

# EXPONENTIAL SENSITIVITY AND TURNPIKE ANALYSIS FOR LINEAR QUADRATIC OPTIMAL CONTROL OF GENERAL EVOLUTION EQUATIONS\*

LARS GRÜNE <sup>†</sup>, MANUEL SCHALLER<sup>‡</sup>, AND ANTON SCHIELA<sup>§</sup>

**Abstract.** We analyze the sensitivity of linear quadratic optimal control problems governed by general evolution equations with bounded or admissible control operator. We show, that if the problem is stabilizable and detectable, the solution of the extremal equation can be bounded by the right-hand side including initial data with the bound being independent of the time horizon. Consequently, the influence of perturbations of the extremal equations decays exponentially in time. This property can for example be used to construct efficient space and time discretizations for a Model Predictive Control scheme. Furthermore, a turnpike property for unbounded but admissible control of general semigroups can be deduced.

**Key words.** Sensitivity Analysis, Turnpike Property, Model Predictive Control

**AMS subject classifications.** 49K20, 49K40, 93D20

**1. Introduction.** In this paper we consider optimal control of linear evolution equations in infinite dimensional spaces. It seems intuitively clear that perturbations of the data that occur in the far future affect the optimal control at present time only slightly. This might serve as an explanation, why closed-loop control techniques, which only take present information into account, often work in a nearly optimal way, even in the long run. The aim of this work is to give a rigorous and quantitative justification to this statement.

More specifically, we analyze the sensitivity of the solutions to linear quadratic optimal control problems, characterized by the first order optimality conditions with respect to the problem data, i.e., source terms and initial data of the state and adjoint equation. In the first part of our paper we are concerned with the question on how the behavior of perturbations of the source terms over time influences the solutions' behavior over time. It turns out that if the dynamics are stabilizable and detectable, the perturbations' influence decays exponentially as the distance in time grows.

Our paper generalizes recent results in this direction from parabolic equations [14], which enjoy very favorable smoothing properties, to general evolution equations with much less structure, which in particular comprises hyperbolic equations. This makes the analysis much more delicate. For example, in [14], the boundary control or boundary observation case could be dealt with in a straightforward fashion, as the state and adjoint belong to a Sobolev space for a.a. times on which boundary control or observation operators are bounded. This is no longer the case when moving to general evolution equations. Instead the concept of unbounded, but admissible control and observation operators as described e.g. in [26] has to be employed to cover the case of boundary control. Many general results, such as Riccati theory, are not available in their full strength within this broader scope.

The analysis in this work is based on analyzing the extremal equations of the optimal control problem, avoiding the use of Riccati equations. Compared to [14] we use weaker assumptions and thus obtain estimates in weaker norms, concerning spatial regularity, which is to be expected when going from parabolic to e.g. hyperbolic equations. Concerning temporal regularity, we show uniform estimates as well as  $L_2$ -type estimates for perturbations in  $L_2$  and in  $L_1$ .

---

\*This work was supported by the DFG Grants GR 1569/17-1 and SCHI 1379/5-1.

<sup>†</sup>Mathematical Institute, University of Bayreuth, 95447 Bayreuth, Germany, ([lars.gruene@uni-bayreuth.de](mailto:lars.gruene@uni-bayreuth.de)).

<sup>‡</sup>Mathematical Institute, University of Bayreuth, 95447 Bayreuth, Germany, ([manuel.schaller@uni-bayreuth.de](mailto:manuel.schaller@uni-bayreuth.de)).

<sup>§</sup>Mathematical Institute, University of Bayreuth, 95447 Bayreuth, Germany, ([anton.schiela@uni-bayreuth.de](mailto:anton.schiela@uni-bayreuth.de)).

One important motivation for our sensitivity analysis is Model Predictive Control (MPC). MPC is a control technique, where the solution of an optimal control problem on an indefinite or infinite horizon is split into the successive solution of problems with short time horizons  $T$ . However, only a first part up to a time  $\tau$ , where  $\tau \ll T$  is implemented and the process is repeated, with the resulting state as initial condition. Consequently, only the first part of the control has to be computed accurately. The fact that perturbations decay exponentially in time allows to discretize and compute solutions with reduced accuracy towards the end of the time horizon, as discretization errors are perturbations of the extremal equations. We refer to [13] for an in-depth introduction to MPC and to [11] for an overview of its approximation properties.

As a second application of our technique we show an exponential turnpike property. This property is important to understand the behavior of solutions to optimal control problems on large time horizons, which remain close to a steady state solution, the so called turnpike, for the majority of the time. There is a large literature on turnpike behavior, cf. [1, 7, 17, 27, 28, 12, 9, 16, 15]. A particular kind of turnpike behavior is the so called exponential turnpike property, cf. [25, 24, 21, 22, 6, 5]. The proofs establishing the turnpike property in the recent works [5, 24, 14] are based on a stabilizability and detectability assumption on the system. Turnpike theorems in Hilbert spaces were given in [24] for general semigroups with bounded control and observation operators, and for boundary controlled parabolic equations. The proofs in [24], however, make use of the Algebraic Riccati Equation, a theory, which is well established for admissible boundary control of parabolic equations, but not for general evolution equations. Here, we show a turnpike result for unbounded but admissible control of non-parabolic equations which was not available up to now. This is possible, since we avoid using Riccati theory in our approach. Our analysis also sheds light on the close connection of exponential sensitivity analysis and the turnpike property, both emerging from the boundedness of the operator corresponding to the extremal equations. This becomes clear by comparing the abstract scaling results in [Theorem 3.2](#) and [Theorem 6.2](#).

This paper is structured as follows. First, in [section 2](#) we present our theoretical framework, the optimal control problem and its optimality conditions. In [section 3](#), we derive a general result on the error propagation over time in [Theorem 3.2](#), under the assumption that various norms of the extremal equations' solution operator, which itself may indeed depend on the horizon  $T$ , can be bounded independently of  $T$ . Then, in [section 4](#) we show that such a  $T$ -independent bound on these operator norms holds, if the dynamics are exponentially stabilizable and detectable. This assumption enters into the construction of special test functions, similar to [21] and [24], that decay exponentially. In [section 5](#), we will extend the results to unbounded but admissible control operators and discuss the necessary modifications to the proofs. In [section 6](#) we show our turnpike result [Theorem 6.2](#) for general evolution equations. Here, too, we allow for unbounded but admissible control operators and general  $C_0$ -semigroups. Last, we present two examples that fulfill the assumptions of our analysis, namely the interior control of a heat equation, and the Dirichlet boundary control of a wave equation.

**2. Setting and preliminaries.** We consider linear quadratic optimal control problems in a Hilbert space  $X$  of the following type.

*Problem 2.1.*

$$(2.1) \quad \begin{aligned} \min_{(y,u)} \frac{1}{2} \int_0^T \|C(y(t) - y_d)\|_Y^2 + \|R(u(t) - u_d)\|_U^2 dt \\ y' - \Lambda y - Bu = f \\ y(0) = y_0, \end{aligned}$$

where  $Y$  and  $U$  are Hilbert spaces for the output and the control and  $T > 0$  is the optimization horizon. The dynamics are given by an evolution equation where  $\Lambda : D(\Lambda) \subset X \rightarrow X$  is a (possibly unbounded) generator of a  $C_0$ -semigroup  $(\mathcal{T}(t))_{t \geq 0}$  on  $X$ ,  $f \in L_1(0, T; X)$  is a source term and  $y_0 \in X$  is the initial datum. Moreover,  $R \in L(U, U)$  is an elliptic operator, leading to convexity of the problem. Finally, the control and observation operators are bounded, i.e.,  $B \in L(U, X)$  and  $C \in L(X, Y)$ .

Only in the first part of this paper we assume that  $B$  and  $C$  are bounded, in order not to hide the main steps of our proof behind technical details. In [section 5](#) we will generalize the result to unbounded but admissible control operators.

Whenever we consider a solution to an evolution equation, we refer to the mild solution  $y \in C(0, T; X)$ , which for [\(2.1\)](#) can be defined by

$$y(t) = \mathcal{T}(t)y_0 + \int_0^t \mathcal{T}(t-s)(Bu(s) + f(s)) ds.$$

We refer to [\[20, Section 4.2\]](#) for an introduction to inhomogeneous abstract Cauchy problems. In the following,  $\|\cdot\|$  resp.  $\langle \cdot, \cdot \rangle$  denote the norm resp. the scalar product in the Hilbert spaces  $X, U$  and  $Y$ . The space  $C(0, T; X)$  with norm  $\|x\|_{C(0, T; X)} := \max_{t \in [0, T]} \|x(t)\|$  will be denoted by  $C(X)$  and  $L_p(0, T; X)$  with norm  $\|x\|_{L_p(0, T; X)}^p := \int_0^T \|x(t)\|^p dt$  will be denoted by  $L_p(X)$  for  $p \in \{1, 2\}$  and  $L_\infty(0, T; X)$  with norm  $\|x\|_{L_\infty(0, T; X)} := \text{ess sup}_{t \in [0, T]} \|x(t)\|$  by  $L_\infty(X)$ . We will assume that there exists at least one minimizer of [Problem 2.1](#). For the existence of minimizers, we refer to [\[19, Chapter 3\]](#). Let  $(y, u) \in C(X) \times L_2(U)$  be a minimizer of [Problem 2.1](#). Then, by the Pontryagin Maximum Principle, see [\[19, Chapter 4\]](#), there exists  $\lambda \in C(X)$  such that

$$(2.2) \quad \begin{aligned} C^*Cy - \lambda' - \Lambda^*\lambda &= C^*Cy_d \\ R^*Ru - B^*\lambda &= R^*Ru_d \\ y' - \Lambda y - Bu &= f \end{aligned}$$

in a mild sense along  $[0, T]$  with initial resp. terminal conditions  $y(0) = y_0$  and  $\lambda(T) = 0$ . We denote by  $\Lambda^*$  the adjoint operator associated with  $\Lambda$  with domain  $D(\Lambda^*)$ . Defining  $Q := R^*R$  and eliminating the control via the second equation with  $u = Q^{-1}B^*\lambda + u_d$  leads to the linear system of equations

$$(2.3) \quad \underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=:M} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} C^*Cy_d \\ 0 \\ Bu_d + f \\ y_0 \end{pmatrix},$$

where  $E_0 y := y(0)$  and  $E_T \lambda := \lambda(T)$ . The operator  $M$  corresponds to the two abstract inhomogeneous evolution equations, i.e., the adjoint equation in the first two rows and the state equation in the last two rows and allows for a brief notation of those. The solution operator of this system, which we denote by  $M^{-1}$ , maps initial values and source terms for the state and the adjoint equation to the solution. This mapping is well defined as a mapping from  $(L_1(X) \times X)^2$  to  $C(X)^2$ , as for right-hand sides with this regularity, the state and adjoint equations have a unique solution in  $C(X)$  each. A central question in the following will be the dependence of this inverse operator's norm on the time  $T$ .

**3. A sensitivity result.** In this section, we present a first sensitivity result which extends the result of [14, Theorem 3.1]. For this, we will refer to the solution  $(y, \lambda) \in C(X)^2$  of (2.3) as the *exact solution*. We will now assume that there is a second pair of variables  $(y_p, \lambda_p) \in C(X)^2$  that satisfies the perturbed system

$$\begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} y_p \\ \lambda_p \end{pmatrix} = \begin{pmatrix} C^*C y_d \\ 0 \\ B u_d + f \\ y_0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ 0 \\ \varepsilon_2 \\ 0 \end{pmatrix}$$

for  $(\varepsilon_1, \varepsilon_2) \in L_1(X)^2$ . This solution will be referred to as the *perturbed solution*. The terms  $\varepsilon_1$  and  $\varepsilon_2$  are perturbations of the dynamics which could stem from discretization errors in time or space. In this subsection we will give an estimate for the norm of the difference of  $(y_p, \lambda_p)$  and  $(y, \lambda)$ . It follows by linearity that the difference between exact and perturbed solution  $(\delta y, \delta \lambda) := (y_p - y, \lambda_p - \lambda)$  solves

$$(3.1) \quad \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ 0 \\ \varepsilon_2 \\ 0 \end{pmatrix}.$$

The question we want to answer is the following. How does the behavior of the perturbations  $\varepsilon_1$  and  $\varepsilon_2$  over time influence the behavior of  $\delta y$  and  $\delta \lambda$ ? [Theorem 3.2](#) gives an answer towards that direction. We will prove a scaling result for evolution equations first. It is a well-known result that, if a semigroup  $(\mathcal{T}(t))_{t \geq 0}$  has generator  $\Lambda$ , the scaled semigroup  $(e^{-\mu t} \mathcal{T}(t))_{t \geq 0}$  has generator  $\Lambda - \mu I$  [8, p.60] with the same domain as  $\Lambda$ , as the domain does not change under bounded perturbations, cf. [8, Chapter III]. In our context, we want to put emphasis on the impact of scaling on the source terms  $\varepsilon_1$  and  $\varepsilon_2$ .

LEMMA 3.1. *Assume  $\delta y, \delta \lambda \in C(X)$  solve the abstract Cauchy problems*

$$\begin{aligned} \delta y' - \Lambda \delta y &= \varepsilon_1 \\ -\delta \lambda' - \Lambda^* \delta \lambda &= \varepsilon_2 \end{aligned}$$

*with initial resp. terminal conditions  $\delta y(0) = \delta y_0$ ,  $\delta \lambda(T) = \delta \lambda_T$ . Then  $\widetilde{\delta y}(t) = e^{-\mu t} \delta y(t)$  and  $\widetilde{\delta \lambda}(t) = e^{-\mu t} \delta \lambda(t)$  solve*

$$(3.2) \quad \begin{aligned} \widetilde{\delta y}' - (\Lambda - \mu I) \widetilde{\delta y} &= e^{-\mu t} \varepsilon_1 \\ -\widetilde{\delta \lambda}' - (\Lambda + \mu I)^* \widetilde{\delta \lambda} &= e^{-\mu t} \varepsilon_2 \end{aligned}$$

*with  $\widetilde{\delta y}(0) = \delta y_0$  and  $\widetilde{\delta \lambda}(T) = e^{-\mu T} \delta \lambda_T$ .*

*Proof.* Using the equivalence of the concept of weak and mild solution, see [19, Section 5.1] and [2], we will use the weak formulation in this case, which is for  $\varphi \in D(\Lambda^*)$  given by

$$\frac{d}{dt} \langle \delta y(t), \varphi \rangle - \langle \delta y(t), \Lambda^* \varphi \rangle = \langle \varepsilon_1(t), \varphi \rangle$$

for a.e.  $t \in [0, T]$ . Inserting  $\delta y = e^{\mu t} \widetilde{\delta y}$  and applying the chain rule yields the result for the state equation. The equation for the scaled adjoint variable follows analogously.  $\square$

The following theorem gives a preliminary result on the propagation of perturbations. It gives an estimate on the scaled difference of exact and perturbed solution under the assumption that different solution operator norms can be bounded independently of  $T$ .

**THEOREM 3.2.** *Assume  $(\delta y, \delta \lambda) \in C(X)^2$  solves (3.1), where  $\varepsilon_1, \varepsilon_2 \in L_1(X)$ . Assume the solution operator's norms*

$$(3.3) \quad \begin{aligned} \|M^{-1}\|_{(L_1(X) \times X)^2 \rightarrow C(X)^2}, & \quad \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow C(X)^2}, \\ \|M^{-1}\|_{(L_1(X) \times X)^2 \rightarrow L_2(X)^2}, & \quad \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2} \end{aligned}$$

can be bounded independently of  $T$ . Then there is a constant  $c \geq 0$  and a scaling factor  $0 < \mu$ , both independent of  $T$ , such that defining

$$\rho := \|e^{-\mu t} \varepsilon_1(t)\|_E + \|e^{-\mu t} \varepsilon_2(t)\|_E$$

for  $E := L_1(X)$  or  $E := L_2(X)$ , we have

$$(3.4) \quad \begin{aligned} \|e^{-\mu t} \delta y\|_{L_2(X)} + \|e^{-\mu t} \delta \lambda\|_{L_2(X)} &\leq c\rho, \\ \|e^{-\mu t} \delta u\|_{L_2(U)} &\leq c\rho \end{aligned}$$

and

$$(3.5) \quad \begin{aligned} \|e^{-\mu t} \delta y\|_{C(X)} + \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho, \\ \|e^{-\mu t} \delta u\|_{L_\infty(U)} &\leq c\rho. \end{aligned}$$

*Proof.* For  $\mu > 0$  we introduce scaled variables  $\widetilde{\delta y}(t) := e^{-\mu t} \delta y(t)$ ,  $\widetilde{\delta \lambda}(t) := e^{-\mu t} \delta \lambda(t)$ ,  $\widetilde{\varepsilon}_1(t) := e^{-\mu t} \varepsilon_1(t)$  and  $\widetilde{\varepsilon}_2(t) := e^{-\mu t} \varepsilon_2(t)$  and apply Lemma 3.1. This yields

$$(3.1) \quad \iff \left( \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} + \underbrace{\mu \begin{pmatrix} 0 & -I \\ 0 & 0 \\ I & 0 \\ 0 & 0 \end{pmatrix}}_{=:P} \right) \begin{pmatrix} \widetilde{\delta y} \\ \widetilde{\delta \lambda} \end{pmatrix} = \begin{pmatrix} \widetilde{\varepsilon}_1 \\ 0 \\ \widetilde{\varepsilon}_2 \\ 0 \end{pmatrix}.$$

Introducing  $\widetilde{z} := (\widetilde{\delta y}, \widetilde{\delta \lambda})$ ,  $\widetilde{\varepsilon} := (\widetilde{\varepsilon}_1, 0, \widetilde{\varepsilon}_2, 0)$  we compute formally

$$(3.6) \quad (M + \mu P)\widetilde{z} = \widetilde{\varepsilon} \quad \Rightarrow \quad (I + \mu PM^{-1})M\widetilde{z} = \widetilde{\varepsilon} \quad \Rightarrow \quad \widetilde{z} = M^{-1}(I + \mu PM^{-1})^{-1}\widetilde{\varepsilon}.$$

Next we expand  $(I + \mu PM^{-1})^{-1}$  into a Neumann series, cf. [18, Theorem 2.14]. In the following denote  $W = (L_2(X) \times X)^2$ . We choose  $\mu > 0$ , such that  $\beta := \mu \|M^{-1}\|_{W \rightarrow L_2(X)^2} < 1$ . By

assumption  $\|M^{-1}\|_{W \rightarrow L_2(X)^2}$  is bounded independently of  $T$ , so we can choose  $\mu > 0$  independently of  $T$ .

Since  $\|P\|_{L_2(X)^2 \rightarrow W} \leq 1$ , it follows  $\|\mu PM^{-1}\|_{W \rightarrow W} \leq \beta < 1$ . Neumann series expansion of  $(I + \mu PM^{-1})^{-1}$  yields

$$(3.7) \quad \|(I + \mu PM^{-1})^{-1}\|_{W \rightarrow W} \leq \sum_{i=0}^{\infty} \|(\mu PM^{-1})^i\|_{W \rightarrow W} \leq \sum_{i=0}^{\infty} \beta^i = \frac{1}{1 - \beta}.$$

This argumentation already yields the desired result in the  $L_2$  case:

$$\|\tilde{z}\|_{L_2(X)}^2 \leq \|M^{-1}\|_{W \rightarrow L_2(X)^2} \|(I + \mu PM^{-1})^{-1}\|_{W \rightarrow W} \|\tilde{\varepsilon}\|_W.$$

To prove the general case we use the following alternative representation, which can be verified by premultiplication with  $(I + \mu PM^{-1})$ :

$$(I + \mu PM^{-1})^{-1} = I - (I + \mu PM^{-1})^{-1} \mu PM^{-1}.$$

Now we estimate the operator norm for  $S = C(X)^2$  or  $S = L_2(X)^2$  via

$$\begin{aligned} \|M^{-1}(I + \mu PM^{-1})^{-1}\|_{(E \times X)^2 \rightarrow S} &= \|M^{-1} - M^{-1}(I + \mu PM^{-1})^{-1} \mu PM^{-1}\|_{(E \times X)^2 \rightarrow S} \\ &\leq \|M^{-1}\|_{(E \times X)^2 \rightarrow S} + \|M^{-1}\|_{W \rightarrow S} \|(I + \mu PM^{-1})^{-1}\|_{W \rightarrow W} \|\mu PM^{-1}\|_{(E \times X)^2 \rightarrow W} \\ &\leq \|M^{-1}\|_{(E \times X)^2 \rightarrow S} + \frac{\mu \|M^{-1}\|_{W \rightarrow S} \|M^{-1}\|_{(E \times X)^2 \rightarrow L_2(X)^2}}{1 - \beta} \end{aligned}$$

using  $\|P\|_{L_2(X)^2 \rightarrow W} \leq 1$  and (3.7). Thus, it follows by (3.6) with taking norms, that

$$(3.8) \quad \begin{aligned} \|\tilde{z}\|_{L_2(X)^2} &\leq \left( \|M^{-1}\|_{(E \times X)^2 \rightarrow L_2(X)^2} + \frac{\mu \|M^{-1}\|_{W \rightarrow L_2(X)^2} \|M^{-1}\|_{(E \times X)^2 \rightarrow L_2(X)^2}}{1 - \beta} \right) \|\tilde{\varepsilon}\|_{E^2}, \\ \|\tilde{z}\|_{C(X)^2} &\leq \left( \|M^{-1}\|_{(E \times X)^2 \rightarrow C(X)^2} + \frac{\mu \|M^{-1}\|_{W \rightarrow C(X)^2} \|M^{-1}\|_{(E \times X)^2 \rightarrow L_2(X)^2}}{1 - \beta} \right) \|\tilde{\varepsilon}\|_{E^2}. \end{aligned}$$

Using the assumption on the  $T$ -independent bound on the operator norms and the definition  $\rho := \|\tilde{\varepsilon}\|_{E^2}$ , we obtain the result for the state and the adjoint by going back to the original variables via  $\tilde{z} = (e^{-\mu t} \delta y, e^{-\mu t} \delta \lambda)$ . For the control we conclude

$$(3.9) \quad \begin{aligned} \|e^{-\mu t} \tilde{\delta u}\|_{L_2(U)} &= \|e^{-\mu t} Q^{-1} B^* \tilde{\delta \lambda}\|_{L_2(U)} \leq \|Q^{-1} B^*\|_{X \rightarrow U} \|\tilde{\delta \lambda}\|_{L_2(X)} \leq c \rho, \\ \|e^{-\mu t} \tilde{\delta u}(t)\| &= \|e^{-\mu t} Q^{-1} B^* \tilde{\delta \lambda}(t)\| \leq \|Q^{-1} B^*\|_{X \rightarrow U} \|\delta \lambda(t)\| \leq c \rho \end{aligned}$$

for a.a.  $t \in [0, T]$ , where we used the bound on  $\|e^{-\mu t} \delta \lambda\|_{L_2(X)}$  and  $\|e^{-\mu t} \delta \lambda\|_{C(X)}$  and the fact that  $B$  and  $Q$  do not involve time derivatives. This yields (3.4) and (3.5).  $\square$

*Remark 3.3.* We will briefly comment on the Neumann series occurring in (3.7). The operator  $(I + \mu PM^{-1})^{-1}$  can be represented by its Neumann series, i.e.,  $(I + \mu PM^{-1})^{-1} = \sum_{i=0}^{\infty} (-\mu PM^{-1})^i$ , see [18, Theorem 2.14]. We provide an illustration for the summand for  $k = 2$ , i.e.,  $(\mu PM^{-1})^2 = \mu PM^{-1} \mu PM^{-1}$ . The application of this operator can be interpreted as the following.  $M^{-1}$  solves the corresponding Cauchy problems with right-hand side including initial and terminal condition. Afterwards, the operator  $\mu P$  maps the solutions to source terms scaled by  $\mu$ , i.e.,  $\mu P(\delta y, \delta \lambda) = (-\mu \delta \lambda, 0, \mu \delta y, 0)$ . This right-hand side then enters  $M^{-1}$  again, the Cauchy problems are solved with zero initial data and source terms  $-\mu \delta \lambda$  and  $\mu \delta y$  and the process is repeated.

The crucial assumption in [Theorem 3.2](#) is that the operator norms in [\(3.3\)](#) can be bounded independently of  $T$ . In [section 4](#), we will show, that this is indeed the case, if the dynamics are exponentially stabilizable and detectable. It will turn out that all bounds in [\(3.3\)](#) can be shown simultaneously.

Before we head to the next part, we recall an integration by parts result for solutions expressed via semigroups.

LEMMA 3.4. *Consider the abstract Cauchy problems*

$$\begin{aligned} x_1' &= \Lambda x_1 + f_1, & x_1(0) &= x_{10}, \\ -x_2' &= \Lambda^* x_2 + f_2, & x_2(0) &= x_{20}, \end{aligned}$$

where  $x_{10}, x_{20} \in X$  and  $f_1, f_2 \in L_1(X)$ .  $\Lambda : D(\Lambda) \rightarrow X$  is the generator of a  $C_0$ -semigroup  $(\mathcal{T}(t))_{t \geq 0}$  on  $X$ . Consider the associated mild solutions  $x_1, x_2 \in C(0, T; X)$  given by

$$\begin{aligned} x_1(t) &= \mathcal{T}(t)x_{10} + \int_0^t \mathcal{T}(t-s)f_1(s) ds & \forall t \in [0, T], \\ x_2(t) &= \mathcal{T}^*(T-t)x_{20} + \int_t^T \mathcal{T}^*(s-t)f_2(s) ds & \forall t \in [0, T]. \end{aligned}$$

Then, for all  $0 \leq s \leq t \leq T$ ,

$$\langle x_1(t), x_2(t) \rangle - \langle x_1(s), x_2(s) \rangle = \int_s^t (\langle x_2, f_1 \rangle - \langle x_1, f_2 \rangle) ds.$$

*Proof.* For a proof of this property, we refer to [[19](#), Proposition 5.7, p.69]. □

**4.  $T$ -independent bounds for the solution operator.** In this section we will derive  $T$ -independent bounds on the norm of the solution operator  $M^{-1}$ , which is a central assumption in the abstract scaling result of [Theorem 3.2](#). Since  $[0, T]$  is bounded, we have the continuous embeddings

$$C(0, T; X) \hookrightarrow L_2(0, T; X) \hookrightarrow L_1(0, T; X).$$

Hence, we may equip  $L_1(0, T; X)$  with the equivalent norm

$$\|v\|_{1\vee 2} := \min\{\|v\|_{L_1(X)}, \|v\|_{L_2(X)}\}$$

(setting  $\|v\|_{L_2(X)} = \infty$  if  $v \notin L_2(0, T; X)$ ), which satisfies

$$\min\left\{1, \frac{1}{\sqrt{T}}\right\} \|v\|_{L_1(X)} \leq \|v\|_{1\vee 2} \leq \|v\|_{L_1(X)}.$$

Likewise, we equip  $C(0, T; X)$  with the equivalent norm

$$\|v\|_{2\wedge\infty} := \max\{\|v\|_{L_2(X)}, \|v\|_{C(X)}\},$$

satisfying

$$\|v\|_{C(X)} \leq \|v\|_{2\wedge\infty} \leq \max\{1, \sqrt{T}\} \|v\|_{C(X)}.$$

Obviously, the equivalence of norms deteriorates for  $T \rightarrow \infty$ . We have the following Hölder-like inequality for these norms:

LEMMA 4.1. *Let  $v \in C(X)$  and  $w \in L_1(X)$ . Then,*

$$\int_0^T \langle v(s), w(s) \rangle ds \leq \|v\|_{2\wedge\infty} \|w\|_{1\vee 2}.$$

*Proof.*

$$\begin{aligned} \int_0^T \langle v(s), w(s) \rangle ds &\leq \min\{\|v\|_{C(X)} \|w\|_{L_1(X)}, \|v\|_{L_2(X)} \|w\|_{L_2(X)}\} \\ &\leq \min\{\|v\|_{2\wedge\infty} \|w\|_{L_1(X)}, \|v\|_{2\wedge\infty} \|w\|_{L_2(X)}\} \leq \|v\|_{2\wedge\infty} \|w\|_{1\vee 2}. \quad \square \end{aligned}$$

The main result of this section will be a  $T$ -independent bound for

$$\|M^{-1}\|_{((L_1(X), \|\cdot\|_{1\vee 2}) \times X)^2 \rightarrow (C(X), \|\cdot\|_{2\wedge\infty})}.$$

This implies all desired  $T$ -independent bounds in (3.3), required by Theorem 3.2. To this end, consider mild solutions  $(y, \lambda)$  of the system

$$(4.1) \quad \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} l_1 \\ \lambda_T \\ l_2 \\ y_0 \end{pmatrix}$$

in  $[0, T]$ , where  $l_1, l_2 \in L_1(X)$  and  $y_0, \lambda_T \in X$  are given. In short we may write again  $z = M^{-1}r$  with  $z = (y, \lambda)$  and  $r = (l_1, \lambda_T, l_2, y_0)$ .

Next, we introduce our main assumption. A semigroup  $(\mathcal{T}(t))_{t \geq 0}$  is called exponentially stable, if there exist  $M, \mu > 0$ , such that  $\|\mathcal{T}(t)\|_{L(X, X)} \leq Me^{-\mu t}$  for all  $t > 0$ .

*Assumption 4.2.*

- A.1**  $(\Lambda, C)$  is exponentially detectable, i.e., there exists  $K_C \in L(Y, X)$  such that the semigroup generated by  $\Lambda^* + C^*K_C^*$  is exponentially stable.
- A.2**  $(\Lambda, B)$  is exponentially stabilizable, i.e., there exists  $K_B \in L(X, U)$  such that the semigroup generated by  $\Lambda + BK_B$  is exponentially stable.

The approach in Lemmas 4.3 and 4.4 is inspired by the stability estimate in [24, Lemma 2] and [21, Lemma 3.5].

LEMMA 4.3. *Let  $\varphi \in C(0, t; X)$  solve*

$$(4.2) \quad \begin{aligned} -\varphi' &= (\Lambda^* + C^*K_C^*)\varphi \quad \text{on } [0, t] \\ \varphi(t) &= y(t) \end{aligned}$$

where  $K_C^*$  is a stabilizing feedback for  $(\Lambda^*, C^*)$ . Then, there are constants  $M_\varphi, k_\varphi > 0$  such that for test functions  $v \in L_2(0, t; X)$

$$(4.3) \quad \int_0^t |\langle v(s), \varphi(s) \rangle| ds \leq \|y(t)\| \frac{M_\varphi}{\sqrt{k_\varphi}} \sqrt{\int_0^t \|v(s)\|^2 e^{-k_\varphi(t-s)} ds}.$$

Let  $\psi \in C(t, T; X)$  solve

$$(4.4) \quad \begin{aligned} \psi' &= (\Lambda + BK_B)\psi \quad \text{on } [t, T] \\ \psi(t) &= \lambda(t) \end{aligned}$$



where  $K_B$  is a stabilizing feedback for  $(\Lambda, B)$ . Then, for test functions  $v \in L_2(t, T; X)$

$$(4.5) \quad \int_t^T |\langle v(s), \psi(s) \rangle| ds \leq \|\lambda(t)\| \frac{M_\psi}{\sqrt{k_\psi}} \sqrt{\int_t^T \|v(s)\|^2 e^{-k_\psi(s-t)} ds}.$$

*Proof.* We will first prove (4.3). By exponential stability of the semigroup, there exist  $M_\varphi, k_\varphi > 0$ , such that

$$\|\varphi(s)\| \leq M_\varphi e^{-k_\varphi(t-s)} \|y(t)\| \quad 0 \leq s \leq t.$$

Using this exponential stability, we get for  $v \in L_2(0, t; X)$ ,

$$\int_0^t |\langle v(s), \varphi(s) \rangle| ds \leq \int_0^t \|v(s)\| \|\varphi(s)\| ds \leq \|y(t)\| \int_0^t \|v(s)\| M_\varphi e^{-k_\varphi(t-s)} ds.$$

The last term can be estimated via:

$$\begin{aligned} \int_0^t \|v(s)\| M_\varphi e^{-k_\varphi(t-s)} ds &= \int_0^t \|v(s)\| M_\varphi e^{-\frac{k_\varphi}{2}(t-s)} \cdot e^{-\frac{k_\varphi}{2}(t-s)} ds \\ &\leq \sqrt{\int_0^t \|v(s)\|^2 M_\varphi^2 e^{-k_\varphi(t-s)} ds} \cdot \underbrace{\sqrt{\int_0^t e^{-k_\varphi(t-s)} ds}}_{< \frac{1}{\sqrt{k_\varphi}}}. \end{aligned}$$

The estimate (4.5) follows analogously.  $\square$

By using  $\varphi$  and  $\psi$  from (4.2) and (4.4), respectively, as test functions for (4.1), we obtain the following pointwise in time identities:

LEMMA 4.4. *Let  $(y, \lambda)$  solve (4.1). If  $\varphi$  solves (4.2), then*

$$(4.6) \quad \|y(t)\|^2 = \int_0^t -\langle K_C C y(s), \varphi(s) \rangle + \langle R^{-*} B^* \lambda(s), R^{-*} B^* \varphi(s) \rangle + \langle l_2(s), \varphi(s) \rangle ds + \langle y_0, \varphi(0) \rangle$$

for all  $0 \leq t \leq T$ . If  $\psi$  solves (4.4), then

$$(4.7) \quad \|\lambda(t)\|^2 = \int_t^T -\langle K_B^* B^* \lambda(s), \psi(s) \rangle - \langle C y(s), C \psi(s) \rangle + \langle l_1(s), \psi(s) \rangle ds + \langle \lambda_T, \psi(T) \rangle$$

for all  $0 \leq t \leq T$ .

*Proof.* We begin with the proof of (4.6). Testing the state equation with  $\varphi(s)$  solving (4.2), integration on  $[0, t]$  and integration by parts in the sense of Lemma 3.4 on  $[0, t]$  with  $x_1 = y$ ,  $f_1 = BQ^{-1}B^*\lambda + l_2$ ,  $x_2 = \varphi$ ,  $f_2 = C^*K_C^*\varphi$  yields

$$\langle y(t), \varphi(t) \rangle - \langle y(0), \varphi(0) \rangle = \int_0^t \langle \varphi(s), BQ^{-1}B^*\lambda(s) + l_2(s) \rangle - \langle y(s), C^*K_C^*\varphi(s) \rangle ds.$$

Rearranging the terms, using the terminal condition  $\varphi(t) = y(t)$  and  $Q^{-1} = (R^*R)^{-1} = R^{-1}R^{-*}$ , we get

$$\|y(t)\|^2 = \int_0^t -\langle K_C C y(s), \varphi(s) \rangle + \langle R^{-*} B^* \lambda(s), R^{-*} B^* \varphi(s) \rangle + \langle l_2(s), \varphi(s) \rangle ds + \langle y_0, \varphi(0) \rangle.$$

The formula (4.7) follows analogous by testing the adjoint equation with  $\psi$  solving (4.4).  $\square$

Based on (4.6) and (4.7) we will derive norm estimates for  $M^{-1}$  as a mapping into  $L_2(X)^2$  and  $C(X)^2$ . While the latter turns out to be rather straightforward (see below), the  $L_2$ -estimate requires to integrate (4.6) and (4.7) over  $[0, T]$ . The crucial observation is that the integrals on the right hand side of (4.6) and (4.7) can be converted into convolutions with exponentially decaying functions. This will allow us to derive an  $L_2$ -estimate without any constants depending on the time  $T$  with the help of the following general lemma:

LEMMA 4.5. *For  $w \in L_1(0, T; 0, \infty)$ , consider*

$$h_1(t) := \int_0^t w(s)e^{-k_\varphi(t-s)} ds, \text{ where } k_\varphi > 0,$$

$$h_2(t) := \int_t^T w(s)e^{-k_\psi(s-t)} ds, \text{ where } k_\psi > 0.$$

Then, there is a constant  $c \geq 0$  independent of  $T$ , such that

$$\|h_i\|_{L_p(0, T)} \leq c\|w\|_{L_1(0, T)} \quad \text{for } i = 1, 2 \text{ and } 1 \leq p \leq \infty$$

*Proof.* Extending  $w$  by 0 from  $[0, T]$  to  $\mathbb{R}$  and defining  $g_1(x) := e^{-k_\varphi x}$  for  $x \geq 0$  and  $g_1(x) = 0$  otherwise, we can write  $h_1$  as a convolution:

$$h_1(t) = (g_1 * w)(t) = \int_{\mathbb{R}} g_1(t-s)w(s) ds,$$

and apply Young's inequality:

$$\|h_1\|_{L_p(0, T)} = \|g_1 * w\|_{L_p(\mathbb{R})} \leq \|g_1\|_{L_p(\mathbb{R})}\|w\|_{L_1(\mathbb{R})} \leq c\|w\|_{L_1(0, T)}$$

because  $\|g_1\|_{L_p(\mathbb{R})} = \|e^{-k_\varphi t}\|_{L_p(\mathbb{R}_+)} \leq c(k_\varphi)$ .

For  $h_2$  the estimate follows in the same way, setting  $g_2(x) = e^{k_\psi x}$  to  $x \leq 0$  and 0 otherwise.  $\square$

In the following, we will denote a generic constant by  $c$  and rename it accordingly over the course of a proof. However, the constants in the proofs will never depend on the horizon  $T$ . Also, we tacitly use equivalence of norms in  $\mathbb{R}^2$ :  $\max\{|a|, |b|\} \approx \sqrt{|a|^2 + |b|^2} \approx |a| + |b|$ .

LEMMA 4.6. *Let  $(y, \lambda)$  solve (4.1) and let Assumption 4.2 A.1 and A.2 hold. Then there exists a constant  $c \geq 0$  independent of  $T$ , such that*

$$(4.8) \quad \|y\|_{2\wedge\infty}^2 + \|\lambda\|_{2\wedge\infty}^2 \leq c \left( \|Cy\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda\|_{L_2(U)}^2 + \|r\|_{1\vee 2}^2 \right).$$

where  $\|r\|_{1\vee 2}^2 := \|l_1\|_{1\vee 2}^2 + \|l_2\|_{1\vee 2}^2 + \|y_0\|^2 + \|\lambda_T\|^2$ .

*Proof.* Our first step will be to derive an estimate for  $\|y(t)\|$  from (4.6). By Lemma 4.3 we estimate the terms occurring in (4.6) as follows:

$$(4.9) \quad \int_0^t |\langle K_C Cy(s), \varphi(s) \rangle| ds \leq \|y(t)\| \frac{M_\varphi \|K_C\|}{\sqrt{k_\varphi}} \sqrt{\int_0^t \|Cy(s)\|^2 e^{-k_\varphi(t-s)} ds},$$

$$(4.10) \quad \int_0^t |\langle R^{-*}B^*\lambda(s), R^{-*}B^*\varphi(s) \rangle| ds \leq \|y(t)\| \frac{M_\varphi \|BR^{-1}\|}{\sqrt{k_\varphi}} \sqrt{\int_0^t \|R^{-*}B^*\lambda(s)\|^2 e^{-k_\varphi(t-s)} ds},$$

$$(4.11) \quad \int_0^t \langle l_2(s), \varphi(s) \rangle ds \leq \|y(t)\| \frac{M_\varphi}{\sqrt{k_\varphi}} \sqrt{\int_0^t \|l_2(s)\|^2 e^{-k_\varphi(t-s)} ds},$$

or alternatively:

$$(4.12) \quad \int_0^t \langle l_2(s), \varphi(s) \rangle ds \leq \int_0^t \|l_2(s)\| \|\varphi(s)\| ds \leq \|y(t)\| M_\varphi \int_0^t \|l_2(s)\| e^{-k_\varphi(t-s)} ds,$$

and finally:

$$(4.13) \quad \langle y_0, \varphi(0) \rangle \leq \|y_0\| \|y(t)\| M_\varphi e^{-k_\varphi t} \leq \|y_0\| \|y(t)\| M_\varphi \sqrt{e^{-k_\varphi t}}.$$

Now we substitute all estimates (4.9)-(4.13) into (4.6) while taking the minimum over (4.11) and (4.12) and cancel  $\|y(t)\|$  on both sides. Taking squares on both side and using the simple inequality

$$a, b, c, d \geq 0 \Rightarrow (a + b + c + d)^2 \leq 4(a^2 + b^2 + c^2 + d^2)$$

we obtain the following pointwise estimate for  $y$ :

$$(4.14) \quad \|y(t)\|^2 \leq c \left( \int_0^t (\|Cy(s)\|^2 + \|R^{-*}B\lambda(s)\|^2) e^{-k_\varphi(t-s)} dt + \|y_0\|^2 e^{-k_\varphi t} \right) + c \min \left\{ \int_0^t \|l_2(s)\|^2 e^{-k_\varphi(t-s)} ds, \left( \int_0^t \|l_2(s)\| e^{-k_\varphi(t-s)} ds \right)^2 \right\}.$$

To derive an estimate for  $\|y\|_{C(X)}^2$  we replace all exponential functions by 1, extend the domains of integration from  $[0, t]$  to  $[0, T]$ , and take the maximum over all  $t \in [0, T]$ :

$$(4.15) \quad \|y\|_{C(X)}^2 \leq c(\|Cy\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda\|_{L_2(U)}^2) + \min\{\|l_2\|_{L_2(X)}^2, \|l_2\|_{L_1(X)}^2\} + \|y_0\|^2.$$

Similar via (4.7) we obtain:

$$(4.16) \quad \|\lambda\|_{C(X)}^2 \leq c(\|Cy\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda\|_{L_2(U)}^2) + \min\{\|l_1\|_{L_2(X)}^2, \|l_1\|_{L_1(X)}^2\} + \|\lambda_T\|^2.$$

To obtain an estimate for  $\|y\|_{L_2(X)}^2$  we have to integrate (4.14) over  $[0, T]$  and apply Lemma 4.5 to the integral terms in (4.14). Setting:

$$w(s) := \|l_2(s)\|^2 \quad \Rightarrow \quad h_1(t) = \int_0^t \|l_2(s)\|^2 e^{-k_\varphi(t-s)} ds,$$

we obtain with Lemma 4.5:

$$\int_0^T \int_0^t \|l_2(s)\|^2 e^{-k_\varphi(t-s)} ds dt = \|h_1\|_{L_1(0,T)} \leq c\|w\|_{L_1(0,T)} = c\|l_2\|_{L_2(X)}^2$$

and similarly:

$$\int_0^T \int_0^t (\|Cy(s)\|^2 + \|R^{-*}B^*\lambda(s)\|^2) e^{-k_\varphi(t-s)} ds dt \leq c(\|Cy\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda\|_{L_2(U)}^2).$$

If we set

$$w(s) := \|l_2(s)\| \quad \Rightarrow \quad h_1(t) = \int_0^t \|l_2(s)\| e^{-k_\varphi(t-s)} ds,$$

we obtain

$$\int_0^T \left( \int_0^t \|l_2(s)\| e^{-k_\varphi(t-s)} \right)^2 ds dt = \|h_1\|_{L_2(0,T)}^2 \leq c \|w\|_{L_1(0,T)}^2 = c \|l_2\|_{L_1(X)}^2.$$

This yields the desired  $L_2$ -estimate:

$$(4.17) \quad \|y\|_{L_2(X)}^2 \leq c(\|Cy(s)\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda(s)\|_{L_2(U)}^2 + \min\{\|l_2\|_{L_2(X)}^2, \|l_2\|_{L_1(X)}^2\} + \|y_0\|^2).$$

In the same way we obtain

$$(4.18) \quad \|\lambda\|_{L_2(X)}^2 \leq c(\|Cy(s)\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda(s)\|_{L_2(U)}^2 + \min\{\|l_1\|_{L_2(X)}^2, \|l_1\|_{L_1(X)}^2\} + \|\lambda_T\|^2).$$

Now we take the maximum of (4.15) and (4.17) and add it to the maximum of (4.16) and (4.18). Using the definition of the norms

$$\|v\|_{2\wedge\infty}^2 = \max\{\|v\|_{L_2(X)}^2, \|v\|_{C(X)}^2\}, \quad \|w\|_{1\vee 2}^2 = \min\{\|w\|_{L_1(X)}^2, \|w\|_{L_2(X)}^2\}$$

our result follows.  $\square$

The first two terms on the right-hand side of (4.8) still depend on the state and the adjoint. We therefore present the following representation formula, motivated by [24, Proof of Theorem 1].

LEMMA 4.7. *Let  $(y, \lambda)$  solve (4.1). Then*

$$(4.19) \quad \|Cy\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda\|_{L_2(U)}^2 = \langle \lambda_T, y(T) \rangle - \langle y_0, \lambda(0) \rangle + \int_0^T \langle l_2(s), \lambda(s) \rangle - \langle l_1(s), y(s) \rangle ds \\ \leq c(\|\lambda_T\| \|y(T)\| + \|y_0\| \|\lambda(0)\| + \|l_2\|_{1\vee 2} \|\lambda\|_{2\wedge\infty} + \|l_1\|_{1\vee 2} \|y\|_{2\wedge\infty}).$$

*Proof.* We apply Lemma 3.4 to the state and adjoint equation, which yields

$$\langle \lambda_T, y(T) \rangle - \langle y_0, \lambda(0) \rangle = \int_0^T \langle l_2(s), \lambda(s) \rangle - \langle l_1(s), y(s) \rangle - \|Cy(s)\|^2 - \|R^{-*}B^*\lambda(s)\|^2 ds$$

Rearranging the terms yields the result.  $\square$

THEOREM 4.8. *Let Assumption 4.2 hold. Then there is  $c \geq 0$  independent of  $T$  such that*

$$\|M^{-1}\|_{((L_1(X), \|\cdot\|_{1\vee 2}) \times X)^2 \rightarrow (C(X), \|\cdot\|_{2\wedge\infty})^2} \leq c.$$

*Proof.* Consider  $z := (y, \lambda) \in C(X)^2$  and  $r := (l_1, \lambda_T, l_2, y_0) \in (L_1(X) \times X)^2$  which satisfy (4.1). Thus, as shown in Lemma 4.6 the estimate (4.8) applies. We substitute (4.19) into (4.8) and apply the Hölder-like inequality of Lemma 4.1 to the integral terms:

$$\|z\|_{2\wedge\infty}^2 := \|y\|_{2\wedge\infty}^2 + \|\lambda\|_{2\wedge\infty}^2 \\ \leq c(\|\lambda_T\| \|y(T)\| + \|y_0\| \|\lambda(0)\| + \|l_2\|_{1\vee 2} \|\lambda\|_{2\wedge\infty} + \|l_1\|_{1\vee 2} \|y\|_{2\wedge\infty} + \|r\|_{1\vee 2}^2) \\ \leq c((\|\lambda_T\| + \|l_1\|_{1\vee 2}) \|y\|_{2\wedge\infty} + (\|y_0\| + \|l_2\|_{1\vee 2}) \|\lambda\|_{2\wedge\infty} + \|r\|_{1\vee 2}^2) \\ \leq c(\|r\|_{1\vee 2} \|z\|_{2\wedge\infty} + \|r\|_{1\vee 2}^2).$$

Application of the simple estimate  $c\|r\|_{1\vee 2} \|z\|_{2\wedge\infty} \leq \frac{1}{2}(c^2\|r\|_{1\vee 2}^2 + \|z\|_{2\wedge\infty}^2)$  yields

$$\|M^{-1}r\|_{2\wedge\infty} = \|z\|_{2\wedge\infty} \leq c\|r\|_{1\vee 2}$$

and thus the desired result.  $\square$

**5. Extension to boundary control with admissible control operators.** In this section, we extend the results of sections 3 and 4 to the case of a control operator  $B$  which is unbounded as mapping into  $X$  but admissible for the semigroup generated by  $\Lambda$  in the sense of [26, Chapter 4]. We refer to [26, Chapter 4] for an in-depth introduction to this topic. The case of an unbounded operator often arises in the case of boundary control. As a consequence of this the operator norm  $\|B\|_{U \rightarrow X} = \|B^*\|_{X \rightarrow U}$  is no longer finite. Inspection of the proofs in sections 3 and 4 yields that these operator norms are only used in inequality (4.10). In addition, the estimate of the control in Theorem 3.2 is performed via  $R^*Ru = B^*\lambda$  using  $\|B^*\|_{X \rightarrow U}$  in (3.9). All remaining estimates and constants do not involve norms of  $B$ .

The goal of this section is to replace boundedness of  $B$  by a weaker assumption, which is known as *admissibility*. Our strategy of proof will be to show surrogates for (4.10) and (3.9), which allow us to generalize our main results to the case of admissible  $B$ . The remaining steps of the proofs remain unchanged. The sensitivity results in Theorem 3.2 for the state and adjoint directly carry over as stated in Theorem 5.4, whereas the estimate for the control in the proof of Theorem 3.2 involves the norm of the control operator. We therefore present a modification of this in the proof of Theorem 5.4. However, one only obtains an integral estimate but no uniform estimate on the control in Theorem 3.2. This is the only price to pay for going from bounded to unbounded but admissible control operators.

**5.1. Admissible control and observation operators.** For ease of reading we recall the definition of admissible control and observation operators. Let  $\Lambda : D(\Lambda) \subset X \rightarrow X$  be the generator of a semigroup  $(\mathcal{T}(t))_{t \geq 0}$  on  $X$ . Moreover, let  $\Lambda^*$  be the adjoint operator of  $\Lambda$  with domain  $D(\Lambda^*)$ . Let  $X_1$  be  $D(\Lambda)$  equipped with the norm  $\|\cdot\|_1 = \|(\beta I - \Lambda) \cdot\|$  for  $\beta \in \rho(\Lambda)$ , where  $\rho(\Lambda) := \{\beta \in \mathbb{C} \mid \beta I - \Lambda \text{ is invertible and } (\beta I - \Lambda)^{-1} \in L(X, X)\}$  is the resolvent set of  $\Lambda$ . Furthermore, again for  $\beta \in \rho(\Lambda)$ , we define  $X_{-1}$  to be the completion of  $X$  with respect to the norm  $\|\cdot\|_{-1} = \|(\beta I - \Lambda)^{-1} \cdot\|$ . We note that the norms  $\|\cdot\|_1$  for different  $\beta$  are equivalent, see [26, Proposition 2.10.1], and the same also holds true for  $\|\cdot\|_{-1}$ , see [26, Proposition 2.10.2]. Furthermore, by e.g. [26, Proposition 2.10.4], the semigroup  $(\mathcal{T}(t))_{t \geq 0}$  can be extended to a semigroup on  $X_{-1}$ .

DEFINITION 5.1. ([26, Definition 4.2.1, Definition 4.3.1])

- i)  $B \in L(U, X_{-1})$  is called an *admissible control operator* for the semigroup  $(\mathcal{T}(t))_{t \geq 0}$ , if for some  $\tau > 0$ ,  $\text{Ran } \Phi_\tau \subset X$ , where for  $u \in L_2(0, \infty; U)$ ,

$$\Phi_\tau u := \int_0^\tau \mathcal{T}(\tau - s)Bu(s) ds.$$

- ii)  $C \in L(X_1, Y)$  is called an *admissible observation operator* for the semigroup  $(\mathcal{T}(t))_{t \geq 0}$  if for some  $\tau > 0$ ,  $\Psi_\tau$  has a continuous extension to  $X$ , where for  $z_0 \in X_1$ ,

$$(\Psi_\tau z_0)(t) := \begin{cases} C\mathcal{T}(t)z_0 & \text{for } t \in [0, \tau] \\ 0 & \text{for } t > \tau. \end{cases}$$

Note, that if i) and ii) in Definition 5.1 are satisfied for one  $\tau \geq 0$ , they hold for all  $\tau \geq 0$ , see [26, Proposition 4.2.2, Proposition 4.3.2]. We briefly recall properties of admissible control operators, which are important in the remainder of this section.

PROPOSITION 5.2. Let  $B$  be an admissible control operator for the semigroup  $(\mathcal{T}(t))_{t \geq 0}$ . Then,

- i)  $B^*$  is an admissible observation operator for the adjoint semigroup  $(\mathcal{T}^*(t))_{t \geq 0}$ .

ii) For all  $t \geq 0$  and  $y(t) \in X$  there exists a constant  $K_t \geq 0$ , such that

$$(5.1) \quad \int_0^t \|B^* \mathcal{T}^*(t-s)y(t)\|^2 ds \leq K_t^2 \|y(t)\|^2.$$

iii) If  $(\mathcal{T}^*(t))_{t \geq 0}$  is exponentially stable, the constant  $K_t$  can be chosen independently of  $t$ .

*Proof.* i) follows by the duality result [26, Theorem 4.4.3]. For ii), see [26, Definition 4.3.1]. The fact that the bound can be chosen independently of  $t$ , as stated in iii), follows by [26, Remark 4.3.5].  $\square$

In the following, we will assume throughout that  $B$  is an admissible control operator for the semigroup generated by  $\Lambda$  and that [Assumption 4.2](#) holds. It follows by i) that  $B^*$  is an admissible control operator for the semigroup generated by  $\Lambda^*$ . By a perturbation result, cf. [26, Theorem 5.4.2],  $B^*$  is also admissible for the exponentially stable semigroup  $(\mathcal{T}^*(t))_{t \geq 0}$  generated by  $\Lambda^* + C^* K_C$ , since  $K_C C \in L(X, X)$ . Hence, we may apply (5.1) to  $B^*$  with  $K_t$  independent of  $t$ . This will turn out as an appropriate replacement for the assumption  $\|B^*\| < \infty$ .

**5.2. Bounds for the solution operator.** We will now use these properties to derive the desired bound on the operator norm in the case of unbounded but admissible control operators.

**THEOREM 5.3.** *Consider [Problem 2.1](#) except that  $B$  is only an admissible control operator for the semigroup generated by  $\Lambda$ . Let [Assumption 4.2](#) hold. Then there is  $c \geq 0$  independent of  $T$ , such that*

$$\|M^{-1}\|_{((L_1(X), \|\cdot\|_{1 \vee 2}) \times X)^2 \rightarrow (C(X), \|\cdot\|_{2 \wedge \infty})^2} \leq c.$$

*Proof.* As already noted, the only step in the proof of [Theorem 4.8](#) where the operator norm of  $B^*$  is needed, is the proof of (4.10). It is thus sufficient to show a modification of this inequality which circumvents this operator norm estimate by using exponential stability of the test functions and the fact that  $B$  is an admissible control operator.

Let  $\varphi$  solve (4.2), i.e.,  $\varphi(s) = \mathcal{T}^*(t-s)y(t)$  where, as above,  $(\mathcal{T}^*(t))_{t \geq 0}$  is the semigroup generated by  $(\Lambda^* + C^* K_C)$ . We show that the critical estimate (4.10) still holds with different constants which do not involve the operator norm of  $B$ . First, a simple calculation using the fact that  $R^{-*}$  is bounded from  $U$  to  $U$  and the Cauchy-Schwarz inequality twice yields

$$(5.2) \quad \begin{aligned} \int_0^t |\langle R^{-*} B^* \lambda(s), R^{-*} B^* \varphi(s) \rangle| ds &\leq c \int_0^t e^{-\frac{k_\varphi}{2}(t-s)} \|R^{-*} B^* \lambda(s)\| \|B^* e^{\frac{k_\varphi}{2}(t-s)} \varphi(s)\| ds \\ &\leq c \sqrt{\int_0^t e^{-k_\varphi(t-s)} \|R^{-*} B^* \lambda(s)\|^2 ds} \sqrt{\int_0^t \|B^* e^{\frac{k_\varphi}{2}(t-s)} \varphi(s)\|^2 ds}. \end{aligned}$$

Using the exponential stability  $\|\mathcal{T}^*(t-s)\| \leq M_\varphi e^{-k_\varphi(t-s)}$ , the scaled semigroup  $(e^{\frac{k_\varphi}{2}(t-s)} \mathcal{T}^*(t))_{t \geq 0}$  is still exponentially stable, and hence we employ [Proposition 5.2](#) ii) and iii) and get

$$\sqrt{\int_0^t \|B^* e^{\frac{k_\varphi}{2}(t-s)} \varphi(s)\|^2 ds} = \sqrt{\int_0^t \|B^* e^{\frac{k_\varphi}{2}(t-s)} \mathcal{T}^*(t-s)y(t)\|^2 ds} \leq K \|y(t)\|$$

with  $K$  independent of  $t$ . Inserting this into (5.2), we conclude

$$\int_0^t |\langle R^{-*} B^* \lambda(s), R^{-*} B^* \varphi(s) \rangle| ds \leq c \|y(t)\| \sqrt{\int_0^t e^{-k_\varphi(t-s)} \|R^{-*} B^* \lambda(s)\|^2 ds},$$

which yields the desired replacement for (4.10) with a different constant independent of the norm of  $B$ . As the remaining results in section 4, namely Lemmas 4.6 and 4.7 and Theorem 4.8 do not use the boundedness of  $B$ , we can conclude the result analogously to the bounded case.  $\square$

As a consequence of Theorem 5.3, the estimates in Theorem 3.2 also hold true in the case of unbounded control operators with constants independent of the horizon  $T$ , except for the uniform estimate for the control. This is the statement of the following theorem.

**THEOREM 5.4.** *Consider Problem 2.1 except that  $B$  is an unbounded but admissible control operator for the semigroup generated by  $\Lambda$ . Let Assumption 4.2 hold. Assume  $(\delta y, \delta \lambda) \in C(X)^2$  solves (3.1) with  $\varepsilon_1, \varepsilon_2 \in L_1(X)$ . Then there is a constant  $c \geq 0$  and a scaling factor  $\mu > 0$ , both independent of  $T$ , such that defining*

$$\rho := \|e^{-\mu t} \varepsilon_1(t)\|_E + \|e^{-\mu t} \varepsilon_2(t)\|_E$$

for  $E := L_1(X)$  or  $E := L_2(X)$ , we have

$$\begin{aligned} \|e^{-\mu t} \delta y\|_{L_2(X)} + \|e^{-\mu t} \delta u\|_{L_2(U)} + \|e^{-\mu t} \delta \lambda\|_{L_2(X)} &\leq c\rho, \\ \|e^{-\mu t} \delta y\|_{C(X)} + \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho. \end{aligned}$$

*Proof.* First, choosing  $\mu$ , such that  $\mu \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2} < 1$ , we conclude by the same reasoning as in the proof of Theorem 3.2 the estimates for the state and adjoint

$$(5.3) \quad \begin{aligned} \|e^{-\mu t} \delta y\|_{L_2(X)} + \|e^{-\mu t} \delta \lambda\|_{L_2(X)} &\leq c\rho, \\ \|e^{-\mu t} \delta y\|_{C(X)} + \|e^{-\mu t} \delta \lambda\|_{C(X)} &\leq c\rho, \end{aligned}$$

with  $c \geq 0$  independent of  $T$  as the occurring operator norms can be bounded by a constant  $c \geq 0$  independently of  $T$  by Theorem 5.3. To estimate the control, we compute, setting  $\widetilde{\delta y}(t) := e^{-\mu t} \delta y(t)$ ,  $\widetilde{\delta \lambda}(t) := e^{-\mu t} \delta \lambda(t)$  and  $\widetilde{\delta u}(t) := e^{-\mu t} \delta u$ , that

$$\|\widetilde{\delta u}\|_{L_2(U)}^2 = \int_0^T \|Q^{-1} B^* \widetilde{\delta \lambda}(t)\|^2 dt \leq c \int_0^T \|C \widetilde{\delta y}(t)\|^2 + \|R^{-*} B^* \widetilde{\delta \lambda}(t)\|^2 dt.$$

Similar to Lemma 4.7, now for the scaled system (3.2), using Lemma 3.4 for the state and adjoint equation, we obtain with  $\widetilde{\varepsilon}_i(t) = e^{-\mu t} \varepsilon_i(t)$ ,  $i = 1, 2$  using the Hölder inequality of Lemma 4.1

$$\begin{aligned} \int_0^T \|C \widetilde{\delta y}(t)\|^2 + \|R^{-*} B^* \widetilde{\delta \lambda}(t)\|^2 dt &= \int_0^T -\langle \widetilde{\varepsilon}_1(t), \widetilde{\delta y}(t) \rangle + \langle \widetilde{\varepsilon}_2(t), \widetilde{\delta \lambda}(t) \rangle - 2\mu \langle \widetilde{\delta y}(t), \widetilde{\delta \lambda}(t) \rangle dt \\ &\leq (\|\widetilde{\varepsilon}_1\|_{1 \vee 2} + \|\widetilde{\varepsilon}_2\|_{1 \vee 2}) (\|\widetilde{\delta y}\|_{2 \wedge \infty} + \|\widetilde{\delta \lambda}\|_{2 \wedge \infty}) + 2\mu \|\widetilde{\delta y}\|_{L_2(X)} \|\widetilde{\delta \lambda}\|_{L_2(X)} \\ &\leq c\rho^2 \end{aligned}$$

where in the last estimate we used the assumption on  $\widetilde{\varepsilon}_i$ ,  $i = 1, 2$  and the estimate on the state and the adjoint (5.3). Taking the square root yields the result for the control.  $\square$

*Remark 5.5.* In a similar fashion, one could allow for an admissible observation operator which is unbounded on  $X$  and bounded control operators  $B \in L(U, X)$  in this section. The case where both  $C$  and  $B$  are unbounded but admissible cannot be included in all generality. This is due to the non-existence of perturbation results for this case, i.e., not every admissible observation operator for  $\Lambda$  is admissible for  $\Lambda + BK$ , where  $B$  is an admissible control operator and  $K$  is a bounded

feedback operator. Mixed perturbation results of this kind are a very delicate matter and we refer to [26, Proposition 5.5.2, Example 5.5.3, Proposition 10.1.10] and the discussion after [26, Corollary 5.5.1] for results on that subject. Moreover, we refer to the example in [26, Section 10.8] where such a perturbation result holds.

*Remark 5.6.* If we have additional structure, the results of [Theorem 4.8](#) and [Theorem 5.3](#) can be refined by bootstrapping arguments. For instance, in the parabolic case there is an additional, more regular space with the continuous dense embedding  $V \hookrightarrow X$ , such that  $\Lambda$  can be extended to an operator  $\bar{\Lambda} \in L(V, V^*)$  and we have a Gårding inequality for  $y \in L_2(0, T; V)$ :

$$(5.4) \quad \exists \omega \in \mathbb{R}, \alpha > 0 : \quad \alpha \|y\|_{L_2(V)}^2 \leq \int_0^T -\langle \bar{\Lambda}y(t), y(t) \rangle_{V^* \times V} dt + \omega \|y\|_{L_2(X)}^2.$$

In this particular case, the state equation yields the regularity  $y \in L_2(0, T; V)$ ,  $y' \in L_2(0, T; V^*)$  and we compute, using integration by parts on  $\langle y', y \rangle_{V^* \times V}$ , [\(4.19\)](#), and assuming boundedness of  $\|B^*\|_{V \rightarrow U}$  (which is weaker than boundedness of  $\|B^*\|_{X \rightarrow U}$ ):

$$\begin{aligned} \int_0^T -\langle \bar{\Lambda}y(t), y(t) \rangle_{V^* \times V} dt &= \int_0^T -\langle y'(t), y(t) \rangle_{V^* \times V} + \langle l_2(t) + B(R^*R)^{-1}B^*\lambda(t), y(t) \rangle_X dt \\ &\leq \frac{1}{2} (\|y(0)\|^2 - \|y(T)\|^2) + \|l_2\|_{1V2} \|y\|_{2\wedge\infty} + \|R^{-*}B^*\lambda\|_{L_2(U)} \|R^{-*}B^*y\|_{L_2(U)} \\ &\leq \frac{1}{2} \|y_0\|^2 + \|r\|_{1V2} \|z\|_{2\wedge\infty} + c(\|r\|_{1V2} + \|z\|_{2\wedge\infty}) \|R^{-*}B^*\|_{V \rightarrow U} \|y\|_{L_2(V)}. \end{aligned}$$

By [Theorem 5.4](#) we may estimate  $\|z\|_{2\wedge\infty}$  by  $\|r\|_{1V2}$  and substitute the result into [\(5.4\)](#). A short computation yields  $\|y\|_{L_2(V)} \leq c\|r\|_{1V2}$  and similarly  $\|\lambda\|_{L_2(V)} \leq c\|r\|_{1V2}$ . Hence, there is a  $T$ -independent bound:

$$\|M^{-1}\|_{(L_1(X), \|\cdot\|_{1V2} \times X)^2 \rightarrow L_2(V)^2} \leq c.$$

By further bootstrapping via:

$$y' = \bar{\Lambda}y + B(R^*R)^{-1}B^*\lambda + l_2 \quad \text{in } L_2(V^*)$$

we obtain  $\|y'\|_{L_2(V^*)} \leq c(\|r\|_{1V2} + \|l_2\|_{L_2(X)})$  and similarly an estimate for  $\|\lambda'\|_{L_2(V^*)}$ . Thus, also for the parabolic space  $W([0, T])$ , which is equipped with the norm  $\|v\|_{W([0, T])} = \|v\|_{L_2(V)} + \|v'\|_{L_2(V^*)}$ , we get the  $T$ -independent bound:

$$\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow W([0, T])^2} \leq c.$$

These additional estimates can be used to obtain results in [Theorem 3.2](#) and [Theorem 5.4](#), and also in [Theorem 6.2](#) below in the respective additional norms.

**6. An exponential turnpike result.** In this section, we extend the scaling approach employed in [Theorem 3.2](#) to deduce a turnpike result. It is an adaption of existing results for parabolic equations in [14, Theorem 5.2, Corollary 5.3] and [5, Lemma 15]. In the case of optimal control problems governed by general evolution equations in Hilbert spaces, turnpike theorems were given in [24] for  $C_0$ -semigroups with bounded control and observation operators, and for boundary controlled parabolic equations. Here we extend these results to the boundary control of non-parabolic



equations with admissible control operators. First, we introduce the solution of a steady state optimization problem, namely  $(\bar{y}, \bar{\lambda}) \in X^2$  being a minimizer of

$$(6.1) \quad \begin{aligned} \min_{\bar{y}, \bar{u}} \quad & \frac{1}{2} \|C(\bar{y} - y_d)\|_Y^2 + \frac{1}{2} \|R(\bar{u} - u_d)\|_U^2 \\ \text{s.t.} \quad & -\Lambda \bar{y} - B\bar{u} = 0, \end{aligned}$$

or, equivalently, as the problem is convex, by coercivity of  $R$ ,  $(\bar{y}, \bar{\lambda})$  solve the corresponding first order conditions

$$(6.2) \quad \begin{pmatrix} C^*C & -\Lambda^* \\ -\Lambda & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\lambda} \end{pmatrix} = \begin{pmatrix} C^*C y_d \\ B u_d \end{pmatrix},$$

where  $Q = R^*R$  and we eliminated the control via  $\bar{u} = Q^{-1}B^*\bar{\lambda} + u_d$ .

LEMMA 6.1. *Let  $(y, u, \lambda)$  solve [Problem 2.1](#) or, equivalently, [\(2.2\)](#) with  $f = 0$ , where  $B$  is allowed to be an unbounded but admissible operator for the semigroup generated by  $\Lambda$ . Moreover, let  $(\bar{y}, \bar{u}, \bar{\lambda})$  solve the corresponding steady state problem [\(6.1\)](#). Then  $(\delta y, \delta \lambda) := (y - \bar{y}, \lambda - \bar{\lambda})$  solves*

$$(6.3) \quad \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ -\bar{\lambda} \\ 0 \\ y_0 - \bar{y} \end{pmatrix},$$

where  $E_0 \delta y := \delta y(0)$  and  $E_T \delta \lambda := \delta \lambda(T)$  and  $\delta u := u - \bar{u} = Q^{-1}B^*\delta \lambda$ .

*Proof.* Using [\(6.2\)](#) and  $\frac{d}{dt} \bar{y} = \frac{d}{dt} \bar{\lambda} = 0$  yields that  $(\bar{y}, \bar{\lambda})$  satisfies

$$(6.4) \quad \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\lambda} \end{pmatrix} = \begin{pmatrix} C^*C y_d \\ \bar{\lambda} \\ B u_d \\ \bar{y} \end{pmatrix}$$

Considering the linear system for  $(y, u, p)$  solving [\(2.2\)](#) with  $f = 0$ , i.e.,

$$(6.5) \quad \begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} C^*C y_d \\ 0 \\ B u_d \\ y_0 \end{pmatrix}$$

we conclude the result by subtracting [\(6.4\)](#) from [\(6.5\)](#).  $\square$

THEOREM 6.2. *Assume  $(y, u, \lambda)$  solves [\(2.2\)](#), where  $B$  is allowed to be an unbounded but admissible control operator for the semigroup generated by  $\Lambda$ . Moreover, let  $(\bar{y}, \bar{u}, \bar{\lambda})$  solve the corresponding steady state problem [\(6.1\)](#). Let [Assumption 4.2 A.1](#) and [A.2](#) hold. Then, defining  $(\delta y, \delta u, \delta \lambda) := (y - \bar{y}, u - \bar{u}, \lambda - \bar{\lambda})$ , there exist  $\mu > 0$  and a constant  $c \geq 0$ , both independent of  $T$ , such that*

$$(6.6) \quad \begin{aligned} \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta y \right\|_{L_2(X)} + \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta u \right\|_{L_2(U)} + \\ + \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta \lambda \right\|_{L_2(X)} \leq c (\|y_0 - \bar{y}\| + \|\bar{\lambda}\|), \end{aligned}$$

$$(6.7) \quad \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta y \right\|_{C(X)} + \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta \lambda \right\|_{C(X)} \leq c (\|y_0 - \bar{y}\| + \|\bar{\lambda}\|).$$

*Proof.* By Lemma 6.1,  $(\delta y, \delta u, \delta \lambda)$  solves (6.3). By Theorem 5.4,  $\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2}$  can be bounded independently of  $T$ . If we introduce a scaling factor  $0 < \mu < \frac{1}{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2}}$  and scaled variables  $\widetilde{\delta y} := \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta y$ ,  $\widetilde{\delta u} := \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta u$  and  $\widetilde{\delta \lambda} := \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta \lambda$ , we can compute

$$(6.3) \quad \Leftrightarrow \left( \underbrace{\begin{pmatrix} C^*C & -\frac{d}{dt} - \Lambda^* \\ 0 & E_T \\ \frac{d}{dt} - \Lambda & -BQ^{-1}B^* \\ E_0 & 0 \end{pmatrix}}_{=:M} + \mu \underbrace{\begin{pmatrix} 0 & F \\ 0 & 0 \\ -F & 0 \\ 0 & 0 \end{pmatrix}}_{=:P} \right) \begin{pmatrix} \widetilde{\delta y} \\ \widetilde{\delta \lambda} \end{pmatrix} = \frac{1}{1 + e^{-\mu T}} \begin{pmatrix} 0 \\ -\bar{\lambda} \\ 0 \\ y_0 - \bar{y} \end{pmatrix}$$

where  $F := \frac{e^{-\mu(T-t)} - e^{-\mu t}}{e^{-\mu t} + e^{-\mu(T-t)}}$  and the factor  $\frac{1}{1 + e^{-\mu T}}$  arises due to the scaling of the initial values. The proof for the estimate of the state and the adjoint in (6.6) and (6.7) is similar to the one of Theorem 3.2. Defining  $\tilde{z} := (\widetilde{\delta y}, \widetilde{\delta \lambda})$  and  $\tilde{r} := \frac{1}{1 + e^{-\mu T}} (0, -\bar{\lambda}, 0, y_0 - \bar{y})$ , we get

$$(6.8) \quad (M + \mu \tilde{P})\tilde{z} = \tilde{r} \quad \Rightarrow \quad (I + \mu \tilde{P}M^{-1})M\tilde{z} = \tilde{r} \quad \Rightarrow \quad \tilde{z} = M^{-1}(I + \mu \tilde{P}M^{-1})^{-1}\tilde{r}.$$

Again, as in the proof of Theorem 3.2 by a standard Neumann series argument, cf. [18, Theorem 2.14], as we chose  $\mu$  such that  $\beta := \mu \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2} < 1$ ,  $(I + \mu \tilde{P}M^{-1})$  is invertible as  $\|F\|_{L_2(X) \rightarrow L_2(X)} \leq 1$  and thus  $\|\tilde{P}\|_{L_2(X)^2 \rightarrow (L_2(X) \times X)^2} \leq 1$ . Therefore, using the Neumann series representation of  $(I + \mu \tilde{P}M^{-1})^{-1}$  yields

$$(6.9) \quad \|(I + \mu \tilde{P}M^{-1})^{-1}\|_{(L_2(X) \times X)^2 \rightarrow (L_2(X) \times X)^2} \leq \sum_{i=0}^{\infty} \|(\mu \tilde{P}M^{-1})^i\|_{(L_2(X) \times X)^2 \rightarrow (L_2(X) \times X)^2} \leq \sum_{i=0}^{\infty} \beta^i = \frac{1}{1 - \beta}.$$

Thus, we conclude with (6.8) and (6.9)

$$(6.10) \quad \begin{aligned} \|\tilde{z}\|_{L_2(X)^2} &\leq \frac{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2}}{1 - \beta} \|\tilde{r}\|_{(L_2(X) \times X)^2}, \\ \|\tilde{z}\|_{C(X)^2} &\leq \frac{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow C(X)^2}}{1 - \beta} \|\tilde{r}\|_{(L_2(X) \times X)^2}. \end{aligned}$$

Finally,  $\|\tilde{r}\|_{(L_2(X) \times X)^2} \leq \|y_0 - \bar{y}\| + \|\bar{\lambda}\|$  and going back to the original variables yields the result (6.7).

The remainder of this proof will now consist of estimating the control to conclude (6.6). To this end, we recall the approach taken in the proof of Theorem 5.4 and compute

$$(6.11) \quad \|\widetilde{\delta u}\|_{L_2(U)}^2 = \int_0^T \|Q^{-1}B^*\widetilde{\delta \lambda}(t)\|^2 dt \leq c \int_0^T \|C\widetilde{\delta y}(t)\|^2 + \|R^{-*}B^*\widetilde{\delta \lambda}(t)\|^2 dt.$$

Similar to Lemma 4.7, using Lemma 3.4 for the scaled state and adjoint equation of system, we

obtain

$$\begin{aligned}
& \int_0^T \|C\widetilde{\delta y}(t)\|^2 + \|B^*\widetilde{\delta\lambda}(t)\|^2 dt \\
&= \langle \widetilde{\delta y}(T), \widetilde{\delta\lambda}(T) \rangle - \langle \widetilde{\delta y}(0), \widetilde{\delta\lambda}(0) \rangle + \mu \int_0^T \langle F\widetilde{\delta y}(t), \widetilde{\delta\lambda}(t) \rangle + \langle \widetilde{\delta y}(t), F\widetilde{\delta\lambda}(t) \rangle dt \\
&\leq \|\widetilde{\delta y}(T)\| \|\widetilde{\delta\lambda}(T)\| + \|\widetilde{\delta y}(0)\| \|\widetilde{\delta\lambda}(0)\| + 2\mu \|\widetilde{\delta y}\|_{L_2(X)} \|\widetilde{\delta\lambda}\|_{L_2(X)} \\
(6.12) \quad &\leq (\|\widetilde{\delta y}(T)\| + \|\widetilde{\delta\lambda}(0)\|)(\|\widetilde{\delta y}(0)\| + \|\widetilde{\delta\lambda}(T)\|) + \mu(\|\widetilde{\delta y}\|_{L_2(X)}^2 + \|\widetilde{\delta\lambda}\|_{L_2(X)}^2),
\end{aligned}$$

where we used  $\|F\|_{L_2(X) \rightarrow L_2(X)} \leq 1$ . In order to estimate the end time value of the state and the initial value of the adjoint we compute:

$$\begin{aligned}
\|\widetilde{\delta y}(T)\| + \|\widetilde{\delta\lambda}(0)\| &= \frac{1}{1 + e^{-\mu T}} (\|\delta y(T)\| + \|\delta\lambda(0)\|) \leq \|\delta y\|_{C(X)} + \|\delta\lambda\|_{C(X)} \\
&\leq \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow C(X)^2} (\|y_0 - \bar{y}\| + \|\bar{\lambda}\|).
\end{aligned}$$

Inserting this into (6.12) and (6.11), and using the estimate for the state and adjoint (6.10), we get

$$\begin{aligned}
\|\widetilde{\delta u}\|_{L_2(U)}^2 &\leq c \left( \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow C(X)^2} (\|y_0 - \bar{y}\|^2 + \|\bar{\lambda}\|^2) \right. \\
&\quad \left. + \mu \frac{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2}^2}{(1 - \beta)^2} (\|y_0 - \bar{y}\|^2 + \|\bar{\lambda}\|^2) \right),
\end{aligned}$$

where taking the square root and using  $\mu \|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2} < 1$  yields

$$\|\widetilde{\delta u}\|_{L_2(U)} \leq c \left( \sqrt{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow C(X)^2}} + \frac{\sqrt{\|M^{-1}\|_{(L_2(X) \times X)^2 \rightarrow L_2(X)^2}}}{(1 - \beta)^2} \right) (\|y_0 - \bar{y}\| + \|\bar{\lambda}\|).$$

**Theorem 5.3** yields bounds for the operator norms independently of  $T$ . Together with (6.10), this gives the estimate (6.6).  $\square$

*Remark 6.3.* Similarly to **Theorem 3.2**, one could deduce a bound on  $\|\frac{1}{e^{-\mu t} + e^{-\mu(T-t)}}(u(t) - \bar{u})\|_{L_\infty(U)}$ , if the control operator is bounded.

We give a short interpretation of the two estimates given in **Theorem 6.2**. For the first inequality, i.e., (6.6) consider a fixed  $0 < \varepsilon < \frac{1}{2}$  and  $t \in [\varepsilon T, (1 - \varepsilon)T]$ . Then, if  $T \rightarrow \infty$ , the two scaling terms  $e^{-\mu t}$  and  $e^{-\mu(T-t)}$  approach zero exponentially fast and we estimate

$$\frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \geq \frac{1}{e^{-\mu\varepsilon T} + e^{-\mu(T-(1-\varepsilon)T)}} = \frac{2}{e^{-\mu\varepsilon T}}$$

and hence e.g. for the difference of state  $y$  and its turnpike  $\bar{y}$ ,

$$\int_0^T \left\| \frac{y(t) - \bar{y}}{e^{-\mu t} + e^{-\mu(T-t)}} \right\|^2 dt \geq \int_{\varepsilon T}^{(1-\varepsilon)T} \left\| \frac{y(t) - \bar{y}}{e^{-\mu t} + e^{-\mu(T-t)}} \right\|^2 dt \geq \frac{4}{e^{-2\mu\varepsilon T}} \int_{\varepsilon T}^{(1-\varepsilon)T} \|y(t) - \bar{y}\|^2 dt$$

which, using (6.6) of **Theorem 6.2**, implies

$$\int_{\varepsilon T}^{(1-\varepsilon)T} \|y(t) - \bar{y}\|^2 dt \leq ce^{-2\mu\varepsilon T} (\|y_0 - \bar{y}\| + \|\bar{\lambda}\|)^2.$$

Proceeding analogously for the adjoint and the control, we have

$$\|y - \bar{y}\|_{L_2(\varepsilon T, (1-\varepsilon)T; X)}, \|u - \bar{u}\|_{L_2(\varepsilon T, (1-\varepsilon)T; U)}, \|\lambda - \bar{\lambda}\|_{L_2(\varepsilon T, (1-\varepsilon)T; X)} \rightarrow 0, \quad \text{if } T \rightarrow \infty$$

i.e.,  $L_2$ -convergence on a part of the whole time interval  $[0, T]$ . The convergence rate is exponential and the part is growing linearly in  $T$ , as its length is  $(1 - 2\varepsilon)T$ . Moreover, its fraction of the whole interval is constant, as  $\frac{(1-2\varepsilon)T}{T} = 1 - 2\varepsilon$ . Thus, on a fixed percentage of the interval  $[0, T]$ , the state, control and adjoint converge to the turnpike in the  $L_2$  norm, as the horizon  $T$  goes to infinity.

For the second inequality, i.e., (6.7), rewriting the pointwise estimate, we have

$$\|y(t) - \bar{y}\| + \|\lambda(t) - \bar{\lambda}\| \leq c(e^{-\mu t} + e^{-\mu(T-t)})$$

for every  $t \in [0, T]$ . Therefore, if we fix  $0 < \varepsilon < \frac{1}{2}$  and take the maximum over all  $t \in [\varepsilon T, (1 - \varepsilon)T]$ , we get

$$\|y - \bar{y}\|_{C(\varepsilon T, (1-\varepsilon)T; X)} + \|\lambda - \bar{\lambda}\|_{C(\varepsilon T, (1-\varepsilon)T; X)} \leq 2ce^{-\mu\varepsilon T}.$$

The right-hand side approaches zero exponentially fast as  $T \rightarrow \infty$ . Therefore, for each  $\varepsilon \in ]0, \frac{1}{2}[$ , we obtain uniform exponential convergence on the interval  $[\varepsilon T, (1 - \varepsilon)T]$  of  $y$  and  $\lambda$  to  $\bar{y}$  and  $\bar{\lambda}$  respectively as  $T \rightarrow \infty$ . Again, similar to i), with  $\frac{(1-2\varepsilon)T}{T} = (1 - 2\varepsilon)$ , we conclude that on a fixed fraction of the whole interval  $[0, T]$ , the state and adjoint converge to the turnpike in the maximum norm.

**7. Examples.** Finally we provide examples for which the error estimation result for bounded control operators, i.e., Theorem 3.2, the error estimation results for unbounded but admissible control operators, i.e., Theorem 5.4 and the turnpike result Theorem 6.2 hold. First, we consider the interior control of a heat equation and secondly, the boundary control of a wave equation.

*Example 7.1.* (Interior control of an unstable heat equation) Let  $\Omega \subset \mathbb{R}^n$  with smooth boundary and  $\omega_c \subset \Omega$  be non-empty. For  $T > 0$ , we consider the evolution equation

$$\begin{aligned} \frac{\partial y}{\partial t} &= (\Delta + c^2 I)y + \chi_{\omega_c} u && \text{in } \Omega \times (0, T) \\ y &= 0 && \text{in } \partial\Omega \times (0, T) \\ y(x, 0) &= y_0 && \text{in } \Omega, \end{aligned}$$

where  $\chi_{\omega_c}$  is the characteristic function of the control domain  $\omega_c$ ,  $U = L_2(\omega_c)$  and  $y_0 \in L_2(\Omega)$ . Moreover, we consider an observation operator  $C = \chi_{\omega_o}$  for non-empty  $\omega_o \subset \Omega$ . As  $\Delta$  generates a semigroup on  $X = L_2(\Omega)$  with domain  $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$ , we obtain by classical perturbation results, cf. [8, Theorem 1.3], that  $(\Delta + c^2 I)$  generates a semigroup with the same domain as  $\Delta$ . If  $c^2 > \lambda_1$ , where  $\lambda_1$  is the smallest eigenvalue of the negative Laplacian, the uncontrolled system is unstable. Defining  $B : L_2(\omega_c) \rightarrow L_2(\Omega)$  via  $Bu := \chi_{\omega_c} u$ , the pair  $(\Delta, B)$  is exactly null controllable in finite time [3, 10], thus stabilizable. Analogously, it follows that  $(\Delta, C)$  is detectable. As  $B$  is bounded on  $L_2(X)$ , one could apply the sensitivity result Theorem 3.2 or the turnpike result Theorem 6.2 to optimal control problems governed by these operators. For moderate instability parameters  $c > 0$  with respect to the control and observation region, one can also obtain a turnpike and sensitivity result in the  $W([0, T])$ -norm with [14, Theorem 3.1, Theorem 5.2] and we refer to [14, Example 3.5] for this matter. However, if  $c$  is large, the results of [14] cannot be applied, while the results of the present paper still hold.

*Example 7.2* (Dirichlet control of a wave equation). Second, we provide an example of a hyperbolic PDE with unbounded but admissible control operator presented in [26, Section 10.9]. We consider the model of a vibrating membrane on  $\Omega \subset \mathbb{R}^2$ , where  $\Omega$  is a bounded  $C^2$  domain and we can take action through Dirichlet boundary control on a part  $\Gamma \subset \Omega$  of the boundary. Moreover, we assume that  $(\Omega, \Gamma, T)$  fulfill the so called Geometric Control Condition (GCC), which ensures that all geometric optics have to enter the control domain in a time smaller than  $T$ . A consequence of this condition is exact controllability, see [4, 23], and hence stabilizability. We consider the wave equation

$$\begin{aligned} \frac{\partial^2 w}{\partial t^2} &= \Delta w && \text{on } \Omega \times (0, T) \\ w &= 0 && \text{on } \partial\Omega \setminus \Gamma \times (0, T) \\ w &= u && \text{on } \Gamma \times (0, T) \\ w(x, 0) &= f(x), \quad \frac{\partial w}{\partial t}(x, 0) = g(x) && x \in \Omega, \end{aligned}$$

where  $f \in L_2(\Omega)$ ,  $g \in H^{-1}(\Omega)$  and  $u \in L_2(L_2(\Gamma))$ . It was shown in [26, Proposition 10.9.1], that one can deduce a corresponding well-posed boundary control system on  $X = L_2(\Omega) \times H^{-1}(\Omega)$  with generator  $\Lambda = \begin{pmatrix} 0 & I \\ -\Lambda_0 & 0 \end{pmatrix}$ ,  $D(\Lambda) = H_0^1(\Omega) \times L_2(\Omega)$ , where  $\Lambda_0$  is the Dirichlet Laplacian and a control operator  $B$  defined by

$$\begin{aligned} Bv &= \begin{pmatrix} 0 \\ \Lambda_0 Dv \end{pmatrix} && \forall v \in U = L_2(\Gamma) \\ B^* \begin{pmatrix} \varphi \\ \psi \end{pmatrix} &= -\frac{\partial}{\partial \nu} (\Lambda_0^{-1} \psi)|_{\Gamma} && \forall (\varphi, \psi) \in D(\Lambda), \end{aligned}$$

where  $D$  is the Dirichlet map and  $\frac{\partial}{\partial \nu}$  the outward normal derivative. We refer the reader to [26, Section 10.6, Section 10.9] for details. In particular,  $B$  is an admissible control operator for the semigroup generated by  $\Lambda$ . Moreover as the GCC is satisfied, the pair  $(\Lambda, B)$  is exponentially stabilizable. If we now consider any observation region  $\omega_o \subset \Omega$ , such that  $(\Omega, \omega_o, T)$  satisfy the GCC, then the pair  $(\Lambda, C)$  is exponentially detectable. As in this example the control operator is unbounded on  $X$ , one cannot apply [Theorem 3.2](#). Yet, [Theorem 5.4](#) and [Theorem 6.2](#) are applicable to the optimal control problems governed by this equation.

**Acknowledgments.** The authors are grateful to Enrique Zuazua, Felix L. Schwenninger and Christopher Straub for valuable discussions.

#### REFERENCES

- [1] B. D. ANDERSON AND P. V. KOKOTOVIC, *Optimal control problems over large time intervals*, Automatica, 23 (1987), pp. 355 – 363.
- [2] J. BALL, *Strongly continuous semigroups, weak solutions, and the variation of constants formula*, Proceedings of the American Mathematical Society, 63 (1977), pp. 370–373.
- [3] V. BARBU, *Controllability of parabolic and Navier-Stokes equations*, Scientiae Mathematicae Japonicae, 56 (2002), pp. 143–211.
- [4] C. BARDOS, G. LEBEAU, AND J. RAUCH, *Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary*, SIAM Journal on Control and Optimization, 30 (1992), pp. 1024–1065.

- [5] T. BREITEN AND L. PFEIFFER, *On the turnpike property and the receding-horizon method for linear-quadratic optimal control problems*. 2018, <https://arxiv.org/abs/1811.02421>.
- [6] 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 (2014), p. 19351957.
- [7] R. DORFMAN, P. A. SAMUELSON, AND R. M. SOLOW, *Linear programming and economic analysis*, Courier Corporation, 1958.
- [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 (2017), p. 297304.
- [10] A. FURSIKOV AND O. IMANUVILOV, *Controllability of evolution equations*, Suhak kangüirok, Seoul National University, 1996.
- [11] L. GRÜNE, *Approximation properties of receding horizon optimal control*, Jahresbericht Deutsche Mathematiker-Vereinigung, 118 (2016), pp. 3–37.
- [12] L. GRÜNE AND M. A. MÜLLER, *On the relation between strict dissipativity and turnpike properties*, Systems & Control Letters, 90 (2016), p. 4553.
- [13] L. GRÜNE AND J. PANNEK, *Nonlinear Model Predictive Control: Theory and Algorithms*, Springer, 2016.
- [14] L. GRÜNE, M. SCHALLER, AND A. SCHIELA, *Sensitivity analysis of optimal control for a class of parabolic PDEs motivated by model predictive control*. submitted, October 2018.
- [15] M. GUGAT AND F. HANTE, *On the turnpike phenomenon for optimal boundary control problems with hyperbolic systems*. 2018, <https://arxiv.org/abs/1806.04438>.
- [16] 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 (2016), p. 6170.
- [17] A. HAURIE, *Optimal control on an infinite time horizon: The turnpike approach*, Journal of Mathematical Economics, 3 (1976), pp. 81 – 102.
- [18] R. KRESS, *Linear integral equations*, Springer, 2014.
- [19] X. LI AND J. YONG, *Optimal Control Theory for Infinite Dimensional Systems*, Birkhäuser, 1995.
- [20] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York, NY, 1983.
- [21] A. PORRETTA AND E. ZUAZUA, *Long time versus steady state optimal control*, SIAM Journal on Control and Optimization, 51 (2013), pp. 4242–4273.
- [22] A. PORRETTA AND E. ZUAZUA, *Remarks on long time versus steady state optimal control*, Mathematical Paradigms of Climate Science, (2016), pp. 67–89.
- [23] J. RAUCH, M. TAYLOR, AND R. PHILLIPS, *Exponential decay of solutions to hyperbolic equations in bounded domains*, Indiana university Mathematics journal, 24 (1974), pp. 79–86.
- [24] 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 (2018), pp. 1222–1252.
- [25] E. TRÉLAT AND E. ZUAZUA, *The turnpike property in finite-dimensional nonlinear optimal control*, Journal of Differential Equations, 258 (2015), pp. 81–114.
- [26] M. TUCSNAK AND G. WEISS, *Observation and control for operator semigroups*, Springer Science & Business Media, 2009.
- [27] M. ZANON, L. GRÜNE, AND M. DIEHL, *Periodic optimal control, dissipativity and MPC*, IEEE Transactions on Automatic Control, 62 (2017), pp. 2943–2949.
- [28] A. J. ZASLAVSKI, *Turnpike theory of continuous-time linear optimal control problems*, Springer, 2016.