Time structure of model error in data assimilation

Javier Amezcua
Peter Jan van Leeuwen
Haonan Ren

May 2019

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

Outline

1. **Some basics of data assimilation**

2. **Time-structure in model error**

3. **Model error and sources**

4. **Summary and current work**

Outline

1. Some basics of data assimilation

2. Time-structure in model error

3. Model error and sources

4. Summary and current work

DA problem setup

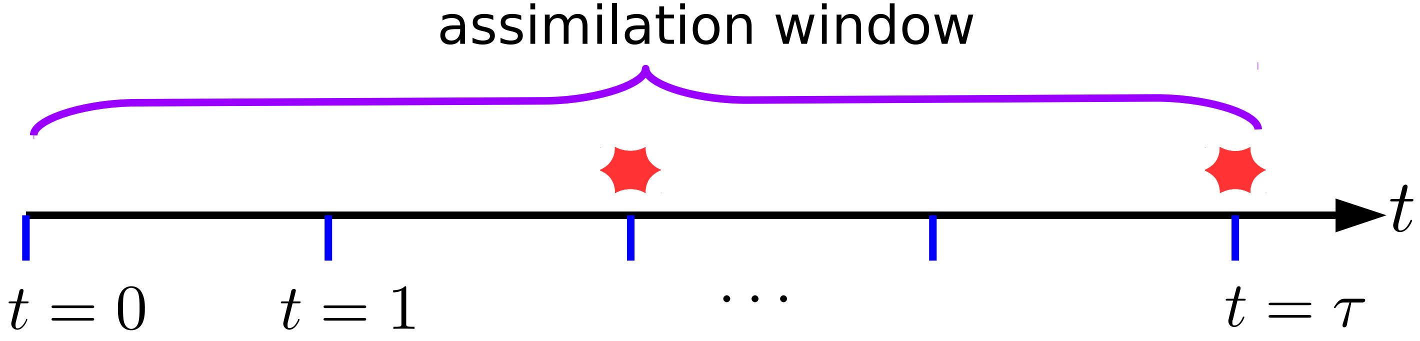
$\mathbf{x}^t \in \mathcal{R}^{N_x}$ **Model variables**

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

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

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

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

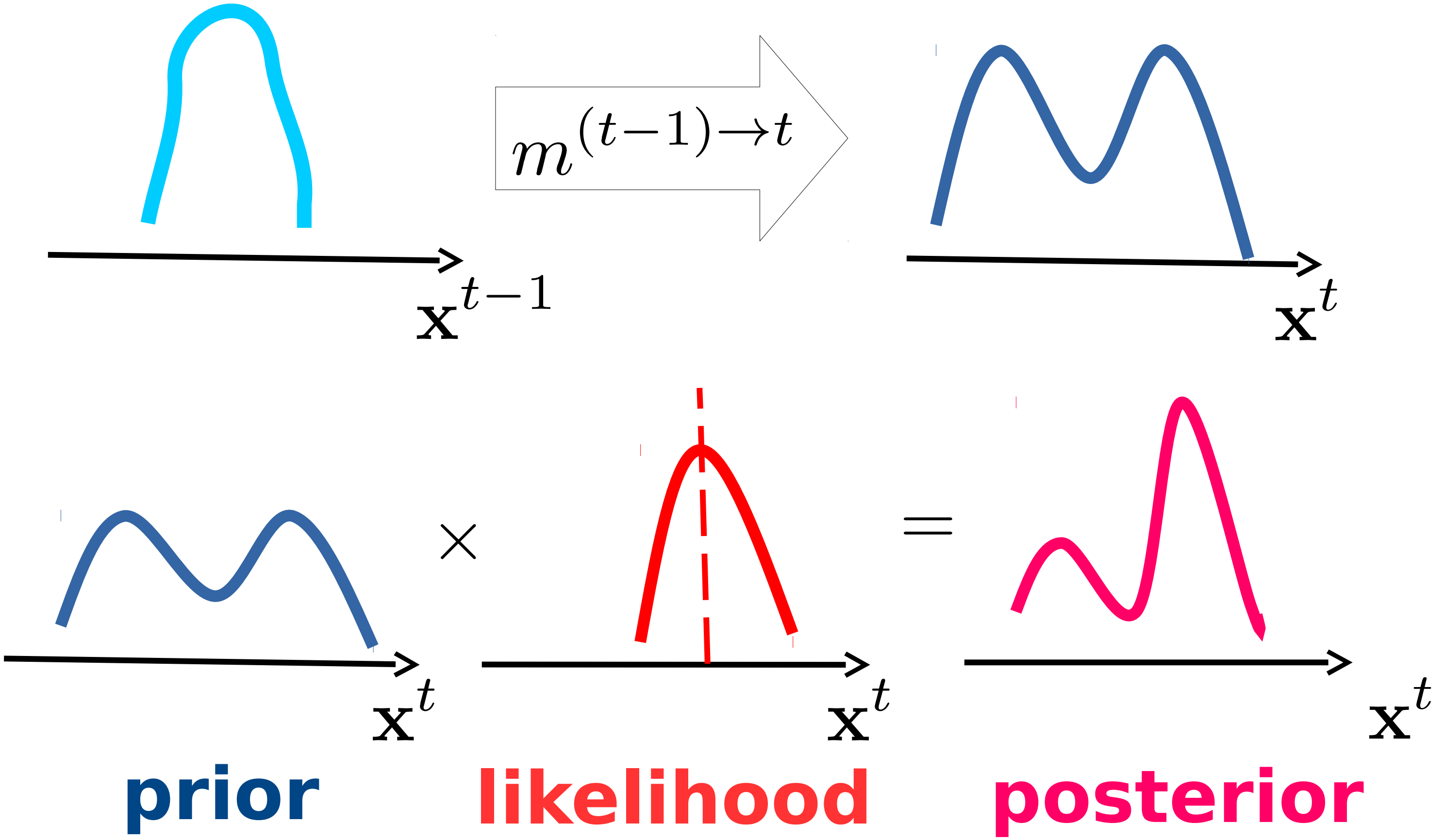


Working with pdf's

Consider the following 1-step scenario:

Forecast:

Analysis:



Bayes theorem

Likelihood of the observations over the assimilation window.

Prior distribution of the state variable over the assimilation window.

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

Marginal distribution of the **observations**.

Posterior probability distribution of the state variables given the observations over the assimilation window.

How to get these elements?

Outline

1. Some basics of data assimilation

2. Time-structure in model error

3. Model error and sources

4. Summary and current work

KF smoothing problem setup

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

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

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

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

For the model error:

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

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

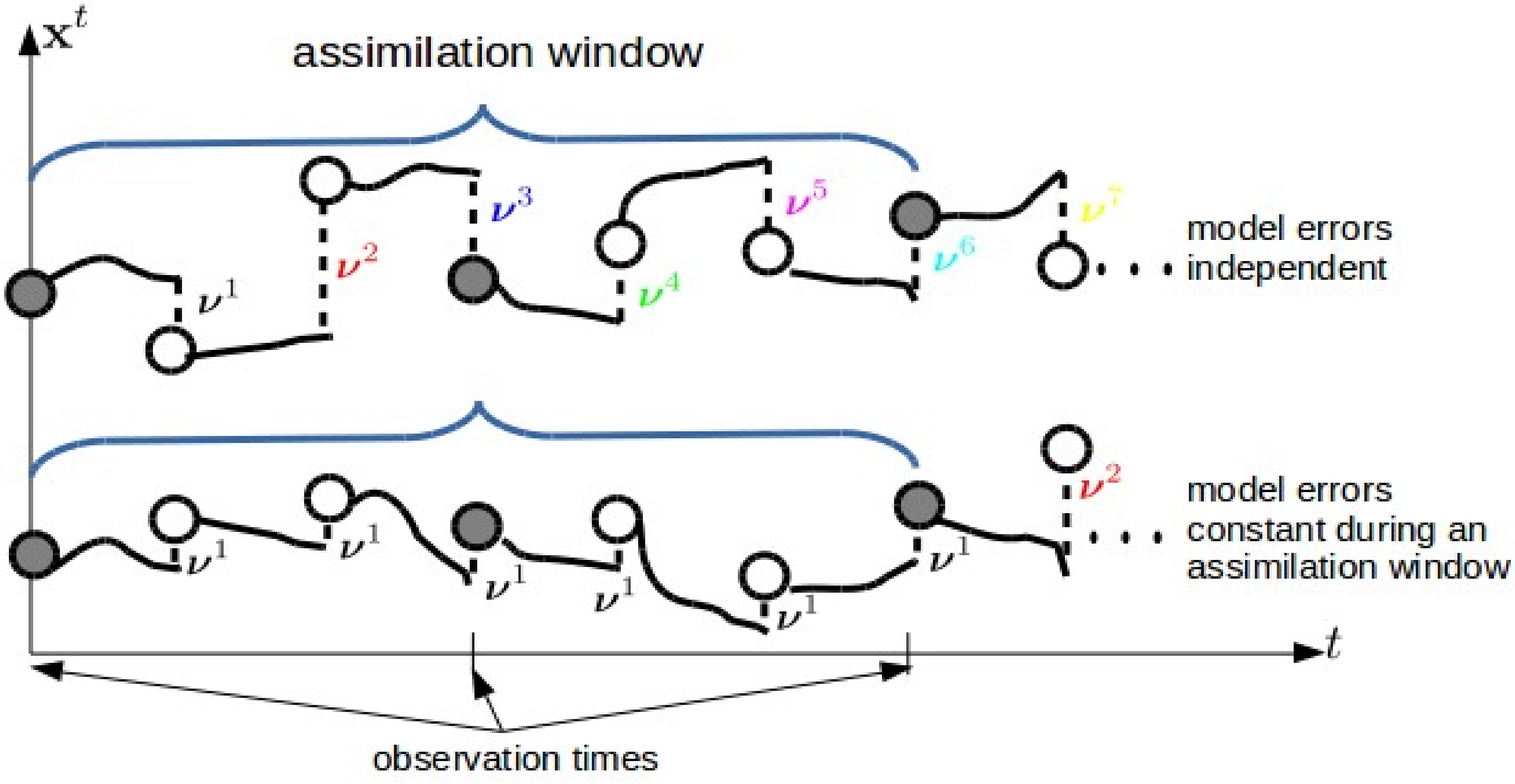
e.g. $$\phi\left(|i-j|, \omega, k\right) = e^{-\frac{|i-j|^k}{\omega}}$$

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

Model error

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

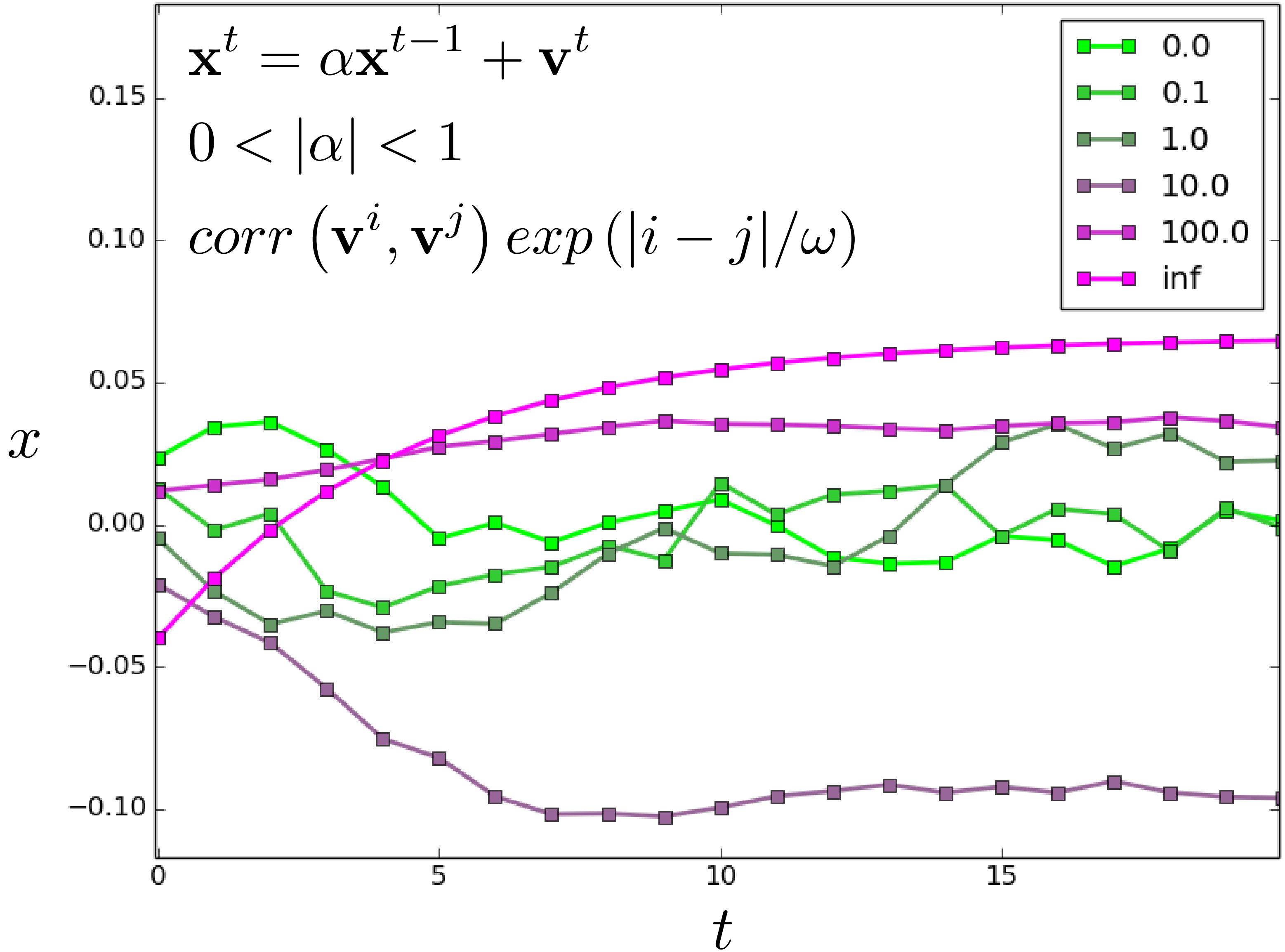
Example: two limiting cases.



$\phi = 0, \omega = 0$

$\phi = 1, \omega \to \infty$

AR1 process with noise



The WC solution to the KS

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

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

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

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

Extended **background/model error covariance**:

With **moments**:

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

$$\mathbf{A}_z^{0:\tau} = \left(\mathbf{I} - \mathbf{K}_z^{0:\tau}\mathbf{H}\mathbf{M}^{0:\tau}\right)\mathbf{D}^{0:\tau} \longrightarrow \mathbf{D}^{0:\tau} = \begin{bmatrix}\mathbf{B} & 0 \\ 0 & \mathbf{Q}^{1:\tau}\end{bmatrix}$$

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

Reminder

The fundamental way in which the Kalman Filter introduce effects from observations is through the following analysis equation.

$$\mathbf{x}^a = \left(\mathbf{I} - \mathbf{KH}\right)\mathbf{x}^b + \mathbf{Ky}$$

analysis background observation

With gain:

$$\mathbf{K} = \mathbf{BH^T}\left(\mathbf{HBH^T} + \mathbf{R}\right)^{-1}$$

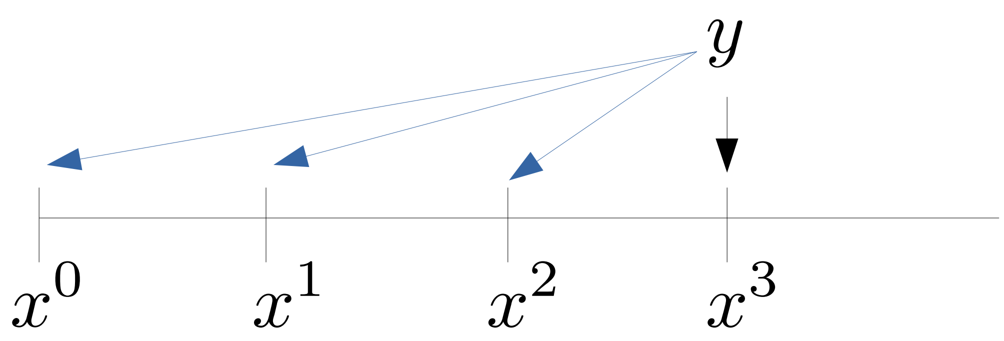
In the Kalman smoother the model error covariance would also appear.

Unidimensional example

In a simple **uni-dimensional case**, it is easy to visualize the behaviour of the **gains** at different time steps.

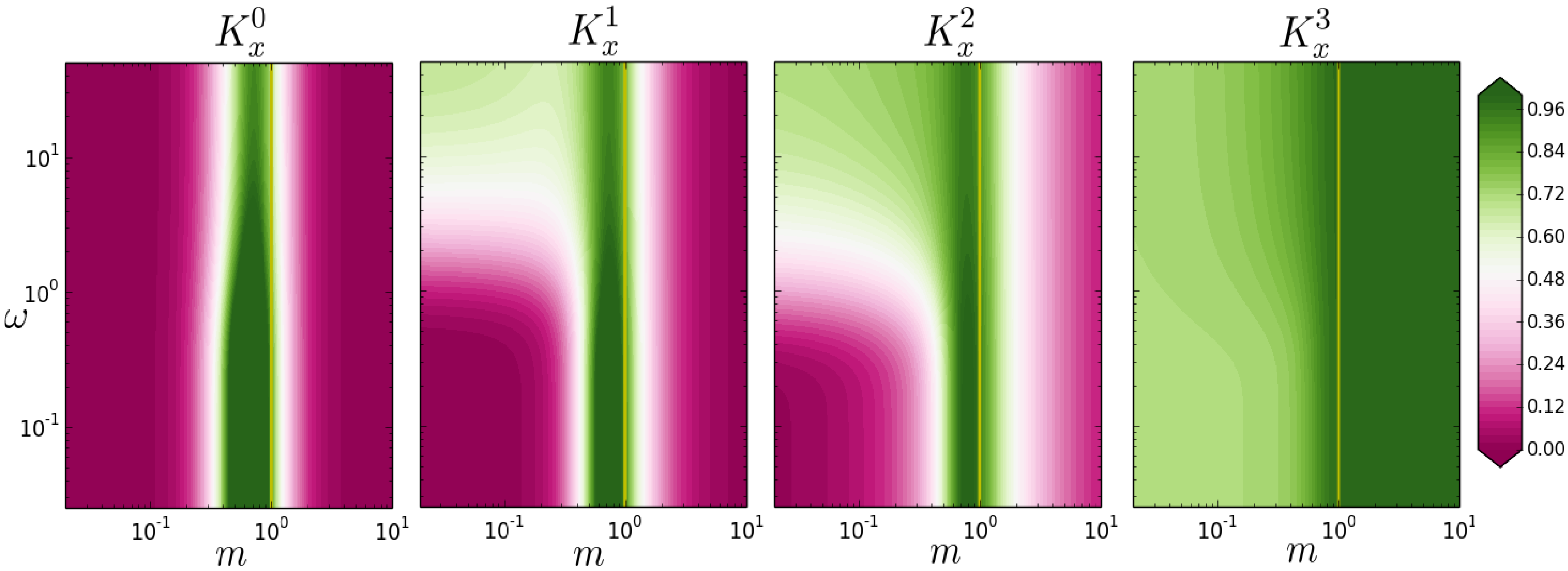
$$\mathbf{x}^a = (\mathbf{I} - \underline{\mathbf{KH}})\mathbf{x}^b + \underline{\mathbf{K}}\mathbf{y}$$

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



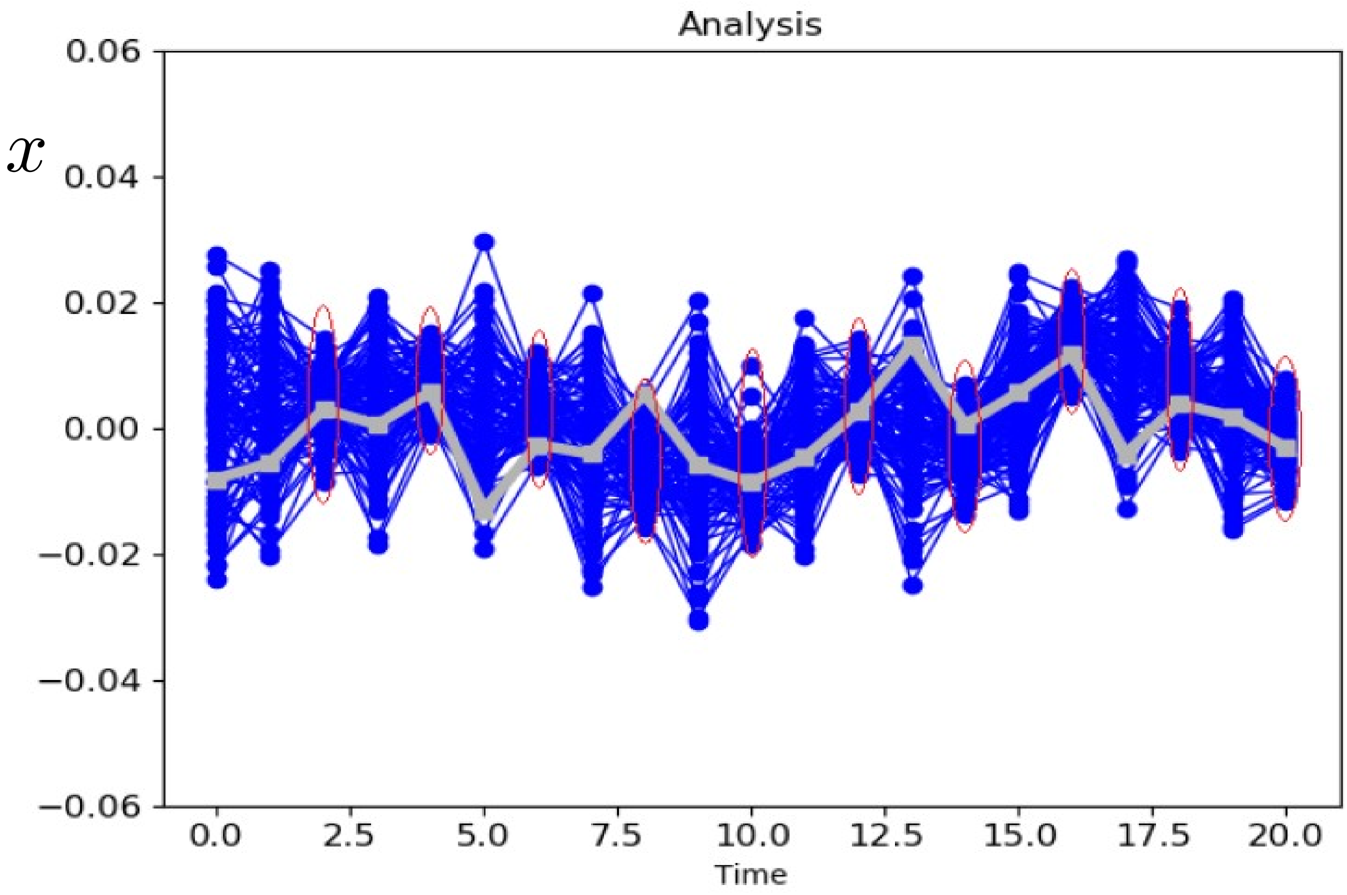
Gains

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



Green regions is where observation dominates the analysis, magenta regions is where background dominates the analysis.

Cycling experiments



DA experiments with cycling using a Stochastic EnKF with large ensemble size.

20-time steps assimilation windows.

Observations every
- 2 steps
- 10 steps

Computed the overall (time-averaged) ratio:

$$\mathrm{ratRS} = \frac{\mathrm{RMSE}_a}{\mathrm{spread}_a}$$

For a healthy DA system this is close to 1.

Cycling experiments

Observations every 2 steps.

	$\omega_g = 0.0$	$\omega_g = 0.1$	$\omega_g = 0.5$	$\omega_g = 1.0$	$\omega_g = 2.0$	$\omega_g = 10$
$\omega = 0.0$	1.0000	0.943089	0.9000	0.8112	0.8333	0.8431
$\omega = 0.1$	0.9675	0.9672	0.8846	0.8252	0.7815	0.8431
$\omega = 0.5$	1.0082	1.0569	0.9538	0.8741	0.8819	0.8881
$\omega = 1.0$	1.1229	1.1148	1.0615	0.9763	0.9583	0.9673
$\omega = 2.0$	1.1545	1.1557	1.0154	0.9792	0.9812	0.9539
$\omega = 10$	1.1382	1.1544	1.0923	0.9650	0.9375	0.9743

Observations every 10 steps.

	$\omega_g = 0.0$	$\omega_g = 0.1$	$\omega_g = 0.5$	$\omega_g = 1.0$	$\omega_g = 2.0$	$\omega_g = 10$
$\omega = 0.0$	0.9848	0.9697	0.9846	1.0667	1.2037	0.9722
$\omega = 0.1$	0.9552	0.9701	0.9692	1.0667	1.1818	0.9583
$\omega = 0.5$	0.9403	0.9254	0.9846	1.0500	1.0500	0.9306
$\omega = 1.0$	0.8806	0.8657	0.9063	0.9667	1.0926	0.9444
$\omega = 2.0$	0.8484	0.8235	0.8308	0.8833	0.9744	0.9306
$\omega = 10$	0.8806	0.8955	0.8769	0.9500	1.0741	0.9722

RMSEa < SPREAD_a

RMSEa = SPREAD_a

RMSEa > SPREAD_a

The **observational frequency has an effect**. Why? Haonan.

Outline

1. Some basics of data assimilation

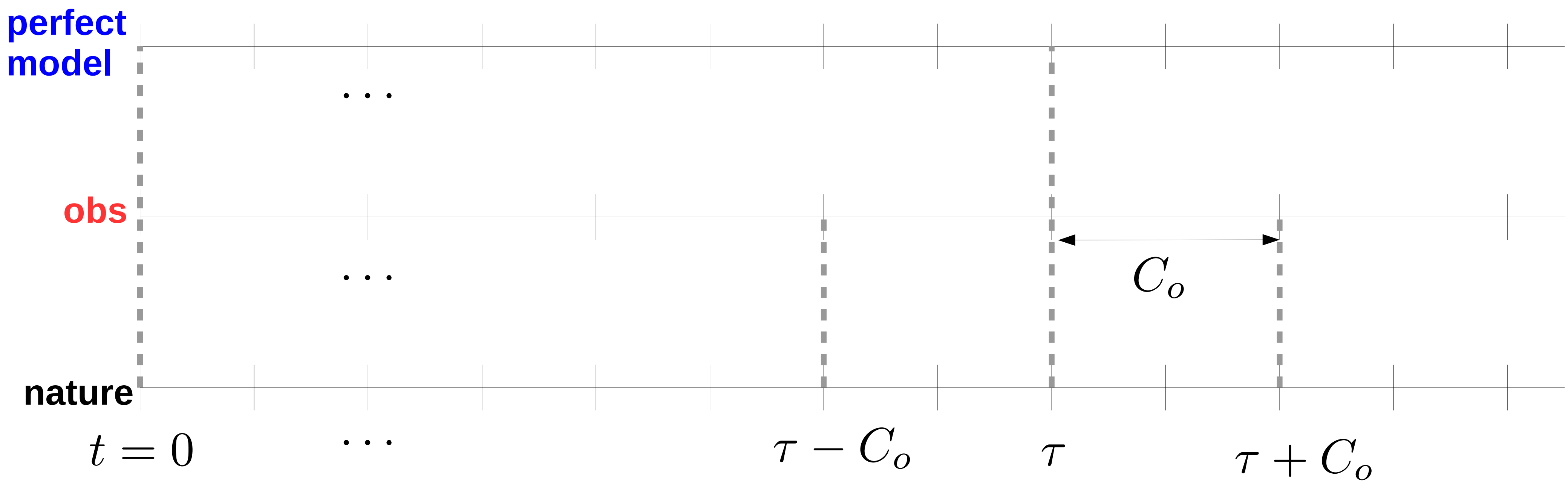
2. Time-structure in model error

3. Model error and sources

4. Summary and current work

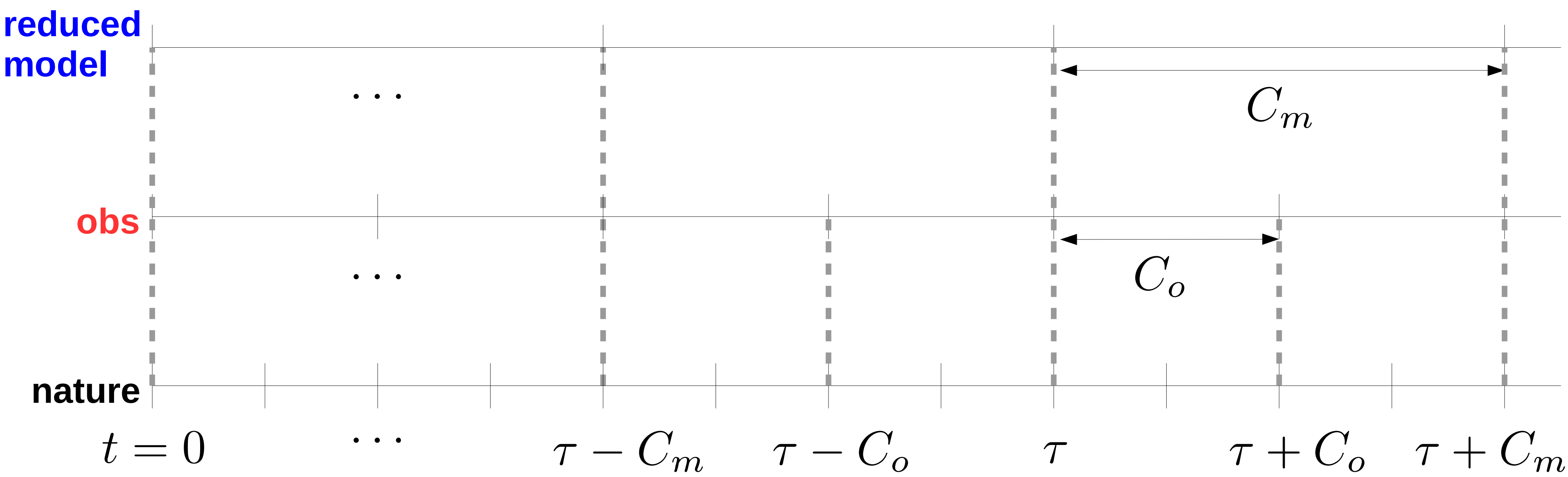
Representing all scales

(a) A **perfect model** resolving **all the temporal scales**. $\mathbf{v}^t = \mathbf{0} \ \forall \ t$



Reduced model

(b) A **reduced model** resolving only **slow temporal scales**. $\mathbf{v}^t \neq \mathbf{0}$



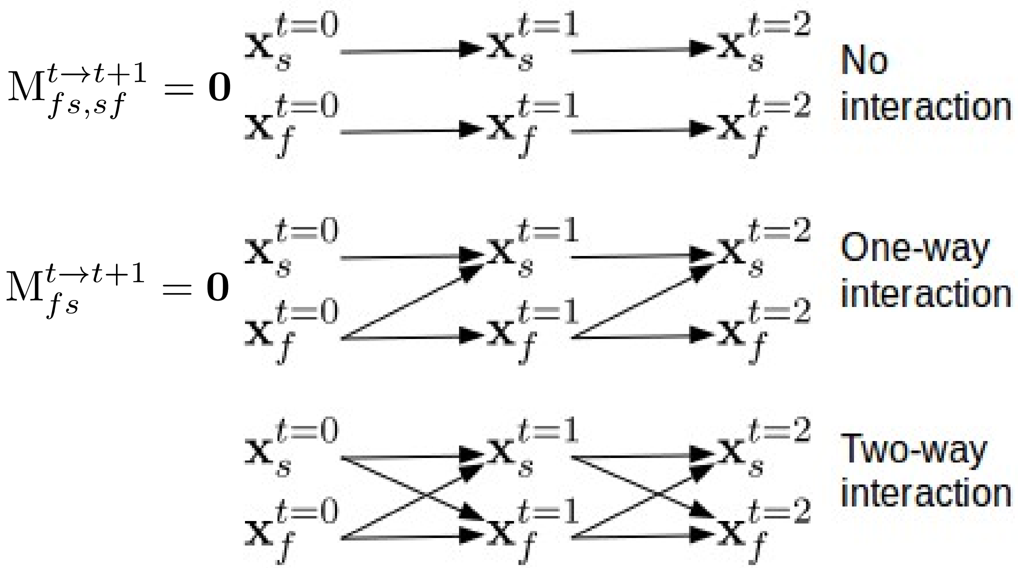
Let's take a look at how model error results from the **interaction** between scales.

Two-scale linear model

$$\mathbf{x}^{t+1} = \mathbf{M}^{t\to t+1}\mathbf{x}^{t}$$

Consider we can partition the state variable into slow and fast components:

$$\mathbf{x}^{t} = \begin{bmatrix}\mathbf{x}_s^t \\ \mathbf{x}_f^t\end{bmatrix} \qquad \mathbf{M}^{t\to t+1} = \begin{bmatrix}\mathbf{M}_{ss}^{t\to t+1} & \mathbf{M}_{sf}^{t\to t+1} \\ \mathbf{M}_{fs}^{t\to t+1} & \mathbf{M}_{ff}^{t\to t+1}\end{bmatrix}$$



Understanding the components

Once we cycle the model we get:

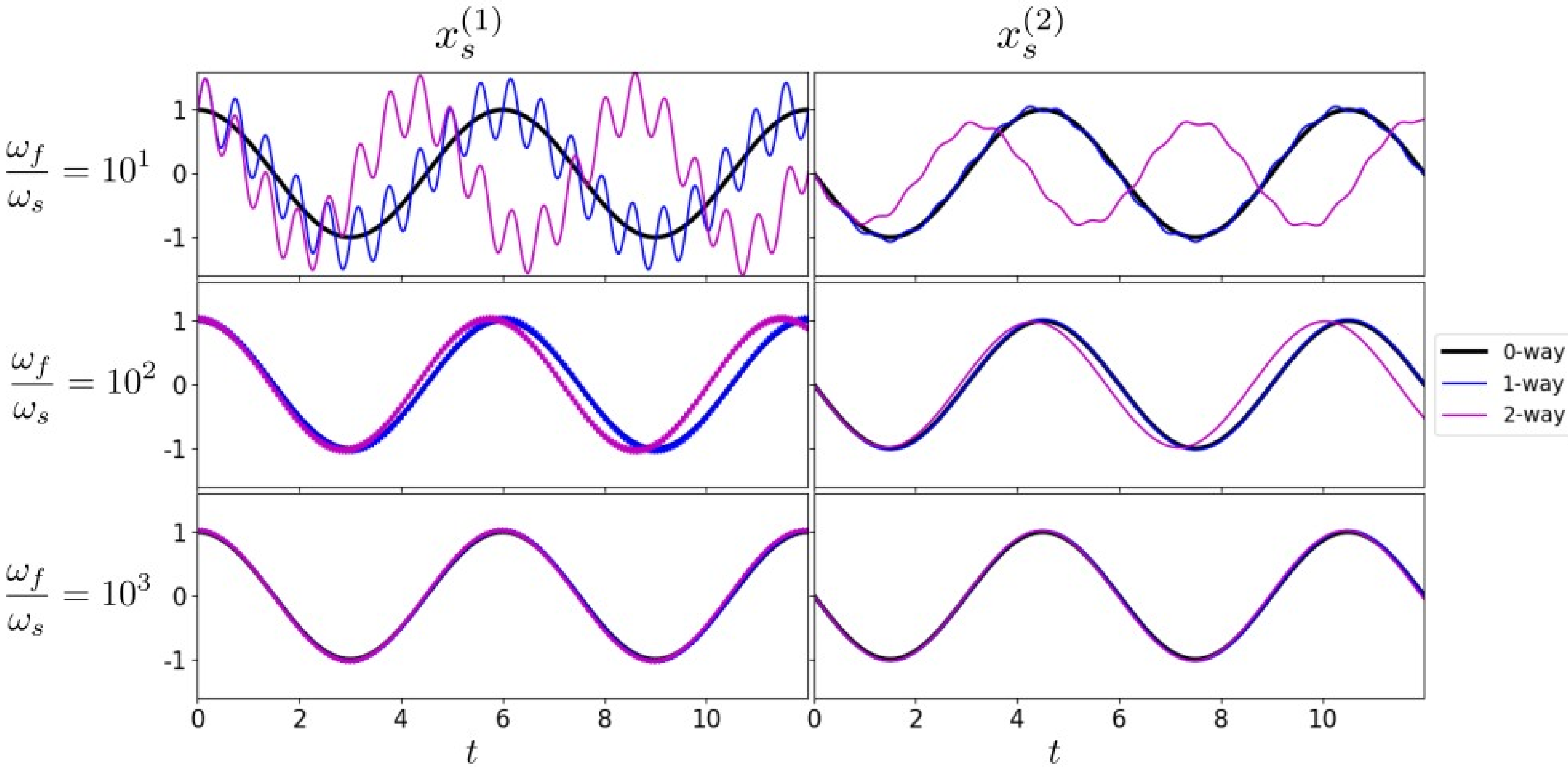
$$\begin{bmatrix} \mathbf{x}_s^t \\ \mathbf{x}_f^t \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{ss}^{0\to t} & \mathbf{M}_{fs}^{0\to t} \\ \mathbf{M}_{sf}^{0\to t} & \mathbf{M}_{ff}^{0\to t} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_s^0 \\ \mathbf{x}_f^0 \end{bmatrix}$$

Independent $\mathrm{M}^{t\to t+1} = \begin{bmatrix} \mathbf{M}_{ss}^{t\to t+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{ff}^{t\to t+1} \end{bmatrix}$ $\mathbf{x}_s^t = \hat{\mathbf{M}}_{ss}^{0\to t}\mathbf{x}_s^0$

Fast to slow $\mathrm{M}^{t\to t+1} = \begin{bmatrix} \mathbf{M}_{ss}^{t\to t+1} & \mathbf{M}_{sf}^{t\to t+1} \\ \mathbf{0} & \mathbf{M}_{ff}^{t\to t+1} \end{bmatrix}$ $\mathbf{x}_s^t = \hat{\mathbf{M}}_{ss}^{0\to t}\mathbf{x}_s^0 + \mathbf{M}_{sf,1w}^{0\to t}\mathbf{x}_f^0$

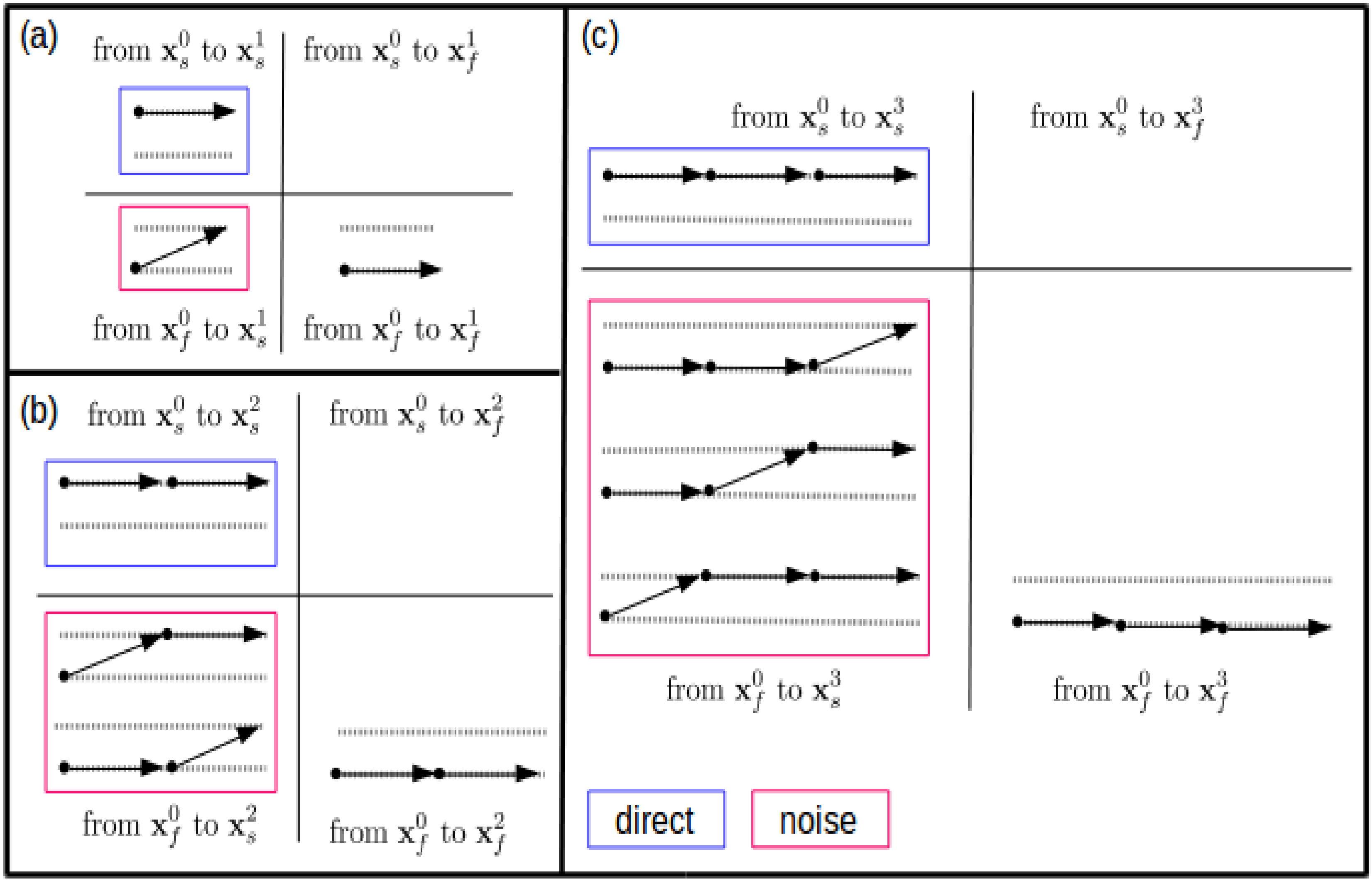
Two-way $\mathrm{M}^{t\to t+1} = \begin{bmatrix} \mathbf{M}_{ss}^{t\to t+1} & \mathbf{M}_{sf}^{t\to t+1} \\ \mathbf{M}_{fs}^{t\to t+1} & \mathbf{M}_{ff}^{t\to t+1} \end{bmatrix}$ $\mathbf{x}_s^t = \hat{\mathbf{M}}_{ss}^{0\to t}\mathbf{x}_s^0 + \tilde{\mathbf{M}}_{ss}^{0\to t}\mathbf{x}_s^0 + \mathbf{M}_{sf,1w}^{0\to t}\mathbf{x}_f^0 + \mathbf{M}_{sf,2w}^{0\to t}\mathbf{x}_f^0$

Evolution of slow variables



Fast-to-slow interaction only

$$\mathbf{x}_s^\tau = \hat{\mathbf{x}}_s^\tau\left(\mathbf{x}_s^0\right) + \mathbf{z}^\tau\left(\mathbf{x}_f^0\right)$$



Fast-to-slow interaction

Process by which noise is generated.

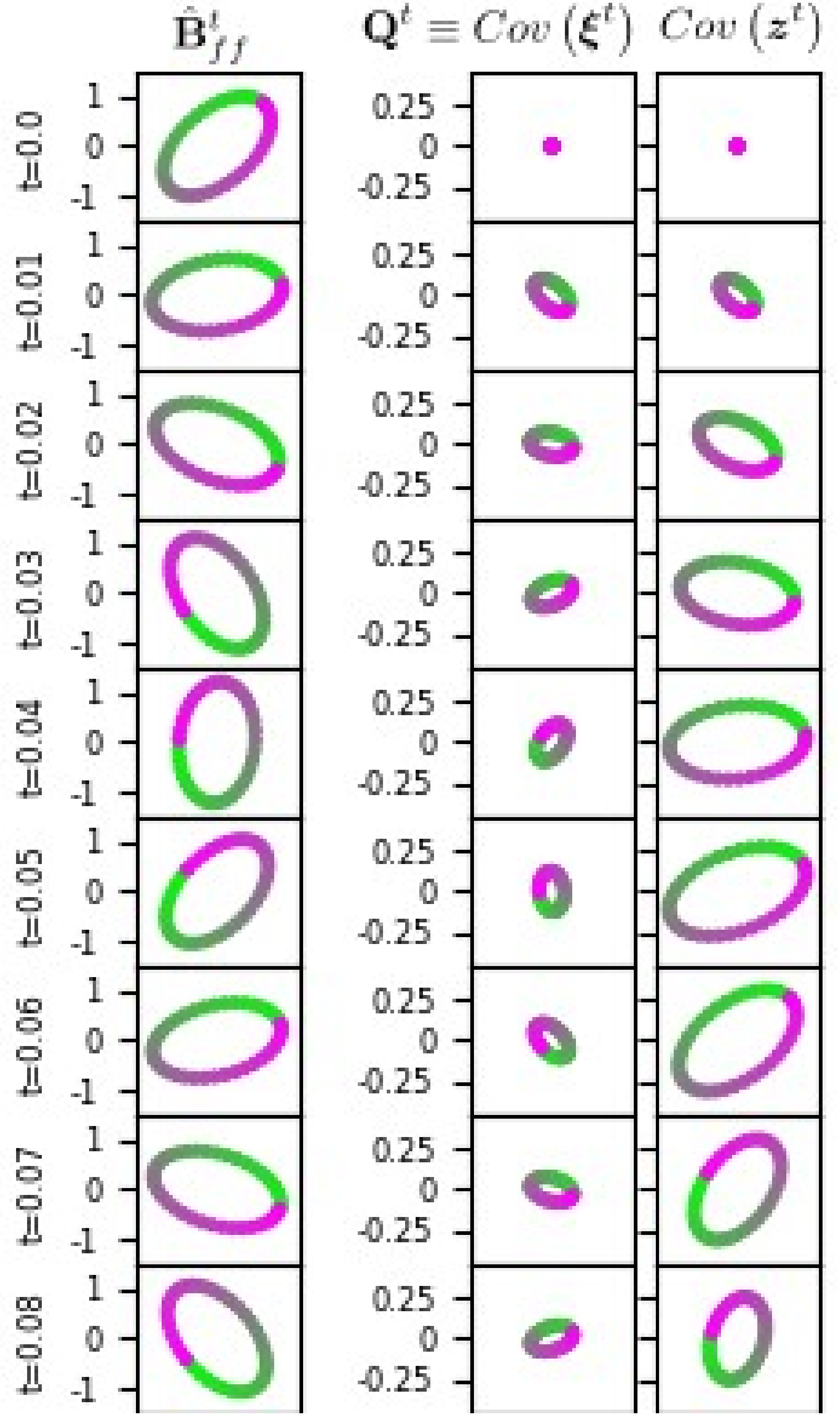
$$\mathbf{x}_s^\tau = \hat{\mathbf{x}}_s^\tau\left(\mathbf{x}_s^0\right) + \mathbf{z}^\tau\left(\mathbf{x}_f^0\right)$$

$$\mathrm{z}^{\ \tau} = \sum_{t=0}^{\tau} \mathbf{M}_{ss}^{t\to\tau}\boldsymbol{\zeta}^t$$

$$\boldsymbol{\zeta}^t = \mathbf{M}_{sf}^{t-1\to t}\mathbf{M}_{ff}^{t\to t-1}\mathbf{z}_f^0$$

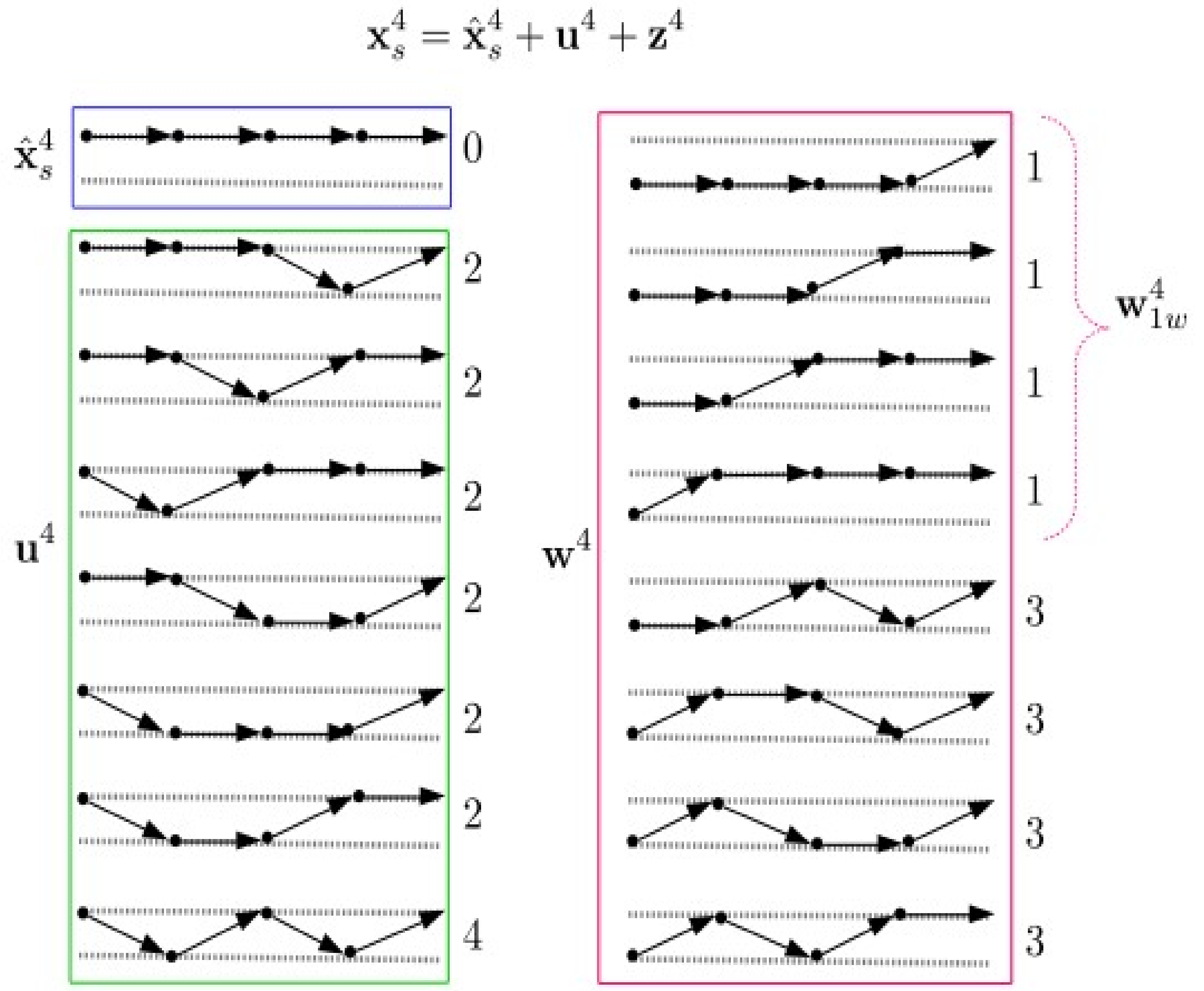
Common lag-1 representation of the dynamics.

$$\mathrm{x}_s^t = \mathbf{M}_{ss}^{t-1\to t}\mathbf{x}_s^{t-1} + \boldsymbol{\zeta}^t$$



Two-way interactions

The components of the noise are not the same as those in the in the 1-way case.



Example 2. Solve using KF

The Kalman filter operates with **mean and covariance**.

Let's consider the **evolution** of an **initial mean and an initial covariance** under the action of this map.

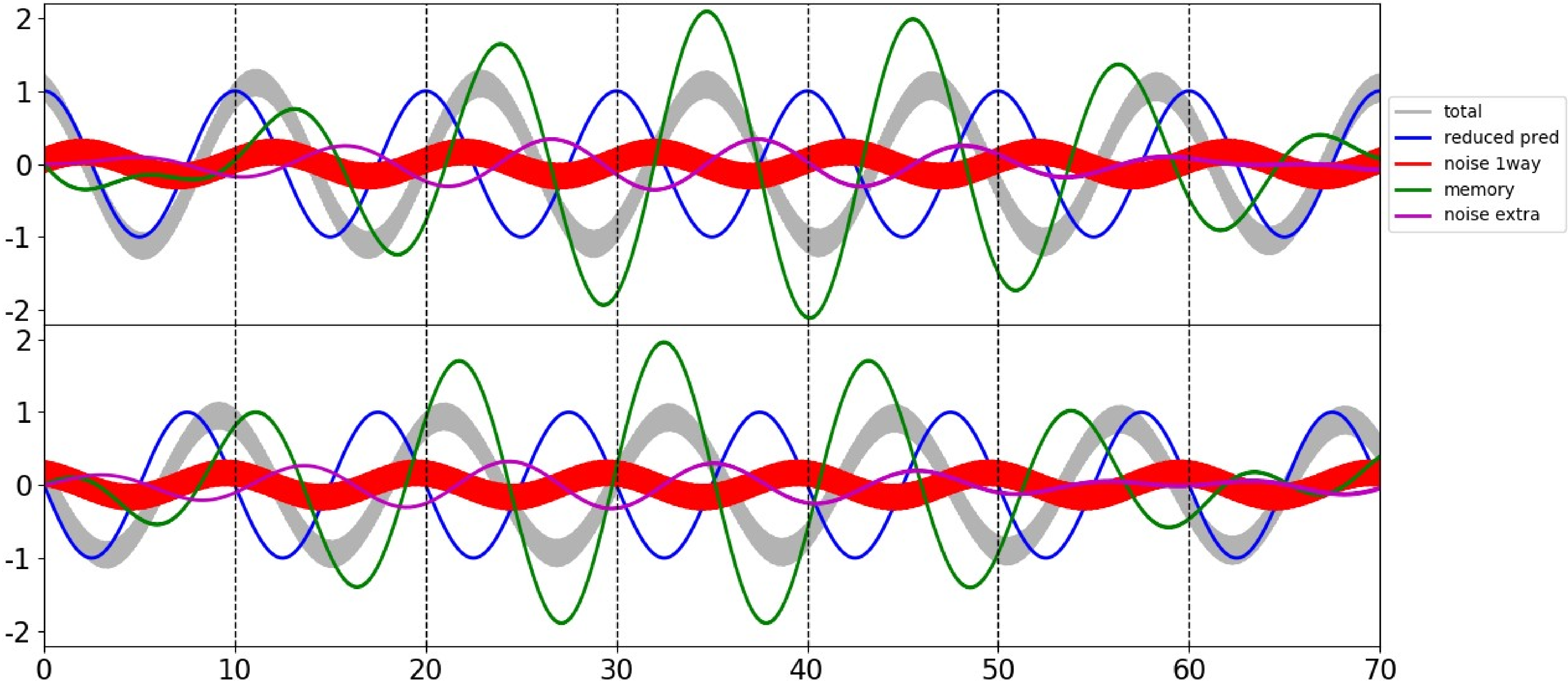
$$\bar{\mathbf{x}}^t = \mathbf{M}^{0\to t}\hat{\mathbf{x}}^0$$

$$\mathbf{B}^t = \mathbf{M}^{0\to t}\mathbf{B}^0\left(\mathbf{M}^{0\to t}\right)^{\mathbf{T}}$$

Separate the effects coming from:

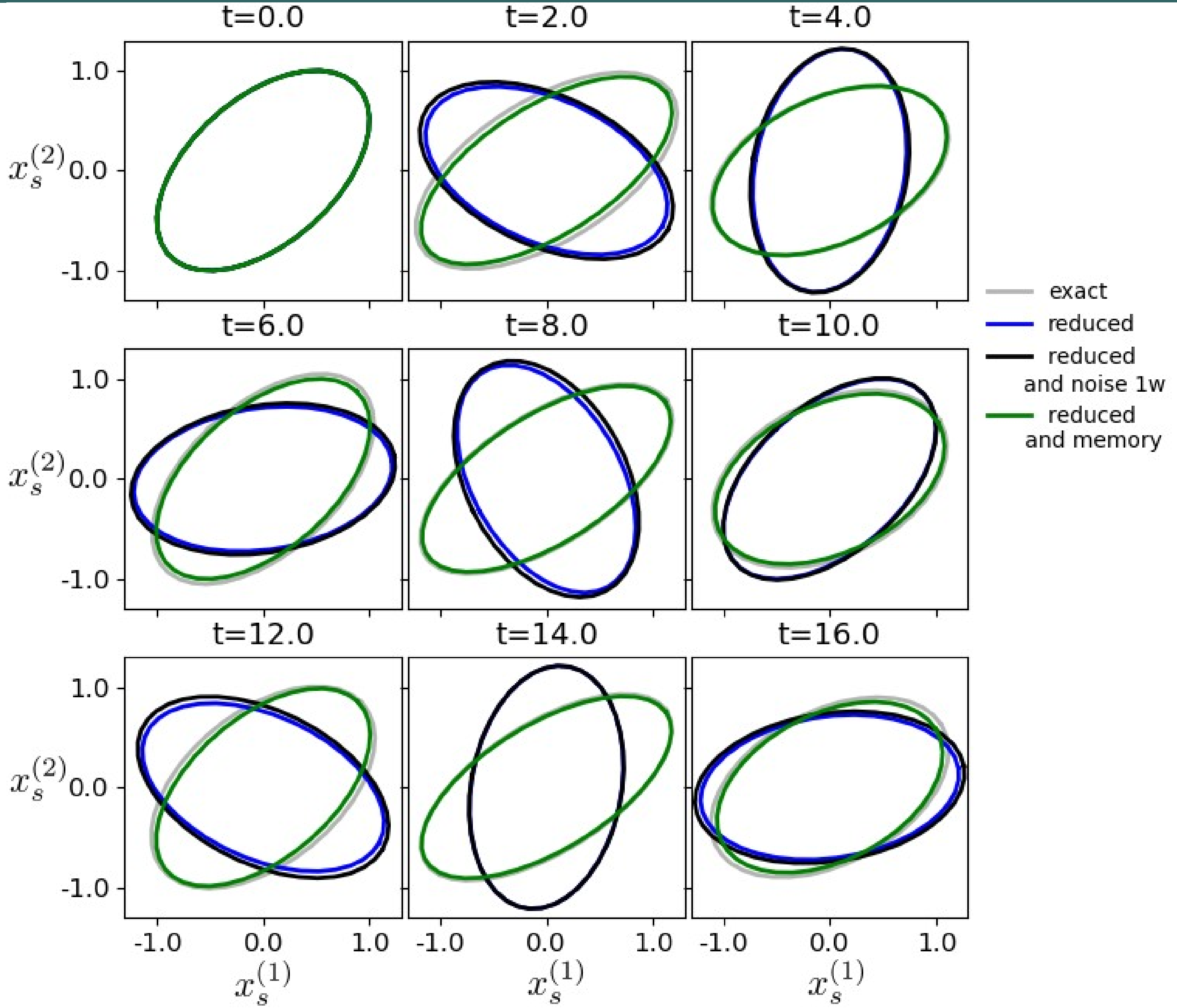
- **reduced prediction**
- **1-way noise**
- **>1 way noise**
- **memory**

The effect in the mean



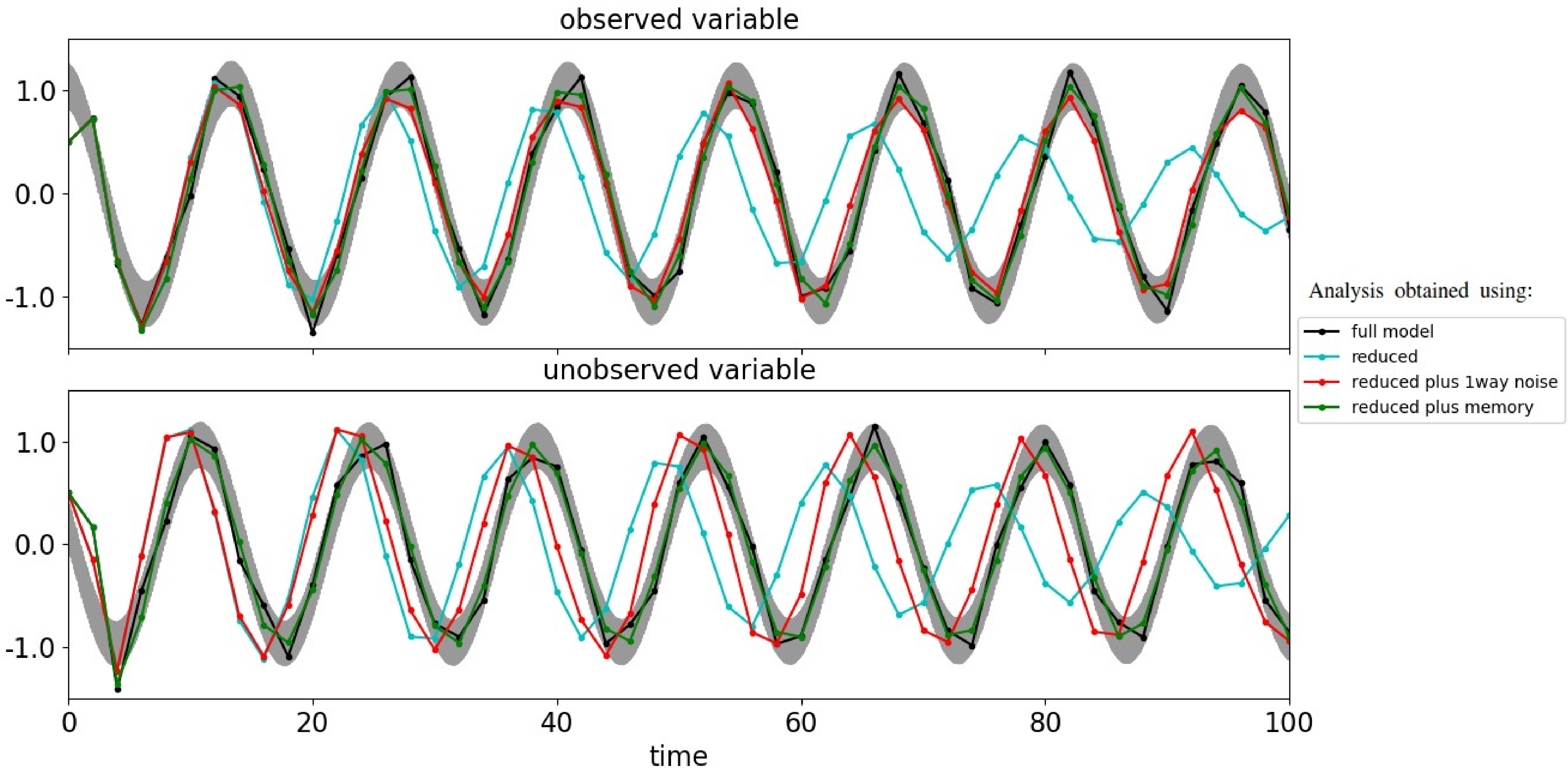
Different components of model error have a very different behaviour.

The effect in the covariance



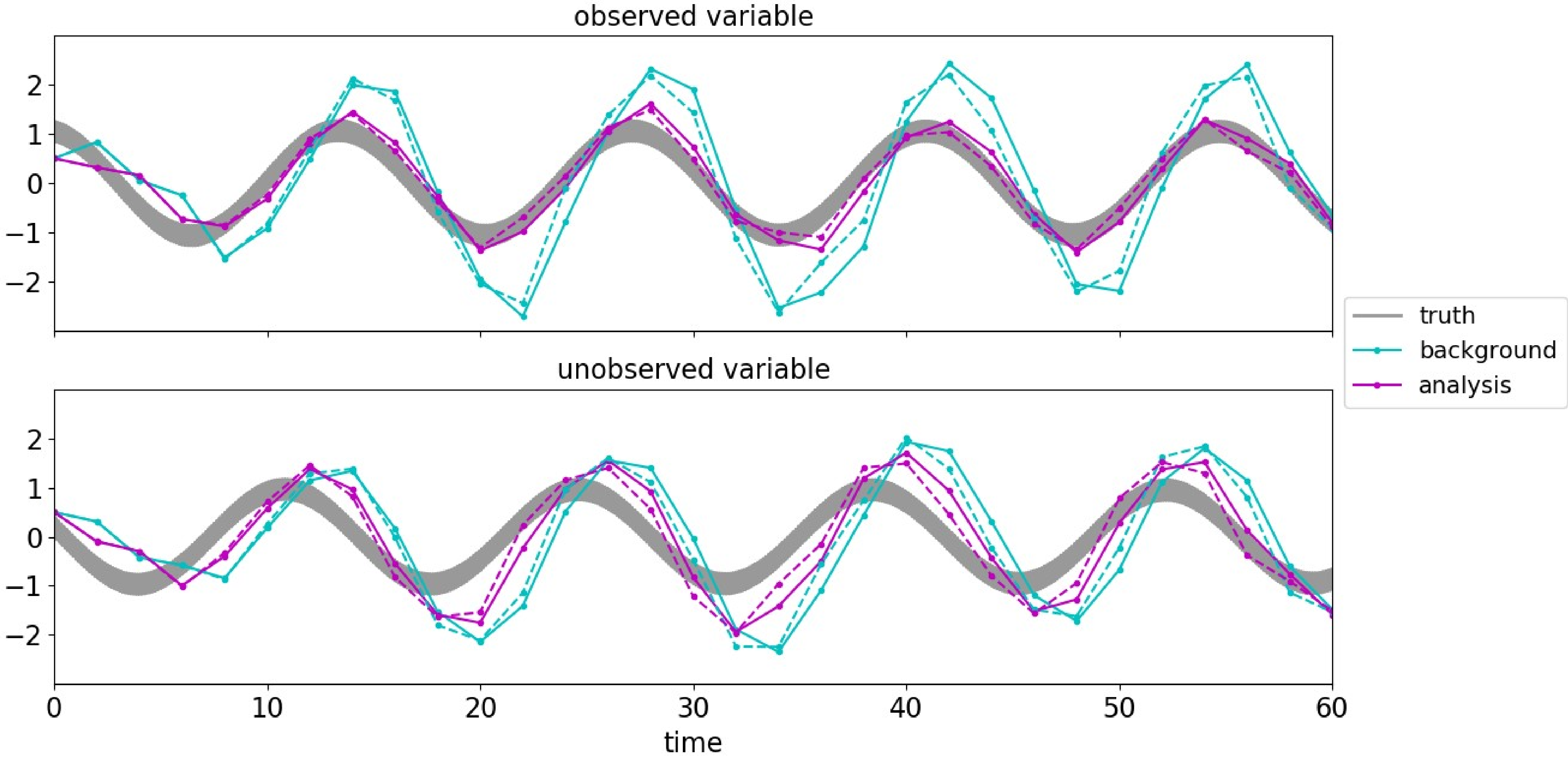
Assimilating only with the reduced model

Now we assimilate observations using different components.



Assimilating with a partial representation of extra components

We use a ‘summarized’ representation of 1w-noise, and partial reconstruction of the memory evolution matrix.



Outline

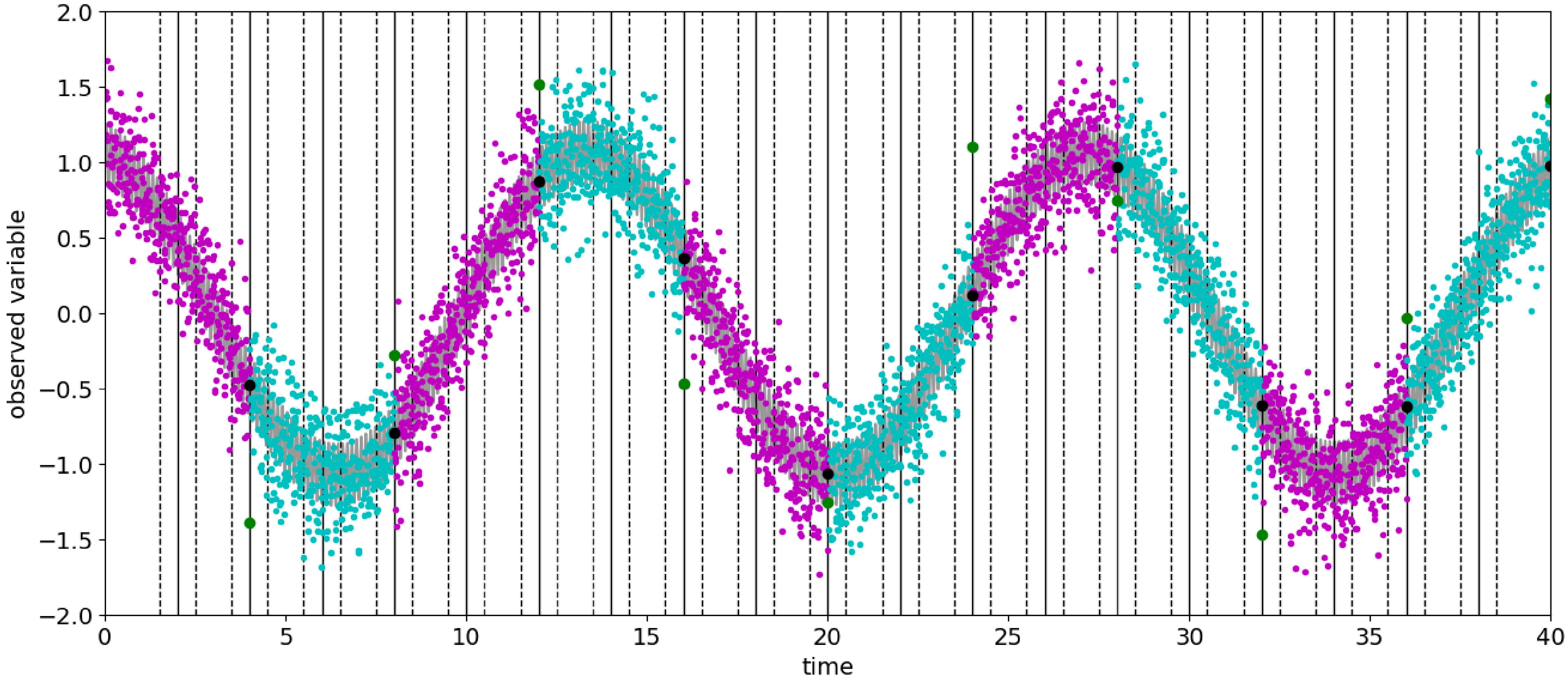
1. Some basics of data assimilation

2. Time-structure in model error

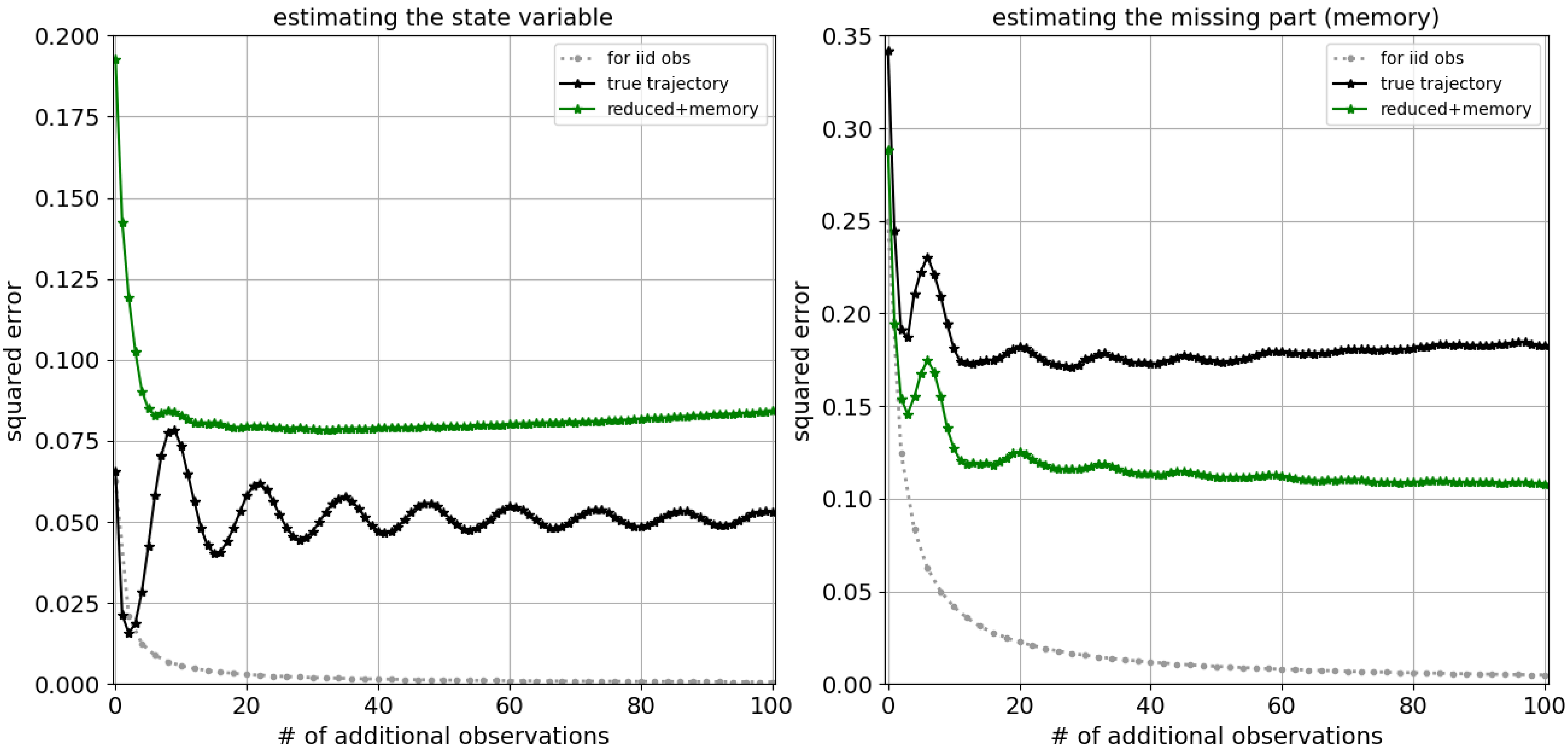
3. Model error and sources

4. Summary and current work

Using asynchronous fast observations



Using asynchronous fast observations



Summary and current 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 'memory' matters considerably**.

- We have illustrated ‘physical’ **origins for model error of different nature**, and the **different behaviour of the components.**

- Can we **use the slow-varying error (memory) in the assimilation.** Can this be deduced from observations?