot), \end{aligned} \qquad \begin{aligned} -\Delta p_j^n &= y_d - y_j^n \ \text{ in } \Omega_j, \\ p_j^n(\cdot,0) &= 0,\ p_j^n(\cdot,\widehat{L}) = 0, \\ p_j^n(a_j,\cdot) &= p_{j-1}^{n-1}(a_j,\cdot), \\ p_j^n(b_j,\cdot) &= p_{j+1}^{n-1}(b_j,\cdot). \end{aligned}$$

The errors $\mathrm{y}_j^n := y - y_j^n$ and $\mathrm{p}_j^n := p - p_j^n$ satisfy the equations

$$\begin{aligned} -\Delta \mathrm{y}_j^n &= \mathrm{p}_j^n/\nu \ \text{ in } \Omega_j, \\ \mathrm{y}_j^n(\cdot,0) &= 0,\ \mathrm{y}_j^n(\cdot,\widehat{L}) = 0, \\ \mathrm{y}_j^n(a_j,\cdot) &= \mathrm{y}_{j-1}^{n-1}(a_j,\cdot), \\ \mathrm{y}_j^n(b_j,\cdot) &= \mathrm{y}_{j+1}^{n-1}(b_j,\cdot), \end{aligned} \qquad \begin{aligned} -\Delta \mathrm{p}_j^n &= -\mathrm{y}_j^n \ \text{ in } \Omega_j, \\ \mathrm{p}_j^n(\cdot,0) &= 0,\ \mathrm{p}_j^n(\cdot,\widehat{L}) = 0, \\ \mathrm{p}_j^n(a_j,\cdot) &= \mathrm{p}_{j-1}^{n-1}(a_j,\cdot), \\ \mathrm{p}_j^n(b_j,\cdot) &= \mathrm{p}_{j+1}^{n-1}(b_j,\cdot). \end{aligned}$$

Expand the errors $\mathrm{y}_j^n := y - y_j^n$ and $\mathrm{p}_j^n := p - p_j^n$ in Fourier sine series ($k_m = \pi m/\widehat{L}$):

$$\mathrm{y}_j^n(x_1,x_2) = \sum_{m=1}^{\infty} \widehat{\mathrm{y}}_j^n(x_1,k_m)\sin(k_m x_2) \quad \text{and} \quad \mathrm{p}_j^n(x_1,x_2) = \sum_{m=1}^{\infty} \widehat{\mathrm{p}}_j^n(x_1,k_m)\sin(k_m x_2).$$

The Fourier coefficients $\widehat{\mathrm{y}}_j^n(\cdot,k_m)$ and $\widehat{\mathrm{p}}_j^n(\cdot,k_m)$ satisfy

$$\begin{aligned} -(\widehat{\mathrm{y}}_j^n)'' + k_m^2\widehat{\mathrm{y}}_j^n &= \widehat{\mathrm{p}}_j^n/\nu \ \text{ in } (a_j,b_j), \\ \widehat{\mathrm{y}}_j^n(a_j) &= \widehat{\mathrm{y}}_{j-1}^{n-1}(a_j), \\ \widehat{\mathrm{y}}_j^n(b_j) &= \widehat{\mathrm{y}}_{j+1}^{n-1}(b_j), \end{aligned} \qquad \begin{aligned} -(\widehat{\mathrm{p}}_j^n)'' + k_m^2\widehat{\mathrm{p}}_j^n &= -\widehat{\mathrm{y}}_j^n \ \text{ in } (a_j,b_j), \\ \widehat{\mathrm{p}}_j^n(a_j) &= \widehat{\mathrm{p}}_{j-1}^{n-1}(a_j), \\ \widehat{\mathrm{p}}_j^n(b_j) &= \widehat{\mathrm{p}}_{j+1}^{n-1}(b_j). \end{aligned}$$

Using the adjoint equation $-(\widehat{\mathrm{y}}_j^n)'' + k_m^2\widehat{\mathrm{p}}_j^n = -\widehat{\mathrm{y}}_j^n$, we obtain that $\widehat{\mathrm{p}}_j^n$ solves

$$\begin{aligned} (\widehat{\mathrm{p}}_j^n)^{IV} - 2k_m^2(\widehat{\mathrm{p}}_j^n)'' + (k^4 + 1/\nu)\widehat{\mathrm{p}}_j^n &= 0, \ \text{ in } (a_j,b_j), \\ \widehat{\mathrm{p}}_j^n(a_j) &= \widehat{\mathrm{p}}_{j-1}^{n-1}(a_j), \\ \widehat{\mathrm{p}}_j^n(b_j) &= \widehat{\mathrm{p}}_{j+1}^{n-1}(b_j), \\ (\widehat{\mathrm{p}}_j^n)''(a_j) &= (\widehat{\mathrm{p}}_{j-1}^{n-1})''(a_j), \\ (\widehat{\mathrm{p}}_j^n)''(b_j) &= (\widehat{\mathrm{p}}_{j+1}^{n-1})''(b_j). \end{aligned}$$

The solution to this fourth-order ODE has the form

$$\widehat{\mathrm{p}}_j^n(x) = A_j^n\sinh(\lambda_1(x-a_j)) + B_j^n\sinh(\lambda_2(x-a_j)) + C_j^n\sinh(\lambda_1(x-b_j)) + D_j^n\sinh(\lambda_2(x-b_j)),$$

where $\lambda_1 = g_1 + ig_2$, $\lambda_2 = g_1 - ig_2$, $g_1 = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{\nu}+k_m^4}+k_m^2}$ and $g_2 = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{\nu}+k_m^4}-k_m^2}$.

Using the transmission conditions and defining the vector

$$\mathbf{v}_{k_m}^n := \left[A_1^n , B_1^n , C_1^n , D_1^n , \cdots , A_N^n , B_N^n , C_N^n , D_N^n \right]^\top$$

we obtain the following equivalent form of the PSM

$$\mathbf{v}_{k_m}^n = G_{k_m} \mathbf{v}_{k_m}^{n-1}.$$

Here, the iteration matrix G_{k_m} is given by

$$G_{k_m} := \begin{bmatrix} & G_R & & & & \\ G_L & & G_R & & & \\ & G_L & & G_R & & \\ & & \ddots & & \ddots & \\ & & & G_L & & G_R \\ & & & & G_L & \end{bmatrix},$$

where

$$G_L := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{s_1(L+\widetilde{L})}{s_1(2L+\widetilde{L})} & 0 & \frac{s_1(L)}{s_1(2L+\widetilde{L})} & 0 \\ 0 & -\frac{s_2(L+\widetilde{L})}{s_2(2L+\widetilde{L})} & 0 & \frac{s_2(L)}{s_2(2L+\widetilde{L})} \end{bmatrix}, \quad G_R := \begin{bmatrix} \frac{s_1(L+\widetilde{L})}{s_1(2L+\widetilde{L})} & 0 & -\frac{s_1(L)}{s_1(2L+\widetilde{L})} & 0 \\ 0 & \frac{s_2(L+\widetilde{L})}{s_2(2L+\widetilde{L})} & 0 & -\frac{s_2(L)}{s_2(2L+\widetilde{L})} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

with $s_1(x) = \sinh(\lambda_1 x)$ and $s_2(x) = \sinh(\lambda_2 x)$.

The iteration

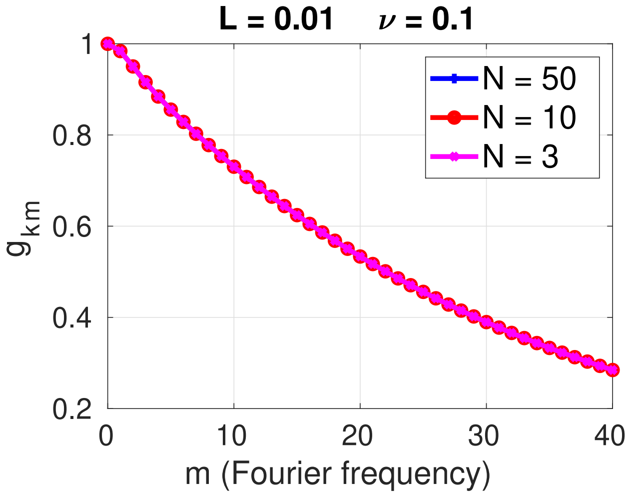
$$\mathbf{v}_{k_m}^n = G_{k_m}\mathbf{v}_{k_m}^{n-1}.$$

converges if and only if $\rho(G_{k_m}) < 1$. Thus, we estimate

$$\rho(G_{k_m}) \leq \|G_{k_m}\|_\infty = g_{k_m}(L, \widetilde{L}, \nu),$$

where

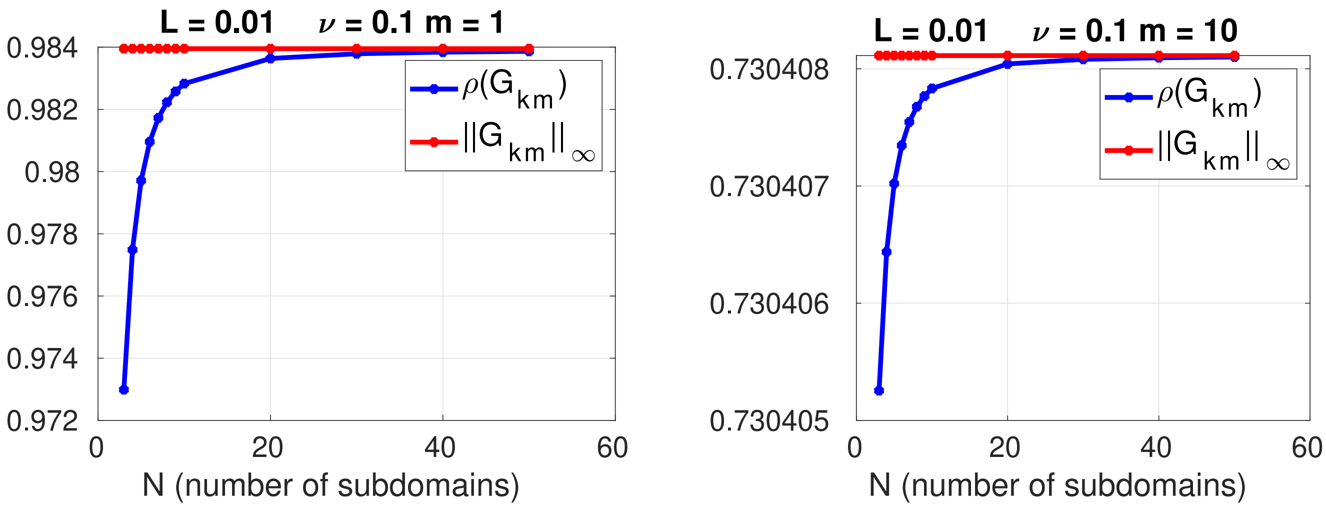
$$g_{k_m}(L, \widetilde{L}, \nu) = \max\left(\frac{|s_1(L+\widetilde{L})| + |s_1(L)|}{|s_1(2L+\widetilde{L})|}, \frac{|s_2(L+\widetilde{L})| + |s_2(L)|}{|s_2(2L+\widetilde{L})|}\right).$$



Thus, $g_{k_m}(L, \widetilde{L}, \nu) \leq g_{k_1}(L, \widetilde{L}, \nu) < 1$ and we obtain convergence.

Remarks:

- The dependence on the parameters ν and L is as in the two-subdomain case.
- The curves in the plot seem to be independent of N, the number of subdomains.



Assume that an algorithm converges geometrically and is stopped at a given tolerance `Tol`:

$$\frac{\|e^n\|}{\|e^0\|} \leq \gamma^n \quad \text{and} \quad \frac{\|e^n\|}{\|e^0\|} \approx \texttt{Tol}.$$

We can then write

$$\texttt{Tol} \approx \frac{\|e^n\|}{\|e^0\|} \leq \gamma^n,$$

and estimate

$$n \leq \frac{|\log \texttt{Tol}|}{|\log \gamma|}.$$

If γ is independent of the number of subdomains N, then the method converges, for N large, in a number of iterations independent of N (weakly scalability).

Remark:

- In our tests the size of the subdomains was kept constant, that is $2L + \widetilde{L} = const$, thus Ω grows as N grows. In this case, the PSM is weakly scalable (in terms of iterations).
- If we keep constant the size of Ω and increase N, then the PSM convergence deteriorates.

This behavior is well understood. See the references:

- *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part I*, Ciaramella and Gander, SIAM J. Numer. Anal. (2017)
- *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part II*, Ciaramella and Gander, SIAM J. Numer. Anal. (2018)
- *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part III*, Ciaramella and Gander, ETNA (2018)
- *On the scalability of classical one-level domain-decomposition methods*, Chaouqui, Ciaramella, Gander and Vanzan, Vietnam J. Math. (2018)
- *On the scalability of the parallel Schwarz method in one-dimension*, Ciaramella, Hassan and Stamm, DD25 proceedings (2020)
- *On the scalability of the Schwarz method*, Ciaramella, Hassan and Stamm, SMAI J. Comp. Math. (2020)
- *Analysis of parallel Schwarz algorithms for time-harmonic problems using block Toeplitz matrices*, Bootland, Dolean, Kyriakis, Pestana, ETNA (2021)
- *On the effect of boundary conditions on the scalability of Schwarz methods*, Ciaramella and Mechelli, DD26 proceedings (2022)
- *On the Schwarz method for unconstrained elliptic optimal control problems*, Ciaramella and Kwok, in preparation (2022)

Restricted additive Schwarz (RAS)

Let us discretize our problem (using, e.g., FD or FE) and write the discrete problem as

$$K_{OC}\mathbf{x} = \mathbf{b},$$

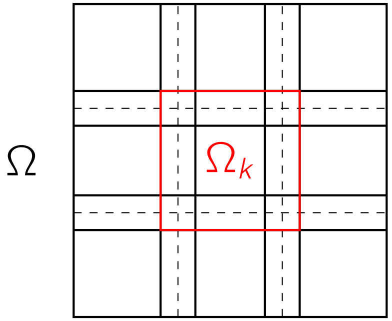
where

$$K_{OC} = \begin{bmatrix} -\frac{1}{\nu}I & A \\ A & I \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbf{p} \\ \mathbf{y} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{f} \\ \mathbf{y}_d \end{bmatrix}.$$

Let us now consider:

- a general domain decomposition,
- R_k and $\widetilde{R}_k$, the usual RAS restriction matrices (with $\widetilde{R}_k$ including a partition of unity),
- the matrices

$$\mathrm{R}_k := \begin{bmatrix} R_k & \\ & R_k \end{bmatrix}, \quad \widetilde{\mathrm{R}}_k := \begin{bmatrix} \widetilde{R}_k & \\ & \widetilde{R}_k \end{bmatrix},$$

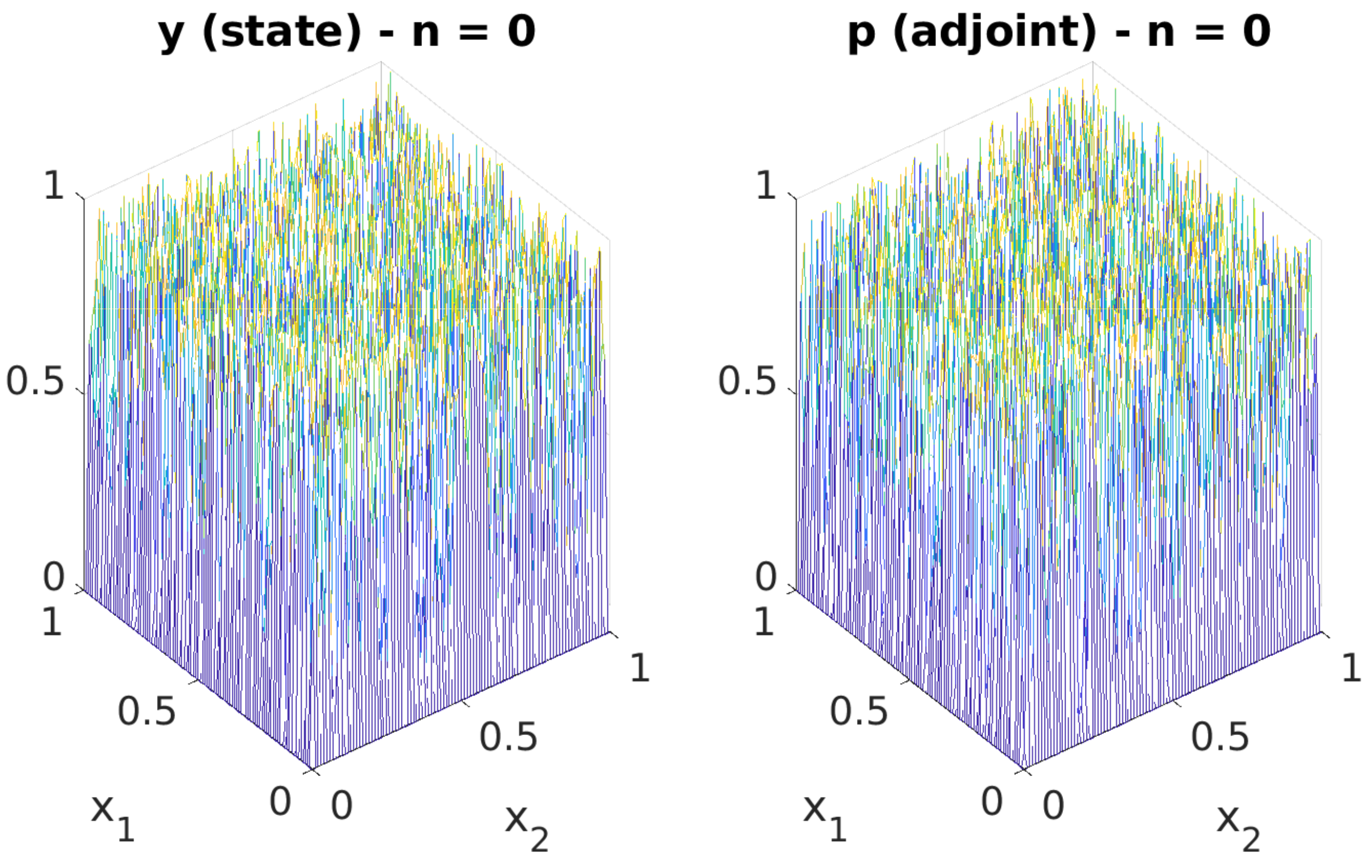


- the local subproblem matrices $K_k := \mathrm{R}_k K_{OC} \mathrm{R}_k^\top$.

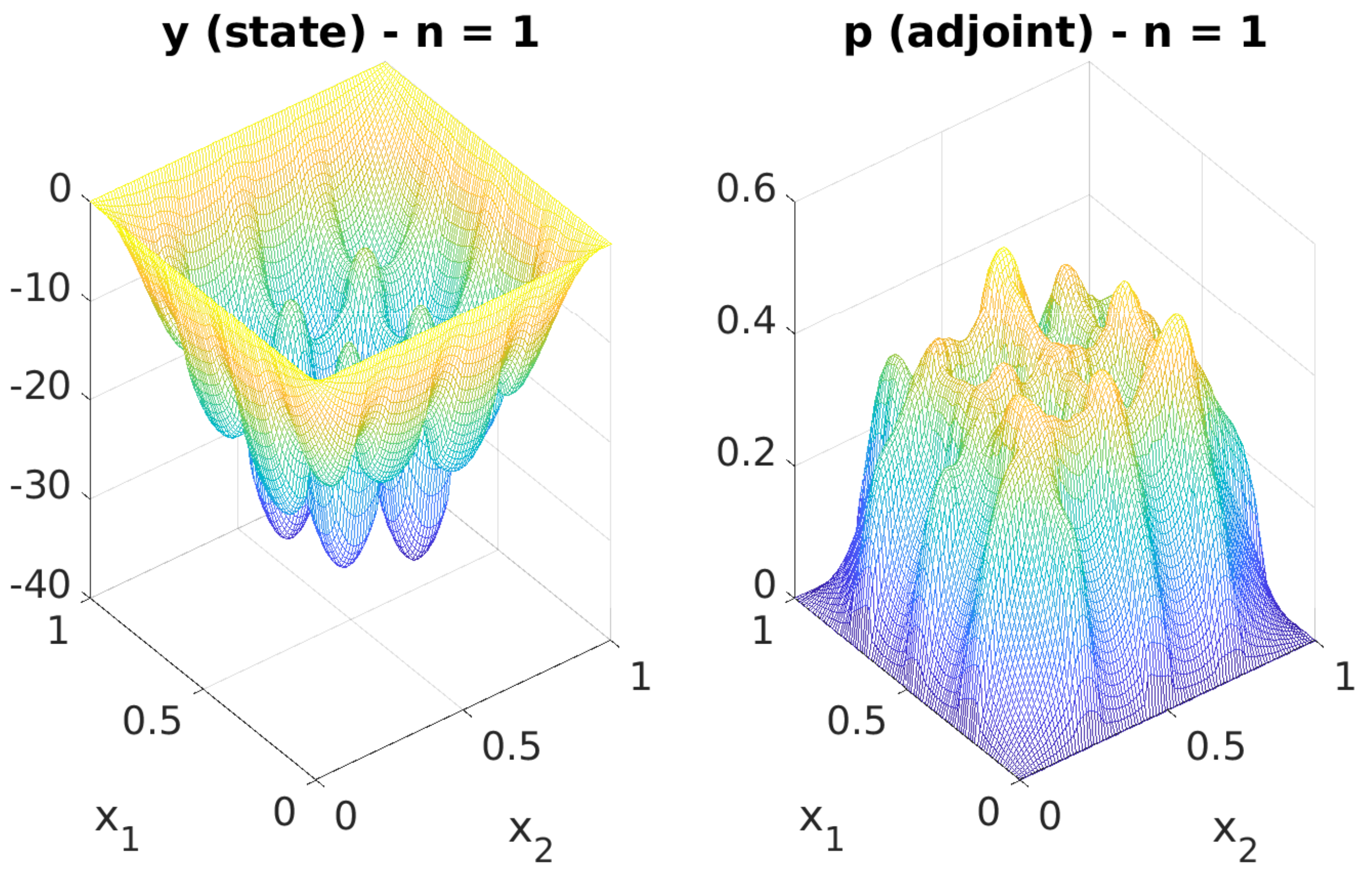
The RAS method can be easily generalized to optimal control problems:

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \sum_{k=1}^{N} \widetilde{\mathrm{R}}_k^\top K_k^{-1} \mathrm{R}_k (\mathbf{b} - K_{OC}\mathbf{x}^n).$$

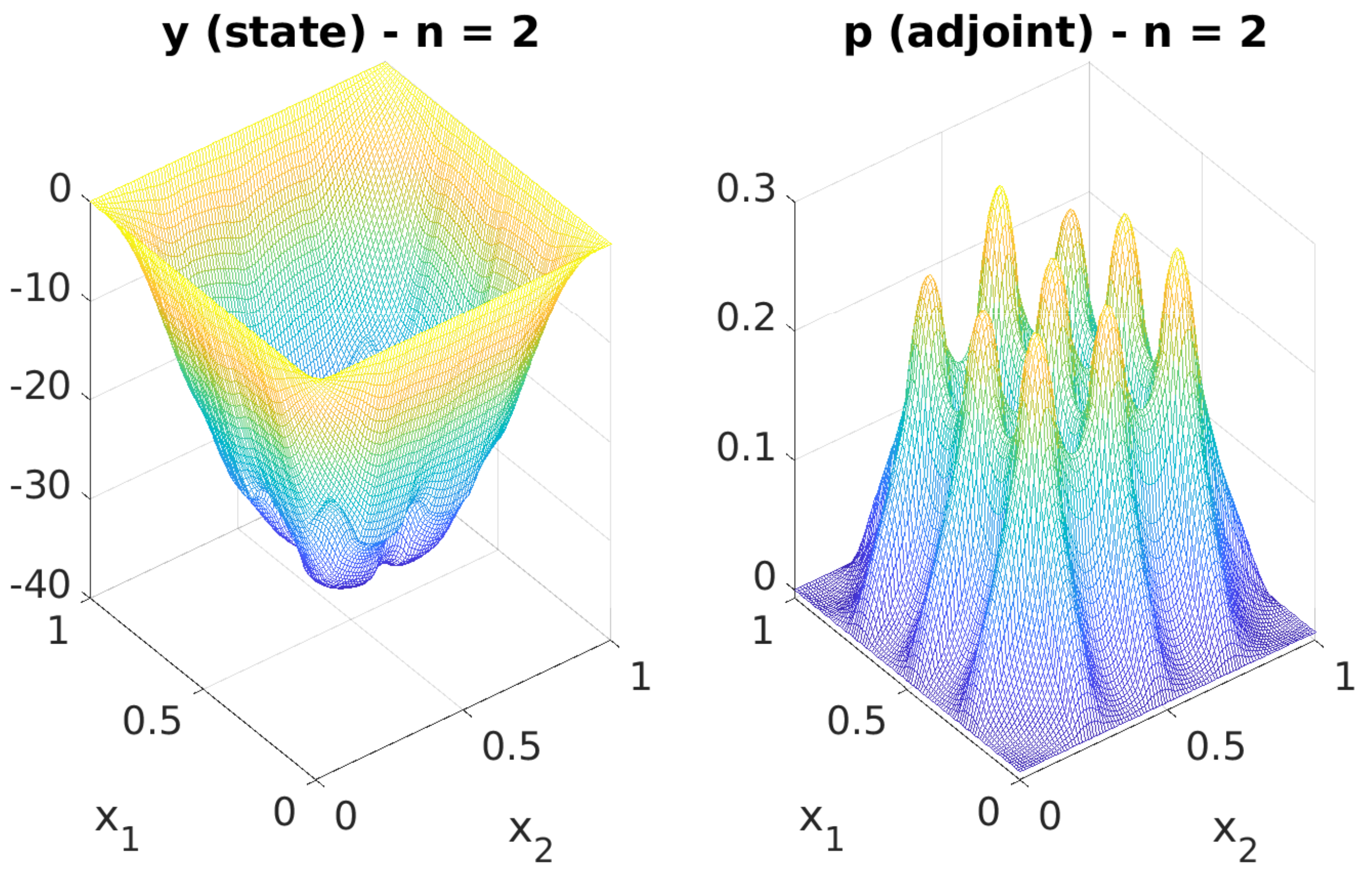
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



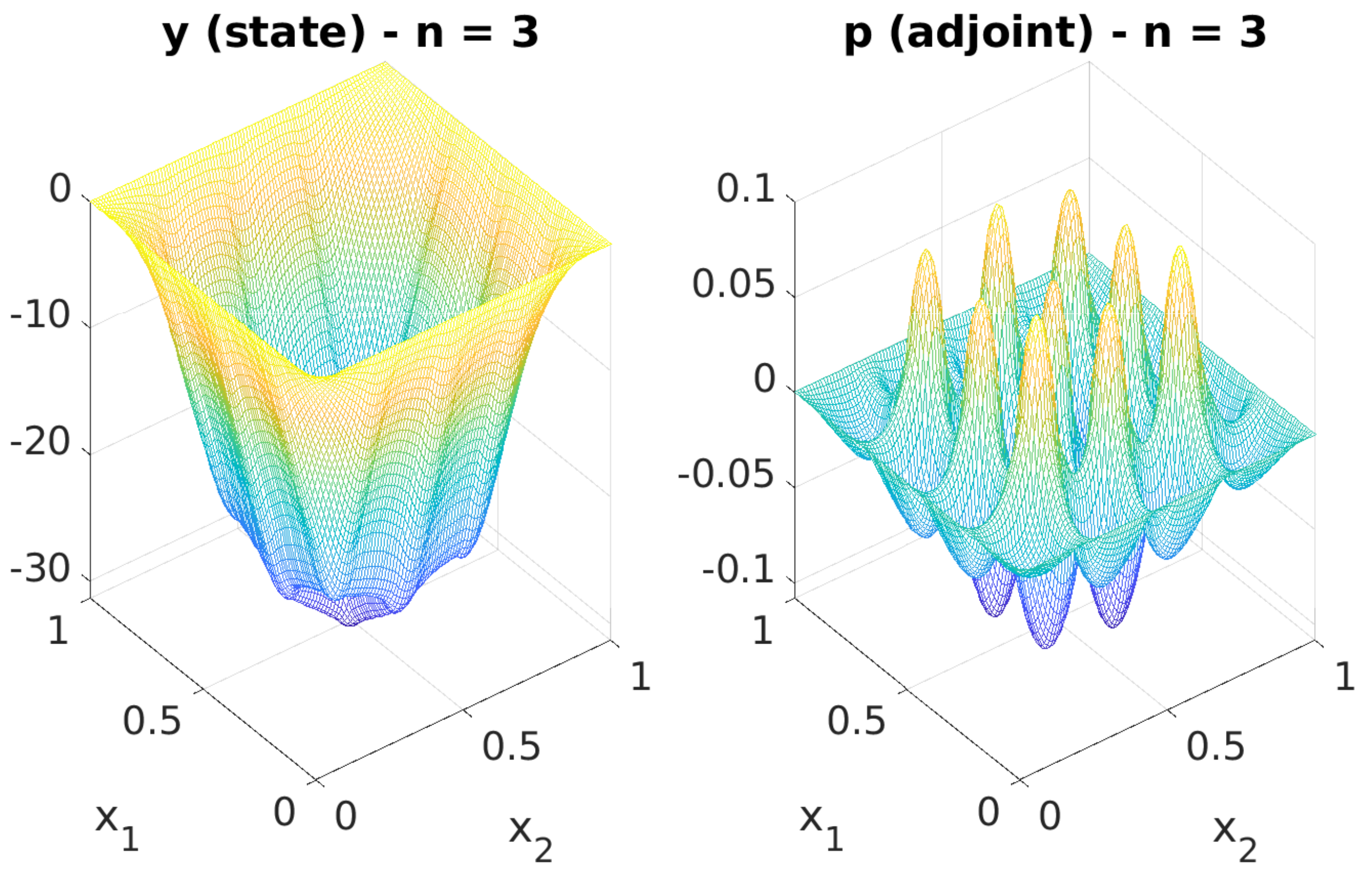
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



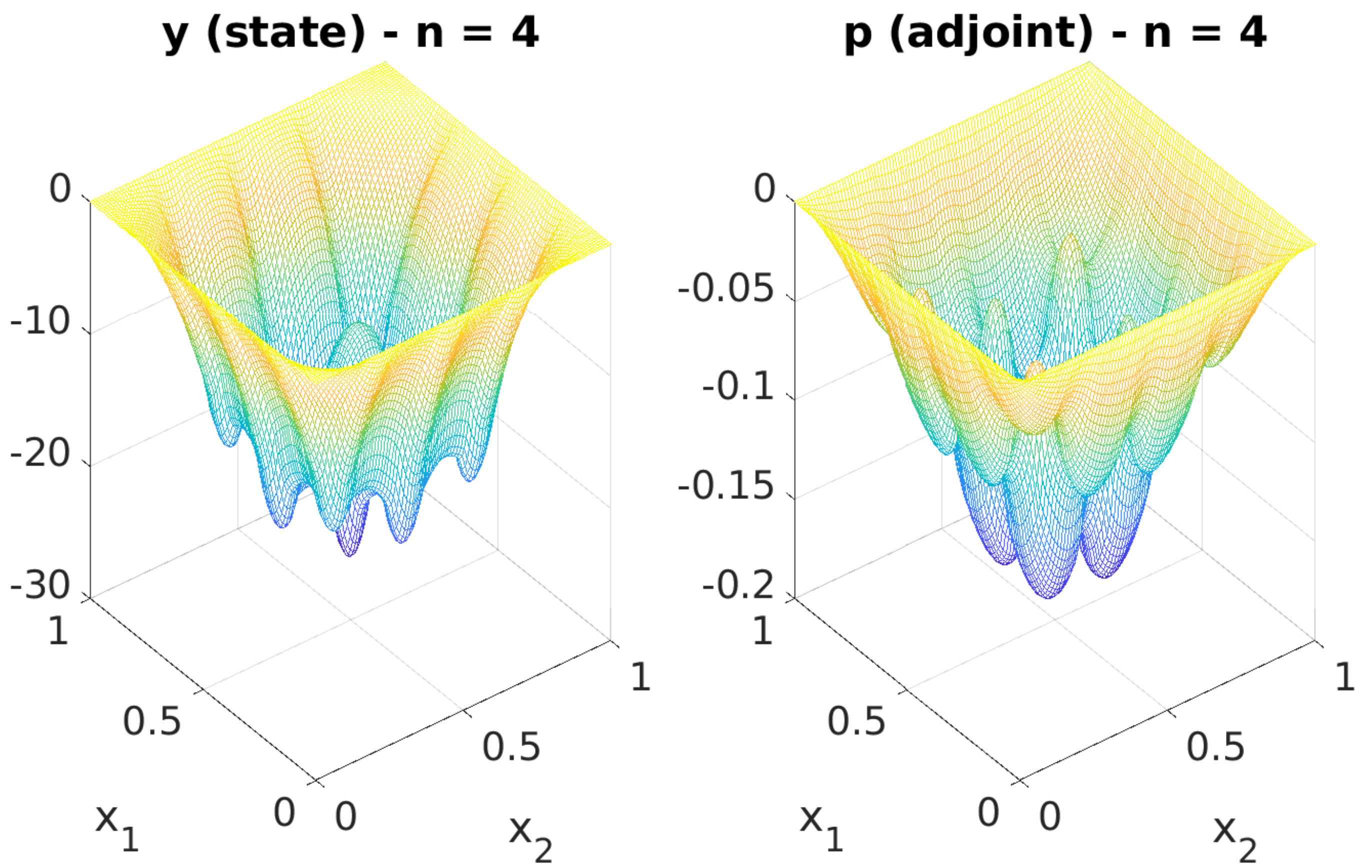
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



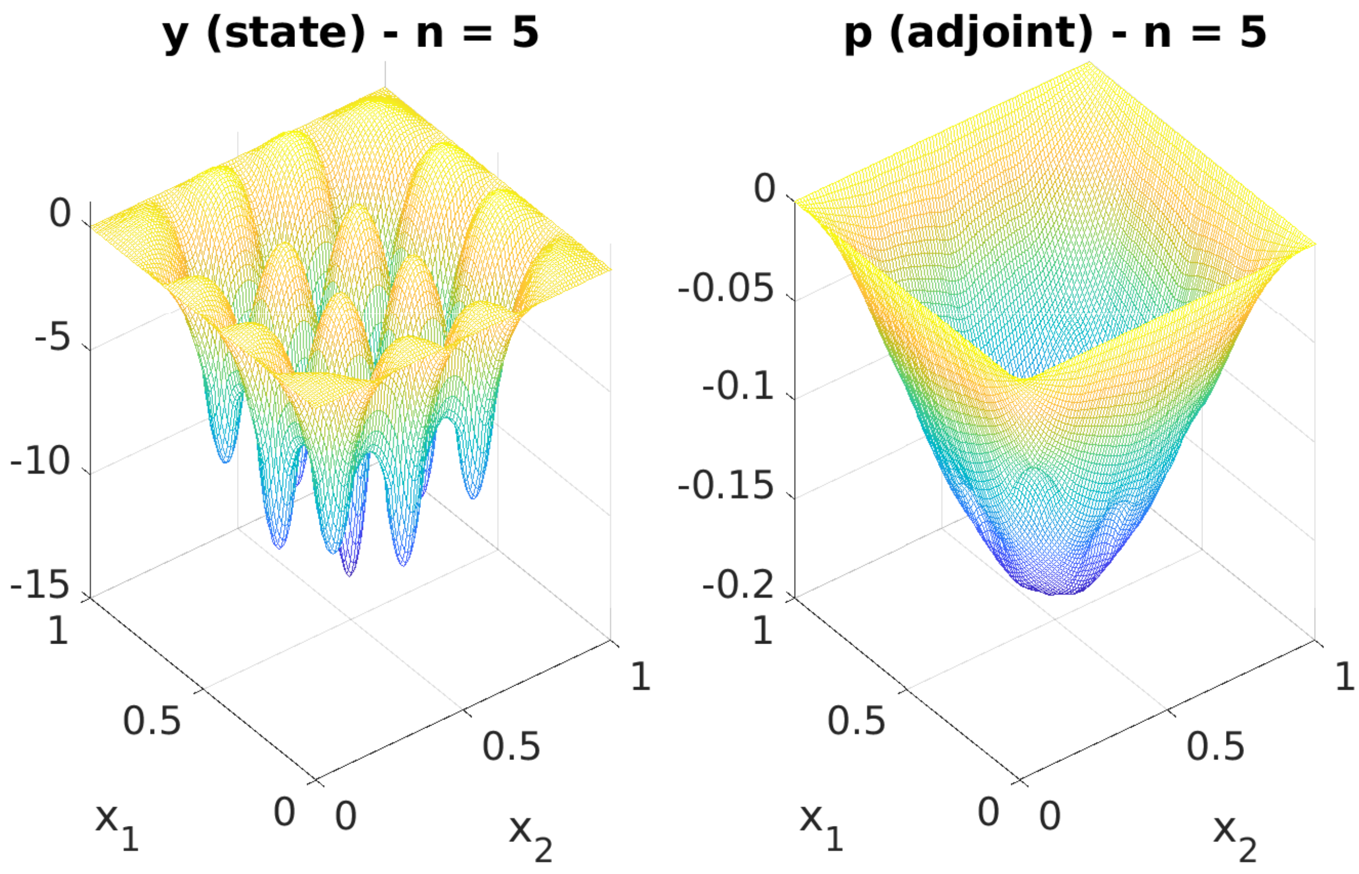
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



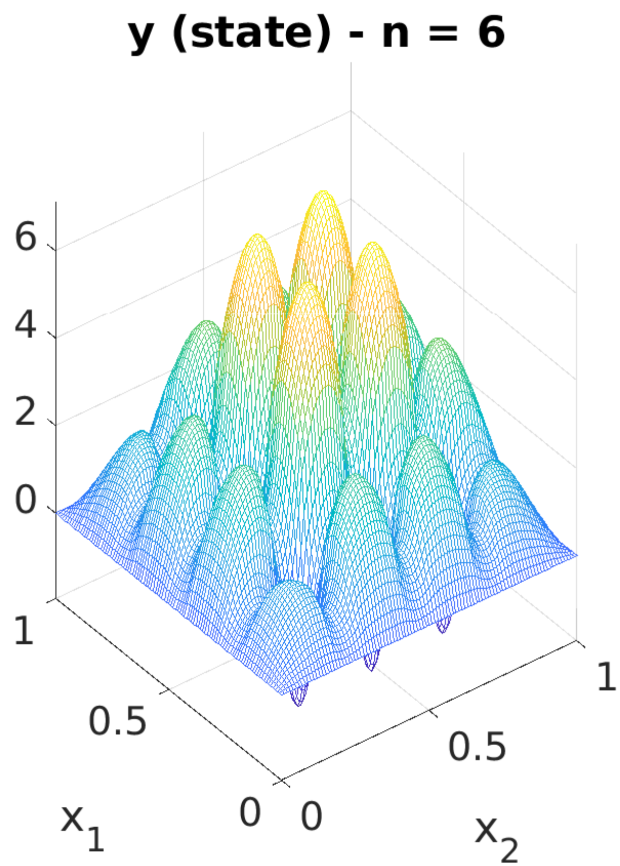
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

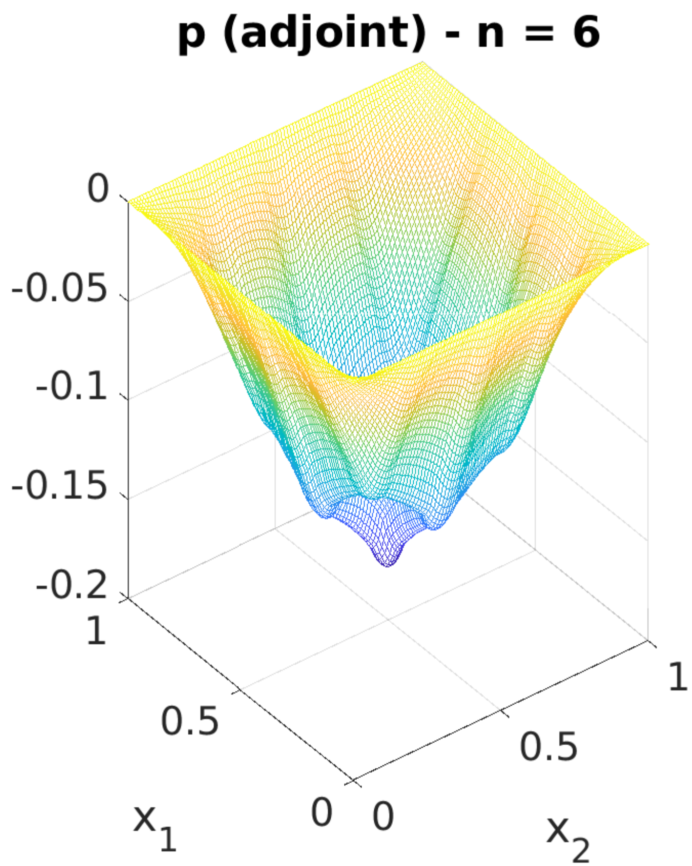


Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

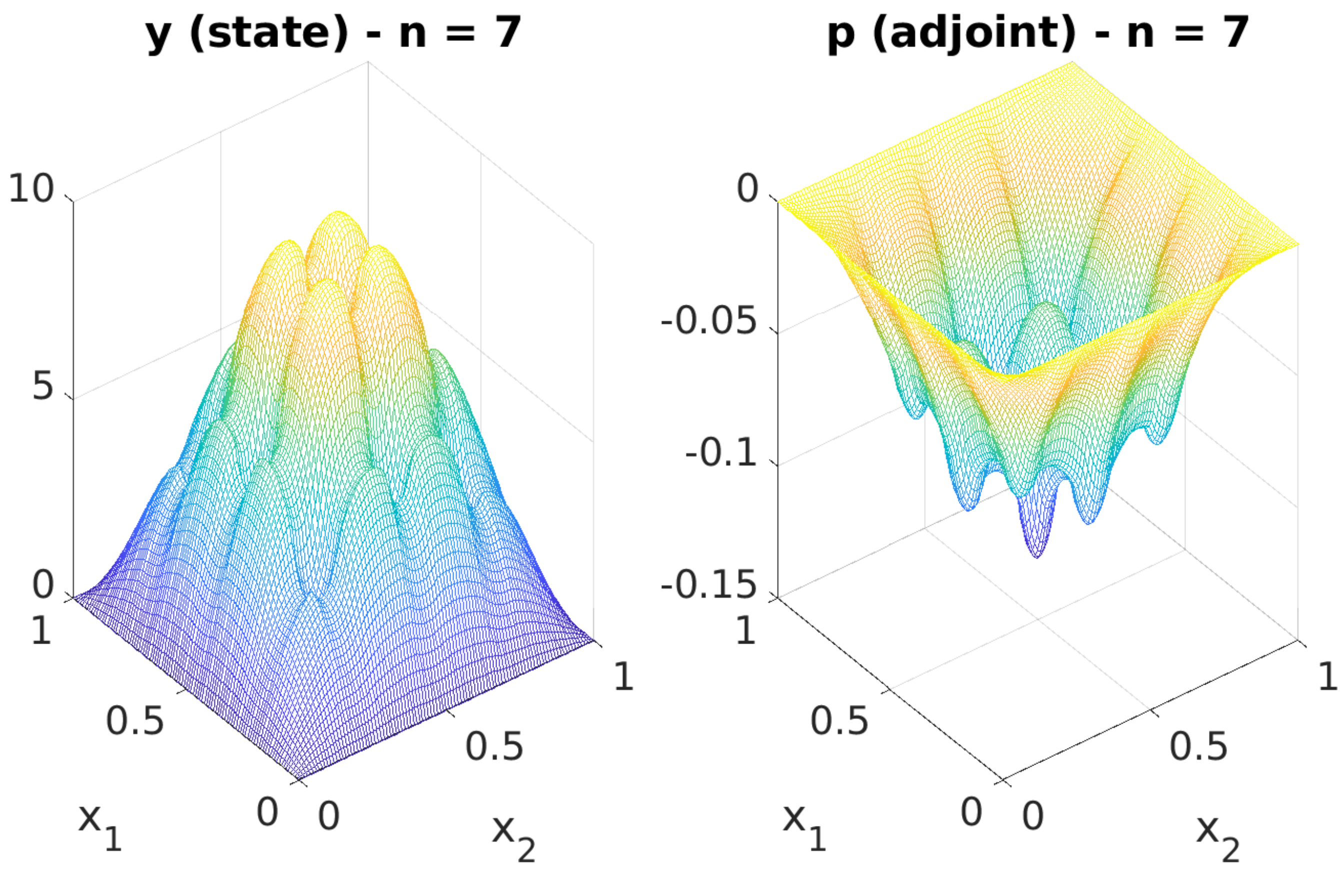


Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

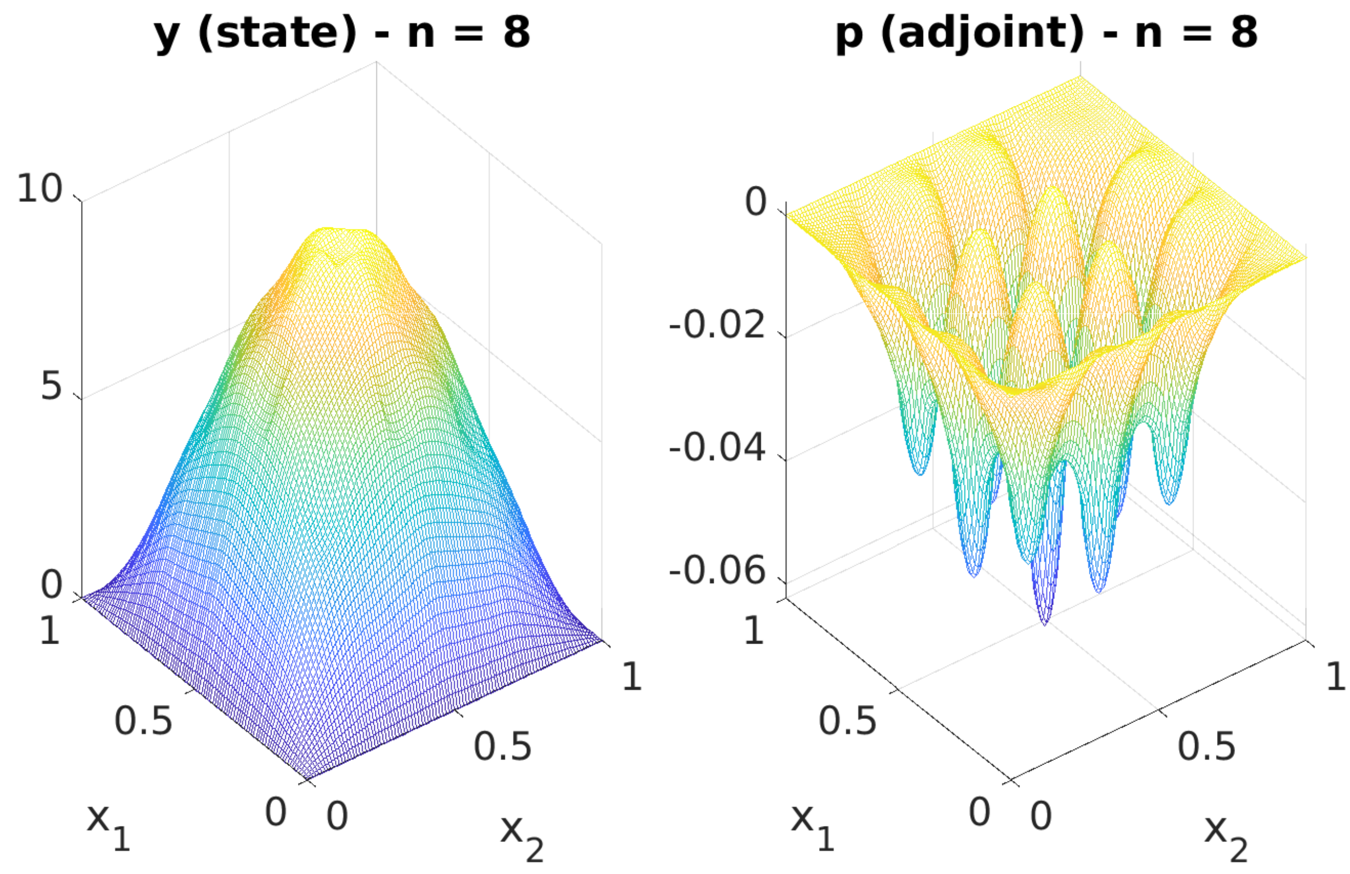




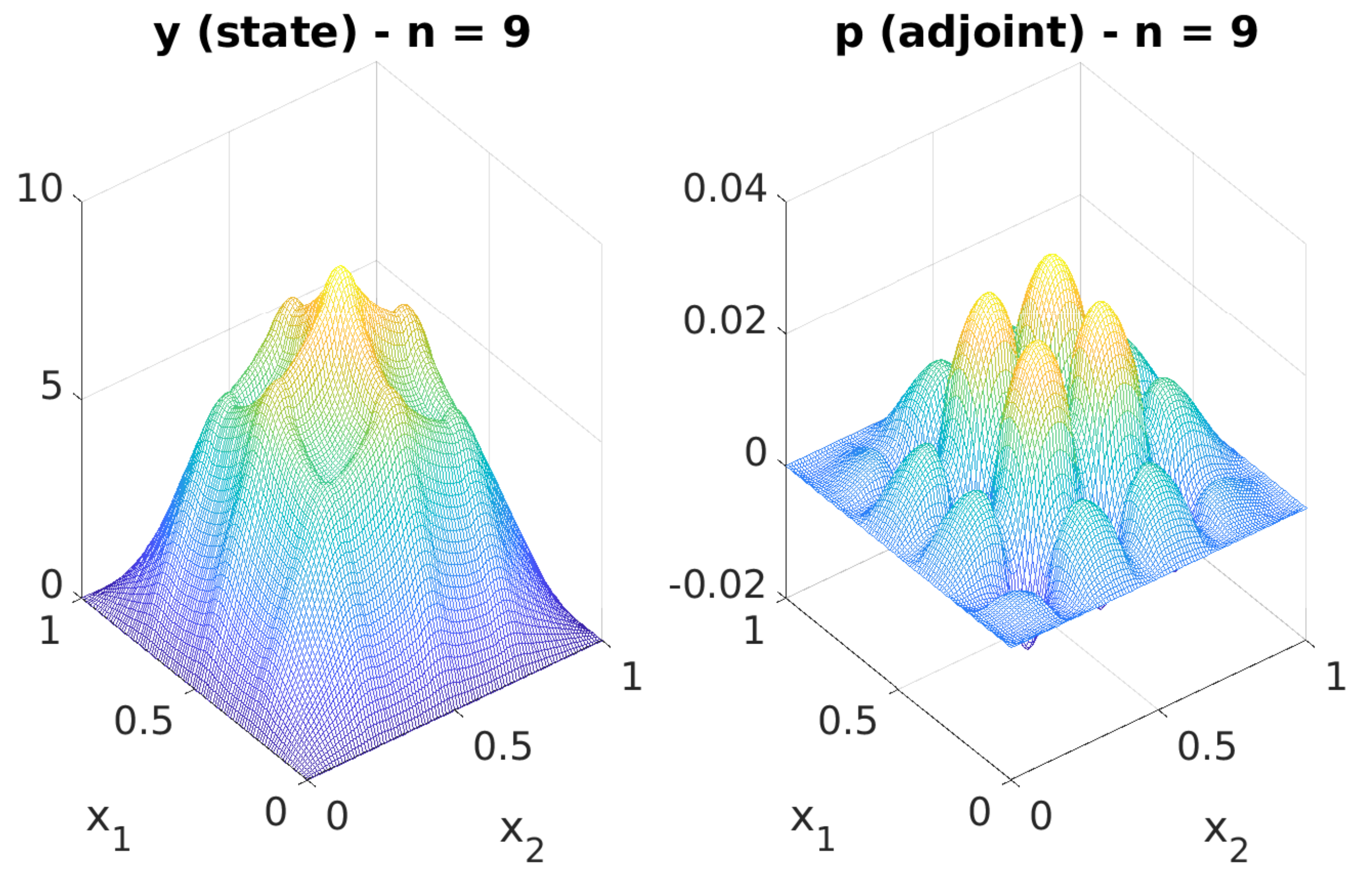
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



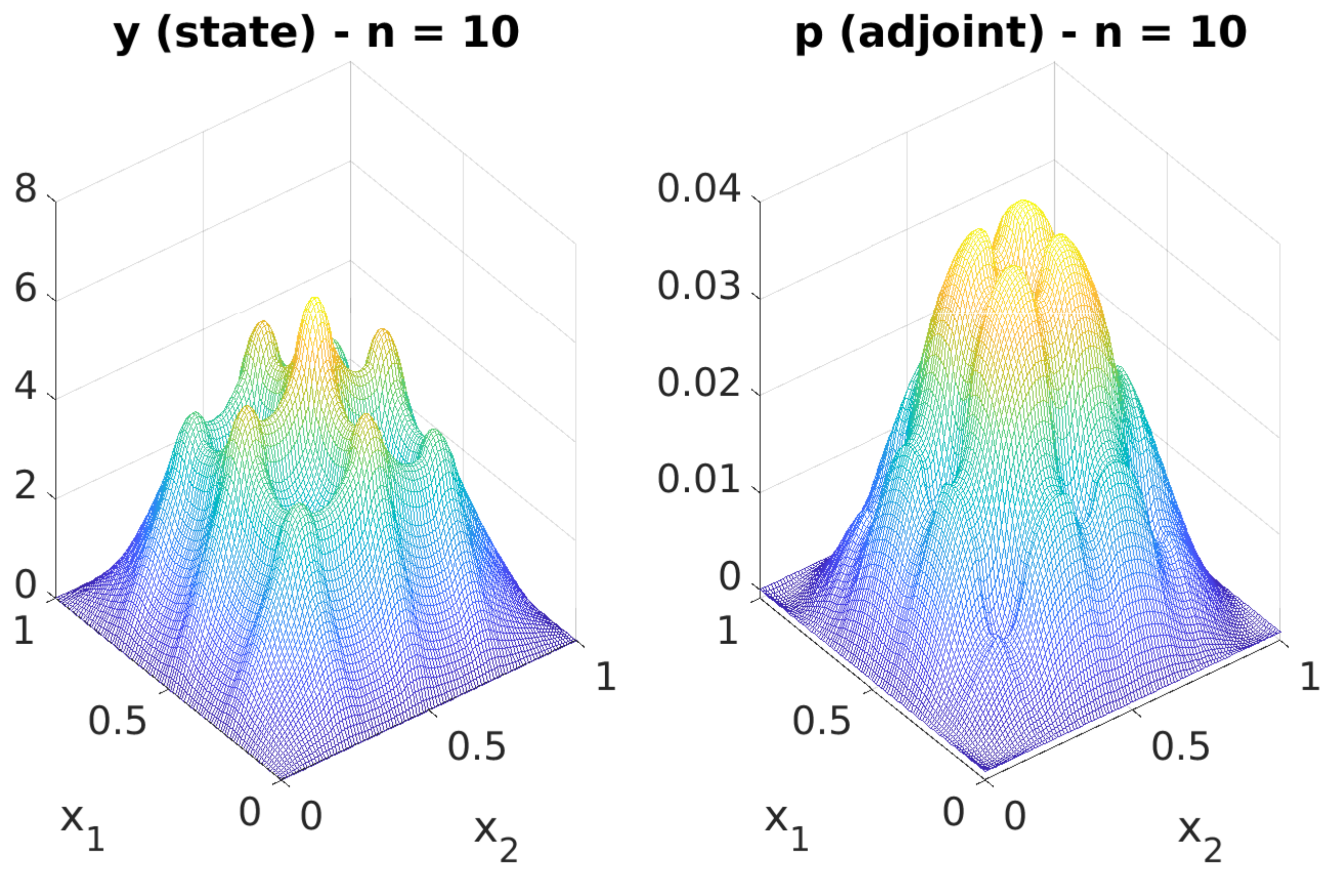
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



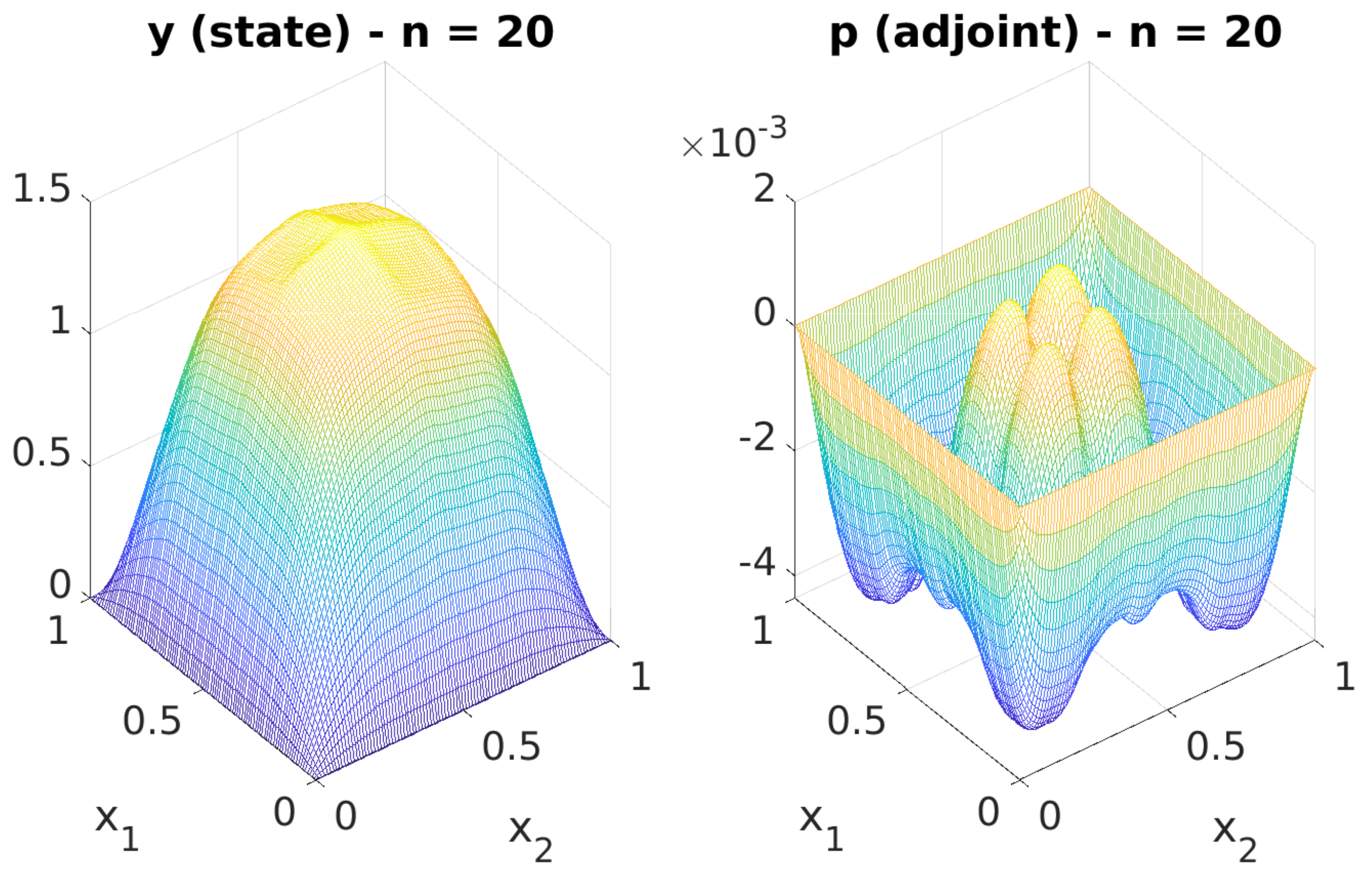
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



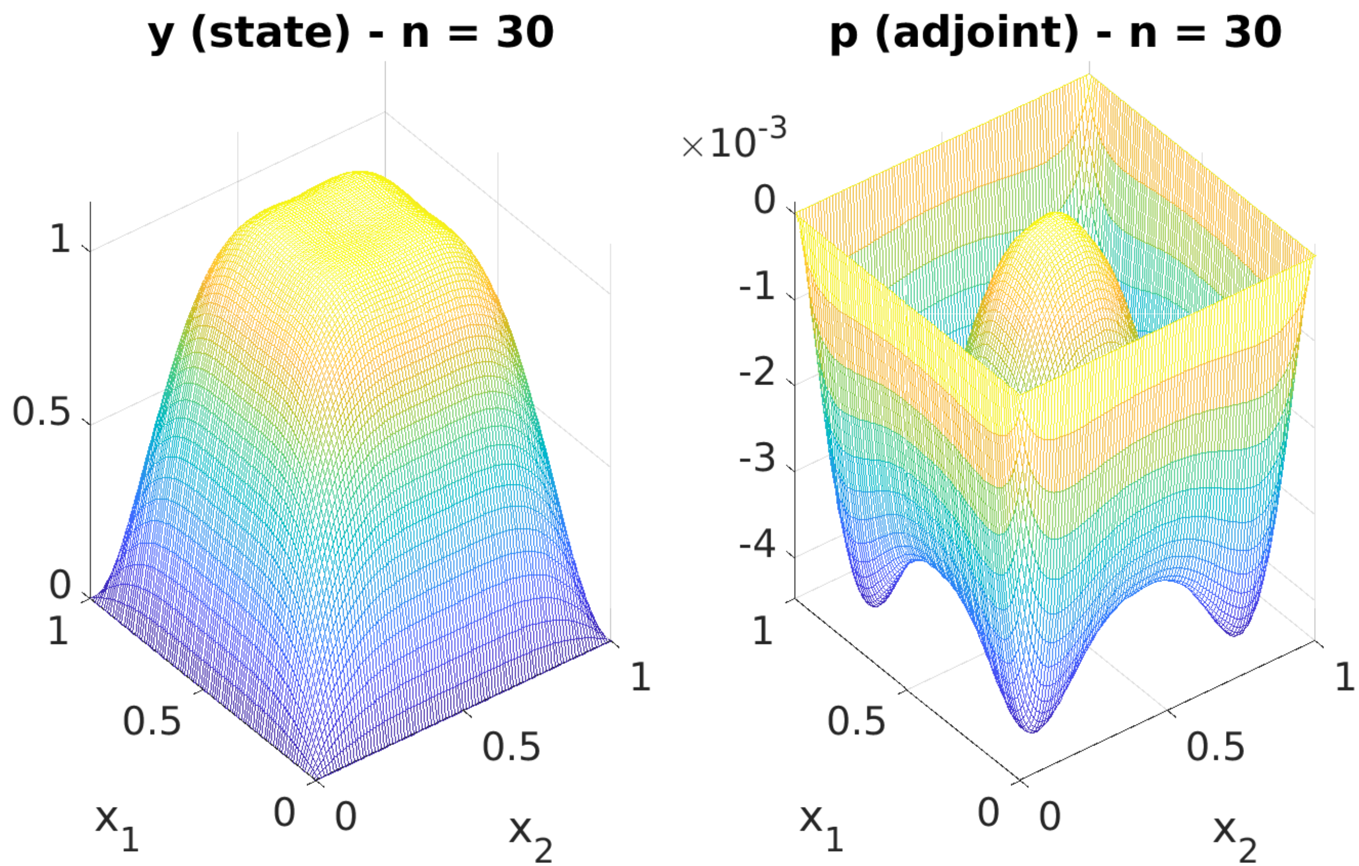
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



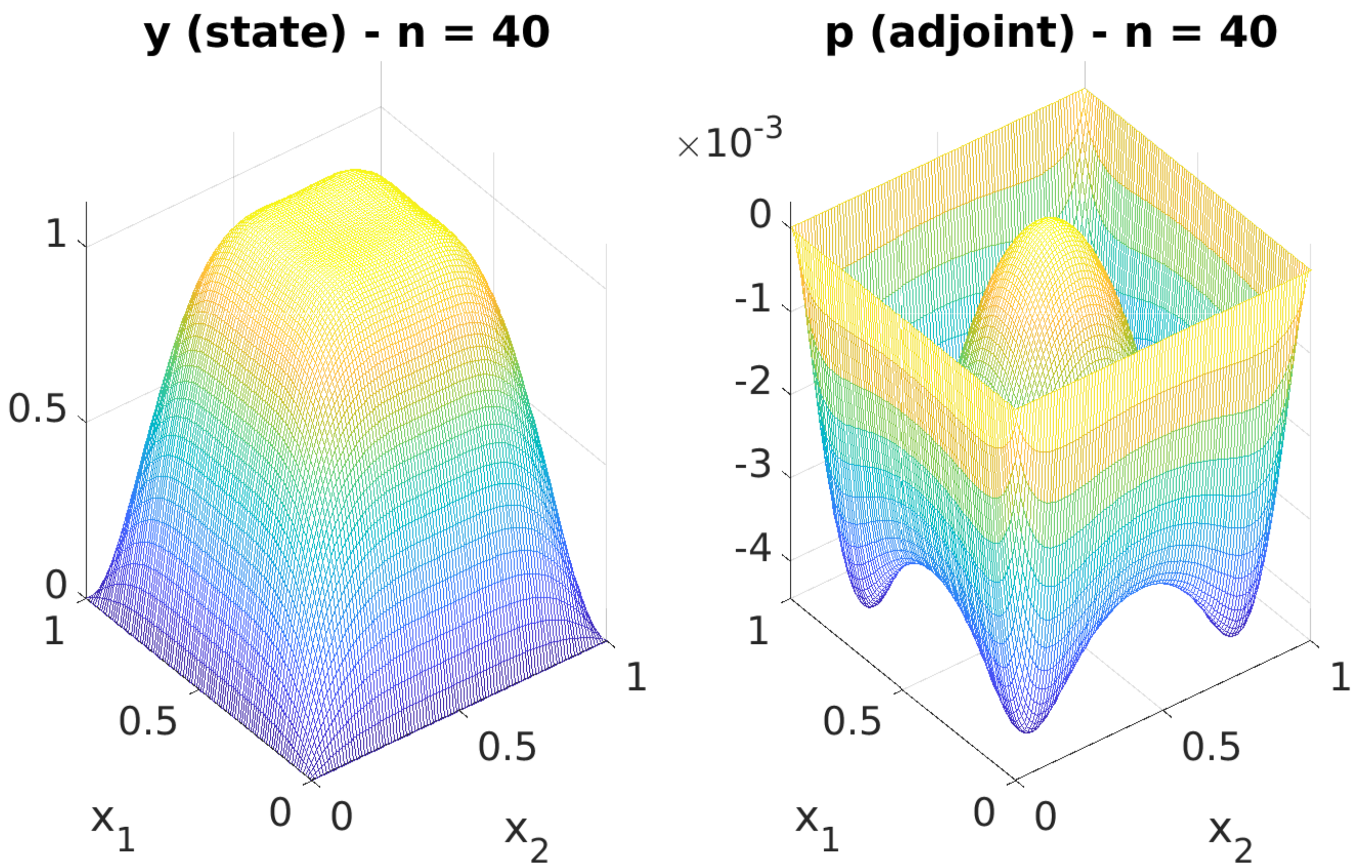
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



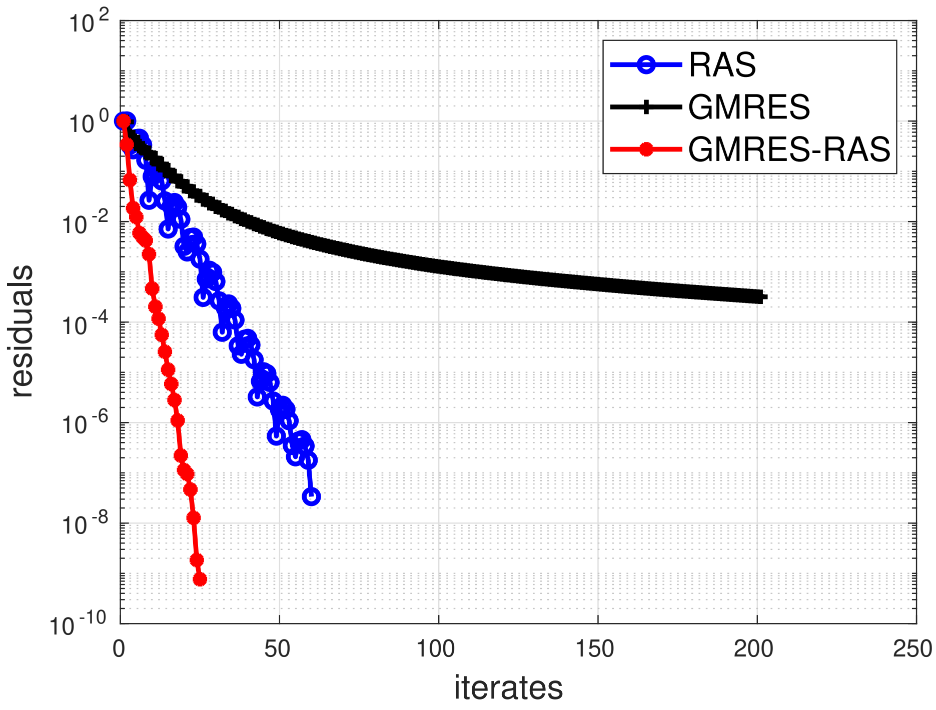
Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.

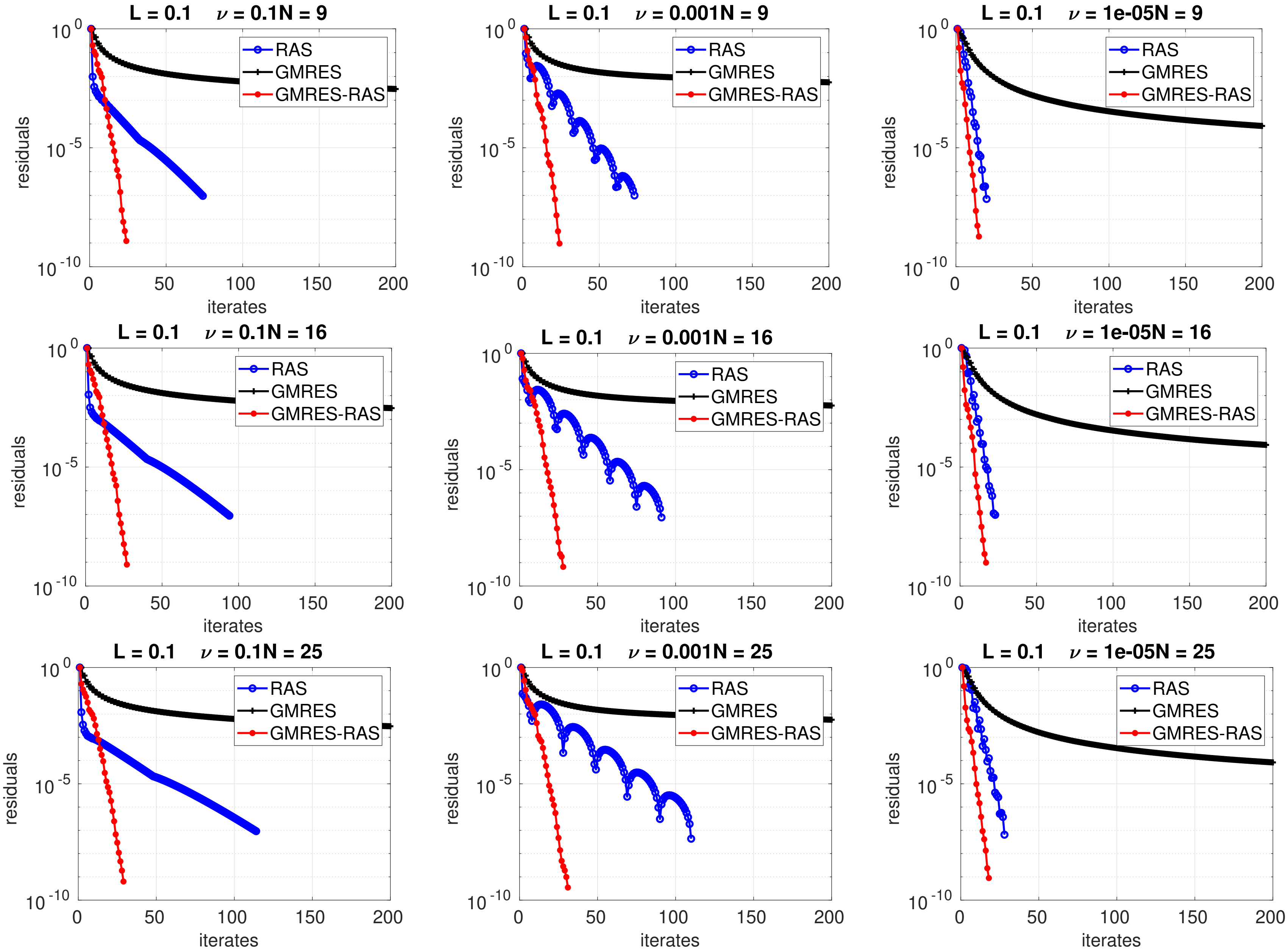


Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.



Numerical experiments: $L = 0.1$, $\nu = 10^{-4}$, $f \equiv 0$, $y_d \equiv 1$.




L = 0.1 ν = 0.1N = 9
L = 0.1 ν = 0.001N = 9
L = 0.1 ν = 1e-05N = 9
L = 0.1 ν = 0.1N = 16
L = 0.1 ν = 0.001N = 16
L = 0.1 ν = 1e-05N = 16
L = 0.1 ν = 0.1N = 25
L = 0.1 ν = 0.001N = 25
L = 0.1 ν = 1e-05N = 25
RAS
GMRES
GMRES-RAS
residuals
iterates
10^0
10^-5
10^-10
0
50
100
150
200

References

- *Iterative Methods and Preconditioners for Systems of Linear Equations*, Ciaramella and Gander, SIAM Fundamentals of Algorithms (2022)
- *On the Schwarz method for unconstrained elliptic optimal control problems*, Ciaramella and Kwok, in preparation (2022)
- *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part I*, Ciaramella and Gander, SIAM J. Numer. Anal. (2017)
- *Schwarz methods over the course of time*, Gander, ETNA (2008)
- *Domain Decomposition Methods - Algorithms and Theory*, Toselli and Widlund, Springer (2005)

Thank you!

3 Part III: Optimized Schwarz methods

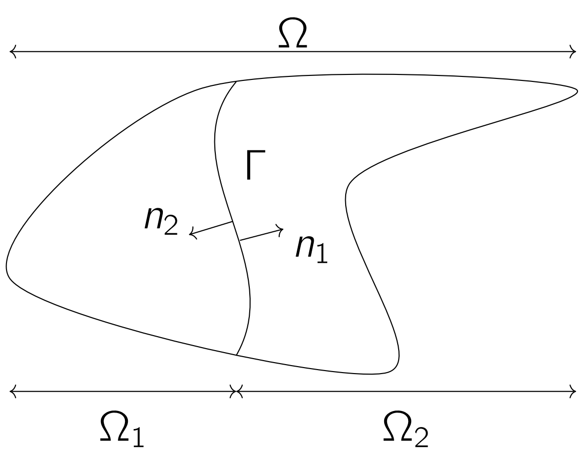
Nonoverlapping Robin Schwarz method

Let us recall our test problem and the corresponding optimality system:

$$\begin{aligned} \min_{y,u}\ & J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2} + \frac{\nu}{2}\|u\|^2_{L^2}, \\ \text{s.t.} & -\Delta y = f + u \quad \text{in } \Omega, \\ & y = 0 \quad \text{on } \partial\Omega. \end{aligned}$$

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p &= y_d - y && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ \nu u &= p && \text{in } \Omega. \end{aligned}$$

We consider a nonoverlapping decomposition of a domain Ω and the OSM:



$$\begin{aligned} -\Delta y_j^n &= f + p_j^n/\nu && \text{in } \Omega_j, \\ y_1^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ \mathcal{B}_j(y_j^n) &= \mathcal{B}_j(y_{3-j}^{n-1}) && \text{on } \Gamma, \\ -\Delta p_j^n &= y_d - y_j^n && \text{in } \Omega_j, \\ p_j^n &= 0 && \text{on } \partial\Omega \cap \partial\Omega_j, \\ \mathcal{B}_j(p_j^n) &= \mathcal{B}_j(p_{3-j}^{n-1}) && \text{on } \Gamma, \end{aligned}$$

where $\mathcal{B}_j(v) = \partial_{n_j} v + qv$, with $q > 0$.

Convergence analysis by energy estimates

The OSM in weak form is, for $j = 1, 2$,

$$\int_{\Omega_j} \nu \nabla y_j^n \cdot \nabla v - p_j^n v \, d\vec{x} - \int_{\Gamma} \nu \partial_{n_j} y_j^n v \, ds = \int_{\Omega_j} \nu f v \, d\vec{x}$$

$\forall v \in H^1(\Omega_j)$ s.t. $v = 0$ on $\partial\Omega_j \setminus \Gamma$ and $\partial_{n_j} y_j^n + q y_j^n = \partial_{n_j} y_{3-j}^{n-1} + q y_{3-j}^{n-1}$ on Γ,

$$\int_{\Omega_j} \nabla p_j^n \cdot \nabla v + y_j^n v \, d\vec{x} - \int_{\Gamma} \partial_{n_j} p_j^n v \, ds = \int_{\Omega_j} y_d v \, d\vec{x}$$

$\forall v \in H^1(\Omega_j)$ s.t. $v = 0$ on $\partial\Omega_j \setminus \Gamma$ and $\partial_{n_j} p_j^n + q p_j^n = \partial_{n_j} p_{3-j}^{n-1} + q p_{3-j}^{n-1}$ on Γ.

The errors $\mathrm{y}_j^n := y_j - y_j^n$ and $\mathrm{p}_j^n := p_j - p_j^n$ satisfy, for $j = 1, 2$,

$$\int_{\Omega_j} \nu \nabla \mathrm{y}_j^n \cdot \nabla v - \mathrm{p}_j^n v \, d\vec{x} = \int_{\Gamma} \nu \partial_{n_j} \mathrm{y}_j^n v \, ds$$

$\forall v \in H^1(\Omega_j)$ s.t. $v = 0$ on $\partial\Omega_j \setminus \Gamma$ and $\partial_{n_j} \mathrm{y}_j^n + q \mathrm{y}_j^n = \partial_{n_j} \mathrm{y}_{3-j}^{n-1} + q \mathrm{y}_{3-j}^{n-1}$ on Γ,

$$\int_{\Omega_j} \nabla \mathrm{p}_j^n \cdot \nabla v + \mathrm{y}_j^n v \, d\vec{x} = \int_{\Gamma} \partial_{n_j} \mathrm{p}_j^n v \, ds$$

$\forall v \in H^1(\Omega_j)$ s.t. $v = 0$ on $\partial\Omega_j \setminus \Gamma$ and $\partial_{n_j} \mathrm{p}_j^n + q \mathrm{p}_j^n = \partial_{n_j} \mathrm{p}_{3-j}^{n-1} + q \mathrm{p}_{3-j}^{n-1}$ on Γ.

Convergence analysis by energy estimates

Let us define the energy

$$E_n := \sum_{j=1}^{2} \nu \|\partial_{n_j} \mathrm{y}_j^n\|_{L^2(\Gamma)}^2 + \nu q^2 \|\mathrm{y}_j^n\|_{L^2(\Gamma)}^2 + \|\partial_{n_j} \mathrm{p}_j^n\|_{L^2(\Gamma)}^2 + q^2 \|\mathrm{p}_j^n\|_{L^2(\Gamma)}^2.$$

We can manipulate E_n and use the OSM transmission conditions (and that $n_1 = -n_2$) to get

$$\begin{aligned}
E_n &= \sum_{j=1}^{2} \Big[\nu \|\partial_{n_j} \mathrm{y}_j^n + q \mathrm{y}_j^n\|_{L^2(\Gamma)}^2 - 2q\nu \langle \partial_{n_j} \mathrm{y}_j^n, \mathrm{y}_j^n \rangle_{L^2(\Gamma)} + \|\partial_{n_j} \mathrm{p}_j^n + q \mathrm{p}_j^n\|_{L^2(\Gamma)}^2 - 2q \langle \partial_{n_j} \mathrm{p}_j^n, \mathrm{p}_j^n \rangle_{L^2(\Gamma)} \Big] \\
&= \sum_{j=1}^{2} \Big[\nu \|\partial_{n_j} \mathrm{y}_{3-j}^{n-1} + q \mathrm{y}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 - 2q\nu \langle \partial_{n_j} \mathrm{y}_j^n, \mathrm{y}_j^n \rangle_{L^2(\Gamma)} + \|\partial_{n_j} \mathrm{p}_{3-j}^{n-1} + q \mathrm{p}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 - 2q \langle \partial_{n_j} \mathrm{p}_j^n, \mathrm{p}_j^n \rangle_{L^2(\Gamma)} \Big] \\
&= \sum_{j=1}^{2} \Big[\nu \|\partial_{n_j} \mathrm{y}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 + \nu q^2 \|\mathrm{y}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 + \|\partial_{n_j} \mathrm{p}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 + q^2 \|\mathrm{p}_{3-j}^{n-1}\|_{L^2(\Gamma)}^2 \Big] \\
&\qquad - 2q \sum_{j=1}^{2} \sum_{m=n-1}^{n} \Big[\nu \langle \partial_{n_j} \mathrm{y}_j^m, \mathrm{y}_j^m \rangle_{L^2(\Gamma)} + \langle \partial_{n_j} \mathrm{p}_j^m, \mathrm{p}_j^m \rangle_{L^2(\Gamma)} \Big] \\
&= E_{n-1} - 2q \sum_{j=1}^{2} \sum_{m=n-1}^{n} \Big[\nu \langle \partial_{n_j} \mathrm{y}_j^m, \mathrm{y}_j^m \rangle_{L^2(\Gamma)} + \langle \partial_{n_j} \mathrm{p}_j^m, \mathrm{p}_j^m \rangle_{L^2(\Gamma)} \Big],
\end{aligned}$$

$$E_n = E_{n-1} - 2q \sum_{j=1}^{2} \sum_{m=n-1}^{n} \Big[\nu \langle \partial_{n_j} \mathrm{y}_j^m, \mathrm{y}_j^m \rangle_{L^2(\Gamma)} + \langle \partial_{n_j} \mathrm{p}_j^m, \mathrm{p}_j^m \rangle_{L^2(\Gamma)} \Big].$$

Now, we use the weak forms of the OSM equations,

$$\int_{\Omega_j} \nu \nabla \mathrm{y}_j^n \cdot \nabla v - \mathrm{p}_j^n v \, d\vec{x} = \int_{\Gamma} \nu \partial_{n_j} \mathrm{y}_j^n v \, ds \;\; \text{and} \;\; \int_{\Omega_j} \nabla \mathrm{p}_j^n \cdot \nabla v + \mathrm{y}_j^n v \, d\vec{x} = \int_{\Gamma} \partial_{n_j} \mathrm{p}_j^n v \, ds,$$

and test them with $v = \mathrm{y}_j^n$ and $v = \mathrm{p}_j^n$, respectively, to obtain

$$\nu \langle \partial_{n_j} \mathrm{y}_j^n, \mathrm{y}_j^n \rangle_{L^2(\Gamma)} = \nu \| \nabla \mathrm{y}_j^n \|_{L^2(\Omega_j)}^2 - \langle \mathrm{p}_j^n, \mathrm{y}_j^n \rangle_{L^2(\Omega_j)},$$

$$\langle \partial_{n_j} \mathrm{p}_j^n, \mathrm{p}_j^n \rangle_{L^2(\Gamma)} = \| \nabla \mathrm{p}_j^n \|_{L^2(\Omega_j)}^2 + \langle \mathrm{y}_j^n, \mathrm{p}_j^n \rangle_{L^2(\Omega_j)}.$$

We insert these into the above energy relation and get

$$E_n = E_{n-1} - 2q \sum_{j=1}^{2} \sum_{m=n-1}^{n} \Big[\nu \| \nabla \mathrm{y}_j^m \|_{L^2(\Omega_j)}^2 + \| \nabla \mathrm{p}_j^m \|_{L^2(\Omega_j)}^2 \Big].$$

Since the second term on the right-hand side is non-negative, we get

$$0 \leq E_n \leq E_{n-1} \;\; \text{for any } n \in \mathbb{N}$$

and hence $E_n \to \ell$ as $n \to \infty$, for some real value $0 \leq \ell < \infty$.

Now, we use again $E_n = E_{n-1} - 2q \sum_{j=1}^2 \sum_{m=n-1}^n \Big[\nu\|\nabla \mathrm{y}_j^m\|_{L^2(\Omega_j)}^2 + \|\nabla \mathrm{p}_j^m\|_{L^2(\Omega_j)}^2\Big]$ to estimate

$$\sum_{j=1}^2 \Big[\nu\|\nabla \mathrm{y}_j^n\|_{L^2(\Omega_j)}^2 + \|\nabla \mathrm{p}_j^n\|_{L^2(\Omega_j)}^2\Big] \leq \sum_{j=1}^2 \sum_{m=n-1}^n \Big[\nu\|\nabla \mathrm{y}_j^m\|_{L^2(\Omega_j)}^2 + \|\nabla \mathrm{p}_j^m\|_{L^2(\Omega_j)}^2\Big] = \frac{1}{2q}(E_{n-1} - E_n).$$

Summing over n both sides of this inequality, we obtain

$$\sum_{n=1}^\infty \sum_{j=1}^2 \Big[\nu\|\nabla \mathrm{y}_j^n\|_{L^2(\Omega_j)}^2 + \|\nabla \mathrm{p}_j^n\|_{L^2(\Omega_j)}^2\Big] \leq \frac{1}{2q}\sum_{n=1}^\infty (E_{n-1} - E_n) = \frac{1}{2q}(E_0 - \ell).$$

Since this series converges the corresponding series converges to zero:

$$\|\nabla \mathrm{y}_j^n\|_{L^2(\Omega_j)} \to 0 \text{ and } \|\nabla \mathrm{p}_j^n\|_{L^2(\Omega_j)} \to 0 \text{ for } j = 1, 2.$$

Using the Poincaré-Friedrichs inequality we get

$$\|\mathrm{y}_j^n\|_{L^2(\Omega_j)} \leq C_y \|\nabla \mathrm{y}_j^n\|_{L^2(\Omega_j)} \text{ and } \|\mathrm{p}_j^n\|_{L^2(\Omega_j)} \leq C_p \|\nabla \mathrm{p}_j^n\|_{L^2(\Omega_j)},$$

for some constants $C_y > 0$ and $C_p > 0$. Hence, we obtain that

$$\|\mathrm{y}_j^n\|_{L^2(\Omega_j)} \to 0 \text{ and } \|\mathrm{p}_j^n\|_{L^2(\Omega_j)} \to 0 \text{ for } j = 1, 2,$$

which then implies

$$\|\mathrm{y}_j^n\|_{H^1(\Omega_j)} \to 0 \text{ and } \|\mathrm{p}_j^n\|_{H^1(\Omega_j)} \to 0 \text{ for } j = 1, 2.$$

Thus, we proved convergence of the OSM in H^1.

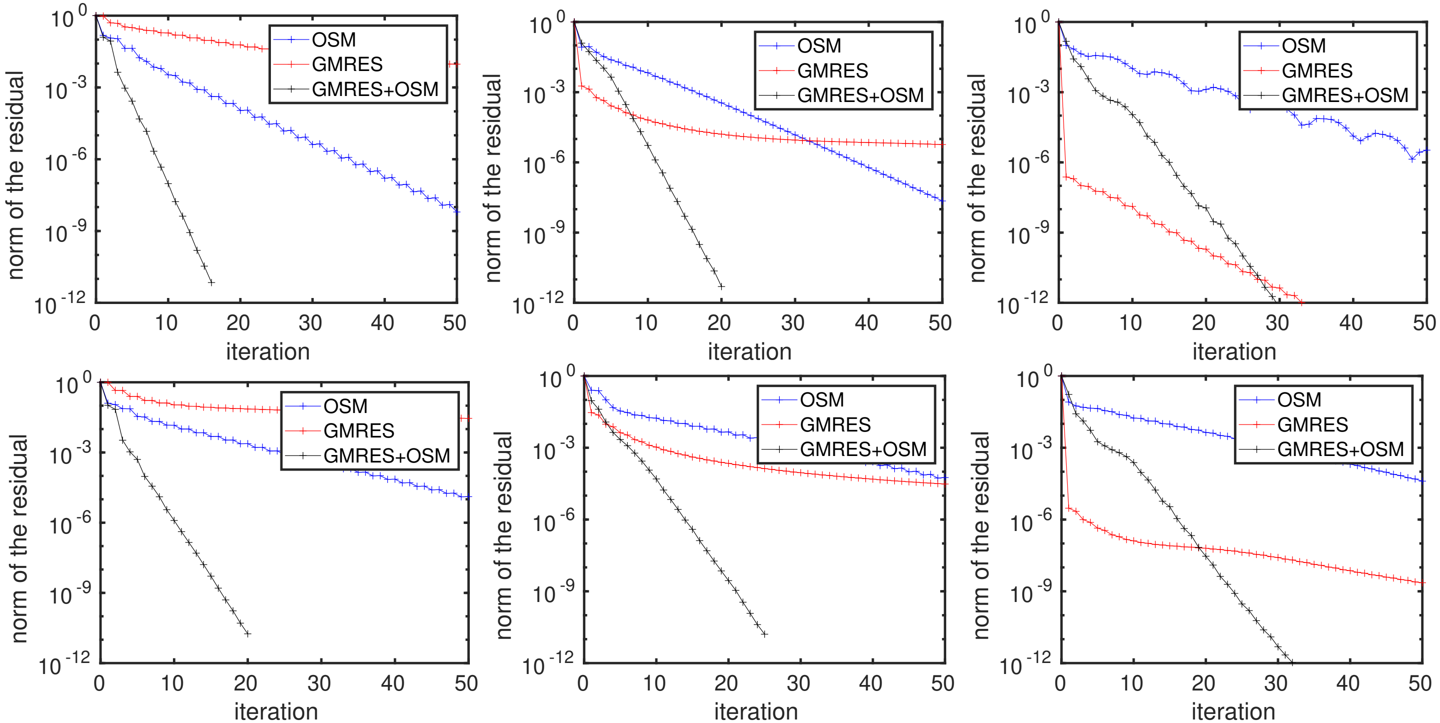
Remarks:

- Convergence is guaranteed also without overlap.
- This analysis was originally proposed by Lions and then used by Benamou for optimal control problems.
- This analysis does not provide information about the speed of convergence of the method.
- A Fourier analysis is also possible (for specific geometries) and
  - it provides information about the convergence of Fourier modes (useful also at the discrete level);
  - Parseval's identity together with the dominated convergence (Lebesgue) theorem need to be used to get convergence in L^2.
- In general, OSM with overlap converges faster than PSM.
- OSM can be written in a RAS form, known as Optimized Restricted Additive Schwarz (ORAS):

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \sum_{k=1}^{N} \widetilde{\mathrm{R}}_k^\top (K_k^{ORAS})^{-1} \mathrm{R}_k (\mathbf{b} - K_{OC}\mathbf{x}^n),$$

where K_k^{ORAS} are modifications of $K_k = \mathrm{R} K_{OC} \mathrm{R}^\top$ to include the Robin transmission conditions.

Numerical experiments



$m = 31$

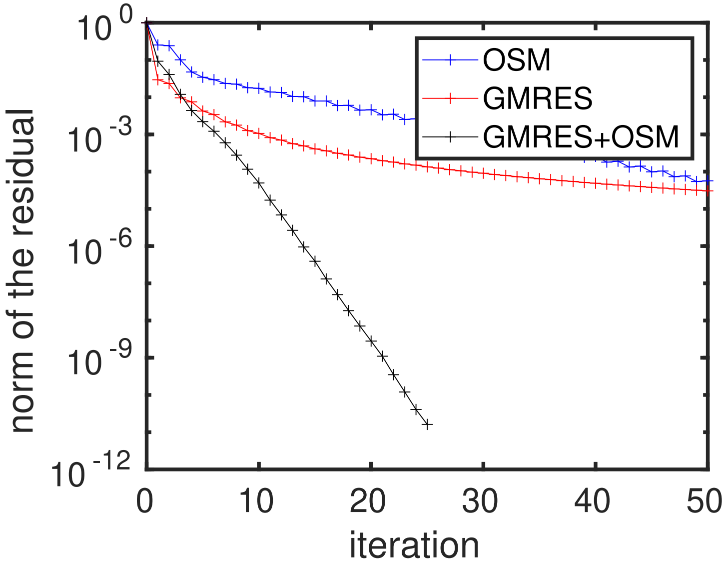
$m = 63$

$\nu = 1e-3$ $\nu = 1e-5$ $\nu = 1e-7$

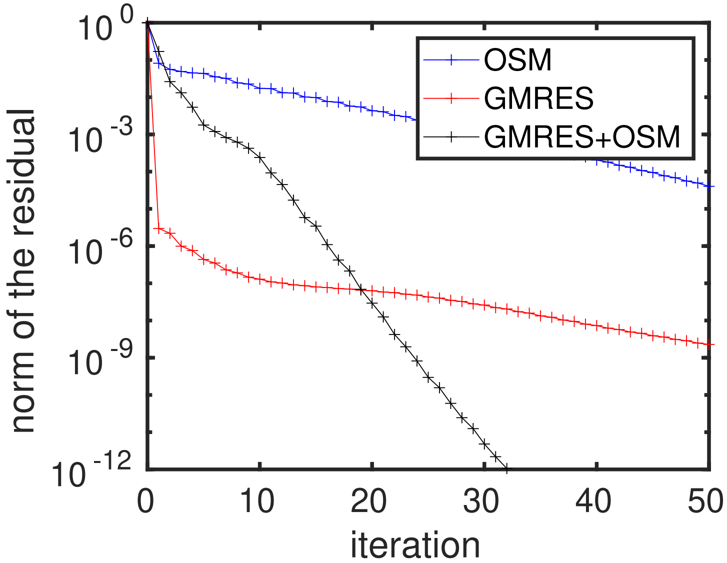
A surprising behavior ...

$$\min_{y,u} J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2} + \frac{\nu}{2}\|u\|^2_{L^2},$$
$$\text{s.t.} -\Delta y = f + u \quad \text{in } \Omega,$$
$$y = 0 \quad \text{on } \partial\Omega.$$

The parameter ν is a regularization parameter: it makes the problem "more convex" and in general "better conditioned". Thus, we expect that (linear) solvers behave better for ν large, but we have that



$\nu = 1e - 5$



$\nu = 1e - 7$

GMRES seems to converge faster when the regularization parameter ν becomes smaller!

A surprising behavior ... and its explanation (C.,Gander 2022)

Consider the matrix $K_{OC} := \begin{bmatrix} -\frac{1}{\nu}I & A \\ A & I \end{bmatrix}$ with $\nu > 0$. The eigenvalues of K_{OC} are

$$\lambda_{j,\pm}(K_{OC}) = \frac{\nu-1}{2\nu} \pm \sqrt{\left(\frac{\nu-1}{2\nu}\right)^2 + \frac{1}{\nu} + \lambda_j(A)^2} \text{ for } j = 1, \dots, m,$$

where $\lambda_j(A)$, $j = 1, \dots, m$, are the eigenvalues of A.

If we consider now $\lambda_{j,\pm}(K_{OC})$ as a function of ν and expand it around zero, we get

$$\lambda_{j,+}(K_{OC}) = 1 + O(\nu),$$
$$\lambda_{j,-}(K_{OC}) = -\frac{1}{\nu} + O(\nu).$$

Thus:

- As $\nu \to 0$, $\lambda_{j,\pm}(K_{OC})$ tend to be split into two clusters:
  - one cluster accumulating around 1,
  - one cluster moving indefinitely toward $-\infty$ and accumulating around $-\frac{1}{\nu}$.
- This explain the observed convergence: the GMRES residual decays very fast in the first two iterations and then the convergence becomes slower again.

References

- *Iterative Methods and Preconditioners for Systems of Linear Equations*,
  Ciaramella and Gander, SIAM Fundamentals of Algorithms (2022)
- *A complex homographic best approximation problem. Application to optimized Robin–Schwarz algorithms, and optimal control problems*,
  Delourme and Halpern, SIAM J. Numer. Anal. (2021)
- *Optimized RAS Methods*,
  Cyr, Gander and Thomas, DD16 Proceedings (2005)
- *Schwarz methods over the course of time*,
  Gander, ETNA (2008)
- *A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations*,
  Benamou, SIAM J. Numer. Anal. (1996)
- *On the Schwarz alternating method. III*,
  Lions, Third DD Conference Papers (1990)

4 Part IV: OSM and nonlinear preconditioning

Nonlinear preconditioning

Consider a linear system of equations

$$F(\mathbf{x}) = A\mathbf{x} - \mathbf{b} = 0.$$

A domain decomposition method (like RAS and ORAS) corresponds to a preconditioner M. Preconditioning consists in transforming $F(\mathbf{x}) = 0$ into

$$T(F(\mathbf{x})) = M^{-1}F(\mathbf{x}) = M^{-1}A\mathbf{x} - M^{-1}\mathbf{b},$$

in a way that the properties of $M^{-1}A$ make a Krylov method converge faster ($M \approx A$).

Nonlinear preconditioning follows the same idea, just recognize the "correspondences":

$$\begin{aligned} \text{System } F(\mathbf{x}) = 0 &: \text{linear} &\longleftrightarrow\quad &\text{nonlinear} \\ \text{Transformation} &: \text{linear dd} &\longleftrightarrow\quad &\text{nonlinear dd} \\ \text{Solver} &: \text{Krylov} &\longleftrightarrow\quad &\text{Newton} \end{aligned}$$

For a nonlinear problem $F(\mathbf{x}) = 0$, we look for a transformation $T \approx F^{-1}$ such that $T(F(\mathbf{x})) = 0$ has the same solution to $F(\mathbf{x}) = 0$ and it guarantees a better behavior of the Newton solver.

Nonsmooth optimal control problems governed by nonlinear PDEs

Consider the nonsmooth optimal control problem

$$\min_{y,u} \; J(y,u) := \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1},$$

subjected to

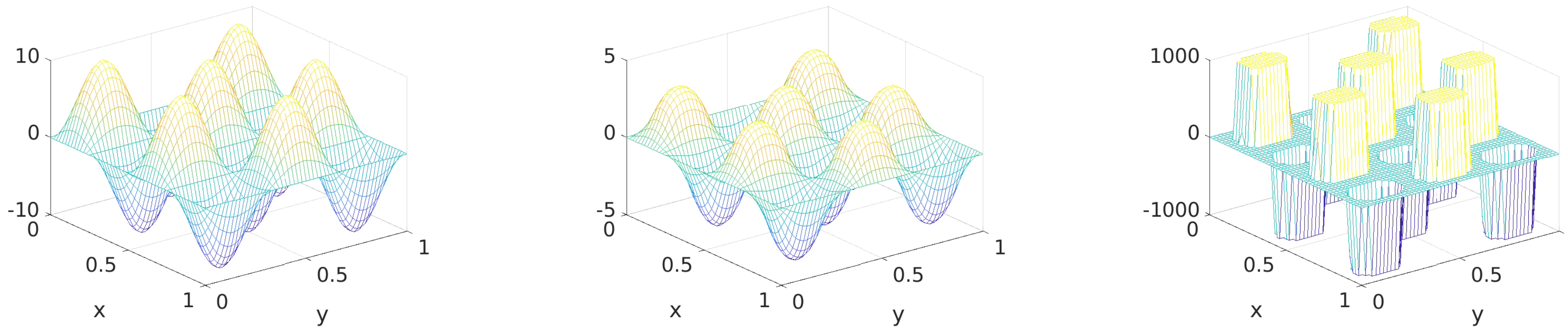
$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + u \text{ in } \Omega, \\ y &= 0 \text{ on } \partial\Omega, \end{aligned}$$

and

$$u \in U_{\mathrm{ad}} := \{v \in L^2(\Omega) \,:\, u_\ell(x) \leq u(x) \leq u_u(x) \text{ in } \Omega\}.$$

Under appropriate regularity assumptions there exists a minimizer $(y,u) \in X \times U_{\mathrm{ad}}$, where $X := H_0^1(\Omega) \cap L^\infty(\Omega)$ (Casas et al. 2012, Tröltzsch 2010).

Solution for $b = 10$, $\nu = 10^{-7}$, $\bar{u} = 10^3$ and $\beta = 10^{-2}$:



Target y_d | Optimal state y | Optimal control u

First-order optimality systems

Let $(y, u) \in X \times L^2(\Omega)$ be a minimizer, there exist an adjoint variable $p \in X$ and a subgradient

$$\lambda \in \partial\|u\|_{L^1} := \{w \in L^2(\Omega) \, : \, \|v\|_{L^1} - \|u\|_{L^1} \geq \langle w, v-u \rangle_{L^2} \text{ for all } v \in L^2(\Omega)\}$$

such that the quadruple (y, u, p, λ) satisfies the system (Casas et al. 2012, Tröltzsch 2010)

$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p + cp + b\varphi'(y)p &= y - y_d && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ \langle \nu u + p + \beta\lambda, v - u \rangle_{L^2} &\geq 0 && \text{for all } v \in U_{\text{ad}}. \end{aligned}$$

This optimality system can be formulated as (Stadler 2009)

$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + u && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p + cp + b\varphi'(y)p &= y - y_d && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega, \\ u &= \mu(p), \end{aligned}$$

where

$$\begin{aligned} \mu(p) := {} & \max(0, (-\beta - p)/\nu) + \min(0, (\beta - p)/\nu) \\ & - \max(0, -u_u + (-p - \beta)/\nu) - \min(0, -u_\ell + (-p + \beta)/\nu). \end{aligned}$$

Replacing the last equation into the first, we get

$$\begin{aligned} -\Delta y + cy + b\varphi(y) &= f + \mu(p) && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega, \\ -\Delta p + cp + b\varphi'(y)(p) &= y - y_d && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega. \end{aligned}$$

This nonlinear and nonsmooth system admits a solution $(y, p) \in X^2$.

Optimality system as a root problem

Let us consider the space $Y := H_0^1(\Omega)^*$ and define the operators

$$
\begin{aligned}
&A : H_0^1(\Omega) \to Y, && A(y)(v) = \int_\Omega \nabla y \cdot \nabla v + cyv \, dx && \forall y, v \in H_0^1(\Omega),\\
&B : H_0^1(\Omega) \to Y, && B(y)(v) = \int_\Omega yv \, dx && \forall y, v \in H_0^1(\Omega),\\
&C : L^\infty(\Omega) \to Y, && C(p)(v) = \int_\Omega \mu(p) v \, dx && \forall p \in L^\infty(\Omega), v \in H_0^1(\Omega),\\
&\Phi : L^\infty(\Omega) \to Y, && \Phi(y)(v) = \int_\Omega b\varphi(y) v \, dx && \forall y \in L^\infty(\Omega), v \in H_0^1(\Omega),\\
&\Psi : (L^\infty(\Omega))^2 \to Y, && \Psi(y,p)(v) = \int_\Omega b\partial_y \varphi(y) pv \, dx && \forall y, p \in L^\infty(\Omega), v \in H_0^1(\Omega),\\
&\ell_f : L^2(\Omega) \to Y, && \ell_f(v) = \int_\Omega fv \, dx && \forall v \in H_0^1(\Omega),\\
&\ell_{y_d} : L^2(\Omega) \to Y, && \ell_{y_d}(v) = \int_\Omega y_d v \, dx && \forall v \in H_0^1(\Omega).
\end{aligned}
$$

If one introduces the map $\mathcal{F} : X^2 \to Y^2$,

$$
\mathcal{F}(y,p) = \begin{bmatrix} A(y) + \Phi(y) - C(p) - \ell_f \\ A(p) + \Psi(y,p) - B(y) - \ell_{y_d} \end{bmatrix},
$$

then the (weak form of the) optimality system is equivalent to

$$
\mathcal{F}(y,p) = 0.
$$

Semismooth Newton method

For a given initialization $(y_0, p_0) \in X^2$, we consider the Newton-type method

$$\begin{aligned} \text{solve} \qquad & D\mathcal{F}(y_k, p_k)(dy, dp) = -\mathcal{F}(y_k, p_k), \\ \text{update} \qquad & (y_{k+1}, p_{k+1}) = (y_k, p_k) + (dy, dp), \end{aligned}$$

for $k = 0, 1, 2, \dots$, where

$$D\mathcal{F}(y,p)(\widetilde{y}, \widetilde{p}) = \begin{bmatrix} A(\widetilde{y}) + \Psi(y, \widetilde{y}) - \widetilde{C}(p)(\widetilde{p}) \\ A(\widetilde{p}) + D_y\Psi(y,p)(\widetilde{y}) + D_p\Psi(y,p)(\widetilde{p}) - B(\widetilde{y}) \end{bmatrix},$$

for any $(\widetilde{y}, \widetilde{p}) \in X^2$, where

$$D_y\Psi(y,p)(\widetilde{y})(v) = \int_\Omega \varphi''(y)[p, \widetilde{y}](v) \text{ for any } v \in H_0^1(\Omega)$$

and $D_p\Psi(y,p)(\widetilde{p}) = \Psi(y, \widetilde{p})$.

We must discuss the term $\widetilde{C}(p)(\widetilde{p})$, since it represents the generalized derivative of

$$C(p)(v) = \int_\Omega \mu(p) v \, dx,$$

where

$$\begin{aligned} \mu(p) :=& \max(0, (-\beta - p)/\nu) + \min(0, (\beta - p)/\nu) \\ & - \max(0, -u_u + (-p - \beta)/\nu) - \min(0, -u_\ell + (-p + \beta)/\nu). \end{aligned}$$

Generalized derivatives

Consider the operators

$$\mathcal{F}_{\max}, \mathcal{F}_{\min} : L^r(\Omega) \to L^s(\Omega),$$

$$\mathcal{F}_{\max}(v) := \max(0, v) \ \text{ and } \ \mathcal{F}_{\min}(v) := \min(0, v).$$

For any $1 \leq s < r \leq \infty$, these are slanty differentiable with generalized derivatives

$$\mathcal{G}_{\max}(v)(x) = \begin{cases} 1 & \text{if } v(x) > 0, \\ 0 & \text{if } v(x) \leq 0, \end{cases} \qquad \mathcal{G}_{\min}(v)(x) = \begin{cases} 1 & \text{if } v(x) \leq 0, \\ 0 & \text{if } v(x) > 0. \end{cases}$$

Hence

$$D\mu(p)(\widetilde{p}) = \frac{1}{\nu}\Big[-\mathcal{G}_{\max}(-\beta - p) - \mathcal{G}_{\min}(\beta - p) + \mathcal{G}_{\max}(-p - \beta - \nu u_u) + \mathcal{G}_{\min}(-p + \beta - \nu u_\ell)\Big]\widetilde{p},$$

or equivalently (using a characteristic function $\chi_{\mathcal{A}(p)}$)

$$D\mu(p)(\widetilde{p})(x) = -\frac{1}{\nu}\chi_{\mathcal{A}(p)}(x)\widetilde{p}(x) \text{ a.e. in } \Omega.$$

Therefore, the operator $\widetilde{C}(p) : L^\infty(\Omega) \to L^2(\Omega)$ is

$$\widetilde{C}(p)(\widetilde{p})(v) = \int_\Omega -\frac{1}{\nu}\chi_{\mathcal{A}(p)}(x)\widetilde{p}(x)v(x)\, dx,$$

for any $v \in L^2(\Omega)$.

Nonlinear optimized Schwarz method

$$
\begin{aligned}
-\Delta y + cy + b\varphi(y) &= f + \mu(p) && \text{in } \Omega \text{ with } y = 0 \text{ on } \partial\Omega,\\
-\Delta p + cp + b\varphi'(y)(p) &= y - y_d && \text{in } \Omega \text{ with } p = 0 \text{ on } \partial\Omega.
\end{aligned}
$$

Consider a non-overlapping decomposition $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$ (with interface Γ) and a parameter $q > 0$. The OSM is

$$
\begin{aligned}
-\Delta y_1^k + cy_1^k + b\varphi(y_1^k) &= f + \mu(p_1^k) && \text{in } \Omega_1,\\
y_1^k &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma,\\
\partial_{n_1} y_1^k + q\, y_1^k &= \partial_{n_1} y_2^{k-1} + q\, y_2^{k-1} && \text{on } \Gamma,\\
-\Delta p_1^k + cp_1^k + b\varphi'(y_1^k)(p_1^k) &= y_1^k - y_d && \text{in } \Omega_1,\\
p_1^k &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma,\\
\partial_{n_1} p_1^k + q\, p_1^k &= \partial_{n_1} p_2^{k-1} + q\, p_2^{k-1} && \text{on } \Gamma,
\end{aligned}
$$

and

$$
\begin{aligned}
-\Delta y_2^k + cy_2^k + b\varphi(y_2^k) &= f + \mu(p_2^k) && \text{in } \Omega_2,\\
y_2^k &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma,\\
\partial_{n_2} y_2^k + q\, y_2^k &= \partial_{n_2} y_1^{k-1} + q\, y_1^{k-1} && \text{on } \Gamma,\\
-\Delta p_2^k + cp_2^k + b\varphi'(y_2^k)(p_2^k) &= y_2^k - y_d && \text{in } \Omega_2,\\
p_2^k &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma,\\
\partial_{n_2} p_2^k + q\, p_2^k &= \partial_{n_2} p_1^{k-1} + q\, p_1^{k-1} && \text{on } \Gamma.
\end{aligned}
$$

Nonlinear OSM preconditioner

Let $\mathbf{y}_1 := (y_1, p_1)$ and $\mathbf{y}_2 := (y_2, p_2)$. The nonlinear OSM subproblems are in the limit given by

$$\mathcal{F}_1(\mathbf{y}_1, \mathbf{y}_2) := \begin{bmatrix} A_1(y_1) + \Phi_1(y_1) - C_1(p_1) + G_L(y_1, y_2) - \ell_{f,1} \\ A_1(p_1) + \Psi_1(y_1, p_1) - B_1(y_1) + G_L(p_1, p_2) - \ell_{y_d,1} \end{bmatrix} = 0,$$

$$\mathcal{F}_2(\mathbf{y}_2, \mathbf{y}_1) := \begin{bmatrix} A_2(y_2) + \Phi_1(y_2) - C_2(p_2) + G_R(y_2, y_1) - \ell_{f,2} \\ A_2(p_2) + \Psi_1(y_2, p_2) - B_2(y_2) + G_R(p_2, p_1) - \ell_{y_d,2} \end{bmatrix} = 0,$$

where G_L and G_R are

$$G_L(y_1, y_2)(v) = \int_\Gamma (-q y_1 + \partial_{n_1} y_2 + q y_2) v,$$

and

$$G_R(y_2, y_1)(v) = \int_\Gamma (-q y_2 + \partial_{n_2} y_1 + q y_1) v.$$

If we introduce the (local) solution mappings

$$S_1(\mathbf{y}_2) := \mathbf{y}_1 \ \text{ solution of } \mathcal{F}_1(\mathbf{y}_1, \mathbf{y}_2) = 0 \text{ for given } \mathbf{y}_2,$$
$$S_2(\mathbf{y}_1) := \mathbf{y}_2 \ \text{ solution of } \mathcal{F}_2(\mathbf{y}_2, \mathbf{y}_1) = 0 \text{ for given } \mathbf{y}_1.$$

Hence, we can rewrite the above system as ("$T(F(\mathbf{x})) = 0$")

$$\mathcal{F}_{\mathrm{P}}(\mathbf{y}_1, \mathbf{y}_2) := \begin{bmatrix} \mathbf{y}_1 - S_1(\mathbf{y}_2) \\ \mathbf{y}_2 - S_2(\mathbf{y}_1) \end{bmatrix} = 0.$$

Preconditioned Newton

The Newton method for $\mathcal{F}_{\mathrm{P}}(\mathbf{y}_1, \mathbf{y}_2) = 0$ is given by

$$\begin{aligned} D\mathcal{F}_{\mathrm{P}}(\mathbf{y}_1^k, \mathbf{y}_2^k)(\mathbf{d}_1, \mathbf{d}_2) &= -\mathcal{F}_{\mathrm{P}}(\mathbf{y}_1^k, \mathbf{y}_2^k), \\ (\mathbf{y}_1^{k+1}, \mathbf{y}_2^{k+1}) &= (\mathbf{y}_1^k, \mathbf{y}_2^k) + (\mathbf{d}_1, \mathbf{d}_2). \end{aligned}$$

Here, we have

$$D\mathcal{F}_{\mathrm{P}}(\mathbf{y}_1^k, \mathbf{y}_2^k)(\mathbf{d}_1, \mathbf{d}_2) = \begin{bmatrix} \mathbf{d}_1 - DS_1(\mathbf{y}_2^k)(\mathbf{d}_2) \\ \mathbf{d}_2 - DS_2(\mathbf{y}_1^k)(\mathbf{d}_1) \end{bmatrix},$$

where $DS_2(\mathbf{y}_1)(\mathbf{d}_1) = (\widetilde{y}_2, \widetilde{p}_2)$ solves

$$\begin{aligned} -\Delta \widetilde{y}_2 + c\widetilde{y}_2 + b\varphi'(y_2)\widetilde{y}_2 + \chi_{\mathcal{A}(p_2)}\widetilde{p}_2 &= 0 && \text{in } \Omega_2, \\ \widetilde{y}_2 &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma, \\ \partial_{n_2}\widetilde{y}_2 + q\,\widetilde{y}_2 &= \partial_{n_2} d_{1,y} + q\, d_{1,y} && \text{on } \Gamma, \\ -\Delta \widetilde{p}_2 + c\widetilde{p}_2 + b\varphi''(y_2)[p_2, \widetilde{y}_2] &= \widetilde{y}_2 && \text{in } \Omega_2, \\ \widetilde{p}_2 &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma, \\ \partial_{n_2}\widetilde{p}_2 + q\,\widetilde{p}_2 &= \partial_{n_2} d_{1,p} + q\, d_{1,p} && \text{on } \Gamma, \end{aligned}$$

and similarly for $DS_1(\mathbf{y}_2)(\mathbf{d}_2) = (\widetilde{y}_1, \widetilde{p}_1)$.

Overall algorithm

Require: Initial guess $\mathbf{y}^0$, tolerance ϵ, maximum number of iterations $k_{\max}$.

Compute $S_1(\mathbf{y}_2^0)$ and $S_2(\mathbf{y}_1^0)$.
Set $k = 0$ and assemble $\mathcal{F}_{\mathrm{P}}(\mathbf{y}^0)$.
while $\|\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)\| \geq \epsilon$ and $k \leq k_{\max}$ **do**
    Compute $\mathbf{d}^k$ by solving $D\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)(\mathbf{d}^k) = -\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)$ using a matrix-free Krylov method, where the action of $D\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)$ on a vector $\mathbf{d}$ is computed as above.
    Update $\mathbf{y}^{k+1} = \mathbf{y}^k + \mathbf{d}^k$.
    Set $k = k + 1$.
    Compute $S_1(\mathbf{y}_2^k)$ and $S_2(\mathbf{y}_1^k)$.
    Assemble $\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)$.
end while

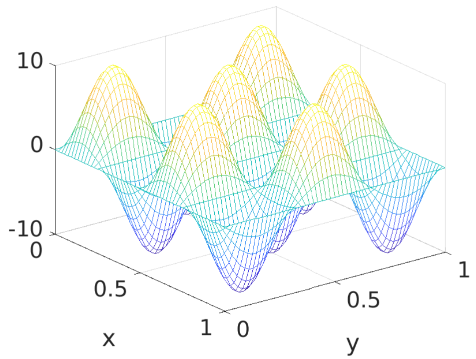
Remarks:

- At each Newton step, one must evaluate $\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k) \Longrightarrow$ solve subdomain problems.
- The linearized BVPs needed for the action of $D\mathcal{F}_{\mathrm{P}}(\mathbf{y}^k)$, are the same as the ones that appear in the inner Newton iterations of the subdomain solves.
- What we hope to gain from OSM preconditioner:
    - Easier subdomain solves (fewer degrees of freedom).
    - “Less nonlinearity” in the coupling between subdomains.
    - Fast convergence in the Krylov method (for well chosen parameter q).

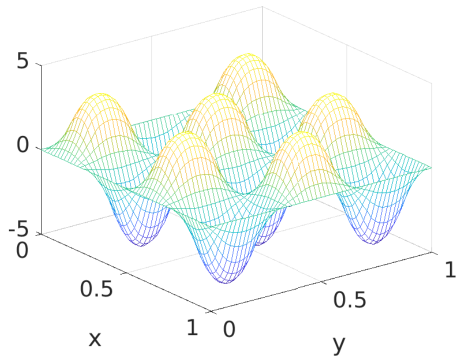
Numerical experiments

Test case: two subdomains (for multisubdomain see C., Müller, Kwok 2022):

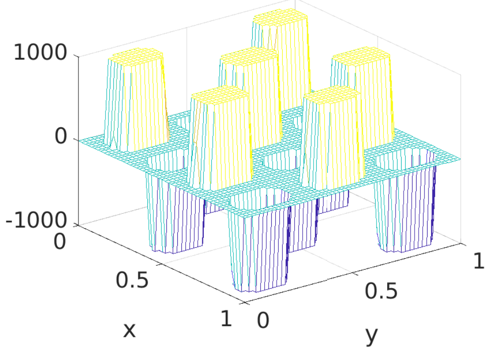
- Constraint PDE: $-\Delta y + cy + b\varphi(y) = u$ in $\Omega = (0,1)^2$.
- Nonlinear term $\varphi(y) = y + \exp(y)$ to be turned on or off ($b = 10$ or $b = 0$).
- Study convergence by varying β (L^1 reg.) and ν (L^2 reg.) and $\bar{u}$ (control bound).
- Solution for $b = 10$, $\nu = 10^{-7}$, $\bar{u} = 10^3$ and $\beta = 10^{-2}$:



Target y_d



Optimal state y



Optimal control u

Nonlinear solution strategies

We compare

- Semismooth Newton on the optimized Schwarz-preconditioned fixed point equation,
- Semismooth Newton on the global (unpreconditioned) system.

Each method is executed

- without continuation on ν,
- with continuation:
  - Start with $\nu = 0.1$.
  - Solve Jacobian system once and update solution; DO NOT iterate to convergence.
  - Replace ν by $\nu/4$ and repeat the process, until the desired ν is reached.

Outer Newton iterations

	q	b	$\bar{u} = 10^3$			$\bar{u} = \infty$		
			$\nu = 10^{-3}$	$\nu = 10^{-5}$	$\nu = 10^{-7}$	$\nu = 10^{-3}$	$\nu = 10^{-5}$	$\nu = 10^{-7}$
$\beta = 0$	1	0	4-5-2-4	6-9-11-11	4-11-41-12	3-5-2-4	3-9-2-9	3-12-3-12
	10	0	4-5-2-4	6-9-11-11	8-11-41-12	3-5-2-4	3-8-2-9	3-11-3-12
	100	0	3-5-2-4	6-9-11-11	×-11-41-12	3-5-2-4	3-9-2-9	3-11-3-12
	1	10	6-6-4-7	×-10-12-12	×-12-38-16	6-6-4-7	×-10-22-23	×-15-×-15
	10	10	5-6-4-7	7-10-12-12	×-12-38-16	5-6-4-7	×-10-22-23	×-14-×-15
	100	10	4-6-4-7	6-10-12-12	×-13-38-16	4-6-4-7	6-10-22-23	×-13-×-15
$\beta = 10^{-2}$	1	0	5-5-3-6	6-9-8-11	×-12-43-13	4-5-3-6	5-9-6-10	×-12-8-15
	10	0	4-5-3-6	6-9-8-11	×-11-43-13	4-5-3-6	4-9-6-10	×-12-8-15
	100	0	4-5-3-6	6-10-8-11	11-12-43-13	4-5-3-6	5-9-6-10	7-12-8-15
	1	10	6-6-4-6	×-11-10-12	×-12-×-17	6-6-4-6	×-10-18-×	×-13-×-15
	10	10	5-6-4-6	×-11-10-12	×-13-×-17	5-6-4-6	×-10-18-×	×-14-×-15
	100	10	4-6-4-6	6-11-10-12	9-13-×-17	4-6-4-6	6-10-18-×	×-13-×-15

Legend: OSM Prec. (no cont. - cont.) - Global Newton (no cont. cont.)

Inner Newton iterations

	q	b	$\bar{u} = 10^3$, $\nu = 10^{-3}$	$\bar{u} = 10^3$, $\nu = 10^{-5}$	$\bar{u} = 10^3$, $\nu = 10^{-7}$	$\bar{u} = \infty$, $\nu = 10^{-3}$	$\bar{u} = \infty$, $\nu = 10^{-5}$	$\bar{u} = \infty$, $\nu = 10^{-7}$
$\beta = 0$	1	0	6 - 5	31 - 12	$\times$ - 18	2 - 5	3 - 8	3 - 11
	10	0	5 - 5	26 - 11	96 - 19	2 - 5	3 - 8	3 - 11
	100	0	2 - 5	18 - 13	$\times$ - 19	2 - 5	2 - 8	3 - 11
	1	10	$\times$ - 17	$\times$ - 35	$\times$ - 47	27 - 17	$\times$ - 34	$\times$ - 60
	10	10	21 - 14	$\times$ - 31	103 - 43	21 - 14	$\times$ - 32	$\times$ - 53
	100	10	8 - 14	26 - 32	$\times$ - 43	8 - 14	45 - 30	$\times$ - 47
$\beta = 10^{-2}$	1	0	13 - 8	32 - 16	84 - 25	8 - 8	10 - 14	$\times$ - 25
	10	0	10 - 8	22 - 17	33 - 23	7 - 8	11 - 15	$\times$ - 24
	100	0	7 - 6	15 - 15	104 - 20	7 - 6	12 - 13	$\times$ - 22
	1	10	$\times$ - 17	$\times$ - 33	$\times$ - 45	28 - 17	$\times$ - 32	$\times$ - 47
	10	10	20 - 14	$\times$ - 33	$\times$ - 48	20 - 14	$\times$ - 30	$\times$ - 46
	100	10	10 - 14	23 - 30	125 - 44	10 - 14	40 - 26	$\times$ - 44

Legend: OSM Prec. (no cont. - cont.)

GMRES iterations

- Global Newton system preconditioned by $-\Delta + cI$ on state and adjoint.
- No need of OSM prec. in “easy” cases ($\nu = 10^{-3}$ and $\beta = b = 0$).
- For harder cases, OSM prec. with $q = 10$ is almost always best!

	q	b	$\bar{u} = 10^3$, $\nu = 10^{-3}$	$\bar{u} = 10^3$, $\nu = 10^{-5}$	$\bar{u} = 10^3$, $\nu = 10^{-7}$	$\bar{u} = \infty$, $\nu = 10^{-3}$	$\bar{u} = \infty$, $\nu = 10^{-5}$	$\bar{u} = \infty$, $\nu = 10^{-7}$
$\beta = 0$	1	0	143 - 128 - 26 - 17	279 - 163 - 303 - 123	139 - 175 - 1255 - 155	100 - 128 - 26 - 17	170 - 140 - 69 - 46	$\times$ - 149 - 371 - 66
	10	0	90 - 75 - 26 - 17	179 - $\boxed{101}$ - 303 - 123	254 - $\boxed{108}$ - 1255 - 155	64 - 75 - 26 - 17	114 - 88 - 69 - 46	$\times$ - 92 - 371 - 66
	100	0	73 - 90 - 26 - 17	177 - 114 - 303 - 123	$\times$ - 125 - 1255 - 155	73 - 90 - 26 - 17	60 - 88 - 69 - 46	54 - 93 - 371 - 66
	1	10	266 - 204 - 49 - 66	$\times$ - 251 - 397 - 255	$\times$ - 268 - 1479 - 457	256 - 204 - 49 - 66	$\times$ - 240 - 2000 - 1172	$\times$ - 379 - $\times$ - 928
	10	10	124 - 88 - 49 - 66	226 - $\boxed{129}$ - 397 - 255	$\times$ - $\boxed{168}$ - 1479 - 255	117 - 88 - 49 - 66	$\times$ - $\boxed{152}$ - 2000 - 1172	$\times$ - $\boxed{190}$ - $\times$ - 928
	100	10	122 - 155 - 49 - 66	139 - 239 - 397 - 255	$\times$ - 274 - 1479 - 457	122 - 155 - 49 - 66	161 - 234 - 2000 - 1172	$\times$ - 250 - $\times$ - 928
$\beta = 10^{-2}$	1	0	226 - 164 - 31 - 42	290 - 198 - 187 - 123	$\times$ - 223 - 1065 - 168	188 - 164 - 31 - 42	246 - 168 - 183 - 109	$\times$ - 218 - 522 - 380
	10	0	111 - 95 - 31 - 42	178 - $\boxed{121}$ - 187 - 123	$\times$ - $\boxed{130}$ - 1065 - 168	115 - 95 - 31 - 42	143 - 124 - 183 - 109	$\times$ - $\boxed{145}$ - 522 - 380
	100	0	135 - 118 - 31 - 42	179 - 158 - 187 - 123	333 - 175 - 1065 - 168	135 - 118 - 31 - 42	145 - 147 - 183 - 109	165 - 173 - 522 - 380
	1	10	273 - 235 - 49 - 54	$\times$ - 238 - 299 - 233	$\times$ - 228 - $\times$ - 416	261 - 235 - 49 - 54	$\times$ - 254 - 1362 - $\times$	$\times$ - 311 - $\times$ - 752
	10	10	139 - 124 - 49 - 54	$\times$ - 158 - 299 - 233	$\times$ - $\boxed{164}$ - $\times$ - 416	138 - 124 - 49 - 54	$\times$ - $\boxed{161}$ - 1362 - $\times$	$\times$ - $\boxed{179}$ - $\times$ - 752
	100	10	122 - 162 - 49 - 54	$\boxed{141}$ - 251 - 299 - 233	215 - 300 - $\times$ - 416	122 - 162 - 49 - 54	167 - 219 - 1362 - $\times$	$\times$ - 303 - $\times$ - 752

Legend: OSM Prec. (no cont. - cont.) - Global Newton (no cont. cont.)

References

- *Nonlinear optimized Schwarz preconditioner for elliptic optimal control problems*, Ciaramella, Kwok, Müller, DD26 Proceedings (2022)
- *Nonlinear Preconditioning: how to use a Nonlinear Schwarz Method to Precondition Newton's Method*, Dolean, Gander, Kheriji, Kwok and Masson, SIAM J. Sci. Comp. (2016)
- *Optimality Conditions and Error Analysis of Semilinear Elliptic Control Problems with L1 Cost Functional*, Casas, Herzog and Wachsmuth, SIAM J. Optim. (2012)
- *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, Tröltzsch, American Mathematical Society (2010)
- *Elliptic optimal control problems with L1-control cost and applications for the placement of control devices*, Stadler, Comput. Optim.Appl. (2006)

Thank you!