To appear in *Optimization*
Vol. 00, No. 00, Month 20XX, 1–20

Turnpike property for functionals involving L^1-norm⋆

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

Keywords: optimal control; parabolic equations; convex optimization

AMS Subject Classification: **49M05; 90C30**

1. Introduction

We introduce the following notation: $L^2 = L^2(\Omega)$, $L^2_T = L^2(\Omega \times (0,T))$,

$$\langle u, v\rangle = \int_\Omega u(x)v(x)\ dx;$$
$$\langle u, v\rangle^2_T = \int_0^T \int_\Omega u(x,t)v(x,t)\ dx\ dt$$

and the correspondent norms $\|\cdot\| = \langle\cdot,\cdot\rangle$ and $\|\cdot\|_T = \langle\cdot,\cdot\rangle_T$. Moreover, we define the norms

$$\|v\|_1 = \int_\Omega |\ v(x)\ |\ \ dx;$$
$$\|v\|_{1,T} = \int_0^T \int_\Omega |\ v(x,t)\ |\ \ dx\ dt.$$

We want to study the following optimal control problem:

$$(\mathcal{P}) \qquad \hat{u} \in \arg\min_{u\in L^2_T} \left\{ J(u) = \alpha_c\|u\|_{1,T} + \frac{\beta}{2}\|u\|^2_T + \alpha_s\|Lu\|_{1,T} + \frac{\gamma}{2}\|Lu - z\|^2_T \right\},$$

where $L:\ L^2_T \to L^2_T$ is defined by

$$Lu = y$$

⋆ This research was partially developed while the first two authors were visiting the BCAM under MINECO Grant MTM2011-29306. It was also partly supported by Croatian Science Foundation under Project WeConMApp/HRZZ-9780, University of Dubrovnik through the Erasmus+ Programme, Fondecyt Grant 1140829, Conicyt Anillo ACT-1106, ECOS Project C13E03, Millenium Nucleus ICM/FIC RC130003, MATH-AmSud Project 15MATH-02, Conicyt Redes 140183, and Basal Project CMM Universidad de Chile.

and y is the solution of the PDE given by

$$\begin{cases} y' + Ay = Bu & (\Omega \times (0,T)) \\ y = 0 & (\partial\Omega \times (0,T)) \\ y(0) = 0 & (\Omega)\,. \end{cases}$$

Notice that, by integration by parts, $L^*\mu = B^*p$, where φ is solution of the adjoint equation:

$$\begin{cases} -p' + A^*p = \mu & (\Omega \times (0,T)) \\ p = 0 & (\partial\Omega \times (0,T)) \\ p(T) = 0 & (\Omega)\,. \end{cases}$$

2. Sparse control: $\alpha_c > 0$ ($\alpha_s = 0$)

2.1. *The stationary problem*

$$(\mathcal{SP}_c) \qquad \bar{u} \in \arg\min_{u \in L^2} \left\{ J_s(u) = \alpha_c \|u\|_1 + \frac{\beta}{2}\|u\|^2 + \frac{\gamma}{2}\|y - z\|^2 : \quad Ay = Bu \right\}.$$

2.1.1. Optimality conditions

$$\begin{cases} A\bar{y} = B\ shrink(-B^*\bar{p}, \frac{\alpha_c}{\beta}) & (\Omega) \\ A^*\bar{p} = \gamma(\bar{y} - z) & (\Omega) \\ \bar{y} = 0,\ \bar{p} = 0 & (\partial\Omega). \end{cases}$$

2.1.2. Numerical algorithm

In order to compute a numerical solution of problem ($\mathcal{SP}_c$), after a discretization by finite differences, we use a prox-prox splitting: first write the state as $y = A^{-1}Bu$, then

- Proximal-point step:

$$\begin{aligned} \tilde{u}_k &= \arg\min_{u \in L^2} \left\{ \frac{\beta}{2}\|u\|^2 + \frac{\gamma}{2}\|A^{-1}Bu - z\|^2 + \frac{1}{2\lambda_k}\|u - u_k\|^2 \right\} \\ &= \left[\left(\beta + \frac{1}{\lambda_k}\right) I + \gamma B^* A^{-*} A^{-1} B \right]^{-1} \left(\frac{1}{\lambda_k} u_k + \gamma B^* A^{-*} z \right). \end{aligned}$$

- Proximal-point step:

$$\begin{aligned} u_{k+1} &= \arg\min_{u \in L^2} \left\{ \alpha_c \|u\|_{1,T} + \frac{1}{2\lambda_k}\|u - \tilde{u}_k\|_T^2 \right\} \\ &= shrink(\tilde{u}_k, \alpha_c \lambda_k). \end{aligned}$$

Remark 2.1 Notice that, when $\alpha_s = 0$, the solution of ($\mathcal{P}_s^c$) is simply given by

$$\bar{u} = \gamma \left[\beta I + \gamma B^* A^{-*} A^{-1} B \right]^{-1} B^* A^{-*} z.$$

2.2. *Evolutionary problem*

$$(\mathcal{P}_c) \qquad \hat{u} \in \arg\min_{u \in L^2_T} \left\{ J(u) = \alpha_c \|u\|_{1,T} + \frac{\beta}{2}\|u\|^2_T + \frac{\gamma}{2}\|Lu - z\|^2_T \right\}.$$

2.2.1. Optimality conditions

Define the classical Lagrangian

$$\mathcal{L}(u, y, p) = J(u) + \langle p, Bu - y' - Ay \rangle_T.$$

By integration by parts, we have

$$\begin{aligned}\mathcal{L}(u, y, p) = \alpha_c \|u\|_{1,T} + \frac{\beta}{2}\|u\|^2_T + \frac{\gamma}{2}\|y - z\|^2_T + \langle B^* p, u \rangle_T \\ + \langle p' - A^* p, y \rangle_T + \langle p(0), y(0) \rangle - \langle p(T), y(T) \rangle.\end{aligned}$$

Deriving with respect to the three variables (u, y, p), we obtain the optimality system:

$$\begin{cases} \hat{y}' + A\hat{y} = B\hat{u} & (\Omega \times (0,T)) \\ -\hat{p}' + A^*\hat{p} = \gamma (y - z) & (\Omega \times (0,T)) \\ \hat{y} = 0, \ \hat{p} = 0 & (\partial\Omega \times (0,T)) \\ \hat{y}(0) = 0, \ \hat{p}(T) = 0 & (\Omega), \end{cases}$$

where the relation between the optimal control and the dual state is given by

$$0 \in \alpha_c \, \partial \| \cdot \|_{1,T}(\hat{u}) + \beta \hat{u} + B^* \hat{p}.$$

The latter is equivalent to

$$\begin{aligned}\hat{u} &= \left(\beta I + \alpha_c \, \partial \| \cdot \|_{1,T}\right)^{-1}(-B^*\hat{p}) \\ &= \arg\min_{v \in L^2_T} \left\{ \alpha_c \|v\|_{1,T} + \frac{1}{2\beta}\|v + B^*\hat{p}\|^2_T \right\} \\ &= shrink(-B^*\hat{p}, \frac{\alpha_c}{\beta}),\end{aligned}$$

where the operator of $soft - shrinkage$ is defined by

$$shrink(t, \alpha) = \begin{cases} t + \alpha & (t < -\alpha) \\ 0 & (-\alpha \le t \le \alpha) \\ t - \alpha & (t > \alpha). \end{cases}$$

Finally,

$$\begin{cases} \hat{y}' + A\hat{y} = B\ shrink(-B^*\hat{p}, \frac{\alpha_c}{\beta}) & (\Omega \times (0,T)) \\ -\hat{p}' + A^*\hat{p} = \gamma\,(y - z) & (\Omega \times (0,T)) \\ \hat{y} = 0,\ \hat{p} = 0 & (\partial\Omega \times (0,T)) \\ \hat{y}(0) = 0,\ \hat{p}(T) = 0 & (\Omega)\,. \end{cases}$$

2.2.2. Numerical algorithm

In order to compute a numerical solution of problem $(\mathcal{P}_c)$, after a discretization by finite differences, we use a grad-prox splitting:

- Gradient step:

$$\begin{aligned}\tilde{u}_k &= u_k - \lambda_k \nabla_u \left[\frac{\beta}{2}\|u\|_T^2 + \frac{\gamma}{2}\|Lu - z\|_T^2\right](u_k) \\ &= u_k - \lambda_k \left[\beta u_k + \gamma L^* \left(Lu_k - z\right)\right] \\ &= u_k - \lambda_k \left[\beta u_k + \gamma B^* p_k\right],\end{aligned}$$

where

$$\begin{cases} y_k' + Ay_k = Bu_k & (\Omega \times (0,T)) \\ y_k = 0 & (\partial\Omega \times (0,T)) \\ y_k(0) = 0 & (\Omega) \end{cases}$$

and

$$\begin{cases} -p_k' + A^* p_k = y_k - z & (\Omega \times (0,T)) \\ p_k = 0 & (\partial\Omega \times (0,T)) \\ p_k(T) = 0 & (\Omega)\,. \end{cases}$$

- Proximal-point step:

$$\begin{aligned}u_{k+1} &= \underset{u \in L_T^2}{\arg\min} \left\{ \alpha_c \|u\|_{1,T} + \frac{1}{2\lambda_k}\|u - \tilde{u}_k\|_T^2 \right\} \\ &= shrink(\tilde{u}_k, \alpha_c \lambda_k).\end{aligned}$$

Remark 2.2 Another possibility is to include the term $\frac{\beta}{2}\|u\|_T^2$ in the proximal step.

Remark 2.3 Notice that, for

$$f(u) = \frac{\beta}{2}\|u\|_T^2 + \frac{\gamma}{2}\|Lu - z\|_T^2,$$

then ∇f is Lipschitz continuous. Indeed, for $u_i \in L_T^2$ $(i = 1,2)$, then

$$\nabla f(u_i) = \beta u_i + \gamma B^* p_i,$$

where

$$\begin{cases} y_i' + Ay_i = Bu_i & (\Omega \times (0,T)) \\ y_i = 0 & (\partial\Omega \times (0,T)) \\ y_i(0) = 0 & (\Omega) \end{cases}$$

and

$$\begin{cases} -p_i' + A^* p_i = y_i - z & (\Omega \times (0,T)) \\ p_i = 0 & (\partial\Omega \times (0,T)) \\ p_i(T) = 0 & (\Omega)\,. \end{cases}$$

By linearity $\delta y = y_2 - y_1$ and $\delta p = p_2 - p_1$ solve the same equations with right-hand-sides $B(u_2 - u_1)$ and δy, respectively. Then

$$\begin{aligned} \|\nabla f(u_2) - \nabla f(u_1)\| &\leq \beta \|u_2 - u_1\|_T + \gamma \|B^*\| \|\delta p\|_T \\ &\leq \beta \|u_2 - u_1\|_T + \gamma\ C_{adj} \|B\| \|\delta y\|_T \\ &\leq \beta \|u_2 - u_1\|_T + \gamma\ C_{adj} C\ \|B\| \|B(u_2 - u_1)\|_T \\ &\leq L \|u_2 - u_1\|_T, \end{aligned}$$

where we defined

$$L = \beta + \gamma\ C_{adj} C\ \|B\|^2.$$

In order the prox-grad method to converge, the restriction on the step size is given by

$$0 < \lambda \leq \lambda_k \leq \Lambda < \frac{2}{L}.$$

3. Sparse state: $\alpha_s > 0$ ($\alpha_c = 0$)

$$(\mathcal{P}_s) \qquad \hat{u} \in \underset{u\in L^2_T}{\arg\min} \left\{ J\left(u\right) = \frac{\beta}{2}\|u\|_T^2 + \alpha_s\|Lu\|_{1,T} + \frac{\gamma}{2}\|Lu - z\|_T^2 \right\}.$$

3.1. *The stationary problem*

$$(\mathcal{SP}_s) \qquad \bar{u} \in \underset{u\in L^2}{\arg\min} \left\{ J_s\left(u\right) = \alpha_c\|u\|_1 + \frac{\beta}{2}\|u\|^2 + \frac{\gamma}{2}\|y - z\|^2 : \quad Ay = Bu \right\}.$$

3.1.1. Optimality conditions

$$\begin{cases} A\bar{y} = -\frac{1}{\beta}BB^*\hat{p} & (\Omega\times(0,T)) \\ \hat{y} = shrink(A^*\hat{p} + \gamma z, \frac{\alpha_s}{\gamma}) & (\Omega\times(0,T)) \\ \bar{y} = 0,\ \bar{p} = 0 & (\partial\Omega\times(0,T))\,. \end{cases}$$

Finally, we obtain a single equation in the dual variable p:

$$\begin{cases} A\ shrink(A^*\bar{p} + \gamma z, \frac{\alpha_s}{\gamma}) = -\frac{1}{\beta}BB^*\bar{p} & (\Omega\times(0,T)) \\ \bar{p} = 0 & (\partial\Omega\times(0,T))\,. \end{cases}$$

3.1.2. Numerical algorithm

In order to compute a numerical solution of problem $(\mathcal{P}_s)$, after a discretization by finite differences, we use a prox-prox splitting on the Augmented Energy: first write the state as $y = A^{-1}Bu$, then

- Proximal-point step:

$$\begin{aligned} u_{k+1} &= \underset{u\in L^2}{\arg\min} \left\{ \frac{\beta}{2}\|u\|^2 + \frac{\gamma}{2}\|A^{-1}Bu - z\|^2 + \frac{\delta}{2\lambda_k}\|A^{-1}Bu - y_k\|^2 + \frac{1}{2\lambda_k}\|u - u_k\|^2 \right\} \\ &= \left[\left(\beta + \frac{1}{\lambda_k}\right) I + \left(\gamma + \frac{\delta}{\lambda_k}\right) B^*A^{-*}A^{-1}B \right]^{-1} \left[\frac{1}{\lambda_k}u_k + B^*A^{-*}\left(\gamma z + \frac{\delta}{\lambda_k}y_k\right) \right]. \end{aligned}$$

- Proximal-point step:

$$\begin{aligned} y_{k+1} &= \underset{y\in L^2}{\arg\min} \left\{ \alpha_s\|y\|_1 + \frac{\delta}{2\lambda_k}\|y - A^{-1}Bu_{k+1}\|^2 + \frac{1}{2\lambda_k}\|y - y_k\|_T^2 \right\} \\ &= shrink(\tilde{y}_k, \tilde{\lambda}_k), \end{aligned}$$

  where we defined

$$\begin{aligned} \tilde{y}_k &= \frac{y_k + \delta A^{-1}Bu_{k+1}}{1+\delta}; \\ \tilde{\lambda}_k &= \frac{\alpha_s\lambda_k}{1+\delta}. \end{aligned}$$

Remark 3.1 Notice that again, when $\alpha_s = 0$, the solution of $(\mathcal{P}_s)$ is simply given by

$$\bar{u} = \gamma \left[\beta I + \gamma B^* A^{-*} A^{-1} B\right]^{-1} B^* A^{-*} z.$$

3.2. ***Evolutionary problem***

3.2.1. Optimality conditions

Define the classical Lagrangian

$$\mathcal{L}(u, y, p) = J(u) + \langle p, Bu - y' - Ay \rangle_T.$$

By integration by parts, we have

$$\begin{aligned}\mathcal{L}(u, y, p) = &\frac{\beta}{2}\|u\|_T^2 + \alpha_s \|y\|_{1,T} + \frac{\gamma}{2}\|y - z\|_T^2 + \langle B^* p, u \rangle_T \\ &+ \langle p' - A^* p, y \rangle_T + \langle p(0), y(0) \rangle - \langle p(T), y(T) \rangle.\end{aligned}$$

Deriving with respect to the three variables (u, y, p), we obtain the optimality system:

$$\begin{cases} \hat{y}' + A\hat{y} = B\hat{u} & (\Omega \times (0,T)) \\ -\hat{p}' + A^*\hat{p} \in \gamma(\hat{y} - z) + \alpha_s \, \partial\| \cdot \|_{1,T}(\hat{y}) & (\Omega \times (0,T)) \\ \hat{y} = 0, \ \hat{p} = 0 & (\partial\Omega \times (0,T)) \\ \hat{y}(0) = 0, \ \hat{p}(T) = 0 & (\Omega), \end{cases}$$

where the relation between the optimal control and the dual state is given by

$$\hat{u} = -\frac{1}{\beta} B^* \hat{p}.$$

The adjoint equation is equivalent to

$$\begin{aligned}\hat{y} &= (\gamma I + \alpha_s \, \partial\| \cdot \|_{1,T})^{-1} \left(-\hat{p}' + A^*\hat{p} + \gamma z\right) \\ &= shrink(-\hat{p}' + A^*\hat{p} + \gamma z, \frac{\alpha_s}{\gamma}).\end{aligned}$$

Finally,

$$\begin{cases} \hat{y}' + A\hat{y} = -\frac{1}{\beta} B B^* \hat{p} & (\Omega \times (0,T)) \\ \hat{y} = shrink(-\hat{p}' + A^*\hat{p} + \gamma z, \frac{\alpha_s}{\gamma}) & (\Omega \times (0,T)) \\ \hat{y} = 0, \ \hat{p} = 0 & (\partial\Omega \times (0,T)) \\ \hat{y}(0) = 0, \ \hat{p}(T) = 0 & (\Omega), \end{cases}$$

3.2.2. Numerical algorithm

In order to compute a numerical solution of problem $(\mathcal{P}_s)$, after a discretization by finite differences, we use a grad-prox splitting on the following Augmented Energy:

$$\mathcal{L}_\lambda(u,y) = \frac{\beta}{2}\|u\|_T^2 + \alpha_s\|y\|_{1,T} + \frac{\gamma}{2}\|Lu - z\|_T^2 + \frac{\delta}{2\lambda}\|Lu - y\|_T^2.$$

Then,

- Gradient step:

$$\begin{aligned} u_{k+1} &= u_k - \lambda_k \nabla_u \left[\frac{\beta}{2}\|u\|_T^2 + \frac{\gamma}{2}\|Lu - z\|_T^2 + \frac{\delta}{2\lambda}\|Lu - y\|_T^2\right](u_k) \\ &= u_k - \lambda_k \left[\beta u_k + \gamma L^*(Lu_k - z) + \frac{\delta}{\lambda_k} L^*(Lu_k - y_k)\right] \\ &= (1 - \beta\lambda_k)\, u_k - B^* p_k, \end{aligned}$$

  where

$$\begin{cases} y'_{u_k} + Ay_{u_k} = Bu_k & (\Omega \times (0,T)) \\ y_{u_k} = 0 & (\partial\Omega \times (0,T)) \\ y_{u_k}(0) = 0 & (\Omega) \end{cases}$$

  and

$$\begin{cases} -p'_k + A^* p_k = (\gamma\lambda_k + \delta)\, y_{u_k} - \gamma\lambda_k z - \delta y_k & (\Omega \times (0,T)) \\ p_k = 0 & (\partial\Omega \times (0,T)) \\ p_k(T) = 0 & (\Omega)\,. \end{cases}$$

- Proximal-point step:

$$\begin{aligned} y_{k+1} &= \underset{y \in L_T^2}{\arg\min} \left\{\alpha_s\|y\|_{1,T} + \frac{\delta}{2\lambda_k}\|y - Lu_{k+1}\|_T^2 + \frac{1}{2\lambda_k}\|y - y_k\|_T^2\right\} \\ &= shrink(\tilde{y}_k, \tilde{\lambda}_k), \end{aligned}$$

  where we defined

$$\tilde{y}_k = \frac{y_k + \delta L u_{k+1}}{1 + \delta};$$
$$\tilde{\lambda}_k = \frac{\alpha_s \lambda_k}{1 + \delta}.$$

Remark 3.2 Another possibility is to consider

$$\mathcal{L}_\lambda(u,y) = \frac{\beta}{2}\|u\|_T^2 + \alpha_s\|y\|_{1,T} + \frac{\gamma}{2}\|y - z\|_T^2 + \frac{\delta}{2\lambda}\|Lu - y\|_T^2.$$

4. Computational experiments

In the following, we present the setting for the numerical experiments.

- Spacial domain: $\Omega = (0, 1)$;
- Time interval: $[0, T]$, with $T = 1$;
- Weight-parameters: $\alpha_c = [0, 0.01]$, $\alpha_s = [0, 0.65]$, $\beta = 0.0001$ and $\gamma = 1$;
- Trajectory target:

$$z(x) = \mathcal{I}_{[x_a, x_b]},$$

  where $x_a = 1.7/3$, $x_b = 3.5/4$;
- Control operator: for $x_1 = 1/7$ and $x_2 = 4/5$,

$$B = \mathcal{I}_{[x_1, x_2]};$$

- A is the finite difference discretization of $-\Delta$;
- Numerical grid: $N_x = 300$ in space, $N_t = 100$ in time.

4.1. *Stationary solutions*

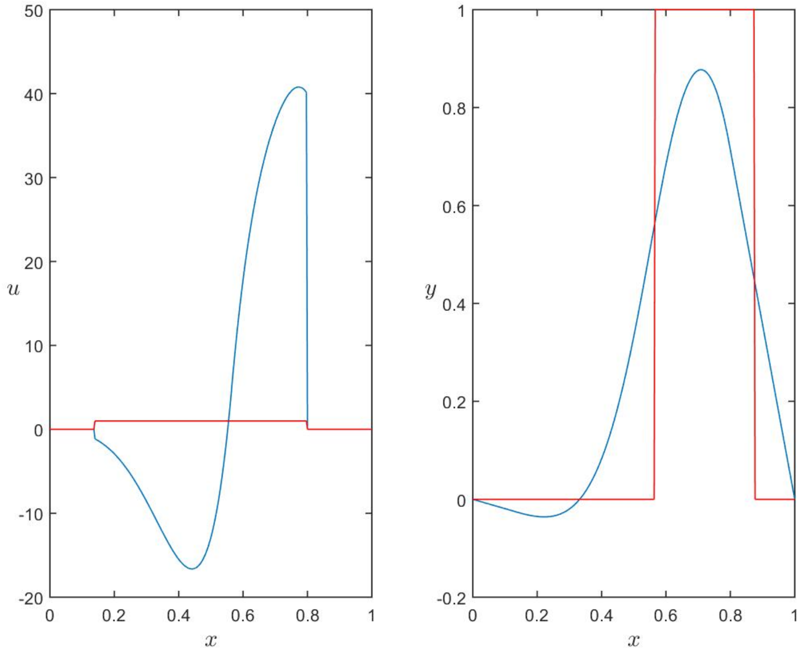

Figure 1.: $\alpha_c = \alpha_s = 0$.

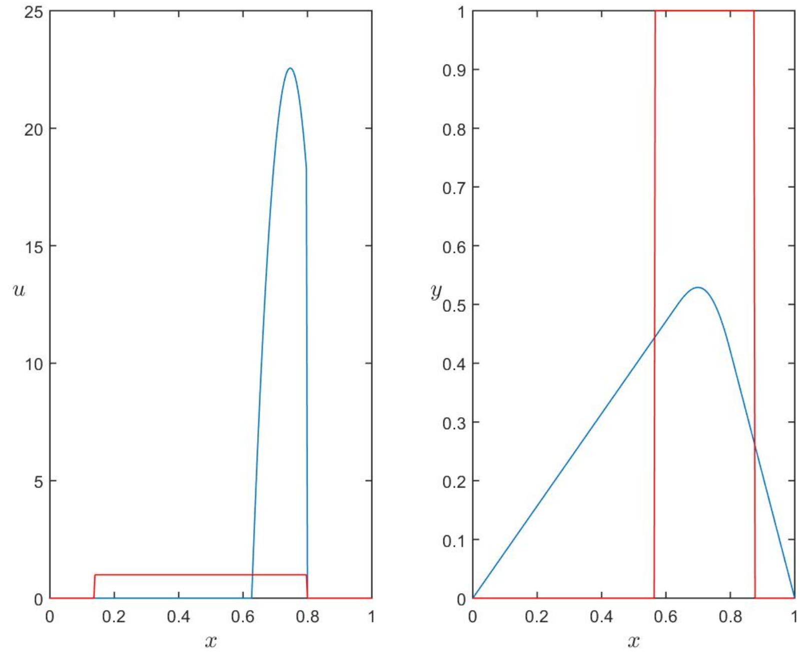

Figure 2.: $\alpha_c = 0.01$, $\alpha_s = 0$.

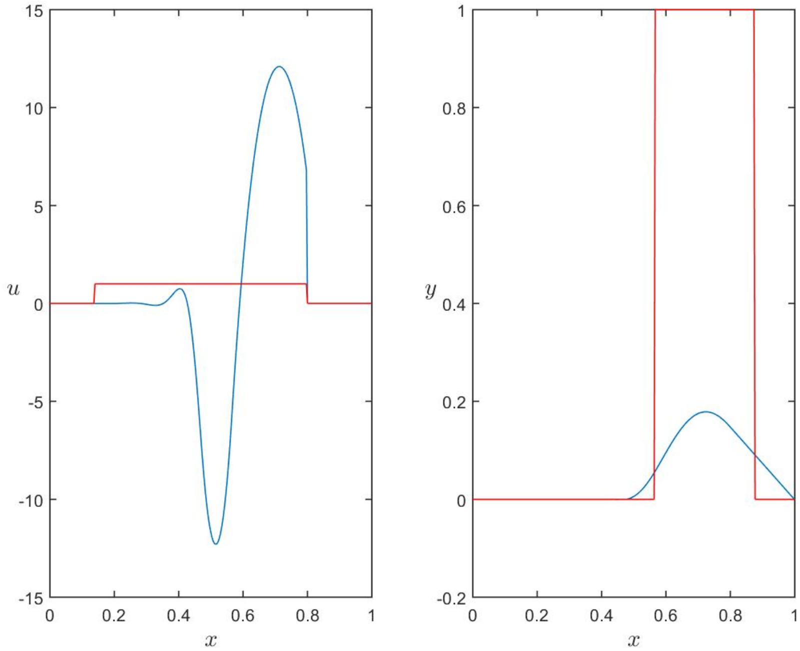

Figure 3.: $\alpha_c = 0$, $\alpha_s = 0.65$.

4.2. *Evolutionary problem*

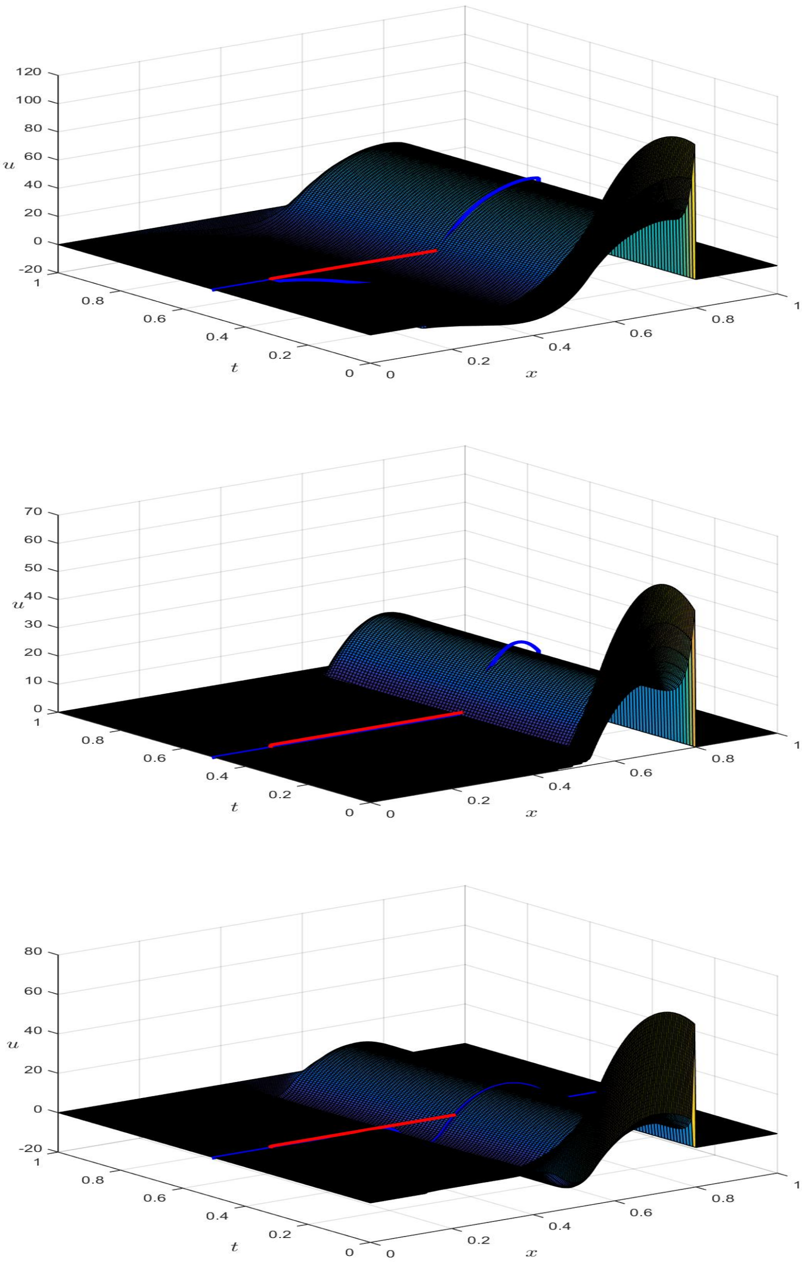

Figure 4.: Optimal control for $\alpha_c = \alpha_s = 0$ (TOP), $\alpha_c = 0.01$, $\alpha_s = 0$ (MIDDLE) and $\alpha_c = 0$, $\alpha_s = 0.65$ (BOTTOM). In red, the controllable subdomain; in blue, the stationary optimal controls.

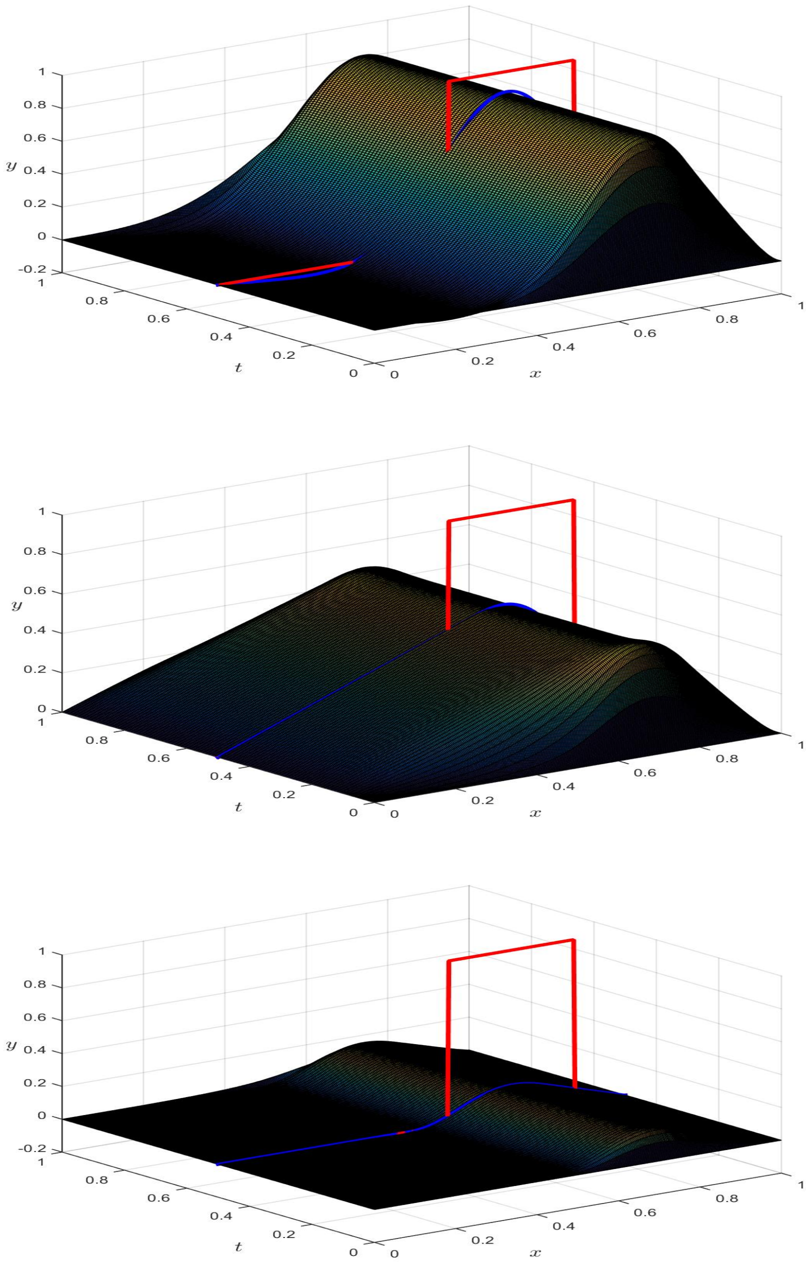

Figure 5.: Optimal state for $\alpha_c = \alpha_s = 0$ (TOP), $\alpha_c = 0.01$, $\alpha_s = 0$ (MIDDLE) and $\alpha_c = 0$, $\alpha_s = 0.65$ (BOTTOM). In red, the target z; in blue, the stationary optimal states.

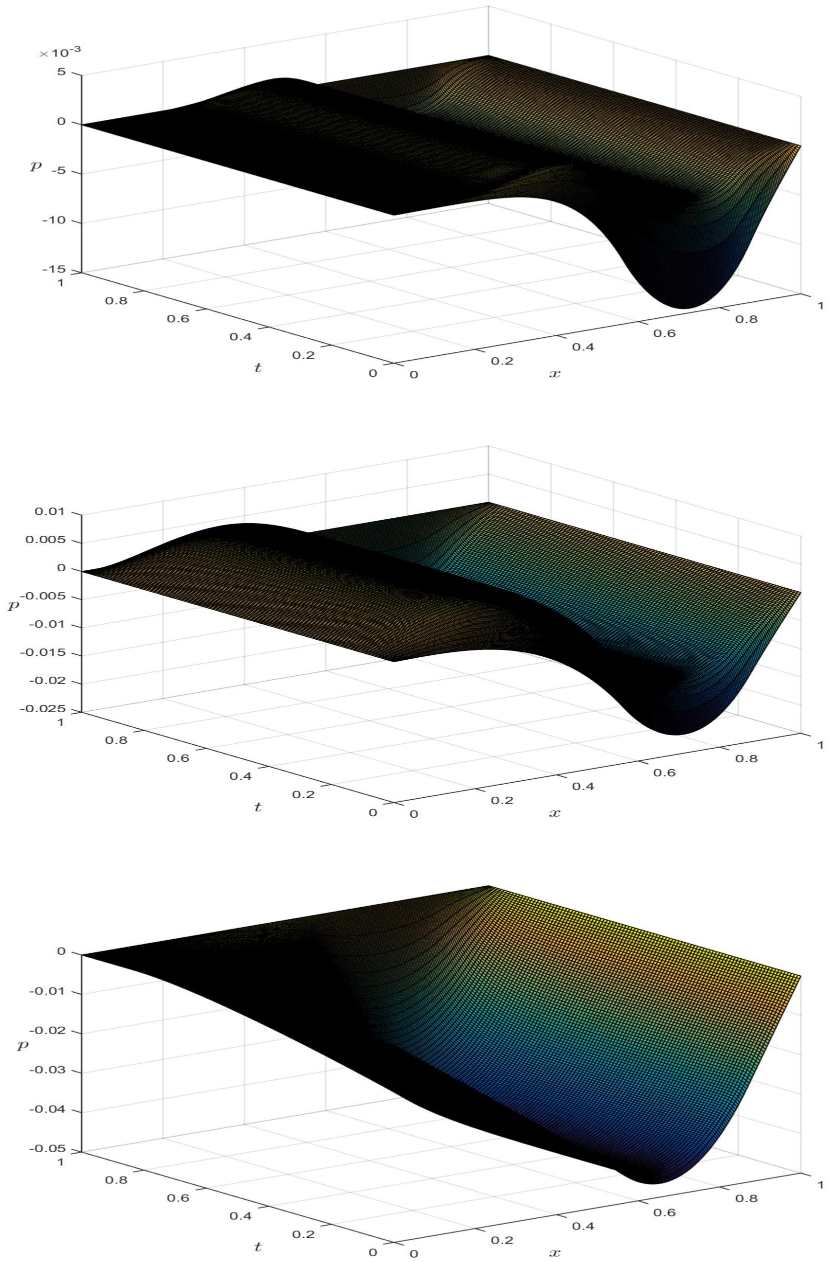

Figure 6.: Optimal adjoint for $\alpha_c = \alpha_s = 0$ (TOP), $\alpha_c = 0.01$, $\alpha_s = 0$ (MIDDLE) and $\alpha_c = 0$, $\alpha_s = 0.65$ (BOTTOM).

References

[1] Peypouquet, J. Convex optimization in normed spaces: theory, methods and examples. With a foreword by Hedy Attouch. Springer Briefs in Optimization. Springer, Cham, 2015. xiv+124 pp.