Turnpike in Lipschitz–nonlinear optimal control

Carlos Esteve-Yagüe, Borjan Geshkovski, Dario Pighin, Enrique Zuazua

HAL Id: hal-03018529

https://hal.science/hal-03018529v3

Preprint submitted on 14 Mar 2022

TURNPIKE IN LIPSCHITZ–NONLINEAR OPTIMAL CONTROL

CARLOS ESTEVE-YAGÜE, 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 system. Our strategy combines the construction of quasi-turnpike controls 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 applications in machine learning through deep residual neural networks, which may be fit within our setting. 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

1.1. **Motivation.** 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 $\mathcal{O}\left(e^{-\mu t} + e^{-\mu(T-t)}\right)$–close to the optimal static solution $\overline{y}$, for $t \in [0, T]$ and for some $\mu > 0$.

1.1.1. *Background.* 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 ([39, 20]). 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 ([40, 46]). The smallness conditions on the initial data can be removed in some specific cases ([36]), 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 [44]). 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.

1.1.2. *A question raised by machine learning.* There has been an ever-increasing need, brought by applications in machine learning via residual neural networks (ResNets, [11, 12, 24]), of turnpike results for nonlinear optimal control problems without smallness conditions on the initial data or the running target, and for systems with globally Lipschitz-continuous but possibly nonsmooth nonlinearities.

In (supervised) machine learning, one looks for a map which interpolates a dataset

$$\left\{x^{(i)}, y^{(i)}\right\}_{i \in \{1, \ldots, n\}} \subset \mathbb{R}^{d_x} \times \mathbb{R}^{d_y},$$

and which gives accurate predictions on unknown points $x \in \mathbb{R}^{d_x}$ ([29]). Such a task may (in many cases) be accomplished by solving

$$\inf_{\substack{u=(w,b) \in L^2(0,T;\mathbb{R}^{d_u}) \\ \mathbf{x}_i \text{ solves } (1.2)}} \sum_{i=1}^{n} \int_0^T \left\|P\mathbf{x}_i(t) - y^{(i)}\right\|^2 \mathrm{d}t + \int_0^T \|u(t)\|^2 \, \mathrm{d}t, \qquad (1.1)$$

where $P : \mathbb{R}^{d_x} \to \mathbb{R}^{d_y}$ is a given surjective map (possibly nonlinear, see Section 8), and the constraint is given by the continuous-time residual neural network[1]

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

with $w(t) \in \mathbb{R}^{d_x \times d_x}$ and $b(t) \in \mathbb{R}^{d_x}$ designate the controls (thus $d_u = d_x^2 + d_x$), whereas $\sigma \in \mathrm{Lip}(\mathbb{R})$ with $\sigma(0) = 0$ is a scalar nonlinear function, defined componentwise in (1.2). The most frequently used nonlinearities in practice are *rectifiers*: $\sigma(x) = \max\{\alpha x, x\}$

---

[1]Also referred to as a neural ODE [7].

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

$$\begin{cases} \dot{\mathbf{x}}_i(t) = w(t)\sigma(\mathbf{x}_i(t)) + b(t) & \text{in } (0, T) \\ \mathbf{x}_i(0) = x^{(i)}. \end{cases} \tag{1.3}$$
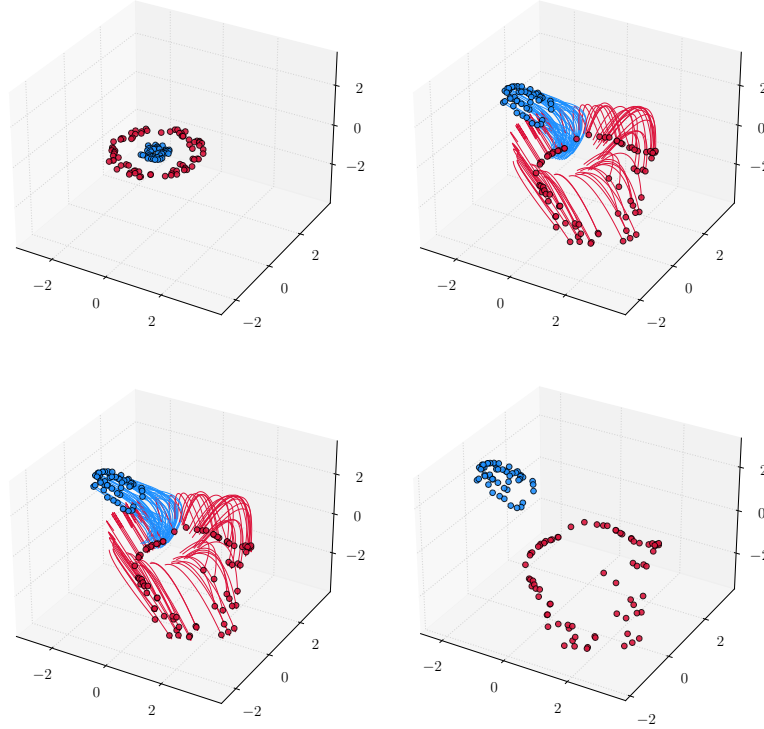
FIGURE 1. *Binary classification task.* One aims to separate the data $\left\{x^{(i)}\right\}_{i \in \{1, \ldots, n\}} \subset \mathbb{R}^3$ (top left) with respect to their color ($y^{(i)} = \pm 1$ for red, blue) by using the flow of (1.3), found by minimizing (1.1). We plot the trajectories $\mathbf{x}(t) := \{\mathbf{x}_i(t)\}_{i \in \{1, \ldots, n\}}$ of (1.3) for $t \leqslant 2$ (top right), $t \leqslant 5$ (mid left), and in time $T = 5$ (mid right). We see stabilization for the projections, for the controls to 0, and hence also for the trajectories to some points $\overline{\mathbf{x}}_i \in P^{-1}(\{y^{(i)}\})$, which are steady states (bottom).
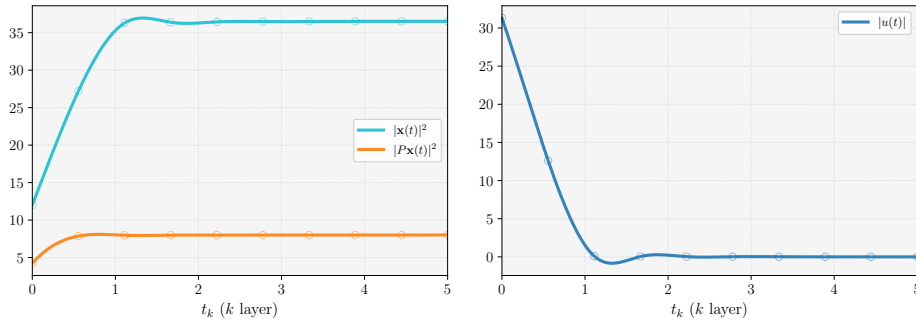
Combinations and variants of (1.2) and (1.3) may also be used ([30]). Optimizing u over $n \gg 1$ different initial data establishes robustness, so that the neural networks (1.2) and (1.3) may correctly perform future predictions on unknown points (Figure 5).

One notes a feature specific to the dynamics $f(\mathbf{x}_i, u)$ in (1.2) and (1.3) for fixed i: any constant vector in $\mathbb{R}^{d_x}$ is a steady state with control $u = (w, b) \equiv 0$. Whence, the solutions to the optimal steady problem

$$\inf_{\substack{(u, \mathbf{x}_i) \in \mathbb{R}^{d_u} \times \mathbb{R}^{d_x} \\ f(\mathbf{x}_i, u) = 0}} \sum_{i=1}^{n} \left\| P\mathbf{x}_i - y^{(i)} \right\|^2 + \|u\|^2$$

with $f(\mathbf{x}_i, u)$ as in (1.2) or (1.3), are precisely given by $(0, \overline{\mathbf{x}}_i)$, where $\overline{\mathbf{x}}_i \in P^{-1}\left(\left\{y^{(i)}\right\}\right)$ for $i \in \{1, \ldots, n\}$ (the preimage of P might not be a singleton if $d_x \neq d_y$). In Figure 1, we see that not only the projections $P\mathbf{x}_i(t)$, but also the optimal trajectories $\mathbf{x}_i(t)$ stabilize to such points: $\overline{\mathbf{x}}_i \in P^{-1}\left(\left\{y^{(i)}\right\}\right)$, for $i \in \{1, \ldots, n\}$, which are, as said above, steady states of the underlying system without control (i.e. with 0 control). Existing turnpike results do not immediately apply to explain this artifact, as discussed in Section 1.1.1, due to the use of nonsmooth nonlinearities and the lack of smallness assumptions on the targets, which would be unrealistic. This motivates the setting of our study (even for more general dynamics), namely, the consideration running targets which are steady states of the underlying dynamics without control (see (1.4) below).

1.1.3. *Interpretation.* The practical interest of this stabilization property is regarding the approximation capacity of ResNets, which are the forward Euler discretization of (1.2) and (1.3) with fixed time-step $\triangle t = {}^T\!/n_t$. Here n_t is the *number of layers*, and when $n_t \gg 1$, one is said to be doing *deep learning*. As $\triangle t$ is fixed, exponential decay would provide a quantitative estimate of the number of layers needed to fit the data, whilst keeping the controls small (thus possibly ensuring generalization). Such an estimate would actually indicate that the time horizon (or number of layers) ought not be big at all so that the approximation error (the first term in (1.1)) reaches 0 with controls of small amplitude. In other words, any layers beyond a certain stopping time T^* can be dropped[2] from the optimization scheme. In our numerical experiment (see Section 8 for detail and further illustrations), we use $T = 5$ and $\triangle t = {}^1\!/2$, and stability occurs beyond $T^* \sim 2$.

1.2. **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.

- In Section 2, we consider optimal control problems consisting of minimizing[3]

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \overline{y}\|^2 \, \mathrm{d}t + \int_0^T \|u(t) - \overline{u}\|^2 \, \mathrm{d}t \tag{1.4}$$

---

[2]It should to be said that a sharp and applicable conclusion would depend on the "complexity" of the dataset, as well as the number n of datapoints, which we do not specifically take into account. But an exponential decay estimate would give a rough idea on how to design methods for numerically estimating the stopping time T^*. We refer the reader to [12, 13, 16, 18, 19] for further detail. Similar conclusions have been drawn in the context of $L^1(0, T; \mathbb{R}^{d_u})$ control penalties in [48].

[3]While not precisely the same as considering (1.1) for $d_x \neq d_y$ and a non-invertible map P, we believe that this setting is a first step towards a complete understanding of (1.1). See Section 9 for a discussion.

subject to $\dot{y} = f(y, u)$, where f is of control-affine form. Under the assumption that the running target $(\overline{u}, \overline{y})$ is a steady control-state pair, namely $f(\overline{y}, \overline{u}) = 0$, 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 [46] do not apply, as they require smallness assumptions 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.

- 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 [21, 36, 40, 54] are not applicable.

1.3. **Outline.** The paper is organized as follows. **Section 2** contains statements of our main results in the setting of finite-dimensional, control-affine systems (namely Theorem 2.1, and Corollary 2.1, Corollary 2.2). We also provide a sketch of our new, purely nonlinear strategy in Section 2.2.1. **Section 3** states the extensions of the finite-dimensional results to the setting of the semilinear wave and heat equation (Theorem 3.1 and Theorem 3.2 respectively). **Section 4** provides some preliminary, but key lemmas, mainly to ensure appropriate L_t^∞ bounds for the discrepancy between a trajectory $y(t)$ and the target steady state $\overline{y}$. **Section 5** provides our proofs of the main results in the finite-dimensional case, namely Theorem 2.1 and Corollary 2.2. **Section 6** provides our proof of Theorem 3.1, namely the main result for the semilinear wave equation, which is a straightforward adaptation of the arguments in the finite-dimensional case. **Section 7** presents our proof of Theorem 3.2, namely the main result for the semilinear heat equation. The same proof applies to Corollary 2.1. **Section 9** concludes the paper with a selection of open problems.

1.4. **Notation.** We denote by $\| \cdot \|$ the standard euclidean norm, and $\mathbb{N} := \{1, 2, \ldots\}$. We denote by $\mathrm{Lip}(\mathbb{R})$ the set of functions $f : \mathbb{R} \to \mathbb{R}$ which are globally Lipschitz continuous.

2. Finite-dimensional systems

2.1. **Setup.** Let $d \geqslant 1$ and $m \geqslant 1$. 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

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

with a nonlinearity f of the form

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

where the vector fields $f_0, \ldots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ are only assumed to be globally Lipschitz continuous. This formulation includes (1.3) – see Remark 6 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

$$y(t) = y^0 + \int_0^t f(y(s), u(s)) \, \mathrm{d}s.$$

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

$$J_T(u) := \phi(y(T)) + \int_0^T \|y(t) - \overline{y}\|^2 \, \mathrm{d}t + \int_0^T \|u(t)\|^2 \, \mathrm{d}t, \qquad (2.3)$$

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 $\overline{y} \in \mathbb{R}^d$ is a given running target which we select as an *uncontrolled steady state* of the nonlinear dynamics, namely

$$f_0(\overline{y}) = 0. \qquad (2.4)$$

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 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

$$\inf_{\substack{(y, u) \in \mathbb{R}^d \times \mathbb{R}^m \\ \text{subject to} \ \ f(y,u)=0}} \|y - \overline{y}\|^2 + \|u\|^2. \qquad (2.5)$$

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

Remark 1 (Controlled steady states)**.** *The choice of the running target $\overline{y}$ 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 $\mathcal{L}(u, y) = \|y - \overline{y}\|^2 + \|u - \overline{u}\|^2$ equals zero when evaluated at the optimal steady state. In fact, we could more generally consider the functional*

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

where $(\overline{u}, \overline{y}) \in \mathbb{R}^m \times \mathbb{R}^d$ is chosen so that $f(\overline{y}, \overline{u}) = 0$ (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 $u_T - \overline{u}$). We have taken $\overline{u} = 0$ for presentational simplicity.

In the context of nonlinear optimal control, such turnpike results have been shown in [46] (see also [44]) 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 [46] 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.** 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 a radius $r > 0$ *and a constant* $C(T_0) > 0$ *such that*

$$\inf_{\substack{u \\ such\ that \\ y(0)=y^0, y(T_0)=\overline{y}}} \|u\|_{L^2(0, T_0; \mathbb{R}^m)} \leqslant C(T_0) \left\| y^0 - \overline{y} \right\|, \tag{2.6}$$

and

$$\inf_{\substack{u \\ such\ that \\ y(0)=\overline{y}, y(T_0)=y^1}} \|u\|_{L^2(0, T_0; \mathbb{R}^m)} \leqslant C(T_0) \left\| y^1 - \overline{y} \right\|, \tag{2.7}$$

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

We discuss the feasibility of this assumption later on, in Remark 3. Note that this is not a smallness assumption – it merely stipulates that, inside some ball centered at $\overline{y}$, the cost of controlling from y^0 to $\overline{y}$ and from $\overline{y}$ to y^1 can be estimated by means of the distance of y^0 and y^1 to $\overline{y}$. We may now state our first main result.

Theorem 2.1 (Turnpike)**.** *Assume that* $f_0, \ldots, f_m \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ *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* $\overline{y} \in \mathbb{R}^d$ *be as in* (2.4). *Then there exists a time* $T^* > 0$, *and constants* $C > 0$ *and* $\mu > 0$, *all depending on* $T_0, y^0, \overline{y}$, *such that for any* $T > T^*$, *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*

$$\|y_T(t) - \overline{y}\| \leqslant C \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \tag{2.8}$$

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

$$\|u_T\|_{L^2(0, T; \mathbb{R}^m)} \leqslant C. \tag{2.9}$$

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 $f_1, \ldots, f_m$ are nonzero constants. This is due to the form of the constant provided by Grönwall inequality-based arguments (e.g. in Lemma 4.1 and Lemma 5.2). We delay a discussion of the specific form of the constants to Section 2.3.

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 $y_T(t) - \overline{y}$. 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 decays exponentially in Corollary 2.2, by using the homogeneity of the system with respect to the control.*

Before proceeding with further remarks, which we postpone to Section 2.3, 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 $\overline{y}$.

Corollary 2.1 (Stabilization). *Suppose that $\phi \equiv 0$ in (2.3). Under the assumptions of Theorem 2.1, there exists a time $T^* > 0$, and constants $C > 0$ and $\mu > 0$, all depending on $T_0, y^0, \overline{y}$, such that for any $T > T^*$, 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*

$$\|y_T(t) - \overline{y}\| \leqslant Ce^{-\mu t} \tag{2.10}$$

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

Strictly speaking, we see Corollary 2.1 as a consequence of the strategy of proof of Theorem 2.1, rather than a direct corollary of the statement. Corollary 2.1 may be proven independently of Theorem 2.1 by a simple adaptation of the proof strategy presented in Section 2.2.1, so we omit the proof. This adaptation is transparent in the proof of Theorem 3.2, for which we provide greater detail, albeit for more specific dynamics (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 $\overline{y} \in \mathbb{R}^d$ is an admissible running target for J_T, since $f(\overline{y}, 0) = 0$ for any $\overline{y} \in \mathbb{R}^d$.

Corollary 2.2 (Control decay). *Suppose that $f_0 \equiv 0$ in (2.2) and $\phi \equiv 0$ in (2.3). Under the assumptions of Theorem 2.1, there exists a time $T^* > 0$, and constants $C > 0$ and $\mu > 0$, all depending on $T_0, y^0, \overline{y}$, such that for any $T > T^*$, 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 (2.10) as well as*

$$\|u_T(t)\| \leqslant Ce^{-\mu t} \tag{2.11}$$

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 6 for (1.2)). The proof of Corollary 2.2 (see Section 5.3) 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

$$\frac{1}{2} \int_t^{t+h} \|u_T(s)\|^2 \, \mathrm{d}s \leqslant \int_t^{t+h} \|y_T(s) - \overline{y}\|^2 \, \mathrm{d}s$$

for h small enough, an estimate which, chained with Corollary 2.1 and the Lebesgue differentiation theorem, will suffice to conclude.

2.2.1. *Sketch of the proof of Theorem 2.1.* The proof of Theorem 2.1 may be found in Section 5.2. It roughly follows the following scheme (see Figure 2).

- **Uniform bound of $J_T(u_T)$.** In Lemma 5.1 we show that there exists a constant $C_0 > 0$ independent of T (but depending on $y^0, \overline{y}, f, \phi$) such that

$$J_T(u_T) \leqslant C_0 \tag{2.12}$$

  holds for all $T > 0$. As $\overline{y}$ is a steady state, (2.12) can be shown easily. Indeed, using controllability (without the estimates on the cost) one finds a control $u^\dagger$ such that the solution $y^\dagger$ to (2.1) on $[0, T_0]$ satisfies $y^\dagger(T_0) = \overline{y}$. Setting $u^{\mathrm{aux}}(t) := u^\dagger(t) 1_{[0,T_0]}(t)$ for $t \in [0, T]$, one sees that $y^{\mathrm{aux}}(t) = \overline{y}$ for $t \in [T_0, T]$, whence $J_T(u^{\mathrm{aux}}) = J_{T_0}(u^\dagger)$, and using the inequality $J_T(u_T) \leqslant J_T(u^{\mathrm{aux}})$ yields (2.12).

- **L_t^∞ bound of $y(t) - \overline{y}$.** In Lemma 4.1, we show that

$$\sup_{t \in [0,T]} \|y(t) - \overline{y}\| \leqslant C\Big(\|y(0) - \overline{y}\| + \|y - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)} + \|u\|_{L^2(0,T;\mathbb{R}^m)}\Big) \tag{2.13}$$

  holds for some constant $C > 0$ depending on T solely through $\|u\|_{L^2(0,T;\mathbb{R}^m)}$, in a continuous and increasing manner. The estimate holds for any, not necessarily optimal u. The globally Lipschitz assumption on the dynamics in (2.1) is used precisely for this estimate. Combined with (2.12), estimate (2.13) yields

$$\sup_{t \in [0,T]} \|y_T(t) - \overline{y}\|^2 + J_T(u_T) \leqslant C_1^2 \tag{2.14}$$

  for some constant $C_1 > 0$ independent of T (but depending on $y^0, \overline{y}, f, \phi$).

- **Turnpike away from the middle of $[0, T]$.** Estimate (2.14) yields turnpike for $t \in [0, \tau + T_0] \cup [T - (\tau + T_0), T]$, where $\tau > 0$ is a degree of freedom, independent of T, to be chosen later on, while T_0 is the controllability time for (2.1). Indeed, for $t \in [0, \tau + T_0]$, from (2.14) one sees that

$$\|y_T(t) - \overline{y}\| \leqslant C_1 e^{\mu t}\, e^{-\mu t} \leqslant C_1 e^{\mu(\tau+T_0)}\Big(e^{-\mu t} + e^{-\mu(T-t)}\Big)$$

  holds for all $\mu > 0$, with $C_1 > 0$ as in (2.14) (thus independent of T, τ). A similar computation can be repeated for $t \in [T - (\tau + T_0), T]$. At this point, one already notes that T needs to be chosen sufficiently large, namely,

$$T > 2(\tau + T_0). \tag{2.15}$$

  Actually, $T^* := 2(\tau + T_0)$ in Theorem 2.1.

- **Turnpike in the middle of $[0, T]$.** To obtain the exponential estimate for $t \in [\tau + T_0, T - (\tau + T_0)]$, it is critical to choose $\tau > 0$ large enough. The clue is to first prove that there exists some constant $C_* > 0$ independent of both T and τ such that

$$\sup_{t \in [n\tau, T - n\tau]} \|y_T(t) - \overline{y}\| \leqslant \left(\frac{C_*}{\sqrt{\tau}}\right)^n \tag{2.16}$$

  holds for all integers $1 \leqslant n \leqslant \frac{1}{\tau}\left(\frac{T}{2} - T_0\right)$. See Figure 2 for a graphical depiction. By virtue of (2.15) we have $\frac{1}{\tau}\left(\frac{T}{2} - T_0\right) > 1$, and this upper bound on n will

become clear in the next step. Suppose that (2.16) does indeed hold. Then from (2.16)

$$\|y_T(t) - \overline{y}\| \leqslant \exp\left(-n \log\left(\frac{\sqrt{\tau}}{C_*}\right)\right) \tag{2.17}$$

holds for all $t \in [n\tau, T - n\tau]$ and n as above. Choosing $\tau > C_*^2$, which up to this point was arbitrary, yields the positivity of the logarithm appearing in (2.17):

$$\sigma := \log\left(\frac{\sqrt{\tau}}{C_*}\right) > 0.$$

Now fix $t \in [\tau + T_0, T - (\tau + T_0)]$. We look to choose the integer $n = n(t)$ as to have $t \in [n(t)\tau, T - n(t)\tau]$ as well as $1 \leqslant n(t) \leqslant \frac{1}{\tau}(\frac{T}{2} - T_0)$, so that estimate (2.17) also holds for t fixed as just before. After some elementary computations, one can see that for both these conditions to hold, it is necessary and sufficient for $n(t)$ to be such that

$$1 \leqslant n(t) \leqslant \frac{t}{\tau + T_0} \qquad \text{and} \qquad 1 \leqslant n(t) \leqslant \frac{T - t}{\tau + T_0}.$$

This leads us to set

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

With t fixed as above, and $n(t)$ set as such, one sees that (2.17) holds. Namely, we have

$$\|y_T(t) - \overline{y}\| \leqslant \exp\left(-n(t)\sigma\right). \tag{2.18}$$

But furthermore, one of either

$$n(t) \geqslant \frac{t}{\tau + T_0} - 1 \qquad \text{or} \qquad n(t) \geqslant \frac{T - t}{\tau + T_0} - 1$$

holds by definition of $n(t)$, and so

$$\exp\left(-n(t)\sigma\right) \leqslant \exp(\sigma)\left(\exp\left(-\frac{t}{\tau + T_0}\sigma\right) + \exp\left(-\frac{T - t}{\tau + T_0}\sigma\right)\right). \tag{2.19}$$

Since $t \in [\tau + T_0, T - (\tau + T_0)]$ was arbitrary, chaining (2.18) and (2.19) leads us to the desired turnpike inequality in $[\tau + T_0, T - (\tau + T_0)]$:

$$\|y_T(t) - \overline{y}\| \leqslant e^{\sigma}\left(e^{-\mu t} + e^{-\mu(T-t)}\right),$$

where

$$\mu := \frac{\sigma}{\tau + T_0} = \frac{\log\left(\frac{\sqrt{\tau}}{C_*}\right)}{\tau + T_0} > 0.$$

Thus, the proof would be complete once (2.16) is shown to hold.

- **Proving** (2.16) **by induction.** To show (2.16), one performs an induction over $1 \leqslant n \leqslant \frac{1}{\tau}(\frac{T}{2} - T_0)$. The upper bound on n appears so that there are always a couple of disjoint intervals of length T_0 within every interval of the form $[n\tau, T - n\tau]$. Let us sketch the proof of the initialization stage $n = 1$. Arguing by contradiction (see Lemma 5.3), one can first find $\tau_1 \in [0, \tau)$ and $\tau_2 \in (T - \tau, T]$ such that

$$\|y_T(\tau_j) - \overline{y}\| \overset{\text{Lem.5.3}}{\leqslant} \frac{\|y_T - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)}}{\sqrt{\tau}} \overset{(2.14)}{\leqslant} \frac{C_1}{\sqrt{\tau}}. \tag{2.20}$$
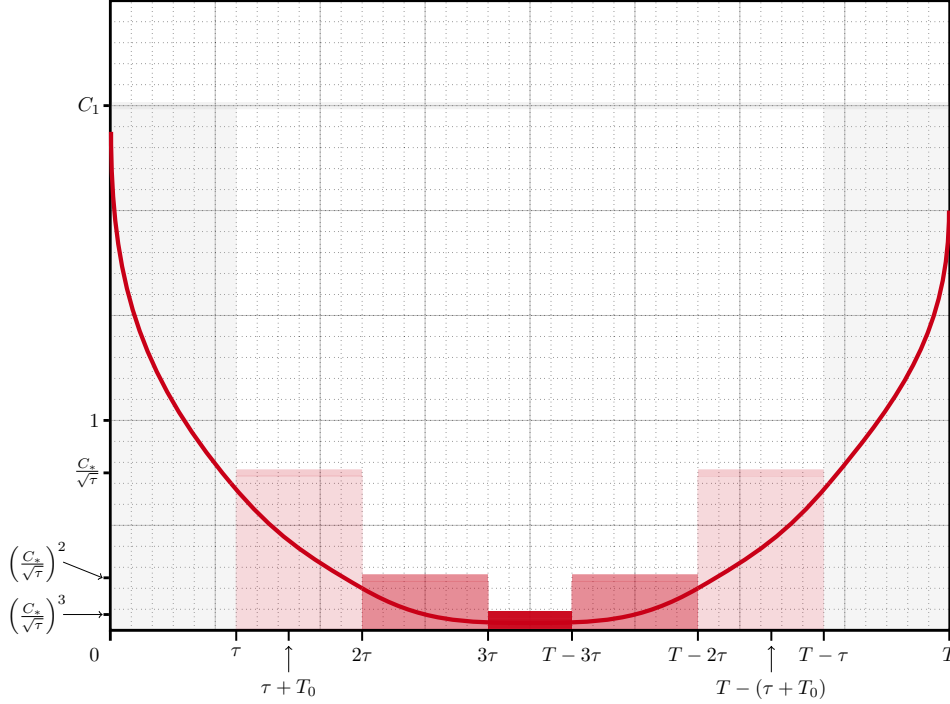
FIGURE 2. The strategy is inspired by the shape of $t \mapsto e^{-t} + e^{-(T-t)}$ on $[0, T]$, which attains its minimum at $t = \frac{T}{2}$. Over successively smaller intervals of the form $[n\tau, T - n\tau]$, with $1 \leqslant n \leqslant \frac{1}{\tau}(\frac{T}{2} - T_0)$, one sees that the upper bound in (2.16), which is uniform in T, decreases exponentially and captures the double-arc exponential of $\|y_T(t) - \overline{y}\|$ (*dotted*) more accurately.

This allows us to recover the desired factor of an inverse of τ. Then, restricting u_T to the subinterval $[\tau_1, \tau_2]$, one sees that it is a solution to

$$\inf_{\substack{u \in L^2(\tau_1, \tau_2; \mathbb{R}^m) \\ \dot{y} = f(y, u) \text{ in } (\tau_1, \tau_2) \\ y(\tau_1) = y_T(\tau_1) \\ y(\tau_2) = y_T(\tau_2)}} \int_{\tau_1}^{\tau_2} \|y(t) - \overline{y}\|^2 \, \mathrm{d}t + \int_{\tau_1}^{\tau_2} \|u(t)\|^2 \, \mathrm{d}t. \qquad (2.21)$$

Just as in the first step of the proof, we look to estimate the functional minimized in (2.21), but this time, not only uniformly in T, τ_1, τ_2, rather also in terms of the distance of $y_T(\tau_j)$ to $\overline{y}$. More specifically, we show that there exists some constant $C(r) > 0$ independent of T, τ_1, τ_2, τ (but depending on r, T_0, f) such that

$$\|y_T(t) - \overline{y}\| \leqslant C(r) \Big(\|y_T(\tau_1) - \overline{y}\| + \|y_T(\tau_2) - \overline{y}\| \Big) \qquad (2.22)$$

holds for all $t \in [\tau_1, \tau_2]$. To this end, we invoke Lemma 4.1, and thus it suffices to bound the functional in (2.21) by means of the right-hand-side in (2.22). To achieve this, since $\tau_2 - \tau_1 \geqslant 2T_0$, we may first control starting from $y_T(\tau_1)$ at time $t = \tau_1$ to $\overline{y}$ in time $t = \tau_1 + T_0$; we stay at $\overline{y}$ by switching off the control until time $t = \tau_2 - T_0$, from which we control to $y_T(\tau_2)$ in time $t = \tau_2$.

By chaining this strategy into one single control[4] u^{aux}, we note that u^{aux} is precisely bounded by the right-hand-side in (2.22) via Assumption 1 (we are in the desired ball by selecting $\tau > C_1^2/r^2$ in (2.20)), while the state tracking term can subsequently be covered by the Grönwall inequality. Estimate (2.22) combined with (2.20) leads to

$$\|y_T(t) - \overline{y}\| \leqslant \frac{2C_1 \cdot C(r)}{\sqrt{\tau}}$$

for all $t \in [\tau_1, \tau_2]$, and thus also for all $t \in [\tau, T - \tau]$. This is precisely (2.16) for $n = 1$. The above argument can then be repeated by induction on smaller intervals $[n\tau, T - n\tau]$ to obtain (2.16) in the general case.

2.3. **Discussion.**   Several pertinent remarks are in order.

Remark 3 (On Assumption 1).   • *In the driftless case ($f_0 = 0$ in (2.2)), the Chow-Rashevskii theorem ([8, 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 [8, 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 [3] 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.*

• *While we suppose that the underlying system is controllable for arbitrarily large data, through (2.6) – (2.7) we solely assume that the cost is proportionate to the distance from the chosen steady state $\overline{y}$ in some, possibly arbitrarily small ball around this steady state. As the estimates (2.6) – (2.7) are more commonly encountered in the linear systems setting $\dot{y} = Ay + Bu$, they thus also hold for semilinear systems where controllability is obtained by perturbation arguments. In such contexts, it is well-known (see e.g. [53, Remark 2.2]) that the minimal L^2–norm control u satisfies*

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

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

• *While there exist necessary and sufficient conditions for ensuring the exact controllability of driftless systems: $\dot{y}(t) = \sum_{j=1}^m u_j(t) f_j(y(t))$, we cannot ensure the validity of estimates (2.6) – (2.7) in the underactuated regime, namely when $m < d$. This is due to the so-called ball-box theorem in sub-Riemannian geometry ([1]), for smooth vector fields $f_1, \ldots, f_m$. This theorem states the following. Suppose that the vector fields $f_1, \ldots, f_m$ satisfy the Hörmander condition, namely that the iterated Lie brackets of these vector fields at any*

---

[4]Such controls are referred to as *quasi-turnpike controls*, since they (and their associated trajectories) look like rough approximations of exact turnpikes (see Figure 4).

point span $\mathbb{R}^d$. *Denote* $\triangle^1(x) := \mathrm{span}\{f_1(x), \ldots, f_m(x)\}$ *for* $x \in \mathbb{R}^d$, *and* $\triangle^{k+1} := \triangle^1 + [\triangle^k, \triangle^1]$ *for* $k \geqslant 1$. *Then by virtue of the Hörmander condition, there exists some* $\kappa \geqslant 1$ *such that* $\triangle^\kappa(x) = \mathbb{R}^d$ *for all* x. *Furthermore, by the ball-box theorem, for* y^0 *close enough to* y^1, *an estimate of the form* $\|y^0 - y^1\| \lesssim d_{\mathrm{SR}}(y^0, y^1) \lesssim \|y^0 - y^1\|^{1/\kappa}$ *holds, where* $d_{\mathrm{SR}}(y^0, y^1)$ *is the sub-Riemannian distance of* y^0 *to* y^1, *equal (modulo a scalar multiple depending on* T_0) *to the* inf *defined in* (2.6) – (2.7). *Herein, one sees that if* $m \geqslant d$, *it may happen to find at least* d *among* m *vector fields which are linearly independent, thus ensuring that* $\kappa = 1$, *as desired; this is quite simply impossible when* $m < d$. *This exact constraint is also encountered in* [12, Theorem 5.1], *where the estimates* (2.6) – (2.7) *are shown to hold for* $m \geqslant d$ *in the driftless setting. In view of this, generalizing the assumptions* (2.6) – (2.7) *to fractional powers of the upper bounds appearing therein is an important open problem. Further clarity regarding this issue is also needed for general control-affine systems beyond semilinear systems, namely those for which linearization techniques might not apply. We refer to* [25, 41] *for developments in this direction.*

Remark 4 (On the time T^*). *Reading the sketch of proof, one notes that* $T^* = 2(T_0 + \tau)$, *where* T_0 *is the controllability time, and* $\tau > 0$ *is chosen sufficiently large. Reading even further, one sees that* τ *needs to be at least larger than* C_1^2/r^2, *where the constant* $C_1 > 0$ *appears in* (2.12) *and* (2.14). *The latter constant is independent of* T *and* r, *but does depend on* T_0 (*and the data* $y^0, \overline{y}$) *through the map*

$$T_0 \mapsto \mathfrak{C}(T_0) := \inf_{\substack{u \\ y(0)=y^0 \\ y(T_0)=\overline{y}}} \left(\|u\|_{L^2(0,T_0;\mathbb{R}^m)} + \|y - \overline{y}\|_{L^2(0,T_0;\mathbb{R}^d)} \right).$$

Actually, due to the innate Grönwall-based argument in Lemma 4.1, C_1 *will roughly be of the form* $C_1 = \mathfrak{C}(T_0) \exp(\mathfrak{C}(T_0))$. *If the system is controllable in any time, the cost* $\mathfrak{C}(T_0)$ *typically explodes as* $T_0 \searrow 0$, *and is bounded for* $T_0 \gg 1$ (*all relative to the distance of* y^0 *to* $\overline{y}$). *Therefore, according to our strategy,* T^* *should increase at least linearly with* $T_0 \gg 1$, *and should explode with* $\mathfrak{C}(T_0)$ *when* $T_0 \searrow 0$. *This discussion also indicates the dependence of* T^* *with respect to the radius* r *of the ball in which the estimates of Assumption 1 hold.*

Remark 5 (On the constants C and μ). *Once again, by reading the sketch of proof, one can see that the constants* $C > 0$ *and* $\mu > 0$ *appearing in the turnpike estimate* (2.8) *are explicit (albeit rather compound). Per the sketch and* (5.33), $\mu > 0$ *is given by*

$$\mu := \frac{\log\left(\frac{\sqrt{\tau}}{4C_\bullet^2}\right)}{\tau + T_0} = \frac{\log\left(\frac{\tau}{16C_\bullet^4}\right)}{T_*},$$

where $C_\bullet = C_\bullet(r, T_0) > 0$ *is the constant appearing in* (2.22); *more specifically, the constant stemming from Lemma 5.2. Moreover,* $\tau > 16C_\bullet^4 + \frac{C_1^2 + 4C_1^2 C_\bullet^2}{r^2}$ *is arbitrary but fixed (as seen in* (5.18)*, where* C_1 *is the same as in Remark 4). On the other hand, per* (5.34), *the constant* $C > 0$ *appearing in* (2.8) *takes the form*

$$C := \max\left\{C_1, \frac{\sqrt{\tau}}{8C_\bullet^2} \max\left\{1, \frac{C_\bullet}{C_1}\right\}\right\}.$$

Remark 6 (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*

$$f(y, u) = \sum_{j=1}^{m} f_j(u_j y) \qquad for \ \ (y, u) \in \mathbb{R}^d \times \mathbb{R}^m \qquad (2.23)$$

where the vector fields $f_1, \ldots, f_m \in \mathrm{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ are additionally assumed to be positively homogeneous of degree 1. Such nonlinearities are motivated by (1.2). Due to the homogeneity of the vector fields in (2.23), the corresponding optimal steady states coincide with those of the driftless case, namely $(u_s, y_s) = (0, \overline{y})$ for any $\overline{y} \in \mathbb{R}^d$.

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 controls. 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

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

Here $f \in \mathrm{Lip}(\mathbb{R})$, $\omega \subset \Omega$ is open (with geometric assumptions given in (3.5)), 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

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

and of the corresponding solution y_T to (3.1). Here $\phi : L^2(\Omega) \to \mathbb{R}_+$ is a given continuous functional, while $\overline{y} \in H_0^1(\Omega)$ is a running target which we select as an uncontrolled steady state of (3.1), namely we assume that $\overline{y}$ is some solution[5] to

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

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 [6, 32] for further results):

---

[5]There is no need for the solution of (3.3) to be unique.

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

$$|f(s)| \leqslant C(1 + |s|^p) \qquad \text{for all } s \in \mathbb{R}$$

$$-\int_0^s f(\zeta)\,\mathrm{d}\zeta \leqslant \frac{\sigma}{2}\, s^2 \qquad \text{for } |s| \ll 1$$

$$0 < -\theta \int_0^s f(\zeta)\,\mathrm{d}\zeta \leqslant -s \int_0^s f(\zeta)\,\mathrm{d}\zeta \qquad \text{for } |s| \gg 1$$

then a nontrivial solution $\overline{y} \in H_0^1(\Omega)$, $\overline{y} \not\equiv 0$ also exists. We refer to [6, 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 [6, Theorem 1.2.3].

The case of a controlled steady state (namely adding $\overline{u}\mathbf{1}_\omega$ 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 $\overline{y}$ is fixed as above, the pair $(u_s, y_s) \equiv (0, \overline{y})$ is the unique solution to the steady optimal control problem

$$\inf_{\substack{(y,u)\in H_0^1(\Omega)\times L^2(\omega) \\ y \text{ solves } (3.4)}} \|y - \overline{y}\|_{H_0^1(\Omega)}^2 + \|u\|_{L^2(\omega)}^2$$

where the steady equation is

$$\begin{cases} -\Delta y + f(y) = u\mathbf{1}_\omega & \text{in } \Omega \\ y = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.4}$$

This is because the functional in the expression above attains its minimum, equal to 0, precisely at $(0, \overline{y})$, 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 $d \geqslant 2$. For any fixed $x_\circ \in \mathbb{R}^d \setminus \overline{\Omega}$, we define

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

where $\nu(x)$ denotes the outward unit normal at $x \in \partial\Omega$. The set $\Gamma(x_\circ)$ coincides with the subset of the boundary arising usually in the context of the multiplier method [31]. We will suppose that for some $\delta > 0$ and $x_\circ \in \mathbb{R}^d \setminus \overline{\Omega}$,

$$\omega = \mathcal{O}_\delta(\Gamma(x_\circ)) \cap \Omega, \tag{3.5}$$

where $\mathcal{O}_\delta(\Gamma(x_\circ)) := \left\{x \in \mathbb{R}^d : \|x - x'\| < \delta \text{ for some } x' \in \Gamma(x_\circ)\right\}$. It is known that, under these geometric assumptions on ω, and since $f \in \mathrm{Lip}(\mathbb{R})$, the wave equation (3.1) is exactly controllable in any time $T > T_{\min}(\omega, \Omega)$, where

$$T_{\min}(\omega, \Omega) = 2 \max_{x \in \overline{\Omega}} \|x - x_\circ\|. \tag{3.6}$$

(See [14, 51, 52], and also the introduction of [26] for an ample survey of controllability results for semilinear wave equations.) We may now state our main result in the context of the wave equation.

Theorem 3.1 (Turnpike). *Suppose that $f \in \mathrm{Lip}(\mathbb{R})$ and $f(0) = 0$. Let $\overline{y} \in H_0^1(\Omega)$ be any solution to (3.3). Let $\phi \in \mathcal{L}(L^2(\Omega); \mathbb{R}_+)$, and suppose that ω is as in (3.5). Then 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 $C > 0$ and $\mu > 0$, such that for any $T > T^*$, 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*

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

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

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

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 bounds through quasi-turnpike controls, wherein one uses the Duhamel formula for mild solutions in view of applying an integral Grönwall inequality-based argument, in the spirit of the ODE setting. The assumption $f(0) = 0$ is of technical nature, and is clarified in Remark 10.

Remark 7 (On the choice of J_T). *We note that in existing turnpike results for the wave equation, e.g. [23, 45, 54], a slightly weaker functional is sometimes considered. For instance, in [54] for the linear wave equation, only the $L^2(0, T; H_0^1(\Omega))$–norm of $y - \overline{y}$ 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), the strategy for proving Theorem 3.1 can be slightly tweaked to obtain an exponential stabilization property for the optimal states.

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 $C > 0$ and $\mu > 0$, such that for any $T > T^*$, 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*

$$\|y_T(t) - \overline{y}\|_{H_0^1(\Omega)} + \|\partial_t y_T(t)\|_{L^2(\Omega)} \leqslant C e^{-\mu t}$$

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

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

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

$$\begin{cases} \partial_t y - \Delta y + f(y) = u\mathbf{1}_\omega & \text{in } (0, T) \times \Omega \\ y = 0 & \text{on } (0, T) \times \partial\Omega \\ y|_{t=0} = y^0 & \text{in } \Omega, \end{cases} \tag{3.7}$$

were $f \in \mathrm{Lip}(\mathbb{R})$, $\omega \subset \Omega$ 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 (3.7). We will again study global minimizers $u_T \in L^2((0,T) \times \omega)$ to nonnegative functionals of the form

$$J_T(u) := \int_0^T \|y(t) - \overline{y}\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \int_0^T \|u(t)\|_{L^2(\omega)}^2 \, \mathrm{d}t, \tag{3.8}$$

and the corresponding solution y_T to (3.7) in the regime $T \gg 1$. Once again, $\overline{y} \in L^2(\Omega)$ 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 (3.8) follows by the direct method in the calculus of variations.

Theorem 3.2 (Stabilization). *Suppose that $f \in \mathrm{Lip}(\mathbb{R})$ and $\Omega \subset \mathbb{R}^d$ are such that (3.3) admits at least one solution, and let $\overline{y} \in H_0^1(\Omega)$ be any such solution. Then for any $y^0 \in L^2(\Omega)$, there exist constants $C, \mu > 0$ (depending solely on f, ω) such that for any $T > 0$, any global minimizer $u_T \in L^2((0,T) \times \omega)$ of J_T defined in (3.8) and corresponding optimal state y_T solution to (3.7) satisfy*

$$\|y_T(t) - \overline{y}\|_{L^2(\Omega)} \leqslant C e^{-\mu t} \|y^0 - \overline{y}\|_{L^2(\Omega)}$$

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

$$\|u_T\|_{L^2((0,T) \times \omega)} \leqslant C \|y^0 - \overline{y}\|_{L^2(\Omega)}.$$

We refer to Section 7 for the proof. We consider the heat equation, in addition to the wave equation, because of the validity (or rather, the partial lack thereof) of the PDE analog of Assumption 1. The heat equation is exactly controllable to controlled trajectories, namely solutions $\widehat{y}$ to (3.7) for given controls $\widehat{u}$. Instead of an estimate such as (2.7), one has

$$\|u - \widehat{u}\|_{L^2((0,T_0) \times \omega)} \leqslant C(T_0) \|y^0 - \widehat{y}(0)\|_{L^2(\Omega)}$$

for minimal L^2–norm controls u steering y to $\widehat{y}$ in time T_0 (see [38, Lemma 8.3] and the references therein). 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. The constants $C > 0$ and $\mu > 0$ are actually explicit (see (7.15), (7.16)) precisely due to the global validity of this estimate (i.e. (7.1)); the factor $\|y^0 - \overline{y}\|$ also appears because of this and due to the absence of a final cost ϕ.

The semilinear heat equation is a commonly used benchmark for nonlinear turnpike results, thus this example serves to compare with existing results. For instance, while we assume that the running targets are steady states, we make no smallness assumptions on the targets or on the initial data, unlike [21, 40]. Furthermore, since we do not use (or thus linearize) the optimality system, we may work with solely globally Lipschitz nonlinearities, in which case the techniques of [21, 36, 40] do not apply.

Remark 8 (On the nonlinearity). *The assumption that f is globally Lipschitz in (3.1) and (3.7) 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 [36] under smallness assumptions on the target. We refer to the end of Section 9.1 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) – (3.7) *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 $\|y(t) - \overline{y}\|$ for both the nonlinear ODE and PDE setting, solely by means of $\|y^0 - \overline{y}\|$ 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 $\|y(t) - \overline{y}\|$. Let us begin with the ODE estimate.

Lemma 4.1. *Let $T > 0$, and let $\overline{y} \in \mathbb{R}^d$ be as in* (2.4). *For any $u \in L^2(0, T; \mathbb{R}^m)$ 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 $C_1 = C_1(f, \overline{y}) > 0$ and $C_2 = C_2(f)$ independent of T such that*

$$\sup_{t \in [0,T]} \|y(t) - \overline{y}\| \leqslant C \left(\|y^0 - \overline{y}\| + \|u\|_{L^2(0,T;\mathbb{R}^m)} + \|y - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)} \right)$$

holds, 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 inequality-based 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 on $[0, t]$, namely writing

$$y(t) - \overline{y} = y^0 - \overline{y} + \int_0^t \left(f_0(y) + \sum_{j=1}^m u_j f_j(y) \right) \mathrm{d}\tau$$

$$= y^0 - \overline{y} + \int_0^t (f_0(y) - f_0(\overline{y})) \, \mathrm{d}\tau + \int_0^t \sum_{j=1}^m u_j \left(f_j(y) - f_j(\overline{y}) \right) \mathrm{d}\tau$$

$$+ \int_0^t \sum_{j=1}^m u_j f_j(\overline{y}) \, \mathrm{d}\tau,$$

we see that, by using the fact that $f_0, \dots, f_m \in \mathrm{Lip}(\mathbb{R}^d; \mathbb{R}^d)$ and the Cauchy-Schwarz inequality for the sums,

$$\|y(t) - \overline{y}\| \leqslant \|y^0 - \overline{y}\| + C(f) \int_0^t \left(1 + \|u(\tau)\| \right) \|y(\tau) - \overline{y}\| \, \mathrm{d}\tau + C_1(f, \overline{y}) \int_0^t \|u(\tau)\| \, \mathrm{d}\tau$$

holds for some constant $C_1(f, \overline{y}) > 0$ independent of T. Here and henceforth, $C(f) > 0$ designates the largest among the Lipschitz constants of all $f_0, \dots, f_m$. Now applying the Cauchy-Schwarz inequality for the last term, the fact that $t \leqslant 1$, and the Grönwall inequality, in conjunction to the estimate just above, lead us to

$$\|y(t) - \overline{y}\| \leqslant C_2 \exp \left(C(f) \sqrt{1 + \int_0^t \|u(\tau)\|^2 \, \mathrm{d}\tau} \right) \left(\|y^0 - \overline{y}\| + \|u\|_{L^2(0,T;\mathbb{R}^m)} \right),$$

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

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

To this end, we argue by contradiction. Suppose that

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

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

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

which is a contradiction. 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

$$y(t) - \overline{y} = y(t^*) - \overline{y} + \int_{t^*}^t \left(f_0(y) + \sum_{j=1}^m u_j f_j(y) \right) \mathrm{d}\tau$$

$$= y(t^*) - \overline{y} + \int_{t^*}^t (f_0(y) - f_0(\overline{y})) \, \mathrm{d}\tau + \int_{t^*}^t \sum_{j=1}^m u_j \Big(f_j(y) - f_j(\overline{y}) \Big) \mathrm{d}\tau$$

$$+ \int_{t^*}^t \sum_{j=1}^m u_j f_j(\overline{y}) \, \mathrm{d}\tau,$$

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

$$\|y(t) - \overline{y}\| \leqslant \|y(t^*) - \overline{y}\| + C(f) \int_{t^*}^t \Big(1 + \|u(\tau)\| \Big) \|y(\tau) - \overline{y}\| \, \mathrm{d}\tau + C_1(f, \overline{y}) \int_{t^*}^t \|u(\tau)\| \, \mathrm{d}\tau.$$

Now applying the Cauchy-Schwarz inequality for the last term, the fact that $t - t^* \leqslant 1$, (4.1), and the Grönwall inequality, in conjunction to the estimate just above, we obtain

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

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

Remark 9. *Let us make two brief observations.*

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

$$\sup_{t \in [0,T]} \|y(t) - \overline{y}\| \leqslant C \left(\|y^0 - \overline{y}\| + \|u - \overline{u}\|_{L^2(0,T;\mathbb{R}^m)} + \|y - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)} \right)$$

  with $C \sim \exp \left(\|u - \overline{u}\|_{L^2(0,T;\mathbb{R}^m)} \right)$. Obtaining a dependence of the constant C with respect to $\|u - \overline{u}\|_{L^2(0,T;\mathbb{R}^m)}$ 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 $f_0(\overline{y}) + \overline{u} f_1(\overline{y}) = 0$,

$$y(t) - \overline{y} = y(t^*) - \overline{y} + \int_{t^*}^t (f_0(y) - f_0(\overline{y}))\, \mathrm{d}s + \int_{t^*}^t (u - \overline{u})(f_1(y) - f_1(\overline{y}))\, \mathrm{d}s$$

$$+ \int_{t^*}^t (u - \overline{u}) f_1(\overline{y})\, \mathrm{d}s + \int_{t^*}^t \overline{u} (f_1(y) - f_1(\overline{y}))\, \mathrm{d}s.$$

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 $f_1, \ldots, f_m$ 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 (3.7).

Lemma 4.2. *Let $T > 0$ be given, and let $\overline{y}$ 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 (3.7). Then there exists a constant $C = C(f) > 0$ independent of T such that*

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

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

The proof is almost identical to the ODE case, but we sketch it for the sake of clarity.

Proof of Lemma 4.2. The proof closely follows that of Lemma 4.1. We first note that by uniqueness, $y - \overline{y}$ can be shown (see [2]) to coincide with the unique mild solution to

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

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

$$y(t) - \overline{y} = e^{t\Delta}(y^0 - \overline{y}) + \int_0^t e^{(t-s)\Delta} u(s)\mathbf{1}_\omega\, \mathrm{d}s - \int_0^t e^{(t-s)\Delta}\left(f(y) - f(\overline{y})\right) \mathrm{d}s, \quad (4.2)$$

where $\left\{e^{t\Delta}\right\}_{t\geqslant 0}$ denotes the heat semigroup on $L^2(\Omega)$ generated by the Dirichlet Laplacian $-\Delta : H^2(\Omega) \cap H_0^1(\Omega) \to L^2(\Omega)$. 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 $0 < t \leqslant 1$. Using the well-known decay $\left\|e^{t\Delta}\right\|_{\mathcal{L}(L^2(\Omega))} \leqslant e^{-\lambda_1(\Omega)t} \leqslant 1$ of the heat semigroup, where $\lambda_1(\Omega) > 0$ denotes the first eigenvalue of $-\Delta$, and the Lipschitz character of f, we find using (4.2) that

$$\|y(t) - \overline{y}\|_{L^2(\Omega)} \leqslant \left\|e^{t\Delta}\left(y^0 - \overline{y}\right)\right\|_{L^2(\Omega)} + \int_0^t \left\|e^{(t-s)\Delta} u(s)\right\|_{L^2(\omega)} \mathrm{d}s$$

$$+ \int_0^t \left\|e^{(t-s)\Delta}\left(f(y(s)) - f(\overline{y})\right)\right\|_{L^2(\Omega)} \mathrm{d}s$$

$$\leqslant \|y^0 - \overline{y}\|_{L^2(\Omega)} + \int_0^t \|u(s)\|_{L^2(\omega)} \mathrm{d}s$$

$$+ C_0 \int_0^t \|y(t) - \overline{y}\|_{L^2(\Omega)} \mathrm{d}s,$$

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

$$\|y(t^*) - \overline{y}\|_{L^2(\Omega)} \leqslant \|y - \overline{y}\|_{L^2((0,T) \times \Omega)} \tag{4.3}$$

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

$$y(t) - \overline{y} = e^{t\Delta}(y(t^*) - \overline{y}) + \int_{t^*}^t e^{(t-s)\Delta} u(s) \mathbf{1}_\omega \, ds - \int_{t^*}^t e^{(t-s)\Delta}\Big(f(y) - f(\overline{y})\Big) \, ds$$

we see just as before that

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

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

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 $\overline{y}$ 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*

$$\|y(t) - \overline{y}\|_{H_0^1(\Omega)} + \|\partial_t y(t)\|_{L^2(\Omega)} \leqslant C\Big(\big\|y_1^0 - \overline{y}\big\|_{H_0^1(\Omega)} + \big\|y_2^0\big\|_{L^2(\Omega)} + \|u\|_{L^2((0,T) \times \omega)}$$
$$+ \|y - \overline{y}\|_{L^2(0,T; H_0^1(\Omega))} + \|\partial_t y\|_{L^2((0,T) \times \Omega)} \Big)$$

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

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

where $D(\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$. The operator A is skew-adjoint and thus generates a strongly continuous semigroup $\{e^{tA}\}_{t \geqslant 0}$ in X by virtue of the Stone-Lumer-Phillips theorem (see e.g. [47, Theorem 3.8.6]). We now denote

$$\mathbf{y} := \begin{bmatrix} y \\ \partial_t y \end{bmatrix}, \qquad \overline{\mathbf{y}} := \begin{bmatrix} \overline{y} \\ 0 \end{bmatrix}.$$

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

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

for $t > 0$ is the unique mild solution to the equation satisfied by the perturbation $\mathbf{y} - \overline{\mathbf{y}}$. 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 $t \geqslant 0$ 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. $\qquad\square$

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 $\overline{y}$ as a steady state of the nonlinear system without control, and on the 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 of 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 $C = C(f, \phi, T_0, \overline{y}, y^0) > 0$ independent of $T > 0$ such that*

$$\|u_T\|_{L^2(0,T;\mathbb{R}^m)} + \|y_T - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)} + \|y_T(t) - \overline{y}\| \leqslant C \qquad (5.1)$$
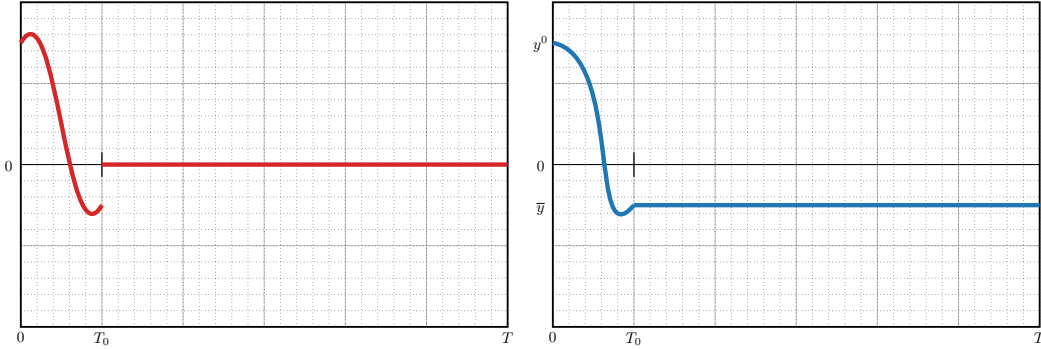
holds for all $t \in [0, T]$.

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 construct a quasi-turnpike control u^{aux} (red), for which the corresponding state y^{aux} (blue) coincides with $\overline{y}$ over (T_0, T). In this way, as $J_T(u_T) \leqslant J_T(u^{\mathrm{aux}})$, and $J_T(u^{\mathrm{aux}})$ is independent of T, we can conclude. The estimate of the third term then follows from Lemma 4.1.

Proof of Lemma 5.1. Case 1). We begin by considering the case $T \geqslant T_0$. 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

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

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

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

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

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

As the right-hand side in the above inequality is clearly independent of T, and depends solely on the $L^2(0, T_0)$ cost of controlling from y^0 to $\overline{y}$ in time T_0, 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 $T > 0$.

Case 2). Now suppose that $T \leqslant T_0$. In this case, we use $\phi \geqslant 0$ and the optimality inequality $J_T(u_T) \leqslant J_T(u_{T_0})$ with the effect of obtaining

$$\|y_T - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)}^2 + \|u_T\|_{L^2(0,T;\mathbb{R}^m)}^2 \leqslant \phi\left(y_{T_0}(T)\right) + \|y_{T_0} - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)}^2 + \|u_{T_0}\|_{L^2(0,T;\mathbb{R}^m)}^2$$

Now $y_{T_0} \in C^0([0, T_0]; \mathbb{R}^d)$ is uniformly bounded with respect to $T \in [0, T_0]$, whence, using the continuity of ϕ, as well as $T \leqslant T_0$, we may conclude that

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

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

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

$$J_{\tau_1,\tau_2}(u) := \int_{\tau_1}^{\tau_2} \|y(t) - \overline{y}\|^2 \, \mathrm{d}t + \int_{\tau_1}^{\tau_2} \|u(t)\|^2 \, \mathrm{d}t \tag{5.2}$$

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

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

and

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

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 $\overline{y}$. This estimate will be the cornerstone of the bootstrap argument performed in the proof of Theorem 2.1.

Lemma 5.2. *Let $\overline{y} \in \mathbb{R}^d$ 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, let $0 \leqslant \tau_1 < \tau_2 \leqslant T$ be fixed such that $\tau_2 - \tau_1 \geqslant 2T_0$, and let $y^{\tau_1}, y^{\tau_2} \in \mathbb{R}^d$ be such that*

$$\|y^{\tau_i} - \overline{y}\| \leqslant r$$

for $i = 1, 2$. Suppose $u_T \in U_{\mathrm{ad}}$ is 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 $C = C(f, T_0, \overline{y}, r) > 0$ independent of $T, \tau_1, \tau_2 > 0$ such that

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

holds for all $t \in [\tau_1, \tau_2]$. Moreover, the map $r \longmapsto C(f, T_0, \overline{y}, r)$ 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 a quasi-turnpike control (steering the corresponding trajectory from y^{τ_1} to y^{τ_2} in time $\tau_2 - \tau_1$, whilst remaining at $\overline{y}$ 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.
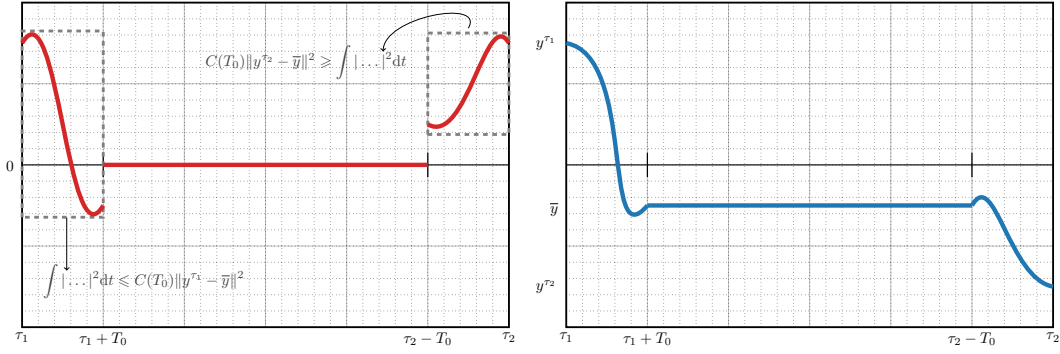

FIGURE 4. **Proof of Lemma 5.2.** The first two terms appearing in the estimate implied by Lemma 5.2 also appear in the functional $J_{\tau_1, \tau_2}(u_T)$. We construct a quasi-turnpike control u^{aux} (red), for which the corresponding state y^{aux} (blue) coincides with $\overline{y}$ over $(\tau_1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

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

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

$$\left\| u^\dagger \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}^2 \leqslant C(T_0) \left\| y^{\tau_1} - \overline{y} \right\|^2, \qquad (5.4)$$

for some $C(T_0) > 0$ independent of y^{τ_1}, τ_1, and which is such that the corresponding solution $y^\dagger$ to

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

satisfies $y^\dagger(\tau_1 + T_0) = \overline{y}$. By integrating (5.5), and using the Lipschitz character of $f_0, \ldots, f_m$, the Grönwall inequality, the Cauchy-Schwarz inequality, and

(5.4), we see that

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

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 $f_0(\overline{y}) = 0$, the Cauchy-Schwarz inequality and (5.6), we moreover see that

$$
\begin{aligned}
\left\| y^\dagger(t) - \overline{y} \right\| &\leqslant \left\| y^{\tau_1} - \overline{y} \right\| + \int_{\tau_1}^t \sum_{j=1}^m \left| u_j^\dagger(s) \right| \left\| f_j(y^\dagger) \right\| \, \mathrm{d}s + \int_{\tau_1}^t \left\| f(y^\dagger) - f(\overline{y}) \right\| \, \mathrm{d}s \\
&\leqslant \left\| y^{\tau_1} - \overline{y} \right\| + C_3 \left\| u^\dagger \right\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^m)} + C(f) \int_{\tau_1}^t \left\| y^\dagger(s) - \overline{y} \right\| \, \mathrm{d}s
\end{aligned}
\tag{5.7}
$$

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

$$
\left\| y^\dagger(t) - \overline{y} \right\| \leqslant C_4 \exp \left(C(f) T_0 \right) \left\| y^{\tau_1} - \overline{y} \right\|
\tag{5.8}
$$

for some $C_4(f, T_0, \overline{y}, r) > 0$ independent of T, τ_1 and $\tau_2 > 0$, 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

$$
\left\| u^\ddagger \right\|_{L^2(\tau_1, \tau_1+T_0; \mathbb{R}^m)}^2 \leqslant C(T_0) \left\| \overline{y} - y^{\tau_2} \right\|^2,
\tag{5.9}
$$

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

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

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

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

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 $f_0(\overline{y}) = 0$, the Cauchy-Schwarz

inequality and (5.11), we moreover see that

$$
\left\| y^{\ddagger}(t) - \overline{y} \right\| \leqslant \int_{\tau_1}^{t} \sum_{j=1}^{m} \left| u_j^{\ddagger}(s) \right| \left\| f_j(y^{\ddagger}) \right\| \, \mathrm{d}s + \int_{\tau_1}^{t} \left\| f\left(y^{\ddagger}\right) - f(\overline{y}) \right\| \, \mathrm{d}s
$$

$$
\leqslant C_7 \left\| u^{\ddagger} \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + C(f) \int_{\tau_1}^{t} \left\| y^{\ddagger}(s) - \overline{y} \right\| \, \mathrm{d}s \tag{5.12}
$$

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

$$
\left\| y^{\ddagger}(t) - \overline{y} \right\| \leqslant C_8 \exp\left(C(f)T_0\right) \left\| y^{\tau_2} - \overline{y} \right\| \tag{5.13}
$$

for some $C_8(f, T_0, \overline{y}, r) > 0$ 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

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

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

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

and thus

$$
y^{\mathrm{aux}}(t) = \overline{y} \qquad \text{in } [\tau_1 + T_0, \tau_2 - T_0], \tag{5.14}
$$

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

$$
J_{\tau_1, \tau_2}(u^{\mathrm{aux}}) = \left\| u^{\dagger} \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)} + \left\| u^{\ddagger} \right\|_{L^2(\tau_1, \tau_1 + T_0; \mathbb{R}^m)}
$$

$$
+ \int_{\tau_1}^{\tau_1 + T_0} \left\| y^{\dagger}(t) - \overline{y} \right\|^2 \, \mathrm{d}t + \int_{\tau_1}^{\tau_1 + T_0} \left\| y^{\ddagger}(t) - \overline{y} \right\|^2 \, \mathrm{d}t
$$

$$
\leqslant C_9 \left(\left\| \overline{y} - y^{\tau_1} \right\|^2 + \left\| \overline{y} - y^{\tau_2} \right\|^2 \right) \tag{5.15}
$$

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

$$
\left\| y_T - \overline{y} \right\|^2_{L^2(\tau_1, \tau_2; \mathbb{R}^d)} + \left\| u_T \right\|^2_{L^2(\tau_1, \tau_2; \mathbb{R}^m)} \leqslant J_{\tau_1, \tau_2}(u_T) \leqslant J_{\tau_1, \tau_2}(u^{\mathrm{aux}})
$$

$$
\leqslant C_9 \left(\left\| \overline{y} - y^{\tau_1} \right\|^2 + \left\| \overline{y} - y^{\tau_2} \right\|^2 \right).
$$

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. $\qquad \square$

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 $\tau \leqslant \frac{T}{2}$, there exist $t_1 \in [0, \tau)$ and $t_2 \in (T - \tau, T]$ such that*

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

Proof of Lemma 5.3. Denote

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

We argue by contradiction. Assume that either

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

or

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

hold. Then we have

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

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. $\qquad \square$

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 from Lemma 5.1. We thus concentrate on proving (2.8), and we split the proof in two parts. Before proceeding, let us first note that by Lemma 5.1, there exists a constant $C_1 > 0$, depending only on $f, T_0, \overline{y}, y^0, \phi$, such that

$$\sup_{t \in [0,T]} \|y_T(t) - \overline{y}\| + \|y_T - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)} \leqslant C_1 \tag{5.16}$$

holds for any $T > 0$ and pair (u_T, y_T) which is optimal for (2.3). Let $r > 0$ be the radius provided by Assumption 1. By Lemma 5.2, we also know that there exists a constant $C_2 > 0$, depending only on $r, f, T_0, \overline{y}$, such that whenever $T \geqslant 2T_0$, for any $\tau_1, \tau_2 \in [0, T]$ such that $\tau_2 - \tau_1 \geqslant 2T_0$ and

$$\|y_T(\tau_i) - \overline{y}\| \leqslant r,$$

the estimate

$$\sup_{t \in [\tau_1, \tau_2]} \|y_T(t) - \overline{y}\| + \|y_T - \overline{y}\|_{L^2(\tau_1, \tau_2; \mathbb{R}^d)} \leqslant C_2 \Big(\|y_T(\tau_2) - \overline{y}\| + \|y_T(\tau_1) - \overline{y}\| \Big) \tag{5.17}$$

holds for any pair (u_T, y_T) which is optimal for (5.2). Now fix

$$\tau > 16 C_2^4 + \frac{C_1^2}{r^2} + \frac{4 C_1^2 C_2^2}{r^2}, \tag{5.18}$$

and let

$$T > 2(\tau + T_0) := T^*$$

be fixed. Let (u_T, y_T) thus be an optimal pair for (2.3) with T as such. The choice of the buffer time τ will become clear in what follows.

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

$$\|y_T(t) - \overline{y}\| \leqslant C_1 e^{\mu t} \, e^{-\mu t}$$
$$\leqslant C_1 e^{\mu(\tau + T_0)} \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \tag{5.19}$$

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

$$\|y_T(t) - \overline{y}\| \leqslant C_1 e^{\mu(T-t)} \, e^{-\mu(T-t)}$$
$$\leqslant C_1 e^{\mu(\tau + T_0)} \left(e^{-\mu t} + e^{-\mu(T-t)} \right) \tag{5.20}$$

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

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

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

$$\|y_T(\tau_i) - \overline{y}\| \leqslant \frac{\|y_T - \overline{y}\|_{L^2(0,T;\mathbb{R}^d)}}{\sqrt{\tau}} \overset{(5.16)}{\leqslant} \frac{C_1}{\sqrt{\tau}}. \tag{5.21}$$

Note that, by virtue of the choice of τ in (5.18), we have that $\frac{C_1}{\sqrt{\tau}} \leqslant r$ and thus

$$\|y_T(\tau_i) - \overline{y}\| \leqslant r \tag{5.22}$$

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 minimizer[6] 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

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

in view of (5.22), we may use (5.17) to find that

$$\|y_T(t) - \overline{y}\| \leqslant C_2 \Big(\|y_T(\tau_1) - \overline{y}\| + \|y_T(\tau_2) - \overline{y}\| \Big) \tag{5.23}$$

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

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

and applying (5.21) to inequality (5.23), we deduce that

$$\|y_T(t) - \overline{y}\| \leqslant \frac{2C_1 C_2}{\sqrt{\tau}} \tag{5.24}$$

$$\leqslant \frac{\kappa}{2} \frac{4C_2^2}{\sqrt{\tau}} \tag{5.25}$$

holds for all $t \in [\tau_1, \tau_2]$. As $\tau_1 \leqslant \tau$ and $T - \tau \leqslant \tau_2$, estimates (5.24) and (5.25) clearly hold for all $t \in [\tau, T - \tau]$.

Step 2): Bootstrap. Inequality (5.25) motivates performing a bootstrap – we will show that for any $n \in \mathbb{N}$ satisfying

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

---

[6]This can be shown by contradiction.

one has

$$\sup_{t \in [n\tau, T - n\tau]} \|y_T(t) - \overline{y}\| \leqslant \frac{\kappa}{2} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n. \tag{5.26}$$

The choice of n is done as to guarantee that $T - 2n\tau \geqslant 2T_0$ in view of a repeated application of Lemma 5.2 (namely (5.17)). Note that (5.24), combined with the choice of τ in (5.18), also implies that

$$\|y_T(t) - \overline{y}\| \leqslant r \tag{5.27}$$

for all $t \in [\tau, T - \tau]$. To prove (5.26), we proceed by induction. The case $n = 1$ clearly holds by (5.25). Thus, assume that (5.26) holds – we aim to show that (5.26) holds at step $n + 1$. To this end, suppose that

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

This is equivalent to $T - 2n\tau - 2T_0 \geqslant 2\tau$ (and recall that $\tau > 0$ is fixed), and it also clearly implies that

$$\tau \leqslant \frac{T - 2n\tau}{2}. \tag{5.28}$$

Since $T - 2n\tau \geqslant 2T_0$, as in Step 1, it can be shown 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 (5.27) holds[7], we can apply Lemma 5.3 on $[n\tau, T - n\tau]$ (noting (5.28)), 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

$$\|y_T(t_i) - \overline{y}\| \leqslant \frac{\|y_T - \overline{y}\|_{L^2(n\tau, T - n\tau; \mathbb{R}^d)}}{\sqrt{\tau}} \leqslant \frac{C_2}{\sqrt{\tau}} \left(\|y_T(n\tau) - \overline{y}\| + \|y_T(T - n\tau) - \overline{y}\| \right).$$

We now use the induction hypothesis (5.26) in the above inequality to obtain

$$\|y_T(t_i) - \overline{y}\| \leqslant \kappa \frac{C_2}{\sqrt{\tau}} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n \tag{5.29}$$

Now since

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

and since $u_T|_{[t_1, t_2]}$ is a global minimizer of J_{t_1, t_2} defined in (5.2), combining Lemma 5.2[8] and (5.29) we are led to deduce that

$$
\begin{aligned}
\|y_T(t) - \overline{y}\| &\leqslant C_2 \Big(\|y_T(t_1) - \overline{y}\| + \|y_T(t_2) - \overline{y}\| \Big) \\
&\leqslant \frac{\kappa}{2} \frac{4C_2^2}{\sqrt{\tau}} \left(\frac{4C_2^2}{\sqrt{\tau}} \right)^n
\end{aligned}
\tag{5.30}
$$

for $t \in [t_1, t_2]$. Since $t_1 < (n+1)\tau$ and $T - (n+1)\tau < t_2$, estimate (5.30) clearly also holds for $t \in [(n+1)\tau, T - (n+1)\tau]$. Identity (5.26) is thus proven.

Step 3): Conclusion. We now look to use (5.26) as to conclude the proof. Suppose that $t \in [\tau + T_0, T - (\tau + T_0)]$. We set

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

---

[7]Note that $n\tau \geqslant \tau$ and $T - n\tau \leqslant T - \tau$, so (5.27) also holds for $t \in [n\tau, T - n\tau]$, hence Lemma 5.2 is applicable.

[8]May be applied once again since (5.27) holds for $t = t_1 \geqslant \tau$ and $t = t_2 \leqslant T - \tau$.

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

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

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

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

We may then apply (5.26) to obtain

$$\|y_T(t) - \overline{y}\| \leqslant \frac{\kappa}{2}\left(\frac{4C_2^2}{\sqrt{\tau}}\right)^{n(t)}. \tag{5.31}$$

By virtue of the choice of τ in (5.18), we see that

$$\frac{4C_2^2}{\sqrt{\tau}} < 1.$$

Moreover, since either

$$n(t) \geqslant \frac{t}{\tau + T_0} - 1 \qquad \text{or} \qquad n(t) \geqslant \frac{T-t}{\tau + T_0} - 1,$$

we may rewrite (5.31) to obtain

$$\|y_T(t) - \overline{y}\| \leqslant \frac{\kappa}{2}\exp\left(-n(t)\log\left(\frac{\sqrt{\tau}}{4C_2^2}\right)\right)$$

$$\leqslant \frac{\kappa}{2}\frac{\sqrt{\tau}}{4C_2^2}\left(\exp\left(-\frac{\log\left(\frac{\sqrt{\tau}}{4C_2^2}\right)}{\tau + T_0}t\right) + \exp\left(-\frac{\log\left(\frac{\sqrt{\tau}}{4C_2^2}\right)}{\tau + T_0}(T-t)\right)\right). \tag{5.32}$$

Looking at (5.32), we see that estimate (2.8) thus holds for all $t \in [\tau + T_0, T - (\tau + T_0)]$, with

$$C := \frac{\kappa\sqrt{\tau}}{8C_2^2} > 0,$$

and

$$\mu := \frac{\log\left(\frac{\sqrt{\tau}}{4C_2^2}\right)}{\tau + T_0} > 0. \tag{5.33}$$

By virtue of (5.19), (5.20) and (5.32), we deduce that (2.8) holds for all $t \in [0,T]$, with $T^* := 2(\tau + T_0)$,

$$C := \max\left\{C_1, \frac{\kappa\sqrt{\tau}}{8C_2^2}\right\} > 0, \tag{5.34}$$

and $\mu > 0$ as in (5.33). This concludes the proof. $\qquad\square$

5.3. 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

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

We recall that $f_1, \ldots, f_m \in \mathrm{Lip}(\mathbb{R}^d; \mathbb{R}^d)$. We begin with the following result.

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

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

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

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

and

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

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

This sort of time-scaling in the context of driftless control affine systems is commonly used – a canonical example is the proof of the Chow-Rashevskii theorem ([8, 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

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

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

Proof of Corollary 2.2. Fix any $t \in [0, T)$ and $0 < h \ll 1$ so that $t + 2h \in [0, T]$, and set

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

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

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

By means of simple changes of variables, and using the suboptimality of u^{aux}, we can readily see that

$$
\begin{aligned}
J_T\left(u_T\right) \leqslant J_T\left(u^{\mathrm{aux}}\right) &= \int_0^T \left\|u^{\mathrm{aux}}(s)\right\|^2 \mathrm{d}s + \int_0^T \left\|y^{\mathrm{aux}}(s) - \overline{y}\right\|^2 \mathrm{d}s \\
&= \int_0^{T-h} \left\|u_T(s)\right\|^2 \mathrm{d}s - \frac{1}{2} \int_t^{t+h} \left\|u_T(s)\right\|^2 \mathrm{d}s \\
&\quad + \int_0^{T-h} \left\|y_T(s) - \overline{y}\right\|^2 + \int_t^{t+h} \left\|y_T(s) - \overline{y}\right\|^2 \mathrm{d}s \\
&\leqslant \int_0^T \left\|u_T(s)\right\|^2 \mathrm{d}s - \frac{1}{2} \int_t^{t+h} \left\|u_T(s)\right\|^2 \mathrm{d}s \\
&\quad + \int_0^T \left\|y_T(s) - \overline{y}\right\|^2 + \int_t^{t+h} \left\|y_T(s) - \overline{y}\right\|^2 \mathrm{d}s. \qquad (5.37)
\end{aligned}
$$

From (5.37), one sees that

$$
\frac{1}{2} \int_t^{t+h} \left\|u_T(s)\right\|^2 \mathrm{d}s \leqslant \int_t^{t+h} \left\|y_T(s) - \overline{y}\right\|^2 \mathrm{d}s. \qquad (5.38)
$$

We combine (5.38) with (2.11) to deduce that

$$
\begin{aligned}
\frac{1}{h} \int_t^{t+h} \left\|u_T(s)\right\|^2 \mathrm{d}s \leqslant \frac{2}{h} \int_t^{t+h} \left\|y_T(s) - \overline{y}\right\|^2 \mathrm{d}s &\leqslant \frac{2C}{h} \int_t^{t+h} e^{-2\mu s} \mathrm{d}s \\
&\leqslant \frac{2C}{h} \int_t^{t+h} e^{-2\mu t} \mathrm{d}s \\
&= 2C e^{-2\mu t}. \qquad (5.39)
\end{aligned}
$$

By the Lebesgue differentiation theorem, using (5.39) we deduce that

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

for a.e. $t \in (0, T)$, as desired. This concludes the proof. $\qquad \square$

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. $\qquad \square$

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* $C = C(f, \phi, \omega, \Omega, \overline{y}, \mathbf{y}^0) > 0$ *independent of* $T > 0$ *such that*

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

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 $T_0 := T_{\min}(\omega, \Omega) + 1$ where $T_{\min}(\omega, \Omega) > 0$ is the minimal controllability time for the semilinear wave equation, defined in (3.6).

Case 1). 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 $y^\dagger(T_0) = \overline{y}$ and $\partial_t y^\dagger(T_0) = 0$ (in $L^2(\Omega)$, and thus a.e.). Now set

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

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

$$y^{\mathrm{aux}}(t) = \overline{y} \quad \text{and} \quad \partial_t y^{\mathrm{aux}}(t) = 0 \qquad \text{for } t \in [T_0, T].$$

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

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

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

Case 2). Now suppose that $T \leqslant T_0$. We use $J_T(u_T) \leqslant J_T(u_{T_0})$ to obtain

$$J_T(u_T) \leqslant \phi(y_{T_0}(T)) + \|y_{T_0} - \overline{y}\|_{L^2(0, T; H_0^1(\Omega))}^2 + \|\partial_t y_{T_0}\|_{L^2((0, T) \times \Omega)}^2 + \|u_{T_0}\|_{L^2((0, T) \times \omega)}^2.$$

Now $y_{T_0} \in C^0([0, T_0]; L^2(\Omega))$ is bounded uniformly with respect to $T \in [0, T_0]$. Hence, using the fact that $\phi \in \mathcal{L}(L^2(\Omega); \mathbb{R}_+)$ and $T \leqslant T_0$, we deduce that

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

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

Since (3.1) is a Lipschitz perturbation of an exactly controllable linear system, the following claim holds.

Claim 6.1 (Cost estimate). *Let* $T_0 > T_{\min}(\omega, \Omega)$, *where* $T_{\min}(\omega, \Omega) > 0$ *is defined in* (3.6). *There exists* $r > 0$ *and* $C = C(T_0, \omega, f) > 0$ *such that*

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

and

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

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

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

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

Remark 10 (Regarding Claim 6.1). *Let us provide more detail regarding Claim 6.1, following* [51, 52], *and also the proofs of* [50, Theorem 2.1] *and* [14, Theorem 2.2]. *For showing the controllability of the semilinear wave equation, one typically proceeds by considering*

$$\partial_t^2 y - \Delta y + g(\zeta)y = f(0) \qquad in\ \Omega \times (0,T),$$

where $\zeta \in L^2((0,T) \times \Omega)$ *is given, and* $g(s) = \frac{f(s)-f(0)}{s}$ *for* $s \in \mathbb{R}$ *is bounded and continuous. It can be shown that the above system is controllable in time* $T > T_{\min}(\omega, \Omega)$ *with continuous dependence of the minimal* L^2-*norm control with respect to the data and* $f(0)$. *The result is transferred to the semilinear system by Schauder's fixed point theorem. To have precisely the same estimates as in Claim 6.1, namely, to remove the dependence of the minimal* L^2-*norm control with respect to* $f(0)$, *we assume that* $f(0) = 0$.

As in the finite-dimensional case, the second "quasi-turnpike" result is one for an auxiliary control problem with fixed endpoints. For $0 \leqslant \tau_1 < \tau_2 \leqslant T$ 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

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

over all $u \in U_{\mathrm{ad}}$, 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

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

and where

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

We recall that $f \in \mathrm{Lip}(\mathbb{R})$. 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. *Fix* $T_0 > T_{\min}(\omega, \Omega)$. *Suppose* $T > 0$ *and* $0 \leqslant \tau_1 < \tau_2 \leqslant T$ *are fixed such that* $\tau_2 - \tau_1 \geqslant 2T_0$. *Let* $r > 0$ *be as in Claim 6.1, and let* $\mathbf{y}^{\tau_1}, \mathbf{y}^{\tau_2}$ *be such that*

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

for $i = 1, 2$. Let $u_T \in U_{\mathrm{ad}}$ 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 $C = C(f, T_0, \Omega, \omega) > 0$ independent of $T, \tau_1, \tau_2, \mathbf{y}^{\tau_i}, r$, such that

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

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

$$\mathbf{y} := \begin{bmatrix} y \\ \partial_t y \end{bmatrix}, \qquad \overline{\mathbf{y}} := \begin{bmatrix} \overline{y} \\ 0 \end{bmatrix}.$$

We also recall the definition of the skew-adjoint operator

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

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

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

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

$$\left\| u^\dagger \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \leqslant C_0 \left\| \mathbf{y}^{\tau_1} - \overline{\mathbf{y}} \right\|_X^2, \tag{6.4}$$

for some $C_0 = C_0(T_0, \omega, f) > 0$, and such that the corresponding solution $\mathbf{y}^\dagger = \begin{bmatrix} y^\dagger \\ \partial_t y^\dagger \end{bmatrix}$ 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 $\mathbf{y}^\dagger(\tau_1 + T_0) = \overline{\mathbf{y}}$ in X. By writing the Duhamel formula for $\mathbf{y}^\dagger - \overline{\mathbf{y}}$, and using the conservative character of $\{e^{tA}\}_{t>0}$ in X, the Cauchy-Schwarz

inequality, the Lipschitz character of f and the Poincaré inequality, we see that

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

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

$$
\left\| \mathbf{y}^\dagger(t) - \overline{\mathbf{y}} \right\|_X \leqslant C_1 \exp\left(C(f, \Omega) T_0 \right) \| \mathbf{y}^{\tau_1} - \overline{\mathbf{y}} \|_X
\tag{6.6}
$$

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 $u^\ddagger \in L^2((\tau_1, \tau_1 + T_0) \times \omega)$ satisfying

$$
\left\| u^\ddagger \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)}^2 \leqslant C_0 \| \overline{\mathbf{y}} - \mathbf{y}^{\tau_2} \|_X^2,
\tag{6.7}
$$

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

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

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

$$
\begin{aligned}
\left\| \mathbf{y}^\ddagger(t) - \overline{\mathbf{y}} \right\|_X &\leqslant \int_{\tau_1}^t \left\| e^{(t-s)A} \begin{bmatrix} 0 \\ u^\ddagger(s)\mathbf{1}_\omega \end{bmatrix} \right\|_X \mathrm{d}s + \int_{\tau_1}^t \left\| e^{(t-s)A} \begin{bmatrix} 0 \\ (f(y^\ddagger) - f(\overline{y})) \end{bmatrix} \right\|_X \mathrm{d}s \\
&\leqslant \sqrt{T_0} \left\| u^\ddagger \right\|_{L^2((\tau_1, \tau_1 + T_0) \times \omega)} + C(f, \Omega) \int_{\tau_1}^t \left\| \mathbf{y}^\ddagger(s) - \overline{\mathbf{y}} \right\|_X \mathrm{d}s,
\end{aligned}
\tag{6.8}
$$

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

$$
\left\| \mathbf{y}^\ddagger(t) - \overline{\mathbf{y}} \right\|_X \leqslant C_2 \exp\left(C(f, \Omega) T_0 \right) \| \mathbf{y}^{\tau_2} - \overline{\mathbf{y}} \|_X
\tag{6.9}
$$

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

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

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

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

and thus

$$\mathbf{y}^{\mathrm{aux}}(t) = \overline{\mathbf{y}} \qquad \text{in } [\tau_1 + T_0, \tau_2 - T_0], \tag{6.10}$$

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

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

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

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

An application of Lemma 4.3 suffices to conclude. $\qquad \square$

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 $f \in \mathrm{Lip}(\mathbb{R})$, as presented in [38, Lemma 8.3] (and the references therein), given any $T_0 > 0$, $y^0 \in L^2(\Omega)$ and $\overline{y} \in H_0^1(\Omega)$ solution to (3.3), there exists a control $u \in L^2((0, T_0) \times \omega)$ such that the unique solution y to (3.7) satisfies $y(T_0) = \overline{y}$, and

$$\| u \|_{L^2((0, T_0) \times \omega)} \leqslant C(T_0, \omega, f) \left\| y^0 - \overline{y} \right\|_{L^2(\Omega)} \tag{7.1}$$

for some $C(T_0, \omega, f) > 0$ (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 $z := y - \overline{y}$, and the control u steering z to 0 in time T is the same as that steering y to $\overline{y}$ in time T. But then, $\| u \|_{L^2((0, T_0) \times \omega)} \leqslant C(T_0, \omega, f) \| z(0) \|_{L^2(\Omega)}$ from the linear system and a fixed-point argument. Let $T > 0$ and $0 \leqslant \tau_1 < T$ be fixed, and suppose $y^{\tau_1} \in L^2(\Omega)$ is given. Consider

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

where y solves

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

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. *Let $T > 0$ and $\tau_1 \geqslant 0$ be given such that $T > \tau_1$, and let $y^{\tau_1} \in L^2(\Omega)$. 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 $C = C(f, \omega) > 0$ independent of $T, \tau_1 > 0$ and y^{τ_1} such that*

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

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

Proof of Lemma 7.1. Case 1). Let us first suppose that $T \geqslant \tau_1 + 1$. By controllability to the steady state $\overline{y}$ (see the discussion around (7.1)), we know that exists a control $u^\dagger \in L^2((\tau_1, \tau_1 + 1) \times \omega)$ satisfying

$$\left\|u^\dagger\right\|_{L^2((\tau_1, \tau_1+1) \times \omega)} \leqslant C_1 \|y^{\tau_1} - \overline{y}\|_{L^2(\Omega)} \tag{7.4}$$

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

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

satisfies $y^\dagger(\tau_1 + 1) = \overline{y}$. Arguing as in the proof of Lemma 4.2, we see that

$$\left\|y^\dagger(t) - \overline{y}\right\|_{L^2(\Omega)} \leqslant \|y^{\tau_1} - \overline{y}\|_{L^2(\Omega)} + \left\|u^\dagger\right\|_{L^2((\tau_1, \tau_1+1) \times \omega)}$$
$$+ C(f) \int_{\tau_1}^t \left\|y^\dagger(s) - \overline{y}\right\|_{L^2(\Omega)} \, \mathrm{d}s \tag{7.5}$$

for $t \in (\tau_1, \tau_1 + 1)$, with $C(f) > 0$ being the Lipschitz constant of f. Applying the Grönwall inequality to (7.5) and using (7.4), we deduce that

$$\left\|y^\dagger(t) - \overline{y}\right\|_{L^2(\Omega)} \leqslant C_2 \exp\left(C(f)\right) \|y^{\tau_1} - \overline{y}\|_{L^2(\Omega)} \tag{7.6}$$

for some $C_2(f, \omega) > 0$ independent of T, τ_1, and for every $t \in (\tau_1, \tau_1 + 1)$. Now set

$$u^{\mathrm{aux}}(t) := \begin{cases} u^\dagger(t) & \text{in } (\tau_1, \tau_1 + 1) \\ 0 & \text{in } (\tau_1 + 1, T) \end{cases}$$

and let y^{aux} be the corresponding solution to (3.7). Clearly

$$y^{\mathrm{aux}}(t) = \overline{y} \qquad \text{for } t \in [\tau_1 + 1, T].$$

Hence, using $J_{\tau_1, T}(u_T) \leqslant J_{\tau_1, T}(u^{\mathrm{aux}})$, (7.6) and (7.4), we see that

$$J_{\tau_1, T}(u_T) \leqslant \left\|y^\dagger - \overline{y}\right\|_{L^2((\tau_1, \tau_1+1) \times \Omega)}^2 + \left\|u^\dagger\right\|_{L^2((\tau_1, \tau_1+1) \times \omega)}^2$$
$$\leqslant C_3 \|y^{\tau_1} - \overline{y}\|_{L^2(\Omega)}^2$$

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

Case 2). Now suppose that $\tau_1 < T < \tau_1 + 1$. We may then use the optimality inequality $J_{\tau_1, T}(u_T) \leqslant J_{\tau_1, T}(u_{\tau_1+1})$, as well as $J_{\tau_1, T}(u_{\tau_1+1}) \leqslant J_{\tau_1, \tau_1+1}(u_{\tau_1+1})$, and since by the previous step, we know that

$$J_{\tau_1, \tau_1+1}(u_{\tau_1+1}) \leqslant C_3 \|y^{\tau_1} - \overline{y}\|_{L^2(\Omega)}^2,$$

where $C_3 = C_3(f, \omega) > 0$ is independent of T, τ_1, we deduce

$$J_{\tau_1,T}(u_T) \leqslant C_3 \left\| y^{\tau_1} - \overline{y} \right\|_{L^2(\Omega)}^2 . \tag{7.7}$$

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

Proof of Theorem 3.2. The proof is of the same spirit[9] 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. Let $T_0 > 0$ be arbitrary, and fix

$$\tau > C_1^4$$

where $C_1 = C_1(f, \omega) > 0$ is the (square root of the) constant appearing in Lemma 7.1. We note that if $T \leqslant 2\tau + T_0$, then the desired estimate clearly follows by arguing as in previous proofs. We thus suppose that

$$T > 2\tau + T_0$$

is 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 $2\tau < T$ and thus $\tau \leqslant \frac{T}{2}$, by Lemma 5.3 there exists a $\tau_1 \in [0, \tau)$ such that

$$\left\| y_T(\tau_1) - \overline{y} \right\|_{L^2(\Omega)} \leqslant \frac{\left\| y_T - \overline{y} \right\|_{L^2((0,T)\times\Omega)}}{\sqrt{\tau}} \leqslant \frac{C_1}{\sqrt{\tau}} \left\| y^0 - \overline{y} \right\|_{L^2(\Omega)}, \tag{7.8}$$

where we used Lemma 7.1 for the second estimate. The control $u_T|_{[\tau_1,T]}$ can be shown to minimize $J_{\tau_1,T}$ 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),

$$\left\| y_T(t) - \overline{y} \right\|_{L^2(\Omega)} \leqslant C_1 \| y_T(\tau_1) - \overline{y} \|_{L^2(\Omega)} \leqslant \frac{C_1^2}{\sqrt{\tau}} \left\| y^0 - \overline{y} \right\|_{L^2(\Omega)} \tag{7.9}$$

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

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

the estimate

$$\sup_{t \in [n\tau, T]} \| y_T(t) - \overline{y} \|_{L^2(\Omega)} \leqslant \left(\frac{C_1^2}{\sqrt{\tau}} \right)^n \left\| y^0 - \overline{y} \right\|_{L^2(\Omega)} \tag{7.10}$$

---

[9] Actually, as already commented below the statement of Corollary 2.1, the proof presented below also roughly applies to show Corollary 2.1, where one also would need to account for the constants which should also depend on the radius $r > 0$. In fact, just as for the heat equation, one could first adapt Lemma 5.2 to a functional of the form (7.2); an adaptation which would hold for initial data in a ball of radius $r > 0$ around $\overline{y}$, and then use the global estimate of Lemma 5.1 and argue as in the beginning of the proof of Theorem 2.1 to fit within this ball, where one bootstraps forward in time only (namely, over intervals of the form $[n\tau, T]$).

holds. 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

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

This clearly implies that

$$\tau \leqslant \frac{T - 2n\tau}{2}. \tag{7.11}$$

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

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

We may apply the induction hypothesis (7.10) to deduce

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

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

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

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 $n(t) := \left\lfloor \frac{t}{\tau + T_0} \right\rfloor$. Clearly $n(t) \geqslant 1$, $t \geqslant n(t)\tau$ and $n(t) \leqslant \frac{T}{2\tau}$ due to the choice of T_0. We may then apply (7.10) to find that

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

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

$$\|y_T(t) - \overline{y}\|_{L^2(\Omega)} \leqslant \exp\left(-n(t) \log\left(\frac{\sqrt{\tau}}{C_1^2}\right)\right) \left\|y^0 - \overline{y}\right\|_{L^2(\Omega)}$$

$$\leqslant \frac{\sqrt{\tau}}{C_1^2} \exp\left(-\frac{\log\left(\frac{\sqrt{\tau}}{C_1^2}\right)}{\tau + T_0} t\right) \left\|y^0 - \overline{y}\right\|_{L^2(\Omega)}$$

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

$$\mu := \frac{\log\left(\frac{\sqrt{\tau}}{C_1^2}\right)}{\tau + T_0} > 0 \tag{7.15}$$

and

$$C := \frac{\sqrt{\tau}}{C_1^2} > 0. \tag{7.16}$$

This concludes the proof. $\qquad \square$

8. Numerics

We briefly comment on the setting of the numerical experiment shown in Figure 1. We make use of the neural ODE (1.3) with $\sigma \equiv \tanh$, and discretize with an explicit midpoint rule with $\triangle t = 1/2$. The operator $P : \mathbb{R}^3 \to [-1, 1]$ appearing in (1.1) is defined as $Px = \mathtt{hardtanh}(p_1 x + p_2)$, through the nonlinear thresholding operator $\mathtt{hardtanh}(z) = 1_{\{s \geqslant 1\}}(z) + z 1_{s \in (-1,1)}(z) + 1_{\{s \leqslant 1\}}(z)$, and the parameters $p_1 \in \mathbb{R}^{3 \times 3}$ and $p_2 \in \mathbb{R}^3$ which are randomly sampled[10] from a normal distribution. We use $n = 2400$ points for training, and 600 points for testing (see Figure 5). We originally consider a dataset of points in $\mathbb{R}^2$, but we embed them in $\mathbb{R}^3$ by adding a 0 to each point. This is to avoid the intersection of trajectories in $\mathbb{R}^2$, which takes place due to uniqueness ([9]). This is why we actually plot the predictor as a map $\mathbb{R}^2 \to [-1, 1]$ in Figure 5. The code for reproducing all figures is available at https://github.com/borjanG/dynamical.systems.
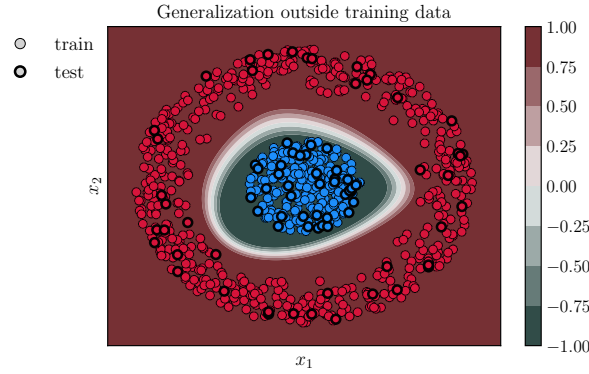
FIGURE 5. We see that the trained predictor, plotted on $[-2, 2]^2$, has generalized the shape of the dataset, as desired.
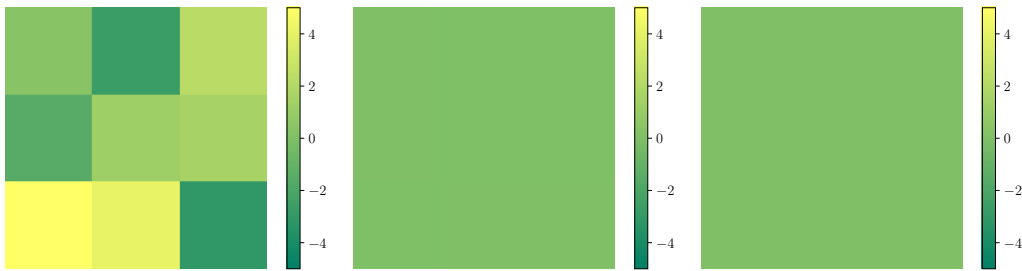

FIGURE 6. To further corroborate the graph showing the decay of optimal controls $\|u(t)\|$ in Figure 1, where $u(t) = (w(t), b(t))$, we display the values of the matrix $w(t) \in \mathbb{R}^{3 \times 3}$ in times $t \in \{0.67, 1.33, 5\}$ (left to right).

---

[10]As a byproduct of the Johnson–Lindenstrauss lemma ([43, Lemma 23.4]), such random projections are of low distortion with respect to the Euclidean distance, in the sense that distances between points are nearly preserved after projecting.

9. 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

- 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.
- 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.

The motivation behind the consideration of steady state running targets in (2.3) was the link with machine learning applications, namely problem (1.1) (although, we saw that our results also apply to many contexts which arise naturally in mechanics). While we see a turnpike phenomenon in the numerical simulations of (1.1), our analysis done for (2.3) strictly applies to (1.1) only when $d_x = d_y$ and P is the identity. On another hand, since P is surjective, we can see (2.3) as a relaxed version of (1.1), in which case, we select the running targets in the functional (2.3) as $\overline{\mathbf{x}}_i \in P^{-1}\left(\left\{y^{(i)}\right\}\right)$ for $i \in \{1, \ldots, n\}$, which are steady states of the underlying system with 0 control (as are all constants, actually). The numerical simulations in Figure 1 are all the more curious because of the fact that 1). we see the stabilization of the full state, while we solely penalize projections of it, and 2). the projection $x \mapsto Px$ is actually not coercive with respect to x (see Section 8). For problems manifesting a lack of observability of certain components of the state in the tracking term, the turnpike property for the observed components and the full controls has been shown in [37], in the setting of linear systems. Should such a property also hold for (1.1), then the stability of the full state seen in Figure 1 can be explained through the decay of the optimal controls to 0 and the specific form of the dynamics. One could envision a fusion of our strategy with Loyasewicz-type inequalities to provide theoretical guarantees, but this remains an open problem.

9.1. **Outlook.** Let us conclude with a select list of additional open problems.

- **Necessity of assuming that $\overline{y}$ is a steady state.** The assumption that the running target $\overline{y}$ in (2.3) is a steady state of the dynamics allows us to easily obtain quasi-turnpike controls 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 $\overline{y}$ associated to a prescribed control $\overline{u}$ can readily be addressed by penalizing $u - \overline{u}$ 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 $\overline{y}$, the minimal L^2–norm control steering the system from y^0 to $\overline{y}$ may be estimated by $\left\| y^0 - \overline{y} \right\|$, and similarly for that from $\overline{y}$ 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.

- **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 [33] 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. [28, 35, 38, 42] and the references therein.

- **More general control systems.** We have considered homogeneous Dirichlet boundary conditions in (3.1) and (3.7) 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. 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, and this semigroup ought to be uniformly bounded for all times.

    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 (apparent) 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 [15] and the references therein). Similarly, boundary control systems may pose technical difficulties, since they require for the introduction of *admissible control operators* (a general functional framework for lifting the trace on the boundary – see [47, Chapter 4]) to be written in a canonical first order form, and consequently, to admit a Duhamel formula representation for the solution. The particular issue for boundary control systems is that there is no guarantee that the inferred control operator would be bounded with respect to T, which is of paramount importance to our strategy. We leave these extensions open to future studies.

- **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 [4, 5, 10, 34] for recent developments). Notwithstanding, our results should be applicable to a system of the form (see [4])

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

where u is a scalar control and f is an appropriate nonlinearity (see [4] 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 $u_T - \overline{u}$) for minimizers u_T of

$$J_T(u) := \int_0^T \|y(t) - \overline{y}\|^2 \, \mathrm{d}t + \int_0^T \|u - \overline{u}\|^2 \, \mathrm{d}t$$

where y_T is the unique solution to

$$\begin{cases} \partial_t y - \Delta y + y^3 = u\mathbf{1}_\omega & \text{in } (0,T) \times \Omega \\ y = 0 & \text{on } (0,T) \times \partial\Omega \\ y|_{t=0} = y^0 & \text{in } \Omega, \end{cases} \tag{9.1}$$

and $\overline{y} \in H_0^1(\Omega)$ is a controlled steady state associated to some $\overline{u} \in L^2(\omega)$ (the case $\overline{u} \equiv 0$ is somewhat trivial due to the inherent stabilization to $\overline{y} \equiv 0$). 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 $u_T - \overline{u}$). To this end, first of all, for any $u \in L^2((0,T) \times \omega)$, using the variational formulation and standard arguments including the Cauchy-Schwarz, Young and Poincaré inequalities, one can find

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

for a.e. $t \in [0,T]$, where $\epsilon > \frac{C(\Omega)}{4}$, whereas y solves (9.1), and thus

$$\|y\|_{C^0([0,T];L^2(\Omega))} \leqslant C_1(\Omega) \left(\|u\|_{L^2((0,T)\times\omega)} + \|y^0\|_{L^2(\Omega)} \right). \tag{9.2}$$

Following the proof of Lemma 4.2 for $f(s) = s^3$ and using (9.2), we may find

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

where now

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

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 $\|u - \overline{u}\|_{L^2((0,T)\times\omega)}$.

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 [49] 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 [27] 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 [22] for a related study) or free boundary problems (see [17] for a control perspective).

References

[1] AGRACHEV, A., BARILARI, D., AND BOSCAIN, U. *A comprehensive introduction to sub-Riemannian geometry*, vol. 181. Cambridge University Press, 2019.

[2] AMANN, H. *Linear and quasilinear parabolic problems*, vol. 1. Springer, 1995.

[3] 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.

[4] 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.

[5] 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.

[6] CAZENAVE, T. *An introduction to semilinear elliptic equations*, vol. 164. Editora do Instituto de Matemática, Universidade Federal do Rio de Janeiro, 2006.

[7] CHEN, T. Q., RUBANOVA, Y., BETTENCOURT, J., AND DUVENAUD, D. K. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems* (2018), pp. 6571–6583.

[8] CORON, J.-M. *Control and nonlinearity*. No. 136. American Mathematical Soc., 2007.

[9] DUPONT, E., DOUCET, A., AND TEH, Y. W. Augmented Neural ODEs. *Advances in Neural Information Processing Systems 32* (2019), 3140–3150.

[10] DUPREZ, M., AND LISSY, P. Bilinear local controllability to the trajectories of the Fokker-Planck equation with a localized control. In *Annales de l'Institut Fourier* (2021).

[11] E, W. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics 5*, 1 (2017), 1–11.

[12] ESTEVE, C., GESHKOVSKI, B., PIGHIN, D., AND ZUAZUA, E. Large-time asymptotics in deep learning. *arXiv preprint arXiv:2008.02491* (2020).

[13] FAULWASSER, T., HEMPEL, A.-J., AND STREIF, S. On the turnpike to design of deep neural nets: Explicit depth bounds. *arXiv preprint arXiv:2101.03000* (2021).

[14] 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.

[15] GESHKOVSKI, B. Null-controllability of perturbed porous medium gas flow. *ESAIM: Control, Optimisation and Calculus of Variations 26* (2020), 85.

[16] GESHKOVSKI, B. *Control in moving interfaces and deep learning.* PhD thesis, Universidad Autónoma de Madrid, 2021.

[17] GESHKOVSKI, B., AND ZUAZUA, E. Controllability of one-dimensional viscous free boundary flows. *SIAM Journal on Control and Optimization 59*, 3 (2021), 1830–1850.

[18] GRÜNE, L., SCHALLER, M., AND SCHIELA, A. Sensitivity analysis of optimal control for a class of parabolic PDEs motivated by model predictive control. *SIAM Journal on Control and Optimization 57*, 4 (2019), 2753–2774.

[19] GRÜNE, L., SCHALLER, M., AND SCHIELA, A. Efficient MPC for parabolic PDEs with goal oriented error estimation. *arXiv preprint arXiv:2007.14446* (2020).

[20] 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.

[21] GRÜNE, L., SCHALLER, M., AND SCHIELA, A. Abstract nonlinear sensitivity and turnpike analysis and an application to semilinear parabolic PDEs. *ESAIM: Control, Optimisation and Calculus of Variations 27* (2021), 56.

[22] 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.

[23] 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.

[24] 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.

[25] JEAN, F., AND PRANDI, D. Complexity of control-affine motion planning. *SIAM Journal on Control and Optimization 53*, 2 (2015), 816–844.

[26] 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.

[27] 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).

[28] LE BALC'H, K. Local controllability of reaction-diffusion systems around nonnegative stationary states. *ESAIM: Control, Optimisation and Calculus of Variations 26* (2020), 55.

[29] LECUN, Y., BENGIO, Y., AND HINTON, G. Deep learning. *Nature 521*, 7553 (2015), 436–444.

[30] 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.

[31] LIONS, J.-L. Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. tome 1. *RMA 8* (1988).

[32] LIONS, P.-L. On the existence of positive solutions of semilinear elliptic equations. *SIAM Review 24*, 4 (1982), 441–467.

[33] 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).

[34] MAZARI, I., AND RUIZ-BALET, D. A fragmentation phenomenon for a nonenergetic optimal control problem: Optimization of the total population size in logistic diffusive models. *SIAM Journal on Applied Mathematics 81*, 1 (2021), 153–172.

[35] MAZARI, I., RUIZ-BALET, D., AND ZUAZUA, E. Constrained control of gene-flow models. *arXiv preprint arXiv:2005.09236* (2020).

[36] PIGHIN, D. The turnpike property in semilinear control. *ESAIM: Control, Optimisation & Calculus of Variations 26* (2021).

[37] PIGHIN, D., AND SAKAMOTO, N. The turnpike with lack of observability. *arXiv preprint arXiv:2007.14081* (2020).

[38] PIGHIN, D., AND ZUAZUA, E. Controllability under positivity constraints of semilinear heat equations. *Mathematical Control & Related Fields 8*, 3&4 (2018), 935.

[39] PORRETTA, A., AND ZUAZUA, E. Long time versus steady state optimal control. *SIAM Journal on Control and Optimization 51*, 6 (2013), 4242–4273.

[40] 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.

[41] PRANDI, D. Hölder equivalence of the value function for control-affine systems. *ESAIM: Control, Optimisation and Calculus of Variations 20*, 4 (2014), 1224–1248.

[42] 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.

[43] SHALEV-SHWARTZ, S., AND BEN-DAVID, S. *Understanding machine learning: From theory to algorithms*. Cambridge University Press, 2014.

[44] TRÉLAT, E. Linear turnpike theorem. *arXiv preprint arXiv:2010.13605* (2020).

[45] 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.

[46] TRÉLAT, E., AND ZUAZUA, E. The turnpike property in finite-dimensional nonlinear optimal control. *Journal of Differential Equations 258*, 1 (2015), 81–114.

[47] TUCSNAK, M., AND WEISS, G. *Observation and control for operator semigroups*. Springer Science & Business Media, 2009.

[48] YAGÜE, C. E., AND GESHKOVSKI, B. Sparse approximation in learning via neural ODEs. *arXiv preprint arXiv:2102.13566* (2021).

[49] ZAMORANO, S. Turnpike property for two-dimensional Navier–Stokes equations. *Journal of Mathematical Fluid Mechanics 20*, 3 (2018), 869–888.

[50] ZHANG, X. Exact controllability of semilinear plate equations. *Asymptotic Analysis 27*, 2 (2001), 95–125.

[51] ZHANG, X., AND ZUAZUA, E. Exact controllability of the semi-linear wave equation. *Unsolved Problems in Mathematical Systems and Control Theory* (2004), 173.

[52] ZUAZUA, E. Exact boundary controllability for the semilinear wave equation. *Nonlinear Partial Differential Equations and their Applications 10* (1991), 357–391.

[53] 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.

[54] ZUAZUA, E. Large time control and turnpike properties for wave equations. *Annual Reviews in Control 44* (2017), 199–210.

Carlos Esteve-Yagüe, Borjan Geshkovski, Dario Pighin
Chair of Computational Mathematics
Fundación Deusto
Av. de las Universidades, 24
48007 Bilbao, Basque Country, Spain
Email address: {carlos.esteve, borjan.geshkovski, dario.pighin}@deusto.es

Enrique Zuazua
Chair in Dynamics, Control, and Numerics
Alexander von Humboldt-Professorship
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