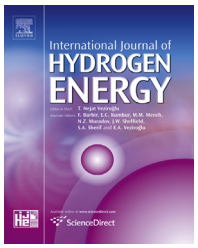
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



A nonlinear principal component analysis approach for turbulent combustion composition space

Hessam Mirgolbabaei [a], *Tarek Echekki* [a,*], *Nejib Smaoui* [b]

[a] *Department of Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC, USA*
[b] *Department of Mathematics, Kuwait University, Kuwait*

ARTICLE INFO

Article history:
Received 12 October 2013
Received in revised form
26 December 2013
Accepted 28 December 2013
Available online 31 January 2014

Keywords:
Turbulent combustion
Nonlinear principal component analysis
Autoignition

ABSTRACT

An approach for the determination of principal components using nonlinear principal component analysis (NLPCA) is proposed in the context of turbulent combustion. NLPCA addresses complex data sets where the contours of the inherent principal directions are curved in the original manifold. Thermo-chemical scalars' statistics are reconstructed from the optimally derived moments. The tabulation of the scalars is then implemented, using artificial neural networks (ANN). The approach is implemented on numerical data generated for the stand-alone one-dimensional turbulence (ODT) simulation of hydrogen autoignition in a turbulent jet with preheated air. It is found that 2 nonlinear principal components are sufficient to capture thermo-chemical scalars' profiles. For some of the scalars, a single principal component reasonably captures the scalars' profiles as well.

1. Introduction

Traditional modeling approaches to address turbulence-chemistry interactions in combustion have largely been based on the transport of representative moments from which thermo-chemical scalars' statistics can be deduced. The choice of these moments can be influenced by a set of factors including the combustion mode (e.g. premixed vs. non-premixed), combustion regime (e.g. flamelet vs. distributed reaction) or the competition of the various time scales (e.g. turbulent mixing vs. chemistry). Optimal strategies based on principal component analysis (PCA) have been proposed in recent years to determine moments that best represent the composition space for different problems [1–7]. Classical PCA addresses primarily the potential co-linearity between the variables. Therefore, it may or may not be adequate for complex data, where the dependencies are nonlinear.

As a nonlinear alternative, nonlinear PCA (NLPCA) [8] is adapted in the present work, using bottleneck neural networks (BNN). The method has been popular in numerous applications (see for example, Refs. [9–14]), including combustion and turbulence, two topics of interest in the present study (see Refs. [13,14]). NLPCA performs an identity mapping where the output of the network is the same vector as the input vector. Between the input and output, the *encoding* and *decoding* are performed using 2 hidden layers and a principal component(s) layer. The output of the component(s) layer represents the principal components (PCs).

* *Corresponding author.*
E-mail address: techekk@ncsu.edu (T. Echekki).
0360-3199/$ – see front matter
http://dx.doi.org/10.1016/j.ijhydene.2013.12.195

Within the context of combustion, the authors have recently investigated the use of Kernel PCA as a nonlinear extension of PCA [15]. The KPCA approach is based on a mapping of the original data into a higher dimensional space, called *feature space*. This mapping improves the linear dependencies in feature space where a PCA is implemented.

In the present study, we attempt to determine if NLPCA can adequately represent the composition space of turbulent combustion. This is a first step towards implementing a method for systematic moment formulation for problems the combustion system cannot be represented using linear manifolds. Although, NLPCA has been demonstrated as a robust method to approximate nonlinear manifolds of complex data, it is useful to explore whether it can effectively do so for turbulent combustion data in composition space. This data may be represented by thermo-chemical scalars that play different roles in the various combustion zones or during the transient evolution of a combustion process, such as during ignition or extinction events. Nonlinear principal components (NLPCs) or simply PCs are derived based on data from stand-alone one-dimensional turbulence (ODT) simulations of hydrogen autoignition in a turbulent jet with preheated air [16]. This is a flow configuration that exhibits complex interactions between mixing and chemistry. Moreover, it also exhibits interesting differential diffusion effects of the fuel relative to other species in the reacting mixture. The configuration is characterized by different regimes of combustion starting with induction followed by autoignition, and the formation of a non-premixed flame [16]. The composition space is reconstructed using artificial neural networks (ANN) [17]. The tabulated and the original data are compared to investigate the validity of the proposed approach.

2. Transported equations for NLPCs

The NLPCA analysis starts with detailed experimental and computational data of thermo-chemical scalars' vector, such as what is adopted here based on temperature and composition $\boldsymbol{\theta} = (T, Y_1, \cdots, Y_{N-1})^T$ where T is the mixture temperature and the Y's are the species mass fractions in a multi-component mixture of $N - 1$ components. The original thermo-chemical scalars are subject to a set of transport equations:

$$\frac{D\boldsymbol{\theta}}{Dt} = \nabla\cdot\mathbf{J}_{\boldsymbol{\theta}} + \mathbf{S}_{\boldsymbol{\theta}} \tag{1}$$

where $\mathbf{J}_{\boldsymbol{\theta}}$ and $\mathbf{S}_{\boldsymbol{\theta}}$ correspond to the diffusive fluxes and the chemical source terms for the original thermo-chemical scalars' vector, $\boldsymbol{\theta}$, respectively. NLPCA analysis results in a smaller vector of principal components: $\boldsymbol{\Psi} = (\Psi_1, \Psi_2, \cdots, \Psi_{N_{PC}})$ where the Ψ's are the principal components and N_{PC} is the dimension of the principal components' vector. We expect N_{PC} to be much smaller than N for the reduction process to be meaningful. Similarly, we model the transport equations for the PCs as follows:

$$\frac{D\boldsymbol{\psi}}{Dt} = \nabla\cdot\mathbf{J}_{\boldsymbol{\psi}} + \mathbf{S}_{\boldsymbol{\psi}} \tag{2}$$

where $\mathbf{J}_{\boldsymbol{\Psi}}$ and $\mathbf{S}_{\boldsymbol{\Psi}}$ correspond to the diffusive fluxes and the chemical source terms for the PCs, respectively. Reconstruction of these terms has been already discussed in details for linear and Kernel PCA, that can be easily generalized to any types of PCA [15]. As discussed in Ref. [15], for NLPCA also, these terms can be reconstructed using an analogy similar to Linear PCA.

Based on the analysis presented in Ref. [15], we defined Jacobian matrices for nonlinear PCs, which are expressed for full set of PCs and the reduced set of PCs as:

$$\widehat{\mathbf{Q}}^T \equiv \frac{\partial\boldsymbol{\psi}}{\partial\boldsymbol{\theta}} \text{ and } \widehat{\mathbf{A}}^T \equiv \frac{\partial\boldsymbol{\psi}^{red}}{\partial\boldsymbol{\theta}}. \tag{3}$$

Matrix $\widehat{\mathbf{Q}}$ as having a dimension N by N (i.e. a full matrix). It may be evaluated numerically using small perturbations of the thermo-chemical scalars vector. Matrix $\widehat{\mathbf{A}}$ is a truncation corresponding to the first N_{PC} rows of $\widehat{\mathbf{Q}}$. From these definitions, the source terms for the PCs may be expressed as follows: $\mathbf{S}_{\boldsymbol{\psi}} = \widehat{\mathbf{Q}}^T\ \mathbf{S}_{\boldsymbol{\theta}}$ [15]; while the diffusion flux term in Eq. (2) is expressed in terms of an effective diffusion coefficient and PCs gradients as: $\mathbf{J}_{\boldsymbol{\psi}} = \rho\mathbf{D}_{\boldsymbol{\psi}}\nabla\boldsymbol{\psi}$ where $\mathbf{D}_{\boldsymbol{\psi}} = \widehat{\mathbf{Q}}^T\mathbf{D}_{\boldsymbol{\theta}}(\widehat{\mathbf{Q}}^T)^{-1}$ and $\mathbf{J}_{\boldsymbol{\theta}} = \rho\mathbf{D}_{\boldsymbol{\theta}}\nabla\boldsymbol{\theta}$ [15]. According to this definition, $\mathbf{D}_{\boldsymbol{\theta}}$ and $\mathbf{D}_{\boldsymbol{\psi}}$ correspond to the diagonal matrix of diffusion coefficients for the original thermo-chemical scalars and the PCs, respectively. The clear advantage of the proposed formulation for the source and the molecular transport terms in Ref. [15] is that both the source term and the diffusion coefficients for the transported PCs also can be tabulated. A potential shortcoming of the definition for the diffusion coefficient for the PC is that the matrix does not have to be symmetric or contain all positive diffusion coefficients, a potential risk for numerical instability with counter-gradient diffusion [15]. However, we have been exploring the potential for rotating the PCs (an in turn, rotating the diffusion coefficients) to reduce the magnitude of the off-diagonal terms of the diffusion coefficient matrix and maintain positive diffusion coefficients along the diagonal where diffusion effects are expected to be important. Results of this work are not shown here.

Having defined a pathway to implement a transport equation for principal components, it is important to explore the effectiveness of a nonlinear PCA reduction. This effort is described here. A follow-up study will be carried out to implement, *a posteriori*, the proposed procedure by transporting the PCs.

3. Bottleneck neural network-PCA

NLPCA is motivated by Cover's theorem, which demonstrates that a nonlinear data structure in the input space is more likely to be linear after high-dimensional nonlinear mapping [8]. The same motivation of mapping onto a higher dimensional space is used in a similar approach based on Kernel PCA [18]. In Kernel PCA, the original manifold is mapped into a higher dimensional space, called the *feature space*, where the classical PCA is performed. In a general topology of BNN, the network consists of two parts: compression or encoding and reconstruction or decoding. In the compression part, the original dimensionality is nonlinearly mapped into components' manifold (i.e. the PCs), through a first hidden layer