Approximation of ND systems with multiple dependent variables

F. van Belzen and S. Weiland

Abstract— **Multi-variable distributed systems describe the evolution of multiple dependent variables over a domain of independent variables. This paper considers model reduction for this type of systems. The method of Proper Orthogonal Decompositions (POD) is adapted using concepts from tensors and tensor decompositions. The result is a model reduction framework that is applicable to systems with an arbitrary number of dependent and independent variables.**

I. Introduction

Multi-variable distributed systems describe the evolution of multiple dependent variables over a domain of independent variables. Such evolutions are usually described by systems of Partial Differential Equations (PDEs). Since analytical solutions of such systems can typically not be inferred, model reduction is particularly relevant to obtain insight in the dynamical behavior of such systems. The technique of combining Proper Orthogonal Decompositions (POD) of signals with Galerkin projections of equation residuals is among the most popular model reduction method, especially in the field of computational fluid dynamics. The technique consists of two steps. Firstly, all signal evolutions are approximated by finite spectral expansions. Secondly, Galerkin projections are employed to project equation residuals on finite dimensional projection spaces so as to represent the dynamic behavior of spectral coefficients.

In its existing form, application of POD to systems with multiple dependent and independent variables leads to a number of problems. First, the definition of the projection spaces requires that all spatial variable variables may be lumped. This way, the structure that is present in the original system description is discarded. Second, the projection spaces used in POD are usually empirical projection spaces, derived from measured or simulated data. When multiple dependent variables are present, the projection spaces may become very sensitive to the scaling of these variables. Third, for large scale systems which require large grid sizes, the basis functions describing the projection spaces may become very high-dimensional since their dimension is in general equal to the number of grid-points. This is usually not desired.

Therefore, this paper aims to adapt the POD method to overcome these drawbacks. It is our specific purpose to derive a framework that is able to handle any number of dependent and independent variables. In this paper, this is achieved by combining POD and the discipline of tensors and tensor decompositions.

This paper is organized as follows. First, the necessary concepts are introduced. Section II discusses tensors and tensor decompositions, whereas Section III provides a short introduction to Proper Orthogonal Decompositions. Then, the concept of Partial Difference Equations is generalized to tensors. It is shown how a tensor can be associated with a physical signal and the shift operator is defined for tensors. Then, we show that a tensor operation employing these shifts is equivalent to a PDE in terms of signals. Finally, all elements are combined to define reduced models. We end with conclusions.

II. Preliminaries

Tensors and tensor decompositions are formally introduced in this section. The concepts discussed here will be used throughout the remainder of the paper.

A. Tensors

An *order-N tensor* is a multi-linear functional

$$T : \mathcal{W}_1 \times \ldots \times \mathcal{W}_N \to \mathbb{R}$$

defined on vector spaces $\mathcal{W}_1, \ldots, \mathcal{W}_N$. That is, T is a linear functional in each of its N arguments. In this paper we will consider the special case where $\mathcal{W}_1, \ldots \mathcal{W}_N$ are separable Hilbert spaces. Elements of T are specified as $t_{\ell_1 \cdots \ell_N}$ where ℓ_k ranges from 1 till the dimension of $\mathcal{W}_k$, denoted L_k, and k ranges from 1 till N. Elements of T are commonly encoded in the N-way array $[[t_{\ell_1 \cdots \ell_N}]] \in \mathbb{R}^{L_1 \times \cdots \times L_N}$ which, especially in signal processing, is taken as a (coordinate-dependent) definition of a tensor. The elements $t_{\ell_1 \cdots \ell_N}$ represent T with respect to a specific collection of bases

$$\{e_1^{(\ell_1)},\ \ell_1 = 1, \ldots, L_1\}\ , \ldots ,\ \{e_N^{(\ell_N)},\ \ell_N = 1, \ldots, L_N\} \tag{1}$$

of $\mathcal{W}_1, \ldots, \mathcal{W}_N$, respectively, in the sense that $t_{\ell_1 \cdots \ell_N} = T(e_1^{(\ell_1)}, \ldots, e_N^{(\ell_N)})$.

Throughout, the set of all order-N tensors on $\mathcal{W}_1 \times \ldots \times \mathcal{W}_N$ is denoted by $\mathcal{T}_N$ which becomes a vector space when equipped with the standard definitions of addition and scalar multiplication.

To define approximations to tensors we will need a *norm* on the space $\mathcal{T}_N$. For this let $\| \cdot \|_k$ denote the induced norm corresponding to the inner product $\langle \cdot, \cdot \rangle_k$ of $\mathcal{W}_k$. We assume this structure for $k = 1, \ldots, N$. The *inner product* of two tensors $S, T \in \mathcal{T}_N$ with elements $s_{k_1 \cdots k_N}$ and $t_{\ell_1 \cdots \ell_N}$, both

Supported by the Dutch Technologiestichting STW under project number EMR.7851.

Department of Electrical Engineering, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands f.v.belzen@tue.nl, s.weiland@tue.nl

ISBN 978-963-311-370-7

defined with respect to the bases (1), is given by

$$\langle S,T\rangle := \sum_{k_1}\cdots\sum_{k_N}\sum_{\ell_1}\cdots\sum_{\ell_N} s_{k_1\cdots k_N} t_{\ell_1\ldots\ell_N} \langle e_1^{(k_1)}, e_1^{(\ell_1)}\rangle\cdots\langle e_N^{(k_N)}, e_N^{(\ell_N)}\rangle.$$

It is immediate that the right-hand side of this expression is invariant under unitary basis transformations (i.e., transformations $Q_k : \mathcal{W}_k \to \mathcal{W}_k$ for which $\|Q_k x\|_k = \|x\|_k$ for all $x \in \mathcal{W}_k$) and so $\mathcal{T}_N$ becomes a well defined inner product space.

For fixed elements $u_k \in \mathcal{W}_k$, $k = 1,\ldots,N$, the functional

$$U(w_1,\ldots,w_N) := \langle u_1, w_1\rangle_1 \cdots \langle u_N, w_N\rangle_N = \Pi_{k=1}^N \langle u_k, w_k\rangle_k$$

defines an order-N tensor which will be denoted by $U = u_1 \otimes \cdots \otimes u_N$. Whenever non-zero, such a tensor will be referred to as a *rank-1 tensor*. With respect to the bases (1), the elements of U are $u_{\ell_1\cdots\ell_N} = u_1^{\ell_1}\cdots u_N^{\ell_N}$ where $u_k^{\ell_k} = \langle u_k, e_k^{(\ell_k)}\rangle_k$ is the coefficient of u_k with respect to the basis vector $e_k^{\ell_k n}$. We have that $\|U\| = \Pi_{k=1}^N \|u_k\|_k$. Every tensor can be represented as a weighted sum of rank-one tensors as follows

$$T = \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N} e_1^{(\ell_1)} \otimes\cdots\otimes e_N^{(\ell_N)} \tag{2}$$

The following lemma proves useful and relates tensor evaluations with tensorial inner products.

Lemma II.1 *Let $T \in \mathcal{T}_N$, $T : \mathcal{W}_1 \times\cdots\times \mathcal{W}_N \to \mathbb{R}$, with $\mathcal{W}_n$ inner product spaces, possibly infinite dimensional, and $w_n \in W_n$ for $n = 1,\ldots,N$. Then*

1)
$$T(w_1,\cdots,w_N) = \langle T, w_1\otimes\cdots\otimes w_N\rangle.$$

2)
$$T(w_1,\cdots,w_N) = \langle w_N, \cdots\langle w_2, \langle w_1, T\rangle_1\rangle_2\cdots\rangle_N.$$

Proof:

1) Let $\{\xi_n^{(\ell_n)}\}_{\ell_n=1}^\infty$ be an orthonormal basis for $\mathcal{W}_n$, $n = 1,\ldots,N$. T can be represented with respect to these bases as $T = \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N}\xi_1^{(\ell_1)} \otimes\cdots\otimes \xi_N^{(\ell_N)}$. The tensor evaluation can be written as $T(w_1,\cdots,w_N) = \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N}\langle w_1, \xi_1^{\ell_1}\rangle\cdots\langle w_N, \xi_N^{\ell_N}\rangle$. Let $U := w_1\otimes\cdots\otimes w_N$. U can be represented as $U = \sum_{\ell_1}\cdots\sum_{\ell_N} u_{\ell_1\cdots\ell_N}\xi_1^{(\ell_1)}\otimes\cdots\otimes\xi_N^{(\ell_N)}$ with $u_{\ell_1\cdots\ell_N} = \prod_{i=1}^N \langle w_i, \xi_i^{(\ell_i)}\rangle_i$. Then,

$$\begin{aligned}\langle T,U\rangle &= \sum_{k_1}\cdots\sum_{k_N}\sum_{\ell_1}\cdots\sum_{\ell_N} t_{k_1\cdots k_N} u_{\ell_1\cdots\ell_N} \\ &\qquad\cdot \underbrace{\langle \xi_1^{k_1}, \xi_1^{\ell_1}\rangle}_{0 \text{ unless } k_1=\ell_1} \cdots \langle \xi_N^{k_N}, \xi_N^{\ell_N}\rangle \\ &= \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N} u_{\ell_1\cdots\ell_N} \\ &= \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N}\langle \xi_1^{\ell_1}, w_1\rangle\cdots\langle \xi_N^{\ell_N}, w_N\rangle\end{aligned}$$

which is the tensor evaluation.

2) To prove the second statement, we first show that $\langle w_1, T(\cdot, v_2,\ldots,v_N)\rangle_1 = T(w_1, v_2,\ldots,v_N)$ for some $v_n \in W_n$, $n = 2,\ldots,N$. Let $\{\xi_n^{(\ell_n)}\}_{\ell_n=1}^\infty$ be an orthonormal basis for $\mathcal{W}_n$, $n = 1,\ldots,N$. T can be represented with respect to these bases as $T = \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N}\xi_1^{(\ell_1)}\otimes\cdots\otimes\xi_N^{(\ell_N)}$. Then

$$T(w_1, v_2,\ldots,v_N) = \sum_{\ell_1}\cdots\sum_{\ell_N} t_{\ell_1\cdots\ell_N}\langle \xi_1^{(\ell_1)}, w_1\rangle \prod_{k=2}^N \langle \xi_k^{(\ell_k)}, v_k\rangle.$$

On the other hand, we can write w_1 as $w_1 = \sum_{k=1}^\infty \langle w_1, \xi_1^{(k)}\rangle \xi_1^{(k)}$. Then

$$\begin{aligned}\langle w_1, W(\cdot, v_2,\ldots,v_N)\rangle_1 &= \sum_{k_1}\langle w_1, \xi_1^{(k_1)}\rangle_1 \, T(\xi_1^{(k_1)}, v_2,\ldots,v_N) \\ &= \sum_{k_1}\langle w_1, \xi_1^{(k_1)}\rangle_1 \sum_{\ell_2}\cdots\sum_{\ell_N} t_{k_1\ell_2\cdots\ell_N}\prod_{k=2}^N\langle \xi_k^{(\ell_k)}, v_k\rangle \\ &= T(w_1, v_2,\ldots,v_N)\end{aligned}$$

Thus, we have that $\langle w_1, W(\cdot, v_2,\ldots,v_N)\rangle_1 = T(w_1, v_2,\ldots,v_N)$. Since tensors are multilinear functionals, this completes the proof.

■

B. Tensor Decompositions

In Eq. (2) a representation of a tensor T was given in terms of the basis functions $\{e_k^{(\ell_k)},\ k = 1,\ldots L_k\}$, $k = 1,\ldots,N$. In this section, we will define a change of basis for T. More specifically, we define basis transformations that will give a good approximation of the tensor when only a limited number of basis functions is used.

More specifically, introduce sets of basis functions $\{\varphi_k^{(\ell_k)},\ k = 1,\ldots,L_k\}$. The representation of T with respect to these basis functions is given by

$$T = \sum_{\ell_1}\cdots\sum_{\ell_N} \tilde{t}_{\ell_1\cdots\ell_N}\varphi_1^{(\ell_1)}\otimes\cdots\otimes\varphi_N^{(\ell_N)}$$

with $\tilde{t}_{\ell_1\cdots\ell_N}$ defined as $\tilde{t}_{\ell_1\cdots\ell_N} = T(\varphi_1^{(\ell_1)},\ldots,\varphi_N^{(\ell_N)})$.

A truncation of T to degree $r = (r_1,\ldots,r_N)$ can be defined as

$$T_r = \sum_{\ell_1=1}^{r_1}\cdots\sum_{\ell_N=1}^{r_N} \tilde{t}_{\ell_1\cdots\ell_N}\varphi_1^{(\ell_1)}\otimes\cdots\otimes\varphi_N^{(\ell_N)} \tag{3}$$

The tensor is now represented by a (coordinate-dependent) N-way array $[[\tilde{t}_{\ell_1\cdots\ell_N}]] \in \mathbb{R}^{r_1\times\cdots\times r_N}$. The discipline of tensor decompositions [6],[7] concerns itself with finding basis functions $\{\varphi_k^{(\ell_k)},\ k = 1,\ldots,r_k\}$, $k = 1,\ldots,N$ such that the error

$$\|T - T_r\| \tag{4}$$

in the norm induced by the tensor inner product is minimized
.

For order-2 tensors, matrices, this approximation problem is solved in the 2-norm by truncation of the Singular Value Decomposition (SVD). In general, there is no optimal solution and several methods to compute tensor decompositions exist. These methods include the Higher-Order Singular Value Decomposition [4], HOSVD, and the Tensor Singular Value Decomposition, TSVD [3]. Each of these methods has different approximation properties, no method is yet known that gives optimal approximations to tensors.

III. Proper Orthogonal Decompositions

In this section a brief introduction to the method of Proper Orthogonal Decompositions is given. More details can be found in for example [1], [2], [5].

We consider a system with N independent variables, namely $x_1, \ldots, x_N$, and n dependent variables $w_1, \ldots, w_n$. The dependent variables will be stacked in a vector as follows

$$\underline{w}(x_1, \ldots, x_N) = \begin{bmatrix} w_1(x_1, \ldots, x_N) \\ \vdots \\ w_n(x_1, \ldots, x_N) \end{bmatrix} \in W \subseteq \mathbb{R}^n. \quad (5)$$

$\underline{w}$ is the solution trajectory of an arbitrary distributed system described by the Partial Differential Equation (PDE)

$$R(\partial_{x_1}, \ldots, \partial_{x_N})\underline{w} = \underline{0} \quad (6)$$

where $R \in \mathbb{R}^{\cdot \times n}[\xi_1, \ldots, \xi_N]$ and the partial differential operators $\frac{\partial}{\partial x_k}$, $1 \leq k \leq N$ are denoted by ∂_{x_k}. (6) is a mapping $\mathbb{D} \subset \mathbb{R}^N \to W$ and forms a PDE in the vector-valued signal $\underline{w}$ that evolves over N independent variables.

The domain of the signal $\underline{w}$, $\mathbb{D}$, is assumed to have the Cartesian structure $\mathbb{D} = \mathbb{X}' \times \mathbb{X}''$, which is typically the product of a spatial and a temporal domain. A Hilbert space $\mathcal{H}$ of square integrable functions on $\mathbb{X}'$ is introduced. $\mathcal{H}$ is assumed to be separable, which means that a countable orthonormal basis $\{\varphi_k,\ k = 1, 2, \ldots\}$ for $\mathcal{H}$ exists. Solutions $\underline{w}$ of (6) with $\underline{w}(\cdot, x'') \in \mathcal{H}$ can now be represented by a spectral expansion as follows

$$\underline{w}(x', x'') = \sum_k a_k(x'')\underline{\varphi}(x') \quad (7)$$

in which the modal coefficients a_k are uniquely determined by $a_k = \langle w, \varphi_k \rangle$. For $r > 0$ reduction of the signal space $\mathcal{H}$ can be defined by the truncation

$$\underline{w}_r(x', x'') = \sum_{k=1}^{r} a_k(x'')\underline{\varphi}(x') \quad (8)$$

For $r > 0$, the reduced order model is then defined by the collection of solutions $\underline{w}_r(x', x'') = \sum_{k=1}^{r} a_k(x'')\underline{\varphi}_k(x')$ that satisfy the Galerkin projection [2]

$$\langle R\left(\partial_{x_1}, \ldots, \partial_{x_N}\right) \underline{w_r}, \underline{\varphi} \rangle = 0 \quad \forall \underline{\varphi} \in \mathcal{H}_r \quad (9)$$

where $\mathcal{H}_r$ is the finite dimensional projection space $\mathcal{H}_r = \mathrm{span}\{\varphi_1, \ldots, \varphi_r\}$. If the spectral expansion of $\underline{w}_r$ is substituted in (9) and $\mathbb{X}'' \subseteq \mathbb{R}$, then (9) becomes a system of r ordinary differential equations in the modal coefficients a_k, $k = 1, \ldots, r$.

Clearly, the quality of the reduced order model entirely depends on the choice of basis functions $\{\varphi_k\}$. In the POD method, the orthonormal basis functions φ_k of $\mathcal{H}$ depend on data that have been either measured or inferred from the model (6). Specifically, for given data w with $w(\cdot, x'') \in \mathcal{H}$, the basis functions φ_k are the ordered normalized eigenfunctions of the data correlation operator $\Phi : \mathcal{H} \to \mathcal{H}$ that is defined as

$$\langle \psi_1, \Phi\psi_2 \rangle := \int_{\mathbb{X}''} \langle \psi_1, w(\cdot, x'') \rangle \cdot \langle w(\cdot, x''), \psi_2 \rangle \mathrm{d}\ x'' \qquad \psi_1, \psi_2 \in \mathcal{H}.$$

The data correlation operator Φ is a well defined linear, bounded, self-adjoint and non-negative operator on $\mathcal{H}$. That is, the basis functions φ_k satisfy $\Phi\varphi_k = \lambda_k \varphi_k$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$.

In applications, the PDE (6) is discretized by sampling the domain $\mathbb{X} = \mathbb{X}' \times \mathbb{X}''$ and by discretizing the PDE according to

$$D(\varsigma_1, \ldots, \varsigma_N)\underline{w} = 0 \quad (10)$$

where $D \in \mathbb{R}^{\cdot \times n}[\xi_1, \ldots, \xi_N]$ is a real matrix-valued polynomial in N indeterminates and ς_k is the forward shift operator acting on the spatial discretization in the kth mode according to $\varsigma_k s(x_1^{(\ell_1)}, \ldots, x_k^{(\ell_k)}, \ldots, x_N^{(\ell_N)}) = s(x_1^{(\ell_1)}, \ldots, x_k^{(\ell_{k+1})}, \ldots, x_N^{(\ell_N)})$. We will assume that (10) is an accurate representation of (6) and refer to solutions of (10) as Finite Element solutions. For the discretized model (10) the solution space $\mathcal{H}$ becomes finite, but large, dimensional, provided that $\mathbb{X}'$ is finite. Then Φ becomes a symmetric non-negative definite matrix and the calculation of POD basis functions becomes an algebraic eigenvalue or singular value decomposition problem.

IV. Interpretation of PDE in terms of tensors

Throughout the remainder of this paper the Partial Difference Equation (PDE) (10) will be considered only. We view $D(\varsigma_1, \ldots, \varsigma_n)$ as an operator mapping signal $\underline{w}$ to the residual $\underline{r} := D(\varsigma_1, \ldots, \varsigma_N)\underline{w}$. It is the purpose of this section of the paper to show that the mapping form $\underline{w}$ to $\underline{r}$ via the PDE (10) is equivalent to mapping a tensor W associated with $\underline{w}$ to another tensor R via a tensor operation which is equivalent to a PDE. This is depicted in the following diagram

$$\begin{array}{ccc} \underline{w}(x_1, \ldots, x_N) & \xrightarrow{D(\varsigma_1, \ldots, \varsigma_N)} & \underline{r}(x_1, \ldots, x_N) \\ \Updownarrow & & \Updownarrow \\ W \in \mathcal{T}_N & \xrightarrow{D(\mathfrak{S}_1, \ldots, \mathfrak{S}_N)} & R \in \mathcal{T}_N \end{array} \quad (11)$$

The reason for showing that a PDE in terms of signals is equivalent to a tensor operation is that it will allow for model reduction to be carried out entirely using tensor techniques. This provides more flexibility for the model reduction framework as will be demonstrated in the next section.

In this section the focus is on obtaining the equivalence of the diagram (11). For this, we will first demonstrate how to associate a tensor with a signal. To be able to define a tensor operator associated with a PDE, we introduce shift operators acting on tensors. The last subsection contains the proof that this tensor operator is indeed equivalent to an ordinary PDE in terms of signals.

A. Associate a tensor with a signal

Throughout the remainder of this paper we will assume that the domain of the PDE (10) has a Cartesian structure.This means we assume that $\mathbb{D} = \mathbb{X}_1 \times \cdots \times \mathbb{X}_N$. Define $\mathcal{X}_k = \mathbb{R}^{L_k}$ with standard inner product and $\{e_k^{(\ell_k)}, \ell_k = 1, \ldots, L_k\}$ an orthonormal basis for $\mathcal{X}_k$, for $1 \leq k \leq N$. Suppose we have a solution trajectory of the PDE, which will be denoted by $\underline{w}(x_1, \ldots, x_N)$. We will associate a tensor $\mathcal{W} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times W \to \mathbb{R}$ with $\underline{w}(x_1, \ldots, x_N)$. $\mathcal{W}$ can be defined as follos

$$\mathcal{W} := \sum_{\ell_1=1}^{\infty} \cdots \sum_{\ell_N=1}^{\infty} \sum_{\ell_{N+1}=1}^{n} w_{\ell_1 \cdots \ell_{N+1}} e_1^{(\ell_1)} \otimes \cdots \otimes e_{N+1}^{(\ell_{N+1})} \tag{12}$$

with coefficients $w_{\ell_1 \cdots \ell_{N+1}}$

$$w_{\ell_1 \cdots \ell_{N+1}} := \langle e_{N+1}^{(\ell_{N+1})}, \cdots \langle e_1^{(\ell_1)}, \underline{w}(x_1, \ldots, x_N) \rangle_1 \cdots \rangle_{N+1} \tag{13}$$

with $1 \leq \ell_k \leq L_k$ for $k = 1, \ldots, N$ and $1 \leq \ell_{N+1} \leq n$. Eq. (12) will be the starting point to define signal approximations and compute POD basis functions.

In the remainder of this paper, the tensor $\mathcal{W}$ will be used to define spectral expansions and a reduced model. However, at some point in the model reduction process it may be interesting to derive the physical signal that is associated with a tensor. This is explained here. Consider a tensor $\mathcal{W} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times W \to \mathbb{R}$, the physical signal $\underline{w}(x_1, \ldots, x_N)$ associated with it is derived as follows.

When associating a tensor with the physical signal $\underline{w}(x_1, \ldots, x_N)$ a link was made between the point in the domain $x_k = x_{p_k}$ and the basis vector $e_k^{(p_k)}$. Similarly, the k-th component of $\underline{w}(x_1, \ldots, x_N)$, $w_k(x_1, \ldots, x_N)$, was linked to the k-th basis vector of W. The same association is used to go back from the tensor to the physical signal. The k-th component of $\underline{w}(x_1, \ldots, x_N)$ at position $(x_{p_1}, \ldots, x_{p_N})$ of the discrete spatial and temporal domain is defined by a tensor evaluation as follows

$$w_k(x_{p_1}, \ldots, x_{p_N}) := \mathcal{W}(e_1^{(p_1)}, \ldots, e_N^{(p_N)}, e_{N+1}^{(k)}).$$

This can be generalized to obtain the full vector as follows

$$\begin{aligned}\underline{w}(x_{p_1}, \ldots, x_{p_N}) &:= \sum_{\ell_1=1}^{L_1} \cdots \sum_{\ell_{N+1}=1}^{n} w_{\ell_1 \cdots \ell_{N+1}} \langle e_1^{(\ell_1)}, e_1^{(p_1)} \rangle \cdots \\ &\quad \langle e_N^{(\ell_N)}, e_N^{(p_N)} \rangle \begin{bmatrix} \langle e_{N+1}^{(\ell_{N+1})}, e_{N+1}^{(1)} \rangle \\ \vdots \\ \langle e_{N+1}^{(\ell_{N+1})}, e_{N+1}^{(n)} \rangle \end{bmatrix} \\ &= \begin{bmatrix} w_{p_1 \cdots p_N 1} \\ \vdots \\ w_{p_1 \cdots p_N n} \end{bmatrix}. \end{aligned} \tag{14}$$

B. Definition of shift for tensors

This section introduces the shift operation for tensors and gives a toy-example of the shift applied both to order-2 tensors and 2-D signals.

Definition IV.1 Consider the rank-one tensor $U = e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{\ell_N}$. The *shift operation in the kth mode*, $\mathfrak{S}_k$, is defined as

$$\begin{aligned}&\mathfrak{S}_k(e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{(\ell_N)}) = \\ &e_1^{(\ell_1)} \otimes \cdots \otimes e_{k-1}^{(\ell_{k-1})} \otimes e_k^{(\ell_k+1)} \otimes e_{k+1}^{(\ell_{k+1})} \otimes \cdots \otimes e_N^{(\ell_N)}.\end{aligned} \tag{15}$$

This operator is *non-cyclic*, i.e.

$$\mathfrak{S}_k(e_1^{(\ell_1)} \otimes \cdots \otimes e_{k-1}^{\ell_{k-1}} \otimes e_k^{(L_k)} \otimes e_{k+1}^{(\ell_{k+1})} \otimes \cdots \otimes e_N^{(\ell_N)}) = 0. \tag{16}$$

The *cyclic shift operation in the kth mode*, $\mathfrak{S}_k^\circ$, is defined as in (15), however

$$\begin{aligned}&\mathfrak{S}_k^\circ(e_1^{(\ell_1)} \otimes \cdots \otimes e_{k-1}^{(\ell_{k-1})} \otimes e_k^{(L_k)} \otimes e_{k+1}^{(\ell_{k+1})} \otimes \cdots \otimes e_N^{(\ell_N)}) \\ &= e_1^{(\ell_1)} \otimes \cdots \otimes e_{k-1}^{(\ell_{k-1})} \otimes e_k^{(1)} \otimes e_{k+1}^{(\ell_{k+1})} \otimes \cdots \otimes e_N^{(\ell_N)}.\end{aligned} \tag{17}$$

Since tensors are multi-linear functionals, it is immediate that for any $T \in \mathcal{T}_N$ $\mathfrak{S}_k \mathfrak{S}_m(T) = \mathfrak{S}_m \mathfrak{S}_k(T)$. This also defines $\mathfrak{S}_k^\tau$ to be

$$\begin{aligned}&\mathfrak{S}_k^\tau(e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{(\ell_N)}) = \\ &e_1^{(\ell_1)} \otimes \cdots \otimes e_{k-1}^{(\ell_{k-1})} \otimes e_k^{(\ell_k+\tau)} \otimes e_{k+1}^{(\ell_{k+1})} \otimes \cdots \otimes e_N^{(\ell_N)}\end{aligned} \tag{18}$$

i.e. operation of $\mathfrak{S}_k^\tau$ is equivalent to τ operations by $\mathfrak{S}_k$.

For a rank-one tensor $T = \varphi_1 \otimes \cdots \otimes \varphi_N$ not defined with respect to the standard bases, shift operators can only be applied after T is converted to the standard bases according to

$$T = \sum_{\ell_1=1}^{L_1} \cdots \sum_{\ell_N=1}^{L_N} t_{\ell_1 \cdots \ell_N} e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{(\ell_N)}$$

with coefficients

$$t_{\ell_1 \cdots \ell_N} = \prod_{k=1}^{N} \langle \varphi_k, e_k^{(\ell_k)} \rangle$$

Example IV.2 Consider

$$w_1(i,j) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

The ij-th entry of w_1 denotes the signal value at position x_i and time t_j. Define $w_2(i,j) = \varsigma_1(w_1(i,j))$ and $w_3(i,j) = \varsigma_1^\circ(w_1(i,j))$. $w_2(i,j)$ and $w_3(i,j)$ are given by the following matrices

$$w_2(i,j) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix};\ w_3(i,j) = \begin{pmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

We will now associate a tensor W_1 with w_1 and show that application of $\mathfrak{S}_1$ and $\mathfrak{S}_1^\circ$ to W_1 yields tensors W_2, W_3 that correspond to w_2 and w_3. W_1 can be represented with respect to the standard bases as follows

$$W_1 = \sum_{\ell_1=1}^{3} \sum_{\ell_2=1}^{3} w_{\ell_1\ell_2} e_1^{(\ell_1)} \otimes e_2^{(\ell_2)}$$

where $w_{\ell_1\ell_2}$ is the (ℓ_1, ℓ_2)-th entry of the matrix representing w_1. We can define two more tensors, $W_2 := \mathfrak{S}_1(W_1)$ and $W_3 := \mathfrak{S}_1^\circ(W_1)$. W_2 has the following expression

$$W_2 = e_1^{(2)} \otimes e_2^{(1)} + 2e_1^{(2)} \otimes e_2^{(2)} + 3e_1^{(2)} \otimes e_2^{(3)} \\ 4e_1^{(3)} \otimes e_2^{(1)} + 5e_1^{(3)} \otimes e_2^{(2)} + 6e_1^{(3)} \otimes e_2^{(3)}. \quad (19)$$

W_2 is a tensor with respect to the standard bases given by

$$W_2 = \sum_{\ell_1} \sum_{\ell_2} w_{\ell_1\ell_2}^{(2)} e_1^{(\ell_1)} \otimes e_2^{(\ell_2)}.$$

The coefficients of W_2 are given by

$$[[w_{\ell_1\ell_2}^{(2)}]] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

which is exactly what w_2 was. Similarly to Eq. (19), an expression can be formulated to describe W_3. It can be shown that W_3 is a tensor with respect to the standard bases given by

$$W_3 = \sum_{\ell_1} \sum_{\ell_2} w_{\ell_1\ell_2}^{(3)} e_1^{(\ell_1)} \otimes e_2^{(\ell_2)}.$$

The coefficients of W_3 are given by

$$[[w_{\ell_1\ell_2}^{(3)}]] = \begin{pmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

which corresponds to the matrix representation of w_3.

C. Equivalence

Now that the shift operation is defined, we can show that the arrows in (11) make sense. Specifically, we will show in the following theorem that the residual $\underline{r}(x_1, \ldots, x_N)$ obtained by operating the PDE on $\underline{w}(x_1, \ldots, x_N)$ is equivalent to the tensor R which is obtained by applying the tensor operator associated with the PDE to the tensor W associated with $\underline{w}$. Note that the definition (18) provides a well-defined tensor operator $D(\mathfrak{S}_1, \ldots, \mathfrak{S}_N)$.

Theorem IV.3 *Let* $\underline{r}(x_1, \ldots, x_N) := D(\varsigma_1, \ldots, \varsigma_N)\underline{w}$ *be the residual obtained by operating* (10) *on* $\underline{w}$*. Furthermore, let* $R \in \mathcal{T}_N$ *be the tensor associated with* $\underline{r}$*. The tensor associated with* $\underline{w}$ *will be denoted by* W *and let* $\hat{R} := D(\mathfrak{S}_1, \ldots, \mathfrak{S}_N)W$ *be the result of applying the tensor-PDE* $D(\mathfrak{S}_1, \ldots, \mathfrak{S}_N)$ *to* W*. We have*

$$R = \hat{R} \quad (20)$$

whenever $D(\mathfrak{S}_1, \ldots, \mathfrak{S}_N)$ *is equivalent to* $D(\varsigma_1, \ldots, \varsigma_N)$*.*

Proof: The operator $D(\varsigma_1, \ldots, \varsigma_N)$ can be characterized in general as

$$D = \sum_{|k|=0}^{K} d_{k_1,\ldots,k_N} \varsigma_1^{k_1} \cdots \varsigma_N^{k_N}$$

where $|k| = \sum k_i$ and K is the degree of the operator. Since we only consider linear PDEs, it suffices to take an arbitrary term of the sum describing D and show that the theorem holds for this term. Furthermore, also due to linearity it is sufficient to only consider scalar $\underline{w}$. We will consider the operator

$$d_{k_1\cdots k_N} \varsigma_1^{k_1} \cdots \varsigma_N^{k_N}$$

where $\sum k_i \leq K$ and $k_p \leq L_p$ for $p = 1, \ldots, N$. We will first apply this operator to $w(x_1, \ldots, x_N)$

$$\left(d_{k_1\cdots k_N} \varsigma_1^{k_1} \cdots \varsigma_N^{k_N} w\right)(x_1, \ldots, x_N) := \\ d_{k_1\ldots k_N} w(x_{1+k_1}, \ldots, x_{N+k_N}) \\ = r(x_1, \ldots, x_N).$$

$w(x_1, \ldots, x_N)$ defines a tensor W

$$W = \sum_{\ell_1=1}^{L_1} \cdots \sum_{\ell_N=1}^{L_N} w_{\ell_1\cdots\ell_N} e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{(\ell_N)}$$

with coefficients $w_{\ell_1\cdots\ell_N}$ defined as

$$w_{\ell_1\cdots\ell_N} = \langle e_N^{(\ell_N)}, \cdots \langle e_1^{(\ell_1)}, w\rangle_1 \cdots \rangle_N.$$

$r(x_1, \ldots, x_N)$ defines a tensor R

$$R = \sum_{\ell_1=1}^{L_1} \cdots \sum_{\ell_N=1}^{L_N} r_{\ell_1\cdots\ell_N} e_1^{(\ell_1)} \otimes \cdots \otimes e_N^{(\ell_N)}$$

with coefficients

$$r_{\ell_1\cdots\ell_N} = \langle e_N^{(\ell_N)}, \cdots \langle e_1^{(\ell_1)}, r\rangle_1 \cdots \rangle_N \\ = \langle e_N^{(\ell_N+k_N)}, \cdots \langle e_1^{(\ell_1+k_1)}, d_{k_1\cdots k_N} \\ w(x_{1+k_1}, \ldots, x_{N+k_N})\rangle_1 \cdots \rangle_N.$$

Conversely, translating $d_{k_1\cdots k_N} \varsigma_1^{k_1} \cdots \varsigma_N^{k_N}$ to the tensor domain gives $d_{k_1\cdots k_N} \mathfrak{S}_1^{k_1} \cdots \mathfrak{S}_N^{k_N}$. Let

$$\hat{R} = d_{k_1\cdots k_N} \mathfrak{S}_1^{k_1} \cdots \mathfrak{S}_N^{k_N} W \\ = \sum_{\ell_1=1}^{L_1} \cdots \sum_{\ell_N=1}^{L_N} d_{k_1\cdots k_N} w_{(\ell_1+k_1)\cdots(\ell_N+k_N)} \\ e_1^{(\ell_1+k_1)} \otimes \cdots \otimes e_N^{(\ell_N+k_N)}$$

Using the definition of the coefficients

$$d_{k_1\cdots k_N} w_{(\ell_1+k_1)\cdots(\ell_N+k_N)} =$$
$$d_{k_1\cdots k_N} \langle e_N^{(\ell_N+k_N)}, \cdots \langle e_1^{(\ell_1+k_1)}, w(x_{1+k_1}, \ldots, x_{N+k_N})\rangle_1 \cdots \rangle_N$$

which is exactly the expression derived for the coefficients of R. ■

The theorem shows that the tensor operator associated with a PDE is indeed equivalent to the original PDE in that it yields the same residuals. This allows us to do model reduction in the tensor domain entirely, as will be shown in the next section.

V. Model reduction using Galerkin projections

Using the material introduced so far, this section describes the proposed model reduction strategy. It is an adapted form of the POD method discussed in Sec. III. As mentioned in the introduction of this paper, the model reduction strategy is entirely based on the use of tensors instead of signals. This implies that we will associate a tensor $\mathcal{W}$ with the signal $\underline{w}$ and carry out the whole procedure using $\mathcal{W}$ and the tensor interpretation of the PDE (10). This section is organized as follows. First, an approximation $\mathcal{W}_r$ to $\mathcal{W}$ is introduced. Then, given a set of measured or simulated data, POD basis functions can be computed. In the second subsection, a Galerkin projection based on tensors is used to define the reduced model.

A. Signal approximation and POD basis computation

In Section III a truncated spectral expansion was used to approximate $\underline{w}$ by $\underline{w}_r$, see Eq. 8. Here, a tensor $\mathcal{W}$ is associated with $\underline{w}$ as demonstrated in Sec. IV-A. $\mathcal{W}$ is a tensor $\mathcal{W} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times W \to \mathbb{R}$, see Eq. (12). Given basis functions $\{\varphi_k^{(\ell_k)}(x_k),\ \ell_k = 1, \ldots, r_k\}$, $k = 1, \ldots, N+1$, we define approximations of $\mathcal{W}$, denoted by $\mathcal{W}_r$ as

$$\mathcal{W}_r = \sum_{\ell_1=1}^{r_1} \cdots \sum_{\ell_N=1}^{r_N} \sum_{\ell_{N+1}=1}^{r_{N+1}} \tilde{w}_{\ell_1\cdots\ell_{N+1}} \varphi_1^{(\ell_1)} \otimes \cdots \otimes \varphi_{N+1}^{(\ell_{N+1})} \tag{21}$$

It may not be desirable to approximate time and the dependent variables. Therefore, the spaces corresponding to time and the dependent variables may be kept intact, by choosing $r_N = L_N$ and $r_{N+1} = n$. $\mathcal{W}_r$ is the tensor equivalent to (8).

As mentioned in Section III, where the concept of POD was introduced, the quality of reduced order models depends entirely on the quality of the basis functions used to construct the reduced order model. This observation remains true when using POD models based on tensor decompositions. Assume a set of measured or simulated data of a solution trajectory of the PDE (10) is available and will be denoted by $\underline{w}(x_1, \ldots, x_N)$. As above, we associate a tensor $\mathcal{W} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times W \to \mathbb{R}$ with $\underline{w}(x_1, \ldots, x_N)$. We will define POD basis functions through approximations of the tensor $\mathcal{W}$.

Definition V.1 The vector of integers $r = (r_1, \ldots, r_{N+1})$ with $r_k \le L_k$ is said to achieve a relative approximation error $\epsilon > 0$ if

$$\frac{\|\mathcal{W} - \mathcal{W}_r\|}{\|\mathcal{W}\|} \le \epsilon \tag{22}$$

In that case, we say that the basis functions $\{\varphi_k^{(1)}, \ldots, \varphi_k^{(r_n)}\}$ for $k = 1, \ldots, N+1$ constitute a POD basis for the model (10) derived from the tensor W.

B. Model reduction through Galerkin projections

So far, we have derived the tensor equivalent of a spectral expansion and shown how POD basis functions can be computed. In this section, we present a Galerkin projection that matches the concepts from these previous sections and can be used to derive a reduced model. The PDE (10) is translated into a tensor-PDE and the reduced model is defined in terms of tensors.

Given the original PDE, (10), three types of solutions can be defined.

Definition V.2 A *strong solution* $\mathcal{W}$ of (10) satisfies

$$D(\mathfrak{S}_1, \ldots \mathfrak{S}_N)\mathcal{W} = 0. \tag{23}$$

Definition V.3 A *weak solution* $\mathcal{W}$ of (10) satisfies

$$\langle D(\mathfrak{S}_1, \ldots \mathfrak{S}_N)\mathcal{W}, \Phi\rangle = 0. \tag{24}$$

for all $\Phi \in \mathcal{T}_N$.

Definition V.4 A *approximate weak solution* $\mathcal{W}_r$ of (10) satisfies

$$\langle D(\mathfrak{S}_1, \ldots \mathfrak{S}_N)\mathcal{W}_r, \Phi\rangle = 0. \tag{25}$$

for all $\Phi \in \mathcal{F}$, where $\mathcal{F} \subset \mathcal{T}_N$.

The reduced model is now defined as the set of tensors that form the approximate weak solutions of (10).

VI. Conclusions and future works

This paper deals with model reduction for systems that prescribe the evolution of a set of physical variables, referred to as dependent variables, over a domain of independent variables. Proper Orthogonal Decompositions (POD) is a method that is commonly used for model reduction of these types of systems. In this paper we have adapted the POD method to explicitly take multiple dependent and independent variables into account while maintaining the structure of the original system. Main tool to achieve this was to use multilinear functionals, tensors.

The proposed model reduction framework is based on tensor analysis and manipulations. In the first step of the methodology, a tensor is associated with a physical signal. Then, this tensor is approximated and used to compute empirical projection spaces. Finally, a tensorial Galerkin projection prescribes the reduced model. To show that this approach is feasible, we introduced a tensor operation which is equivalent to a partial difference equation. One of the main results of this paper shows that this tensor operation is indeed equivalent to a traditional PDE in terms of signals.

REFERENCES

[1] M. Kirby. *Geometric data analysis*. John Wiley, New York, 2001.
[2] V. Thomee. *Galerkin finite element methods for parabolic problems*. Springer, Berlin, 1997.
[3] S. Weiland and F. van Belzen. Singular Value Decompositions and Low-Rank Approximations of Tensors. *IEEE Trans. Signal Proc.*, 58(3):1171–1182, 2010.
[4] L. de Lathauwer et al. A Multilinear Singular Value Decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.
[5] K. Kunisch and S. Volkwein, "Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics," SIAM J. Numer. Anal., 40:492–515, 2002.
[6] T.G. Kolda, "Orthogonal tensor decompositions," SIAM J. Matrix Analysis and Applications, Vol. 23, No. 1, pp. 243-255, 2001.
[7] T. Zhang and G.H. Golub, "Rank one approximation to high order tensors," SIAM J. on Matrix Analysis and Applications, Vol. 23, No. 2, pp. 534-550, 2001.