Proceedings of Machine Learning Research vol 144:1–13, 2022

A Proper Generalized Decomposition Algorithm for Backward Reachable Sets and Tubes.

Olalekan Ogunmolu LEKANMOLU@MICROSOFT.COM
Microsoft Research, 300 Lafayette Street, New York City, NY 10012, USA.

Ian Abraham IAN.ABRAHAM@YALE.EDU
Department of Mechanical and Aerospace Engineering, Yale University, New Haven, Connecticut, USA.

Sylvia Herbert SHERBERT@ENG.UCSD.EDU
Department of Mechanical and Aerospace Engineering, UC San Diego, La Jolla, CA, USA.

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 partial core*: an iterative quasi-proper generalized decomposition (PGD) 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.

Keywords: Control Systems, Hamilton-Jacobi-Isaacs, Optimal Control, Systems Verification, Reachability Analysis, Galerkin Methods, Reduced Order Models

1. 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 sys-

tems 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)

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

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

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 stability guarantees;
- stability in well-defined regions of the state space where approximation is guaranteed to work;
- 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 § 2; § 3 describes the concepts and topics we will build upon in describing our proposal in § 4; we present results and insights from experiments in § 5. We conclude the paper in § 6. 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.

2. 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 bold-face 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 $\{0, 1, \cdots, N-1\}$.

2.1. Vectors, Matrices, and Tensors.

2.1.1. 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 $\{\mathbf{e}_i\}$.

2.1.2. 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 3, 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 3. We refer readers to

1. Or the dyadic product.

Table 1: **Common Notations**

Table 2: *

	Notation	Description
Tensor Operations	$\mathbb{T}_n$	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 $\mathbf{S}$.

Table 3: *

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.

2.2. 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 int Ω; 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.

3. Background and Preliminaries.

3.1. 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

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 $\boldsymbol{v}(\cdot)$

3.2. 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. on $t \leq \bar{t} \leq \tau$ implies $\beta[\boldsymbol{u}](\bar{t}) = \beta[\hat{\boldsymbol{u}}](\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$. 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

$$\begin{aligned} \boldsymbol{V}^-(\boldsymbol{x}, t) &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} V(\boldsymbol{u}, \beta[\boldsymbol{u}]) \\ &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} \int_t^T l(\tau, \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \beta[\boldsymbol{u}](\tau)) d\tau + g\left(\boldsymbol{x}(T)\right). \end{aligned} \tag{9}$$

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

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

These non-local PDEs i.e. (9) and (10) 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

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

We now establish two Lemmas from

3.4. Reachability for Systems Verification.

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

• *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,

BRTs are popularly analyzed on a game of two vehicles with non-stochastic dynamics

3.4.1. 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

Recent works have started exploring scaling up the Cauchy-type HJ problem for guaranteeing safety of higher-dimensional physical systems: the authors of

In

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?
• Can we provide rational decomposition schemes that preserve the numerical stability of monotone Lax-Friedrichs and essentially non-oscillatory 3.4.2. 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

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t; t_0, \boldsymbol{x}_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) \tag{11}$$

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

$$\boldsymbol{\xi}(t_0) = \boldsymbol{\xi}(t_0; t_0, \boldsymbol{x}_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) = \boldsymbol{x}_0. \tag{12}$$

Read (11): 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 (11)

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

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), or (10) 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 (**??**) such that the solution to (8) enters the unsafe set. Following the backward reachability formulation of 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. (**??**)); – 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. (**??**)); – II has regular controls, $\boldsymbol{u}$, drawn from a Lebesgue measurable set, $\mathcal{U}$ (c.f. (9)).

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)

TO-DO: We essentially obtain a *pseudo iterative dynamic game*

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

$$\frac{\partial \boldsymbol{V}^-}{\partial t}(\boldsymbol{x}, t) + \min\{0, \boldsymbol{H}^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{V}^-_{\boldsymbol{x}})\} = 0 \tag{13a}$$

$$\boldsymbol{V}^-(\boldsymbol{x}, 0) = g(\boldsymbol{x}), \tag{13b}$$

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

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

$$\boldsymbol{V}(\boldsymbol{x}, 0) = g(x), \quad \boldsymbol{x} \in \mathbb{R}^n \tag{14b}$$

with lower Hamiltonian,

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

4. 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.

4.1. ROM Construction

We consider separated representations of the value function $\boldsymbol{V}(\boldsymbol{x}, t)$ (c.f. (**??**)), parameterized by basis functions $\boldsymbol{\phi} \in \boldsymbol{\Phi}$[3] 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 (**??**).

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

$$\langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_{\mathcal{F}} = \int_{\mathcal{S}} \langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_{\mathcal{F}} d\boldsymbol{x}. \tag{16}$$

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

$$L^2(\mathcal{S}; \mathcal{F}) = \left\{ \boldsymbol{\phi} : \mathcal{S} \to \mathcal{F}; \int_{\mathcal{S}} \|\boldsymbol{\phi}(\boldsymbol{x})\|_{\mathcal{F}}^2 \, d\boldsymbol{x} < +\infty \right\} \tag{17}$$

and associate with it the dual space $\breve{\mathcal{F}} = H^{-1}(\Omega)$[4]. **The decomposition problem is to find** $\boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi}) \in \mathcal{F} \otimes \mathcal{S}$ such that

$$\frac{\partial \boldsymbol{V}}{\partial t}(\boldsymbol{x}, t) \in L^2(\breve{\mathcal{F}} \otimes \mathcal{S}) \tag{18}$$

and

$$\amalg(\boldsymbol{V}, \boldsymbol{\phi}) = \uplus(\boldsymbol{\phi}), \ \forall \, \boldsymbol{\phi} \in \mathcal{F} \otimes \mathcal{S}. \tag{19}$$

where $\amalg(\cdot, \cdot)$ and $\uplus(\cdot)$ are respectively multilinear transforms (to be introduced shortly). The solution of (19) satisfies the boundary conditions to the initial value problem of c.f. (13) in a weak sense

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. ,

$$\boldsymbol{V}(\boldsymbol{x}, t) \approx \boldsymbol{V}(\boldsymbol{x}, t; \boldsymbol{\phi}) \tag{20a}$$

$$\equiv \sum_{i=0}^{\infty} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_s, \ \boldsymbol{\phi}_i \in \mathcal{F}, \ \boldsymbol{\psi}_i \in \mathcal{S} \tag{20b}$$

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 § 3.3.*

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

$$\boldsymbol{V}_r(\boldsymbol{x}, t) = \sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_{\mathcal{F}}, \ \boldsymbol{\phi}_i \in \mathcal{F}, \ \boldsymbol{\psi}_i \in \mathcal{S} \tag{21}$$

3. Note that $\boldsymbol{\Phi} \subset \mathcal{F} \in \mathbb{R}^{d_\phi}$.
4. Following (1), when we write a norm in the form of (16), we shall adopt the Frobenius inner product norm throughout.

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

$$\|\boldsymbol{V} - \boldsymbol{V}_r\|_F^2 = \min_{\substack{\{\boldsymbol{\psi}_i\}_{i=1}^{\infty} \in \mathcal{S} \\ \{\boldsymbol{\phi}_i\}_{i=1}^{\infty} \in \mathcal{F}}} \|\boldsymbol{V} - \sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i(t), \boldsymbol{\phi}_i(\boldsymbol{x}) \rangle_{\mathcal{F}}\|_F^2. \tag{22}$$

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. § 2) so that the full problem described in Lemma **??**, corresponds to the following parametric P.D.E

$$\frac{\partial \mathbb{V}}{\partial t}(\boldsymbol{x}, t; \boldsymbol{\phi}) + \mathbb{H}(\boldsymbol{x}, t, p; \boldsymbol{\phi}) = 0, \; t \in [T, 0], x \in \Omega, \boldsymbol{\phi} \in \boldsymbol{\Phi} \tag{23a}$$

$$\mathbb{V}(\boldsymbol{x}, T; \boldsymbol{\phi}) = g(\boldsymbol{x}; \boldsymbol{\phi}) \quad \boldsymbol{x} \in \{\bar{\Omega} \backslash \text{int } \Omega\}, \; \boldsymbol{\phi} \in \boldsymbol{\Phi} \tag{23b}$$

with corresponding Hamiltonian,

$$\mathbb{H}(\boldsymbol{x}, t, p; \boldsymbol{\phi}) = \min \left\{ 0, \max_{\boldsymbol{u} \in \mathcal{U}} \min_{\boldsymbol{v} \in \mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), p \rangle \right\}. \tag{24}$$

The decomposition of (21) can be considered a pseudo-eigenvalue problem (this is established in § 4.5) which proves efficient for separated representations in many applications including stochastic nonlinear PDEs

4.2. 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 § 4.5) 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

$$\begin{aligned} \|\mathbb{V} - \mathbb{V}_r\|_F^2 &= \|\mathbb{V}\|_F^2 - \sigma(\phi_i(\boldsymbol{x})) \\ &\equiv \|\mathbb{V}\|_F^2 - \|\breve{\mathbb{V}}\|_F^2, \end{aligned} \tag{25}$$

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 § 4.5). The optimality proof of (25) is given in appendix **??** via (**??**).

Suppose that the optimal or quasi-optimal[5] order r for decomposing $\mathbb{V}$ in (23b) has been found. The quasi-optimal Galerkin-approximation of (23b) is given as

$$\frac{\partial \mathbb{V}}{\partial t}(\boldsymbol{x}, t; \boldsymbol{\phi}) + \mathbb{H}(\boldsymbol{x}, t, p; \boldsymbol{\phi}) = 0 \; x \in \Omega, \boldsymbol{\phi} \in \mathcal{F} \tag{26}$$

or

$$\begin{aligned} -\frac{\partial}{\partial t}\left(\sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i, \boldsymbol{\phi}_i \rangle_{\mathcal{F}}\right) = \min\Bigg\{0, \max_{\boldsymbol{u}\in\mathcal{U}} \min_{\boldsymbol{v}\in\mathcal{V}} \langle f(t; x, \boldsymbol{u}, \boldsymbol{v}), \\ \frac{\partial}{\partial \boldsymbol{x}}\left(\sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i, \boldsymbol{\phi}_i \rangle_{\mathcal{F}}\right)\Bigg\rangle\Bigg\} \end{aligned} \tag{27}$$

$$\begin{aligned} -\sum_{i=0}^{r-1} \langle \boldsymbol{\psi}_i, \frac{\partial \boldsymbol{\phi}_i}{\partial t} \rangle_{\mathcal{F}} = \min\Bigg\{0, \max_{\boldsymbol{u}\in\mathcal{U}} \min_{\boldsymbol{v}\in\mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), \\ \sum_{i=0}^{r-1} \left\langle \boldsymbol{\psi}_i \frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{x}} \right\rangle_{\mathcal{F}} \Bigg\rangle\Bigg\}. \end{aligned} \tag{28}$$

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

$$\boldsymbol{\Psi}_r^T = [\boldsymbol{\psi}_0, \boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_{r-1}], \; \boldsymbol{\Phi}_r = [\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_{r-1}]^T. \tag{29}$$

Thus, we can rewrite (28) as

$$-\boldsymbol{\Psi}_r^T \dot{\boldsymbol{\Phi}}_r = \min\left\{0, \max_{\boldsymbol{u}\in\mathcal{U}} \min_{\boldsymbol{v}\in\mathcal{V}} \langle f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}), \langle \boldsymbol{\Psi}_r^T, \boldsymbol{\Phi}_{\boldsymbol{x}} \rangle_{\mathcal{F}} \rangle\right\}. \tag{30}$$

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 (30) 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 (30) admits separable solutions on finite meshes, rendering solution of the viscosity problem (13) easily computable with the usual high precision and accuracy that Lax-Friedrichs schemes afford

4.3. 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,

$$\dot{\boldsymbol{x}} = f(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}) \tag{31}$$

$$= f_{\boldsymbol{x}}\boldsymbol{x}(t) + f_{\boldsymbol{u}}(t)\boldsymbol{u}(t), f_{\boldsymbol{v}}\boldsymbol{v}(t), \tag{32}$$

5. The rationale for introducing quasi-optimality is discussed in section 4.5 where the Galerkin basis functions and coefficients are resolved with sequentially-truncated higher order singular value decomposition of $\mathbb{V}(\boldsymbol{x}, t)$.

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[6].

Putting (32) into (30), we find that

$$-\boldsymbol{\Psi}_r^T \dot{\boldsymbol{\Phi}}_r = \min \Big\{ 0, \max_{\boldsymbol{u}\in\mathcal{U}} \min_{\boldsymbol{v}\in\mathcal{V}} \langle f_x x(t) + f_u(t)\boldsymbol{u}(t), f_v \boldsymbol{v}(t), \big\langle \boldsymbol{\Psi}_r^T, \boldsymbol{\Phi}_{\boldsymbol{x}} \big\rangle_{\mathcal{F}} \big\rangle \Big\}, \tag{33}$$

so that the term on the right hand side becomes

$$\min \Big\{ 0, \big\langle f_x \boldsymbol{x}(t), \big\langle \boldsymbol{\Psi}_r^T, \boldsymbol{\Phi}_{\boldsymbol{x}} \big\rangle_{\mathcal{F}} \big\rangle + \max_{\boldsymbol{u}\in\mathcal{U}} \langle f_u(t)\boldsymbol{u}(t), \big\langle \boldsymbol{\Psi}_r^T, \boldsymbol{\Phi}_{\boldsymbol{x}} \big\rangle_{\mathcal{F}} \big\rangle + \min_{\boldsymbol{v}\in\mathcal{V}} \big\langle f_v \boldsymbol{v}(t), \big\langle \boldsymbol{\Psi}_r^T, \boldsymbol{\Phi}_{\boldsymbol{x}} \big\rangle_{\mathcal{F}} \big\rangle \Big\}. \tag{34}$$

Similar to (30), (34) can be resolved on a mesh. However, when the dynamics are separable as in the foregoing, the saddle point necessary condition i.e. (**??**) 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?

4.4. 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 (30) and (34). 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. ,

$$\amalg(\boldsymbol{V}_{r-1} + \phi\psi, \phi\psi_1^\star + \phi_1^\star\psi, \psi_2^\star \cdots \psi_r^\star) = \uplus(\phi^\star\psi), \forall\, \phi^\star \in \mathcal{F}. \tag{35}$$

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

$$\amalg(\boldsymbol{V}_{r-1} + \phi\psi_0, \phi\psi_1, \cdots \phi\,\psi_r) = \uplus(\phi^\star\psi), \forall\, \phi^\star \in \mathcal{F}. \tag{36}$$

TO-DO: we establish a Lemma due to

. In our treatment, we resort to higher-order singular value decomposition (HOSVD)

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

Therefore, the approximation error, $\|\mathbb{V} - \hat{\mathbb{V}}\|^2 = \|\mathbb{V}\|^2 - \|\breve{\mathbb{V}}\|^2$ (as shown in Appendix **??**). It now suffices to define the projection approximation error of the value function. A truncated

6. Even when the separation in (32) 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

Algorithm 1 Iterative PGD Scheme for Computing BRS/BRTs

[1] IterativeBRT$\boldsymbol{\psi}_0, \boldsymbol{\phi}_0, r, r_{max}$ r_{max}: max. iterations. Set $\mathbb{V}_{r-1} = \langle \psi_0, \phi_0 \rangle$ **while** $r < r_{max}$ **do**

end

Compute $\boldsymbol{\psi}_r$ from (39). Compute $\boldsymbol{\phi}_r$ from (40). Set $\mathbb{V}_r \leftarrow \langle \mathbb{V}_{r-1} + \psi_r \phi_r \rangle_{\mathcal{F}}$ Update $\mathbb{V}_r$. Set $\langle \boldsymbol{\Psi}_r^T, \dot{\boldsymbol{\Phi}}_r \rangle_F \leftarrow \mathbb{V}_r$ c.f. (30) or (34) $\mathbb{L}_r \leftarrow \int_{\mathcal{S}} \langle \boldsymbol{\Psi}_r^T, \dot{\boldsymbol{\Phi}}_r \rangle_{\mathcal{F}} d\boldsymbol{x}$. Compute Partial BRT, $\mathbb{L}_r$ $r = r + 1$. Advance the decomposition.

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

$$\breve{\mathbb{V}}_r = \mathbb{V} \otimes_0 \mathbf{U}_0^T \otimes_1 \mathbf{U}_1^T \otimes_2 \mathbf{U}_2^T \cdots \otimes_{r-1} \mathbf{U}_r^T, \tag{37}$$

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

$$\mathbb{V}_r = \breve{\mathbb{V}}_r \otimes_0 \mathbf{U}_0 \otimes_1 \mathbf{U}_1 \otimes_2 \mathbf{U}_2 \cdots \otimes_{r-1} \mathbf{U}_r \tag{38a}$$

$$\equiv \mathbb{V} \otimes_0 \mathbf{U}_0 \mathbf{U}_0^T \otimes_1 \cdots \otimes_{r-1} \mathbf{U}_r \mathbf{U}_r^T, \, r \in [R]\,. \tag{38b}$$

where (38b) is a result of putting (37) into (38a).

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 (25) is to set

$$\boldsymbol{\psi}_r = \mathbb{V} \otimes_0 \mathbf{U}_0^T \cdots \otimes_{r-1} \mathbf{U}_r^T \tag{39}$$

and

$$\boldsymbol{\phi}_r = \breve{\mathbb{V}}_r \otimes_0 \mathbf{U}_0 \cdots \otimes_{r-1} \mathbf{U}_r. \tag{40}$$

Algorithm 1 describes how we compose the separable value function that satisfies (19). Lines 1 and 1 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 1 of the algorithm, the integral is solved using Total Variation Diminishing (TVD) Runge-Kutta scheme as described in

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

$$\| \mathbb{V} \otimes_n \left(\mathbf{I} - \mathbf{U}_n \mathbf{U}_n^T \right) \|_F^2 \leq \delta^2 \frac{\| \mathbb{V} \|_F^2}{N}. \tag{41}$$

Algorithm 2 Incrementally Truncated Higher-Order SVD.

[1] Truncated$\mathbb{V}$, ϵ ϵ: Desired accuracy. $\mathbb{P} \leftarrow \mathbb{V}\,\mathbb{P}$: c.f. (3). **for** $n = 0, 1, \cdots, N-1$ **do**

end

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

Whilst appealing, algorithm 1 utilizes the full value function at each step (c.f. (39)); 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. **4.5. Informed Incremental Value Function Decompositions**

Next, we will leverage the sequentially-truncated high-order SVD of **5. Results and Discussion.**

6. Conclusion.

References