THE INTERVAL TURNPIKE PROPERTY FOR ADJOINTS

TIMM FAULWASSER, LARS GRÜNE, JUKKA-PEKKA HUMALOJA,
AND MANUEL SCHALLER

ABSTRACT. In this work we derive an interval turnpike result for adjoints of finite- and infinite-dimensional nonlinear optimal control problems under the assumption of an interval turnpike on states and controls. We consider stabilizable dynamics governed by a generator of a semigroup with finite-dimensional unstable part satisfying a spectral decomposition condition and show the desired turnpike property under continuity assumptions on the first-order optimality conditions. We further give stronger estimates for analytic semigroups and provide a numerical example with a boundary controlled semilinear heat equation to illustrate the results.

1. INTRODUCTION

The turnpike property is a particular feature of optimal solutions of dynamic optimal control problems (OCPs). It is characterized by the phenomenon of optimal solutions to long-horizon OCPs staying close a specific steady state, the so-called turnpike, for the majority of the time. First described in a paper by von Neumann in the middle of the 20th century [27], turnpike behavior has since received vast interest, cf. the recent works [16, 18, 19, 20, 29, 32, 34, 38]. For nonlinear problems, a way to derive turnpike properties is linearization of the extremal equations, analysis of the linearization and a smallness assumption, cf. [4, 17, 35, 36]. Another possibility is to assume a particular notion of dissipativity, cf. [7, 9, 14, 15], which has the advantage to allow for global turnpike properties on state and control, i.e., without a smallness condition on, e.g., the initial distance to the turnpike. In that context, however, up to now there were no results on the behavior of the corresponding adjoints. Thus, in this paper, we will show that the turnpike behavior of state and control induces turnpike behavior of the adjoints without smallness assumptions. To this end, we analyze the first-order necessary optimality conditions and, loosely speaking, show for problems governed by general evolution equations that continuity

of the nonlinearities and convergence of state and control imply convergence of the adjoints. While our results are formulated in an infinite-dimensional setting, the results are new also for finite-dimensional systems, which form a special case of our setting. Besides being an important structural property of the optimal triplet, turnpike properties can be leveraged in design of numerical methods. For example [36] suggests to exploit them in indirect shooting methods, in [16, 17] it is used for tailored discretization of infinite dimensional OCPs in a receding-horizon setting and [10] hinges on them in mixed-integer OCPs.

After introducing the optimal control problem, the first-order conditions and the functional analytic setting in Section 2, we deduce the desired turnpike property of the adjoints in Section 3 for exponentially stable and exactly controllable systems. Assuming that the underlying operator satisfies a spectrum decomposition assumption, we prove the result in Section 4 for stabilizable systems with finite-dimensional unstable part. In Section 5 we provide stronger estimates assuming more structure of the problem, i.e., that the underlying semigroup is analytic and hence the dynamics are given by a parabolic evolution equation. Further in Section 6 we discuss and give sufficient conditions for our main assumption, i.e., interval turnpike behavior of state and control. Finally in Section 7 we present a numerical example with a boundary controlled semilinear heat equation on a two-dimensional domain.

2. Setting and preliminaries

Let $(X, \| \cdot \|_X)$ be a Banach space and $(U, \langle \cdot, \cdot \rangle_U)$ be a Hilbert space with corresponding norm $\| \cdot \|_U$. Consider the optimal control problem

$$
(2.1) \quad \begin{aligned}
&\min_{u \in L_2(0,T;U)} \ \int_0^T J(x(t), u(t)) \, dt \\
&\text{s.t.} \quad \dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t) + f(x(t), u(t)) \\
&\qquad x(0) = x_0,
\end{aligned}
$$

where $J : X \times U \to \mathbb{R}$ is sufficiently smooth, $x_0 \in X$, $\mathcal{A} : D(\mathcal{A}) \subset X \to X$ generates a strongly continuous semigroup on X, $\mathcal{B} \in L(U, X)$, and $f : X \times U \to X$ is a sufficiently smooth, locally Lipschitz nonlinearity. We will assume that the above problem has at least one optimal solution $(x, u) \in C(0, T; X) \times L_2(0, T; U)$, cf. [24, Chap. 3]. Additionally, we consider $(\bar{x}, \bar{u}) \in X \times U$ to be an optimal solution of the corresponding steady state system, i.e., $(\bar{x}, \bar{u})$ solves

$$
(2.2) \quad \begin{aligned}
&\min_{u \in U} \ J(x, u) \\
&\text{s.t.} \quad 0 = \mathcal{A}x + \mathcal{B}u + f(x, u),
\end{aligned}
$$

Our goal in this paper is to find conditions under which interval turnpike behavior of the states and control inputs implies interval turnpike behavior of

the adjoints. Our basic assumption on the behavior of the optimal solutions is thus the following.

Assumption 2.1 (Interval turnpike property for states and controls)**.** Let (x, u) be an optimal solution to (2.1). We assume there are strictly monotonously increasing functions $t_1, t_2 : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ with $0 \leq t_1(T) \leq t_2(T) \leq T$ such that $\nu(T) := t_2(T) - t_1(T)$ is strictly monotonously increasing and unbounded and for each $\varepsilon > 0$ there is $T_0 > 0$ such that

$$\|x(t) - \bar{x}\|_X + \|u(t) - \bar{u}\|_U \leq \varepsilon \qquad \forall t \in [t_1(T), t_2(T)], \; T \geq T_0.$$

Note that this bound immediately implies $u \in L_\infty(t_1(T), t_2(T); U)$. We also note that our definition of the turnpike property implies uniform convergence of x to $\bar{x}$ and u to $\bar{u}$ on $[t_1(T), t_2(T)]$. This means that in general the lengths of the intervals $[0, t_1(T)]$ and $[t_2(T), T]$ have to tend to infinity as $T \to \infty$, as these intervals contain the approaching and the leaving arc of the optimal trajectory, respectively, outside an ε-neighborhood of the equilibrium with $\varepsilon \to 0$ as $T \to \infty$, see Remark 2.2 for an example. We emphasize that we do not require any bound on the relation between the lengths of the near-turnpike interval $[t_1(T), t_2(T)]$ and the overall interval $[0, T]$. However, if we have such a bound, then we can take it into account in our analysis, see Remarks 3.3 and 3.8, below.

Remark 2.2. We say that $(x, u) \in C(0, T; X) \times L_2(0, T; U)$ satisfies the exponential turnpike property, if there is a constant $c > 0$ and a decay parameter $\mu > 0$, both independent of T such that we have

$$\|x(t) - \bar{x}\|_X + \|u(t) - \bar{u}\|_U \leq c \left(e^{-\mu t} + e^{-\mu(T-t)} \right).$$

If this inequality holds, it can be easily seen, cf. [18, Discussion after Rem. 6.3] that we can choose $\delta \in (0, \frac{1}{2})$ such that for each $\varepsilon > 0$ there is a horizon T such that

$$\|x(t) - \bar{x}\|_{L_2(\delta T, (1-\delta)T; X)} + \|u(t) - \bar{u}\|_{L_2(\delta T, (1-\delta)T; U)} \leq \varepsilon$$

and

$$\|x(t) - \bar{x}\|_{C(\delta T, (1-\delta)T; X)} + \|u(t) - \bar{u}\|_{L_\infty(\delta T, (1-\delta)T; U)} \leq \varepsilon,$$

i.e., L_2 and uniform convergence on a fixed part of the time interval $[0, T]$ for $T \to \infty$. Thus, Assumption 2.1 is satisfied with $t_1(T) = \delta T$, $t_2(T) = (1-\delta)T$ and $\nu(T) = (1 - 2\delta)T$.

By the corresponding necessary optimality conditions of above problem (2.1), cf. [24, Chap. 4], there is an adjoint state $\lambda \in C(0, T; X)$ such that

$$
\begin{aligned}
\dot{\lambda}(t) &= -(\mathcal{A} + f_x(x(t), u(t)))^* \lambda(t) + J_x(x(t), u(t)) \\
0 &= (\mathcal{B} + f_u(x(t), u(t)))^* \lambda(t) + J_u(x(t), u(t)) \\
\dot{x}(t) &= \mathcal{A}x(t) + \mathcal{B}u(t) + f(x(t), u(t)),
\end{aligned}
$$

(2.3)

where $x(0) = x_0$ and $\lambda(T) = 0$. Analogously, by the necessary optimality conditions of the steady state problem there is an adjoint state $\bar{\lambda} \in X$ such that

$$0 = -(\mathcal{A} + f_x(\bar{x}, \bar{u}))^* \bar{\lambda} + J_x(\bar{x}, \bar{u})$$
$$0 = (\mathcal{B} + f_u(\bar{x}, \bar{u}))^* \bar{\lambda} + J_u(\bar{x}, \bar{u})$$
$$0 = \mathcal{A}\bar{x} + \mathcal{B}\bar{u} + f(\bar{x}, \bar{u}).$$

Our goal in this paper is to show the interval turnpike property of the adjoint λ.

Definition 2.3 (Interval turnpike property for adjoint states)**.** We say that the adjoint λ satisfies the interval turnpike property at the steady state adjoint $\bar{\lambda}$, if there are strictly monotonously increasing functions $s_1, s_2 : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ with $0 \leq s_1(T) \leq s_2(T) \leq T$ such that $\theta(T) := s_2(T) - s_1(T)$ is strictly monotonously increasing and unbounded and for each $\varepsilon > 0$ there is $T_0 > 0$ such that

$$\|\lambda(t) - \bar{\lambda}\|_Y \leq \varepsilon \qquad \forall t \in [s_1(T), s_2(T)], \, T \geq T_0.$$

For our analysis, we define the remainder terms

$$r_f(t) := f(x(t), u(t)) - f(\bar{x}, \bar{u})$$

and for $\star \in \{x, u\}$

$$r_{f_\star}(t) := f_\star(x(t), u(t)) - f_\star(\bar{x}, \bar{u}),$$
$$r_{J_\star}(t) := J_\star(x(t), u(t)) - J_\star(\bar{x}, \bar{u}).$$

Thus, denoting $A := \mathcal{A} + f_x(\bar{x}, \bar{u})$, $B := \mathcal{B} + f_u(\bar{x}, \bar{u})$, and $(\delta x, \delta u, \delta \lambda) := (x - \bar{x}, u - \bar{u}, \lambda - \bar{\lambda})$, we have that

(2.4) $$\dot{\delta\lambda}(t) = -A^* \delta\lambda(t) - r_{f_x}(t)^* \lambda(t) + r_{J_x}(t)$$

(2.5) $$0 = B^* \delta\lambda(t) + r_{f_u}(t)^* \lambda(t) + r_{J_u}(t)$$

(2.6) $$\dot{\delta x}(t) = \mathcal{A}\delta x + \mathcal{B}\delta u + r_f(t)$$

with $\delta x(0) = x_0 - \bar{x}$ and $\delta\lambda(T) = -\bar{\lambda}$. It is clear that the solutions x, u, and λ of eq. (2.3) depend on T and hence also δx, δu, and $\delta\lambda$ do. However, for the sake of readability, we do not explicitly indicate this dependence. We note that, using the definition of δx, δu, and $\delta\lambda$, the inequalities from Assumption 2.1 and Definition 2.3 can be written as $\|\delta x(t)\|_X + \|\delta u(t)\|_U \leq \varepsilon$ and $\|\delta\lambda(t)\|_Y \leq \varepsilon$, respectively.

Remark 2.4 (Extension to box constraints on the control)**.** We briefly discuss how the case of box constraints on the control can be treated analogously to the unconstrained case, if one assumes that the turnpike lies in the interior of the constraints. To this end, in the dynamic problem (2.1) we add the constraints

(2.7) $$u_a \leq u(t) \leq u_b$$

for a.e. $t \in [0, T]$ and in the steady state problem (2.2) we add the constraints

$$(2.8) \qquad u_a \leq \bar{u} \leq u_b$$

for $u_a, u_b \in U$. In the case of box constraints, standard assumptions and the classical methods of calculus of variations, cf., e.g., [37, Chap. 5], [21] or [25], ensure that, besides λ, there exist two additional multipliers $0 \leq \mu_a, \mu_b \in L_2(0, T; U)$ such that the stationarity condition, i.e., the second line of (2.3) becomes

$$0 = (\mathcal{B} + f_u(x(t), u(t)))^* \lambda(t) + J_u(x(t), u(t)) + \mu_a(t) - \mu_b(t),$$
$$(2.9) \qquad 0 = \langle \mu_a, u - u_a \rangle_{L_2(0, T; U)} \qquad 0 = \langle \mu_b, u_b - u \rangle_{L_2(0, T; U)},$$
$$u_a \leq u(t) \leq u_b,$$

for a.e. $t \in [0, T]$. If we assume that the constraint is not active at an optimal solution of the steady state problem (2.2), i.e., the inequalities in (2.8) are strict, then Assumption 2.1 assures that we can choose $\varepsilon > 0$ such that also the solution of the dynamic OCP does not touch the constraints on $[t_1(T), t_2(T)]$, i.e., (2.7) is also strict on this interval. In that case, the complementarity condition (2.9) assures that $\mu_a(t) = \mu_b(t) = 0$ for a.e. $t \in [t_1(T), t_2(T)]$ and the dynamics reduce to the unconstrained case (2.3) on the subinterval $[t_1(T), t_2(T)]$.

In our subsequent analysis, we will exploit that, due to Assumption 2.1 and continuity, the remainder terms r_f, $r_{f_\star}$ and $r_{J_\star}$ defined above approach zero for $t \in [t_1(T), t_2(T)]$. In order to make this property rigorous in the appropriate function spaces, we introduce superposition operators. Intuitively, a superposition operator is a nonlinear map between function spaces induced by a given nonlinear function defined on, e.g., finite-dimensional spaces by superposition. We refer the interested reader to [37, Sec. 4.3.3] for a short introduction and [2, 11] for an in-depth treatment of these topics in Sobolev and Lebesgue spaces of abstract functions. In order to not hide the main steps behind technical details, we only consider the case of scalar nonlinearities here.

Definition 2.5 (Superposition operator). Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$. Consider a mapping $\varphi : \mathbb{R} \to \mathbb{R}$. Then the mapping Φ defined by

$$\Phi(x)(\omega) = \varphi(x(\omega)) \qquad \text{for } \omega \in \Omega$$

assigns to a function $x : \Omega \to \mathbb{R}$ a new function $z : S \to \mathbb{R}$ via the relation $z(\omega) = \varphi(x(\omega))$ for $\omega \in \Omega$ and is called a Nemytskij operator or superposition operator.

An immediate question that arises is the following: Given a function $x \in L_p(\Omega)$, which integrability does the image $\underline{\Phi}(x)$ have? It turns out that in case $p < \infty$, this is coupled to growth assumptions on the underlying nonlinearity. It is to be expected as, e.g., for $\varphi(x) = x^3$, the corresponding superposition operator maps $L_{3p}(\Omega)$ to $L_p(\Omega)$.

Proposition 2.6. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous. For $1 \leq p, q < \infty$ let*

$$(2.10) \qquad |\varphi(s)| \leq c_1 + c_2 |s|^{\frac{p}{q}} \qquad \forall\, s \in \mathbb{R}$$

for constants $c_1 \in \mathbb{R}$ and $c_2 \geq 0$. Then the corresponding superposition operator Φ maps $L_p(\Omega)$ into $L_q(\Omega)$. Additionally, it is continuous as a nonlinear map from $L_p(\Omega)$ to $L_q(\Omega)$, i.e., if $\|x - z\|_{L_p(\Omega)} \to 0$, we have that

$$\|\Phi(x) - \Phi(z)\|_{L_q(\Omega)} \to 0.$$

Proof. See [11, Thm. 1] and [11, Thm. 4]. $\qquad\qquad\qquad\qquad\qquad\square$

It can also be shown that the assumptions of Proposition 2.6 are not only sufficient for continuity, but also necessary, cf. [11, Thm. 3.1]. Thus, e.g., for a cubic nonlinearity $f(x, u) = -x^3$ and assuming that the state and control approach the turnpike in some L_p-norm, the remainder term r_f will vanish in the $L_{p/3}$-norm. We now formulate the main assumption considering the continuity of the remainder terms.

Assumption 2.7. We assume that there is a real Hilbert space $(Y, \langle \cdot, \rangle_Y)$ with corresponding norm $\|\cdot\|_Y$ such that the superposition operators induced by the remainder terms $r^*_{f_x}(t)$ and $r_{J_x}(t)$ for any $t \in [0, T]$ are continuous from X to $L(X, Y)$ and X to Y respectively. Additionally, we assume that the superposition operators corresponding to the remainder terms $r^*_{f_u}(t)$ and $r_{J_u}(t)$ for any $t \in [0, T]$ are continuous from U to $L(U, Y)$ and U to $U^* \simeq U$ respectively.

Remark 2.8. In the finite-dimensional setting with $X = Y = \mathbb{R}^n$ and $U = \mathbb{R}^m$, $n, m \in \mathbb{N}$, the superposition operator concept is not needed and the subsequent results will hold for all Lipschitz nonlinearities. In the infinite-dimensional setting the assumption on continuity of the superposition operators corresponding to $r_{f_u}(t)$ and $r_{J_u}(t)$ allows, e.g., for $Y = L_2(\Omega)$ if the cost functional is quadratic in the control and the dynamics include a polynomial nonlinearity in the control, if U is embedded in a regular L_p space with large p. The continuity of the superposition operators corresponding to $r_{f_x}(t)$ and $r_{f_u}(t)$ can be verified if the state space X is sufficiently regular and embedded into an L_p-space with p large and the nonlinearity is, e.g., polynomial in x. Additionally if the superposition operator corresponding to $f_x(\bar{x}, \bar{u})^*$ can be extended to a compact operator from the domain of $\mathcal{A}^*$ in Y to Y and if the semigroup generated by $\mathcal{A}^*$ is exponentially stable, the perturbed operator $A^* = (\mathcal{A} + f_x(\bar{x}, \bar{u}))^*$ generates a semigroup on Y, cf. [8, Sec. III.2] and Section 7.

We assume that $A^* = (\mathcal{A} + f_x(\bar{x}, \bar{u}))^*$ generates a strongly continuous semigroup $(\mathcal{T}^*(t))_{t \geq 0}$ on Y, $B \in L(U, Y)$ and that $\bar{\lambda} \in Y$ and whenever we refer to a solution of (2.4), we mean it in the mild sense, i.e., for the adjoint, we have the variation of constants formula, cf. [28, Sec. 4.2],

$$(2.11) \quad \delta\lambda(t) = \mathcal{T}^*(T - t)\delta\lambda(T) + \int_t^T \mathcal{T}^*(s - t)\left(r_{f_x}(s)^* \lambda(s) + r_{J_x}(s)\right)\, ds.$$

3. Stable or exactly controllable systems

We first analyze two particular cases, to which we will reduce more general systems in Section 4: On the one hand the case where A^* generates an exponentially stable semigroup on Y and on the other hand the case of (A, B) being exactly controllable.

Theorem 3.1 (Adjoint turnpike for exponentially stable A^*). *Let Assumption 2.7 hold. Let (x, u) satisfy the interval turnpike property of Assumption 2.1 with the intervals $[t_1(T), t_2(T)]$ and assume that the adjoints satisfy $\rho := \sup_{T \geq 0} \|\lambda\|_{C(t_1(T), t_2(T); X)} < \infty$. Assume that A^* generates an exponentially stable semigroup $(\mathcal{T}^*(t))_{t \geq 0}$ on Y. Then λ satisfies the interval turnpike property from Definition 2.3.*

Proof. First, we write the adjoint equation, i.e., the first equation of (2.4), by means of the variation of constants formula (2.11) for $t \in [t_1(T), t_2(T)]$ on $[t, t_2(T)]$

$$\delta\lambda(t) = \mathcal{T}^*(t_2(T) - t)\delta\lambda(t_2(T)) + \int_t^{t_2(T)} \mathcal{T}^*(s - t) \left(r_{f_x}(s)^* \lambda(s) + r_{J_x}(s) \right) \, ds.$$

By exponential stability of the semigroup there is $M \geq 1$ and $\mu > 0$ such that $\|\mathcal{T}^*(t)\|_{L(Y,Y)} \leq Me^{-\mu t}$ for all $t \geq 0$. This implies the existence of $c > 0$ such that the estimate

$$
\begin{aligned}
\|\delta\lambda(t)\|_Y &\leq Me^{-\mu(t_2(T)-t)}\|\delta\lambda(t_2(T))\|_Y \\
&\quad + c \left(\|r_{f_x}\|_{C(t,t_2(T); L(X,Y))}\rho + \|r_{J_x}\|_{C(t,t_2(T);Y)} \right)
\end{aligned}
$$

holds for all $t \in [t_1(T), t_2(T)]$. Setting $s_1(T) := t_1(T)$, $s_2(T) = t_1(T) + t_2(T) - t_1(T))/2$, and recalling that $t_2(T) - t_1(T) \to \infty$ as $T \to \infty$, we obtain for sufficiently large T that $Me^{-\mu(t_2(T)-t)} \leq 1/2$ for all $t \in [s_1(T), s_2(T)]$. This implies

$$\|\delta\lambda(t)\|_Y \leq 2c \left(\|r_{f_x}\|_{C(t,t_2(T); L(X,Y))}\rho + \|r_{J_x}\|_{C(t,t_2(T);Y)} \right)$$

for all $t \in [s_1(T), s_2(T)]$. The assertion follows since $\|r_{f_x}\|_{C(t,t_2(T); L(X,Y)} \to 0$ and $\|r_{J_x}\|_{C(t,t_2(T);Y)} \to 0$ as $T \to \infty$ due to Assumptions 2.1 and 2.7. $\square$

Remark 3.2. If we add a term $\sigma_T(t)$ with $\|\sigma_T\|_{C(t,t_2(T);Y)} < \infty$ on the right hand side of (2.4), then a straightforward extension of the proof shows that for all sufficiently large T we obtain

$$\|\delta\lambda(t)\|_Y \leq \varepsilon + 2c\|\sigma_T\|_{C(t,t_2(T);Y)}$$

for all $t \in [s_1(T), s_2(T)]$.

Remark 3.3. If the times $t_1(T)$ and $t_2(T)$ in the interval turnpike property for state and control satisfy $(t_2(T) - t_1(T))/T \geq C(T)$ for a function $C : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, then the times $s_1(T)$ and $s_2(T)$ of the adjoint turnpike property satisfy $(s_2(T) - s_1(T))/T \geq C(T)/2$.

Next, we discuss the case of (A, B) being exactly controllable.

Definition 3.4 (Exact and approximate controllability, [6, Def. 4.1.3])**.** For any $\tau \in [0, T]$, we define the controllability map $\phi_\tau : L_2(0, \tau; U) \to Y$ by

$$\phi_\tau u := \int_0^\tau \mathcal{T}(\tau - s) B u(s) \, ds.$$

We call (A, B) exactly controllable in time $t_c > 0$ if $\text{ran } \phi_{t_c} = Y$. Similarly, we call (A, B) approximately controllable in time t_c if $\overline{\text{ran } \phi_{t_c}} = Y$.

It is clear that exact and approximate controllability coincide in finite-dimensions. An important characterization of controllability is the following observability inequality, which was proven first in the seminal paper [26] with the Hilbert Uniqueness Method.

Theorem 3.5 ([6, Thm. 4.1.7])**.** (A, B) *is exactly controllable in time* $t_c > 0$ *if and only if there is* $\alpha_{t_c} > 0$ *such that*

$$\int_0^{t_c} \|B^* \mathcal{T}^*(s) x_0\|_U^2 \, ds \geq \alpha_{t_c} \|x_0\|_Y^2 \quad \forall x_0 \in Y.$$

Using substitution in the previous estimate we immediately obtain that

(3.1)
$$\int_{t-t_c}^t \|B^* \mathcal{T}^*(t - s) \delta\lambda(t)\|_U^2 \, ds \geq \alpha_{t_c} \|\delta\lambda(t)\|_Y^2 \quad \forall \delta\lambda(t) \in Y, \quad t \in [t_c, T]$$

This enables us to derive the following bound on $\|\delta\lambda(t)\|_Y$.

Theorem 3.6. *Let* (A, B) *be exactly controllable in time* $t_c > 0$. *Then there is* $c > 0$ *independent of* T, *such that*

$$\|\delta\lambda(t)\|_Y^2 \leq c \int_{t-t_c}^t \|r_{f_u}(s)^* \lambda(s) + r_{J_u}(s)\|_U^2 + \| - r_{f_x}(s)^* \lambda(s) + r_{J_x}(s)\|_Y^2 \, ds.$$

Proof. The proof of this estimate is inspired by [30, Proof of Rem. 2.1], where the finite-dimensional case is considered. We decompose $\delta\lambda = \delta\lambda_1 + \delta\lambda_2$, where for any $s < t$

$$\delta\lambda_1'(s) = -A^* \delta\lambda_1(s), \qquad\qquad\qquad \delta\lambda_1(t) = \delta\lambda(t),$$
$$\delta\lambda_2'(s) = -A^* \delta\lambda_2(s) - r_{f_x}(s)^* \lambda(s) + r_{J_x}(s), \qquad \delta\lambda_2(t) = 0,$$

and apply the observability estimate (3.1) to $\delta\lambda_1(s) = \mathcal{T}^*(t - s) \delta\lambda(t)$. This yields

$$\alpha_{t_c} \|\delta\lambda(t)\|_Y^2 \leq \int_{t-t_c}^t \|B^* \delta\lambda_1(s)\|_U^2 \, ds \leq \int_{t-t_c}^t \|B^* \delta\lambda(s)\|_U^2 + \|B^* \delta\lambda_2(s)\|_U^2 \, ds.$$

Further, we estimate

$$\int_{t-t_c}^t \|B^* \delta\lambda_2(s)\|_U^2 \, ds \leq \int_{t-t_c}^t \|B^* \int_s^t \mathcal{T}^*(\tau - s)(-r_{f_x}(\tau)^* \lambda(\tau)$$
$$+ r_{J_x}(\tau)) \, d\tau\|_U^2 \, ds$$
$$\leq c(t_c) \int_{t-t_c}^t \| - r_{f_x}(s)^* \lambda(s) + r_{J_x}(s)\|_Y^2 \, ds.$$

Finally, by (2.5), we have that

$$\int_{t-t_c}^{t} \|B^* \delta\lambda(s)\|_U^2 = \int_{t-t_c}^{t} \|r_{f_u}(s)^*\lambda(s) + r_{J_u}(s)\|_U^2 \, ds,$$

which concludes the proof. □

Since the right hand side of the inequality from Theorem 3.6 obviously tends to zero if the integrands tend to zero, we immediately obtain the following corollary.

Corollary 3.7. *Let Assumption 2.7 hold and let (A, B) be exactly controllable in time $t_c > 0$. Let (x, u) satisfy the turnpike property of Assumption 2.1 with the intervals $[t_1(T), t_2(T)]$ and assume that the adjoints satisfy $\rho := \sup_{T \geq 0} \|\lambda\|_{C(t_1(T), t_2(T); Y)} < \infty$. Then λ satisfies the interval turnpike property from Definition 2.3 with $s_1(T) = t_1(T) + t_c$ and $s_2(T) = t_2(T)$.*

Proof. Follows immediately from Theorem 3.6. □

Remark 3.8. Under the assumptions of Remark 3.3, for the times from Corollary 3.7 the estimate $(s_2(T) - s_1(T))/T \geq C(T) - t_c/T$ holds.

Remark 3.9. Similar to Remark 3.2, it is easily seen from the proof of Theorem 3.6 that if we add a term $\sigma_T(t)$ with $\|\sigma_T\|_{C(t-t_c, t; Y)} \leq \bar{\sigma}_T < \infty$ and $\rho_T(t)$ on the right hand sides of (2.4) and (2.5), respectively, then the result of Theorem 3.6 changes to

$$(3.2) \qquad \|\delta\lambda(t)\|_Y^2 \leq c \int_{t-t_c}^{t} \|r_{f_u}(s)^*\lambda(s) + r_{J_u}(s) + \sigma_T(s)\|_U^2$$

$$+ \| - r_{f_x}(s)^*\lambda(s) + r_{J_x}(s) + \rho_T(s)\|_Y^2 \, ds.$$

We then obtain as a counterpart for the inequality in Definition 2.3 the bound

$$\|\delta\lambda(t)\|_Y \leq \varepsilon + c(\bar{\sigma}_T + \bar{\rho}_T) \qquad \forall t \in [s_1(T), s_2(T)], \, T \geq T_0.$$

4. Stabilizable systems with finite-dimensional unstable part

In this section we extend our results to exponentially detectable (A^*, B^*), where the unstable part of A^* is finite-dimensional and B^* has finite rank. We note that this includes all finite-dimensional systems with stabilizable linearization. In order to define the correct setting for infinite-dimensional systems, we present the spectrum decomposition assumption as follows.

Definition 4.1 ([6, Def. 5.2.5]). Denoting $\sigma^+(A) := \sigma(A) \cap \{s \in \mathbb{C} : \mathrm{Re}\, s \geq 0\}$ and $\sigma^-(A) = \sigma(A) \cap \{s \in \mathbb{C} : \mathrm{Re}\, s < 0\}$, an operator A satisfies the spectral decomposition assumption if $\sigma^+(A)$ is bounded and separated from $\sigma^-(A)$ in such a way that a rectifiable, simple, closed curve Γ can be drawn so as to enclose an open set containing $\sigma^+(A)$ in its interior and $\sigma^-(A)$ in its exterior.

Remark 4.2. Classes of operators satisfying the spectrum decomposition assumption include, e.g., delay equations [6, Sec. 2.4] and Riesz-spectral operators with a pure point spectrum and only finitely many eigenvalues in $\sigma^+(A)$. More concrete examples of the latter are compact perturbation of the Laplace operator, i.e., $A = \Delta + c^2 I$ for $c \in \mathbb{R}$ or models of damped vibrations such as

$$A = \begin{bmatrix} 0 & I \\ -A_0 & -D \end{bmatrix}$$

where A_0 is a positive operator and D is an unbounded damping operator (see, e.g.,[22] and the Euler-Bernoulli example with Kelvin-Voigt damping).

If A^* satisfies the decomposition assumption, by [6, Lem. 2.5.7] the decomposition of the spectrum induces a corresponding decomposition of Y. Defining the spectral projection P by

$$P y_0 := \frac{1}{2\pi i} \int_\Gamma (sI - A^*)^{-1} y_0 \, ds$$

for $y_0 \in Y$, where Γ from Definition 4.1 is traversed once in the positive direction, we obtain the decomposition $Y = Y_u \oplus Y_s$, where $Y_u = PY$ and $Y_s = (I - P)Y$. Moreover, the spectral projection yields a linear coordinate transform such that the pair (A^*, B^*) can be transformed into the form

$$(4.1) \qquad \widetilde{A^*} = \begin{bmatrix} A_u^* & 0 \\ 0 & A_s^* \end{bmatrix}, \qquad \widetilde{B^*} = \begin{bmatrix} B_u^* & B_s^* \end{bmatrix}$$

where $A_u^*, B_u^*, A_s^*, B_s^*$ are restrictions of A^* and B^* to Y_u and Y_s, respectively. Note that A_u^* and B_u^* are bounded operators. We impose the following assumption on A^*.

Assumption 4.3. A^* satisfies the spectrum decomposition assumption such that it has the decomposition according to (4.1), where A_u^* is finite-dimensional and A_s^* is exponentially stable.

If we split up the transformed adjoint accordingly via

$$(4.2) \qquad \widetilde{\delta\lambda} = \begin{pmatrix} \delta\lambda_u \\ \delta\lambda_s \end{pmatrix}, \widetilde{\lambda} = \begin{pmatrix} \lambda_u \\ \lambda_s \end{pmatrix},$$

then the equations (2.4) and (2.5) attain the form

$$(4.3) \qquad \delta\dot{\lambda}_u = -A_u^* \delta\lambda_u - \tilde{r}_1^* \lambda_u - \tilde{r}_2^* \lambda_s + \tilde{r}_3$$

$$(4.4) \qquad \delta\dot{\lambda}_s = -A_s^* \delta\lambda_s - \tilde{r}_7^* \lambda_u - \tilde{r}_8^* \lambda_s + \tilde{r}_9$$

$$(4.5) \qquad 0 = B_u^* \delta\lambda_u + B_s^* \delta\lambda_s + \tilde{r}_4 \lambda_u + \tilde{r}_5 \lambda_s + \tilde{r}_6.$$

Here, the terms $\tilde{r}_j$ are derived via coordinate transformation and splitting from the remainder terms in (2.4)–(2.6) and—up to multiplication by appropriate constants—satisfy the same estimates as these remainder terms. Using this decomposition, we can prove the following theorem.

Theorem 4.4 ((A^*, B^*) exponentially detectable, B^* has finite rank). *Let Assumption 2.7 hold. Let (x, u) satisfy the turnpike property of Assumption 2.1 on $[t_1(T), t_2(T)]$ and assume that $\rho := \sup_{T \geq 0} \|\lambda\|_{C(t_1(T), t_2(T); Y)} < \infty$. Let Assumption 4.3 hold and further assume that B^* has finite rank and (A^*, B^*) is exponentially detectable. Then λ satisfies the interval turnpike property from Definition 2.3.*

Proof. First note that the claimed property holds for $\delta\lambda$ if and only if it holds for the transformed adjoint $\widetilde{\delta\lambda}$ from (4.2). The property for $\widetilde{\delta\lambda}$, in turn, holds if and only if it holds for the two components $\delta\lambda_u$ and $\delta\lambda_s$. Moreover, note that the assumed exponential detectability of (A^*, B^*) implies that the finite-dimensional pair (A_u^*, B_u^*) is (exponentially) detectable, and hence, (exactly) observable by Hautus rank condition [6, Def. 1.2.6].

We start by applying the extension of Theorem 3.1 described in Remark 3.2 to $\delta\lambda_s$, with $\sigma_T = -\tilde{r}_7^* \lambda_u$. We note that the fact that equation (4.5) contains additional terms compared to equation (2.5) does not affect the applicability of Theorem 3.1 and Remark 3.2, because equation (2.5) is not used in its proof. Due to the uniform boundedness of λ which implies uniform boundedness of λ_u, σ_T tends to zero as $T \to \infty$ on $[t_1(T), t_2(T)]$. Hence, we obtain the desired property for $\delta\lambda_s$ on an interval $[\tilde{s}_1(T), \tilde{s}_2(T)]$. We note that by the construction in the proof of Theorem 3.1 we obtain $[\tilde{s}_1(T), \tilde{s}_2(T)] \subset [t_1(T), t_2(T)]$.

Now for $\delta\lambda_u$ we use the extension of Theorem 3.6 described in Remark 3.9 with $\sigma_T = -\tilde{r}_2^* \lambda_s$ and $\rho_T = B_s^* \delta\lambda_s - \tilde{r}_5^* \lambda_s$, on $[\tilde{s}_1(T), \tilde{s}_2(T)]$. Since all terms become arbitrarily small on $[\tilde{s}_1(T), \tilde{s}_2(T)]$ as $T \to \infty$, we obtain the desired estimate for $\delta\lambda_u$ on $[s_1(T), s_2(T)]$ with $s_1(T) = \tilde{s}_1(T) + t_c$ and $s_2(T) = \tilde{s}_2(T)$. $\square$

Remark 4.5. Note that if B^* has finite rank and Assumption 4.3 is satisfied, (A^*, B^*) is exponentially detectable in particular if (A^*, B^*) is approximately observable: in that case both (A_s^*, B_s^*) and (A_u^*, B_u^*) are approximately observable [33, Lem. 9.7.2], which for the finite-dimensional pair (A_u^*, B_u^*) coincides with exact observability.

Remark 4.6. For linear quadratic problems, detectability and stabilizability imply an exponential turnpike property and thus an interval turnpike property for states, controls and adjoints for a very general class of infinite dimensional systems [18, Thm. 17].

Remark 4.7. The assumptions on (A^*, B^*) in Theorem 4.4 are in particular satisfied if (A^*, B^*) is finite-dimensional and detectable. In that case, the result can be alternatively proven by using the decomposition into the observable and nonobservable subspaces of (A^*, B^*), which is of the form

$$\widetilde{A^*} = \begin{bmatrix} A_1^* & 0 \\ A_2^* & A_3^* \end{bmatrix}, \qquad \widetilde{B^*} = \begin{bmatrix} B_1^* & 0 \end{bmatrix},$$

where (A_1^*, B_1^*) is (exactly) observable and A_3^* is (exponentially) stable (the unstable subspace is contained in the observable subspace by the Hautus

rank condition). Then the equations for $\widetilde{\delta\lambda}$ become

$$(4.6) \qquad \dot{\delta\lambda_1} = -A_1^*\delta\lambda_1 - \tilde{r}_1^*\lambda_1 - \tilde{r}_2^*\lambda_2 + \tilde{r}_3$$

$$(4.7) \qquad \dot{\delta\lambda_2} = -A_2^*\delta\lambda_1 - A_3^*\delta\lambda_2 - \tilde{r}_7^*\lambda_1 - \tilde{r}_8^*\lambda_2 + \tilde{r}_9$$

$$(4.8) \qquad 0 = B_1^*\delta\lambda_1 + \tilde{r}_4\lambda_1 + \tilde{r}_5\lambda_2 + \tilde{r}_6.$$

For proving Theorem 4.4 we can then proceed similarly as above, but in reverse order: We first obtain the desired estimate for $\delta\lambda_1$ using Theorem 3.6 and Remark 3.9 and then obtain the estimate for $\delta\lambda_2$ using Theorem 3.1, Remark 3.2 and the estimate for $\delta\lambda_1$ from the first part of the proof.

Note that the observability decomposition has very limited use for infinite-dimensional systems as the nonobservable subspace is the largest $T^*(t)$-invariant subspace in Y contained in $\ker B^*$ [6, Lem. 4.1.18]. For example, if (A^*, B^*) is approximately observable, then the nonobservable subspace is the trivial subspace $\{0\}$ and the observability decomposition is redundant.

5. Particular case of an analytic semigroup

We briefly present a refined version of Theorem 3.1 for the case where A^* generates an analytic semigroup. In this case, we can improve the estimate in Definition 2.3 by using stronger norms. To this end, we define the space

$$W^{A^*}(0,T) := \left\{ v \in L_2(0,T;Y) : v' \in L_2(0,T;Y) \right\} \cap L_2(0,T;D(A^*)).$$

endowed with the norm

$$\|v\|_{W^{A^*}(0,T)} := \|v'\|_{L_2(0,T;Y)}^2 + \|v\|_{L_2(0,T;Y)}^2 + \|A^*v\|_{L_2(0,T;Y)}^2.$$

It was shown in, e.g., [3, Pt. II-1, Rem. 4.2] that

$$(5.1) \qquad W^{A^*}(0,T) \hookrightarrow C(0,T;(D(A^*),Y)_{\frac{1}{2}}),$$

with embedding constant $c_E > 0$ independently of T, where $(D(A^*),Y)_{\frac{1}{2}}$ denotes the real interpolation space as defined in [3, Pt. II-1, Sec. 4.3]. If A^* generates an exponentially stable semigroup, this interpolation space can be shown to be isomorphic to the domain of the fractional power $(A^*)^{\frac{1}{2}}$ in many applications, cf. [23, Sec. 0.2.1].

Theorem 5.1. *Let the assumptions of Theorem 3.1 hold. Assume additionally that the semigroup $(\mathcal{T}^*(t))_{t\geq 0}$ is analytic. Suppose that for each $\tilde{\varepsilon} > 0$ there is $\tilde{T}_0 > 0$ such that*

$$\int_{t_1(T)}^{t_2(T)} \|r_{f_x}(s)\|_{L(X,Y)} + \|r_{J_x}(s)\|_Y \, ds \leq \tilde{\varepsilon} \qquad \forall T \geq \tilde{T}_0.$$

Then the interval turnpike property for adjoints from Definition 2.3 holds and, in addition, for each $\varepsilon > 0$ there is $T_0 > 0$ such that the adjoints satisfy

$$\|\delta\lambda\|_{W^{A^*}(t,t_2(T))} + \|\delta\lambda(t)\|_{(D(A^*),Y)_{\frac{1}{2}}} \leq \varepsilon \qquad \forall t \in [s_1(T), s_2(T)], T \geq T_0.$$

Proof. By exponential stability with decay rate $\mu > 0$, the scaled semigroup $(e^{\frac{\mu}{2}(t_2(T)-t)}\mathcal{T}^*(t_2(T)-t))_{t\geq 0}$ is still exponentially stable and its generator has the same domain as $\bar{A}^*$, cf. [8, Chap. III, Thm. 1.3]. Hence we can apply a well-known estimate for exponentially stable analytic semigroups, cf. [3, Pt. II-1, Prop. 3.7, Thm. 3.1] and for $\delta\lambda$ solving (2.4) we obtain that

$$\|\delta\lambda\|_{W^{A^*}(t,t_2(T))}$$

$$\leq c\left(\|-r_{f_x}^*\lambda + r_{J_x}\|_{L_2(t,t_2(T);Y)} + Me^{-\frac{\mu}{2}(t_2(T)-t)}\|\delta\lambda(t_2(T))\|_{(D(A^*),Y)_{\frac{1}{2}}}\right).$$

Using the boundedness of $\delta\lambda$ and the integral convergence of r_{f_x} and r_{J_x}, the first term approaches zero for $T \to \infty$. For the second term, we set $s_1(T) := t_1(T)$ and $s_2(T) = t_1(T) + (t_2(T) - t_1(T))/2$ and recalling that $t_2(T) - t_1(T) \to \infty$ as $T \to \infty$, there is T such that $Me^{-\frac{\mu}{2}(t_2(T)-t)} \leq \frac{c_E}{2c}$ for all $t \in [s_1(T), s_2(T)]$. Thus we can estimate

$$Me^{-\frac{\mu}{2}(t_2(T)-t)}\|\delta\lambda(t_2(T))\|_{(D(A^*),Y)_{\frac{1}{2}}} \leq \frac{c_E}{2c}\|\delta\lambda(t_2(T)))\|_{(D(A^*),Y)_{\frac{1}{2}}}.$$

Using the embedding (5.1) with T-independent embedding constant c_E and subtracting $\frac{c_E}{2}\|\delta\lambda(t_2(T)))\|_{(D(A^*),Y)_{\frac{1}{2}}}$ on both sides yields the result. $\square$

Remark 5.2. The assumption of the L_2-convergence of the remainder terms in Theorem 5.1 is satisfied if the convergence of state and control to the turnpike is, e.g., exponential as in Remark 2.2.

6. Discussion of assumptions

In this part we will briefly give sufficient conditions to render the interval turnpike of the primal variables as our main assumption Assumption 2.1 fulfilled. These conditions can be guaranteed by a combination of strict dissipativity, controllability, and stabilizability. We further provide an example where the boundedness of the adjoint $\sup_{T\geq 0}\|\lambda\|_{C(t_2(T),T;H^1(\Omega))} < \infty$ assumed in Theorem 3.1, Theorem 4.4, and Corollary 3.7 holds. First, we give a theorem stating that under suitable assumptions, measure turnpike implies interval turnpike.

Theorem 6.1. *Assume that the following conditions hold:*
   (i) *The system has the measure turnpike property, i.e., for each $\varepsilon > 0$ the Lebesgue measure of the set of times $t \in [0,T]$ for which*

$$\|x(t) - \bar{x}\|_X + \|u(t) - \bar{u}\|_U > \varepsilon$$

   holds is bounded independent of T.
   (ii) *The system is stabilizable at $\bar{x}$ with cost proportional to the initial distance to $\bar{x}$, i.e., there exists a constant $C > 0$, a neighborhood N_1 of $\bar{x}$ and a function $\eta \in \mathcal{K}$ such that for all $x_0 \in N_1$ and each $T > 0$ there is $u \in L_2(0,T;U)$ with $x(T) \in B_{C\|x_0-\bar{x}\|}(\bar{x})$ and $\int_0^T J(x,u)dt \leq TJ(\bar{x},\bar{u}) + \eta(\|x_0 - \bar{x}\|)$.*

(iii) *The optimal value function*

$$V_T(x_0) := \min_{u \in L_2(0,T;U)} \int_0^T J(x,u)dt$$

is approximately continuous in $\bar{x}$ uniformly in T in the following sense: there are $\gamma \in \mathcal{K}$, $\rho \in \mathcal{L}$ and a neighborhood N_2 of $\bar{x}$ such that

$$|V_T(x) - V_T(\bar{x})| \le \gamma(\|x - \bar{x}\|_X) + \rho(T)$$

for all $x \in N_2$ and all $T \ge 0$.

(iv) *Excursions from the turnpike are more expensive than staying in the turnpike, i.e., there is a function $\sigma \in \mathcal{K}_\infty$ and a neighborhood N_3 of $\bar{x}$ such that for each $T > 0$ and each trajectory x and control u satisfying $x(0), x(T) \in N_3$ the inequality*

$$\int_0^T J(x,u)dt \ge TJ(\bar{x},\bar{u}) - \sigma(\|x(0) - \bar{x}\|_X) - \sigma(\|x(T) - \bar{x}\|_X)$$
$$+ \max_{t \in [0,T]} \sigma(\|x(t) - \bar{x}\|_X)$$

holds.

Then the interval turnpike property from Assumption 2.1 holds.

Proof. For each $\tilde{\varepsilon} > 0$ that is sufficiently small such that $B_{\tilde{\varepsilon}}(\bar{x}) \subset N_1$ and $B_{C\tilde{\varepsilon}}(\bar{x}) \subset N_2 \cap N_3$ denote by $\mu(\tilde{\varepsilon})$ the bound on the Lebesgue measure from (i) for $\varepsilon = \tilde{\varepsilon}$. Then, for any $T \ge 8\mu(\tilde{\varepsilon})$ there exist times $\hat{t}_1(T,\tilde{\varepsilon}) \in [T/8, 3T/8]$ and $\hat{t}_2(T,\tilde{\varepsilon}) \in [5T/8, 7T/8]$ such that

$$\|x(t) - \bar{x}\|_X + \|u(t) - \bar{u}\|_U \le \tilde{\varepsilon}$$

for $t = \hat{t}_1(T,\tilde{\varepsilon})$ and $t = \hat{t}_2(T,\tilde{\varepsilon})$. The overall cost of the optimal solution can be written as

$$\int_0^T J(x,u)dt = \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(x,u)dt + \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(x,u)dt + \int_{\hat{t}_2(T,\tilde{\varepsilon})}^T J(x,u)dt,$$

where we omit the argument "(t)" in the integrands in order to shorten the notation. In addition, we consider the control $\hat{u}$ and corresponding trajectory $\hat{x}$ that is constructed as follows: it follows the optimal control until time $\hat{t}_1(T,\tilde{\varepsilon})$, then uses the control from (ii) with $x_0 = \hat{x}(\hat{t}_1(T,\tilde{\varepsilon}))$ from $\hat{t}_1(T,\tilde{\varepsilon})$ until $\hat{t}_2(T,\tilde{\varepsilon})$, and finally it uses the optimal control for horizon $T - \hat{t}_2(T,\tilde{\varepsilon})$ and initial condition $\hat{t}_2(T,\tilde{\varepsilon})$. The overall cost of this trajectory can be decomposed in the same way

$$\int_0^T J(\hat{x},\hat{u})dt = \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt + \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt + \int_{\hat{t}_2(T,\tilde{\varepsilon})}^T J(\hat{x},\hat{u})dt.$$

Now, by construction of $\hat{x}$ and $\hat{u}$ as well as (ii) and (iii) and the fact that tails of optimal trajectories are optimal trajectories, we get

(6.1) $$\int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt = \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(x,u)dt,$$

$$(6.2) \qquad \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt \le (\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) + \eta(\tilde{\varepsilon})$$

and

$$(6.3) \qquad \left| \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(\hat{x},\hat{u})dt - \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(x,u)dt \right| \le \gamma(C\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \rho(T - \hat{t}_2(T,\tilde{\varepsilon})).$$

Since x and u are optimal, we moreover obtain

$$\int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(x,u)dt + \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(x,u)dt + \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(x,u)dt$$

$$\le \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt + \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt + \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(\hat{x},\hat{u})dt,$$

implying

$$\int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(x,u)dt \le \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt - \int_0^{\hat{t}_1(T,\tilde{\varepsilon})} J(x,u)dt$$

$$+ \int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(\hat{x},\hat{u})dt$$

$$+ \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(\hat{x},\hat{u})dt - \int_{\hat{t}_2(T,\tilde{\varepsilon})}^{T} J(x,u)dt.$$

Inserting (6.1)–(6.3) into this inequality, we arrive at

$$\int_{\hat{t}_1(T,\tilde{\varepsilon})}^{\hat{t}_2(T,\tilde{\varepsilon})} J(x,u)dt$$

$$\le (\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) + \eta(\tilde{\varepsilon}) + \gamma(C\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \rho(T - \hat{t}_2(T,\tilde{\varepsilon}))$$

$$\le (\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) + \eta(\tilde{\varepsilon}) + \gamma(C\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \rho(T/8).$$

From (iv) we thus obtain

$$(\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) + \eta(\tilde{\varepsilon}) + \gamma(C\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \rho(T/8)$$

$$\ge (\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) - \sigma(\|x(\hat{t}_1(T,\tilde{\varepsilon})) - \bar{x}\|_X)$$

$$- \sigma(\|x(\hat{t}_2(T,\tilde{\varepsilon})) - \bar{x}\|_X) + \max_{t\in[\hat{t}_1(T,\tilde{\varepsilon}),\hat{t}_2(T,\tilde{\varepsilon})]} \sigma(\|x(t) - \bar{x}\|_X)$$

$$\ge (\hat{t}_2(T,\tilde{\varepsilon}) - \hat{t}_1(T,\tilde{\varepsilon}))J(\bar{x},\bar{u}) - \sigma(\tilde{\varepsilon})$$

$$- \sigma(C\tilde{\varepsilon}) + \max_{t\in[\hat{t}_1(T,\tilde{\varepsilon}),\hat{t}_2(T,\tilde{\varepsilon})]} \sigma(\|x(t) - \bar{x}\|_X).$$

Using the bounds on $\hat{t}_1(T,\tilde{\varepsilon})$ and $\hat{t}_2(T,\tilde{\varepsilon})$ this implies

$$\max_{t\in[3T/8,5T/8]} \|\hat{x}(t) - \bar{x}\|_X \le \sigma^{-1}(\eta(\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \sigma(\tilde{\varepsilon}) + \gamma(C\tilde{\varepsilon}) + \sigma(C\tilde{\varepsilon}) + \rho(T/8)).$$

Now, defining $t_1(T) = 3T/8$ and $t_2(T) = 5T/8$, the property from Assumption 2.1 follows for every $\varepsilon > 0$ by setting $\tilde{\varepsilon} > 0$ and $T \ge 8\mu(\varepsilon)$ so large that $\sigma^{-1}(\eta(\tilde{\varepsilon}) + \gamma(\tilde{\varepsilon}) + \sigma(\tilde{\varepsilon}) + \gamma(C\tilde{\varepsilon}) + \sigma(C\tilde{\varepsilon}) + \rho(T/8)) < \sigma.$ $\qquad\square$

Remark 6.2. The properties needed in the assumption of Theorem 6.1 follow from other well known properties of the optimal control problem under consideration:

- Condition (i) follows from strict dissipativity and Condition (ii). Strict dissipativity demands the existence of a storage function $S : X \to \mathbb{R}$, bounded from below, and a function $\alpha \in \mathcal{K}_\infty$ such that

$$S(x(t)) \le S(x(0)) + \int_0^t J(x(\tau), u(\tau)) - \alpha(\|x(\tau) - \bar{x}\|_X + \|u(\tau) - \bar{u}\|_U)d\tau$$

  holds along all solutions. In finite dimensions this is shown using exponential reachability of $\bar{x}$, which is implied by stabilizability, in [9, Theorem 2]. In infinite dimensions the implication strict dissipativity $\Rightarrow$ measure turnpike is analyzed in [34, Theorem 2].
- Condition (ii) follows from exponential stabilizability of $\bar{x}$ by an affine linear feedback $u = Kx + \bar{u}$ and continuity of J. This can be seen straightforwardly by integrating J along the exponentially stable closed loop solution.
- Condition (iii) follows from strict dissipativity and (exact) local controllability around $\bar{x}$. A proof in discrete time can be found in [13, Sec. 6]. This proof easily carries over to the continuous time setting of this paper.
- Condition (iv) follows from strict dissipativity if the α in the strict dissipativity formulation grows quickly enough and the storage function is continuous in $\bar{x}$. Continuity implies that $S(x(0))$ and $S(x(T))$ can be bounded by the σ-terms and the growth condition on α implies the inequality

$$\int_0^t \alpha(\|x(\tau) - \bar{x}\|_X + \|u(\tau) - \bar{u}\|_U)d\tau \ge \max_{t \in [0,T]} \sigma(\|x(t) - \bar{x}\|_X).$$

All theorems in this paper provide the interval turnpike property for the adjoints if the adjoints are bounded for horizon T tending to infinity. In other words, the theorems state that the adjoints are either unbounded or they satisfy the turnpike property. It is thus necessary to establish a bound on the adjoints in order to conclude the turnpike property. We end this section by providing an example where such a bound can be deduced.

Example 6.3 (Boundedness of adjoint). We give an example with $X = H_0^1(\Omega)$ and $Y = L_2(\Omega)$, where the bound $\sup_{T \ge 0} \|\lambda\|_{C(t_2(T),T;H^1(\Omega))} < \infty$ assumed in Theorems 3.1 and 4.4 and Corollary 3.7 holds. Consider the running cost $J(x,u) = \frac{1}{2}\|x - x_d\|_{L_2(\Omega_o)}^2 + \frac{1}{2}\|u - u_d\|_{L_2(\Omega_c)}^2$, $\mathcal{A} = \Delta$ endowed with Dirichlet boundary conditions and that $f(x,u) = f(x)$ is monotonously non-increasing, i.e., $f'(x) \le 0$ with $f(0) = 0$. Throughout this example, we consider by $c \ge 0$ a T-independent generic constant. By optimality of (x,u)

we obtain for any (x_r, u_r) satisfying the dynamics, that

$$\frac{1}{2}\|x - x_{\mathrm{d}}\|^2_{L_2((0,T)\times\Omega_o)} \leq \int_0^T J(x(t), u(t))\, dt \leq \int_0^T J(x_r(t), u_r(t))\, dt$$

$$= \frac{1}{2}\left(\|x_r - x_d\|^2_{L_2((0,T)\times\Omega_o)} + \|u_r - u_d\|^2_{L_2((0,T)\times\Omega_c)}\right)$$

Inserting this into the adjoint equation yields, using parabolic regularity and $f'(x) \leq 0$, that

$$(6.4) \qquad \|\lambda(t)\|_{H^1(\Omega)} \leq c\|x - x_{\mathrm{d}}\|_{L_2((0,T)\times\Omega_o)}$$

$$\leq c\left(\|x_r - x_d\|_{L_2((0,T)\times\Omega_o)} + \|u_r - u_d\|_{L_2((0,T)\times\Omega_c)}\right)$$

for all $t \in [0, T]$. This yields the result if

$$\int_0^T J(x_r, u_r) = \frac{1}{2}\left(\|x_r - x_d\|^2_{L_2((0,T)\times\Omega_o)} + \|u_r - u_d\|^2_{L_2((0,T)\times\Omega_c)}\right)$$

can be bounded independently of T for any $u_r \in L_2((0,T) \times \Omega_c)$ and corresponding state x_r. The same argumentation also carries over to the case of Neumann boundary control. The case of boundary observation can not be included, as we do not have the required regularity of the adjoint to deduce (6.4).

7. Numerical example

We present an example with a semilinear heat equation with Neumann boundary control:

$$x' - \Delta x + x^3 = 0 \qquad\qquad \text{in } [0, T] \times \Omega,$$

$$\frac{\partial x}{\partial \nu} = u \qquad\qquad \text{in } [0, T] \times \partial\Omega,$$

$$x(0) = x_0 \qquad\qquad \text{in } \Omega,$$

where $\frac{\partial x}{\partial \nu}$ is the outward unit normal derivative. As a cost function, we consider

$$\int_0^T J(x(t), u(t))dt = \int_0^T \frac{1}{2}\|x(t) - x_d\|^2_{L_2(\Omega)} + \frac{1}{2}\|u(t)\|^2_{L_2(\partial\Omega)}\, dt.$$

We choose $\Omega = [0, 3] \times [0, 1]$ as the spatial domain and the horizon $T = 10$. Additionally, we set $x_0 = 0$ and the reference trajectory defined by

$$x_d(\omega) := g\left(\frac{10}{3}\left\|\omega - \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}\right\|\right),$$

$$\text{where} \qquad g(s) := \begin{cases} 10e^{1 - \frac{1}{1 - s^2}} & s < 1 \\ 0 & \text{else.} \end{cases}$$

18   T. FAULWASSER, L. GRÜNE, J.-P. HUMALOJA, AND M. SCHALLER

This static reference is depicted in Figure 1. The optimal control problem is solved with the C++-library for vector space algorithms $Spacy^1$ using the finite element library $Kaskade\gamma$ [12].
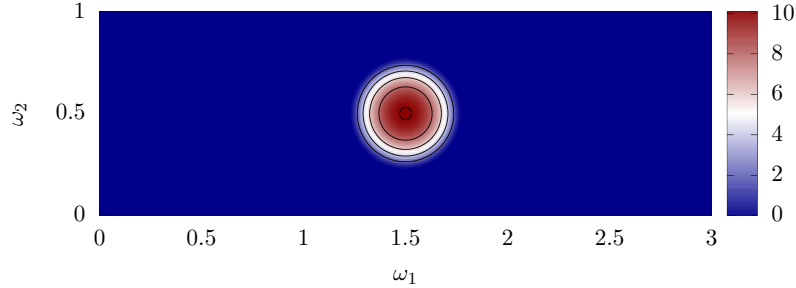

FIGURE 1. Depiction of the static reference trajectory $x_d(\omega_1, \omega_2)$.

For in-depth analysis of semilinear parabolic optimal control problems we refer the reader to [31] or [37, Chap. 5]. Considering the regularity of the static adjoint, for sufficiently smooth data we obtain that $\bar{\lambda} \in C(\bar{\Omega})$, cf. [5]. We set $\mathcal{A} = \Delta$ and $\varphi(x) = x^3$ and denote the superposition operator corresponding to $\varphi'(x) = 3x^2$ by Φ. We numerically verify that the turnpike property of Assumption 2.1 for the optimal state and control holds in $X = H^1(\Omega)$ and $U = L_2(\partial\Omega)$, cf. Figure 2. Thus, we depict in Figure 2 the norm of state and adjoint over time and in Figure 3 a snapshot of the dynamic solution and compare it to the steady state solution.
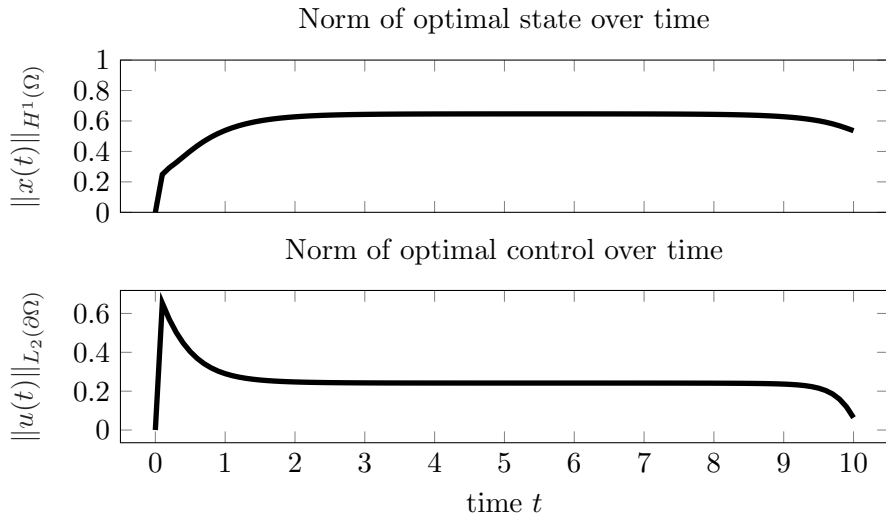

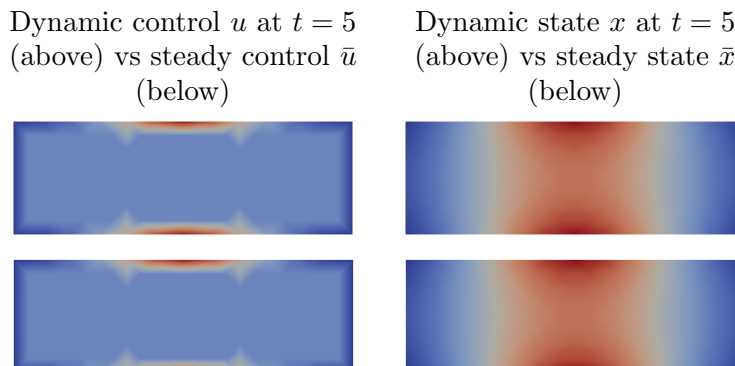FIGURE 2. Turnpike property for optimal state and control.

Dynamic control u at $t = 5$
(above) vs steady control $\bar{u}$
(below)

Dynamic state x at $t = 5$
(above) vs steady state $\bar{x}$
(below)

FIGURE 3. Comparison of dynamic solution (x, u) in the middle of the time interval with the steady state solution $(\bar{x}, \bar{u})$ with the same coloring as used in Figure 2.

While the turnpike property of the state and adjoints is verified numerically, the remaining assumptions can be checked analytically as follows. By the classical embeddings $H^1(\Omega) \hookrightarrow L_p(\Omega)$ for any $1 \leq p < \infty$ for $\Omega \subset \mathbb{R}^2$, cf. [37, Sec. 7.1] or [1, Chap. V] and as the nonlinearity is cubic, the occurring superposition operators satisfy Assumption 2.7 for $Y = L_2(\Omega)$. Moreover, as $D(\mathcal{A}^*) = \{v \in H^2(\Omega) \,|\, \frac{\partial}{\partial \nu} v = 0\} \hookrightarrow H^1(\Omega) = X$ compactly, the perturbation $\Phi(\bar{x}) \in L(X, Y)$ is $\mathcal{A}^*$-compact and $A^* = \mathcal{A}^* + \Phi(\bar{x})^*$ generates an analytic semigroup on $L_2(\Omega)$, and $D(A^*) = D(\mathcal{A}^*)$, cf. [8, Chap. III, Thm. 2.10]. Thus, assuming additionally a boundedness condition of the adjoint, Theorem 4.4 applies and we obtain the turnpike property for the adjoint in $Y = L_2(\Omega)$. Additionally by analyticity of the semigroup, the improved estimate of Theorem 5.1 holds and we get the estimate also in the space $W^{A^*}(0, T)$ with $D(A^*) = \{v \in H^2(\Omega) \,|\, \frac{\partial}{\partial \nu} v = 0\}$ and $(D(A^*), L_2(\Omega))_{\frac{1}{2}} \subset H^1(\Omega)$. In Figure 4 we observe the turnpike property for the adjoint in the $H^1(\Omega)$-norm and also in the $L_2(\Omega)$-norm.
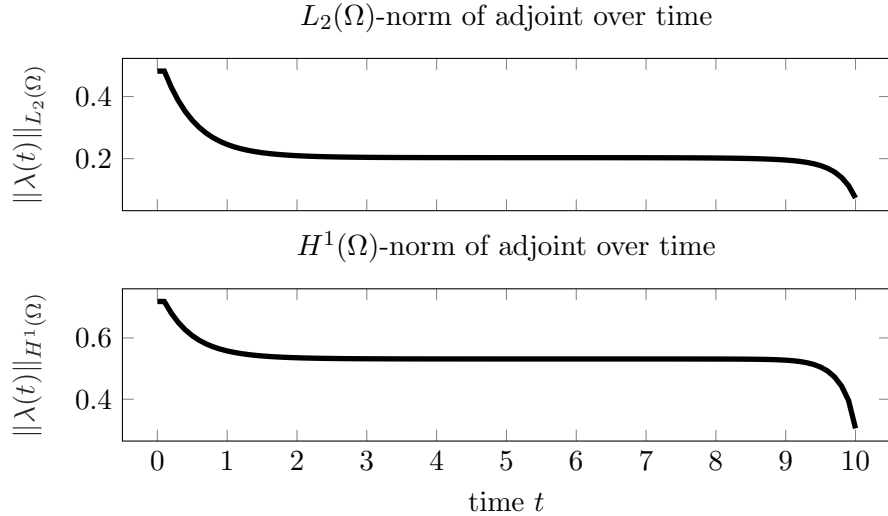
$L_2(\Omega)$-norm of adjoint over time

$H^1(\Omega)$-norm of adjoint over time

time t

FIGURE 4. Turnpike property for the adjoint.

8. SUMMARY AND OUTLOOK

This paper presented an interval turnpike result for the adjoints of finite- and infinite-dimensional nonlinear optimal control problems under the assumption of an interval turnpike on states and controls. We analyzed the case of stabilizable dynamics governed by a generator of a semigroup with finite dimensional unstable part satisfying a spectral decomposition condition. We have shown the desired turnpike property under continuity assumptions on the first-order optimality conditions. We illustrated our findings drawing upon a boundary controlled semilinear heat equation.

We briefly discuss possible directions of further research. A central assumption in the results of this work is the boundedness condition on the adjoint, i.e., $\sup_{T \geq 0} \|\lambda\|_{C(t_2(T),T;X)} < \infty$. We have presented a very particular case in Example 6.3 where this bound holds, however, under the strict condition that the cost at the equilibrium has to vanish. It is desirable to prove the bound under milder conditions.

In Section 7 we verified the turnpike property of state and control only numerically. This is due to the fact that the verification of assumptions of global turnpike theorems via dissipativity, cf. Remark 6.2, is a highly non-trivial issue. In that context, suitable storage functions need to be constructed and we refer to [14] for a promising approach in that direction.

An inspection of the proofs of Sections 3 and 4 shows that, under suitable assumptions, it should be possible replace the control operator $\mathcal{B}$ by an unbounded operator, which is, e.g., the case for boundary control. In that context, a concept like admissibility or assuming that solutions are classical solutions is a suitable replacement for boundedness of $\mathcal{B}$.

References

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, Amsterdam, Boston, 1975.

[2] J. Appell and P. P. Zabrejko. *Nonlinear Superposition Operators*. Cambridge Tracts in Mathematics. Cambridge University Press, 1990.

[3] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. Mitter. *Representation and control of infinite dimensional systems*. Springer Science & Business Media, 2007.

[4] T. Breiten and L. Pfeiffer. On the turnpike property and the receding-horizon method for linear-quadratic optimal control problems. *SIAM Journal on Control and Optimization*, 58(2):1077–1102, 2020.

[5] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM Journal on Control and Optimization*, 31(4):993–1006, 1993.

[6] R. F. Curtain and H. Zwart. *An introduction to infinite-dimensional linear systems theory*, volume 21. Springer Science & Business Media, 1995.

[7] T. Damm, L. Grüne, M. Stieler, and K. Worthmann. An exponential turnpike theorem for dissipative discrete time optimal control problems. *SIAM Journal on Control and Optimization*, 52(3):1935–1957, 2014.

[8] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag New York, 2000.

[9] T. Faulwasser, M. Korda, C. N. Jones, and D. Bonvin. On turnpike and dissipativity properties of continuous-time optimal control problems. *Automatica*, 81:297–304, 2017.

[10] T. Faulwasser and A. Murray. Turnpike properties in discrete-time mixed integer optimal control. *IEEE Control Systems Letters*, 4:704–709, 2020.

[11] H. Goldberg, W. Kampowsky, and F. Tröltzsch. On Nemytskij operators in Lp-spaces of abstract functions. *Mathematische Nachrichten*, 155(1):127–140, 1992.

[12] S. Götschel, A. Schiela, and M. Weiser. Kaskade 7 – A flexible finite element toolbox. *Computers & Mathematics with Applications*, 2020.

[13] L. Grüne. Economic receding horizon control without terminal constraints. *Automatica*, 49(3):725–734, 2013.

[14] L. Grüne and R. Guglielmi. On the relation between turnpike properties and dissipativity for continuous time linear quadratic optimal control problems. *Math. Control Rel. Fields*, 2020. To appear.

[15] L. Grüne and M. A. Müller. On the relation between strict dissipativity and turnpike properties. *Systems & Control Letters*, 90:45–53, 2016.

[16] L. Grüne, M. Schaller, and A. Schiela. Sensitivity analysis of optimal control for a class of parabolic PDEs motivated by model predictive control. *SIAM Journal on Control and Optimization*, 57(4):2753–2774, 2019.

[17] L. Grüne, M. Schaller, and A. Schiela. Exponential sensitivity analysis for model predictive control of PDEs. urn:nbn:de:bvb:703-epub-4590-3, 2020.

[18] L. Grüne, M. Schaller, and A. Schiela. Exponential sensitivity and turnpike analysis for linear quadratic optimal control of general evolution equations. *Journal of Differential Equations*, 268(12):7311–7341, 2020.

[19] M. Gugat and F. Hante. On the turnpike phenomenon for optimal boundary control problems with hyperbolic systems. *SIAM Journal on Control and Optimization*, 57(1):264–289, 2019.

[20] M. Gugat, E. Trélat, and E. Zuazua. Optimal Neumann control for the 1D wave equation: Finite horizon, infinite horizon, boundary tracking terms and the turnpike property. *Systems & Control Letters*, 90:61–70, 2016.

[21] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

[22] B. Jacob, C. Trunk, and M. Winklmeier. Analyticity and Riesz basis property of semigroups associated to damped vibrations. *J. Evol. Equ.*, 8(2):263–281, 2008.

[23] I. Lasiecka and R. Triggiani. *Control Theory for Partial Differential Equations: Volume 1, Abstract Parabolic Systems: Continuous and approximation theories*, volume 1. Cambridge University Press, 2000.

[24] X. Li and J. Yong. *Optimal Control Theory for Infinite Dimensional Systems*. Birkhäuser, 1995.

[25] J.-L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag Berlin-Heidelberg, 1971.

[26] J.-L. Lions. Exact controllability, stabilization and perturbations for distributed systems. *SIAM review*, 30(1):1–68, 1988.

[27] J. v. Neumann. A model of general economic equilibrium. *The Review of Economic Studies*, 13(1):1–9, 1945.

[28] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, NY, 1983.

[29] A. Porretta. On the turnpike property for mean field games. *Minimax Theory and its Applications*, 3(2):285–312, 2018.

[30] A. Porretta and E. Zuazua. Long time versus steady state optimal control. *SIAM Journal on Control and Optimization*, 51(6):4242–4273, 2013.

[31] J. P. Raymond and H. Zidani. Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Applied Mathematics and Optimization*, 39(2):143–177, 1999.

[32] N. Sakamoto, D. Pighin, and E. Zuazua. The turnpike property in nonlinear optimal control — A geometric approach. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 2422–2427, 2019.

[33] O. J. Staffans. *Well-Posed Linear Systems*. Cambridge University Press, 2005.

[34] E. Trélat and C. Zhang. Integral and measure-turnpike properties for infinite-dimensional optimal control systems. *Mathematics of Control, Signals, and Systems*, 30(1):3, 2018.

[35] E. Trélat, C. Zhang, and E. Zuazua. Steady-state and periodic exponential turnpike property for optimal control problems in Hilbert spaces. *SIAM Journal on Control and Optimization*, 56(2):1222–1252, 2018.

[36] E. Trélat and E. Zuazua. The turnpike property in finite-dimensional nonlinear optimal control. *Journal of Differential Equations*, 258(1):81–114, 2015.

[37] F. Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications*, volume 112. American Mathematical Soc., 2010.

[38] A. J. Zaslavski. *Turnpike theory of continuous-time linear optimal control problems*. Springer, 2016.

(T. Faulwasser) TU Dortmund University, Institute of Energy Systems, Energy Efficiency and Energy Economics, Germany
  Email address: `timm.faulwasser@ieee.org`

(L. Grüne) University of Bayreuth, Department of Mathematics, Germany
  Email address: `lars.gruene@uni-bayreuth.de`

(M. Schaller) Technische Universität Ilmemau, Institute for Mathematics, Germany
  Email address: `manuel.schaller@tu-ilmenau.de`

(J.-P. Humaloja) Tampere University, Mathematics, Computing Sciences, Finland

Email address: jukka-pekka.humaloja@tuni.fi