

Strategies for model reduction: comparing different optimal bases

D.T. Crommelin and A.J. Majda

29th March 2004

Courant Institute of Mathematical Sciences
and Center for Atmosphere Ocean Science,
New York University
251 Mercer Street
New York, NY 10012
email: crommelin@cims.nyu.edu

Revised version

Submitted to J. Atmos. Sci.

Abstract

Several different ways of constructing optimal bases for efficient dynamical modelling are compared: Empirical Orthogonal Functions (EOFs), Optimal Persistence Patterns (OPPs) and Principal Interaction Patterns (PIPs). Past studies on fluid-dynamical topics have pointed out that EOF-based models can have difficulties reproducing behaviour dominated by irregular transitions between different dynamical states. We address this issue in a geophysical context, by assessing the ability of these strategies for efficient dynamical modelling to reproduce the chaotic regime transitions in a simple atmosphere model. The atmosphere model is the well-known Charney-DeVore model, a 6-dimensional truncation of the equations describing barotropic flow over topography in a beta-plane channel geometry. This model is able to generate regime transitions for well-chosen parameter settings. The models based on PIPs are found to be superior to the EOF- and OPP-based models, in spite of some undesirable sensitivities inherent to the PIP method.

1 Introduction

Within the climate research community, a need is felt for efficient models which produce realistic dynamics using as few degrees of freedom as possible. The investigation of the coupled atmosphere-ocean system is limited by the unfortunate fact that spectral atmosphere models have to be quite detailed in order to generate the reasonably realistic behaviour needed in many studies of the climate system. A detailed, complex spectral atmosphere model slows down the numerical integration of coupled atmosphere-ocean models enormously, thus hampering the study of the climate system on long timescales. It is known, however, that spectral models are very inefficient: their dynamics can be generated by far simpler models, if only a suitable model basis is chosen. Finding such suitable, or optimal, bases is therefore of obvious interest.

A number of studies have been devoted to the construction and use of optimal bases for atmospheric modelling. Typically, these studies arrive at a reduced model for atmospheric flow in two steps. First, an optimal basis is chosen and calculated, and the atmosphere model to be reduced is transformed to the new basis. Empirical Orthogonal Functions (EOFs) are the most common choice for the optimal basis (Rinne and Karhilla 1975; Schubert 1985, 1986; Selten 1993, 1995, 1997a,b; Achatz and Branstator 1999; D’Andrea and Vautard 2001; Achatz and Opsteegh 2003a,b), but other choices such as Principal Interaction Patterns (Achatz et al. 1995; Kwasniok 1996, 2004) and Optimal Persistence Patterns (DelSole 2001) have also been made. In the second step the transformed model is truncated, and some kind of closure scheme is applied to the truncated model to account for the effect of the unresolved degrees of freedom on the still resolved modes. The closure ranges from adding extra damping (Selten 1995) to empirical fitting of the forcing and

linear terms (Achatz and Branstator 1999) and from calculating linear, nonlinear and stochastic correction terms (Majda et al. 1999, 2003) to optimization of deterministic model coefficients (Kwasniok 2003).

A separate class of reduced models consists of linear models with stochastic forcing, in which the use of an optimal basis is combined with a closure that uses linear and stochastic correction terms to represent not only unresolved modes but also nonlinear processes (e.g. Branstator and Haupt 1998; Winkler et al. 2001).

Under the name Proper Orthogonal Decomposition (POD), or Karhunen-Loève (KL) expansion, the technique of EOFs is also used in other fluid-dynamical contexts to arrive at reduced models (e.g. Aubry et al. 1988, Sirovich 1989, Cazemier et al. 1994). It is there that a serious shortcoming of EOF- (or POD-) models was first noticed. The amount of variance of a system represented by the leading n EOFs is often taken as an indication of the quality of a reduced model using those first n EOFs. If n EOFs describe, say, 99% of the variance, one (naively) expects the reduced model using n EOFs to be nearly perfect. These expectations were severely contradicted in a study by Aubry et al. (1993), in which POD-models of the Kuramoto-Sivashinsky equation were studied. They found that a model based on the leading 6 POD modes could not reproduce the right dynamics, even though those 6 POD modes represent 99.9995% of the variance. Similar problems with models based on POD-modes were reported by Armbruster et al. (1992) in a study of Kolmogorov flow in a regime of bursting behaviour. Modes representing only a tiny amount of variance can be crucial in the generation of certain types of dynamics. In particular systems that exhibit sudden transitions between different states (i.e. bursting behaviour) will be susceptible to this kind of problems when trying to model them using

EOFs or POD-modes. The modes excited during the transitions do not represent a large amount of variance, yet they are crucial in generating the right dynamics.

Although the atmosphere does not possess extreme forms of bursting behaviour, it is nevertheless marked by episodes of more and less turbulent behaviour. The large-scale circulation can be caught for a while in some flow configuration (or regime), before it makes a relatively swift transition to another state. Having in mind the dramatic failure of EOF-models as reported by Aubry et al. one can wonder whether EOF-models of the atmosphere will similarly have problems reproducing atmospheric regime behaviour, and whether other choices of optimal bases will perform differently. In this paper we will look into this issue by comparing various optimal bases in their ability to reproduce the regime behaviour generated by a simple atmosphere model. The model we shall use is the well-known Charney-DeVore model (Charney and DeVore 1979, CDV hereafter). Long considered to be a model possessing regime-like steady states but unable to produce transitions between these regimes, Crommelin et al. (2004) show that the CDV model can, by itself, generate regime transitions at realistic parameter settings. They find the transitions to be guided by a perturbed heteroclinic cycle connecting the steady states of the CDV model. This cycle is due to the interaction of barotropic and topographic instabilities, the two instability mechanisms present in the model.

After introducing the model and its regime behaviour in section 2, we compare 3 different types of optimal bases: Empirical Orthogonal Functions in section 3, Optimal Persistence Patterns in section 4 and Principal Interaction Patterns in section 5. The focus is on the ability of the reduced models, formulated in terms of these optimal bases, to reproduce the chaotic regime transitions of the CDV model at the parameter settings described in

section 2. All reduced models are obtained by projecting the CDV model onto one of the optimal bases, and then truncating to the desired number of basis patterns. No closure scheme of any kind is applied, since we want to study the quality of optimal bases, not of closure schemes. Moreover, there is not much of a physical rationale for applying a closure scheme to reduced versions of the CDV model: being low-order models, they do not possess a cascade or small-scale processes, so there is no argument for closures based on e.g. eddy viscosity. Thus, in this paper we study only bare truncations.

2 Charney-DeVore model with regime transitions

The starting point for the comparison of various optimal bases will be the 6-dimensional truncation of the equations for barotropic flow in a β -plane channel with orography that is known as the Charney-DeVore model (Charney and DeVore 1979, CDV hereafter). The formulation of the model as it is used here was presented by De Swart (1988, 1989), who used a slightly different scaling and a more general zonal forcing profile than CDV. The model can show rapid transitions between flow regimes if the parameters of the model are carefully chosen (Crommelin et al. 2004).

The set of ordinary differential equations that makes up the 6-dimensional model is as

follows:

$$\begin{aligned}
\dot{x}_1 &= \gamma_1^* x_3 - C(x_1 - x_1^*) \\
\dot{x}_2 &= -(\alpha_1 x_1 - \beta_1) x_3 - C x_2 - \delta_1 x_4 x_6 \\
\dot{x}_3 &= (\alpha_1 x_1 - \beta_1) x_2 - \gamma_1 x_1 - C x_3 + \delta_1 x_4 x_5 \\
\dot{x}_4 &= \gamma_2^* x_6 - C(x_4 - x_4^*) + \varepsilon(x_2 x_6 - x_3 x_5) \\
\dot{x}_5 &= -(\alpha_2 x_1 - \beta_2) x_6 - C x_5 - \delta_2 x_4 x_3 \\
\dot{x}_6 &= (\alpha_2 x_1 - \beta_2) x_5 - \gamma_2 x_4 - C x_6 + \delta_2 x_4 x_2
\end{aligned} \tag{2.1}$$

The model coefficients are given by

$$\begin{aligned}
\alpha_m &= \frac{8\sqrt{2}}{\pi} \frac{m^2}{4m^2 - 1} \frac{b^2 + m^2 - 1}{b^2 + m^2}, & \beta_m &= \frac{\beta b^2}{b^2 + m^2}, \\
\delta_m &= \frac{64\sqrt{2}}{15\pi} \frac{b^2 - m^2 + 1}{b^2 + m^2}, & \gamma_m^* &= \gamma \frac{4m}{4m^2 - 1} \frac{\sqrt{2}b}{\pi}, \\
\varepsilon &= \frac{16\sqrt{2}}{5\pi}, & \gamma_m &= \gamma \frac{4m^3}{4m^2 - 1} \frac{\sqrt{2}b}{\pi(b^2 + m^2)}.
\end{aligned} \tag{2.2}$$

A timestep of $\Delta t = 1$ is interpreted as 1 day. For a detailed account of the derivation of this model, see De Swart (1988, 1989).

In the model equations, one can recognise advection by the zonal flow with components x_1, x_4 (terms with α_i, δ_i and ε), the β -effect (terms with β_i), topographic interaction terms (γ_i, γ_i^*), Ekman damping (C -terms) and zonal forcing (x_1^*, x_4^*). The free parameters in the model determine the damping timescale (C), the zonal forcing (x_1^* and x_4^*), the topographic height (γ), the beta-effect (β) and the length-width ratio of the beta-channel (b).

Crommelin et al. (2004) found that the model shows transitions between flow regimes if the parameters are set to $(x_1^*, x_4^*, C, \beta, \gamma, b) = (0.95, -0.76095, 0.1, 1.25, 0.2, 0.5)$. For a discussion and interpretation of these parameter settings, see their paper. Here we only point out that the regime behaviour in this model, present at realistic parameter values, is

due to the combination of topographic and barotropic instabilities. The latter is the result of the more general zonal forcing than was used in CDV.

In figure 1 (top) the datapoints are shown of one 4000-day long integration of the model (output every 0.1 day, initial transient of 500 days not shown). In order to assess the size of the basin of attraction of the attractor to which the system is drawn in this one particular integration, 40 000 integrations were made, each starting from randomly chosen initial conditions (each $x_i(t = 0)$ drawn from a uniform distribution on the interval $(-1, 1)$). Each integration is 2000 days long. All 40 000 end points are shown in figure 1 (bottom). Together they give a good coverage of the attractor shown in the top panel, so the basin of attraction has considerable size, and may very well be the entire phase space.

In figure 2 (bottom), a piece of one model integration projected onto its first EOF is shown, making the regime transitions clearly visible. In Crommelin et al. (2004) this regime behaviour was related to the formation of a heteroclinic cycle between steady states, at parameter settings for which the onset of barotropic instability coincides with the onset of topographic instability. The structure of the heteroclinic cycle is such that the minimum number of phase-space dimensions needed for the embedding of this cycle is three. We therefore cannot expect deterministic reduced models with less than three dimensions to be able to faithfully reproduce the regime behaviour. Conversely, the cycle structure suggests that 3 degrees of freedom should be enough for a reduced model to reproduce the cycle.

The swift regime transitions make this model a good test case for model reduction strategies and optimal bases used in atmospheric science, as the model has (far) more geophysical relevance than the Kuramoto-Sivashinsky equation or Kolmogorov flow equations. After all, the model describes atmospheric flow, albeit severely simplified. In the following

sections three different types of optimal bases will be tested, to see if they can reproduce the chaotic regime transitions described above.

3 Empirical Orthogonal Functions

The technique of calculating EOFs and using them as a model basis is well known and will not be explained in detail here. Having chosen a metric M_k , the EOFs \mathbf{p}_i are simply the eigenvectors of the eigenvalue problem

$$C M_k \mathbf{p}_i = \lambda_i^2 \mathbf{p}_i \quad (3.1)$$

in which C is the covariance matrix:

$$C_{ij} = \overline{(x_i - \bar{x}_i)(x_j - \bar{x}_j)}. \quad (3.2)$$

If the time-mean $\bar{\mathbf{x}}$ is not subtracted when calculating C , the leading EOF will to a large extent coincide with the time-mean state. Leaving out the subtraction of $\bar{\mathbf{x}}$ did not improve the performance of the EOF models; those results are therefore not shown here.

Table 1 gives the variance spectra associated with the EOFs, using alternately the streamfunction L2-norm M_0 and the kinetic energy norm M_1 . These norms are defined by

$$\int \psi^2 = \mathbf{x}^T M_0 \mathbf{x}, \quad (3.3)$$

$$\frac{1}{2} \int \psi \Delta \psi = \mathbf{x}^T M_1 \mathbf{x}, \quad (3.4)$$

in which \mathbf{x} is the vector of variables of the CDV model, ψ the associated streamfunction field, and the integration runs over the physical domain of the model. The EOFs were calculated from a 105000-day integration of the original model, 1 datapoint each day, leaving

out an initial transient period of 5000 days. The cumulative variance associated with the leading three EOFs is very high: 97 % or more. In spite of the minimal variance associated with the trailing EOF modes, their inclusion in the model can be crucial in reproducing the regime transitions with an EOF model. A 5 EOF model could not reproduce the regime transitions, independent of the norm used. Interestingly, a 4 EOF model with EOFs calculated using the M_1 norm. was able to generate regime transitions. Using the M_0 norm, no regime transitions were observed with 4 EOFs. Models using 3 EOFs could not reproduce regime transitions.

The 4 EOF model that was able to generate transitions did so in a much too regular way: it reproduced the transitions between different parts of the attractor seen in the original CDV model, but not their irregular or chaotic appearance. Rather, its behaviour was periodic. A really faithful reproduction of the dynamics of the CDV model, including its chaotic nature, was not seen for any of the EOF models. The EOF models that could not reproduce the transitions were either drawn to some stable fixed point, or to a periodic solution not having any characteristic of regime-like behaviour. To summarize the asymptotic behaviour of the various models, their properties are listed in table 2. To be able to do so, ensemble integrations were made for all models, in much the same way as for the original CDV model previously: 2000 integrations, each 4000 days long, starting from randomly chosen initial conditions.

Figure 2 gives an impression of the periodic regime behaviour generated by the 4 EOF model, the EOFs of which were calculated using the M_1 norm. The upper left panel shows the result of a 2000-day integration of that model, projected onto the plane spanned by the leading two EOFs. For comparison, the output of the CDV model is projected onto

the same plane, see upper right panel. In the lower two panels timeseries are plotted, both for the 4 EOF model and the CDV model. It is clear that the regime transitions generated by the reduced EOF model are much too regular.

4 Optimal Persistence Patterns

The previous section made clear that models based on EOFs have indeed difficulties reproducing transitional behaviour, as observed earlier by Aubry et al. (1993). Even using 5 EOFs, representing up to 99.8 % of the variance, no model could reproduce the chaotic regime behaviour of the original CDV model. In this section models using another optimal basis will be investigated, to see whether they will do better. This time, the Optimal Persistence Patterns (OPPs) proposed by DelSole (2001) will serve as a basis for the models. These patterns are chosen to maximize either of two measures of persistence, both related to the decorrelation time of the system:

$$T_1 = \int_0^\infty \rho(\tau) d\tau \quad \text{or} \quad T_2 = 2 \int_0^\infty \rho^2(\tau) d\tau, \quad (4.1)$$

in which $\rho(\tau)$ is the correlation function depending on time lag τ . The idea is, given a dataset $\mathbf{g}(t)$ in some phase space \mathcal{P} , to find a vector $\mathbf{e}_1 \in \mathcal{P}$ such that the time series $v_1(t) = \mathbf{e}_1^T \mathbf{g}(t)$ has maximal T_1 (or T_2), then a second vector \mathbf{e}_2 , orthogonal in some sense to \mathbf{e}_1 , that again maximizes T_1 or T_2 , and so on. The ordering of the patterns based on their persistence or correlation time makes the OPPs an interesting type of optimal basis. If one aims to reproduce the long timescale behaviour of a system, a set of patterns with maximal correlation times is a natural candidate for the basis of a reduced model.

The system under investigation, the CDV model in a dynamical regime of chaotic

transitions, has significant oscillatory aspects to its behaviour. We therefore use T_2 to calculate optimals rather than T_1 , as the former measure is more suitable to use on systems with oscillatory correlation functions. The dataset, generated by the CDV model, that is used for the calculations is the same as the one used for the calculation of the EOFs in the previous section: 100 000 datapoints, each 1 time unit apart, after an initial transient period of 5000 days. It will be denoted by \mathbf{x}^{CDV} instead of $\mathbf{g}(t)$ from now on. The set of optimals is calculated by maximizing T_2 under the constraint that the patterns \mathbf{e}_i are mutually orthogonal in time. That is, they are orthogonal using the lag-zero covariance matrix C_0 as metric: $\mathbf{e}_j^T C_0 \mathbf{e}_i = 0$ if $i \neq j$.

DelSole refers to the patterns \mathbf{e}_i as filter patterns, and identifies the actual optimal persistence patterns as $\mathbf{r}_i = C_0 \mathbf{e}_i$. The maximally persistent time series $v_i(t)$ are the expansion coefficients for the expansion of the dataset \mathbf{x}^{CDV} in terms of the OPPs:

$$\mathbf{x}^{\text{CDV}}(t) = \sum_{i=1}^L \mathbf{r}_i v_i(t), \quad v_i(t) = \mathbf{e}_i^T \mathbf{x}^{\text{CDV}}(t). \quad (4.2)$$

Orthogonality in time implies that the covariance of two timeseries $v_i(t), v_j(t)$ is zero:

$$\langle v_i(t) v_j(t) \rangle = \mathbf{e}_i^T C_0 \mathbf{e}_j = 0, \quad i \neq j, \quad (4.3)$$

with $\langle \rangle$ denoting time-average. For details of the calculation, see DelSole (2001).

The resulting six optimal patterns, and their associated values of T_2 , are robust when varying the initial guesses needed for the minimization routine; however, the order in which we find them is not. That is, the set of patterns is robust but not always by itself entirely ordered according to descending T_2 . This may be due to the existence of local maxima for T_2 . The six patterns break up into three groups: one pattern with $T_2 = 92.8$, two

with $T_2 = 15.3, 13.8$ and three with $T_2 = 7.5, 6.3, 5.7$. We arrange the OPPs in order of descending T_2 (i.e. the first OPP has largest T_2).

Figure 3 shows autocorrelation functions for the CDV model dataset $\mathbf{x}^{\text{CDV}}(t)$ and for the timeseries $\{\mathbf{v}_i\}$ obtained using (4.2). For comparison, the autocorrelation functions of the EOFs (norm M_1 used) are also shown. As can be seen, the OPP technique works well in isolating the pattern (the first OPP) responsible for the long timescale correlation in the system. Several of the original variables project significantly onto the first OPP (in particular x_1, x_3, x_4) and have therefore rather long ‘tails’ in their correlation functions.

Having determined the set of filter patterns $\{\mathbf{e}_i\}$ and OPPs $\{\mathbf{r}_i\}$ from the dataset $\mathbf{x}^{\text{CDV}}(t)$, we use them as a new basis for the CDV model. This means we expand the CDV model variables $\mathbf{x}(t)$ on the OPP basis to arrive at equations describing the time evolution of the expansion coefficients $\mathbf{v}(t)$:

$$\dot{v}_i(t) = \mathbf{e}_i^T \dot{\mathbf{x}}(t) \quad (4.4)$$

This defines a projection from the 6-dimensional phase-space of the CDV model into the L -dimensional phase space of a (reduced) OPP-model ($L \leq 6$). The coefficients $v_i(t)$, $i = 1, \dots, L$, are the variables of the OPP-based model.

Projection onto the leading OPPs results in reduced OPP-based models none of which is able to reproduce the chaotic regime transitions observed in the CDV model. Using 5 OPPs, the model is attracted to either a fixed point or a periodic solution. With 4 OPPs, the model is unbounded; the 3 OPP model is attracted to one fixed point. These conclusions are drawn based on ensemble integrations with 2000 members and randomly chosen initial conditions.

Although the regime behaviour of the CDV model can be considered a long timescale

phenomenon, it is possible that some short timescale processes are indispensable for generating transitions. Other combinations of OPPs than those with largest T_2 may therefore give better results. Trying all possible combinations of 3, 4 and 5 OPPs did not yield a single OPP-based reduced model able to reproduce the chaotic regime transitions. Most of them were unbounded or were attracted to a fixed point. Only one combination was partially able to generate periodic regime transitions: the 5-OPP model consisting of variables v_1, v_2, v_3, v_5, v_6 was drawn either to a periodic solution with a structure resembling the CDV model attractor, or to a nearby fixed point. Figure 4 shows the endstates from 40000 integrations with this 5-OPP model, starting from random initial states. The results were projected back onto the original CDV model variables, so they can be easily compared with the results of the CDV model itself as depicted in figure 1. For the projection shown in figure 4, the fixed point is located at $(x_1, x_4) \approx (0.725, -0.320)$. It attracted more than half of the 40000 initial states.

5 Principal Interaction Patterns

5.1 Outline of the method

The technique of Principal Interaction Patterns, or PIPs, was introduced by Hasselmann (1988) and refined by Kwasniok (1996, 1997, 2001, 2004). A technique similar to PIPs is presented by Wu (1996). The calculation of PIPs takes into account the dynamics of the model for which one tries to find an efficient description; PIP-based models therefore can be expected to be more suitable than EOF-based models to reproduce the behaviour of some complex, high-dimensional model. Extensive accounts of the way to calculate PIPs

can be found in the papers by Kwasniok; in particular Kwasniok (1997) gives many details. Here we only give a brief review of the method.

Assume we have a model represented by an n -dimensional set of coupled ODEs,

$$\dot{x} = F(x), \quad x^T = (x_1, \dots, x_n). \quad (5.1)$$

Let x_p denote the projection of x onto a number of PIPs. The projection of system (5.1) onto the PIPs yields a reduced system:

$$\dot{x}_p = F_p(x_p) \quad (5.2)$$

If we integrate the system (5.1) over time τ , starting from initial state x^0 , we end up in x^τ . Projection of the initial state yields x_p^0 . Now we integrate the PIP system (5.2) from this projected initial state x_p^0 and end up in x_p^τ . The difference at time τ between the state of the PIP model (5.2) and that of the original model (5.1) is denoted by

$$d^\tau = x_p^\tau - x^\tau. \quad (5.3)$$

We can integrate the norm of the difference:

$$Q = \int_0^{\tau_{\max}} [d^\tau, d^\tau] d\tau. \quad (5.4)$$

(A slight variation is made in Kwasniok (2004), where Q is defined as $Q = [d^{\tau_{\max}}, d^{\tau_{\max}}]$, i.e. as the difference between the endpoints rather than the integrated difference.) Q depends on the details of the projection P , on the initial state x^0 and on the integration time τ_{\max} . Taking the ensemble average of Q over all initial states x^0 on the attractor results in the error function χ :

$$\chi(\tau_{\max}, P) = \langle Q(x^0, \tau_{\max}, P) \rangle \quad (5.5)$$

Finding PIPs boils down to minimizing χ under variation of P . In principle, one could also consider the parameters of the PIP model (5.2) as independent variables in the minimization procedure, rather than have them determined by projection of system (5.1) according to P . This is done in Kwasniok (2004). In this way, the minimization procedure not only generates the optimal patterns but also acts as a kind of closure algorithm. However, for the current study we did not use this extended type of PIP calculation. Our results were obtained with PIP models that resulted from projection and truncation, without the application of a closure scheme.

The integration time τ_{\max} remains undetermined; it can be chosen on the basis of some physical argument or other consideration. The resulting PIP system can be quite sensitive to the choice of τ_{\max} , see Kwasniok (2004). We will come back to this issue later on.

For the actual computation of the PIPs the gradient of the error function χ with respect to the PIP-coefficients is needed. An expression for that gradient is derived in Kwasniok (1997). Furthermore, we need to choose τ_{\max} , a metric M defining our inner product $[\cdot, \cdot]$ and some constraints on the PIPs. The latter is necessary since each PIP-model allows for a linear transformation of the basis vectors (i.e. the PIPs) resulting in an equivalent PIP-model. To remove this ambiguity, Kwasniok imposes constraints on the set of PIPs: the patterns must be orthonormal, and their amplitudes mutually uncorrelated.

5.2 Results

The PIP method was applied to the CDV model, using again the dataset consisting of 100 000 datapoints also used for the calculation of EOFs and OPPs. The metric was chosen to be the kinetic energy metric. As initial guess for the minimization of χ we used the EOFs

calculated using the same metric. The ensemble average was taken over N_{int} ensemble members (i.e. the PIP model was integrated N_{int} times for each calculation of χ).

In practice, the behaviour of the PIP model resulting from the minimization procedure not only depends on the choice of τ_{max} , but also on N_{int} . To reduce the computation time, usually not the full dataset of 100 000 points was used, but only a segment of it. In other words: usually $N_{\text{int}}\tau_{\text{max}} < 100000$. As we will see, the best results were obtained using only the first 1% of the datapoints of the full dataset.

It is not a priori clear what should be the measure of performance of a PIP model. We could use the error χ as a measure, but that does not give an indication whether the PIP model can reproduce the chaotic regime behaviour of the original 6-dim. model. This may not be a big surprise, as the regime behaviour is a low-frequency phenomenon, whereas χ measures the error developing in a time-interval $[0, \tau_{\text{max}}]$. Thus, χ measures the “predictive skill” of the PIP model rather than its ability to reproduce the “climate statistics” of the original model. Choosing τ_{max} to be very large (e.g. 1000) is unlikely to solve this problem: Kwasniok (2004) notes that PIP models derived with large τ_{max} have too little variance. Besides, we do not require the PIP model to be able to follow the orbit of the original model for a long time (that would be a too stringent requirement), but rather to show *grosso modo* the right low-frequency behaviour.

For lack of a better criterion, we judge the performance of the PIP models (that is, their ability to reproduce the chaotic regime behaviour of the original CDV model) qualitatively by eye and quantitatively by the power spectra of their time series. We found models with 5, 4 and 3 PIPs able to reproduce the regime behaviour. Timeseries resulting from integrations of these models are shown in figures (5)-(7). The 5-PIP model was obtained

using $\tau_{\max} = 20$, $N_{\text{int}} = 50$ and yielded $\chi = 0.014$; for the 4-PIP model $\tau_{\max} = 10$, $N_{\text{int}} = 100$ was used, giving $\chi = 0.19$. The 3-PIP model was calculated with $\tau_{\max} = 25$, $N_{\text{int}} = 40$ and yielded $\chi = 0.37$. Similar to Kwasniok (1996, 1997) we found the differences between PIPs and EOFs to show up in the trailing PIPs. The first few PIPs were almost identical to the leading EOFs.

For these three different PIP models generating chaotic regime behaviour, power spectral densities (PSDs) are shown in figure 8. The PSDs were calculated from the timeseries of the first PIPs (shown in figures 5 - 7). For comparison, the PSD resulting from projecting the CDV model data onto the same first PIP of the 5-PIP model is also shown (using the first PIP from the 4- or 3-PIP models gave almost identical PSDs). As can be seen, all three PIP models reproduce the main features of the CDV data PSD : similar spectral power at low frequencies, and a rapid decrease of power for frequencies higher than 0.1 days^{-1} . The 3- and 4-PIP models show extra spectral power between 0.2 and 0.3 days^{-1} , not present in the 6-dim. model. The maximum of the PSD of the CDV model data, between 0.003 and 0.004 days^{-1} , is less prominent in the PIP models; there the PSD maxima are at somewhat higher frequencies ($0.008 - 0.009 \text{ days}^{-1}$).

5.3 Method sensitivities

The PIP method works well for the test case under consideration in this paper. However, the results depend sensitively on a few parameter choices that must be made for doing the PIP calculation. The problem of choosing τ_{\max} is known, and has been discussed previously by Kwasniok (2004). For our test case, we have found no other way of determining a suitable τ_{\max} than by trial-and-error. No a priori identifiable, physically relevant timescale

was available, nor have we found significant coherence among the values of τ_{\max} giving good results for the 5-, 4- and 3-PIP models (τ_{\max} ranging from 10 to 25).

In table 3 an overview is given of a number of different PIP models, with varying τ_{\max} and N_{int} . As can be seen there, the succesful 5-PIP model with $\tau_{\max} = 20, N_{\text{int}} = 50$ is somewhat robust under changes in τ_{\max} and N_{int} . In contrast, the 4-PIP models cannot reproduce the chaotic regime behaviour anymore when modest changes from $\tau_{\max} = 10$ or $N_{\text{int}} = 100$ are made. The 3-PIP model still generates the right behaviour when changing N_{int} from 40 to 30 or 50, but not when changing τ_{\max} away from 25.

Another interesting aspect to be seen in table 3 is the disappointing performance of models calculated with large N_{int} . A large number of ensemble members in the calculation of the error χ does not guarantee a good PIP model - on the contrary. We suspect that a too large N_{int} leads to a kind of overdetermination in the PIP calculation. Something similar is likely to happen when τ_{\max} is too large. Kwasniok (2004) reports that his models are too much damped when τ_{\max} is too large. In table 3 it can be seen that the 4- and 3-PIP models calculated with $\tau_{\max} = 100$ and $N_{\text{int}} = 10$ end up in a fixed point. However, the 5-PIP model with those parameters still produces the right behaviour.

A final sensitivity of the PIP calculation relates to the way the ensemble members are chosen. For our PIP calculations we took N_{int} consecutive segments, each of length τ_{\max} , from the dataset of the CDV model (recall the CDV model dataset consists of one long forward numerical integration). Thus, those segments are not completely independent of each other. Redoing the calculations with more independent segments (by having a considerable amount of time, about 1000 timesteps, in between those segments), for 4 PIPs with $\tau_{\max} = 10, N_{\text{int}} = 100$ and for 3 PIPs with $\tau_{\max} = 25, N_{\text{int}} = 40$ resulted in

models showing periodic behaviour without regime characteristics. The model with 5 PIPs and $\tau_{\max} = 20$, $N_{\text{int}} = 50$, calculated using independent segments, still generated chaotic regime transitions. Thus, fitting the PIP model to a dataset containing a few complete instances of the regime transition cycles apparently works better than fitting it to a dataset with numerous little pieces of those cycles.

It is likely that this sampling sensitivity is to some extent related to the dynamics of the CDV model. At the chosen parameter settings, the CDV model is only weakly mixing, reflected in its long timescale correlations, and its attractor is highly inhomogeneous. Because of this inhomogeneity, two sample sets can give different outcomes when used in the PIP calculation: the two sets represent various regions of the attractor in different ways. Realistic systems have attractors that are generally more homogeneous than the CDV model attractor, and the sampling sensitivity reported here can well be less severe for such systems. Nevertheless, even high-dimensional, realistic systems seldom have completely homogeneous attractors, and may therefore be prone to some form of sampling sensitivity when constructing reduced models.

6 Conclusion

Three different types of optimal bases have been tested in this study: Empirical Orthogonal Functions, Optimal Persistence Patterns and Principal Interaction Patterns. All three types have been used to construct reduced model versions of a simple atmosphere model generating transitions between flow regimes (the Charney-DeVore model with a zonal forcing allowing barotropic instability). Previous studies on reduced models of the Kuramoto-

Sivashinsky equation and of Kolmogorov flow have made clear that EOF-based models can have difficulties to reproduce behaviour involving quick transitions between different dynamical states. These problems were also encountered in this study: EOF-based reduced models were not able to reproduce the chaotic regime behaviour of the original CDV model, although the EOFs represented up to 99.8% of the CDV model variance. OPPs did not turn out to be a good alternative for EOFs in this case, as the OPP-based models performed worse than the EOF-based models. One model, based on 5 OPPs, produced an attracting periodic solution resembling the structure of the CDV model attractor, but it also had a steady state with a large basin of attraction. None of the other models based on 3, 4 or 5 OPPs was able to produce regime transitions, periodic nor chaotic; many of them even gave unbounded solutions. The OPP technique works well, however, for identifying patterns with long timescale correlations.

Models based on PIPs performed best in our test case. It was possible to find 5-, 4- and even 3-PIP models able to generate the chaotic regime behaviour seen in the CDV model. These models reproduced the main features of the power spectral density of the CDV model data projected onto the leading PIP. Nevertheless, the PIP technique involves some sensitivities that are difficult to understand. The best choice for the time integration upper limit τ_{\max} is hard to make a priori, something also noted by Kwasniok (2003). Furthermore, the data sampling involved in the PIP calculation is a bit troublesome, as the usual rule “more is better” does not apply here. Rather, a risk of overdetermination seems to enter the stage if too much data is used. Also, the extent to which the data-segments used in the PIP calculation are (in)dependent influences the outcome, again with a somewhat counterintuitive result: dependent samples work better. This sampling sensitivity

may be related to the inhomogeneity of the CDV model attractor.

Acknowledgments. This research is funded through NSF-CMG grant DMS-0222133.

References

- Achatz, U. and G. Branstator, 1999: A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability. *J. Atmos. Sci.*, **56**, 3140–3160.
- Achatz, U. and J. D. Opsteegh, 2003a: Primitive-equation-based low-order models with seasonal cycle. part I: Model construction. *J. Atmos. Sci.*, **60**, 465–477.
- Achatz, U. and J. D. Opsteegh, 2003b: Primitive-equation-based low-order models with seasonal cycle. part II: Application to complexity and nonlinearity of large-scale atmosphere dynamics. *J. Atmos. Sci.*, **60**, 478–490.
- Achatz, U., G. Schmitz, and K.-M. Greisinger, 1995: Principal interaction patterns in baroclinic wave life cycles. *J. Atmos. Sci.*, **52**, 3201–3213.
- Armbruster, D., R. Heiland, E.J. Kostelich, and B. Nicolaenko, 1992: Phase-space analysis of bursting behavior in Kolmogorov flow. *Physica D*, **58**, 392–401.
- Aubry, N., P. Holmes, J. L. Lumley, and E. Stone, 1988: The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.*, **192**, 115–173.
- Aubry, N., W.-Y. Lian, and E. S. Titi, 1993: Preserving symmetries in the proper orthogonal decomposition. *SIAM J. Sci. Comput.*, **14**, 483–505.
- Branstator, G. and S.E. Haupt, 1998: An empirical model of barotropic atmospheric dynamics and its response to tropical forcing. *J. Climate*, **11**, 2645–2667.

- Cazemier, W., R.W.C.P. Verstappen, and A.E.P. Veldman, 1998: Proper orthogonal decomposition and low-dimensional models for driven cavity flows. *Phys. Fluids*, **10**, 1685–1699.
- Charney, J. G. and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1205–1216.
- Crommelin, D. T., J. D. Opsteegh, and F. Verhulst, 2004. A mechanism for atmospheric regime behaviour. *J. Atmos. Sci.*, to appear.
- D’Andrea, F. and R. Vautard, 2001: Extratropical low-frequency variability as a low-dimensional problem. I: A simplified model. *Quart. J. R. Met. Soc.*, **127**, 1357–1374.
- De Swart, H. E., 1988: Low-order spectral models of the atmospheric circulation: a survey. *Acta Appl. Math.*, **11**, 49–96.
- De Swart, H. E., 1989: Analysis of a six-component atmospheric spectral model: Chaos, predictability and vacillation. *Physica D*, **36**, 222–234.
- DelSole, T., 2001: Optimally persistent patterns in time-varying fields. *J. Atmos. Sci.*, **58**, 1341–1356.
- Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11015–11021.
- Kwasniok, F., 1996: The reduction of complex dynamical systems using principal interaction patterns. *Physica D*, **92**, 28–60.

- Kwasniok, F., 1997: Optimal Galerkin approximations of partial differential equations using principal interaction patterns. *Phys. Rev. E*, **55**, 5365–5375.
- Kwasniok, F., 2001: Low-dimensional models of the Ginzburg-Landau equation. *SIAM J. Appl. Math.*, **61**, 2063–2079.
- Kwasniok, F., 2004: Empirical low-order models of barotropic flow. *J. Atmos. Sci.*, **61**, 235–245.
- Majda, A.J., I. Timofeyev, and E. Vanden-Eijnden, 1999: Models for stochastic climate prediction. *Proc. Natl. Acad. Sci. USA*, **96**, 14687–14691.
- Majda, A.J., I. Timofeyev, and E. Vanden-Eijnden, 2003: Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.*, **60**, 1705–1722.
- Rinne, J. and V. Karhilla, 1975: A spectral barotropic model in horizontal empirical orthogonal functions. *Quart. J. R. Met. Soc.*, **101**, 365–382.
- Schubert, S.D., 1985: A statistical-dynamical study of empirically determined modes of atmospheric variability. *J. Atmos. Sci.*, **42**, 3–17.
- Schubert, S.D., 1986: The structure, energetics and evolution of the dominant frequency-dependent three-dimensional atmospheric modes. *J. Atmos. Sci.*, **43**, 1210–1237.
- Selten, F. M., 1993: Toward an optimal description of atmospheric flow. *J. Atmos. Sci.*, **50**, 861–877.
- Selten, F. M., 1995: An efficient description of the dynamics of barotropic flow. *J. Atmos. Sci.*, **52**, 915–936.

- Selten, F. M., 1997a: Baroclinic empirical orthogonal functions as basis functions in an atmospheric model. *J. Atmos. Sci.*, **54**, 2100–2114.
- Selten, F. M., 1997b: A statistical closure of a low-order barotropic model. *J. Atmos. Sci.*, **54**, 1086–1093.
- Sirovich, L., 1989: Chaotic dynamics of coherent structures. *Physica D*, **37**, 126–145.
- Winkler, C.R., M. Newman, and P.D. Sardeshmukh, 2001: An linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate*, **14**, 4474–4494.
- Wu, C.-J., 1996: Large optimal truncated low-dimensional dynamical systems. *Discr. Cont. Dyn. Syst.*, **2**, 559–583.

List of Tables

1	EOF variance spectra, using L2 norm M_0 and kinetic energy norm M_1 . Shown are the cumulative variances of the CDV model data.	30
2	Summary of dynamics of various EOF models	30
3	Results of various PIP models	31

List of Figures

- 1 Top: one integration of 6-dim. barotropic model, 4000 days long. Bottom: endpoints of 40 000 integrations of the same model, each 2000 days long, starting from randomly chosen initial conditions. Shown are projections onto the (x_1, x_4) plane. 32
- 2 Results from 2000 day integrations using 4 EOF model and CDV model. Upper left: 4 EOF model, M_1 norm, projection onto EOF 1,2 plane. Upper right: CDV model, same projection. Middle: 4 EOF model, EOF 1 versus time. Bottom: CDV model, EOF 1 versus time. 33
- 3 Autocorrelation functions for the CDV model dataset $\mathbf{x}^{\text{CDV}}(t)$ (top), for the timeseries obtained by projection of $\mathbf{x}^{\text{CDV}}(t)$ onto the EOFs (middle) and for the timeseries v_i obtained by expanding $\mathbf{x}^{\text{CDV}}(t)$ on the OPP basis (bottom). On the abscissa the time lag in days. 34
- 4 Endpoints of 40000 integrations of the 5-OPP model using OPPs 1,2,3,5,6, starting from random initial states. The datapoints are projected back onto the variables of the CDV model in order to make comparison with figure 1 easy. Shown here is the projection onto the plane (x_1, x_4) . A fixed point attracting more than half of the integrations is located at $(x_1, x_4) \approx (0.725, -0.320)$ 35
- 5 Timeseries $z(t)$ from the reduced model using 5 PIPs, calculated with $\tau_{\text{max}} = 20$, $N_{\text{int}} = 50$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 5 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time. . . 36

- 6 Timeseries $z(t)$ from the reduced model using 4 PIPs, calculated with $\tau_{\max} = 10, N_{\text{int}} = 100$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 4 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time. 37
- 7 Timeseries $z(t)$ from the reduced model using 3 PIPs, calculated with $\tau_{\max} = 25, N_{\text{int}} = 40$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 3 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time. . . 38
- 8 Power spectral densities (PSDs) of various PIP models. Depicted are PSDs made from the amplitude timeseries of the first PIP, as generated by the PIP models. The first PIP is almost identical for all 3 PIP models. Upper left: 5-PIP model ($\tau_{\max} = 50, N_{\text{int}} = 20$). Lower left: 4-PIP model ($\tau_{\max} = 10, N_{\text{int}} = 100$). Lower right: 3-PIP model ($\tau_{\max} = 25, N_{\text{int}} = 40$). Also shown (upper right) is the PSD resulting from projecting the data generated by the original CDV model onto the first PIP of the 5 PIP model. 39

No. of EOF	Cumul. variance, norm M_0	Cumul. variance, norm M_1
1	0.67954	0.65968
2	0.93427	0.94665
3	0.97576	0.98780
4	0.98849	0.99422
5	0.99611	0.99844
6	1.00000	1.00000

Table 1: EOF variance spectra, using L2 norm M_0 and kinetic energy norm M_1 . Shown are the cumulative variances of the CDV model data.

Model	Norm	Dynamics
Original CDV	–	chaotic, regimes
5 EOFs	M_0	fixed point
5 EOFs	M_1	fixed point
4 EOFs	M_0	fixed point
4 EOFs	M_1	periodic, regimes
3 EOFs	M_0	fixed point
3 EOFs	M_1	periodic, no regimes

Table 2: Summary of dynamics of various EOF models

No. PIPs	τ_{\max}	N_{int}	χ	Variance	Dynamics
5	20	40	0.014	0.997	chaotic, regimes
5	20	50	0.014	0.997	chaotic, regimes
5	20	60	0.015	0.997	chaotic, regimes
5	20	100	0.016	0.997	chaotic, regimes
5	20	1000	0.016	0.996	periodic, regimes
5	15	50	0.01	0.996	chaotic, regimes
5	25	50	0.02	0.996	chaotic, regimes
5	40	50	0.11	0.996	periodic, regimes
5	100	10	0.46	0.998	chaotic, regimes
4	10	90	0.18	0.987	periodic, regimes
4	10	100	0.19	0.985	chaotic, regimes
4	10	110	0.07	0.968	periodic, no regimes
4	10	1000	0.07	0.967	periodic, no regimes
4	7	100	0.05	0.966	periodic, no regimes
4	15	100	0.11	0.965	fixed point
4	20	100	0.13	0.969	periodic, no regimes
4	100	10	0.85	0.929	fixed point
3	25	30	0.35	0.939	chaotic, regimes
3	25	40	0.37	0.933	chaotic, regimes
3	25	50	0.36	0.937	chaotic, regimes
3	25	100	0.50	0.911	periodic, no regimes
3	25	1000	0.58	0.950	periodic, no regimes
3	30	40	0.53	0.939	periodic, no regimes
3	20	40	0.28	0.958	chaotic, no regimes
3	100	10	0.87	0.748	fixed point

Table 3: Results of various PIP models

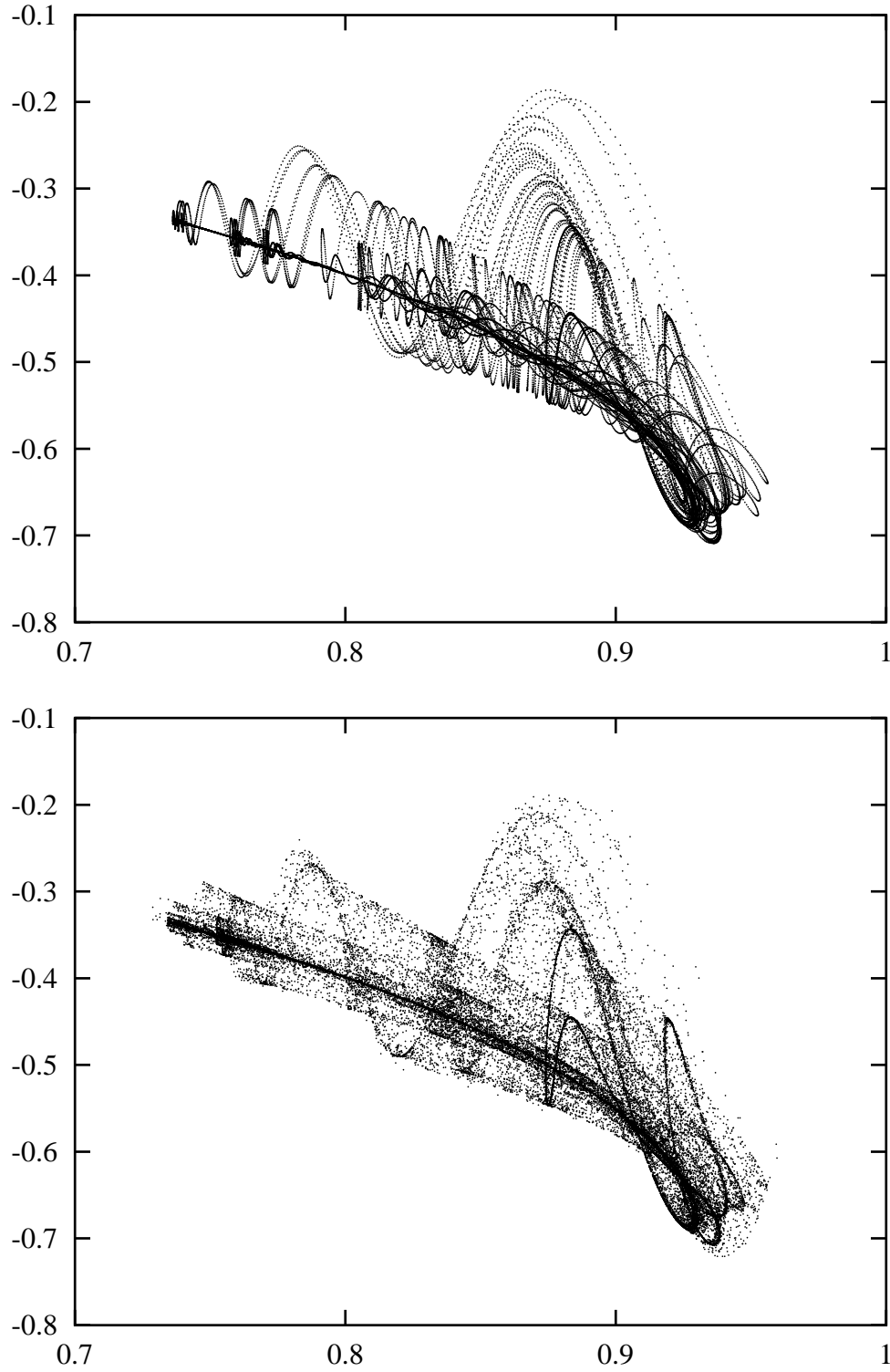


Figure 1: Top: one integration of 6-dim. barotropic model, 4000 days long. Bottom: endpoints of 40 000 integrations of the same model, each 2000 days long, starting from randomly chosen initial conditions. Shown are projections onto the (x_1, x_4) plane.

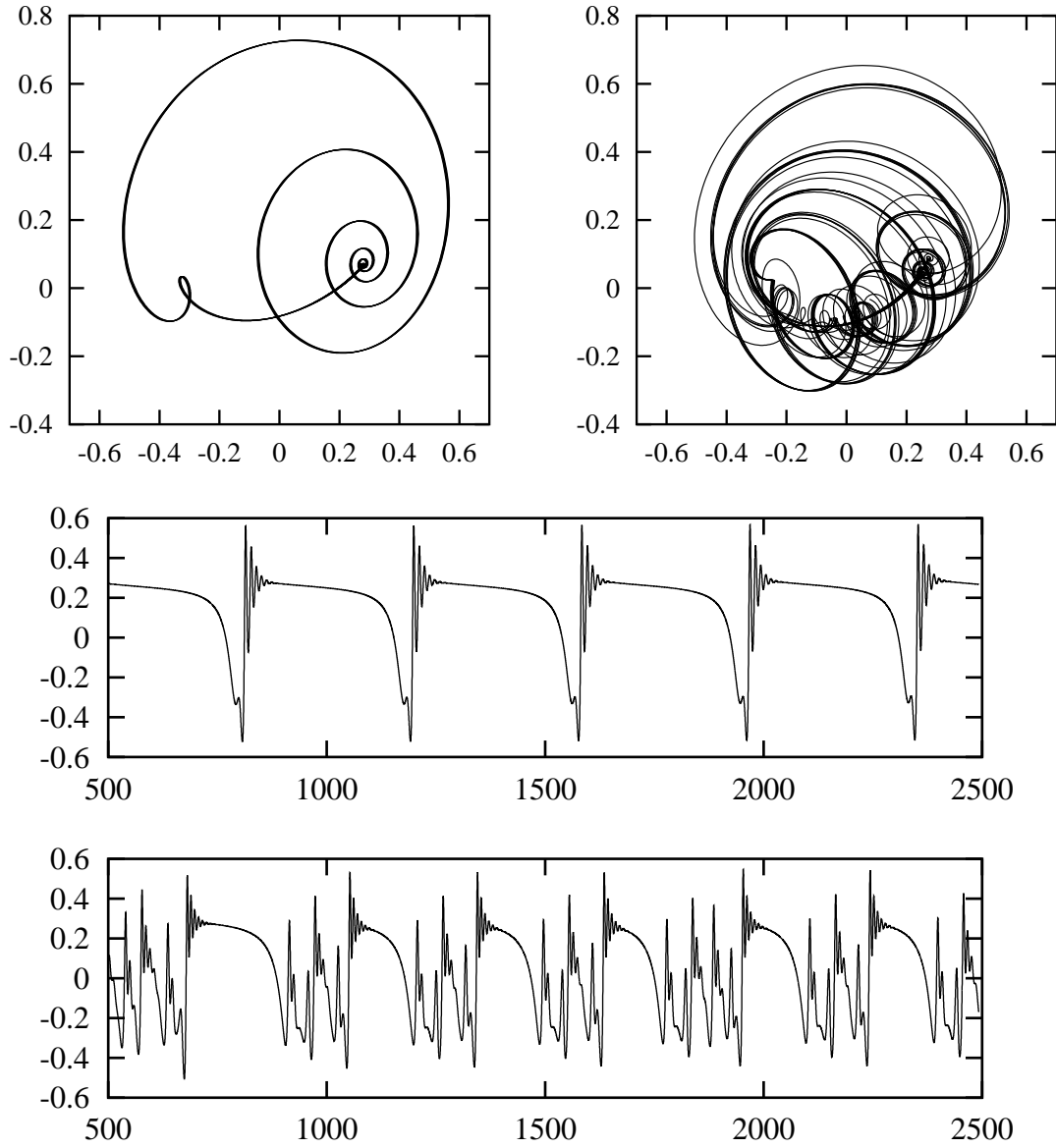


Figure 2: Results from 2000 day integrations using 4 EOF model and CDV model. Upper left: 4 EOF model, M_1 norm, projection onto EOF 1,2 plane. Upper right: CDV model, same projection. Middle: 4 EOF model, EOF 1 versus time. Bottom: CDV model, EOF 1 versus time.

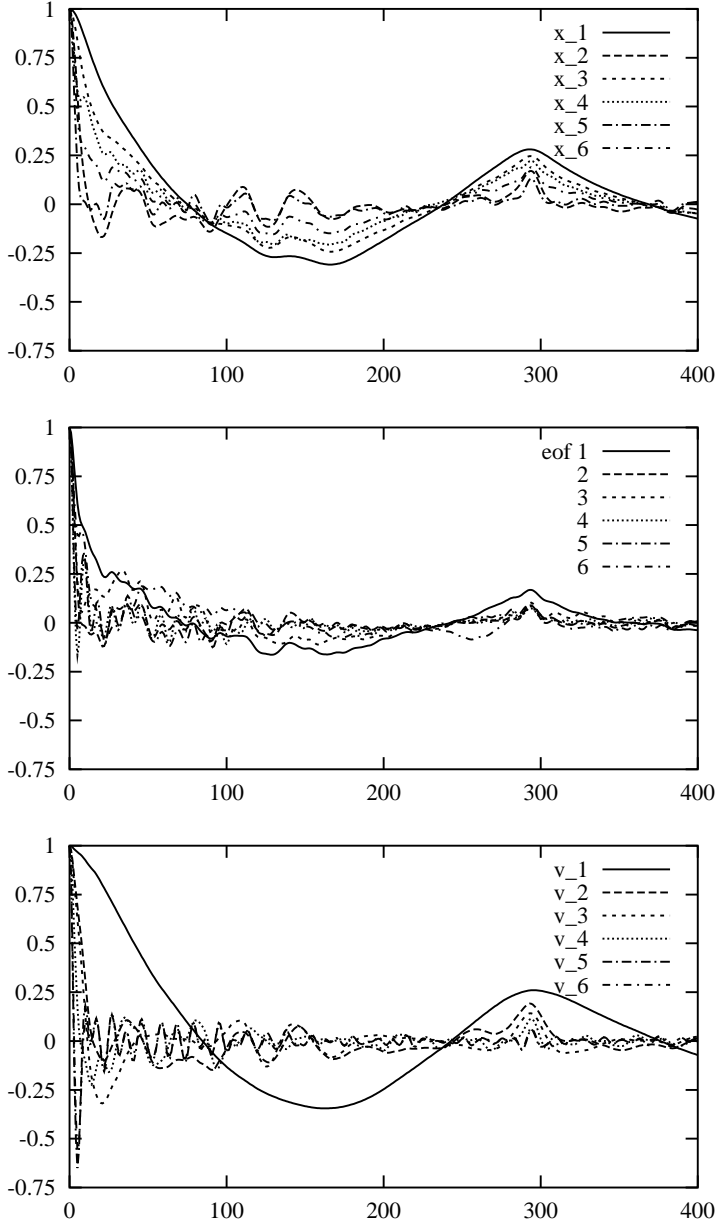


Figure 3: Autocorrelation functions for the CDV model dataset $\mathbf{x}^{\text{CDV}}(t)$ (top), for the timeseries obtained by projection of $\mathbf{x}^{\text{CDV}}(t)$ onto the EOFs (middle) and for the timeseries v_i obtained by expanding $\mathbf{x}^{\text{CDV}}(t)$ on the OPP basis (bottom). On the abscissa the time lag in days.

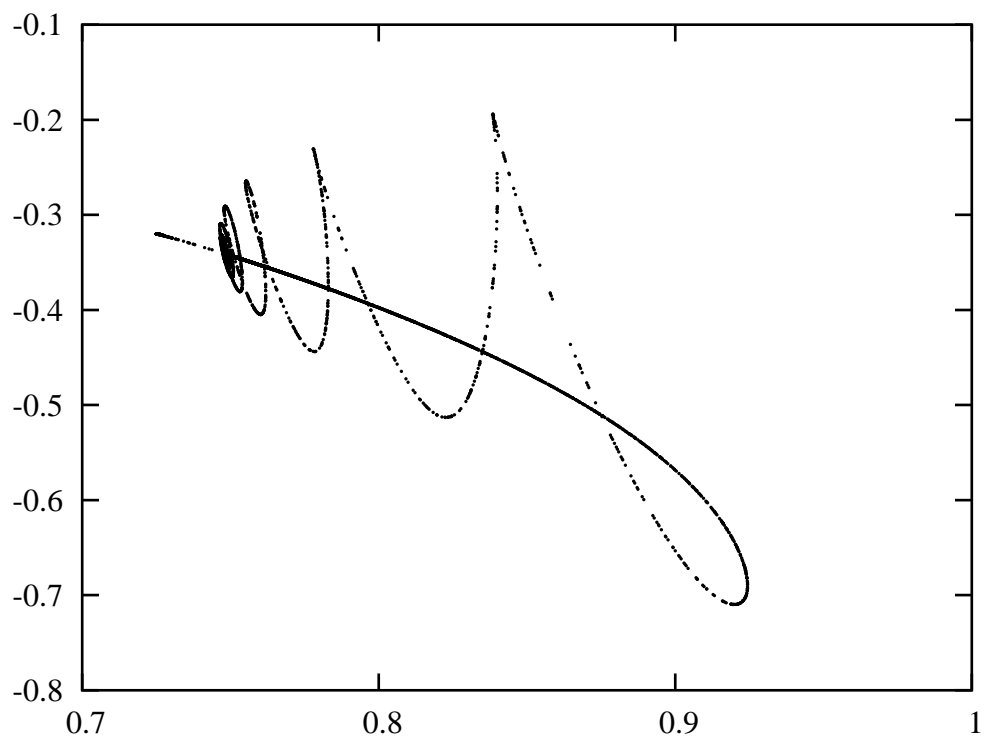


Figure 4: Endpoints of 40000 integrations of the 5-OPP model using OPPs 1,2,3,5,6, starting from random initial states. The datapoints are projected back onto the variables of the CDV model in order to make comparison with figure 1 easy. Shown here is the projection onto the plane (x_1, x_4) . A fixed point attracting more than half of the integrations is located at $(x_1, x_4) \approx (0.725, -0.320)$.

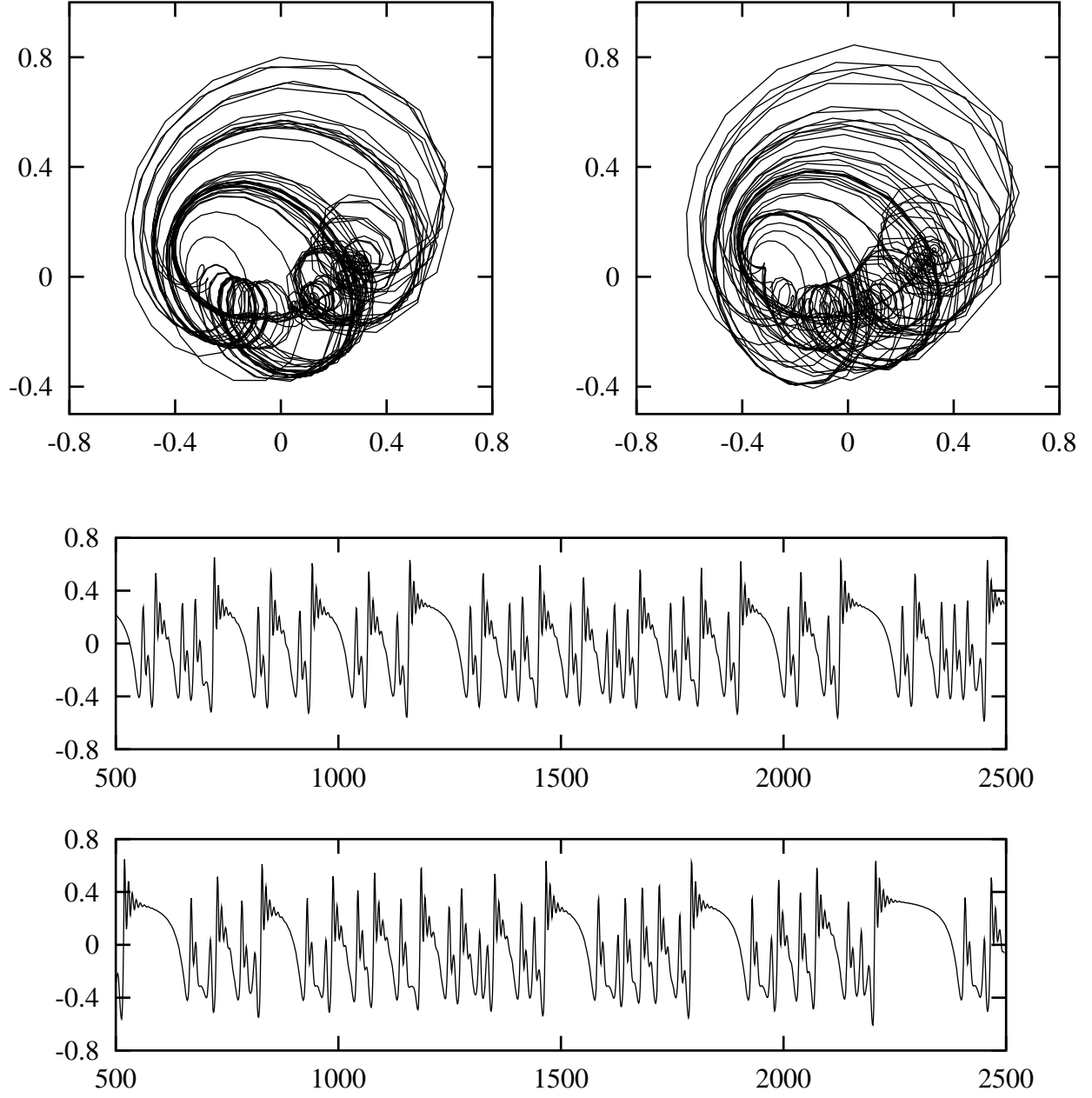


Figure 5: Timeseries $z(t)$ from the reduced model using 5 PIPs, calculated with $\tau_{\max} = 20$, $N_{\text{int}} = 50$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 5 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time.

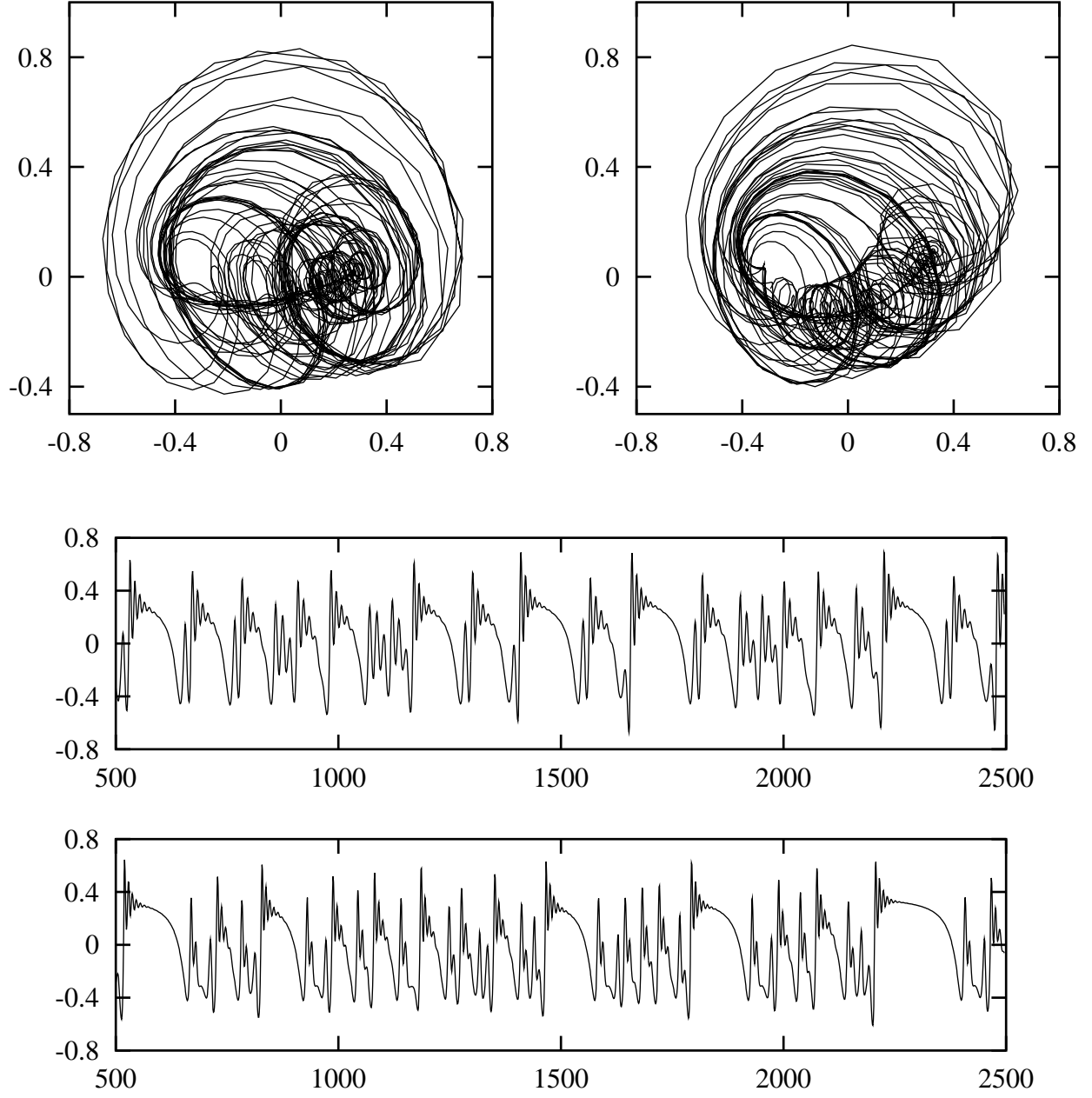


Figure 6: Timeseries $z(t)$ from the reduced model using 4 PIPs, calculated with $\tau_{\max} = 10$, $N_{\text{int}} = 100$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 4 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time.

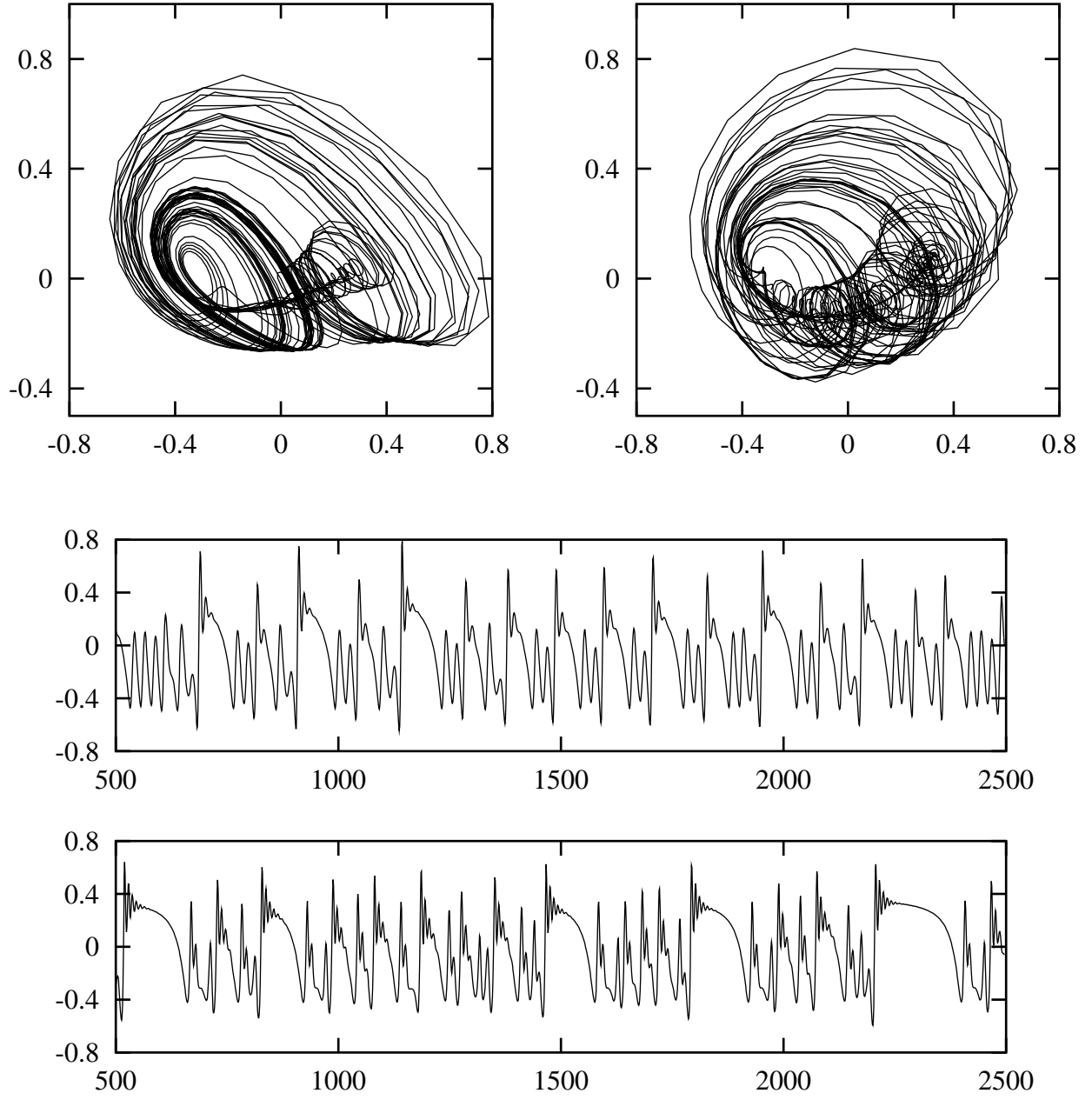


Figure 7: Timeseries $z(t)$ from the reduced model using 3 PIPs, calculated with $\tau_{\max} = 25$, $N_{\text{int}} = 40$. For comparison, the timeseries $z^{\text{ref}}(t)$ of the full 6-dim. model, projected onto the same 3 PIPs, are added. Top, left: z_1 versus z_2 . Top, right: z_1^{ref} versus z_2^{ref} . Middle: z_1 versus time. Bottom: z_1^{ref} versus time.

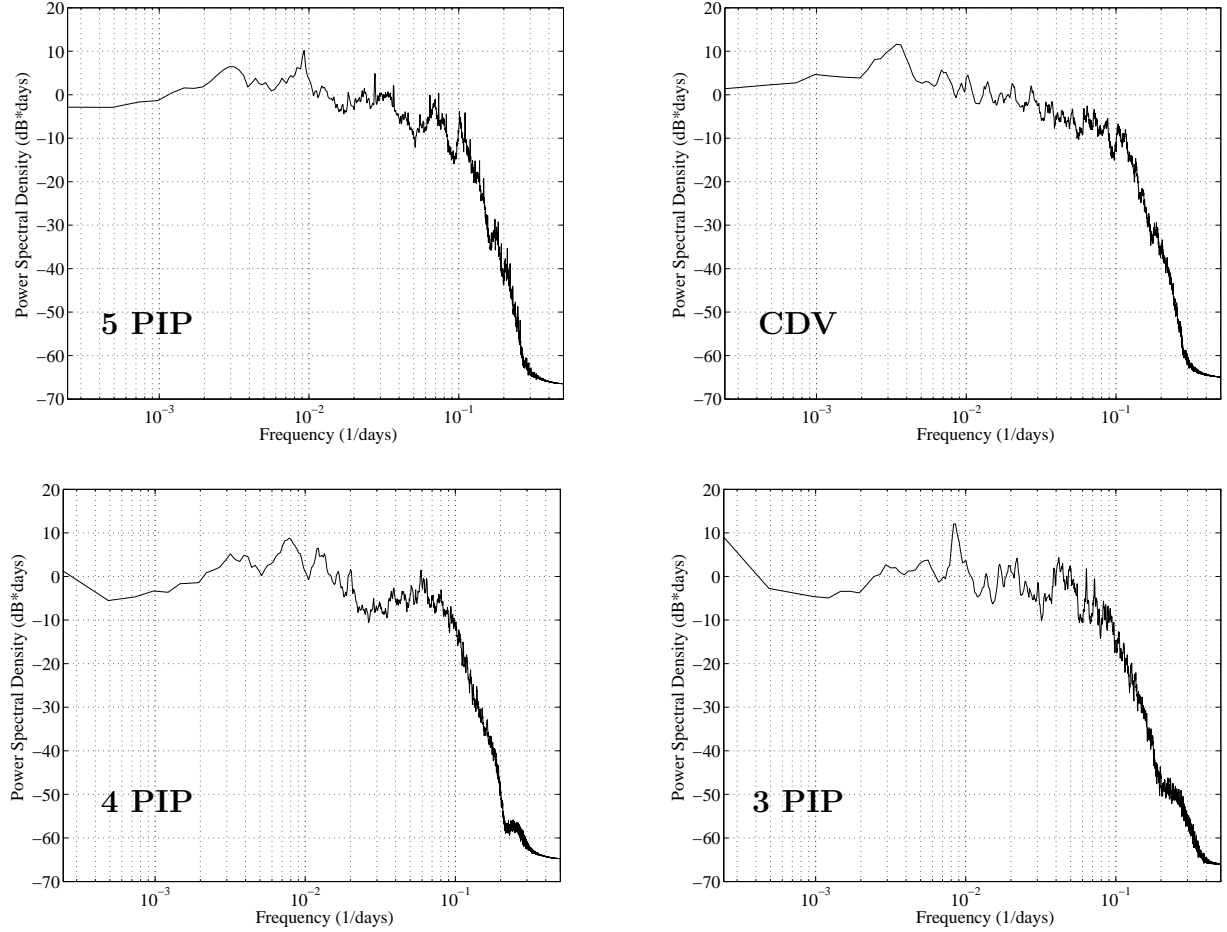


Figure 8: Power spectral densities (PSDs) of various PIP models. Depicted are PSDs made from the amplitude timeseries of the first PIP, as generated by the PIP models. The first PIP is almost identical for all 3 PIP models. Upper left: 5-PIP model ($\tau_{\max} = 50$, $N_{\text{int}} = 20$). Lower left: 4-PIP model ($\tau_{\max} = 10$, $N_{\text{int}} = 100$). Lower right: 3-PIP model ($\tau_{\max} = 25$, $N_{\text{int}} = 40$). Also shown (upper right) is the PSD resulting from projecting the data generated by the original CDV model onto the first PIP of the 5 PIP model.