Optimal control of parabolic equations by spectral decomposition

Martin Lazar
University of Dubrovnik

Osijek, 2018
Seminar za optimizaciju i primjene

Joint work with: C. Molinari,
Universidad Técnica Federico Santa María, Valparaíso, Chile

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

The problem framework

An optimal control problem for an abstract heat equation:

$$\begin{cases} -y'(t) = \mathcal{A}y(t) & \text{for } t > 0 \\ y(0) = u \in H. \end{cases} \tag{1}$$

H - Hilbert space
$\mathcal{A}$ – positive semidefinite, self-adjoint unbounded operator on H,
– with dense domain $D(\mathcal{A})$,
– with compact resolvent.
Keynote example: $\mathcal{A} = -\Delta$, the Dirichlet Laplacian in $L^2(\Omega)$.

$\{\mathcal{S}_t\}_{t\geq 0}$ – strongly continuous semigroup of non-expansive linear operators generated by $-\mathcal{A}$

Control u – initial datum aiming to:

(1) steer the solution (arbitrarily closed) to a desired target in a given time horizon,
(2) minimise a given energy functional.

The system (1) is controllable to a *target state* $y^T \in H$ in time $T > 0$ if there is $u \in H$ such that

$$\mathcal{S}_T u = y^T.$$

In general, system (1) is NOT controllable to an arbitrary target.
E.g. $\mathcal{A} = -\Delta$, the Dirichlet Laplacian in $L^2(\Omega)$

$$(\forall\, t > 0) \quad S_t H \subseteq D(\mathcal{A}) = \mathrm{H}^2(\Omega) \cap \mathrm{H}^1_0(\Omega)$$

– no target state $y^T \in H \setminus D(\mathcal{A})$ can be attained in any time.

System (1) is approximately controllable:
for every target time $T > 0$, target state y^T, tolerance $\varepsilon > 0$,
there exists an initial datum $u \in H$ such that

$$\|\mathcal{S}_T u - y^T\|_H \leq \varepsilon.$$

The problem

Given a tolerance $\varepsilon > 0$, a control time $T > 0$, and a target state y^T, find

$$(\mathcal{P}) \qquad \hat{u} = \underset{u \in H}{\operatorname{argmin}} \left\{ J(u) : \|\mathcal{S}_T u - y^T\|_H \leq \varepsilon \right\},$$

where

$$J(u) = \frac{\alpha}{2} \|u\|_H^2 + \frac{1}{2} \int_0^T \beta(t) \, \|\mathcal{S}_t u - y^d(t)\|_H^2 \, dt,$$

with

- $y_d \in L^2(0, T; H)$, target trajectory;
- $\alpha > 0$, weight of the control cost;
- $0 \leq \beta \in L^2(0, T)$, weight of the control on the trajectory.

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

The solution is unique!

Indeed, by means of the *indicator function*:

$$I_C(y) = \begin{cases} 0 & \text{if } y \in C \\ +\infty & \text{else.} \end{cases}$$

the problem $(\mathcal{P})$ is restated as

$$\min_{u \in H} \{J(u) + I_{\bar{B}}(\mathcal{S}_T u)\},$$

where $\bar{B} = \overline{B(y^T;\varepsilon)}$.

$J + I_{\bar{B}} \circ \mathcal{S}_{\mathcal{T}}$ is proper, strongly convex and lower-semicontinuous

$\Longrightarrow$ problem $(\mathcal{P})$ has a unique solution $\hat{u}$.
Let $\tilde{u}$ denote the unique solution of the unconstrained problem

$$\tilde{u} = \operatorname*{argmin}_{u \in H} \; J(u),$$

and $\tilde{y} = \mathcal{S}_T \tilde{u}$ the corresponding final state.

Proposition

If $\|\tilde{y} - y^T\|_H \leq \varepsilon$, then $\hat{u} = \tilde{u}$.
Otherwise, the optimal final state verifies $\hat{y} \in \partial \bar{B}$.

Standard approaches to the problem

$\beta = 0$ (no desired trajectory)
HUM (Hilbert Uniqueness Method), some penalised version:

- based on the dual problem,
- discretisation of the system,
- approximation by a finite dimensional problem,
- iterative algorithm for getting the control.

Glowinski, R., Lions, J. L. Exact and approximate controllability for distributed parameter systems, Acta Numer. (1994), 269-378.

Boyer, F. On the penalized HUM approach and its application to the numerical approximation of null-controls for parabolic problems, ESAIM: Proceedings (2013), no. 41, 15-58.

$\beta \neq 0$
– even a more complex numerical treatment (convex optimisation techniques)

We present a different approach based on spectral decomposition of the solution by eigenfunctions of $\mathcal{A}$,

Geometrical interpretation

Denote

$$\psi = (J \circ S_{-T}),$$

i.e. $\psi(y) = J(u)$ for $u = S_{-T}(y)$.

Introduce sublevel sets of $\psi = (J \circ S_{-T})$:

$$\begin{aligned} W_c &= \{y \in H : \psi(y) \leq c\} \\ &= \{y \in H : \; y = \mathcal{S}_T u \text{ for some } u \in H \text{ with } J(u) \leq c\}. \end{aligned}$$

W_c is empty for $c < \tilde{c} = J(\tilde{u})$.

$(W_c)_{c \geq \tilde{c}}$ – a nested family of nonempty closed convex sets centred at $\tilde{y}$, that increases with c.

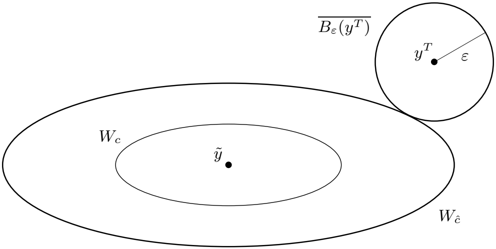

Figure: Sublevel sets W_c and the target ball.

The target ball is hit for the first time by $W_{\hat{c}}$, where $\hat{c} = J(\hat{u})$.
The intersection $\hat{y}$ is the optimal final state.

$$\hat{y} - y^T = -\hat{\gamma} \nabla \psi \left(\hat{y} \right),$$

for some $\hat{\gamma} > 0$.
Together with

$$\|\hat{y} - y^T\|_H = \varepsilon$$

we get a fully determined system for $\hat{\gamma}, \hat{y}$.

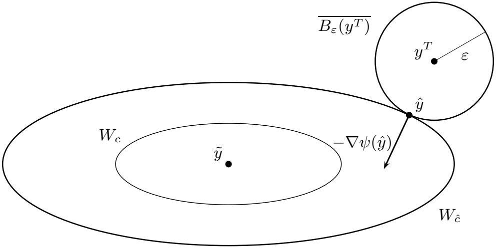

Figure: Optimal final state $\hat{y}$ as intersection of $W_{\hat{c}}$ and $\overline{B_\varepsilon(y^T)}$

Spectral decomposition

Denote:

$(\varphi_n)_{n\in\mathbf{N}}$ – an orthonormal basis of H, consisting of eigenfunctions of $\mathcal{A}$
$(\lambda_n)_{n\in\mathbf{N}}$ – a sequence of corresponding (nonnegative) eigenvalues λ_n,

$$\lim_n \lambda_n = +\infty$$

y_n – the n-th Fourier coefficient of $y \in H$.

The ellipsoids W_c can be now characterised as

$$W_c = \left\{ \sum_n y_n \varphi_n : \underbrace{\sum_n \left(a_n y_n^2 + b_n y_n + c_n \right)}_{\psi(y)=J(u)} \leq c \right\}, \tag{2}$$

where

$$a_n = \left(\frac{\alpha}{2} + \frac{1}{2} \int_0^T \beta(t) e^{-2\lambda_n t} dt \right) e^{2\lambda_n T};$$

$$b_n = -e^{\lambda_n T} \int_0^T \beta(t) e^{-\lambda_n t} y_n^d(t)\, dt;$$

$$c_n = \frac{1}{2} \int_0^T \beta(t) \left(y_n^d(t) \right)^2 dt.$$

The geometrical interpretation

$$\hat{y} - y^T = -\hat{\gamma}\nabla\psi\left(\hat{y}\right),$$

together with

$$\left(\nabla\psi(y)\right)_n = 2a_n y_n + b_n, \quad n \in \mathbf{N},$$

we get an explicit formula for the Fourier coefficients of the optimal final state $\hat{y}$:

$$\hat{y}_n = \frac{y_n^T - \hat{\gamma} b_n}{1 + 2\hat{\gamma} a_n}. \tag{3}$$

It remains to determine the constant $\hat{\gamma} > 0$.

Condition $\|\hat{y} - y^T\| = \varepsilon$ together with (3) implies

$$G(\hat{\gamma}) := \sum_n \left(\frac{\hat{\gamma}\ \left(2a_n y_n^T + b_n\right)}{1 + 2\gamma a_n}\right)^2 = \varepsilon^2. \tag{4}$$

G – strictly increasing,
 – G(0)=0,
 – $\lim_{\gamma\to\infty} = \|\tilde{y} - y^T\| =: \varepsilon_c$.

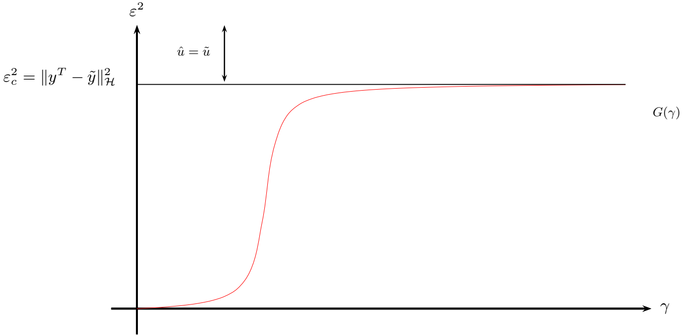

Figure: Function G

The equation

$$G(\hat{\gamma}) = \varepsilon^2$$

has the unique solution for every $\varepsilon \in \big(0, \|\tilde{y} - y^T\|_H\big)$.

Theorem [Generalized HUM]*

The solution of the optimal control problem $(\mathcal{P})$ is given by

$$\hat{u} = \sum_n \hat{u}_n \varphi_n = \sum_n e^{\lambda_n T} \hat{y}_n \varphi_n,$$

where the Fourier coefficients $(\hat{y}_n)_{n\in\mathbf{N}}$ of the final state $\hat{y}$ are given by:

$$\hat{y}_n = \frac{y_n^T - \hat{\gamma} b_n}{1 + 2\hat{\gamma} a_n},$$

$$G(\hat{\gamma}) := \sum_n \left(\frac{\hat{\gamma} \ \left(2 a_n y_n^T + b_n\right)}{1 + 2\gamma a_n} \right)^2 = \varepsilon^2.$$

* L., M., Molinari, C., Peypouquet, J.: Optimal control of parabolic equations by spectral decomposition. Optimization 66 (8), 1359-1381 (2017)

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

Obtained explicit formulas incorporate infinite series.
Truncation required.
Introduce the truncated approximation of the optimal final state, $\hat{y}^N$:

$$\hat{y}^N = \sum_{n=1}^{N} \hat{y}_n^N \varphi_n, \quad \text{with} \quad \hat{y}_n^N = \frac{y_n^T - \hat{\gamma}_N b_n}{1 + 2\hat{\gamma}_N a_n}. \tag{5}$$

Truncated Fourier series with approximate coefficients.

Theorem

The following estimate holds

$$\|\hat{y}^N - \hat{y}\|_H^2 \leq 4\|y^T - y^{T,N}\|_H^2 + \frac{4\|\beta\|_{L^2(0,T)}^2}{\alpha^2 e^{2\lambda_N T}} \|y^d - y^{d,N}\|_{L^2(0,T;H)}^2,$$

where $y^{T,N} = \sum_{n=0}^{N} y_n^T \varphi_n$ and $y^{d,N}(t) = \sum_{n=0}^{N} y_n^d(t)\, \varphi_n$ the truncated series representation of the target final state and the reference trajectory in the distributed cost, respectively.

Numerical algorithm

– produces the approximate optimal final state $\hat{y}^N$, with precision ρ.

Step 1. Determine N such that

$$\max\left\{\|y^T - y^{T,N}\|_H^2, \frac{\|\beta\|_{L^2(0,T)}^2}{\alpha^2 e^{2\lambda_N T}}\|y^d - y^{d,N}\|_{L^2(0,T;H)}^2\right\} \leq \frac{\rho^2}{8}.$$

Step 2. Compute $\varepsilon_N := \lim_{\gamma\to\infty} G_N(\gamma)$.
Step 3. For $\varepsilon \geq \varepsilon_N$ the approximative solution is $\tilde{u}^N = \sum_{n=1}^N -\frac{b_n}{2a_n} e^{\lambda_n T}\varphi_n$.
Otherwise, proceed to Step 4.
Step 4. Solve equation $G_N(\gamma) = \varepsilon^2$ numerically to find $\hat{\gamma}_N$ (bisection method or other).
Step 5. Compute the approximate optimal final state $\hat{y}^N$ using (5), and the approximate optimal control $\hat{u}^N$ by

$$\hat{u}^N = \sum_{n=1}^N \hat{u}_n^N \varphi_n = \sum_{n=1}^N e^{\lambda_n T}\hat{y}_n^N \varphi_n.$$

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

Example 1 - Energy minimisation in 2D

The heat equation on $\Omega = (0,1) \times (0,1)$

$$\begin{cases} \frac{d}{dt} y - \Delta y = 0 & \Omega \times (0,T) \\ y = 0 & \partial\Omega \times (0,T) \\ y(0) = u & (0,\pi). \end{cases} \tag{6}$$

Target time $T = 0.001$
We use the eigenfunctions of the Dirichlet Laplacian on the rectangle

$$\varphi_{j,k}(x_1, x_2) = \sin(j\pi x_1)\sin(k\pi x_2), \qquad j,k = 1,2,\ldots,$$

with corresponding eigenvalues

$$\lambda_{j,k} = (j\pi)^2 + (k\pi)^2 .$$

Example 1 - Energy minimisation in 2D

$\alpha = 1$
$\beta = 0$ - no prescribed reference trajectory.

$$(\mathcal{P}) \qquad \hat{u} \in \arg \min_{u \in L^2(0,\pi)} \left\{ \frac{1}{2} \|u\|_{L^2}^2 : \ \mathcal{S}_T u \in \overline{B_\varepsilon \left(y^T\right)} \right\}.$$

We choose a reachable final target y^T.
We introduce

$$u(x_1, x_2) = \exp\left(-\left(x_1^2 + x_2^2\right)\right) \cdot \sin\left(5\pi x_1^3\right) \cdot \sin\left(5\pi x_2^7\right),$$

and we set $u^T = u^N$, the Fourier representation of u using the first 15×15 coefficients.

The final target, given as $y^T = S_T u^T$, has a finite series representation.

The aim: Explore the differences between the initial datum u^T that generates the target and the solution $\hat{u}$ for various values of ε.

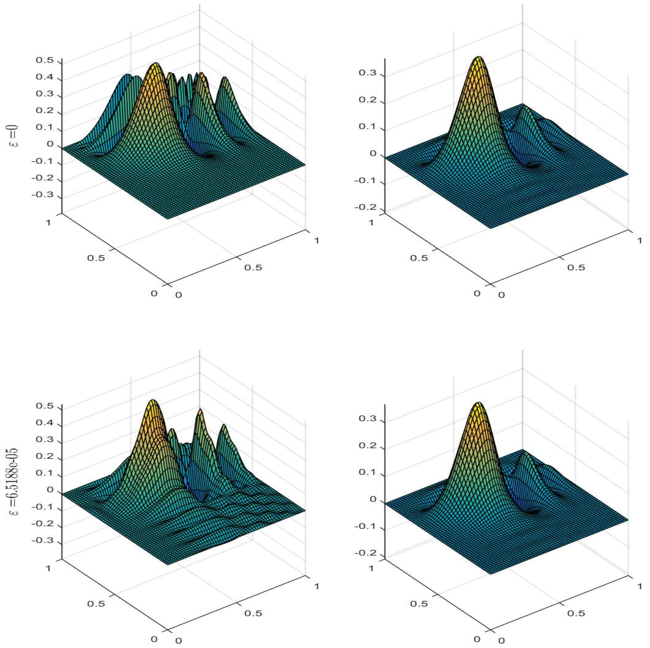

Figure: Initial data (left) and final state (right) for different values of ε.

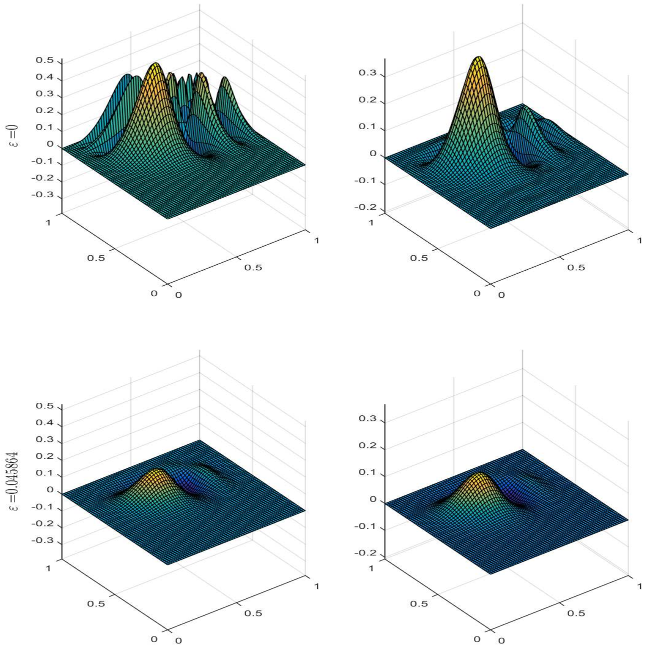

Figure: Initial data (left) and final state (right) for different values of ε.

Example 2 - Energy minimisation and trajectory regulation, 1D

The heat equation on $\Omega = (0, \pi)$, $T = 0.01$

- $\alpha = 10^{-4}$;
- $\beta(t) = \mathbb{1}_{[t_1, t_2]}(t)$, with $t_1 = T/3$ and $t_2 = 2T/3$;
- $y^d(x,t)$ as a smoothing regularisation (through classical mollifier) of the function $x \longmapsto \mathbb{1}_{[x_1, x_2]}(x)$, with $x_1 = \pi/5$ and $x_2 = 2\pi/5$;
- $y^T(x)$ as a smoothing regularisation (again, through mollifier) of the function $x \longmapsto \mathbb{1}_{[x_3, x_4]}(x)$, with $x_3 = 3\pi/5$ and $x_4 = 4\pi/5$.

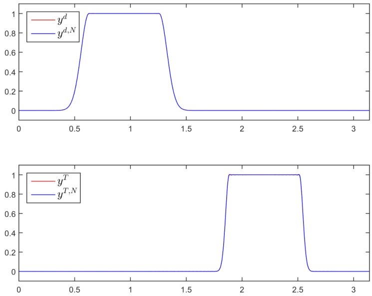

Figure: TOP: reference trajectory y^d for the distributed cost. BOTTOM: target final state y^T, in comparison with their reconstructions after Fourier decomposition with N=185 coefficients (indistinguishable).

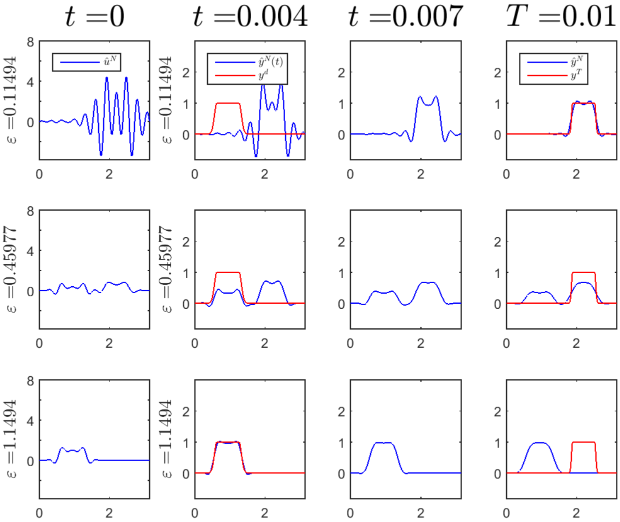

Figure: For the three values of ε: evolution of the solution in time and comparison with the reference trajectory $y^d(t)$ $(t = 0.004)$, and with the target y^T $(T = 0.01)$.

Outline

- Problem formulation
- Characterisation of the solution
- Numerical recovery
- Numerical examples
- Distributed control

The problem framework*

The constrained minimisation problem

$$(\mathcal{P}) \qquad \min_u \left\{ J(u) : \quad y(T) \in \overline{B_\varepsilon\left(y^T\right)} \right\},$$

where:
– J is a given cost functional
– y^T is a given target
– y the solution of

$$(\mathcal{E}) \qquad \begin{cases} \frac{d}{dt} y(t) + \mathcal{A} y(t) = \mathcal{B}_t u(t) & \text{for } t \in (0,T) \\ y(0) = 0. \end{cases}$$

H1 The functional J is strictly convex, coercive and lower-semicontinuous.

H2 The unbounded linear operator $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ is positive semidefinite, selfadjoint with dense domain $D(\mathcal{A})$ and compact resolvent.

H3 The operator $\mathcal{B}_t$ belongs to $\mathcal{L}\left(\mathcal{U}, \mathcal{H}\right)$ for each time $t \in (0,T)$; moreover the pair (A, B_t) is approximately controllable in time T.

U, H - real Hilbert space

* L, M., Molinari, C.: Optimal distributed control of the heat-type equations by spectral decomposition, submitted.

Existing numerical approaches

Optimal control problems involving distributed control:

- usually restricted to a null control problem
- approximation by an auxilliary problem (penalisation one)
- Fenchel-Rockafeller duality
- iterative algorithm for getting the control,
- (finite element) discretisation of the system (convergence issues!),

Fernandez-Cara, Munch: Numerical exact controllability of the 1D heat equation: duality and Carleman weights, JOTA (2014).

Fernandez-Cara, Munch: Strong convergent approximations of null controls for the 1D heat equation, SEMA (2013).

Labbé, Trélat: Uniform controllability of semidiscrete approximations of parabolic control systems, SCL (2006).

Characterisation of the solution by the dual problem

We introduce the Fenchel conjugate $J^\star$ of the functional J:

$$J^\star(u^\star) = \sup_{u \in L^2_{T,\mathcal{U}}} \{ \langle u^\star, u \rangle_{T,\mathcal{U}} - J(u) \} \quad \text{for } u^\star \in L^2_{T,\mathcal{U}}.$$

Theorem [Generalized HUM]

Let $\bar{y} \in \mathcal{H}$ be a reachable state.
Then

$$\bar{u} \in \operatorname*{argmin}_{u \in \mathcal{U}} \{ J(u) : \ \mathcal{T} u = \bar{y} \}.$$

is of the form $\bar{u} = \nabla J^\star \left(-\mathcal{T}^* \bar{\varphi}^T\right)$, where

$$\bar{\varphi}^T \in \arg\min_{\varphi^T \in \mathcal{H}} \left\{ J^\star(-\mathcal{T}^* \varphi^T) + \langle \bar{y}, \varphi^T \rangle_{\mathcal{H}} \right\}.$$

$\mathcal{T} : L^2_{T,\mathcal{U}} \to \mathcal{H}$ is the operator that takes the distributed control and gives the corresponding final state

$$\mathcal{T} u = y(T).$$

$$\mathcal{T}^* \varphi^T = \mathcal{B}^* \varphi,$$

where φ is the solution to the dual problem satisfying $\varphi(T) = \varphi^T$.

Characterisation of the solution by the dual problem

It is enough to restrict minimisation problem $(\mathcal{P})$ to controls of form $u = \nabla J^\star \left(-\mathcal{T}^* \varphi^T\right)$.
For such u

$$J(u) = F(\varphi^T),$$

where

$$F(\varphi^T) = -\left[\langle \nabla J^\star \left(-\mathcal{T}^* \varphi^T\right),\ \mathcal{T}^* \varphi^T \rangle_{L^2_{T,\mathcal{U}}} + J^\star \left(-\mathcal{T}^* \varphi^T\right)\right].$$

Theorem

The solution of problem $(\mathcal{P})$ is

$$\hat{u} = \nabla J^\star \left(-\mathcal{T}^* \hat{\varphi}^T\right),$$

where $\hat{\varphi}^T$ is a solution of

$$\min_{\varphi^T \in \mathcal{H}} \left\{ F(\varphi^T) : \ \ \|y(T) - y^T\|_{\mathcal{H}} = \varepsilon. \right\}. \tag{7}$$

Optimal control constructive characterisation

For

$$J(u) = \frac{1}{2}\int_0^T \alpha(t)\ \|u(t)\|^2_{L^2(\omega)}\ dt\ +\ \frac{1}{2}\int_0^T \beta(t)\ \|y_u(t) - y^d(t)\|^2_{L^2(\omega')}\ dt.$$

we obtain

$$\hat{\varphi}^T = \arg\min_{\varphi^T\in\mathcal{H}} \left\{\ F(\varphi^T)\right\} = \arg\min_{\varphi^T\in\mathcal{H}} \left\{\ \langle M_T\ \varphi^T, \varphi^T\rangle_{\mathcal{H}}\right\},$$

where $M_t:\ \mathcal{H}\to\mathcal{H}$ is given by:

$$M_t\left(\varphi^T\right) = \int_0^t e^{(s-t)\mathcal{A}}\mathcal{B}\left\{\left[(C^*C)^{-1}\left(\mathcal{B}^* e^{(\cdot - T)\mathcal{A}^*}\varphi^T\right)\right](s)\right\}\ ds, \qquad (8)$$

and $C = (C_1, C_2)$:

$$(C_1 u)(t) = \sqrt{\alpha(t)}\ u(t)\ \mathbb{1}_\omega;$$
$$(C_2 u)(t) = \sqrt{\beta(t)}\ y_u(t)\ \mathbb{1}_{\omega'}.$$
$$y(T) = -M_T\varphi^T + \tilde{y}(T).$$

Consequently, the original problem $(\mathcal{P})$ is equivalent to

$$(\mathcal{P}') \qquad \min_{\varphi^T\in\mathcal{H}}\left\{\ \langle M_T\ \varphi^T, \varphi^T\rangle_{\mathcal{H}}:\ \ \|\underbrace{M_T\ \varphi^T - \tilde{y}(T)}_{-y(T)} + y^T\|_{\mathcal{H}} = \varepsilon\ \right\}.$$