Some recent results on time domain decomposition methods for PDE-constrained optimization

Numerical study of scalability for heat control problems

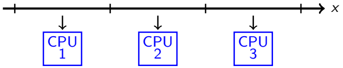
Liu-Di LU

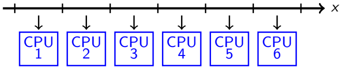
Section of Mathematics
University of Geneva

Milano, June 26th, 2025

1D Poisson equation

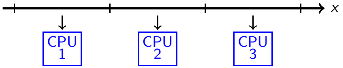
Example: $-\partial_{xx} u = 0$, $u(0) = u(1) = 0$, with initial guess $u^0 = 1$ and parallel Schwarz algorithm.

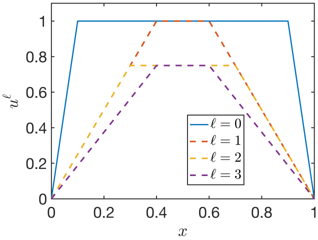


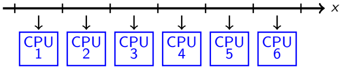


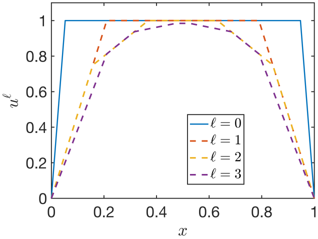
1D Poisson equation

Example: $-\partial_{xx} u = 0$, $u(0) = u(1) = 0$, with initial guess $u^0 = 1$ and parallel Schwarz algorithm.

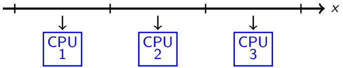


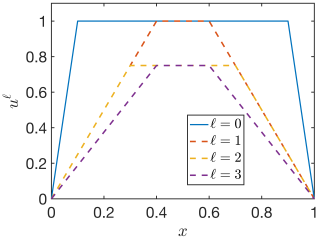


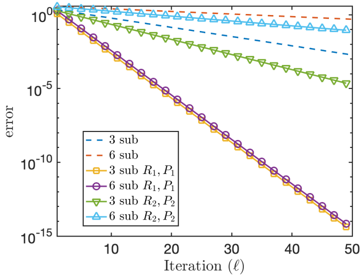


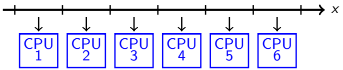


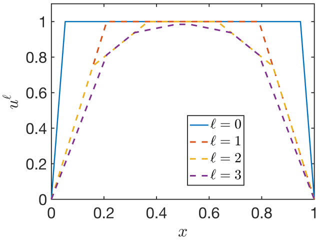
Example: $-\partial_{xx} u = 0$, $u(0) = u(1) = 0$, with initial guess $u^0 = 1$ and parallel Schwarz algorithm.











PDE-constrained optimization

Optimization problem: For $\hat{y} \in L^2(Q)$, $\gamma \geq 0$, $\nu > 0$ and $\Omega \subset \mathbb{R}^n$, minimize the cost functional

$$J(y,u) := \frac{1}{2}\|y - \hat{y}\|^2_{L^2(Q)} + \frac{\gamma}{2}\|y(T) - \hat{y}(T)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)},$$

subject to

$$\partial_t y - \Delta y = u \ \text{ in } Q := (0,T)\times\Omega, \quad y = 0 \ \text{ on } \Sigma := (0,T)\times\partial\Omega, \quad y = y_0 \ \text{ on } \Sigma_0 := \{0\}\times\Omega.$$

PDE-constrained optimization

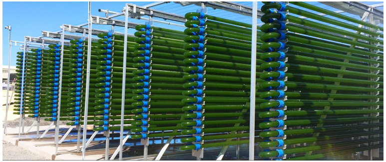
Optimization problem: For $\hat{y} \in L^2(Q)$, $\gamma \geq 0$, $\nu > 0$ and $\Omega \subset \mathbb{R}^n$, minimize the cost functional

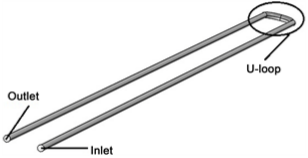
$$J(y,u) := \frac{1}{2}\|y - \hat{y}\|^2_{L^2(Q)} + \frac{\gamma}{2}\|y(T) - \hat{y}(T)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)},$$

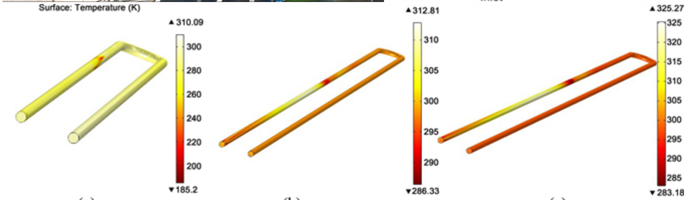
subject to

$$\partial_t y - \Delta y = u \ \text{ in } Q := (0,T)\times\Omega, \quad y = 0 \ \text{ on } \Sigma := (0,T)\times\partial\Omega, \quad y = y_0 \ \text{ on } \Sigma_0 := \{0\}\times\Omega.$$

Example: control temperature in photobioreactors.







(a) (b) (c)

PDE-constrained optimization

Optimization problem: For $\hat{y} \in L^2(Q)$, $\gamma \geq 0$, $\nu > 0$ and $\Omega \subset \mathbb{R}^n$, minimize the cost functional

$$J(y,u) := \frac{1}{2}\|y - \hat{y}\|^2_{L^2(Q)} + \frac{\gamma}{2}\|y(T) - \hat{y}(T)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)},$$

subject to

$$\partial_t y - \Delta y = u \text{ in } Q := (0,T)\times\Omega, \quad y = 0 \text{ on } \Sigma := (0,T)\times\partial\Omega, \quad y = y_0 \text{ on } \Sigma_0 := \{0\}\times\Omega.$$

First-order optimality system (forward-backward):

$$\begin{aligned} \partial_t y - \Delta y &= \nu^{-1}\lambda && \text{in } Q, & \partial_t \lambda + \Delta\lambda &= y - \hat{y} && \text{in } Q,\\ y &= 0 && \text{in } \Sigma, & \lambda &= 0 && \text{in } \Sigma,\\ y &= y_0 && \text{in } \Sigma_0, & \lambda &= -\gamma(y - \hat{y}) && \text{in } \Sigma_T := \{T\}\times\Omega, \end{aligned}$$

Optimization problem: For $\hat{y} \in L^2(Q)$, $\gamma \geq 0$, $\nu > 0$ and $\Omega \subset \mathbb{R}^n$, minimize the cost functional

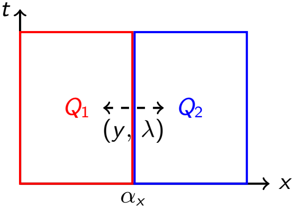
$$J(y,u) := \frac{1}{2}\|y - \hat{y}\|^2_{L^2(Q)} + \frac{\gamma}{2}\|y(T) - \hat{y}(T)\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)},$$

subject to

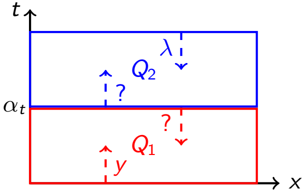
$$\partial_t y - \Delta y = u \ \text{ in } Q := (0,T) \times \Omega, \quad y = 0 \ \text{ on } \Sigma := (0,T) \times \partial\Omega, \quad y = y_0 \ \text{ on } \Sigma_0 := \{0\} \times \Omega.$$

First-order optimality system (forward-backward):

$$\begin{aligned} \partial_t y - \Delta y &= \nu^{-1}\lambda && \text{in } Q, & \partial_t \lambda + \Delta \lambda &= y - \hat{y} && \text{in } Q,\\ y &= 0 && \text{in } \Sigma, & \lambda &= 0 && \text{in } \Sigma,\\ y &= y_0 && \text{in } \Sigma_0, & \lambda &= -\gamma(y - \hat{y}) && \text{in } \Sigma_T := \{T\} \times \Omega, \end{aligned}$$

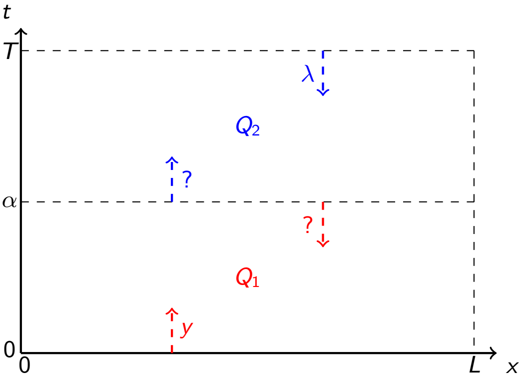


Waveform iterations



Time domain decomposition

Example: Control heat distribution w.r.t. a target $\hat{y}$.

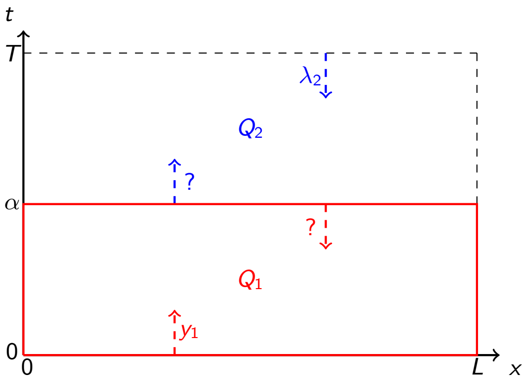


Subdomains: $Q_1 = (0, L) \times (0, \alpha)$ and $Q_2 = (0, L) \times (\alpha, T)$

$$\begin{aligned} \partial_t y - \partial_{xx} y &= \nu^{-1}\lambda, & \partial_t \lambda + \partial_{xx}\lambda &= y - \hat{y}, \\ y(0,t) &= 0, & \lambda(0,t) &= 0, \\ y(L,t) &= 0, & \lambda(L,t) &= 0, \\ y(x,0) &= y_0(x), & \lambda(x,T) &= 0. \end{aligned}$$

Idea of time domain decomposition

Example: Control heat distribution w.r.t. a target $\hat{y}$.

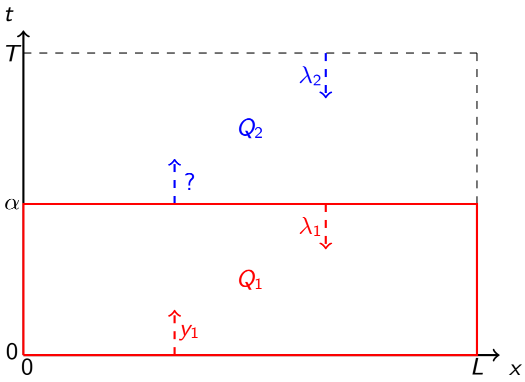


Subdomain: $Q_1 = (0, L) \times (0, \alpha)$

$$\begin{aligned} \partial_t y_1^\ell - \partial_{xx} y_1^\ell &= \nu^{-1}\lambda_1^\ell, \\ y_1^\ell(0,t) &= 0, \\ y_1^\ell(L,t) &= 0, \\ y_1^\ell(x,0) &= y_0(x), \end{aligned}$$

$$\begin{aligned} \partial_t \lambda_1^\ell + \partial_{xx} \lambda_1^\ell &= y_1^\ell - \hat{y}_1, \\ \lambda_1^\ell(0,t) &= 0, \\ \lambda_1^\ell(L,t) &= 0, \end{aligned}$$

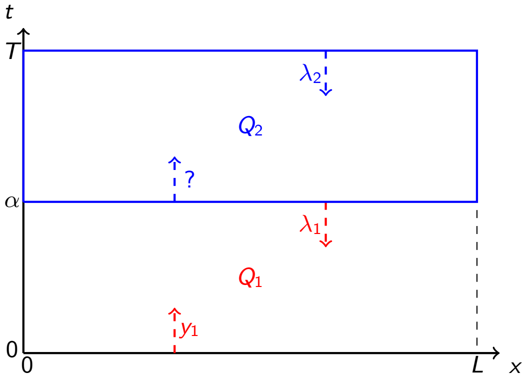
Example: Control heat distribution w.r.t. a target $\hat{y}$.



Subdomain: $Q_1 = (0, L) \times (0, \alpha)$

$$\begin{aligned} \partial_t y_1^\ell - \partial_{xx} y_1^\ell &= \nu^{-1}\lambda_1^\ell, & \partial_t \lambda_1^\ell + \partial_{xx}\lambda_1^\ell &= y_1^\ell - \hat{y}_1, \\ y_1^\ell(0,t) &= 0, & \lambda_1^\ell(0,t) &= 0, \\ y_1^\ell(L,t) &= 0, & \lambda_1^\ell(L,t) &= 0, \\ y_1^\ell(x,0) &= y_0(x), & \lambda_1^\ell(x,\alpha) &= \lambda_2^{\ell-1}(x,\alpha). \end{aligned}$$

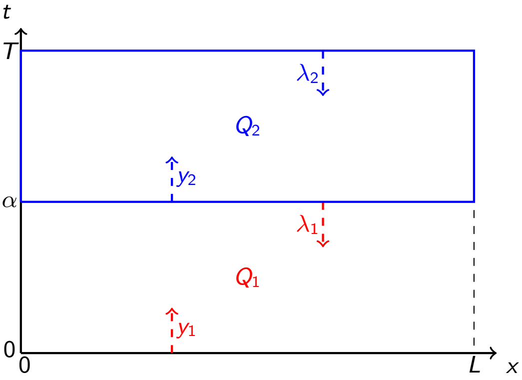
Example: Control heat distribution w.r.t. a target $\hat{y}$.



Subdomain: $Q_2 = (0, L) \times (\alpha, T)$

$$\begin{aligned} \partial_t y_2^\ell - \partial_{xx} y_2^\ell &= \nu^{-1} \lambda_2^\ell, \\ y_2^\ell(0, t) &= 0, \\ y_2^\ell(L, t) &= 0, \end{aligned} \qquad \begin{aligned} \partial_t \lambda_2^\ell + \partial_{xx} \lambda_2^\ell &= y_2^\ell - \hat{y}_2, \\ \lambda_2^\ell(0, t) &= 0, \\ \lambda_2^\ell(L, t) &= 0, \\ \lambda_2^\ell(x, T) &= 0. \end{aligned}$$

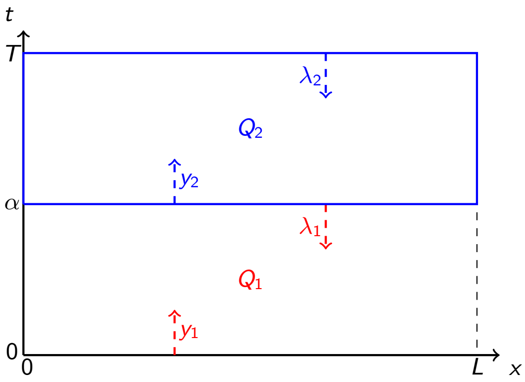
Example: Control heat distribution w.r.t. a target $\hat{y}$.



Subdomain: $Q_2 = (0, L) \times (\alpha, T)$

$$\begin{aligned}
\partial_t y_2^\ell - \partial_{xx} y_2^\ell &= \nu^{-1}\lambda_2^\ell, & \partial_t \lambda_2^\ell + \partial_{xx}\lambda_2^\ell &= y_2^\ell - \hat{y}_2,\\
y_2^\ell(0,t) &= 0, & \lambda_2^\ell(0,t) &= 0,\\
y_2^\ell(L,t) &= 0, & \lambda_2^\ell(L,t) &= 0,\\
y_2^\ell(x,\alpha) &= y_1^\ell(x,\alpha), & \lambda_2^\ell(x,T) &= 0.
\end{aligned}$$

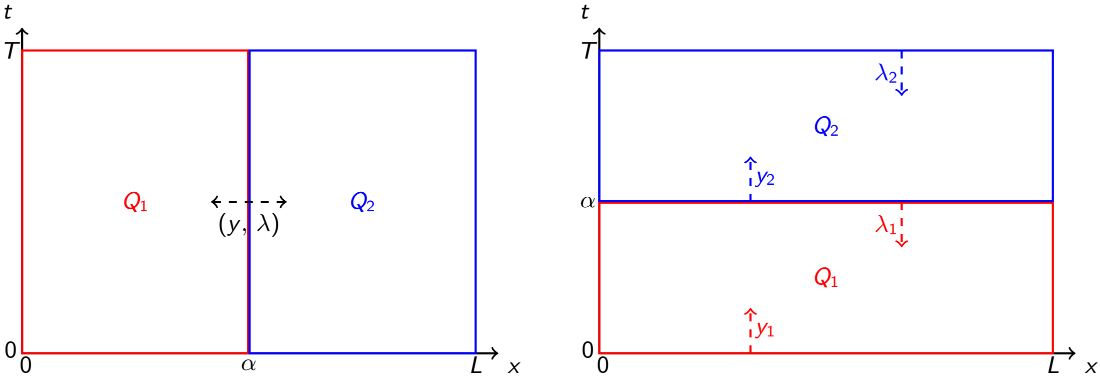
Example: Control heat distribution w.r.t. a target $\hat{y}$.



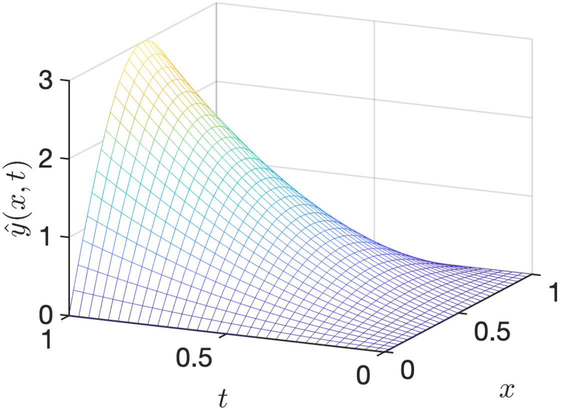
Subdomain: $Q_2 = (0, L) \times (\alpha, T)$

$$
\begin{aligned}
\partial_t y_2^\ell - \partial_{xx} y_2^\ell &= \nu^{-1}\lambda_2^\ell, & \partial_t \lambda_2^\ell + \partial_{xx}\lambda_2^\ell &= y_2^\ell - \hat{y}_2,\\
y_2^\ell(0,t) &= 0, & \lambda_2^\ell(0,t) &= 0,\\
y_2^\ell(L,t) &= 0, & \lambda_2^\ell(L,t) &= 0,\\
y_2^\ell(x,\alpha) &= y_1^{\ell-1}(x,\alpha), & \lambda_2^\ell(x,T) &= 0.
\end{aligned}
$$

Space vs time decomposition

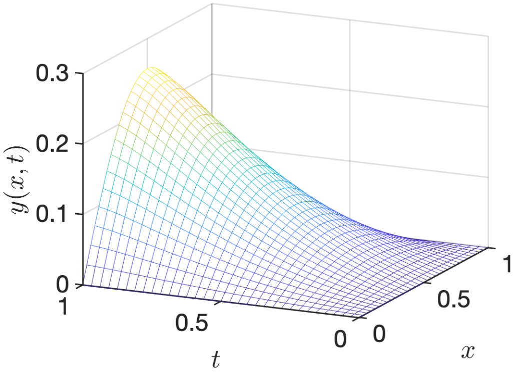


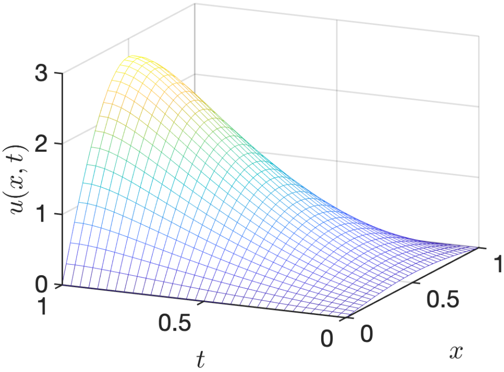
Numerical example: consider the target function $\hat{y}(x,t) = \sin(\pi x)(2t^2+2)$.



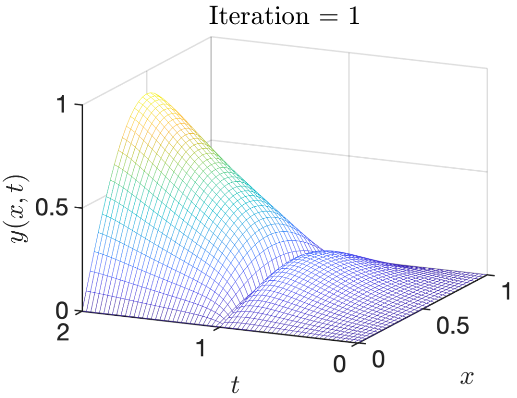
Test case

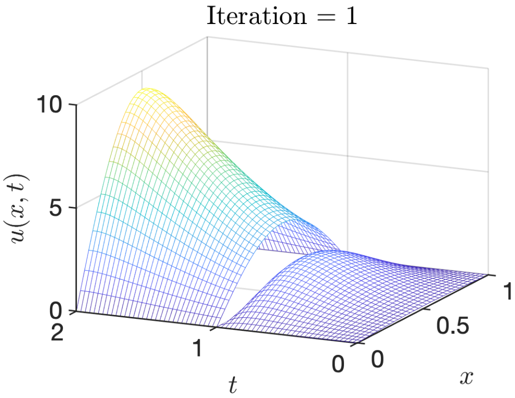
Numerical solutions: Crank-Nicolson with mesh size $h_t = h_x = \frac{1}{32}$ and penalization parameters: $\nu = 0.1$, $\gamma = 0.1$.



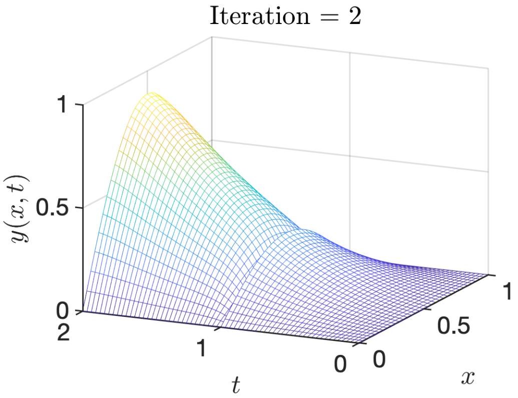


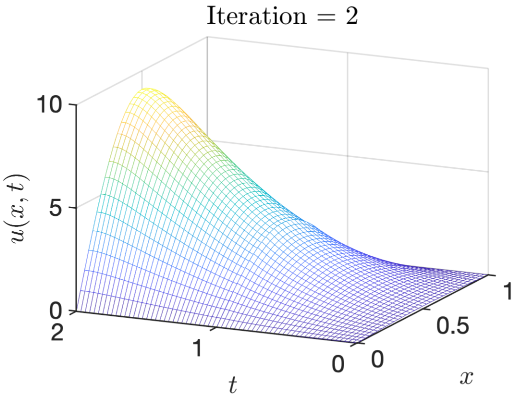
Two subdomains:



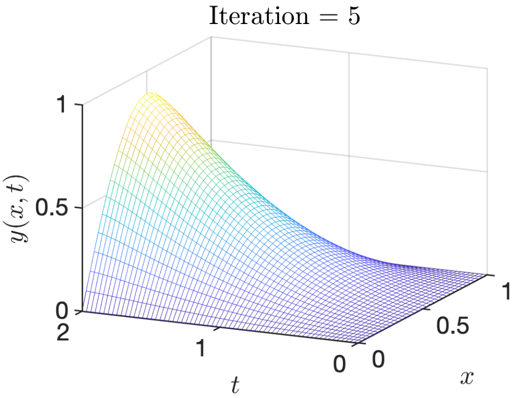


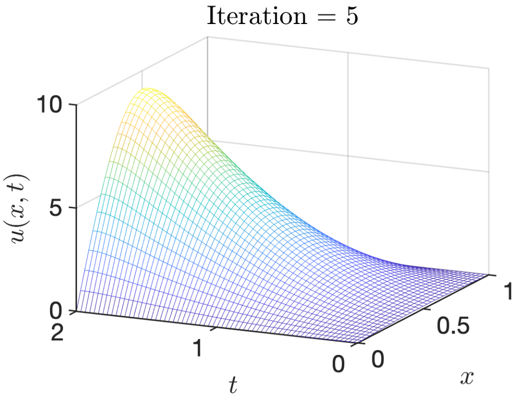
Two subdomains:



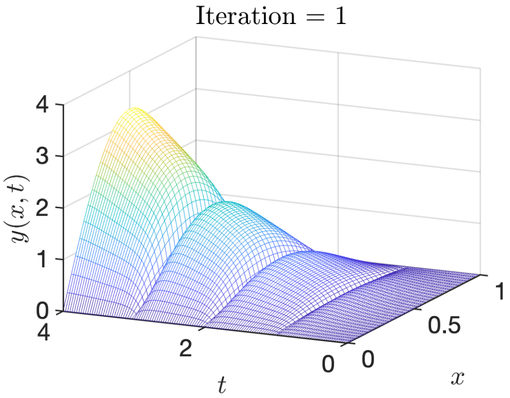


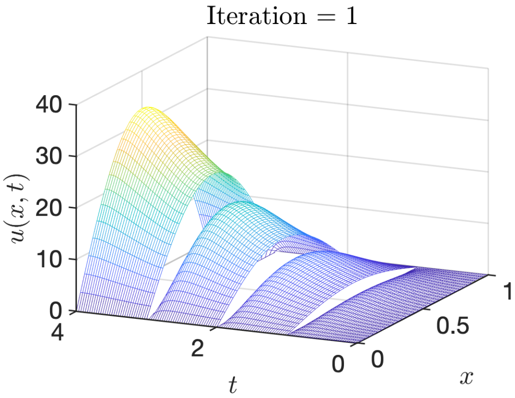
Two subdomains:



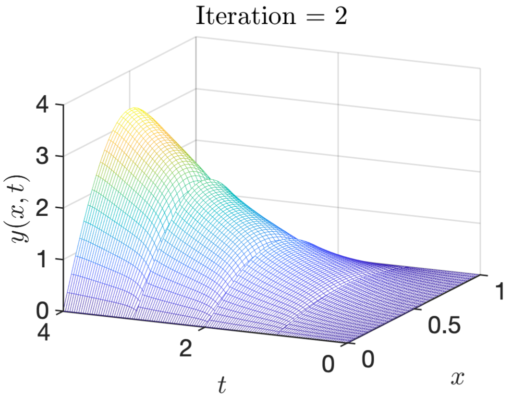


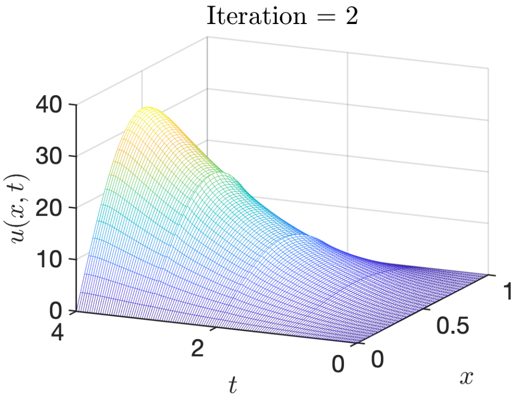
Four subdomains:



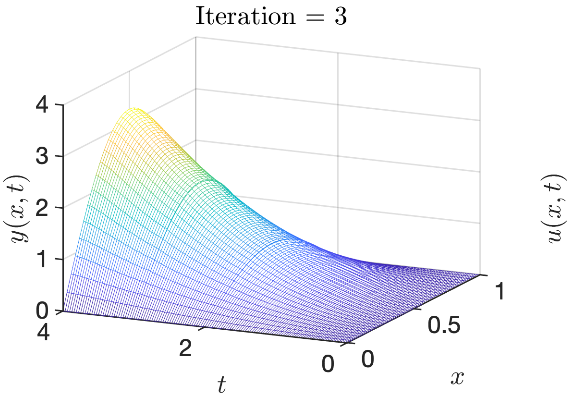


Four subdomains:

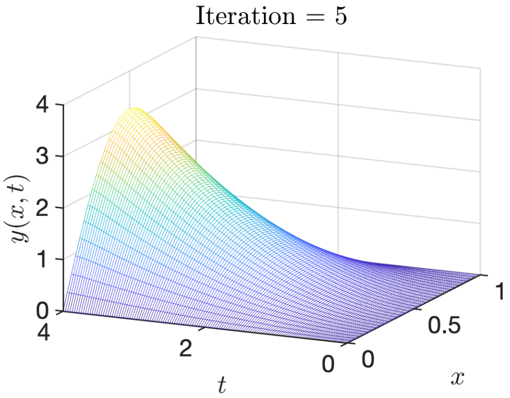


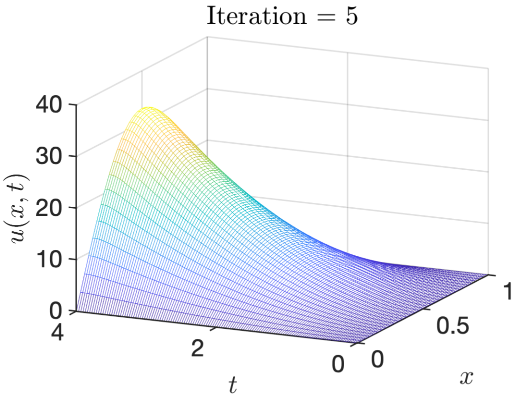


Four subdomains:

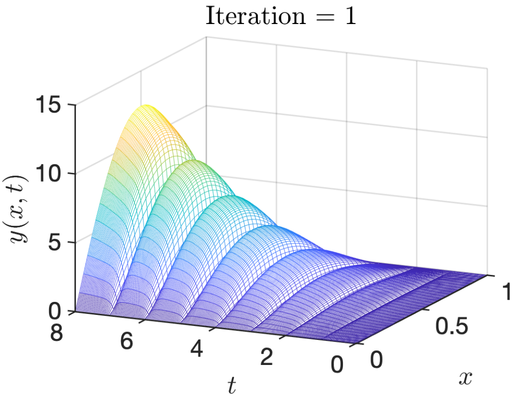


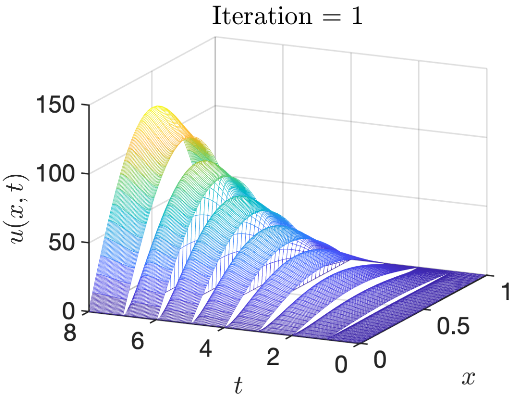
Four subdomains:



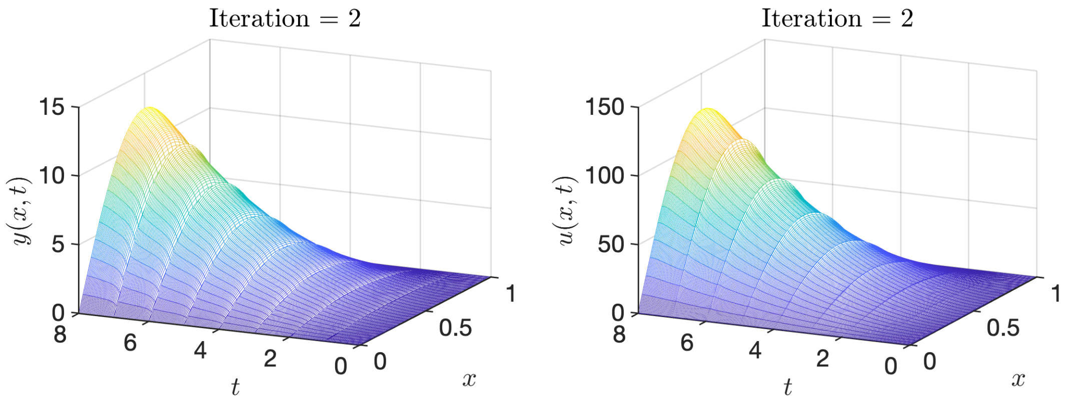


Eight subdomains:

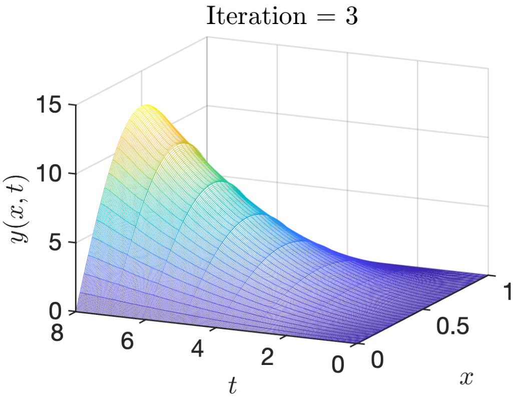


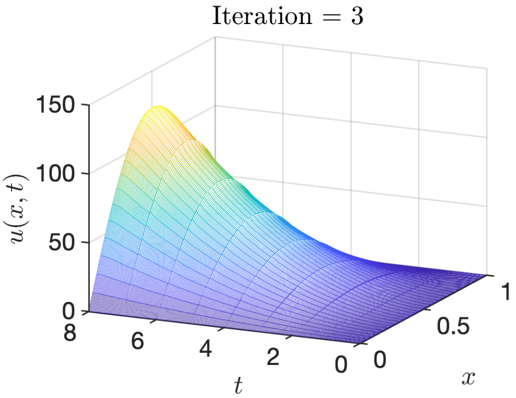


Eight subdomains:

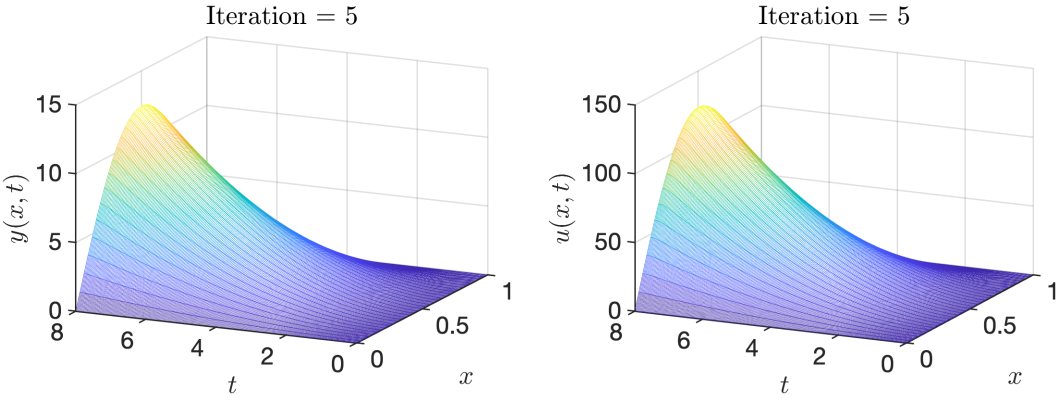


Eight subdomains:

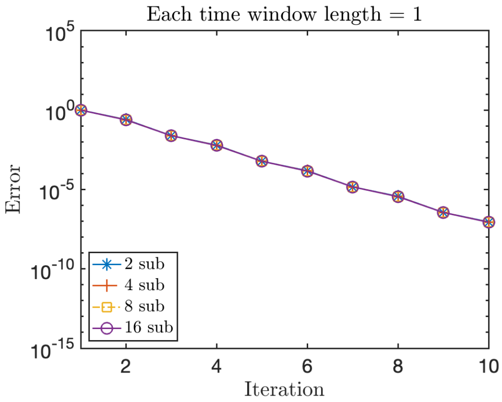




Eight subdomains:

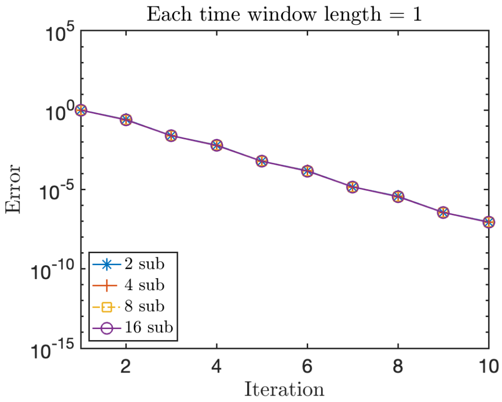


Error decay

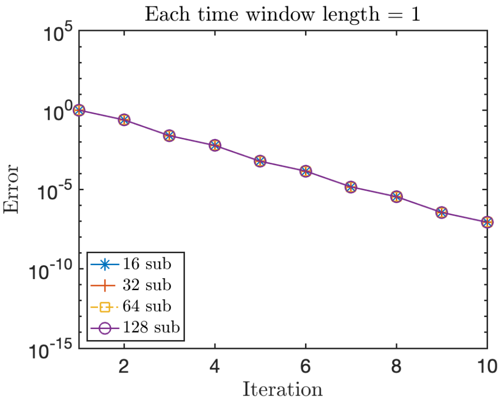


Each time window length = 1
10^5
10^0
10^{-5}
10^{-10}
10^{-15}
Error
2
4
6
8
10
Iteration
2 sub
4 sub
8 sub
16 sub

Error decay

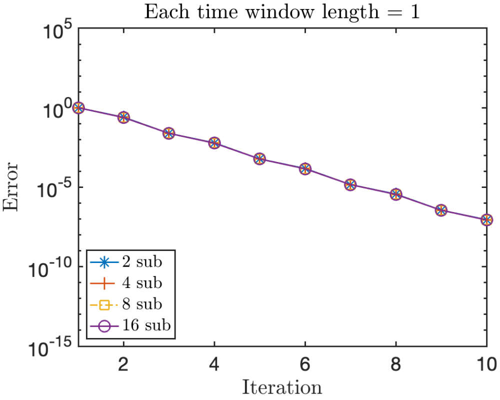


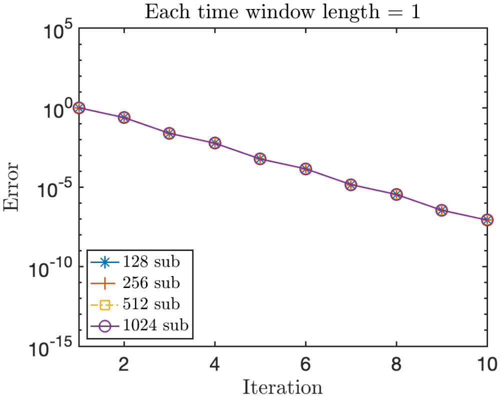
Each time window length = 1
Error
Iteration
2 sub
4 sub
8 sub
16 sub



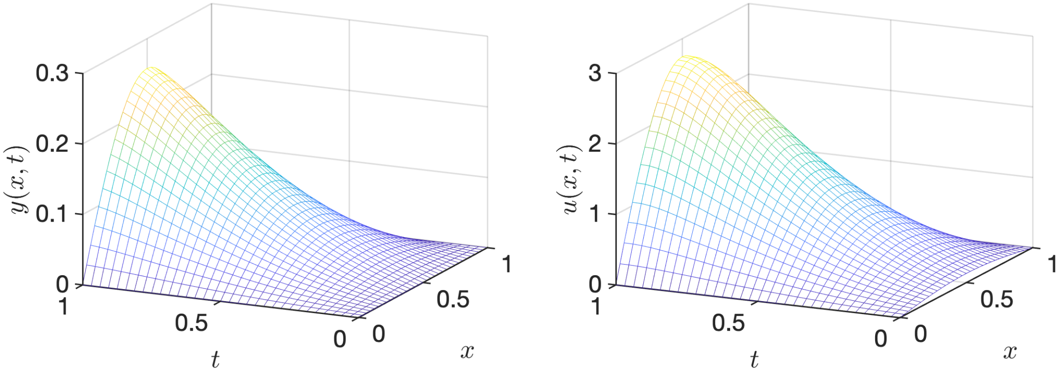
Each time window length = 1
Error
Iteration
16 sub
32 sub
64 sub
128 sub

Error decay

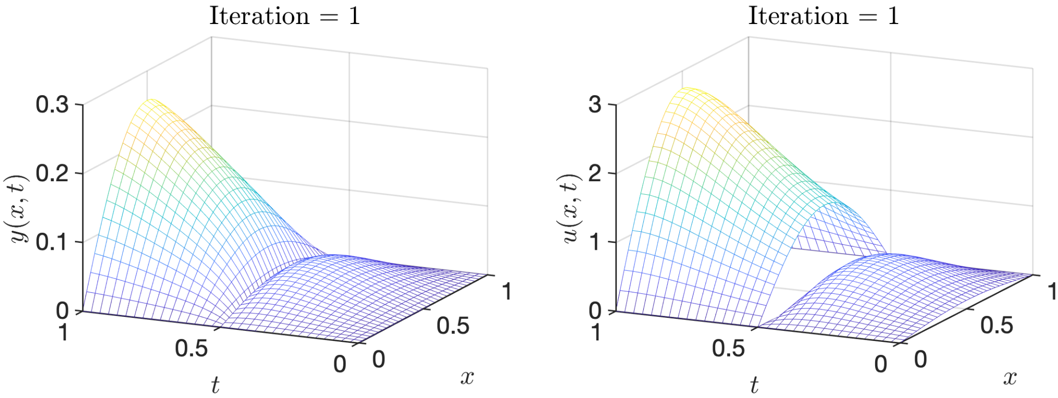

Each time window length = 1
Error
Iteration
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub


Each time window length = 1
Error
Iteration
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
128 sub
256 sub
512 sub
1024 sub

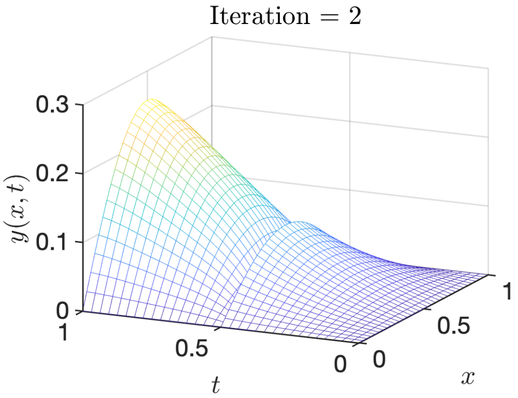
Numerical solutions: Crank-Nicolson with mesh size $h_t = h_x = \frac{1}{32}$ and penalization parameters: $\nu = 0.1$, $\gamma = 0.1$.

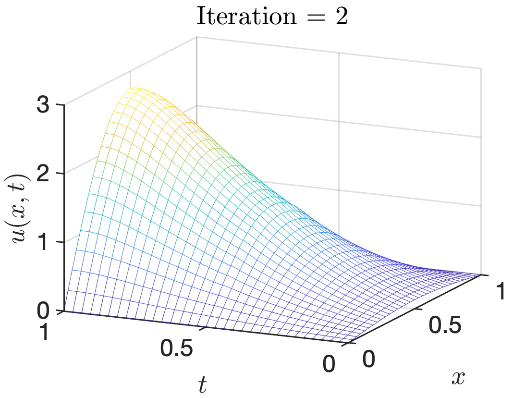


Two subdomains:

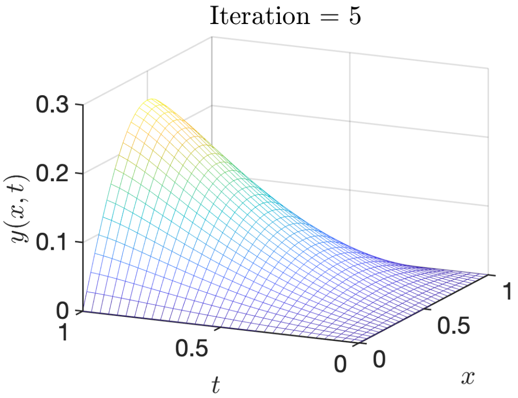


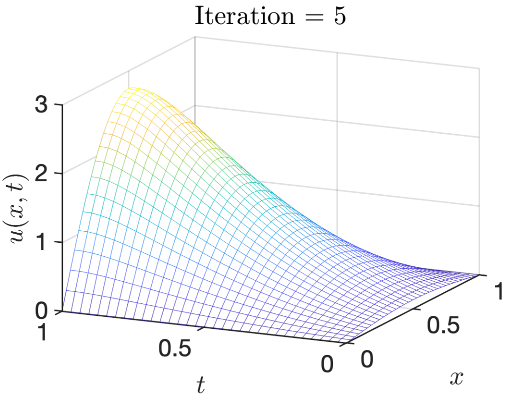
Two subdomains:



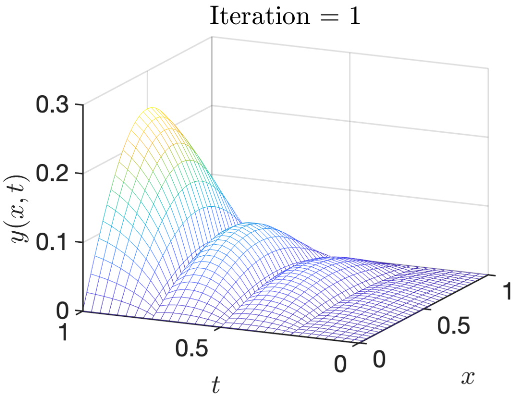


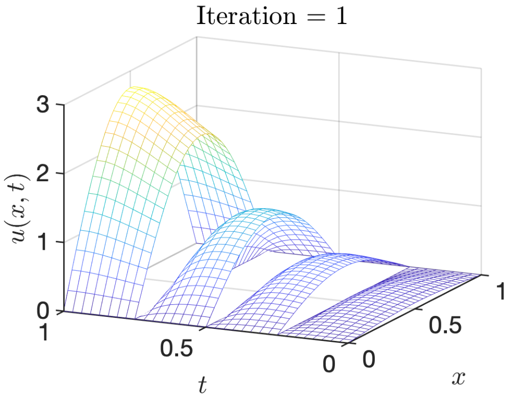
Two subdomains:



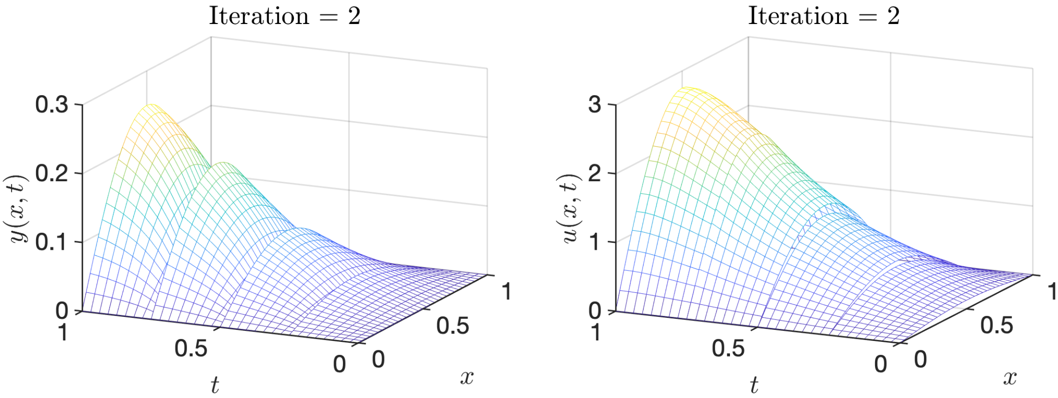


Four subdomains:

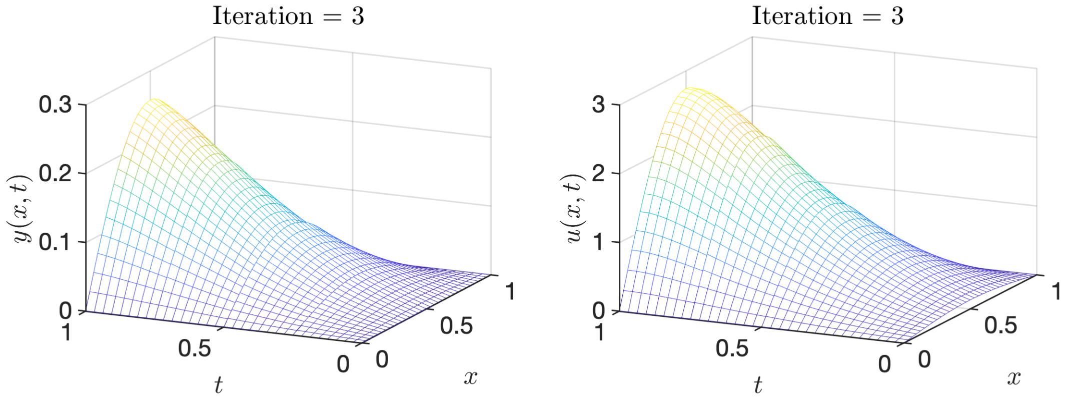




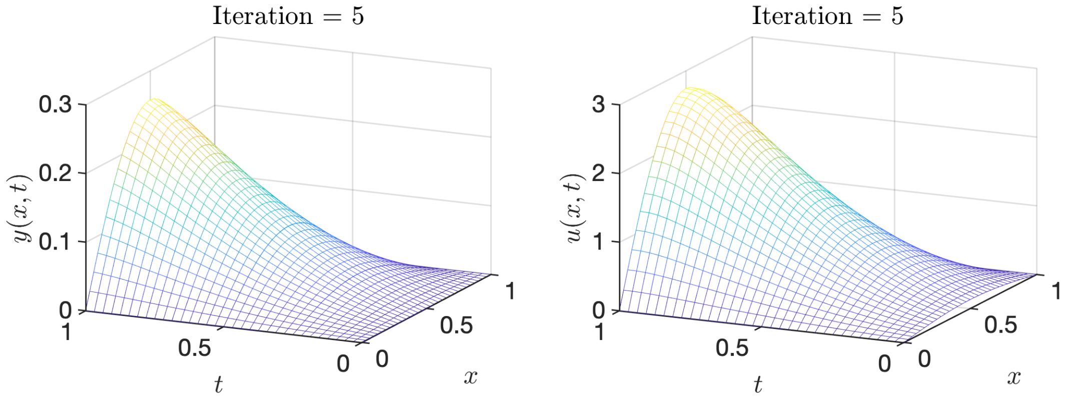
Four subdomains:



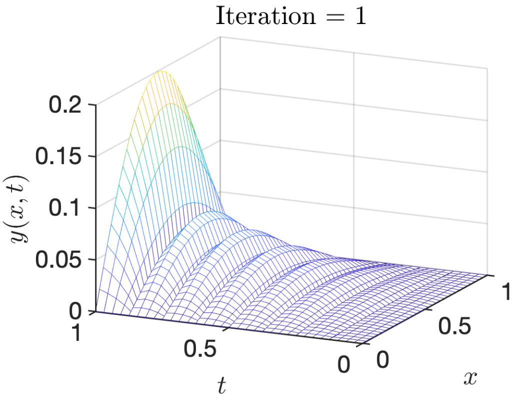
Four subdomains:

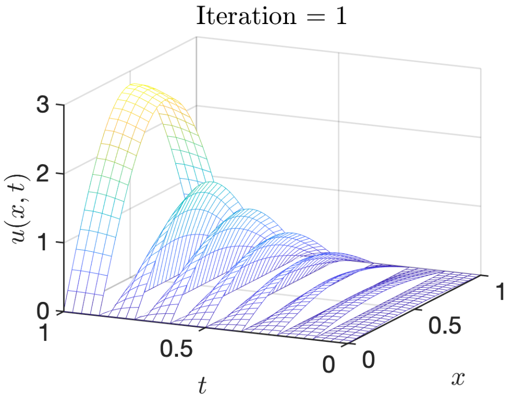


Four subdomains:

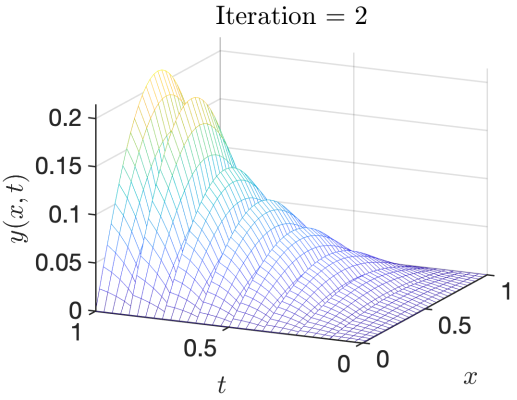


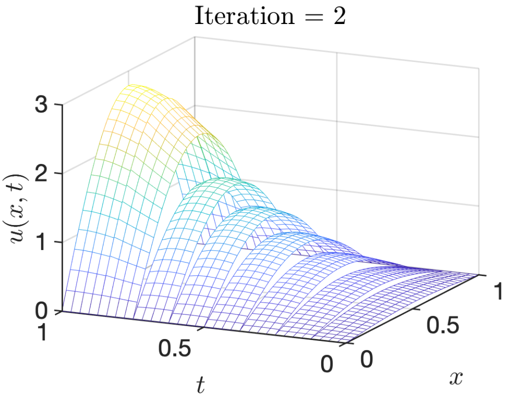
Eight subdomains:



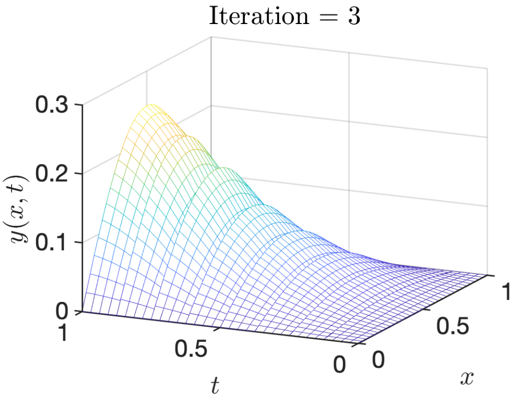


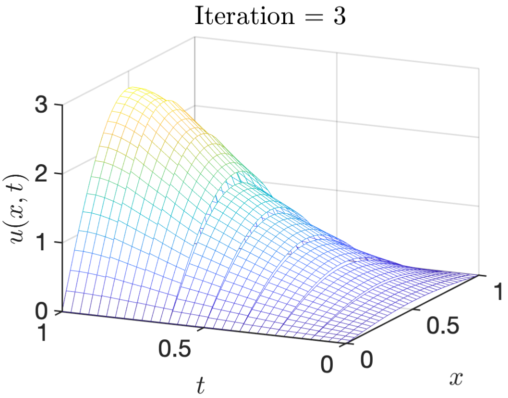
Eight subdomains:





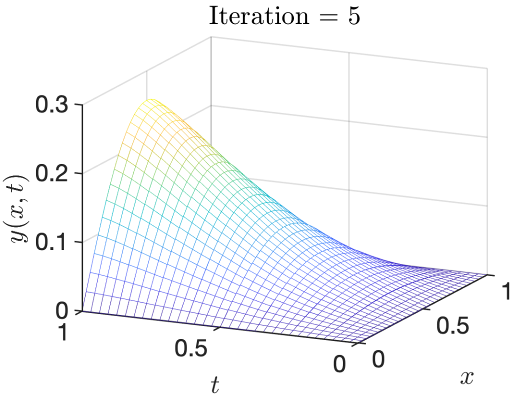
Eight subdomains:

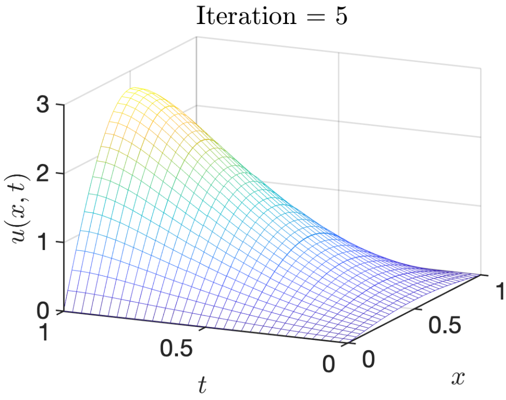




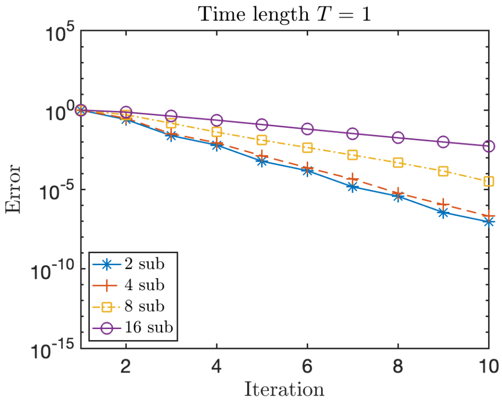
Strong scalability

Eight subdomains:



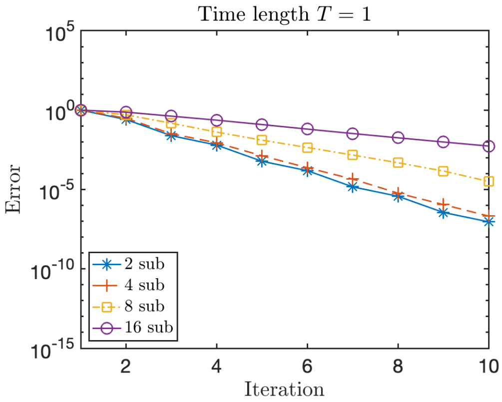


Error decay

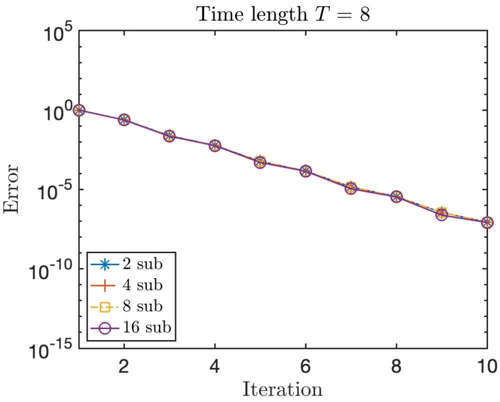


Time length $T = 1$
Error
Iteration
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub

Error decay

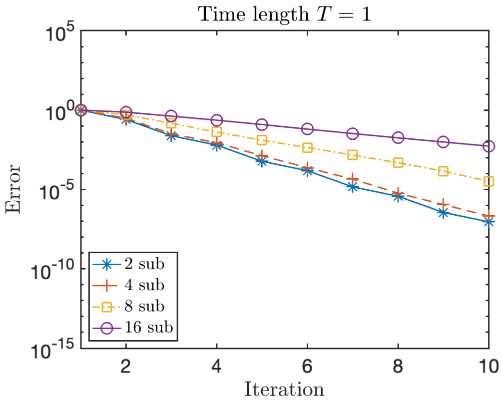


Time length $T = 1$
Error
Iteration
10^5
10^0
10^-5
10^-10
10^-15
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub

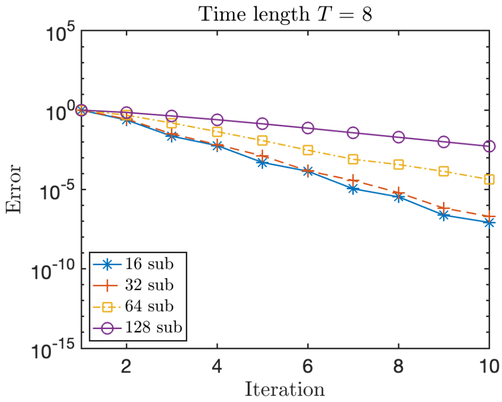


Time length $T = 8$
Error
Iteration
10^5
10^0
10^-5
10^-10
10^-15
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub

Error decay

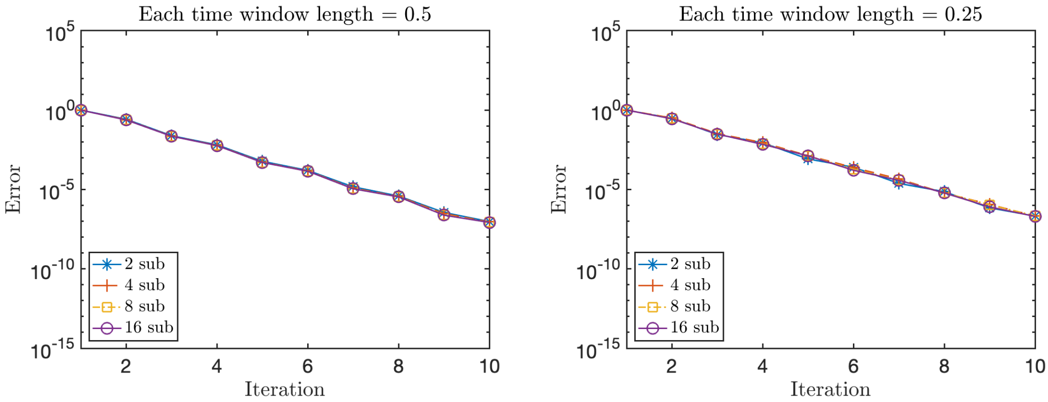


Time length $T = 1$
Error
Iteration
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub



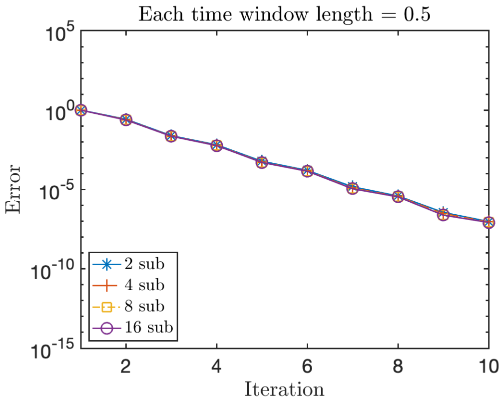
Time length $T = 8$
Error
Iteration
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
16 sub
32 sub
64 sub
128 sub

Error decay

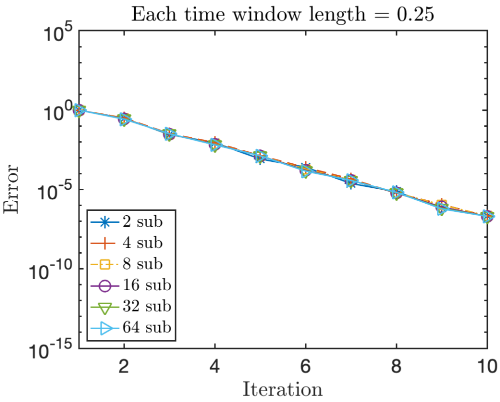


Each time window length = 0.5
Error
Iteration
10^5
10^0
10^-5
10^-10
10^-15
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub
Each time window length = 0.25
Error
Iteration
10^5
10^0
10^-5
10^-10
10^-15
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub

Error decay

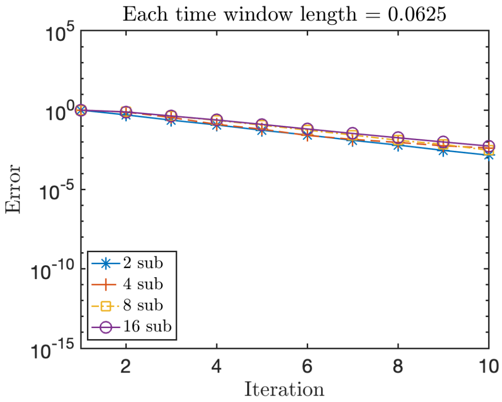


Each time window length = 0.5
Error
Iteration
10^5
10^0
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub

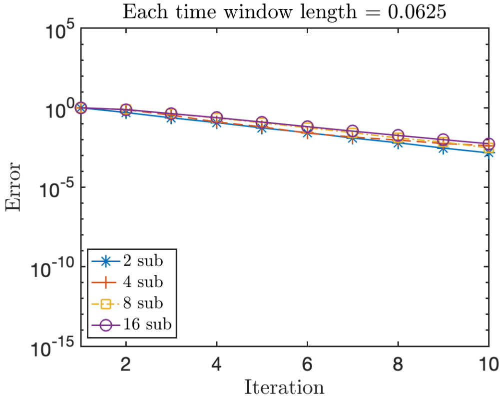


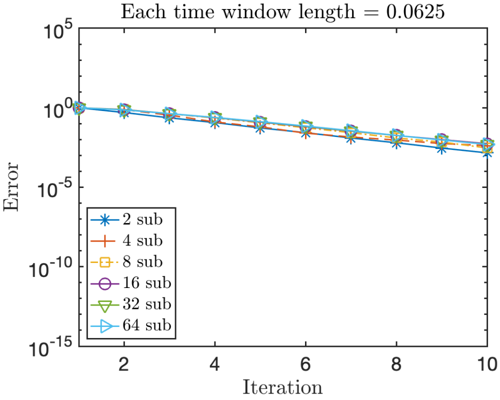
Each time window length = 0.25
Error
Iteration
10^5
10^0
10^{-5}
10^{-10}
10^{-15}
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub
32 sub
64 sub

Error decay

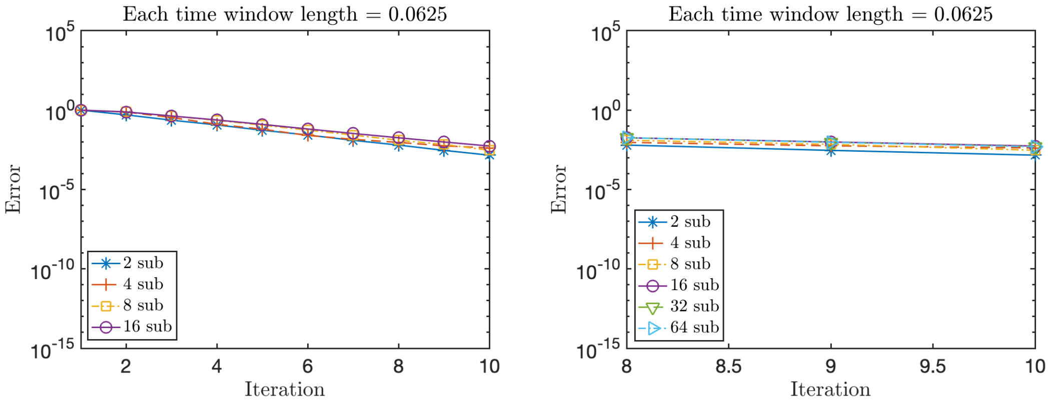


Each time window length = 0.0625
Error
Iteration
10^5
10^0
10^-5
10^-10
10^-15
2
4
6
8
10
2 sub
4 sub
8 sub
16 sub


Each time window length = 0.0625
Error
Iteration
2 sub
4 sub
8 sub
16 sub

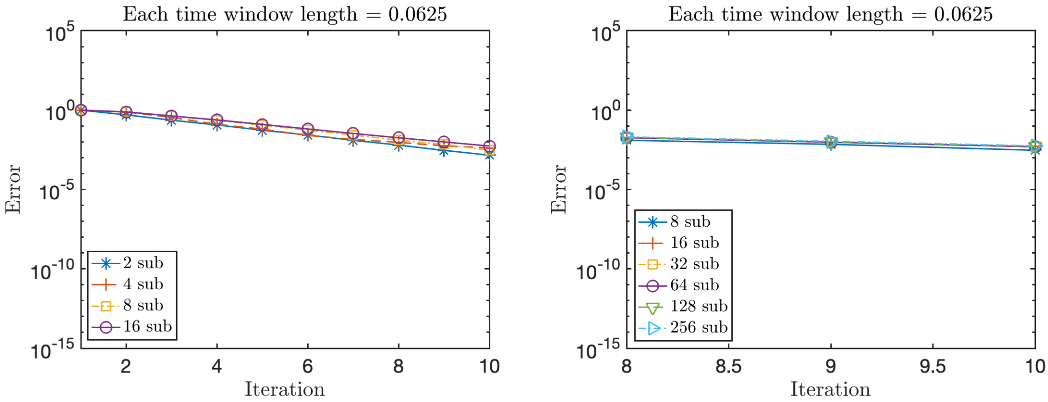

Each time window length = 0.0625
Error
Iteration
2 sub
4 sub
8 sub
16 sub
32 sub
64 sub

Error decay



Each time window length = 0.0625
10^5
10^0
10^-5
10^-10
10^-15
Error
2
4
6
8
10
Iteration
2 sub
4 sub
8 sub
16 sub
Each time window length = 0.0625
8
8.5
9
9.5
10
2 sub
4 sub
8 sub
16 sub
32 sub
64 sub

Error decay



Each time window length = 0.0625
Error
Iteration
2 sub
4 sub
8 sub
16 sub
Each time window length = 0.0625
Error
Iteration
8 sub
16 sub
32 sub
64 sub
128 sub
256 sub

Take away message:

- Time parallel Schwarz method is weak scalable.
- Time parallel Schwarz method can be strong scalable if the time length is not too small.

Conclusion

Take away message:

- Time parallel Schwarz method is weak scalable.
- Time parallel Schwarz method can be strong scalable if the time length is not too small.

Need to be done:

- Theoretical analysis for multi-subdomains, intuition: $\tanh(\sigma_i \frac{T}{N})$.
- Numerical tests for different variants of time domain decomposition.

Take away message:

- Time parallel Schwarz method is weak scalable.
- Time parallel Schwarz method can be strong scalable if the time length is not too small.

Need to be done:

- Theoretical analysis for multi-subdomains, intuition: $\tanh(\sigma_i \frac{T}{N})$.
- Numerical tests for different variants of time domain decomposition.

Thank you for your attention !