Time-correlated model error in (ensemble) Kalman smoothers

Javier Amezcua
Peter Jan van Leeuwen

8th EnKF Workshop. Mont Gabriel, Canada. May 2018.

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL
Data Assimilation Research Centre
University of Reading

Outline

1. Introduction and motivation
2. Error in model error
3. Weak-constraint solution to Kalman Smoother
4. Uni-dimensional example
5. Introducing errors
6. Small error approximation
7. Conclusions, future work

1. Setup: smoothing problem

$\mathbf{x}^t \in \mathcal{R}^{N_x}$ **Model variables**

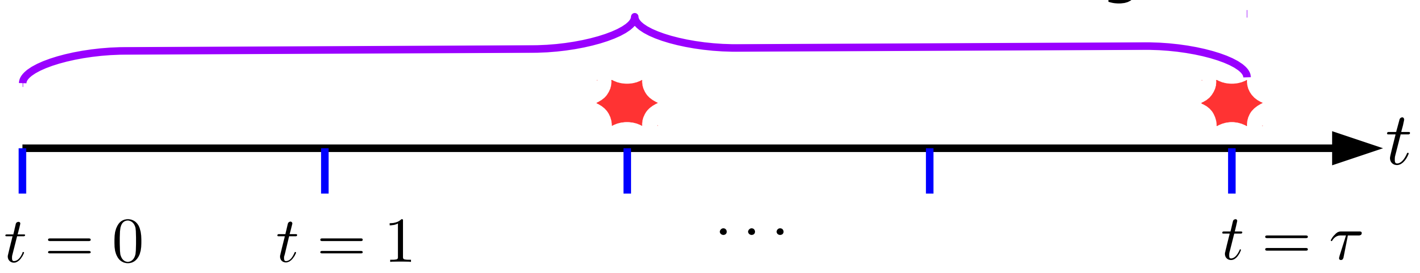
$\mathbf{y}^l \in \mathcal{R}^{N_y}$ **Observations**

$$\mathbf{x}^t = m^{(t-1)\to t}\left(\mathbf{x}^{t-1}\right) + \mathbf{v}^t$$

$$\mathbf{y}^l = h^l\left(\mathbf{x}^{t=l}\right) + \boldsymbol{\eta}^l$$

$$\{\mathbf{x}^0, \mathbf{v}^t, \boldsymbol{\eta}^l\}\ r.v.,\ \ \mathbf{x}^0 \perp \mathbf{v}^t \perp \boldsymbol{\eta}^l$$

assimilation window: **smoothing**



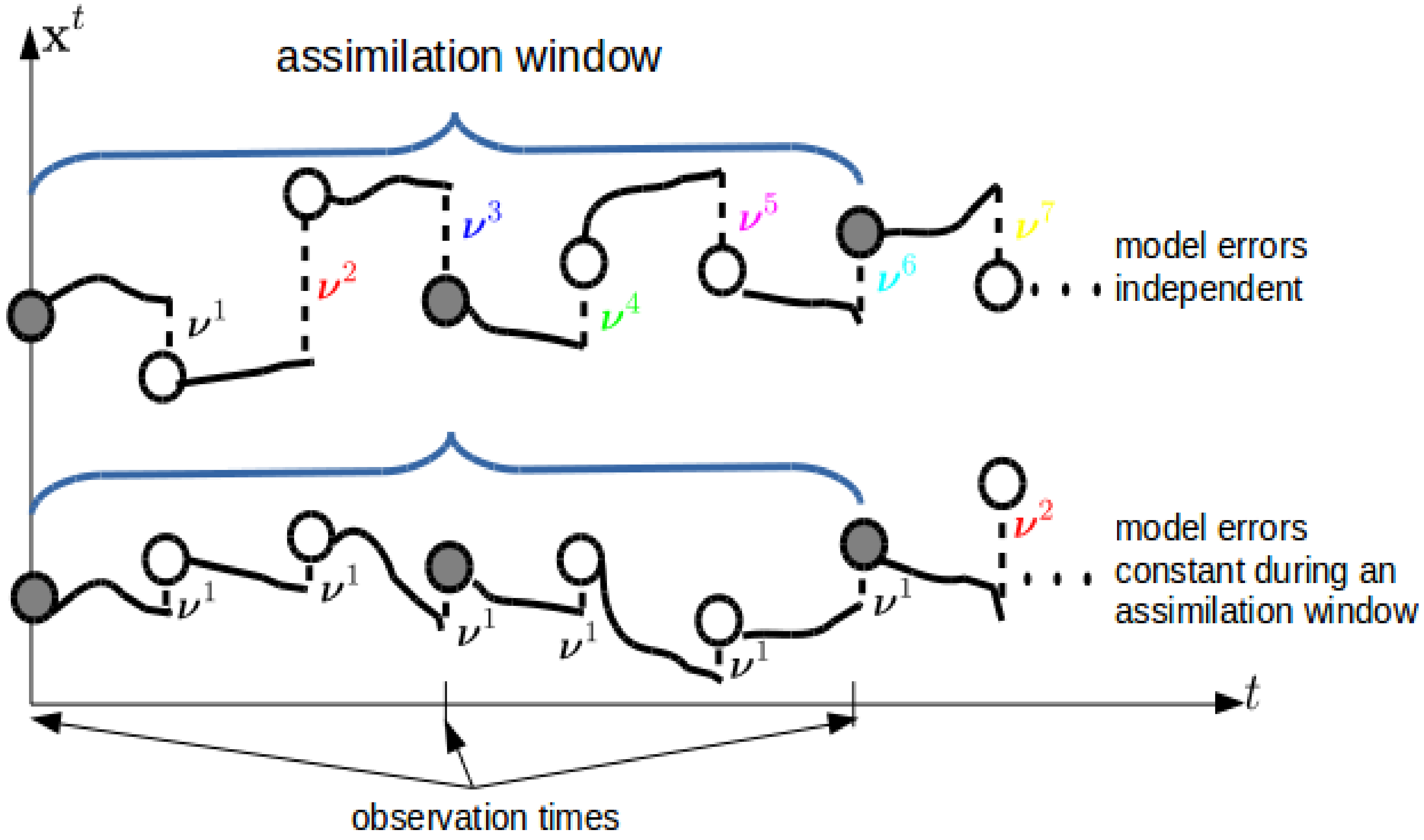
To obtain the **posterior** pdf we can use **Bayes' theorem**.

$$p\left(\mathbf{x}^{0:\tau} | \mathbf{y}^{1,L}\right) = \frac{p\left(\mathbf{y}^{1:L} | \mathbf{x}^{0:\tau}\right) p\left(\mathbf{x}^{0:\tau}\right)}{p\left(\mathbf{y}^{1:L}\right)}$$

2. Model error

$p\left(\mathbf{x}^{0:\tau}\right)$

Consider the two **limiting cases**:



2. Model error

$$\mathbf{x}^t = m^{(t-1)\to t}\left(\mathbf{x}^{t-1}\right) + \mathbf{v}^t$$

In general: $cov(\mathbf{v}^i, \mathbf{v}^j) = \phi\left(|i-j|, \omega\right)\mathbf{Q}$

e.g. $\phi\left(|i-j|, \omega\right) = e^{-\frac{|i-j|}{\omega}}$

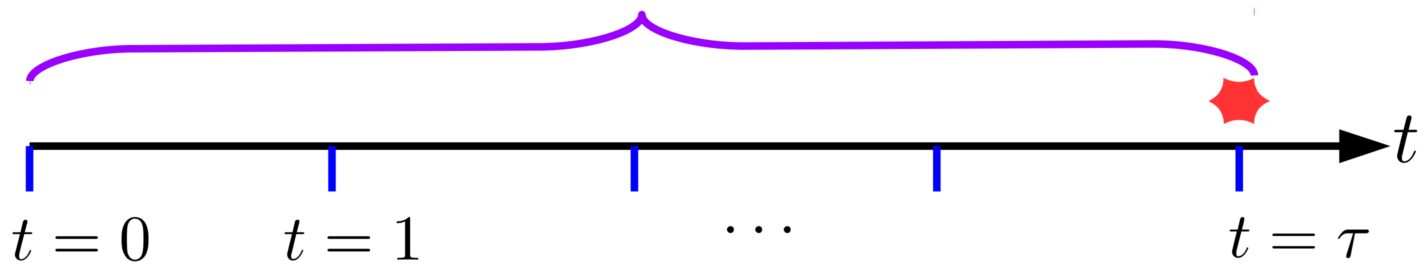
What happens if we use a **wrong time-scale** in DA?

Forecast model vs. **real (imperfect) model**

ω_g ω

More fundamentally, what is the **effect of the memory of the error?**

Simpler problem: Kalman Smoother



Initial conditions: $\mathbf{x}^0 \sim N\left(\mu_x^{0,b}, \mathrm{B}\right)$

Dynamics: $\mathbf{x}^t = \mathrm{M}^{(t-1)\rightarrow t}\mathbf{x}^{t-1} + \mathbf{v}^t \rightarrow \mathbf{v}^t \sim N\left(\mathbf{0}, \mathbf{Q}\right)$

$cov\left(\mathbf{v}^i, \mathbf{v}^j\right) = \phi\left(|i-j|, \omega\right)\mathbf{Q}$

Observation: $\mathbf{y}^l = \mathbf{H}^l\mathbf{x}^t + \boldsymbol{\eta}^l \rightarrow \boldsymbol{\eta}^l \sim N\left(\mathbf{0}, \mathbf{R}\right)$

The WC solution to the KS

Write this as an **extended** problem. $\mathbf{z}^{0:\tau} = \left[(\mathbf{x}^0)^\mathbf{T}, (\mathbf{v}^{1:\tau})^\mathbf{T}\right]^\mathbf{T}$

$$\mathbf{M}^{0:\tau} = \left[\mathbf{M}^{0\to\tau}, \mathbf{M}^{1\to\tau}, \mathbf{M}^{2\to\tau}, \cdots, \mathbf{M}^{(\tau-1)\to\tau}, \mathbf{I}\right]$$

At the time of the **observation**: $\mathbf{x}^\tau = \mathbf{M}^{0:\tau}\mathbf{z}^{0:\tau}$

The **analysis**: $\underline{\mathbf{z}^{0:\tau}|\mathbf{y} \sim N\left(\boldsymbol{\mu}_z^{0:\tau,a}, \mathbf{A}_z^{0:\tau}\right)}$

With **moments**:

$$\boldsymbol{\mu}_z^{0:\tau,a} = \left(\mathbf{I} - \mathbf{K}_z^{0:\tau}\mathbf{H}\mathbf{M}^{0:\tau}\right)\boldsymbol{\mu}_z^{0:\tau,b} + \mathbf{K}_z^{0:\tau}\mathbf{y}$$

$$\mathbf{A}_z^{0:\tau} = \left(\mathbf{I} - \mathbf{K}_z^{0:\tau}\mathbf{H}\mathbf{M}^{0:\tau}\right)\mathbf{D}^{0:\tau}$$

Extended **background/model error covariance**:

$$\mathbf{D}^{0:\tau} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{1:\tau} \end{bmatrix}$$

Where does the time-scale appear?

The **gain**:

$$\mathrm{K}_z^{0:\tau} = \mathrm{D}^{0:\tau}\left(\mathrm{M}^{0:\tau}\right)^{\mathbf{T}}\mathrm{H}^{\mathbf{T}}\left(\Gamma^{\tau}\right)^{-1}$$

Note this is a block-matrix, with elements corresponding to the **different time steps**

The **total covariance** (acts as 'denominator' in the gain):

$$\Gamma^{\tau} = \mathrm{H}\mathrm{B}^{\tau}\mathrm{H}^{\mathbf{T}} + \mathrm{H}\Lambda^{\tau}\mathrm{H}^{\mathbf{T}} + \mathrm{R} \longrightarrow \begin{cases} \mathrm{B}^{\tau} = \mathrm{M}^{0\to\tau}\mathrm{B}\left(\mathrm{M}^{0\to\tau}\right)^{\mathbf{T}} \\ \Lambda^{\tau} = \mathrm{M}^{1:\tau}\mathrm{Q}^{1:\tau}\left(\mathrm{M}^{1:\tau}\right)^{\mathbf{T}} \end{cases}$$

Contribution of **model error** to **total covariance** depends on **memory**:

$$\Lambda^{\tau} = \sum_{i=1}^{\tau}\sum_{j=1}^{\tau}\mathrm{M}^{i\to\tau}\mathrm{Q}(\mathrm{M}^{j\to\tau})^{\mathbf{T}}\phi\left(|i-j|,\omega\right)$$

$$\begin{cases} \lim\limits_{\omega\to 0}\Lambda^{\tau} = \sum_{j=1}^{\tau}\mathrm{M}^{j\to\tau}\mathrm{Q}(\mathrm{M}^{j\to\tau})^{\mathbf{T}}, \\ \lim\limits_{\omega\to\infty}\Lambda^{\tau} = \tilde{\mathrm{M}}^{\tau}\mathrm{Q}\left(\tilde{\mathrm{M}}^{\tau}\right)^{\mathbf{T}} \end{cases}$$

Impacts at different times:

The **gain** for **initial conditions**:

$$\mathbf{K}_x^0 = \mathbf{B}\left(\mathbf{M}^{0\to\tau}\right)^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}\left(\boldsymbol{\Gamma}^{\tau}\right)^{-1}$$

The **gain** for **model-error jumps**:

$$\mathbf{K}_v^j = \mathbf{Q}\left(\sum_{i=1}^{\tau}\left(\mathbf{M}^{i\to\tau}\right)^{\mathbf{T}}\phi(|i-j|,\omega)\right)\mathbf{H}^{\mathbf{T}}\left(\boldsymbol{\Gamma}^{\tau}\right)^{-1}
\left\{\begin{array}{l}
\lim\limits_{\omega\to 0}\mathbf{K}_v^j = \mathbf{Q}\left(\mathbf{M}^{j\to\tau}\right)^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}\left(\boldsymbol{\Gamma}^{\tau}\right)^{-1}, \\[2ex]
\lim\limits_{\omega\to\infty}\mathbf{K}_v^j = \left(\tilde{\mathbf{M}}^{\tau}\right)^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}\left(\boldsymbol{\Gamma}^{\tau}\right)^{-1}
\end{array}\right.$$

The gain for the **state variables at different times**:

$$\mathbf{K}_x^t = \mathbf{M}^{0\to t}\mathbf{K}_x^0 + \sum_{j=1}^{t}\mathbf{M}^{j\to t}\mathbf{K}_v^j$$

This is a **composed** gain.

4. Unidimensional example

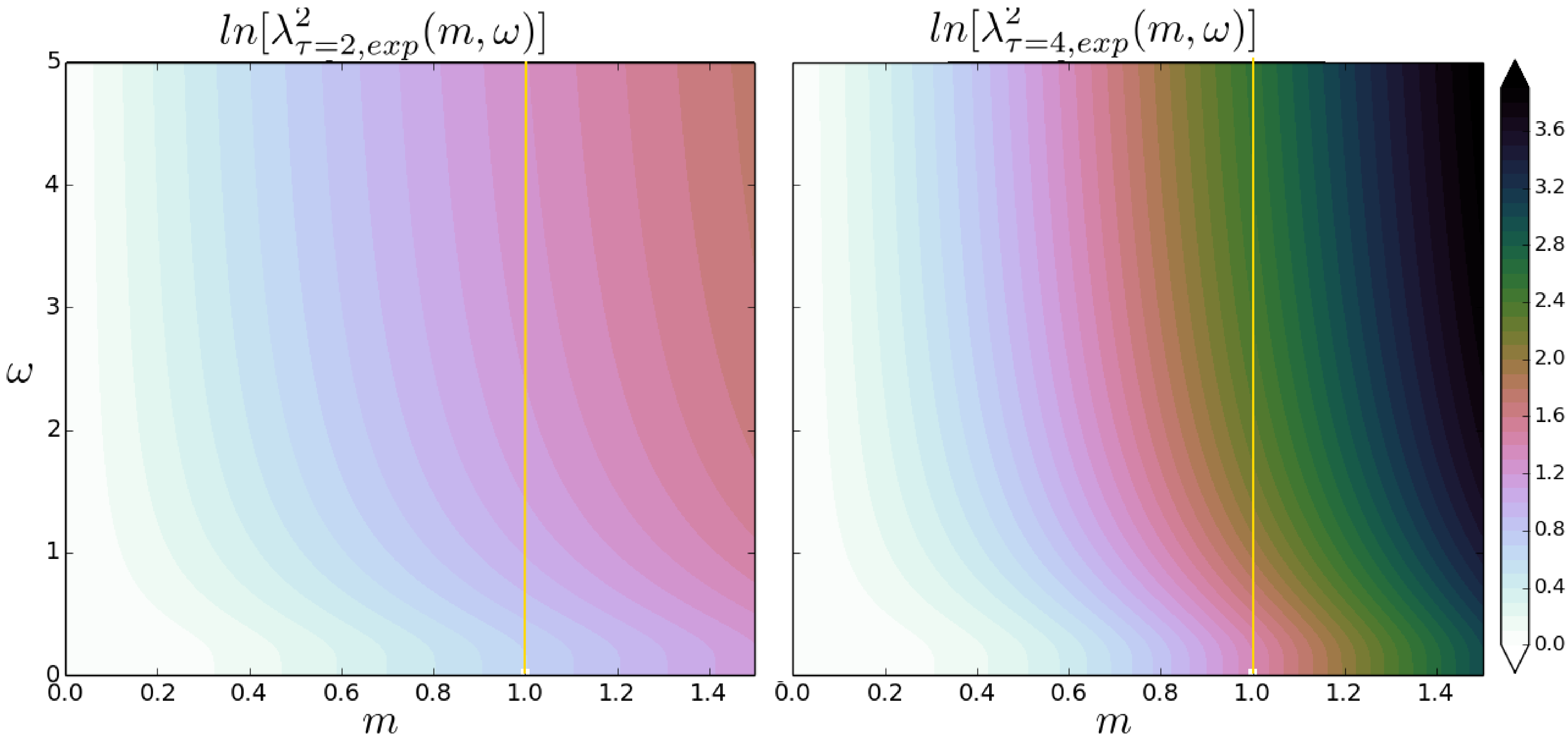
In a simple **uni-dimensional case**, it is easy to visualize the behaviour of

- **model error contribution** to the **total covariance**
- the **gains** at different time steps.

This scalars are plotted as a function of the **model and the time-autocorrelation scale.**

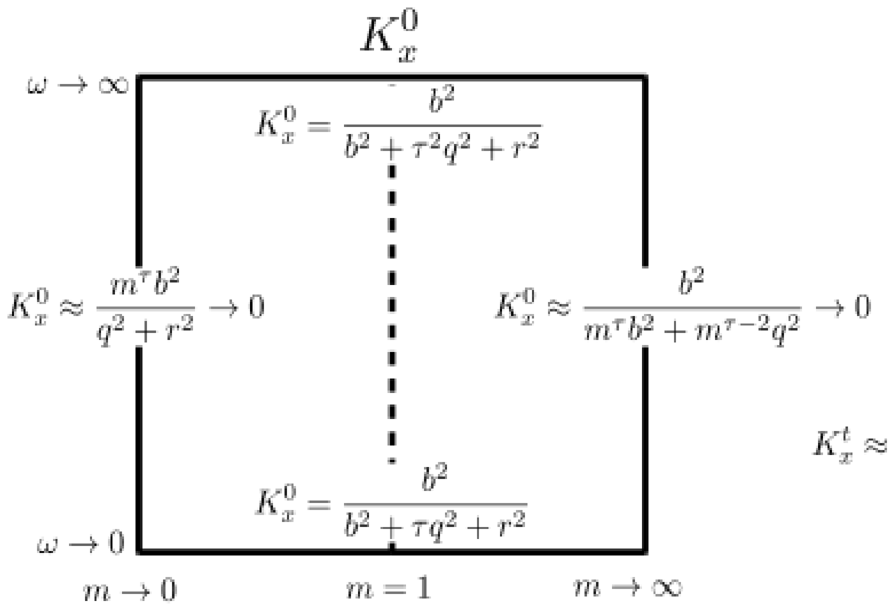
Unidimensional example

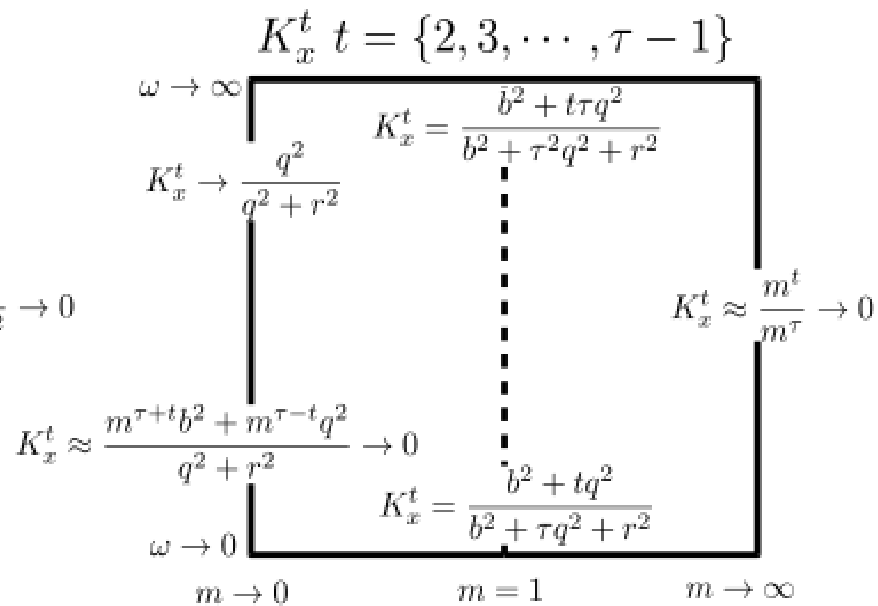
Contribution from model error to the total error covariance.

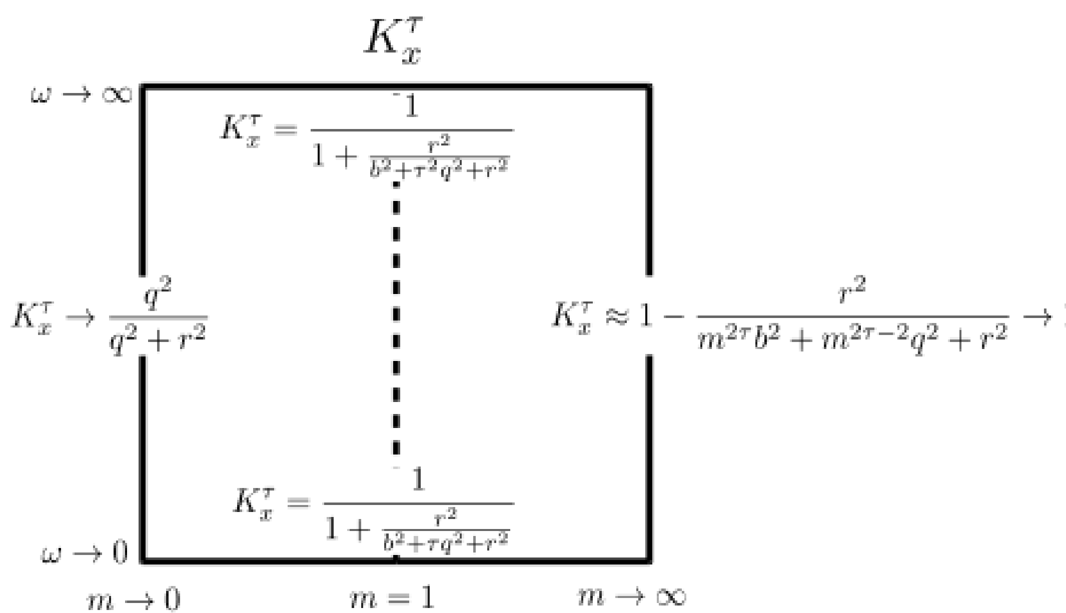


Gains

$$\mathbf{K}_x^t = \mathbf{M}^{0\to t}\mathbf{K}_x^0 + \sum_{j=1}^{t} \mathbf{M}^{j\to t}\mathbf{K}_v^j$$

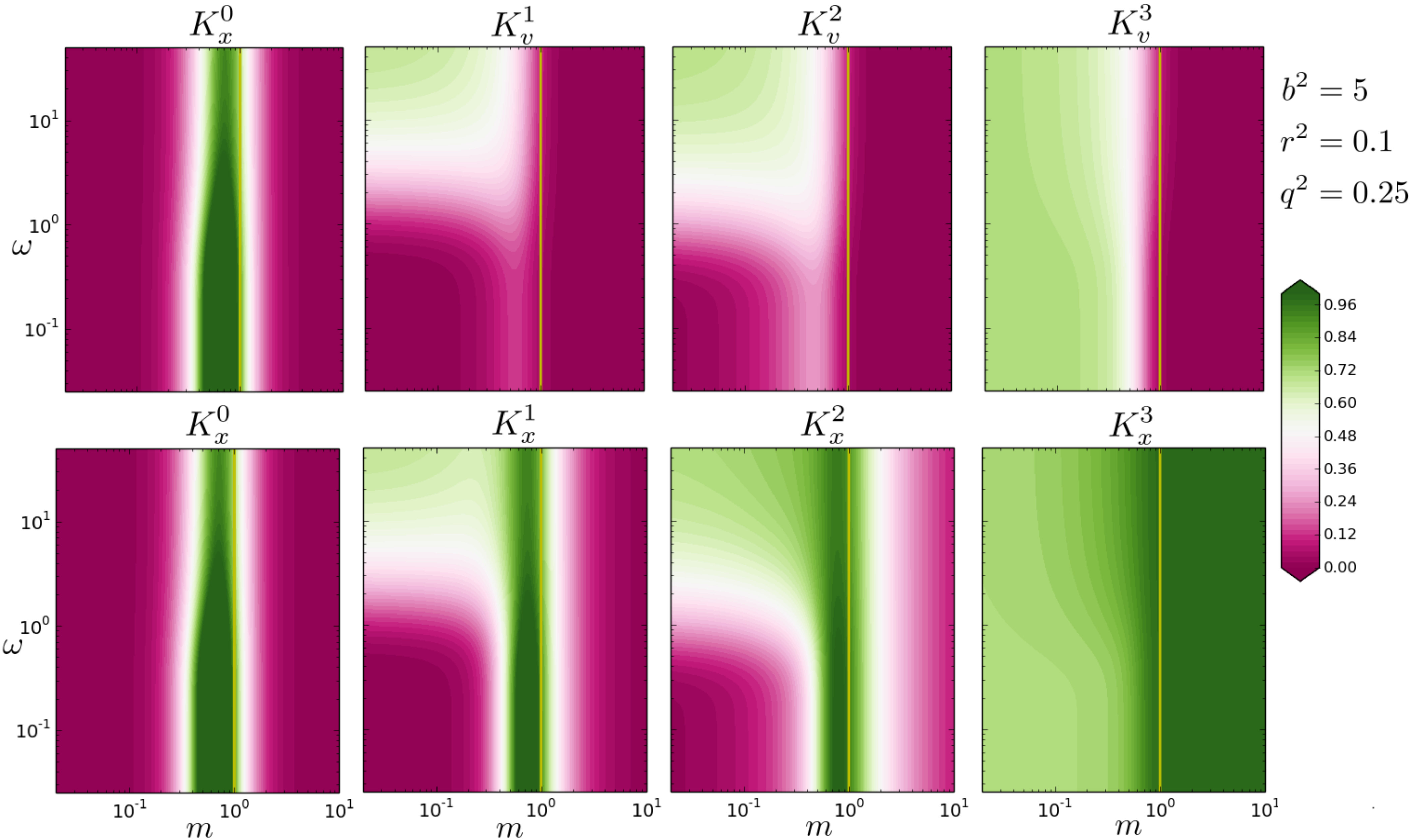






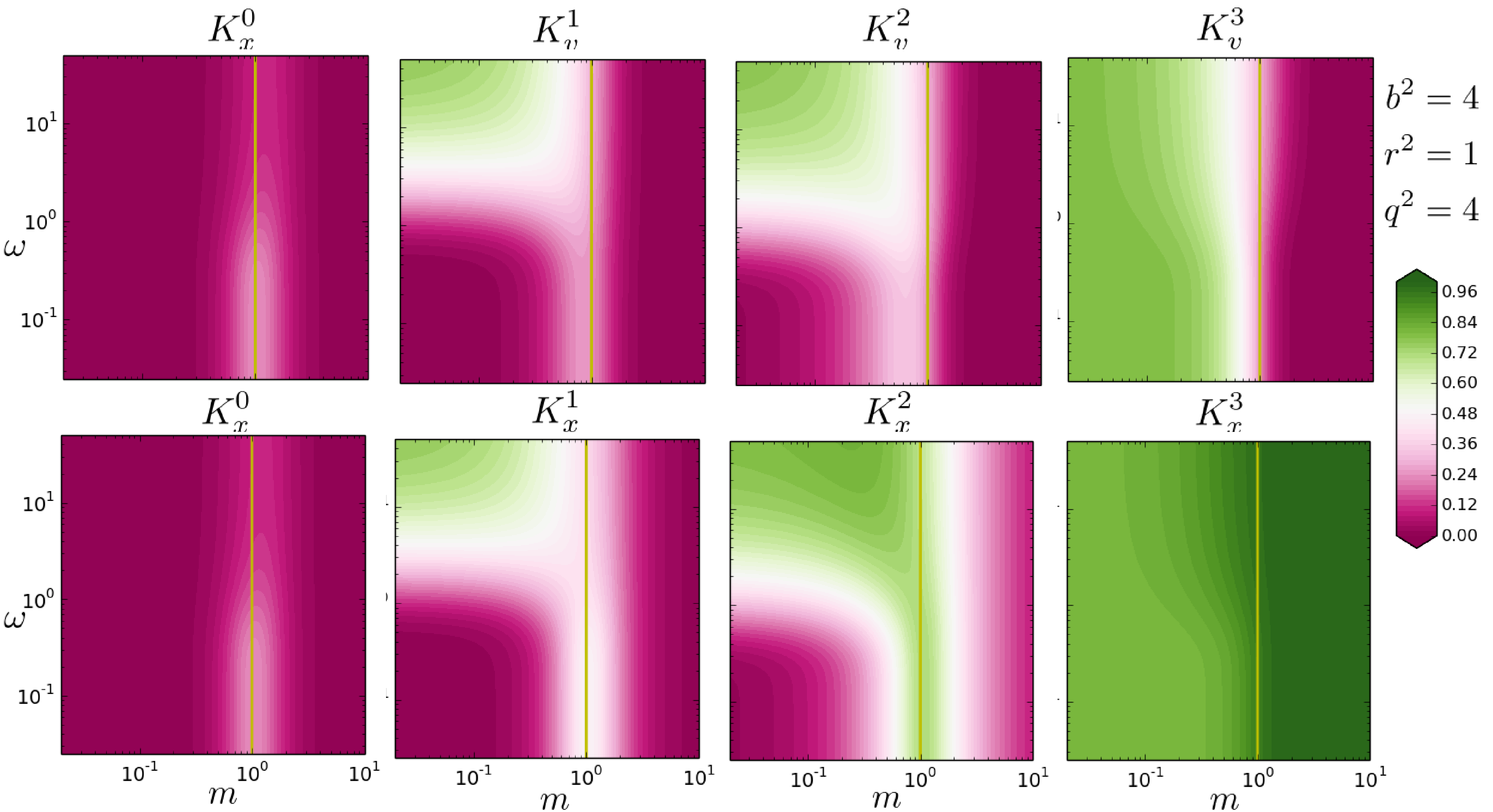
Gains

$$\mathbf{K}_x^t = \mathbf{M}^{0\to t}\mathbf{K}_x^0 + \sum_{j=1}^{t} \mathbf{M}^{j\to t}\mathbf{K}_v^j$$



Gains

$$\mathbf{K}_x^t = \mathrm{M}^{0\to t}\mathbf{K}_x^0 + \sum_{j=1}^{t} \mathrm{M}^{j\to t}\mathbf{K}_v^j$$



5. Introducing errors

The extended background covariance:

$$\mathbf{D}_e^{0:\tau} = \begin{bmatrix} \underset{\uparrow}{\overset{\text{sampling}}{\mathbf{B}_e}} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_g^{1:\tau} \end{bmatrix}$$

Wrong prescription

The total effect of (**indirect**) **sampling errors and a wrong prescription** of the time-scale?

5. Introducing errors

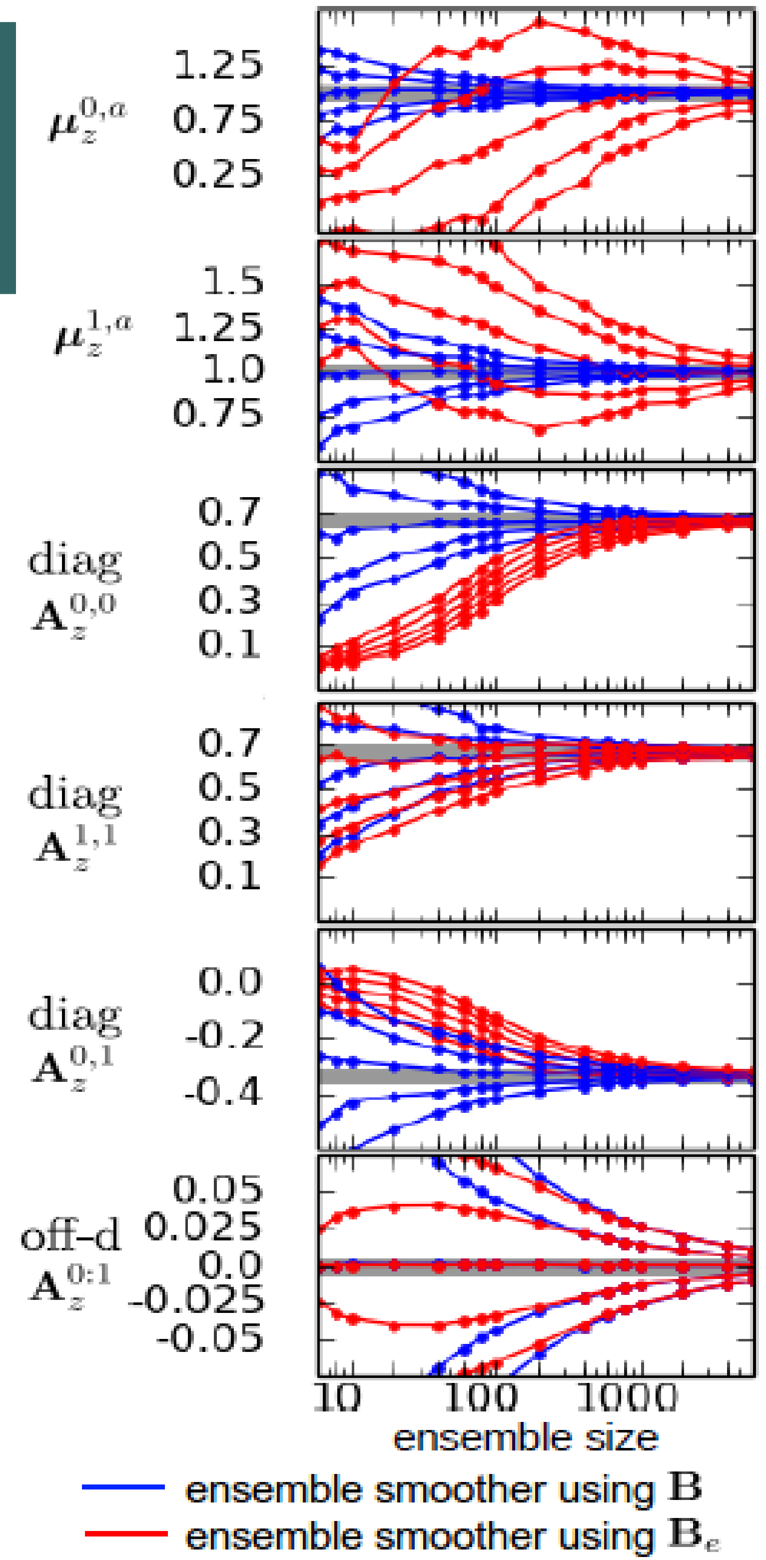
The extended background covariance:

sampling

$$\mathbf{D}_e^{0:\tau} = \begin{bmatrix} \mathbf{B}_e & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_g^{1:\tau} \end{bmatrix}$$

Wrong prescription

The total effect of (**indirect**) **sampling errors and a wrong prescription** of the time-scale?



The error-ridden matrices

$$\mathbf{D}_{ge}^{0:\tau}, \mathbf{\Gamma}_{ge}^{\tau}, \mathbf{K}_{ge}^{0:\tau},$$

Once there is an error in the **background error** matrix, this impacts the **total covariance** and the **gain**.

Solution for each analysis member:

$$\underline{\mathbf{z}_{g,n_e}^{0:\tau,a} = \left(\mu_z^{0:\tau,b} + \mathbf{K}_{z,ge}^{0:\tau}\mathbf{d}\right) + \left(\zeta_{g,n_e}^{0:\tau,b} + \mathbf{K}_{z,ge}^{0:\tau}\delta_{g,n_e}\right)}$$

mean perturbation

The exact expressions for the 'ge' expressions are:

$$\mathbf{\Gamma}_{ge}^{\tau} = \mathbf{\Gamma}^{\tau} + \mathbf{\Gamma}_{g\epsilon}^{\tau} = \left(\mathbf{I} + \mathbf{\Gamma}_{g\epsilon}^{\tau}\left(\mathbf{\Gamma}^{\tau}\right)^{-1}\right)\mathbf{\Gamma}^{\tau}$$

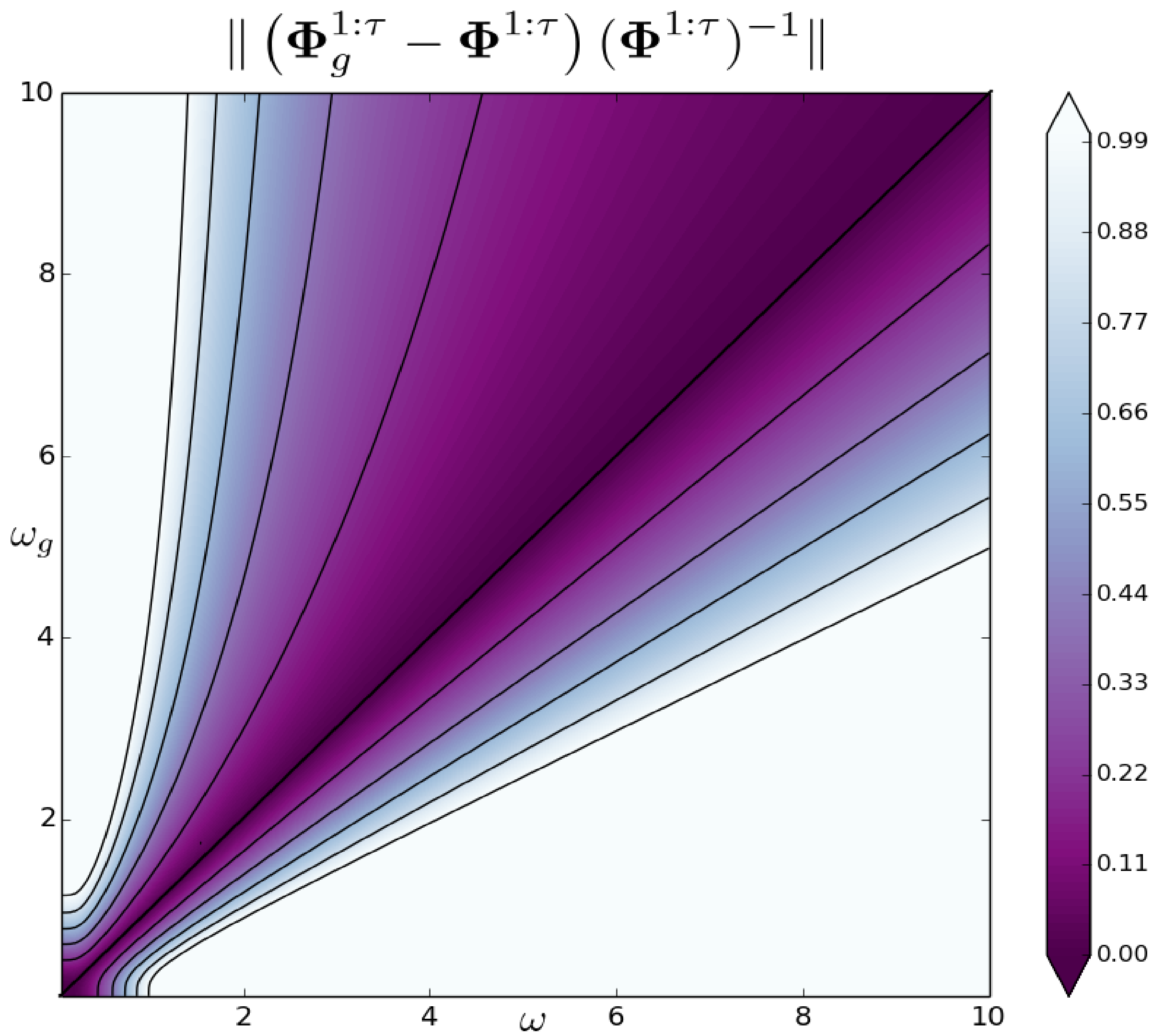
$$\underline{\mathbf{K}_{z,ge}^{0:\tau} = \left(\mathbf{I} + \epsilon_D^{0:\tau}\left(\mathbf{D}^{0:\tau}\right)^{-1}\right)\mathbf{K}_z^{0:\tau}\left(\mathbf{I} + \mathbf{\Gamma}_{g\epsilon}^{\tau}\left(\mathbf{\Gamma}^{\tau}\right)^{-1}\right)^{-1}}$$

6. Small error approximation

$$\|\epsilon_D^{0:\tau}\left(\mathrm{D}^{0:\tau}\right)^{-1}\| \ll 1,$$

Sampling error is 'small'.

Prescription error is 'small'.



How does the **error in the memory affect the norm**?

Small error approximation

The inverse can be linearised.

$$\mathbf{K}^{0:\tau}_{z,ge} \approx \left(\mathbf{I} + \epsilon^{0:\tau}_{D} \left(\mathbf{D}^{0:\tau}\right)^{-1}\right) \mathbf{K}^{0:\tau}_{z} \left(\mathbf{I} - \mathbf{\Gamma}^{\tau}_{g\epsilon} \left(\mathbf{\Gamma}^{\tau}\right)^{-1}\right)$$

And then things become additive:

$$\underline{\mathbf{K}^{0:\tau}_{z,ge} \approx \mathbf{K}^{0:\tau}_{z} + \mathbf{K}^{0:\tau}_{z,g\epsilon}}$$

To first order, the ensemble-based wrong-memory gain is the **sum** of the **correct gain** and an **error part.**

$$\underline{\mathbf{K}^{0:\tau}_{z,g\epsilon} = \epsilon^{0:\tau}_{D} \left(\mathbf{D}^{0:\tau}\right)^{-1} \mathbf{K}^{0:\tau}_{z} - \mathbf{K}^{0:\tau}_{z} \mathbf{\Gamma}^{\tau}_{g\epsilon} \left(\mathbf{\Gamma}^{\tau}\right)^{-1}}$$

Small error approximation

The inverse can be linearised.

$$\mathbf{K}^{0:\tau}_{z,ge} \approx \left(\mathbf{I} + \epsilon^{0:\tau}_D \left(\mathbf{D}^{0:\tau}\right)^{-1}\right) \mathbf{K}^{0:\tau}_z \left(\mathbf{I} - \mathbf{\Gamma}^{\tau}_{g\epsilon} \left(\mathbf{\Gamma}^{\tau}\right)^{-1}\right)$$

And then things become additive:

$$\underline{\mathbf{K}^{0:\tau}_{z,ge} \approx \mathbf{K}^{0:\tau}_z + \mathbf{K}^{0:\tau}_{z,g\epsilon}}$$

To first order, the ensemble-based wrong-memory gain is the **sum** of the **correct gain** and an **error part.**

$$\underline{\mathbf{K}^{0:\tau}_{z,g\epsilon} = \epsilon^{0:\tau}_D \left(\mathbf{D}^{0:\tau}\right)^{-1} \mathbf{K}^{0:\tau}_z - \mathbf{K}^{0:\tau}_z \mathbf{\Gamma}^{\tau}_{g\epsilon} \left(\mathbf{\Gamma}^{\tau}\right)^{-1}}$$

The components of the **sampling estimators.**

	exact part	direct sampling error	indirect errors error (linear)	indirect sampling error (non-linear)
$\bar{\mathbf{z}}^{0:\tau,a}_{ge}$	$\mu^{0:\tau}_z$	$+\bar{\zeta}^{0:\tau,a}_g$	$+\mathbf{K}^{0:\tau}_{z,g\epsilon}\mathbf{d}$	$+\mathbf{K}^{0:\tau}_{z,g\epsilon}\bar{\delta}_g$
$\mathbf{A}^{0:\tau}_{z,ge}$	$\mathbf{A}^{0:\tau}_z$	$+\left(\mathbf{I} - \mathbf{K}^{0:\tau}_z \mathbf{H}\mathbf{M}^{0:\tau}\right)\epsilon^{0:\tau}_D$	$-\mathbf{K}^{0:\tau}_{z,g\epsilon}\mathbf{H}\mathbf{M}^{0:\tau}\mathbf{D}^{0:\tau}$	$-\left(\mathbf{H}\mathbf{M}^{0:\tau}\epsilon^{0:\tau}_D\right)^{\mathbf{T}}\left(\mathbf{\Gamma}^{\tau}\right)^{-1}\mathbf{H}\mathbf{M}^{0:\tau}\epsilon^{0:\tau}_D$

7. Conclusions, future work

- We have explored the effect of time **auto-correlated model error** in the **Kalman Smoother**.

- The impact of the observations over a window depend strongly upon the magnitude of the model. For **shrinking models**, the **magnitude of the 'ememory' matters considerably**.

- We have explored the errors created by: **sampling estimators of the background error covariance**, and the **error in prescribing the model error covariance**. Their interactions are not trivial, and they affect both mean and covariance.

- This can be useful when one has to use **fixed (or slow-evolving) model errors for computational reasons**.

- What happens with **cycling** and with **non-linear models**? (current MSc/PhD Student)

Shameless advertising

Two professor-level position (grade 9) in Reading:

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Met Office

University of Reading