TRANSFORMS AND PRECONDITIONING IN THE MET OFFICE 3D VAR SCHEME

Ross Bannister
D.A.R.C.

Cost Function in w-space (reduced resolution, real space):

$$J(\vec{w}') = J_b(\vec{w}') + J_o(\vec{w}')$$

$$J_b(\vec{w}') = \frac{1}{2}(\vec{w}'^b - \vec{w}')^T \mathbf{B}^{-1}(\vec{w}'^b - \vec{w}')$$

$$J_o(\vec{w}') = \frac{1}{2}(\vec{y}^o - H(\vec{w}', \vec{w}^g))^T (\mathbf{E} + \mathbf{F})^{-1}(\vec{y}^o - H(\vec{w}', \vec{w}^g))$$

Gradients w.r.t. w':

$$\frac{\mathrm{d}J_b}{\mathrm{d}\vec{w}'} = -\mathbf{B}^{-1}(\vec{w}'^b - \vec{w}')$$

$$\frac{\mathrm{d}J_o}{\mathrm{d}\vec{w}'} = -\mathbf{H}^T(\mathbf{E} + \mathbf{F})^{-1}(\vec{y}^o - H(\vec{w}', \vec{w}^g))$$

Hessian,

$$\frac{\mathrm{d}^2J_b}{\mathrm{d}\vec{w}'^2} = \mathbf{B}^{-1}$$

$$\frac{\mathrm{d}^2J_o}{\mathrm{d}\vec{w}'^2} = \mathbf{H}^T(\mathbf{E} + \mathbf{F})^{-1}\mathbf{H}$$

Contruct a control variable, $\vec{v}$:

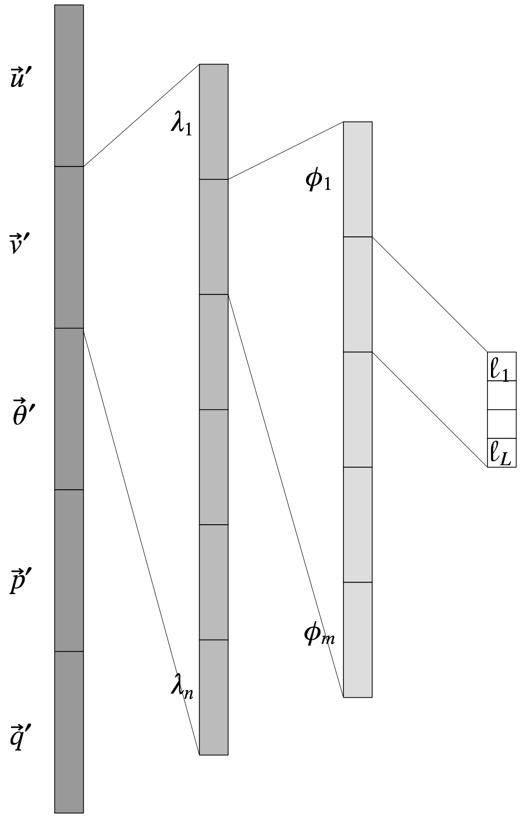
Transform $\vec{w}' = U(\vec{v})$

Inverse $\vec{v} = T(\vec{w}')$

$$\frac{\mathrm{d}^2J_b}{\mathrm{d}\vec{v}^2} = \mathbf{I}$$

Such that

What is the basis of a $\vec{w}'$ vector?



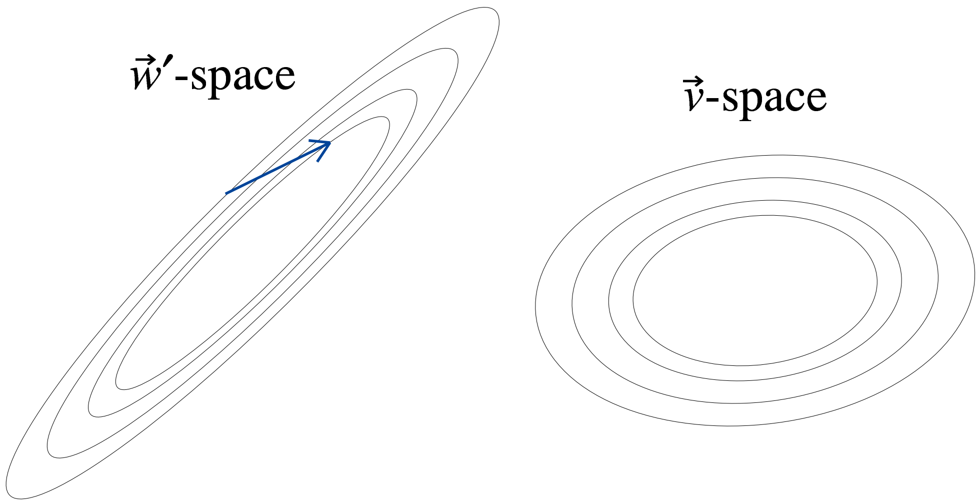
Perturbations are from a guess state:

$$\vec{w}' = \vec{w} - \vec{w}_g$$

$$\vec{w}'^b = \vec{w}^b - \vec{w}_g$$

Why do we want to make a transformation to $\vec{v}$-space?

- Makes the problem manageable.
- Preconditioning - makes the minimization process more efficient and accurate.
- No preconditioning: (Largest e.v.)/(smallest e.v.) ~ 10^{10} !
- Makes the scheme more complicated to understand.
- Balance problems?



What is the principle of the preconditioning transform?

To make the 'weight' of each control variable equal.

$$\vec{w}' = \mathbf{U}\vec{v} \qquad \vec{v} = \mathbf{T}\vec{w}'$$

$$\mathbf{U} = \mathbf{T}^{-1}$$

$$J_b = \frac{1}{2}(\mathbf{U}\vec{v}^b - \mathbf{U}\vec{v})^T \mathbf{B}^{-1} (\mathbf{U}\vec{v}^b - \mathbf{U}\vec{v})$$

$$= \frac{1}{2}(\vec{v}^b - \vec{v})^T \mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} (\vec{v}^b - \vec{v})$$

Choose **U** such that $\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} = \mathbf{I}$

$$\mathbf{U}^{-1}\mathbf{B}\left(\mathbf{U}^T\right)^{-1} = \mathbf{I} \qquad \Rightarrow \qquad \mathbf{B} = \mathbf{U}\mathbf{U}^T$$

$$\mathbf{T}\mathbf{B}\mathbf{T}^T = \mathbf{I}$$

U is not a unitary or orthogonal transform, instead it is like the square-root of **B**.

<u>The information regarding the covariances is transferred into the transformation itself (and inverts **B**!).</u>

This is done in two steps:

$$\text{Let } \mathbf{U} = \mathbf{U}_2\mathbf{U}_1$$

$$\mathbf{U}_1^{-1}\mathbf{U}_2^{-1}\mathbf{B}\left(\mathbf{U}_2^T\right)^{-1}\left(\mathbf{U}_1^T\right)^{-1} = \mathbf{I} \qquad *$$

Consider **B** afresh. Diagonalize with a transform $\mathbf{Y}^T$:

$$\mathbf{Y}^T\mathbf{B}\mathbf{Y} = \Lambda$$

eigenfunctions, rows of $\mathbf{Y}^T$

eigenvalues, diagonal matrix Λ

There are, by definition, no co-variances between the eigenmodes. Can now 'remove' the variance by:

$$\Lambda^{-1/2}\mathbf{Y}^T\mathbf{B}\mathbf{Y}\Lambda^{-1/2} = \mathbf{I}$$

c.f. (*) to show that:

$$\mathbf{U}_1 = \Lambda^{1/2} \quad \Rightarrow \quad \mathbf{T}_1 = \Lambda^{-1/2}$$

$$\mathbf{U}_2 = \mathbf{Y} \quad \Rightarrow \quad \mathbf{T}_2 = \mathbf{Y}^T$$

Problem: **B** is too large to work with (even at half resolution).

# fields	# long. points	# lat. points	# levels	# elements **B**
5	216	163	30	$> 10^{13}$
5	48	37	42	$> 10^{11}$

The solution in three easy stages ...

Assume that:

- We can choose an alternative set of physical parameters which are only weakly correlated,
- The covariances within each parameter can be 'removed' separately (e.g. vertical and horizontal parts normalized independently).
- We can use the last section as a guide.

(i) The first stage of the **T**-transform (parameter transform).

(ii) The second stage is a vertical transformation.

(iii) The third stage is a horizontal transformation.

$$\mathbf{T} = \mathbf{T}_h \mathbf{T}_v \mathbf{T}_p$$

(i) The parameter transform

Parameter	Eqs.
ψ	$\nabla^2 \psi = \vec{k} \cdot \nabla \times \vec{u}$
χ	$\nabla^2 \chi = \nabla \cdot \vec{u}$
${}^A p$	$p = {}^G p + {}^A p$
μ	q / q_{sat}

	ψ	χ	${}^A p$	μ
ψ	$\mathbf{B}_\psi$	0	0	0
χ	0	$\mathbf{B}_\chi$	0	0
${}^A p$	0	0	$\mathbf{B}_{{}^A p}$	0
μ	0	0	0	$\mathbf{B}_\mu$

$\mathbf{B}^{-1}$ has a similar structure

Background term is written (in terms of parameter perturbations),

$$J_B(v_p) = \frac{1}{2}(\vec{v}_p^b - \vec{v}_p)^T \mathbf{B}^{-1}(\vec{v}_p^b - \vec{v}_p)$$

$$= \frac{1}{2}(\vec{\psi}^b - \vec{\psi})^T \mathbf{B}_\psi^{-1}(\vec{\psi}^b - \vec{\psi})$$

$$+ \frac{1}{2}(\vec{\chi}^b - \vec{\chi})^T \mathbf{B}_\chi^{-1}(\vec{\chi}^b - \vec{\chi}) + \ \ldots$$

$$\vec{v}_p = \mathbf{T}_p \vec{w}' = \begin{pmatrix} \vec{\psi} \\ \vec{\chi} \\ {}^A\vec{p} \\ \vec{\mu} \end{pmatrix}$$

Cov. matrices for each parameter (e.g. ψ) - *outer* or *tensor* product:

$$\text{Cov} = \overline{(\vec{\psi} - \vec{\psi}_t)(\vec{\psi} - \vec{\psi}_t)^T}$$

<u>The size of these covariance matrices is still too large.</u>

Do remaining transformations for each parameter separately.

(ii) Vertical transform

Aim (each parameter):
Reformulate (i) the state variable and (ii) the cov. matrix in terms of modes which are **uncorrelated** in the vertical, each of which having **unit variance**.

We can calculate the vertical covariances for each column:

$$\mathrm{Cov}(\lambda, \phi; \ell, \ell') = \overline{(\psi(\lambda, \phi, t; \ell) - \psi_t(\lambda, \phi, t; \ell)) \times (\psi(\lambda, \phi, t; \ell') - \psi_t(\lambda, \phi, t; \ell'))}$$

Average over all (λ, ϕ) to form global covariance matrix:

$$\text{Let} \quad \mathbf{B}_\psi^{vert}(\ell, \ell') = \langle \mathrm{Cov}(\lambda, \phi; \ell, \ell') \rangle$$

Decompose this such that $\mathbf{B}_\psi^{vert} = \mathbf{I}$, as before (with weighting).

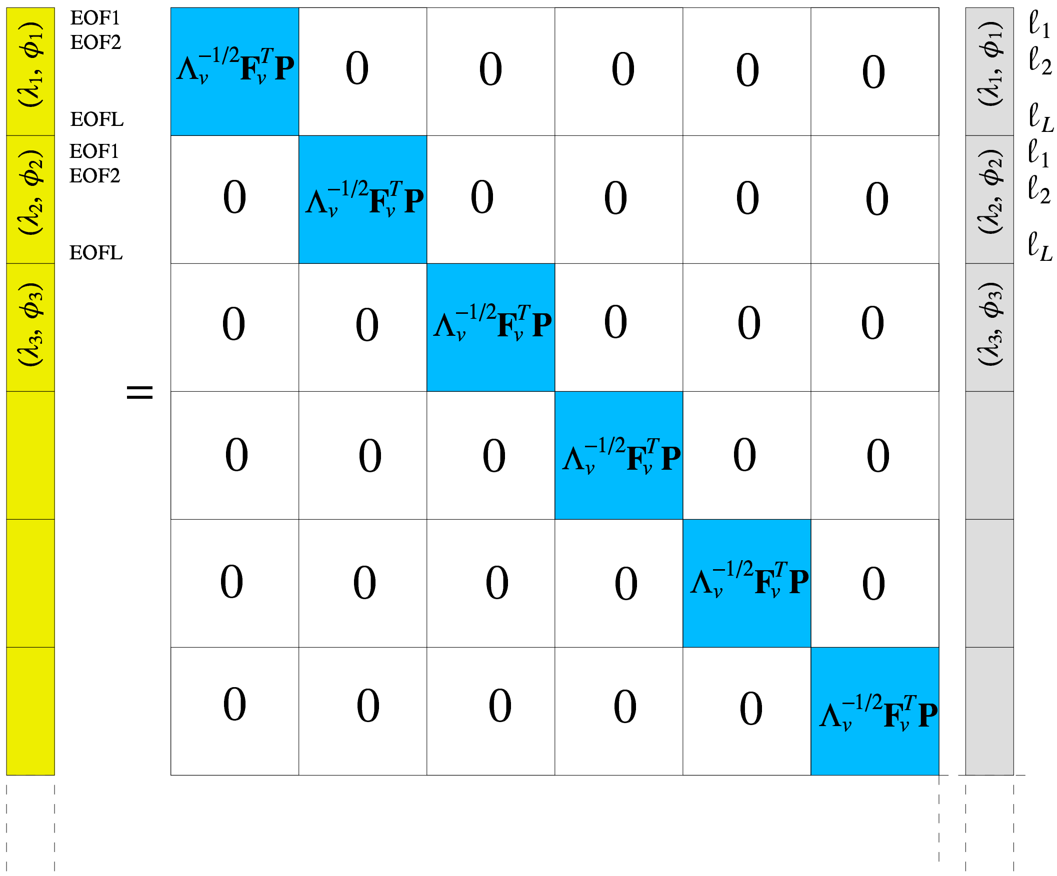
$$\Lambda_v^{-1/2}\mathbf{F}_v^T\mathbf{P}\mathbf{B}_\psi^{vert}\mathbf{P}\mathbf{F}_v\Lambda_v^{-1/2} = \mathbf{I}$$

$$\mathbf{T}_v^{vert}\,(global\ av.) = \Lambda_v^{-1/2}\mathbf{F}_v^T\mathbf{P}$$

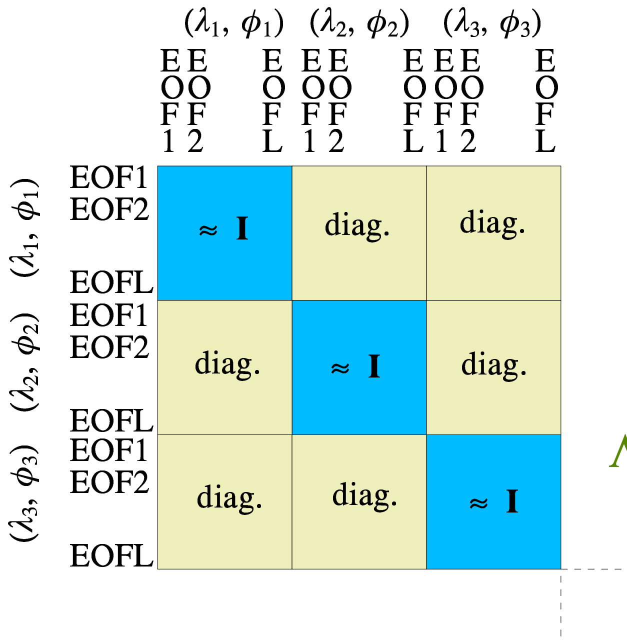
How can we make this into a transform acting on $\vec{\psi}$?

Transforming the state vector ($\vec{\psi}$):

$$\vec{\psi}_{EOF} \quad = \quad \mathbf{T}_v \quad \vec{\psi}$$



Transforming the error covariance matrix: $\mathbf{T}_v\mathbf{B}_\psi\mathbf{T}_v^T$

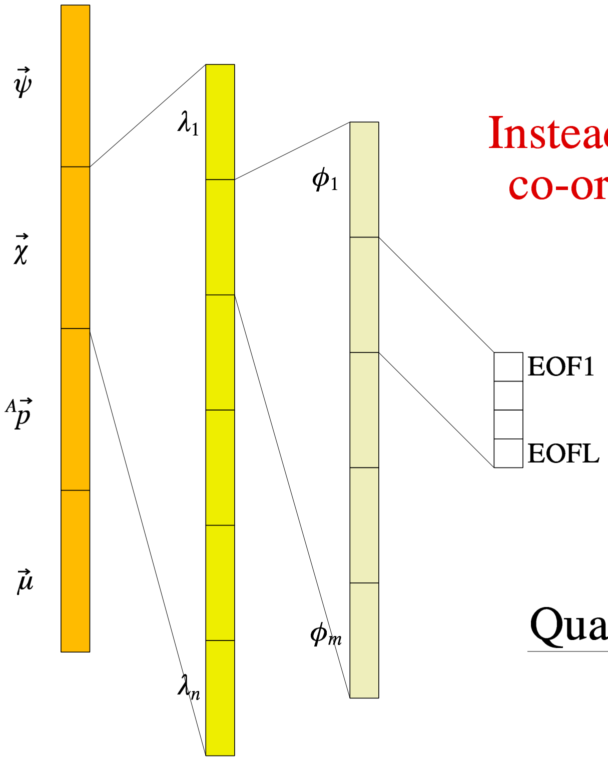


Include lat. variation

$$\Lambda_v \rightarrow \Lambda_v(\phi)$$

$$\Lambda_v(\phi) = \mathbf{F}_v^T\mathbf{P}\mathbf{B}_\psi^{vert}(\phi)\mathbf{P}\mathbf{F}_v$$

Transforming the state vector: $\vec{v}_v = \mathbf{T}_v \vec{v}_v$



Instead of labelling the vertical co-ordinate with 'level', label with 'EOF index'

Quantities are perturbations

Think of surfaces of constant vertical EOF index.

This is the result of the vertical transform.

Horizontal transform

(iii) Horizontal Transform

Aim (each parameter):
Reformulate (i) the state variable and (ii) the cov. matrix in terms of modes which are **uncorrelated** in the horizontal, each of which having **unit variance**.

Decompose into modes which we assume are uncorrelated.
Effectively (for one ψ-EOF suface):

$$\Lambda_h^{-1/2}\mathbf{F}_h^T\mathbf{P}\mathbf{B}_\psi^{hor}\mathbf{P}\mathbf{F}_h\Lambda_h^{-1/2} = \mathbf{I}$$

$$\mathbf{T}_h^{hor} = \Lambda_h^{-1/2}\mathbf{F}_h^T\mathbf{P}$$

$\mathbf{B}_\psi^{hor}$ is not explicitly calculated. Choose:

$\mathbf{P}$ as a weight matrix, different from before,
$\mathbf{F}_v^T$ as a horizontal spectral transform, and
$\Lambda_h^{1/2}$ as the correlation spectrum of the modes.

What does the horizontal transform look like (acts on $\vec{\psi}_{EOF}$)?

Transforming the state vector ($\vec{\psi}$):

$$\vec{\psi}_{v-space} = \mathbf{T}_h \vec{\psi}_{EOF}$$

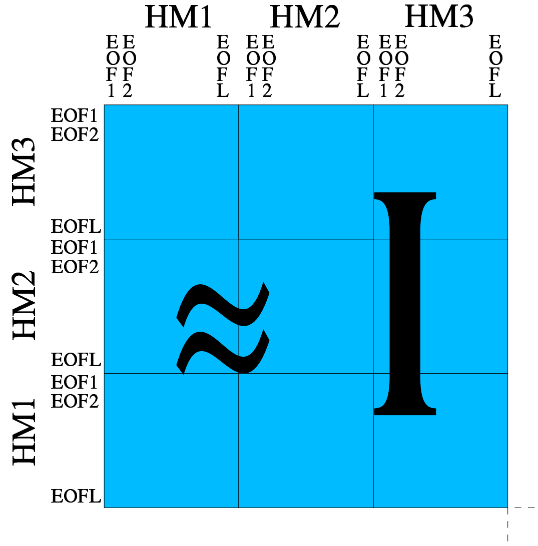
$\mathbf{T}_h^{hor}$ Transform associated with surface EOF1

$\mathbf{T}_h^{hor}$ Transform associated with surface EOF2

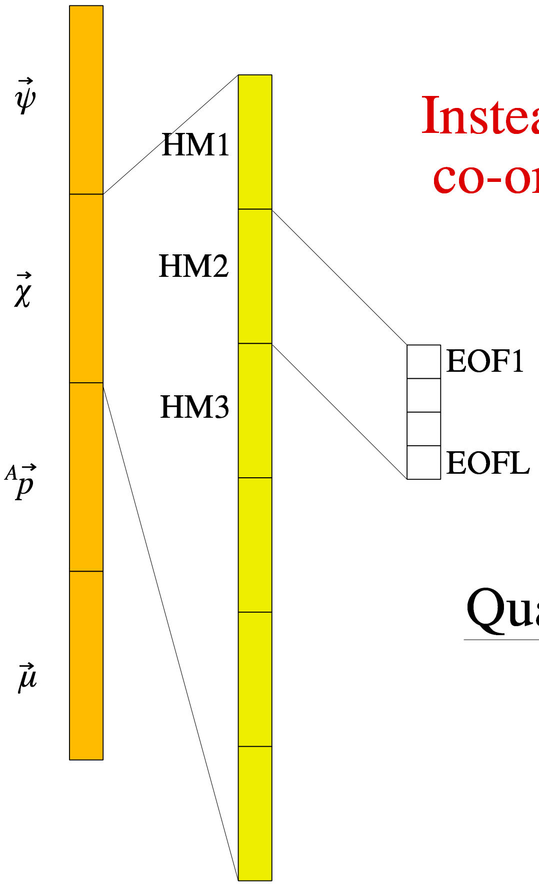
$\mathbf{T}_h^{hor}$ Transform associated with surface EOF3, etc

Transforming the error covariance matrix:

$$\mathbf{T}_h \mathbf{T}_v \mathbf{B}_\psi \mathbf{T}_v^T \mathbf{T}_h^T \approx \mathbf{I}$$



The transformed state vector: $\vec{v} = \mathbf{T}_h \vec{v}_v$



Instead of labelling the horizontal co-ordinates with 'long/lat', label with 'HM index'

Quantities are perturbations

This is the result of all three transforms

Perform descent algorithm in this space

Summary of Equations

1. $\vec{w}$-space formulation

$$J(\vec{w}') = \frac{1}{2}(\vec{w}'^b - \vec{w}')^T \mathbf{B}^{-1} (\vec{w}'^b - \vec{w}') +$$

$$\frac{1}{2}(\vec{y}^o - H(\vec{w}', \vec{w}^g))^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y}^o - H(\vec{w}', \vec{w}^g))$$

$$\frac{\mathrm{d}J}{\mathrm{d}\vec{w}'} = -\mathbf{B}^{-1}(\vec{w}'^b - \vec{w}') - \mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y}^o - H(\vec{w}', \vec{w}^g))$$

$$\frac{\mathrm{d}^2 J_b}{\mathrm{d}\vec{w}'^2} = \mathbf{B}^{-1} + \mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} \mathbf{H}$$

$$\vec{w} = \text{model state } (' \text{ pertbtn}, \ b \text{ backgrnd}, \ g \text{ guess})$$

$$H(\vec{w}', \vec{w}^g) \approx H(\vec{w}^g) + \mathbf{H}\vec{w}'$$

2. $\vec{v}$-space formulation

$$\vec{v} = \mathbf{T}\vec{w}' \qquad \vec{w}' = \mathbf{U}\vec{v} \qquad \mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} = \mathbf{I}$$

$$J(\vec{v}) = \frac{1}{2}(\vec{v}^b - \vec{v})^T \mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} (\vec{v}^b - \vec{v}) +$$

$$\frac{1}{2}(\vec{y}^o - H(\mathbf{U}\vec{v}, \vec{w}^g))^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y}^o - H(\mathbf{U}\vec{v}, \vec{w}^g))$$

$$\frac{\mathrm{d}J}{\mathrm{d}\vec{v}} = -\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U}(\vec{v}^b - \vec{v}) - \mathbf{U}^T\mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y}^o - H(\mathbf{U}\vec{v}, \vec{w}^g))$$

$$\frac{\mathrm{d}^2 J_b}{\mathrm{d}\vec{v}^2} = \mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} + \mathbf{U}^T\mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} \mathbf{H}\mathbf{U}$$