On the Estimation of Posterior Error Bars

Martin Losch[1], Alexey Yaremchuk[1,2], Maxim Yaremchuk[3], Jens Schröter[1], Bernadette Sloyan[1]

[1]Alfred-Wegener-Institute for Polar- and Marine Resarch, Bremerhaven, Germany
[2]Andreyev Acoustics Institute, Moscow, Russia
[3]International Pacific Research Center (IPRC), University of Hawaii, USA

(email: mlosch@awi-bremerhaven.de)

Introduction

Posterior error estimates are important parts of an inverse method solution. Unfortunately, their evaluation can become computationally quite demanding since this task always involves some sort of inverse of the Hessian of the problem. However, some observables for which the error estimates are required can be represented by a relatively small part of the Hessian spectrum and thus do not need its full inverse for estimating their errors.
We propose a polynomial approximation of the Hessian spectrum which enables us to obtain reliable error estimates for some observables. We compare the results to “standard” methods and show examples of applications for large problems.

Methods

1. Full inversion

The inverse of the Hessian of the cost function is the covariance of the deviation of the control variables φ from the optimal state φ^*. An observable Φ_α, α=1,2,... can be expanded into powers of $(\varphi - \varphi^*)$:

$$\Phi_\alpha = \Phi_\alpha(\varphi^*) + \phi_\alpha^T(\varphi - \varphi^*) + \ldots$$

One finds that near the optimum state the covariance of the observables can be expressed as $\phi_\alpha{}^T H^{-1}\phi_\beta$.

2. Conjugate gradients

When the dimension D of the problem is very large, it is still possible to calculate the vector $H\psi$ for any ψ. In order to find $\phi_\alpha{}^T H^{-1}\phi_\beta$, the equation $HX_\beta = \phi_\beta$ is solved for X_β with a conjugate gradients method to obtain the covariance $\phi_\alpha{}^T X_\beta$.

3. Polynomial Approximation

The covariance $\phi_\alpha{}^T H^{-1}\phi_\beta$ is written in a more symmetric form $\psi_\alpha{}^T\psi_\beta$ with $\psi = H^{-1/2}\phi$. The latter can be expressed as a series of Chebyshev's Polynomials $U_n(x)$:

$$H^{-\frac{1}{2}}\phi = \frac{16}{\pi}\sum_{n=0}^{\infty}\left(\frac{n+1}{4(n+1)^2-1}u_n\right)$$

with $u_n = U_n(1-2H)\,\phi$. Truncuation of the series corresponds to smoothing the spectrum of the Hessian H. Let $\varepsilon_k = \sin^2(\theta_k/2)$ be the eigenvalues of H and e_k the eigenvectors, then the spectral function

$$\rho_{\alpha\beta} = \sum_{k=1}^{D}\delta(\theta-\theta_k)(\phi_\alpha^T e_k)(\phi_\beta^T e_k)$$

$$= \sin\theta\frac{2}{\pi}\sum_{k=1}^{\infty}\sin(n\theta)(\phi_\alpha^T u_{n-1}^{(\beta)})$$

is the spectral function of the Hessian $\delta(\theta-\theta_k)$ weighted by $(\phi_\alpha{}^T e_k)(\phi_\beta{}^T e_k)$. Truncuation results again in smoothing the spectrum. From this spectrum we can decide what part of the spectrum of H is needed for the correct evaluation of $\phi_\alpha{}^T H^{-1}\phi_\beta$ and choose the inverting window accordingly.

4. Comparision

For the Southern Ocean WOCE section SR3 a small geostrophic toy model with 136 control variables (tracers and bottom velocities) illustrates the three methods described. Figures 1 and 2 show the temperature and velocity field after inversion. the posterior errors and their correlation of the total mass and heat transports through the section obtained by the three different methods show a reasonable agreement (Table 1)

Table 1:
Comparision of errors obtained by the three different methods

Method	# of iterations	Mass Transport Error [Sv]	Heat Transport Error [PW]	Correlation of Errors
Full inversion		32	1.6	0.27
conjugate gradients	520	32	1.7	0.17
polynomial approximation	500	32	1.4	0.28

Figure 1: Temperature of toy model for SR3. Note the coarse grid.

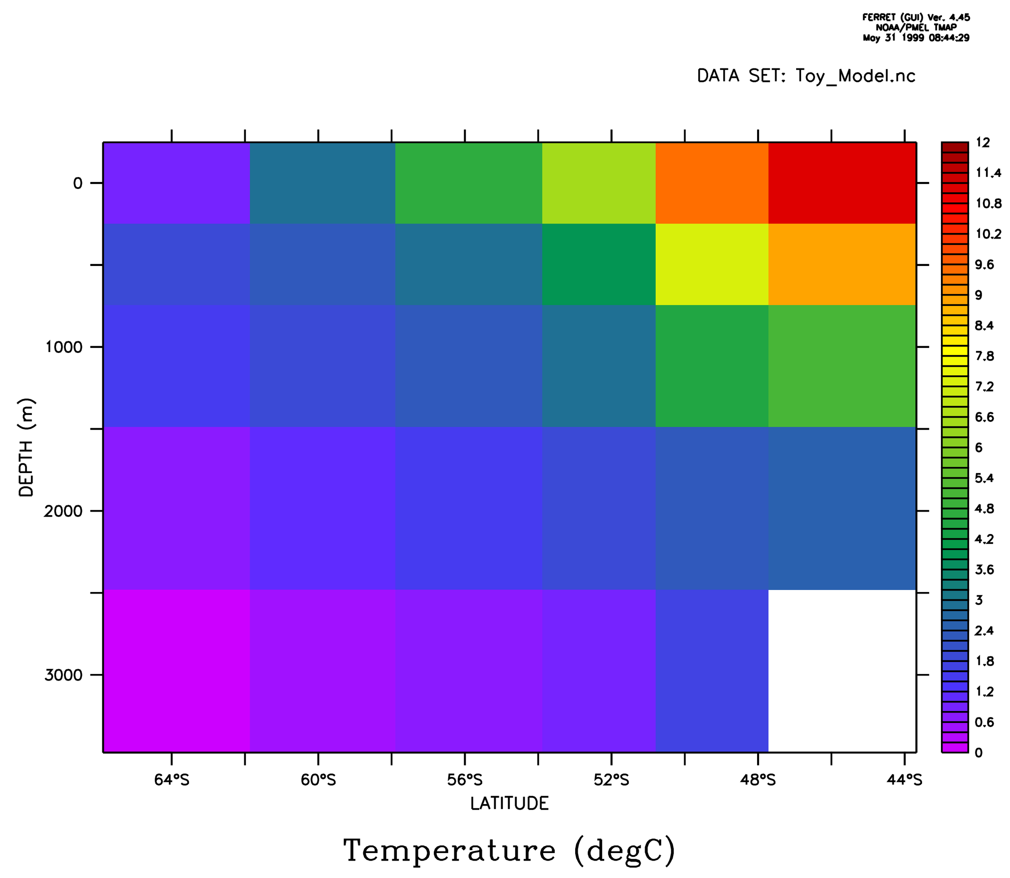


Figure 2: Velocity normal to section (cm/s) of the toy model of SR3.

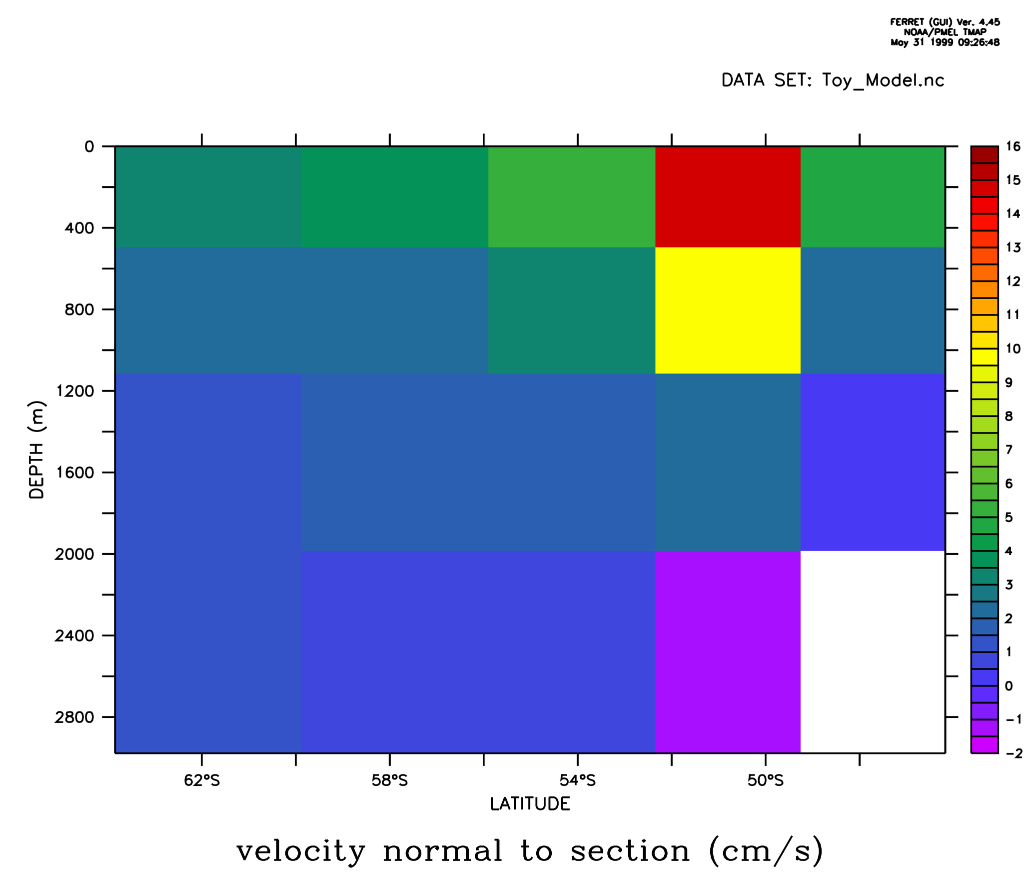
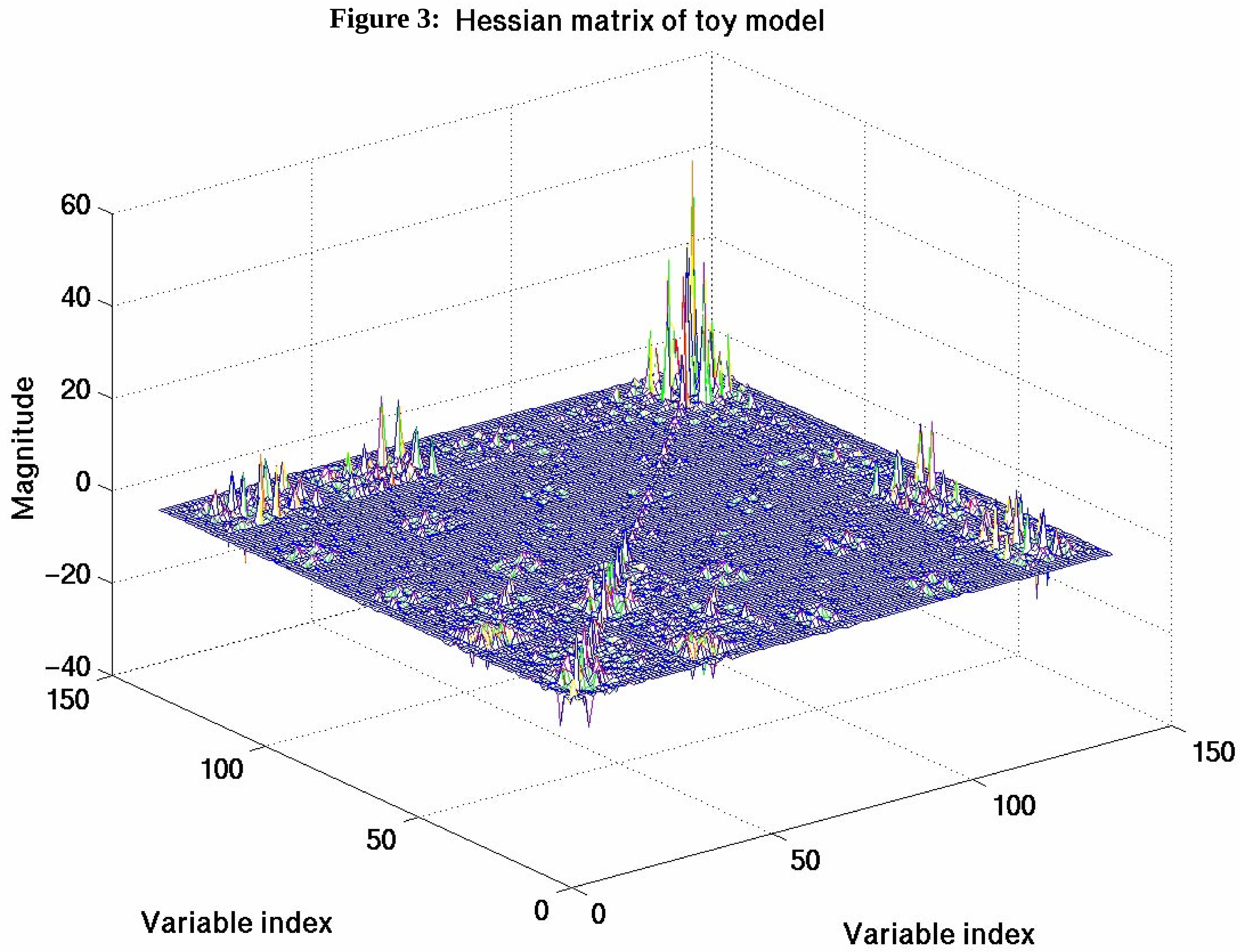


Figure 3: Hessian matrix of toy model



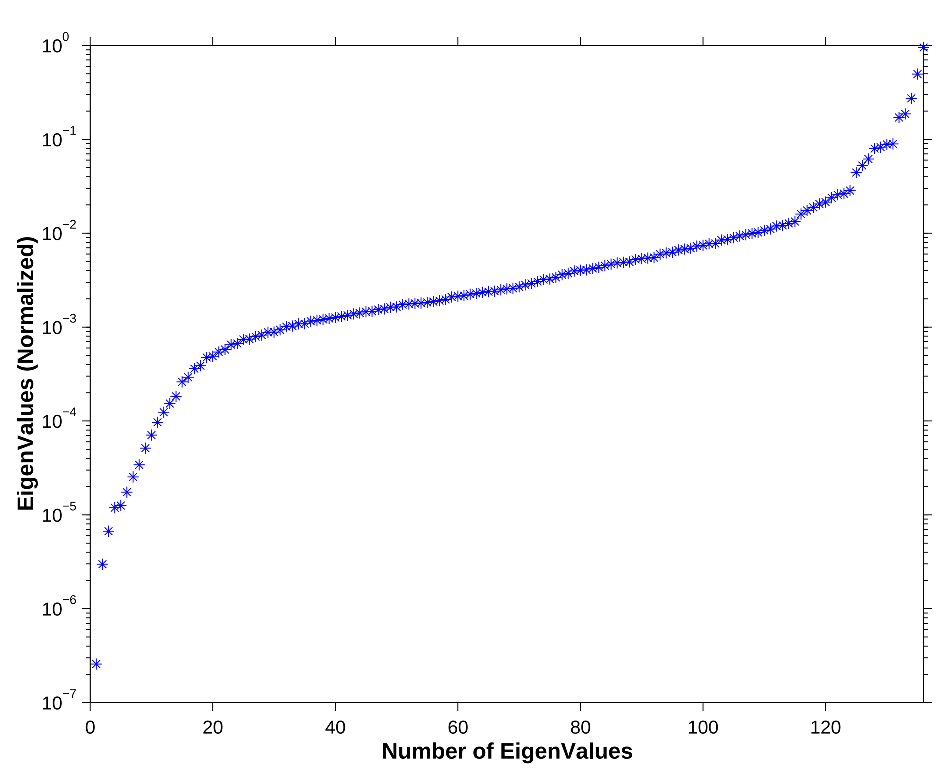


Figure 4: Eigenvaluespectrum of the Hessian matrix of the toy model. The condition number is 3×10^{-6}

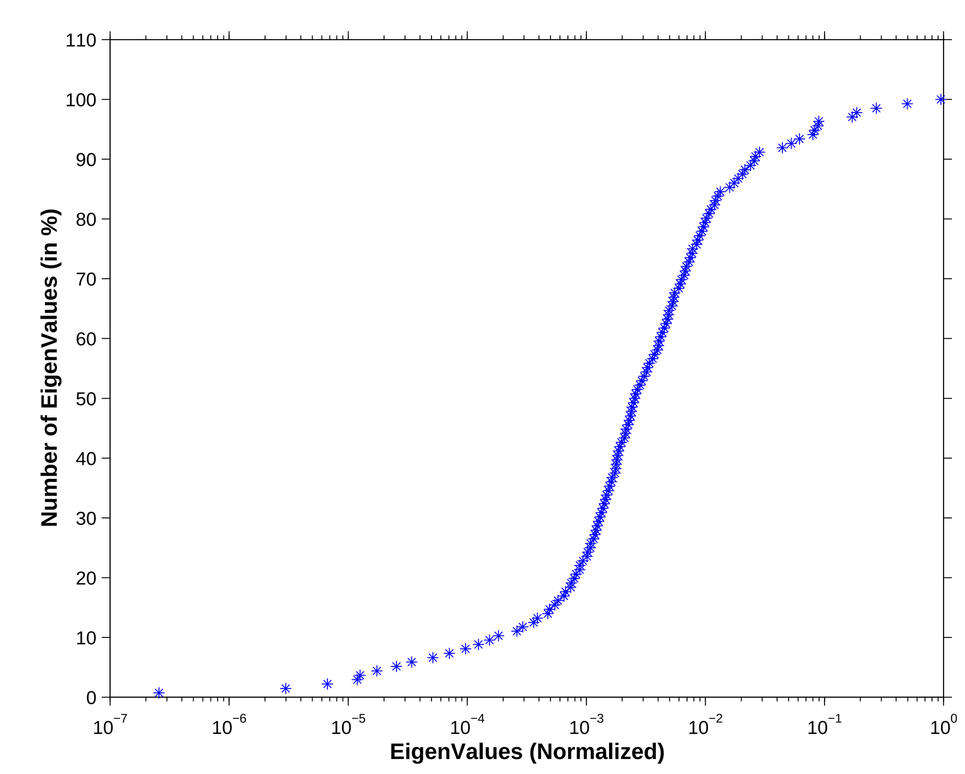


Figure 5: Same eigenvaluespectrum of the Hessian matrix of the toy model for SR3 as in Figure 4. The representation has been changed to match that of figures 6-8 (lower panels). On the abscissa, there are the normalized magnitude of the eigenvalues, on the ordinate the index number of the eigenvalues relative to the dimension of the problem (D = 136)

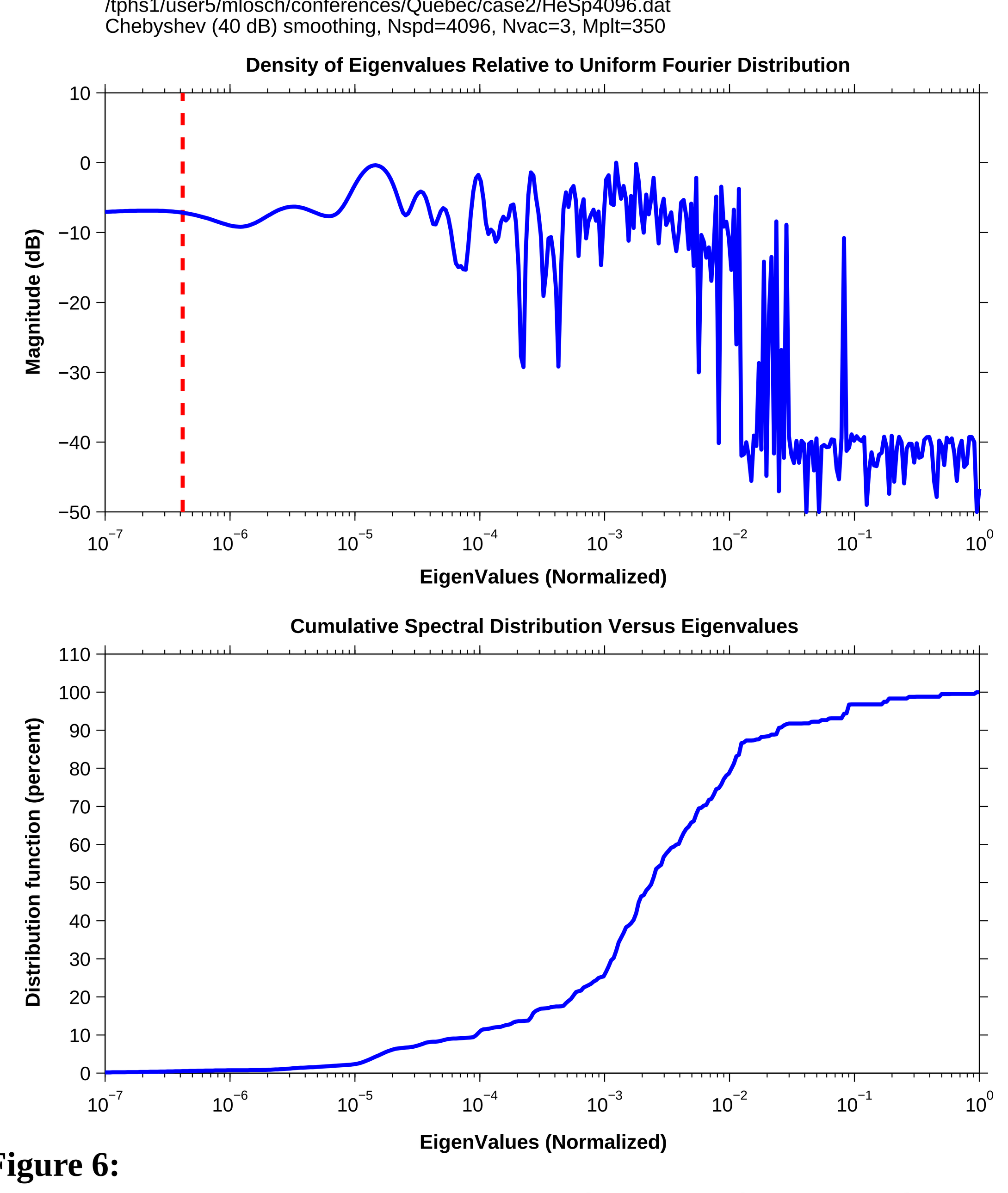


Figure 6:

- Upper panel: smoothed spectral density of Hessian of toy model for SR3. Number of coefficients $u_n(x)$ = 4096.
- Lower panel: cumulative distribution of the same Hessian spectrum. Going from right to left you can see how many eigenvalues (in %) are resolved. Some of the peaks in the upper part of the spectrum are so sharp that they cannot be plotted, but their contribution shows in the cumlative distribution.

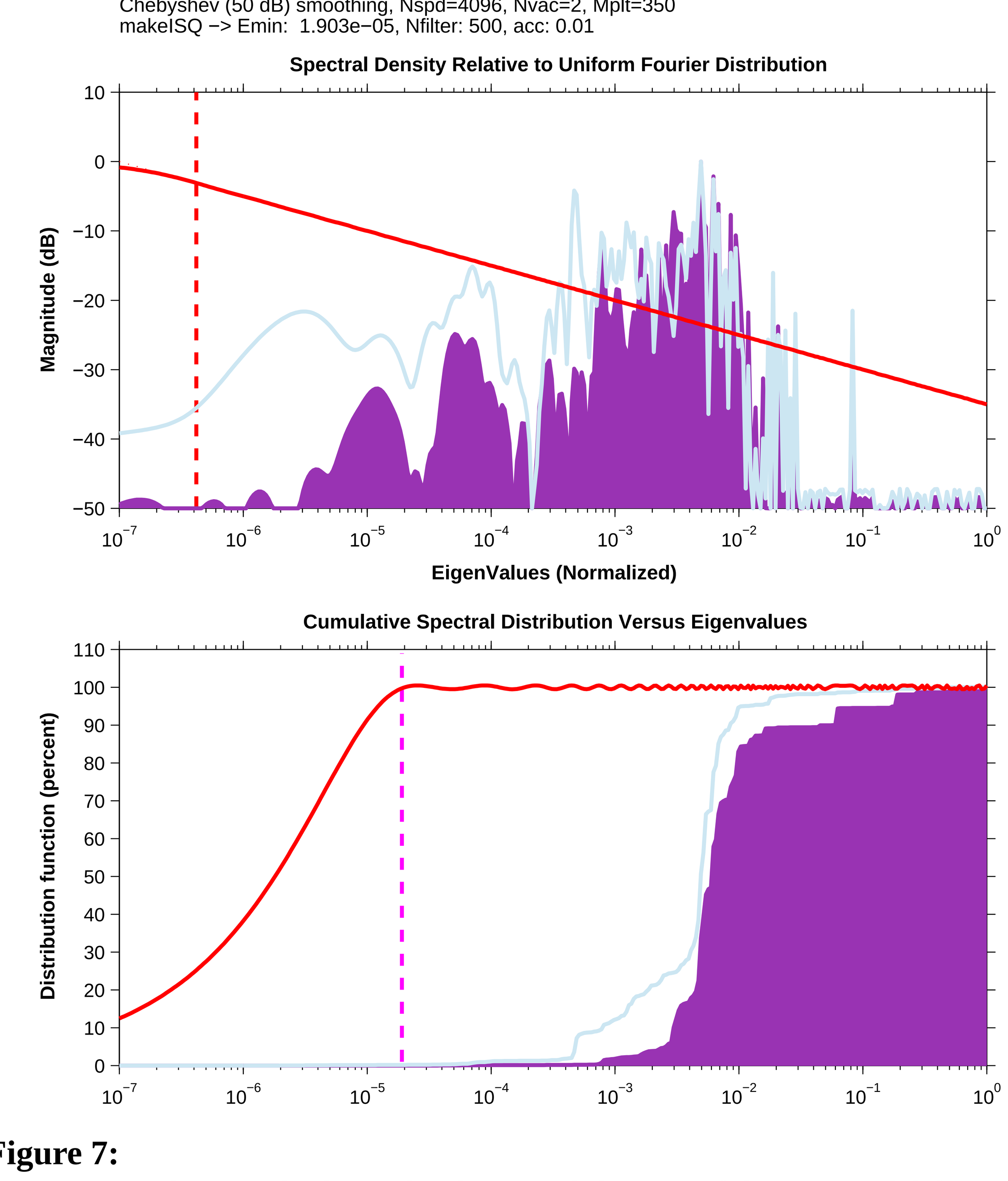


Figure 7:

- Upper panel: smoothed weighted spectral density of the toy model for mass (cyan line) and heat transport (purple filled contour). Number of coefficients $u_n(x)$ = 4096.
- Lower panel: cumulative distribution of the weighted spectrum. Since the distribution function drops to zero at about $2x10^{-5}$ of the maximum eigenvalue for both properties (cyan line: volume transport, purple filled contour: heat transport), the inverting window (red line) can be chosen such that only 500 coefficients $u_n(x)$ are needed for the inversion, thus reducing the computational cost considerably. As in figure 6 some eigenvalues in the upper part of the spectrum cannot be seen in the upper panel but they show in the cumulative distribution.

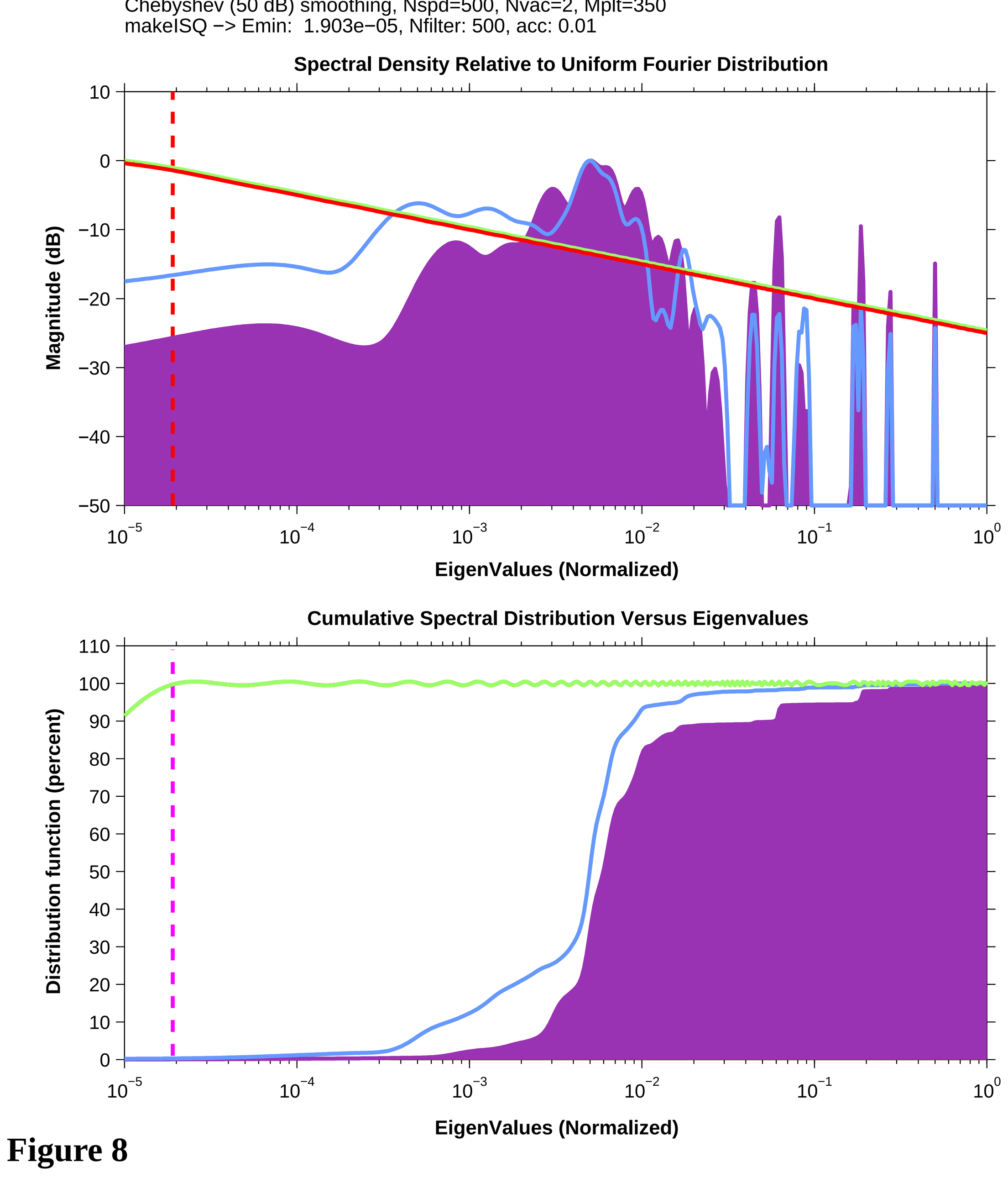


Figure 8

Same weighted spectral density (upper panel) and cumulative spectral distribution (lower panel) for mass (cyan line) and heat transport (purple filled contour) as in figure 7, but with a smaller number of coefficients $u_n(x)$ = 500. The smaller number of coefficients leads to a higher degree of smoothing of the spectral density.

Examples of Applications

1. Box model of the Southern Ocean

An inverse box model of the Southern Ocean estimates heat and volume fluxes per layer through hydrographic sections. The control space has 1600 dimensions. This makes the full inversion of the Hessian possible. As an example of the results, the mass flux and its errors for the zonal Indian Ocean section along 32°S (Ind32) are shown in figure 9. Table 2 lists the correlation of the flux errors of the different water masses

Table 2: Correlation of mass flux errors for Indian 32°S (SW = Surface/Thermocline Water, IW = Intermediate Water, U/LDW = Upper/Lower Deep Water, BW = Bottom Water)

	SW	IW	UDW	LDW	BW
SW	1	0.07	0.06	-0.05	-0.21
IW	0.07	1	-0.51	-0.09	0.05
UDW	0.06	-0.51	1	-0.47	-0.07
LDW	-0.05	-0.09	-0.47	1	-0.59
BW	-0.21	0.05	-0.07	-0.59	1

2. Nonlinear section inverse of SR3

A nonlinear geostrophic inverse model similar to the one used as the toy model (see above) is applied to the meridional Southern Ocean WOCE section SR3 between Tasmania and the Antartic continent, to estimate mass and property transports. As the number of control variables is quite large (~17000), the Hessian cannot be calculated nor inverted explicitly. The errors in Table 3 are estimated using the conjugate gradients approach, where for each variance estimate several thousand iterations were necessary for the algorithm to converge.

Table 3: Transport estimates through SR3

Mass [10^6 m^3/s]	Heat [PW]	Salt [10^8 m^3/s psu]	Oxygen [Mmol/s]	Silicate [10^2 kmol/s]
146 +/- 29	1.72 +/- 0.33	50 +/- 10	34 +/- 7	82+/-25

3. Nonlinear section inverse -- a twin experiment

The same model as in the previous section but augmented with satellite altimetry data is used to invert an artifical hydrographic section data set produced by a general circulation model (1/3° North Atlantic Model of the FLAME group). The zonal section is located in the North Atlantic along 24.5°N. Again the number of control variables is fairly large (~12000) and the Hessian cannot be computed completely. Here for the variance estimation of the mass and heat transports, the polynomial approximation approach is employed. Figure 10 and 11 show the spectral density of the Hessian of the model and the weighted spectral density for mass and heat transports through one small part of the section (Florida Strait). Note that the number of coefficients used to represent the Hessian spectrum is not bigger than for the toy model, although the problem is much larger, neither does the number of coefficients needed for the inversion depend directly on the dimension of the problem. The transports and errors obtained by this method are for mass 28.2 +/- 8.2 Sv and for heat 2.10 +/- 0.57. The correlation is 0.87.

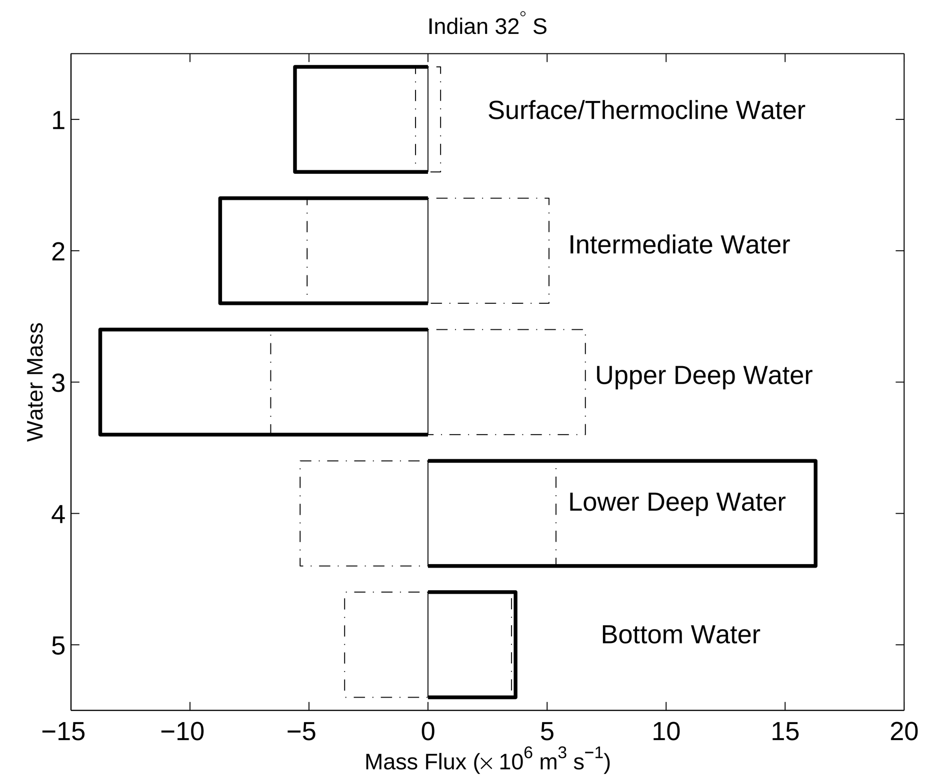


Figure 9:
Mass Transports (solid) and posterior errors (dash-dotted), divided into water masses for a zonal section in the Indian Ocean at 32°S.

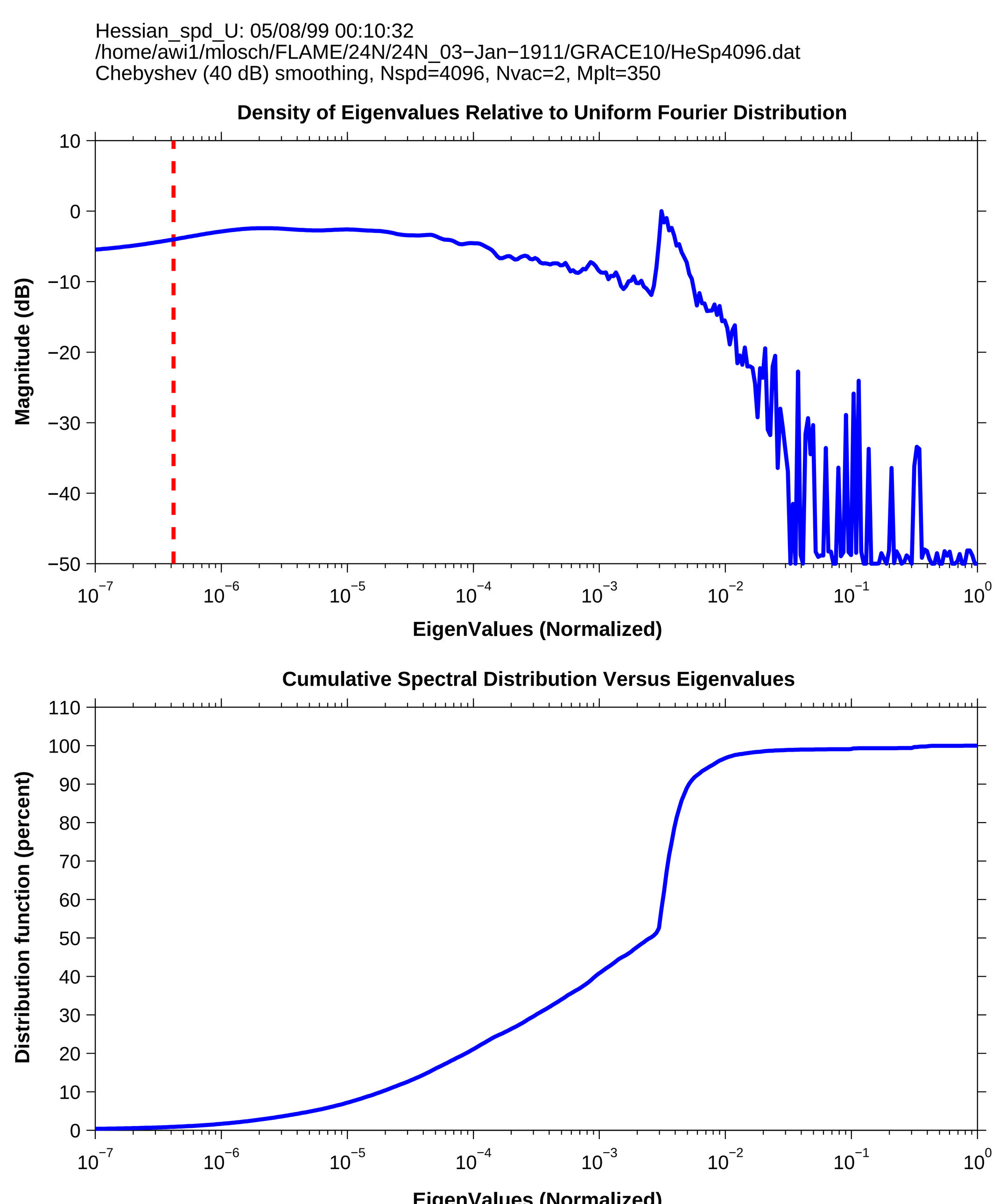


Figure 10:
Smoothed spectral density (upper panel) and cumulative spectral distribution (lower panel) of the Hessian of the 24°N problem.

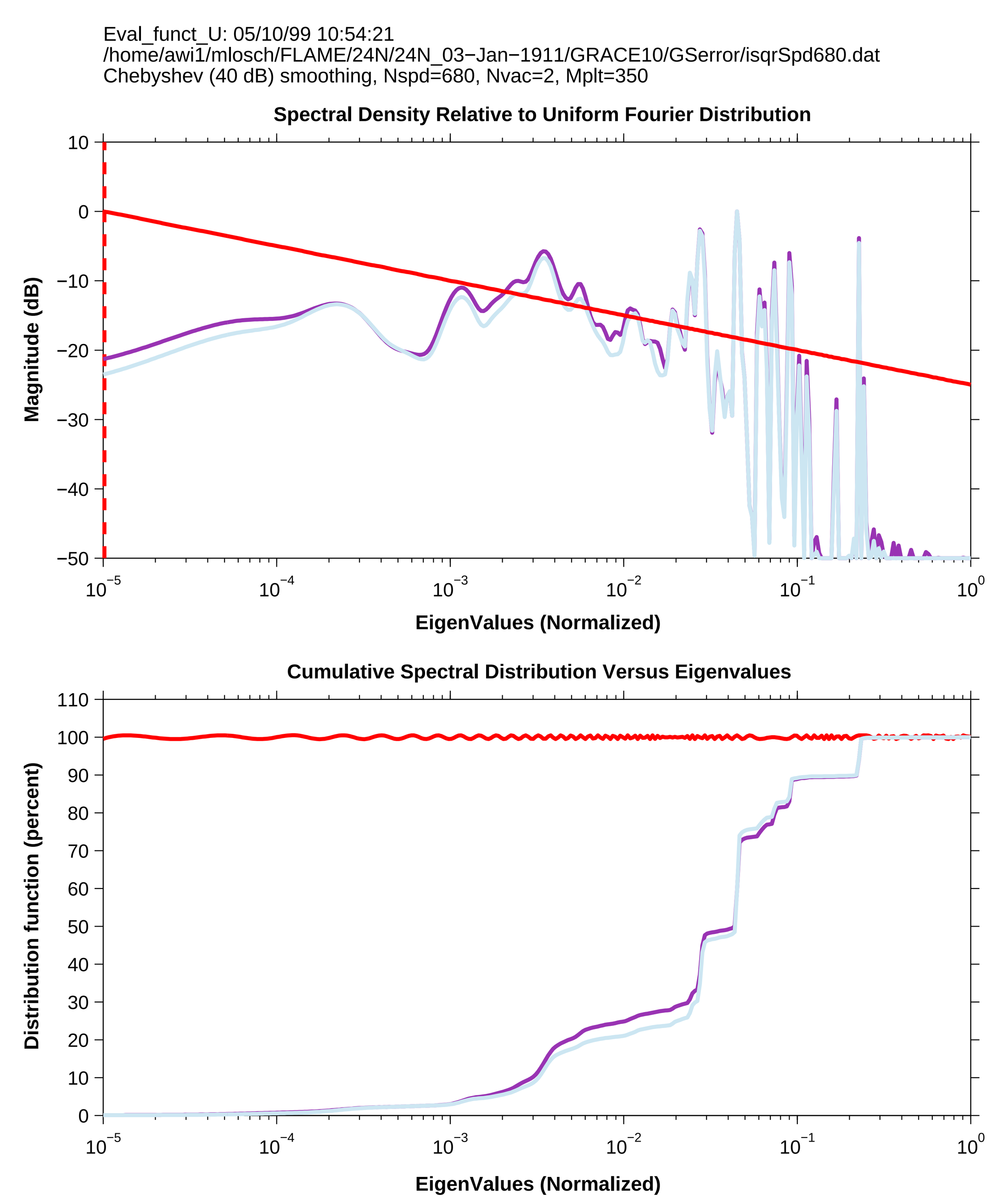


Figure 11:
Part of the smoothed spectral density weighted for mass (cyan line) and heat transport (purple line) through Florida Strait and its cumulative distribution. Although the dimension of the problem is approximately 12000, only 680 coefficients $u_n(x)$ are needed for the inversion to calculate the posterior errors of the mass and heat transport through Florida Strait.

Conclusions

Three methods are presented for the calculation of posterior errors of inverse solutions. Of these methods, the full inversion of the Hessian is the most accurate but it is only applicable to small and moderate size problems. Of the other two methods the conjugate gradients approach leads to reliable results, but is computationally more expensive than finding the solution to the inverse problem itself. The polynomial approximation on the other hand yields also reliable estimates of the errors, but it involves an approximation of the Hessian spectrum by Chebyshev's Polynomial. This drawback is at the same time a great advantage over the conjugate gradient method because it decreases computing time by an order of magnitude.

References

- Yaremchuk, A. I. and Schröter, J., 1998: Spectral Analysis of Symmetric Operators: Application to the Laplace Tidal Model. *Journal of Computational Physics*, **147**,1-27.
- Sloyan, B. M., 1997: The circulation of the Southern Ocean and the adjacent ocean basins determined by inverse methods. *Ph.D. thesis*. Institute of Antartic and Southern Ocean Studies, University of Tasmania.
- Yaremchuk, M. et. al., 1999: On the zonal and meridional circulation an ocean transports between Tasmania and Antartica, submitted.
- Schröter, J., Sloyan,B. M. and Losch, M., 1999: Impact of the GOCE mission for ocean circulation, Midterm report, Task 4.