45th AIAA Aerospace Sciences Meeting and Exhibit 8 - 11 Jan 2007, Reno, Nevada

Low Order Galerkin Models for the Actuated Flow Around 2-D Airfoils

Gilead Tadmor* and Maria D. Centuori†
Communication & Digital Signal Processing Center
Northeastern University, 440 Dana Research Building, Boston, MA 02115, U.S.A.

Bernd R. Noack‡, Mark Luchtenburg§ and Oliver Lehmann¶
Institute of Fluid Dynamics and Technical Acoustics
Berlin University of Technology Straße des 17. Juni, S-10623 Berlin, Germany

Marek Morzyński‖
Institute of Combustion Engines and Transportation
Poznań University of Technology, Ulica Piotrowo 3, 60-965 Poznań, Poland

The objective of this paper is twofold. One is to explore extensions of the generalized mean-field empirical Galerkin model, previously developed for wake instabilities[1] to singularly actuated 2D airfoils, including a high lift configuration and a single airfoil at a high angle of attack. We present a minimum order mean field model, explore the role of the mean field as a mediator between actuation and dominant instabilities, and illustrate the need for richer mode-sets, as operating conditions change and during sharp transients. The second objective is to develop computational tools for effective extraction of empirical modes, for such models. We present two such tools. The first is the utilization of temporal harmonic analysis to separate the empirical reference into sub-reference, each representing a single (time varying) coherent flow structures' category. The second is a two-step approximate POD procedure, intended to drastically reduce the computational cost of POD approximations of very long / rich reference ensembles. It is based on a first step reference partitioning and a second step global analysis, with an a priori guaranteed resolution level.

I. Introduction

Owing to their (optimally) high resolution and relative computational simplicity, *proper orthogonal decomposition* (POD) Galerkin models[2,3] offer an attractive framework for model based flow control. Their effective utilization, however, is still often hampered by persistent stumbling blocks. Those include a commonly poor capability to capture transient dynamics, difficulties to incorporate singular, unsteady actuation, and, when very large, very high dimensional reference ensembles are involved, the computational burden associated with POD analysis.

*Professor, CDSP Director
†Graduate student
‡Ass. Professor
§Graduate student
¶Student
‖Professor

This note contributes to an evolving toolkit, addressing these issues. Its objective is twofold. One is to explore extensions of the generalized mean-field empirical Galerkin model, previously developed for wake instabilities[1] to singularly actuated 2D airfoils. The second and subsequent objective concerns computational tools for effective extraction of empirical modes, for such models. We use two benchmark configurations to demonstrate these developments: The separated flow over a NACA 0012 airfoil at a 30^o angle of attack, and the flow over a high lift configuration. These two configurations and their control have been studied extensively e.g.[4–6] and references therein are merely a random selection from the voluminous literature and significant related work. It is therefore stressed that the role of the two benchmarks in the current presentation is primarily that of demonstration tools for the objectives mentioned above.

The common use of attractors as references for POD mode extraction is one impediment to transient representation. Such references do not provide information on changes in the mean-field and deformation in leading vortical structures. In[7,8] we suggested the interpolation between fixed structure / fixed (low) order, parameter varying models to account for the slow deformation in leading modes. (See[9] for a related approach.) We addressed the first issue with the introduction of a generalized mean-field model, in[1] (see also[1,10–13]), where a *shift mode* captures the mean-field interaction with the leading instability. Here we explore an extension of these approaches to the separated flows over 2D airfoils with a singular boundary actuator. In particular, we explore the challenging representation of singular boundary actuation with dedicated *actuation modes*,[14] and the role of the dynamic mean-field *shift mode* as a mediator between the actuation and the shedding instability.

A natural feature of the remedies outlined above is to include sufficiently rich natural and actuated transients in POD references. Such references, however, tend to blend effects of multiple spatial and temporal frequencies (hence distinct flow structures), to the detriment of effective low order representations. We suggest a spatio-temporal analysis where the reference is first partitioned as a sum of single frequency trajectories, and independent POD analysis of each of them. This resembles *dynamic phasors* representations[15–17], common in power engineering. The difference is that here each phasor is a distributed vector field.

The analysis of long, transient references, offsets the computational advantages of indirect POD analysis, using the method of snapshots[2,3]. Indeed, its computational feasibility may require an approximation scheme. The preliminary frequency separation is a costly addition to this computational burden. At the same time, the joint compression of long / rich references is a powerful tool that can lead to truly dramatic subsequent computational savings. For example, it is indeed dramatically cheaper to explore the structure of tunable mode sets, in the spirit of[7], when an a priori compressed reference is used, instead of the original reference. As an illustration, in an example used in this paper a dimension reduction to about 0.6% of the original maintained near perfect (>98%) resolution of the original data, immensely facilitating subsequent analysis. We therefore propose a two step, approximate POD procedure, to reduce computational complexity, *with a priori guaranteed error bounds*, similar to the original POD procedure, but at a small fraction of the original computational cost. It includes preliminary POD of temporal domain partitions, and application of the frequency separation algorithm to an already compressed, high resolution data set. Examples of very recent related work include a complementary spatial partitioning algorithm[18] and a POD-of-POD method[19]. Like the algorithm suggested here, the paper[19] presents a two step procedure, starting with PODs of a time-partitioned reference. The distinguishing aspect of the algorithm suggested here is that the global POD, in the second step, is applied to an ensemble whose statistics is guaranteed to approximate the original reference. It therefore provides guaranteed error bounds in the final approximation. This issue may become critical for the efficiency of the algorithm when modes with substantially different energy content are involved (as opposed to a single dominant mode pair per partition, as in[19]).

The paper is organized as follows. Section II briefly reviews numerical, dynamical and modeling aspects of the two benchmarks. Section III summarizes our finding concerning low dimensional approximations and modeling, in these benchmarks. We discuss the spatio-temporal approach to POD approximation, in Section IV. Finally, Section V reviews two variants of the proposed two step POD algorithm.

II. The Two Dimensional Airfoil Benchmarks

A. The High Lift Configuration Benchmark

The two-dimensional high lift configuration represents the swept constant chord half (SCCH) model, see fig. 1, used in several studies targeting flow control,[20,21] The main section angle of attack is 6^o, and the flap

and slat angles are as indicated in fig. 1. The Reynolds number is $Re = 10^6$. At these operating conditions, the flow over the slat and wing is fully attached. In contrast, the flow over the flap is separated, with typical periodic vortex shedding. Zero net flux actuation, perpendicular to the flap surface, at its leading edge, is used to attenuate this instability and to partially reattach the flow. The actuation frequency in the simulations discussed here is about twice the natural shedding frequency. In this case, attenuated shedding in the controlled attractor locks on the actuation frequency.

The numerical solver is based on a finite-volume scheme for solution of the unsteady Reynolds averaged Navier-Stokes (URANS) equations. The solver is of second order accuracy in space and time. The LLR $k-\omega$ model by Rung[22] is used for turbulence modelling.

The computational domain for the flow extends 15 chords upstream, upwards, below and 25 chords downstream with respect to the leading edge of the wing. The grid is structured, locally refined at the the actuation position, and consists of approximately 93000 points. The zero net flux actuation is modeled as a periodic boundary condition (fluctuating velocity).

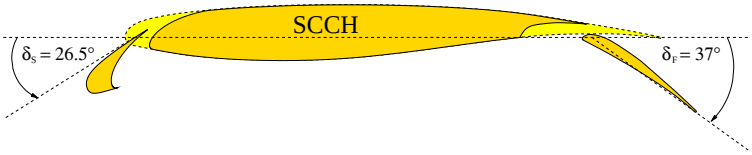

Figure 1. The swept constant chord half configuration.

B. The NACA0012 Benchmark

The 2D NACA0012 airfoil is considered at and angle of attack of 30^o, as seen in Figure 2 (left). The Reynolds number is $Re = 200$, normalized with respect to a chord length of $c = 2$ and an incoming flow velocity of $U_\infty = 1$. The computational domain is a rectangle, $\mathbf{\Omega} = \{\mathbf{x} = (x, y) : -5 \leq x \leq 15, -5 \leq y \leq 5\}$. Direct numerical simulations were run with a triangular grid of 3115 points (also shown in Figure 2 (right)) making this model readily available for coarse but rapid testing and prototyping. The steady, fully attached flow is unstable, and the natural flow is, again, characterized by periodic vortex shedding, forming a von Kármán vortex street. Figure 3 depicts the attractor's mean, the 4 leading POD modes of the attractor and the POD eigenvalues, indicating that some 95% of the attractor's perturbation kinetic energy is resolved by the first mode pair. This is therefore a typical 2D, unstable wake flow. Following[1,11], a good qualitative approximation of natural transient dynamics would be by a generalized mean-field model, with modes representing the first vortex shedding harmonic and a *shift mode* mode, capturing slow changes in the mean flow.

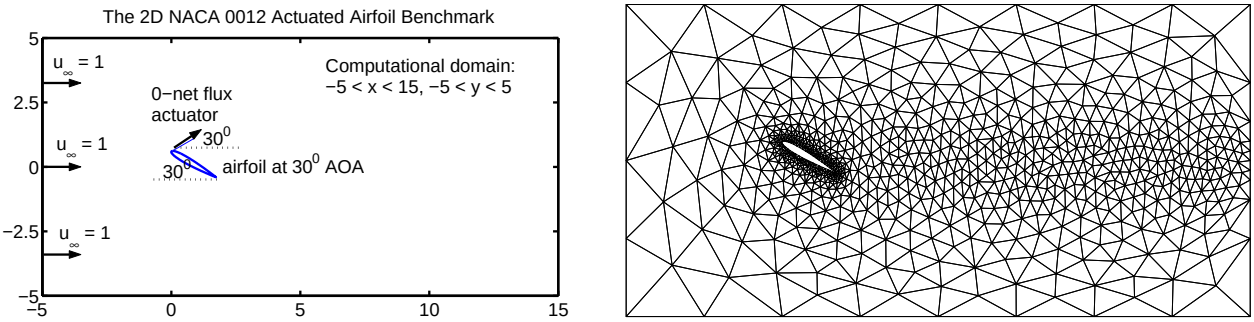

Figure 2. Left: The computational domain and boundary conditions of the NACA0012 benchmark. Right: The 3115 points computational grid, used in DNS simulations.

While a 3 states model provides a qualitative approximation, mode deformation along longer transients is associated with deformation of leading flow structures[7,9,19,23]. Good approximation may therefore require to compensate for such deformations. The approach proposed in[7,8] is to parameterize the fixed size family of modes. As in the cited papers, that would be a 3 modes family, where $\mathbf{u}_i^\kappa$, $i = 1, 2$ represent the first shedding harmonic, $\mathbf{u}_3^\kappa$ capture changes in the base flow, and κ is a parameter representing the operating condition, along the modeled manifold, which can be dynamically estimated from a sensor reading.

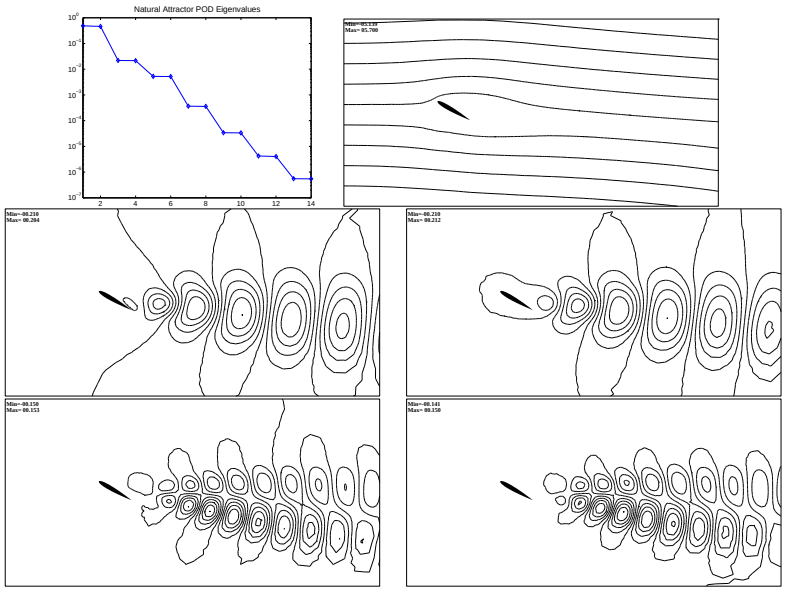

Figure 3. Natural attractor analysis for the NACA 0012 benchmark. Top row, Left: POD Eigenvalues. The remaining 5 plots depict the mean flow and leading 4 POD modes of the natural attractor, represented by streamlines. Top Row, right: The mean flow. Middle row: first and second POD modes, representing the main shedding harmonic. Bottom row, the third and fourth POD modes, representing the second harmonic.

Zero net flux synthetic jet actuation is modeled by a controlled boundary value condition near the front tip point. The jet is pointing downstream and upwards relative to the horizon, at an angle of 30^o (that is, 60^o with respect to the airfoil chord), as schematically shown in Figure 2. Open loop actuation was by sinusoidal waveform at frequency of $\omega_a = 66.6$, which is roughly 10 times the natural shedding frequency. As will be seen later, this open loop actuation attenuates the instability. We shall explore means to incorporate the zero net flux actuator in a Galerkin modeling framework, and identify the role of the shift mode as an effective representation of the physical mediation between actuation and the dominant instability.

III. Benchmark Modeling and Approximation

Based on URANS simulations with constant / stepped actuation, Section A is focused on the structure of a minimum order, generalized mean-field model for the actuated flow, and on calibrated with respect to empirical POD data. In Section B we demonstrate the utilization of much longer time trajectories of DNS data, to explore the flow structure subject to time varying actuation. In particular, we use that section to demonstrate the computational tools of the following, Sections IV and V, including frequency partitioning of the reference and a two step, approximate POD procedure.

A. A Generalized Mean-Field Model for the High Lift Configuration

The data comprise a trajectory that starts with a transient from the natural attractor to an attenuated attractor, under constant actuation, and then returns to the natural attractor, once actuation stop. The simulation lasts approximately 30 convective time units (tU_∞/c).

Following our previous work[1, 11, 24], a minimum order mean field model is proposed for the unactuated (natural) flow. This model is a 3 states ODE, consisting of a pair of Fourier coefficients, representing the (natural) dominant vortex shedding and a coefficient that reflects the change in mean flow. This mean field model is extended to include the actuation frequency, roughly the second harmonic of the natural shedding frequency, that is excited by open loop actuation. This addition leads to a minimum order, 5 states, quadratically nonlinear ordinary differential equation. The model accommodates the key physical phenomenon, that when forcing is switched on, the coefficients representing the natural shedding undergo rapid decay and the coefficients corresponding to the actuated, smaller structures, grow from their natural, low level steady state. A principal sketch, illustrating the extended mean field model, is shown in fig. 4.

The least order two frequency model is expressed is written in terms of the state $a(t)$ of the Galerkin

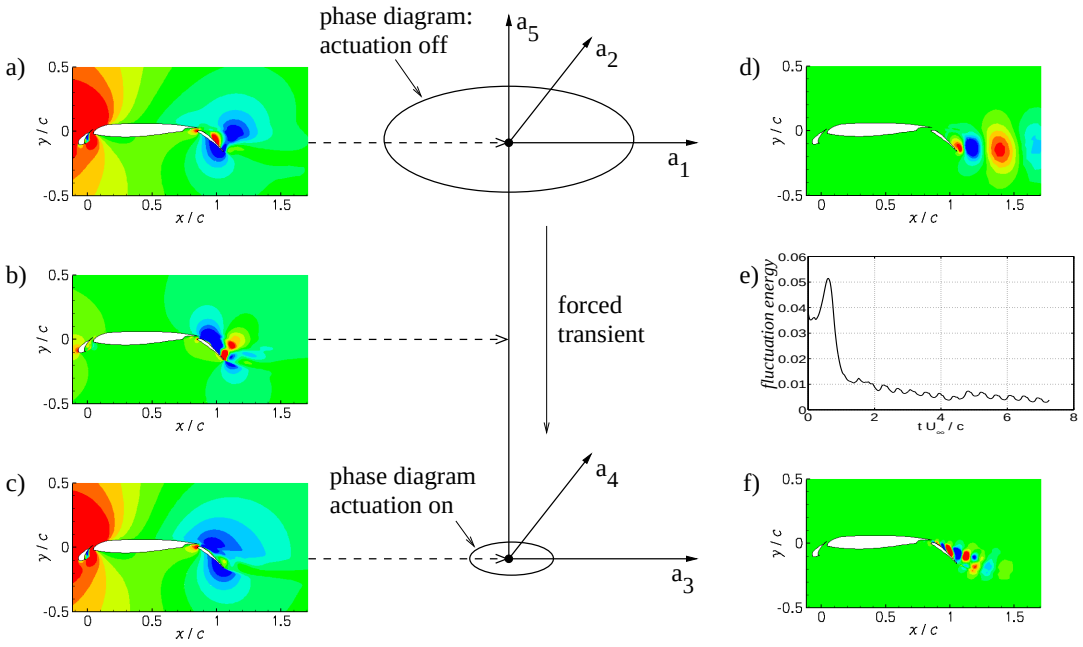

Figure 4. Principal sketch of the two frequency model of an open loop forced flow around a high lift configuration. Left: mean flow of the natural (a) and actuated (c) flow. The shift mode (b) describes the difference among them. Right: first POD mode of the natural (d) and actuated flow (f). The flow fields show the velocity field in vertical direction, red is positive (upwards) and blue negative. Middle column: phase diagram of the 5 state model. The limit cycle of the natural flow is spanned by a_1 and a_2, the actuated flow by a_3 and a_4. The decay of the fluctuation energy during a forced transient is shown in (e).

approximation, and the sinusoidal actuation command $b(t)$.

$$\begin{bmatrix} \dot{a}_1 \\ \dot{a}_2 \\ \dot{a}_3 \\ \dot{a}_4 \\ \dot{a}_\Delta \end{bmatrix} = \begin{bmatrix} 0 & -\omega_1 & 0 & 0 & -\beta_{\omega_1} a_1 \\ \omega_1 & 0 & 0 & 0 & -\beta_{\omega_1} a_2 \\ 0 & 0 & -\sigma_{\omega_2} & -\omega_2 & -\beta_{\omega_2} a_3 \\ 0 & 0 & \omega_2 & -\sigma_{\omega_2} & -\beta_{\omega_2} a_4 \\ \alpha_{\omega_1} a_1 & \alpha_{\omega_1} a_2 & \alpha_{\omega_2} a_3 & \alpha_{\omega_2} a_4 & -\sigma_\Delta \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_\Delta \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ g_3 \\ g_4 \\ 0 \end{bmatrix} + \dot{b} \begin{bmatrix} 0 \\ 0 \\ g_3{}' \\ g_4{}' \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ c_{\omega_2} \\ c_{\omega_2} \\ c_\Delta \end{bmatrix} \quad (1)$$

(Clearly, the role of $\dot{b}(t)$ in this simple model is merely as a representation of the quadrature signal, with 90^o phase lag with respect to $b(t)$.) A Galerkin projection is problematic with URANS data, and the free parameters are calibrated using empirical data.

One key aspect of the structure of the model (1), is that the mean-field serves the critical role, as a mediator between the actuation and the natural shedding instability. This is supported by the structure of the perturbations, as seen in Figure 4, which suggest no direct interaction between the modes associated with the two frequencies, and is in the basic tenets of generalized mean-field theory[10, 11, 25, 26]. In the following section we shall use the NACA 0012 benchmark as an illustration of the more common, analytical derivation of the Galerkin ODE, via a Galerkin projection, leading to the same conclusion regarding the mean field role as a mediator between the actuation and the instability.

A comparison of the qualitative behavior of the Fourier coefficients from URANS data and integration of (1) are shown in fig. 5. A steady state match is achieved by parameter tuning via harmonic balancing. A good agreement of the first transient is achieved by tuning the rate coefficients. While qualitatively similar, a larger mismatch is conspicuous during the second transient, once actuation is switched off. The model (1) clearly stipulates that the shift mode must first undershoot into negative values in order to excite the

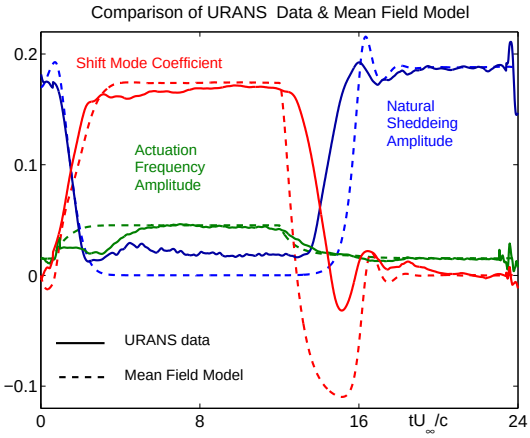

Figure 5. Comparison of the (low pass filtered) oscillation amplitudes and shift mode coefficient from URANS data (solid lines) and their counterpart from simulation of the minimum order mean field model, (1) (dashed lines). Constant actuation at twice the natural shedding frequency is switched on and then off, triggering the two transients: from the natural attractor to the attenuated attractor, early on, and back to the natural attractor, later on. The shift mode coefficient vanishes on the natural attractor, and the shedding harmonic, on the controlled attractor, where it is substituted by elevated oscillation level at the actuation frequency.

natural harmonic instability. Qualitatively, such an undershot is present in the empirical data, but it is much smaller, and far too late, to explain the early rise in the natural harmonic oscillations. The reason will clarify in the more detailed discussion of our second benchmark, in the following section: Base flow dynamic becomes much richer during sharp transients, and requires multiple base-flow modes, orthogonal to the long term dominant, shift mode.

B. Modeling and Approximation for the Actuated NACA 0012 Airfoil

The reference simulation used to generate data for this study starts at a perturbed approximation of the steady solution, under an actuation of the form $b(t) = A_a(t)\sin(\omega_a t)$. The rectilinear amplitude $A_a(t)$ was proportional to the line marked as "mode 1", in the bottom left plot, in Figure 8. (The precise meaning of the figure will be explained shortly.) The entire simulation is captured by 6400 snapshots.

Preliminary Compression

A preliminary application of the approximate POD algorithm, detailed in Section V, was used to compress the initial simulation data to a 37-dimensional sate space, with near-perfect resolution (an average of 99.85% and a minimum of 99.3% of the TKE). The use of the approximate POD algorithm leads to a nearly 20-fold reduction of the computational burden when compared with a standard POD reduction of the entire reference. From this point on, the entire analysis is carried out in the compressed state space, which is isometric to $\mathbb{R}^{37}$, which is a mere 0.06% of the original state space dimensionality of 6230, leading to drastic computational relief in subsequent analysis.

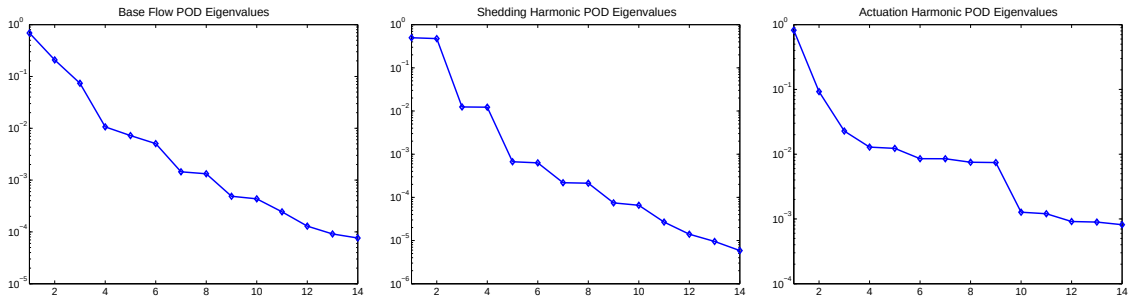

Figure 6. POD Eigenvalues of the frequency partitioned actuated reference trajectories in the NACA 0012 benchmark. Left: The base-flow. Middle: The shedding harmonic. Right: The actuation harmonic.

Frequency Separation

We are interested in an approximation focused on three dominant flow components, distinguished by their frequency content: A non-oscillatory component, representing time variations in the base-flow, oscillations at the (slowly time varying) shedding frequency, and direct response to the actuation, at the fixed actuation frequency. Using the algorithm described in Section IV, the approximated reference trajectory was partitioned accordingly. Leading POD modes where then extracted from each of these three sub-references. Since now the state space is low dimensional, that was done using the direct method. The respective POD eigenvalues are depicted in Figure 6. In what follows, modes representing base flow transients are denoted $\mathbf{u}_{b,k}$, modes representing the shedding harmonic are denoted $\mathbf{u}_{h,k}$, and modes representing the actuation frequency are denoted $\mathbf{u}_{a,k}$. As expected the leading base flow mode is nearly identical to the normalized difference between the natural attractor and the controlled attractor mean, which is a natural *shift mode*, following the ideas in[1] . The (non orthogonal) Galerkin approximation is thus of the form

$$\mathbf{u}(\mathbf{x},t) \approx \mathbf{u}_0(\mathbf{x}) + \sum_k a_{b,k}(t)\mathbf{u}_{b,k}(\mathbf{x}) + \sum_k a_{h,k}(t)\mathbf{u}_{h,k}(\mathbf{x}) + \sum_k a_{a,k}(t)\mathbf{u}_{a,k}(\mathbf{x}) \tag{2}$$

where $\mathbf{u}_0$ is the simulation mean and where the coefficients are obtained by an oblique projection.

A Galerkin approximation that employs the leading base flow mode as a single shift mode, the leading pair of shedding harmonic modes, and the leading two actuation harmonic modes captures an average of 91.3% of the perturbation energy (excluding the initial transient, where the mean-field transient is quite different). For a higher 94.3% resolution of the TKE one uses two modes representing base flow transients, four modes (two pairs), at the shedding harmonic, and, again, two modes for the actuation harmonic. These modes are depicted in Figure 7.

One conspicuous fact, in Figure 6, is that while POD modes of the shedding harmonic sub-reference appear in typical, nearly balanced pairs, that is not the case for modes representing oscillations at the actuation frequency. The simple explanation is that much of the energy stored in that frequency, in close proximity to the actuation point, is transferred to other flow components, as the distance from the actuation point grows. The modes representing the actuation frequency, of which two are shown in Figure 7, capture the effects over growing ranges, and thus, at increasingly diminished intensities.

Relation to Interpolated Models

Figure 8 shows some key aspects of the time trajectories. The two leading base flow coefficients are shown in the left plot of fig. 8. A useful observation is that the relations between the two coefficients tends to change on a much longer time constant than the coefficients themselves. The exception is the sharp transient when the actuation is rapidly ramped down for $t_n \in [240, 285]$ (where $t_n = tU_\infty/c$ is measured in convective time units). This means that over shorter time intervals, changes in the mean flow are strongly dominated by a single *local shift mode*. That local shift mode can be computed by an extremely cheap application of the POD procedure to the short-time trajectory of the coefficients vector $[a_{b,k}(t)]_{k=1}^K$ in a high fidelity approximation of the base flow transient. We leave the obvious details out, to prevent further digression of the present discussion. A similar observation applies to the shedding harmonic oscillations where a single mode pair dominates this frequency over relatively long time intervals. This naturally leads to the concept of tuned, interpolated models[7] , mentioned earlier.

Actuation Model

Ideally, each of the coefficients $a_{h,k}(t)$ and $a_{a,k}(t)$ is defines a zero mean oscillatory signal:

$$a_{h,k}(t) = A_{h,k}(t)\sin(\phi_{h,k}(t)) \quad \text{and} \quad a_{a,k}(t) = A_{a,k}(t)\sin(\omega_a t - \theta_{a,k}(t))$$

with slowly varying amplitudes $A_{h,k}$ and $A_{a,k}$, shedding frequency $\dot{\phi}_h$ and actuation phase lag θ_a. Focusing on the fixed actuation frequency, a FIR filter readily extracts the amplitude $A_{a,k}(t)$ and the phase lag (relative to the actuation command) $\theta_{a,k}(t)$. These are the quantities shown in the bottom two plots in Figure 8. The two modes respond in distinctly different manners: While the response of $a_{a,2}(t)$ is clearly nonlinear and dynamic, the response of $a_{a,1}(t)$ is nearly linearly dependent on the actuation command: its (filtered) amplitude is proportional to the commanded amplitude, and is phase lag is nearly constant. The

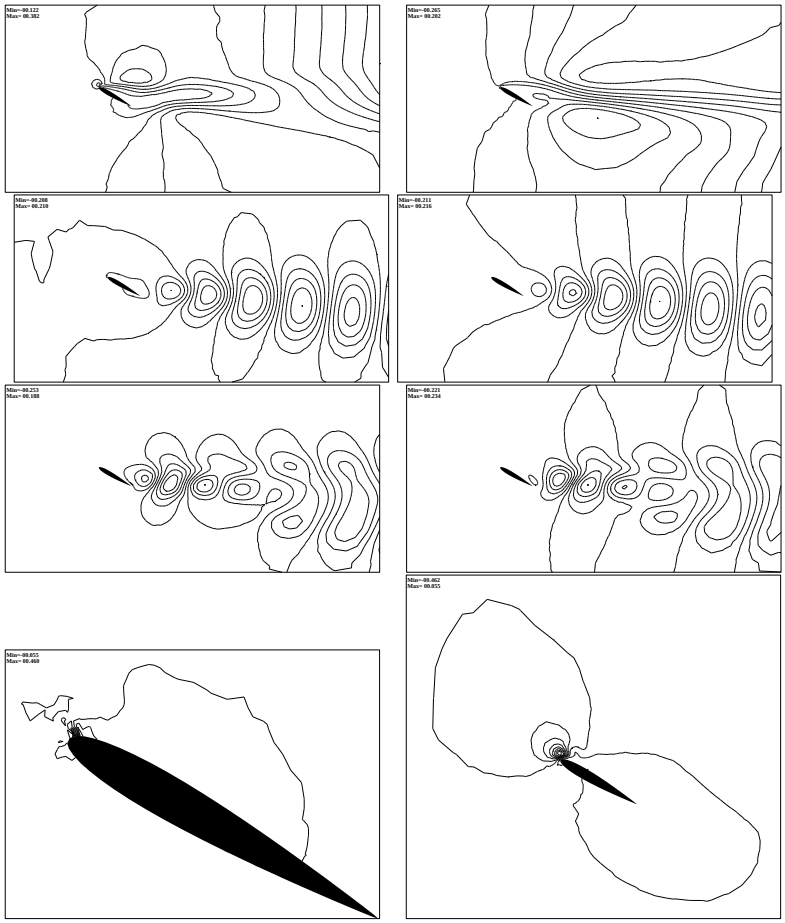

Figure 7. Leading POD modes of each of the harmonic sub-references in the NACA 0012 benchmark, visualized by streamlines. Top, left to right: Leading base flow modes. Second row: Leading shedding frequency mode pair. Third row: Secondary mode pair at the shedding frequency. Bottom: Leading actuation harmonic modes. The latter two are local in nature and the plots are zoomed in the area of highest fluctuation.

reason for this difference is due to the fact that $\mathbf{u}_{a,1}$ represent the response in the immediate vicinity of the actuator, while $\mathbf{u}_{a,2}$ represent the dynamic response, further afield, as can be seen in Figure 7.

The coefficient $a_{a,1}(t)$ will thus serve as a proxy for the actuation command, and the mode $\mathbf{u}_{a,1}$ may serve as an *actuation mode* in the sense that $a_{a,1}(t)\mathbf{u}_{a,1}$ is an imposed velocity component and time derivative $\dot{a}_{a,1}(t)\mathbf{u}_{a,1}$ is an imposed volume-force equivalent of the actuation. The response of $a_{a,2}(t)$, on the other hand, is dynamic, representing fluid dynamic effects, further afield from the actuation point, and it must be included as a state variable in a reduced order model.

Actuation Response and The Galerkin System

The amplitude of shedding harmonic oscillations is easier to compute due to the fact that these modes come in nearly balanced pairs. The time trajectory of that amplitude is depicted in the top middle plot, in Figure 8.

The top right plot, in Figure 8, is provided to facilitate a visualization of the actuation response of the dominant base flow mode and the oscillations at the shedding frequency. The red and green curves are, respectively, normalized forms of the leading base flow coefficient, $a_{b,1}(t)$ and of the oscillation amplitude at the shedding frequency. The normalization is such that the respective values of 1 on the natural attractor, and 0 on the attenuated, controlled attractor. The blue plot is that of $-a_{a,1}(t)$, normalized the same way. (Thus, the blue plot equals 1 where the actuation vanishes, and 0, where it is maximal.) We recall[1] that in natural transients, the shift mode is slaved to the shedding instability, and thus changes in the shift mode coefficient lag behind changes in the oscillation amplitude. The converse is observed here. This is in agreement to with the observation made in the calibrated model, presented for the high lift configuration,

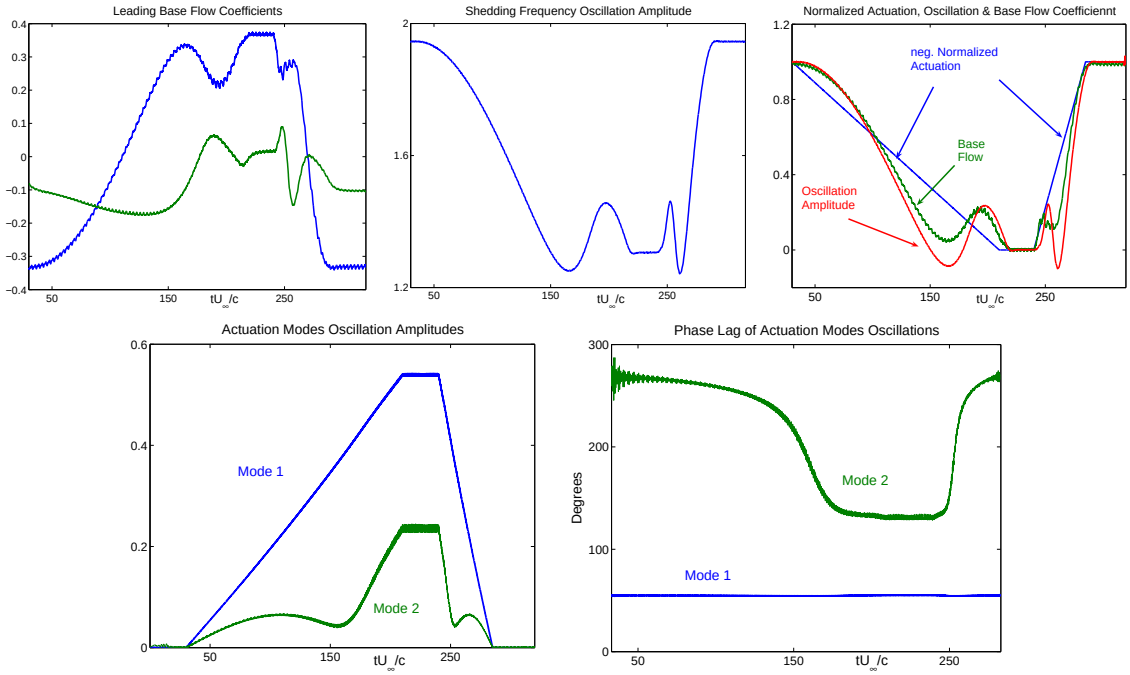

Figure 8. Analysis of the actuated references in the NACA 0012 benchmark. Top, Left: The two leading base flow coefficients. Top, Middle: Oscillation amplitude at the shedding frequency, as captured by the leading two mode pairs. Top, Right: Normalized plot of the firs base flow coefficient, the shedding harmonic oscillation and the sign-inverted first actuation mode coefficient. Bottom, Left: Oscillation amplitude of the time coefficients of the two actuation modes. Bottom, Right: Phase lag (in degrees, mod 360^o) of these coefficients' oscillations, relative to the actuation command. The phase lag is defined only for the actuated interval.

in Section A, and with the Galerkin projection, in (3), below.

$$M\dot{a} = A(a)a + c \tag{3}$$

where the state and the proxy for the actuation command (i.e., $a_{a,1}$) are lumped in the vector and the need for the correlation matrix of the respective modes, M, is due to lack of perfect orthogonality of the modes. A formal Galerkin projection computation leads to the following approximate structure:

$$a = \begin{bmatrix} a_{h,1} \\ a_{h,2} \\ a_{b,1} \\ a_{a,1} \\ a_{a,2} \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0.06 \\ 0 & 1 & 0 & 0 & 0.02 \\ 0 & 0 & 1 & 0 & -0.03 \\ 0 & 0 & 0 & 1 & 0.07 \\ 0.06 & 0.02 & -0.03 & 0.07 & 1 \end{bmatrix} \quad c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -0.06 \\ -0.08 \end{bmatrix}$$

and

$$A(a) = \begin{bmatrix} 0.02 & 1 + 0.01a_{b,1} & 0.03a_{h,1} \\ -1 - 0.01a_{b,1} & 0.02 & 0.03a_{h,2} \\ -0.03a_{h,1} & -0.03a_{h,2} & -0.49 - 0.12a_{b,1} \\ 0 & 0 & 1 + 0.62a_{b,1} \\ 0 & 0 & 0.48 + 1.5a_{b,1} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.4 + 31.5a_{a,1} + 3.6a_{a,2} + 0.1a_{b,1} & 0.04 - 2.61a_{a,2} - 1.5a_{b,1} \\ -22.4200 - 173a_{a,1} + 1.23a_{a,2} - 1.15a_{b,1} & 0.76 - 4.95a_{a,2} - 2.73a_{b,1} \\ -2.21 - 19.6a_{a,1} + 3.27a_{a,2} - 2.2a_{b,1} & -1.56 - 1.76a_{a,2} + 0.8a_{b,1} \end{bmatrix}$$

While this system is not expected to produce the precise time traces of the DNS data, its structure reveals the energy exchange patterns between the modes included in the Galerkin approximation[27]. In particular, the following points are noted in summary:

- Excluding actuation, our model is essentially identical to the mean-field field model in[1], identifying the current $a_{b,1}$ and $a_{h,k}$, $k = 1, 2$, with a constant shift of $-a_\Delta$ and with a_k, respectively.
- High fidelity representation requires continuous deformation in both the local expansion modes for the periodic instability and the shift mode, in line with the interpolated model proposed in[7]. We have not included these computations here, but stress that they can easily be derived from a higher order Galerkin approximation, of the form (2), applying a computationally cheap POD procedure to the time coefficients trajectories of each of the three harmonic categories.
- We have identified $\mathbf{u}_{a,1}$ and an *actuation mode*, representing the actuation command in a low order Galerkin approximation. A second, dynamic actuation, mode provides an added dynamic mediation between the actuation and the actuated flow.
- Energy exchange between the actuation and the shedding harmonic is mediated by the base flow.

IV. POD of a Frequency Separated Reference

A pre-requisite from very low order models is the ability to focus on physically meaningful flow structures, including vortices and associated spatial and temporal frequencies. POD approximations are successful in satisfying this requirement in simple, symmetric configurations, such as the 2D cylinder wake flow example[1,28]. Yet in non-symmetric and 3D configurations, POD modes often fail to cleanly capture vortical structures and single out distinct temporal harmonics, as illustrated in the left plot, in Figure 9.

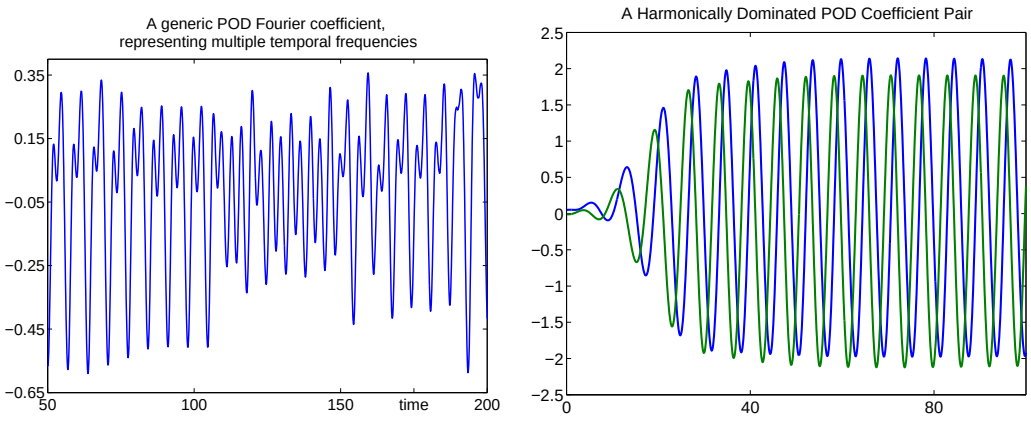

Figure 9. Left: A generic transient time trace of a secondary Fourier coefficient of a POD mode. A single mode captures multiple harmonic components. Right: A generic dominant POD pair for a dominant harmonic. Such a pair can be used to determine the instantaneous phase of such oscillations, for the entire distributed flow state, helping to avoid the generation of modes such as the one on the left.

The idea underlying the remedy suggested here is simple. The reference trajectory $\mathbf{u}(\mathbf{x}, t)$, is first partitioned as the sum of a fixed base flow, $\mathbf{u}_0$, several fixed-frequency components $\mathbf{u}_f(\mathbf{x}, t)$, $f = 1, \ldots, F$, and a resolvent $\mathbf{u}_{res}(\mathbf{x}, t)$ (for the purpose of error assessment):

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{u}_0 + \sum_{f=0}^{F} \mathbf{u}_f(\mathbf{x}, t) + \mathbf{u}_{res}(\mathbf{x}, t) \tag{4}$$

The partition into a base flow, shedding frequency and actuation frequency components, in Section III, is an example. POD approximation is extracted separately, for each $\mathbf{u}_f(\mathbf{x}, t)$. The combined dominant modes are used in the Galerkin approximation. Since mutual orthogonality is not imposed on the sub-references, the extracted mode set would typically *not* be orthogonal and the time coefficients of the modes are computed

by an oblique projection of the original reference $\mathbf{u}(\mathbf{x}, t)$ on the calculated mode set, and TKE posteriori resolution needs to be evaluated a posteriori.

The heuristic rational for this procedure is twofold. First, it is expected that POD modes extracted from a single frequency trajectory will be dominated by flow structures that are associated with that frequency. Second, well-posedness of the problem suggests linear independence of state space components associated with distinct flow structures. Implementation of this approach gives rise to several challenges:

Computational complexity: Both temporal harmonic analysis of a distributed parameter signal, and POD approximation of multiple sub-references, instead of a single trajectory are computationally expansive. The suggested solution to both issues is a preliminary, high resolution POD (or use a approximate POD, as described in Section V) approximation of the original reference. Generically, a near perfect resolution of the original reference can be achieved with multiple orders of magnitudes of reduction in data dimensionality, making both tasks outlined in this section inexpensive.

Evaluating a time varying frequency / period: A meaningful and robustly quantifiable definition of the instantaneous frequency (or phase), needs to be extracted, and used as a basis for local harmonic approximation. There is no clear, a priori definition of the instantaneous period, and we explored several basic approaches, of which we describe here just two.

Optimal local harmonic approximation: The instantaneous frequency at the time t is defined as the dominant frequency of a test signal over a symmetric, fixed length, time window $[t-0.5\tau, t+0.5\tau]$, where τ is an upper bound on expected period lengths. Using a filter bank, it is identified by yielding best $L_2[t-0.5\tau, t+0.5\tau]$ resolution of the test signal with a postulated length harmonic expansions. The scalar test signal would be extracted from the data. For example, even when some Fourier coefficients of the POD approximation resemble fig. 9, overall near periodicity typically implies the existence of a pair of modes whose coefficients are strongly dominated by a single frequency, such as the pair depicted in the right plot, in Figure 9, which can be used as test signals.

Using instantaneous dynamic phasors: Here we borrow from the conceptual framework of synchronous electric machines and the representation of 3-phase AC currents by their *direct* and *quadrature* components. Let $(a_k(t),\, a_{k+1}(t))$ be the time trajectories of the harmonically Galerkin coefficients mentioned earlier, such as those depicted in the right plot, in Figure 9. We postulate a representation

$$a_k(t) = A_0(t) + R(t)\cos(\phi(t)), \quad a_{k+1}(t) = B_0(t) + R(t)\sin(\phi(t))$$

where the time variations in $A_0(t)$, $B_0(t)$, $R(t)$ and $\omega(t) \doteq \dot{\phi}(t)$ are assumed slow relative to the instantaneous period length $\tau(t) \doteq 2\pi/\omega(t)$. The *dc* components $A_0(t)$ and $B_0(t)$ are estimated by an window averaging FIR filter, and the instantaneous phase, $\phi(t)$, is evaluated as

$$\phi(t) = \angle(a_k(t) - A_0(t)) + \imath(a_{k+1}(t) - B_0(t)))$$

An FIR filter, extracting the locally near-linear components of the graph of ϕ will help to remove noise driven artifacts. Having determined the unwrapped phase ϕ, it is used as an indexing parameter for the reference trajectory, instead of time, and POD analysis is performed for sampling at equal ϕ intervals.

Of these two approaches (and several variants, left out for space considerations), we found the latter easier and more robust. The results for the NACA 0012 airfoil, in Section III.B, were obtained, using this method.

Non-commensurate frequencies. One expects a lock-in to the actuation frequency when it is sufficiently close either to the natural shedding frequency or to a close harmonic. This is the case in the high lift configuration benchmark, discussed here, where actuation is at the 2^{nd} harmonic of the natural shedding frequency. When the actuation frequency is much higher, the shedding frequency will remain dynamic, while the actuation frequency is deterministic and fixed, resulting in the presence of non-commensurate frequencies. This, however, is expected to arise only when the gap between the shedding and actuation frequency is large, justifying an independent extraction of the two components.

V. Reduced Computational Complexity: Two Steps Approximate POD

The goal to obtain unified state compressions for very long reference trajectories, has been motivated earlier, and we repeat it here briefly: Long data sets, covering multiple operating conditions are necessary

to be able to extract the modes necessary to produce a reduced order model for transients covering a wider dynamic range. While it is possible to obtain the modes separately, for each operating condition, it is far easier to analyze local mode deformation and tunable mode-sets, for the entire range, using a single, low dimensional, compressed reference. As has been noted earlier, a mere 37 dimensional compression ($\approx 0.5\%$ of the original dimension) provides a near perfect resolution of the perturbation energy in a rich reference, in the NACA 0012 example. Yet as the computational cost of proper orthogonal decomposition / principal component analysis is cubic, it becomes exorbitant when both the data set and the dimension are very large. Methods to alleviate this burden are therefore not merely desirable, but eventually become critical enablers. As is commonly the case, key to such methods is domain partitioning.

This section reviews two variants of such a method. It trades a (typically small) reduction in optimality for a drastic reduction in computational complexity. Two significant aspects of the proposed algorithm include the fact that it is associated with *provable error bounds*, like the standard procedure, and that *it does not require multiple iterations*, and is complete in two steps. It comprises a preliminary model reduction step, applied to a partitioning of the original ensemble, and a subsequent global POD procedure, of an approximate, smaller ensemble. Heuristic implementation guidelines aim to tilt the tradeoff in a way that will minimize efficiency losses, without affecting the computational savings. One intended use of the algorithm, is as a high fidelity compression of the original ensemble, which will ease the computational cost of subsequent analysis, such as the frequency partitioning, discussed earlier. Since the compressed representation is used only as an intermediate product, computational savings and the quality of the approximation (i.e., a global error bound) are valued more than strict optimality in the sense that a minimal mode set is used. It is noted that while the detailed algorithm outlined here is exploits partitioning of the time interval (more generally, of the samples's index set) essentially the same approach applies to spatial partitioning.

A. Problem Setting

Both space and ensemble parameterization (e.g., time in the POD context, or an indexing set, in other principal component analysis applications) are assumed discretized from the outset. The POD/PCA procedure is applied to an ensemble

$$\mathbf{\Phi} = \{\Phi_i \,:\, i \in \mathcal{I}\} \subset \mathbf{X}, \quad \Phi_0 = mean(\mathbf{\Phi}), \qquad \tilde{\mathbf{\Phi}} = \left\{\tilde{\Phi}_i \doteq \Phi_i - \Phi_0, : \, i \in \mathcal{I}\right\}$$

where $\mathbf{X}$ is a Hilbert space, $\mathcal{I}$ is an index set (e.g., discrete time), and both $M \doteq dim(\mathbf{X})$ and $N \doteq size(\mathcal{I})$, are very large. By a dual use of notations, we shall identify ensembles, such as $\tilde{\mathbf{\Phi}}$, with matrix operators

$$\tilde{\mathbf{\Phi}} \;\; \mathbf{X} \mapsto \mathbb{R}^N, \quad \tilde{\mathbf{\Phi}}\xi = \sum_{i\in\mathcal{I}} \xi_i \tilde{\Phi}_i$$

Given a $\rho \in (0,1)$, the PCA/POD approximation identifies a lowest dimension subspace $\mathbf{Y} \subset \mathbf{X}$ for which

$$\rho \leq \frac{\langle \|\pi_Y \tilde{\Phi}_i\|^2 \rangle}{\langle \|\tilde{\Phi}_i\|^2 \rangle} = \frac{tr\, \pi_Y \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Phi}}' \pi_Y}{tr\, \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Phi}}'} \tag{5}$$

where "*tr*" indicate a matrix trace, π_Y is the orthogonal projection onto $\mathbf{Y}$, $\|\ \|$ is the norm in $\mathbf{X}$, associated with an inner product $(\ ,\ )$, and $\langle\ \rangle$ is the ensemble average. Specifically, given the singular value decomposition (SVD)

$$\tilde{\mathbf{\Phi}} = \mathbf{USV}' = \sum_{i\in\mathcal{I}} \sigma_i U_i V_i'$$

the optimal subspace is

$$\mathbf{Y} \doteq span\{U_i\}_{i=1}^{L}, \qquad L \doteq \min\left\{\tilde{L} \,:\, \frac{\sum_{i=1}^{\tilde{L}} \sigma_i^2}{\sum_{i=1}^{N} \sigma_i^2} \geq \rho\right\}, \qquad \pi_Y = \sum_{i=1}^{L} U_i U_i'$$

The common, indirect *method of snapshots*[3] addresses the generic situation where $N << M$: Then the right singular vectors V_i are computed first as eigenvectors of the $N \times N$ correlation matrix $\tilde{\mathbf{\Phi}}'\tilde{\mathbf{\Phi}}$. The formalism of the ensuing discussion is rather that of the direct method, and the analysis of $\mathbf{\Phi}\mathbf{\Phi}'$. It is stressed that this

is done *only* to facilitate the evaluation of error bounds. Actual computations, and indeed, the true benefits of the algorithm presented here, are expected to be realized by the equivalent method of snapshots.

The computational complexity of the POD procedure, including the computation of the covariance matrix, is (assuming the typical $M > N$) $O(M \cdot N^2)$. The primary intent of the algorithms discussed below is to reduce the computational burden via a partition of the ensemble index, $\mathcal{I} = \cup_{k=1}^{K} \mathcal{I}_k$, and a preliminary POD dimension reduction in each of the subsets

$$\mathbf{\Phi}_k = \{\Phi_i \,:\, i \in \mathcal{I}_k\} \subset \mathbf{U}, \quad k = 1 : K$$

For simplicity we shall assume that all index sub-sets are of the same size: $N_0 \doteq size(\mathcal{I}_k) = N/K$. The challenge is to integrate the information obtained by local PODs into an approximation of the ideal, global POD, with guaranteed error bounds.

The focus here is on partitioning the index set. An example of a recent discussion on the complementary use of spatial partitioning can be found in[18] .

B. Time Partition: A Shared Mean Algorithm

This procedure is practical when available RAM suffices to store the entire ensemble, Φ, whereby the global average can be computed without multiple data loading from a hard disk. A lower bound $0 < \rho < 1$ is specified for the requisite resolution, as defined in (5).

Step 0: Compute the global mean, Φ_0 and the perturbation samples, $\tilde{\Phi}_i$, $i \in \mathcal{I}$.

Step 1: Compute the SVD for each of the operators corresponding to the sub-ensembles $\tilde{\mathbf{\Phi}}_k$

$$\tilde{\mathbf{\Phi}}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k' = \sum_{i=1}^{N_0} \sigma_{k,i} U_{k,i} V_{k,i}' \tag{6}$$

Step 2: Truncate the SVD at $N_k \le N_0$, substituting $\tilde{\mathbf{\Phi}}_k$ by

$$\bar{\mathbf{\Phi}}_k \doteq \bar{\mathbf{U}}_k \bar{\mathbf{S}}_k \bar{\mathbf{V}}_k' \doteq \sum_{i=1}^{N_k} \sigma_{k,i} U_{k,i} V_{k,i}' \quad \text{where} \quad N_k \doteq \min\left\{ \tilde{N} \,:\, \frac{\sum_{i=1}^{\tilde{N}} \sigma_{k,i}^2}{\sum_{i=1}^{N_0} \sigma_{k,i}^2} \ge \sqrt{\rho} \right\}, \qquad \pi_{Y_k} \doteq \sum_{i=1}^{N_k} U_{k,i} U_{k,i}'$$

Note that $\bar{\mathbf{\Phi}}_k = \pi_{Y_k} \tilde{\mathbf{\Phi}}_k$.

Step 3: Define the alternative sub-ensembles and associated operators

$$\mathbf{\Psi}_k \doteq \{\sigma_{k,i} U_{k,i}\}_{i=1}^{N_k} \qquad \Leftrightarrow \qquad \mathbf{\Psi}_k \xi = \bar{\mathbf{U}}_k \bar{\mathbf{S}}_k \xi = \sum_{i=1}^{N_k} \sigma_{k,i} \xi_i U_{k,i}$$

Using operator notations, the selection of N_k provides for

$$tr\, \mathbf{\Psi}_k \mathbf{\Psi}_k' = tr\, \bar{\mathbf{U}}_k \bar{\mathbf{S}}_k^2 \bar{\mathbf{U}}_k' = tr\, \bar{\mathbf{U}}_k \bar{\mathbf{S}}_k \bar{\mathbf{V}}_k' \bar{\mathbf{V}}_k \bar{\mathbf{S}}_k \bar{\mathbf{U}}_k' = tr\, \bar{\mathbf{\Phi}}_k \bar{\mathbf{\Phi}}_k' = tr\, \pi_{Y_k} \tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k' \pi_{Y_k} \ge \sqrt{\rho}\, tr\, \tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k'$$

The advantage of $\mathbf{\Psi_k}$ is that it holds only N_k sample vectors, compared with $N_0 >> N_k$ vectors in $\bar{\mathbf{\Phi}}_\mathbf{k}$.

Step 4: Define the truncated *global* set whose size is $\bar{N} \doteq \sum_{k=1}^{K} N_k \le N$, M

$$\mathbf{\Psi} \doteq \cup_{k=1}^{K} \mathbf{\Psi_k} = \cup_{k=1}^{K} \{\sigma_{k,i} U_{k,i}\}_{i=1}^{N_k}$$

and apply the POD procedure to $\mathbf{\Psi}$ with the resolution bound $\sqrt{\rho}$

$$\mathbf{\Psi} = \bar{\mathbf{U}} \bar{\mathbf{S}} \bar{\mathbf{V}}' \doteq \sum_{i=1}^{\bar{N}} \bar{\sigma}_i \bar{U}_i \bar{V}_i', \quad L \doteq \min\left\{ \tilde{L} \,:\, \frac{\sum_{i=1}^{\tilde{L}} \bar{\sigma}_i^2}{\sum_{i=1}^{\bar{N}} \bar{\sigma}_i^2} \ge \sqrt{\rho} \right\}, \qquad \mathbf{Y} \doteq span\{\bar{U}_i\}_{i=1}^{L}, \quad \pi_Y \doteq \sum_{i=1}^{L} \bar{U}_i \bar{U}_i'$$

Denote $\bar{\mathbf{\Psi}} \doteq \pi_Y \mathbf{\Psi}$. The associated Galerkin approximation of the original ensemble is then

$$\bar{\mathbf{\Phi}} \doteq \pi_Y \tilde{\mathbf{\Phi}} = \left\{ \bar{\Phi}_i \doteq \sum_{j=1}^{N} f_{i,\,j} \bar{U}_j \,:\, f_i \doteq \bar{U}_i' \tilde{\mathbf{\Phi}} \in \mathbb{R}^N \right\}_{i=1}^{L}$$

Observation V.1 *The projected ensemble satisfies the resolution bound ρ:*

$$\frac{tr\,\pi_Y\tilde{\mathbf{\Phi}}\tilde{\mathbf{\Phi}}'\pi_Y}{tr\,\tilde{\mathbf{\Phi}}\tilde{\mathbf{\Phi}}'} \;\geq\; \rho$$

Proof. The following follows from the constructions in Steps 1 – 4s

$$\begin{aligned} tr\,\pi_Y\tilde{\mathbf{\Phi}}\tilde{\mathbf{\Phi}}'\pi_Y &= tr\,\pi_Y\left(\sum_{k=1}^K\tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Phi}}_k'\right)\pi_Y \geq tr\,\pi_Y\left(\sum_{k=1}^K\pi_{Y_k}\tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Phi}}_k'\pi_{Y_k}\right)\pi_Y \\ &= tr\,\pi_Y\left(\sum_{k=1}^K\mathbf{\Psi}_k\mathbf{\Psi}_k'\right)\pi_Y = tr\,\pi_Y\mathbf{\Psi}\mathbf{\Psi}'\pi_Y = tr\,\bar{\mathbf{\Psi}}\bar{\mathbf{\Psi}}' \geq \sqrt{\rho}\,tr\mathbf{\Psi}\mathbf{\Psi}' \\ &= \sqrt{\rho}\,tr\,\sum_{k=1}^K\mathbf{\Psi}_k\mathbf{\Psi}_k' = \sqrt{\rho}\,tr\,\sum_{k=1}^K\bar{\mathbf{\Phi}}_k\bar{\mathbf{\Phi}}_k' \geq \rho\,tr\,\sum_{k=1}^K\tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Phi}}_k' = \rho\,tr\,\tilde{\mathbf{\Phi}}\tilde{\mathbf{\Phi}}' \quad \square \end{aligned} \tag{7}$$

Concluding comments on computational savings. The algorithm trades a single POD procedure of a set of N samples, by K POD computations for sub-ensembles of size $N_0 = N/K$, and a final POD computation for an ensemble of size $\bar{N} = \sum_{k=1}^K \bar{N}_k$. The computational savings, in general, and an evaluation for the NACA0012 benchmark example we discussed earlier, are summarized in the Table 1.

	General Case	NACA 0012, 99.9% Resolution
State dimension	M	6230
Reference length	N	6400
Number of partitions	K	20
Snapshots per partition	N/K	320
Snapshots in 2^{nd} step	$\bar{\text{N}}$	243
Standard POD	$c\cdot M\cdot N^2$	$c\cdot 2.55189E11$
Approx. POD	$c\cdot M\cdot N^2\left(\frac{1}{K}+\frac{\bar{N}^2}{N^2}\right)$	$c\cdot 1.31269E10$
Savings Factor	$\frac{KN^2}{KN^2+N^2}$	19.4 (i.e, 5.14% of POD)
Compressed dimension	L	37 $=0.06\%$ of M at 99.85% TKE resolution

Table 1. Computational savings of the approximate POD procedure: When the factor K is selected with an anticipation of $\bar{N} < N_0$, the saving factor is at least $\frac{K^2}{K+1}$. In our example, it requires about 5.14% of the computational load of a full POD computation, and provides a 168 fold dimension reduction in the compressed data.

The constant c depends on the specifics of the numerical implementation but is of no consequence at the level of this discussion.It is noticed that a reasonable a priori estimate $\bar{N}_0$ of N_k is often feasible, identically for all k. In that case $\bar{N} \approx KN_0$. Using this approximation, it is easy to see that an optimal selection of K, to maximize the savings factor, is $K_* \doteq \frac{N}{N_0}$. An a posteriori inspection of the NACA 0012 example shows that using $K = 20$, or even higher, would have likely reduced the computational burden even more, reaching a savings of nearly 20 folds, as compared with a traditional POD.

The efficiency of the algorithm is also a function of the intermediate inequalities (7), and, in turn of the non-commutativity of the global π_Y with the local π_{Y_k} . Efficiency is therefore expected to improve under one of two extremes: Either the sub-ensembles $\mathbf{\Phi}_k$ are mutually orthogonal, which is unlikely in fluid applications, or the statistics of individual sub-ensembles are similar to that of the global ensemble. When $\mathbf{\Phi}$ represents multiple operating conditions, it will thus be advantageous to partition the index set $\mathcal{I}$ in a way that each sub-ensemble will represent, more or less equally, all the operating conditions.

C. Time Partition: A Local Mean Algorithm

Again, the starting point is a partition of the ensemble $\mathbf{\Phi}$ into K sub-ensembles and a required resolution bound $\rho \in (0, 1)$, as described in Section A. The variant suggested here concerns the case where the memory is insufficient to store the entire ensemble $\mathbf{\Phi}$. Computation of the global mean, in Step 0, would thus require an additional, time consuming loading of each sample from the hard disk. To avoid this, the local mean is computed first for each sub-ensemble, and is used in the local POD procedure, instead of the global mean. The global mean is evaluated and used only in a counterpart of Step 4.

As above, the effect of this added component on the efficiency of the algorithm depends on the quality of local estimates of the global mean. Pertinent a priori knowledge can thus be exploited in an effective partitioning of the global ensemble.

Step 1 For each $k = 1, \ldots, K$, compute the local sub-ensemble mean $\Phi_{k,0} = mean\,(\mathbf{\Phi_k})$, the perturbation sub-ensemble

$$\widehat{\mathbf{\Phi}}_k = \left\{ \widehat{\Phi}_{k,i} \doteq \Phi_i - \Phi_{k,0} \,:\, i \in \mathcal{I}_k \right\}$$

and an SVD of the corresponding perturbation operator

$$\widehat{\mathbf{\Phi}}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k' = \sum_{i=1}^{N_0} \sigma_{k,i} U_{k,i} V_{k,i}'.$$

Step 2: Truncate the SVD of Step 1 at $N_k \leq N_0$, substituting the original sample sub-ensemble $\widehat{\mathbf{\Phi}}_k$ by the columns of the matrix

$$\breve{\mathbf{\Phi}}_k \doteq \sum_{i=1}^{N_k} \sigma_{k,i} U_{k,i} V_{k,i}' \quad \text{where} \quad N_k \doteq \min \left\{ \tilde{N} \,:\, \frac{\sum_{i=1}^{\tilde{N}} \sigma_{k,i}^2}{\sum_{i=1}^{N_0} \sigma_{k,i}^2} \geq \sqrt{\rho} \right\}, \qquad \pi_{Y_k} \doteq \sum_{i=1}^{N_k} U_{k,i} U_{k,i}'$$

Let $\Phi_0 = mean(\mathbf{\Phi})$ be the yet unknown global mean, let $\tilde{\Phi}_{k,0} \doteq \Phi_{k,0} - \Phi_0$, let $\mathbf{Y}_k \doteq span\{U_{k,i}\}_{i=1}^{N_k}$, and denote $\eta = [1, 1, \ldots, 1]' \in \mathbb{R}^{N_0}$. Given that $0 = mean(\widehat{\mathbf{\Phi}}_\mathbf{k}) = \frac{\mathbf{1}}{\mathbf{N_0}} \widehat{\mathbf{\Phi}}_\mathbf{k} \eta$, the following is a routine probabilistic equality

$$\tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k' = (\widehat{\mathbf{\Phi}}_k + \tilde{\Phi}_{k,0}\eta')(\widehat{\mathbf{\Phi}}_k + \tilde{\Phi}_{k,0}\eta')' = \widehat{\mathbf{\Phi}}_k \widehat{\mathbf{\Phi}}_k' + N_0 \tilde{\Phi}_{k,0} \tilde{\Phi}_{k,0}'$$

The truncated version of $\tilde{\mathbf{\Phi}}_k$ is denoted

$$\bar{\mathbf{\Phi}}_k \doteq \breve{\mathbf{\Phi}}_k + \tilde{\Phi}_{k,0}\eta'$$

Since $mean(\breve{\mathbf{\Phi}}_k) = mean(\pi_{Y_k} \widehat{\mathbf{\Phi}}_\mathbf{k}) = \mathbf{0}$, hence $\breve{\mathbf{\Phi}}_k \eta = 0$, a repetition of the equalities right above, leads to

$$\begin{aligned} tr\, \bar{\mathbf{\Phi}}_k \bar{\mathbf{\Phi}}_k' = tr\, \left(\breve{\mathbf{\Phi}}_k + \tilde{\Phi}_{k,0}\eta' \right) \left(\breve{\mathbf{\Phi}}_k + \tilde{\Phi}_{k,0}\eta' \right)' &= tr\, \left(\breve{\mathbf{\Phi}}_k \breve{\mathbf{\Phi}}_k' + N_0 \tilde{\Phi}_{k,0} \tilde{\Phi}_{k,0}' \right) \\ &\geq \sqrt{\rho}\, tr\, \left(\widehat{\mathbf{\Phi}}_k \widehat{\mathbf{\Phi}}_k' + N_0 \tilde{\Phi}_{k,0} \tilde{\Phi}_{k,0}' \right) = \sqrt{\rho}\, tr\, \tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k' \end{aligned} \tag{8}$$

Step 3: Let $\eta_k \doteq [1, 1, \ldots, 1]' \in \mathbb{R}^{N_k+1}$, and let $\{W_{k,i}\}_{i=1}^{N_k} \subset \mathbb{R}^{N_k+1}$ be an orthonormal basis for $\eta_k^\perp$ (that is, let it be a maximal orthonormal set of zero-mean vectors in $\mathbb{R}^{N_k+1}$). Let $\breve{\mathbf{U}}_k \doteq [U_{k,i}]_{i=1}^{N_k}$, $\breve{\mathbf{\Sigma}}_k \doteq diag\{\sigma_{k,i}\}_{i=1}^{N_k}$ and $\breve{\mathbf{W}}_k = [W_{k,i}]_{i=1}^{N_k}$, and define the alternative truncated sub-ensemble as the columns of the matrix

$$\breve{\mathbf{\Psi}}_k \doteq \breve{\mathbf{U}}_k \breve{\mathbf{\Sigma}}_k \breve{\mathbf{W}}_k' = \sum_{i=1}^{N_k} \sigma_{k,i} U_{k,i} W_{k,i}' \quad \Rightarrow \quad \breve{\mathbf{\Psi}}_k \breve{\mathbf{\Psi}}_k' = \breve{\mathbf{\Phi}}_k \breve{\mathbf{\Phi}}_k'$$

and its (unknown) counterpart

$$\mathbf{\Psi}_k = \breve{\mathbf{\Psi}}_k + \sqrt{\frac{N_0}{N_k + 1}} \Phi_{k,0} \eta_k'$$

In these terms there holds

$$tr\, \mathbf{\Psi}_k \mathbf{\Psi}_k' = tr\, \left(\breve{\mathbf{\Psi}}_k \breve{\mathbf{\Psi}}_k' + N_0 \Phi_{k,0} \Phi_{k,0}' \right) = tr\, \left(\breve{\mathbf{\Phi}}_k \breve{\mathbf{\Phi}}_k' + N_0 \Phi_{k,0} \Phi_{k,0}' \right) \geq \sqrt{\rho}\, tr\, \tilde{\mathbf{\Phi}}_k \tilde{\mathbf{\Phi}}_k'$$

Step 4: The added component, when compared with Step 4 of Section B, is the computation of the global mean from the previously computed sub-ensemble means (here we explicitly use the fact that the sub-ensembles are equal in size[a]):

$$\Phi_0 = mean\, \{\Phi_{k,0}\}_{k=1}^K$$

Having computed Φ_0, the values of the sub-ensembles $\mathbf{\Psi_k}$ becomes explicit. The completion of Step 4, and thus, of the algorithm, therefore proceed precisely as in Section B, leading to the same conclusion:

Observation V.2 *The projected ensemble satisfies the resolution bound ρ:*

$$\frac{tr\, \pi_Y \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Phi}}' \pi_Y}{tr\, \tilde{\mathbf{\Phi}} \tilde{\mathbf{\Phi}}'} \;>\; \rho$$

[a] Alternative approaches, based on sub-ensemble partitioning with variable size and weight will lead to obvious counterparts.

Acknowledgments

The work has been funded by the Deutsche Forschungsgemeinschaft (DFG) under grants 258/1-1 and 258/2-3, by the US National Science Foundation (NSF) under grants 0524070 and 0410246, and by the US Air Force Office of Scientific Research (AFOSR) under grants FA95500510399 and FA95500610373. The authors acknowledge funding and excellent working conditions of the Collaborative Research Centre (Sfb 557) "Control of complex turbulent flow" which is supported by the DFG and hosted at the Technical University Berlin. Stimulating discussions with Laurent Cordier, Rudibert King, Michael Schlegel, Mark Pastoor, Jon Scouten and Tino Weinkauf are acknowledged. We are grateful for outstanding hardware and software support by Lars Oergel and Martin Franke at TU Berlin.

References

[1]B.R. Noack, K. Afanasiev, M. Morzyński, G. Tadmor, and F. Thiele. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.*, 497:335–363, 2003.

[2]P. Holmes, J.L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge, 1998.

[3]L. Sirovich. Turbulence and the dynamics of coherent structures, Parts I–III. *Quart. Appl. Math.*, XLV:561–590, 1987.

[4]J. Seidel, S. Siegel, K. Cohen, and T. McLaughlin. POD based separation control on the NACA0015 airfoil. In *43rd AIAA Aerospace Sciences Meeting and Exhibit*, 2005. paper 2005-297.

[5]J. Ausseur, J. Pinier, and M. Glauser. Flow separation control using a convection based POD approach. In *3rd AIAA Flow Control Conference*, 2006. paper 2006-3017.

[6]Y. Tian, Q. Song, and L. Cattafesta. Adaptive feedback control of flow separation. In *3rd AIAA Flow Control Conference*, 2006. paper 2006-3016.

[7]O. Lehmann, M. Luchtenburg, B.R. Noack, R. King, M. Morzynski, and G. Tadmor. Wake stabilization using POD Galerkin models with interpolated modes. MoA15.2 In *44th IEEE Conference on Decision and Control and European Control Conference ECC, Seville, Spain*, 12.-15. December 2005.

[8]M. Morzynski, W. Stankiewicz, B. R. Noack, F. Thiele, and G. Tadmor. Generalized meanfield model with continuous mode interpolation for flow control. In *3rd AIAA Flow Control Conference, San Francisco, CA*, 5-8 June 2006. paper AIAA-2006-3488, Invited.

[9]J. Borggaard, A. Hay, and D. Pelletier. Interval-based reduced order models for unsteady fluid flow. *Int. J. Num. Anal. & Modeling*, , 2006. submitted.

[10]A Maurel, V Pagneux, and J.E. Wesfreid. Mean-flow correction as a nonlinear saturation mechanism. *Europhysics Letters*, 32:217 – 222, 1995.

[11]J. Gerhard, M. Pastoor, R. King, B.R. Noack, A. Dillmann, M. Morzyński, and G. Tadmor. Model-based control of vortex shedding using low-dimensional Galerkin models. In *33rd AIAA Fluids Conference and Exhibit*, Orlando, Florida, U.S.A., June 23–26, 2003, 2003. Paper 2003-4262.

[12]S. Siegel, K. Cohen, and T. McLaughlin. Feedback control of a circular cylinder wake in experiment and simulation. In *33rd AIAA Fluids Conference and Exhibit*, Orlando, Florida, U.S.A., June 23–26, 2003, 2003. Paper No 2003-3571.

[13]G. Tadmor, B. R. Noack, O. Lehmann, M. Morzynski, W. Stankiewicz, and J. Gonzalez. Shift modes and transient dynamics in low order, design oriented Galerkin models. In *45th AIAA Aerospace Sciences Meeting and Exhibit*, 8 - 11 January, 2007. Paper AIAA 2007-111.

[14]B.R. Noack, G. Tadmor, and M. Morzyński. Low-dimensional models for feedback flow control. Part I: Empirical Galerkin models. In *2nd AIAA Flow Control Conference*, Portland, Oregon, U.S.A., June 28 – July 1, 2004. AIAA Paper 2004-2408 (invited contribution).

[15]C.L. DeMarco and G.C. Verghese. Bringing phasor dynamics into the power system load flow. In *North American Power Symposium*, pages 463–469, 1993.

[16]V. Venkatasubramanian, H. Schättler, and J. Zaborsky. Fast time-varying phasor analysis in the balanced three phase large electric power system. *IEEE Transactions on Automatic Control*, TAC – 40:1975 – 1982, 1995.

[17]G. Tadmor. On approximate phasor models in dissipative, bilinear systems. *IEEE Transaction on Circuits & Systems I: Fundamental Theory*, 49:1167 – 1179, 2002.

[18]C. Beattie, J. Borggaard, S. Gugercin, and T. Iliescu. A domain decomposition approach to POD. In *Proc. of the 45th IEEE Conference on Decision and Control*, 2006.

[19]K. Cohen, S. Siegel, J. Seidel, and T. McLaughlin. Reduced order modeling for closed-loop control of three dimensional wakes. In *3rd AIAA Flow Control Conference*, 2006. paper 2006-3356.

[20]M. Schatz, B. Günther, and F. Thiele. Computational investigation of separation control over high-lift airfoil flows. In R. King, editor, *Active Flow Control Conference, Berlin*. Springer-Verlag, 2006.

[21]K. Kaepernick, L. Koop, and K. Ehrenfried. Investigation of the unsteady flow field inside a leading edge slat cove. In *11th AIAA/CEAS Aeroacoustics Conference*, 2005. Paper 2005-2813.

[22]T. Rung and F. Thiele. Computational modelling of complex boundary layer flows. In *9th int. Symp. on transport phenomena in thermal-fluid engineering*, pages 321–326, 1996.

[23]M. Bergmann, L. Cordier, and J.-P. Brancher. Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced order model. *Phys. Fluids*, 17:097101–097121, 2005.

[24]G. Tadmor, B.R. Noack, M. Morzyński, and S. Siegel. Low-dimensional models for feedback flow control. Part II: Controller design and dynamic estimation. In *2nd AIAA Flow Control Conference*, Portland, Oregon, U.S.A., June 28 – July 1, 2004. AIAA Paper 2004-2409 (invited contribution).

[25]B Protas and J.E. Wesfreid. On the relation between the global modes and the spectra of drag and lift in periodic wake flows. *Competes Rendus Mecanique*, 331:49 – 54, 2003.

[26]M. Morzyński, S. Goujon-Durand, B.R. Noack, J.E. Wesfreid, and B. Thiria. Numerical analysis of the wake control behind a circular cylinder with oscillatory rotation. In *Proceedings of XXI International Congress of Theoretical and Applied Mechanics 2004*, Kluwer/Springer, Warschau, Polen, 15.–21. August 2004, 2005.

[27]B.R. Noack, P. Papas, and P.A. Monkewitz. The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.*, 523:339–365, 2005.

[28]A.E. Deane, I.G. Kevrekidis, G.E. Karniadakis, and S.A. Orszag. Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders. *Phys. Fluids A*, 3(10):2337–2354, 1991.