SENSITIVITY ANALYSIS OF OPTIMAL CONTROL FOR A CLASS OF PARABOLIC PDES MOTIVATED BY MODEL PREDICTIVE CONTROL[*]

LARS GRÜNE [†], MANUEL SCHALLER[‡], AND ANTON SCHIELA[§]

Abstract. We analyze the sensitivity of the extremal equations that arise from the first order optimality conditions for time dependent optimization problems. More specifically, we consider parabolic PDEs with distributed or boundary control and a linear quadratic performance criterion. We prove the solution's boundedness with respect to the right-hand side of the first order optimality condition which includes initial data. If the system fulfills a particular stabilizability and detectability assumption, the bound is independent of the time horizon. As a consequence, the influence of a perturbation of the right-hand side decreases exponentially backward in time. We use this property for the construction of efficient numerical discretizations in a Model Predictive Control scheme. Moreover, a quantitative turnpike theorem in the $W([0,T])$-norm is derived.

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

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

1. Introduction. Model predictive control (MPC) is a control method in which the solution of optimal control problems on infinite or indefinitely long horizons is split up into the successive solution of optimal control problems (OCPs) on finite time horizons T. Only a first part with given length τ of this solution is implemented, where usually $\tau \ll T$. The resulting state is then set as an initial condition and the process is repeated. This means that in a numerical solution of the optimal control problem only the first part of the optimal control must be computed accurately. Motivated by this observation, we will study the effect of perturbations near the end of the optimization horizon on the initial part of the control. It will turn out that under suitable assumptions, their influence decays exponentially in time. Thus, they are indeed negligible if the horizon is long enough. In this paper, we consider linear quadratic optimal control problems, more precisely tracking type objective functionals and dynamics described by linear parabolic differential equations, e.g. a linear heat equation.

An in depth introduction to model predictive control can be found in [15]. A central assumption for the approximation of the infinite horizon solution by a solution controlled with an MPC-feedback is that the dynamics of the optimal control problem exhibit turnpike behavior, i.e. the time dependent solution of the optimal control problem remains close to a constant value for the majority of the time [13]. This value is called the turnpike of the OCP. There are several different turnpike characterizations and a wide literature thereon, cf. [2, 6, 18, 30, 31, 14, 9, 17, 16]. A particular kind of turnpike behavior is the so called exponential turnpike property, cf. [26, 25, 22, 23, 4]. The approach in this paper is inspired by the idea of establishing exponential turnpike behavior by analysis of the extremal equations as done in recent works [26, 25] in order to provide a rigorous error propagation result.

In the works [26, 25, 22, 23], the proofs of the turnpike property exploit the fact that the dynamics of the so called approaching arc and the leaving arc can be decoupled. Due to this, as depicted in Figure 1, one might conjecture that if the dynamics of an optimal control problem imply

turnpike behavior, perturbations at the end of the trajectory do not affect the initial part if the time horizon T is large.
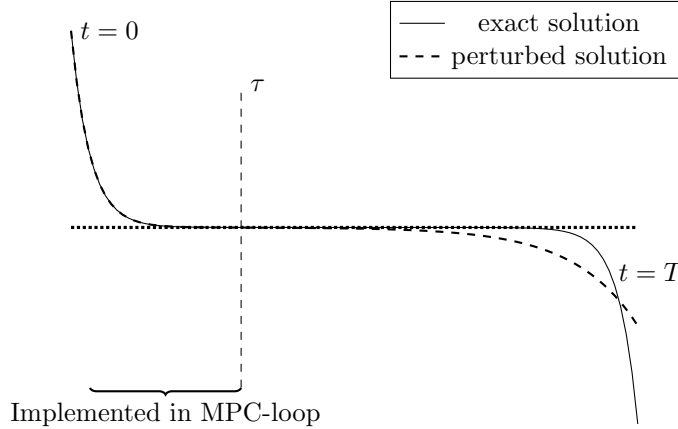

Fig. 1: Turnpike behavior of the OCP's solution and possible perturbation ε at the end of the optimization horizon.

In this paper, we will show that this property can be proven rigorously for linear-quadratic optimal control problems with dynamics satisfying a particular kind of stabilizability and detectability assumption.

We emphasize that in an optimal control context it is not a priori clear how perturbations or source terms occuring at times $t \gg 0$ influence the solution at the beginning of the time horizon. This is due to the structure of the optimality system, involving a backwards-in-time equation, i.e. the adjoint equation, which could transport perturbations from the end of the horizon to the initial part. Hence, a particular structure of the optimal control problem is needed in order to ensure that this backwards-in-time error propagation does not occur. To this end, the stabilizability and detectability assumption on the problem data is needed. As we will see, this assumption also implies the turnpike property and the same mechanism that drives the optimal solutions to the turnpike also causes the exponential damping of perturbations occuring at times $t \gg 0$.

The paper is organized as follows. In Section 2, we introduce the involved function spaces, the weak time derivative, and the resulting formulation of the PDE. Moreover, we define a linear quadratic optimal control problem with dynamics governed by a parabolic PDE and derive optimality conditions.

Section 3 contains two central results regarding the error estimation motivated by MPC. We first present an estimate in Theorem 3.1 which proves that the influence of perturbations decays exponentially backwards in time. However, it involves the norm of an operator mapping the perturbations to the absolute error of the variables. To allow for arbitrary large time horizons T we prove in Corollary 3.16 that under the stabilizability and detectabiliy assumption the operator norm occurring in Theorem 3.1 is bounded independently of T.

A numerical example is presented in Section 4. In an MPC-feedback loop, the theoretical results established in Section 3 motivate the use of temporal grids for the computation of the

optimal controls that are fine at the beginning and get coarser exponentially towards the end time. Consequently, the MPC-feedbacks can be computed with higher accuracy for a given computational cost, as more grid points are used in the area of interest (the interval $[0, \tau]$) or we can speed up the computation for a given desired accuracy. As a result, the cost functional value of the resulting MPC closed loop trajectory can be reduced by several orders of magnitude. Moreover, for unstable PDEs with reaction terms, the required number of grid points to achieve stabilization by MPC is significantly lower when using an exponential grid in contrast to a conventional uniform grid.

Eventually, we present an exponential turnpike theorem in Theorem 5.2 similar to the result presented in [25], where we, however, also conclude a turnpike result for the space and time derivatives.

2. Setting and preliminaries. In this section, we will briefly introduce the involved spaces and formulations for the linear quadratic optimal control problem of interest. We will recall some functional analytic results from the literature which are important in our context. We refer to [10], [20] and [32, Section 23.2ff] for an in-depth treatment of these topics.

2.1. Vector valued functions and generalized time derivatives. Assume $(V, \| \cdot \|_V)$ is a separable and reflexive Banach space and let $[0, T]$ be a bounded proper interval on $\mathbb{R}$. By $L_p(0, T; V)$ we denote the space of Bochner integrable functions $f : [0, T] \to V$ endowed with the norm

$$\|f\|_{L_p(0,T;V)} := \left(\int_0^T \|f(t)\|_V^p \, dt \right)^{\frac{1}{p}}.$$

If V is reflexive, then $L_p(0, T, V)$ is, and if $p < \infty$, it holds that

$$L_p(0, T; V)^* \cong L_{p'}(0, T; V^*), \qquad \text{where } \frac{1}{p} + \frac{1}{p'} = 1.$$

We will assume, that $V \hookrightarrow L_2(\Omega)$ continuously and densely. Together with V^*, the dual space of V, these spaces then form a so called Gelfand triple (or evolution triple) $V \hookrightarrow L_2(\Omega) \hookrightarrow V^*$. For a function $v \in L_p(0, T; V)$ we denote the time derivative in the sense of V-valued distributions by v' or $\frac{d}{dt} v$. By $W([0, T])$ we mean the space of functions in $L_2(0, T; V)$ with weak derivative in $L_2(0, T; V^*)$, i.e. $W([0, T]) := \{v \in L_2(0, T; V) \mid v' \in L_2(0, T; V^*)\}$. A well-known result is that $W([0, T]) \hookrightarrow C([0, T], L_2(\Omega))$ continuously (cf. [32, Proposition 23.23]). Note that for functions $v, w \in W([0, T])$ integration by parts is allowed and

$$\langle v(s), w(s) \rangle_{L_2(\Omega)} + \int_s^t v'(\tau)(w(\tau)) \, d\tau = \langle v(t), w(t) \rangle_{L_2(\Omega)} - \int_s^t w'(\tau)(v(\tau)) \, d\tau.$$

Again, we refer to [10] and [32] for these results. For brevity of notation, we will denote $L_2(0, T; V)$ by $L_2(V)$ and $L_2(0, T; L_2(\Omega))$ by $L_2(L_2(\Omega))$:

To formulate parabolic PDEs as abstract operator equations, we define the time differential operator

$$D : W([0, T]) \to L_2(V)^* \times L_2(\Omega)^*$$

$$(Dy)(v, v_0) := \int_0^T y'(t)(v(t)) \, dt + \langle y(0), v_0 \rangle_{L_2(\Omega)} \qquad \forall (v, v_0) \in L_2(V) \times L_2(\Omega).$$

With this operator, the abstract evolution equation

$$y' = f \quad \text{in} \quad L_2(V^\star)$$
$$y(0) = y_0 \quad \text{in} \quad L_2(\Omega)$$

where $f \in L_2(V^\star)$ and $y_0 \in L_2(\Omega)$ can be formulated via

$$(Dy)(v, v_0) = f(v) + \langle y_0, v_0 \rangle \qquad \forall (v, v_0) \in L_2(V) \times L_2(\Omega).$$

The corresponding adjoint operator is given by

$$D^* : L_2(V^*) \times L_2(\Omega) \to W([0, T])^*$$

$$(D^*(v, v_0))(y) = \int_0^T y'(t)(v(t)) \, dt + \langle y(0), v_0 \rangle_{L_2(\Omega)} \qquad \forall y \in W([0, T]).$$

As shown in [24, Section 3.2], for $\lambda, y \in W([0, T])$, the adjoint operator corresponds to a backwards-in-time differential operator $(D_-\lambda)(y, y(T)) := -\int_0^T \lambda'(t)(y(t)) \, dt + \langle \lambda(T), y(T) \rangle_{L_2(\Omega)}$ via the identity

$$(D^*(\lambda, \lambda(0)))(y) = (D_-\lambda)(y, y(T)) \qquad \forall \lambda, y \in W([0, T])$$

2.2. Parabolic PDEs. After having defined the time derivative, we introduce a linear and bounded operator

$$\Lambda : L_2(V) \to L_2(V^*)$$

which could represent, e.g., a Laplacian in weak form, i.e., $(\Lambda y)\lambda = \int_0^T \int_\Omega \nabla y \cdot \nabla \lambda \, d\omega dt$.

An operator corresponding to a parabolic PDE in weak form can then be defined by $A : W([0, T]) \to L_2(V^*) \times L_2(\Omega)^*$, where

$$(2.1) \qquad (Ay)(v, v_0) := (Dy)(v, v_0) + (\Lambda y)(v) = \int_0^T y'(t)(v(t)) \, dt + \langle y(0), v_0 \rangle_{L_2(\Omega)} + (\Lambda y)(v),$$

for $(v, v_0) \in L_2(V) \times L_2(\Omega)$, with adjoint operator $A^* : L_2(V) \times L_2(\Omega) \to W([0, T])^*$.

A possible problem formulation could then be given by

$$(2.2) \qquad (Ay)(v, v_0) = \int_0^T \int_\Omega f(t, \omega) v(t, \omega) \, d\omega dt + \langle y_0, v_0 \rangle_{L_2(\Omega)} \qquad \forall (v, v_0) \in L_2(V) \times L_2(\Omega),$$

where $f \in L_2(V^*)$ is a source term and $y_0 \in L_2(\Omega)$ the initial datum.

The formulation (2.2) represents a so called variational approach which is treated extensively in [8, 20, 29, 32] and is often used in the context of Galerkin methods. Another classical approach is the semigroup theory, cf. [3, 28, 21, 7]. Both, the variational and the semigroup approach are presented in [5, Chapter XVII f.]. Both approaches offer a powerful framework for the analysis

of linear and nonlinear PDEs and both lead, for initial values $y_0 \in L_2(\Omega)$ and inhomogenities $f \in L_2(L_2(\Omega))$ to solutions in $C(L_2(\Omega))$. In this work, we chose the variational approach, as less regular inhomogenities $f \in L_2(V^\star)$ are naturally included, and lead, together with initial value $y_0 \in L_2(\Omega)$ to solutions $y \in W([0,T])$. This allows us to deduce estimates in the strongest possible norm in a straight-forward fashion. For non-parabolic equations, where this regularity property of the state, i.e. $y \in L_2(V)$ does not hold, the semigroup approach can be considered and we refer to section 6 for an outlook to future work.

Concerning the variational approach, solvability of problems of above type (2.2) was discussed e.g. in [24] and in the above cited literature. Under a (weak) ellipticity assumption on Λ it can be shown that A and A^* are continuously invertible [24, Theorem 3.4]. The following proposition is also a central result of the aforementioned reference and characterizes the regularity of solutions of adjoint equations with particular right-hand sides.

PROPOSITION 2.1. *[24, Proposition 3.8]*
For a given $(\lambda, \lambda_0) \in L_2(V) \times L_2(\Omega)$, the following assertions are equivalent:
1. *There exist $(l, l_0, l_T) \in L_2(V^*) \times L_2(\Omega) \times L_2(\Omega)$ such that for all $w \in W([0,T])$ it holds that*

$$(A^*(\lambda, \lambda_0))(w) = \int_0^T l(t)(w(t)) \, dt + \langle l_0, w(0) \rangle_{L_2(\Omega)} + \langle l_T, w(T) \rangle_{L_2(\Omega)}.$$

2. *$\lambda \in W([0,T])$.*
If one, and hence both of these assertions hold, then $\lambda(T) = l_T$ and $\lambda_0 - \lambda(0) = l_0$.

Therefore, if we restrict the range of the adjoint operator to $L_2(V^*) \times L_2(\Omega)$, we obtain a Lagrange multiplier $\lambda \in W([0,T])$ with a prescribed end-time value. As a consequence, the rule of integration by parts holds and the adjoint equation can be interpreted as a backwards-in-time equation (see [24, Section 3.2]).

2.3. Optimization problems with parabolic PDEs. In this section, we will move to optimization problems governed by parabolic PDEs. We recall the Gelfand or evolution triple $V \hookrightarrow L_2(\Omega) \hookrightarrow V^*$. The variable's spaces are given by
- $W([0,T])$ for the state,
- $L_2(0,T;U)$ for the control, where U is a Hilbert space. Possible choices could be $L_2(\Omega)$ for distributed control or $L_2(\partial\Omega)$ in the case of boundary control.
- $L_2(0,T;Y)$ for the output space, where Y is a Hilbert space. As for the control, one could choose $L_2(\Omega)$ for distributed observation or $L_2(\partial\Omega)$ for boundary observation.
- $L_2(V)$ for the test functions of the dynamics,
- $L_2(\Omega)$ for the test functions of the initial value.

For the definition of the objective function, we use a linear bounded state observation operator

$$C : L_2(V) \to L_2(Y)$$

and a linear, bounded and $L_2(U)$-elliptic operator for the weighting of the control

$$R : L_2(U) \to L_2(U).$$

For the definition of the dynamics, we assume $A : W([0,T]) \to L_2(V) \times L_2(\Omega)$ of the form $A = D + \Lambda$ with D and Λ as introduced in the previous section, i.e. for test functions $v, v_0 \in L_2(V) \times L_2(\Omega)$

we have

$$(Ay)(v, v_0) = (Dy)(v, v_0) + (\Lambda y)(v) = \int_0^T y'(t)(v(t)) \, dt + \langle y(0), v_0 \rangle_{L_2(\Omega)} + (\Lambda y)(v)$$

Moreover, the control is incorporated by a linear bounded operator

$$B : L_2(U) \to L_2(V^*)$$

For given $y_0 \in L_2(\Omega)$, the optimization problem is stated as following:

(2.3) $$\min_{(y,u)} J(y, u) \quad \text{s.t.} \quad Ay - Bu = \langle y_0, \cdot \rangle$$

where the cost functional is given by

$$J(y, u) := \frac{1}{2} \|C(y - y_d)\|^2_{L_2(Y)} + \|R(u - u_d)\|^2_{L_2(U)}$$

for $y_d \in V$ and $u_d \in U$. We note that source terms in the PDE could be incorporated straightfor-wardly. However, in order to simplify the presentation, we only consider the case without source terms.

A central tool for the analysis in this work is the characterization of minimizers via the first order necessary optimality conditions. Due to convexity of the problem, these conditions are also sufficient. The following proposition states the optimality conditions at a minimizer.

PROPOSITION 2.2. *Let* $(y, u) \in W([0, T]) \times L_2(U)$ *be a minimizer of problem* (2.3). *Then if* A *is closed and surjective, there exists* $(\lambda, \lambda_0) \in L_2(V) \times L_2(\Omega)$, *such that*

$$\langle Cy, Cv \rangle_{L_2(Y)} + (A^*(\lambda, \lambda_0))(v) = \langle Cy_d, Cv \rangle_{L_2(Y)} \quad \forall v \in W([0, T])$$
$$\langle Ru, Rw \rangle_{L_2(U)} - (B^*\lambda)(w) = \langle Ru_d, Rw \rangle_{L_2(U)} \quad \forall w \in L_2(U)$$
$$(Ay)(p, p_0) - (Bu)(p) = \langle y_0, p_0 \rangle_{L_2(\Omega)} \quad \forall (p, p_0) \in L_2(V) \times L_2(\Omega)$$

Proof. [24, Theorem 1.1, Remark 1.2] □

Writing this as an equation in the respective dual spaces, we get

$$\begin{pmatrix} C^*M_yC & 0 & A^* \\ 0 & R^*M_uR & -B^* \\ A & -B & 0 \end{pmatrix} \begin{pmatrix} y \\ u \\ (\lambda, \lambda_0) \end{pmatrix} = \begin{pmatrix} C^*M_yCy_d \\ R^*M_uRu_d \\ \langle y_0, \cdot \rangle \end{pmatrix},$$

where M_y, M_u are the Riesz isomorphisms from $L_2(Y)$ to $L_2(Y)^*$ and $L_2(U)$ to $L_2(U)^*$ respectively. Defining

(2.4) $$Q := R^*M_uR,$$

we can eliminate the control via $u = Q^{-1}B^*\lambda + u_d$ and obtain

(2.5) $$\begin{pmatrix} C^*M_yC & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} y \\ (\lambda, \lambda_0) \end{pmatrix} = \begin{pmatrix} 0 \\ \langle y_0, \cdot \rangle \end{pmatrix} + \begin{pmatrix} C^*M_yCy_d \\ Bu_d \end{pmatrix}.$$

In the following, we will refer to (2.5) as the extremal equations or the extremal system. First, we present a regularity result for the adjoint state. The first line of (2.5) corresponding to the adjoint equation maps into $W([0,T])^*$. As a consequence, we obtain a Lagrange multiplier $(\delta\lambda, \delta\lambda_0) \in L_2(V) \times L_2(\Omega)$. If we restrict ourselves to more regular right-hand sides, we obtain an improved regularity of the Lagrange multiplier, as the following lemma shows.

LEMMA 2.3. *Let* $(y, (\lambda, \lambda_0)) \in (W([0,T]) \times (L_2(V) \times L_2(\Omega)))$ *solve*

$$(C^* M_y Cy + (A^*(\lambda, \lambda_0)))w = l_1 w + \langle \lambda_T, w(T) \rangle \quad \forall w \in W([0,T])$$

where $l_1 \in L_2(V^*)$ *and* $\lambda_T \in L_2(\Omega)$. *Then* $\lambda \in W([0,T])$ *and* $\lambda(0) = \lambda_0$.

Proof. Follows from Proposition 2.1 with $l(t) = l_1(t) + C^* M_y Cy(t)$ for a.e. $t \in [0,T]$, $l_0 = 0$ and $l_T = \lambda_T$. □

For brevity of notation, we define a linear operator corresponding to the extremal equations by

$$M : W([0,T])^2 \to (L_2(V^*) \times L_2(\Omega))^2$$

(2.6) $$M := \begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} = \begin{pmatrix} C^* M_y C & (D+\Lambda)^* \\ (D+\Lambda) & -BQ^{-1}B^* \end{pmatrix}$$

Note, that we only admit right-hand sides in $L_2(V^*) \times L_2(\Omega)$ for the adjoint equation and hence obtain a more regular Lagrange multiplier. This allows us to omit the second argument of the Lagrange multiplier (λ, λ_0), as $\lambda(0) = \lambda_0$. Additionally, the improved regularity $\lambda \in W([0,T])$ enables us to use integration by parts not only for the state, but also for the adjoint state, which will be crucial in the remainder.

3. Error estimation motivated by model predictive control. In this section, we present the main result of this paper. We aim to quantify, how perturbations of the right-hand side of the extremal equations influence the solution. The motivation is that these perturbations could stem from discretization errors, i.e. they may resemble the residual of the discretization scheme which corresponds to the meshsize. We show that perturbations that increase exponentially in time only influence the initial part of the solution by an amount that corresponds to the size of the perturbation at the initial time. This follows from two results we shall present in the following, namely Theorem 3.1 and Corollary 3.16. This result motivates the coarsening of grids in time in an MPC-context, as it suggests that the interesting part of the control, i.e. the control on an initial interval which actually is implemented in an MPC-controller, is only affected negligibly by larger discretization errors towards the end of the time horizon. Again, we want to stress that this is not clear a priori, as the backwards-in-time adjoint equation could propagate perturbations from close to the end time T to the initial part. In the following we will denote the perturbation of the adjoint equation by $\varepsilon_1 \in L_2(V^*)$ and the perturbation of the state equation by $\varepsilon_2 \in L_2(V^*)$.

3.1. An abstract error estimation result. First, we present a preliminary result. For this, we will refer to the solution (y, λ) of (2.5) as the *exact solution*. We will now assume that we have a second pair of variables $(\tilde{y}, \tilde{\lambda}) \in W([0,T])^2$ solving the perturbed system

$$\begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \tilde{y} \\ \tilde{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ \langle y_0, \cdot \rangle \end{pmatrix} + \begin{pmatrix} C^* M_y Cy_d \\ Bu_d \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

for $(\varepsilon_1, \varepsilon_2) \in L_2(V^*)^2$. This solution will be referred to as the *perturbed solution*. The terms ε_1 and ε_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 $(\tilde{y}, \tilde{\lambda})$ and (y, λ), which represents the absolute error. It follows by linearity that the difference between exact and perturbed solution $\begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} := \begin{pmatrix} \tilde{y} - y \\ \tilde{\lambda} - \lambda \end{pmatrix}$ fulfills the system of equations

$$(3.1) \qquad \begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

The main question in the remaining part of this paper is now the following: How does the behaviour of the perturbations ε_1 and ε_2 over time influence the behaviour of the error δy and $\delta \lambda$? The first central result reads as follows:

THEOREM 3.1. *Assume* $(\delta y, \delta \lambda) \in W([0,T])$ *solves* (3.1) *for given* $\varepsilon_1, \varepsilon_2 \in L_2(V^*)$ *and choose* $\mu > 0$ *and* $\rho > 0$ *such that*

$$0 \le \mu < \frac{1}{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}$$

and

$$\|e^{-\mu t}\varepsilon_1(t)\|_{L_2(V^*)} + \|e^{-\mu t}\varepsilon_2(t)\|_{L_2(V^*)} \le \rho.$$

Then it holds that

$$\|e^{-\mu t}\delta y\|_{W([0,T])} + \|e^{-\mu t}\delta \lambda\|_{W([0,T])} \le \rho \frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1 - \beta},$$

$$\|e^{-\mu t}\delta u\|_{L_2(U)} \le \rho \|R^{-1}B^*\|_{L_2(V^*) \to L_2(U)} \frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1 - \beta},$$

where $\beta := \mu \|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2} < 1$ *and* $\delta u = R^{-1}B^* \delta \lambda$.

Proof. First, we introduce scaled variables $\widetilde{\delta y} := e^{-\mu t}\delta y \in W([0,T])$ and $\widetilde{\delta \lambda} := e^{-\mu t}\delta \lambda \in W([0,T])$ and compute, for $(v, v_0) \in L_2(V) \times L_2(\Omega)$ and scaled test function $\tilde{v} := e^{\mu t}v$:

$$(D\delta y)(v, v_0) = (D(e^{\mu t}\widetilde{\delta y}))(v, v_0) = \int_0^T \frac{d}{dt}(e^{\mu t}\widetilde{\delta y})(v)\, dt + \langle \widetilde{\delta y}(0), v_0 \rangle$$

$$= \mu \int_0^T \tilde{v}\widetilde{\delta y}\, dt + \int_0^T \frac{d}{dt}\widetilde{\delta y}(\tilde{v})\, dt + \langle \widetilde{\delta y}(0), v_0 \rangle$$

$$= \left((\mu I + D)\widetilde{\delta y} \right)(\tilde{v}, v_0)$$

and as $\delta\lambda \in W([0,T])$, $\delta\lambda(0) = \delta\lambda_0$ we proceed for the adjoint

$$(D^*(\delta\lambda, \delta\lambda(0)))v = (Dv)(\delta\lambda, \delta\lambda(0))$$

$$= \int_0^T \frac{d}{dt}(\delta\lambda)\, dt + \langle v(0), \delta\lambda(0)\rangle = -\int_0^T \frac{d}{dt}\delta\lambda(v)\, dt + \langle v(T), \delta\lambda(T)\rangle$$

$$= -\int_0^T \frac{d}{dt}(e^{\mu t}\widetilde{\delta\lambda})(v)\, dt + \langle \tilde{v}(T), \widetilde{\delta\lambda}(T)\rangle$$

$$= -\mu \int_0^T \tilde{v}\widetilde{\delta\lambda}\, dt - \int_0^T \frac{d}{dt}\widetilde{\delta\lambda}(\tilde{v})\, dt + \langle \tilde{v}(T), \widetilde{\delta\lambda}(T)\rangle$$

$$= -\mu \int_0^T \tilde{v}\widetilde{\delta\lambda}\, dt + \int_0^T \frac{d}{dt}\tilde{v}\widetilde{\delta\lambda}\, dt + \langle \tilde{v}(0), \widetilde{\delta\lambda}(0)\rangle$$

$$= \left((-\mu I + D)\tilde{v}\right)(\widetilde{\delta\lambda}, \widetilde{\delta\lambda}(0)) = \left((-\mu I + D)^*(\widetilde{\delta\lambda}, \widetilde{\delta\lambda}(0))\right)(\tilde{v}).$$

Defining $\tilde{\varepsilon}_1 := e^{-\mu t}\varepsilon_1$, $\tilde{\varepsilon}_2 := e^{-\mu t}\varepsilon_2$, we obtain an equivalent representation of the extremal system

$$\begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \delta y \\ \delta\lambda \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

$$\iff \quad \left(\begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} + \mu \underbrace{\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}}_{=:P}\right) \begin{pmatrix} \widetilde{\delta y} \\ \widetilde{\delta\lambda} \end{pmatrix} = \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \end{pmatrix}$$

Introducing $\tilde{z} := \begin{pmatrix} \widetilde{\delta x} \\ \widetilde{\delta\lambda} \end{pmatrix}$, $\tilde{\varepsilon} := \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \end{pmatrix}$ we compute

$$(M + \mu P)\tilde{z} = \tilde{\varepsilon} \qquad \text{in } L_2(V^*)^2$$

(3.2)
$$(I + \mu M^{-1}P)\tilde{z} = M^{-1}\tilde{\varepsilon} \quad \text{in } W([0,T])^2$$

By a standard Neumann argument, cf. [19, Theorem 2.14], if we choose $\mu > 0$, such that $\beta := \mu\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2} < 1$, we get invertibility of $(I + \mu M^{-1}P)$ as operator from $W([0,T])^2$ to $W([0,T])^2$, as clearly $\|P\|_{W([0,T])^2 \to L_2(V^*)^2} = 1$. Moreover, using the Neumann series representation of $(I + \mu M^{-1}P)^{-1}$ yields

$$\|(I + \mu M^{-1}P)^{-1}\|_{W([0,T])^2 \to W([0,T])^2} \leq \sum_{i=0}^{\infty} \|(\mu M^{-1}P)^i\|_{W([0,T])^2 \to W([0,T])^2} \leq \sum_{i=0}^{\infty} \beta^i = \frac{1}{1-\beta}.$$

Thus, we conclude

$$\tilde{z} = \underbrace{(I - \mu M^{-1}P)^{-1}}_{W([0,T])^2 \to W([0,T])^2} \underbrace{M^{-1}}_{L_2(V^*)^2 \to W([0,T])^2} \tilde{\varepsilon}$$

which implies the estimate

$$\|\tilde{z}\|_{W([0,T])^2} \leq \frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1-\beta}\|\tilde{\varepsilon}\|_{L_2(V^*)^2} \leq \frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1-\beta}\rho,$$

where $\|\tilde{\varepsilon}\|_{L_2(V^*)^2} \leq \rho$ follows from our assumption on the bound of ε_1 and ε_2. Writing this in the original variables yields

$$\|e^{-\mu t}\delta y\|_{W([0,T])} + \|e^{-\mu t}\delta\lambda\|_{W([0,T])} \leq \frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1-\beta}\rho.$$

For the control we conclude

$$\|e^{-\mu t}\delta u\|_{L_2(U)} = \|e^{-\mu t}R^{-1}B^*\delta\lambda\|_{L_2(U)} = \|R^{-1}B^*e^{-\mu t}\delta\lambda\|_{L_2(U)}$$
$$\leq \|R^{-1}B^*\|_{L_2(V)\to L_2(U)}\frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1-\beta}\rho,$$

where we used that R and B do not involve time derivatives. □

COROLLARY 3.2. *Let the assumptions of Theorem 3.1 hold. Then there exists $K > 0$ independent of T such that*

$$\|e^{-\mu t}\delta y\|_{C(0,T;L_2(\Omega)))} + \|e^{-\mu t}\delta\lambda\|_{C(0,T;L_2(\Omega))} \leq K\rho\frac{\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2}}{1-\beta},$$

Proof. This follows by $W([0,T]) \hookrightarrow C(0,T;L_2(\Omega))$. For the fact, that the embedding constant K is independent of T, see e.g. the proof of [10, Satz 1.17]. □

The parameter of exponential decay μ and the right-hand side of the estimates depend on the operator norm $\|M^{-1}\|_{(L_2(V^*))^2 \to W([0,T])^2}$. To obtain estimates independent of the end time T, we will show in the next section, that under a stabilizability and detectability assumption this norm can be bounded independently of T.

Remark 3.3. We observe the dependence of the right-hand side of the estimate for the control on $\|B^*\|_{L_2(V)\to L_2(U^*)}$. It is therefore desirable to bound this term independent on the end time T, to obtain estimates which do not deteriorate for large T. Two examples of operators $L : L_2(X_1) \to L_2(X_2)$ on Hilbert spaces X_1, X_2 for which this is the case are

- $Ly = \int\limits_0^T \underline{L}y\, dt$, where $\underline{L} : X_1 \to X_2$, as then $\|L\| \leq \|\underline{L}\|$,
- $Ly = \int\limits_0^T \underline{L}(t)y\, dt$, where $\underline{L} : [0,T] \to L(X_1, X_2)$ is a continuous mapping and
  $\sup\limits_{t\in[0,\infty[}\|\underline{L}(t)\|_{L(X_1,X_2)} < \infty.$

3.2. Boundedness of the extremal equations' solution operator. In this section, we will deduce a bound on the extremal equations' solution operator norm $\|M^{-1}\|_{(L_2(V^*)\times L_2(\Omega))^2 \to W([0,T])^2}$. We include initial data in the domain of M^{-1}, i.e. the right-hand side, which allows us to deduce a turnpike result in the last section. The operator norm occuring in Theorem 3.1 can easily be

estimated via $\|M^{-1}\|_{L_2(V^*)^2 \to W([0,T])^2} \leq \|M^{-1}\|_{(L_2(V^*) \times L_2(\Omega))^2 \to W([0,T])^2}$. First, we rewrite the extremal equations for an arbitrary right-hand side $l_1, l_2 \in L_2(V^*)$ and $y_0, \lambda_T \in L_2(\Omega)$:

$$(3.3) \qquad \begin{pmatrix} C^* M_y C & A^* \\ A & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} + \begin{pmatrix} \langle \lambda_T, \cdot \rangle \\ \langle y_0, \cdot \rangle \end{pmatrix}$$

We will now introduce a particular notion of stability, characterized by an ellipticity condition.

DEFINITION 3.4. *An operator $S : L_2(V) \to L_2(V^*)$ is called $L_2(V)$-elliptic, if there exists $\alpha > 0$ independent of T such that*

$$(3.4) \qquad (Sv)(v) \geq \alpha \|v\|_{L_2(V)}^2$$

Remark 3.5.
- If S fulfills the $L_2(V)$-ellipticity condition (3.4) for $\alpha > 0$, it can be shown by a simple scaling argument, that the solution of

$$Dv + Sv = \langle v_0, \cdot \rangle$$

  fulfills $\|e^{\mu t} v\|_{L_2(V)} \leq \frac{1}{\sqrt{\alpha - \mu}} \|v_0\|_{L_2(\Omega)}$ for $\mu < \alpha$ and also $\|v(t)\|_{L_2(\Omega)} \leq e^{-\alpha t} \|v(0)\|_{L_2(\Omega)}$ for $t \geq 0$, i.e., we obtain immediate energy dissipation. If $S = (\Lambda + BK)$ for a control operator B and feedback operator K, the latter estimate clearly excludes wave-like equations with finite speed of propagation and control or observation not on the whole domain. The condition may also not be satisfied as the operator might be only elliptic in a weaker space, i.e. the space of the initial condition. We will further discuss this issue in section 6.
- It is clear, that S is $L_2(V)$-elliptic if and only if S^* is.
- An example of a $H_0^1(\Omega)$-elliptic operator is the Laplacian in weak form

$$(Sv)(v) = \int_0^T \int_\Omega \nabla y \cdot \kappa(t, \omega) \nabla v \, d\omega dt,$$

  where $\kappa(t, \omega)$ is a uniformly bounded measurable function from $[0, T]$ into the set of real matrices, satisfying an uniform ellipticity condition $v \cdot \kappa(t, \omega) v \geq c|v|^2$.

The exponential estimates in the previous remark motivate the notion of the following definition, namely V-exponential stabilizabilty.

DEFINITION 3.6.
- *A pair of operators (Λ, B), where $\Lambda : L_2(V) \to L_2(V^*)$ and $B : L_2(U) \to B : L_2(V^*)$ is called V-exponentially stabilizable, if there exists a feedback operator $K_B : L_2(V) \to L_2(U)$, such that $(\Lambda + BK_B)$ is $L_2(V)$-elliptic, i.e. fulfills condition (3.4).*
- *A pair of operators (Λ, C) where $\Lambda : L_2(V) \to L_2(V^*)$ and $C : L_2(V) \to L_2(Y)$ is called V-exponentially detectable, if (Λ^*, C^*) is V-exponentially detectable, i.e. there exists $K_C : L_2(Y) \to L_2(V^*)$ such that $(\Lambda^* + C^* K_C^*) = (\Lambda + K_C C)^*$ is V-exponentially stable.*

To illustrate this property, we present two examples that fulfill Definition 3.6, which are given by an unstable heat equation with distributed and Neumann boundary control, respectively.

Example 3.7. Let $\Omega \subset \mathbb{R}^d$ be a $C^{0,1}$-domain, where $d = 1, 2$ or 3, $V = \{v \in H^1(\Omega) \,|\, \partial_\nu v = 0\}$. For the control we consider $U = L_2(\Omega_c)$ where $\Omega_c \subseteq \Omega$ is of measure larger than zero. For $y, v \in$

$L_2(V)$ let $(\Lambda y)v := \int\limits_0^T \int\limits_\Omega \nabla y \cdot \kappa(t,\omega)\nabla v - \gamma yv\, d\omega dt$ for $\gamma > 0$, where $\kappa(t,\omega)$ is a uniformly bounded measurable function from $[0,T]$ into the set of real matrices, satisfying the uniform ellipticity condition $v \cdot \kappa(t,\omega)v \geq c|v|^2$. For arbitrary small $\gamma > 0$, the solutions of $(D + \Lambda)y = y_0$ would be unstable. Firstly, we observe that

$$(\Lambda y)(y) + (\gamma + c)\|y\|^2_{L_2(L_2(\Omega))} \geq c\|y\|_{L_2(H_0^1(\Omega))},$$

therefore by [24, Theorem 3.4, Corollary 3.5] $A := D + \Lambda$ and A^* are continuously invertible. To stabilize the system V-exponentially, we consider a control operator $(Bu)v := \int\limits_0^T \int\limits_{\Omega_c} uv\, d\omega dt$ and observation on the whole domain, i.e. C is the identity on $L_2(V)$ and $Y = V$. Defining the feedback-operator $K_B : L_2(V) \to L_2(L_2(\Omega_c))$ by $K_By(t) := -Ky(t)_{|\Omega_c}$ for $K > 0$ and a.e. t, we conclude

$$(\Lambda + BK_B)(y)(y) \geq \int\limits_0^T \int\limits_\Omega c\|\nabla y\|^2 - \gamma y^2 + \chi_{\Omega_c} K y^2\, d\omega dt \geq C(\gamma, c, K, \Omega)\|y\|^2_{L_2(V)}.$$

For given c, Ω, Ω_c, positivity of $C(\gamma, c, K, \Omega)$ can be assured if $\gamma > 0$ is small enough. This follows by the generalized Poincaré inequality, cf. [27, Lemma 2.5]. If $\Omega_c = \Omega$, we note that K_B defined above is V-exponentially stabilizing for every $\gamma > 0$. Thus, for K_C we can choose γI where I is the identity on $L_2(V)$, which yields the $L_2(V)$-ellipticity of $(\Lambda + K_C C)$ for all $\gamma > 0$.

Example 3.8. A similar result holds, if we replace the distributed control in Example 3.7 by Neumann boundary control of the form $\partial_\nu y = u$ on $\Gamma = \partial\Omega$. Defining $V = H^1(\Omega)$ yields the bounded control operator $(Bu)(v) = \int\limits_0^T \int\limits_{\Gamma_c} u\,\mathrm{tr}(v)\, dsdt$ where $\mathrm{tr} : H^1(\Omega) \to L_2(\Omega)$ is the trace operator. Therefore, a feedback K_B can be defined via $K_By = -K\,\mathrm{tr}\,y$ for $K > 0$ leading to $H^1(\Omega)$ ellipticty of $(\Lambda + BK_B)$ if γ is moderate. This follows by the generalized Friedrichs inequality cf. [27, Lemma 2.5].

Remark 3.9. In the above examples the instability constant γ has to be moderate to show V-exponential stability using the Friedrichs and Poincaré inequality if the control and observation region is not the whole domain. Classical exponential stability, i.e. $\|y(t)\| \leq Me^{-\mu t}\|y_0\|$ can be shown, however, for arbitrary γ, cf. [1, Section 3.4.1]. We recall that if an operator is V-elliptic, the solutions satisfy the property $\|y(t)\| \leq M\|y_0\|e^{-\mu t}$ with $M = 1$. Considering the case of Neumann boundary control, it was shown that the equation is exponentially stabilizable for arbitrary $\gamma > 0$ with overshoot constants $M = 1$ for the case where γ is smaller than the constant of the generalized Friedrichs or Poincaré inequality and with $M > 1$ for arbitrary large γ. For this fact, we again refer to [1, Section 3.4.1].

After having illustrated this notion of stabilizability and detectability, we will use it to show a bound on the solution operators norm. We start with a preliminary result on the time derivative operator $D : W([0,T]) \to L_2(V)^* \times L_2(\Omega)^*$, cf. [24, Section 2.3].

PROPOSITION 3.10. *For* $w \in W([0,T])$ *it holds that*

$$(Dw)(w, w(0)) = \frac{1}{2}\left(\|w(T)\|^2_{L_2(\Omega)} + \|w(0)\|^2_{L_2(\Omega)}\right)$$

Proof. Follows by $(Dw)(w, w(0)) = \int_0^T w'(t)(w(t))\, dt + \|w(0)\|^2_{L_2(\Omega)}$ and integration by parts. $\square$

We summarize the needed properties in the following assumption.

Assumption 3.11. We assume the following for the optimal control problem 2.3 and the feedback operators:

- (Λ, B) is V-exponentially stabilizable,
- (Λ, C) is V-exponentially detectable,
- all operator norms that appear in the estimates, i.e. of the problem formulation $\|\Lambda\|$, $\|B\|$, $\|R\|$, $\|C\|$ and of the stabilizing feedbacks $\|K_B\|$ and $\|K_C\|$, cf. Definition 3.6, are bounded independent of T. For examples where this is the case, we refer to Remark 3.3.

LEMMA 3.12. *Assume* $(y, \lambda) \in W([0, T])$ *solves* (3.3) *and let Assumption* 3.11 *hold. Then there are constants* $c_1, c_2 > 0$ *independent of* T, *such that*

$$\|y(T)\|^2_{L_2(\Omega)} + \|y\|^2_{L_2(V)} \le c_1 \left(\|Cy\|^2_{L_2(Y)} + \|R^{-*}B^*\lambda\|^2_{L_2(U)} + \|l_2\|^2_{L_2(V^*)} + \|y_0\|^2_{L_2(\Omega)} \right)$$

$$\|\lambda(0)\|^2_{L_2(\Omega)} + \|\lambda\|^2_{L_2(V)} \le c_2 \left(\|Cy\|^2_{L_2(Y)} + \|R^{-*}B^*\lambda\|^2_{L_2(U)} + \|l_1\|^2_{L_2(V^*)} + \|\lambda_T\|^2_{L_2(\Omega)} \right)$$

Proof. For the result on the state, we test the second equation of (3.3) with (y, y_0) and obtain

$$(Dy)(y, y_0) + (\Lambda y)(y) = (BQ^{-1}B^*\lambda)(y) + l_2(y) + \|y_0\|^2_{L_2(\Omega)}$$

Adding the term $(K_C C y)(y)$, where K_C is a stabilizing feedback for (Λ, C) from Definition 3.6 on both sides and Proposition 3.10 yields

$$\frac{1}{2} \left(\|y(T)\|^2 + \|y(0)\|^2 \right) + (\Lambda + K_C C)(y)(y) = (K_C C y)(y) + (BQ^{-1}B^*\lambda)(y) + l_2(y) + \|y_0\|^2_{L_2(\Omega)}$$

Using the V-exponential stability of $\Lambda + K_C C$ which follows by V-exponential stability of $(\Lambda + K_C C)^*$, cf. Remark 3.5, we obtain $\alpha_1 > 0$ such that

$$\frac{1}{2}\|y(T)\|^2_{L_2(\Omega)} + \alpha_1\|y\|^2_{L_2(V)} \le \|K_C\|\|Cy\|\|y\| + \|BR^{-1}\|\|R^{-*}B^*\lambda\|\|y\| + \|l_2\|\|y\| + \frac{1}{2}\|y_0\|^2$$

First, we bound the left-hand side from below by $\min\{\frac{1}{2}, \alpha_1\}(\|y(T)\|^2_{L_2(\Omega)} + \|y\|^2_{L_2(V)})$. Then, computing

$$\|K_C\|\|Cy\|\|y\| \le \frac{4}{\min\{\frac{1}{2}, \alpha_1\}}\|K_C\|^2\|Cy\|^2 + \frac{\min\{\frac{1}{2}, \alpha_1\}}{4}\|y\|^2$$

and similar for the three remaining terms involving $\|y\|$ yields

$$\frac{\min\{\frac{1}{2}, \alpha_1\}}{4}(\|y(T)\|^2_{L_2(\Omega)} + \|y\|^2_{L_2(V)}) \le$$
$$\frac{4}{\min\{\frac{1}{2}, \alpha_1\}}(\|K_C\|^2 + \|BR^{-1}\|^2 + 1)(\|Cy\|^2 + \|R^{-*}B^*\lambda\|^2 + \|l_2\|^2 + \|y_0\|^2)$$

The claim for the state follows with $c_1 := \frac{16\left(\|K_C\|^2 + \|BR^{-1}\|^2 + 1\right)}{\min\{\frac{1}{4}, \alpha_1^2\}}$. For the adjoint, we test the first equation of (3.3) with $(\lambda, \lambda(T))$ and add the term $(BK_B\lambda)(\lambda)$ on both sides. The constant in this case is given by $c_2 := \frac{16\left(\|K_B\|^2\|R^*\|^2 + \|C^*\|^2 + 1\right)}{\min\{\frac{1}{4}, \alpha_2^2\}}$. $\square$

The estimates in Lemma 3.12 still include right-hand sides depending on y and λ. This dependence can be eliminated with the following lemma.

LEMMA 3.13. *Let* (y, λ) *solve* (3.3). *Then*

$$\|Cy(s)\|_{L_2(Y)}^2 + \|R^{-*}B^*\lambda(s)\|_{L_2(U)}^2 = -\langle y_0, \lambda(0)\rangle_{L_2(\Omega)} + \langle \lambda_T, y(T)\rangle_{L_2(\Omega)} - l_2(y) + l_1(\lambda)$$

$$\leq \left(a\left(\|\lambda(0)\|^2 + \|y(T)\|^2 + \|y\|^2 + \|\lambda\|^2\right) + \frac{\|y_0\|^2 + \|\lambda_T\|^2 + \|l_2\|^2 + \|l_1\|^2}{a} \right)$$

for arbitrary $a > 0$.

Proof. Testing the first equation of (3.3) with $(y, y(T))$ and the second equation of (3.3) with $(\lambda, \lambda(0))$ and subtracting the former from the latter yields the result. The second estimate follows from the classical estimate $xy \leq ax^2 + \frac{y^2}{a}$ for all $x, y \in \mathbb{R}$ and $a > 0$. $\qquad\square$

Eventually, we obtain the following stability estimate.

PROPOSITION 3.14. *Assume* $(y, \lambda) \in W([0, T])^2$ *solves* (3.3) *and let* Assumption 3.11 *hold. Then, there exists a constant* $c_3 > 0$ *independent of* T, *such that the following stability estimate holds*

$$\|\lambda(0)\|_{L_2(\Omega)}^2 + \|y(T)\|_{L_2(\Omega)}^2 + \|y\|_{L_2(V)}^2 + \|\lambda\|_{L_2(V)}^2$$

$$\leq c_3(\|l_1\|_{L_2(V^*)}^2 + \|l_2\|_{L_2(V^*)}^2 + \|y_0\|_{L_2(\Omega)}^2 + \|\lambda_T\|_{L_2(\Omega)}^2)$$

Proof. Adding the stability estimates from Lemma 3.12 and using the estimate of Lemma 3.13 with $a = \frac{1}{(c_1+c_2)}$ yields the result for $c_3 = 4(c_1 + c_2)^2$, as $c_1, c_2 \geq 1$ and hence $(c_1 + c_2)^2 \geq c_1 + c_2$. $\square$

Eventually, the stability estimate for the variables in the $L_2(V)$-norm allows for the deduction of an estimate in the $W([0, T])$-norm.

THEOREM 3.15. *Assume* $(y, \lambda) \in W([0, T])$ *solves* (3.3) *and let* Assumption 3.11 *hold. Then, there exists a constant* $c_4 > 0$ *independent of* T *such that*

$$\|y\|_{W([0,T])}^2 + \|\lambda\|_{W([0,T])}^2 \leq c_4(\|l_1\|_{L_2(V^*)}^2 + \|l_2\|_{L_2(V^*)}^2 + \|y_0\|_{L_2(\Omega)}^2 + \|\lambda_T\|_{L_2(\Omega)}^2)$$

Proof. With (3.3) and $A = D + \Lambda$ we deduce for $v \in L_2(0, T; V)$

$$(Dy)(v, 0) = -(\Lambda y)(v) - (BQ^{-1}B^*\lambda)(v) + l_2(v)$$

implying

$$\|y'\|_{L_2(V^*)} \leq \|\Lambda\|\|y\|_{L_2(V)} + \|BQ^{-1}B^*\|\|\lambda\|_{L_2(V)} + \|l_2\|_{L_2(V^*)}.$$

Proceeding analogously for the adjoint and using the estimate on $\|y\|_{L_2(V)}^2 + \|\lambda\|_{L_2(V)}^2$ from Proposition 3.14 this yields the result. $\qquad\square$

COROLLARY 3.16. *Let* Assumption 3.11 *hold. Then, with the constant* $c_4 > 0$ *independent of* T *defined in* Theorem 3.15 *it holds that*

$$\|M^{-1}\|_{(L_2(V^*)\times L_2(\Omega))^2 \to W([0,T])^2} \leq c_4$$

As a consequence, all estimates in Theorem 3.1 *and the choice of the exponential decay parameter* μ *are independent of the horizon* T.

To conclude this theoretical part, we recall the most important result. Using the abstract scaling result Theorem 3.1 together with Corollary 3.16, we showed for systems fulfilling Assumption 3.11, and thus especially Example 3.7 and Example 3.8, that perturbations of the extremal equation's right-hand side growing exponentially in time only lead to errors in the variables that are growing exponentially in time. In particular, perturbations, that are small at the initial part of the time interval lead to disturbances of the solutions that are small at the initial part. This result is of special interest in the context of an MPC-controller, as the control on the initial part of the time interval is used as a feedback there.

4. Performance analysis of exponential grids in an MPC-feedback loop. In this section, we use the theoretical results to construct temporal grids for a numerical discretization that are tailored to the application in an MPC-feedback loop. To this end, we interpret the perturbations $\varepsilon_1(t)$, $\varepsilon_2(t)$ of the right hand side of the extremal equation as numerical errors. The coarser the meshsize, the larger the perturbations will be and vice versa. For systems of ordinary differential equations, an analysis of the relation between such perturbations and the meshsizes of an implicit Euler scheme was given in [11]. All computations were performed using the C++-library Spacy[1] and the finite element C++-library Kaskade7 [12].

We first introduce the MPC-algorithm for an optimal control problem.

Algorithm 4.1 Standard MPC Algorithm

1: Initialize $T > \tau > 0$, $k = 0$ and $y(0)$
2: **while** $k < K$ **do**
3:    Solve OCP on $[0, T]$ with initial data $y(k\tau)$ save optimal control in u^k
4:    Simulate model on $[0, T]$ with source u^k and initial data $y(k\tau)$, save in y^k.
5:    Build trajectory: $y\big|_{[k\tau,(k+1)\tau]} := y^k\big|_{[0,\tau]}$, $u\big|_{[k\tau,(k+1)\tau]} := u^k\big|_{[0,\tau]}$
6:    $k = k + 1$
7: **end while**

As stated in the introduction of this paper, the standard MPC-algorithm only implements the first part of the trajectory until a time τ, see Algorithm 4.1, Line 4. Therefore, we are especially interested in a high accuracy of the computed control on $[0, \tau]$. Since the estimate from Theorem 3.1 allows for an exponential increase of the perturbations, the specialized grids' meshsize grows exponentially in time. In the following, we will compare the cost of closed loop trajectories generated by Algorithm 4.1 for uniform grids and the specialized, exponential grids for the same number of total grid points.

As a model problem, we consider the optimization problem on a unit square $\Omega := [0, 1]^2$:

$$\min_{(y,u)} \quad \frac{1}{2}\|(y - y_d)\|^2_{L_2([0,T]\times\Omega)} + \frac{0.01}{2}\|u\|^2_{L_2([0,T]\times\Omega)}$$

subject to the parabolic PDE with parameters $d > 0$, $\mu \geq 0$, given in classical form

$$y' + d\Delta y - \mu y = u \quad \text{on} \quad]0, T] \times \Omega$$
$$y = 0 \quad \text{on} \quad]0, T] \times \partial\Omega$$
$$y(0) = 0 \quad \text{on} \quad \Omega$$

---
[1]https://spacy-dev.github.io/Spacy/

or in weak form, where the natural space in this case is $V = H_0^1(\Omega)$:

$$(Dy)(v, v_0) + d \int\limits_0^T \int\limits_\Omega \nabla y \cdot \nabla v \, d\omega dt - \mu \int\limits_0^T \int\limits_\Omega yv \, d\omega dt - \int\limits_0^T \int\limits_\Omega uv \, d\omega dt = 0$$

for all test functions $(v, v_0) \in L_2(V) \times L_2(\Omega)$. We note that for $\mu = 0$, this equation models a linear heat equation with zero initial condition which is stable due to the ellipticity of $\nabla \cdot \nabla \cdot$ in $H_0^1(\Omega)$. The parameter μ allows us to reduce the stability, or even obtain an unstable problem as soon as μ is chosen larger than d times the smallest eigenvalue of $\nabla \cdot \nabla \cdot$, given by 2π on the unit square. It can be easily seen, that the problem is V-exponentially stabilizable for arbitrary μ by choosing a feedback $Ky = (\mu + d)y$, leading to $(\Lambda + BK)(y)(y) = d\|y\|_{L_2(H_0^1(\Omega))}$

We apply Algorithm 4.1 with different grids for the solution of the OCP, see Algorithm 4.1, Line 3. The simulation of the model (Line 4) is always performed on a very fine temporal mesh on the interval $[k\tau, (k+1)\tau]$ for every iteration index k. The spatial mesh was kept constant for all timesteps. Table 1 shows the parameters for the four different problem settings for which the performance was tested. Figure 2 depicts the performance measured by the cost functional values of

	d	μ	T	τ	MPC-loops
Stable problem 1	0.01	0	3	0.5	4
Stable problem 2			2	1	
Unstable problem 1	0.1	3	2	1	8
Unstable problem 2		4	4		

Table 1: Parameters of the different problem settings.

the closed loop trajectory for the MPC-feedback computed with the exponential grid and the MPC-feedback computed with the uniform grid. For comparison, the horizontal lines in each of the four figures depict the closed loop cost for and MPC-feedback computed on a very fine grid as reference. For the stable problems we observe a monotone decrease of the closed loop cost if we increase the number of grid points. However, one can see that the particular distribution of grid points, i.e., using exponentially many points at the beginning of the time interval, leads to significant decrease of the cost. Moreover, when moving to unstable dynamics, in the case of the exponential grid, a very low cost functional value is achieved for very few gridpoints. Also, non-monotone behavior of the cost functional value with respect to the number of grid points is observed, which could be due to non-uniform refinement. Moreover, on the bottom left and right, we observe some kind of saturation effect, i.e. using even more grid points in $[0, \tau]$ does not increase the performance. In the last example on the bottom right, for 8 grid points, the feedback computed with the uniform grid led to an unstable behaviour of the resulting closed loop trajectory and very high cost functional value. Therefore, it was not depicted in this figure. The exponential grid feedback however, stabilized the system and moreover was able to achieve a very low cost functional value.

5. An exponential turnpike property. Another consequence of the bounded operator norm of M^{-1}, see Corollary 3.16, is an exponential turnpike property, which will be derived in this section and is similar to Theorem 3.1. The difference is that we will not have an $L_2(V^*)$-perturbation of the right-hand side, but initial conditions not equal to zero. Moreover, we will not estimate the distance to a perturbed solution, but to a steady state solution.
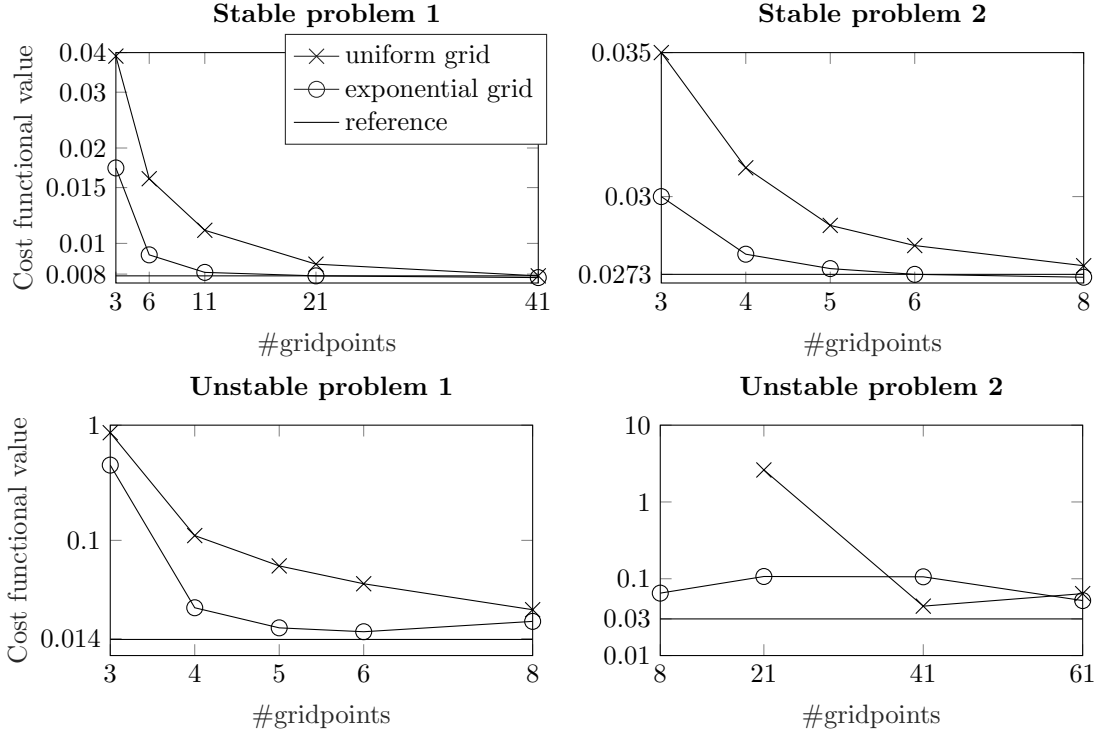
Fig. 2: Illustration of the cost functional value of the closed loop trajectory generated by MPC-feedbacks computed with different numbers of temporal gridpoints.

A similar result was proven in [26] for nonlinear finite dimensional problems. This was extended to a general Hilbert spaces setting in [25]. Moreover, the interested reader is referred to [22, 23]. The novelty of our result compared to these references is that the algebraic Riccati equation is used in these references. In contrast to this, our result on the operator norm bound enables us to derive a turnpike theorem without the use of the Riccati Theory.

Note that the previous results of this paper allowed for time dependent operators in the cost function or the dynamics. To define a steady state problem corresponding to (2.3) we will now restrict ourselves to time independent operators of the type $(\Lambda y)(v) = \int_0^T (\underline{\Lambda} y)(v)\, dt$, where $\Lambda : L_2(V) \to L_2(V^*)$ and $\underline{\Lambda} : V \to V^*$. Similarly, we assume B to be derived from $\underline{B}$, C from $\underline{C}$ and R from $\underline{R}$. The first pair of variables $(y, \lambda) \in W([0, T])^2$ we consider is the solution of the extremal equations (2.5). Secondly, we introduce the solution of a steady state optimization problem, namely $(\bar{y}, \bar{\lambda}) \in V^2$ being a minimizer of

$$\min_{\bar{y}, \bar{u}} \quad \frac{1}{2} \|C(\bar{y} - y_d)\|_Y^2 + \|R(\bar{u} - u_d)\|_U^2$$

$$\text{s.t.} \quad \underline{\Lambda}\bar{y} - \underline{B}\bar{u} = 0.$$

or, equivalently, a solution of the corresponding first order conditions

$$(5.1) \qquad \begin{pmatrix} \underline{C^*C} & \underline{\Lambda}^* \\ \underline{\Lambda} & -\underline{BQ}^{-1}\underline{B}^* \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\lambda} \end{pmatrix} = \begin{pmatrix} \underline{C^*C}y_d \\ \underline{B}u_d \end{pmatrix},$$

where $\underline{Q} = R^*R$.

LEMMA 5.1. *Let (y, λ) and $(\bar{y}, \bar{\lambda})$ solve (2.5) and (5.1) respectively. Then $\begin{pmatrix} \delta y \\ \delta\lambda \end{pmatrix} := \begin{pmatrix} y - \bar{y} \\ \lambda - \bar{\lambda} \end{pmatrix}$ solves*

$$(5.2) \qquad \underbrace{\begin{pmatrix} C^*M_yC & (D+\Lambda)^* \\ (D+\Lambda) & -BQ^{-1}B^* \end{pmatrix}}_{=M} \begin{pmatrix} \delta y \\ \delta\lambda \end{pmatrix} = \begin{pmatrix} -\langle\bar{\lambda}, \cdot\rangle \\ \langle y_0 - \bar{y}, \cdot\rangle \end{pmatrix}.$$

Proof. Using (5.1), integrating both equations over $[0, T]$ yields that $(\bar{y}, \bar{\lambda})$ satisfies

$$\begin{pmatrix} C^*M_yC & \Lambda^* \\ \Lambda & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\lambda} \end{pmatrix} = \begin{pmatrix} C^*M_yCy_d \\ Bu_d \end{pmatrix}.$$

As $\bar{y}$ and $\bar{\lambda}$ are independent of the time, we obtain $D\bar{y} = D^*\bar{\lambda} = 0$ and thus

$$\begin{pmatrix} C^*M_yC & (D+\Lambda)^* \\ (D+\Lambda) & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\lambda} \end{pmatrix} = \begin{pmatrix} \langle\bar{\lambda}, \cdot\rangle \\ \langle\bar{y}, \cdot\rangle \end{pmatrix} + \begin{pmatrix} C^*M_yCy_d \\ Bu_d \end{pmatrix}.$$

Now the result follows by linearity. □

In the following, whenever we denote $\|M^{-1}\|$, we mean the norm as an operator from $(L_2(V^*) \times L_2(\Omega))^2 \to W([0, T])^2$. For the sake of brevity, we will omit the explicit notation.

THEOREM 5.2. *Assume $(\delta y, \delta\lambda) \in W([0, T])^2$ solves (5.2). Then, for $0 \leq \mu < \frac{1}{\|M^{-1}\|}$*

$$\left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta y \right\|_{W([0,T])} + \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta\lambda \right\|_{W([0,T])} \leq \frac{\|M^{-1}\|}{1-\beta} (\|y_0 - \bar{y}\|_{L_2(\Omega)} + \|\bar{\lambda}\|_{L_2(\Omega)})$$

$$\left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta u \right\|_{L_2(U)} \leq \|R^{-1}B^*\|_{L_2(V) \to L_2(U)} \frac{\|M^{-1}\|}{1-\beta} (\|y_0 - \bar{y}\|_{L_2(\Omega)} + \|\bar{\lambda}\|_{L_2(\Omega)}),$$

where $\beta := \mu\|M^{-1}\| < 1$ and $\delta u = R^{-*}B^*\delta\lambda$. If, moreover, (Λ, B) is V-exponentially stabilizable and (Λ, C) is V-exponentially detectable in the sense of Definition 3.6, then $\|M^{-1}\|$ is independent of T.

Proof. We introduce scaled variables $\widetilde{\delta y} := \frac{1}{(e^{-\mu t} + e^{-\mu(T-t)})} \delta y$ and $\widetilde{\delta\lambda} := \frac{1}{(e^{-\mu t} + e^{-\mu(T-t)})} \delta\lambda$. For

$v \in W([0,T])$ and $\tilde{v} = (e^{-\mu t} + e^{-\mu(T-t)})v$, $\tilde{v}_0 = (1 + e^{-\mu T})v_0$ we compute

$$
\begin{aligned}
(D\delta y)(v, v_0) &= (D((e^{-\mu t} + e^{-\mu(T-t)})\widetilde{\delta y}))(v, v_0) \\
&= \int_0^T \frac{d}{dt}((e^{-\mu t} + e^{-\mu(T-t)})\widetilde{\delta y})v \, dt + \langle \delta y(0), v_0 \rangle \\
&= \mu \int_0^T \widetilde{\delta y}(e^{-\mu(T-t)} - e^{-\mu t})v \, dt + \int_0^T \frac{d}{dt}\widetilde{\delta y}\tilde{v} \, dt + (1 + e^{-\mu T})\langle \widetilde{\delta y}(0), v_0 \rangle \\
&= \left((\mu F + D)\widetilde{\delta y} \right)(\tilde{v}, \tilde{v}_0),
\end{aligned}
$$

where $(Fv)w = \int_0^T \frac{(e^{-\mu(T-t)} - e^{-\mu t})}{(e^{-\mu t} + e^{-\mu(T-t)})}v(t)w(t) \, dt$. Similarly,

$$
\begin{aligned}
(D^*\delta\lambda)(v) &= (Dv)(\delta\lambda) = \int_0^T \frac{d}{dt}v(\delta\lambda) \, dt + \langle v(0), \delta\lambda(0) \rangle \\
&= -\int_0^T \frac{d}{dt}\delta\lambda(v) + \langle v(T), \delta\lambda(T) \rangle = -\int_0^T \frac{d}{dt}((e^{-\mu t} + e^{-\mu(T-t)})\widetilde{\delta\lambda})(v) \, dt + \langle v(T), \delta\lambda(T) \rangle \\
&= \int_0^T \mu \left((e^{-\mu t} - e^{-\mu(T-t)})\widetilde{\delta\lambda}v - \frac{d}{dt}\widetilde{\delta\lambda}((e^{-\mu t} + e^{-\mu(T-t)})(v)) \right) dt + \langle \tilde{v}(T), \widetilde{\delta\lambda}(T) \rangle \\
&= -\mu(F\widetilde{\delta\lambda})(\tilde{v}) - \int_0^T \frac{d}{dt}\widetilde{\delta\lambda}(\tilde{v}) \, dt + \langle \tilde{v}(T), \widetilde{\delta\lambda}(T) \rangle \\
&= -\mu(F\widetilde{\delta\lambda})(\tilde{v}) + \int_0^T \frac{d}{dt}\tilde{v}(\widetilde{\delta\lambda}) \, dt + \langle \tilde{v}(0), \widetilde{\delta\lambda}(0) \rangle \\
&= \left((-\mu F + D)\tilde{v} \right)(\widetilde{\delta\lambda}) = \left((-\mu F + D)^*\widetilde{\delta\lambda} \right)(\tilde{v}).
\end{aligned}
$$

The substitution of variables now yields for the initial resp. terminal data

$$
i_1(v_0) := \langle \delta y_0, v_0 \rangle = \langle \delta y_0, \frac{1}{1 + e^{-\mu T}}\tilde{v}_0 \rangle = \frac{1}{1 + e^{-\mu T}}\langle \delta y_0, \tilde{v}_0 \rangle,
$$

$$
i_2(v_T) := \langle \delta\lambda_T, v_T \rangle = \langle \delta\lambda_T, \frac{1}{e^{-\mu T} + 1}\tilde{v}_T \rangle = \frac{1}{1 + e^{-\mu T}}\langle \delta\lambda_T, \tilde{v}_T \rangle,
$$

where $\tilde{v}_T = (e^{-\mu T} + 1)v_T$. Using this substitution, we transform the extremal system

$$\begin{pmatrix} C^* M_y C & (D+\Lambda)^* \\ (D+\Lambda) & -BQ^{-1}B^* \end{pmatrix} \begin{pmatrix} \delta y \\ \delta \lambda \end{pmatrix} = \begin{pmatrix} i_1 \\ i_2 \end{pmatrix}$$

$$\Longleftrightarrow \quad \left(\begin{pmatrix} C^* M_y C & (D+\Lambda)^* \\ (D+\Lambda) & -BQ^{-1}B^* \end{pmatrix} - \mu \underbrace{\begin{pmatrix} 0 & F \\ -F & 0 \end{pmatrix}}_{=:P} \right) \begin{pmatrix} \widetilde{\delta x} \\ \widetilde{\delta \lambda} \end{pmatrix} = \frac{1}{1+e^{-\mu T}} \begin{pmatrix} i_1 \\ i_2 \end{pmatrix}.$$

Defining $z := \begin{pmatrix} \widetilde{\delta y} \\ \widetilde{\delta \lambda} \end{pmatrix}$, $\tilde{i} := \frac{1}{1+e^{-\mu T}} \begin{pmatrix} i_1 \\ i_2 \end{pmatrix}$, we compute

$$(M - \mu P)z = \tilde{i} \Longleftrightarrow (I - \mu M^{-1}P)z = M^{-1}\tilde{i}.$$

If we now use $\beta = \mu \|M^{-1}\| < 1$, we get invertibility of $(I - \mu M^{-1}P)$ by [19, Theorem 2.14], as $\|P\|_{W([0,T])^2 \to L_2(V^*)^2} \leq 1$ (this will be shown at the end of this proof), and using the Neumann series representation of $(I - \mu M^{-1}P)^{-1}$ yields

$$\|(I - \mu M^{-1}P)^{-1}\|_{W([0,T])^2 \to W([0,T])^2} \leq \sum_{i=0}^{\infty} \|(\mu M^{-1}P)^i\|_{W([0,T])^2 \to W([0,T])^2} \leq \sum_{i=0}^{\infty} \beta^i = \frac{1}{1-\beta},$$

$$z = (I - \mu M^{-1}P)^{-1} M^{-1}\tilde{i},$$

$$\|z\|_{X^2} \leq \frac{\|M^{-1}\|}{1-\beta} \|\tilde{i}\|_{(L_2(\Omega)^*)^2} \leq \frac{\|M^{-1}\|}{1-\beta} \|i\|_{(L_2(\Omega)^*)^2}.$$

where $\|\tilde{i}\|_{(L_2(\Omega)^*)^2} = \underbrace{\frac{1}{1+e^{-\mu T}}}_{<1} \|i\|_{(L_2(\Omega)^*)^2} \leq \|i\|_{(L_2(\Omega)^*)^2}$ was used. Writing this in the original variables yields

$$\left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta y \right\|_{W([0,T])} + \left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta \lambda \right\|_{W([0,T])} \leq \frac{\|M^{-1}\|}{1-\beta} (\|y_0 - \bar{y}\|_{L_2(\Omega)} + \|\bar{\lambda}\|_{L_2(\Omega)}).$$

Eventually, for the control we arrive at

$$\left\| \frac{1}{e^{-\mu t} + e^{-\mu(T-t)}} \delta u \right\|_{L_2(U)} = \|e^{-\mu t} R^{-*} B^* \delta p\|_{L_2(U)}$$

$$\leq \|R^{-1}B^*\|_{L_2(V) \to L_2(U)} \frac{\|M^{-1}\|}{1-\beta} (\|y_0 - \bar{y}\|_{L_2(\Omega)} + \|\bar{\lambda}\|_{L_2(\Omega)})$$

It now remains to show the assumption $\|P\|_{W([0,T])^2 \to L_2(V^*)^2} \leq 1$. This follows immediately from

$$P = \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix}$$

and

$$|(Fv)(w)| = |\int\limits_{0}^{T} \underbrace{\frac{(e^{-\mu(T-t)} - e^{-\mu t})}{(e^{-\mu t} + e^{-\mu(T-t)})}}_{|\cdot| < 1} v(t)w(t)\, dt| \le \int\limits_{0}^{T} v(t)w(t)\, dt \le \|v\|_{W([0,T])}\|w\|_V$$

for $v \in W([0,T])$ and $w \in L_2(V)$. $\qquad\square$

This result allows us to conclude a turnpike result similar to [25, Theorem 1].

COROLLARY 5.3. *Let the assumptions of Theorem 5.2 hold. Then there exists $c > 0$ such that for all $t \in [0,T]$*

$$\|y(t) - \bar{y}\|_{L_2(\Omega)} + \|\lambda(t) - \bar{\lambda}\|_{L_2(\Omega)} \le c(e^{-\mu T} + e^{-\mu(T-t)})\left(\|y_0 - \bar{y}\|_{L_2(\Omega)} + \|\bar{\lambda}\|_{L_2(\Omega)}\right)$$

If (Λ, B) is V-exponentially stabilizable and (Λ, C) is V-exponentially detectable in the sense of Definition 3.6, the constants c and μ can be chosen independent of T.

Proof. Analogous to the proof of Corollary 3.2 this follows by $W([0,T]) \hookrightarrow C(0,T; L_2(\Omega))$ and the embedding constant being independent of T. $\qquad\square$

6. Conclusion and Outlook. In this work, we considered linear quadratic optimization problems with parabolic PDEs and showed under a particular stabilizability and detectability assumption, that the norm of the extremal equation's solution operator can be bounded independently of the end time T. This allowed us to conclude two results. Firstly, we obtained that the influence of perturbations decays exponentially in time, which is of particular interest in the context of Model Predictive Control as it allows for efficient grid generation. This was illustrated by a numerical example. Secondly, the bound on the solution operator's norm enabled us to conclude a turnpike result in the $W([0,T])$-norm.

We note the similarity of Corollary 5.3 to the turnpike theorem in [25, Theorem 1], where the authors proved the same result under classical stabilizability and detectability assumptions, i.e. the closed-loop semigroups being exponentially stable. These assumptions are weaker than our assumption of $L_2(V)$-ellipticity of the closed loop operator, as $L_2(V)$-ellipticity implies $L_2(L_2(\Omega))$-ellipticity which implies $\|y(t)\|_{L_2(\Omega)} \le e^{-\mu t}\|y(0)\|_{L_2(\Omega)}$ for the solution y, i.e. classical exponential stability with overshoot constant 1. In the context of parabolic PDEs, however, the $L_2(V)$-ellipticity enabled us to conclude a turnpike result in the $W([0,T])$-norm in Theorem 5.2, i.e., an integral turnpike property on the time and space derivatives. Another advantage of using $L_2(V)$-ellipticity of the closed loop operator is that boundary control and observation can be easily incorporated using the regularity of the solutions, as $y, \lambda \in L_2(V)$ and boundary control and observation operators are bounded on this space.

The approach taken in this paper, i.e. the abstract scaling result of the error estimate Theorem 3.1 or the turnpike result Theorem 5.2 can be generalized to general, hence in particular hyperbolic, evolution equations on a Hilbert space X. In this context, it has to be shown that under classical stabilizability assumptions, i.e. that there is a feedback operator such that the closed loop is exponentially stable, that the operator norm $\|M^{-1}\|_{(L_2(X)\times L_2(\Omega))^2 \to L_2(X)^2}$ and $\|M^{-1}\|_{(L_2(X)\times L_2(\Omega))^2 \to C(X)^2}$ can be bounded independently of T. As a consequence, this would allow for a generalization of the error estimation result to hyperbolic evolution equations. Likewise, one could deduce turnpike results for general evolution equations with admissible boundary control or observation operators under stabilizability and detectability conditions, a setting for which, to

the authors best knowledge, there are no turnpike results for space dimension 2 or 3. For boundary controlled hyperbolic equations in one space dimension, turnpike results are given in [17, 16]. For parabolic equations with boundary control or observation, a turnpike result was shown in [25, Theorem 6]. The restriction of [25, Section 2.4] to parabolic equations in this case of boundary control or observation is due to the utilization of Algebraic Riccati theory, which is well established for analytic semigroups if, but not for general semigroups if the control or observation is unbounded. The abstract scaling approach given in Theorem 5.2 could overcome this difficulty. Details will be subject to further research.

Acknowledgments. The authors are grateful to Enrique Zuazua, Felix L. Schwenninger and Hannes Meinlschmidt for valuable discussions.

REFERENCES

[1] N. ALTMÜLLER, *Model Predictive Control for Partal Differential Equations*, PhD thesis, Universität Bayreuth, 2014.

[2] B. D. ANDERSON AND P. V. KOKOTOVIC, *Optimal control problems over large time intervals*, Automatica, 23 (1987), pp. 355 – 363, https://doi.org/https://doi.org/10.1016/0005-1098(87)90008-2, http://www.sciencedirect.com/science/article/pii/0005109887900082.

[3] R. F. CURTAIN AND H. ZWART, *An introduction to infinite-dimensional linear systems theory*, vol. 21, Springer Science & Business Media, 1995.

[4] 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. 1935–1957, https://doi.org/10.1137/120888934, http://dx.doi.org/10.1137/120888934.

[5] R. DAUTRAY AND J.-L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 5 Evolution Problems I*, Springer, 2000.

[6] R. DORFMAN, P. A. SAMUELSON, AND R. M. SOLOW, *Linear programming and economic analysis*, Courier Corporation, 1958.

[7] K.-J. ENGEL AND R. NAGEL, *One-Parameter Semigroups for Linear Evolution Equations*, Springer-Verlag New York, 2000.

[8] L. C. EVANS, *Partial differential equations*, American Mathematical Society, Providence, R.I., 2010.

[9] T. FAULWASSER, M. KORDA, C. N. JONES, AND D. BONVIN, *On turnpike and dissipativity properties of continuous-time optimal control problems*, CoRR, abs/1509.07315 (2015), http://arxiv.org/abs/1509.07315, https://arxiv.org/abs/1509.07315.

[10] H. GAJEWSKI, K. GRÖGER, AND K. ZACHARIAS, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, vol. 67, Akademie-Verlag Berlin, 1974, https://doi.org/10.1002/mana.19750672207, http://dx.doi.org/10.1002/mana.19750672207.

[11] C. W. GEAR, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1971.

[12] S. GÖTSCHEL, M. WEISER, AND A. SCHIELA, *Solving optimal control problems with the Kaskade 7 finite element toolbox*, 2010.

[13] L. GRÜNE, *Approximation properties of receding horizon optimal control*, Jahresbericht Deutsche Mathematiker-Vereinigung, 118 (2016), pp. 3–37.

[14] L. GRÜNE AND M. A. MÜLLER, *On the relation between strict dissipativity and turnpike properties*, Systems & Control Letters, 90 (2016), p. 45–53, https://doi.org/10.1016/j.sysconle.2016.01.003, http://dx.doi.org/10.1016/j.sysconle.2016.01.003.

[15] L. GRÜNE AND J. PANNEK, *Nonlinear Model Predictive Control: Theory and Algorithms*, Springer, 2016.

[16] M. GUGAT AND F. HANTE, *On the turnpike phenomenon for optimal boundary control problems with hyperbolic systems.* https://arxiv.org/abs/1806.04438.

[17] 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. 61–70, https://doi.org/10.1016/j.sysconle.2016.02.001, http://dx.doi.org/10.1016/j.sysconle.2016.02.001.

[18] A. HAURIE, *Optimal control on an infinite time horizon: The turnpike approach*, Journal of Mathematical Economics, 3 (1976), pp. 81 – 102, https://doi.org/https://doi.org/10.1016/0304-4068(76)90007-0, http:

//www.sciencedirect.com/science/article/pii/0304406876900070.

[19] R. Kress, V. Maz'ya, and V. Kozlov, *Linear integral equations*, vol. 17, Springer, 1989.

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

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

[22] A. Porretta and E. Zuazua, *Long time versus steady state optimal control*, SIAM Journal on Control and Optimization, 51 (2013), pp. 4242–4273, https://doi.org/10.1137/130907239, http://dx.doi.org/10.1137/130907239, https://arxiv.org/abs/http://dx.doi.org/10.1137/130907239.

[23] A. Porretta and E. Zuazua, *Remarks on long time versus steady state optimal control*, Springer IN-dAM Series, (2016), p. 67–89, https://doi.org/10.1007/978-3-319-39092-5_5, http://dx.doi.org/10.1007/978-3-319-39092-5_5.

[24] A. Schiela, *A concise proof for existence and uniqueness of solutions of linear parabolic PDEs in the context of optimal control*, System & Control Letters, 62 (2013), pp. 895–901.

[25] 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, https://doi.org/10.1137/16M1097638, https://doi.org/10.1137/16M1097638, https://arxiv.org/abs/https://doi.org/10.1137/16M1097638.

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

[27] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*, Vieweg und Teubner, 2009.

[28] M. Tucsnak and G. Weiss, *Observation and control for operator semigroups*, Springer Science & Business Media, 2009.

[29] J. Wloka, *Partial Differential Equations*, Cambridge University Press, 1987, https://doi.org/10.1017/CBO9781139171755.

[30] M. Zanon, L. Grüne, and M. Diehl, *Periodic optimal control, dissipativity and mpc*, IEEE Transactions on Automatic Control, 62 (2017), pp. 2943–2949, https://doi.org/10.1109/TAC.2016.2601881.

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

[32] E. Zeidler, *Nonlinear Functional Analysis and its Applications- II/A: Linear Monotone Operators*, Springer-Verlag New York, 1990.