Schwarz-in-Time Methods for Parabolic Optimal Control Problems

Jean-Morlet Chair Research School, CIRM 2022

Felix Kwok (joint work with Martin J. Gander)

September 5, 2022

Département de mathématiques et de statistique, Université Laval

Outline

- Optimization Under DE Constraints
- Time-domain decomposition for control
- Analysis by Diagonalization
- Analysis by Energy Estimates

Optimization Under DE Constraints

Optimal Control

According to Glowinski & Lions,

> *"At* **a given time horizon** *we want the system under study to behave* **exactly** *as we wish (or in a manner arbitrarily close to it)."*

[1] R. Glowinski & J.L. Lions, Exact and approximate controllability for distributed parameter systems, *Acta Numerica*, 1994.

Optimal Control

- Optimal control problem : minimize

$$J[y,u] = \frac{1}{2}\|Dy(T) - y_{\text{target}}\|^2 + \frac{\nu}{2}\int_0^T \|u(t)\|^2 \, dt$$

subject to the (non)-linear ODE constraint

$$\dot{y}(t) = f(y(t)) + Bu(t), \qquad t \in (0,T).$$

- Assumptions :
  1. No state or control constraints
  2. Control entering additively
- Includes cases where f is the discretization of a partial differential operator

Problem with Tracking

- Minimize

$$J[y,u] = \frac{1}{2}\int_0^T \|Cy(t) - \hat{y}(t)\|^2\, dt + \frac{\nu}{2}\int_0^T \|u(t)\|^2\, dt$$

subject to the (non)-linear ODE constraint

$$\dot{y}(t) = f(y(t)) + Bu(t), \qquad t \in (0,T).$$

- Can be formally transformed into a problem with no tracking by introducing additional state variable $z(t)$ satisfying

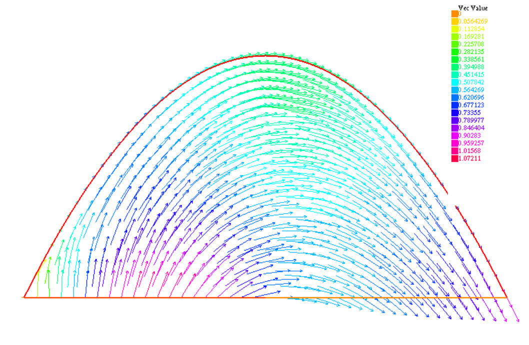
$$\frac{d}{dt}(z^2) = \|Cy(t) - \hat{y}(t)\|^2$$

- In this talk, we will derive first-order optimality conditions directly, without using this nonlinear transformation
- Problem may have both tracking and target terms

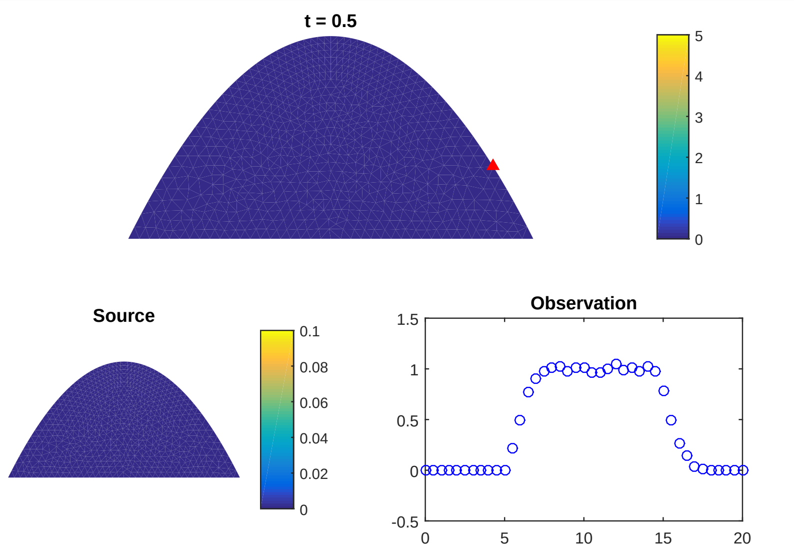
Example : Contaminant Tracking

Find source term u that best match observation $y_d(t)$, subject to the advection-diffusion equation

$$\frac{\partial y}{\partial t} + \nabla \cdot (\mathbf{v}y - \nu \nabla y) = Bu$$



Example : Contaminant Tracking



Applications

- Weather prediction : assimilation of measurements into prediction model, cf. 4DVar
- Aeronautics : Aircraft design for reduction of noise due to boundary layer separation (He-Glowinski-Metcalfe-Periaux 1998, Dandois 2007, Borel-Halpern-Ryan 2010, ...)
- Bio-medicine : Drug administration in chemotherapy (Jackson & Byrne 2000, Rockne et al. 2010, Corwin et al. 2013,...)
- Oil & Gas : Oil field management optimization, data assimilation, history matching, ...
- **ALLOWAPP project** with L. Halpern, B. Delourme, J. Salomon and H.-Y. Liu (French ANR/RGC Hong Kong) : control for wave propagation problems with applications to wave localization and data assimilation
- ...

Optimality System (for linear constraint PDE)

- For the problem

$$\min J[y,u] = \frac{1}{2}\int_0^T \|Cy(t) - \hat{y}\|^2\,dt + \frac{\gamma}{2}\|Dy(T) - y_T\|^2 + \frac{\nu}{2}\int_0^T \|u(t)\|^2\,dt$$

$$\text{s.t.} \qquad \dot{y}(t) + Ay(t) = Bu(t), \quad t \in (0,T).$$

- Derive first-order optimality conditions formally using Lagrange multipliers λ :

$$L(y,\lambda,u) = J(y,u) + \langle \lambda, \dot{y} + Ay - Bu\rangle.$$

- We choose the inner product $\langle u, v\rangle = \int_0^T u^T v\,dt$.

Optimality System

- Since the optimal solution is a stationary point of $L(y, \lambda, u)$, we have

$$\frac{\partial}{\partial \varepsilon} L(y + \varepsilon z, \lambda, u) = 0 \qquad \text{for all } z \in V,$$

which gives

$$0 = \langle Cy - \hat{y}, Cz \rangle + \gamma(Dy(T) - y_T, Dz(T)) + \int_0^T (\lambda, \dot{z} + Az)\, dt.$$

- Integration by parts gives

$$\begin{aligned} 0 = {} & \langle C^T(Cy - \hat{y}), z \rangle + \gamma(D^T(Dy(T) - y_T), z(T)) \\ & + (\lambda(T), z(T)) - (\lambda(0), z(0)) + \int_0^T (-\dot{\lambda} + A^T \lambda, z)\, dt. \end{aligned}$$

Optimality System

UNIVERSITÉ LAVAL

$$
\begin{aligned}
0 &= \langle C^T(Cy - \hat{y}), z\rangle + \gamma(D^T(Dy(T) - y_T), z(T)) \\
&\quad + (\lambda(T), z(T)) - \underbrace{(\lambda(0), z(0))}_{=0} + \int_0^T (-\dot{\lambda} + A^T\lambda, z)\,dt.
\end{aligned}
$$

- This equation must be satisfied for all z with $z(0) = 0$, so we get the *adjoint problem*

$$
\begin{aligned}
\dot{\lambda} - A^T\lambda &= C^T(Cy - \hat{y}) \qquad \text{on } (0, T), \\
\lambda(T) &= -\gamma D^T(Dy(T) - y_T).
\end{aligned}
$$

- Taking a variation with respect to u gives the algebraic constraint $u = \nu^{-1}B^T\lambda$.

Optimality System : Summary

- First order optimality system (using Lagrange multipliers) :

$$\begin{cases} \dot{y} + Ay = \nu^{-1} BB^T \lambda, \\ y(0) = y_0, \end{cases}$$

Forward problem

$$\begin{cases} \dot{\lambda} - A^T \lambda = C^T (Cy - \hat{y}), \\ \lambda(T) = -\gamma D^T (Dy(T) - \hat{y}(T)), \end{cases}$$

Adjoint problem

- Coupled two-point boundary value problem !
- If we then discretize the ODE system in time, we get the "Optimize-then-discretize" approach
- We could also first discretize the state ODE and the objective function before deriving the optimality conditions $\implies$ "Discretize-then-optimize"
- Either way, get a huge linear system ($d+1$-dimensional problem with $N_x \times N_t$ unknowns) !

Parallelization

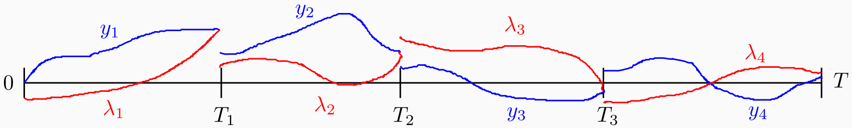
- Fastest supercomputers in the world (June 2022) :
  - Frontier (ORNL, USA, 8,730,112 cores, 1,685 Pflops/s)
  - Fugaku (RIKEN, Japan, 7,630,848 cores, 537 Pflops/s)

Parallelization

- Gradient descent methods require solving forward and backward problems repeatedly
- Use existing parallelization techniques for IVP
  - Discretize in time + DD in space
  - Multigrid (Hackbusch 1984, Horton & Vandewalle 1995, . . .)
  - Waveform relaxation (Gander & Stuart 1998, Giladi & Keller 2001, Heinkenschloss & Herty 2007, . . .)
  - **Parareal** (Lions, Maday & Turinici 2001, Mathew, Sarkis & Schaerer 2010, Ulbrich 2015, . . .)
- BUT : does not exploit structure of the control problem
- Our approach : **Time-domain decomposition** on coupled forward-backward problem
- Related approach : ParaOpt (cf. talks by J. Salomon)
- For maximal scalability, use **in combination with DD in space** (cf. talks by V. Dolean, G. Ciaramella, B.C. Mandal, . . .)

Time-domain decomposition for control

Time-domain decomposition for control



- Divide time horizon $(0, T)$ into "subdomains" $I_i = (T_{i-1}, T_i)$
- Subdomain problem $(y_i(t), \lambda_i(t))$ on I_i well defined (and easier to solve) when $y(T_{i-1})$ and $\lambda(T_i)$ are given
- Interface states $Y_i = y(T_i)$ and $\Lambda_i = \lambda(T_i)$ satisfy **continuity conditions** :

$$y_i(T_i) = y_{i+1}(T_i), \qquad \lambda_i(T_{i+1}) = \lambda_{i+1}(T_{i+1})$$

- If the subdomains do not overlap, this is essentially a multiple shooting problem, which we want to solve iteratively.

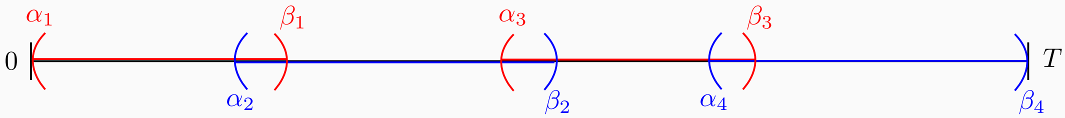
Overlapping Schwarz (Barker & Stoll (2015))

- Use overlapping subintervals (α_j, β_j)
- Solve the coupled forward-backward PDE on each subinterval in parallel

$$\dot{y}_j^k + A y_j^k = \nu^{-1}\lambda_j^k, \qquad \dot{\lambda}_j^k - A^T\lambda_j^k = y_j^k - \hat{y}$$

- Initial and final conditions from neighbours at previous iterate :

$$y_j^k(\alpha_j) = y_{j-1}^{k-1}(\alpha_j), \qquad \lambda_j^k(\beta_j) = \lambda_{j+1}^{k-1}(\beta_j).$$




SWITZERLAND
0 km 12,5 25 37,5 km
© 2009 Ezilon.com All Right Reserved
National Capital (128,000 in 2007)
over 350,000
over 100,000
over 50,000
over 20,000
other main city
other city
Geneva
Ascona
Workshop
Lugano
DD22
GERMANY
AUSTRIA
FRANCE
ITALY
LIECHTENSTEIN
Basel
Zürich
Luzern
BERN/BERNE
Lausanne
Genève (Geneva)
Lac Léman
Sankt Gallen
Winterthur
Schaffhausen
Bodensee
Lake Maggiore
Lago di Lugano
Como
Sondrio
Annecy
Besançon
Mulhouse
Konstanz
Friedrichshafen
Ravensburg
Bregenz
Vaduz
Chur
Davos
Bellinzona
Locarno
Verbania
Aosta
Berner Alpen
Glarner Alpen
Penninische Alpen
Lepontine
Silvretta
Albula Alpen

Overlapping Schwarz (Barker & Stoll (2015))

They observe experimentally that :

- Fast convergence for Dirichlet problems
- Convergence **even when subdomains do not overlap**
- For fixed overlap size, convergence is nearly independent of the spatial and temporal grid size
- Convergence may slow down when we increase the number of subintervals

Can we understand this behaviour ?


Geneva
Ascona
Workshop
DD22
SWITZERLAND
FRANCE
FRANCE
ITALY
Mulhouse
Schaffhausen
Ravensburg
Basel
Zürich
Besançon
Biel
BERN/BERNE
Neuchâtel
Yverdon-les-Bains
Fribourg
Thun
Lausanne
Lac Léman
Genève (Geneva)
Berner Alpen
Annecy
Verbania
Lago Maggiore
Como
Aosta

Optimized Schwarz Method (Gander & K., DD22 proceedings, 2016)

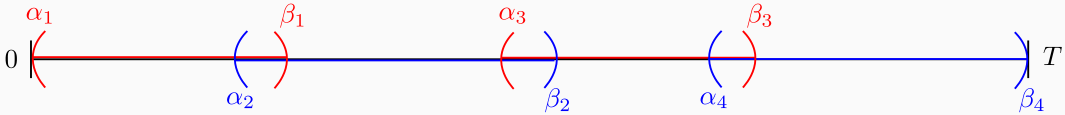
For $k = 1, 2, \ldots$, solve on each (α_j, β_j)

$$\begin{cases} \dot{y}_j^k + Ay_j^k = \nu^{-1}\lambda_j^k & \text{on } (\alpha_j, \beta_j), \\ \dot{\lambda}_j^k - A^T\lambda_j^k = y_j^k - \hat{y}_j, \end{cases}$$

with boundary conditions (cf. Lagnese & Leugering 2003)

$$y_j^k(\alpha_j) - q_j\lambda_j^k(\alpha_j) = y_{j-1}^{k-1}(\alpha_j) - q_j\lambda_{j-1}^{k-1}(\alpha_j),$$

$$\lambda_j^k(\beta_j) + p_j y_j^k(\beta_j) = \lambda_{j+1}^{k-1}(\beta_j) + p_j y_{j+1}^{k-1}(\beta_j).$$



Optimized Schwarz Method (Gander & K., DD22 proceedings, 2016)

Université Laval

For $p, q \neq 0$, this is equivalent to

$$\min \frac{1}{2}\int_{\alpha_j}^{\beta_j} \|y(t;u) - \hat{y}\|^2 + \frac{\nu}{2}\int_{\alpha_j}^{\beta_j} \|u\|^2 + \frac{p_j}{2}\|y(\beta_j;u) - p_j^{-1} g_{j+1}^{k-1}\|^2 + \frac{1}{2q_j}\|y(\alpha_j;u) - h_{j-1}^{k-1}\|^2$$

where

$$g_{j+1}^{k-1} = \lambda_{j+1}^{k-1}(\beta_j) + p_j y_{j+1}^{k-1}(\beta_j), \qquad h_{j-1}^{k-1} = y_{j-1}^{k-1}(\alpha_j) - q_j \lambda_{j-1}^{k-1}(\alpha_j)$$

- For $p = q = 0$, this reduces to Dirichlet transmission conditions
- Minimization problem with small changes in boundary conditions $\implies$ solvers available !

Subdomain solves

- **A simple shooting method :** for a given initial condition y_0 and control, consider the mapping $F(y_0, u)$ as follows :
  1. Integrate $\dot{y} + Ay = Bu$, $y(0) = y_0$ forwards to $t = T$
  2. Let $\lambda(T) = h - py(T)$
  3. Integrate $\dot{\lambda} - A^T y = C^T(Cy - \hat{y})$ backwards to $t = 0$.
  4. $F(y_0, u) = (y_0 - q\lambda(0) - g, \nu u - B^T \lambda)$
- Then
$$F(y_0, u) = F(0,0) + K \begin{pmatrix} y_0 \\ u \end{pmatrix}$$
is an affine mapping, so we can solve $F(y_0, u) = 0$ using e.g. GMRES
- Alternatively, use an all-at-once approach (Rees, Stoll & Wathen (2010), Pearson, Stoll & Wathen (2012), Pearson (2016), ...), or any other solver for a single time interval.

Optimized Schwarz Method (Gander & K., DD22 proceedings, 2016)

For $k = 1, 2, \ldots$, solve on each (α_j, β_j)

$$\begin{cases} \dot{y}_j^k + A y_j^k = \nu^{-1} \lambda_j^k & \text{on } (\alpha_j, \beta_j), \\ \dot{\lambda}_j^k - A^T \lambda_j^k = y_j^k - \hat{y}_j, \end{cases}$$

with boundary conditions

$$\begin{aligned} y_j^k(\alpha_j) - q_j \lambda_j^k(\alpha_j) &= y_{j-1}^{k-1}(\alpha_j) - q_j \lambda_{j-1}^{k-1}(\alpha_j), \\ \lambda_j^k(\beta_j) + p_j y_j^k(\beta_j) &= \lambda_{j+1}^{k-1}(\beta_j) + p_j y_{j+1}^{k-1}(\beta_j). \end{aligned}$$

- Convergence for which values of p_j and q_j ?
- How to choose p_j and q_j to optimize convergence ?

Analysis by Diagonalization

Convergence Analysis

- Diagonalization
  - + Explicit formula for contraction rate
  - + With or without overlap
  - – Assumes $A = A^T$
- Energy estimates
  - Integration by parts
  - + General setting ($A \neq A^T$, boundary control, etc.)
  - + Multiple subdomains
  - – No overlap

Convergence Analysis

- Diagonalization
  - \+ Explicit formula for contraction rate
  - \+ With or without overlap
  - − Assumes $A = A^T$
- Energy estimates
  - • Integration by parts
  - \+ General setting ($A \neq A^T$, boundary control, etc.)
  - \+ Multiple subdomains
  - − No overlap

Analysis for two subdomains

- Subdomain problems :

$$\begin{cases} \begin{bmatrix} \dot{y}_1^k \\ \dot{\lambda}_1^k \end{bmatrix} + \begin{bmatrix} A & -\nu^{-1} I \\ -I & -A^T \end{bmatrix} \begin{bmatrix} y_1^k \\ \lambda_1^k \end{bmatrix} = \begin{bmatrix} 0 \\ -\hat{y} \end{bmatrix} \qquad \text{on } I_1 = (0, \beta), \\ y_1^k(0) = y_0, \\ \lambda_1^k(\beta) + p y_1^k(\beta) = \lambda_2^{k-1}(\beta) + p y_2^{k-1}(\beta), \end{cases}$$

$$\begin{cases} \begin{bmatrix} \dot{y}_2^k \\ \dot{\lambda}_2^k \end{bmatrix} + \begin{bmatrix} A & -\nu^{-1} I \\ -I & -A^T \end{bmatrix} \begin{bmatrix} y_2^k \\ \lambda_2^k \end{bmatrix} = \begin{bmatrix} 0 \\ -\hat{y} \end{bmatrix} \qquad \text{on } I_2 = (\alpha, T), \\ y_2^k(\alpha) - q\lambda_2^k(\alpha) = y_1^{k-1}(\alpha) - q\lambda_1^{k-1}(\alpha), \\ \lambda_2^k(T) = -\gamma(y_2^k(T) - \hat{y}(T)). \end{cases}$$

Analysis for two subdomains

- Assume $A = A^T$ and diagonalize : $y \to z$, $\lambda \to \mu$:

$$\begin{cases} \begin{bmatrix} \dot{z}_1^k \\ \dot{\mu}_1^k \end{bmatrix} + \begin{bmatrix} D & -\nu^{-1} I \\ -I & -D \end{bmatrix} \begin{bmatrix} z_1^k \\ \mu_1^k \end{bmatrix} = \begin{bmatrix} 0 \\ -\hat{z} \end{bmatrix} \qquad \text{on } I_1 = (0, \beta), \\ z_1^k(0) = z_0, \\ \mu_1^k(\beta) + p z_1^k(\beta) = \mu_2^{k-1}(\beta) + p z_2^{k-1}(\beta), \end{cases}$$

- The ODE system decouples into n independent 2×2 subsystems :

$$\dot{z}_j^{(i),k} + d_i z_j^{(i),k} - \nu^{-1} \mu_j^{(i),k} = 0,$$
$$\mu_1^{(i),k} - z_j^{(i),k} - d_i \mu_j^{(i),k} = -\hat{z}^{(i)},$$

- For subdomain $I_2 = (\alpha, T)$, we have the same ODE system, but with the boundary conditions

$$z_2^k(\alpha) - q\mu_2^k(\alpha) = z_1^{k-1}(\alpha) - q\mu_1^{k-1}(\alpha),$$
$$\mu_2^k(T) = -\gamma(z_2^k(T) - \hat{z}(T)).$$

Analysis for two subdomains

UNIVERSITÉ LAVAL

- Eliminating μ : the ODE in z gives

$$\mu_j^{(i),k} = \nu(\dot{z}_j^{(i),k} + d_i z_j^{(i),k}),$$

so substituting into the adjoint $\mu_1^{(i),k} - z_j^{(i),k} - d_i \mu_j^{(i),k} = -\hat{z}^{(i)}$ yields

$$\ddot{z}_j^{(i),k} - (d_i^2 + \nu^{-1}) z_j^{(i),k} = -\nu^{-1} \hat{z}^{(i)}.$$

- For subdomain I_1, we also get the boundary conditions

$$z_1^{(i),k}(0) = z_0^{(i)}(0)$$

$$\dot{z}_1^{(i),k} + (d_i + p\nu^{-1}) z_1^{(i),k}\Big|_{t=\beta} = \dot{z}_2^{(i),k-1} + (d_i + p\nu^{-1}) z_2^{(i),k-1}\Big|_{t=\beta}.$$

- Even for $p = 0$, this corresponds to Robin conditions !

Theorem (Gander & K., 2016)

The parallel Schwarz method converges whenever $\rho < 1$, where

$$\rho^2 = \max_{d_i \in \lambda(A)} \left| \frac{\sigma_i q \cosh(\sigma_i \alpha) + (q d_i - \nu^{-1}) \sinh(\sigma_i \alpha)}{\sigma_i \cosh(\sigma_i \beta) + (d_i + p\nu^{-1}) \sinh(\sigma_i \beta)} \right.$$
$$\left. \cdot \frac{\nu^{-1/2} \left[p \cosh(\sigma_i (T-\beta) + \theta_i) - \gamma \cosh(\sigma_i (T-\beta) - \theta_i) \right] - (1 - \nu^{-1} p \gamma) \sinh(\sigma_i (T-\beta))}{\nu^{-1/2} \left[\cosh(\sigma_i (T-\alpha) + \theta_i) + q\gamma \cosh(\sigma_i (T-\alpha) - \theta_i) \right] + (q + \nu^{-1}\gamma) \sinh(\sigma_i (T-\alpha))} \right|,$$

with

- $d_i = i$th eigenvalue of A,
- $\sigma_i = \sqrt{d_i^2 + \nu^{-1}} > d_i \geq 0$,
- $\theta_i = \tanh^{-1}(d_i/\sigma_i)$.

Dirichlet Case ($p = q = 0$)

The convergence rate simplifies to

$$\rho^2 = \max_i \left(\frac{\sinh(\sigma_i \alpha)}{\cosh(\sigma_i \beta + \theta_i)} \cdot \frac{\nu^{1/2} \sinh(\sigma_i (T - \beta)) + \gamma \cosh(\sigma_i (T - \beta) - \theta_i)}{\gamma \sinh(\sigma_i (T - \alpha)) + \nu^{1/2} \cosh(\sigma_i (T - \alpha) + \theta_i)} \right).$$

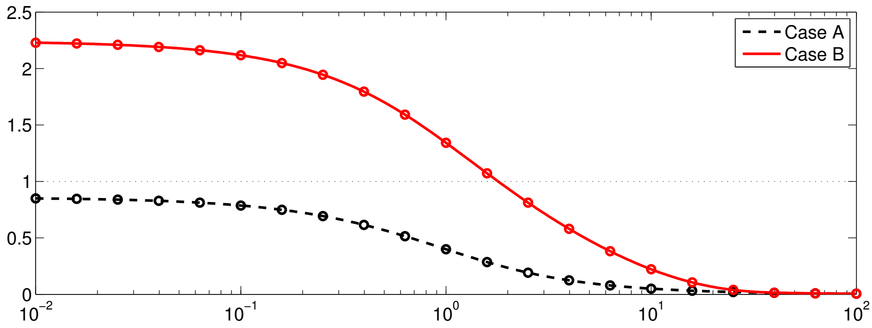
Theorem : ($\gamma = 0$, no target state) For two subdomains with overlap $L \geq 0$, the parallel Schwarz method for two subdomains converges with the estimate

$$\rho \leq \frac{e^{-L\sqrt{d_{\min}^2 + \nu^{-1}}}}{\sqrt{1 + \nu d_{\min}^2} + \nu^{1/2} d_{\min}},$$

where $d_{\min} > 0$ is the smallest eigenvalue of A.

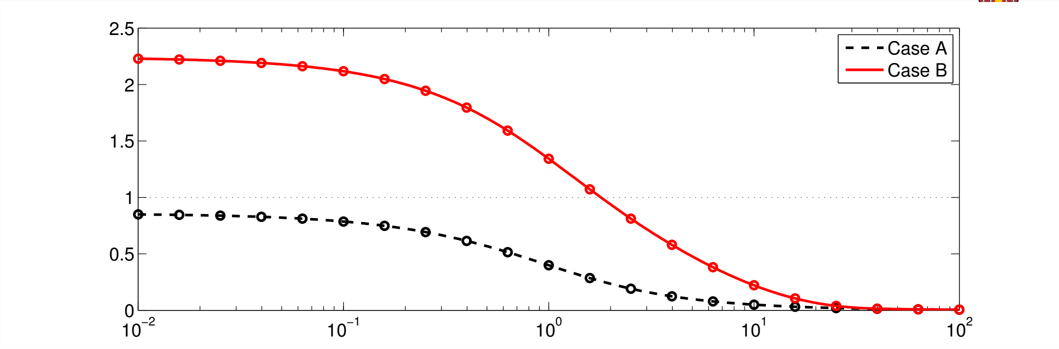
- Method converges even without overlap
- Convergence independent of the spatial mesh parameter !

Dirichlet Case ($p = q = 0$)



- Case A : $\Omega_1 = (0,1),\ \Omega_2 = (1,3),\ \gamma = 0$
- Case B : $\Omega_1 = (0,2.9),\ \Omega_2 = (2.9,3),\ \gamma = 10$

Dirichlet Case ($p = q = 0$)



- Case A converges for all positive definite matrices
- Convergence slow if $d_{\min} \ll 1$
- Case B diverges if $d_{\min} \lesssim 2$ (e.g. Neumann boundary)

Optimized case, $p = q$

- If $\gamma = 0$, the expression simplifies to

$$\rho^2 = \max_{d_i \in \lambda(A)} \left| \frac{\sigma_i p \cosh(\sigma_i \alpha) + (p d_i - \nu^{-1}) \sinh(\sigma_i \alpha)}{\sigma_i \cosh(\sigma_i \beta) + (d_i + p \nu^{-1}) \sinh(\sigma_i \beta)} \cdot \frac{p \sigma_i \cosh(\sigma_i (T - \beta)) + (p d_i - 1) \sinh(\sigma_i (T - \beta))}{\sigma_i \cosh(\sigma_i (T - \alpha)) + (p + d_i) \sinh(\sigma_i (T - \alpha))} \right|.$$

Optimized case, $p = q$

- If $\gamma = 0$, the expression simplifies to

$$\rho^2 = \max_{d_i \in \lambda(A)} \left| \frac{\sigma_i p \cosh(\sigma_i \alpha) + (p d_i - \nu^{-1}) \sinh(\sigma_i \alpha)}{\sigma_i \cosh(\sigma_i \beta) + (d_i + p\nu^{-1}) \sinh(\sigma_i \beta)} \cdot \frac{p\sigma_i \cosh(\sigma_i (T-\beta)) + (p d_i - 1) \sinh(\sigma_i (T-\beta))}{\sigma_i \cosh(\sigma_i (T-\alpha)) + (p + d_i) \sinh(\sigma_i (T-\alpha))} \right|.$$

- For high frequencies and no overlap, we have

$$\rho \longrightarrow p \cdot \underbrace{\lim_{d_i \to \infty} \left(\frac{\cosh(\sigma_i \alpha + \theta_i) \cosh(\sigma_i (T - \alpha) + \theta_i)}{\cosh(\sigma_i \alpha + \theta_i) \cosh(\sigma_i (T - \alpha) + \theta_i)} \right)^{1/2}}_{=1}.$$

So convergence cannot occur unless $p \in [0, 1)$.

Optimized case, $p = q$

- If $\gamma = 0$, the expression simplifies to

$$\rho^2 = \max_{d_i \in \lambda(A)} \left| \frac{\sigma_i p \cosh(\sigma_i \alpha) + (pd_i - \nu^{-1}) \sinh(\sigma_i \alpha)}{\sigma_i \cosh(\sigma_i \beta) + (d_i + p\nu^{-1}) \sinh(\sigma_i \beta)} \cdot \frac{p\sigma_i \cosh(\sigma_i(T-\beta)) + (pd_i - 1) \sinh(\sigma_i(T-\beta))}{\sigma_i \cosh(\sigma_i(T-\alpha)) + (p + d_i) \sinh(\sigma_i(T-\alpha))} \right|.$$

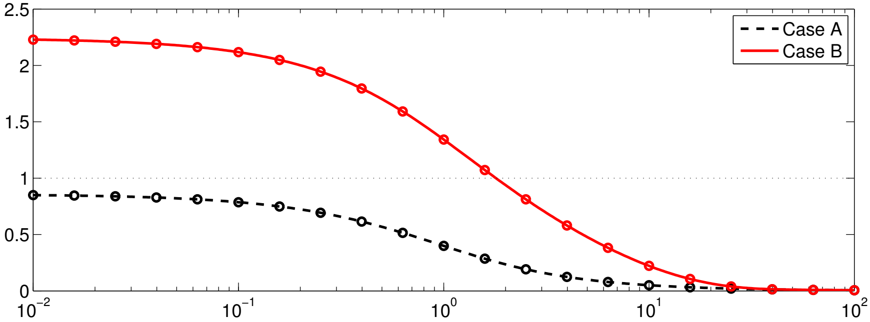
- For high frequencies and no overlap, we have

$$\rho \longrightarrow p \cdot \underbrace{\lim_{d_i \to \infty} \left(\frac{\cosh(\sigma_i \alpha + \theta_i) \cosh(\sigma_i(T-\alpha) + \theta_i)}{\cosh(\sigma_i \alpha + \theta_i) \cosh(\sigma_i(T-\alpha) + \theta_i)} \right)^{1/2}}_{=1}.$$

- Optimal p obtained by equioscillation : find p^* such that

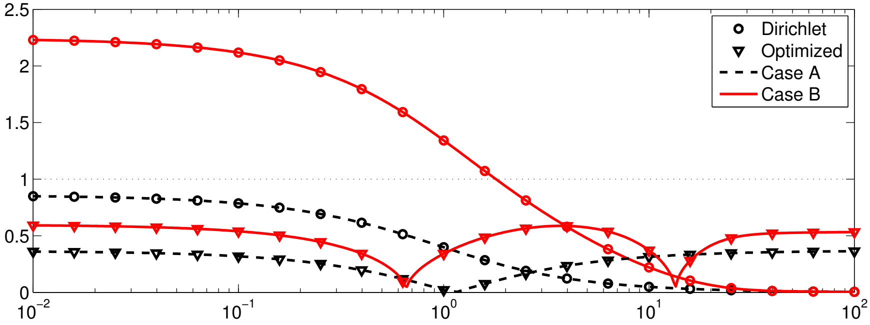
$$\lim_{d_i \to 0} \rho(p^*) = \lim_{d_i \to \infty} = p^*.$$

Dirichlet Case ($p = q = 0$)



- Case A : $\Omega_1 = (0,1)$, $\Omega_2 = (1,3)$, $\gamma = 0$
- Case B : $\Omega_1 = (0,2.9)$, $\Omega_2 = (2.9,3)$, $\gamma = 10$

Optimized case, $p = q$

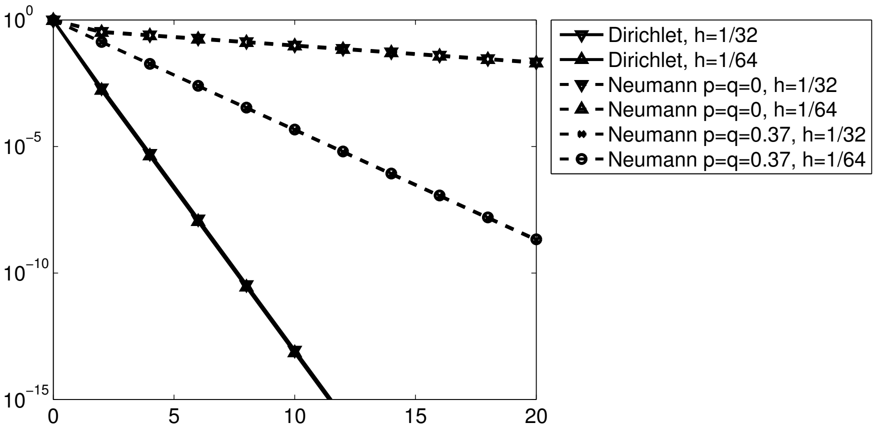


- Case A : $\Omega_1 = (0, 1)$, $\Omega_2 = (1, 3)$, $\gamma = 0$
- Case B : $\Omega_1 = (0, 2.9)$, $\Omega_2 = (2.9, 3)$, $\gamma = 10$
- Convergence for all frequencies

Numerical Example 1

- Governing PDE : $u_t = u_{xx}$ in $(x,t) \in (0,1) \times (0,3)$
- Discretization : Crank–Nicolson with $h = 1/32$ and $h = 1/64$
- Dirichlet or Neumann boundary conditions in space
- Two temporal subdomains : $\Omega_1 = (0,1)$, $\Omega_2 = (1,3)$

Numerical Example 1



- Mesh independent convergence
- Optimized conditions beneficial for Neumann case

Analysis by Energy Estimates

Convergence Analysis

- Diagonalization
  - \+ Explicit formula for contraction rate
  - \+ With or without overlap
  - – Assumes $A = A^T$
- Energy estimates
  - Integration by parts
  - \+ General setting ($A \neq A^T$, boundary control, etc.)
  - \+ Multiple subdomains
  - – No overlap

Motivation : Optimized Schwarz for Laplace Equation

- Consider solving the 1D Laplace equation using optimized Schwarz on two subdomains :

$$-\frac{d^2u_i^k}{dx^2} = 0 \qquad \text{on } \Omega_i,\ i = 1, 2,$$
$$\frac{du_i^k}{dn_i} + pu_i^k = -\frac{du_{3-i}^{k-1}}{dn_{3-i}} + pu_{3-i}^{k-1} \qquad \text{on } \Gamma,$$
$$u_1^k(0) = u_2^k(1) = 0.$$

- Idea of energy estimate : multiply DE on Ω_1 by u_1^k and integrate by parts :

$$0 = -\int_0^{x_\Gamma} u_1^k \frac{d^2u_1^k}{dx^2}\,dx = \int_0^{x_\Gamma} \left(\frac{du_1^k}{dx}\right)^2 dx - u_1^k(x_\Gamma)\frac{du_1^k}{dx}(x_\Gamma)$$

Note that the cross term $u_1^k(x_\Gamma)\dfrac{du_1^k}{dx}(x_\Gamma)$ can be written as

$$\begin{aligned} u_1^k(x_\Gamma)\frac{du_1^k}{dx}(x_\Gamma) &= \frac{1}{4p}\left(\frac{du_1^k}{dx}(x_\Gamma)+pu_1^k(x_\Gamma)\right)^2 - \frac{1}{4p}\left(-\frac{du_1^k}{dx}(x_\Gamma)+pu_1^k(x_\Gamma)\right)^2 \\ &= \underbrace{\frac{1}{4p}\left(\frac{du_2^{k-1}}{dx}(x_\Gamma)+pu_2^{k-1}(x_\Gamma)\right)^2}_{\text{input trace}} - \underbrace{\frac{1}{4p}\left(-\frac{du_1^k}{dx}(x_\Gamma)+pu_1^k(x_\Gamma)\right)^2}_{\text{output trace}}. \end{aligned}$$

Doing this also for Ω_2, we obtain for the two subdomains

$$\begin{aligned} \frac{1}{4p}\left(\frac{du_2^{k-1}}{dx}(x_\Gamma)+pu_2^{k-1}(x_\Gamma)\right)^2 &= \int_0^{x_\Gamma}\left(\frac{du_1^k}{dx}\right)^2 dx + \frac{1}{4p}\left(-\frac{du_1^k}{dx}(x_\Gamma)+pu_1^k(x_\Gamma)\right)^2 \\ \underbrace{\frac{1}{4p}\left(-\frac{du_1^{k-1}}{dx}(x_\Gamma)+pu_1^{k-1}(x_\Gamma)\right)^2}_{\text{input traces}} &= \underbrace{\int_{x_\Gamma}^1\left(\frac{du_2^k}{dx}\right)^2 dx}_{\text{internal energy}} + \underbrace{\frac{1}{4p}\left(\frac{du_2^k}{dx}(x_\Gamma)+pu_2^k(x_\Gamma)\right)^2}_{\text{output traces}}. \end{aligned}$$

- Summing over both subdomains, we have

$$R^{k-1} - R^k = E^k,$$

where

$$E^k = \sum_{i=1}^{2} \int_{\Omega_i} \left(\frac{du_i^k}{dx} \right)^2 dx, \quad R^k = \frac{1}{4p} \left[\left(-\frac{du_1^k}{dx}(x_\Gamma) + p u_1^k(x_\Gamma) \right)^2 + \left(\frac{du_2^k}{dx}(x_\Gamma) + p u_2^k(x_\Gamma) \right)^2 \right].$$

- Summing over k leads to a bounded telescoping sum :

$$\sum_{k=0}^{K} E^k = R^0 - R^K \leq R^0 < \infty \qquad \text{for all } K,$$

so $E^k \to 0$ as $k \to \infty$. Together with $u_1(0) = u_2(1) = 0$, this implies $u_i^k \to 0$ as $k \to \infty$.
- Argument also works for multiple subdomains, 2D or 3D problems, etc.

Energy Estimates for Control

- For the control problem, we have transmission conditions of the form

$$\lambda + py = h, \qquad y - q\lambda = g.$$

- To mimic the elliptic case, we need to
  1. Multiply equations and integrate by parts,
  2. Ensure the internal energy term has the right sign,
  3. Write boundary terms as a difference of transmission traces, i.e., as

$$c_1\|\lambda + py\|^2 - c_2\|y - q\lambda\|^2.$$

Energy Estimates for Control

- By linearity, subtract the exact solution to obtain the error equations

$$\dot{y} + Ay = \nu^{-1}\lambda, \qquad \dot{\lambda} - A^T\lambda = y.$$

- We no longer assume that A is symmetric, but we want its symmetric part $H = \frac{1}{2}(A + A^T)$ to be positive semi-definite
- Consider the change of variables

$$\begin{pmatrix} z \\ \mu \end{pmatrix} = \underbrace{\begin{bmatrix} 1 & r \\ -s & 1 \end{bmatrix}}_{=B} \begin{pmatrix} y \\ \lambda \end{pmatrix} \quad \Longleftrightarrow \quad \begin{pmatrix} y \\ \lambda \end{pmatrix} = \frac{1}{1+rs} \begin{bmatrix} 1 & -r \\ s & 1 \end{bmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix},$$

where $r, s > 0$ are to be chosen as a function of p and q.

Energy Estimates : Necessary Conditions

If we multiply the transformed system by (μ^T, z^T) and integrate, we obtain on Ω_1

$$0 = \mu(\alpha)^T z(\alpha) - \mu(0)^T z(0) + \frac{1}{1+rs}\int_0^\alpha \mu^T (r^2 - 2rH - \nu^{-1})\mu + \frac{1}{1+rs}\int_0^\alpha z^T (s^2\nu^{-1} - 2sH - 1)z$$

with $H = \frac{1}{2}(A + A^T) \geq 0$.

Energy Estimates : Necessary Conditions

If we multiply the transformed system by (μ^T, z^T) and integrate, we obtain on Ω_1

$$0 = \mu(\alpha)^T z(\alpha) - r\|\lambda(0)\|^2 + \frac{1}{1+rs}\int_0^\alpha \mu^T(r^2 - 2rH - \nu^{-1})\mu + \frac{1}{1+rs}\int_0^\alpha z^T(s^2\nu^{-1} - 2sH - 1)z$$

with $H = \frac{1}{2}(A + A^T) \geq 0$. We want to choose r and s such that

- $r, s > 0$,
- $r^2 - 2rH - \nu^{-1}$ and $s^2\nu^{-1} - 2sH - 1$ are *negative definite*,
- $\mu^T z = (\lambda - sy)^T(y + r\lambda) = c_1\|\lambda + py\|^2 - c_2\|y - q\lambda\|^2$.

Energy Estimates

With this choice, we obtain for the kth iteration

$$c_1\|\lambda_1^k(\alpha)+py_1^k(\alpha)\|^2-c_2\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2=r\|\lambda_1^k(0)\|^2+\frac{1}{1+rs}\int_0^\alpha\langle\text{pos. terms}\rangle$$

Energy Estimates

With this choice, we obtain for the kth iteration

$$c_1\|\lambda_2^{k-1}(\alpha)+py_2^{k-1}(\alpha)\|^2-c_2\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2=r\|\lambda_1^k(0)\|^2+\frac{1}{1+rs}\int_0^\alpha\langle\text{pos. terms}\rangle$$

Energy Estimates

With this choice, we obtain for the kth iteration

$$c_1\|\lambda_2^{k-1}(\alpha)+py_2^{k-1}(\alpha)\|^2-c_2\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2=r\|\lambda_1^k(0)\|^2+\frac{1}{1+rs}\int_0^\alpha\langle\text{pos. terms}\rangle$$

Similarly, for Ω_2, we have

$$-\hat{c}_1\|\lambda_2^k(\alpha)+py_2^k(\alpha)\|^2+\hat{c}_2\|y_2^k(\alpha)-q\lambda_2^k(\alpha)\|^2=\hat{s}\|y_2^k(T)\|^2+\frac{1}{1+\hat{r}\hat{s}}\int_\alpha^T\langle\text{pos. terms}\rangle$$

Energy Estimates

With this choice, we obtain for the kth iteration

$$c_1\|\lambda_2^{k-1}(\alpha)+py_2^{k-1}(\alpha)\|^2-c_2\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2=r\|\lambda_1^k(0)\|^2+\frac{1}{1+rs}\int_0^\alpha\langle\text{pos. terms}\rangle$$

Similarly, for Ω_2, we have

$$-\hat{c}_1\|\lambda_2^k(\alpha)+py_2^k(\alpha)\|^2+\hat{c}_2\|y_1^{k-1}(\alpha)-q\lambda_1^{k-1}(\alpha)\|^2=\hat{s}\|y_2^k(T)\|^2+\frac{1}{1+\hat{r}\hat{s}}\int_\alpha^T\langle\text{pos. terms}\rangle$$

Energy Estimates

With this choice, we obtain for the kth iteration

$$c_1\|\lambda_2^{k-1}(\alpha)+py_2^{k-1}(\alpha)\|^2-c_2\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2=r\|\lambda_1^k(0)\|^2+\frac{1}{1+rs}\int_0^\alpha\langle\text{pos. terms}\rangle$$

Similarly, for Ω_2, we have

$$-\hat{c}_1\|\lambda_2^k(\alpha)+py_2^k(\alpha)\|^2+\hat{c}_2\|y_1^{k-1}(\alpha)-q\lambda_1^{k-1}(\alpha)\|^2=\hat{s}\|y_2^k(T)\|^2+\frac{1}{1+\hat{r}\hat{s}}\int_\alpha^T\langle\text{pos. terms}\rangle$$

Thus, we have the two-step convergence estimate

$$\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|^2\leq\frac{c_1\hat{c}_2}{c_2\hat{c}_1}\|y_1^{k-2}(\alpha)-q\lambda_1^{k-2}(\alpha)\|^2,$$

so $\|y_1^k(\alpha)-q\lambda_1^k(\alpha)\|\to 0$ if $\frac{c_1\hat{c}_2}{c_2\hat{c}_1}<1$, and likewise for $\|\lambda_2^k(\alpha)+py_2^k(\alpha)\|$. This then implies convergence of μ and z inside the subdomains.

Energy Estimates : Satisfying the Constraints

$$\|y_1^k(\alpha) - q\lambda_1^k(\alpha)\|^2 \leq \frac{c_1\hat{c}_2}{c_2\hat{c}_1}\|y_1^{k-2}(\alpha) - q\lambda_1^{k-2}(\alpha)\|^2.$$

- Our constraints :
  1. $r, s > 0$,
  2. $r^2 - 2rH - \nu^{-1}$ and $s^2\nu^{-1} - 2sH - 1$ must be negative definite,
  3. $\mu^T z = (\lambda - sy)^T(y + r\lambda) = c_1|\lambda + py|^2 - c_2|y - q\lambda|^2$.
- There is only one equation (but two unknowns) per subdomain, so we can use the other unknown to mimimize $(c_1\hat{c}_2)/(c_2\hat{c}_1)$.
- The other constraints give bounds on r and s (and hence p and q).

Energy Estimates

- **Constraint 2 :** we need $r^2 - 2rH - \nu^{-1}$ and $s^2\nu^{-1} - 2sH - 1$ to be negative definite.
- If $d_i > 0$ are the eigenvalues of H, then $r^2 - 2rd_i - \nu^{-1}$ are eigenvalues of $r^2 - 2rH - \nu^{-1}$.
- So we need

$$r^2 - 2rd_i - \nu^{-1} < 0 \iff 0 < r < d_i + \sqrt{d_i^2 + \nu^{-1}} \qquad \text{for all } i.$$

  We therefore need

$$0 < r < r_{\max} := d_{\min} + \sqrt{d_{\min}^2 + \nu^{-1}}.$$

- Similarly, we need $0 < s < s_{\max}$, where

$$0 < s < \nu d_{\min} + \sqrt{\nu^2 d_{\min}^2 + \nu} := s_{\max} \qquad \text{for all } i.$$

- Note that $r_{\max} s_{\max} > 1$!

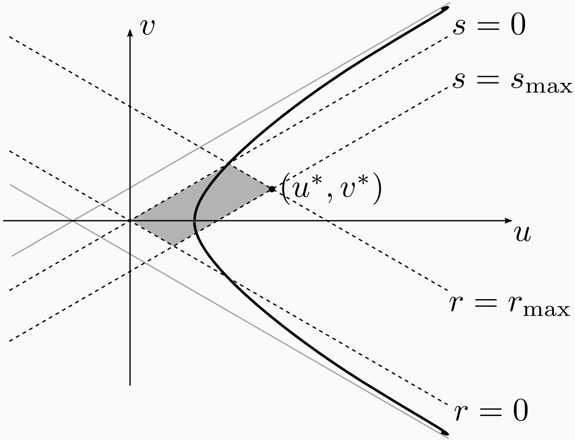
- **Constraint 3 :** we need $\mu^T z = (\lambda - sy)^T(y + r\lambda) = c_1\|\lambda + py\|^2 - c_2\|y - q\lambda\|^2$.
- We consider the case of $0 < pq < 1$ (the other cases are simpler to analyze). Equating coefficeints for $y^T y$, $\lambda^T \lambda$ and $y^T \lambda$ gives

$$r = c_1 - c_2 q^2, \qquad s = c_2 - p^2 c_1, \qquad 1 - rs = 2c_1 p + 2c_2 q.$$

- If we let $c_1 = q(u + v)$ and $c_2 = p(u - v)$, we obtain

$$\begin{aligned} \frac{r}{q} &= (1 - pq)u + (1 + pq)v \ \in \ [0, r_{\max}/q], \\ \frac{s}{p} &= (1 - pq)u - (1 + pq)v \ \in \ [0, s_{\max}/p], \\ 1 - rs &= \boxed{1 - pq[(1 - pq)u^2 - (1 + pq)v^2] = 4pqu.} \end{aligned}$$

- The boxed equation corresponds to a hyperbola in the uv-plane!



- One can show that $r_{\max} s_{\max} > 1 \implies (u^*, v^*)$ lies to the right of the hyperbola

  $\implies$ Solutions exist for any choice of p and q, as long as $0 < pq < 1$!

Energy Estimates

- After some algebra, one can show that

$$c_1 = \frac{r + sq^2}{1 - p^2q^2}, \qquad c_2 = \frac{s + rp^2}{1 - p^2q^2},$$

and the hyperbola leads to a compatibility condition between r and s

$$(1 - pq)(1 - rs) = 2(pr + qs). \tag{$*$}$$

- This allows us to eliminate either r or s from c_1/c_2 to obtain

$$\frac{c_1}{c_2} = \left(\frac{q + r}{1 - pr}\right)^2 = \left(\frac{1 - qs}{p + s}\right)^2 \implies \rho = \left(\frac{c_1\hat{c}_2}{c_2\hat{c}_1}\right)^{1/2} = \frac{1 - qs}{p + s} \cdot \frac{1 - p\hat{r}}{q + \hat{r}}$$

- Minimize ρ subject to $pq < 1$, $0 \le r \le r_{\max}$, $0 \le \hat{s} \le s_{\max}$ and $(*)$!

Theorem : (Convergence for two subdomains) Let $\gamma = 0$ (no target state). If $p > 0$ and $q > 0$ are such that $pq < 1$, and assume that $H = \frac{1}{2}(A + A^T)$ is positive semidefinite with smallest eigenvalue $d_{\min} \geq 0$. Then Then the two-subdomain OSM converges with

$$\rho \leq \max\left\{q, \frac{1 - q\nu r_{\max}}{p + \nu r_{\max}}\right\} \cdot \max\left\{p, \frac{1 - p r_{\max}}{q + r_{\max}}\right\} < 1,$$

where $r_{\max} = d_{\min} + \sqrt{d_{\min}^2 + \nu^{-1}}$.

Theorem : (Optimal p and q for $d_{\min} = 0$) Under the same hypotheses as the previous theorem, the choice of

$$p = \nu^{-1/2}(\sqrt{2} - 1), \qquad q = \nu^{1/2}(\sqrt{2} - 1)$$

minimizes the contraction factor for $d_{\min} = 0$. The resulting contraction factor is

$$\rho = 3 - 2\sqrt{2} \approx 0.1716.$$

Energy Estimates : Multiple subdomains

- For multiple subdomains, one writes the relation in μ_j and z_j on each subdomain Ω_j :

$$0 = \mu_j^k(\alpha_j)^T z_j^k(\alpha_j) - \mu_j^k(\alpha_{j-1})^T z_j^k(\alpha_{j-1}) + \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (\mu_j^k)^T (r^2 - 2rH - \nu^{-1})\mu_j^k$$
$$+ \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (z_j^k)^T (s^2\nu^{-1} - 2sH - 1) z_j^k$$

$$\frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} \left[(\mu_j^k)^T M_1 \mu_j^k + (z_j^k)^T M_2 z_j^k\right] = c_1\|\lambda_j^k(\alpha_j) + p y_j^k(\alpha_j)\|^2 - c_2\|y_j^k(\alpha_j) - q\lambda_j^k(\alpha_j)\|^2$$
$$- c_1\|\lambda_j^k(\alpha_{j-1}) + p y_j^k(\alpha_{j-1})\|^2 + c_2\|y_j^k(\alpha_{j-1}) - q\lambda_j^k(\alpha_{j-1})\|^2$$

Energy Estimates : Multiple subdomains

UNIVERSITÉ LAVAL

- For multiple subdomains, one writes the relation in μ_j and z_j on each subdomain Ω_j :

$$0 = \mu_j^k(\alpha_j)^T z_j^k(\alpha_j) - \mu_j^k(\alpha_{j-1})^T z_j^k(\alpha_{j-1}) + \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (\mu_j^k)^T (r^2 - 2rH - \nu^{-1})\mu_j^k + \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (z_j^k)^T (s^2\nu^{-1} - 2sH - 1) z_j^k$$

$$\frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} \left[(\mu_j^k)^T M_1 \mu_j^k + (z_j^k)^T M_2 z_j^k\right] = c_1\|\lambda_{j+1}^{k-1}(\alpha_j) + p y_{j+1}^{k-1}(\alpha_j)\|^2 - c_2\|y_j^k(\alpha_j) - q\lambda_j^k(\alpha_j)\|^2 - c_1\|\lambda_j^k(\alpha_{j-1}) + p y_j^k(\alpha_{j-1})\|^2 + c_2\|y_{j-1}^{k-1}(\alpha_{j-1}) - q\lambda_{j-1}^{k-1}(\alpha_{j-}$$

Energy Estimates : Multiple subdomains

- For multiple subdomains, one writes the relation in μ_j and z_j on each subdomain Ω_j :

$$0 = \mu_j^k(\alpha_j)^T z_j^k(\alpha_j) - \mu_j^k(\alpha_{j-1})^T z_j^k(\alpha_{j-1}) + \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (\mu_j^k)^T(r^2 - 2rH - \nu^{-1})\mu_j^k$$
$$+ \frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} (z_j^k)^T(s^2\nu^{-1} - 2sH - 1)z_j^k$$

$$\frac{1}{1+rs}\int_{\alpha_{j-1}}^{\alpha_j} \left[(\mu_j^k)^T M_1 \mu_j^k + (z_j^k)^T M_2 z_j^k\right] = c_1\|\lambda_{j+1}^{k-1}(\alpha_j) + p y_{j+1}^{k-1}(\alpha_j)\|^2 - c_2\|y_j^k(\alpha_j) - q\lambda_j^k(\alpha_j)\|^2$$
$$- c_1\|\lambda_j^k(\alpha_{j-1}) + p y_j^k(\alpha_{j-1})\|^2 + c_2\|y_{j-1}^{k-1}(\alpha_{j-1}) - q\lambda_{j-1}^{k-1}(\alpha_{j-}$$

- Summing over all j, we obtain

$$E^k \leq R^{k-1} - R^k,$$

where E^k is the sum of the internal energies, and R^k is the sum of the kth Robin traces.

Energy Estimates : Multiple subdomains

- For multiple subdomains, one needs to choose the same c_1 and c_2 for all subdomains to get the telescoping argument to work
- Nonetheless, one can obtain a contraction estimate if one can find constants K_1 and K_2 such that

$$K_1 R^k \leq E^k \leq K_2 R^k$$

Energy Estimates : Multiple subdomains

- For multiple subdomains, one needs to choose the same c_1 and c_2 for all subdomains to get the telescoping argument to work
- Nonetheless, one can obtain a contraction estimate if one can find constants K_1 and K_2 such that

$$K_1 R^k \leq E^k \leq K_2 R^k$$

> **Theorem :** (Multiple subdomains) Let $\gamma = 0$ (no target state). Then there exists $p, q > 0$ such that $pq < 1$ and OSM with N subdomains converges.

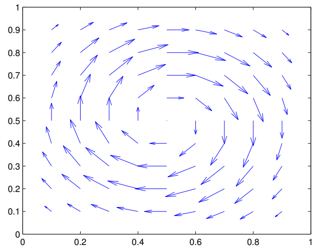
- A scaling argument shows that as H decreases, the contraction factor behaves in the worst case like $\rho \approx 1 - cH$, so a coarse grid is needed in general.
- Results also available when the control and/or observations only occur on a subset of Ω, see preprint (Gander & K., 2022)

Numerical Example 2

- 2D advection-diffusion equation on $\Omega = (0,1) \times (0,1)$

$$y_t - \nabla \cdot (\nabla y + \mathbf{b} y) = u$$

$$\mathbf{b} = \sin \pi x \sin \pi y \begin{pmatrix} y - 0.5 \\ 0.5 - x \end{pmatrix}$$



- $T = 3$, split into two subdomains at $\alpha = 1$
- Neumann conditions, no target state
- Upwind discretization, $h = 1/16$ and $h = 1/32$
- Transmission conditions : $p = q = \sqrt{2} - 1$

Numerical Example 2

- Predicted convergence factor : 0.1716

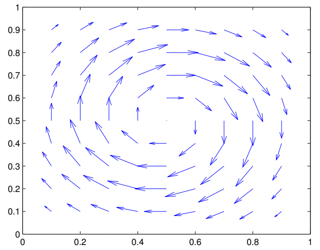
Its	$h = 1/16$		$h = 1/32$	
	Error	Ratio	Error	Ratio
1	9.9908e-001		9.9977e-001	
2	1.3762e-001	0.1378	1.3810e-001	0.1381
3	2.0115e-002	0.1462	2.0266e-002	0.1468
4	3.0901e-003	0.1536	3.1234e-003	0.1541
5	4.9302e-004	0.1595	4.9936e-004	0.1599
6	8.0785e-005	0.1639	8.1899e-005	0.1640
7	1.3474e-005	0.1668	1.3659e-005	0.1668
8	2.2729e-006	0.1687	2.3023e-006	0.1686
9	3.8599e-007	0.1698	3.9046e-007	0.1696
10	6.5653e-008	0.1701	6.6306e-008	0.1698

Numerical Example 3

- 2D advection-diffusion equation on $\Omega = (0,1) \times (0,1)$

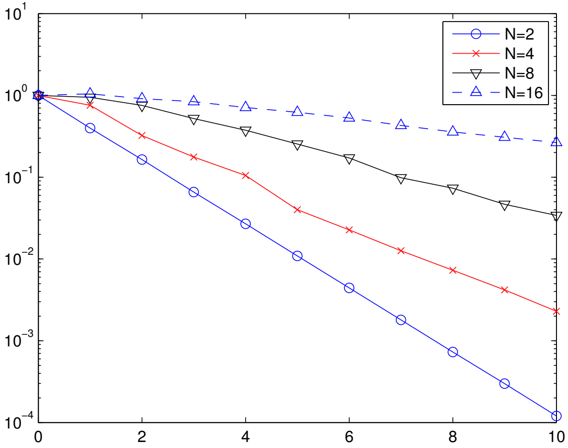
$$y_t - \nabla \cdot (\nabla y + \mathbf{b} y) = u$$

$$\mathbf{b} = \sin \pi x \sin \pi y \begin{pmatrix} y - 0.5 \\ 0.5 - x \end{pmatrix}$$



- $T = 4$, split into 2, 4, 8, 16 equal subdomains
- Neumann conditions, no target state
- Upwind discretization, $h = 1/16$
- Transmission conditions : $p = q = \sqrt{2} - 1$

Numerical Example 3



Numerical Example 3

We expect $\rho = 1 - CH$:

H	ρ	$1-\rho$	$H^{-1}(1-\rho)$
1/2	0.4063	0.5937	1.1864
1/4	0.5659	0.4341	1.7364
1/8	0.6653	0.3347	2.6776
1/16	0.8409	0.1591	2.5456

Observation and Control Over Subsets

- If the control and/or observation is only supported on a subset of Ω (i.e., if $B \neq I$ or $C \neq I$), then the ODE system becomes

$$\begin{bmatrix} \dot{y} \\ \dot{\lambda} \end{bmatrix} + \begin{bmatrix} A & -\nu^{-1} B B^T \\ -C^T C & -A^T \end{bmatrix} \begin{bmatrix} y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ -C^T \hat{y} \end{bmatrix}.$$

- Using the same calculation as before, we see that convergence requires

$$z^T (s^2 \nu^{-1} B B^T - 2sH - C^T C) z \leq 0, \qquad \mu^T (r^2 C^T C - 2rH - \nu^{-1} B B^T) \mu \leq 0$$

  for all z and μ.
- The condition on s is satisfied if $\ker(C) \cap \ker(H) \subset \ker(B^T)$ and if

$$0 \leq s \leq s^* = \nu \min_{B^T z \neq 0} \frac{z^T H z}{\|B^T z\|^2} + \sqrt{\left(\frac{z^T H z}{\|B^T z\|^2} \right)^2 + \frac{\|Cz\|^2}{\nu \|B^T z\|^2}}$$

Theorem : Let $\gamma = 0$ (no target state). Suppose that

$$\ker(C) \cap \ker(H) = \ker(B^T) \cap \ker(H) = \{0\}.$$

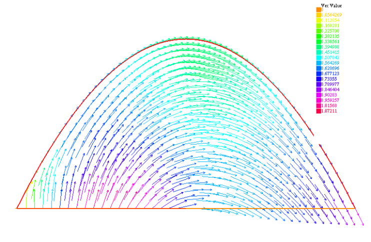
Then there exist $p, q > 0$ such that OSM with N subdomains converges.

- A good choice of s (and similarly for r) is given by twice the smallest eigenvalue of the generalized eigenvalue problem

$$B^T H B v = \lambda (B^T B)^2 v.$$

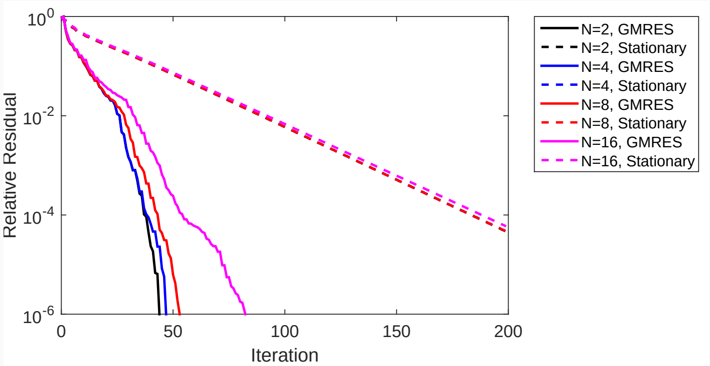
Numerical Example 4

- 2D advection-diffusion equation
- Flow field obtained by Stokes equation
- Finite volume method as in Bermúdez et al (1998)



- Source (control) at centre of domain, observation at one point on boundary
- 736 dof in space, 64 time steps
- $T = 32$, split into 2, 4, 8, 16 equal subdomains
- Transmission conditions : $p = q = 0.8563$

Numerical Example 4



Conclusion

Summary

- Optimized Schwarz method for control :
  - Converges with or without overlap
  - Choose Robin parameters to optimize convergence
  - Analysis by diagonalization or energy estimates
  - Global communication needed for scalability, cf. ParaOpt (Gander, Kwok & Salomon SISC 2020)
- Ongoing work :
  - Control for transport and wave propagation problems (ALLOWAP project, with L. Halpern, B. Delourme and J. Salomon)
  - Preconditioning for local subproblems
  - Control constraints

References

M. J. Gander and F. Kwok.
Schwarz Methods for the Time-Parallel Solution of Parabolic Control Problems
Domain Decomposition Methods in Computational Science and Engineering XXII, pp.207-216, Springer-Verlag 2016.

F. Kwok.
On the Time-Domain Decomposition of Parabolic Optimal Control Problems
Domain Decomposition Methods in Computational Science and Engineering XXIII, pp.55-67, Springer-Verlag 2017.

M. J. Gander and F. Kwok.
Optimized Schwarz-in-Time Methods for Parabolic Control Problems
In preparation (2022).