Differential Dynamic Programming for Backward Reachability Analysis.

Author Names Omitted for Anonymous Review. Paper-ID [add your ID here]

***Abstract*—Backward reachability analysis verifies a robustness metric that guarantees system safety. However, it is premised on solving implicitly-constructed value functions on spatio-temporal grids to verify a robustness metric that guarantees system safety – up to a specified time bound. However, as state dimensions increase, time-space discretization methods become impractical owing to their exponential complexity. Approximation schemes in global value function space fail to preserve the robustness guarantees of basic backward reachability theory. We present *safe to the last truncation*: an iterative decomposition scheme that incrementally truncates a high-dimensional value function to the minimum low-rank tensor necessary for computing reachable sets, tubes and reach-avoid with guarantee to a local saddle-point extrema. This paper presents an initial evaluation of our proposal on the backward reachable sets and a classical time-optimal bang-bang control time-to-reach problem.**

I. Introduction

Recent developments in cyberphysical systems (CPS) have created complex entanglements with interactions that are difficult to analyze. The "physical" and "cyber" couplings of such systems due to complex interconnection of control systems, sensors, and software make planning and executing in real-time, safety-critical scenarios like collision avoidance in uneven terrains, or sensing efficiently in the presence of multiple agents—all require deep integration and the any actions of system components must be planned meticulously. Therefore, the safety analysis of combined CPS systems in the presence of sensing, control, and learning becomes timely and crucial. Differential optimal control theory and games offer a powerful paradigm for resolving the safety of multiple agents interacting over a shared space. Both problems rely on a resolution of the Hamilton-Jacobi-Bellman (HJB) or the Hamilton-Jacobi-Isaacs (HJI) equation in order to solve the control problem [**?**]. As HJ-type equations have no classical solution for almost all *practical* problems, stable numerical and computational methods need to be brought to bear in order to produce solutions with (approximately) optimal guarantees.

With essentially non-oscillating (ENO) [46] Lax-Friedrichs [12] schemes applied to numerically resolve HJ Hamiltonians [15], we can now obtain unique (viscosity) solutions to HJ-type equations with high accuracy and precision *on a mesh*. Employing meshes for resolving inviscid Euler equations whose solutions are the derivatives of HJ equations, these methods scale exponentially with state dimensions, making them ineffective for complex systems—a direct consequence of *curse of dimensionality* [6]. Truncated power series methods [23, 24, 14, 30] are successive approximations of HJ value functions; however, these limit the stability region of the resulting approximate controller, and require a careful tuning of the approximate controller such that it has a direct effect on the original optimal control problem. In addition, stability is not easily guaranteed for series approximated HJ value functions where it is generally assumed that the highest-ordered terms in the series truncation dominate neglected higher-order terms.

a) Contributions: Therefore in subject matter and emphasis, this paper reflects the influences described in the foregoing. As a result, we focus on computational techniques because almost all *practical* problems cannot be analytically resolved. To analyze safety, we cast our problem formulation within the framework of *Cauchy-type* HJ equations [11], and we specifically resolve the scalable safety problem by solving the terminal value in the HJ PDE within the framework of Mitchell's *robustly controlled backward reachable tubes* [35]. In this sentiment, new computational techniques are introduced including (i) iterative Galerkin approximation of large value functions; (ii) finite difference approximation schemes with error estimates (essentially, an extension of [9] on reduced Hilbertian spaces); and (iii) analytic saddle solutions to approximated HJ value functions. to synthesize approximately optimal control laws (essentially, saddle-point solutions)

The rest of this paper is organized as follows: **??** describes the concepts and § III topics that we will build upon in describing our proposal in **??**; we present results and insights from experiments in § V. We conclude the paper in § VII. This work is the first to systematically provide a rational incremental decomposition scheme that provides approximation guarantees on regions of the state space where approximate HJ control laws are valid as well as provide a rational analysis for high-dimensional verification of nonlinear systems with guarantees.

II. Related Work

a) Dynamic Programming and Two-Person Games.: The formal relationships between the dynamic programming (DP) optimality condition for the *value* in differential two-person zero-sum games, and the solutions to PDEs that solve "min-max" or "max-min" type nonlinearity (the Isaacs' equation) was presented in [22]. Essentially, Isaacs' claim was that if the *value* functions are smooth enough, then they solve certain first-order partial differential equations (PDE) problems with "max-min" or "min-max"-type nonlinearity. However, the DP value functions are seldom regular enough to admit a solution in the classical sense. "Weaker" solutions on the other hand [31, 15, 9, 12, 16] provide generalized "viscosity" solutions to HJ PDEs under relaxed regularity conditions; these viscosity solutions are not necessarily differentiable anywhere in the state space, and the only regularity prerequisite in the definition is continuity [11]. However, wherever they are differentiable, they satisfy the upper and lower values of HJ PDEs in a classical sense. Thus, they lend themselves well to many real-world problems existing at the interface of discrete, continuous, and hybrid systems [32, 45, 35, 16, 37]. Viscosity Solutions to *Cauchy-type* HJ Equations admit usefulness in backward reachability analysis [37]. In scope and focus, this is the bulwark upon which we build our formulation in this paper.

b) Reachability for Systems Verification.: Reachability analysis is one of many verification methods that allows us to reason about (cpntrol-affine) dynamical systems. The verification problem may consist in finding a *set of reachable states* that lie along the trajectory of the solution to a first order nonlinear partial differential equation that originates from some initial state $\boldsymbol{x}_0 = \boldsymbol{x}(0)$ up to a specified time bound, $t = t_f$. *From a set of initial and unsafe state sets, and a time bound, the time-bounded safety verification problem is to determine if there is an initial state and a time within the bound that the solution to the PDE enters the unsafe set*. Reachability could be analyzed in a (i) *forward* sense, whereupon system trajectories are examined to determine if they enter certain states from an *initial set*; (ii) *backward* sense, whereupon system trajectories are examined to determine if they enter certain *target sets*; (iii) *reach set* sense, in which they are examined to see if states reach a set at a *particular time*; or (iv) *reach tube* sense, in which they are evaluated that they reach a set at a point *during a time interval*.

Backward reachable sets (BRS) or tubes (BRTs) are popularly analyzed as a game of two vehicles with non-stochastic dynamics [34]. Such BRTs possess discontinuity at cross-over points (which exist at edges) on the surface of the tube, and may be non-convex. Therefore, treating the end-point constraints under these discontinuity characterizations need careful consideration and analysis when switching control laws if the underlying PDE does not have continuous partial derivatives (we discuss this further in **??**.

c) Global Mesh-based Methods and Up-Scaling Reachability Analysis: Consider a reachability problem defined in a space of dimension $D = 12$ based on the non-incremental time-space discretization of each space coordinate. For $N = 100$ nodes, the total nodes required is 10^{120} on the volumetric grid[1]. The curse of dimensionality [6] greatly incapacitates current uniform grid discretization methods for guaranteeing the robustness of backward reachable sets (BRS) and tubes (BRTs) [37] of complex systems. Recent works have started exploring scaling up the Cauchy-type HJ problem for guaranteeing safety of higher-dimensional physical systems: the authors of [2] provide local updates to BRS in unknown static environments with obstacles that may be unknown *a priori* to the agent; using standard meshing techniques for time-space uniform discretization over the entire physical space, and only updating points traversed locally, a safe navigation problem was solved in an environment assumed to be static. This makes it non-amenable to *a priori* unknown *dynamic* environments where the optimal value to the min-max HJ problem may need to be adaptively updated based on changing dynamics.

In [21], the grid was naively refined along the temporal dimension, leveraging local decomposition schemes together with warm-starting optimization of the value function from previous solutions in order to accelerate learning for safety under the assumption that the system is either completely decoupled, or coupled over so-called "self-contained subsystems". While the empirical results of [4] demonstrate the feasibility of optimizing for the optimal value function in backward reachability analysis for up to ten dimensions for a system of Dubins vehicles, there are no guarantees that are provided. An analysis exists for a 12 dimensional systems [27] with up to a billion data points in the state space, that generates robustly optimal trajectories. However, this is restricted to linear systems. Other associated techniques scale reachability with function approximators [17, 18] in a reinforcement learning framework; again these methods lose the hard safety guarantees owing to the approximation in value function space.

d) Reduced Order Modeling (ROM): Multilinear compositions of linear forms are an efficient way of manipulating complex systems. Higher-order tensors, in particular, are increasingly playing crucial roles in the storage, analysis, and use of high-dimensional data. Applications range from deep learning, higher-order statistics, chemometrics, psychometrics to signal processing inter alia. Evidence abounds that linearized nonlinear system dynamics, truncated at a reduced order (e.g. in linearized power series expansions[43, 24, 38, 33]) admit higher precision and accuracy of the approximation of the underlying nonlinear system since the moments and accumulations of higher-order dynamics are equivalent to the power series expansion coefficients. In what follows, we introduce a multilinear decomposition scheme aimed at decomposition of large backward reachable tubes in order to alleviate the exponential complexity of mesh constraints; it is an iterative scheme that generates separable reduced order models (ROM) of the original value function, which are respectively compactly representable on a mesh – making our method amenable to resolving terminal value functions using level sets methods.

III. Preliminaries

TO-DO: potentially add in the stuff about HJI here ?

A. Hamilton-Jacobi Reachability: Problem Setup

Consider the dynamical system

$$\dot{\boldsymbol{x}}(t) = f(t, \boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{v}(t)) \quad T \le t \le 0 \tag{1a}$$

$$\boldsymbol{x}(T) = \boldsymbol{x}, \tag{1b}$$

where $\boldsymbol{x}$ is the state that evolves from some initial negative time T to final time 0, and $\boldsymbol{u}(\cdot)$ and $\boldsymbol{v}(\cdot)$ are respectively the control and disturbance signals. Here $f(t, \cdot, \cdot, \cdot)$ and $\boldsymbol{x}(\cdot)$

[1]Whereas, there are only 10^{97} baryons in the observable universe (excluding dark matter)!

are assumed to be bounded and Lipschitz continuous. This bounded Lipschitz continuity property assures uniqueness of the system response $\boldsymbol{x}(\cdot)$ to controls $\boldsymbol{u}(\cdot)$ and $\boldsymbol{v}(\cdot)$ [16].

For a state $\boldsymbol{x} \in \Omega$ and a fixed time t: $T \leq t < 0$, suppose that the set of all controls for players $\boldsymbol{P}$ and $\boldsymbol{E}$ are respectively

$$\bar{\mathcal{U}} \equiv \{\boldsymbol{u} : [t,0] \to \mathcal{U} | \boldsymbol{u} \text{ measurable}, \mathcal{U} \in \mathbb{R}^m\}, \tag{2}$$
$$\bar{\mathcal{V}} \equiv \{\boldsymbol{v} : [t,0] \to \mathcal{V} | \boldsymbol{v} \text{ measurable}, \mathcal{V} \subset \mathbb{R}^p\}. \tag{3}$$

For any admissible control-disturbance pair $(\boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot))$ and initial phase $(\boldsymbol{x}, T)$, Crandall [11] and Evan's [15] claim is that there exists a unique function

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t; T, \boldsymbol{x}, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) \tag{4}$$

that satisfies (1) a.e. with the property that

$$\boldsymbol{\xi}(T) = \boldsymbol{\xi}(T; T, \boldsymbol{x}, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) = \boldsymbol{x}. \tag{5}$$

Read (4): the motion of (1) passing through phase $(\boldsymbol{x}, T)$ under the action of control $\boldsymbol{u}$, and disturbance $\boldsymbol{v}$, and observed at a time t afterwards.

For any optimal control problem a value function is constructed based on the optimal cost (or payoff) of any input phase $(\boldsymbol{x}, T)$. In reachability analysis, typically this is defined using a terminal cost function $g(\cdot) : \mathbb{R}^n \to \mathbb{R}$ that satisfies

$$|g(\boldsymbol{x})| \leq k \tag{6a}$$
$$|g(\boldsymbol{x}) - g(\hat{\boldsymbol{x}})| \leq k|\boldsymbol{x} - \hat{\boldsymbol{x}}| \tag{6b}$$

for constant k and all $T \leq t \leq 0$, $\hat{\boldsymbol{x}}, \boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{u} \in \mathcal{U}$ and $\boldsymbol{v} \in \mathcal{V}$. The zero sublevel set of $g(\boldsymbol{x})$ i.e.

$$\mathcal{L}_0 = \{\boldsymbol{x} \in \bar{\Omega} \,|\, g(\boldsymbol{x}) \leq 0\}, \tag{7}$$

is the *target set* in the phase space $\Omega \times \mathbb{R}$ (proof in [37]). This target set can represent the failure set (to avoid) or a goal set (to reach) in the state space. Note that the target set, $\mathcal{L}_0$, is a closed subset of $\mathbb{R}^n$ and is in the closure of Ω. Typically $\mathcal{L}_0$ is user-defined, and $g(x)$ is a signed distance function, that is negative inside the target set and positive elsewhere.

Reachability analysis seeks to capture all initial conditions from which trajectories of the system may enter the target set. This could be desirable (in the case where the target set is a goal) or undesirable (where the target set represents the failure set). Frequently in reachability analysis one seeks to measure the *minimum cost* over time of trajectories of the system:

$$\min_t g(\boldsymbol{\xi}(t; T, \boldsymbol{x}_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot))). \tag{8}$$

If this minimum cost is negative, then the trajectory entered the target set *at some time* $t \in [T, 0]$ over the time horizon. If the minimum cost is positive, then the trajectory will never enter the target set in the time horizon.

B. Hamilton-Jacobi Reachability: Definition and Construction of the Value Function

Rather than computing the minimum cost for every possible trajectory of the system, in safety analysis it is sufficient to consider the minimum cost under optimal behavior from both players. The optimal behavior of each player depends on whether the target set represents a goal or a failure set. For a safety (avoiding a failure set) problem setup, the evader $\boldsymbol{E}$ is seeking to maximize the minimum cost (keeping the system out of the target set) and the pursuer $\boldsymbol{P}$ seeks to minimize it. Suppose that the pursuer's mapping strategy (starting at t) is $\beta : \bar{\mathcal{U}}(t) \to \bar{\mathcal{V}}(t)$ provided for each $t \leq \tau \leq T$ and $\boldsymbol{u}, \hat{\boldsymbol{u}} \in \bar{\mathcal{U}}(t)$; then $\boldsymbol{u}(\bar{t}) = \hat{\boldsymbol{u}}(\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$ implies $\beta[\boldsymbol{u}](\bar{t}) = \beta[\hat{\boldsymbol{u}}](\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$. The differential game's (lower) value for a solution $\boldsymbol{x}(t)$ that solves (1) for $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t) = \beta[\boldsymbol{u}](\cdot)$ is

$$V(\boldsymbol{x}, t) = \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} \min_{t \in [T,0]} g(\boldsymbol{x}(T)). \tag{9}$$

For a goal-satisfaction (liveness) problem setup, the behavior of the evader and pursuer are reversed.

Optimal trajectories emanating from initial phases $(\boldsymbol{x}, T)$ where the value function is non-negative will maintain non-negative cost over the entire time horizon, and therefore will never enter the target set. Optimal trajectories from initial phases where the value function is negative will enter the target set at some point within the time horizon. For the safety problem setup in (9) we can define the corresponding *robustly controlled backward reachable tube* for $\tau \in [T, 0]$[2] in this way, i.e. the closure of the open set

$$\mathcal{L}([\tau, 0], \mathcal{L}_0) = \{\boldsymbol{x} \in \Omega \,|\, \exists\, \beta \in \bar{\mathcal{V}}(t) \,\forall\, \boldsymbol{u} \in \mathcal{U}(t), \exists\, \bar{t} \in [T, 0], \\ \boldsymbol{\xi}(\bar{t}) \in \mathcal{L}_0\}, \, \bar{t} \in [T, 0]. \tag{10}$$

Read: the set of states from which the strategies of $\boldsymbol{P}$, and for all controls of $\boldsymbol{E}$ imply that we reach the target set within the interval $[T, 0]$. More specifically, following Lemma 2 of [37], the states in the reachable set admit the following properties w.r.t the value function $\boldsymbol{V}$

$$\boldsymbol{x} \in \mathcal{L} \implies \boldsymbol{V}^-(\boldsymbol{x}, t) \leq 0 \tag{11a}$$
$$\boldsymbol{V}^-(\boldsymbol{x}, t) \leq 0 \implies \boldsymbol{x} \in \mathcal{L}. \tag{11b}$$

Observe:

- The goal of the pursuer, or $\boldsymbol{P}$, is to drive the system's trajectories into the unsafe set i.e., $\boldsymbol{P}$ has $\boldsymbol{u}$ at will and aims to minimize the termination time of the game (c.f. (7));
- The evader, or $\boldsymbol{E}$, seeks to avoid the unsafe set i.e., $\boldsymbol{E}$ has controls $\boldsymbol{v}$ at will and seeks to maximize the termination time of the game (c.f. (7));
- $\boldsymbol{E}$ has regular controls, $\boldsymbol{u}$, drawn from a Lebesgue measurable set, $\mathcal{U}$ (c.f. (9)).
- $\boldsymbol{P}$ possesses *nonanticipative strategies* (c.f. (9)) i.e. $\beta[\boldsymbol{u}](\cdot)$ such that for any of the ordinary controls, $\boldsymbol{u}(\cdot) \in \mathcal{U}$ of $\boldsymbol{E}$, $\boldsymbol{P}$ knows how to optimally respond to $\boldsymbol{E}$'s inputs.

This is a classic reachability problem on the resolution of the infimum-supremum over the *strategies* of $\boldsymbol{P}$ and *controls* of $\boldsymbol{E}$ with the time of capture resolved as an extremum of a cost functional) over a time interval.

For goal-satisfaction (liveness) problem setups, the strategies are flipped and the backward reachable tube instead marks

[2]The (backward) horizon T is negative.

the states from which the evader $\boldsymbol{E}$ can successfully reach the target set despite worst-case efforts of the pursuer $\boldsymbol{P}$.

Computing the value function is in general challenging and non-convex. Additionally, the value function is hardly smooth throughout the state space, so it lacks classical solutions even for smooth Hamiltonian and boundary conditions. However, the value function is a "viscosity" (generalized) solution [31, 11] of the associated HJ-Isaacs (HJI) PDE, i.e. solutions which are *locally Lipschitz* in $\Omega \times [0, T]$, and with at most first-order partial derivatives in the Hamiltonian. The HJI PDE is as follows:

$$\frac{\partial V}{\partial t}(\boldsymbol{x}, t) + \min\{0, \boldsymbol{H}^{-}(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}_{\boldsymbol{x}}^{-})\} = 0 \tag{12a}$$
$$V(\boldsymbol{x}, 0) = g(\boldsymbol{x}), \tag{12b}$$

where the vector field $V_{\boldsymbol{x}}$ is known in terms of the game's terminal conditions so that the overall game is akin to a two-point boundary-value problem. For more details on the construction of this PDE, see [37]. To solve for the value function one can discretize the state space and apply the HJI PDE using dynamic programming over the global mesh. Note that this PDE must be applied at every grid point in the state space and at every instant of time within the time horizon. As the system scales in dimension, this computation scales exponentially.

Once computed, the value function provides a safety certificate (defined by its zero level set) and corresponding safety controller (defined by the spatial gradients along the zero level set). SH: don't know if we should go into the online control portion of this or not

In addition to producing a value function, the dynamic programming process can be used to generate and minimum time-to-reach (TTR) function. This is a function that maps initial conditions to the minimum time horizon required to reach the target set. This can be computed by "stacking" the zero level sets of the value function as it propagates backwards in time. SH: cite Ian Mitchell, maybe Insoon

IV. Power Series Approximation of the Terminal Cost

Let us first to introduce the quadratic approximation scheme of the value function on a reduced basis before discussing the learning method for the optimal Galerkin spectral decomposition.

Suppose an (optimal) reduced basis with order r has been found that admits the most energetic modes of $\boldsymbol{V}$. Call the cost on this basis $\boldsymbol{V}_r$. Define the state, control, and disturbance on the reduced basis as $\boldsymbol{x}_r(t)$, $\boldsymbol{u}_r(t)$, and $\boldsymbol{v}_r(t)$ respectively, where $t \in [T, 0]$. When we decompose the system into the reduced $\boldsymbol{x}_r(t)$, $\boldsymbol{u}_r(t)$, and $\boldsymbol{v}_r(t)$, the dynamics no longer describes the original states and controls, but rather the variation of the state and controls on the reduced basis from the state and the control pairs on the nonlinear system of equation (1) i.e. $\delta\boldsymbol{x}(t)$, $\delta\boldsymbol{u}(t)$, and $\delta\boldsymbol{v}(t)$ respectively[3]. It follows that we can write the following relations

$$\boldsymbol{x}(t) = \boldsymbol{x}_r(t) + \delta\boldsymbol{x}(t), \;\; \boldsymbol{u}(t) = \boldsymbol{u}_r(t) + \delta\boldsymbol{u}(t), \tag{13a}$$
$$\boldsymbol{v}(t) = \boldsymbol{v}_r(t) + \delta\boldsymbol{v}(t), \;\; t \in [-T, 0]\,. \tag{13b}$$

For convenience' sake, let us drop the templated time arguments in (13) so that our canonical problem becomes

$$(1) \implies \frac{d}{dt}(\boldsymbol{x}_r + \delta\boldsymbol{x}) = f(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v}_r + \delta\boldsymbol{v}), \tag{14a}$$
$$\boldsymbol{x}_r(0) + \delta\boldsymbol{x}(0) = \boldsymbol{x}(0). \tag{14b}$$

(**??**) implies

$$-\frac{\partial \boldsymbol{V}}{\partial t}(\boldsymbol{x}_r + \delta\boldsymbol{x}, t) = \min\left\{0, \max_{\delta\boldsymbol{u}\in\mathcal{U}} \min_{\delta\boldsymbol{v}\in\mathcal{V}} \left\langle f(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v}_r + \delta\boldsymbol{v}), \frac{\partial \boldsymbol{V}}{\partial \boldsymbol{x}}(\boldsymbol{x}_r + \delta\boldsymbol{x}, t)\right\rangle\right\}.$$
$$(\mathbf{??}) \implies \boldsymbol{V}(\boldsymbol{x}_r, 0) = g(0; \boldsymbol{x}_r(0) + \delta\boldsymbol{x}(0)); \tag{15}$$

and

$$(4) \implies \boldsymbol{\xi}(t) = \boldsymbol{\xi}(t; t_0, \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u} + \delta\boldsymbol{u}, \boldsymbol{v} + \delta\boldsymbol{v}). \tag{16}$$

In particular, on the reduced order basis (ROB), the state dynamics now become

$$\dot{\boldsymbol{x}}_r(\tau) = f(t; \boldsymbol{x}_r(\tau), \boldsymbol{u}_r(\tau), \boldsymbol{v}_r(\tau)), \quad \tau \in [-T, 0] \tag{17a}$$
$$\boldsymbol{x}_r(0) = \boldsymbol{x}. \tag{17b}$$

Let the optimal cost for using the optimal control $\boldsymbol{u}^\star(\tau) = \boldsymbol{u}_r(\tau) + \delta\boldsymbol{u}^\star(\tau)$ when $\tau \in [t, 0]$ on the phase $(\boldsymbol{x}_r, t)$ be denoted $\boldsymbol{V}^\star(\boldsymbol{x}_r, t)$; and the ROM cost for using $\boldsymbol{u}_r(\tau)$; $\tau \in [t, 0]$ be $\boldsymbol{V}_r(\boldsymbol{x}_r, t)$. Suppose further that we denote the difference between these two costs on the phase $(\boldsymbol{x}_r, t)$ by $\tilde{\mathbf{V}}^\star$, then we have

$$\tilde{\mathbf{V}}^\star = \boldsymbol{V}^\star(\boldsymbol{x}_r, t) - \boldsymbol{V}_r(\boldsymbol{x}_r, t). \tag{18}$$

Theorem 1. *The HJI variational inequality c.f. (12) admits the following approximated expansion on the reduced model:*

$$-\frac{\partial \boldsymbol{V}_r}{\partial t} - \frac{\partial \tilde{\mathbf{V}}}{\partial t} - \left\langle \frac{\partial \boldsymbol{V}_{\boldsymbol{x}}}{\partial t}, \delta\boldsymbol{x}\right\rangle - \frac{1}{2}\left\langle \delta\boldsymbol{x}, \frac{\partial \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}}{\partial t}\delta\boldsymbol{x}\right\rangle = \min\left\{\mathbf{0}, \max_{\delta\boldsymbol{u}\in\mathcal{U}} \min_{\delta\boldsymbol{v}\in\mathcal{V}} \left\langle f^T(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v}_r + \delta\boldsymbol{v}), \boldsymbol{V}_{\boldsymbol{x}} + \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}\,\delta\boldsymbol{x}\right\rangle\right\}. \tag{19}$$

Furthermore, this expansion is bounded by $O(\delta\boldsymbol{x}^3)$.

Corollary 1. *If the viscosity solution obtained via a Lax-Friedrichs integration scheme for solving (19) converges to a local optimum, then the backward reachable tube will converge*

[3]Note that $\delta\boldsymbol{x}(t)$, $\delta\boldsymbol{u}(t)$, and $\delta\boldsymbol{v}(t)$ are respectively measured with respect to $\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{v}(t)$ and are not necessarily small. However, our case is very much helped when they are small and we conjecture that our decomposition scheme favors the smallness in the values of these variations.

to a locally optimal solution. In addition, if we overapproximate the resulting numerical solution, the reachable set or tube will converge to an optimal region in the state space.

Proof: The singular value decomposition of $\boldsymbol{V}$ is

$$\boldsymbol{V} = \boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{\Theta}^T, \tag{20}$$

where, $\boldsymbol{\Upsilon} \in \mathbb{C}^{n\times r}$, $\boldsymbol{\Lambda} \in \mathbb{C}^{r\times r}$, $\boldsymbol{\Theta} \in \mathbb{C}^{m\times r}$, and $r \le m$ can be an approximate or exact rank of $\boldsymbol{V}$. The modes of the reduced basis are the columns of $\boldsymbol{\Upsilon}$ which are ideally orthonormal i.e. $\boldsymbol{\Upsilon}^\star \boldsymbol{\Upsilon} = \mathbf{I}$. We are concerned with the leading eigen values and eigenvectors of $\boldsymbol{V}$; therefore, we project $\boldsymbol{V}$ onto the proper orthogonal decomposition (POD) modes in $\boldsymbol{\Upsilon}$ according to

$$\boldsymbol{V}_r = \boldsymbol{\Upsilon}^T \boldsymbol{V} \boldsymbol{\Upsilon}. \tag{21}$$

This reduced model is the Galerkin projection onto the semidiscrete ordinary differential equations (o.d.e.):

$$\frac{d\boldsymbol{V}_r}{dt} = \boldsymbol{\Upsilon}^T \frac{d\boldsymbol{V}}{dt}\boldsymbol{\Upsilon}. \tag{22}$$

For the moment, let us focus on the l.h.s. of (15). Our derivations closely follow that of Jacobson [23]. The major difference is that our choice of $\boldsymbol{x}_r$ is guaranteed to be close to that of $\boldsymbol{x}$ so that we need not prescribe stringent conditions for when local control laws are valid on the nonlinear system. Suppose the optimal terminal cost, $\boldsymbol{V}^\star$, is sufficiently smooth to allow a power series expansion in the state variation $\delta\boldsymbol{x}$ about reduced state, $\boldsymbol{x}_r$, we find that

$$\boldsymbol{V}^\star(\boldsymbol{x}_r + \delta\boldsymbol{x}, t) = \boldsymbol{V}^\star(\boldsymbol{x}_r, t) + \langle \boldsymbol{V}_{\boldsymbol{x}}, \delta\boldsymbol{x}\rangle + \frac{1}{2}\langle \delta\boldsymbol{x}, \boldsymbol{V}^\star_{\boldsymbol{x}\boldsymbol{x}}\delta\boldsymbol{x}\rangle + \text{h.o.t.} \tag{23}$$

Here, h.o.t. signifies higher order terms. This expansion scheme is consistent with Volterra-series model order reduction methods [19] or differential dynamic programming schemes that decompose nonlinear systems as a summation of Taylor series expansions [24]. Using (18), (23) becomes

$$\boldsymbol{V}^\star(\boldsymbol{x}_r + \delta\boldsymbol{x}, t) = \boldsymbol{V}_r(\boldsymbol{x}_r, t) + \tilde{\mathbf{V}}^\star + \langle \boldsymbol{V}_{\boldsymbol{x}}, \delta\boldsymbol{x}\rangle + \frac{1}{2}\langle \delta\boldsymbol{x}, \boldsymbol{V}^\star_{\boldsymbol{x}\boldsymbol{x}}\delta\boldsymbol{x}\rangle + \text{h.o.t.} \tag{24}$$

The expansion in (24) may be more costly than solving for the original value function owing to the large dimensionality of the states as higher order terms are expanded. However, consider:

- $\boldsymbol{V}_r(\boldsymbol{x}_r, t)$ already contains the dominant modes of $\boldsymbol{V}(\boldsymbol{x}, t)$ as a result of the singular value decomposition scheme; therefore w.l.o.g. states in the reduced order basis (ROB), $\boldsymbol{V}_r(\boldsymbol{x}_r, t)$, will be sufficiently close to those that originate in (1);
- If the above is true, the state variation $\delta\boldsymbol{x}$ will be sufficiently small owing to the fact that $\boldsymbol{x} \approx \boldsymbol{x}_r$ c.f. (13).

Therefore, we can avoid the infinite data storage requirement by truncating the expansion in (24) at, say, the quadratic (second-order) terms in $\delta\boldsymbol{x}$. Seeing that $\delta\boldsymbol{x}$ is sufficiently small, the second-order cost terms will dominate higher order terms, and this new cost will result in an $O(\delta\boldsymbol{x}^3)$ approximation error, affording us realizable control laws that can be executed on the system (1). From (23), we have

$$\boldsymbol{V}^\star(\boldsymbol{x}_r + \delta\boldsymbol{x}, t) = \boldsymbol{V}_r + \tilde{\mathbf{V}}^\star + \langle \boldsymbol{V}_{\boldsymbol{x}}, \delta\boldsymbol{x}\rangle + \frac{1}{2}\langle \delta\boldsymbol{x}, \boldsymbol{V}^\star_{\boldsymbol{x}\boldsymbol{x}}\delta\boldsymbol{x}\rangle. \tag{25}$$

Denoting by $\boldsymbol{V}^\star_{\boldsymbol{x}}$ the co-state on the r.h.s of (15), we can similarly expand it up to second order terms as follows

$$\boldsymbol{V}^\star_{\boldsymbol{x}}(\boldsymbol{x}_r + \delta\boldsymbol{x}, t) = \frac{\partial \boldsymbol{V}^\star_r}{\partial \boldsymbol{x}}(\boldsymbol{x}_r, t) + \langle \boldsymbol{V}^\star_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}_r, t), \delta\boldsymbol{x}\rangle. \tag{26}$$

Note that *the co-state in* (26) *and parameters on the r.h.s. of* (25) *are evaluated on the reduced model, specifically at the phase* $(\boldsymbol{x}_r, t)$. Substituting (25) and (26) into (15), abusing notation by dropping the superscripts and the templated phase arguments, we find that

$$-\frac{\partial \boldsymbol{V}_r}{\partial t} - \frac{\partial \tilde{\mathbf{V}}}{\partial t} - \left\langle \frac{\partial \boldsymbol{V}_{\boldsymbol{x}}}{\partial t}, \delta\boldsymbol{x}\right\rangle - \frac{1}{2}\left\langle \delta\boldsymbol{x}, \frac{\partial \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}}{\partial t}\delta\boldsymbol{x}\right\rangle = \min\Big\{0, \max_{\delta\boldsymbol{u}}\min_{\delta\boldsymbol{v}} \langle f^T(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v}_r + \delta\boldsymbol{v}), \boldsymbol{V}_{\boldsymbol{x}} + \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}\delta\boldsymbol{x}\rangle\Big\}. \tag{27}$$

Observe that $\boldsymbol{V}_r + \tilde{\mathbf{V}}$, $\boldsymbol{V}_{\boldsymbol{x}}$, and $\boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}$ are all functions of the phase $(\boldsymbol{x}, r)$ so that

$$\frac{d}{dt}\left(\boldsymbol{V}_r + \tilde{\mathbf{V}}\right) = \frac{\partial}{\partial t}\left(\boldsymbol{V}_r + \tilde{\mathbf{V}}\right) + \langle f^T(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r), \boldsymbol{V}_{\boldsymbol{x}}\rangle \tag{28a}$$

$$\dot{\boldsymbol{V}}_{\boldsymbol{x}} = \frac{\partial \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}}{\partial t} + \langle f^T(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r), \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}\rangle \tag{28b}$$

$$\dot{\boldsymbol{V}}_{\boldsymbol{x}\boldsymbol{x}} = \frac{\partial \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}}{\partial t}. \tag{28c}$$

■

The left hand side of (27) admits a quadratic form, so that we can regress a quadratic form to fit the functionals and derivatives of the optimal structure of the ROB. The r.h.s. can be similarly expanded as above. Define

$$\boldsymbol{H}(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}_{\boldsymbol{x}}) = \langle \boldsymbol{V}_{\boldsymbol{x}}, f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v})\rangle \tag{29}$$

so that (27) becomes

$$-\frac{\partial \boldsymbol{V}_r}{\partial t} - \frac{\partial \tilde{\mathbf{V}}}{\partial t} - \left\langle \frac{\partial \boldsymbol{V}_{\boldsymbol{x}}}{\partial t}, \delta\boldsymbol{x}\right\rangle - \frac{1}{2}\left\langle \delta\boldsymbol{x}, \frac{\partial \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}}{\partial t}\delta\boldsymbol{x}\right\rangle = \min\Big\{\mathbf{0}, \max_{\delta\boldsymbol{u}}\min_{\delta\boldsymbol{v}} [\boldsymbol{H}(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v} + \delta\boldsymbol{v}, \boldsymbol{V}_{\boldsymbol{x}}) + \langle \boldsymbol{V}_{\boldsymbol{x}\boldsymbol{x}}\delta\boldsymbol{x}, f(t; \boldsymbol{x}_r + \delta\boldsymbol{x}, \boldsymbol{u}_r + \delta\boldsymbol{u}, \boldsymbol{v}_r + \delta\boldsymbol{v})\rangle]\Big\}. \tag{30}$$

Expanding the r.h.s. about $\boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r$ up to second-order only[4], we find that

$$\min\Big\{\mathbf{0}, \max_{\delta \boldsymbol{u}} \min_{\delta \boldsymbol{v}} [\boldsymbol{H} + \langle \boldsymbol{H_x} + \boldsymbol{V_{xx}} f, \delta \boldsymbol{x}\rangle + \langle \boldsymbol{H_u}, \delta \boldsymbol{u}\rangle + \langle \boldsymbol{H_v}, \delta \boldsymbol{v}\rangle + \langle \delta \boldsymbol{u}, (\boldsymbol{H_{ux}} + f_{\boldsymbol{u}}^T \boldsymbol{V_{xx}})\delta \boldsymbol{x}\rangle + \left\langle \delta \boldsymbol{v}, (\boldsymbol{H_{vx}} + f_{\boldsymbol{v}}^T \boldsymbol{V_{xx}})\delta \boldsymbol{x}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{u}, \boldsymbol{H_{uv}} \delta \boldsymbol{v}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{v}, \boldsymbol{H_{vu}} \delta \boldsymbol{u}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{u}, \boldsymbol{H_{uu}} \delta \boldsymbol{u}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{v}, \boldsymbol{H_{vv}} \delta \boldsymbol{v}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{x}, \left(\boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}}\right) \delta \boldsymbol{x}\right\rangle\Big]\Big\}. \tag{31}$$

Let us recall that when capture[5] occurs, we must have the Hamiltonian of the value function be zero as a necessary condition for the players' saddle-point controls [34, 22] i.e.

$$\boldsymbol{H_u}(t; \boldsymbol{x}_r, \boldsymbol{u}_r^\star, \boldsymbol{v}_r, \boldsymbol{V_x}) = 0; \ \boldsymbol{H_v}(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r^\star, \boldsymbol{V_x}) = 0. \tag{32}$$

where $\boldsymbol{u}_r^\star$ and $\boldsymbol{v}_r^\star$ respectively represent the optimal control laws for both players at time t.

A state-control relationship of the following form is sought:

$$\delta \boldsymbol{u} = \boldsymbol{k_u} \delta \boldsymbol{x}, \quad \delta \boldsymbol{v} = \boldsymbol{k_v} \delta \boldsymbol{x} \tag{33}$$

so that (31) in the context of (32) yields

$$\boldsymbol{H_u} + \boldsymbol{H_{uu}} \delta \boldsymbol{u} + \left(\boldsymbol{H_{ux}} + f_{\boldsymbol{u}}^T \boldsymbol{V_{xx}}\right) \delta \boldsymbol{x} + \frac{1}{2} \boldsymbol{H_{uv}} \delta \boldsymbol{v} = 0 \tag{34a}$$

$$\boldsymbol{H_v} + \boldsymbol{H_{vv}} \delta \boldsymbol{v} + \left(\boldsymbol{H_{vx}} + f_{\boldsymbol{v}}^T \boldsymbol{V_{xx}}\right) \delta \boldsymbol{x} + \frac{1}{2} \boldsymbol{H_{vu}} \delta \boldsymbol{u} = 0. \tag{34b}$$

Using (32) and equating like terms in the resulting equation to those in (33), we have the following for the state gains:

$$\boldsymbol{k_u} = -\frac{1}{2} \boldsymbol{H_{uu}}^{-1} \left[\boldsymbol{H_{uv}} \boldsymbol{k_v} + 2\left(\boldsymbol{H_{ux}} + f_{\boldsymbol{u}}^T \boldsymbol{V_{xx}}\right)\right], \text{ and} \tag{35}$$
$$\boldsymbol{k_v} = -\frac{1}{2} \boldsymbol{H_{vv}}^{-1} \left[\boldsymbol{H_{vu}} \boldsymbol{k_u} + 2\left(\boldsymbol{H_{vx}} + f_{\boldsymbol{v}}^T \boldsymbol{V_{xx}}\right)\right].$$

Putting the maximizing $\delta \boldsymbol{u}$ and the minimizing $\delta \boldsymbol{v}$ into (31), whilst neglecting terms in $\delta \boldsymbol{x}$ beyond second-order, we have

$$\min\left\{\mathbf{0}, \left[\boldsymbol{H} + \left\langle\left(\boldsymbol{H_x} + \boldsymbol{V_{xx}} f + \boldsymbol{k_u}^T \boldsymbol{H_u} + \boldsymbol{k_v}^T \boldsymbol{H_v}\right), \delta \boldsymbol{x}\right\rangle + \frac{1}{2}\left\langle \delta \boldsymbol{x}, \left(\boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}} + \boldsymbol{k_u}^T \boldsymbol{H_{uu}} \boldsymbol{k_u} + \boldsymbol{k_v}^T \boldsymbol{H_{vv}} \boldsymbol{k_v}\right) \delta \boldsymbol{x}\right\rangle\right]\right\}. \tag{36}$$

Now, we can compare coefficients with the l.h.s. of (30) and find the quadratic expansion of the reduced value function admits the following analytical solution on its right hand side:

$$-\frac{\partial \boldsymbol{V}_r}{\partial t} - \frac{\partial \tilde{\boldsymbol{V}}}{\partial t} = \min\{\mathbf{0}, \boldsymbol{H}\} \tag{37a}$$

$$-\frac{\partial \boldsymbol{V_x}}{\partial t} = \min\left\{\mathbf{0}, \boldsymbol{H_x} + \boldsymbol{V_{xx}}\, f + \boldsymbol{k_u}^T \boldsymbol{H_u} + \boldsymbol{k_v}^T \boldsymbol{H_v}\right\} \tag{37b}$$

$$-\frac{\partial \boldsymbol{V_{xx}}}{\partial t} = \min\left\{\mathbf{0}, \boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}} + \boldsymbol{k_u}^T \boldsymbol{H_{uu}} \boldsymbol{k_u} + \boldsymbol{k_v}^T \boldsymbol{H_{vv}} \boldsymbol{k_v}\right\}. \tag{37c}$$

Furthermore, comparing the above with (28) and noting that $-\dot{\boldsymbol{V}}_r = 0$[6], we find that

$$-\dot{\tilde{\boldsymbol{V}}} = -\frac{\partial \tilde{\boldsymbol{V}}}{\partial t} \triangleq \min\{\mathbf{0}, \boldsymbol{H} - \boldsymbol{H}(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r, \boldsymbol{V_x})\} \tag{38a}$$

$$-\dot{\boldsymbol{V}}_{\boldsymbol{x}} = \min\left\{\mathbf{0}, \boldsymbol{H_x} + \boldsymbol{V_{xx}}\left(f - f(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r)\right) \right. \tag{38b}$$
$$\left. + \boldsymbol{k_u}^T \boldsymbol{H_u} + \boldsymbol{k_v}^T \boldsymbol{H_v}\right\} \tag{38c}$$

$$-\frac{\partial \boldsymbol{V_{xx}}}{\partial t} = \min\left\{\mathbf{0}, \boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}} + \boldsymbol{k_u}^T \boldsymbol{H_{uu}} \boldsymbol{k_u} + \boldsymbol{k_v}^T \boldsymbol{H_{vv}} \boldsymbol{k_v}\right\} \tag{38d}$$

where $\boldsymbol{k_u}$ and $\boldsymbol{k_v}$ are as defined in (33). Note that at a saddle point, the first-order necessary condition for optimality c.f. (32) implies

$$-\dot{\tilde{\boldsymbol{V}}} = \min\{0, \boldsymbol{H} - \boldsymbol{H}(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r, \boldsymbol{V_x})\} \tag{39a}$$

$$-\dot{\boldsymbol{V}}_{\boldsymbol{x}} = \min\left\{0, \boldsymbol{H_x} + \boldsymbol{V_{xx}}\left(f - f(\boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r)\right)\right\} \tag{39b}$$

$$-\frac{\partial \boldsymbol{V_{xx}}}{\partial t} = \min\left\{\mathbf{0}, \boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}} + \boldsymbol{k_u}^T \boldsymbol{H_{uu}} \boldsymbol{k_u} + \boldsymbol{k_v}^T \boldsymbol{H_{vv}} \boldsymbol{k_v}\right\} \tag{39c}$$

whereupon every quantity in (39) is evaluated at $\boldsymbol{x}_r, \boldsymbol{u}^\star$.

The boundary conditions for (39) at $t = 0$ is

$$\boldsymbol{V}(\boldsymbol{x}_r, 0) = \boldsymbol{g}(0; \boldsymbol{x}_r(0)); \tag{40}$$

so that

$$\tilde{\boldsymbol{V}}(0) = 0 \tag{41a}$$
$$\boldsymbol{V_x}(0) = \boldsymbol{g_x}(0; \boldsymbol{x}_r(0)) \tag{41b}$$
$$\boldsymbol{V_{xx}}(0) = \boldsymbol{g_{xx}}(0; \boldsymbol{x}_r(0)). \tag{41c}$$

The following control laws are then applied

$$\boldsymbol{u} = \boldsymbol{u}_r + \boldsymbol{k_u} \delta \boldsymbol{x}, \tag{42}$$
$$\boldsymbol{v} = \boldsymbol{v}_r + \boldsymbol{k_v} \delta \boldsymbol{x}. \tag{43}$$

Therefore, at any time on a ROB of the value function, a local approximation of $\boldsymbol{V}$ consists in employing the TO-DO: Lax-Friedrichs scheme on the following system

$$-\left[\boldsymbol{E} + \boldsymbol{F} \delta \boldsymbol{x} + \frac{1}{2} \delta \boldsymbol{x} \boldsymbol{G} \delta \boldsymbol{x}\right] = \min\left\{\mathbf{0}, \boldsymbol{H} - \boldsymbol{H}(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r, \boldsymbol{V_x}) + \boldsymbol{H_x} + \boldsymbol{V_{xx}}\left(f - f(t; \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{v}_r)\right) + \boldsymbol{H_{xx}} + f_{\boldsymbol{x}}^T \boldsymbol{V_{xx}} + \boldsymbol{V_{xx}} f_{\boldsymbol{x}} + \boldsymbol{k_u}^T \boldsymbol{H_{uu}} \boldsymbol{k_u} + \boldsymbol{k_v}^T \boldsymbol{H_{vv}} \boldsymbol{k_v}\right\} \tag{44}$$

where $\boldsymbol{E}, \boldsymbol{F},$ and $\boldsymbol{G}$ are appropriately defined.

[4]This is because the l.h.s. was truncated at second order expansion previously. Ultimately, the $\delta \boldsymbol{u}, \delta \boldsymbol{v}$ terms will be quadratic in $\delta \boldsymbol{x}$ if we neglect h.o.t.

[5]A capture occurs when $\boldsymbol{E}$'s separation from $\boldsymbol{P}$ becomes less than a specified e.g. capture radius.

[6]The stage cost is zero from (??).

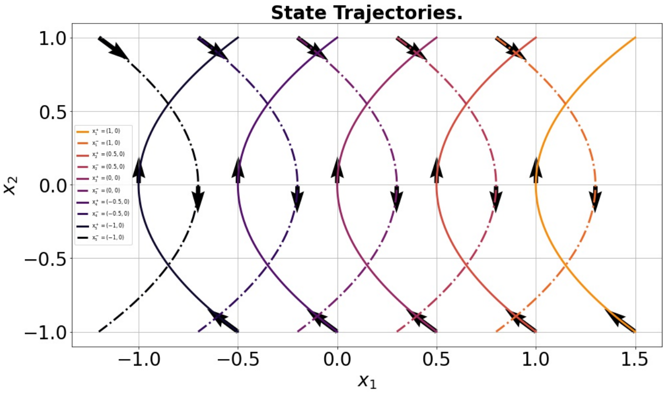

Fig. 1: State trajectories of the double integral plant. The solid curves are trajectories generated for $\boldsymbol{u} = +1$ while the dashed curves are trajectories for $\boldsymbol{u} = -1$.

V. Results and Discussion.

We now provide results and analysis of the proposed numerical algorithm on benchmark control problems.

A. Time Optimal Control of the Double Integrator

Here, we analyze our proposal on a time-optimal control problem. Specifically, we consider the double integral plant which has the following second-order dynamics

$$\ddot{\boldsymbol{x}}(t) = \boldsymbol{u}(t). \tag{45}$$

This admits bounded control signals $\mid \boldsymbol{u}(t) \mid \leq 1$ for all t. After a change of variables,we have the following system of first-order differential equations

$$\dot{\boldsymbol{x}}_1(t) = \boldsymbol{x}_2(t), \quad \dot{\boldsymbol{x}}_2(t) = \boldsymbol{u}(t), \quad \mid \boldsymbol{u}(t) \mid \leq 1.$$

The *reachability problem is to address the possibility of reaching all points in the state space in a* ***transient*** *manner*. Therefore, we set the running cost to zero, so that the Hamiltonian is $H = p_1\dot{\boldsymbol{x}}_1 + p_2\dot{\boldsymbol{x}}_2$. The necessary optimality condition stipulates that the minimizing control law is $\boldsymbol{u}(t) = -\operatorname{sign}(p_2(t))$. On a finite time interval, say, $t \in [t_0, t_f]$, the time-optimal $\boldsymbol{u}(t)$ is a constant k so that for initial conditions $\boldsymbol{x}_1(t_0) = \boldsymbol{\xi}_1$ and $\boldsymbol{x}_2(t_0) = \boldsymbol{\xi}_2$, it can be verified that the state trajectories obey the relation

$$\boldsymbol{x}_1(t) = \boldsymbol{\xi}_1 + \frac{1}{2}k\left(\boldsymbol{x}_2^2 - \boldsymbol{\xi}_2^2\right), \text{ where, } t = k\left(\boldsymbol{x}_2(t) - \boldsymbol{\xi}_2\right). \tag{46}$$

The trajectories traced out over a finite time horizon $t = [-1, 1]$ on a state space and under the control laws $\boldsymbol{u}(t) = \pm 1$ is depicted in Fig. 1. The curves with arrows that point upwards denote trajectories under the control law $\boldsymbol{u} = +1$; call these trajectories $\boldsymbol{\gamma}_+$; while the trajectories with dashed curves and downward pointing arrows were executed under $\boldsymbol{u} = -1$; call these trajectories $\boldsymbol{\gamma}_-$. The time to go from any point on any of the intersections to the origin on the state trajectories of Fig. 1 is our approximation problem. This minimum time

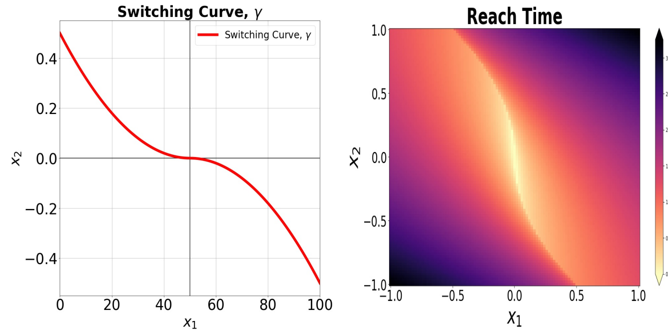

Fig. 2: (L) Switching curve for the double integral plant. (R) Analytical time to reach the origin on the state grid, $(\mathbb{R} \times \mathbb{R})$; the switching curve corresponds to the bright orange coloration for states on $(0, 0)$.

admits an analytical solution [1] given by

$$t^\star(\boldsymbol{x}_1, \boldsymbol{x}_2) = \begin{cases} \boldsymbol{x}_2 + \sqrt{4\boldsymbol{x}_1 + 2\boldsymbol{x}_2^2} & \text{if } \boldsymbol{x}_1 > \frac{1}{2}\boldsymbol{x}_2|\boldsymbol{x}_2| \\ -\boldsymbol{x}_2 + \sqrt{-4\boldsymbol{x}_1 + 2\boldsymbol{x}_2^2} & \text{if } \boldsymbol{x}_1 < -\frac{1}{2}\boldsymbol{x}_2|\boldsymbol{x}_2| \\ \mid \boldsymbol{x}_2 \mid & \text{if } \boldsymbol{x}_1 = \frac{1}{2}\boldsymbol{x}_2|\boldsymbol{x}_2|. \end{cases} \tag{47}$$

Let us define R_+ as the portions of the state space above the curve $\boldsymbol{\gamma}$ and R_- as the portions of the state space below the curve $\boldsymbol{\gamma}$. The confluence of the locus of points on $\boldsymbol{\gamma}_+$ and $\boldsymbol{\gamma}_-$ is the switching curve, depicted on the left inset of Fig. 2, and given as

$$\boldsymbol{\gamma} \triangleq \boldsymbol{\gamma}_+ \cup \boldsymbol{\gamma}_- = \left\{ (\boldsymbol{x}_1, \boldsymbol{x}_2) : \boldsymbol{x}_1 = \frac{1}{2}\boldsymbol{x}_2 \mid \boldsymbol{x}_2 \mid \right\}. \tag{48}$$

We now state the **time-optimal control problem**: *The control problem is to find the control law that forces* (46) *to the origin* $(0, 0)$ *in the* ***shortest possible time***. The time-optimal control law, $\boldsymbol{u}^\star$, that solves this problem is unique and is

$$\begin{aligned} \boldsymbol{u}^\star &= \boldsymbol{u}^\star(\boldsymbol{x}_1, \boldsymbol{x}_2) = +1 \quad \forall\, (\boldsymbol{x}_1, \boldsymbol{x}_2) \in \boldsymbol{\gamma}_+ \cup \mathbb{R}+ \\ \boldsymbol{u}^\star &= \boldsymbol{u}^\star(\boldsymbol{x}_1, \boldsymbol{x}_2) = -1 \quad \forall\, (\boldsymbol{x}_1, \boldsymbol{x}_2) \in \boldsymbol{\gamma}_- \cup \mathbb{R}-. \end{aligned} \tag{49}$$

We compare our approximated terminal value solution using our proposal against (i) the numerical solution found via level sets methods [44] and (ii) the analytical solution of the *time to reach (TTR) the origin* problem.

VI. Parallelization Scheme for Resolving the Separated Hamiltonians

In this section, we describe the alternating direction method of multipliers that

A. Bregman Alternating Method of Multipliers

TO-DO: Methinks we should have a BADMM alorithm for the fast numerical resolution of the Hamiltonians. Who wants to write this section.

VII. Conclusion.

Acknowledgments

References

[1] Michael Athans and Peter L Falb. *Optimal Control: An Introduction to the Theory and its Applications*. Courier Corporation, 2013. 7

[2] Andrea Bajcsy, Somil Bansal, Eli Bronstein, Varun Tolani, and Claire J Tomlin. An Efficient Reachability-based Framework for Provably Safe Autonomous Navigation in Unknown Environments. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 1758–1765. IEEE, 2019. 2

[3] Grey Ballard, Alicia Klinvex, and Tamara G. Kolda. TuckerMPI: A Parallel C++/MPI Software Package for Large-scale Data Compression via the Tucker Tensor Decomposition. *ACM Transactions on Mathematical Software*, 46(2), 2020. ISSN 15577295. doi: 10.1145/3378445.

[4] Somil Bansal and Claire J Tomlin. DeepReach : A Deep Learning Approach to High-Dimensional Reachability. 2

[5] Tamer Basar and Geert Jan Olsder. *Dynamic noncooperative game theory*, volume 23. Siam, 1999.

[6] Richard Bellman. *Dynamic programming*. Princeton University Press, 1957. ISBN 0-486-42809-5. 1, 2

[7] Stephen Boyd, Neal Parikh, and Eric Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[8] Francisco Chinesta, Pierre Ladeveze, and Elias Cueto. A Short Review on Model Order Reduction based on Proper Generalized Decomposition. *Archives of Computational Methods in Engineering*, 18(4):395–404, 2011.

[9] M. G. Crandall, L. C. Evans, and P. L. Lions. Some Properties of Viscosity Solutions of Hamilton-Jacobi Equations. *Transactions of the American Mathematical Society*, 282(2):487, 1984. ISSN 00029947. doi: 10.2307/1999247. 1

[10] Michael Crandall and Andrew Majda. The method of fractional steps for conservation laws. *Numerische Mathematik*, 34(3):285–314, 1980. ISSN 0029599X. doi: 10.1007/BF01396704.

[11] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983. 1, 2, 3, 4

[12] Michael G Crandall and Andrew Majda. Monotone difference approximations for scalar conservation laws. *Mathematics of Computation*, 34(149):1–21, 1980. 1

[13] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. A multilinear singular value decomposition. *SIAM journal on Matrix Analysis and Applications*, 21 (4):1253–1278, 2000.

[14] W.F. Denham. Book Reviews: Differential Dynamic Programming. *Review of Social Economy*, 36(2):228–229, 1978. ISSN 1470-1162. 1

[15] L.C. Evans and Panagiotis E. Souganidis. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. *Indiana Univ. Math. J*, 33(5):773–797, 1984. ISSN 0022-2518. 1, 3

[16] L.C. Evans and Panagiotis E. Souganidis. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. *Indiana Univ. Math. J*, 33(5):773–797, 1984. ISSN 0022-2518. 1, 2, 3

[17] Jaime F. Fisac, Anayo K. Akametalu, Melanie N. Zeilinger, Shahab Kaynama, Jeremy Gillula, and Claire J. Tomlin. A General Safety Framework for Learning-Based Control in Uncertain Robotic Systems. *IEEE Transactions on Automatic Control*, 64(7):2737–2752, 2019. ISSN 15582523. doi: 10.1109/TAC.2018.2876389. 2

[18] Jaime F. Fisac, Neil F. Lugovoy, Vicenc Rubies-Royo, Shromona Ghosh, and Claire J. Tomlin. Bridging hamilton-jacobi safety analysis and reinforcement learning. *Proceedings - IEEE International Conference on Robotics and Automation*, 2019-May:8550–8556, 2019. ISSN 10504729. doi: 10.1109/ICRA.2019.8794107. 2

[19] Chenjie Gu. Qlmor: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 30(9):1307–1320, 2011. 5

[20] Wolfgang Hackbusch. *Tensor Spaces and Numerical Tensor Calculus*. Springer-Verlag, 2012.

[21] Sylvia Herbert, Jason J Choi, Suvansh Sanjeev, Marsalis Gibson, Koushil Sreenath, and Claire J Tomlin. Scalable learning of safety guarantees for autonomous systems using hamilton-jacobi reachability. *arXiv preprint arXiv:2101.05916*, 2021. 2

[22] R Isaacs. Differential games 1965. *Kreiger, Huntigton, NY*. 1, 6

[23] David H Jacobson. New second-order and first-order algorithms for determining optimal control: A differential dynamic programming approach. *Journal of Optimization Theory and Applications*, 2(6):411–440, 1968. 1, 5

[24] David H. Jacobson and David Q. Mayne. *Differential Dynamic Programming*. American Elsevier Publishing Company, Inc., New York, NY, 1970. 1, 2, 5

[25] L.V. Kantarovich and G.P. Akilov. *Functional Analysis in Normed Spaces*. Pergamon Press, 1964.

[26] Arie Kapteyn, Heinz Neudecker, and Tom Wansbeek. An Approach to n-mode Components Analysis. *Psychometrika*, 51(2):269–275, 1986.

[27] Shahab Kaynama, Ian M. Mitchell, Meeko Oishi, and Guy A. Dumont. Scalable safety-preserving robust control synthesis for continuous-time linear systems. *IEEE Transactions on Automatic Control*, 60(11):3065–3070, 2015. 2

[28] Tamara G. Kolda and Brett W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009. ISSN 00361445.

[29] Pierre Ladevèze. *Nonlinear computational structural mechanics: new approaches and non-incremental methods*

of calculation. Springer Science & Business Media, 2012.

[30] Weiwei Li and Emanuel Todorov. Iterative linear quadratic regulator design for nonlinear biological movement systems. In *ICINCO (1)*, pages 222–229. Citeseer, 2004. 1

[31] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69. London Pitman, 1982. 1, 4

[32] John Lygeros. On reachability and minimum cost optimal control. *Automatica*, 40(6):917–927, 2004. 2

[33] Stephen R. McReynolds. The successive sweep method and dynamic programming. *Journal of Mathematical Analysis and Applications*, 19(3):565–598, 1967. ISSN 10960813. 2

[34] AW Merz. The game of two identical cars. *Journal of Optimization Theory and Applications*, 9(5):324–343, 1972. 2, 6

[35] Ian Mitchell. A Robust Controlled Backward Reach Tube with (Almost) Analytic Solution for Two Dubins Cars (Presentation). 74:242–224, 2020. doi: 10.29007/mx3f. 1, 2

[36] Ian M Mitchell. A Toolbox of Level Set Methods. *UBC Department of Computer Science Technical Report TR-2007-11*, 2007.

[37] Ian M. Mitchell, Alexandre M. Bayen, and Claire J. Tomlin. A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. *IEEE Transactions on Automatic Control*, 50(7):947–957, 2005. ISSN 00189286. doi: 10.1109/TAC.2005.851439. 2, 3, 4

[38] S. K. Mitter. Successive approximation methods for the solution of optimal control problems. *Automatica*, 3(3-4):135–149, 1966. ISSN 00051098. 2

[39] ROYUD Nishino, Crissman Loomis, and Shohei Hido. Cupy: A numpy-compatible library for nvidia gpu calculations. *31st confernce on neural information processing systems*, 151, 2017.

[40] Anthony Nouy. Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms. *Computer Methods in Applied Mechanics and Engineering*, 197(51-52):4718–4736, 2008.

[41] Anthony Nouy. A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1603–1626, 2010. ISSN 00457825.

[42] Anthony Nouy and Olivier P Le Maıˆtre. Generalized spectral decomposition for stochastic nonlinear problems. *Journal of Computational Physics*, 228(1):202–235, 2009.

[43] Olalekan Ogunmolu, Nicholas Gans, and Tyler Summers. Minimax iterative dynamic game: Application to nonlinear robot control tasks. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6919–6925. IEEE, 2018. 2

[44] S Osher and R Fedkiw. Level Set Methods and Dynamic Implicit Surfaces. *Applied Mechanics Reviews*, 57(3):B15–B15, 2004. ISSN 0003-6900. 7

[45] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988. 2

[46] Stanley Osher and Chi-Wang Shu. High-Order Essentially Nonoscillatory Schemes for Rotary-Wing Wake Computations. *Journal of Aircraft*, 41(2):258–267, 2004. ISSN 15333868. doi: 10.2514/1.9320. 1

[47] Ledyard R Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31(3):279–311, 1966.

[48] Nick Vannieuwenhoven, Raf Vandebril, and Karl Meerbergen. A New Truncation Strategy for the Higher-Order Singular Value Decomposition. *SIAM Journal on Scientific Computing*, 34(2):A1027–A1052, 2012.