An Iterative Galerkin Approximation Scheme for Certain Backward Reachability Problems.

***Abstract*—Backward reachability analysis is premised on solving implicitly-constructed value functions on spatio-temporal grids in order 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 particle* ore: an iterative quasi-proper generalized decomposition (QPGD) scheme that incrementally truncates a high-dimensional value function to the minimum low-rank tensor necessary for computing reachable sets and tubes with quasi-optimality guarantees. Our scheme is essentially a proper generalized decomposition (PGD) *a priori* model order reduction technique that allows us to tame the curse of dimensionality, and preserve reachability analyses guarantees up to a projection approximation error in numerical verification of complex systems with Cauchy-type Hamilton-Jacobi Isaacs terminal values. This paper presents an initial evaluation of our proposal on the BRS/BRTs of multiple pursuer-evader system of non-stochastic dynamical systems.**

I. INTRODUCTION.

Designed cyberphysical systems (CPS) are a complex interconnection of control systems, sensors, and their software whose communication protocols have created complex entanglements with interactions that are difficult to analyze. CPS are traditionally engineered to sense and interact with the physical world "smartly". Modern cyberphysical systems may include modern manufacturing assembly lines where humans and machines jointly work to deliver products to a supply chain controlled by computer software resources, personalized interoperable medical devices, autonomous cars on a highway, (almost unmanned) long-hauled passenger flights, or general logistics inter alia.

The "physical" and "cyber" couplings of such systems is critical in modern CPS infrastructure: generating control laws – where dynamics may be complex and not necessarily known in advance; planning and executing in real-time collision avoidance schemes in uneven terrains, or sensing efficiently in the presence of multiple agents – all require deep integration and the 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) [1] Lax-Friedrichs [2] schemes applied to numerically resolve HJ Hamiltonians [3], 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* [4]. Truncated power series methods [5]–[8] 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.

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 [9], 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* [10]. Extensions to *Dirichlet-type* HJ equations are straightforward.

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 [11] on reduced Hilbertian spaces); and (iii) analytic saddle solutions to approximated HJ value functions. to synthesize approximately optimal control laws (essentially, saddle-point solutions) TODO: with stability guarantees for resolving the terminal value in the viscosity solutions to HJI value functions.

In order to analyze the safety of emerging CPS systems given the computational and memory drawbacks of level sets methods, it is the opinion of the authors that

- easily implementable approximation schemes with -ity guarantees;
- stability in well-defined regions of the state space where approximation is guaranteed to work;

[1] Microsoft Research, 300 Lafayette Street, New York, NY10012, USA. [2]Department of Mechanical and Aerospace Engineering, Yale University, New Haven, Connecticut, USA; [3]Department of Mechanical and Aerospace Engineering, UC San Diego, La Jolla, CA, USA. lekanmolu@microsoft.com, ian.abraham@yale.edu, sherbert@eng.ucsd.edu.

- and low run-time computation and memory requirements that address the jugular of the curse of dimensionality; whilst
- providing bounds on the error in the approximation,

are the best means for tackling this problem.

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

II. Notations and Definitions.

Throughout this article, time variables e.g. t, t_0, τ, T will always be integers. Vectors will be denoted by small boldface letters such as $\mathbf{e}, \mathbf{u}, \mathbf{v}$ e.t.c. An n-dimensional vector will be the set $\{x_1, x_2, \ldots, x_n\}$. Unless otherwise noted, vector elements will be column-wise stacked. When we refer to a row-vector, we shall introduce the transpose as a superscript i.e. $\boldsymbol{x}^T$. Matrices and tensors will respectively be denoted by bold-math upper case Roman and double stroke font letters e.g. $\mathbf{T}, \mathbf{S}$ (resp. $\mathbb{T}$, $\mathbb{S}$). We designate uppercase letters I, J, K, L, M, N, P, R for tensor sizes (the total number of elements encompassed along a dimension of a tensor), and lowercase letters i, j, k, l, m, n, p, r for corresponding tensor indices. We adopt zero-indexing for matrix and tensor operations throughout such that if index i corresponds to size I, we write $i = 0, 1, \cdots, I-1$. Lastly, for a size N, we denote by $[N]$ the set $\{0, 1, \cdots, N-1\}$.

A. Vectors, Matrices, and Tensors.

1) Vectors: We define the *direction cosines* of the orthonormal basis $\{\mathbf{e}'_i\}$ oriented with respect to $\{\mathbf{e}_j\}$ as $\mathbf{Q}_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j$. so that by orthonormality and by $\mathbf{e}'_i = \mathbf{Q}_{ik}\mathbf{e}_k \; \forall i = (1, 2, 3)$, we have $\delta_{ij} = \mathbf{e}'_i \cdot \mathbf{e}'_j = \mathbf{Q}_{ik}\,\mathbf{e}_k \cdot \mathbf{e}'_j = \mathbf{Q}_{ik}\mathbf{Q}_{jk}$, where δ_{ij} is the Krönecker delta symbol. The *triple scalar product* $(\mathbf{u} \wedge \mathbf{v}) \cdot \mathbf{w}$ is $(\mathcal{E}_{ijp} u_i v_j \mathbf{e}_p) \cdot (w_k \mathbf{e}_k) = \mathcal{E}_{ijk} u_i v_j w_k$, where $\mathcal{E}_{ijk}$ is the *alternating symbol*. For two vectors $\mathbf{u}$ and $\mathbf{v}$ moving between bases $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$, their components' product $u_i v_j$ transform according to the tensor product[1], $(\mathbf{u} \otimes \mathbf{v})_{ij} = u'_i v'_j = \mathbf{Q}_{ip}\mathbf{Q}_{jq} u_p v_q$. Thus, $I = \delta_{ij} \mathbf{e}_i \otimes \mathbf{e}_j := \mathbf{e}_i \otimes \mathbf{e}_j$ for an arbitrary orthonormal basis $\{\mathbf{e}_i\}$.

2) Tensor Algebra: We refer to the *mode-n unfolding* (or *matricization*) of a tensor, $\mathbb{T}$, as the rearrangement of its N elements into a matrix, $\mathbb{T}_{(n)} \in \mathbb{R}^{I_n \times \Pi_{k \neq n}^{n-1} I_k}$ where $n \in \{0, 1, \cdots, N-1\}$. The *multilinear rank* of $\mathbb{T} \in \mathbb{R}^{I_0 \times I_1 \cdots \times I_{N-1}}$ is an N-tuple with elements that correspond to the rank of the mode-n vector space i.e., $(R_0, R_1, \cdots, R_{N-1})$. The *Frobenius inner product* of two tensors $\mathbb{T}_1, \mathbb{T}_2 \in \mathbb{R}^{I_0 \times I_1 \times I_{n-1} \cdots \times I_{N-1}}$ is

$$\begin{aligned} \langle \mathbb{T}_1, \mathbb{T}_2 \rangle_F &= \text{trace}\left(\mathbb{T}^T_{2_{(n)}}, \mathbb{T}_{1_{(n)}}\right) \\ &= \text{trace}\left(\mathbb{T}^T_{1_{(n)}}, \mathbb{T}_{2_{(n)}}\right) \\ &= \langle \mathbb{T}_2, \mathbb{T}_1 \rangle_F. \end{aligned} \tag{1}$$

By the *norm of a tensor* with dimension N, we shall mean the square root of the sum of squares of all its elements. This is equivalent to the Frobenius norm along any n-mode unfolding, $\mathbb{T}_{(n)}$, of tensor $\mathbb{T}$. Thus,

$$\|\mathbb{T}\|_F^2 := \langle \mathbb{T}, \mathbb{T} \rangle_F = \|\mathbb{T}_{(n)}\|_F^2 \tag{2}$$

for any n-mode unfolding of the tensor. We may otherwise refer to $\|\cdot\|_F$ as the Hilbert-Schmidt norm.

Following the convention delineated in Table I, we define the product of tensor $\mathbb{T}$ (of size $I_0 \times I_1 \times I_{n-1} \cdots \times I_{N-1}$) and a matrix $\mathbf{U}$ (of size $J \times I_n$) as

$$\mathbb{P} = \mathbb{T} \otimes_n \mathbf{U} \implies \mathbb{P}_{(n)} = \mathbf{U}\mathbb{T}_{(n)}. \tag{3}$$

For different modes, the ordering of the modes is not consequential so that

$$\mathbb{T} \otimes_n \mathbf{U} \otimes_m \mathbf{V} = \mathbb{T} \otimes_m \mathbf{V} \otimes_n \mathbf{U} \quad \forall\, m \neq n. \tag{4}$$

However, in the same mode, order matters so that $\mathbb{T} \otimes_n \mathbf{U} \otimes_n \mathbf{V} = \mathbb{T} \otimes_n \mathbf{V}\,\mathbf{U}$. The *multilinear orthogonal projection* from a tensor space with dimension $I_0 \times \cdots I_{n-1} \times I_n \times I_{n+1} \cdots \times I_{N-1}$ onto the subspace $I_0 \times \cdots I_{n-1} \times U_n \times I_{n+1} \cdots \times I_{N-1}$ is the orthogonal projection along mode n given by

$$\pi_n \mathbb{T} := \mathbb{T} \otimes_n \left(\mathbf{I} - \mathbf{U}_n \mathbf{U}_n^T\right). \tag{5}$$

The rest of the notations we use for tensor operations in this article are described in Table I. We refer readers to [12], [13] for a detailed description of other tensor algebraic notations and multilinear operations.

B. Controls and Sets.

We define Ω as the open set in $\mathbb{R}^n$. To avoid the cumbersome phrase "the state $\boldsymbol{x}$ at time t", we will associate the pair $(\boldsymbol{x}, t)$ to the *phase* of the system for a state $\boldsymbol{x}$ at time t. Furthermore, we associate the Cartesian product of Ω and the space $T = \mathbb{R}^1$ of all time values as the *phase space*. The interior of Ω is denoted by $\text{int}\,\Omega$; whilst the closure of Ω is denoted $\bar{\Omega}$. We denote by $\delta\Omega\,(:= \bar{\Omega} \backslash \text{int}\,\Omega)$ the boundary of the set Ω. Unless otherwise stated, vectors $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t)$ are reserved for admissible control (resp. disturbance) at time t. We say $\boldsymbol{u}(t)$ (resp. $\boldsymbol{v}(t)$) is piecewise continuous in t, if for each t, $\boldsymbol{u} \in \mathcal{U}$ (resp. $\boldsymbol{v} \in \mathcal{V}$), $\mathcal{U}$, and V are Lebesgue measurable and compact sets.

III. Background and Preliminaries.

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

[1] Or the dyadic product.

TABLE I: **Common Notations**

Tensor Operations

Notation	Description
$\mathbb{T}_n$	n-mode unfolding of $\mathbb{T}$.
$\mathbf{G} = \mathbb{T}_n \mathbb{T}_n^T$	Gram matrix.
$[N] = \{0, 1, \cdots, N-1\}$	Total number of modes in $\mathbb{T}$.
$\|\mathbb{T}\|_F$	The Hilbert-Schmidt norm of $\mathbb{T}$.
$\mathbb{T} \otimes_n \mathbf{U}$	n-mode product of $\mathbb{T}$ with matrix $\mathbf{U}$.
$\mathbb{T} \hat{\otimes}_n \mathbf{v}$	n-mode product of $\mathbb{T}$ with vector $\mathbf{v}$.
$\mathbb{T} \circledast \mathbf{S}$	Kronecker product of $\mathbb{T}$ with matrix $\mathbf{S}$.
$\mathbb{T} \odot \mathbf{S}$	Khatri-Rao product of $\mathbb{T}$ with matrix $\mathbf{S}$.

Differential Optimal Control and Games

a.e.	Almost everywhere.
$\boldsymbol{\xi}$	System trajectory.
I, II	Players I and II respectively.
$\boldsymbol{V}(t, \boldsymbol{x})$	Value function of the differential game.
$\boldsymbol{V}_{\boldsymbol{x}}(t, \boldsymbol{x}), \boldsymbol{V}_t(t, \boldsymbol{x})$	Spatial derivative (resp. time derivative) of $\boldsymbol{V}$.
$\boldsymbol{V}^-(t, \boldsymbol{x}), \boldsymbol{V}^+(t, \boldsymbol{x})$	Lower and upper values of the differential game.
$\boldsymbol{H}^-(t; \cdot), \boldsymbol{H}^+(t; \cdot)$	A game's lower and upper Hamiltonians.
$\bar{\mathcal{U}}$, $\bar{\mathcal{V}}$	Controls set for I and II respectively.
$\mathcal{A}(t)$, $\mathcal{B}(t)$	Strategies set for I and II respectively, starting at t.
$\mathcal{F}(t, \boldsymbol{x}; \cdot)$	A separable Hilbert-space where $\boldsymbol{x}$ is defined.
$\mathcal{F}^\star(t, \boldsymbol{x}; \cdot)$	Dual of the separable Hilbert-space, $\mathcal{F}(\cdot)$.
$\mathcal{L}_0(\tau)$	A differential game's target set.
$\mathcal{L}([t, 0], \tau)$	A differential game's backward reachable set.

"min-max" or "max-min" type nonlinearity (the Isaacs' equation) was presented in [14]. 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 [2], [3], [11], [15], [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 [9]. 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 [10], [16]–[19].

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

$$\bar{\mathcal{U}} \equiv \{\boldsymbol{u} : [t, T] \to \mathcal{U} | \boldsymbol{u} \text{ measurable}\} \tag{6}$$
$$\bar{\mathcal{V}} \equiv \{\boldsymbol{v} : [t, T] \to \mathcal{V} | \boldsymbol{v} \text{ measurable}\}, \tag{7}$$

where $\boldsymbol{u} \subset \mathbb{R}^m$ and $\boldsymbol{v} \subset \mathbb{R}^p$ are given admissible controls. We are concerned with the differential equation,

$$\dot{\boldsymbol{x}}(\tau) = f(\tau, \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \boldsymbol{v}(\tau)) \quad t \leq \tau \leq T \tag{8a}$$
$$\boldsymbol{x}(t) = \boldsymbol{x}, \tag{8b}$$

where $f(\tau, \cdot, \cdot, \cdot)$ and $\boldsymbol{x}(\cdot)$ are 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]. Associated with (8) is the payoff functional

$$\begin{aligned} \boldsymbol{V}(\boldsymbol{u}, \boldsymbol{v}) &= \boldsymbol{V}_{t,\boldsymbol{x}}(\boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) \\ &= \int_t^T l(\tau, \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \boldsymbol{v}(\tau)) d\tau + g(\boldsymbol{x}(T)), \end{aligned} \tag{9}$$

with $g(\cdot)$ being bounded from above and Lipschitz continuous. In addition, the flow field $l : [0, T] \times \mathbb{R}^n \times \mathcal{U} \times \mathcal{V} \to \mathbb{R}$ is bounded, and uniformly continuous over $0 \leq t \leq T$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{u} \in \mathcal{U}$ and $\boldsymbol{u} \in \mathcal{V}$. Player I's goal is to maximize the payoff (9) and II's goal is to minimize it. Henceforward, we refer to $\boldsymbol{V}$ as the value.

B. Upper and Lower Values of the Differential Game.

Suppose that player II's mapping strategy (starting at t) is $\beta : \bar{\mathcal{V}}(t) \to \bar{\mathcal{U}}(t)$ 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. $t \leq \bar{t} \leq \tau$ implies $\beta[\boldsymbol{u}](\bar{t}) = \beta[\hat{\boldsymbol{u}}](\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$. And the differential game's lower value for a solution $x(t)$ that solves (8) for $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t) = \beta[u](\cdot)$ is

$$\begin{aligned} \boldsymbol{V}^-(\boldsymbol{x}, t) &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} V(\boldsymbol{u}, \beta[\boldsymbol{u}]) \\ &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} \int_t^T l(\tau, \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \beta[\boldsymbol{u}](\tau)) d\tau + g\left(\boldsymbol{x}(T)\right). \end{aligned} \tag{10}$$

Similarly, suppose that player I's mapping strategy (starting at t) is $\alpha : \bar{\mathcal{V}}(t) \to \bar{\mathcal{U}}(t)$ provided for each $t \leq \tau \leq T$ and $\boldsymbol{v}, \hat{\boldsymbol{v}} \in \bar{\mathcal{V}}(t)$; then $\boldsymbol{v}(\bar{t}) = \hat{\boldsymbol{v}}(\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$ implies $\alpha[\boldsymbol{v}](\bar{t}) = \alpha[\hat{\boldsymbol{v}}](\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$. And the differential game's upper value for a solution $\boldsymbol{x}(t)$ that solves (8) for $\boldsymbol{u}(t) = \alpha[\boldsymbol{v}](\cdot)$ and $\boldsymbol{v}(t)$ is

$$\begin{aligned} \boldsymbol{V}^+(\boldsymbol{x}, t) &= \sup_{\alpha \in \mathcal{A}(t)} \inf_{\boldsymbol{v} \in \mathcal{V}(t)} V(\alpha[\boldsymbol{v}], \boldsymbol{v}) \\ &= \sup_{\alpha \in \mathcal{A}(t)} \inf_{v \in \mathcal{V}(t)} \int_t^T l(\tau, \boldsymbol{x}(\tau), \alpha[\boldsymbol{v}](\tau), \boldsymbol{v}(\tau)) d\tau + g\left(\boldsymbol{x}(T)\right). \end{aligned} \tag{11}$$

These non-local PDEs i.e. (10) and (11) are hardly smooth throughout the state space so that they lack classical solutions even for smooth Hamiltonian and boundary conditions. However, these two values are "viscosity" (generalized) solutions [9], [15] of the associated HJ (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. In what follows, we introduce the notion of viscosity solutions to the HJI payoff/value functionals in (11), and (10).

C. Viscosity Solution of HJ-Isaac's Equations.

We now establish two Lemmas from [16] that will aid the construction of our main contribution in this paper.

Lemma 1. *The lower value* $\boldsymbol{V}^-$ *in* (10) *is the viscosity solution to the lower Isaac's equation*

$$\frac{\partial \boldsymbol{V}^-}{\partial t} + \boldsymbol{H}^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}^-_{\boldsymbol{x}}) = 0, \ t \in [0, T] \ \boldsymbol{x} \in \mathbb{R}^n \quad (12a)$$

$$\boldsymbol{V}^-(\boldsymbol{x}, T) = g(x(T)), \quad \boldsymbol{x} \in \mathbb{R}^m \quad (12b)$$

with lower Hamiltonian,

$$\boldsymbol{H}^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, p) = \max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), p \rangle. \quad (13)$$

where p*, the co-state, is the spatial derivative of* $\boldsymbol{V}^-$ *w.r.t* $\boldsymbol{x}$*.*

Lemma 2. *The upper value* $\boldsymbol{V}^+$ *in* (11) *is the viscosity solution of the upper Isaac's equation*

$$\frac{\partial \boldsymbol{V}^+}{\partial t} + \boldsymbol{H}^+(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}^+_{\boldsymbol{x}}) = 0, \ t \in [0, T], \ \boldsymbol{x} \in \mathbb{R}^n \quad (14a)$$

$$\boldsymbol{V}^+(\boldsymbol{x}, T) = g(\boldsymbol{x}(T)), \quad \boldsymbol{x} \in \mathbb{R}^n \quad (14b)$$

with upper Hamiltonian,

$$\boldsymbol{H}^+(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, p) = \min_{\boldsymbol{v} \in \mathcal{V}} \max_{\boldsymbol{u} \in \mathcal{U}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), p \rangle, \quad (15)$$

with p *being appropriately defined.*

Corollary 1. *(i)* $\boldsymbol{V}^- \leq \boldsymbol{V}^+$ *over* $(t \in [0, T] \ \boldsymbol{x} \in \mathbb{R}^n)$ *(ii) if for all* $t \in [0, T], (\boldsymbol{x}, p) \in \mathbb{R}^n$*, the minimax condition is satisfied i.e.* $\boldsymbol{H}^+(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, p) = \boldsymbol{H}^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, p)$*, then* $\boldsymbol{V}^- \equiv \boldsymbol{V}^+$*.*

D. Reachability for Systems Verification.

Reachability analysis is one of many verification methods that allows us to reason about online 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

- *forward* sense, whereupon system trajectories are examined to determine if they enter certain states from an *initial set*;
- *backward* sense, whereupon system trajectories are examined to determine if they enter certain *target sets*;
- *reach set* sense, in which they are examined to see if states reach a set at a *particular time*; or
- *reach tube* sense, in which they are evaluated that they reach a set at a point during a time interval.

Backward reachability consists in avoiding an unsafe set of states under the worst-possible disturbance at all times; relying on nonanticipative control strategies, [19]'s construction does not necessarily use a state feedback control law during games. However, this worst-possible disturbance assumption is not formally inculcated in the backward reachability analyses that are used. In a sense, it is reasonable to ignore nonlinear $\mathcal{H}^2$ or $\mathcal{H}^\infty$ analyses for Dubins vehicle [20] dynamics with constant inputs that only vary in sign for either player [21] since the worst possible is known by default. In other problem domains, this is not sufficient, and in our analyses we provide an $\mathcal{H}^\infty$ scheme [22]'s in constructing an appropriate *worst-possible* disturbance that *actually* guarantees robustness in continuous control applications.

BRTs are popularly analyzed on a game of two vehicles with non-stochastic dynamics [21]. Such BRTs possess discontinuity at cross-over points (which exist at edges) on the surface of the tube, 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 P.D.E does not have continuous partial derivatives (we discuss this further in section IV).

1) Insufficiency of Global Mesh-based Methods.: 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[2]. The curse of dimensionality [4] greatly incapacitates current uniform grid discretization methods for guaranteeing the robustness of backward reachable sets (BRS) and tubes (Tubes) [19] 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 [23] 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 [24], 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 [25] 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 [26] 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 [27], [28] in a reinforcement learning framework; again these methods lose the hard safety guarantees owing to the approximation in value function space.

In these sentiments, we seek to answer the following questions for high-dimensional systems:

- What role does sparsity play in the representation of BRTs and BRS for high-order systems?

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

- Can we provide rational decomposition schemes that preserve the numerical stability of monotone Lax-Friedrichs and essentially non-oscillatory [1] gradient methods to the HJ values and Hamiltonians?
- What is the relevance of local state space partitioning into subsystems [29] to problems with possibly high dimensions?
- With projection to reduced order systems, can we relax the strong assumptions made in local decomposition [29], [30] e.g. about the dynamics of the global system consisting of separable subsystems?

We briefly answer the first question. As long as value functions are implicitly defined as signed distance value functions on a grid, there is no possibility of exploiting sparsity for high-dimensional value functions. This is because these SDF-based value functions are constructed by choosing an interface on the grid. The value function is positive within the interface and negative outside the interface. Therefore, the representation of such values are completely dense. Unless we can find methods to sparsely represent the value function on a grid, exploiting sparsity of the value function is hopeless. In the sections that follow, we seek to answer the other questions posed above.

2) Reachability from Differential Games Optimal Control: For any admissible control-disturbance pair $(\boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot))$ and initial phase $(\boldsymbol{x}_0, t_0)$, Crandall and Evan's claim is that there exists a unique function

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t; t_0, \boldsymbol{x}_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) \tag{16}$$

that satisfies (8) a.e. with the property that

$$\boldsymbol{\xi}(t_0) = \boldsymbol{\xi}(t_0; t_0, \boldsymbol{x}_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) = \boldsymbol{x}_0. \tag{17}$$

Read (16): the motion of (8) passing through phase $(\boldsymbol{x}_0, t_0)$ under the action of control $\boldsymbol{u}$, and disturbance $\boldsymbol{v}$, and observed at a time t afterwards. One way to design a system verification problem is compute the reachable set of states that lie along the trajectory (16) such that we evade the unsafe sets up to a time e.g. t_f within a given time bound e.g. $[t_0, t_f]$. In this regard, we discard the *cost-to-go*, $l(t; \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \boldsymbol{v}(\tau))$ in (9), (10), or (11) and certify safety as resolving the terminal value, $g(\boldsymbol{x}(T))$.

Therefore, we can cast a target set as the time-resolved terminal value $\boldsymbol{V}^-(\boldsymbol{x}, T) = g(\boldsymbol{x}(T))$ so that given a time bound, and an unsafe set of states, the time-bounded safety verification problem consists in certifying that there is no phase within the target set (18) such that the solution to (8) enters the unsafe set. Following the backward reachability formulation of [19], we say the zero sublevel set of $g(\cdot)$ in (12) i.e.

$$\mathcal{L}_0 = \{\boldsymbol{x} \in \bar{\Omega} \,|\, g(\boldsymbol{x}) \leq 0\}, \tag{18}$$

is the *target set* in the product of space $\mathcal{S}^0 \times \mathbb{R}^+$ for a backward reachability problem (proof in [19]). Note that the target set, $\mathcal{L}_0$, is a closed subset of $\mathbb{R}^n$ and is in the closure of Ω. And the *robustly controlled backward reachable tube* for $\tau \in [T, 0]$[3] is the closure of the open set

$$\mathcal{L}([\tau, 0], \mathcal{L}_0) = \{\boldsymbol{x} \in \Omega \,|\, \exists\, \beta \in \bar{\mathcal{V}}(t)\, \forall\, \boldsymbol{u} \in \mathcal{U}(t), \exists\, \bar{t} \in [T, 0], \boldsymbol{\xi}(\bar{t}) \in \mathcal{L}_0\},\ \bar{t} \in [T, 0]\,. \tag{19}$$

Read: the set of states from which the strategies of II, and for all controls of I imply that we reach the target set within the interval $[T, 0]$. In backward reachability analysis, the lower value of the differential game i.e. (10) is used in constructing an analysis of the backward reachable set (or tube). More specifically, following Lemma 2 of [19], the states in the reachable set value function $\boldsymbol{V}$

$$x \in \mathcal{L}_0 \implies \boldsymbol{V}^-(\boldsymbol{x}, t) \leq 0 \tag{20a}$$
$$\boldsymbol{V}^-(\boldsymbol{x}, t) \leq 0 \implies \boldsymbol{x} \in \mathcal{L}_0. \tag{20b}$$

Observe:

- The pursuer or player I's goal is to drive the system into the unsafe set i.e., I aims to minimize the termination time of the game (c.f. (18));
- The evader, or player II, seeks to avoid the unsafe set i.e., II seeks to maximize the termination time of the game (c.f. (18));
- II has regular controls, $\boldsymbol{u}$, drawn from a Lebesgue measurable set, $\mathcal{U}$ (c.f. (10)).

This is a classic reachability problem on the resolution of the infimum-supremum over the strategies and controls of both players in the resolution of an extremum over a time interval of a cost function (nay, value functional) [17]. Let us recall that the open-loop policy pair $\{\boldsymbol{u}(t), \boldsymbol{v}(t)\}$ constitute a saddle-point solution to the differential game (8), (9), with final time $t_f = \inf\{t \in \mathbb{R}^+ : (\boldsymbol{x}(t), t) \in \mathcal{L}_0\}$ if

$$J(\boldsymbol{u}^\star(t), \boldsymbol{v}(t)) \leq J(\boldsymbol{u}^\star(t), \boldsymbol{v}^\star(t)) \leq J(\boldsymbol{u}(t), \boldsymbol{v}^\star(t)) \tag{21}$$

When capture[4] occurs, we must needs have the Hamiltonian of the value function be zero as a necessary condition for the players' saddle-point controls [14], [21] i.e. ,

$$\boldsymbol{H}_u(t; \boldsymbol{x}, \boldsymbol{u}^\star, \boldsymbol{v}, p) = 0,\ \boldsymbol{H}_v(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}^\star, p) = 0, \tag{22}$$

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

TO-DO: We essentially obtain a *pseudo iterative dynamic game* [31], albeit in open-loop settings, where either player infers the current state useful enough for generating closed-loop input control laws. An implicit surface function, $\{\boldsymbol{V}^-(\boldsymbol{x}, t) : [-T, 0] \times \mathcal{X} \to \mathbb{R},\ \forall\, t > 0\}$ i.e. the terminal value $\boldsymbol{V}^-(\boldsymbol{x}, t)$, that characterizes the target set $\mathcal{L}_0$ is the viscosity solution to the HJI PDE

$$\frac{\partial \boldsymbol{V}^-}{\partial t}(\boldsymbol{x}, t) + \min\{0, \boldsymbol{H}^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}^-_{\boldsymbol{x}})\} = 0 \tag{23a}$$
$$\boldsymbol{V}^-(\boldsymbol{x}, 0) = g(\boldsymbol{x}), \tag{23b}$$

where the vector field $\boldsymbol{V}^-_{\boldsymbol{x}}$ is known in terms of the game's terminal conditions so that the overall game is akin to a

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

[4] A capture occurs when II's separation from I becomes less than a specified e.g. capture radius.

two-point boundary-value problem. Henceforward, for ease of readability, we will remove the minus superscript on the lower value and Hamiltonian (13) so that we have

$$\frac{\partial \boldsymbol{V}}{\partial t} + \min\{0, \boldsymbol{H}(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}_{\boldsymbol{x}})\} = 0, \ t \in [T, 0] \, , \ \boldsymbol{x} \in \mathbb{R}^n \tag{24a}$$

$$\boldsymbol{V}(\boldsymbol{x}, 0) = g(x), \quad \boldsymbol{x} \in \mathbb{R}^n \tag{24b}$$

with lower Hamiltonian,

$$\boldsymbol{H}(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, p) = \max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), p \rangle. \tag{25}$$

IV. Decomposition Scheme.

In what follows, we introduce a multilinear decomposition scheme aimed at rational decomposition of large backward reachable tubes and to alleviate the exponential scaling 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.

A. ROM Construction

We consider separated representations of the value function $\boldsymbol{V}(\boldsymbol{x}, t)$ (c.f. (12)), parameterized by basis functions $\boldsymbol{\phi} \in \boldsymbol{\Phi}$[5] i.e. $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi})$, where, $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi})$, defined over a region $(\omega \times \mathbb{R})$ $(\omega \subset \Omega)$, possesses real values on the Hilbert space $\mathcal{F} = H_0^1(\Omega) \in \mathbb{R}^{d_\phi}$ that satisfy the boundary conditions of (12b).

Let us first introduce the notations that will enable us to construct the decomposition problem. We let $\langle \cdot ; \cdot \rangle_{\mathcal{F}}$ be the inner product associated with a norm, $\| \cdot \|_{\mathcal{F}}$, induced on the tensor product space $\mathcal{F} \otimes \mathcal{S}$, which we define as

$$\langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_{\mathcal{F}} = \int_{\mathcal{S}} \langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_{\mathcal{F}} d\boldsymbol{x}. \tag{26}$$

We define the function space $L^2(\mathcal{S}; \mathcal{F})$ as

$$L^2(\mathcal{S}; \mathcal{F}) = \left\{ \boldsymbol{\phi} : \mathcal{S} \to \mathcal{F}; \int_{\mathcal{S}} \|\boldsymbol{\phi}(\boldsymbol{x})\|_{\mathcal{F}}^2 \, d\boldsymbol{x} < +\infty \right\} \tag{27}$$

and associate with it the dual space $\breve{\mathcal{F}} = H^{-1}(\Omega)$[6]. **The decomposition problem is to find $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi}) \in \mathcal{F} \otimes \mathcal{S}$ such that**

$$\frac{\partial \boldsymbol{V}}{\partial t}(\boldsymbol{x}, t) \in L^2(\breve{\mathcal{F}} \otimes \mathcal{S}) \tag{28}$$

and

$$\amalg(\boldsymbol{V}, \boldsymbol{\phi}) = \uplus(\boldsymbol{\phi}), \ \forall \, \boldsymbol{\phi} \in \mathcal{F} \otimes \mathcal{S}. \tag{29}$$

where $\amalg(\cdot, \cdot)$ and $\uplus(\cdot)$ are respectively multilinear transforms (to be introduced shortly). The solution of (29) satisfies the boundary conditions to the initial value problem of c.f. (23) in a weak sense [32].

[5]Note that $\boldsymbol{\Phi} \subset \mathcal{F} \in \mathbb{R}^{d_\phi}$.

[6]Following (1), when we write a norm in the form of (26), we shall adopt the Frobenius inner product norm throughout.

In this sentiment, $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi})$ is the sum of the tensor products of free parameters $\{\boldsymbol{\psi}_i\}_{i=0}^{\infty} \in \mathcal{S}$ and basis functions $\{\boldsymbol{\phi}_i\}_{i=0}^{\infty} \in \mathcal{F}$ i.e. ,

$$\boldsymbol{V}(\boldsymbol{x}, t) \approx \boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi}) \tag{30a}$$

$$\equiv \sum_{i=0}^{\infty} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_s, \ \boldsymbol{\phi}_i \in \mathcal{F}, \ \boldsymbol{\psi}_i \in \mathcal{S} \tag{30b}$$

where $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi}) : \mathbb{R}^n \times \mathbb{R} \to \mathcal{F} \otimes \mathcal{S}$; $\{\boldsymbol{\phi}_i\}_{i=0}^{\infty}$ are basis functions in $\mathcal{F}$, and $\{\{\boldsymbol{\psi}_i\}_{i=0}^{\infty} \subset \mathcal{S}\}$ represent the influence of the *principal components* of $\boldsymbol{V}$ on the respective bases, $\boldsymbol{\phi}_i$. *These summands over tensor products constitute the Galerkin decomposition of the viscosity solution $\boldsymbol{V}(\boldsymbol{x}, t)$ mentioned in § III-C.*

A separable representation of $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi})$ of order r therefore can be defined as the function

$$\boldsymbol{V}_r(\boldsymbol{x}, t) = \sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_{\mathcal{F}}, \ \boldsymbol{\phi}_i \in \mathcal{F}, \ \boldsymbol{\psi}_i \in \mathcal{S} \tag{31}$$

that admits an optimal solution $\boldsymbol{V} \in \mathcal{F} \otimes \mathcal{S}$ up to an approximation error $\epsilon > 0$, established via the Galerkin orthogonality criteria

$$\|\boldsymbol{V} - \boldsymbol{V}_r\|_F^2 = \min_{\substack{\{\boldsymbol{\psi}_i\}_{i=1}^{\infty} \in \mathcal{S} \\ \{\boldsymbol{\phi}_i\}_{i=1}^{\infty} \in \mathcal{F}}} \|\boldsymbol{V} - \sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_{\mathcal{F}}\|_F^2. \tag{32}$$

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 to higher-order statistics, chemometrics, psychometrics, and signal processing. Evidence abounds that linearized dynamics of nonlinear systems, which are truncated at an n-th order (e.g. in power series linearized expansions) admit higher precision and accuracy of the approximation of the underlying system; this is because the moments and accumulations of higher-order dynamics are equivalent to the power series expansion coefficients.

Let $\mathbb{V} \in \mathbb{R}^{I_0 \times I_1 \times \cdots \times I_{N-1}}$ be a high-dimensional value function for a system of multiple interacting agents, each with dynamics $\dot{\boldsymbol{x}}_0, \dot{\boldsymbol{x}}_1, \cdots \dot{\boldsymbol{x}}_n$ that span the full rank of $\mathbb{V}$. Since the value functions for such an interactive CPS system are of high order, we replace the value function $\boldsymbol{V}(\boldsymbol{x}, t)$ with its tensor representation, $\mathbb{V}(\boldsymbol{x}, t)$ (c.f. § II) so that the full problem described in Lemma 1, corresponds to the following parametric P.D.E

$$\frac{\partial \mathbb{V}}{\partial t}(\boldsymbol{x}, t; \boldsymbol{\phi}) + \mathbb{H}(\boldsymbol{x}, t, p; \boldsymbol{\phi}) = 0, \ t \in [T, 0] \, , x \in \Omega, \boldsymbol{\phi} \in \boldsymbol{\Phi} \tag{33a}$$

$$\mathbb{V}(\boldsymbol{x}, T; \boldsymbol{\phi}) = g(\boldsymbol{x}; \boldsymbol{\phi}) \quad \boldsymbol{x} \in \{\bar{\Omega} \backslash \text{int } \Omega\}, \ \boldsymbol{\phi} \in \boldsymbol{\Phi} \tag{33b}$$

with corresponding Hamiltonian,

$$\mathbb{H}(\boldsymbol{x}, t, p; \boldsymbol{\phi}) = \min \left\{ 0, \max_{\boldsymbol{u} \in \mathcal{U}} \min_{\boldsymbol{v} \in \mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), p \rangle \right\}. \tag{34}$$

The decomposition of (31) can be considered a pseudo-eigenvalue problem (this is established in § IV-E) which proves

efficient for separated representations in many applications including stochastic nonlinear PDEs [33], [34] and finite element methods [35]. The state space can be split into disjoint regions where the value function is continuously differentiable in each region. TO-DO: The *singular surfaces* [14] that separate the respective disjoint value functions constitute manifolds which have discontinuous derivative properties and we follow [36, Theorem 8.2]'s manifold resolution strategy i.e. these manifolds satisfy the saddle equilibrium strategies

$$\mathbb{H}(t;\boldsymbol{x},\boldsymbol{u}^\star,\boldsymbol{v}) \leq \mathbb{H}(t;\boldsymbol{x},\boldsymbol{u}^\star,\boldsymbol{v}^\star) \leq \mathbb{H}(t;\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}^\star). \tag{35}$$

We defer the treatment of *dispersal surfaces* [14] to a future work. TO-DO: When the value function is not continuously differentiable; or the value function becomes discontinuous, we resort to classical fractional steps in finite differencing schemes for conservation laws [37] applied on a dimension-by-dimension basis to a mesh on which a separated composition is defined [1].

B. Galerkin approximation of the Variational HJI Problem

We now derive the Galerkin approximation of the viscosity solution to the terminal HJI problem. Assume that a decomposition V_r of order r is already known (this could be obtained by a partial truncation of the value function as described in § IV-E) or randomly initialized. For the next order r, a new couple $(\boldsymbol{\psi},\boldsymbol{\phi})$ is optimal if it satisfies the Galerkin orthogonal metric on the induced norm of the tensor product space $\mathcal{S}\otimes\mathcal{F}$ such that

$$\begin{aligned}\|\mathbb{V}-\mathbb{V}_r\|_F^2 &= \|\mathbb{V}\|_F^2 - \sigma(\phi_i(\boldsymbol{x})) \\ &\equiv \|\mathbb{V}\|_F^2 - \|\breve{\mathbb{V}}\|_F^2,\end{aligned} \tag{36}$$

where $\sigma(\boldsymbol{\phi}_i(\boldsymbol{x}))$ denotes an eigen decomposition of $\boldsymbol{\phi}_i(\boldsymbol{x})$ and $\breve{\mathbb{V}}$ is the *core tensor* of $\boldsymbol{V}$, obtained from a higher order singular value decomposition of $\boldsymbol{V}$ (We discuss this in § IV-E). The optimality proof of (36) is given in appendix A via (50).

Suppose that the optimal or quasi-optimal[7] order r for decomposing $\mathbb{V}$ in (33b) has been found. The quasi-optimal Galerkin-approximation of (33b) is given as

$$\frac{\partial \mathbb{V}}{\partial t}(\boldsymbol{x},t;\boldsymbol{\phi}) + \mathbb{H}(\boldsymbol{x},t,p;\boldsymbol{\phi}) = 0 \;\; x\in\Omega, \boldsymbol{\phi}\in\mathcal{F} \tag{37}$$

or

$$-\frac{\partial}{\partial t}\left(\sum_{i=0}^{r-1}\langle\boldsymbol{\psi}_i,\boldsymbol{\phi}_i\rangle_{\mathcal{F}}\right) = \min\left\{0, \max_{\boldsymbol{u}\in\mathcal{U}}\min_{\boldsymbol{v}\in\mathcal{V}}\langle f(t;x,\boldsymbol{u},\boldsymbol{v}), \frac{\partial}{\partial\boldsymbol{x}}\left(\sum_{i=0}^{r-1}\langle\psi_i,\phi_i\rangle_{\mathcal{F}}\right)\right\rangle\right\} \tag{38}$$

$$-\sum_{i=0}^{r-1}\langle\boldsymbol{\psi}_i, \frac{\partial\boldsymbol{\phi}_i}{\partial t}\rangle_{\mathcal{F}} = \min\left\{0, \max_{\boldsymbol{u}\in\mathcal{U}}\min_{\boldsymbol{v}\in\mathcal{V}}\langle f(t;\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}), \sum_{i=0}^{r-1}\left\langle\boldsymbol{\psi}_i\frac{\partial\boldsymbol{\phi}_i}{\partial\boldsymbol{x}}\right\rangle_{\mathcal{F}}\right\rangle\right\}. \tag{39}$$

[7]The rationale for introducing quasi-optimality is discussed in section IV-E where the Galerkin basis functions and coefficients are resolved with sequentially-truncated higher order singular value decomposition of $\mathbb{V}(\boldsymbol{x},t)$.

Let us vectorize the coefficients $\{\boldsymbol{\psi}\}_{i=0}^{r-1}\subset\mathcal{S}$ and basis functions in the separable Hilbert space $\{\boldsymbol{\phi}\}_{i=0}^{r-1}\subset\mathcal{F}$ as follows

$$\boldsymbol{\Psi}_r^T = [\boldsymbol{\psi}_0,\boldsymbol{\psi}_1,\cdots,\boldsymbol{\psi}_{r-1}],\;\; \boldsymbol{\Phi}_r = [\boldsymbol{\phi}_0,\boldsymbol{\phi}_1,\cdots,\boldsymbol{\phi}_{r-1}]^T. \tag{40}$$

Thus, we can rewrite (39) as

$$-\boldsymbol{\Psi}_r^T\dot{\boldsymbol{\Phi}}_r = \min\left\{0, \max_{\boldsymbol{u}\in\mathcal{U}}\min_{\boldsymbol{v}\in\mathcal{V}}\left\langle f(t;\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}), \langle\boldsymbol{\Psi}_r^T,\boldsymbol{\Phi}_{\boldsymbol{x}}\rangle_{\mathcal{F}}\right\rangle\right\}. \tag{41}$$

Note that $\boldsymbol{\Phi}_{\boldsymbol{x}}$ are the spatial derivatives of $\{\boldsymbol{\phi}\}_{i=0}^{r-1}$ w.r.t $\boldsymbol{x}$. Furthermore, we choose not to factorize the coefficients $\boldsymbol{\Psi}_r^T$ in (41) because we want to retain the characteristics of the original value function $\mathbb{V}(\boldsymbol{x},t)$ on the respective bases, $\{\langle\boldsymbol{\psi}_i,\boldsymbol{\phi}_i\rangle\}_{i=0}^{r}$. At once, we see that if the optimal decomposition components $\{\boldsymbol{\psi}_i\}_{i=1}^{r}\in\boldsymbol{\Psi}$ and $\{\boldsymbol{\phi}_i\}_{i=0}^{r-1}\in\boldsymbol{\Phi}$ are known, equation (41) admits separable solutions on finite meshes, rendering solution of the viscosity problem (23) easily computable with the usual high precision and accuracy that Lax-Friedrichs schemes afford [1], [2], [11], [37]. We defer the identification of the parameters Φ, Ψ to § IV-E.

C. Galerkin HJI Approximation Under Separable Dynamics

Now, suppose that the dynamics $f(t;\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})$ from (8) is separable into its state, control, and disturbance components in an additive manner as follows,

$$\dot{\boldsymbol{x}} = f(t;\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}) \tag{42}$$

$$= f_{\boldsymbol{x}}\boldsymbol{x}(t) + f_{\boldsymbol{u}}(t)\boldsymbol{u}(t), f_{\boldsymbol{v}}\boldsymbol{v}(t), \tag{43}$$

where f_x, f_u, and f_v are the respective components of the system dynamics for the state, control law, and disturbance. This separable dynamics is typically observed for autonomous systems such as Dubins vehicles in relative coordinates, quadcopters, and many natural systems[8].

Putting (43) into (41), we find that

$$-\boldsymbol{\Psi}_r^T\dot{\boldsymbol{\Phi}}_r = \min\left\{0, \max_{\boldsymbol{u}\in\mathcal{U}}\min_{\boldsymbol{v}\in\mathcal{V}}\langle f_x x(t) + f_u(t)\boldsymbol{u}(t), f_v\boldsymbol{v}(t), \left\langle\boldsymbol{\Psi}_r^T,\boldsymbol{\Phi}_{\boldsymbol{x}}\right\rangle_{\mathcal{F}}\right\rangle\right\}, \tag{44}$$

so that the term on the right hand side becomes

$$\min\left\{0, \left\langle f_x\boldsymbol{x}(t), \left\langle\boldsymbol{\Psi}_r^T,\boldsymbol{\Phi}_{\boldsymbol{x}}\right\rangle_{\mathcal{F}}\right\rangle + \max_{\boldsymbol{u}\in\mathcal{U}}\langle f_u(t)\boldsymbol{u}(t), \left\langle\boldsymbol{\Psi}_r^T,\boldsymbol{\Phi}_{\boldsymbol{x}}\right\rangle_{\mathcal{F}}\right\rangle + \min_{\boldsymbol{v}\in\mathcal{V}}\left\langle f_v\boldsymbol{v}(t), \left\langle\boldsymbol{\Psi}_r^T,\boldsymbol{\Phi}_{\boldsymbol{x}}\right\rangle_{\mathcal{F}}\right\rangle\right\}. \tag{45}$$

Similar to (41), (45) can be resolved on a mesh. However, when the dynamics are separable as in the foregoing, the saddle point necessary condition i.e. (21) allows us to find an analytic solution. TO-DO: We defer this treatment to a future work as follows: TO-DO: Under development. A future paper?

[8]Even when the separation in (43) is not possible globally, we can consider a perturbation $\delta\boldsymbol{x}$ about the state $\boldsymbol{x}$ along a nominal trajectory $\bar{\boldsymbol{x}}$ so that the system's locally linear state is iteratively measured with respect to $\boldsymbol{x}$ as it is commonly done in linear quadratic methods. We defer the treatment of these locally linearized dynamics to a future work. For an in-depth treatment, see [5], [7], [38], [39].

D. PGD Decomposition Scheme

It now remains for us to establish an optimal way to compute the basis and coefficients of the optimal Galerkin decomposition in (41) and (45). TO-DO: First, we introduce the following multilinear mappings:

- Let $S_r : \mathcal{F} \to \mathcal{S}$ be the multilinear transformation from a real-valued state in $\phi \in \mathcal{F}$ to $\psi \in \mathcal{S}$ i.e. ,

$$\amalg(\boldsymbol{V}_{r-1} + \phi\psi, \phi\psi_1^\star + \phi_1^\star\psi, \psi_2^\star \cdots \psi_r^\star) = \uplus(\phi^\star\psi), \forall\, \phi^\star \in \mathcal{F}. \tag{46}$$

- Let $F_r : \mathcal{S} \to \mathcal{F}$ be the multilinear transformation from $\psi \in \mathcal{S}$ to $\boldsymbol{\phi} \in \mathcal{F}$ i.e. ,

$$\amalg(\boldsymbol{V}_{r-1} + \phi\psi_0, \phi\psi_1, \cdots \phi\,\psi_r) = \uplus(\phi^\star\psi), \forall\, \phi^\star \in \mathcal{F}. \tag{47}$$

TO-DO: we establish a Lemma due to [40] that allows every tensor $\mathbb{V}$ to admit a higher-order singular value decomposition. In our treatment, we resort to higher-order singular value decomposition (HOSVD) [41][9], extended to N-way tensors by [42]. This consists in decomposing it into the product of a core tensor, $\breve{\mathbb{V}} \in \mathbb{R}^{R_0 \times R_1 \times \cdots \times R_n}$, $(n <= N-1)$, and the orthonormal matricized tensor modes, $\boldsymbol{U}_{n=0}^{N-1}$, of the components of $\mathbb{V}$ (where $\mathbb{V} \in \mathbb{R}^{I_0 \times I_1 \times \cdots \times I_n}$; and $R_i \leq I_i$) i.e.,

$$\mathbb{V} \approx \hat{\mathbb{V}} \tag{48a}$$

$$= \breve{\mathbb{V}} \otimes_0 \mathbf{U}_0 \otimes_1 \mathbf{U}_1 \otimes_2 \mathbf{U}_2 \cdots \otimes_{N-2} \mathbf{U}_{N-1} \tag{48b}$$

where $\mathbf{U}_0 \in \mathbb{R}^{I_0 \times R_0}$, $\mathbf{U}_1 \in \mathbb{R}^{I_1 \times R_1}, \cdots \mathbf{U}_{N-1} \in \mathbb{R}^{I_{N-1} \times R_{N-1}}$ are the factor matrices (ideally orthonormal in each mode) that represent the influence of the principal components of each mode of $\mathbb{V}$ on the *core tensor* $\breve{\mathbb{V}}$. The decomposition in (48) can be obtained via the following minimization problem

$$\min_{\breve{\mathbb{V}}, \mathbf{U}_0, \cdots, \mathbf{U}_{N-1}} \|\mathbb{V} - \breve{\mathbb{V}} \otimes_0 \mathbf{U}_0 \cdots \otimes_{N-2} \mathbf{U}_{N-1}\| \tag{49}$$

The entries of $\breve{\mathbb{V}}$ denotes level of interaction between the different components $\boldsymbol{U}_n$. If the orthonormal components of (48) are known, then the *optimal core* of $\breve{\mathbb{V}}$ in terms of $\mathbb{V}$ is [12, §4.2]

$$\breve{\mathbb{V}} = \mathbb{V} \otimes_0 \mathbf{U}_0^T \otimes_1 \mathbf{U}_1^T \otimes_2 \mathbf{U}_2^T \cdots \otimes_{N-2} \mathbf{U}_{N-1}^T. \tag{50}$$

Putting (50) into (48), the approximation of $\mathbb{V}$ is essentially a projection of the original value function tensor along its respective n-basis modes on the reduced system, i.e.

$$\hat{\mathbb{V}} = \mathbb{V} \otimes_0 \mathbf{U}_0\mathbf{U}_0^T \cdots \otimes_{N-2} \mathbf{U}_{N-1}\mathbf{U}_{N-1}^T. \tag{51}$$

Therefore, the approximation error, $\|\mathbb{V} - \hat{\mathbb{V}}\|^2 = \|\mathbb{V}\|^2 - \|\breve{\mathbb{V}}\|^2$ (as shown in Appendix A). It now suffices to define the projection approximation error of the value function. A truncated decomposition up to mode-r would consist of a *partial core* $\breve{\mathbb{V}}_r$ and its corresponding orthonormal matrices $\boldsymbol{U}_0, \cdots, \boldsymbol{U}_r$, $r \in [R]$ defined as

$$\breve{\mathbb{V}}_r = \mathbb{V} \otimes_0 \mathbf{U}_0^T \otimes_1 \mathbf{U}_1^T \otimes_2 \mathbf{U}_2^T \cdots \otimes_{r-1} \mathbf{U}_r^T, \tag{52}$$

[9]In his original work, Tucker only prescribed the decomposition of a tensor for up to 3 modes.

Algorithm 1 Iterative PGD Scheme for Computing BRS/BRTs

1: **function** IterativeBRT($\boldsymbol{\psi}_0, \boldsymbol{\phi}_0, r, r_{max}$) ▷ r_{max}: max. iterations.
2: Set $\mathbb{V}_{r-1} = \langle \psi_0, \phi_0 \rangle$
3: **while** $r < r_{max}$ **do**
4: Compute $\boldsymbol{\psi}_r$ from (54).
5: Compute $\boldsymbol{\phi}_r$ from (55).
6: Set $\mathbb{V}_r \leftarrow \langle \mathbb{V}_{r-1} + \psi_r\phi_r \rangle_{\mathcal{F}}$ ▷ Update $\mathbb{V}_r$.
7: Set $\langle \boldsymbol{\Psi}_r^T, \dot{\boldsymbol{\Phi}}_r \rangle_F \leftarrow \mathbb{V}_r$ c.f. (41) or (45)
8: $\mathbb{L}_r \leftarrow \int_{\mathcal{S}} \langle \boldsymbol{\Psi}_r^T, \dot{\boldsymbol{\Phi}}_r \rangle_{\mathcal{F}} d\boldsymbol{x}$. ▷ Compute Partial BRT, $\mathbb{L}_r$
9: $r = r + 1$. ▷ Advance the decomposition.
10: **end while**
11: **end function**

so that the truncated value function (up to mode r) is

$$\mathbb{V}_r = \breve{\mathbb{V}}_r \otimes_0 \mathbf{U}_0 \otimes_1 \mathbf{U}_1 \otimes_2 \mathbf{U}_2 \cdots \otimes_{r-1} \mathbf{U}_r \tag{53a}$$

$$\equiv \mathbb{V} \otimes_0 \mathbf{U}_0\mathbf{U}_0^T \otimes_1 \cdots \otimes_{r-1} \mathbf{U}_r\mathbf{U}_r^T,\, r \in [R]\,. \tag{53b}$$

where (53b) is a result of putting (52) into (53a).

Now, revisiting the functions $\{\boldsymbol{\psi}\}_{i=0}^{r-1}$ and $\{\boldsymbol{\phi}\}_{i=0}^{r-1}$ in the previous two sections, a convenient way to compute the coefficients and basis functions that satisfy the Galerkin orthogonality criterium (36) is to set

$$\boldsymbol{\psi}_r = \mathbb{V} \otimes_0 \mathbf{U}_0^T \cdots \otimes_{r-1} \mathbf{U}_r^T \tag{54}$$

and

$$\boldsymbol{\phi}_r = \breve{\mathbb{V}}_r \otimes_0 \mathbf{U}_0 \cdots \otimes_{r-1} \mathbf{U}_r. \tag{55}$$

Algorithm 1 describes how we compose the separable value function that satisfies (29). Lines 4 and 5 describe the optimal resolution of the decomposition parameters, $\{\boldsymbol{\psi}\}_{i=0}^{r-1}$, and basis functions, $\{\boldsymbol{\phi}\}_{i=0}^{r-1}$. In line 7 of the algorithm, the integral is solved using Total Variation Diminishing (TVD) Runge-Kutta scheme as described in [43, §3.5] (originally implemented in [44], which we re-implement in CuPy [45] as we leverage parallel computation). In addition, having r as an input variable into the algorithm allows us to take advantage of warm-starting schemes, typical in Reinforcement learning schemes [28] so that the algorithm need not be run in one fell-swoop. Partial BRT's and BRS's can be distributively learned on separate CPU/GPU cores, and later assembled on a centralized nodes to aid faster computation. This may for example be similar to the alternating direction method of multipliers.

In algorithm 1, the maximum number of iterations, r_{max} can be chosen in an informed was as highlighted below. First, we restate the following theorem that allows us to provide a bound on the projection error.

Theorem 1. *[Vannieuwenhoven, Vandebril, and Meerbergen] [13, Th. 5.1]. Suppose $\mathbb{V}$ is a tensor of size $I_0 \times I_1 \times \cdots \times I_{N-1}$, approximated by $\hat{\mathbb{V}}$ as in (51), the approximation error of (36) is*

$$\begin{aligned}\|\mathbb{V} - \hat{\mathbb{V}}\|_F^2 = \|\mathbb{V} \otimes_0 \left(\mathbf{I} - \mathbf{U}_0\mathbf{U}_0^T\right)\|_F^2 \\ + \|\tilde{\mathbb{V}}_0 \otimes_1 \left(\mathbf{I} - \mathbf{U}_1\mathbf{U}_1^T\right)\|_F^2 + \cdots \\ + \|\tilde{\mathbb{V}}_{N-2} \otimes_{N-1} \left(\mathbf{I} - \mathbf{U}_{N-1}\mathbf{U}_{N-1}^T\right)\|_F^2.\end{aligned} \tag{56}$$

Furthermore, the approximation error is bounded by

$$\|\mathbb{V} - \hat{\mathbb{V}}\|_F^2 \leq \sum_{n=0}^{N-1} \|\mathbb{V} \otimes_n \left(\mathbf{I} - \mathbf{U}_n \mathbf{U}_n^T\right)\|_F^2. \tag{57}$$

Equation (57) allows us to choose an approximation error that informs us about the level of information we want preserved on the decomposed value $\mathbb{V}$. Therefore, to find a Tucker decomposition $\hat{\mathbb{V}}$ of $\mathbb{V}$ whose relative decomposition error is no greater than a certain positive ϵ, we first unfold the tensor along one of its modes, compute the Gram matrix and then do an eigen decomposition:

$$\mathbf{G} \equiv \mathbb{V}_{(n)} \mathbb{V}_{(n)}^T = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T, \tag{58}$$

where $\mathbf{\Lambda} = \text{diag}\left(\{\lambda_1, \lambda_2, \cdots, \lambda_{I_n}\}\right)$, and $\lambda_1 \geq \lambda_2 \geq \cdots > \lambda_{I_n} \geq 0$, and $\mathbf{V}$ contains the corresponding eigenvectors. Therefore, we can choose the orthonormal matrix $\mathbf{U}_n$ and the low-rank tensor R_{n-1} at mode $n-1$ as

$$\mathbf{U}_n = \mathbf{V}\left[:, 0 : R_{n-1}\right] \tag{59}$$

where

$$\begin{aligned} R_n = &\min_{R \in [I_n]} R \\ \text{subject to } &\sum_{i=R+1}^{I_n} \lambda_i \leq \epsilon^2 \|\mathbb{V}\|^2 / N. \end{aligned} \tag{60}$$

We can iteratively choose R_n from $n = 1$ up to an $n = r$ that guarantees we satisfy the error bound in (57) so that

$$\|\mathbb{V} \otimes_n \left(\mathbf{I} - \mathbf{U}_n \mathbf{U}_n^T\right)\|_F^2 \leq \delta^2 \frac{\|\mathbb{V}\|_F^2}{N}. \tag{61}$$

Whilst appealing, algorithm 1 utilizes the full value function at each step (c.f. (54)); therefore it does not easily lend itself to large-scale problems. In what follows, we propose an incrementally-constructed value function whereupon we work with the *partial cores* of $\boldsymbol{V}$ at each step of the PGD iteration. In essence, we construct the next factor matrix based on the previously computed value core.

E. Informed Incremental Value Function Decompositions

Next, we will leverage the sequentially-truncated high-order SVD of [13] to device an iteratively refined decomposition scheme onto which we will project the high-order value function. Working with partial cores, at step n, we generate the next factor matrix based on $\hat{\mathbb{V}}_{n-1}$. In particular, if the conditions of Theorem 1 hold, then

$$\begin{aligned} \|\mathbb{V} - \hat{\mathbb{V}}\|_F^2 = &\|\mathbb{V} \otimes_0 \left(\mathbf{I} - \mathbf{U}_0 \mathbf{U}_0^T\right)\|_F^2 + \|\tilde{\mathbb{G}}_0 \otimes_1 \\ &\left(\mathbf{I} - \mathbf{U}_1 \mathbf{U}_1^T\right)\|_F^2 \cdots + \|\tilde{\mathbb{G}}_{N-2} \otimes_{N-1} \\ &\left(\mathbf{I} - \mathbf{U}_{N-1} \mathbf{U}_{N-1}^T\right)\|_F^2. \end{aligned} \tag{62}$$

For convenience, we reproduce algorithm I of [13] in Algorithm 2.

Algorithm 2 Incrementally Truncated Higher-Order SVD.

function TruncatedValue($\mathbb{V}$, ϵ) ▷ ϵ: Desired accuracy.
  $\mathbb{P} \leftarrow \mathbb{V}$ ▷ $\mathbb{P}$: c.f. (3).
  for $n = 0, 1, \cdots, N-1$ **do**
    $\mathbf{G} \leftarrow \mathbb{P}_{(n)} \mathbb{P}_{(n)}^T$ ▷ $\mathbf{G}$: Gram matrix.
    $(\lambda, \mathbf{W}) \leftarrow \text{eig}\ (\mathbf{G})$ ▷ eig: Eigen Decomposition.
    $R_n \leftarrow \min R \in [I_n] \,|\, \sum_{i=R+1}^{I_n} \lambda_i \leq \epsilon^2 \|\mathbb{V}\|^2 / N$
    $\mathbf{U}_n \leftarrow \mathbf{W}\left[:, 0 : R_{n-1}\right]$
    $\mathbb{P} \leftarrow \mathbf{G} \otimes_n \mathbf{U}_n^T$
  end for
  $\mathbb{G} \leftarrow \mathbb{P}$
  $\mathbf{U} \leftarrow \{\mathbf{U}_0, \cdots, U_{N-1}\}$
  return $(\mathbb{G}, \mathbf{U})$ ▷ $\mathbb{G}$: Core tensor; $\{\mathbf{U}\}_{i=0}^{N-1}$: Orthonomal matrices.
end function

V. Results and Discussion.

VI. Conclusion.

Appendix

Appendix A
Value Function's ROM Projection Error.

The projection error between the original value function $\mathbb{V}$ and its reduced basis $\hat{\mathbb{V}}$ is

$$\begin{aligned} \|\mathbb{V} - \hat{\mathbb{V}}\|_F^2 &= \|\mathbb{V} - \mathbb{G} \otimes_0 \mathbf{U}_0 \cdots \otimes_{N-1} \mathbf{U}_{N-1}\|_F^2 \\ &= \|\mathbb{V}\|_F^2 - 2\langle \mathbb{V}, \mathbb{G} \otimes_0 \mathbf{U}_0 \cdots \otimes_{N-1} \mathbf{U}_{N-1} \rangle + \\ &\quad \|\mathbb{G} \otimes_0 \mathbf{U}_0 \cdots \otimes_{N-1} \mathbf{U}_{N-1}\|_F^2 \\ &= \|\mathbb{V}\|_F^2 - 2 \underbrace{\langle \mathbb{V} \otimes_0 \mathbf{U}_0^T \cdots \otimes_{N-1} \mathbf{U}_{N-1}^T, \mathbb{G} \rangle}_{\langle \mathbb{G}, \mathbb{G} \rangle} + \|\mathbb{G}\|_F^2 \end{aligned} \tag{63}$$

$$\text{or } \|\mathbb{V} - \hat{\mathbb{V}}\|_F^2 = \|\mathbb{V}\|_F^2 - \|\mathbb{G}\|_F^2. \tag{64}$$

References

[1] S. Osher and C.-W. Shu, "High-Order Essentially Nonoscillatory Schemes for Rotary-Wing Wake Computations," *Journal of Aircraft*, vol. 41, no. 2, pp. 258–267, 2004. 1, 5, 7

[2] M. G. Crandall and A. Majda, "Monotone difference approximations for scalar conservation laws," *Mathematics of Computation*, vol. 34, no. 149, pp. 1–21, 1980. 1, 3, 7

[3] L. Evans and P. E. Souganidis, "Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations," *Indiana Univ. Math. J*, vol. 33, no. 5, pp. 773–797, 1984. 1, 3

[4] R. Bellman, *Dynamic programming*. Princeton University Press, 1957. 1, 4

[5] D. H. Jacobson, "New second-order and first-order algorithms for determining optimal control: A differential dynamic programming approach," *Journal of Optimization Theory and Applications*, vol. 2, no. 6, pp. 411–440, 1968. 1, 7

[6] D. H. Jacobson and D. Q. Mayne, *Differential Dynamic Programming*. American Elsevier Publishing Company, Inc., New York, NY, 1970. 1

[7] W. Denham, "Book Reviews: Differential Dynamic Programming," *Review of Social Economy*, vol. 36, no. 2, pp. 228–229, 1978. 1, 7

[8] W. Li and E. Todorov, "Iterative linear quadratic regulator design for nonlinear biological movement systems." in *ICINCO (1)*. Citeseer, 2004, pp. 222–229. 1

[9] M. G. Crandall and P.-L. Lions, "Viscosity solutions of hamilton-jacobi equations," *Transactions of the American mathematical society*, vol. 277, no. 1, pp. 1–42, 1983. 1, 3

[10] I. Mitchell, "A Robust Controlled Backward Reach Tube with (Almost) Analytic Solution for Two Dubins Cars (Presentation)," vol. 74, pp. 242–224, 2020. 1, 3

[11] 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*, vol. 282, no. 2, p. 487, 1984. 1, 3, 7

[12] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009. 2, 8

[13] N. Vannieuwenhoven, R. Vandebril, and K. Meerbergen, "A New Truncation Strategy for the Higher-Order Singular Value Decomposition," *SIAM Journal on Scientific Computing*, vol. 34, no. 2, pp. A1027–A1052, 2012. 2, 8, 9

[14] R. Isaacs, "Differential games 1965," *Kreiger, Huntigton, NY*. 3, 5, 7

[15] P.-L. Lions, *Generalized solutions of Hamilton-Jacobi equations*. London Pitman, 1982, vol. 69. 3

[16] L. Evans and P. E. Souganidis, "Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations," *Indiana Univ. Math. J*, vol. 33, no. 5, pp. 773–797, 1984. 3

[17] J. Lygeros, "On reachability and minimum cost optimal control," *Automatica*, vol. 40, no. 6, pp. 917–927, 2004. 3, 5

[18] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988. 3

[19] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE Transactions on Automatic Control*, vol. 50, no. 7, pp. 947–957, 2005. 3, 4, 5

[20] L. E. Dubins, "On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents," *American Journal of mathematics*, vol. 79, no. 3, pp. 497–516, 1957. 4

[21] A. Merz, "The game of two identical cars," *Journal of Optimization Theory and Applications*, vol. 9, no. 5, pp. 324–343, 1972. 4, 5

[22] K. Zhou and J. C. Doyle, *Essentials of Robust Control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104. 4

[23] A. Bajcsy, S. Bansal, E. Bronstein, V. Tolani, and C. 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)*. IEEE, 2019, pp. 1758–1765. 4

[24] S. Herbert, J. J. Choi, S. Sanjeev, M. Gibson, K. Sreenath, and C. J. Tomlin, "Scalable learning of safety guarantees for autonomous systems using hamilton-jacobi reachability," *arXiv preprint arXiv:2101.05916*, 2021. 4

[25] S. Bansal and C. J. Tomlin, "DeepReach : A Deep Learning Approach to High-Dimensional Reachability." 4

[26] S. Kaynama, I. M. Mitchell, M. Oishi, and G. A. Dumont, "Scalable safety-preserving robust control synthesis for continuous-time linear systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3065–3070, 2015. 4

[27] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A General Safety Framework for Learning-Based Control in Uncertain Robotic Systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2019. 4

[28] J. F. Fisac, N. F. Lugovoy, V. Rubies-Royo, S. Ghosh, and C. J. Tomlin, "Bridging hamilton-jacobi safety analysis and reinforcement learning," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2019-May, pp. 8550–8556, 2019. 4, 8

[29] M. Chen, S. L. Herbert, M. S. Vashishtha, S. Bansal, and C. J. Tomlin, "Decomposition of Reachable Sets and Tubes for a Class of Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3675–3688, 2018. 5

[30] S. L. Herbert, M. Chen, S. Han, S. Bansal, J. F. Fisac, and C. J. Tomlin, "Fastrack: A modular framework for fast and guaranteed safe motion planning," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 1517–1522. 5

[31] O. Ogunmolu, N. Gans, and T. Summers, "Minimax iterative dynamic game: Application to nonlinear robot control tasks," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 6919–6925. 5

[32] A. Nouy, "A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 23-24, pp. 1603–1626, 2010. 6

[33] A. Nouy and O. P. Le Maıˆtre, "Generalized spectral decomposition for stochastic nonlinear problems," *Journal of Computational Physics*, vol. 228, no. 1, pp. 202–235, 2009. 7

[34] P. Ladevèze, *Nonlinear computational structural mechanics: new approaches and non-incremental methods of calculation*. Springer Science & Business Media, 2012. 7

[35] A. Nouy, "Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 51-52, pp. 4718–4736, 2008. 7

[36] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. Siam, 1999, vol. 23. 7

[37] M. Crandall and A. Majda, "The method of fractional steps for conservation laws," *Numerische Mathematik*, vol. 34, no. 3, pp. 285–314, 1980. 7

[38] S. K. Mitter, "Successive approximation methods for the solution of optimal control problems," *Automatica*, vol. 3, no. 3-4, pp. 135–149, 1966. 7

[39] S. R. McReynolds, "The successive sweep method and dynamic programming," *Journal of Mathematical Analysis and Applications*, vol. 19, no. 3, pp. 565–598, 1967. 7

[40] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000. 8

[41] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966. 8

[42] A. Kapteyn, H. Neudecker, and T. Wansbeek, "An Approach to n-mode Components Analysis," *Psychometrika*, vol. 51, no. 2, pp. 269–275, 1986. 8

[43] S. Osher and R. Fedkiw, "Level Set Methods and Dynamic Implicit Surfaces," *Applied Mechanics Reviews*, vol. 57, no. 3, pp. B15–B15, 2004. 8

[44] I. M. Mitchell, "A Toolbox of Level Set Methods." *UBC Department of Computer Science Technical Report TR-2007-11*, 2007. 8

[45] R. Nishino, C. Loomis, and S. Hido, "Cupy: A numpy-compatible library for nvidia gpu calculations," *31st confernce on neural information processing systems*, vol. 151, 2017. 8