



HAL
open science

Turnpike in Lipschitz-nonlinear optimal control

Carlos Esteve, Borjan Geshkovski, Dario Pighin, Enrique Zuazua

► **To cite this version:**

Carlos Esteve, Borjan Geshkovski, Dario Pighin, Enrique Zuazua. Turnpike in Lipschitz-nonlinear optimal control. 2020. hal-03018529v1

HAL Id: hal-03018529

<https://hal.science/hal-03018529v1>

Preprint submitted on 22 Nov 2020 (v1), last revised 14 Mar 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TURNPIKE IN LIPSCHITZ–NONLINEAR OPTIMAL CONTROL

CARLOS ESTEVE, BORJAN GESHKOVSKI, DARIO PIGHIN, AND ENRIQUE ZUAZUA

ABSTRACT. We present a new proof of the turnpike property for nonlinear optimal control problems, when the running target is a steady control-state pair of the underlying dynamics. Our strategy combines the construction of suboptimal quasi-turnpike trajectories via controllability, and a bootstrap argument, and does not rely on analyzing the optimality system or linearization techniques. This in turn allows us to address several optimal control problems for finite-dimensional, control-affine systems with globally Lipschitz (possibly nonsmooth) nonlinearities, without any smallness conditions on the initial data or the running target. These results are motivated by the large-layer regime of residual neural networks, commonly used in deep learning applications. We show that our methodology is applicable to controlled PDEs as well, such as the semilinear wave and heat equation with a globally Lipschitz nonlinearity, once again without any smallness assumptions.

CONTENTS

| | |
|-----------------------------------|----|
| 1. Introduction | 1 |
| 2. Finite-dimensional systems | 4 |
| 3. Infinite-dimensional systems | 10 |
| 4. Preliminary results | 13 |
| 5. Proof of Theorem 2.1 | 17 |
| 6. Proof of Theorem 3.1 | 33 |
| 7. Proof of Theorem 3.2 | 38 |
| 8. Concluding remarks and outlook | 41 |
| References | 44 |

Keywords. Optimal control; Turnpike; Nonlinear systems; Stabilization; Deep learning.

AMS Subject Classification. 34H05; 34H15; 93C15; 93C20.

1. INTRODUCTION

The *turnpike property* reflects the fact that, for suitable optimal control problems set in a sufficiently large time horizon, any optimal solution thereof remains, during most of the time, close to the optimal solution of a corresponding “static” optimal control problem. This optimal static solution is referred to as *the turnpike* – the name stems from the idea that a turnpike is the fastest route between two points which are far apart, even if it is not the most direct route. In many cases, the turnpike property is described by an exponential estimate – for instance, the optimal trajectory $y_T(t)$ is $\mathcal{O}(e^{-\mu t} + e^{-\mu(T-t)})$ -close to the optimal static solution \bar{y} , for $t \in [0, T]$ and for some $\mu > 0$.

Date: November 22, 2020.

The prevalent (but not exclusive) argument for proving exponential turnpike results relies on a thorough analysis of the optimality system provided by the Pontryagin Maximum Principle. In the context of linear quadratic optimal control problems, under appropriate controllability or stabilizability conditions, turnpike is established via properties of the optimality system characterizing the optimal controls and states through the coupling with the adjoint system, see Porretta & Zuazua [31].

In the case of nonlinear dynamics, this argument thus requires nonlinearities which are continuously differentiable. A linearization argument is used – the linear study and a fixed point argument provide nonlinear results under smallness assumptions on the initial data and the target, see Zuazua et al. [32, 36]. The smallness conditions on the initial data can be removed in some specific cases (see e.g. Pighin [29]), but to the best of our knowledge, the assumptions on the running target have not been as of yet (albeit, they may be removed under restrictive assumptions, such as strict dissipativity, uniqueness of minimizers and C^2 -regular nonlinearities – see [34]). This is due to the lack of tools for showing that the linearized optimality system corresponds to a linear-quadratic control problem satisfying the turnpike property, when the running target of the original nonlinear control problem is large.

There has been an ever-increasing need however, brought by applications in *deep learning* via *residual neural networks* (ResNets) (see [10, 11, 18]), of turnpike results for nonlinear optimal control problems without smallness conditions on the data or the running target, and for systems with globally Lipschitz-continuous but possibly nonsmooth nonlinearities.

In deep learning, one wishes to find a map which interpolates a dataset $\{\vec{x}_i, \vec{y}_i\}_{i=1}^N$ where $\vec{x}_i \in \mathbb{R}^{d_x}$ and $\vec{y}_i \in \mathbb{R}^{d_y}$ and gives accurate predictions on unknown points $\vec{x} \in \mathbb{R}^{d_x}$. Such a task may be accomplished by minimizing

$$\int_0^T \sum_{i=1}^N \|P\mathbf{x}_i(t) - \vec{y}_i\|^2 dt + \int_0^T \|u(t)\|^2 dt, \quad (1.1)$$

where $u := [w, b]^\top$ and $P : \mathbb{R}^{d_x} \rightarrow \mathbb{R}^{d_y}$ is an affine surjective map, subject to

$$\begin{cases} \dot{\mathbf{x}}_i(t) = \sigma(w(t)\mathbf{x}_i(t) + b(t)) & \text{in } (0, T) \\ \mathbf{x}_i(0) = \vec{x}_i, \end{cases} \quad (1.2)$$

with $w \in L^2(0, T; \mathbb{R}^{d_x \times d_x})$ and $b \in L^2(0, T; \mathbb{R}^{d_x})$ designating the controls, whereas $\sigma \in \text{Lip}(\mathbb{R})$ with $\sigma(0) = 0$ is a scalar nonlinear function, defined componentwise in (1.2). The most frequently used nonlinearities in practical applications are *rectifiers*: $\sigma(x) = \max\{\alpha x, x\}$ for $\alpha \in [0, 1)$, and *sigmoids*: $\sigma(x) = \tanh(x)$. The order of the nonlinearity σ and the affine map within may be permuted to obtain a driftless control-affine system

$$\begin{cases} \dot{\mathbf{x}}_i(t) = w(t)\sigma(\mathbf{x}_i(t)) + b(t) & \text{in } (0, T) \\ \mathbf{x}_i(0) = \vec{x}_i. \end{cases} \quad (1.3)$$

Combinations and variants of (1.2) and (1.3) may also be used, see e.g. [23]. Optimizing u over $N \gg 1$ different initial data establishes robustness, so that the neural networks (1.2) and (1.3) may correctly perform future predictions on unknown points.

In Figure 1, we see stabilization for the trajectories to some points $\bar{\mathbf{x}}_i \in P^{-1}(\{\vec{y}_i\})$, which are uncontrolled steady states of (1.2) and (1.3). This motivates the choice of running target as a steady control-state pair we consider in this work ((1.4)), which

would then entail bounds for (1.1) (see [11]). The practical interest of the turnpike and stabilization analysis when $T \gg 1$ presented herein is its link to the large-layer regime and approximation capacity (dual to [7]) of ResNets, which are the forward Euler discretizations of (1.2) and (1.3) (see [10]). This regime is the common setting for many deep learning applications [22]. We refer the reader to [11] for further details.

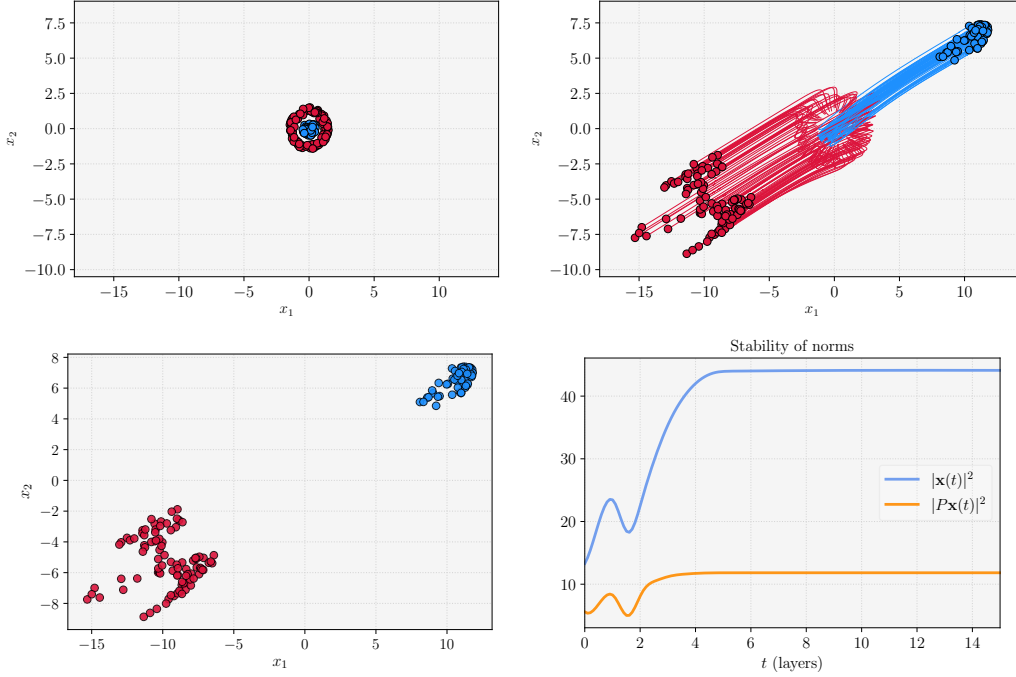
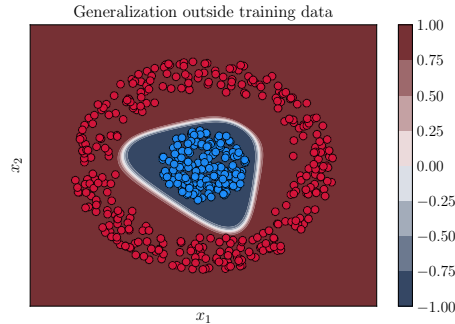


FIGURE 1. *A binary classification task in deep learning.* One aims to separate the data points $\{\vec{x}_i\}_i$ in \mathbb{R}^2 (top left) with respect to their color by using the controlled flow of (1.2) – (1.3) at time $T = 15$, here done by minimizing (1.1) ($\vec{y}_i = \pm 1$ for red/blue). We visualize the evolution of the trajectories of (1.3) (top right) and their output (bottom left). We see a stabilization property for the projections, but also the trajectories to some points $\bar{\mathbf{x}}_i \in P^{-1}(\{\vec{y}_i\})$ (bottom right). Displayed below is the inferred classifier on $[-2, 2]^2$, generalizing the shape of the dataset.



Our contributions. To answer this need, and motivated by problems as those above, in this work we provide a different perspective on the turnpike property in the context of nonlinear dynamics, and we bring forth the following contributions.

1. In Section 2, we consider optimal control problems consisting of minimizing

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \bar{y}\|^2 dt + \int_0^T \|u(t) - \bar{u}\|^2 dt \quad (1.4)$$

subject to $\dot{y} = f(y, u)$, where f is of control-affine form. Under the assumption that the running target (\bar{u}, \bar{y}) is a steady control-state pair, namely $f(\bar{y}, \bar{u}) = 0$, and that the system is controllable with an estimate on the cost (see Assumption 1), in Theorem 2.1 we prove the exponential turnpike property described above. The main novelty lies in the fact that the nonlinearity f is only assumed to be globally Lipschitz continuous, and the result comes without any smallness conditions on the initial data or the specific running target. In this case, existing results such as those presented in Trélat & Zuazua [36, 34] do not apply, as they either require smallness assumptions or uniqueness of minimizers, and C^2 -nonlinearities.

Moreover, whenever the functional to be minimized does not contain a final-time cost (such as $\phi(y(T))$ in J_T above), we can prove (see Corollary 2.1 below) that the exponential arc near the final time $t = T$ disappears, thus entailing an exponential stabilization property for the optimal state to the running target.

2. In Section 3, the finite-dimensional results are extended to analogue optimal control problems for underlying PDE dynamics. This is illustrated in Theorem 3.1, Corollary 3.1 and Theorem 3.2 in the context of the semilinear wave and heat equation with globally Lipschitz-only nonlinearity, once again under the assumption that the running target is a steady control-state pair. We make no smallness assumptions neither on it, nor on the initial data, thus covering some cases where results from [15, 29, 32, 42] are not applicable.

Notation. We denote by $\|\cdot\|$ the standard euclidean norm, and $\mathbb{N} := \{1, 2, \dots\}$. We denote by $\text{Lip}(\mathbb{R})$ (resp. $\text{Lip}_{\text{loc}}(\mathbb{R})$) the set of functions $f : \mathbb{R} \rightarrow \mathbb{R}$ which are globally (resp. locally) Lipschitz continuous.

2. FINITE-DIMENSIONAL SYSTEMS

2.1. **Setup.** Let $d \geq 1$ and $m \geq 1$. We will consider differential control systems where the state $y(t)$ lives in \mathbb{R}^d and the control input $u(t)$ in \mathbb{R}^m . Given $T > 0$, we focus on *control-affine* systems, namely canonical nonlinear systems

$$\dot{y} = f(y, u) \quad \text{in } (0, T) \quad (2.1)$$

with a nonlinearity f of the form

$$f(y, u) = f_0(y) + \sum_{j=1}^m u_j f_j(y) \quad \text{for } (y, u) \in \mathbb{R}^d \times \mathbb{R}^m, \quad (2.2)$$

where the vector fields $f_0, \dots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ are only assumed to be globally Lipschitz continuous. This formulation includes (1.3) – see Remark 3 for possible extensions to (1.2).

For any given initial datum $y^0 \in \mathbb{R}^d$ and control input $u \in L^1(0, T; \mathbb{R}^m)$, system (2.1), with f as in (2.2), admits a unique solution $y \in C^0([0, T]; \mathbb{R}^d)$ with $y(0) = y^0$. This can

be shown by means of a fixed point theorem and the Grönwall inequality applied to the integral formulation

$$y(t) = y^0 + \int_0^t f(y(s), u(s)) \, ds.$$

Given $y^0 \in \mathbb{R}^d$, we will investigate the behavior when $T \gg 1$ of global minimizers $u_T \in L^2(0, T; \mathbb{R}^m)$ to nonnegative functionals of the form

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \bar{y}\|^2 \, dt + \int_0^T \|u(t)\|^2 \, dt, \quad (2.3)$$

and of the corresponding solutions y_T to (2.1) with $y_T(0) = y^0$. Here, $\phi \in C^0(\mathbb{R}^d; \mathbb{R}_+)$ is a given final cost, while $\bar{y} \in \mathbb{R}^d$ is a given running target which we select as an *uncontrolled steady state* of the nonlinear dynamics, namely

$$f_0(\bar{y}) = 0. \quad (2.4)$$

We provide further comments on the specific choice of the running target just below, in Remark 1. Due to the coercivity of J_T and the explicit form of f in (2.2), the existence of a minimizer of J_T follows from the direct method in the calculus of variations.

Due to the presence of the state tracking term in the definition of J_T , which regulates the state over the entire time interval $[0, T]$, the well-known *turnpike property* is expected to hold: over long time horizons, the optimal control-state pair (u_T, y_T) should be "near" the optimal steady control-state pair (u_s, y_s) , namely a solution to the problem

$$\inf_{u \in \mathbb{R}^m} \|y - \bar{y}\|^2 + \|u\|^2 \quad \text{subject to } f(y, u) = 0. \quad (2.5)$$

Now note that, due to the assumption (2.4) on the running target \bar{y} , and the form of the nonlinearity f in (2.2), it can be seen that $(u_s, y_s) \equiv (0, \bar{y})$ designates the unique optimal stationary solution, namely the unique solution to (2.5).

Remark 1 (Controlled steady states). *The choice of the running target \bar{y} in (2.4) is tailored to our proof strategy and the choice of the functional J_T in (2.3). The key feature our methodology requires is that the Lagrangian $\mathcal{L}(u, y) = \|y - \bar{y}\|^2 + \|u - \bar{u}\|^2$ equals zero when evaluated at the optimal steady state. In fact, we could more generally consider the functional*

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \bar{y}\|^2 \, dt + \int_0^T \|u(t) - \bar{u}\|^2 \, dt$$

where $(\bar{u}, \bar{y}) \in \mathbb{R}^m \times \mathbb{R}^d$ is chosen so that $f(\bar{y}, \bar{u}) = 0$ (with f as in (2.2)), as discussed in the introduction. The results presented below could then readily be adapted to this case (by additionally changing (2.9) and Assumption 1 to an L^2 -bound of $u_T - \bar{u}$). We have taken $\bar{u} = 0$ for presentational simplicity.

In the context of nonlinear optimal control, such turnpike results have been shown by Trélat & Zuazua in [36] (see also [34]) for C^2 -regular nonlinearities f . This order of regularity is required due to the proof strategy, which relies on linearizing the optimality system given by the Pontryagin Maximum Principle. As a consequence, the results in [36] are also local, in the sense that smallness conditions are assumed on the initial data and target in view of applying a fixed point argument. In this work, we take a further step and obtain global results for globally Lipschitz nonlinearities.

2.2. Main results. The notion of controllability plays a key role in the context of turnpike. Hence, before proceeding, we state the following assumption.

Assumption 1 (Controllability & cost estimate). *We will assume that (2.1) is controllable in some time $T_0 > 0$, meaning that there exists some time $T_0 > 0$ such that for any $y^0, y^1 \in \mathbb{R}^d$, there exists a control $u \in L^2(0, T_0; \mathbb{R}^m)$ such that the corresponding solution $y \in C^0([0, T_0]; \mathbb{R}^d)$ to (2.1) with $y(0) = y^0$ satisfies $y(T_0) = y^1$.*

We will moreover assume that there exists an $r > 0$ and a constant $C(T_0) > 0$ such that

$$\inf_{\substack{u \\ \text{such that} \\ y(0)=y^0, y(T_0)=\bar{y}}} \|u\|_{L^2(0, T_0; \mathbb{R}^m)} \leq C(T_0) \|y^0 - \bar{y}\|, \quad (2.6)$$

and

$$\inf_{\substack{u \\ \text{such that} \\ y(0)=\bar{y}, y(T_0)=y^1}} \|u\|_{L^2(0, T_0; \mathbb{R}^m)} \leq C(T_0) \|y^1 - \bar{y}\|, \quad (2.7)$$

hold for any $y^0, y^1 \in \{x \in \mathbb{R}^d : \|x - \bar{y}\| \leq r\}$, where $\bar{y} \in \mathbb{R}^d$ is fixed as in (2.4).

We discuss the feasibility of this assumption later on, in Remark 4. Note that this is not a smallness assumption – it merely stipulates that, inside some ball centered at \bar{y} , the cost of controlling from and to \bar{y} can be estimated by means of the distance to \bar{y} .

We may now state our first main result.

Theorem 2.1 (Turnpike). *Assume that $f_0, \dots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ in (2.2), and assume that (2.1) is controllable in some time $T_0 > 0$ in the sense of Assumption 1. Let $y^0 \in \mathbb{R}^d$ be given, and let $\bar{y} \in \mathbb{R}^d$ be as in (2.4). Then there exists a time $T^* > 0$ and constants $C_1, C_2, \mu > 0$ such that for any $T \geq T^*$, any global minimizer $u_T \in L^2(0, T; \mathbb{R}^m)$ to J_T defined in (2.3) and corresponding optimal state y_T solution to (2.1) with $y_T(0) = y^0$ satisfy*

$$\|y_T(t) - \bar{y}\| \leq C_1 \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \quad (2.8)$$

for all $t \in [0, T]$, and

$$\|u_T\|_{L^2(0, T; \mathbb{R}^m)} \leq C_2. \quad (2.9)$$

We sketch the idea of the proof (which may be found in Section 5.2) in Section 2.2.1 below. The rate $\mu > 0$ appearing in (2.8) depends on the datum y^0 due to the multiplicative form of the control, but is uniform with respect to y^0 when the control is additive, namely, when f_1, \dots, f_m are nonzero constants. This is due to the form of the constant provided by Grönwall arguments (e.g. in Lemma 4.1 and Lemma 5.2).

Remark 2 (On (2.9)). *An exponential estimate for the optimal control u_T is a hallmark of turnpike results obtained by analyzing the optimality system. Therein, the optimal control can be characterized explicitly via the adjoint state, which, much like the optimal state, fulfills an exponential estimate. Since in this work we do not use the optimality system, we do not have as much information on $u_T(t)$ as we have on $y_T(t) - \bar{y}$. The latter quantity, in addition to being penalized by J_T , may be further estimated by using the system dynamics. In the context of driftless systems, we show that $u_T(t)$ too is in $\mathcal{O}(e^{-\mu t} + e^{-\mu(T-t)})$ in Corollary 2.2, by using the homogeneity of the system with respect to the control.*

Before proceeding with further remarks, let us state a couple of important corollaries of Theorem 2.1.

Firstly, when one considers an optimal control problem for J_T without a final cost for the endpoint $y(T)$, namely taking $\phi \equiv 0$ in (2.3), Theorem 2.1 can in fact be improved to an *exponential stabilization* estimate to the running target \bar{y} .

Corollary 2.1 (Stabilization). *Suppose that $\phi \equiv 0$ in J_T defined in (2.3). Under the assumptions of Theorem 2.1, there exists a time $T^* > 0$, and constants $C_1, C_2, \mu > 0$ such that for any $T \geq T^*$, any global minimizer $u_T \in L^2(0, T; \mathbb{R}^m)$ to J_T defined in (2.3) and corresponding optimal state y_T solution to (2.1) with $y_T(0) = y^0$ satisfy (2.9) as well as*

$$\|y_T(t) - \bar{y}\| \leq C_1 e^{-\mu t} \quad (2.10)$$

for all $t \in [0, T]$.

We refer to Section 5.3 for a proof. In fact, Corollary 2.1 may be proven independently of Theorem 2.1 by a simple adaptation of the proof strategy. This is illustrated in the proof of Theorem 3.2 in the context of the semilinear heat equation.

On another hand, when the underlying dynamics (2.1) are of *driftless control affine* form (namely, $f_0 \equiv 0$ in (2.2)), we can obtain an exponential decay for the optimal controls as well. Note that in this case, any $\bar{y} \in \mathbb{R}^d$ is an admissible running target for J_T , since $f(\bar{y}, 0) = 0$ for any $\bar{y} \in \mathbb{R}^d$.

Corollary 2.2 (Control decay). *Suppose that $f_0 \equiv 0$ in (2.2). Under the assumptions of Theorem 2.1, there exists a time $T^* > 0$, and constants $C, \mu > 0$ such that for any $T \geq T^*$, any global minimizer $u_T \in L^2(0, T; \mathbb{R}^m)$ to J_T defined in (2.3) and corresponding optimal state y_T solution to (2.1) with $y_T(0) = y^0$ satisfy (2.8) as well as*

$$\|u_T(t)\| \leq C \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \quad (2.11)$$

for a.e. $t \in [0, T]$.

If moreover, $\phi \equiv 0$ in J_T defined in (2.3), in addition to (2.10), there exist constants $C_1, \mu_1 > 0$ independent of T such that

$$\|u_T(t)\| \leq C_1 e^{-\mu_1 t} \quad (2.12)$$

holds for a.e. $t \in [0, T]$.

Corollary 2.1 and Corollary 2.2 are in particular applicable for the continuous time analog (1.3) of ResNets (see Remark 3 for (1.2)).

The proof of Corollary 2.2 (see Section 5.4) will follow by firstly using a specific suboptimal control (constructed using the time-scaling specific to driftless systems) to estimate $J_T(u_T)$ and obtain

$$\int_t^{t+h} \|u_T(s)\|^2 ds \leq \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 ds$$

for h small enough, an estimate which, coupled with the turnpike estimates of Theorem 2.1 – Corollary 2.1 and the Lebesgue differentiation theorem, will suffice to conclude.

The proof of Theorem 2.1 may be found in Section 5.2. It roughly follows the following scheme (see also Figure 2 just below).

2.2.1. *Sketch of the proof of Theorem 2.1.* For simplicity, suppose that $T \geq 2T^*$.

- 1). By controllability, we first construct a suboptimal quasi-turnpike control u^1 which is such that the associated state y^1 satisfies $y^1(T_0) = \bar{y}$, and $u^1(t) = 0$ for $t \in [T_0, T]$. Thus $y^1(t) = \bar{y}$ for $t \in [T_0, T]$. Due to the form of J_T in (2.3), this would imply that $J_T(u^1)$ is independent of T , and by using $J_T(u_T) \leq J_T(u^1)$, would also entail a uniform bound of $J_T(u_T)$ with respect to T . A Grönwall argument ensures that, moreover,

$$\|y_T - \bar{y}\|_{L^2(0,T;\mathbb{R}^d)} + \|y_T(t) - \bar{y}\| \leq C_0 \quad \text{for all } t \in [0, T] \quad (2.13)$$

for some $C_0 > 0$ independent of T . (2.13) alone is enough to obtain the desired exponential estimates for $t \in [0, T^*] \cup [T - T^*, T]$, an interval whose length is independent of T . More details can be found in Lemma 5.1.

- 2). Since $T^* \leq \frac{T}{2}$, by a simple contradiction argument (see Lemma 5.3), there exist $\tau_1 \in [0, T^*]$ and $\tau_2 \in [T - T^*, T]$ such that

$$\|y_T(\tau_i) - \bar{y}\| \leq \frac{\|y_T - \bar{y}\|_{L^2(0,T;\mathbb{R}^d)}}{\sqrt{T^*}} \stackrel{(2.13)}{\leq} \frac{C_0}{\sqrt{T^*}}. \quad (2.14)$$

- 3). On $[\tau_1, \tau_2]$, the optimal control u_T will minimize a functional without the final cost $\phi(y_T(T))$ but with a terminal constraint on the state y_T . By controllability, using a second suboptimal quasi-turnpike control u^2 satisfying estimates as those in Assumption 1, and using $J_T(u_T) \leq J_T(u^2)$ along with a Grönwall argument, one shows an estimate of the form

$$\|y_T(t) - \bar{y}\| \leq C_1 \left(\|y_T(\tau_1) - \bar{y}\| + \|y_T(\tau_2) - \bar{y}\| \right) \quad (2.15)$$

$$\stackrel{(2.14)}{\leq} \frac{2C_1^2}{\sqrt{T^*}} \quad (2.16)$$

for all $t \in [\tau_1, \tau_2]$, thus also for $t \in [T^*, T - T^*] \subset [\tau_1, \tau_2]$ where $C_1 > 0$ is independent of T . Assumption 1 is used precisely in this step, and is essential in obtaining an estimate of the mould of (2.15). For more details, see Lemma 5.2.

- 4). A *bootstrap argument* (Section 5.2): estimate (2.16) can be iterated by shrinking the time interval to obtain an estimate of the form

$$\|y_T(t) - \bar{y}\| \leq \left(\frac{2C_1^2}{\sqrt{T^*}} \right)^n \quad \text{for } [nT^*, T - nT^*] \quad (2.17)$$

for "suitable" $n \geq 1$. Then taking $T^* > 4C_1^4$ and a suitable choice of n in (2.17) will yield the exponential estimate for $t \in [T^*, T - T^*]$.

2.3. **Comments on the main results.** Several pertinent remarks are in order.

Remark 3 (On the nonlinearity). *With little modifications, Theorem 2.1 and Corollary 2.1 also apply to system (2.1) with nonlinearities f of the form*

$$f(y, u) = \sum_{j=1}^m f_j(u_j y) \quad \text{for } (y, u) \in \mathbb{R}^d \times \mathbb{R}^m \quad (2.18)$$

where the vector fields $f_1, \dots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ are additionally assumed to be positively homogeneous of degree 1, and an H^1 -penalization instead of only L^2 of the control appears in the definition of J_T , in order to assert sufficient compactness for proving the

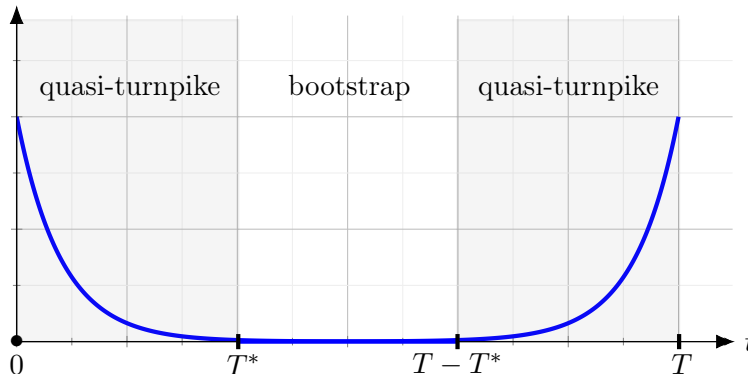


FIGURE 2. A sketch of the scheme. We use a quasi-turnpike control to bound $J_T(u_T)$ uniformly in T , which entails the exponential estimates on $[0, T^*] \cup [T - T^*, T]$. We then perform a bootstrap by iteratively shrinking symmetric intervals within $[T^*, T - T^*]$ in view of obtaining an estimate of the mould of (2.17).

existence of minimizers. Such nonlinearities are motivated by (1.2). Due to the homogeneity of the vector fields in (2.18), the corresponding optimal steady states coincide with those of the driftless case, namely $(u_s, y_s) = (0, \bar{y})$ for any $\bar{y} \in \mathbb{R}^d$.

Remark 4 (On Assumption 1). Both parts of Assumption 1 are needed in our strategy.

- In the driftless case ($f_0 = 0$ in (2.2)), the Chow-Rashevskii theorem (see [6, Chapter 3, Section 3.3]), characterized by iterated Lie brackets, is a necessary and sufficient condition for the global exact controllability of systems with smooth vector fields. But general necessary and sufficient conditions which ensure the exact controllability of control-affine systems are not known to our knowledge – see [6, Chapter 3]. This is mainly due to the drift term f_0 , which affects the geometry of the problem and may pose obstructions to the controllability in arbitrary time – see [2] for a survey on these issues. We do insist however, that we merely require controllability in a possibly large time T_0 , and not necessarily in any arbitrarily small time.
- The assumptions (2.6) – (2.7) are more commonly encountered in the linear systems setting, and thus also for nonlinear systems obtained by perturbation arguments. In such contexts, it is well-known (see e.g. [41, Remark 2.2]) that the minimal L^2 -norm control u satisfies

$$\|u\|_{L^2(0, T_0; \mathbb{R}^m)} \leq C(T_0) (\|y^0\| + \|y^1\|)$$

for some $C(T_0) > 0$. This makes Assumption 1 entirely plausible in the settings mentioned above. Indeed, we consider $z := y - \bar{y}$, then either $z^0 = 0$ (if $y^0 = \bar{y}$) or $z^1 = 0$ (if $y^1 = \bar{y}$). The control u steering y from y^0 to y^1 in time T would then be the same as the one steering z from either 0 to $y^1 - \bar{y}$ or from $y^0 - \bar{y}$ to 0 in time T , and the above estimate would yield the desired assumption.

To complete this discussion, we refer to [11, Theorem 5.2], where in the context of driftless systems motivated by neural networks (see the Introduction), we prove a local controllability result along with estimates (2.6) – (2.7). The

main caveat when comparing to the setting we consider here is that in neural networks, the control is typically a matrix of dimension $d \times d$ (eventhough we find a single control for $N \gg 1$ initial data), allowing us to access the entire state, whereas here it is a vector in \mathbb{R}^m , possibly with $m < d$. Nonetheless, driftless systems motivated by neural networks remain a case where our results apply.

3. INFINITE-DIMENSIONAL SYSTEMS

We illustrate the flexibility of the finite-dimensional arguments and adapt them to the semilinear wave and heat equation. As a matter of fact, the only difference between the finite and infinite dimensional setting is in the proof of uniform control and state bounds by means of quasi-turnpike strategies. The specific proof of turnpike is identical in both cases. We distinguish the case of the wave and heat equation because of the validity of the PDE analog of Assumption 1, as made more precise below.

3.1. Semilinear wave equation. Let $T > 0$ and let $\Omega \subset \mathbb{R}^d$ be a bounded and (at least C^2) regular domain. We will be interested in control systems of the form

$$\begin{cases} \partial_t^2 y - \Delta y + f(y) = u \mathbf{1}_\omega & \text{in } (0, T) \times \Omega \\ y = 0 & \text{on } (0, T) \times \partial\Omega \\ (y, \partial_t y)|_{t=0} = \mathbf{y}^0 & \text{in } \Omega. \end{cases} \quad (3.1)$$

Here $f \in \text{Lip}(\mathbb{R})$, $\omega \subset \Omega$ is open (with geometric assumptions given in (3.4)), whereas $\mathbf{y}^0 = (y_1^0, y_2^0)$ is a given initial datum. It is well-known, by fixed-point arguments, that for any initial data $\mathbf{y}^0 = (y_1^0, y_2^0) \in H_0^1(\Omega) \times L^2(\Omega)$ and for any $u \in L^2((0, T) \times \omega)$, there exists a unique finite-energy solution $y \in C^0([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$ to (3.1).

As in the finite-dimensional case, we will address the behavior when $T \gg 1$ of global minimizers $u_T \in L^2((0, T) \times \omega)$ to nonnegative functionals of the form

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \bar{y}\|_{H_0^1(\Omega)}^2 dt + \int_0^T \|\partial_t y(t)\|_{L^2(\Omega)}^2 dt + \int_0^T \|u(t)\|_{L^2(\omega)}^2 dt, \quad (3.2)$$

and of the corresponding solution y_T to (3.1). Here $\phi \in C^0(L^2(\Omega); \mathbb{R}_+)$ is a given final cost, while $\bar{y} \in H_0^1(\Omega)$ is a running target which we select as an uncontrolled steady state of (3.1), namely we assume that \bar{y} is some solution¹ to

$$\begin{cases} -\Delta \bar{y} + f(\bar{y}) = 0 & \text{in } \Omega \\ \bar{y} = 0 & \text{on } \partial\Omega. \end{cases} \quad (3.3)$$

We henceforth moreover assume that f, Ω are such that a solution to (3.3) exists. This can be ensured in a variety of different cases, including, for instance (see [5, 25] for further results):

¹There is no need for the solution of (3.3) to be unique.

- If $f(0) = 0$, then clearly $\bar{y} \equiv 0$ is one solution. But if moreover there exist $p \in \left(1, \frac{d+2}{d-2}\right)$ ($p \in (1, \infty)$ for $d = 1, 2$), $\nu < \lambda_1(\Omega)$ and $\theta > 2$ such that

$$\begin{aligned} |f(s)| &\leq C(1 + |s|^p) && \text{for all } s \in \mathbb{R} \\ -\int_0^s f(\zeta) \, d\zeta &\leq \frac{\nu}{2} s^2 && \text{for } |s| \text{ small} \\ 0 < -\theta \int_0^s f(\zeta) \, d\zeta &\leq -s \int_0^s f(\zeta) \, d\zeta && \text{for } |s| \text{ large,} \end{aligned}$$

then a nontrivial solution $\bar{y} \in H_0^1(\Omega)$, $\bar{y} \not\equiv 0$ also exists. We refer to [5, Theorem 2.5.6]. This fact is a consequence of the mountain pass theorem. Here $\lambda_1(\Omega)$ denotes the first eigenvalue of the Dirichlet Laplacian $-\Delta$.

- When $d = 1$ and $\Omega = (-R, R)$, then both necessary and sufficient conditions on f can be provided ensuring the existence of nontrivial solutions – see [5, Theorem 1.2.3].

The case of a controlled steady state (namely adding $\bar{u}\mathbf{1}_\omega$ in (3.3)) may also be considered, under the condition that the functional J_T is modified appropriately as discussed in Remark 1. The existence of minimizers to J_T again follows by the direct method in the calculus of variations.

We note that, since \bar{y} is fixed as above, the pair $(u_s, y_s) \equiv (0, \bar{y})$ is the unique solution to the steady-state optimal control problem

$$\inf_{u \in L^2(\omega)} \|y - \bar{y}\|_{H_0^1(\Omega)}^2 + \|u\|_{L^2(\omega)}^2 \quad \text{subject to} \quad \begin{cases} -\Delta y + f(y) = u\mathbf{1}_\omega & \text{in } \Omega \\ y = 0 & \text{on } \partial\Omega. \end{cases}$$

This is because the functional in the expression above attains its minimum, equal to 0, precisely at $(0, \bar{y})$, a pair which satisfies the constraint provided by the elliptic equation.

Before proceeding, we need to define the appropriate geometric setup for ensuring the exact controllability of (3.1) when $d \geq 2$. For any fixed $x_\circ \in \mathbb{R}^d \setminus \bar{\Omega}$, we define

$$\Gamma_\circ := \{x \in \partial\Omega : (x - x_\circ) \cdot \nu(x) > 0\}$$

where $\nu(x)$ denotes the outward unit normal at $x \in \partial\Omega$. The set Γ_\circ coincides with the subset of the boundary arising usually in the context of the multiplier method [24]. We will suppose that for some $\delta > 0$,

$$\omega = \mathcal{O}_\delta(\Gamma_\circ) \cap \Omega, \tag{3.4}$$

where $\mathcal{O}_\delta(\Gamma_\circ) := \{x \in \mathbb{R}^d : |x - x'| < \delta \text{ for some } x' \in \Gamma_\circ\}$. It is known that, under these geometric assumptions on ω , and since $f \in \text{Lip}(\mathbb{R})$, there exists some time $T_{\min} = T_{\min}(\Omega, \omega) > 0$ such that the wave equation (3.1) is exactly controllable in any time $T_0 > T_{\min}$, see [12, 39] and also [9, Section 7.2] (see also the introduction of [19] for an ample survey of controllability results for semilinear wave equations). These results are extensions of the one-dimensional results in [40].

We may now state our main result in the context of the wave equation.

Theorem 3.1 (Turnpike). *Suppose that $f \in \text{Lip}(\mathbb{R})$ and $\Omega \subset \mathbb{R}^d$ are such that (3.3) admits at least one solution, and let $\bar{y} \in H_0^1(\Omega)$ be any such solution. Let $\phi \in C^0(L^2(\Omega); \mathbb{R}_+)$, and suppose that ω is as in (3.4). For any $\mathbf{y}^0 \in H_0^1(\Omega) \times L^2(\Omega)$, there exists a time $T^* > T_{\min}(\omega, \Omega)$ and constants $C_1, C_2 > 0$ and $\mu > 0$, such that for any $T \geq T^*$, any*

global minimizer $u_T \in L^2((0, T) \times \omega)$ to J_T defined in (3.2) and corresponding optimal state y_T solution to (3.1) satisfy

$$\|y_T(t) - \bar{y}\|_{H_0^1(\Omega)} + \|\partial_t y_T(t)\|_{L^2(\Omega)} \leq C_1 \left(e^{-\mu t} + e^{-\mu(T-t)} \right)$$

for all $t \in [0, T]$, and

$$\|u_T\|_{L^2((0, T) \times \omega)} \leq C_2.$$

Moreover, $\mu > 0$ is independent of \mathbf{y}^0 .

The proof of turnpike (see Section 6) is identical to the finite-dimensional case. Some technical adaptations are however needed for obtaining the quasi-turnpike bounds, wherein one uses the Duhamel formula for mild solutions in view of applying an integral Grönwall argument, in the spirit of the ODE setting.

Remark 5 (On the choice of J_T). *We note that in existing turnpike results for the wave equation, e.g. [17, 35, 42], a slightly weaker functional is sometimes considered. For instance, in [42] for the linear wave equation, only the $L^2(0, T; H_0^1(\Omega))$ -norm of $y - \bar{y}$ is penalized, and not the $L^2((0, T) \times \Omega)$ -norm of $\partial_t y$, yet turnpike is shown to hold for the full state $(y, \partial_t y)$. This is justified by the equipartition of energy property, which states that, along a given time interval $[0, T]$, the energy concentrated on the y component in $H_0^1(\Omega)$ and on the $\partial_t y$ component in $L^2(\Omega)$ is comparably the same up to a compact remainder term. We choose to work with a functional penalizing the full state of the system due to the specificity of our proof strategy.*

Similarly to the finite-dimensional case, when $\phi \equiv 0$ in (3.2), Theorem 3.1 entails an exponential stabilization property for the optimal states, namely

Corollary 3.1 (Stabilization). *Suppose that $\phi \equiv 0$ in J_T defined in (3.2). Under the assumptions of Theorem 3.1, there exists a time $T^* > T_{\min}(\omega, \Omega)$ and constants $C_1, C_2, \mu > 0$ such that for any $T \geq T^*$, any global minimizer $u_T \in L^2((0, T) \times \omega)$ to J_T defined in (3.2) and corresponding optimal state y_T solution to (3.1) satisfy*

$$\|y_T(t) - \bar{y}\|_{H_0^1(\Omega)} + \|\partial_t y_T(t)\|_{L^2(\Omega)} \leq C_1 e^{-\mu t}$$

for all $t \in [0, T]$ and

$$\|u_T\|_{L^2((0, T) \times \omega)} \leq C_2.$$

Moreover, $\mu > 0$ is independent of \mathbf{y}^0 .

3.2. Semilinear heat equation. To complete our presentation, we will also discuss control systems of the form

$$\begin{cases} \partial_t y - \Delta y + f(y) = u \mathbf{1}_\omega & \text{in } (0, T) \times \Omega \\ y = 0 & \text{on } (0, T) \times \partial\Omega \\ y|_{t=0} = y^0 & \text{in } \Omega, \end{cases} \quad (3.5)$$

where $f \in \text{Lip}(\mathbb{R})$, $\omega \subset \Omega$ is any open, non-empty subset, whereas y^0 is a given initial datum. It is well-known that for any given $T > 0$, $y^0 \in L^2(\Omega)$ and $u \in L^2((0, T) \times \omega)$, there exists a unique globally-defined solution $y \in C^0([0, T]; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$ to (3.5).

We will again study global minimizers $u_T \in L^2((0, T) \times \omega)$ to nonnegative functionals of the form

$$J_T(u) := \int_0^T \|y(t) - \bar{y}\|_{L^2(\Omega)}^2 dt + \int_0^T \|u(t)\|_{L^2(\omega)}^2 dt, \quad (3.6)$$

and the corresponding solution y_T to (3.5) in the regime $T \gg 1$. Once again, $\bar{y} \in L^2(\Omega)$ is a running target which we select as an uncontrolled steady state, namely a solution to (3.3). The existence of minimizers to J_T defined in (3.6) follows by the direct method in the calculus of variations.

Theorem 3.2 (Stabilization). *Suppose that $f \in \text{Lip}(\mathbb{R})$ and $\Omega \subset \mathbb{R}^d$ are such that (3.3) admits at least one solution, and let $\bar{y} \in H_0^1(\Omega)$ be any such solution. For any $y^0 \in L^2(\Omega)$, there exists $T^* > 0$ and constants $C_1, C_2, \mu > 0$ such that for any $T \geq T^*$, any global minimizer $u_T \in L^2((0, T) \times \omega)$ of J_T defined in (3.6) and corresponding optimal state y_T solution to (3.5) satisfy*

$$\|y_T(t) - \bar{y}\|_{L^2(\Omega)} \leq C_1 e^{-\mu t}$$

for all $t \in [0, T]$, and

$$\|u_T\|_{L^2((0, T) \times \omega)} \leq C_2.$$

Moreover, $\mu > 0$ is independent of y^0 .

We refer to Section 7 for the proof.

We consider the heat equation in addition to the wave equation because of the validity of the PDE analog of Assumption 1. The heat equation is exactly controllable to controlled trajectories, namely solutions \hat{y} to (3.5) for given controls \hat{u} . Instead of an estimate such as (2.7), one has $\|u - \hat{u}\|_{L^2((0, T_0) \times \omega)} \leq C(T_0) \|y^0 - \hat{y}(0)\|_{L^2(\Omega)}$ (see e.g. [30, Lemma 8.3] and the references therein) for minimal L^2 -norm controls u steering y to \hat{y} in time T_0 . Such an estimate does not suffice for applying our methodology, as we clearly need to estimate the minimal L^2 -norm control by means of the distance of the initial data to the target. Nonetheless, we illustrate that the stabilization result can be shown independently of the turnpike result. Indeed, the proof closely follows that of Theorem 2.1, with the exception that we only need to perform the bootstrap forward in time, whence we do not require that the system is controllable to anything else but a steady state. We refer to Section 7 for more details.

The semilinear heat equation is a commonly used benchmark for nonlinear turnpike results, thus this example serves to compare with existing results, such as those in [29].

Remark 6 (On the nonlinearity). *The assumption that f is globally Lipschitz in (3.1) and (3.5) could perhaps be relaxed to a locally Lipschitz f (for which blow-up is avoided and controllability is ensured – for instance, $f(y) = y^3$), under the condition that one can show a uniform $L^\infty((0, T) \times \Omega)$ -estimate of y_T with respect to $T > 0$. Arguments of this sort in the context of turnpike can be found in [29] under smallness assumptions on the target. We refer to the end of Section 8.1 for a discussion of a (possibly technical) impediment encountered in applying our methodology to the cubic heat equation. In addition to the controllability properties it entails for (3.1) – (3.5) as blow-up is avoided, we use the Lipschitz character of f in the estimates in Lemma 6.1, Lemma 6.2 and Lemma 7.1.*

4. PRELIMINARY RESULTS

We begin by presenting a couple of simple but important lemmas, containing bounds of the quantity $\|y(t) - \bar{y}\|$ for both the nonlinear ODE and PDE setting, solely by means of $\|y^0 - \bar{y}\|$ and the tracking terms appearing in the functional J_T . These bounds would thus imply that bounding the functional J_T uniformly in T would entail a bound for the desired quantity $\|y(t) - \bar{y}\|$.

Let us begin with the ODE estimate.

Lemma 4.1. *Let $T > 0$ be given, and let $\bar{y} \in \mathbb{R}^d$ be as in (2.4). For any data $u \in L^2(0, T; \mathbb{R}^m)$ and $y^0 \in \mathbb{R}^d$, let $y \in C^0([0, T]; \mathbb{R}^d)$ be the solution to (2.1) with $y(0) = y^0$. Then there exist constants $C_1 = C_1(f, \bar{y}) > 0$ and $C_2 = C_2(f)$ independent of T such that*

$$\|y(t) - \bar{y}\| \leq C \left(\|y^0 - \bar{y}\| + \|u\|_{L^2(0, T; \mathbb{R}^m)} + \|y - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)} \right)$$

holds for all $t \in [0, T]$, where

$$C := C_1 \exp \left(C_2 \|u\|_{L^2(0, T; \mathbb{R}^m)} \right).$$

As insinuated by the form of the constant in the estimate, the proof follows a Grönwall argument. However, as this constant depends on T only through the L^2 -norm of the control u , we present the proof for the sake of clarity.

Proof of Lemma 4.1. Let us first suppose that $t \in [0, 1]$. By integrating the equation satisfied by y , and using the fact that $f_0, \dots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ and $t \leq 1$, as well as Cauchy-Schwarz, it may be seen that

$$\|y(t) - \bar{y}\| \leq C_0 \left(\|y^0 - \bar{y}\| + \|u\|_{L^2(0, T; \mathbb{R}^m)} \right)$$

for some $C_0 = C_0(f) > 0$.

Now suppose that $t \in (1, T]$. We begin by showing that for any such t , there exists a $t^* \in (t - 1, t]$ such that

$$\|y(t^*) - \bar{y}\| \leq \|y - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}. \quad (4.1)$$

To this end, we argue by contradiction. Suppose that

$$\|y(t^*) - \bar{y}\| > \|y - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}$$

for all $t^* \in (t - 1, t]$. Then

$$\|y - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2 = \int_0^T \|y(t) - \bar{y}\|^2 dt \geq \int_{t-1}^t \|y(\tau) - \bar{y}\|^2 d\tau > \|y - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2,$$

which contradicts the hypothesis. Thus (4.1) holds.

Consequently, we know that there exists $t^* \in (t - 1, t]$ such that (4.1) holds. By integrating the equation satisfied by y in $[t^*, t]$, namely writing

$$\begin{aligned} y(t) - \bar{y} &= y(t^*) - \bar{y} + \int_{t^*}^t \left(f_0(y) + \sum_{j=1}^m u_j f_j(y) \right) d\tau \\ &= y(t^*) - \bar{y} + \int_{t^*}^t (f_0(y) - f_0(\bar{y})) d\tau + \int_{t^*}^t \sum_{j=1}^m u_j (f_j(y) - f_j(\bar{y})) d\tau \\ &\quad + \int_{t^*}^t \sum_{j=1}^m u_j f_j(\bar{y}) d\tau, \end{aligned}$$

we see that, by using the Lipschitz character of f_0, \dots, f_m and Cauchy-Schwarz for the sums,

$$\|y(t) - \bar{y}\| \leq \|y(t^*) - \bar{y}\| + C_0(f) \int_{t^*}^t \left(1 + \|u(\tau)\| \right) \|y(\tau) - \bar{y}\| d\tau + C_1(f, \bar{y}) \int_{t^*}^t \|u(\tau)\| d\tau.$$

Now applying a combination of Cauchy-Schwarz, the fact that $t - t^* \leq 1$, (4.1), and the Grönwall inequality to the inequality just above, we obtain

$$\|y(t) - \bar{y}\| \leq C_2 \exp \left(C_3(f) \sqrt{1 + \int_{t^*}^t \|u(\tau)\|^2 d\tau} \right) \left(\|y - \bar{y}\|_{L^2(0,T;\mathbb{R}^d)} + \|u\|_{L^2(0,T;\mathbb{R}^m)} \right),$$

for some $C_2(f, \bar{y}) > 0$ and $C_3(f) > 0$, from which, using $\sqrt{x^2 + y^2} \leq x + y$ for $x, y > 0$, the desired statement readily follows. \square

Remark 7. *Let us make two brief observations.*

- We note that in the case where the running target is (\bar{u}, \bar{y}) with $f(\bar{y}, \bar{u}) = 0$ and $\bar{u} \neq 0$, and thus we minimize J_T defined in (1.4), we argue as above to obtain a bound of the form

$$\|y(t) - \bar{y}\| \leq C \left(\|y^0 - \bar{y}\| + \|u - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)} + \|y - \bar{y}\|_{L^2(0,T;\mathbb{R}^d)} \right)$$

with $C \sim \exp(\|u - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)})$. Obtaining a dependence of the constant C with respect to $\|u - \bar{u}\|_{L^2(0,T;\mathbb{R}^m)}$ rather than just $\|u\|_{L^2(0,T;\mathbb{R}^m)}$ is important, as by using the functional and optimality arguments, we will be able to obtain a uniform bound with respect to T of the former, which does not necessarily entail a bound on the latter. The argument for deducing such a bound is identical to the proof of Lemma 4.1 – assume that $m = 1$ for notational simplicity, and observe that, since $f_0(\bar{y}) + \bar{u}f_1(\bar{y}) = 0$,

$$\begin{aligned} y(t) - \bar{y} &= y(t^*) - \bar{y} + \int_{t^*}^t (f_0(y) - f_0(\bar{y})) ds + \int_{t^*}^t (u - \bar{u})(f_1(y) - f_1(\bar{y})) ds \\ &\quad + \int_{t^*}^t (u - \bar{u})f_1(\bar{y}) ds + \int_{t^*}^t \bar{u}(f_1(y) - f_1(\bar{y})) ds. \end{aligned}$$

One may then proceed as before.

- It may readily be seen that if the control is of additive rather than multiplicative form, i.e. if f_1, \dots, f_m are nonzero constants, then the constant appearing in the estimate provided by Lemma 4.1 will not depend on the time horizon T .

We state and prove an analogous result for the semilinear heat equation (3.5). The proof is almost identical to the ODE case, but we sketch it for the sake of clarity.

Lemma 4.2. *Let $T > 0$ be given, and let \bar{y} be as in (3.3). For any $u \in L^2((0, T) \times \omega)$ and $y^0 \in L^2(\Omega)$, let $y \in C^0([0, T]; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$ be the unique weak solution to (3.5). Then there exists a constant $C = C(f) > 0$ independent of T such that*

$$\|y(t) - \bar{y}\|_{L^2(\Omega)} \leq C \left(\|y^0 - \bar{y}\|_{L^2(\Omega)} + \|u\|_{L^2((0,T) \times \omega)} + \|y - \bar{y}\|_{L^2((0,T) \times \Omega)} \right)$$

holds for all $t \in [0, T]$.

Proof of Lemma 4.2. The proof closely follows that of Lemma 4.1. We first note that by uniqueness, $y - \bar{y}$ can be shown (see [1]) to coincide with the unique mild solution to

$$\begin{cases} \partial_t z - \Delta z + f(z + \bar{y}) - f(\bar{y}) = u \mathbf{1}_\omega & \text{in } (0, T) \times \Omega \\ z = 0 & \text{on } (0, T) \times \partial\Omega \\ z|_{t=0} = y^0 - \bar{y} & \text{in } \Omega \end{cases}$$

which is given by the Duhamel/variation by constants formula:

$$y(t) - \bar{y} = e^{t\Delta}(y^0 - \bar{y}) + \int_0^t e^{(t-s)\Delta} u(s) \mathbf{1}_\omega \, ds - \int_0^t e^{(t-s)\Delta} (f(y) - f(\bar{y})) \, ds, \quad (4.2)$$

where $\{e^{t\Delta}\}_{t>0}$ denotes the heat semigroup on $L^2(\Omega)$ generated by the Dirichlet Laplacian $-\Delta : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow L^2(\Omega)$. Of course, (4.2) is interpreted as an identity in $L^2(\Omega)$. We may thus proceed and use (4.2) throughout.

First suppose that $0 < t \leq 1$. Using the well-known property $\|e^{t\Delta}\| \leq e^{-\lambda_1(\Omega)t} \leq 1$ of the heat semigroup (where $\lambda_1(\Omega) > 0$ denotes the first eigenvalue of the Dirichlet Laplacian), and the Lipschitz character of f , we find using (4.2) that

$$\begin{aligned} \|y(t) - \bar{y}\|_{L^2(\Omega)} &\leq \|e^{t\Delta}(y^0 - \bar{y})\|_{L^2(\Omega)} + \int_0^t \|e^{(t-s)\Delta} u(s)\|_{L^2(\omega)} \, ds \\ &\quad + \int_0^t \|e^{(t-s)\Delta} (f(y(s)) - f(\bar{y}))\|_{L^2(\Omega)} \, ds \\ &\leq \|y^0 - \bar{y}\|_{L^2(\Omega)} + \int_0^t \|u(s)\|_{L^2(\omega)} \, ds \\ &\quad + C_0 \int_0^t \|y(s) - \bar{y}\|_{L^2(\Omega)} \, ds, \end{aligned}$$

where $C_0 = C_0(f) > 0$ is the Lipschitz constant of f . As $t \leq 1$, we may use Cauchy-Schwarz and Grönwall to conclude.

Now suppose that $t \in (1, T]$. Arguing as in the proof of Lemma 4.1, we know that there exists a $t^* \in (t-1, t]$ such that

$$\|y(t^*) - \bar{y}\|_{L^2(\Omega)} \leq \|y - \bar{y}\|_{L^2((0,T) \times \Omega)} \quad (4.3)$$

holds. By writing the Duhamel formula for $y - \bar{y}$ in $[t^*, t]$, namely writing

$$y(t) - \bar{y} = e^{t\Delta}(y(t^*) - \bar{y}) + \int_{t^*}^t e^{(t-s)\Delta} u(s) \, ds - \int_{t^*}^t e^{(t-s)\Delta} (f(y) - f(\bar{y})) \, ds$$

we see just as before that

$$\|y(t) - \bar{y}\|_{L^2(\Omega)} \leq \|y(t^*) - \bar{y}\|_{L^2(\Omega)} + \int_0^t \|u(s)\|_{L^2(\omega)} \, ds + C_0 \int_{t^*}^t \|y(s) - \bar{y}\|_{L^2(\Omega)} \, ds$$

where $C_0 = C_0(f) > 0$ is the Lipschitz constant of f . Using the fact that $t^* - t \leq 1$ and (4.3), we may, as before, apply Cauchy-Schwarz and Grönwall to conclude. \square

We finally show the analog estimate for the semilinear wave equation, which is, after defining the proper functional setup, identical to the proof of Lemma 4.2.

Lemma 4.3. *Let $T > 0$ be given, and let \bar{y} be as in (3.3). For any $u \in L^2((0, T) \times \omega)$ and $\mathbf{y}^0 = (y_1^0, y_2^0) \in H_0^1(\Omega) \times L^2(\Omega)$, let $y \in C^0([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$ be the unique weak solution to (3.1). Then there exists a constant $C = C(f, \Omega) > 0$ independent of T such that*

$$\begin{aligned} &\|y(t) - \bar{y}\|_{H_0^1(\Omega)} + \|\partial_t y(t)\|_{L^2(\Omega)} \\ &\leq C \left(\|y_1^0 - \bar{y}\|_{H_0^1(\Omega)} + \|y_2^0\|_{L^2(\Omega)} + \|u\|_{L^2((0,T) \times \omega)} + \|y - \bar{y}\|_{H_0^1((0,T) \times \Omega)} + \|\partial_t y\|_{L^2((0,T) \times \Omega)} \right) \end{aligned}$$

holds for all $t \in [0, T]$.

Proof of Lemma 4.3. Once (3.1) is written as a first order evolution equation in an appropriate Hilbert space X , the proof is identical to that of Lemma 4.2. Define the energy space $X := H_0^1(\Omega) \times L^2(\Omega)$, and consider the closed, densely-defined operator

$$A := \begin{bmatrix} 0 & \text{Id} \\ \Delta & 0 \end{bmatrix}, \quad D(A) = D(\Delta) \times H_0^1(\Omega),$$

where $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$. The operator A is skew-adjoint and thus generates a strongly continuous semigroup $\{e^{tA}\}_{t>0}$ in X by virtue of the Stone-Lumer-Phillips theorem (see e.g. [37, Theorem 3.8.6]). We now denote

$$\mathbf{y} := \begin{bmatrix} y \\ \partial_t y \end{bmatrix}, \quad \bar{\mathbf{y}} := \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix}.$$

Analog arguments to those in Lemma 4.2 lead us to deduce that

$$\mathbf{y}(t) - \bar{\mathbf{y}} = e^{tA} (\mathbf{y}^0 - \bar{\mathbf{y}}) + \int_0^t e^{(t-s)A} \begin{bmatrix} 0 \\ u(s)\mathbf{1}_\omega - f(y(s)) + f(\bar{y}) \end{bmatrix} ds \quad (4.4)$$

for $t > 0$ is the unique mild solution to the equation satisfied by the perturbation $\mathbf{y} - \bar{\mathbf{y}}$. Of course, (4.4) is interpreted as an identity in X . By virtue of the conservative character of the semigroup, namely $\|e^{tA}g\|_X = \|g\|_X$ for all $t > 0$ and $g \in X$, we see that one may apply precisely the same arguments as in the proof of Lemma 4.2, this time to the integral formulation (4.4) in X (with an intermediate application of the Poincaré inequality after using the Lipschitz character of f) to conclude. \square

5. PROOF OF THEOREM 2.1

In this section, we present the proof of Theorem 2.1, Corollary 2.1 and Corollary 2.2. The proof of Theorem 2.1 requires a couple of preliminary results. In particular, we will, by means of a quasi-turnpike control strategy, provide bounds – uniform with respect to the time horizon T^- of the tracking terms appearing in the definition (2.3) of the functional J_T for the optimal control-state pairs (u_T, y_T) .

5.1. Quasi-turnpike lemmas. Both of the following results are heavily based on the specific choice of target \bar{y} as a steady state of the nonlinear system with 0 control, and on the (locally) Lipschitz character of the nonlinear terms.

We begin with the following lemma.

Lemma 5.1. *Let $y^0 \in \mathbb{R}^d$ be given, and assume that system (2.1) is controllable in some time $T_0 > 0$. Let $T > 0$ be fixed, and let $u_T \in L^2(0, T; \mathbb{R}^m)$ be a global minimizer to J_T defined in (2.3), with y_T denoting the associated solution to (2.1) with $y_T(0) = y^0$.*

Then, there exists a constant $C = C(f, \phi, T_0, \bar{y}, y^0) > 0$ independent of $T > 0$ such that

$$\|u_T\|_{L^2(0, T; \mathbb{R}^m)} + \|y_T - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)} + \|y_T(t) - \bar{y}\| \leq C \quad (5.1)$$

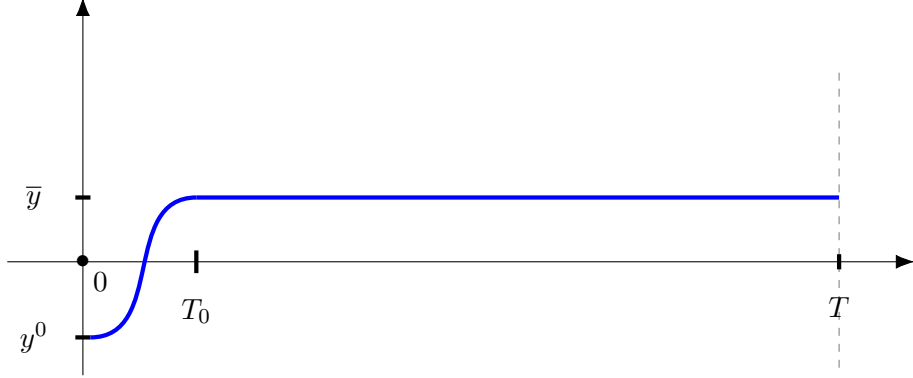
holds for all $t \in [0, T]$.

Proof of Lemma 5.1. We begin by considering the case $T \geq T_0$. Using the controllability assumption, we know that there exists a control $u^\dagger \in L^2(0, T_0; \mathbb{R}^m)$ such that the corresponding solution y^\dagger to

$$\begin{cases} \dot{y}^\dagger = f(y^\dagger, u^\dagger) & \text{in } (0, T_0) \\ y^\dagger(0) = y^0 \end{cases}$$



FIGURE 3. **Proof of Lemma 5.1.** The first two terms appearing in (5.1) also appear in the functional $J_T(u_T)$. We thus construct an admissible quasi-turnpike control u^{aux} (red), for which the corresponding state y^{aux} (blue) coincides with \bar{y} over (T_0, T) . In this way, as $J_T(u_T) \leq J_T(u^{\text{aux}})$, and $J_T(u^{\text{aux}})$ is independent of T , we can conclude. The estimate of the third term then follows from Lemma 4.1.



satisfies $y^\dagger(T_0) = \bar{y}$. Now set

$$u^{\text{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (0, T_0) \\ 0 & \text{in } (T_0, T) \end{cases}$$

and let y^{aux} be the corresponding solution to (2.1) with $y^{\text{aux}}(0) = y^0$. Clearly $y^{\text{aux}}(t) = \bar{y}$ for $t \in [T_0, T]$. Hence, using $\phi \geq 0$ and $J_T(u_T) \leq J_T(u^{\text{aux}})$, we see that

$$\|y_T - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2 + \|u_T\|_{L^2(0, T; \mathbb{R}^m)}^2 \leq \phi(\bar{y}) + \|y^\dagger - \bar{y}\|_{L^2(0, T_0; \mathbb{R}^d)}^2 + \|u^\dagger\|_{L^2(0, T_0; \mathbb{R}^m)}^2.$$

As the right hand side in the above inequality is clearly independent of T , we conclude the proof by applying Lemma 4.1 after noting the uniform boundedness of $\|u_T\|_{L^2(0, T; \mathbb{R}^m)}$ with respect to $T > 0$.

Now suppose that $T \leq T_0$. In this case, we use $\phi \geq 0$ and the optimality inequality $J_T(u_T) \leq J_T(u_{T_0+1})$ with the effect of obtaining

$$\begin{aligned} \|y_T - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2 + \|u_T\|_{L^2(0, T; \mathbb{R}^m)}^2 \\ \leq \phi(y_{T_0+1}(T)) + \|y_{T_0+1} - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2 + \|u_{T_0+1}\|_{L^2(0, T; \mathbb{R}^m)}^2 \end{aligned}$$

Now the trajectory $y_{T_0+1} \in C^0([0, T_0 + 1]; \mathbb{R}^d)$ is uniformly bounded with respect to T by virtue of the case presented just above. Whence, using the continuity of ϕ , as well as $T \leq T_0$, we may conclude that

$$\|y_T - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}^2 + \|u_T\|_{L^2(0, T; \mathbb{R}^m)}^2 \leq C$$

for some $C > 0$ independent of T . We may use Lemma 4.1 to conclude. \square

We will now focus on an auxiliary control problem with *fixed* endpoints. Namely, given $y^{\tau_1}, y^{\tau_2} \in \mathbb{R}^d$, and $0 \leq \tau_1 < \tau_2 \leq T$, this problem consists in minimizing the nonnegative functional

$$J_{\tau_1, \tau_2}(u) := \int_{\tau_1}^{\tau_2} \|y(t) - \bar{y}\|^2 dt + \int_{\tau_1}^{\tau_2} \|u(t)\|^2 dt \quad (5.2)$$

over all $u \in U_{\text{ad}}$, where $y \in C^0([\tau_1, \tau_2]; \mathbb{R}^d)$ denotes the unique solution to

$$\begin{cases} \dot{y} = f(y, u) & \text{in } (\tau_1, \tau_2) \\ y(\tau_1) = y^{\tau_1} \end{cases} \quad (5.3)$$

where

$$U_{\text{ad}} := \{u \in L^2(\tau_1, \tau_2; \mathbb{R}^m) : y(\tau_2) = y^{\tau_2}\}.$$

The following lemma is of key importance in what follows. It ensures that the optimal controls (for J_{τ_1, τ_2}) and trajectories are in fact bounded by means of the distance of the starting point y^{τ_1} and endpoint y^{τ_2} from the running target \bar{y} . This estimate will be the cornerstone of the bootstrap argument performed in the proof of Theorem 2.1.

Lemma 5.2. *Let $\bar{y} \in \mathbb{R}^d$ be as in (2.4), and assume that system (2.1) is controllable in some time $T_0 > 0$ in the sense of Assumption 1. Let $r > 0$ be the radius provided by Assumption 1, and let $y^{\tau_1}, y^{\tau_2} \in \mathbb{R}^d$ be such that*

$$\|y^{\tau_i} - \bar{y}\| \leq r$$

for $i = 1, 2$. Let $0 \leq \tau_1 < \tau_2 \leq T$ be fixed such that $\tau_2 - \tau_1 \geq 2T_0$, and let $u_T \in U_{\text{ad}}$ be a global minimizer to J_{τ_1, τ_2} defined in (5.2), with y_T denoting the associated solution to (5.3) with $y_T(\tau_2) = y^{\tau_2}$.

Then, there exists a constant $C = C(f, T_0, \bar{y}, r) > 0$ independent of $T, \tau_1, \tau_2 > 0$ such that

$$\|u_T\|_{L^2(\tau_1, \tau_2; \mathbb{R}^m)}^2 + \|y_T - \bar{y}\|_{L^2(\tau_1, \tau_2; \mathbb{R}^d)}^2 + \|y_T(t) - \bar{y}\|^2 \leq C \left(\|y^{\tau_1} - \bar{y}\|^2 + \|y^{\tau_2} - \bar{y}\|^2 \right)$$

holds for all $t \in [\tau_1, \tau_2]$. Moreover, the map $r \mapsto C(f, T_0, \bar{y}, r)$ is non-decreasing as a function from \mathbb{R}_+ to \mathbb{R}_+ .

The key idea of the proof of Lemma 5.2 lies in the construction of an auxiliary suboptimal quasi-turnpike control (steering the corresponding trajectory from y^{τ_1} to y^{τ_2} in time $\tau_2 - \tau_1$, whilst remaining at \bar{y} over an interval of length $\tau_2 - \tau_1 - 2T_0$; see the figure just below) in view of estimating each individual addend of $J_{\tau_1, \tau_2}(u_T)$, which is the minimal value of the functional J_{τ_1, τ_2} . This construction will yield the desired result.

Proof of Lemma 5.2. Using the controllability assumption, we know the following.

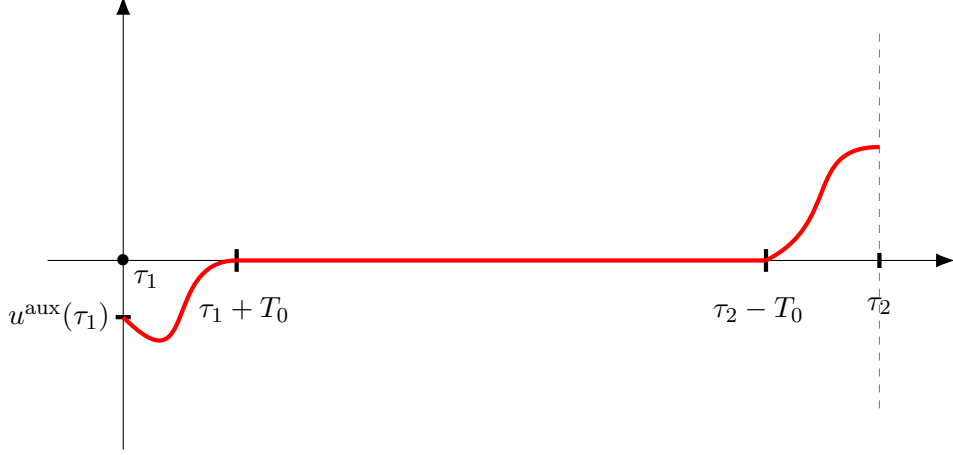
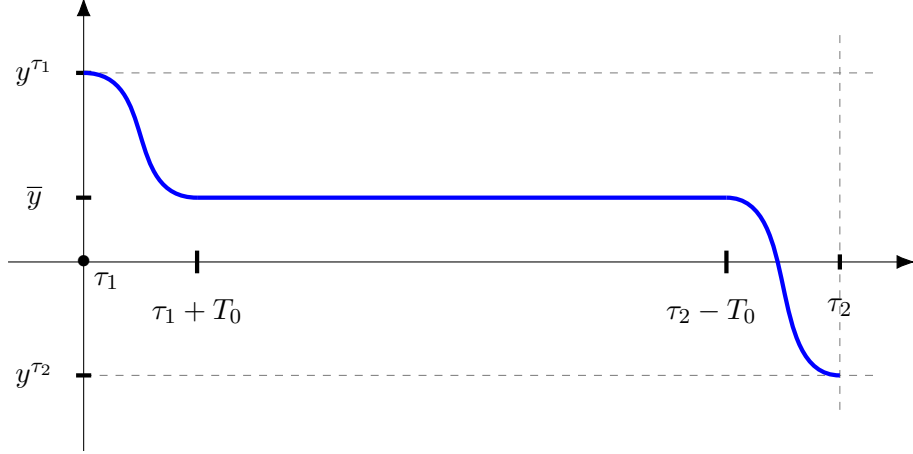


FIGURE 4. **Proof of Lemma 5.2.** The first two terms appearing in the estimate implied by Lemma 5.2 also appear in the functional $J_{\tau_1, \tau_2}(u_T)$. We thus construct an admissible quasi-turnpike control u^{aux} (red), for which the corresponding state y^{aux} (blue) coincides with \bar{y} over $(\tau_1 + T_0, \tau_2 - T_0)$. In this way, as $J_{\tau_1, \tau_2}(u_T) \leq J_{\tau_1, \tau_2}(u^{\text{aux}})$, and $J_{\tau_1, \tau_2}(u^{\text{aux}})$ is independent of T, τ_1, τ_2 , we can conclude. The estimate of the third term then follows from Lemma 4.1.



- There exists a control $u^\dagger \in L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)$ satisfying

$$\|u^\dagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}^2 \leq C(T_0) \|y^{\tau_1} - \bar{y}\|^2, \quad (5.4)$$

for some $C(T_0) > 0$, and which is such that the corresponding solution y^\dagger to

$$\begin{cases} \dot{y}^\dagger = f(y^\dagger, u^\dagger) & \text{in } (\tau_1, \tau_1 + T_0) \\ y^\dagger(\tau_1) = y^{\tau_1} \end{cases} \quad (5.5)$$

satisfies $y^\dagger(\tau_1 + T_0) = \bar{y}$. By integrating (5.5), and using the Lipschitz character of f_0, \dots, f_m , Grönwall's inequality, Cauchy-Schwarz and (5.4), we see that

$$\begin{aligned} \|y^\dagger(t)\| &\leq C_0 \left(\|y^{\tau_1}\| + \|u^\dagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + 1 \right) \exp \left(C_0 \|u^\dagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} \right) \\ &\leq C_1 \left(\|y^{\tau_1}\| + \|y^{\tau_1} - \bar{y}\| + 1 \right) \exp \left(C_1 \|y^{\tau_1} - \bar{y}\| \right) \\ &\leq C_1 \left(\|y^{\tau_1}\| + r + 1 \right) \exp \left(C_1 r \right) \\ &\leq C_2 \left(\|\bar{y}\| + r + 1 \right) \exp \left(C_2 r \right) \end{aligned} \quad (5.6)$$

for some $C_0 = C_0(f, T_0) > 0$, $C_1 = C_1(f, T_0) > 0$, $C_2 = C_2(f, T_0) > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$. Then, by integrating (5.5) once again, and using $f_0(\bar{y}) = 0$, Cauchy-Schwarz and (5.6), we moreover see that

$$\begin{aligned} \|y^\dagger(t) - \bar{y}\| &\leq \|y^{\tau_1} - \bar{y}\| + \int_{\tau_1}^t \sum_{j=1}^m |u_j^\dagger(s)| \|f_j(y^\dagger)\| \, ds + \int_{\tau_1}^t \|f(y^\dagger) - f(\bar{y})\| \, ds \\ &\leq \|y^{\tau_1} - \bar{y}\| + C_3 \|u^\dagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + C(f) \int_{\tau_1}^t \|y^\dagger(s) - \bar{y}\| \, ds \end{aligned} \quad (5.7)$$

for some $C_3(f, T_0, r, \bar{y}) > 0$, with $C(f) > 0$ being the Lipschitz constant of the vector fields f_j . Finally, applying Grönwall's inequality to (5.7) and using (5.4), we deduce that

$$\|y^\dagger(t) - \bar{y}\| \leq C_4 \exp(C(f)T_0) \|y^{\tau_1} - \bar{y}\| \quad (5.8)$$

for some $C_4(f, T_0, \bar{y}, r) > 0$ independent of $T, \tau_1, \tau_2 > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$. Note that in view of (5.6), both C_3 and C_4 are non-decreasing with respect to the parameter $r > 0$.

- There exists a control $u^\ddagger \in L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)$ satisfying

$$\|u^\ddagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}^2 \leq C(T_0) \|\bar{y} - y^{\tau_2}\|^2, \quad (5.9)$$

and which is such that the corresponding solution y^\ddagger to

$$\begin{cases} \dot{y}^\ddagger = f(y^\ddagger, u^\ddagger) & \text{in } (\tau_1, \tau_1 + T_0) \\ y^\ddagger(\tau_1) = \bar{y} \end{cases} \quad (5.10)$$

satisfies $y^\ddagger(\tau_1 + T_0) = y^{\tau_2}$. By integrating (5.10), and using the Lipschitz character of f_0, \dots, f_m , Grönwall's inequality, Cauchy-Schwarz and (5.9), we see that

$$\begin{aligned} \|y^\ddagger(t)\| &\leq C_5 \left(\|\bar{y}\| + \|u^\ddagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + 1 \right) \exp \left(C_5 \|u^\ddagger\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} \right) \\ &\leq C_6 \left(\|\bar{y}\| + \|\bar{y} - y^{\tau_2}\| + 1 \right) \exp \left(C_6 \|\bar{y} - y^{\tau_2}\| \right) \\ &\leq C_6 \left(\|\bar{y}\| + r + 1 \right) \exp \left(C_6 r \right) \end{aligned} \quad (5.11)$$

for some $C_5(f) > 0$ and $C_6(f, T_0) > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$. Then, by integrating (5.10) once again, and using $f_0(\bar{y}) = 0$, Cauchy-Schwarz and (5.11), we moreover see that

$$\begin{aligned} \left\| y^\dagger(t) - \bar{y} \right\| &\leq \int_{\tau_1}^t \sum_{j=1}^m \left| u_j^\dagger(s) \right| \left\| f_j(y^\dagger) \right\| ds + \int_{\tau_1}^t \left\| f(y^\dagger) - f(\bar{y}) \right\| ds \\ &\leq C_7 \left\| u^\dagger \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + C(f) \int_{\tau_1}^t \left\| y^\dagger(s) - \bar{y} \right\| ds \end{aligned} \quad (5.12)$$

for some $C_7(f, T_0, r, \bar{y}) > 0$, with $C(f) > 0$ being the Lipschitz constant of the vector fields f_j . Finally, applying Grönwall's inequality to (5.12) and using (5.9), we deduce that

$$\left\| y^\dagger(t) - \bar{y} \right\| \leq C_8 \exp(C(f)T_0) \|y^{\tau_2} - \bar{y}\| \quad (5.13)$$

for some $C_8(f, T_0, \bar{y}, r) > 0$ independent of $T, \tau_1, \tau_2 > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$. Note that in view of (5.6), both C_7 and C_8 are non-decreasing with respect to the parameter $r > 0$.

Now set

$$u^{\text{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (\tau_1, \tau_1 + T_0) \\ 0 & \text{in } (\tau_1 + T_0, \tau_2 - T_0) \\ u^\dagger(t - (\tau_2 - \tau_1 - T_0)) & \text{in } (\tau_2 - T_0, \tau_2), \end{cases}$$

and let y^{aux} be the corresponding solution to (5.3). By construction, we have

$$y^{\text{aux}}(t) = y^\dagger(t) \quad \text{in } [\tau_1, \tau_1 + T_0],$$

and thus

$$y^{\text{aux}}(t) = \bar{y} \quad \text{in } [\tau_1 + T_0, \tau_2 - T_0], \quad (5.14)$$

whereas we also have $y^{\text{aux}}(\tau_2) = y^{\tau_2}$, whence $u^{\text{aux}} \in U_{\text{ad}}$.

We now evaluate J_{τ_1, τ_2} at u^{aux} , which by virtue of a simple change of variable as well as (5.14), (5.4), (5.8), (5.9) and (5.13), leads us to

$$\begin{aligned} J_{\tau_1, \tau_2}(u^{\text{aux}}) &= \left\| u^\dagger \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}^2 + \left\| u^\dagger \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}^2 \\ &\quad + \int_{\tau_1}^{\tau_1 + T_0} \left\| y^\dagger(t) - \bar{y} \right\|^2 dt + \int_{\tau_1}^{\tau_1 + T_0} \left\| y^\dagger(t) - \bar{y} \right\|^2 dt \\ &\leq C_9 \left(\|\bar{y} - y^{\tau_1}\|^2 + \|\bar{y} - y^{\tau_2}\|^2 \right) \end{aligned} \quad (5.15)$$

where $C_9 = C_9(f, \bar{y}, T_0, r) > 0$ is independent of $T, \tau_1, \tau_2 > 0$, and is non-decreasing with respect to r . Hence $u_T \in U_{\text{ad}}$ is uniformly bounded with respect to $T, \tau_1, \tau_2 > 0$, as in view of (5.15) we have

$$\begin{aligned} \|y_T - \bar{y}\|_{L^2(\tau_1, \tau_2; \mathbb{R}^d)}^2 + \|u_T\|_{L^2(\tau_1, \tau_2; \mathbb{R}^m)}^2 &\leq J_{\tau_1, \tau_2}(u_T) \leq J_{\tau_1, \tau_2}(u^{\text{aux}}) \\ &\leq C_9 \left(\|\bar{y} - y^{\tau_1}\|^2 + \|\bar{y} - y^{\tau_2}\|^2 \right). \end{aligned}$$

An application of Lemma 4.1 combined with the uniform boundedness of $\|u_T\|_{L^2(\tau_1, \tau_2; \mathbb{R}^m)}$ with respect to $T, \tau_2, \tau_1 > 0$ suffices to conclude. \square

Before proceeding with the proof of Theorem 2.1, we will need the following key lemma.

Lemma 5.3. *Let X be a Banach space, $T > 0$ and $f \in C^0([0, T]; X)$. For any $\tau \leq \frac{T}{2}$, there exist $t_1 \in [0, \tau)$ and $t_2 \in (T - \tau, T]$ such that*

$$\|f(t_i)\|_X \leq \frac{\|f\|_{L^2(0, T; X)}}{\sqrt{\tau}} \quad \text{for } i = 1, 2.$$

Proof of Lemma 5.3. Denote

$$\eta(\tau) := \frac{\|f\|_{L^2(0, T; X)}}{\sqrt{\tau}}.$$

We argue by contradiction. Assume that either

$$\|f(t)\|_X > \eta(\tau) \quad \text{for all } t \in [0, \tau)$$

or

$$\|f(t)\|_X > \eta(\tau) \quad \text{for all } t \in (T - \tau, T].$$

hold. Then we have

$$\int_0^T \|f(t)\|_X^2 dt \geq \int_0^\tau \|f(t)\|_X^2 dt + \int_{T-\tau}^T \|f(t)\|_X^2 dt > \tau \eta(\tau)^2.$$

Hence

$$\eta(\tau)^2 < \frac{1}{\tau} \int_0^T \|f(t)\|_X^2 dt = \eta(\tau)^2,$$

which yields a contradiction. This concludes the proof. \square

5.2. Proof of Theorem 2.1. We are now in a position to prove our first main result.

Proof of Theorem 2.1. We begin by noting that (2.9) follows immediately from Lemma 5.1. We thus concentrate on proving (2.8) – we split the proof in two parts.

Before proceeding, let us first note that by Lemma 5.1, there exists a constant $C_1(f, T_0, \bar{y}, y^0) > 0$ such that whenever $T \geq 2T_0$,

$$\|y_T(t) - \bar{y}\| \leq C_1 \quad \text{for all } t \in [0, T]. \quad (5.16)$$

Let $r > 0$ be the radius provided by Assumption 1. By Lemma 5.2, we know that there exists a constant $C_2(f, T_0, \bar{y}, r) > 0$ such that for any $\tau_1, \tau_2 \in [0, T]$ such that $\tau_2 - \tau_1 \geq 2T_0$ and

$$\|y_T(\tau_i) - \bar{y}\| \leq r$$

for $i = 1, 2$, the estimate

$$\|y_T(t) - \bar{y}\| \leq C_2 \left(\|y_T(\tau_2) - \bar{y}\| + \|y_T(\tau_1) - \bar{y}\| \right) \quad \text{for all } t \in [\tau_1, \tau_2]$$

holds. Now let

$$\tau > 16C_2^4 + \frac{C_1^2}{r^2} + \frac{4C_1^2 C_2^2}{r^2} + T_0 \quad (5.17)$$

and let

$$T \geq 2\tau + 2T_0$$

be fixed. The choice of the "buffer" time τ will become clear in what follows (in fact, it will also be seen that $T^* := \frac{\tau + T_0}{2}$ in the statement of the theorem).

Part 1: We note that for $t \in [0, \tau + T_0]$ and $t \in [T - (\tau + T_0), T]$, the desired estimate (2.8) can be obtained without too much difficulty, as the length of both time intervals is independent of T . Indeed, by (5.16), for any $\mu > 0$ we have

$$\begin{aligned} \|y_T(t) - \bar{y}\| &\leq C_1 = C_1 e^{\mu t} e^{-\mu t} \\ &\leq C_1 e^{\mu(\tau+T_0)} \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \end{aligned} \quad (5.18)$$

for $t \in [0, \tau + T_0]$, and

$$\begin{aligned} \|y_T(t) - \bar{y}\| &\leq C_1 = C_1 e^{\mu(T-t)} e^{-\mu(T-t)} \\ &\leq C_1 e^{\mu(\tau+T_0)} \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \end{aligned} \quad (5.19)$$

for $t \in [T - (\tau + T_0), T]$.

Thus, it only remains to be seen what happens when $t \in [\tau + T_0, T - (\tau + T_0)]$. We will address this case by means of a bootstrap argument in Part 2 just below.

Part 2: We now aim to show (2.8) for $t \in [\tau + T_0, T - (\tau + T_0)]$. To this end, we proceed in three steps.

Step 1). **Preparation.** Since $\tau \leq \frac{T}{2}$, by Lemma 5.3 and Lemma 5.1, there exist a couple of time instances $\tau_1 \in [0, \tau]$ and $\tau_2 \in [T - \tau, T]$ such that

$$\|y_T(\tau_i) - \bar{y}\| \leq \frac{\|y_T - \bar{y}\|_{L^2(0, T; \mathbb{R}^d)}}{\sqrt{\tau}} \leq \frac{C_1}{\sqrt{\tau}}. \quad (5.20)$$

Note that, by virtue of the choice of τ in (5.17), we have that $\frac{C_1}{\sqrt{\tau}} \leq r$ and thus

$$\|y_T(\tau_i) - \bar{y}\| \leq r \quad (5.21)$$

also holds. We shall now restrict our analysis onto $[\tau_1, \tau_2]$, and extrapolate onto the subset $[\tau, T - \tau]$. First note that $u_T|_{[\tau_1, \tau_2]}$ is a global minimizer² of J_{τ_1, τ_2} defined in (5.2) with fixed endpoints $y^{\tau_1} = y_T(\tau_1)$ and $y^{\tau_2} = y_T(\tau_2)$, and thus clearly $y_T|_{[\tau_1, \tau_2]}$ solves (5.3). As

$$\tau_2 - \tau_1 \geq T - 2\tau \geq 2T_0,$$

in view of (5.21), we may use Lemma 5.2 with the effect of deducing that

$$\|y_T(t) - \bar{y}\| \leq C_2 \left(\|y_T(\tau_1) - \bar{y}\| + \|y_T(\tau_2) - \bar{y}\| \right) \quad (5.22)$$

holds for all $t \in [\tau_1, \tau_2]$. Setting

$$\kappa := \max \left\{ 1, \frac{C_1}{C_2} \right\},$$

and applying (5.20) to inequality (5.22), we deduce that

$$\|y_T(t) - \bar{y}\| \leq \kappa \frac{2C_2^2}{\sqrt{\tau}} \quad (5.23)$$

holds for all $t \in [\tau_1, \tau_2]$. As $\tau_1 \leq \tau$ and $T - \tau \leq \tau_2$, estimate (5.23) clearly holds for all $t \in [\tau, T - \tau]$.

²This can be shown by contradiction.

Step 2). **Bootstrap.** Inequality (5.23) motivates performing a bootstrap – we will show that for any $n \in \mathbb{N}$ satisfying

$$n \leq \frac{1}{\tau} \left(\frac{T}{2} - T_0 \right),$$

one has

$$\|y_T(t) - \bar{y}\| \leq \frac{\kappa}{2} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n \quad \text{for } t \in [n\tau, T - n\tau]. \quad (5.24)$$

The choice of n is done as to guarantee that $T - 2n\tau \geq 2T_0$ in view of a repeated application of Lemma 5.2. Note that (5.22), combined with the choice of τ in (5.17), also implies that

$$\|y_T(t) - \bar{y}\| \leq r \quad (5.25)$$

for all $t \in [\tau, T - \tau]$.

To prove (5.24), we proceed by induction. The case $n = 1$ clearly holds by (5.23). Thus, assume that (5.24) holds – we aim to show that (5.24) holds at step $n + 1$. To this end, let

$$n + 1 \leq \frac{1}{\tau} \left(\frac{T}{2} - T_0 \right).$$

This clearly implies that

$$\tau \leq \frac{T - 2n\tau}{2}. \quad (5.26)$$

As in Step 1, since $T - 2n\tau \geq 2T_0$, it can be seen that $u_T|_{[n\tau, T - n\tau]}$ is a global minimizer of $J_{n\tau, T - n\tau}$ defined in (5.2). Taking these facts into account, and noting that (5.25) holds³, we can apply Lemma 5.3 on $[n\tau, T - n\tau]$ (note (5.26)), and Lemma 5.2 with $\tau_1 = n\tau$ and $\tau_2 = T - n\tau$, to deduce that there exist a couple of times $t_1 \in [n\tau, (n + 1)\tau]$ and $t_2 \in (T - (n + 1)\tau, T - n\tau]$ such that

$$\begin{aligned} \|y_T(t_i) - \bar{y}\| &\leq \frac{\|y_T - \bar{y}\|_{L^2(n\tau, T - n\tau; \mathbb{R}^d)}}{\sqrt{\tau}} \\ &\leq \frac{C_2}{\sqrt{\tau}} \left(\|y_T(n\tau) - \bar{y}\| + \|y_T(T - n\tau) - \bar{y}\| \right) \end{aligned}$$

We now use the induction hypothesis (5.24) to obtain

$$\|y_T(t_i) - \bar{y}\| \leq \kappa \frac{2C_2}{\sqrt{\tau}} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n \quad (5.27)$$

Now since

$$t_2 - t_1 \geq T - 2(n + 1)\tau \geq 2T_0,$$

³Note that $n\tau \geq \tau$ and $T - n\tau \leq T - \tau$, so (5.25) also holds for $t \in [n\tau, T - n\tau]$, hence Lemma 5.2 is applicable.

and since $u_T|_{[t_1, t_2]}$ is a global minimizer of J_{t_1, t_2} defined in (5.2), combining Lemma 5.2⁴ and (5.27) we are led to deduce that

$$\begin{aligned} \|y_T(t) - \bar{y}\| &\leq C_2 \left(\|y_T(t_1) - \bar{y}\| + \|y_T(t_2) - \bar{y}\| \right) \\ &\leq \frac{\kappa}{2} \frac{4C_2^2}{\sqrt{\tau}} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n \end{aligned} \quad (5.28)$$

for $t \in [t_1, t_2]$. Since $t_1 < (n+1)\tau$ and $T - (n+1)\tau < t_2$, estimate (5.28) clearly also holds for $t \in [(n+1)\tau, T - (n+1)\tau]$. Identity (5.24) is thus proven.

Step 3). **Conclusion.** We now look to use (5.24) as to conclude the proof. Suppose that $t \in [\tau + T_0, T - (\tau + T_0)]$. We set

$$n(t) := \min \left\{ \left\lfloor \frac{t}{\tau + T_0} \right\rfloor, \left\lfloor \frac{T - t}{\tau + T_0} \right\rfloor \right\},$$

where $\lfloor z \rfloor$ denotes the integer part of $z \in \mathbb{R}$. Clearly $n(t) \geq 1$ and

$$n(t)\tau \leq t \leq T - n(t)\tau.$$

Moreover, since $z \mapsto \frac{z-2T_0}{z}$ is non-decreasing,

$$\begin{aligned} n(t) &\leq \frac{T}{2(\tau + T_0)} = \frac{T}{2\tau} \frac{2(\tau + T_0) - 2T_0}{2(\tau + T_0)} \\ &\leq \frac{T}{2\tau} \frac{T - 2T_0}{T} \\ &= \frac{1}{\tau} \left(\frac{T}{2} - T_0 \right). \end{aligned}$$

We may then apply (5.24) to obtain

$$\|y_T(t) - \bar{y}\| \leq \frac{\kappa}{2} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^{n(t)}. \quad (5.29)$$

As $\tau > 16C_2^2$, we see that $\frac{4C_2^2}{\sqrt{\tau}} < 1$. Moreover, since either $n(t) \geq \left\lfloor \frac{t}{\tau + T_0} \right\rfloor - 1$ or $n(t) \geq \left\lfloor \frac{T-t}{\tau + T_0} \right\rfloor - 1$ holds, we may rewrite (5.29) to obtain

$$\begin{aligned} \|y_T(t) - \bar{y}\| &\leq \frac{\kappa}{2} \exp \left(-n(t) \log \left(\frac{\sqrt{\tau}}{4C_2^2} \right) \right) \\ &\leq \frac{\kappa}{2} \frac{\sqrt{\tau}}{4C_2^2} \left(\exp \left(-\frac{\log \left(\frac{\sqrt{\tau}}{4C_2^2} \right)}{\tau + T_0} t \right) + \exp \left(-\frac{\log \left(\frac{\sqrt{\tau}}{4C_2^2} \right)}{\tau + T_0} (T - t) \right) \right). \end{aligned} \quad (5.30)$$

Whence, (2.8) holds for all $t \in [\tau + T_0, T - (\tau + T_0)]$, with

$$C := \frac{\kappa}{2} \frac{\sqrt{\tau}}{4C_2^2}$$

and

$$\mu := \frac{\log \left(\frac{\sqrt{\tau}}{4C_2^2} \right)}{\tau + T_0} > 0. \quad (5.31)$$

⁴May be applied once again since (5.25) holds for $t = t_1 \geq \tau$ and $t = t_2 \leq T - \tau$.

By virtue of (5.18), (5.19) and (5.30), we deduce that (2.8) holds for all $t \in [0, T]$, with $T^* := \tau + T_0$,

$$C := \max \left\{ C_1, \frac{\kappa}{2} \frac{\sqrt{\tau}}{4C_2^2} \right\}$$

and μ as in (5.31). This concludes the proof. \square

5.3. Proof of Corollary 2.1. Let us now provide a proof to Corollary 2.1.

Proof of Corollary 2.1. By Theorem 2.1, with $\phi \equiv 0$, there exist constants $C_1 > 0$ and $\mu_1 > 0$ such that

$$\|y_T(t) - \bar{y}\| \leq C_1 \left(e^{-\mu_1 t} + e^{-\mu_1(T-t)} \right)$$

holds for all $t \in [0, T]$. We now distinguish two cases.

- If $t \in [0, \frac{T}{2}]$, we also have

$$\|y_T(t) - \bar{y}\| \leq C_1 \left(e^{-\mu_1 t} + e^{-\mu_1(T-t)} \right) \leq 2C_1 e^{-\mu_1 \frac{t}{2}}. \quad (5.32)$$

The desired estimates thus holds in this case.

- We now consider the case $t \in [\frac{T}{2}, T]$. First set

$$u^{\text{aux}}(t) := \begin{cases} u_T(t) & \text{for } t \in \left[0, \frac{T}{2}\right] \\ 0 & \text{for } t \in \left[\frac{T}{2}, T\right]. \end{cases}$$

The state y^{aux} , solution to (2.1) with $y^{\text{aux}}(0) = y^0$ associated to u^{aux} is precisely

$$y^{\text{aux}}(t) = \begin{cases} y_T(t) & \text{for } t \in \left[0, \frac{T}{2}\right] \\ y_T\left(\frac{T}{2}\right) & \text{for } t \in \left[\frac{T}{2}, T\right]. \end{cases}$$

Using the inequality $J_T(u_T) \leq J_T(u^{\text{aux}})$ along with (5.32) gives

$$\begin{aligned} \int_{\frac{T}{2}}^T \|u_T(t)\|^2 dt + \int_{\frac{T}{2}}^T \|y_T(t) - \bar{y}\|^2 dt &\leq \int_{\frac{T}{2}}^T \left\| y_T\left(\frac{T}{2}\right) - \bar{y} \right\|^2 dt \\ &\leq \int_{\frac{T}{2}}^T 4C_1^2 \exp(-\mu_1 T) dt \\ &\leq 2C_1^2 T \exp(-\mu_1 T) \\ &\leq C_2 \exp\left(-\mu_1 \frac{T}{4}\right), \end{aligned} \quad (5.33)$$

for some $C_2 > 0$ independent of $T > 0$.

Now by Lemma 4.1, for any $t \in [\frac{T}{2}, T]$,

$$\|y_T(t) - \bar{y}\| \leq C_3 \left(\left\| y_T\left(\frac{T}{2}\right) - \bar{y} \right\| + \|u_T\|_{L^2(\frac{T}{2}, T; \mathbb{R}^m)} + \|y_T - \bar{y}\|_{L^2(\frac{T}{2}, T; \mathbb{R}^d)} \right)$$

for some $C_3 > 0$ of the form

$$C_3 \sim \exp\left(\|u_T\|_{L^2(\frac{T}{2}, T; \mathbb{R}^m)}\right).$$

In view of (2.9), $\|u_T\|_{L^2(\frac{T}{2}, T; \mathbb{R}^m)}$ is bounded uniformly with respect to T , and thus $C_3 > 0$ is independent of T . Combining the above estimate with (5.33) leads us to

$$\|y_T(t) - \bar{y}\| \leq C_4 \exp\left(-\mu_1 \frac{T}{4}\right) \leq C_4 \exp\left(-\mu_1 \frac{t}{4}\right), \quad (5.34)$$

for some $C_4 > 0$ independent of $T > 0$ and for all $t \in [\frac{T}{2}, T]$.

Combining (5.32) and (5.34), we see that for $C_5 := \max\{2C_1, C_4\}$ and $\mu := \frac{\mu_1}{4}$, the stabilization estimate

$$\|y_T(t) - \bar{y}\| \leq C_5 e^{-\mu t},$$

for all $t \in [0, T]$. This concludes the proof. \square

5.4. Proof of Corollary 2.2. We finish this section with the proof of Corollary 2.2, which stipulates an exponential decay of optimal controls in the context of driftless control-affine systems, namely (2.1) with a nonlinearity of the form

$$f(y, u) = \sum_{j=1}^m u_j f_j(y) \quad \text{for } (y, u) \in \mathbb{R}^d \times \mathbb{R}^m. \quad (5.35)$$

We recall that, here, $f_1, \dots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$.

We begin with the following simple result.

Lemma 5.4. *Let $T_0 > 0$, $y^0 \in \mathbb{R}^d$ and $u_{T_0} \in L^2(0, T_0; \mathbb{R}^m)$ be given, and let $y_{T_0} \in C^0([0, T_0]; \mathbb{R}^d)$ be the unique solution to*

$$\begin{cases} \dot{y}_{T_0} = f(y_{T_0}, u_{T_0}) & \text{in } (0, T_0) \\ y_{T_0}(0) = y^0 \end{cases} \quad (5.36)$$

with f as in (5.35). Let $T > 0$, and define

$$u_T(t) := \frac{T_0}{T} u_{T_0}\left(t \frac{T_0}{T}\right) \quad \text{for } t \in [0, T],$$

and

$$y_T(t) := y_{T_0}\left(t \frac{T_0}{T}\right) \quad \text{for } t \in [0, T].$$

Then $y_T \in C^0([0, T]; \mathbb{R}^d)$ is the unique solution to (2.1) with $y_T(0) = y^0$ associated to u_T .

This sort of time-scaling in the context of driftless control affine systems is commonly used in control theoretical contexts – a canonical example is the proof of the Chow-Rashevskii controllability theorem, see [6, Chapter 3, Section 3.3]. We provide the short proof for completeness.

Proof of Lemma 5.4. Using the fact that y_{T_0} is the solution to (7.8) and the change of variable $\tau = s \frac{T}{T_0}$, we see that

$$\begin{aligned} y_T(t) &:= y_{T_0} \left(t \frac{T_0}{T} \right) = y^0 + \int_0^{t \frac{T_0}{T}} f(y_{T_0}(s), u_{T_0}(s)) \, ds \\ &= y^0 + \int_0^t \frac{T_0}{T} f \left(y_{T_0} \left(\tau \frac{T_0}{T} \right), u_{T_0} \left(\tau \frac{T_0}{T} \right) \right) \, d\tau \\ &= y^0 + \int_0^t f(y_T(\tau), u_T(\tau)) \, d\tau. \end{aligned}$$

It follows that y_T solves (2.1) with $y_T(0) = y^0$, and we conclude by uniqueness. \square

Proof of Corollary 2.2. As $t \mapsto \|u_T(t)\|^2$ is in $L^1(0, T)$, by the Lebesgue differentiation theorem, we have

$$\|u_T(t)\|^2 = \lim_{h \searrow 0^+} \frac{1}{h} \int_t^{t+h} \|u_T(s)\|^2 \, ds$$

for almost every $t \in (0, T)$. Hence, we will aim at estimating the integral on the right hand side by constructing an appropriate auxiliary suboptimal auxiliary control, and conclude by passing to the limit as $h \searrow 0^+$. We will split the proof in two parts, namely separate the proof of (2.11) (i.e. $\phi \not\equiv 0$) and (2.12) (i.e. $\phi \equiv 0$); they mainly differ in the construction of the suboptimal auxiliary control required very last estimate (5.45) before concluding.

Part 1: Proof of (2.11). Fix $t \in [0, T)$ and $0 < h \ll 1$ so that $t + 2h^2 + 2h \in [0, T]$. Let us set

$$u^{\text{aux}}(s) := \begin{cases} u_T(s) & \text{for } s \in [0, t] \\ \frac{1}{2} u_T \left(t + \frac{s-t}{2} \right) & \text{for } s \in (t, t + 2h^2] \\ \frac{h+2}{2} u_T \left(\left(\frac{h+2}{2} \right) s - \frac{h+2}{2} (t + 2h^2) + t + h^2 \right) & \text{for } t \in (t + 2h^2, t + 2h^2 + 2h] \\ u_T(s) & \text{for } s \in (t + 2h^2 + 2h, T]. \end{cases}$$

The specific choice of u^{aux} will become clear in what follows – the factor h in the third line will be essential in the subsequent estimates. By Lemma 5.4, the state y^{aux} , solution to (2.1) associated to u^{aux} is precisely

$$y^{\text{aux}}(s) = \begin{cases} y_T(s) & \text{for } s \in [0, t] \\ y_T \left(t + \frac{s-t}{2} \right) & \text{for } s \in (t, t + 2h^2] \\ y_T \left(\left(\frac{h+2}{2} \right) s - \frac{h+2}{2} (t + 2h^2) + t + h^2 \right) & \text{for } t \in (t + 2h^2, t + 2h^2 + 2h] \\ y_T(s) & \text{for } t \in (t + 2h^2 + 2h, T]. \end{cases}$$

Note in particular that $y^{\text{aux}}(T) = y_T(T)$, whence $\phi(y^{\text{aux}}(T)) = \phi(y_T(T))$. Our goal is then to rewrite the simple inequality $J_T(u_T) \leq J_T(u^{\text{aux}})$ as to estimate the infinitesimal average of $\|u_T(t)\|^2$ taken about t by the infinitesimal average of $\|y_T(t) - \bar{y}\|^2$ taken about t , for which we have an exponential estimate by Theorem 2.1.

We proceed as follows.

- On one hand, using the change of variable $\tau := t + \frac{s-t}{2}$ we see that

$$\begin{aligned} \int_t^{t+2h^2} \|u^{\text{aux}}(s)\|^2 ds &= \int_t^{t+2h^2} \left\| \frac{1}{2} u_T \left(t + \frac{s-t}{2} \right) \right\|^2 ds \\ &= \frac{1}{2} \int_t^{t+h^2} \|u_T(\tau)\|^2 d\tau. \end{aligned} \quad (5.37)$$

On another hand, via $\tau := \left(\frac{h+2}{2}\right)s - \frac{h+2}{2}(t+2h^2) + t + h^2$ we see that

$$\begin{aligned} &\int_{t+2h^2}^{t+2h^2+2h} \|u^{\text{aux}}(s)\|^2 ds \\ &= \int_{t+2h^2}^{t+2h^2+2h} \left\| \frac{h+2}{2} u_T \left(\left(\frac{h+2}{2}\right)s - \frac{h+2}{2}(t+2h^2) + t + h^2 \right) \right\|^2 ds \\ &= \frac{h+2}{2} \int_{t+h^2}^{t+h^2+2h} \|u_T(\tau)\|^2 d\tau. \end{aligned} \quad (5.38)$$

Combining (5.37) and (5.38) and since $0 < h \ll 1$, it follows that

$$\begin{aligned} \int_0^T \|u^{\text{aux}}(s)\|^2 ds &= \int_0^t \|u^{\text{aux}}(s)\|^2 ds + \int_t^{t+2h^2} \|u^{\text{aux}}(s)\|^2 ds \\ &\quad + \int_{t+2h^2}^{t+2h^2+2h} \|u^{\text{aux}}(s)\|^2 ds + \int_{t+2h^2+2h}^T \|u^{\text{aux}}(s)\|^2 ds \\ &= \int_0^t \|u_T(s)\|^2 ds + \frac{1}{2} \int_t^{t+h^2} \|u_T(\tau)\|^2 d\tau \\ &\quad + \frac{h+2}{2} \int_{t+h^2}^{t+h^2+2h} \|u_T(\tau)\|^2 d\tau + \int_{t+2h^2+2h}^T \|u_T(\tau)\|^2 d\tau \\ &\leq \int_0^T \|u_T(s)\|^2 ds - \frac{1}{2} \int_t^{t+h^2} \|u_T(s)\|^2 ds \\ &\quad + \frac{h}{2} \int_{t+h^2}^{t+h^2+2h} \|u_T(\tau)\|^2 d\tau. \end{aligned} \quad (5.39)$$

- We now focus on rewriting the state tracking term. On one hand, by means of $\tau := t + \frac{s-t}{2}$ we see that

$$\begin{aligned} \int_t^{t+2h^2} \|y^{\text{aux}}(s) - y\|^2 ds &= \int_t^{t+2h^2} \left\| y_T \left(t + \frac{s-t}{2} \right) - \bar{y} \right\|^2 ds \\ &= 2 \int_t^{t+h^2} \|y_T(\tau) - \bar{y}\|^2 d\tau. \end{aligned} \quad (5.40)$$

On another hand, via $\tau := \left(\frac{h+2}{2}\right)s - \frac{h+2}{2}(t+2h^2) + t+h^2$ we see that

$$\begin{aligned}
& \int_{t+2h^2}^{t+2h^2+2h} \|y^{\text{aux}}(s) - \bar{y}\|^2 ds \\
&= \int_{t+2h^2}^{t+2h^2+2h} \left\| y_T \left(\left(\frac{h+2}{2}\right)s - \frac{h+2}{2}(t+2h^2) + t+h^2 \right) - \bar{y} \right\|^2 ds \\
&= \frac{2}{h+2} \int_{t+h^2}^{t+h^2+2h} \|y_T(\tau) - \bar{y}\|^2 d\tau. \tag{5.41}
\end{aligned}$$

Combining (5.40) and (5.41) and since $0 < h \ll 1$, we obtain

$$\begin{aligned}
\int_0^T \|y^{\text{aux}}(s) - \bar{y}\|^2 ds &= \int_0^t \|y^{\text{aux}}(s) - \bar{y}\|^2 ds + \int_t^{t+2h^2} \|y^{\text{aux}}(s) - \bar{y}\|^2 ds \\
&\quad + \int_{t+2h^2}^{t+2h^2+2h} \|y^{\text{aux}}(s) - \bar{y}\|^2 ds + \int_{t+2h^2+2h}^T \|y^{\text{aux}}(s) - \bar{y}\|^2 ds \\
&= \int_0^t \|y_T(s) - \bar{y}\|^2 ds + 2 \int_t^{t+h^2} \|y_T(\tau) - \bar{y}\|^2 d\tau \\
&\quad + \frac{2}{h+2} \int_{t+h^2}^{t+h^2+2h^2} \|y_T(\tau) - \bar{y}\|^2 d\tau + \int_{t+2h^2+2h}^T \|y_T(\tau) - \bar{y}\|^2 d\tau \\
&\leq \int_0^T \|y_T(s) - \bar{y}\|^2 ds + \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 ds. \tag{5.42}
\end{aligned}$$

We may now proceed with the main argument. Using the optimality of u_T , and applying (5.39) and (5.42), we see that

$$\begin{aligned}
J_T(u_T) &\leq J_T(u^{\text{aux}}) = \phi(y^{\text{aux}}(T)) + \int_0^T \|y^{\text{aux}}(s) - \bar{y}\|^2 ds + \int_0^T \|u^{\text{aux}}(s)\|^2 ds \\
&= \phi(y_T(T)) + \int_0^T \|y_T(s) - \bar{y}\|^2 ds + \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 ds \\
&\quad + \int_0^T \|u_T(s)\|^2 ds - \frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 ds \\
&\quad + \frac{h}{2} \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds. \tag{5.43}
\end{aligned}$$

From (5.43), one clearly sees that

$$\frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 ds \leq \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 ds + \frac{h}{2} \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds. \tag{5.44}$$

We combine estimate (5.44) with (2.8) to deduce that

$$\begin{aligned}
\frac{1}{h} \int_t^{t+h} \|u_T(s)\|^2 ds &\lesssim \frac{1}{h} \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 ds + \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds \\
&\leq \frac{C}{h} \int_t^{t+h} \left(e^{-\mu s} + e^{-\mu(T-s)} \right)^2 ds + \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds \\
&\leq \frac{C}{h} \int_t^{t+h} \left(e^{-\mu t} + e^{-\mu(T-t-h)} \right)^2 ds + \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds \\
&= C \left(e^{-\mu t} + e^{-\mu(T-t-h)} \right)^2 + \int_{t+h^2}^{t+h^2+2h} \|u_T(s)\|^2 ds \quad (5.45)
\end{aligned}$$

for some $C > 0$ independent of T . Thus, by using the Lebesgue differentiation theorem and the Lebesgue dominated convergence theorem (applied to the integrable function $s \mapsto \|u_T(s)\|^2 \mathbf{1}_{(t+h^2, t+h^2+2h)}(s)$) in (5.45), we deduce that

$$\|u_T(t)\| = \lim_{h \searrow 0^+} \left(\frac{1}{h} \int_t^{t+h} \|u_T(s)\|^2 ds \right)^{1/2} \leq C \left(e^{-\mu t} + e^{-\mu(T-t)} \right),$$

as desired.

Part 2: Proof of (2.12). This part is somewhat simpler due to the fact that $\phi \equiv 0$. Hence the suboptimal control has a much simpler structure as we do not require a factor of h to deal with a remainder term as in (5.43). From there on, the arguments are identical to those above.

Fix any $t \in [0, T]$ and $0 < h \ll 1$, so that $t + 2h \in [0, T]$ and set

$$u^{\text{aux}}(s) := \begin{cases} u_T(s) & \text{for } s \in [0, t] \\ \frac{1}{2} u_T \left(t + \frac{s-t}{2} \right) & \text{for } s \in (t, t+2h] \\ u_T(s-h) & \text{for } s \in (t+2h, T]. \end{cases}$$

By Lemma 5.4, the state y^{aux} , solution to (2.1) associated to u^{aux} is precisely

$$y^{\text{aux}}(s) = \begin{cases} y_T(s) & \text{for } s \in [0, t] \\ y_T \left(t + \frac{s-t}{2} \right) & \text{for } s \in (t, t+2h] \\ y_T(s-h) & \text{for } s \in (t+2h, T]. \end{cases}$$

Arguing by means of simple changes of variable just as we did for obtaining (5.39) and (5.42), and using the suboptimality of u^{aux} , we can readily see that

$$\begin{aligned}
J_T(u_T) &\leq J_T(u^{\text{aux}}) = \int_0^T \|u^{\text{aux}}(s)\|^2 \, ds + \int_0^T \|y^{\text{aux}}(s) - \bar{y}\|^2 \, ds \\
&= \int_0^{T-h} \|u_T(s)\|^2 \, ds - \frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 \, ds \\
&\quad + \int_0^{T-h} \|y_T(s) - \bar{y}\|^2 \, ds + \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 \, ds \\
&\leq \int_0^T \|u_T(s)\|^2 \, ds - \frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 \, ds \\
&\quad + \int_0^T \|y_T(s) - \bar{y}\|^2 \, ds + \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 \, ds. \tag{5.46}
\end{aligned}$$

From (5.46), one clearly sees that

$$\frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 \, ds \leq \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 \, ds. \tag{5.47}$$

We combine estimate (5.47) with (2.12) to deduce that

$$\begin{aligned}
\frac{1}{h} \int_t^{t+h} \|u_T(s)\|^2 \, ds &\lesssim \frac{1}{h} \int_t^{t+h} \|y_T(s) - \bar{y}\|^2 \, ds \\
&\leq \frac{C}{h} \int_t^{t+h} e^{-2\mu s} \, ds \\
&\leq \frac{C}{h} \int_t^{t+h} e^{-2\mu t} \, ds \\
&= C e^{-2\mu t}. \tag{5.48}
\end{aligned}$$

Thus by the Lebesgue differentiation theorem, using (5.48) we deduce that

$$\|u_T(t)\| = \lim_{h \searrow 0^+} \left(\frac{1}{h} \int_t^{t+h} \|u_T(s)\|^2 \, ds \right)^{1/2} \leq C e^{-\mu t},$$

as desired. This concludes the proof. \square

6. PROOF OF THEOREM 3.1

In this section, we provide details of the proof of Theorem 3.1. The proof of Corollary 3.1 follows by repeating the proof of Corollary 2.1 in the appropriate functional setting, so we omit it.

Proof of Theorem 3.1. Once (3.1) is written as a first order evolution equation set in $X := H_0^1(\Omega) \times L^2(\Omega)$ (see the proof of Lemma 4.3 for this setup), the only noticeable difference in the proof of Theorem 3.1 with respect to the proof of Theorem 2.1 are the specific quasi-turnpike lemmas one applies in the preparation (Lemma 6.1 in Part 1 & Step 1 of Part 2) and bootstrap (Lemma 6.2 in Step 2). So one simply repeats the proof of Theorem 2.1 whilst applying Lemma 6.1, Lemma 6.2 and Lemma 5.3 with X as above. Whence, the proof follows from these two lemmas, stated and proven just below. \square

Lemma 6.1. *Let $\mathbf{y}^0 = (y_1^0, y_2^0) \in H_0^1(\Omega) \times L^2(\Omega)$ be given. Let $T > 0$ be fixed, and let $u_T \in L^2((0, T) \times \omega)$ be a global minimizer to J_T defined in (3.2), with y_T denoting the associated solution to (3.1). Then, there exists a constant $C = C(f, \phi, \omega, \Omega, \bar{y}, \mathbf{y}^0) > 0$ independent of $T > 0$ such that*

$$J_T(u_T) + \|y_T(t) - \bar{y}\|_{H_0^1(\Omega)}^2 + \|\partial_t y_T(t)\|_{L^2(\Omega)}^2 \leq C$$

holds for all $t \in [0, T]$.

Proof of Lemma 6.1. The proof follows the lines of that of Lemma 5.1, simply adapted to the PDE setting. Fix $T_0 \geq T_{\min}$ where $T_{\min} = T_{\min}(\omega, \Omega) > 0$ is the minimal controllability time for the semilinear wave equation.

We begin by considering the case $T > T_0$. By controllability, we know that exists some control $u^\dagger \in L^2((0, T_0) \times \omega)$ such that the corresponding solution y^\dagger to

$$\begin{cases} \partial_t^2 y^\dagger - \Delta y^\dagger + f(y^\dagger) = u^\dagger \mathbf{1}_\omega & \text{in } (0, T_0) \times \Omega \\ y^\dagger = 0 & \text{on } (0, T_0) \times \partial\Omega \\ (y^\dagger, \partial_t y^\dagger)|_{t=0} = \mathbf{y}^0 & \text{in } \Omega. \end{cases}$$

satisfies $y^\dagger(T_0) = \bar{y}$ and $\partial_t y^\dagger(T_0) = 0$ a.e. in Ω . Now set

$$u^{\text{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (0, T_0) \\ 0 & \text{in } (T_0, T) \end{cases}$$

and let y^{aux} be the corresponding solution to (3.1). Clearly

$$y^{\text{aux}}(t) = \bar{y} \quad \text{and} \quad \partial_t y^{\text{aux}}(t) = 0 \quad \text{for } t \in [T_0, T] \text{ a.e. in } \Omega.$$

Combining this fact with $J_T(u_T) \leq J_T(u^{\text{aux}})$, we see that

$$J_T(u_T) \leq \phi(\bar{y}) + \|y^\dagger - \bar{y}\|_{L^2(0, T_0; H_0^1(\Omega))}^2 + \|\partial_t y^\dagger\|_{L^2((0, T) \times \Omega)}^2 + \|u^\dagger\|_{L^2((0, T_0) \times \omega)}^2.$$

As the right hand side in the above inequality is clearly independent of T , we conclude by applying Lemma 4.3.

Now suppose that $T \leq T_0$. We use the optimality inequality $J_T(u_T) \leq J_T(u_{T_0+1})$ to obtain

$$\begin{aligned} J_T(u_T) &\leq \phi(y_{T_0+1}(T)) + \|y_{T_0+1} - \bar{y}\|_{L^2(0, T; H_0^1(\Omega))}^2 \\ &\quad + \|\partial_t y_{T_0+1}\|_{L^2(0, T; L^2(\Omega))}^2 + \|u_{T_0+1}\|_{L^2((0, T) \times \omega)}^2. \end{aligned}$$

By the previous case addressed just above, the trajectory $y_{T_0+1} \in C^0([0, T_0+1]; L^2(\Omega))$ is bounded uniformly with respect to T . Hence, using the fact that $\phi \in C^0(L^2(\Omega); \mathbb{R}_+)$ and $T \leq T_0$, we deduce that

$$J_T(u_T) \leq C \tag{6.1}$$

for some $C > 0$ independent of T . Combining (6.1) with Lemma 4.3 allows us to conclude. \square

We note that since $f \in \text{Lip}(\mathbb{R})$, following the spirit of our finite-dimensional arguments, since (3.1) is a Lipschitz perturbation of an exactly controllable linear system, the following claim holds.

Claim 6.1 (Cost estimate). *Let $T_0 > T_{\min}$, where $T_{\min} = T_{\min}(\Omega, \omega) > 0$ is the minimal controllability time for (3.1). There exists $r > 0$ and $C = C(T_0, \omega, f) > 0$ such that*

$$\inf_{\substack{u \\ \text{such that} \\ (y, \partial_t y)|_{t=0} = \mathbf{y}^0 \\ \text{and} \\ (y, \partial_t y)|_{t=T_0} = (\bar{y}, 0)}} \|u\|_{L^2((0, T_0) \times \omega)}^2 \leq C \left(\|y_1^0 - \bar{y}\|_{H_0^1(\Omega)}^2 + \|y_2^0\|_{L^2(\Omega)}^2 \right),$$

and

$$\inf_{\substack{u \\ \text{such that} \\ (y, \partial_t y)|_{t=0} = (\bar{y}, 0) \\ \text{and} \\ (y, \partial_t y)|_{t=T_0} = \mathbf{y}^1}} \|u\|_{L^2((0, T_0) \times \omega)}^2 \leq C \left(\|y_1^1 - \bar{y}\|_{H_0^1(\Omega)}^2 + \|y_2^1\|_{L^2(\Omega)}^2 \right),$$

hold for any $\mathbf{y}^0 = (y_1^0, y_2^0)$ and $\mathbf{y}^1 = (y_1^1, y_2^1)$ such that

$$\mathbf{y}^0, \mathbf{y}^1 \in \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in H_0^1(\Omega) \times L^2(\Omega) : \left\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix} \right\|_{H_0^1(\Omega) \times L^2(\Omega)} \leq r \right\},$$

where y solves (3.1) and $\bar{y} \in H_0^1(\Omega)$ is fixed as in (3.3).

As in the finite-dimensional case, the second quasi-turnpike result is one for an auxiliary control problem with fixed endpoints. For $0 \leq \tau_1 < \tau_2 \leq T$ and given $\mathbf{y}^{\tau_1}, \mathbf{y}^{\tau_2} \in H_0^1(\Omega) \times L^2(\Omega)$, this auxiliary problem consists in minimizing the nonnegative functional

$$J_{\tau_1, \tau_2}(u) := \int_{\tau_1}^{\tau_2} \|y(t) - \bar{y}\|_{H_0^1(\Omega)}^2 dt + \int_{\tau_1}^{\tau_2} \|\partial_t y(t)\|_{L^2(\Omega)}^2 + \int_{\tau_1}^{\tau_2} \|u(t)\|_{L^2(\omega)}^2 dt \quad (6.2)$$

over all $u \in U_{\text{ad}}$, where $y \in C^0([\tau_1, \tau_2]; H_0^1(\Omega)) \cap C^1([\tau_1, \tau_2]; L^2(\Omega))$ denotes the unique solution to

$$\begin{cases} \partial_t^2 y - \Delta y + f(y) = u \mathbf{1}_\omega & \text{in } (\tau_1, \tau_2) \times \Omega \\ y = 0 & \text{on } (\tau_1, \tau_2) \times \partial\Omega \\ (y, \partial_t y)|_{t=\tau_1} = \mathbf{y}^{\tau_1} & \text{in } \Omega. \end{cases} \quad (6.3)$$

and where

$$U_{\text{ad}} := \{u \in L^2((\tau_1, \tau_2) \times \omega) : (y, \partial_t y)|_{t=\tau_2} = \mathbf{y}^{\tau_2}\}.$$

We recall that $f \in \text{Lip}(\mathbb{R})$.

We now state and prove the wave equation analog of Lemma 5.2, which we recall, is the cornerstone of the bootstrap argument in our turnpike proof.

Lemma 6.2. *Let $T_0 > 0$ and $r > 0$ be provided by Claim 6.1, and let $\mathbf{y}^{\tau_1}, \mathbf{y}^{\tau_2}$ be such that*

$$\mathbf{y}^{\tau_i} \in \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in H_0^1(\Omega) \times L^2(\Omega) : \left\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix} \right\|_{H_0^1(\Omega) \times L^2(\Omega)} \leq r \right\}$$

for $i = 1, 2$. Let $T > 0$ and $0 \leq \tau_1 < \tau_2 \leq T$ be fixed such that $\tau_2 - \tau_1 \geq 2T_0$, and let $u_T \in U_{\text{ad}}$ be a global minimizer to J_{τ_1, τ_2} defined in (6.2), with y_T denoting the associated solution to (6.3). Then, there exists a constant $C = C(f, T_0, \Omega, \omega) > 0$ independent of $T, \tau_1, \tau_2 > 0$ such that

$$\begin{aligned} J_{\tau_1, \tau_2}(u_T) + \|y_T(t) - \bar{y}\|_{H_0^1(\Omega)}^2 + \|\partial_t y_T(t)\|_{L^2(\Omega)}^2 \\ \leq C \left(\|y_1^{\tau_1} - \bar{y}\|_{H_0^1(\Omega)}^2 + \|y_2^{\tau_1}\|_{L^2(\Omega)}^2 + \|y_1^{\tau_2} - \bar{y}\|_{L^2(\Omega)}^2 + \|y_2^{\tau_2}\|_{L^2(\Omega)}^2 \right) \end{aligned}$$

holds for all $t \in [\tau_1, \tau_2]$.

Proof of Lemma 6.2. The proof follows the lines of that of Lemma 5.2, with some slight technical differences. We provide details for the sake of completeness. For notational purposes, it will be significantly simpler to operate in the canonical first order system framework presented in the proof of Lemma 4.3. For the same reason, we will also drop the subscripts of T .

We set $X := H_0^1(\Omega) \times L^2(\Omega)$, and we denote

$$\mathbf{y} := \begin{bmatrix} y \\ \partial_t y \end{bmatrix}, \quad \bar{\mathbf{y}} := \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix}.$$

We also recall the definition of the skew-adjoint operator

$$A := \begin{bmatrix} 0 & \text{Id} \\ \Delta & 0 \end{bmatrix}, \quad D(A) = D(\Delta) \times H_0^1(\Omega),$$

where $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$. Then the desired estimate simply writes as

$$J_{\tau_1, \tau_2}(u) + \|\mathbf{y}(t) - \bar{\mathbf{y}}\|_X^2 \leq C \left(\|\mathbf{y}^{\tau_1} - \bar{\mathbf{y}}\|_X^2 + \|\mathbf{y}^{\tau_2} - \bar{\mathbf{y}}\|_X^2 \right)$$

for all $t \in [\tau_1, \tau_2]$. We proceed similarly as in the proof of Lemma 5.2. Using Claim 6.1, we know the following.

- There exists a control $u^\dagger \in L^2((\tau_1, \tau_1 + T_0) \times \omega)$ satisfying

$$\|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \leq C_0 \|\mathbf{y}^{\tau_1} - \bar{\mathbf{y}}\|_X^2, \quad (6.4)$$

for some $C_0 = C_0(T_0, \omega, f) > 0$, and such that the corresponding solution

$$\mathbf{y}^\dagger = \begin{bmatrix} y^\dagger \\ \partial_t y^\dagger \end{bmatrix} \text{ to}$$

$$\begin{cases} \partial_t \mathbf{y}^\dagger - A \mathbf{y}^\dagger + \begin{bmatrix} 0 \\ f(y^\dagger) \end{bmatrix} = \begin{bmatrix} 0 \\ u^\dagger \mathbf{1}_\omega \end{bmatrix} & \text{in } (\tau_1, \tau_1 + T_0) \\ \mathbf{y}^\dagger|_{t=\tau_1} = \mathbf{y}^{\tau_1} \end{cases}$$

satisfies $\mathbf{y}^\dagger(\tau_1 + T_0) = \bar{\mathbf{y}}$ in X . By writing the Duhamel formula for $\mathbf{y}^\dagger - \bar{\mathbf{y}}$, and using the conservative character of $\{e^{tA}\}_{t>0}$ in X , Cauchy-Schwarz, the Lipschitz character of f and the Poincaré inequality, we see that

$$\begin{aligned} \|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X &\leq \|e^{tA}(\mathbf{y}^{\tau_1} - \bar{\mathbf{y}})\|_X + \int_{\tau_1}^t \left\| e^{(t-s)A} \begin{bmatrix} 0 \\ u^\dagger(s) \mathbf{1}_\omega \end{bmatrix} \right\|_X ds \\ &\quad + \int_{\tau_1}^t \left\| e^{(t-s)A} \begin{bmatrix} 0 \\ (f(y^\dagger) - f(\bar{y})) \end{bmatrix} \right\|_X ds \\ &\leq \|\mathbf{y}^{\tau_1} - \bar{\mathbf{y}}\|_X + \sqrt{T_0} \|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)} \\ &\quad + C(f, \Omega) \int_{\tau_1}^t \|\mathbf{y}^\dagger(s) - \bar{\mathbf{y}}\|_X ds, \end{aligned} \quad (6.5)$$

with $C(f, \Omega) > 0$ depending solely on the Poincaré constant and the Lipschitz constant of f . Applying Grönwall's inequality to (6.5) and using (6.4), we deduce that

$$\|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X \leq C_1 \exp(C(f, \Omega)T_0) \|\mathbf{y}^{\tau_1} - \bar{\mathbf{y}}\|_X \quad (6.6)$$

holds for some $C_1(f, T_0, \omega) > 0$ independent of $T, \tau_1, \tau_2 > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$.

- There exists a control $u^\dagger \in L^2((\tau_1, \tau_1 + T_0) \times \omega)$ satisfying

$$\|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \leq C_0 \|\bar{\mathbf{y}} - \mathbf{y}^{\tau_2}\|_X^2, \quad (6.7)$$

and which is such that the corresponding solution $\mathbf{y}^\dagger = \begin{bmatrix} y^\dagger \\ \partial_t y^\dagger \end{bmatrix}$ to

$$\begin{cases} \partial_t \mathbf{y}^\dagger - \mathbf{A} \mathbf{y}^\dagger + \begin{bmatrix} 0 \\ f(y^\dagger) \end{bmatrix} = \begin{bmatrix} 0 \\ u^\dagger \mathbf{1}_\omega \end{bmatrix} & \text{in } (\tau_1, \tau_1 + T_0) \\ \mathbf{y}^\dagger|_{t=\tau_1} = \bar{\mathbf{y}} \end{cases}$$

satisfies $\mathbf{y}^\dagger(\tau_1 + T_0) = \mathbf{y}^{\tau_2}$ in X . Arguing just as above, we see that

$$\begin{aligned} \|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X &\leq \int_{\tau_1}^t \|e^{(t-s)A} \begin{bmatrix} 0 \\ u^\dagger(s) \mathbf{1}_\omega \end{bmatrix}\|_X ds + \int_{\tau_1}^t \|e^{(t-s)A} \begin{bmatrix} 0 \\ (f(y^\dagger) - f(\bar{y})) \end{bmatrix}\|_X ds \\ &\leq \sqrt{T_0} \|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)} + C(f, \Omega) \int_{\tau_1}^t \|\mathbf{y}^\dagger(s) - \bar{\mathbf{y}}\|_X ds, \end{aligned} \quad (6.8)$$

with $C(f, \Omega) > 0$ depending solely on the Poincaré constant and the Lipschitz constant of f . Applying Grönwall's inequality to (6.8) and using (6.7), we deduce that

$$\|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X \leq C_2 \exp(C(f, \Omega)T_0) \|\mathbf{y}^{\tau_2} - \bar{\mathbf{y}}\|_X \quad (6.9)$$

holds for some $C_2(f, T_0, \omega) > 0$ independent of $T, \tau_1, \tau_2 > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$.

Now set

$$u^{\text{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (\tau_1, \tau_1 + T_0) \\ 0 & \text{in } (\tau_1 + T_0, \tau_2 - T_0) \\ u^\dagger(t - (\tau_2 - \tau_1 - T_0)) & \text{in } (\tau_2 - T_0, \tau_2), \end{cases}$$

and let $\mathbf{y}^{\text{aux}} = \begin{bmatrix} y^{\text{aux}} \\ \partial_t y^{\text{aux}} \end{bmatrix}$ be the corresponding solution to (6.3). By construction, we have

$$\mathbf{y}^{\text{aux}}(t) = \mathbf{y}^\dagger(t) \quad \text{in } [\tau_1, \tau_1 + T_0],$$

and thus

$$\mathbf{y}^{\text{aux}}(t) = \bar{\mathbf{y}} \quad \text{in } [\tau_1 + T_0, \tau_2 - T_0], \quad (6.10)$$

whereas we also have $\mathbf{y}^{\text{aux}}(\tau_2) = \mathbf{y}^{\tau_2}$, whence $u^{\text{aux}} \in U_{\text{ad}}$.

We now evaluate J_{τ_1, τ_2} at u^{aux} , which by virtue of a simple change of variable as well as (6.10), (6.4), (6.6), (6.7) and (6.9), leads us to

$$\begin{aligned} J_{\tau_1, \tau_2}(u^{\text{aux}}) &= \|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 + \|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \\ &\quad + \int_{\tau_1}^{\tau_1 + T_0} \|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X^2 dt + \int_{\tau_1}^{\tau_1 + T_0} \|\mathbf{y}^\dagger(t) - \bar{\mathbf{y}}\|_X^2 dt \\ &\leq C_3 \left(\|\bar{\mathbf{y}} - \mathbf{y}^{\tau_1}\|_X^2 + \|\bar{\mathbf{y}} - \mathbf{y}^{\tau_2}\|_X^2 \right) \end{aligned} \quad (6.11)$$

where $C_3(f, T_0, \Omega, \omega) > 0$ is independent of $T, \tau_1, \tau_2 > 0$. By virtue of the optimality of u and (6.11), we have

$$J_{\tau_1, \tau_2}(u) \leq J_{\tau_1, \tau_2}(u^{\text{aux}}) \leq C_3 \left(\|\bar{y} - \mathbf{y}^{\tau_1}\|_X^2 + \|\bar{y} - \mathbf{y}^{\tau_2}\|_X^2 \right).$$

An application of Lemma 4.3 suffices to conclude. \square

7. PROOF OF THEOREM 3.2

For the semilinear heat equation, we can adapt the proof strategy of Theorem 2.1 to directly prove the stabilization result stipulated by Theorem 3.2. We provide details of the proof, as it is not an immediate application of that of Theorem 2.1.

We recall that since $f \in \text{Lip}(\mathbb{R})$, as presented in [30, Lemma 8.3] (and the references therein), given any $T > 0$, $y^0 \in L^2(\Omega)$ and $\bar{y} \in H_0^1(\Omega)$ solution to (3.3), there exists a control $u \in L^2((0, T) \times \omega)$ such that the unique solution y to (3.5) satisfies $y(T) = \bar{y}$, and

$$\|u\|_{L^2(\Omega)} \leq C(T, \omega, f) \|y^0 - \bar{y}\|_{L^2(\Omega)} \quad (7.1)$$

for some $C(T, \omega, f) > 0$ (the dependence on f is through the Lipschitz constant which is an upper bound for the potential appearing in the associated linear problem). Indeed, we may consider $z := y - \bar{y}$, and the control u steering z to 0 in time T is the same as that steering y to \bar{y} in time T . But then, $\|u\|_{L^2(\Omega)} \leq C(T, \omega, f) \|z(0)\|_{L^2(\Omega)}$ from the linear system and a fixed-point argument.

Suppose $y^{\tau_1} \in L^2(\Omega)$ is given. Let $T > 0$ and $0 \leq \tau_1 < T$ be fixed. Consider

$$J_{\tau_1, T}(u) := \int_{\tau_1}^T \|y(t) - \bar{y}\|_{L^2(\Omega)}^2 dt + \int_{\tau_1}^T \|u(t)\|_{L^2(\omega)}^2 dt, \quad (7.2)$$

where y solves

$$\begin{cases} \partial_t y - \Delta y + f(y) = u \mathbf{1}_\omega & \text{in } (\tau_1, T) \times \Omega \\ y = 0 & \text{on } (\tau_1, T) \times \partial\Omega \\ y|_{t=\tau_1} = y^{\tau_1} & \text{in } \Omega. \end{cases} \quad (7.3)$$

We will only need the following lemma, which is similar to Lemma 6.2. In fact, the blueprint of the proof below is contained therein.

Lemma 7.1. *Suppose $y^{\tau_1} \in L^2(\Omega)$ is given. Let $T > 0$ and τ_1 be given such that $T > \tau_1$. Let $u_T \in L^2((\tau_1, T) \times \omega)$ be any global minimizer to $J_{\tau_1, T}$ defined in (7.2), with y_T denoting the corresponding solution to (7.3). Then, there exists a constant $C = C(f, \bar{y}, \omega) > 0$ independent of $T, \tau_1 > 0$ and y^{τ_1} such that*

$$J_{\tau_1, T}(u_T) + \|y_T(t) - \bar{y}\|_{L^2(\Omega)}^2 \leq C \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)}^2$$

holds for all $t \in [\tau_1, T]$.

Proof of Lemma 7.1. Fix an arbitrary $T_0 > 0$.

Let us first suppose that $T \geq \tau_1 + T_0$. By controllability to the steady state \bar{y} (see the discussion around (7.1)), we know that exists a control $u^\dagger \in L^2((\tau_1, \tau_1 + T_0) \times \omega)$ satisfying

$$\|u^\dagger\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)} \leq C_1 \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)} \quad (7.4)$$

for some $C_1 = C_1(T_0, \omega, f) > 0$ and such that the corresponding solution y^\dagger to

$$\begin{cases} \partial_t y^\dagger - \Delta y^\dagger + f(y^\dagger) = u^\dagger \mathbf{1}_\omega & \text{in } (\tau_1, \tau_1 + T_0) \times \Omega \\ y^\dagger = 0 & \text{on } (\tau_1, \tau_1 + T_0) \times \partial\Omega \\ y^\dagger|_{t=0} = y^0 & \text{in } \Omega. \end{cases}$$

satisfies $y^\dagger(\tau_1 + T_0) = \bar{y}$ a.e. in Ω . Arguing as in the proof of Lemma 4.2, we see that

$$\begin{aligned} \left\| y^\dagger(t) - \bar{y} \right\|_{L^2(\Omega)} &\leq \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)} + \sqrt{T_0} \left\| u^\dagger \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)} \\ &\quad + C(f) \int_{\tau_1}^t \left\| y^\dagger(s) - \bar{y} \right\|_{L^2(\Omega)} ds \end{aligned} \quad (7.5)$$

for $t \in (\tau_1, \tau_1 + T_0)$, with $C(f) > 0$ being the Lipschitz constant of f . Applying Grönwall's inequality to (7.5) and using (7.4), we deduce that

$$\left\| y^\dagger(t) - \bar{y} \right\|_{L^2(\Omega)} \leq C_2 \exp(C(f)T_0) \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)} \quad (7.6)$$

for some $C_2(f, T_0, \omega) > 0$ independent of $T, \tau_1, \tau_2 > 0$, and for every $t \in (\tau_1, \tau_1 + T_0)$.

Now set

$$u^{\text{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (\tau_1, \tau_1 + T_0) \\ 0 & \text{in } (\tau_1 + T_0, T) \end{cases}$$

and let y^{aux} be the corresponding solution to (3.5). Clearly $y^{\text{aux}}(t) = \bar{y}$ for $t \in [\tau_1 + T_0, T]$, a.e. in Ω . Hence, using $J_{\tau_1, T}(u_T) \leq J_{\tau_1, T}(u^{\text{aux}})$, (7.6) and (7.4), we see that

$$\begin{aligned} J_{\tau_1, T}(u_T) &\leq \left\| y^\dagger - \bar{y} \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \Omega)}^2 + \left\| u^\dagger \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \\ &\leq C_3 \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)}^2 \end{aligned}$$

for some $C_3(f, T_0, \omega) > 0$ independent of $T, \tau_1 > 0$. Applying Lemma 4.2 suffices to conclude.

Now suppose that $\tau_1 < T < T_0 + \tau_1$. We may then use the optimality inequality $J_{\tau_1, T}(u_T) \leq J_{\tau_1, T}(u_{T_0 + \tau_1})$, and since by the previous step, we know that

$$J_{\tau_1, T}(u_{T_0 + \tau_1}) \leq J_{\tau_1, T_0 + \tau_1}(u_{T_0 + \tau_1}) \leq C_3 \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)}^2$$

where $C_3 = C_3(f, T_0, \omega) > 0$ is independent of $T, \tau_1 > 0$, we deduce

$$J_{\tau_1, T}(u_T) \leq C_3 \|y^{\tau_1} - \bar{y}\|_{L^2(\Omega)}^2. \quad (7.7)$$

We may conclude by combining (7.7) with Lemma 4.2. \square

Proof of Theorem 3.2. The proof is of the same spirit as that of Theorem 2.1, the only difference being the fact that we only need to bootstrap "forward" in time due to the lack of final cost, which renders the proof significantly less technical. The control estimate follows from Lemma 7.1. We thus concentrate solely on estimating the state.

Fix

$$\tau > C_1^4$$

where $C_1 = C_1(f, \bar{y}, \omega) > 0$ is the (square root of the) constant appearing in Lemma 7.1, and let $T_0 > \tau$ be arbitrary and fixed⁵. Similarly to the proof of Theorem 2.1, we will

⁵Note that this choice is independent of the one done in the proof of Lemma 7.1.

have $T^* := \tau + T_0$ in the statement. Let

$$T \geq \tau + T_0$$

be fixed.

First note that for $t \in [0, \tau + T_0]$, just as in Part 1 of the proof of Theorem 2.1, the desired estimate can easily be obtained for such t since the length of the time interval is independent of T . Hence, we will solely concentrate on the case $t \in [\tau + T_0, T]$. To this end, we will mimic the steps done in the proof of Theorem 2.1.

Step 1). **Preparation.** Since $2\tau < \tau + T_0 < T$ and thus $\tau \leq \frac{T}{2}$, by Lemma 5.3 there exists a $\tau_1 \in [0, \tau)$ such that

$$\|y_T(\tau_1) - \bar{y}\|_{L^2(\Omega)} \leq \frac{\|y_T - \bar{y}\|_{L^2((0,T) \times \Omega)}}{\sqrt{\tau}} \leq \frac{C_1}{\sqrt{\tau}} \|y^0 - \bar{y}\|_{L^2(\Omega)}. \quad (7.8)$$

Now the control $u_T|_{[\tau_1, T]}$ minimizes $J_{\tau_1, T}$ with initial data $y^{\tau_1} = y_T(\tau_1)$ for (7.3), to which clearly the solution is $y_T|_{[\tau_1, T]}$. So by Lemma 7.1 and (7.8),

$$\|y_T(t) - \bar{y}\|_{L^2(\Omega)} \leq C_1 \|y_T(\tau_1) - \bar{y}\|_{L^2(\Omega)} \leq \frac{C_1^2}{\sqrt{\tau}} \|y^0 - \bar{y}\|_{L^2(\Omega)} \quad (7.9)$$

holds for all $t \in [\tau_1, T]$. Since $\tau_1 < \tau$, (7.9) also holds for all $t \in [\tau, T]$.

Step 2). **Bootstrap.** We bootstrap (7.9) and prove that for any $n \in \mathbb{N}$ satisfying

$$n \leq \frac{T}{2\tau},$$

the estimate

$$\|y_T(t) - \bar{y}\|_{L^2(\Omega)} \leq \left(\frac{C_1^2}{\sqrt{\tau}} \right)^n \|y^0 - \bar{y}\|_{L^2(\Omega)} \quad (7.10)$$

holds for all $t \in [n\tau, T]$. We proceed by induction. The case $n = 1$ holds by (7.9). Thus assume that (7.10) holds at some stage $n \in \mathbb{N}$ and suppose that

$$n + 1 \leq \frac{T}{2\tau}.$$

This clearly implies that

$$\tau \leq \frac{T - 2n\tau}{2}. \quad (7.11)$$

Now the control $u_T|_{[n\tau, T]}$ is a global minimizer of $J_{n\tau, T}$. We can thus apply Lemma 7.1 with $\tau_1 = n\tau$, and Lemma 5.3 (note (7.11)) on $[n\tau, T - n\tau]$, to deduce that there exists $t_1 \in [n\tau, (n+1)\tau)$ such that

$$\|y_T(t_1) - \bar{y}\|_{L^2(\Omega)} \leq \frac{\|y_T - \bar{y}\|_{L^2((n\tau, T) \times \Omega)}}{\sqrt{\tau}} \leq \frac{C_1}{\sqrt{\tau}} \|y_T(n\tau) - \bar{y}\|_{L^2(\Omega)}.$$

So now we apply the induction hypothesis (7.10) to deduce

$$\|y_T(t_1) - \bar{y}\|_{L^2(\Omega)} \leq \frac{C_1}{\sqrt{\tau}} \left(\frac{C_1^2}{\sqrt{\tau}} \right)^n \|y^0 - \bar{y}\|_{L^2(\Omega)}. \quad (7.12)$$

Since $u_T|_{[t_1, T]}$ is a global minimizer of $J_{t_1, T}$, we can apply Lemma 7.1 and use (7.12) to deduce that

$$\|y_T(t) - \bar{y}\|_{L^2(\Omega)} \leq C_1 \|y_T(t_1) - \bar{y}\|_{L^2(\Omega)} \leq \frac{C_1^2}{\sqrt{\tau}} \left(\frac{C_1^2}{\sqrt{\tau}} \right)^n \|y^0 - \bar{y}\|_{L^2(\Omega)} \quad (7.13)$$

holds for all $t \in [t_1, T]$. Clearly, as $t_1 < (n + 1)\tau$, (7.13) also holds for all $t \in [(n + 1)\tau, T]$. This concludes the induction proof, and so (7.10) does indeed hold.

Step 3). **Conclusion.** We now use (7.10) to conclude the proof. Suppose $t \in [\tau + T_0, T]$ is arbitrary and fixed. Set $n(t) := \lfloor \frac{t}{\tau + T_0} \rfloor$. Clearly $n(t) \geq 1$, $t \geq n(t)\tau$ and $n(t) \leq \frac{T}{2\tau}$ due to the choice of T_0 . We may then apply (7.10) to find that

$$\|y_T(t) - \bar{y}\|_{L^2(\Omega)} \leq \left(\frac{C_1^2}{\sqrt{\tau}}\right)^{n(t)} \|y^0 - \bar{y}\|_{L^2(\Omega)} \quad (7.14)$$

Now since $\tau > C_1^4$ and $n(t) \geq \frac{t}{\tau + T_0} - 1$, we can see from (7.14) that

$$\begin{aligned} \|y_T(t) - \bar{y}\|_{L^2(\Omega)} &\leq \exp\left(-n(t) \log\left(\frac{\sqrt{\tau}}{C_1^2}\right)\right) \|y^0 - \bar{y}\|_{L^2(\Omega)} \\ &\leq \frac{\sqrt{\tau}}{C_1^2} \exp\left(-\frac{\log\left(\frac{\sqrt{\tau}}{C_1^2}\right)}{\tau + T_0} t\right) \|y^0 - \bar{y}\|_{L^2(\Omega)} \end{aligned}$$

The desired estimate thus holds for all $t \in [\tau + T_0, T]$, with

$$\mu := \frac{\log\left(\frac{\sqrt{\tau}}{C_1^2}\right)}{\tau + T_0} > 0$$

and

$$C := \frac{\sqrt{\tau}}{C_1^2} \|y^0 - \bar{y}\|_{L^2(\Omega)}.$$

This concludes the proof. \square

8. CONCLUDING REMARKS AND OUTLOOK

We have presented a new methodology for proving the turnpike property for nonlinear optimal control problems set in large time horizons, under the assumption that the running target is a steady control-state pair, and that the system is controllable with a local estimate on the cost. These assumptions allow us to bypass necessary optimality conditions and a study of the adjoint system, and rather relies on calculus of variations–based arguments.

More precisely, we have concluded that

- (1). The exponential turnpike property holds for optimal state trajectories of optimal control problems for nonlinear finite and infinite-dimensional dynamics, whenever the cost functional is coercive with respect to the distance of the state to the target steady state. The nonlinearity may be assumed to be only globally Lipschitz continuous (and thus possibly nonsmooth). The result holds without any smallness assumptions on the initial data.
- (2). The last exponential arc (near $t = T$) can be removed whenever the optimal control problem is considered without a final time cost, and thus entails an exponential stabilization estimate for the optimal state trajectory.

8.1. **Outlook.** Let us conclude with a list of select problems related to our study.

- **Necessity of assuming that \bar{y} is a steady state.** The assumption that the running target \bar{y} in (2.3) is a steady state of the dynamics allows us to easily obtain quasi-turnpike strategies allowing us to obtain the key estimates in Lemma 5.1 and Lemma 5.2 (resp. Lemma 6.1, Lemma 6.2, Lemma 7.1 in the PDE setting). The case of controlled steady states \bar{y} associated to a prescribed control \bar{u} can readily be addressed by penalizing $u - \bar{u}$ over $[0, T]$ instead of solely u as noted in Remark 1. But we were unable to see if this is a necessary assumption in the nonlinear context in the absence of smallness conditions on the target, and whether the controlled steady state case can be covered by solely penalizing u . These questions merit in-depth investigation.
- **Weakening Assumption 1.** An important hypothesis we made throughout is Assumption 1, which required that, at least for data y^0, y^1 in the vicinity of the free steady state \bar{y} , the minimal L^2 -norm control steering the system from y^0 to \bar{y} may be estimated by $\|y^0 - \bar{y}\|$, and similarly for that from \bar{y} to y^1 . This is a hallmark of linear control systems, which is also expected for nonlinear systems for which controllability results are obtained by linearization or perturbation methods and a fixed-point argument. But in the general context of control-affine systems, such an assumption may appear restrictive, even though it is local. It is thus of interest to see how the results and methodology can be pertained whilst weakening Assumption 1.

In fact, more generally, it would be of interest to investigate whether the methodology presented herein can still be applied by only assuming approximate controllability with an adequate estimate on the control cost.

- **Turnpike with state or control constraints.** A problem which has not been extensively covered in the literature is the turnpike property with positivity (or box) constraints on either the state or the control. Slightly weaker integral turnpike results under such constraints have been obtained in [27] by means of *quantitative inequalities*. Such a study would complement the already existent nonlinear *controllability under constraints theory* – a topic covered in several recent works, see e.g. [21, 28, 30, 33] and the references therein.
- **More general control systems.** We have considered homogeneous Dirichlet boundary conditions in (3.1) and (3.5) merely to avoid additional technical details. The proofs of Theorem 3.1 (resp. Theorem 3.2) only require that the underlying dynamics are exactly controllable (resp. controllable to a steady state), thus, the same results hold with Neumann boundary conditions or boundary controls. Similarly, variable coefficients and lower order terms may be considered, as long as these coefficients are time-independent, as we are using a Duhamel formula along with a semigroup representation of the solution.

In fact, we have chosen the wave and heat equation for the sake of presentation, but the respective results could possibly be extended to a more general scenario of exactly controllable semilinear systems with similar assumptions, e.g. dispersive equations (Schrödinger, Korteweg-de Vries), coupled systems, and so on. The necessity of a Duhamel formula may however be an impediment to the extension of our results to the context of quasilinear systems such as the porous medium equation (see [13] and the references therein).

- **Bilinear control systems.** It would also be of interest to establish the turnpike property for bilinear control systems. This would be the somewhat true analog of the control-affine systems presented herein, and under suitable assumptions on the nonlinearity, one could expect that our methodology applies to such cases as well. We have not addressed such systems for the simplicity of presentation and due to the controllability assumptions we make, as the controllability theory for bilinear problems is not complete (albeit, see [3, 4, 8, 26] for recent developments). Notwithstanding, our results should be applicable to a system of the form (see [3])

$$\begin{cases} \partial_t y - \partial_x^2 y = u(t)f(y) & \text{in } (0, T) \times (0, \pi) \\ \partial_x y(t, 0) = \partial_x y(t, \pi) = 0 & \text{in } (0, T) \\ y|_{t=0} = y^0 & \text{in } (0, \pi) \end{cases}$$

where u is a scalar control and f is an appropriate nonlinearity (see [3] for sufficient conditions for ensuring controllability, and globally Lipschitz for applying our methodology).

- **More general nonlinearities.** Finally, it would be of interest to investigate problems where our methodology does not immediately apply, such as the paradigmatic example of the cubic heat equation. This problem consists in seeing whether one may prove Theorem 3.2 (with the estimate on u_T changed by an estimate of $u_T - \bar{u}$) for minimizers u_T of

$$J_T(u) := \int_0^T \|y(t) - \bar{y}\|^2 dt + \int_0^T \|u - \bar{u}\|^2 dt$$

where y_T is the unique solution to

$$\begin{cases} \partial_t y - \Delta y + y^3 = u \mathbf{1}_\omega & \text{in } (0, T) \times \Omega \\ y = 0 & \text{on } (0, T) \times \partial\Omega \\ y|_{t=0} = y^0 & \text{in } \Omega, \end{cases} \quad (8.1)$$

and $\bar{y} \in H_0^1(\Omega)$ is a controlled steady state associated to some $\bar{u} \in L^2(\omega)$ (the case $\bar{u} \equiv 0$ is somewhat trivial due to the inherent stabilization to $\bar{y} \equiv 0$). Let us elaborate on a possible technical impediment in the direct application of our strategy. Clearly, for Theorem 3.2 to hold in this case, it would suffice to prove Lemma 7.1 for $f(s) = s^3$ (while replacing the estimate of u_T by an estimate of $u_T - \bar{u}$). To this end, first of all, for any $u \in L^2((0, T) \times \omega)$, using the variational formulation and standard arguments including Cauchy-Schwarz, Young with ϵ and Poincaré inequalities, one can find

$$\frac{d}{dt} \int_\Omega |y(t, x)|^2 dx \leq \epsilon \int_\omega \|u(t, x)\|^2 dx$$

for a.e. $t \in [0, T]$, where $\epsilon > \frac{C(\Omega)}{4}$, whereas y solves (8.1), and thus

$$\|y\|_{C^0([0, T]; L^2(\Omega))} \leq C_1(\Omega) \left(\|u\|_{L^2((0, T) \times \omega)} + \|y^0\|_{L^2(\Omega)} \right). \quad (8.2)$$

Following the proof of Lemma 4.2 for $f(s) = s^3$ and using (8.2), we may find

$$\|y(t) - \bar{y}\|_{L^2(\Omega)} \leq C \left(\|y^0 - \bar{y}\|_{L^2(\Omega)} + \|u - \bar{u}\|_{L^2((0, T) \times \omega)} + \|y - \bar{y}\|_{L^2((0, T) \times \Omega)} \right),$$

where now

$$C \sim \exp(\|u\|_{L^2((0,T)\times\omega)}).$$

It is precisely at this point where the issue appears, since simply by using the form of the functional, we are not in a position to prove that $\|u\|_{L^2((0,T)\times\omega)}$ is uniformly bounded with respect to T , but rather only $\|u - \bar{u}\|_{L^2((0,T)\times\omega)}$. Should this be possible, then one can expect our methodology to apply to the cubic heat equation as well, but as things stand, turnpike without smallness conditions in this case remains open.

Further examples worth analyzing include the heat equation with a convective nonlinearity $f(y, \nabla y)$, even in one space dimension (e.g. the Burgers equation); along these lines we refer to [38] for a local turnpike result for the 2d Navier-Stokes system. Similar questions can be asked for the semilinear wave equation, where the nonlinearity is sometimes only assumed to be super-linear (see [20] for a subcritical optimal control study) – our methodology a priori applies if the nonlinearity is either truncated by some cut-off, or if one manages to prove uniform estimates of $\|y_T\|_{L^\infty((0,T)\times\Omega)}$ with respect to T . Further nonlinear problems which could be investigated include hyperbolic systems (see [16] for a related study) or free boundary problems (see [14] for a control perspective).

Acknowledgments. B.G. thanks Idriss Mazari (TU Wien) for helpful comments.

Funding: B.G. and E.Z. have received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.765579-ConFlex. D.P., C.E. and E.Z. have received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement NO. 694126-DyCon). The work of E. Z. has been supported by the Alexander von Humboldt-Professorship program, the Transregio 154 Project “Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks” of the German DFG, grant MTM2017-92996-C2-1-R COSNET of MINECO (Spain) and by the Air Force Office of Scientific Research (AFOSR) under Award NO. FA9550-18-1-0242.

REFERENCES

- [1] AMANN, H. *Linear and quasilinear parabolic problems*, vol. 1. Springer, 1995.
- [2] BEAUCHARD, K., AND MARBACH, F. Quadratic obstructions to small-time local controllability for scalar-input systems. *Journal of Differential Equations* 264, 5 (2018), 3704–3774.
- [3] BEAUCHARD, K., AND MARBACH, F. Unexpected quadratic behaviors for the small-time local null controllability of scalar-input parabolic equations. *Journal de Mathématiques Pures et Appliquées* 136 (2020), 22–91.
- [4] CANNARSA, P., FLORIDIA, G., AND KHAPALOV, A. Y. Multiplicative controllability for semilinear reaction–diffusion equations with finitely many changes of sign. *Journal de Mathématiques Pures et Appliquées* 108, 4 (2017), 425–458.
- [5] CAZENAVE, T. *An introduction to semilinear elliptic equations*, vol. 164. Editora do Instituto de Matemática, Universidade Federal do Rio de Janeiro, 2006.
- [6] CORON, J.-M. *Control and nonlinearity*. No. 136. American Mathematical Soc., 2007.
- [7] CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems* (1989), 303–314.
- [8] DUPREZ, M., AND LISSY, P. Bilinear local controllability to the trajectories of the Fokker-Planck equation with a localized control. *arXiv preprint arXiv:1909.02831* (2019).
- [9] DUYSCKAERTS, T., ZHANG, X., AND ZUAZUA, E. On the optimality of the observability inequalities for parabolic and hyperbolic systems with potentials. In *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis* (2008), vol. 25, Elsevier, pp. 1–41.
- [10] E, W. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics* 5, 1 (2017), 1–11.

- [11] ESTEVE, C., GESHKOVSKI, B., PIGHIN, D., AND ZUAZUA, E. Large-time asymptotics in deep learning. *arXiv preprint arXiv:2008.02491* (2020).
- [12] FU, X., YONG, J., AND ZHANG, X. Exact controllability for multidimensional semilinear hyperbolic equations. *SIAM Journal on Control and Optimization* 46, 5 (2007), 1578–1614.
- [13] GESHKOVSKI, B. Null-controllability of perturbed porous medium gas flow. *ESAIM: Control, Optimisation and Calculus of Variations* 26 (2020), 85.
- [14] GESHKOVSKI, B., AND ZUAZUA, E. Controllability of one-dimensional viscous free boundary flows. *Hal preprint* (2019).
- [15] GRÜNE, L., SCHALLER, M., AND SCHIELA, A. Exponential sensitivity and turnpike analysis for linear quadratic optimal control of general evolution equations. *Journal of Differential Equations* 268, 12 (2020), 7311–7341.
- [16] GUGAT, M., AND HANTE, F. M. On the turnpike phenomenon for optimal boundary control problems with hyperbolic systems. *SIAM Journal on Control and Optimization* 57, 1 (2019), 264–289.
- [17] GUGAT, M., TRÉLAT, E., AND ZUAZUA, E. Optimal Neumann control for the 1d wave equation: Finite horizon, infinite horizon, boundary tracking terms and the turnpike property. *Systems & Control Letters* 90 (2016), 61–70.
- [18] HE, K., ZHANG, X., REN, S., AND SUN, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 770–778.
- [19] JOLY, R., AND LAURENT, C. A note on the semiglobal controllability of the semilinear wave equation. *SIAM Journal on Control and Optimization* 52, 1 (2014), 439–450.
- [20] KUNISCH, K., AND MEINLSCHMIDT, H. Optimal control of an energy-critical semilinear wave equation in 3d with spatially integrated control constraints. *Journal de Mathématiques Pures et Appliquées* (2020).
- [21] LE BALC'H, K. Local controllability of reaction-diffusion systems around nonnegative stationary states. *ESAIM: Control, Optimisation and Calculus of Variations* 26 (2020), 55.
- [22] LECUN, Y., BENGIO, Y., AND HINTON, G. Deep learning. *Nature* 521, 7553 (2015), 436–444.
- [23] LIN, H., AND JEGELKA, S. Resnet with one-neuron hidden layers is a universal approximator. In *Advances in Neural Information Processing Systems* (2018), pp. 6169–6178.
- [24] LIONS, J.-L. Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. tome 1. *RMA* 8 (1988).
- [25] LIONS, P.-L. On the existence of positive solutions of semilinear elliptic equations. *SIAM Review* 24, 4 (1982), 441–467.
- [26] MAZARI, I., AND RUIZ-BALET, D. A fragmentation phenomenon for a non-energetic optimal control problem: optimisation of the total population size in logistic diffusive models. *arXiv preprint arXiv:2005.08515* (2020).
- [27] MAZARI, I., AND RUIZ-BALET, D. Quantitative stability for eigenvalues of Schrödinger operator & application to the turnpike property for a bilinear optimal control problem. *Hal preprint* (2020).
- [28] MAZARI, I., RUIZ-BALET, D., AND ZUAZUA, E. Constrained control of gene-flow bistable reaction-diffusion equations: spatially heterogeneous models and infection-dependent models. *Hal preprint* (2020).
- [29] PIGHIN, D. The turnpike property in semilinear control. *arXiv:2004.03269* (2020).
- [30] PIGHIN, D., AND ZUAZUA, E. Controllability under positivity constraints of semilinear heat equations. *Mathematical Control & Related Fields* 8, 3&4 (2018), 935.
- [31] PORRETTA, A., AND ZUAZUA, E. Long time versus steady state optimal control. *SIAM Journal on Control and Optimization* 51, 6 (2013), 4242–4273.
- [32] PORRETTA, A., AND ZUAZUA, E. Remarks on long time versus steady state optimal control. In *Mathematical Paradigms of Climate Science*. Springer, 2016, pp. 67–89.
- [33] RUIZ-BALET, D., AND ZUAZUA, E. Control under constraints for multi-dimensional reaction-diffusion monostable and bistable equations. *Journal de Mathématiques Pures et Appliquées* 143 (2020), 345–375.
- [34] TRÉLAT, E. Linear turnpike theorem. *arXiv preprint arXiv:2010.13605* (2020).
- [35] TRÉLAT, E., ZHANG, C., AND ZUAZUA, E. Steady-state and periodic exponential turnpike property for optimal control problems in Hilbert spaces. *SIAM Journal on Control and Optimization* 56, 2 (2018), 1222–1252.

- [36] TRÉLAT, E., AND ZUAZUA, E. The turnpike property in finite-dimensional nonlinear optimal control. *Journal of Differential Equations* 258, 1 (2015), 81–114.
- [37] TUCSNAK, M., AND WEISS, G. *Observation and control for operator semigroups*. Springer Science & Business Media, 2009.
- [38] ZAMORANO, S. Turnpike property for two-dimensional Navier–Stokes equations. *Journal of Mathematical Fluid Mechanics* 20, 3 (2018), 869–888.
- [39] ZHANG, X., AND ZUAZUA, E. Exact controllability of the semi-linear wave equation. *Unsolved Problems in Mathematical Systems and Control Theory* (2004), 173.
- [40] ZUAZUA, E. Exact controllability for semilinear wave equations in one space dimension. In *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis* (1993), vol. 10, Elsevier, pp. 109–129.
- [41] ZUAZUA, E. Controllability and observability of partial differential equations: some results and open problems. In *Handbook of differential equations: evolutionary equations*, vol. 3. Elsevier, 2007, pp. 527–621.
- [42] ZUAZUA, E. Large time control and turnpike properties for wave equations. *Annual Reviews in Control* 44 (2017), 199–210.

Carlos Esteve, Borjan Geshkovski, Dario Pighin

Departamento de Matemáticas
 Universidad Autónoma de Madrid
 28049 Madrid, Spain

and

Chair of Computational Mathematics
 Fundación Deusto
 Av. de las Universidades, 24
 48007 Bilbao, Basque Country, Spain

Email address: {carlos.esteve, borjan.geshkovski, dario.pighin}@uam.es

Enrique Zuazua

Chair in Applied Analysis, Alexander von Humboldt-Professorship
 Department of Mathematics
 Friedrich-Alexander-Universität Erlangen-Nürnberg
 91058 Erlangen, Germany

and

Chair of Computational Mathematics
 Fundación Deusto
 Av. de las Universidades, 24
 48007 Bilbao, Basque Country, Spain

and

Departamento de Matemáticas
 Universidad Autónoma de Madrid
 28049 Madrid, Spain

Email address: enrique.zuazua@fau.de