Optimal control of parabolic equations: a spectral calculus based approach

Ivica Nakić

University of Zagreb
Joint work with L. Grubišić, M. Lazar and M. Tautenhahn

IX Partial differential equations, optimal design and numeric
Benasque, 24 August 2022

The problem

Inital condition or starting optimal control: solve the following problem

$$\min_{u \in \mathcal{H}} \left\{ J(u) \colon \, \|y(T) - y^*\| \leq \varepsilon \right\},$$

where

$$\begin{cases} \dot{y}(t) + Ay(t) = f(t) & \text{for } 0 \leq t \leq T, \\ y(0) = u, \end{cases}$$

$$J(u) = \frac{\alpha}{2}\|u\|^2 + \frac{1}{2}\int_0^T \beta(t)\|y(t) - w(t)\|^2 \, \mathrm{d}t.$$

The problem

Here we assume:

- $A \geq \kappa$ selfadjoint operator on a Hilbert space $\mathcal{H}$, given as a partial differential operator on the domain $\Omega \subset \mathbb{R}^d$,
- $f \in L^2((0, T); \mathcal{H})$.

Parameters are:

- y^* the target state,
- $\varepsilon > 0$ the tolerance,
- $\alpha > 0$ weight,
- $\beta \in L^\infty((0, T); [0, \infty))$ weight, and
- $w \in L^2((0, T); \mathcal{H})$ the desired trajectory of the system.

Solving the problem

The problem is ill-posed, thus numerically challenging.

We use *first optimize, then discretize* approach. We want

- to construct a closed form expression of the solution of the problem, and
- use a discretization scheme which is tailored to the particular problem.

The solution

The solution u^{opt} of the problem is given by

$$u^{\mathrm{opt}} = (\mu^{\varepsilon} S_{2T} + \Psi)^{-1}(\mu^{\varepsilon} S_T y^* + \psi),$$

$$\Psi = \alpha I + \int_0^T \beta(t) S_{2t} \mathrm{d}t, \quad \psi = \int_0^T \beta(t) S_t w(t) \mathrm{d}t.$$

Here

- $\{S_t\}$ is the semigroup generated by $-A$
- $\mu^{\varepsilon} \geq 0$ is
  - the unique solution of $\Phi(\mu) = \varepsilon$ if $\varepsilon < \|\Psi^{-1} S_T \psi - y^*\|$,
  - and zero otherwise.

The solution

- the function $\Phi\colon [0,\infty) \to [0,\infty)$ is defined by

$$\Phi(\mu) = \|y^* - (\mu S_{2T} + \Psi)^{-1}(\mu S_{2T} y^* + S_T \psi)\|$$

  and is a decreasing function.
- The solution of the unconstrained problem ($\varepsilon = \infty$) is given by the same formula with $\mu^\varepsilon = 0$.

Practical solution

Mathematical models for interacting dynamics on networks

- Let $\phi(\mu) = y^* - (\mu S_{2T} + \Psi)^{-1}(\mu S_{2T} y^* + S_T \psi)$, hence $\Phi(\mu) = \|\phi(\mu)\|$. Then $\phi(\mu) = y^* - x$, where x is the solution of the equation

$$(\mu S_{2T} + \Psi)\, x = \mu S_{2T} y^* + S_T \psi,$$

hence the calculation of $\Phi(\mu)$ reduces to solving a linear equation.

- Note also that the optimal initial state u^{opt} is the solution of the equation

$$(\mu^{\varepsilon} S_{2T} + \Psi)\, x = \mu^{\varepsilon} S_T y^* + \psi.$$

How to calculate Ψ

Mathematical models for interacting dynamics on networks

To calculate Ψ we use the fact that it can be written as a function of A by using

$$\int_0^T \beta(t) S_{2t} \mathrm{d}t = \int_{-\infty}^{\infty} \int_0^T \beta(t) \mathrm{e}^{-2t\lambda} \mathrm{d}t \, \mathrm{d}E(\lambda) = \tilde{\beta}_0(A),$$

where $\tilde{\beta}_0$ is a function given by $\tilde{\beta}_0(\lambda) = \int_0^T \beta(t) \mathrm{e}^{-2t\lambda} \mathrm{d}t$ and E is the spectral measure of A.

How to calculate ψ

Mathematical models for interacting dynamics on networks

We can always find a good approximant for ψ. Let $\tilde{w}(t) = \sum_{i=1}^{N} w_i \chi_{[t_{i-1},t_i]}$ be an approximation of w, where $0 = t_0 < t_1 < \cdots < t_N = T$, $w_i \in \mathcal{H}$, $i = 1, \ldots, N$, and χ_S is the characteristic function of the set S. Then

$$\tilde{\psi} = \sum_{i=1}^{N} \tilde{\beta}_i(A) w_i, \text{ where } \tilde{\beta}_i(\lambda) = \int_{t_{i-1}}^{t_i} \beta(t) \mathrm{e}^{-t\lambda} \mathrm{d}t,$$

is an approximation of ψ.

How to *really* calculate Ψ and ψ: $1/5$

Mathematical models for interacting dynamics on networks

Both Ψ and ψ can be represented as functions of operator A, so we want to *efficiently* calculate $g(A)$ for a given function g.

If g has the form $u_0(x) + u_1(x)\mathrm{e}^{ax}$, where u_0, u_1 are rational functions, we use the fact that such functions can be very efficiently approximated: there exists a rational function $r_{n,n+k}$ such that

$$\|g - r_{n,n+k}\|_\infty \leq C \cdot 9.28903^{-n}$$

How to *really* calculate Ψ and ψ: $2/5$

Mathematical models for interacting dynamics on networks

We use `rkfit` algorithm to find a rational function

$$r_{RK}(x) = r_0 + \frac{r_1}{x - \zeta_1} + \cdots + \frac{r_d}{x - \zeta_d}$$

such that

$$\|r_{RK} - g\|_\infty \leq \texttt{tol}\|g\|_2.$$

Now the operator $r_{RK}(A) := r_0 I + \sum_{i=1}^{d} r_i(A - \zeta_i)^{-1}$ satisfies

$$\|g(A) - r_{RK}(A)\|_{L(\mathcal{H})} \leq \texttt{tol}\|g\|_2.$$

So it remains to approximate the resolvent of the operator A.

How to *really* calculate Ψ and ψ: $3/5$

Mathematical models for interacting dynamics on networks

We approximate the (action of the) resolvent by selecting a finite dimensional subspace $\mathcal{V}_h \subset \mathrm{Dom}(A^{1/2})$ and then forming the Galerkin projection of A onto $\mathcal{V}_h$. The Galerkin projection $A_h : \mathcal{V}_h \to \mathcal{V}_h$ is given by the formula (assuming $A \leq 0$)

$$A_h = (A^{1/2} P_h)^* (A^{1/2} P_h),$$

where P_h is the orthogonal projection onto $\mathcal{V}_h$.

How to *really* calculate Ψ and ψ: $4/5$

Let $\mathcal{V}_h$ be the space of piece-wise linear, for a given triangular tessellation of Ω, and continuous functions on Ω. Let

- h be the maximal diameter of a triangle in the chosen tessellation, and
- ν be a parameter depending on the regularity of the functions in $\mathrm{Dom}(A)$.

The resolvent estimate for A using the Galerkin projection A_h for $h < h_0$ and $v \in V_h$ is

$$\|(A-z)^{-1}v - (A_h - z)^{-1}v\|_2 \leq Ch^{2\nu}\|v\|_2.$$

Here h_0 denotes the minimal level of refinement from which the estimate holds.

How to *really* calculate Ψ and ψ: $5/5$

Finally, let $g(x) = u_0(x) + u_1(x)\mathrm{e}^{ax}$ be a perturbed exponential function. For a given rational function r_{RK} and $v \in V_h$ we have the estimate

$$\|g(A)v - r_{RK}(A_h)v\|_2 \leq \|g - r_{RK}\|_\infty \|v\|_2 + dCh^{2\nu}\|v\|_2.$$

By choosing suitable r_{RK} and h, this estimate ensures a good approximation of $g(A)v$ based on a finite dimensional approximation of the operator A.

The theorem from the following slide ensures that the numerical solution is a good approximation of the solution of the optimal problem.

Sensitivity of the problem

Mathematical models for interacting dynamics on networks

Let us perturb all the parameters of the problem with perturbations $< \nu$, in respective norms, such that that the perturbed problem has the same structure (A still selfadjoint etc.). For the perturbation of the operator A, we allow δA which is relatively bounded with respect to A. Then we have

Theorem

For small enough $\nu > 0$ the optimal solutions of the original and the perturbed problem differ (in norm) by $< C\nu$ where C does not depend on ν.

An example

Mathematical models for interacting dynamics on networks

Let $A < 0$, $f = 0$, $\beta = \chi_{[T/3, 2T/3]}$ and assume that w does not depend on time.

Then

$$\Psi = \alpha I + \frac{1}{2} A^{-1} S_{2T/3}(I - S_{2T/3}),$$
$$\psi = A^{-1} S_{T/3}(I - S_{T/3}) w.$$

Hence

$$\Psi = \alpha I + \int_{\mathbb{R}} \underbrace{\mathrm{e}^{\lambda\ 2T/3}(1/\lambda - \mathrm{e}^{\lambda 2T/3}/\lambda)}_{:= g(\lambda)} \,\mathrm{d}E(\lambda).$$

An example

The value of the function $\Phi(\mu)$ is computed using the rational approximation and the spectral calculus

$$(\mu S_{2T}+\Psi)^{-1}\mu S_{2T}y^* = \int_{-\infty}^{0} \underbrace{\frac{\mu \mathrm{e}^{2T\lambda}}{\mu \mathrm{e}^{2T\lambda}+\alpha+(1/\lambda-\mathrm{e}^{\lambda\ 2T/3}/\lambda)}}_{:=g(\lambda)}\,\mathrm{d}E(\lambda)y^*$$

$$\approx r_0 y^* + \sum_{i=1}^{18} r_i (A-\zeta_i)^{-1} y^*,$$

where $\texttt{tol} = 10^{-15}$.

The solution of $\Phi(\mu)=\varepsilon$ is obtain using the Brent method.

1D heat equation

Mathematical models for interacting dynamics on networks

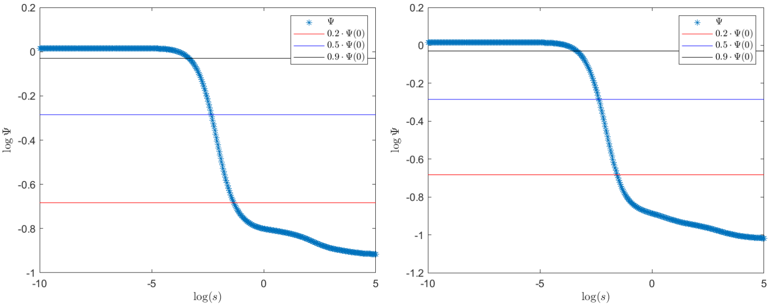
Let

- $A = -\partial_x((1 + a\chi_{[2.2,\pi]})\partial_x)$ on $\Omega = [0, \pi]$, $a \leq 0$,
- $\alpha = 10^{-4}$,
- $T = 0.01$,
- $w = \chi_{[\pi/5, 2\pi/5]}$,
- $y^* = \chi_{[3\pi/5, 4\pi/5]}$,
- $\varepsilon = [0.2, 0.5, 0.9]\Phi(0)$.

$\Phi(0)$ is the distance between the target y^* and $y(T)$ in the case $y(0) = u^{\min}$, $u^{\min}$ being the solution of the corresponding unconstrained problem.

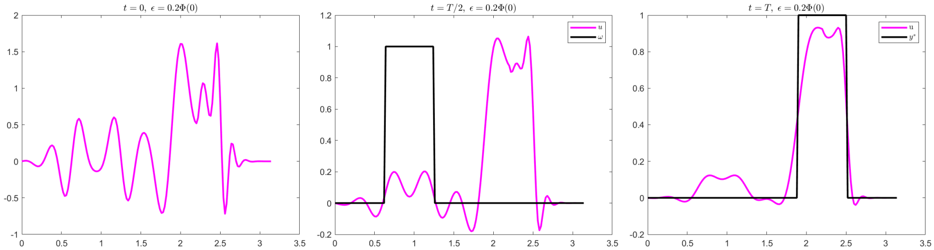
$a = 0$ and $a = -0.8$, corresponding μ_ε

Mathematical models for interacting dynamics on networks

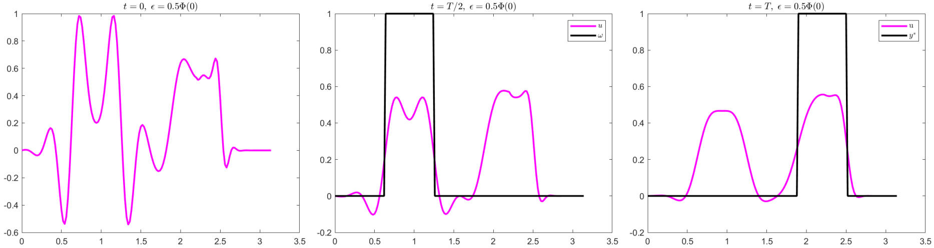


$a = -0.8,\ \varepsilon = 0.2\Phi(0)$

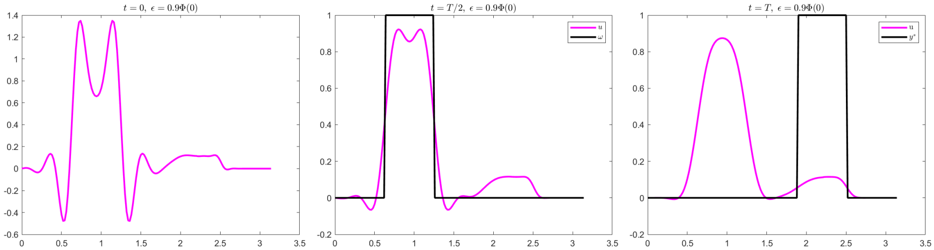
Mathematical models for interacting dynamics on networks



$a = -0.8,\ \varepsilon = 0.5\Phi(0)$



$a = -0.8,\ \varepsilon = 0.9\Phi(0)$

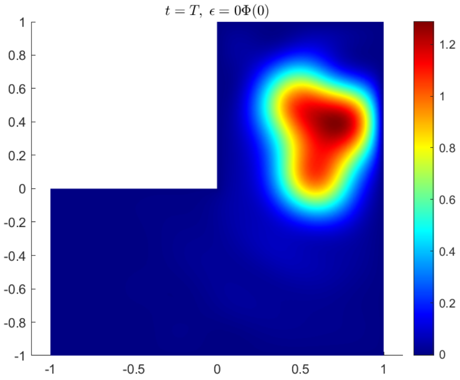


2D heat equation on irregular domain

Let

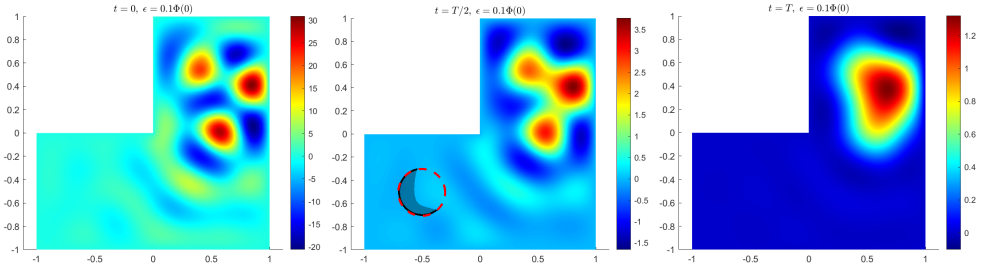
- A be Dirichlet Laplacian on $\Omega = [-1,1]^2 \setminus ([-1,0] \times [0,1])$,
- $\alpha = 10^{-4}$,
- $T = 0.01$,
- $w(x) = \chi_{\|x-x_0\|_1 \leq 0.2}$,
- $y^*(x) = \mathrm{e}^{-20\|x-x_1\|^2} + \mathrm{e}^{-20\|x-x_2\|^2} + \mathrm{e}^{-30\|x-x_3\|^2}$, with $x_0 = (-0.5, -0.5)$, $x_1 = (0.5, 0.5)$, $x_2 = (0.6, 0.1)$ and $x_3 = (0.8, 0.4)$.

The prescribed final target y^*.



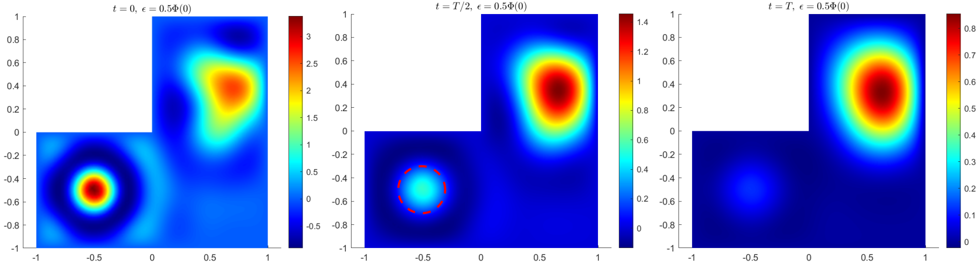
$\varepsilon = 0.1\Phi(0)$

Mathematical models for interacting dynamics on networks

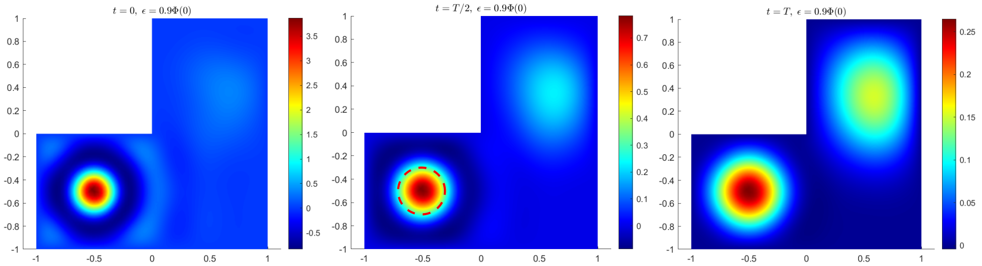


$\varepsilon = 0.5\Phi(0)$

Mathematical models for interacting dynamics on networks



$\varepsilon = 0.9\Phi(0)$



Non-homogeneous boundary condition

Suppose we have

$$\begin{cases} \dot{y}(t) + Ly(t) = 0 \quad \text{for } 0 \leq t \leq T, \\ Gy(t) = g(t), \\ y(0) = u, \end{cases}$$

where (L, G) form a well-posed *boundary control system* on $\mathcal{U}, \mathcal{Z}$ and $\mathcal{X}$.

The optimal control problem is given by

$$J(u) = \frac{\alpha}{2}\|u\|^2 + \frac{1}{2}\int_0^T \beta(t)\|y(t) - w(t)\|_{\mathcal{Z}}^2 \,\mathrm{d}t,$$

$$\min_{u \in \mathcal{Z}} \left\{ J(u) \colon \|y(T) - y^*\|_{\mathcal{Z}} \leq \varepsilon \right\}.$$

.

Non-homogeneous b. c.

Mathematical models for interacting dynamics on networks

We can also solve optimal control problems of this type.

Our approach is to lift the boundary condition to obtain

$$\dot{y} + \hat{A}y = Bg(t), \qquad (*)$$

where $\hat{A} = \left(L_{|\,\mathrm{Ker}\,G}\right)_{-1}$ and $B \in \mathcal{L}(U, \mathcal{X}_{-1})$.

One cannot directly apply the previous result as the the constraints are given in $\mathcal{Z} \subset \mathcal{X}$ and the equation $(*)$ lives in $\mathcal{X}_{-1}$.

Recapitulation

MATYDYNNET Mathematical models for interacting dynamics on networks

- We constructed and implemented a numerical algorithm for a constrained optimal control problem.
- For the numerical implementation we explored efficient Krylov subspace techniques that allowed us to approximate a complex function of an operator by a series of linear problems. We provided a-priori estimates for the approximation that are not sensitive to any particular spatial discretization, and neither to a matrix representation of the operator A.
- We provided a complete quantified sensitivity analysis of the solution with respect to all the data entering the problem.
- This approach can be generalised to other optimal control problems.