Journal of Computational Physics 325 (2016) 380–394

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

An evaluation of noise reduction algorithms for particle-based fluid simulations in multi-scale applications

M.J. Zimoń [a,b,*], R. Prosser [a], D.R. Emerson [b], M.K. Borg [c], D.J. Bray [b], L. Grinberg [d], J.M. Reese [c]

[a] School of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Manchester M13 9PL, UK
[b] Scientific Computing Department, STFC Daresbury Laboratory, Warrington WA4 4AD, UK
[c] School of Engineering, The University of Edinburgh, Edinburgh EH9 3FB, UK
[d] IBM T.J. Watson Research Center, 1 Rogers St, Cambridge, MA 02142, United States

ARTICLE INFO

Article history:
Received 18 May 2016
Accepted 21 August 2016
Available online 26 August 2016

Keywords:
Noise reduction
Molecular dynamics
Windowed proper orthogonal decomposition
Wavelet thresholding
Singular spectrum analysis
Empirical mode decomposition

ABSTRACT

Filtering of particle-based simulation data can lead to reduced computational costs and enable more efficient information transfer in multi-scale modelling. This paper compares the effectiveness of various signal processing methods to reduce numerical noise and capture the structures of nano-flow systems. In addition, a novel combination of these algorithms is introduced, showing the potential of hybrid strategies to improve further the de-noising performance for time-dependent measurements. The methods were tested on velocity and density fields, obtained from simulations performed with molecular dynamics and dissipative particle dynamics. Comparisons between the algorithms are given in terms of performance, quality of the results and sensitivity to the choice of input parameters. The results provide useful insights on strategies for the analysis of particle-based data and the reduction of computational costs in obtaining ensemble solutions.

1. Introduction

Numerical simulation is an essential tool for gaining a better understanding of many physical phenomena that can be difficult to describe with analytical methods or experimental studies. The statistical mechanics of complex systems is often analysed with molecular dynamics (MD) [1], Monte Carlo methods, e.g. direct simulation Monte Carlo (DSMC) [2] or dissipative particle dynamics (DPD) [3]; a comprehensive summary of all the modelling strategies can be found in Karniadakis et al. [4]. These procedures can be used to resolve accurately the dynamics at atomistic, meso- and micro-scales and are widely used to simulate nano/micro flows confined in channels such as carbon nanotubes [5,6]. In addition, information obtained from molecular simulations forms the basis of new and emerging hybrid multi-scale modelling methods for physical and biological applications (see [7] for a review). Examples demonstrating the ubiquity of multi-scale, multi-physics applications include the dynamics of complex fluid flows [8], the classical turbulence problem [9], meteorological predictions [10], chemical and biological reactions [11]. Moreover, there is significant potential to apply multi-scale techniques to sociological problems, such as crowd and traffic flow [12].

* Corresponding author at: Scientific Computing Department, STFC Daresbury Laboratory, Warrington WA4 4AD, UK. Current address: IBM Research, Daresbury Laboratory, Warrington WA4 4AD, UK.
E-mail address: malgorzata.zimon@uk.ibm.com (M.J. Zimoń).

http://dx.doi.org/10.1016/j.jcp.2016.08.021
0021-9991/

The central problems with all particle-based and multi-scale modelling are the computational cost of the simulations and the accurate measurement and transfer of information across disparate length and time scales; there currently exist many sources of uncertainty and noise disturbing this intra-scale transfer, with an associated loss of simulation fidelity. Circumventing this problem often requires large samples and long averaging periods, resulting in computationally expensive calculations.

The objective of this paper is to investigate the capabilities of a number of mathematical algorithms introduced in the literature to assist noise reduction in particle-based modelling. A number of benchmark fluid flow problems, performed with molecular dynamics and dissipative particle dynamics, are used to investigate the usefulness of the considered methods and provide guidelines on how the algorithms can be successfully applied. The main focus of this paper is on novel procedures that provide rapid, adaptive, noise-free coarse-graining of micro-scale phenomena, and can further be employed in molecular-continuum simulations.

In this paper, new algorithms are proposed that combine the strengths of proper orthogonal decomposition (POD) with other techniques, hereafter referred to as POD+ methods, in order to achieve better efficiency in processing time-dependent fields. This work directly tackles the important challenge of extracting information from the data without significant additional computational cost.

The paper is organised as follows: the basic theory for the methods is described in Sec. 2. A comparison of the performance of each technique in de-noising particle-based simulations is presented in Sec. 3, followed by remarks and recommendations for the use of the methods under investigation.

2. Theoretical background

In the following section we briefly review the numerical methods employed. First, we discuss algorithms based on singular value decomposition (or eigenvalue decomposition) and QR decomposition. The second part of the review focuses on strategies using wavelet transforms, wavelet thresholding and the WienerChop filter [13]. We also discuss the application of empirical mode decomposition to noise reduction. Novel couplings of proper orthogonal decomposition to these algorithms are introduced at the end of the Section.

2.1. Noise filtering with singular value decomposition and QR factorisation

2.1.1. Proper orthogonal decomposition

Define an element $A(\tau^s, x)$ of the real $N \times M$ matrix as a measurement from the x-th probe taken at the τ^s-th time instant. Proper orthogonal decomposition can be done either by eigenvalue decomposition (EVD) of the symmetric matrix $C = AA^\dagger$ ($A^\dagger A$ if $N > M$),[1] or by singular value decomposition of A:

$$A = U\Sigma V^\dagger, \tag{1}$$

where, in the case of full SVD, U is an $N \times N$ orthogonal matrix, V is an $M \times M$ orthogonal matrix, and Σ is an $N \times M$ diagonal matrix. Columns of U and V are left and right singular vectors, respectively. The diagonal entries of Σ, called singular values, are the square roots of eigenvalues, $s_n = \sqrt{\lambda_n}$ for $s_1 \geq s_2 \geq \ldots \geq s_n \geq 0$, where λ_n is the n-th eigenvalue of the diagonal matrix.

If A is a collection of measurements corrupted by additive noise, it can be represented in the form $A = A_t + B$, where A_t is a matrix that contains the *true* signals, and B denotes the noise. Given the decomposition in Eq. (1), the rank-k approximation of A can be written in vector form as

$$A_k(\tau^s, x) = \sum_{n=1}^{k} s_n u_n v_n^\dagger, \tag{2}$$

where $1 \leq k \leq min(N, M)$, u_n and v_n are the orthonormal (temporal or spatial) POD modes corresponding to the n-th columns of the matrices U and V, respectively. The key property of POD is that A_k is optimal in the sense that $min\left\{\|A_t - A_k\|_F^2\right\} = \sum_{n=k+1}^{min(N,M)} s_n^2$, where $\|.\|_F$ is the Frobenius norm. In this paper, the rank k is referred to as the number of *dominant* modes. A practical implementation of POD based on time-windows (WPOD) described by Grinberg [14], can be used for particle-based simulations.

The main challenge in estimation of A_t is the choice of the truncation parameter. One possible approach for defining the rank k is to select a cumulative percentage of total variation which modes should contribute. Unfortunately, such cut-off is often insufficient, as important aspects of the observables can be present in the direction of low variance modes. Therefore, in addition to studying the energy content of eigenvalues, most of the reported data is analysed here with the log-eigenvalue diagram (LEV), $\log_{10}(\lambda_n = s_n^2)$, based on the assumption that if higher singular vectors represent uncorrelated noise, then the corresponding λ_n should decay exponentially with increasing n [15]. We also considered the smoothness of the temporal

[1] The superscript † indicates matrix transpose since $A \in \mathbb{R}$.