Initial Uncertainties in the EPS: Singular Vector Perturbations

Training Course 2017
An evolving EPS

- EPS 1992–2010: initial perturbations based on singular vectors (SVs)
- Are SVs optimal?
  - Ideally, SVs should be computed with an initial time norm based on the analysis uncertainty of the day.
  - However, if a good estimate of the analysis uncertainty of the day is available, SVs may not be required any more for the initial perturbations.
- The goal is to obtain a sample of the distribution of initial uncertainty from an ensemble of data assimilations
Outline of this lecture

1. singular vectors?
2. perturbations?
3. some background:
   - perturbation growth etc.
   - norms
   - singular value decomposition
   - tangent-linear system
4. an idealised example: singular vectors in the Eady model
5. SVs in the operational EPS
6. initial condition perturbations
Singular Vectors?

initial SVs evolved SVs

analysis forecast

initial SVs evolved SVs
Perturbations?

\[ t_0 \quad t_1 \]

- \( \text{CF} \)
- \( \text{PF} \)
- \( \text{PF} - \text{CF} \)
Nonlinear Model

Consider the spatially discretised equations describing the atmospheric dynamics and physics written in this form

\[ \frac{d}{dt} \mathbf{x} = F(\mathbf{x}), \quad \text{with} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad F = \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix} \]

where \( \mathbf{x} \in \mathbb{R}^N \) denotes the \( N \)-dimensional state vector and \( F(\mathbf{x}) \in \mathbb{R}^N \) its tendency. Integrate from \( t_0 \) to \( t \) gives the nonlinear model:

\[ \mathbf{x}(t) = \mathcal{F}(\mathbf{x}(t_0)) \]
Tangent-linear system

Let $x_r(t)$ be a solution of

$$\frac{d}{dt} x = F(x),$$

(1)

Then the *tangent-linear system* is given by

$$\frac{d}{dt} x = A(x_r(t)) \cdot x,$$

(2)

where $[A(x)]_{jk} = (\partial F_j / \partial x_k)(x)$ denotes the Jacobi matrix of $F$.

For any solution $x$ of (2), $x_r + \varepsilon x$ approximates a solution of (1) starting at $x_r(t_0) + \varepsilon x(t_0)$ to first order in $\varepsilon$.

The *tangent-linear* propagator from $t_0$ to $t_1$ is the matrix $M_{[t_0,t]}$ such that $M_{[t_0,t]} x_0$ is a solution of (2) for any initial perturbation $x_0$ and where $M_{[t_0,t_0]} = I$. 
Perturbation Growth

Perturbation growth is defined as:

\[ \sigma^2 = \frac{\langle x(t), x(t) \rangle}{\langle x(t_0), x(t_0) \rangle} \]

\[ = \frac{\langle M_{[t_0,t]} x(t_0), M_{[t_0,t]} x(t_0) \rangle}{\langle x(t_0), x(t_0) \rangle} \]

\[ = \frac{\langle M_{[t_0,t]}^T M_{[t_0,t]} x(t_0), x(t_0) \rangle}{\langle x(t_0), x(t_0) \rangle} \]

with inner product \( \langle \cdot, \cdot \rangle \) and growth factor \( \sigma^2 \).

\[ \Rightarrow \] Largest growth is associated with eigenvectors of \( M_{[t_0,t]}^T M_{[t_0,t]} \).

These eigenvectors are determined by a singular value decomposition of \( M_{[t_0,t]} \).
**Norms**

- The definition of singular vectors in the context of ensemble prediction involves norms (based on an inner product or metric). These are required to measure the amplitude of perturbations.

\[ \langle x, x \rangle_C = x^T C x \]

where \( C \) is symmetric (\( C^T = C \)) and positive definite (\( x^T C x > 0 \) for \( x \neq 0 \)).

- For predictability applications, the appropriate choice for the **initial time norm** is the analysis error covariance metric, *i.e.* the norm that is based on the inverse of the initial error covariance matrix (or some estimate thereof).

\[ \|x\|_i^2 = x^T C_0^{-1} x \]

- The **final time norm** \( \|x\|_f \) is a convenient RMS measure of forecast error.

- **Total energy norm** is used both at initial and final time for the operational singular vector computations at ECMWF:

\[ \|x\|_E^2 = x^T E x = \frac{1}{2} \int_{p_0}^{p_1} \int_S \left( u^2 + v^2 + \frac{c_p}{T_r} T^2 \right) dp \, ds + \frac{1}{2} R_d T_r p_r \int_S (\ln \rho_{sfc})^2 ds \]
On the choice of the initial time norm

- The structure of singular vectors depends on the choice of the norm, in particular the initial time norm.

- An enstrophy norm at initial time penalises perturbations with small spatial scales, the initial SVs are planetary-scale structures.

- A streamfunction variance norm at initial time penalises the large scales and favours sub-synoptic scale perturbations.

- With a total energy norm at initial time, the energy spectrum of the initial SVs is “white” and best matches the spectrum of analysis error estimates from analyses differences (Palmer et al. 1998)
singular value decomposition of a matrix

Consider a matrix

\[ Q = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{m1} & \cdots & q_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n} \]

Its singular value decomposition is defined as

\[ Q = U S V^T, \quad (3) \]

where \( U \) and \( V \) are orthogonal \( m \text{-by-} m \) and \( n \text{-by-} n \) matrices. Matrix \( S \) is a diagonal \( m \text{-by-} n \) matrix (\( s_{ij} = 0 \) if \( i \neq j \), \( s_{jj} \equiv \sigma_j \)). The values \( \sigma_j \) on the diagonal of \( S \) are called singular values. The columns \( u_j \) of \( U \) are referred to as left singular vectors and the columns \( v_j \) of \( V \) are referred to as right singular vectors.

Eq. (3) implies that

\[ Q v_j = \sigma_j u_j \]

One can show that the \( v_j \) are the eigenvectors of \( Q^T Q \)

see Golub and Van Loan: *Matrix Computations* for further details
spherical valued decompositions of the propagator

\[ M = USV^T \rightarrow Mv_j = \sigma_j u_j \]

with the (initial) singular vectors \( v_j \) being the eigenvectors and the squared singular values \( \sigma_j^2 \) being the eigenvalues of \( M^TM \). The \( u_j \) are called the evolved singular vectors.

Singular vectors are optimal perturbations in the following sense.

- the ratio of the final time norm to the initial time norm is given by the singular value:
  \[ \frac{\|Mv_j\|_f}{\|v_j\|_i} = \sigma_j \]  
  \[ (4) \]

- Singular vector \( j \) is the direction in phase space that maximises the ratio of norms in the subspace orthogonal (with respect to \( C_0^{-1} \)) to the space spanned by singular vectors \( 1 \ldots j - 1 \).
The tangent-linear model and its adjoint

- For a numerical model with $\sim 10^5 - 10^8$ variables it is not possible to obtain the propagator $M$ as a matrix.
- Instead algorithmic differentiation is used to obtain the first derivative of the numerical algorithm that represents the forecast model. For any initial perturbation $x$, the evolved perturbation $Mx$ is obtained via an integration of the tangent-linear model.
- Then, the numerical algorithm representing $M^T$ the adjoint (transpose) of the propagator is constructed. The adjoint model is integrated backward from $t_1$ to $t_0$.
- The reference solution $x_r(t)$ about which the equations are linearised is referred to as trajectory.
- The time interval the SVs are calculated for is called the optimization interval.

See also related presentations by Angela Benedetti, Marta Janisková and the “Hands on: Coding of Tangent Linear and Adjoints” session in Training Course on Data Assimilation & Use of Satellite Data
SVs in the Eady model

- channel with periodic boundary conditions in the zonal direction
- linear shear of basic state flow $\overline{U} = Sz$

\[ f\text{-plane with } f = 10^{-4}\text{ s}^{-1} \]
\[ \text{Brunt-Vaisala frequency } N = 10^{-2}\text{ s}^{-1} \]
\[ \text{total energy norm at initial and final time} \]
\[ \text{discretisation: 21 levels in the vertical, 16 wavenumbers in the horizontal} \]
SVs in the Eady model: $t_{\text{opt}} = 24$ h

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 6.4$, $\sigma_2 = 6.2$, $\sigma_3 = 6.1$. 
SVs in the Eady model: \( t_{\text{opt}} = 48 \text{ h} \)

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: \( \sigma_1 = 24.4 \), \( \sigma_2 = 22.3 \), \( \sigma_3 = 17.9 \).
Eady model: $t_{\text{opt}} = 48$ h, $t = 0$ h

Streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

Singular values: $\sigma_1 = 24.4$, $\sigma_2 = 22.3$, $\sigma_3 = 17.9$. 
Eady model: \( t_{\text{opt}} = 48 \text{ h}, \quad t = 6 \text{ h} \)

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: \( \sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9. \)
Eady model: $t_{\text{opt}} = 48 \text{ h}, \quad t = 12 \text{ h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9$. 
Eady model: \( t_{\text{opt}} = 48 \text{ h}, \quad t = 18 \text{ h} \)

Streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

Singular values: \( \sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9. \)
Eady model: $t_{\text{opt}} = 48 \text{ h}, \quad t = 24 \text{ h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9$. 
Eady model: $t_{\text{opt}} = 48 \text{ h}, \quad t = 30 \text{ h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9.$
Eady model: $t_{\text{opt}} = 48 \text{ h}, \quad t = 36 \text{ h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9$. 
Eady model: $t_{\text{opt}} = 48 \text{ h}, \quad t = 42 \text{ h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9$. 
Eady model: $t_{opt} = 48 \, \text{h}, \quad t = 48 \, \text{h}$

streamfunction perturbation, SV 1 (top), SV 2 (middle), SV 3 (bottom)

singular values: $\sigma_1 = 24.4, \quad \sigma_2 = 22.3, \quad \sigma_3 = 17.9$. 
Growth mechanisms

- PV unshielding
- Intensification of boundary thermal anomalies through winds associated with interior PV anomalies
- Interaction of waves on upper and lower boundary

See Badger and Hoskins (2001); Morgan and Chen (2002); DeVries and Opsteegh (2005); De Vries et al. (2009)
Singular vectors in the operational EPS

- \( t_{\text{opt}} \equiv t_1 - t_0 = 48 \text{ h} \)
- resolution: T42L62

**Extra-tropics:** 50 SVs for N.-Hem. \((30^\circ \text{N} – 90^\circ \text{N})\)  
+ 50 SVs for S.-Hem.\((30^\circ \text{S} – 90^\circ \text{S})\). Tangent-linear model with vertical diffusion and surface friction only.

**Tropical cyclones:** 5 singular vectors per region targeted on active tropical depressions/cyclones. Up to 6 such regions. Tangent-linear model with representation of diabatic processes (large-scale condensation, convection, radiation, gravity-wave drag, vert. diff. and surface friction).

**Localisation** is required to avoid that too many leading singular vectors are located in the dynamically more active winter hemisphere (Buizza 1994). Also required to obtain (more slowly growing) perturbations associated with tropical cyclones (Puri et al. 2001). In order to optimise perturbations for a specific region simply replace the propagator \( \mathbf{M} \) in the equations by \( \mathbf{P} \mathbf{M} \), where \( \mathbf{P} \) denotes the projection operator which sets the state vector \( (T, u, v, \ln p_{\text{sfc}}) \) in grid-point space to zero outside the region of interest and is the identity inside it.
Singular values $\sigma_j$ — extra-tropics

**Northern Hem.**
- solid: 2005070100
- dashed: 2005092100
- dotted: 2005122100
- chain-dashed: 2006032100

**Southern Hem.**
- solid: 2005070100
- dashed: 2005092100
- dotted: 2005122100
- chain-dashed: 2006032100
Singular vector growth characteristics

average energy of the leading 50 singular vectors

initial time \((\times 50)\), final time \(t = 48 \text{ h} (\times 1)\)

- : total energy; - : kinetic energy

Northern hemisphere extra-tropics, 2006032100

vertical profile

spectrum

<table>
<thead>
<tr>
<th>wave number</th>
<th>wave length</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>8000 km</td>
</tr>
<tr>
<td>10</td>
<td>4000 km</td>
</tr>
<tr>
<td>20</td>
<td>2000 km</td>
</tr>
<tr>
<td>40</td>
<td>1000 km</td>
</tr>
</tbody>
</table>

200 hPa \(\leftrightarrow\) level 20
500 hPa \(\leftrightarrow\) level 35
850 hPa \(\leftrightarrow\) level 48
300 hPa \(\leftrightarrow\) level 27
700 hPa \(\leftrightarrow\) level 42
925 hPa \(\leftrightarrow\) level 52

200 hPa \(\leftrightarrow\) level 20
500 hPa \(\leftrightarrow\) level 35
850 hPa \(\leftrightarrow\) level 48
300 hPa \(\leftrightarrow\) level 27
700 hPa \(\leftrightarrow\) level 42
925 hPa \(\leftrightarrow\) level 52

Initial Uncertainties in the EPS (I)
Regional distribution of Northern Hem. SVs

square root of vertically integrated total energy of SV 1–50 (shading)
500 hPa geopotential (contours)
initial singular vectors, 21 March 2006, 00 UTC

Evolved singular vectors, 23 March 2006, 00 UTC
Singular vector 5: initial time

21 March 2006, 00 UTC

Temperature at \( \approx 700 \text{ hPa} \)
Singular vector 5: final time

23 March 2006, 00 UTC
meridional wind component at ≈ 300 hPa
Initial condition perturbations

- Initial condition uncertainty is represented by a (multi-variate) Gaussian distribution in the space spanned by the leading singular vectors.
- The perturbations based on a set of singular vectors $v_1, \ldots, v_m$ are of the form

$$x_j = \sum_{k=1}^{m} \alpha_{jk} v_k$$  \hspace{1cm} (5)

- The $\alpha_{jk}$ are independent draws from a truncated Gaussian distribution.
- The Gaussian is truncated at $\pm 3$ standard deviations to avoid numerical instabilities for extreme values.
- The width of the distribution is set so that the spread of the ensemble matches the root-mean square error in an average over many cases ($\beta \approx 10$).
Initial condition perturbations (2)

- For the **extra-tropical perturbations**, the leading 50 initial singular vectors (in each hemisphere) are combined with the leading 50 evolved singular vectors (replaced by EDA perturbations since 22 June 2010)

- For each of the (up to 6) optimisation regions targeted on a tropical cyclone, the leading 5 initial singular vectors are combined.

- To make sure that the ensemble mean is centred on the unperturbed analysis a **plus-minus symmetry** has been introduced:
  - coefficients for members 1, 3, 5, ..., 49 are sampled,
  - the perturbation for members 2, 4, 6, ... 50 is set to minus the perturbation of the member $j - 1$ ($x_j = -x_{j-1}$).

Note: The sign of a singular vector itself is arbitrary.
Initial condition perturbation for member 1

Temperature (every 0.2 K); 21 March 2006, 00 UTC

at $\approx 700$ hPa

at $50^\circ$N
Initial condition perturbation for member 2

Temperature (every 0.2 K); 21 March 2006, 00 UTC
at \(\approx 700 \) hPa

at 50\(^\circ\)N
Initial condition perturbation for member 5

Temperature (every 0.2 K); 21 March 2006, 00 UTC
at ≈ 700 hPa

at 50°N
Initial condition perturbation for member 50

Temperature (every 0.2 K); 21 March 2006, 00 UTC
at ≈ 700 hPa

at 50°N
linear algebra

**m-by-n matrix**, $m$ rows and $n$ columns $Q = egin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$

**matrix multiplication**: entry in $i$-th row and $j$-th column is the inner product of row $i$ of $A$ and col. $j$ of $B$

\[ C = AB \iff c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj} \]

**matrix transpose**: swap rows with columns

\[ C = A^T \iff c_{ij} = a_{ji} \quad ; \quad (AB)^T = B^T A^T \]

**column vector**: $m$-by-1 matrix

\[ v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \]

**row vector**: 1-by-$m$ matrix

\[ v^T = (v_1 \; \ldots \; v_m) \]
linear algebra (cont.)

inner product:

\[ x^T y = \sum_{j=1}^{m} x_j y_j \]

or more generally

\[ x^T C y = \sum_{j=1}^{m} \sum_{k=1}^{m} x_j C_{jk} y_k \]

where \( C \) is symmetric (\( C^T = C \)) and positive definite (\( x^T C x > 0 \) for \( x \neq 0 \)).

orthogonal and orthonormal sets of vectors:

orthogonal:

\[ x^T y = 0 \]

orthonormal = orthogonal and normalised:

\[ v_j^T v_k = \delta_{jk} = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases} \]

orthogonal matrix:

row vectors and column vectors are orthonormal sets of vectors

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{pmatrix}
\]

\[ v^T v = I \]
Singular vectors of the propagator

Consider an initial time norm \( \|x\|_i^2 = x^T C_0^{-1} x \)
and a final time norm \( \|x\|_f^2 = x^T D x \),

where \( C_0 \) and \( D \) are positive definite symmetric matrices. Then, we consider the propagator for a fixed time interval \( \mathbf{M} \equiv \mathbf{M}(t_0, t_1) \) and apply scalings depending on the norms (non-dimensionalisation). The reason for this particular scaling of the propagator will become obvious later.

The singular value decomposition of the scaled propagator \( \tilde{\mathbf{M}} \) is

\[
\tilde{\mathbf{M}} \equiv D^{1/2} M C_0^{1/2} = \tilde{\mathbf{U}} \mathbf{S} \tilde{\mathbf{V}}^T
\]

(6)

Here, \( \mathbf{S} \) is the diagonal matrix containing the decreasing singular values \( \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N \); \( \tilde{\mathbf{U}} \) and \( \tilde{\mathbf{V}} \) contain the non-dimensional left and right singular vectors, respectively (see slide 6).
Singular vectors of the propagator (2)

Usually, the left and right singular vectors are transformed to physical space:

\[
\begin{align*}
\text{initial SVs} & \quad V = C_0^{1/2} \tilde{V} \\
\text{normalised evolved SVs} & \quad U = D^{-1/2} \tilde{U}
\end{align*}
\]

The initial SVs are orthonormal with respect to \(C_0^{-1}\):

\[
v_j^T C_0^{-1} v_k = \delta_{jk}.
\]

The normalised evolved SVs are orthonormal with respect to \(D\):

\[
u_j^T D u_k = \delta_{jk}.
\]

The singular value decomposition of the propagator in dimensional form reads

\[
M = USV^T C_0^{-1}.
\] (7)

Let \(v_j\) denote the \(j\)-th column of \(V\). It is referred to as (initial) singular vector \(j\). From (7) it is obvious that

\[
M v_j = \sigma_j u_j
\] (8)

\(\sigma_j u_j\) is the evolved SV.
EOFs of a linear estimate of the FC error covariance matrix

The initial singular vectors are orthonormal with respect to the estimate of the inverse initial error covariance matrix:

\[ \mathbf{V}^T \mathbf{C}_0^{-1} \mathbf{V} = \mathbf{I} \quad \text{(identity matrix)}; \quad \Rightarrow \quad \mathbf{C}_0 = \mathbf{V} \mathbf{V}^T. \]

Using (7), the linear evolution of the covariance estimate from \( t_0 \) to \( t_1 \) can be expressed as

\[ \mathbf{C}_1 = \mathbf{M} \mathbf{C}_0 \mathbf{M}^T = \mathbf{U} \mathbf{S}^2 \mathbf{U}^T. \]  \hspace{1cm} (9)

with scaling of errors with \( \mathbf{D}^{1/2} \):

\[ \mathbf{D}^{1/2} \mathbf{C}_1 \mathbf{D}^{1/2} = \mathbf{D}^{1/2} \mathbf{M} \mathbf{V} \mathbf{V}^T \mathbf{M}^T \mathbf{D}^{1/2} = (\mathbf{D}^{1/2} \mathbf{U}) \mathbf{S}^2 (\mathbf{D}^{1/2} \mathbf{U})^T. \]  \hspace{1cm} (10)

• Eqn. (10) provides the EOF decomposition of the (scaled) forecast error covariance estimate.
• The leading singular vectors evolve into the directions of the leading EOFs of the (scaled) forecast error covariance estimate.
• This property makes the singular vectors an attractive basis for sampling initial condition uncertainties.
References and further readings


**Tangent-linear and adjoint models**

[Links to ECMWF presentations](http://www.ecmwf.int/newsevents/training/meteorological_presentations/pdf/DA/TangLin.pdf)
[Links to ECMWF presentations](http://www.ecmwf.int/newsevents/training/meteorological_presentations/pdf/DA/Param.pdf)
[Links to ECMWF presentations](http://www.ecmwf.int/newsevents/training/meteorological_presentations/pdf/PA/assim123.pdf)