

On the representation of initial  
uncertainties with multiple sets of  
singular vectors optimised for  
different criteria

Martin Leutbecher

Research Department

Submitted to Q. J. R. Meteorol. Soc.

September 2007

*This paper has not been published and should be regarded as an Internal Report from ECMWF.  
Permission to quote from it should be obtained from the ECMWF.*



European Centre for Medium-Range Weather Forecasts  
Europäisches Zentrum für mittelfristige Wettervorhersage  
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: [library@ecmwf.int](mailto:library@ecmwf.int)

©Copyright 2007

European Centre for Medium-Range Weather Forecasts  
Shinfield Park, Reading, RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

Initial uncertainties can be represented effectively in ensemble prediction systems by sampling errors in a subspace spanned by the leading singular vectors of the forecast model's tangent-linear propagator. The initial time metric in the singular vector computation is the inverse of the assumed analysis error covariance matrix  $\mathbf{A}$ . The singular vectors evolve into the leading eigenvectors of the forecast error covariance estimate obtained by linearly propagating  $\mathbf{A}$  to the singular vector optimisation time. In this sense, singular vectors provide an optimal subspace for sampling initial uncertainties. However, the optimality is only guaranteed for the particular optimisation criterion used in the singular vector computation. For instance, it could be suboptimal for forecast ranges that differ from the singular vector optimisation time.

Here, two alternative approaches are discussed that account for several optimisation criteria. The first approach is a simple ortho-normalisation approach applied to multiple sets of singular vectors. Potentially, the ortho-normalisation can yield suboptimal perturbations. In response to the expected deficiency of the ortho-normalisation approach, the second approach has been developed. It yields orthogonal subspaces for different optimisation criteria without compromising optimality. For a given subspace  $L$ , consisting of a set of leading singular vectors optimised for the first criterion (or criteria), singular vectors are computed in the subspace orthogonal to  $L$ . The optimality properties of these *subspace singular vectors* are described and proved.

The leading subspaces obtained with the two approaches are compared in two examples. First, an idealised example based on singular vectors computed for two optimisation times in the Eady model is considered. Then, both techniques are applied to initial perturbations targeted on tropical cyclones in the Ensemble Prediction System (EPS) of the European Centre for Medium-Range Weather Forecasts. The methodologies allow a consistent representation of initial uncertainties during extra-tropical transitions.

## 1 Introduction

Ensemble techniques are being employed to quantify uncertainty in geophysical fluid flow predictions (Lewis 2005; Leutbecher and Palmer 2007). Reliable ensembles require the representation of forecast model uncertainties and initial condition uncertainties; the representation of the latter is the focus of this study. Many geophysical fluid flows tend to exhibit a very selective growth of initial errors. A distribution of initial errors will be strongly stretched in some directions of phase space and squashed in other directions as it evolves in time. Lorenz (1965) showed how singular vectors of the forecast model's tangent-linear propagator can be used to estimate the error growth in the linear regime and to rank the directions in phase space according to the error growth. This property of the singular vectors has been exploited successfully in constructing strategies to selectively sample the distribution of initial errors in those directions that will dominate the forecast errors at later ranges (Buizza and Palmer 1995; Molteni et al. 1996).

Singular vectors depend on a choice of norm. Ehrendorfer and Tribbia (1997), Palmer et al. (1998) and others noted that the appropriate initial time norm (also referred to as metric) is the inverse of the analysis error covariance matrix. Among the simple metrics, the total energy metric appears to be a reasonable approximation of an analysis error covariance metric (Palmer et al. 1998; Lawrence et al. 2007). The choice of the initial metric is not discussed any further as the methodology derived here holds for any analysis error covariance estimate (provided it is a positive definite symmetric matrix).

The study focuses on strategies that represent initial condition errors only in a subspace  $L$  of the model state space  $\mathcal{L}$ . It is envisaged that the dimension of  $L$  is far smaller than the dimension  $n$  of the model state space  $\mathcal{L}$ . Anderson (1997) discusses such strategies for situations in which the distribution of initial errors is known. He suggests to focus on those strategies that are sampling the errors in  $L$  in a manner consistent with the actual initial error distribution. However, in most applications the knowledge about the actual initial error distribution is fairly limited. Therefore, we will demand here that the sampling of initial condition errors is consistent with

the *estimate*  $\mathbf{A}$  of the analysis error covariances. Higher order moments of the initial error distribution will not be considered. A concise definition of the consistency will be given in Sec. 2.1. As will be shown later, a set of  $\mathbf{A}^{-1}$ -ortho-normal vectors is required in order to construct an initial error representation that is consistent with  $\mathbf{A}$ .

Singular vectors computed with an initial norm based on the  $\mathbf{A}^{-1}$  inner product are ortho-normal with respect to  $\mathbf{A}^{-1}$ . Now, consider multiple optimisation criteria (e.g. two different optimisation times  $\tau_1$  and  $\tau_2$ ). In general, the leading singular vectors optimised for the different criteria will not be mutually  $\mathbf{A}^{-1}$ -orthogonal. Thus, a consistent selective sampling strategy requires to construct a set of  $\mathbf{A}^{-1}$ -ortho-normal vectors from the different sets of singular vectors. However, the ortho-normalisation will, in general, not respect the optimality of the perturbations. In this paper, an alternative approach is developed in which a given selective sampling strategy is augmented in an optimal way by computing singular vectors in the subspace  $\mathbf{A}^{-1}$ -orthogonal to the already sampled subspace. Such singular vectors will be referred to as *subspace singular vectors*.

The *ortho-normalisation* approach and the *subspace* approach are first applied to the Eady model. Since the work of Farrell (1988), numerous studies have examined the structure and dynamics of optimal perturbations of quasi-geostrophic baroclinic shear flow of the Eady type (see DeVries and Opsteegh 2005, and references therein). Initial perturbations based on two sets of singular vectors are considered. The first set of singular vectors maximises total energy at 24 h and the second set maximises total energy at 48 h.

Secondly, the two approaches for multiple optimisation criteria are applied to the Ensemble Prediction System (EPS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) in the context of an operational implementation. This application aims at improving the representation of initial uncertainty in the vicinity of tropical cyclones during extra-tropical transitions. In the EPS, perturbations are targeted on tropical cyclones (Puri et al. 2001). However, in the configuration operational until September 2004, these perturbations are based on singular vector computations with optimisation regions that do not extend poleward of a latitude of  $25^\circ$ . The meridional limitation of the optimisation regions was imposed in order to avoid the duplication of perturbations that are already represented in the extra-tropical singular vectors, which are optimised for latitudes from  $30^\circ$  to the pole. In a series of subsequent EPS runs during an extra-tropical transition, the spread of the tropical cyclone tracks tends to exhibit a sudden decrease in this configuration. The drop in spread arises solely from the meridional limitation of the perturbations. Here, a new perturbation methodology is presented that removes this unrealistic decrease in spread by targeting perturbations on tropical cyclones up to latitudes of  $40^\circ$ . In the new methodology, the duplication of perturbations already present in a set of extra-tropical singular vectors is avoided by ensuring that the tropical cyclone perturbations are orthogonal to the extra-tropical perturbations. Perturbations obtained with the ortho-normalisation approach are compared with perturbations obtained with the subspace approach.

The methodology for the two approaches for multiple optimisation criteria is introduced in Sec. 2. Section 3 presents the results for the application in the Eady model while Sec. 4 discusses the application in the ECMWF EPS. Discussion and conclusions follow in Sections 5 and 6, respectively. A proof of the optimality property of the subspace singular vectors is given in the Appendix.

## 2 Methodology

Section 2.1 introduces the concept of sampling initial uncertainty in a subspace consistent with a covariance matrix and explains the need for sampling initial uncertainty in orthogonal subspaces if several optimisation criteria are considered. A formulation of the maximum variance property of singular vectors suitable for the subsequent derivations is given in Sec. 2.2. The representation of initial uncertainty with singular vectors computed for two (or more) optimisation criteria using the ortho-normalisation approach follows in Sec. 2.3.

The subspace approach for representing initial uncertainty with multiple sets of singular vectors is presented in Sec. 2.4.

## 2.1 Consistent sampling of initial uncertainty and decomposition of covariances

Singular vector based initial perturbations represent initial uncertainty by a random vector  $\hat{\mathbf{x}}$  in a subspace  $L$  spanned by the leading singular vectors. [Leutbecher and Palmer \(2007\)](#) define that a random vector  $\hat{\mathbf{x}}$  in subspace  $L$  is consistent with the analysis error covariance estimate  $\mathbf{A}$  if its covariance is given by

$$\mathbf{A}_p(L) \equiv \mathbf{P}(L)\mathbf{A}(\mathbf{P}(L))^{\mathbf{T}}. \quad (1)$$

Here  $\mathbf{P}(L)$  denotes the  $\mathbf{A}^{-1}$ -orthogonal projection onto  $L$ . Superscript  $\mathbf{T}$  is used for matrix transpose. This definition of consistent sampling is motivated by the fact that a random vector  $\mathbf{x}$  with covariance  $\mathbf{A}$  can be uniquely decomposed into uncorrelated components in subspace  $L$  and its  $\mathbf{A}^{-1}$ -orthogonal complement and that the component in  $L$  has covariance  $\mathbf{A}_p(L)$ . Henceforth, the over-line will refer to the orthogonal complement of a subspace with respect to the inner product defined by  $\mathbf{A}^{-1}$

$$\bar{L} \equiv \{\mathbf{v} | \mathbf{v}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{w} = 0, \quad \forall \mathbf{w} \in L\}. \quad (2)$$

The decomposition

$$\mathbf{x} = \mathbf{x}_L + \mathbf{x}_{\bar{L}}, \quad (3)$$

with  $\mathbf{x}_L = \mathbf{P}(L)\mathbf{x} \in L$  and  $\mathbf{x}_{\bar{L}} = \mathbf{P}(\bar{L})\mathbf{x} \in \bar{L}$  yields a decomposition of  $\mathbf{x}$  into uncorrelated components.

In order to get a random vector that is consistent with  $\mathbf{A}$ , one can use a Gaussian sampling technique based on an  $\mathbf{A}^{-1}$ -ortho-normal basis of  $L$ . Such a sampling technique is used in the operational EPS configuration at ECMWF ([Leutbecher and Palmer 2007](#)). Note, that all singular vectors from a set computed with initial norm based on  $\mathbf{A}^{-1}$  for a particular optimisation criterion are already ortho-normal with respect to  $\mathbf{A}^{-1}$ . But two singular vectors from different sets are generally not  $\mathbf{A}^{-1}$ -orthogonal.

For any pair of  $\mathbf{A}^{-1}$ -orthogonal subspaces  $L'$  and  $L''$ , the covariances are additive

$$\mathbf{A}_p(L' + L'') = \mathbf{A}_p(L') + \mathbf{A}_p(L'') \quad (4)$$

because the components in  $\mathbf{A}^{-1}$ -orthogonal subspaces are uncorrelated and  $\mathbf{P}(L' + L'') = \mathbf{P}(L') + \mathbf{P}(L'')$ . In particular, the full covariance matrix is decomposed as

$$\mathbf{A} = \mathbf{A}_p(L) + \mathbf{A}_p(\bar{L}). \quad (5)$$

The decomposition (4) implies that the sum of a set of independent random vectors  $\mathbf{x}_k \in L_k$ ,  $k = 1 \dots r$  each  $\mathbf{x}_k$  consistent with  $\mathbf{A}$  yields a random vector  $\hat{\mathbf{x}} \in \hat{L} = L_1 + \dots + L_r$  which is also consistent with  $\mathbf{A}$  if the subspaces  $L_k$  are mutually  $\mathbf{A}^{-1}$ -orthogonal. The covariance of  $\hat{\mathbf{x}}$  is then given by the sum of the covariances associated with the individual subspaces  $L_k$

$$\mathbf{A}_p(\hat{L}) = \sum_{k=1}^r \mathbf{A}_p(L_k). \quad (6)$$

For any linear operator  $\mathbf{T}$ , the covariance matrix of the transformed random vector  $\mathbf{T}\hat{\mathbf{x}}$  has covariance decompositions given by

$$\mathbf{TAT}^{\mathbf{T}} = \mathbf{TA}_p(L)\mathbf{T}^{\mathbf{T}} + \mathbf{TA}_p(\bar{L})\mathbf{T}^{\mathbf{T}} \quad \text{and} \quad (7)$$

$$\mathbf{TA}_p(\hat{L})\mathbf{T}^{\mathbf{T}} = \sum_{k=1}^r \mathbf{TA}_p(L_k)\mathbf{T}^{\mathbf{T}}. \quad (8)$$

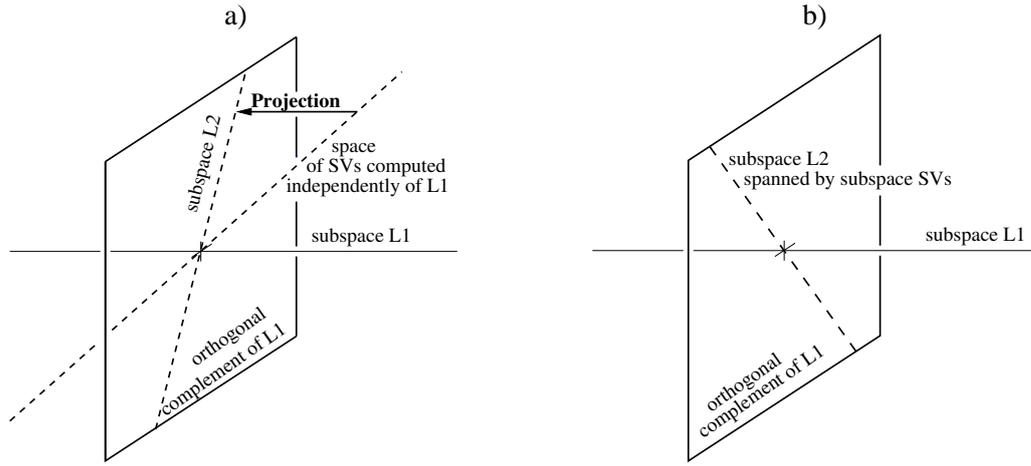


Figure 1: Schematic of two alternative methods to construct orthogonal subspaces for two different optimisation criteria: (a) ortho-normalisation of independently computed sets of singular vectors, (b) singular vectors for the second criterion are computed in the subspace ortho-normal to  $L_1$  rather than in the full state space.

## 2.2 Singular vectors and the maximum variance property

A linear estimate of the forecast error covariance matrix at time  $t$  is given by  $\mathbf{M}\mathbf{M}^T$ , where  $\mathbf{M}$  denotes the tangent-linear propagator from initial time to time  $t$ . This estimate neglects imperfections of the forecast model, nonlinearities as well as other imperfections of the tangent-linear model and it assumes that the covariance of initial uncertainty is given by  $\mathbf{A}$ . It is assumed that forecast error variance is quantified using a symmetric positive semi-definite matrix  $\mathbf{E}$  as metric. This metric could be, for instance, the total energy metric localised to a certain geographical region. In general, the matrix  $\mathbf{E}$  can be different for different optimisation criteria. Then, the trace  $\text{tr}(\mathbf{E}^{1/2}\mathbf{M}\mathbf{M}^T\mathbf{E}^{1/2})$  is an estimate of the total forecast error variance using metric  $\mathbf{E}$ . If we write  $\mathbf{T} = \mathbf{E}^{1/2}\mathbf{M}$  and make use of Equation (7), the total forecast error variance can be decomposed as

$$\text{tr}(\mathbf{T}\mathbf{A}\mathbf{T}^T) = \text{tr}(\mathbf{T}\mathbf{A}_p(L)\mathbf{T}^T) + \text{tr}(\mathbf{T}\mathbf{A}_p(\bar{L})\mathbf{T}^T). \quad (9)$$

Note, that Equation (9) is valid for any matrix  $\mathbf{T}$  not only the tangent-linear propagator scaled by the square root of a metric.

The singular vectors of  $\mathbf{T}\mathbf{A}^{1/2}$  are known to provide a particular decomposition of (9) which will be formulated now. Let

$$\mathbf{T}\mathbf{A}^{1/2} = \mathbf{U}\mathbf{S}\tilde{\mathbf{V}}^T, \quad (10)$$

denote the singular value decomposition, where  $\mathbf{U}, \tilde{\mathbf{V}}$  are orthogonal matrices and  $\mathbf{S}$  is a diagonal matrix containing the non-negative singular values in decreasing order. The singular vectors are the columns of  $\mathbf{U}$  and  $\tilde{\mathbf{V}}$ . The right singular vectors in  $\tilde{\mathbf{V}}$  are non-dimensional. Their scaled counterparts

$$\mathbf{V} = \mathbf{A}^{1/2}\tilde{\mathbf{V}} \quad (11)$$

are usually referred to as *initial* singular vectors when  $\mathbf{T}$  involves the tangent-linear propagator. The symbol  $L^*[\mathbf{T}, k]$  will be used to refer to the subspace spanned by the leading  $k$  initial singular vectors, i.e. the first  $k$  columns of  $\mathbf{V}$ . Furthermore, let  $s_j(\mathbf{T})$  denote the  $j$ -th singular value of operator  $\mathbf{T}$ . The singular vectors are optimal perturbations in the sense that they optimise the following ratio of norms

$$\left\| \mathbf{T}\mathbf{A}^{1/2}\mathbf{x} \right\| / \|\mathbf{x}\|, \quad (12)$$

where  $\|\cdot\|$  denotes the Euclidean norm. Furthermore, the subspace  $L^*[\mathbf{T}, k]$  has the property of being the  $k$ -dimensional subspace that *explains* most variance

$$\max_{L \subset \mathcal{L}, \dim(L)=k} \text{tr}(\mathbf{T}\mathbf{A}_p(L)\mathbf{T}^T) = \text{tr}(\mathbf{T}\mathbf{A}_p(L^*[\mathbf{T}, k])\mathbf{T}^T). \quad (13)$$

Ehrendorfer and Tribbia (1997) discuss the maximum variance property of the singular vectors in detail for the case where  $\mathbf{T}$  consists of the tangent-linear propagator scaled by a metric. In the following, the more general formulation (13) valid for any matrix  $\mathbf{T}$  is required.

### 2.3 Multiple optimisation criteria: ortho-normalisation approach

We consider two optimisation criteria which are associated with matrices  $\mathbf{T}_1$  and  $\mathbf{T}_2$ . They could be composed of propagators for two different optimisation times. Alternatively,  $\mathbf{T}_1$  and  $\mathbf{T}_2$  could also differ in terms of the geographical localisation at final time.

The ortho-normalisation approach considers the singular value decompositions of  $\mathbf{T}_1\mathbf{A}^{1/2}$  and  $\mathbf{T}_2\mathbf{A}^{1/2}$ , i.e. this corresponds to independent singular vector computations. The subspace spanned by the leading  $\ell_1$  singular vectors of  $\mathbf{T}_1\mathbf{A}^{1/2}$  is denoted by  $L_1$ . Then, subspace  $L_2$  is obtained by projecting the singular vectors of  $\mathbf{T}_2\mathbf{A}^{1/2}$  on  $\overline{L_1}$  (Fig. 1a). Formally, we write

$$L_2 = \mathbf{P}(\overline{L_1})L^*[\mathbf{T}_2, \ell_2]. \quad (14)$$

The ortho-normalisation approach is expected to yield, in general, perturbations in  $\overline{L_1}$  that are sub-optimal with respect to the second optimisation criterion.

### 2.4 The subspace approach

The second approach achieves orthogonality without compromising optimality by restricting the singular vector computation for the second criterion to subspace  $\overline{L_1}$  (Fig. 1b). Therefore, we will refer to it as the subspace approach.

As previously, the subspace spanned by the leading singular vectors of  $\mathbf{T}_1\mathbf{A}^{1/2}$  is denoted by  $L_1$ . Note, however, that the following properties are valid for any subspace  $L_1$ . The goal of optimising for  $\mathbf{T}_2$  in the space orthogonal to  $L_1$  is then achieved by considering the singular value decomposition of

$$\mathbf{T}_s = \mathbf{T}_2\mathbf{P}(\overline{L_1}). \quad (15)$$

The singular vectors are solutions of the generalised eigenproblem

$$\mathbf{T}_s^T \mathbf{T}_s \mathbf{x} = \sigma^2 \mathbf{A}^{-1} \mathbf{x}. \quad (16)$$

The singular vectors of  $\mathbf{T}_s\mathbf{A}^{1/2}$  are characterised by the following properties:

- All singular vectors of  $\mathbf{T}_s$  with positive singular value lie in  $\overline{L_1}$ .
- For dimensions  $k$  with positive singular value  $s_1 > \dots > s_k(\mathbf{T}_s) > 0$ , the subspace  $L^*[\mathbf{T}_s, k]$  spanned by the leading  $k$  singular vectors of  $\mathbf{T}_s$  satisfies the following maximum variance criterion

$$\max_{L \subset \overline{L_1}, \dim(L)=k} \text{tr}(\mathbf{T}_2\mathbf{A}_p(L)\mathbf{T}_2^T) = \text{tr}(\mathbf{T}_2\mathbf{A}_p(L^*[\mathbf{T}_s, k])\mathbf{T}_2^T) \quad (17)$$

In other words, subspace  $L^*[\mathbf{T}_s, k]$  is the  $k$ -dimensional subspace orthogonal to  $L_1$  which explains most forecast error variance of  $\mathbf{T}_2\mathbf{A}\mathbf{T}_2^T$ .

A formal proof of statements (a) and (b) is given in Appendix A. We will refer to the singular vectors of  $\mathbf{T}_s$  with positive singular value as *subspace singular vectors*. As a result of the additivity of the forecast error covariances associated with  $\mathbf{A}^{-1}$ -orthogonal subspaces, statement (b) implies that subspace  $L_1 + L^*[\mathbf{T}_s, k]$  is the  $\ell_1 + k$ -dimensional subspace containing  $L_1$  that explains most of the forecast error variance of  $\mathbf{T}_2 \mathbf{A} \mathbf{T}_2^T$ . The subspace approach is sequential and non-commutative, i.e. the space  $L_1 + L_2$  depends on the order of the optimisation criteria as will be demonstrated in Sec. 3.

### 3 Quasi-geostrophic baroclinic shear flow

#### 3.1 Definition of the linear problem

Now, the subspace approach and the ortho-normalisation approach will be illustrated in a low-dimensional, yet non-trivial, example based on the Eady model. We consider a basic state with linear vertical shear  $\bar{U} = Sz$  and constant Brunt-Väisälä frequency  $N$ . The quasi-geostrophic (QG) dynamics is linearised about this basic state on the  $f$ -plane, i.e. constant Coriolis parameter  $f_0$ . Rigid boundaries are assumed at  $z = 0$  and  $z = H$ . In the zonal direction, we consider a channel of length  $L_{\text{ch}}$  with periodic boundary conditions. All perturbation variables are uniform in the meridional direction  $\partial/\partial y = 0$ . Following Farrell and Ioannou (1996), the variables are non-dimensionalised using the height  $H$  as vertical scale, the Rossby deformation radius  $L_d = NH/f_0$  as horizontal scale and  $N/(f_0 S)$  as time scale. The numerical examples use  $H = 10\text{ km}$ ,  $SH = 46.3\text{ ms}^{-1}$ ,  $N = 10^{-2}\text{ s}^{-1}$ ,  $f_0 = 10^{-4}\text{ s}^{-1}$  and  $L_{\text{ch}} = 10^4\text{ km}$ . For these values, a unit non-dimensional time corresponds to 6 h and a unit non-dimensional horizontal length corresponds to 1000 km. Equations (18)–(20) below appear in non-dimensional form (Note, that the same symbols will be used for the non-dimensional variables). The equations describing the evolution of streamfunction perturbations  $\psi$  are given by

$$\left(\frac{\partial}{\partial t} + z\frac{\partial}{\partial x}\right) \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2}\right) = 0 \quad (18)$$

in the interior,  $0 < z < 1$ , and

$$\left(\frac{\partial}{\partial t} + z\frac{\partial}{\partial x}\right) \frac{\partial \psi}{\partial z} - \frac{\partial \psi}{\partial x} = 0 \quad (19)$$

at the top and bottom boundaries,  $z = 0$  and  $z = 1$ . Equations (18) and (19) describe the conservation of QG potential vorticity and the advection of temperature at the boundaries, respectively.

Equations (18) and (19) are discretised in the vertical using 21 equidistant levels. A Fourier representation with 16 wavenumbers is used for the horizontal direction. Thereby, waves with wavelength between  $L/16$  and  $L$  can be described in the channel. The total energy inner product is defined as

$$\langle \psi; \tilde{\psi} \rangle = \frac{1}{2} \int_0^L dx \int_0^1 dz \left( \frac{\partial \psi}{\partial x} \frac{\partial \tilde{\psi}}{\partial x} + \frac{\partial \psi}{\partial z} \frac{\partial \tilde{\psi}}{\partial z} \right). \quad (20)$$

In the examples given below, the total energy metric will be used at initial time and final time. Furthermore, the non-dimensional optimisation times are set to  $\tau_1 = 4$  and  $\tau_2 = 8$ ; these values correspond to dimensional times of 24 h and 48 h. This choice is motivated by an optimisation time of 48 h used in the operational ECMWF EPS and recent experimentation with 24-hour optimisation time singular vectors using a diabatic tangent-linear model (Coutinho et al. 2004; Hoskins and Coutinho 2005; Walser et al. 2006).

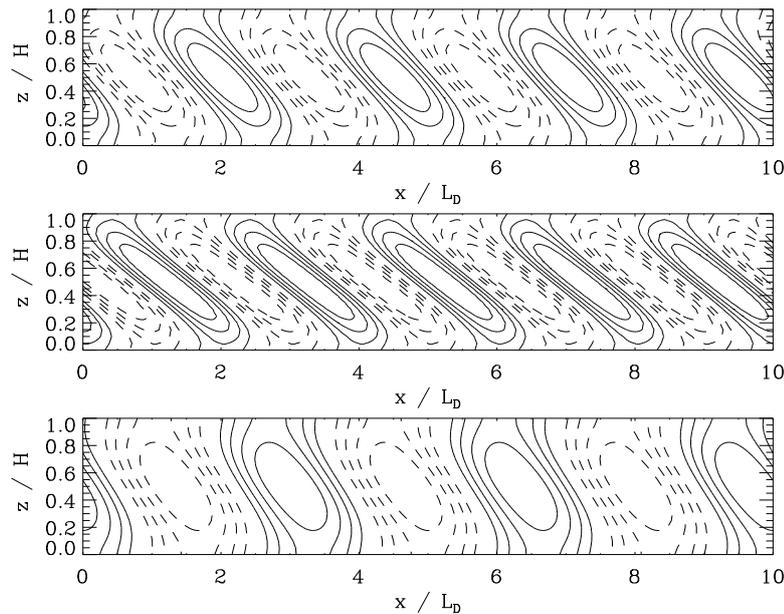


Figure 2: The leading three singular vectors (SVs) for the discretised Eady problem using the total energy metric at initial and final time (top: SV 1, middle: SV 2, bottom: SV 3). The non-dimensional optimisation time  $\tau_1 = 4$  corresponds to a dimensional time of 24 h. Plotted is the streamfunction (Dashed contours: negative values; zero-contour omitted). The discretisation is described in Sec. 3.1; the singular values are given in Table 1.

## 3.2 Results

The leading 3 singular vectors for the non-dimensional optimisation times  $\tau_1 = 4$  and  $\tau_2 = 8$  are plotted in Figures 2 and 3, respectively. We will refer to these singular vectors as  $\tau_1$ -singular vectors and  $\tau_2$ -singular vectors in short. The streamfunction perturbations exhibit the familiar tilt against the vertical shear and resemble the leading singular vectors presented by Mukougawa and Ikeda (1994) and by Morgan and Chen (2002). In their studies, the horizontal wavelength is set as external parameter. Here, it is determined by the singular vector computation itself. The dimensional wavelengths of the leading 3 singular vectors are  $2.50L_d$ ,  $2.00L_d$  and  $3.33L_d$  for  $\tau_1$  and  $3.33L_d$ ,  $2.50L_d$  and  $5.00L_d$  for  $\tau_2$ . In particular, the  $\tau_1$ -singular vectors 1 and 3 have the same wavelength as the  $\tau_2$ -singular vectors 2 and 1. The leading  $\tau_2$ -singular vectors are more confined around mid-height ( $z = 0.5H$ ) and are more tilted than the  $\tau_1$ -singular vectors with the same horizontal wavelength.

Now, we apply the ortho-normalisation approach. For this example, subspace  $L_1$  is chosen to be the subspace spanned by the leading three  $\tau_1$ -singular vectors  $\mathbf{v}_k(\tau_1), k = 1, 2, 3$ . Subspace  $L_2$  is obtained by projecting the subspace spanned by the leading three  $\tau_2$ -singular vectors  $\mathbf{v}_k(\tau_2), k = 1, 2, 3$  into  $\overline{L_1}$ . The matrix of projection coefficients is given by

$$[\mathbf{v}_j^T(\tau_2)\mathbf{A}^{-1}\mathbf{v}_k(\tau_1)]_{jk} = \begin{pmatrix} 0 & 0 & 0.93 + 0.03i \\ 0.77 - 0.06i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where  $i$  denotes the imaginary unit. The analysis is based on complex valued singular vectors (cf. Mukougawa and Ikeda 1994, for mathematical details). Multiplication of the complex singular vector by a complex number of modulus 1 corresponds to a shift of the wave phase. The first  $\tau_2$ -singular vector has a large projection on the third  $\tau_1$ -singular vector and the second  $\tau_2$ -singular vector has a large projection on the first  $\tau_1$ -singular vector whereas the third  $\tau_2$ -singular vector is orthogonal to  $L_1$ .

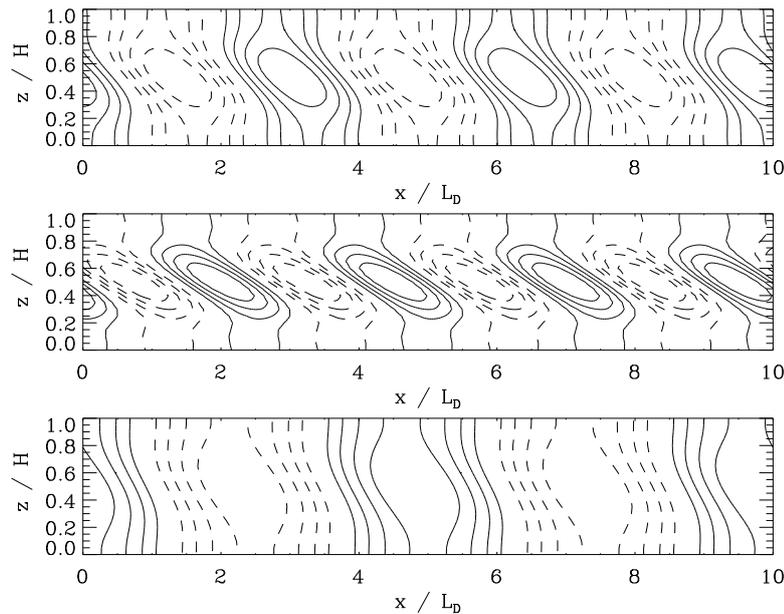


Figure 3: as Fig. 2 but for a non-dimensional optimisation time of  $\tau_2 = 8$ ; this corresponds to 48 h.

The projected and ortho-normalised  $\tau_2$ -singular vectors are a basis of  $L_2$ . They are plotted in Figure 4. As one would expect, the structure of the first two  $\tau_2$ -singular vectors has been significantly altered whereas the third  $\tau_2$ -singular vector remains unchanged.

Next, the subspace approach is applied with  $L_1$  still given by the subspace spanned by the leading three  $\tau_1$ -singular vectors. The optimisation for time  $\tau_2$  is then restricted to  $L_1$  using  $\mathbf{T}_s = \mathbf{E}^{1/2}\mathbf{M}(\tau_2)\mathbf{P}(\overline{L_1})$  in Equation (16). Here,  $\mathbf{M}(\tau_2)$  and  $\mathbf{E}$  denote the propagator from time 0 to  $\tau_2$  and the total energy metric, respectively. The singular values for the  $\tau_2$ -subspace singular vectors are always smaller than the singular values of the  $\tau_2$ -singular vectors (Table 1). This is a general property of the subspace approach because perturbations restricted to a subspace cannot be more optimal than perturbations computed in the full space.

The central question of this work can be addressed now: Are the  $\tau_2$ -subspace singular vectors spanning a different space than the projected  $\tau_2$ -singular vectors? The answer is *yes* because the second  $\tau_2$ -subspace singular vector has a wavelength of  $2.00L_d$  and is therefore orthogonal to the leading three projected  $\tau_2$ -singular vectors, which have different wavelength (see Figs. 4 and 5). However, the first  $\tau_2$ -subspace singular vector is identical to the third (projected)  $\tau_2$ -singular vector and the third  $\tau_2$ -subspace singular vector is similar to the second projected and (ortho-)normalised  $\tau_2$ -singular vector.

The singular vectors optimised in  $\overline{L_1}$  always maximise the explained variance, see Equation (17) in Sec. 2. As the leading three  $\tau_2$ -subspace singular vectors span a different space than the leading three projected  $\tau_2$ -singular vectors, one would expect the former to explain more forecast error variance. Total forecast error variances explained by the spaces spanned by the leading five projected singular vectors and by the leading five subspace singular vectors are given in Table 2. The numbers confirm that the subspace singular vectors explain indeed more forecast error variance than the projected  $\tau_2$ -singular vectors. The results show that the difference between the two approaches is largest for small dimension  $n = 1, 2$  of the space  $L_2$ . In the limit of increasing dimension  $n$ , the covariance matrix  $\mathbf{A}_p(L_2)$  converges towards the covariance matrix  $\mathbf{A}_p(\overline{L_1})$  for both methods and the explained forecast error variances have to become identical. Note, that the leading three  $\tau_2$ -singular vectors explain 22.6% of the total forecast error variance at  $t = \tau_2$ . This value is exceeded by the variance explained by  $L_1 + L_2$  for  $\dim(L_2) = 3$  with the ortho-normalisation approach while the subspace

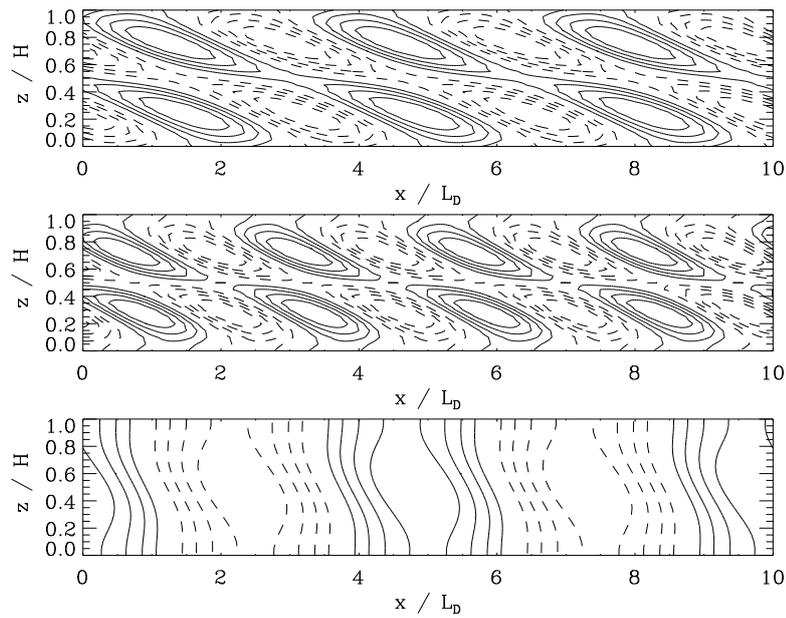


Figure 4: as Figure 3 but for the three vectors that are obtained from the  $\tau_2$ -singular vectors (cf. Fig. 3) by projecting them onto the orthogonal complement of the  $\tau_1$ -singular vectors (cf. Fig. 2) and subsequent ortho-normalisation.

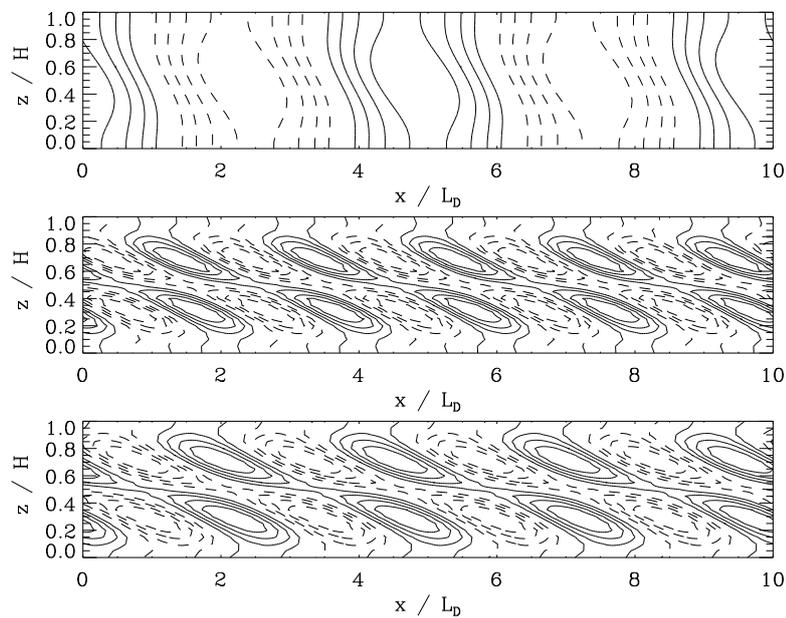


Figure 5: as Figure 3 but for the leading three subspace singular vectors of the Eady model computed for optimisation time  $\tau_2$  in the subspace orthogonal to the leading three  $\tau_1$ -singular vectors (top: SV 1, middle: SV 2, bottom: SV 3). The singular values are given in Tab. 1.

Table 1: Singular values of the leading 5 singular vectors for the discretised Eady problem and optimisation times  $\tau_1, \tau_2$ .

singular vector $j$	1	2	3	4	5
$\tau_1 = 4$	6.39	6.18	6.08	5.83	5.75
$\tau_2 = 8$	24.36	22.25	17.86	16.38	14.27
subspace $\tau_2 = 8$	17.86	16.02	14.91	14.27	13.56

 Table 2: Percentage of total forecast error variance at  $t = \tau_2$  explained by subspaces  $L_2$  and  $L_1 + L_2$  for  $\dim(L_1) = 3$  and  $\dim(L_2) = 1, \dots, 5$ .

method	subspace	$\dim(L_2)$				
		1	2	3	4	5
ortho-norm.	$L_2$	1.6	5.1	10.2	14.3	17.6
subspace	$L_2$	5.1	9.3	12.8	16.1	19.1
ortho-norm.	$L_1 + L_2$	16.0	19.5	24.7	28.7	32.0
subspace	$L_1 + L_2$	19.5	23.7	27.3	30.5	33.5

approach already exceeds 22.6% for  $\dim(L_2) = 2$ .

So far, an example with  $\dim(L_1) = 3$  was considered and  $L_1$  was optimised for  $\tau_1 = 4$  and  $L_2$  for  $\tau_2 = 8$ . Now, the dimensions of  $L_1$  and  $L_2$  are varied between 1 and 5 with  $\dim(L_1) = \dim(L_2)$ . Moreover, the order of the optimisation times is examined, i.e. singular vectors are computed for  $\tau_1 = 4, \tau_2 = 8$  as well as for  $\tau_1 = 8, \tau_2 = 4$ . The forecast error variances explained by  $L_1 + L_2$  at  $t = 4$  and  $t = 8$  are given in Table 3. As predicted by the theory in Sec. 2, more forecast error variance is explained by the subspace approach than the ortho-normalisation approach for all dimensions and both lead times. However, the forecast error variance explained by the ortho-normalisation approach is large; the ratio of variance explained by the ortho-normalisation approach to the variance explained by the subspace approach exceeds 0.8 in all cases.

The subspace  $L_1 + L_2$  obtained with the ortho-normalisation approach does not depend on the order of the optimisation criteria but it does so for the subspace approach. At time  $t = 4$ , the subspace approach with  $\tau_2 = 4$  explains more forecast error variance than with  $\tau_2 = 8$  except for  $\dim(L_1 + L_2) = 4$ . At time  $t = 8$ , the subspace approach with  $\tau_2 = 8$  explains slightly more variance than with  $\tau_2 = 4$ .

 Table 3: Percentage of total forecast error variance explained by subspace  $L_1 + L_2$  at time  $t$  for  $\dim(L_1) = \dim(L_2) = 1, \dots, 5$ .

method	$\tau_1$	$\tau_2$	$t$	$\dim(L_1 + L_2)$				
				2	4	6	8	10
ortho-norm.	4/8	8/4	4	5.5	8.9	11.6	14.9	17.8
subspace	4	8	4	5.5	10.3	11.9	15.3	18.7
subspace	8	4	4	5.5	9.9	14.3	17.5	20.7
ortho-norm.	4/8	8/4	8	14.6	19.2	24.7	29.6	33.2
subspace	4	8	8	14.6	20.8	27.3	31.1	34.8
subspace	8	4	8	14.6	19.4	25.5	30.5	34.9

## 4 Application in the ECMWF EPS

This section describes how the methods developed in the previous section can be exploited to revise the initial perturbations targeted on tropical cyclones in the operational ECMWF Ensemble Prediction System (EPS). In the configuration operational up to September 2004, optimisation regions targeted on tropical cyclones were restricted to latitudes between 25°S and 25°N. This ensured that the associated singular vectors were almost orthogonal to the extra-tropical singular vectors, which are optimised for the regions from 30° latitude to the pole. As a consequence of this latitudinal limitation of the tropical cyclone perturbations, the EPS had the undesirable feature that the spread of tropical cyclone tracks exhibited a sudden decrease during extra-tropical transitions. Users may have falsely interpreted the reduction in the spread of tropical cyclone positions as an increase in predictability.

With the methods developed in this paper, it is possible to extend the perturbations targeted on tropical cyclones further pole-ward without duplicating structures that are already represented in the set of used extra-tropical singular vectors. Here, we compare the perturbation structures obtained with the ortho-normalisation approach and the subspace approach. In the notation of Sec. 2, subspace  $L_1$  refers to the space spanned by the leading 50 extra-tropical singular vectors. Subspace  $L_2$  is the space spanned by the leading 5 singular vectors for a particular tropical cyclone.

Initial perturbations for both approaches have been computed daily during the period 11–24 September 2006. In this period, the tropical cyclones Gordon, Shanshan, Helene and Yagi underwent extra-tropical transitions. The ortho-normalisation approach and the subspace approach will be referred to as experiments O and S, respectively. Apart from the ortho-normalisation approach, both experiments adopt the currently operational configuration for the EPS initial perturbations (Leutbecher and Palmer 2007): All sets of singular vectors are computed with 48-hour optimisation time. The spectral model is triangularly truncated at wavenumber 42 and uses 62 levels. The extra-tropical singular vectors are computed with an adiabatic version of the tangent-linear model whereas the singular vectors targeted on tropical cyclones are computed with a diabatic tangent-linear model, which is computationally more expensive than the adiabatic version. Puri et al. (2001) showed that the representation of diabatic processes in the tangent-linear model is required in order to obtain initial perturbations relevant for tropical cyclone motion. The initial perturbations for each of the extra-tropical regions (90°S–30°S and 30°N–90°N) are constructed from the leading 50 singular vectors whereas the initial perturbations for each region targeted on a tropical cyclone is based on the leading 5 singular vectors. In both experiments, initial perturbations are computed for tropical cyclones between 40°S and 40°N. The algorithm determining the optimisation region for a tropical cyclone takes into account the position of the tropical cyclone as predicted by the operational EPS run initialised 12 hours prior to the singular vector initial time (van der Grijn et al. 2004).

Now, the similarity of the different initial perturbations obtained by experiments O and S is discussed for the tropical cyclones undergoing extra-tropical transitions. Three different subspaces are considered: subspace F spanned by singular vectors computed in the full space (from Exp. O); subspace FP obtained by projecting the subspace F into the orthogonal complement of the extra-tropical Northern Hemisphere singular vectors (from Exp. O); subspace S spanned by the singular vectors computed in the orthogonal complement of the extra-tropical SVs (from Exp. S). Consistent with the operational EPS, results for the spaces spanned by the leading five singular vectors are discussed. Structural differences between the three spaces are quantified using the similarity index introduced by Buizza (1994). It measures the degree of parallelism of subspaces. The index varies between 0 (orthogonal subspaces) and 1 (identical subspaces). It is computed as the average square norm of the projection of the ortho-normal basis vectors of one subspace on the other subspace. Here, the index is computed using the total energy inner product.

In the period 11-24 September, there are 26 sets of singular vectors targeted on tropical cyclones where the optimisation regions significantly overlaps with the extra-tropical optimisation region. For these cases, the

Table 4: Similarity indices between singular vector (SV) subspaces  $F$ ,  $FP$ , and  $S$ : ( $F$ ) SVs computed in full space, ( $FP$ ) space spanned by full space SVs projected into the orthogonal complement of the extra-tropical SVs, ( $S$ ) SVs computed in subspace. The latitude of the tropical cyclone position at optimisation time is given in column  $Lat^*$ .

tropical cyclone	Date	$Lat^*(^{\circ}N)$	FP-S	F-S	FP-F
Gordon	11 Sep	25	0.80	0.75	0.85
Gordon	12 Sep	29	0.98	0.89	0.91
Gordon	13 Sep	31	1.00	0.94	0.94
Gordon	14 Sep	31	0.81	0.69	0.82
Gordon	15 Sep	33	0.94	0.76	0.80
Gordon	16 Sep	37	0.74	0.39	0.45
Gordon	17 Sep	38	0.75	0.31	0.39
Gordon	18 Sep	38	0.87	0.46	0.52
Gordon	19 Sep	–	0.80	0.50	0.58
Gordon	20 Sep	–	0.85	0.65	0.72
Shanshan	14 Sep	27	0.99	0.86	0.87
Shanshan	15 Sep	33	0.97	0.74	0.76
Shanshan	16 Sep	39	0.92	0.50	0.54
Shanshan	17 Sep	43	0.80	0.33	0.42
Shanshan	18 Sep	–	0.96	0.37	0.38
Helene	19 Sep	29	0.99	0.95	0.95
Helene	20 Sep	33	0.95	0.60	0.62
Helene	21 Sep	37	0.80	0.42	0.47
Helene	22 Sep	40	0.84	0.53	0.61
Helene	23 Sep	–	0.94	0.45	0.48
Helene	24 Sep	–	0.93	0.46	0.49
Yagi	20 Sep	26	0.96	0.88	0.91
Yagi	21 Sep	31	0.98	0.78	0.79
Yagi	22 Sep	36	0.74	0.42	0.55
Yagi	23 Sep	–	0.93	0.44	0.46
Yagi	24 Sep	–	0.77	0.32	0.39

similarity indices between the three spaces  $F$ ,  $FP$  and  $S$  are listed in Table 4. In 9 of 26 cases, the similarity index between space  $FP$  and space  $F$  is lower than 0.5. This implies that the singular vectors computed in the full space have a significant projection on the extra-tropical singular vectors. In consequence, the subspace singular vectors differ significantly from the singular vectors computed in the full space for these 9 cases. The similarity index between spaces  $F$  and  $S$  drops below 0.5 in 11 cases. These cases include all cases with similarity index  $FP-F$  lower than 0.5. The projection of space  $F$  into the orthogonal complement of the extra-tropical singular vectors increases the similarity to the subspace singular vectors considerably. The similarity index between spaces  $FP$  and  $S$  is less than 0.8 in only 4 cases and it is larger than 0.7 in all 26 cases. In conclusion, the ortho-normalisation approach and the subspace approach appear to yield very similar spaces in most cases.

## 5 Discussion

The focus has been on two optimisation criteria but both approaches can be applied to more than two optimisation criteria. In the ortho-normalisation approach, the  $k$ -th subspace is obtained by projecting the singular vectors of the  $k$ -th optimisation criterion on  $\overline{L_1 + \dots + L_{k-1}}$ . In the subspace approach, the  $k$ -th singular vector

computation is restricted to  $\overline{L_1 + \dots + L_{k-1}}$ .

Whether there is a need to enforce orthogonality, will depend on the particular choice of the optimisation criteria and also on the dimension of the subspaces. An interesting limiting case occurs if one considers two identical optimisation criteria. Let  $L_1$  be the subspace spanned by the leading  $\ell_1$  singular vectors. Then, the first  $\ell_1$  vectors projected on  $\overline{L_1}$  will be identical to zero and explain no forecast error variance at all. In contrast, the  $j$ -th subspace eigenpair (singular value, singular vector) computed in  $\overline{L_1}$  is identical to the  $(\ell_1 + j)$ -th eigenpair computed in the full state space (This statement assumes that the singular values are all different. If they are not, only the spaces for a given singular value are unique.). This property has been exploited to check the correctness of the implementation of subspace singular vectors in the ECMWF Integrated Forecast System (IFS) and in the Eady model example. It could also be employed to restart a singular vector computation, i.e. to append singular vectors to an existing set. Another interesting limiting case occurs if one considers final time metrics involving projections on disjoint parts of the state space, e.g. local projection operators on geographically separated regions. Then, the leading singular vectors computed for different criteria are expected to be almost mutually orthogonal and there is no need to enforce orthogonality. However, orthogonality becomes an issue if one considers overlapping optimisation regions.

The ortho-normalisation approach and the subspace approach have both been applied to the computation of initial perturbations of the ECMWF EPS. The two methods have been exploited in order to extend pole-wards the perturbations targeted on tropical cyclones. Results of daily computations in the period 11-24 September 2006, which covers the extra-tropical transition of tropical cyclones Gordon, Shanshan, Helene and Yagi, indicate that both approaches tend to yield quite similar subspaces. Note, that the pole-ward extension of perturbations targeted on tropical cyclones has been implemented in the operational EPS in September 2004 using the subspace approach. The decision was based on the fact that both approaches incur about the same computational cost but the subspace approach is, in principle, superior to the ortho-normalisation approach. The results presented here suggest that this superiority may be small in practice, though, and the sample size required to demonstrate the superiority of the subspace approach is likely to be computationally prohibitive. The pole-ward extension of the perturbations targeted on tropical cyclones has resulted in a more consistent spread of tropical cyclone tracks during the extra-tropical transitions.

For regional applications, ensemble prediction systems using singular vectors targeted on the region of interest, e.g. part of Europe, have been considered ([Hersbach et al. 2000](#); [Frogner and Iversen 2001](#)). At forecast ranges beyond the singular vector optimisation time, such targeted ensemble prediction systems might be spread deficient because the perturbations have propagated through the region of interest and initial conditions further upstream were not perturbed. Using one of the approaches discussed here, the initial perturbations could be augmented by perturbations based on a second set of singular vectors that are optimised for a much larger region, say the Northern Hemisphere extra-tropics, which contains the region of interest. It is anticipated that this could extend the forecast range at which such targeted ensemble prediction systems would be useful.

Current operational ensemble prediction systems in which initial uncertainties are represented by singular vectors are using one optimisation time only. This raises the question which value of the optimisation time yields the best ensemble forecasts. Short optimisation times will guarantee that tangent-linear dynamics constitutes a more accurate approximation of finite-amplitude perturbation dynamics. The appropriate scale for the amplitude of the initial perturbations is set by the amplitude of typical analysis errors. Estimates of the time scale up to which the tangent-linear approximation is useful for synoptic-scale atmospheric dynamics range from about 24 h to about 72 h ([Gilmour et al. 2001](#); [Reynolds and Rosmond 2003](#)). However, short optimisation times may yield perturbations that are suboptimal for the longer forecast ranges. Thus, a trade-off may be required between the relevance of the perturbations for the longer forecast ranges and the accuracy of the tangent-linear approximation. Furthermore, one would expect that there is no optimisation time that is optimal for all forecast ranges. Therefore, an ensemble using perturbations based on multiple sets of singular vectors, each computed

for a different optimisation time, may be superior to an ensemble based on perturbations computed for a single optimisation time. The example of singular vectors computed for two different optimisation times in the Eady model illustrated that some of the singular vectors for optimisation time 1 have a large projection on singular vectors for optimisation time 2. Therefore, it is expected that approaches which accommodate multiple optimisation times will require a method to enforce the orthogonality of the different sets of singular vectors. The example has also demonstrated that the subspace singular vector approach explains more forecast error variance than the approach of orthogonalising independently computed singular vectors. However, the difference of the explained variances is rather small. In all examples with the Eady model, the variance explained by the ortho-normalisation approach exceeds 80% of the variance explained by the subspace approach.

The overall computational cost of the two approaches to construct mutually orthogonal subspaces for different optimisation criteria is approximately equal because the main cost is incurred by the tangent-linear and adjoint integrations in the singular vector computations and not by the projections. There may be a difference though for time-critical applications like numerical weather prediction if the tangent-linear and adjoint code is not scaling well on parallel computing architectures. The computation of subspace singular vectors implies a sequential approach: The computation of subspace singular vectors for the second optimisation criterion requires the singular vectors of the first optimisation to define the projection operator. To speed up the sequential singular vector computations, the number of processors used for each computation can be increased. In the limit of code that scales perfectly on a parallel architecture, the independent computation and the sequential computation can be run within the same wall-clock time using the same overall resources. It is worth noting here that the operational constraints on the resources available for singular vector computations in ensemble prediction systems can be relaxed significantly by using a trajectory started from a short-range forecast for the tangent-linear and adjoint integrations rather than a trajectory started from an analysis. [Leutbecher \(2005\)](#) showed for the ECMWF EPS that this change yields almost identical singular vectors and does not appear to affect the skill of the probabilistic forecasts.

When multiple optimisation criteria are considered, there is no unique way how to define an optimal subspace for all of them. Here, a sequential approach has been developed. For each criterion a separate optimisation is performed in order to obtain a subspace of a given dimension. The optimisation is restricted to the orthogonal complement of the sum of the already obtained subspaces. The resulting sum of subspaces depends on the order of the optimisation criteria. For one particular criterion one order may explain more forecast error variance than the other orders. However, for another optimisation criterion, another order may be best. This was illustrated by the Eady model example. The non-commutativity of the optimisation criteria is avoided by an alternative method of accounting for multiple optimisation criteria. It optimises a weighted sum of the forecast error variances associated with the individual criteria (Mark Buehner, personal communication). However, the resulting subspace will depend on the choice of weights. Without carefully choosing the weights, the resulting subspace may be dominated by structures that optimise only one of the criteria.

## 6 Conclusions

Here, the representation of initial uncertainty using multiple sets of singular vectors optimised for different criteria is discussed. Methods which represent initial uncertainties consistent with an analysis error covariance estimate  $\mathbf{A}$  require that subspaces associated with different sets of singular vectors are orthogonal with respect to the inner product based on  $\mathbf{A}^{-1}$ .

This paper introduces the concept of singular vectors optimised in a subspace which is orthogonal to a subspace of an independent set of singular vectors. It is proven that this subspace method is optimal in the sense of augmenting an arbitrary given subspace in order to maximise the explained forecast error variance in the

augmented space.

The subspace approach has been compared with a simple ortho-normalisation approach which can also be employed in order to generate initial perturbations from multiple sets of singular vectors that are not already mutually orthogonal.

The subspace approach and the ortho-normalisation approach have been applied to an idealised example based on the Eady model and to the ECMWF EPS. In both applications, the simple ortho-normalisation approach provides perturbations which are quite similar to the perturbations obtained with the subspace approach.

## Acknowledgements

This work has benefitted from discussions with Tim Palmer. I am grateful to Martin Ehrendorfer for suggesting to consider the special case of two identical optimisation criteria. Comments by Martin Jukes and two anonymous reviewers helped to significantly improve the manuscript. Furthermore, I would like to thank Mark Buehner for a discussion on an alternative way of representing multiple optimisation criteria.

## A Appendix: Proof of the subspace singular vector properties

In order to shorten the notation,  $\mathbf{T}_2$  is replaced by  $\mathbf{T}$  in the following. First, statement (a) in Sec. 2.4 is proved. Let  $\mathbf{w}$  be a singular vector of  $\mathbf{T}_s$  with singular value  $\sigma > 0$ . Therefore,  $\mathbf{w}$  solves the eigenproblem  $\mathbf{T}_s^T \mathbf{T}_s \mathbf{w} = \sigma^2 \mathbf{A}^{-1} \mathbf{w}$ . Now, let  $\mathbf{z} \in L_1$ . This implies

$$\mathbf{z}^T \mathbf{A}^{-1} \mathbf{w} = \sigma^{-2} \mathbf{z}^T \mathbf{T}_s^T \mathbf{T}_s \mathbf{w} = \sigma^{-2} (\mathbf{P}(\overline{L}_1) \mathbf{z})^T \mathbf{T}^T \mathbf{T}_s \mathbf{w} = 0, \quad (21)$$

because  $\mathbf{P}(\overline{L}_1) \mathbf{z} = 0$ . Equation (21) implies  $\mathbf{w} \in \overline{L}_1$  which proves (a).

Now we prove statement (b) in Sec. 2.4. The proof relies on the maximum variance property of operator  $\mathbf{T}_s$ . In order to apply this property we first need to prove the following two equivalences (23) and (24).

For all subspaces  $L \subset \overline{L}_1$ ,  $\mathbf{P}(L) = \mathbf{P}(\overline{L}_1) \mathbf{P}(L)$ . This implies

$$\mathbf{T} \mathbf{A}_p(L) \mathbf{T}^T = \mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T, \quad \forall L \subset \overline{L}_1. \quad (22)$$

Therefore

$$\max_{L \subset \overline{L}_1, \dim(L)=k} \text{tr}(\mathbf{T} \mathbf{A}_p(L) \mathbf{T}^T) = \max_{L \subset \overline{L}_1, \dim(L)=k} \text{tr}(\mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T). \quad (23)$$

Next, it is shown that the maximum in Equation (23) remains unaltered if it is computed over all  $k$ -dimensional subspaces in the statespace  $\mathcal{L}$ , i.e.

$$\max_{L \subset \mathcal{L}, \dim(L)=k} \text{tr}(\mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T) = \max_{L \subset \overline{L}_1, \dim(L)=k} \text{tr}(\mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T). \quad (24)$$

It is obvious that L.H.S.  $\geq$  R.H.S. in Equation (24) as  $\overline{L}_1 \subset \mathcal{L}$ . To show L.H.S.  $\leq$  R.H.S., we consider a subspace  $\tilde{L}$  which is not in  $\overline{L}_1$ . We can write  $\tilde{L} = L' + L''$  with  $L' \subset \overline{L}_1$  and  $L'' \subset L_1$  and  $\dim(L') < k$ . As  $\mathbf{P}(\overline{L}_1) \mathbf{P}(L'') = 0$ ,

$$\text{tr}(\mathbf{T}_s \mathbf{A}_p(\tilde{L}) \mathbf{T}_s^T) = \text{tr}(\mathbf{T}_s \mathbf{A}_p(L') \mathbf{T}_s^T) \leq \max_{L \subset \overline{L}_1, \dim(L)=k} \text{tr}(\mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T) \quad (25)$$

which proves L.H.S.  $\leq$  R.H.S. in Equation (24).

Now, we make use of the maximum variance property of the singular vectors of operator  $\mathbf{T}_s$ , Equation (13). This yields,

$$\max_{L \subset \mathcal{L}, \dim(L)=k} \text{tr}(\mathbf{T}_s \mathbf{A}_p(L) \mathbf{T}_s^T) = \text{tr}(\mathbf{T}_s \mathbf{A}_p(L^*[\mathbf{T}_s, k]) \mathbf{T}_s^T). \quad (26)$$

But as  $L^*[\mathbf{T}_s, k] \subset \overline{L_1}$ , operator  $\mathbf{T}_s$  can be replaced by  $\mathbf{T}$  on the right hand side of Equation (26). Combining Equations (23), (24) and (26) proves Equation (17).

## References

- Anderson JL. 1997. The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. *Mon. Weather Rev.* **125**: 2969–2983.
- Buizza R. 1994. Sensitivity of optimal unstable structures. *Q. J. R. Meteorol. Soc.* **120**: 429–451.
- Buizza R, Palmer TN. 1995. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* **52**: 1434–1456.
- Coutinho MM, Hoskins BJ, Buizza R. 2004. The influence of physical processes on extratropical singular vectors. *J. Atmos. Sci.* **61**: 195–209.
- DeVries H, Opsteegh JD. 2005. Optimal perturbations in the Eady model: Resonance versus PV unshielding. *J. Atmos. Sci.* **62**: 492–505.
- Ehrendorfer M, Tribbia JJ. 1997. Optimal prediction of forecast error covariances through singular vectors. *J. Atmos. Sci.* **54**: 286–313.
- Farrell BF. 1988. Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.* **45**: 163–172.
- Farrell BF, Ioannou PJ. 1996. Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.* **53**: 2025–2040.
- Frogner I-L, Iversen T. 2001. Targeted ensemble prediction for northern Europe and parts of the north Atlantic ocean. *Tellus* **53A**: 35–55.
- Gilmour I, Smith LA, Buizza R. 2001. On the duration of the linear regime: Is 24 hours a long time in weather forecasting? *J. Atmos. Sci.* **58**: 3525–3539.
- Hersbach H, Mureau R, Opsteegh JD, Barkmeijer J. 2000. A short-range to early-medium range ensemble prediction system for the european area. *Mon. Weather Rev.* **128**: 3501–3519.
- Hoskins BJ, Coutinho MM. 2005. Moist singular vectors and the predictability of some high impact european cyclones. *Q. J. R. Meteorol. Soc.* **131**: 581–601.
- Lawrence AR, Leutbecher M, Palmer TN. 2007. Comparison of total energy and Hessian singular vectors: Implications for observation targeting. *Q. J. R. Meteorol. Soc.* (submitted, available at [http://www.ecmwf.int/publications/library/ecpublications/\\_pdf/tm/501-600/tm517.pdf](http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm/501-600/tm517.pdf)).
- Leutbecher M. 2005. On ensemble prediction using singular vectors started from forecasts. *Mon. Weather Rev.* **133**: 3038–3046.

- Leutbecher M, Palmer TN. 2007. Ensemble forecasting. *J. Comp. Phys.* (in press, DOI: 10.1016/j.jcp.2007.02.014, available also as ECMWF Tech. Memo. 514).
- Lewis JM. 2005. Roots of ensemble forecasting. *Mon. Weather Rev.* **133**: 1865–1885.
- Lorenz EN. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus* **17**: 321–333.
- Molteni F, Buizza R, Palmer TN, Petroliagis T. 1996. The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc.* **122**: 73–119.
- Morgan MC, Chen CC. 2002. Diagnosis of optimal perturbation evolution in the eady model. *J. Atmos. Sci.* **59**: 169–185.
- Mukougawa H, Ikeda T. 1994. Optimal excitation of baroclinic waves in the eady model. *J. Meteorol. Soc. Japan* **72**: 499–513.
- Palmer TN, Gelaro R, Barkmeijer J, Buizza R. 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **55**: 633–653.
- Puri K, Barkmeijer J, Palmer TN. 2001. Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Q. J. R. Meteorol. Soc.* **127**: 709–731.
- Reynolds CA, Rosmond TE. 2003. Nonlinear growth of singular-vector-based perturbations. *Q. J. R. Meteorol. Soc.* **129**: 3059–3078.
- van der Grijn G, Paulsen JE, Lalaurette F, Leutbecher M. 2004. Early medium-range forecasts of tropical cyclones. *ECMWF Newsletter* **102**: 7–14.
- Walser A, Arpagaus M, Appenzeller C, Leutbecher M. 2006. The impact of moist singular vectors and horizontal resolution on short-range limited-area ensemble forecasts for two European winter storms. *Mon. Weather Rev.* **134**: 2877–2887.