Convergence Issues in the Estimation of Interchannel Correlated Observation Errors in Infrared Radiance Data

Pierre Gauthier

Department of Earth and Atmospheric science

Université du Québec à Montréal (UQAM)

Co-authors: Ping Du, Sylvain Heilliette and Louis Garand

Meteorology research division

Environment and Climate change Canada (ECCC)

Outline

- Estimation of interchannel error correlation for hyperspectral infrared sounders (e.g., AIRS, IASI,CrIS)
  - Statistical consistency diagnostics (Desroziers *et al.,* 2005).
  - Used by Bormann *et al.* (2010), Stewart *et al.* (2013), Weston *et al.* (2014)
  - Environment Canada (Heilliette and Garand, 2015)
- A simple 1D model based on a complex observation operator
  - True error statistics are known
  - Convergence of an iterative tuning approach
  - Impact of under- and over- estimated background error
  - Can we recover the true observation error covariances?
- Conclusions

Statistical estimation

- Data assimilation blends information from a background state $\mathbf{x}_b$ (e.g., a short-term forecast) to that from observations, $\mathbf{y}$ taking into account their relative errors
  - Background state: $\mathbf{x}_b = \mathbf{x}_{true} + \varepsilon_b$ Background error covariance: $\mathbf{B} = \langle \varepsilon_b \varepsilon_b^T \rangle$
  - Observations: $\mathbf{y} = \mathbf{y}_{true} + \varepsilon_o$ Observation error covariance: $\mathbf{R} = \langle \varepsilon_o \varepsilon_o^T \rangle$
- Best linear unbiased estimate, the analysis $\mathbf{x}_a$, is

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b) = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{x}_b + \mathbf{K}\mathbf{y}$$

  with $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$,the gain matrix and **H**, a linear observation operator (e.g., the Jacobian of a radiance transfer model)
- The assimilation needs **B** and **R** to correctly weigh these two sources of information

Statistical consistency

- Desroziers et al. (2005): what should we obtain if there were a perfect agreement between the estimated and ***a priori*** error statistics
- Observation departures: $\mathbf{d} = (\mathbf{y} - \mathbf{H}\mathbf{x}_b), \quad \mathbf{a} = (\mathbf{y} - \mathbf{H}\mathbf{x}_a)$

$$\langle \mathbf{d}\mathbf{d}^T \rangle \equiv \mathbf{D} = (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) \equiv \mathbf{D}$$

$$\langle \mathbf{a}\mathbf{d}^T \rangle \equiv \tilde{\mathbf{R}} = \mathbf{R}(\mathbf{D}^{-1}\mathbf{D})$$

$$\langle (\mathbf{d} - \mathbf{a})\mathbf{d}^T \rangle \equiv \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T = \mathbf{H}\mathbf{B}\mathbf{H}^T(\mathbf{D}^{-1}\mathbf{D})$$

  - The **blue color** indicates that they are deduced from the innovations **d** but not directly measured.
  - Difficulty is to disentangle the observation and background error from the innovations
  - Additional information is needed (Talagrand, 1999, 2003)

An iterative approach to estimate the error statistics

- Consistency diagnostic is a necessary but not sufficient condition for the background and observation error covariances to be the true one
- Iterative method to solve two matrix equations

$$\left\langle \mathbf{a}\mathbf{d}^{T} \right\rangle \equiv \mathbf{R}_{k+1} = \mathbf{R}_{k}\left(\mathbf{D}_{k}^{-1}\mathbf{D}_{true}\right)$$

$$\left\langle (\mathbf{d}-\mathbf{a})\mathbf{d}^{T} \right\rangle \equiv \left(\mathbf{HBH}^{T}\right)_{k+1} = \left(\mathbf{HBH}^{T}\right)_{k}\left(\mathbf{D}_{k}^{-1}\mathbf{D}\right)$$

- These can be obtained by introducing these estimates in the assimilation to obtain an updated analysis.
- Approach used in Gauthier *et al.* (MWR 2018) was to devise a 1D assimilation for which the true error statistics are known.

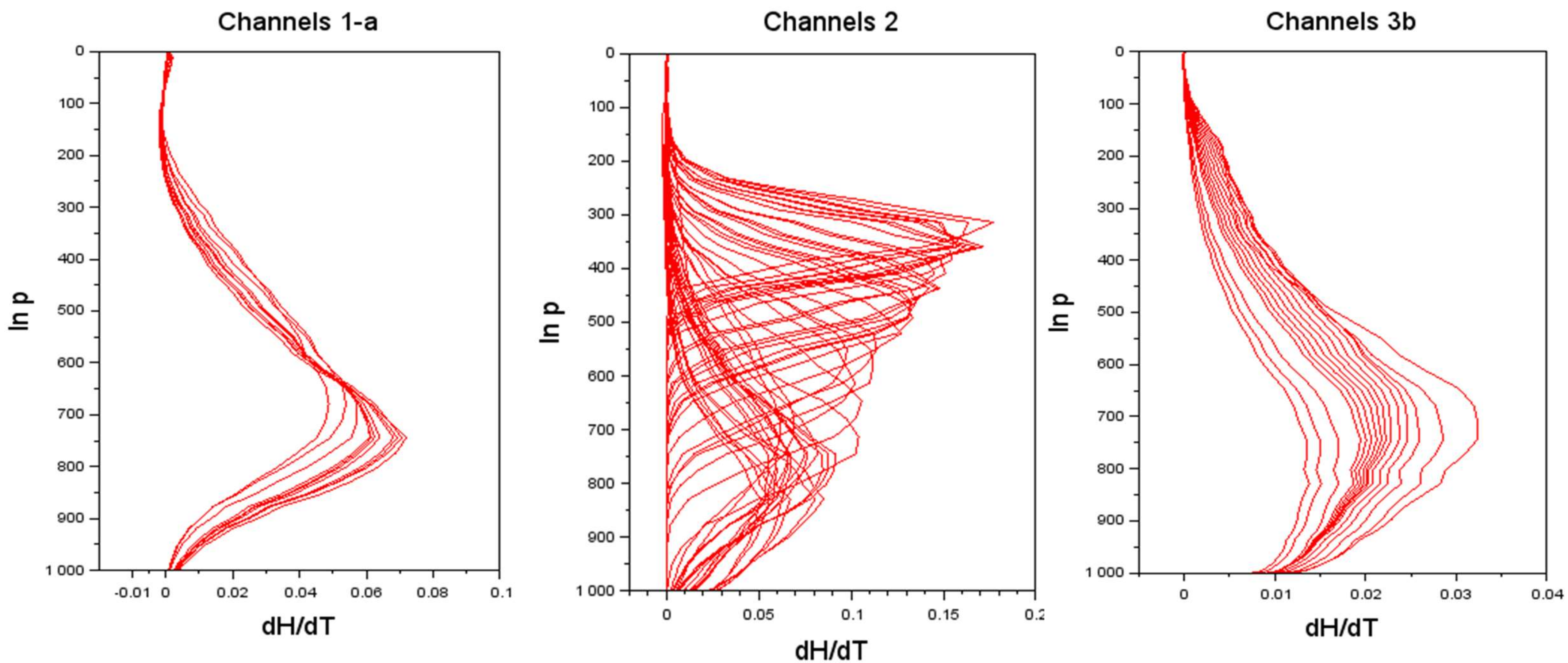
The 1D experiments

- **Observation operator**
  - Used for the assimilation of infrared radiances from hyperspectral sounders (e.g., AIRS, IASI)
  - Radiative transfer (RTTOV) has been linearized around a real background state (T, ln q, p_s , T_s) on 80 levels up to 0.1 hPa
  - The observation operator **H** is therefore linear
- **Background error covariances** (**B**)
  - Used for the climatological component (3D-Var) of the operational EnVar assimilation system of Environment Canada.
- Assume $\mathbf{R}_{true}$ and $\mathbf{B}_{true}$ based on what is used in the operational system
  - The true innovation covariances **D** are known
  - The background error covariances are assumed to be known and well calibrated
  - Only tune the observation error covariances **R** to evaluate the convergence and if the true value is recovered

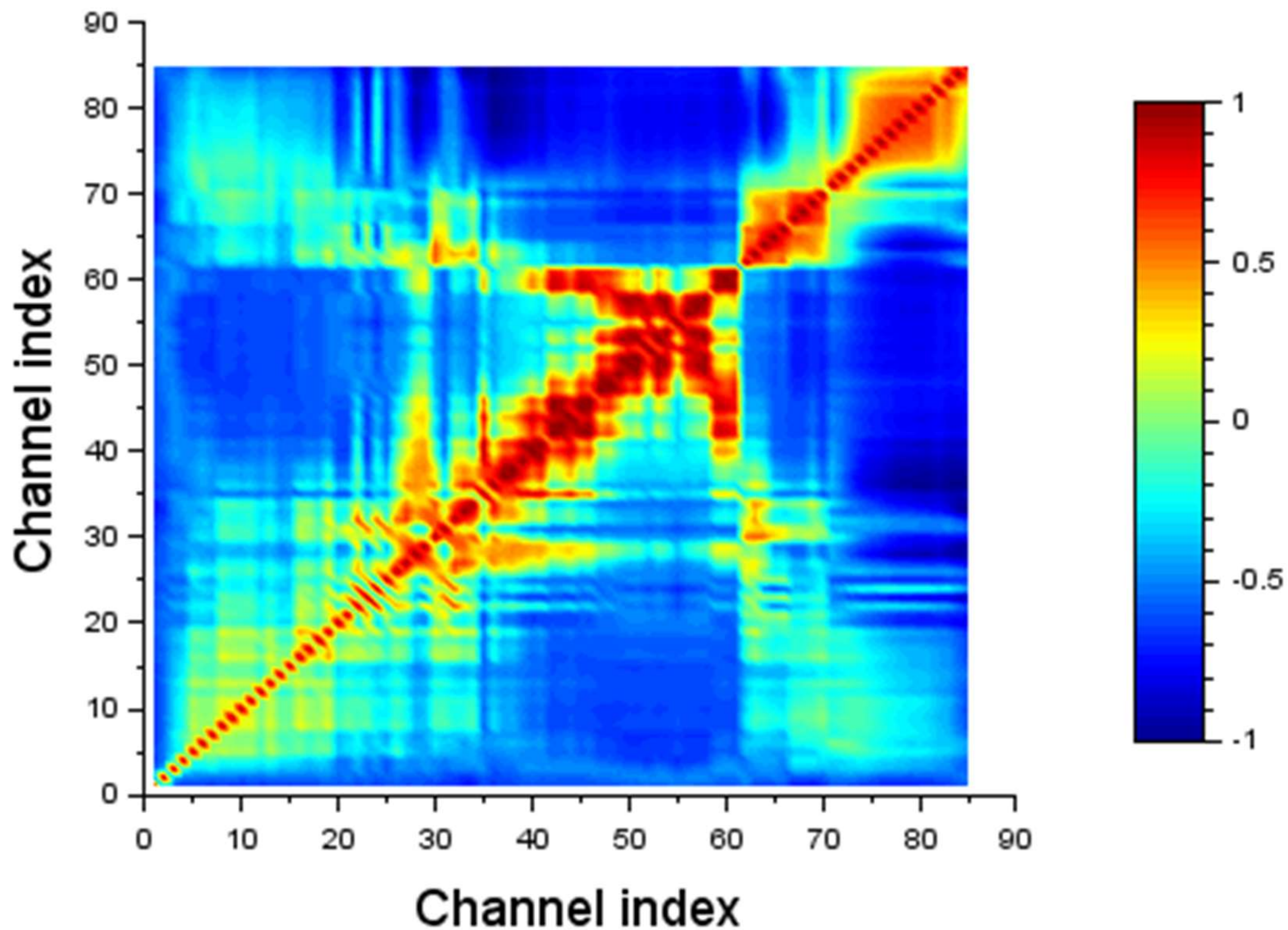
Table 1 Channel groups for AIRS

Spectral range					
Group	**Index**	**Channel No.**	**cm^{-1}**	**µm**	**Main sensitivity**
1-a	5-15	232-310	716-739	14.0-13.5	T, T_s, CO_2
1-b	16-34	333-1382	746-1292	13.4-7.7	T_s, T, H_2O, O_3
2	35-61	1424-1852	1316-1604	7.6-6.2	H_2O, T
3-a	62-70	1865-1911	2182-2224	4.6-4.5	T, T_s, N_2O
3-b	71-85	2112-2142	2391-2420	4.2-4.1	T_s, T, CO_2

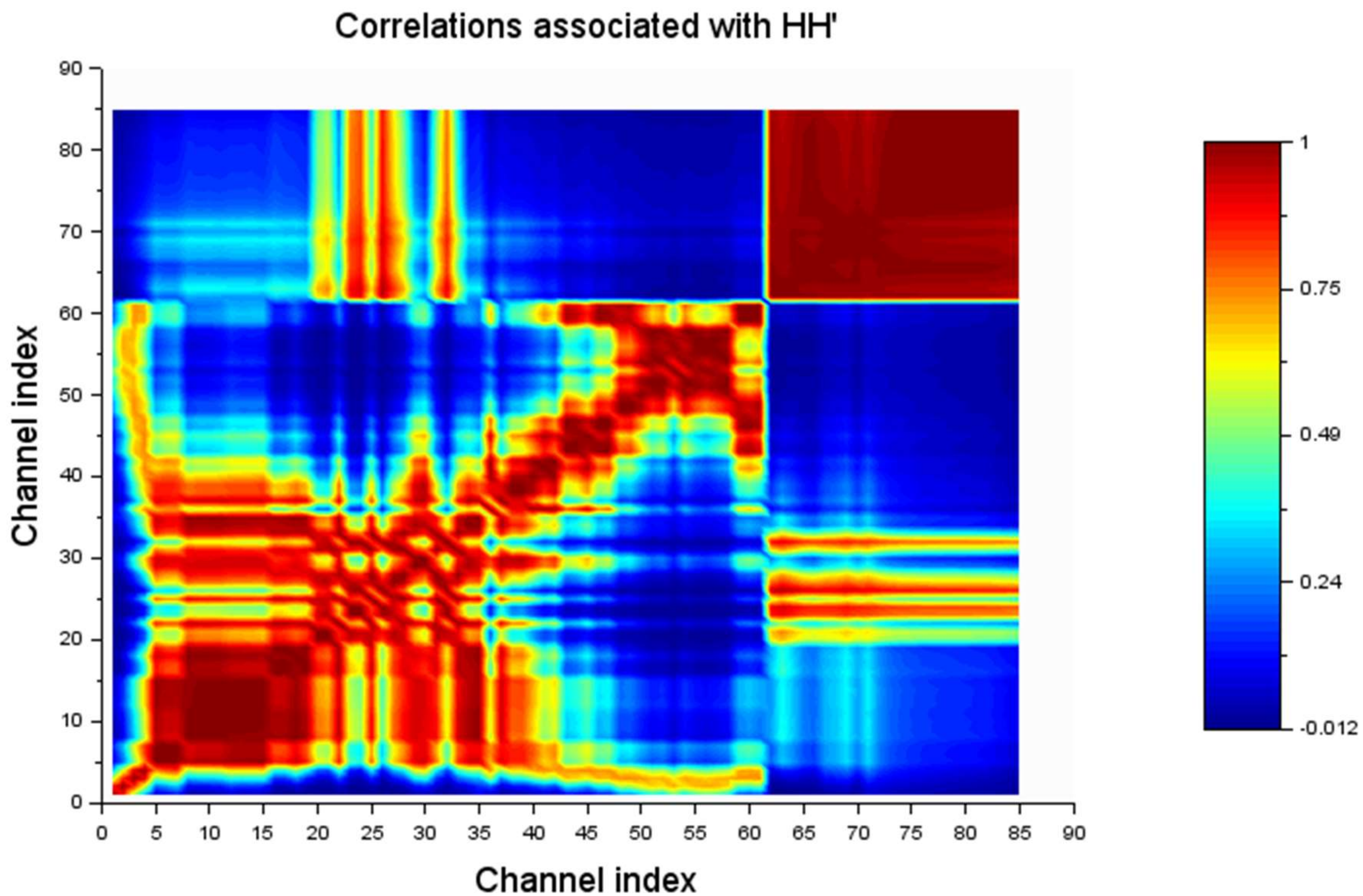
Jacobian with respect to temperature dH/dT



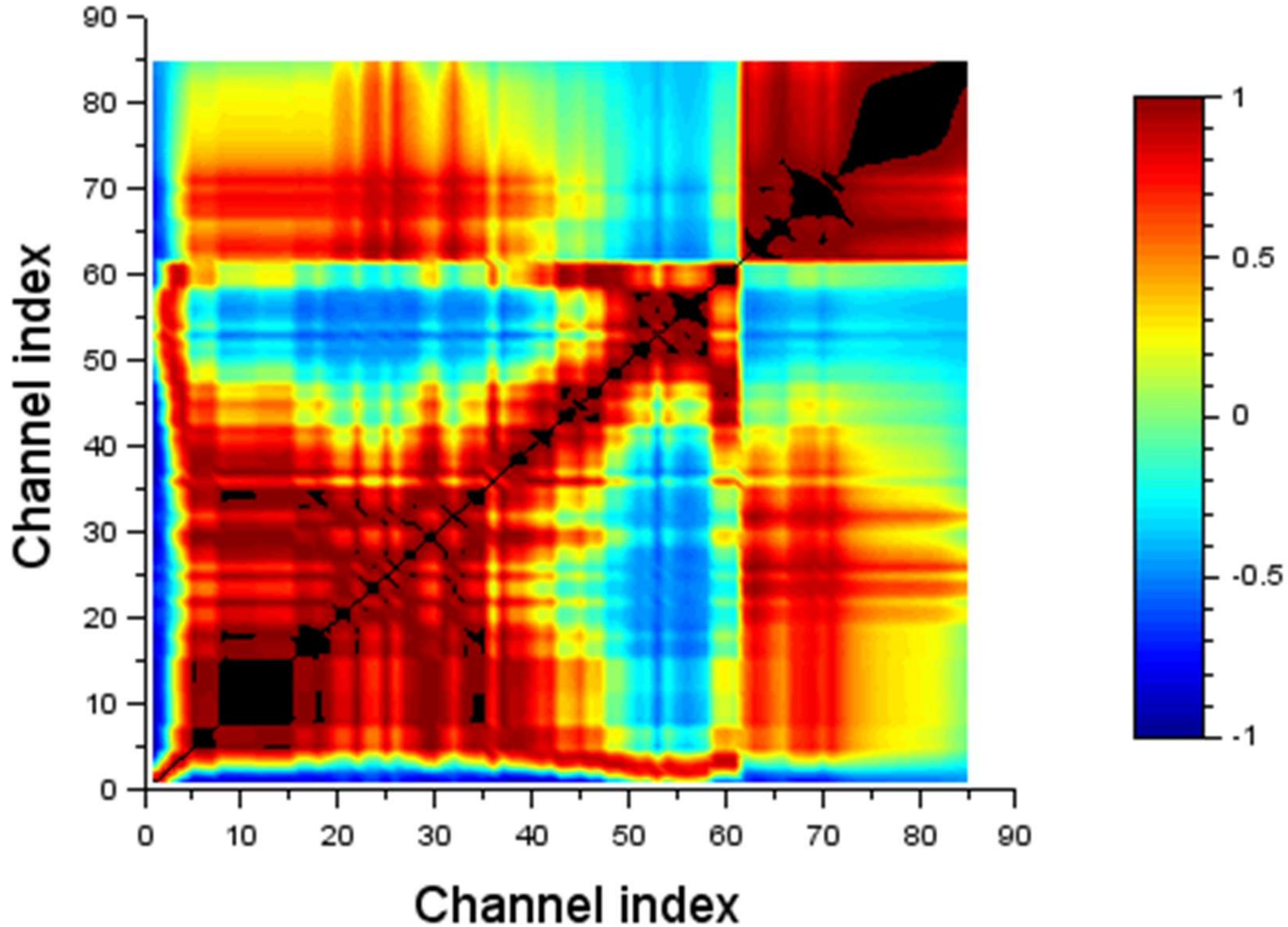
The “true” observation error correlations



Correlations associated with HH^T



Correlations of $HB_{true}H^T$



Tuning both R and HBH^T

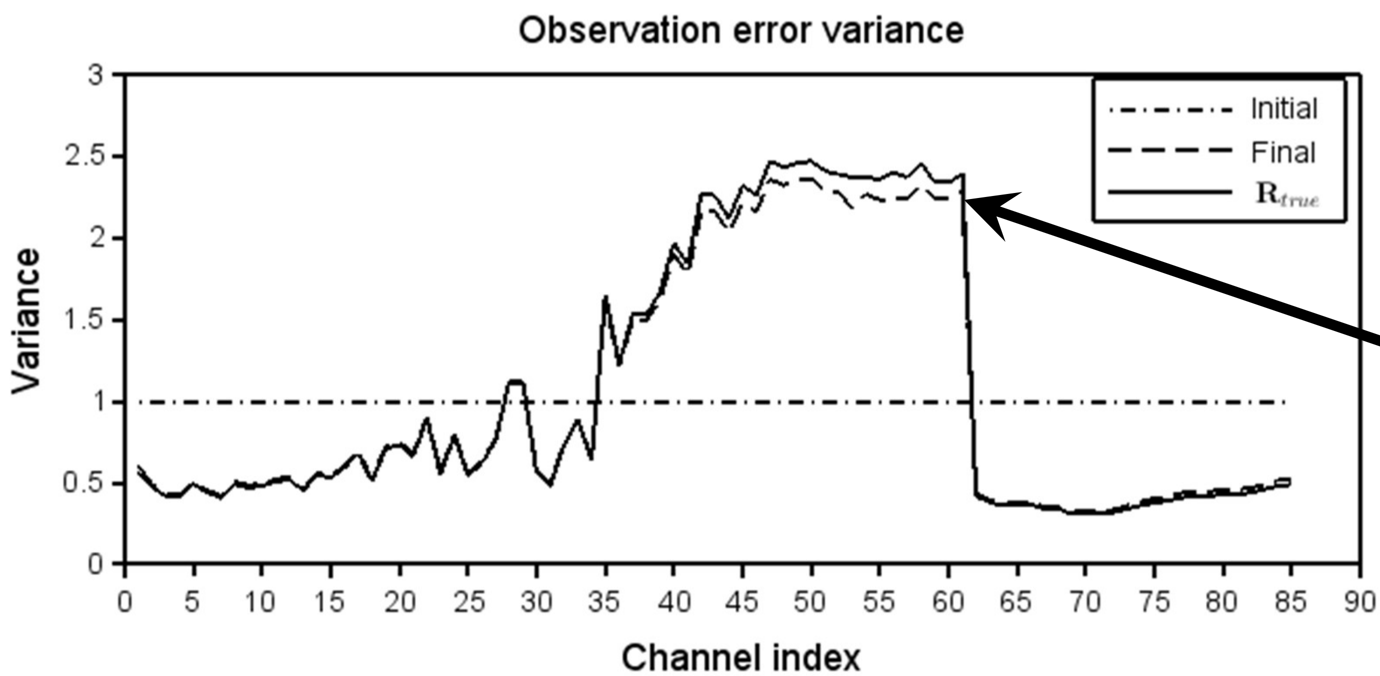
- If both **R** and (**HBH^T**) are corrected, then after one iteration

$$\mathbf{R}_1 = \mathbf{R}_0 \left(\mathbf{D}_0^{-1} \mathbf{D}_{true} \right), \qquad \left(\mathbf{HBH}^T \right)_1 = \left(\mathbf{HBH}^T \right)_0 \left(\mathbf{D}_0^{-1} \mathbf{D}_{true} \right)$$

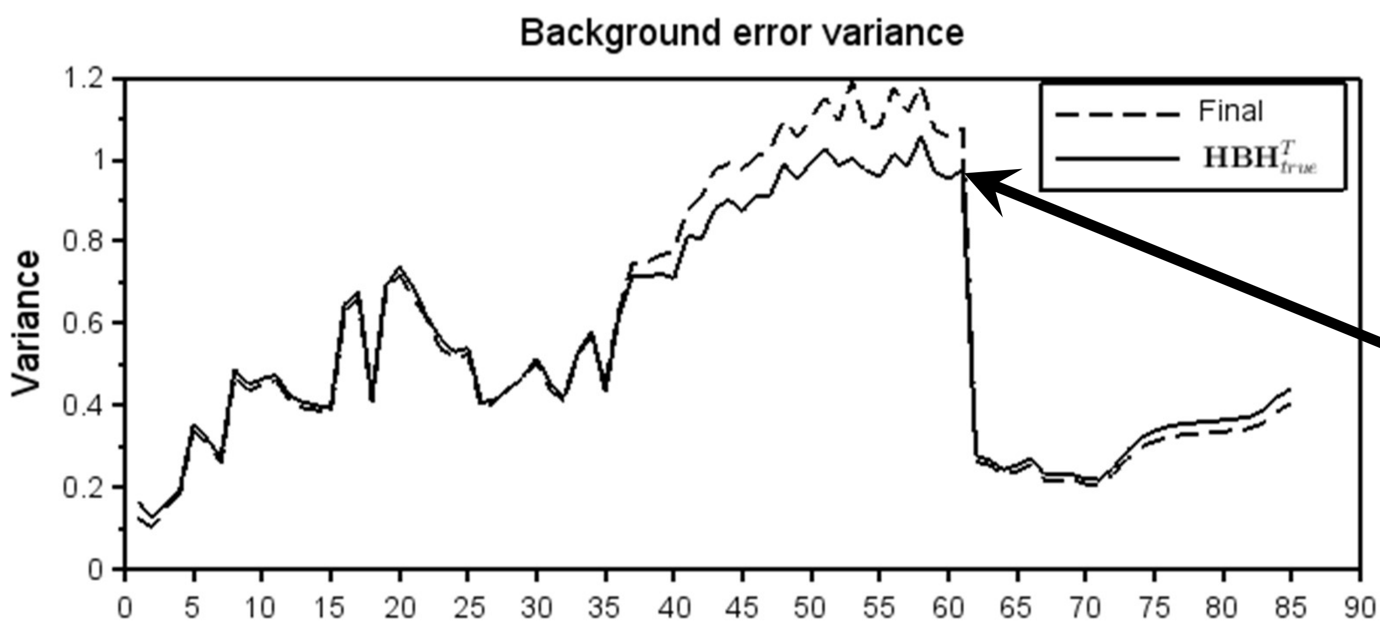
which leads to $\mathbf{D}_1 = \mathbf{D}_{true}$ and $\mathbf{HK}_1 = \mathbf{HK}_0$, the information content remains the same.

- Experiments in which $\mathbf{HBH}^T = \alpha\, \mathbf{HB}_{true}\mathbf{H}^T$, $\alpha > 1$ (<1), **B** is over (under) estimated

Tuning both R and HBH^T



Underestimated



Overestimated

R_t diagonal

$B = \alpha B_t$ $R_0 = Id$

$\alpha = 1$

Perfect fit to the innovation variance

Convergence when HBH^T is kept constant

- Solving iteratively the matrix equation: $\mathbf{R}_{k+1} = \mathbf{R}_k \left(\mathbf{D}_k^{-1}\mathbf{D}_{true}\right)$
- Measuring convergence:
  - Distance between iterates: $\left\|\mathbf{R}_{k+1} - \mathbf{R}_k\right\|_F$ Distance to the true solution: $\left\|\mathbf{R}_{k+1} - \mathbf{R}_{true}\right\|_F$
- Assuming $(\mathbf{HBH}^T) = \alpha \mathbf{HB}_{true}\mathbf{H}^T$ to be fixed, the solution should converge to

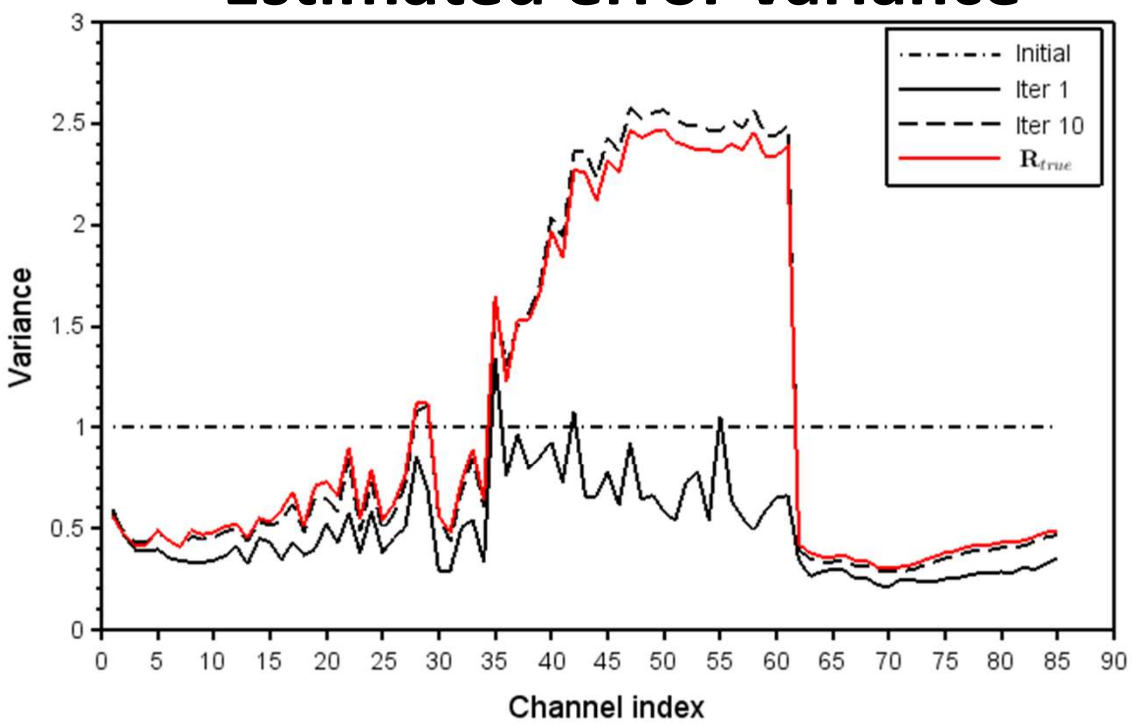
$$\boxed{\mathbf{R}^* = \mathbf{R}_{true} + (1-\alpha)\mathbf{HB}_{true}\mathbf{H}^T}$$

- Remarks:
  - Since for any two symmetric matrices $(\mathbf{MN})^T = \mathbf{NM} \neq \mathbf{MN}$

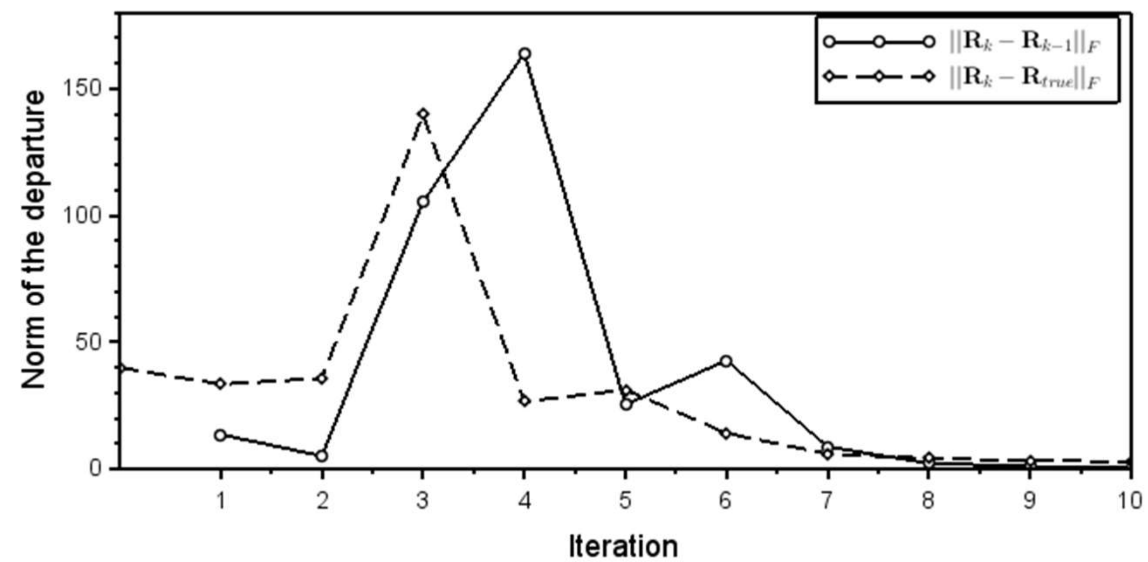
    it is necessary to filter each iterate ($\bar{\mathbf{R}}_{k+1} = \frac{1}{2}(\mathbf{R}_{k+1} + \mathbf{R}_k)$)

Experiment with $HBH^T = HB_{true}H^T$

Estimated error variance

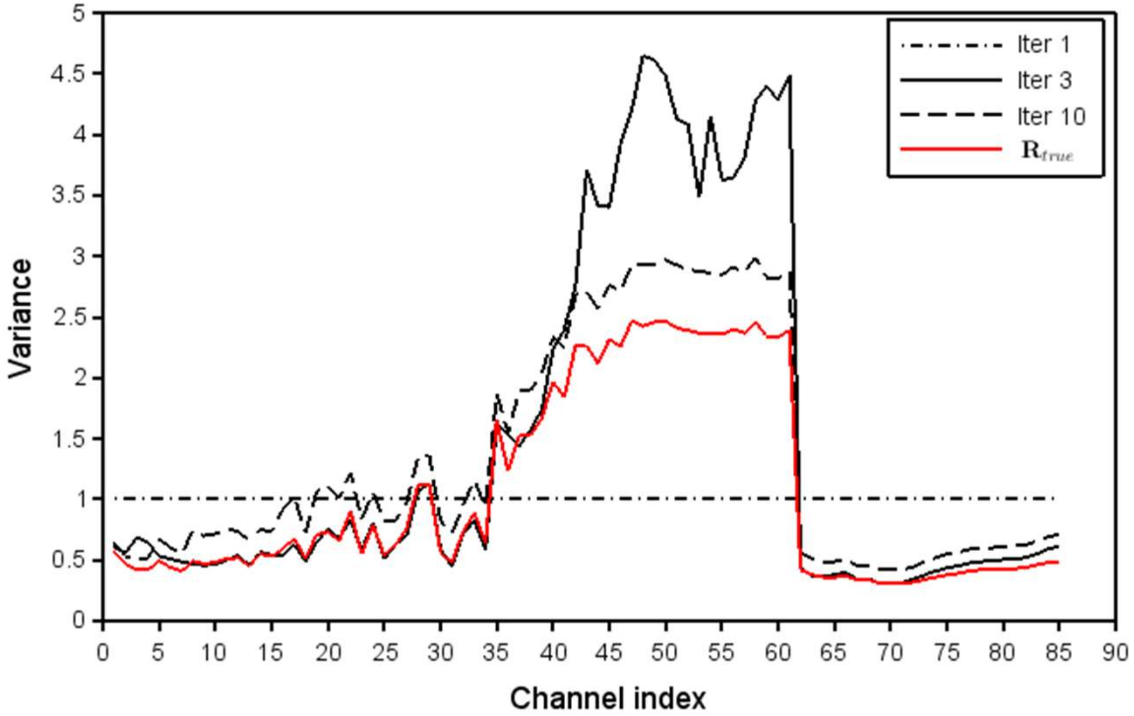


Convergence

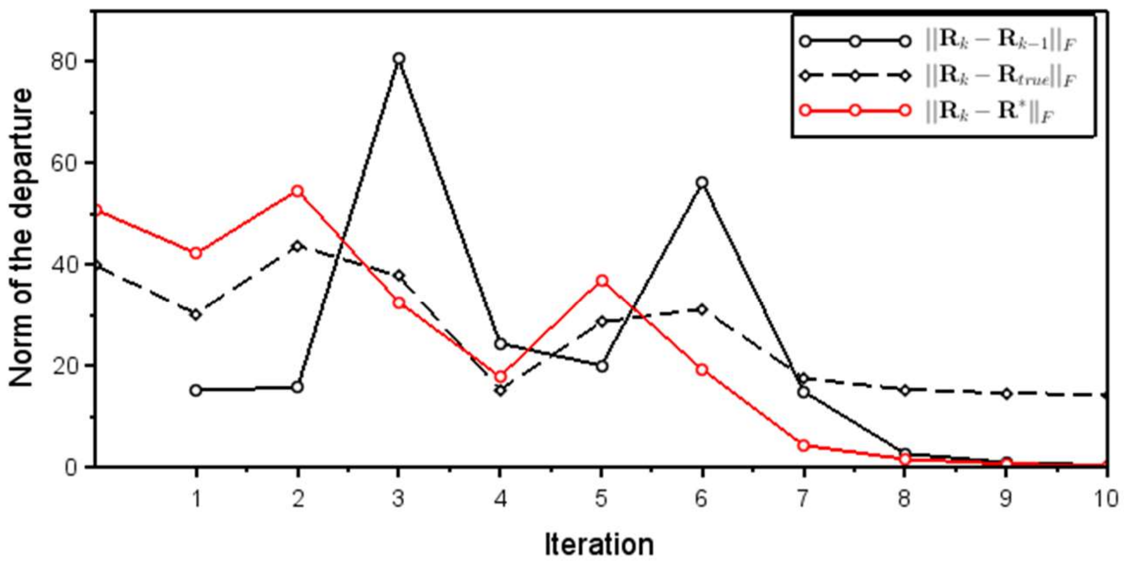


Experiment with $HBH^T = 0.5\ HB_{true}H^T$

Estimated error variance

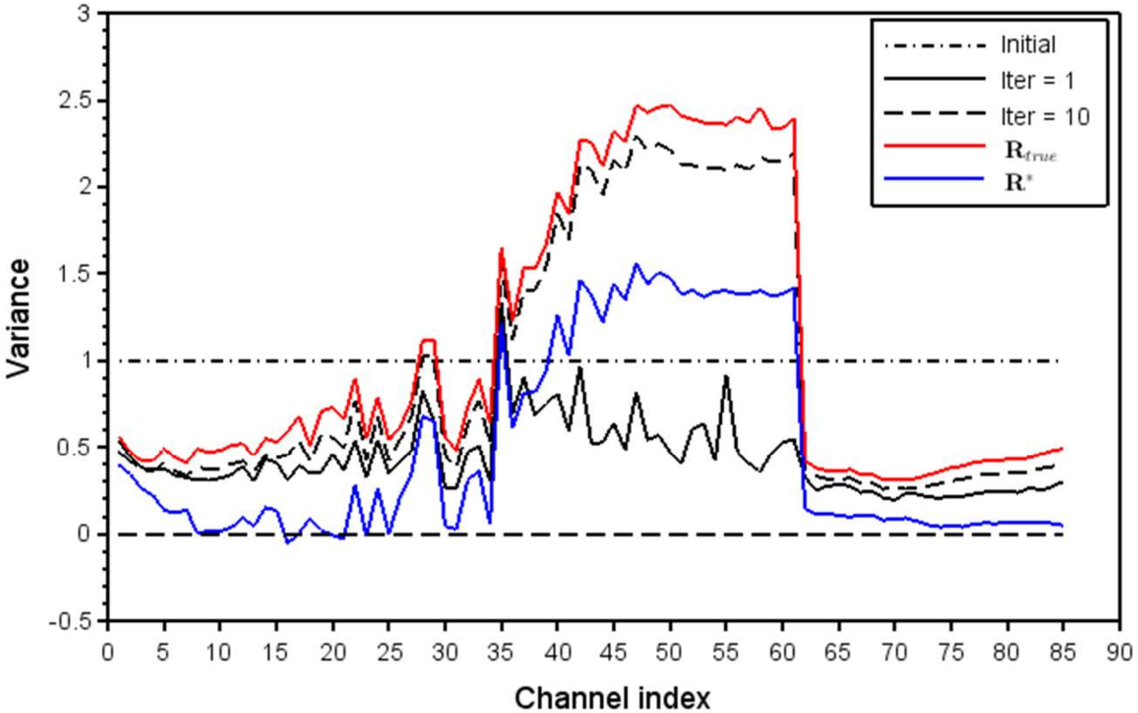


Convergence

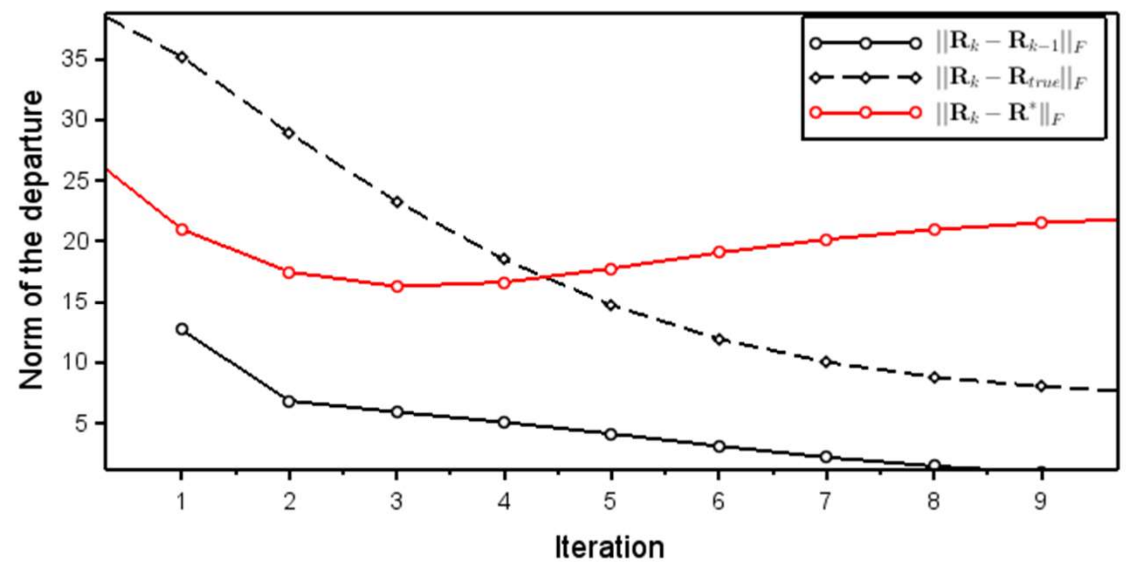


Experiment with $HBH^T = 2.0\ HB_{true}H^T$

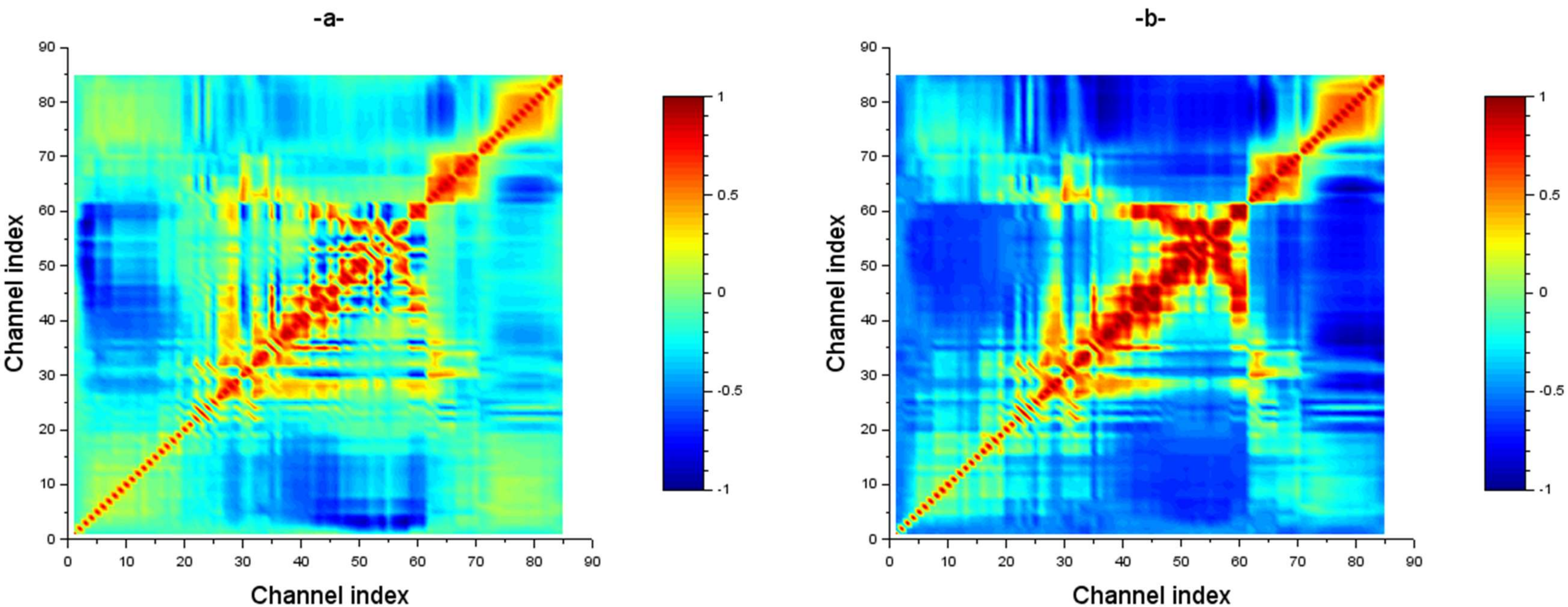
Estimated error variance



Convergence



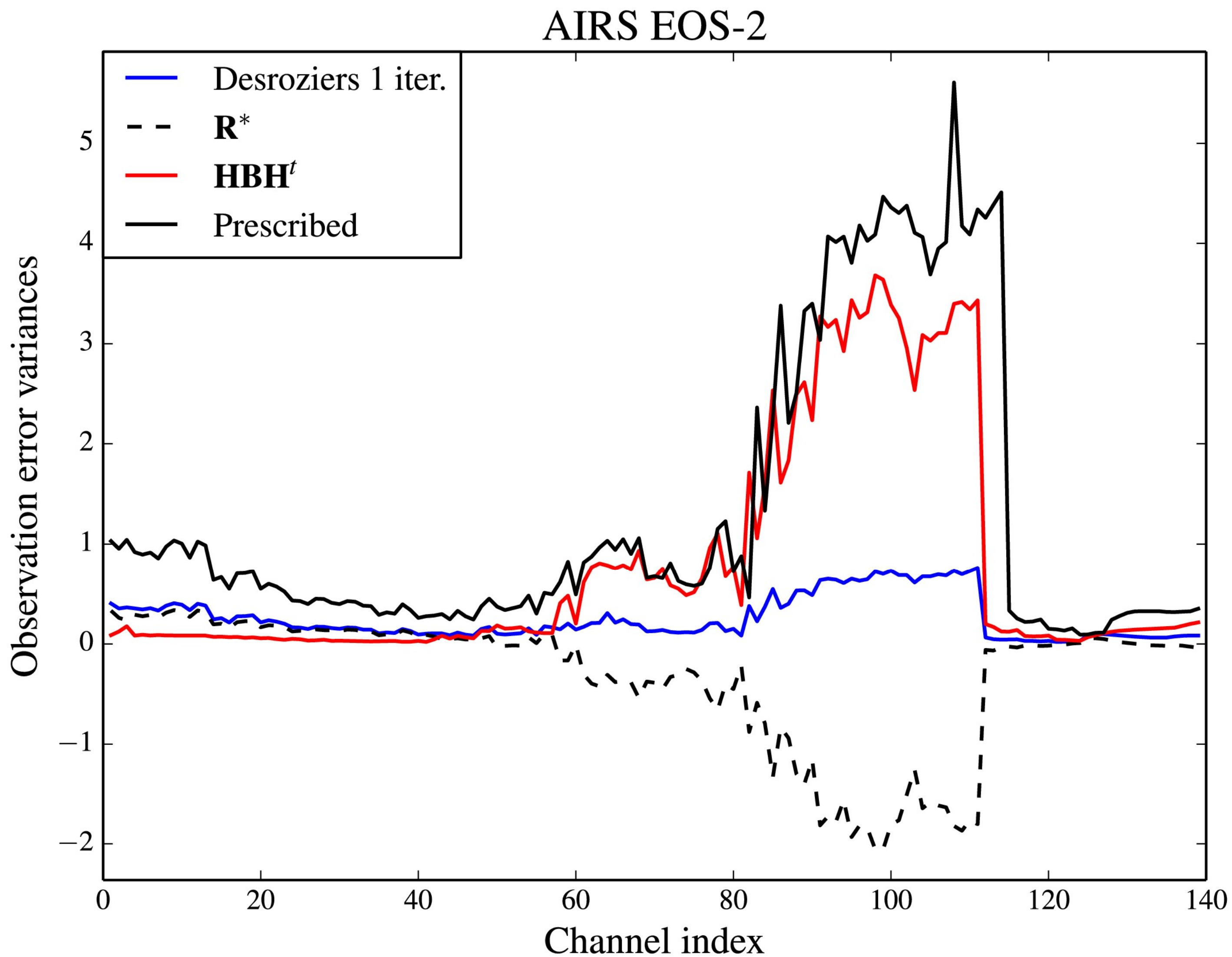
Estimated observation error correlations after a) One iteration, b)10 iterations



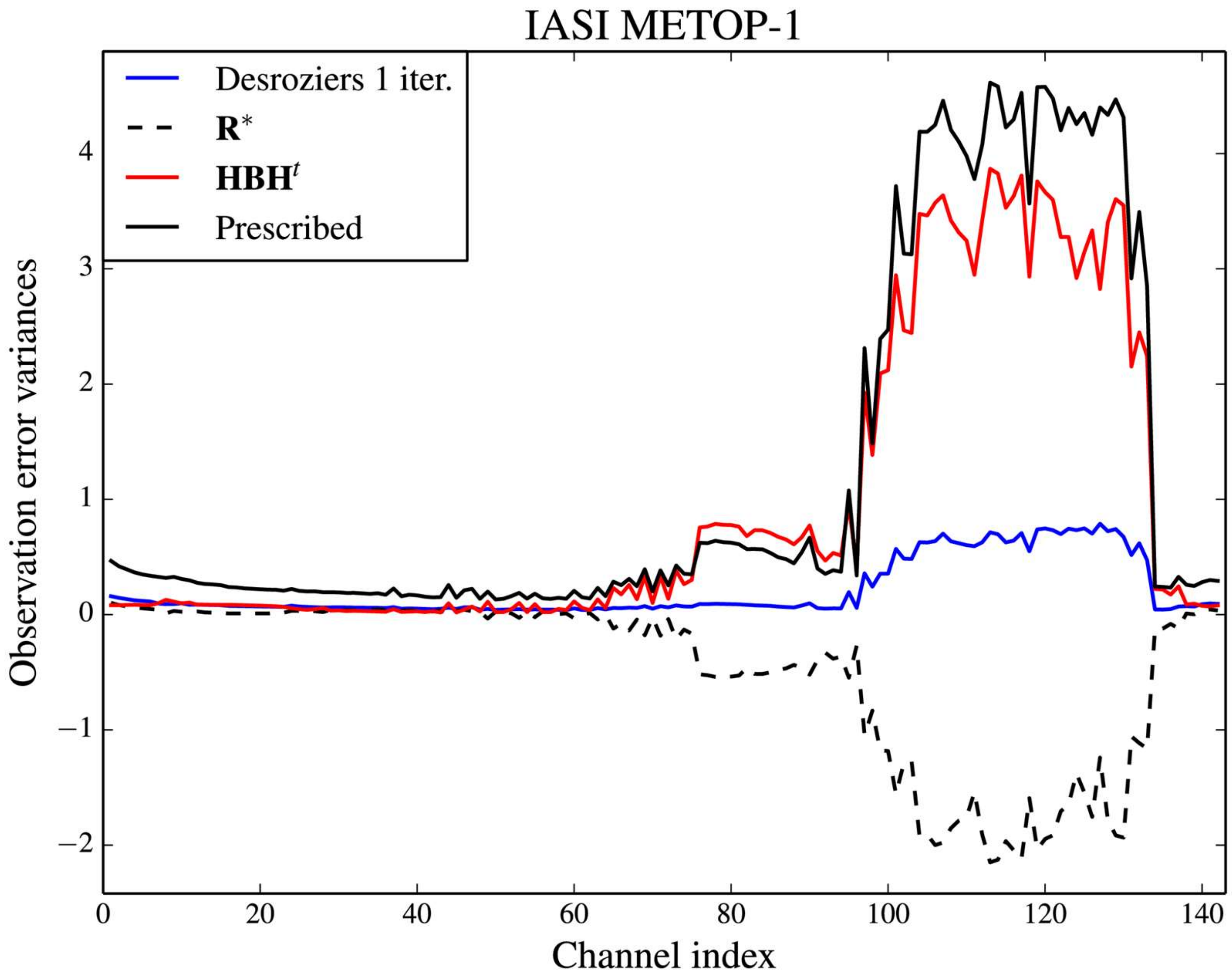
Using R* as a diagnostic

- Under the assumption that $\mathbf{HBH}^T$ correct, the observation error covariance can be obtained for all observations used in a data assimilation system
  - Estimates were computed from results of the Environment Canada DA system

Observation error variance prescribed (black), estimated with one iteration , and what we should have at convergence (dashed)



Observation error variance prescribed (black), estimated with one iteration , and what we should have at convergence (dashed)



Conclusion-1

- Statistical consistency provides a necessary condition for the error statistics to be correct but this is not sufficient to be able to estimate observation error
- An iterative method based on this principle is not converging to the exact solution
  - It does not even produce symmetric matrices
- Additional assumption: the background error statistics to be correct
  - The innovations would provide the exact solution **R***
  - Computing this solution for all observations used in the assimilation system of Environment Canada indicate that this assumption leads to unphysical observation statistics (negative variance)
  - Background error statistics need to be revisited

Conclusion-2

- Current estimate with EnKF has shown that the background error variance exceed that of the innovation
  - What our result show is that the observation error variances would provide a more detailed diagnostic to test if the background error statistics are consistent
- Observation error can be estimated on a physical basis (Chun et al., 2015)
- Caveats
  - Error statistics are Gaussian, unbiased and the observation operator is linear
- Reference: Gauthier, P., P. Du, S. Heilliette and L. Garand, 2018: *Mon.Wea.Rev*, **146**, 3227-3239.

Acknowledgements

- Part of this work was done while on sabbatical at the University of Reading
- Thanks to Profs. Peter van Leeuwen and Nancy Nichols for having me and also to Prof. Sarah Dance and Dr. Joanne Waller for the discussions we have had on this subject.