CERFACS

EUROPEAN CENTRE FOR RESEARCH AND ADVANCED TRAINING IN SCIENTIFIC COMPUTING

Accounting for correlated observation error in variational ocean data assimilation

Olivier Goux[1], Anthony Weaver[1], Selime Gürol[1], Oliver Guillet[2], Youssef Diouane[3]

1: CERFACS, Toulouse

2: Météo-France, Toulouse

3: École Polytechnique, Montréal

DA-TT Meeting, 9-11 May 2023, CNR Rome

* Work supported by the Copernicus Climate Change Service

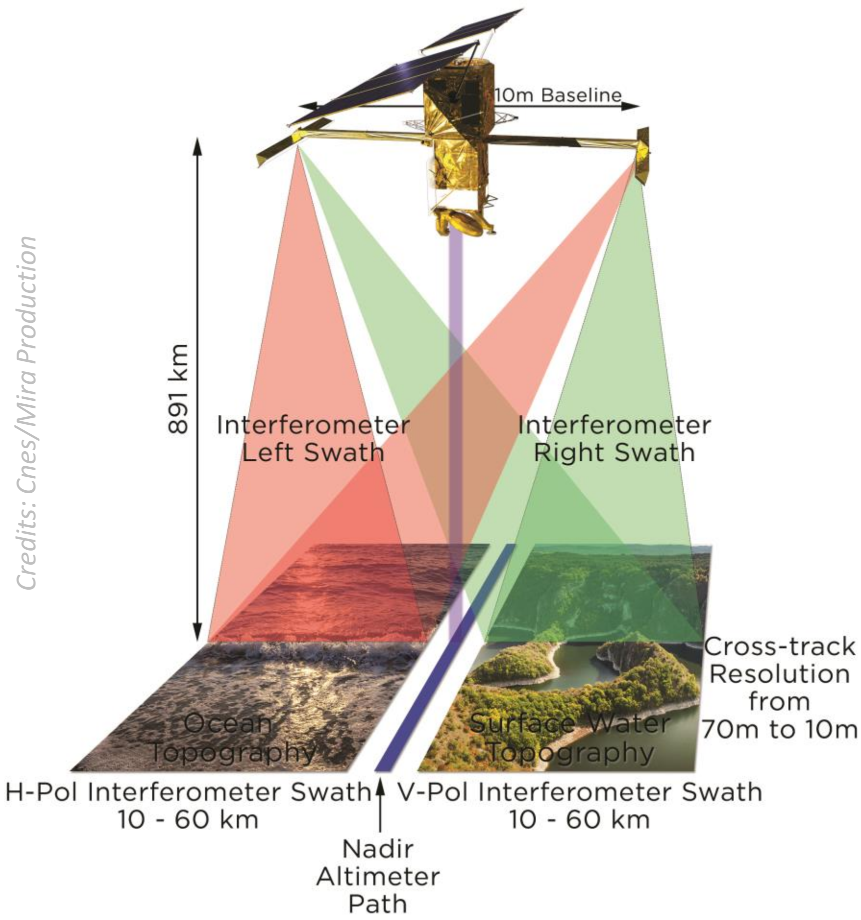
Climate Change Service

climate.copernicus.eu

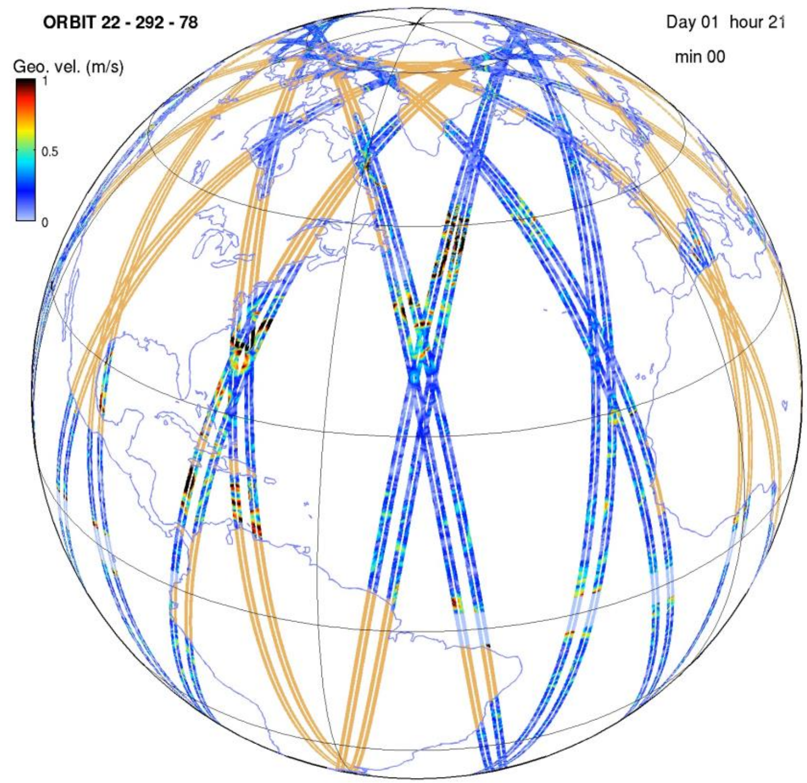
ECMWF

www.cerfacs.fr

Example: the SWOT mission



Credits: Cnes/Mira Production



Credits: JPL/C.Ubelmann

Example: the SWOT mission

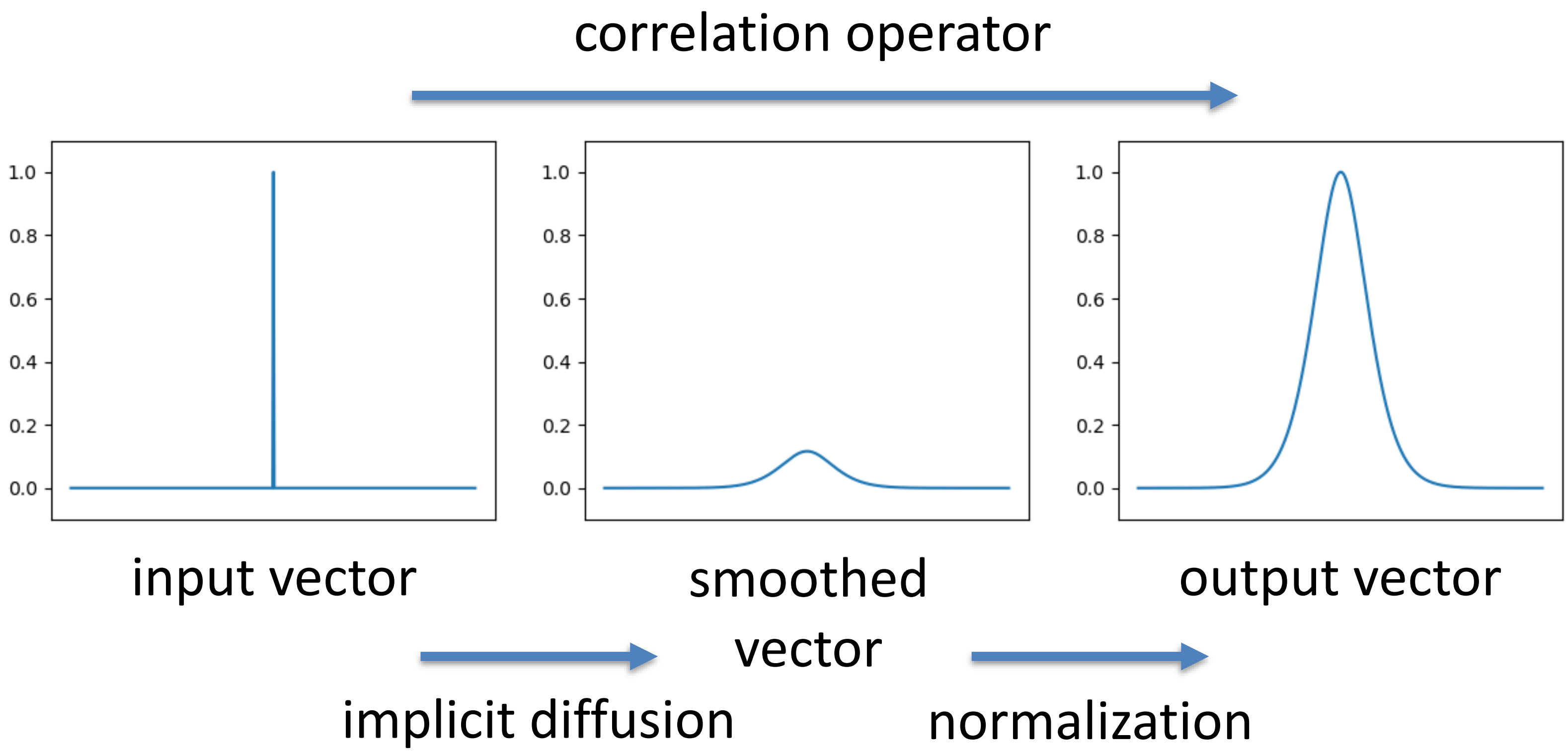
King, R. R. and Martin, M. J. (2021): Assimilating realistically simulated wide-swath altimeter observations in a high-resolution shelf-seas forecasting system, *Ocean Sci.*, 17, 1791–1813, https://doi.org/10.5194/os-17-1791-2021, 2021.

	SSH (m)		**Temperature (K)**		**Salinity (PSU)**		**Surface Current Speed**	
No SWOT (Control)	0.034		0.324		0.053		0.104	
white noise only	0.027	(-21%)	0.340	(+5%)	0.051	(-4%)	0.091	(-13%)
Full error	0.041	(+21%)	0.391	(+21%)	0.054	(+2%)	0.114	(+13%)
Full error, Half swath, Superobbed	0.032	(-6%)	0.324	(0%)	0.052	(-2%)	0.104	(0%)

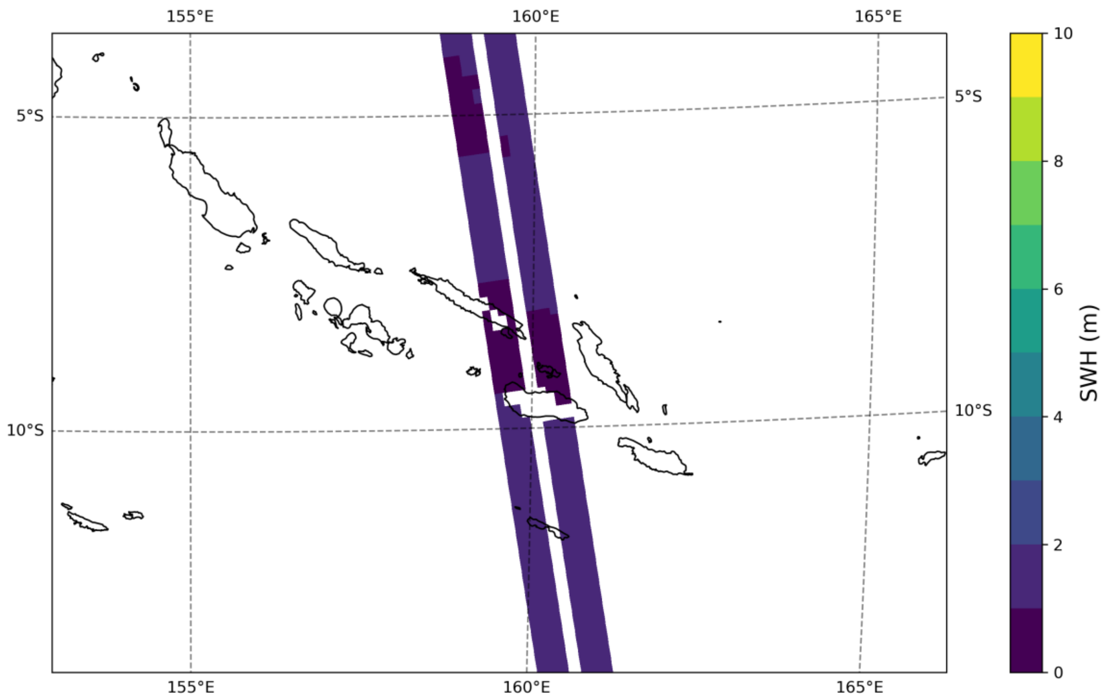
« when correlated errors are included in the full swath SWOT observations, there is a degradation in the sub-surface temperature and salinity, and the SSH and surface currents are degraded with a clear increase in the mean surface currents. While restricting the SWOT data to the inner half of the swath and applying observation averaging with a 5 km radius negated most of the negative impacts, it also severely limited the positive impacts. **»**

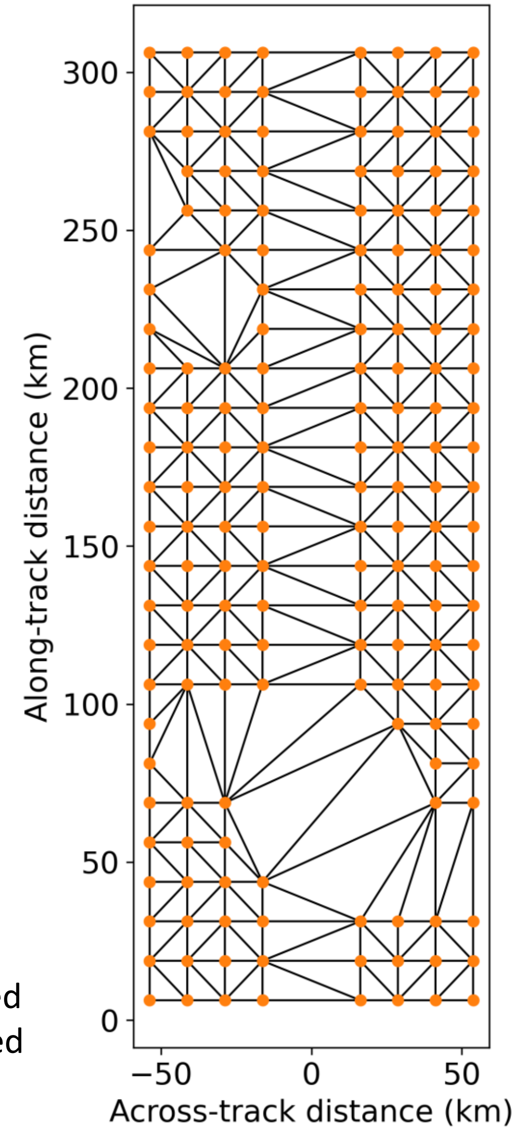
Diffusion operators

- Diffusion operators are already in use in DA to model correlation operators for background error
- With an implicit scheme, their inverse is easily accessible which makes them suitable to model inverse correlation operators for observation error
- The cost of a product with $\boldsymbol{R}^{-1}$ would be much lower than the cost of a product with $\boldsymbol{B}$



Diffusion operators can be discretized on a mesh with a finite element method, which makes them suitable with unstructured observations

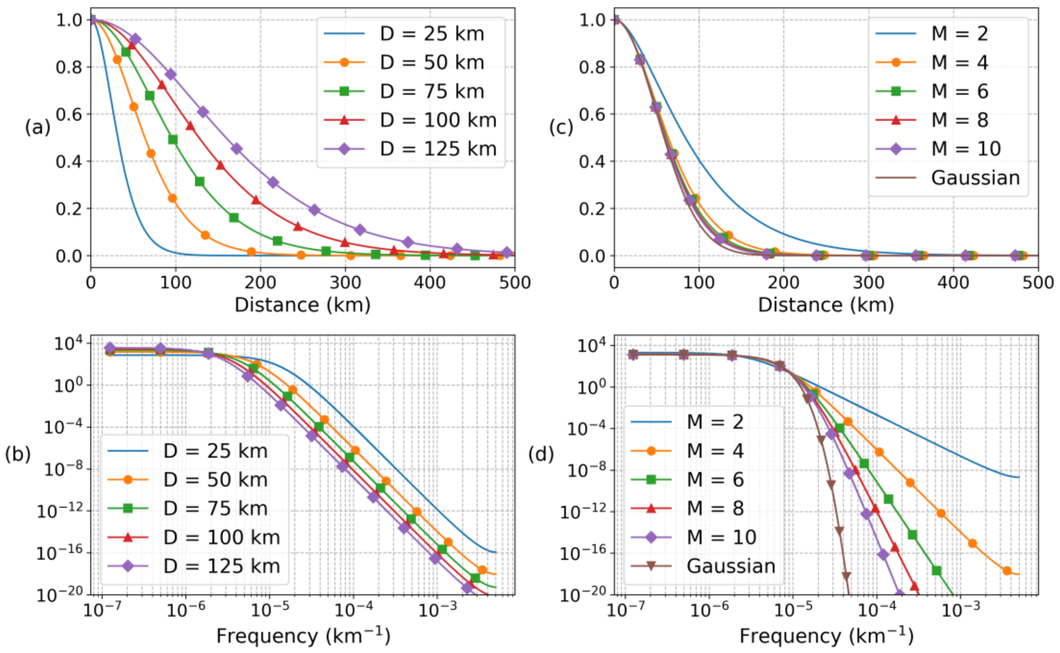




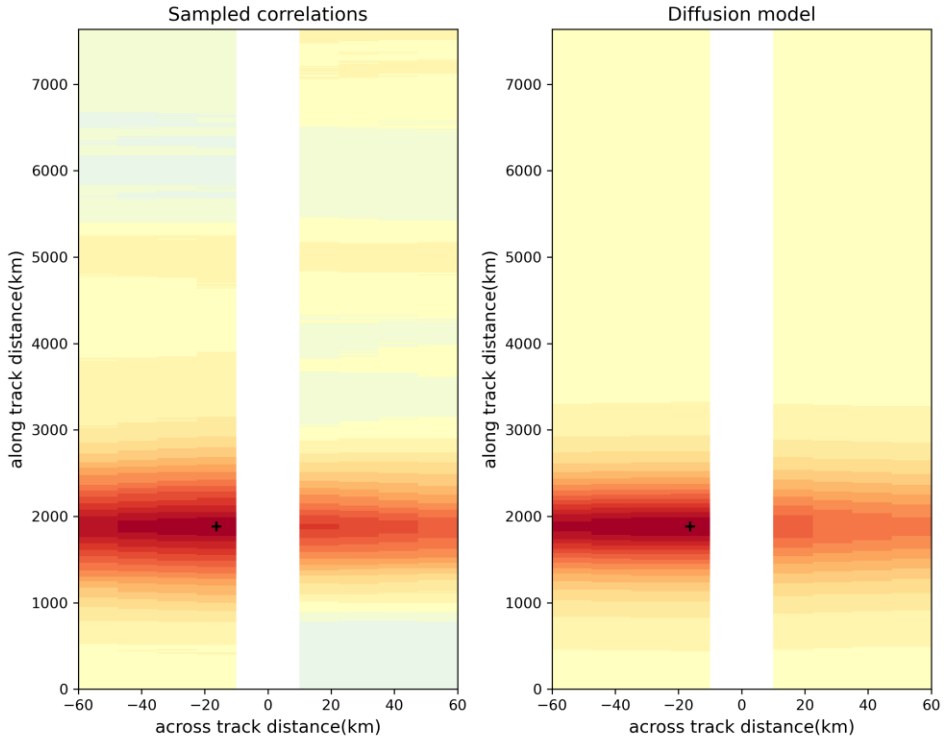
Guillet, O, Weaver, AT, Vasseur, X, Michel, Y, Gratton, S, Gurol, S. Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. *Q J R Meteorol Soc* 2019; 145: 1947– 1967. https://doi.org/10.1002/qj.3537

Diffusion operators

Both the cut-off and roll-off of the error spectrum modelled by diffusion operators can be adjusted using a Daley length scale D, and a smoothness parameter M.



The spatially variable and anisotropic diffusion tensor makes the operator flexible enough to fit estimated observation error correlations



Understanding DA with a non-diagonal R

In variational DA, we need to converge fast and towards an accurate analysis.
The choice of parameters for $\boldsymbol{R}$ has a role to play for both properties. To gain insight on the behaviour of the DA algorithm when both $\boldsymbol{B}$ and $\boldsymbol{R}$ are non-diagonal, we study a very simple system

- 1D periodic domain
- Regular model and observation grids
- Spatially constant covariance parameters

→

- $\boldsymbol{B}$ and $\boldsymbol{R}$, and $\boldsymbol{HBH}^T$ are circulant matrices
- All three are diagonal in a Fourier basis
- Their eigenvalues represent error power spectrums

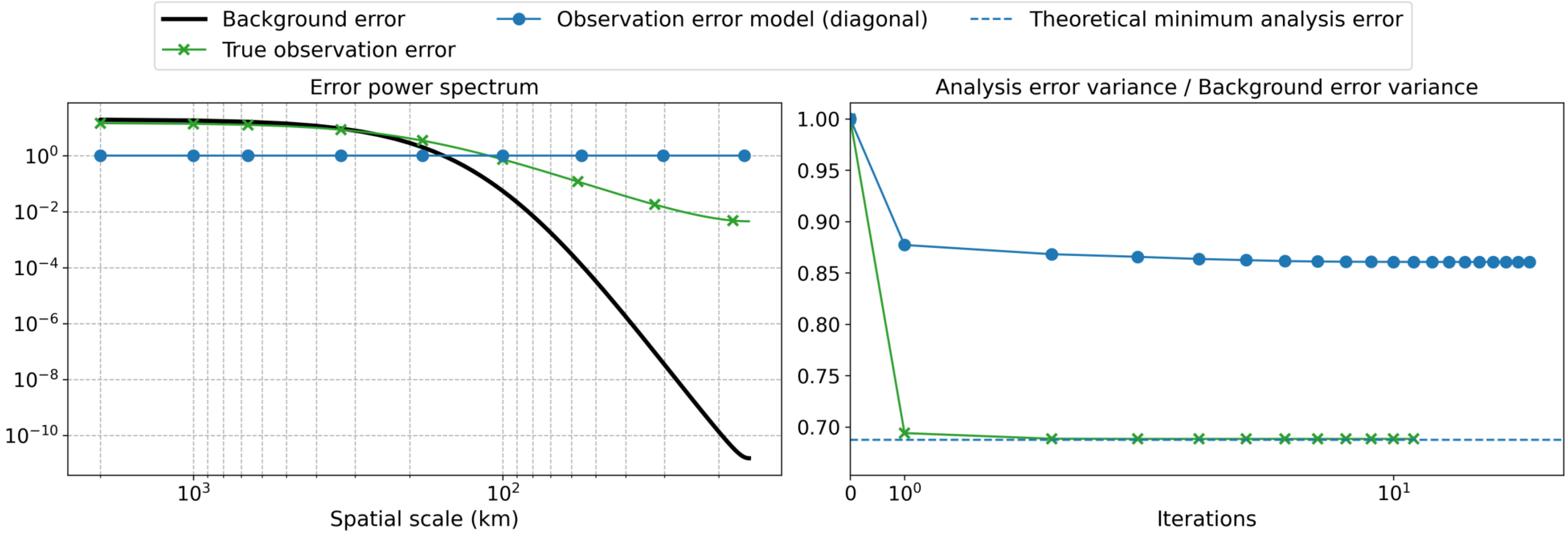
Indicator of the convergence rate: $$\text{Condition number} = \max_i \; 1 + \frac{\lambda_i(\boldsymbol{HBH}^{\mathrm{T}})}{\lambda_i(\boldsymbol{R})}$$

Goux O., Gürol S., Weaver A. T., Guillet O., Diouane Y. (2022). Impact of correlated observation errors on the convergence of the conjugate gradient algorithm in variational data assimilation 2212.02305, arXiv

Convergence experiment

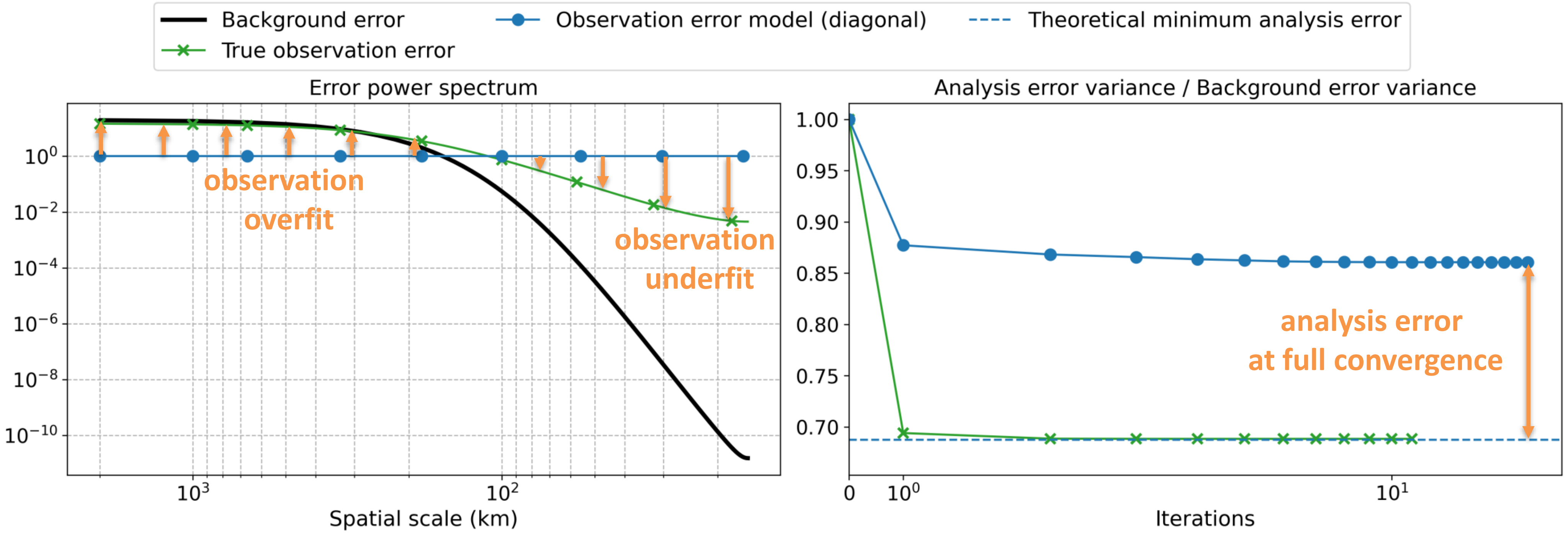
- Ensembles of background and observations are simulated with known error statistics (***B*** and ***R***)
- They are assimilated using ***B*** and $\widetilde{\boldsymbol{R}}$
- The analysis error variance is estimated at each iteration of the B-PCG from the ensemble of solutions

True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 30$ km ; $M_o = 2$



Convergence experiment

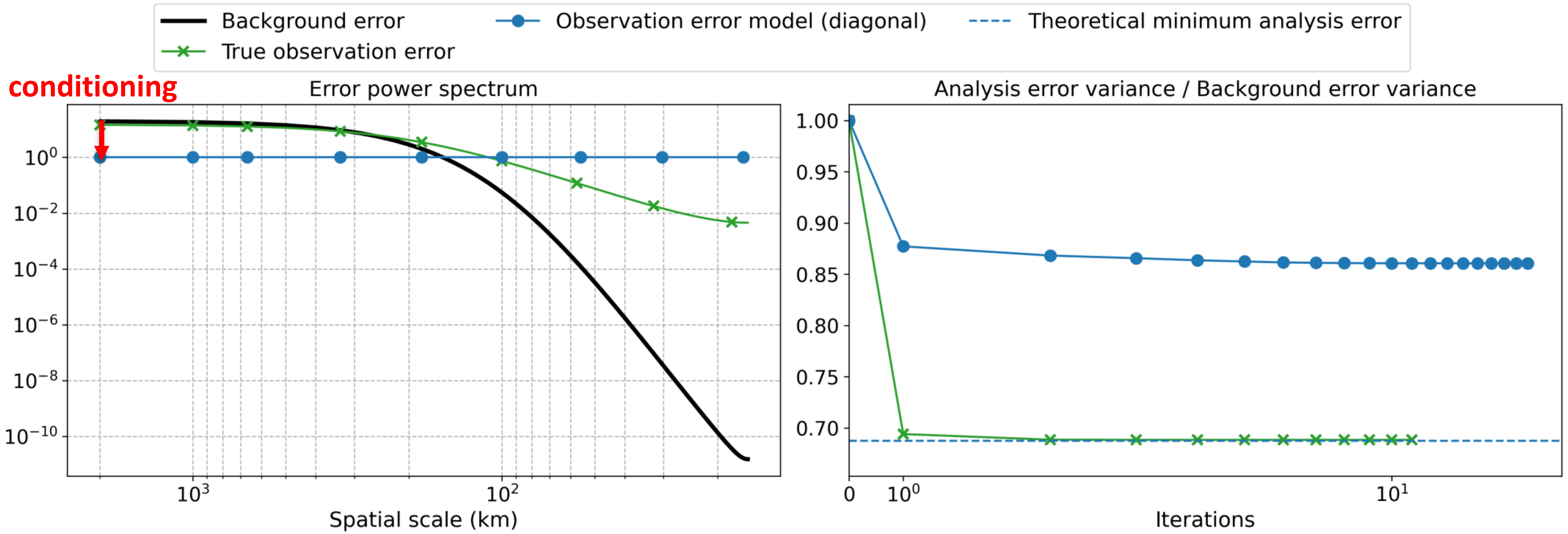
All differences between the true observation error and the observation error model contribute to a sub-optimal analysis error at full convergence.



Using a non-diagonal $\boldsymbol{R}$ induces an overfit of the observations at large spatial scales and an underfit at small spatial scales

Convergence experiment

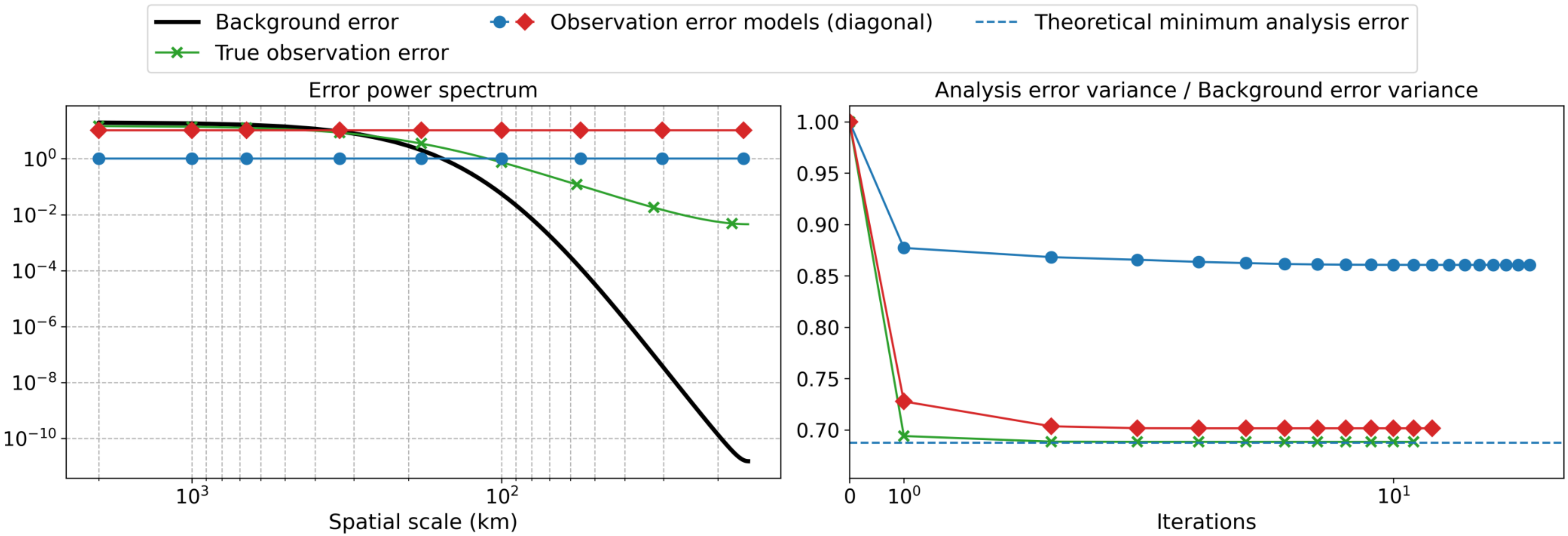
The largest ratio of the eigenvalues of ***B*** and ***R*** determines the condition number and strongly influences the convergence rate of CG



Using a non-diagonal ***R*** compared to a diagonal ***R*** does <u>not</u> necessarily degrade the conditioning or slow down the convergence

Convergence experiment

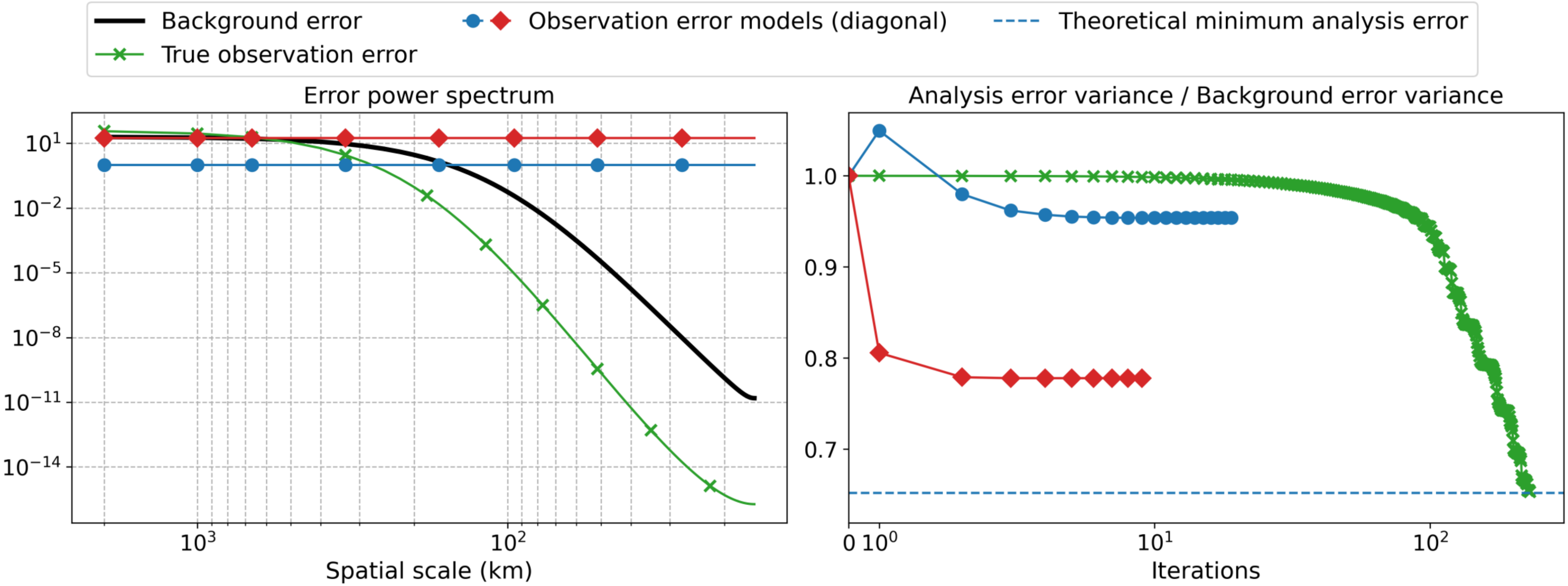
We use variance inflation with the inflation factor that minimizes the analysis error at full convergence to find the best possible diagonal observation error model.



Variance inflation prioritizes reducing the overfit at large spatial scales where errors are large at the expense of small spatial scales.

Convergence experiment

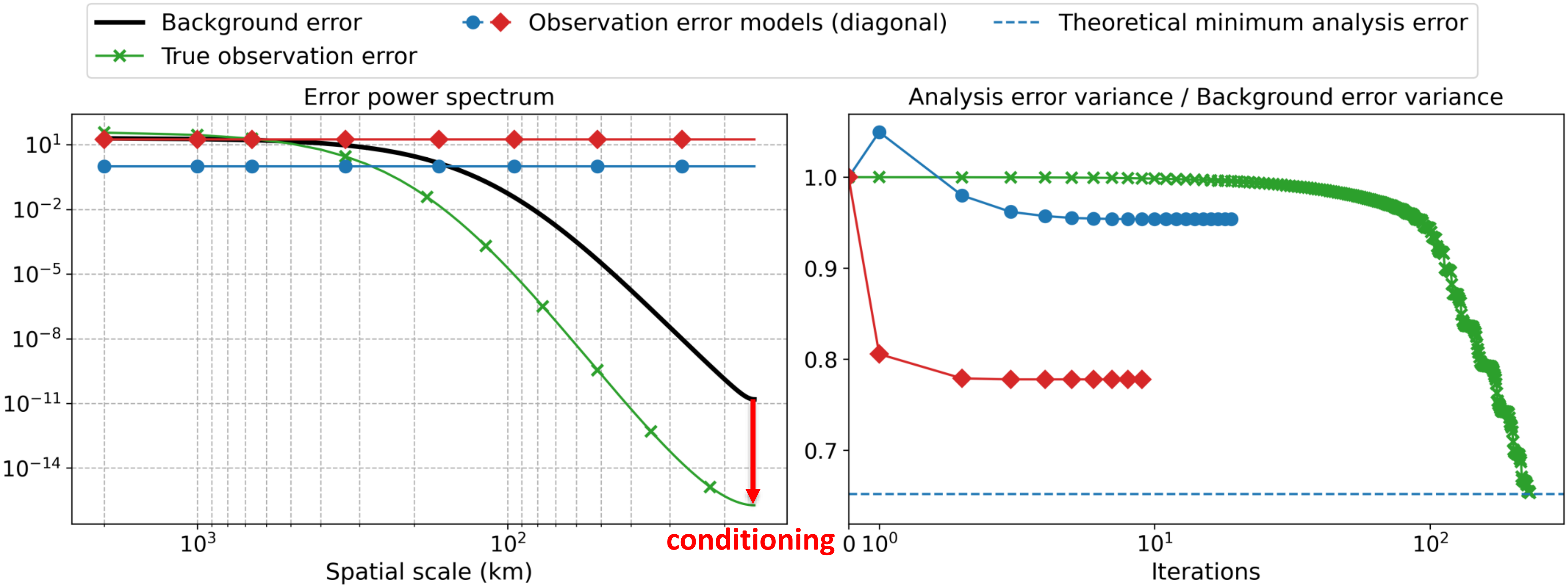
True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$



Using the most accurate parameters for the observation error leads to a very low analysis error but sometimes only after a prohibitive number of iterations (especially if $M_o > M_b$)

Convergence experiment

True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$

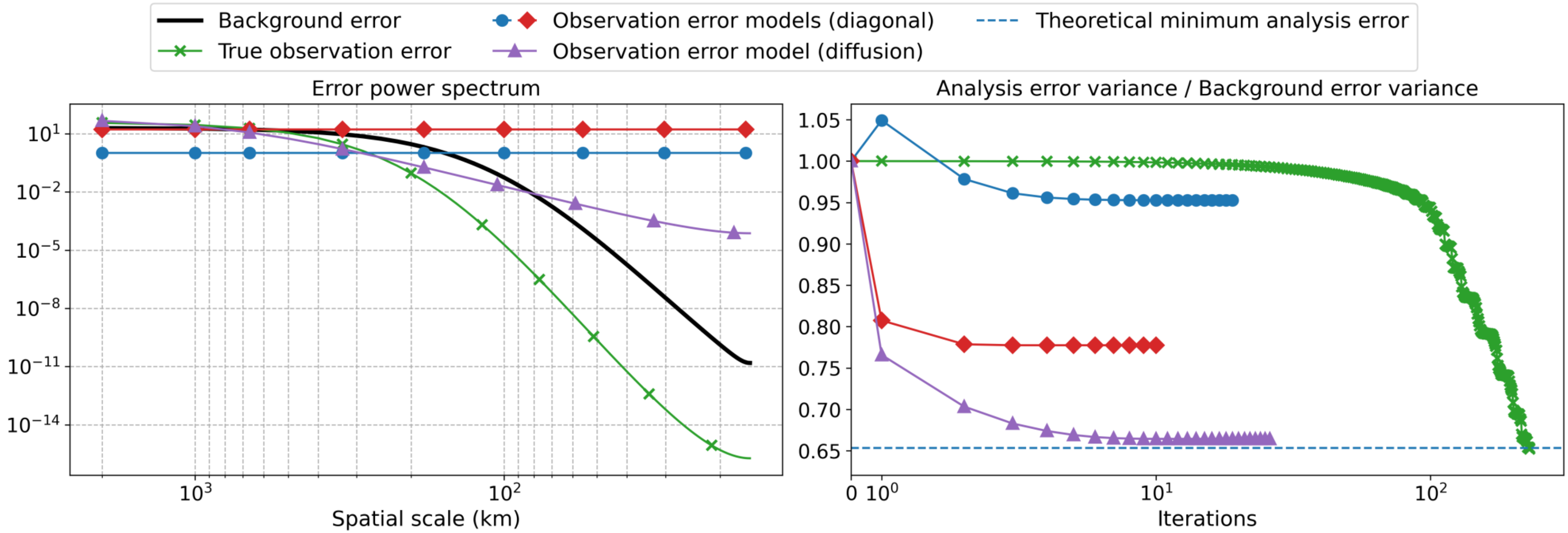


Using the most accurate parameters for the observation error lead to a very low analysis error but sometimes only after a prohibitive number of iterations (especially if $M_o > M_b$)

Convergence experiment

True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$

`Reconditioned' observation error model : $D_o = 120$ km ; $M_o = 2$

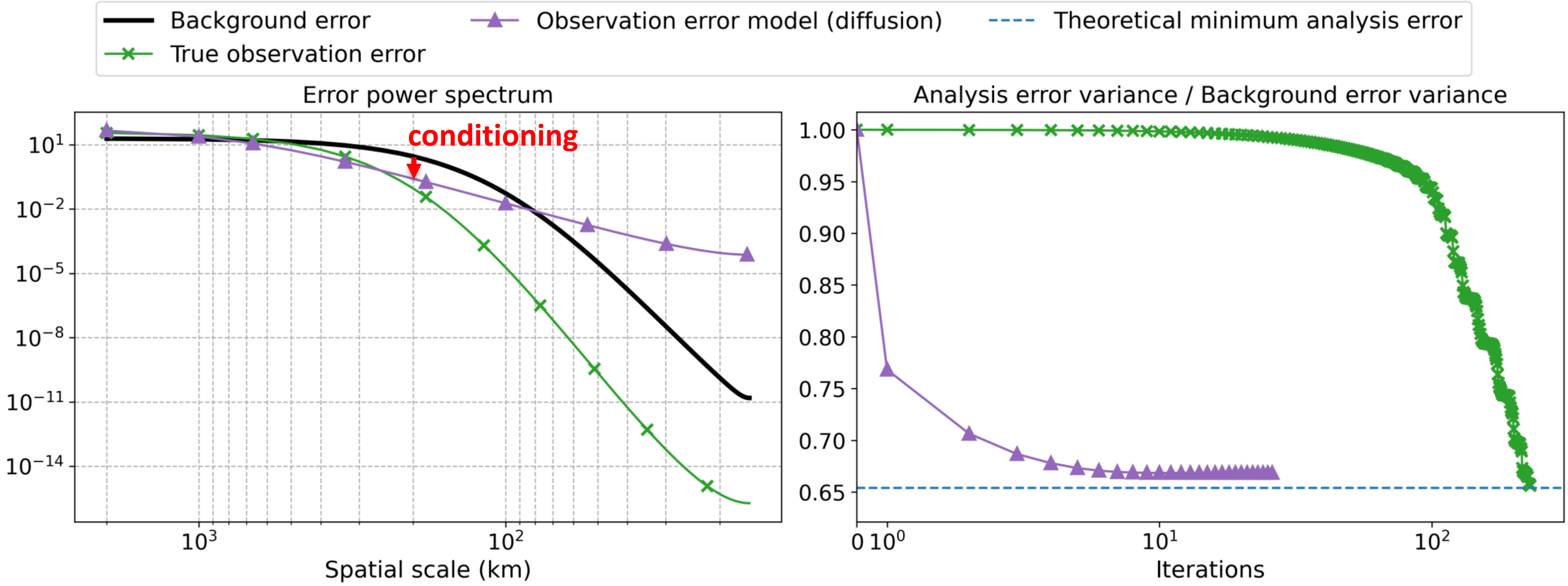


Reducing M_o is a viable strategy to reach a reasonable convergence rate without degrading too much the analysis error at full convergence.

Convergence experiment

True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$

`Reconditioned' observation error model : $D_o = 120$ km ; $M_o = 2$



Reducing M_o increases the observation error power specifically at the smallest spatial scales where both the background and observation error are insignificant, which limits the impact on the analysis error.

Summary

- Diffusion operators could be used in an operational context to model the inverse of the observation error correlation operator.

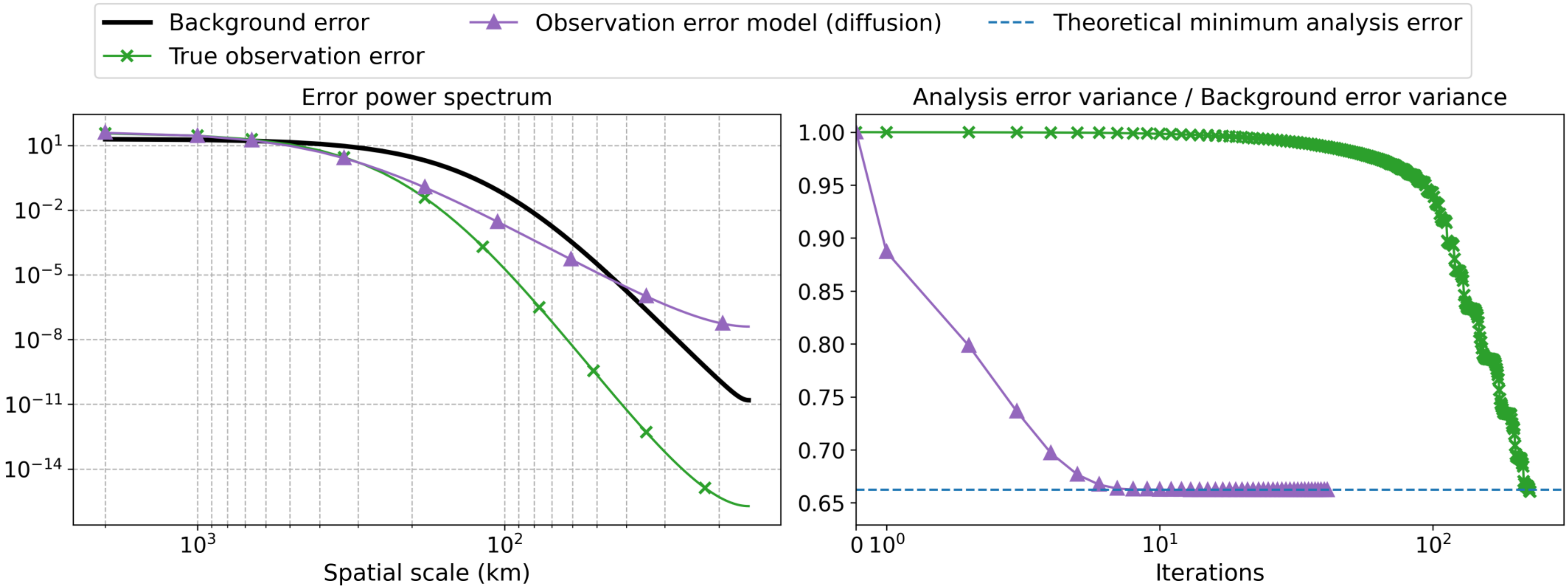
- The choice of parameters of the diffusion operator, or any observation error correlation model, should account for their impact on the convergence rate of the B-PCG, as a non-diagonal R can drastically improve or degrade the convergence rate.

- The conditioning tends to be improved by using a non-diagonal **R** over a diagonal **R** if $M_o \leq M_b$, and enforcing this relation even if it is not representative of the estimated error statistics is a viable `reconditioning' strategy.

The next step is to implement diffusion operators for observation error correlations in NEMOVAR to evaluate the impact of a non-diagonal R in an operational system.

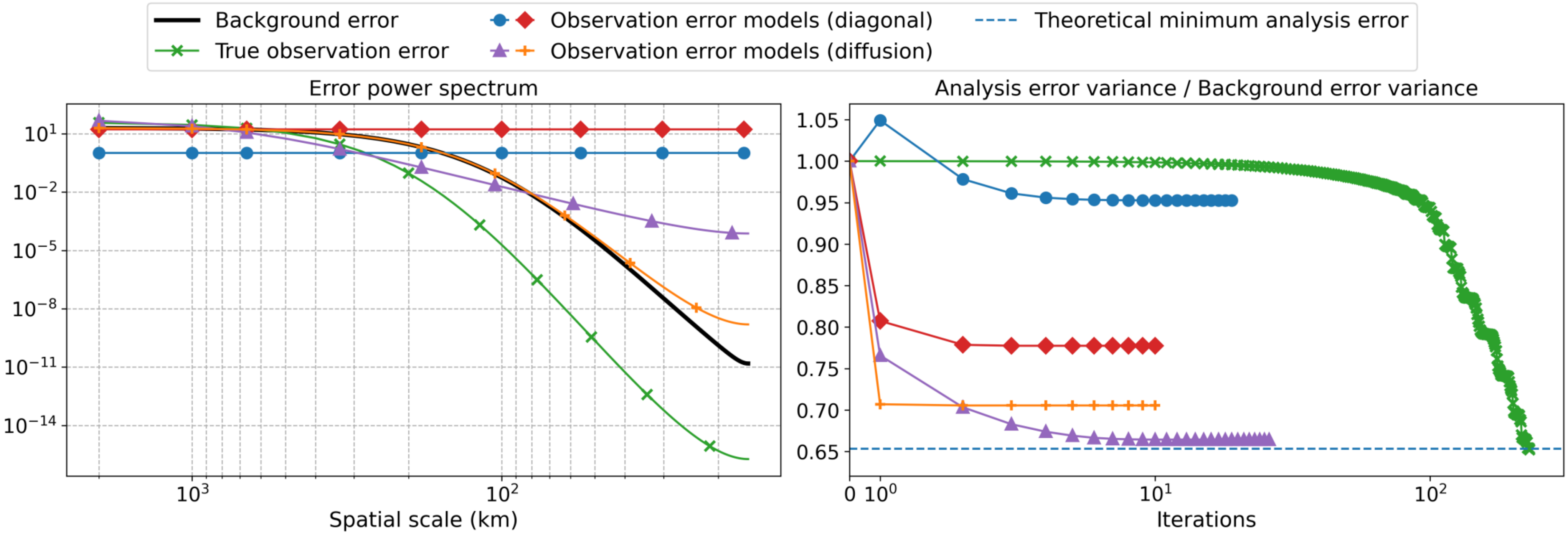
True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$

`Reconditioned' observation error model : $D_o = 120$ km ; $M_o = 4$



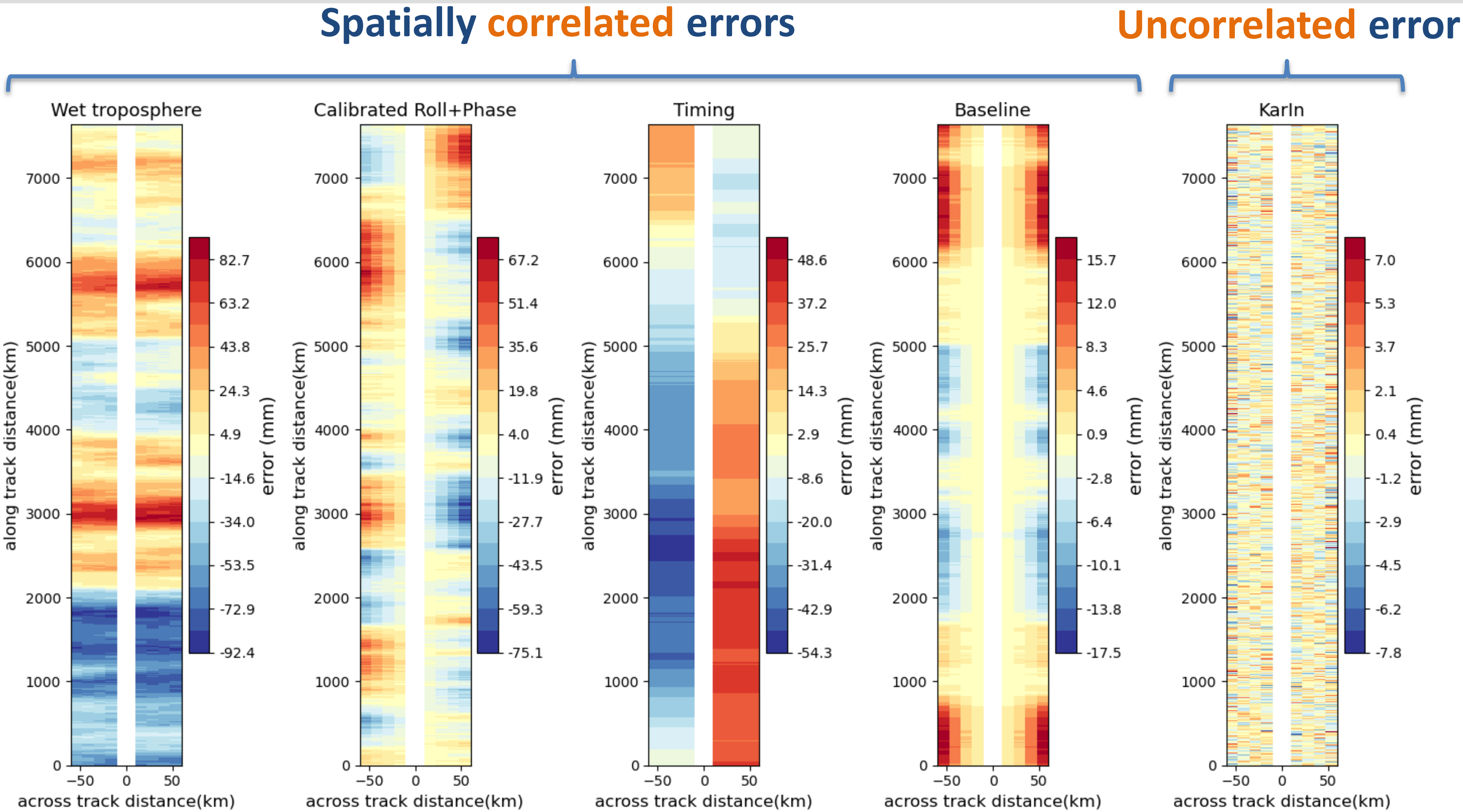
Annex: minimal condition number

True parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$
`Reconditioned' observation error model : $D_o = 120$ km ; $M_o = 2$
`Fastest' observation error model : $D_o = 60$ km ; $M_o = 8$

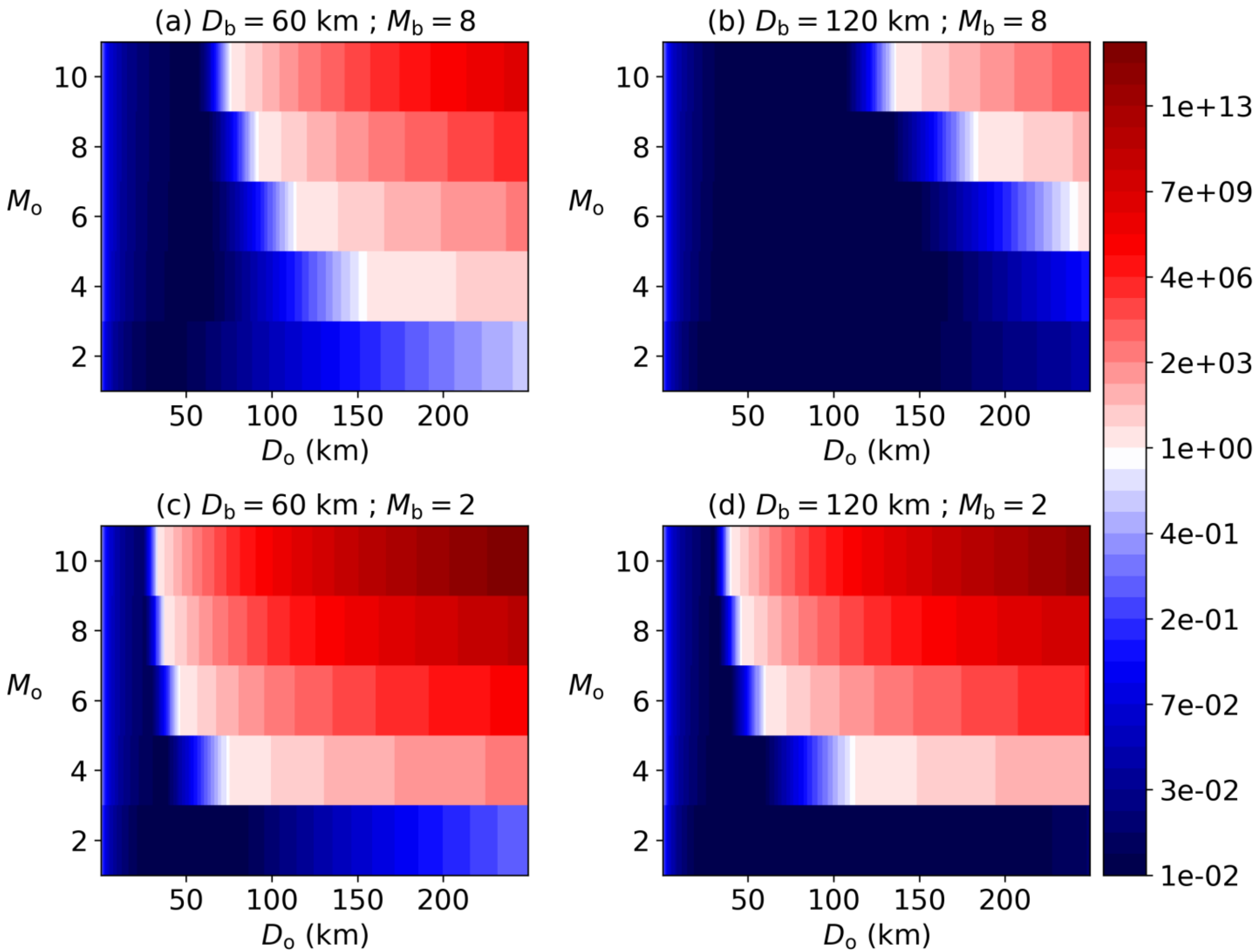


Even if we completely prioritize the convergence speed by specifying an observation error power as close as possible to the background error power, the analysis error at full convergence is still lower than with the best diagonal model.

Anexx: SWOT error budget

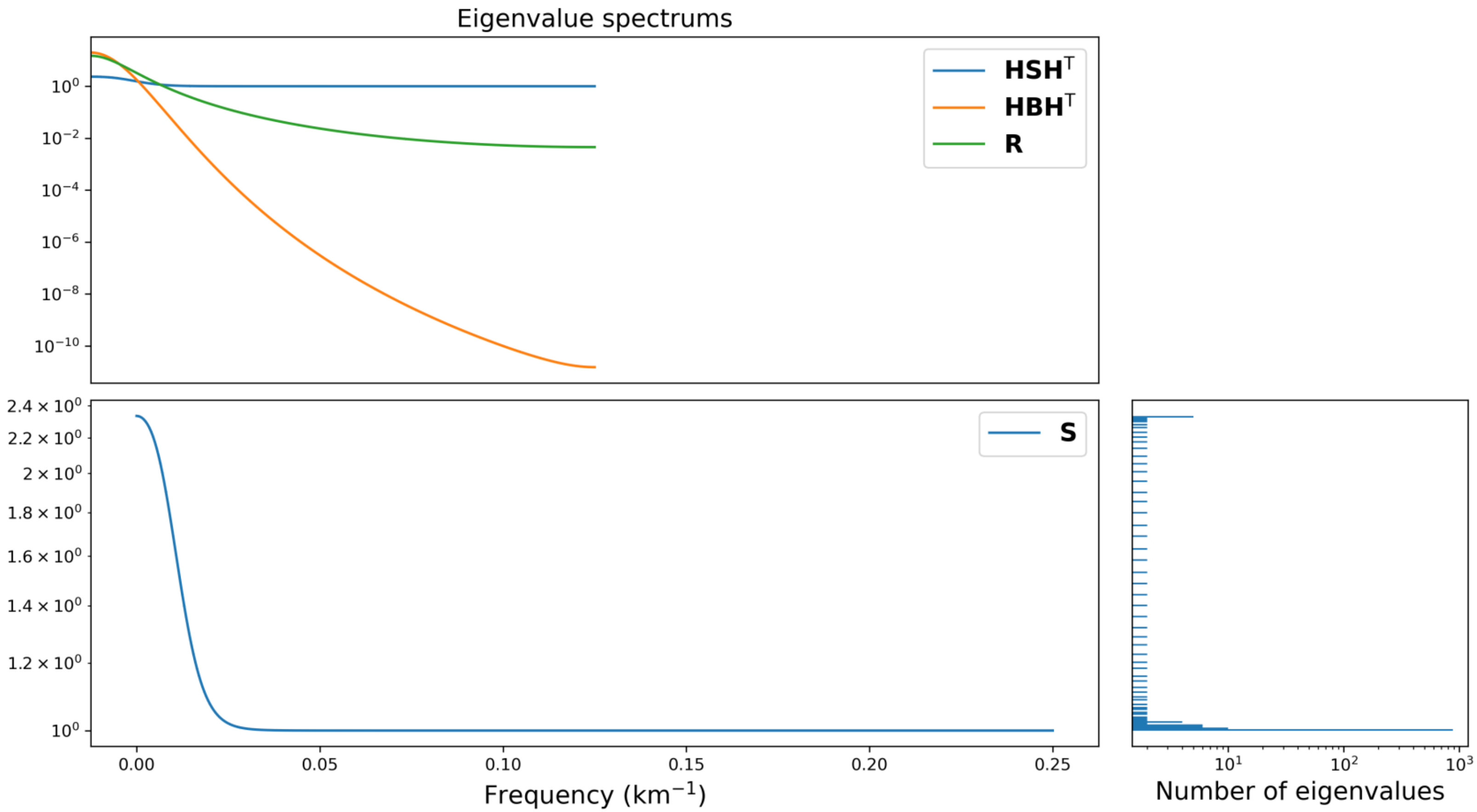


Condition number with a non-diagonal **R** / Condition number with a diagonal **R**



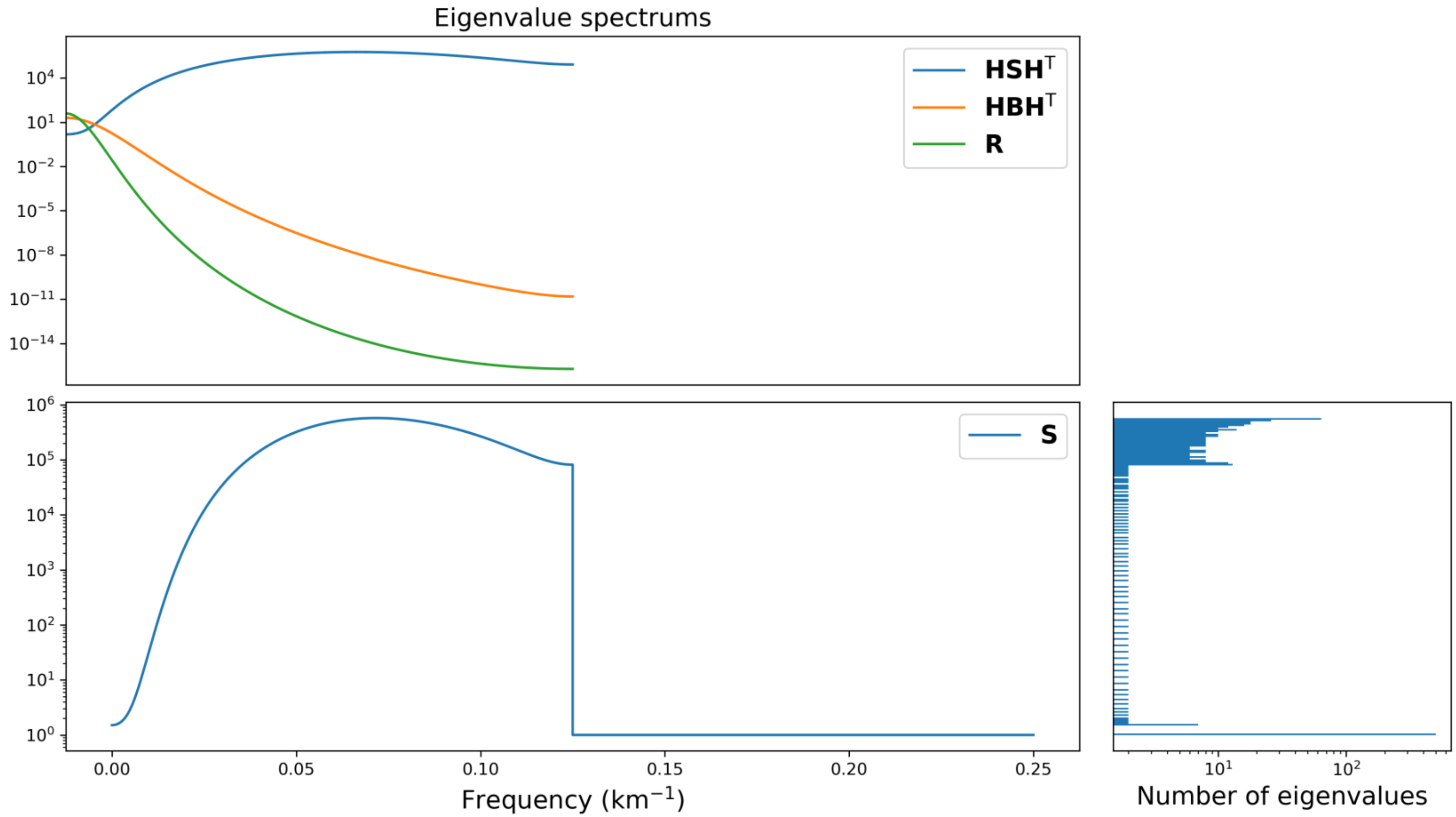
Annex: clustering with small OECs

Parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 30$ km ; $M_o = 2$



Annex: clustering with large OECs

Parameters: $D_b = 60$ km ; $M_b = 8$; $D_o = 120$ km ; $M_o = 10$



Annex: nadir altimer

Nadir altimeter error correlations estimated via the JPL/CNES SWOT simulator