Understanding the Error of Representation

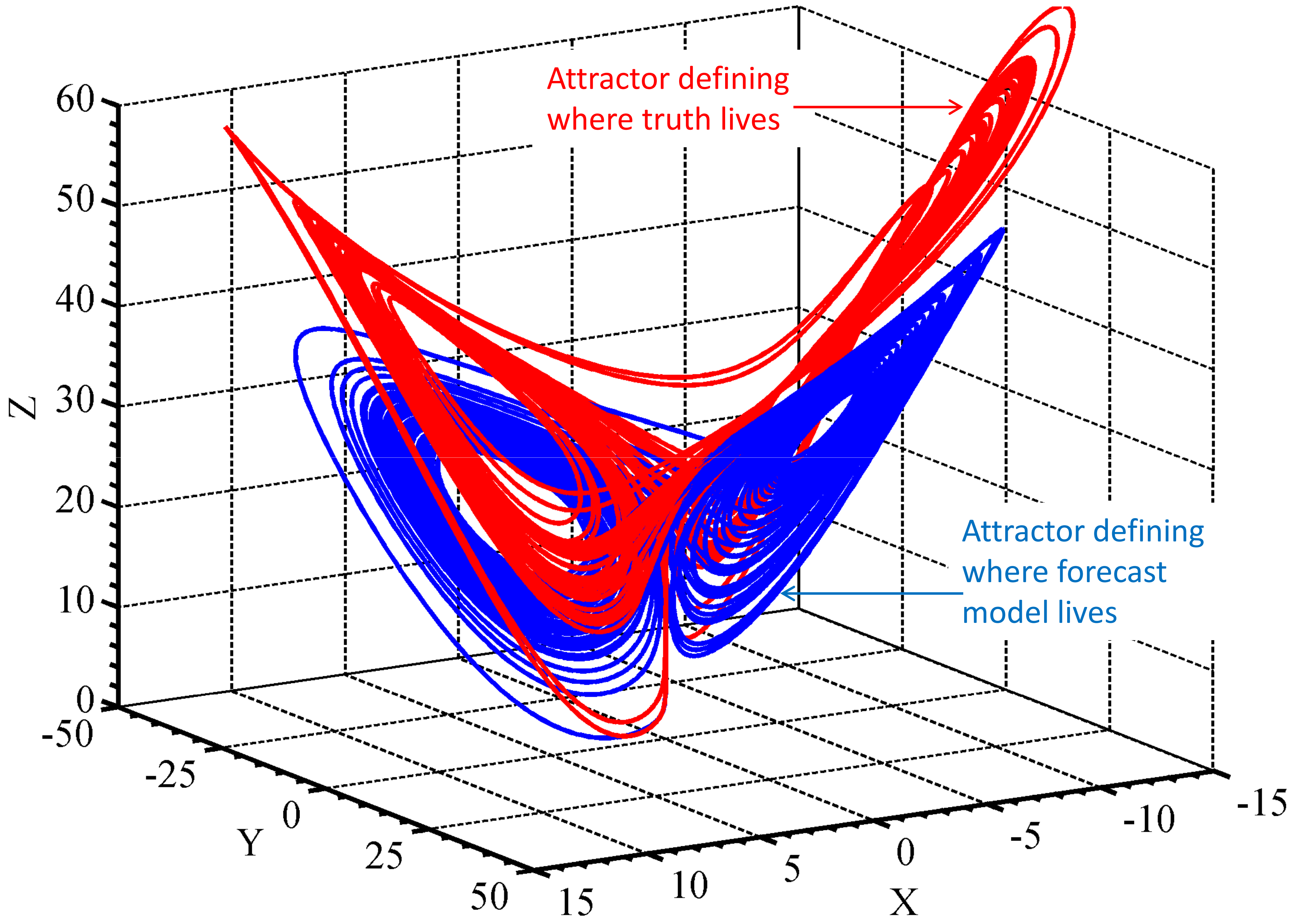
Nancy Nichols

University of Reading

and

Dan Hodyss

US Naval Research Laboratory at Monterey



Attractor defining where truth lives
Attractor defining where forecast model lives
Z
60
50
40
30
20
10
0
Y
-50
-25
0
25
50
X
15
10
5
0
-5
-10
-15

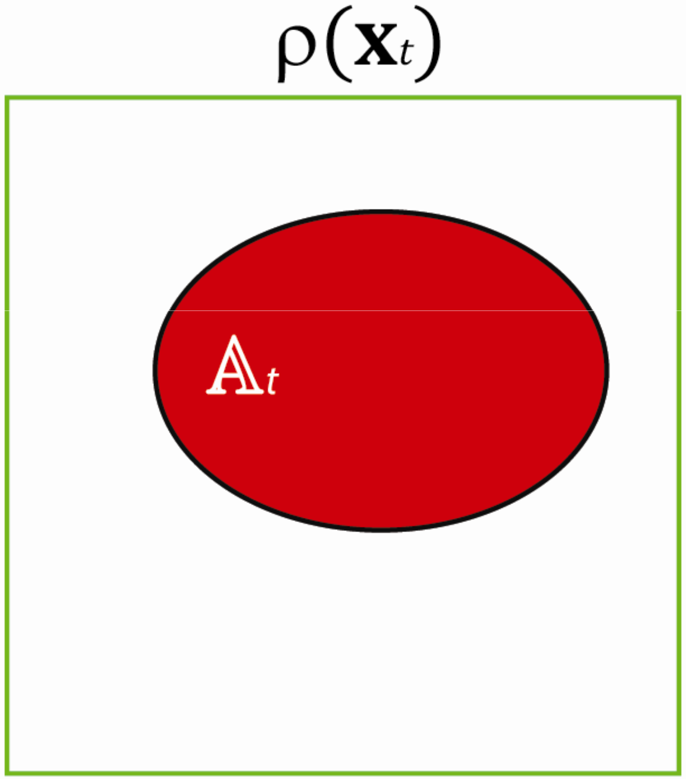
Climatological Distribution

- The distribution of states that results from running the true model for a long time, discretizing phase space, and counting the number of times the trajectory entered each cell of the discretization.

- In the limit as the number of cells and the length of the trajectory approaches infinity we obtain a pdf we label

$$\boxed{\text{climatological distribution} \rightarrow \rho\left(\mathbf{x}_t\right)}$$

Defining the True "Attractor"



The set of states for which

$$\rho\left(\mathbf{x}_t\right) > 0$$

will be labeled A_t

Data Assimilation on the True Attractor

Step j = 0: $\boxed{\rho(\mathbf{x}_t) \quad \text{climatological distribution}}$

$$\downarrow$$

Step j = 1: $\rho(\mathbf{x}_t|\mathbf{y}_1) = C_1\rho(\mathbf{y}_1|\mathbf{x}_t)\rho(\mathbf{x}_t)$

$$\downarrow$$

Step j = 2: $\rho(\mathbf{x}_t|\mathbf{y}_2,\mathbf{y}_1) = C_2\rho(\mathbf{y}_2|\mathbf{x}_t)\rho(\mathbf{x}_t|\mathbf{y}_1)$

$$\vdots$$

Step j = N: $\underbrace{\rho(\mathbf{x}_t|\mathbf{Y})}_{\text{Posterior}} = C_N \underbrace{\rho(\mathbf{y}_N|\mathbf{x}_t)}_{\text{Ob Likelihood}}\underbrace{\rho(\mathbf{x}_t|\mathbf{y}_{N-1},\mathbf{y}_{N-2},\ldots)}_{\text{Prior}}$

where all the observations are collected together: $\mathbf{Y} = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_N]$

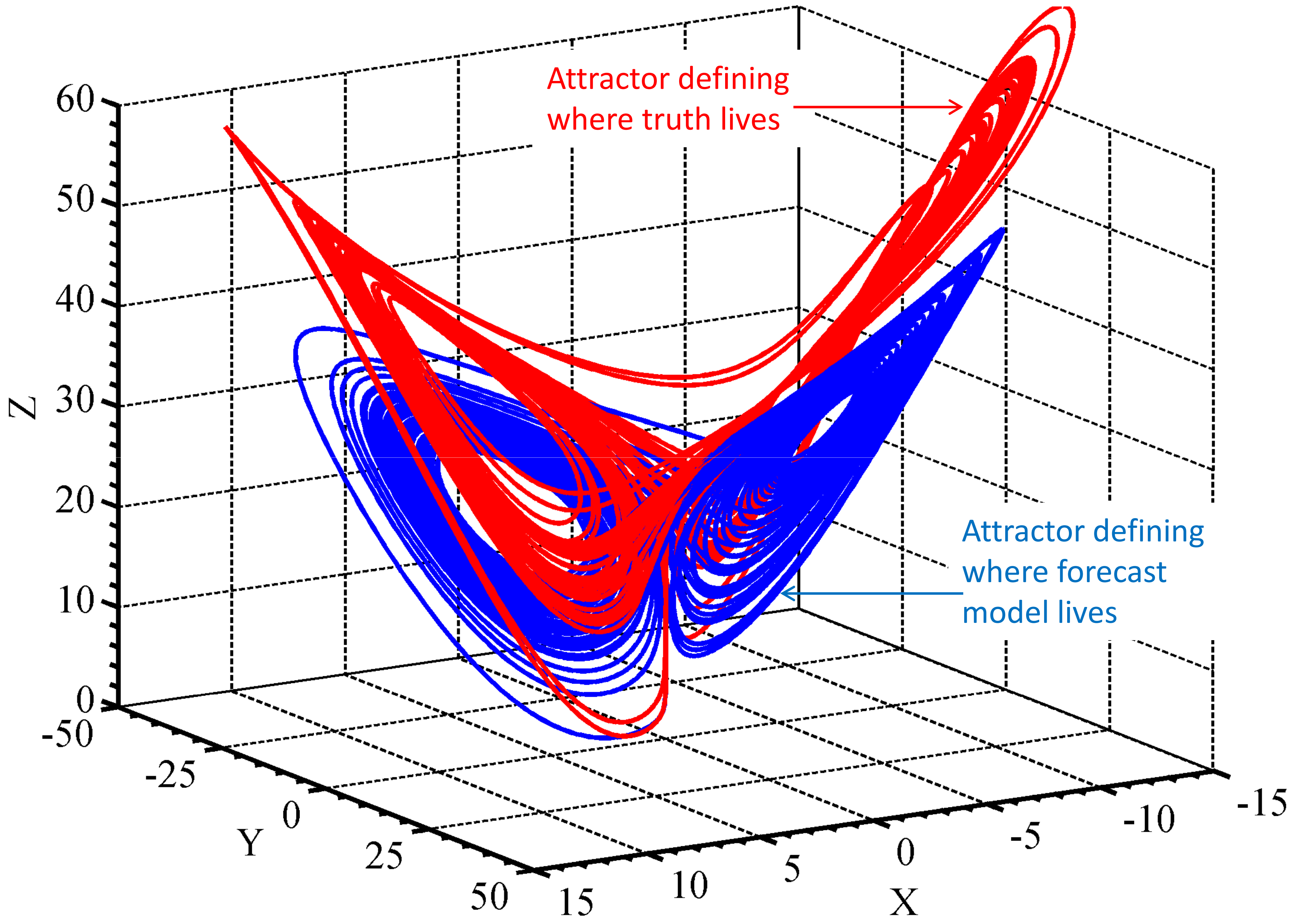
$\mathbb{A}_t$ is invariant to DA

- The data assimilation problem is:

$$\underbrace{\rho\left(\mathbf{x}_t \middle| \mathbf{Y}\right)}_{\text{Posterior}} = C_N \underbrace{\rho\left(\mathbf{y}_N \middle| \mathbf{x}_t\right)}_{\text{Ob Likelihood}} \underbrace{\rho\left(\mathbf{x}_t \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right)}_{\text{Prior}}$$

where the observations are $\mathbf{Y} = \left[\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_N\right]$

- Because $\rho\left(\mathbf{y}_N \middle| \mathbf{x}_t\right) > 0$, the DA steps j = 0 to N do not change the elements of $\mathbb{A}_t$.

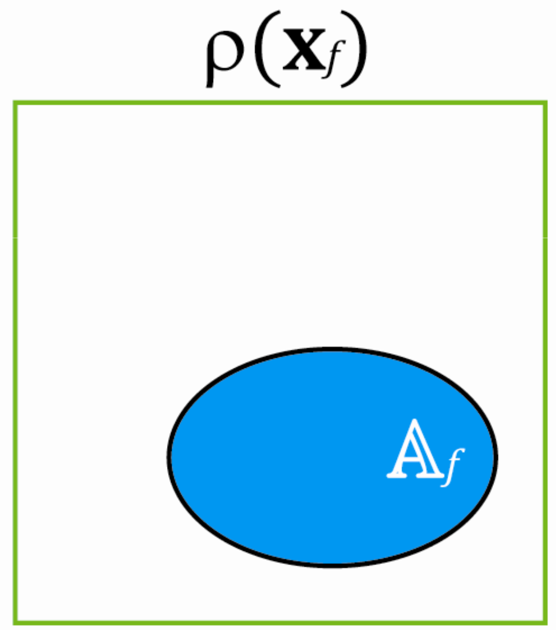

Attractor defining
where truth lives
Attractor defining
where forecast
model lives
Z
Y
X
60
50
40
30
20
10
0
-50
-25
0
25
50
15
10
5
0
-5
-10
-15

Defining the Forecast "Attractor"

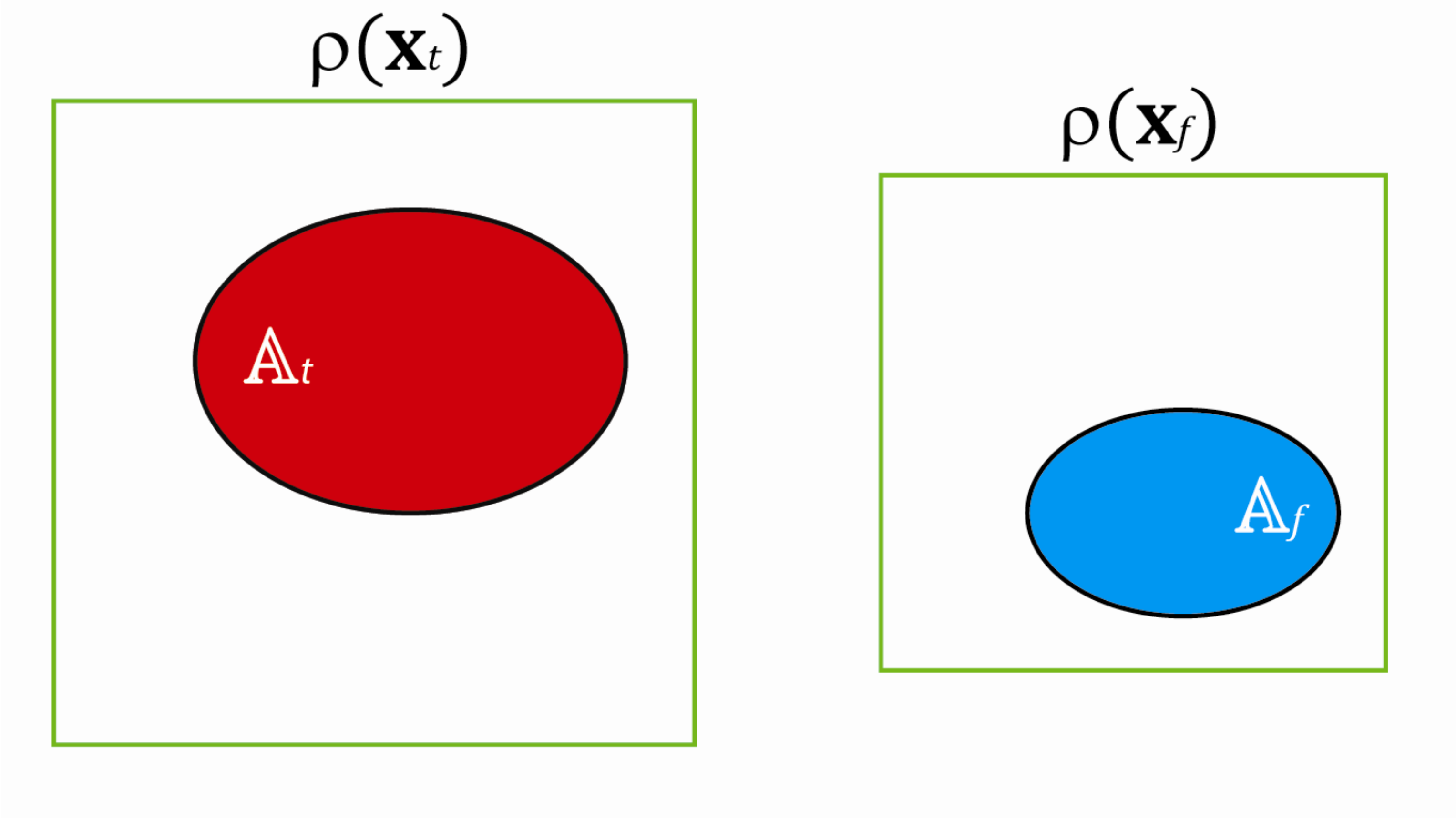
The set of states for which

$$\rho\left(\mathbf{x}_f\right) > 0$$

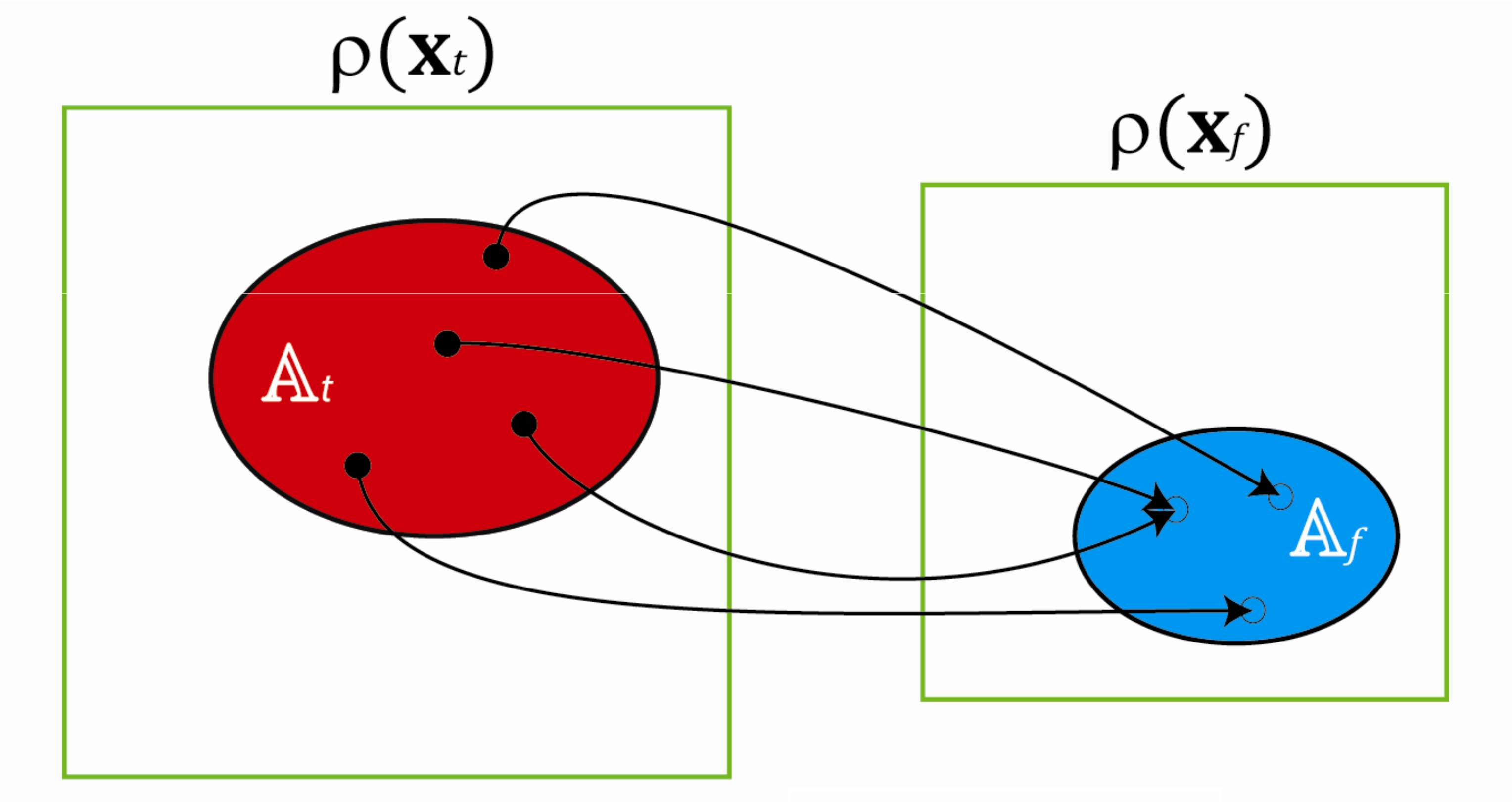
will be labeled A_f



True and Forecast "Attractors"



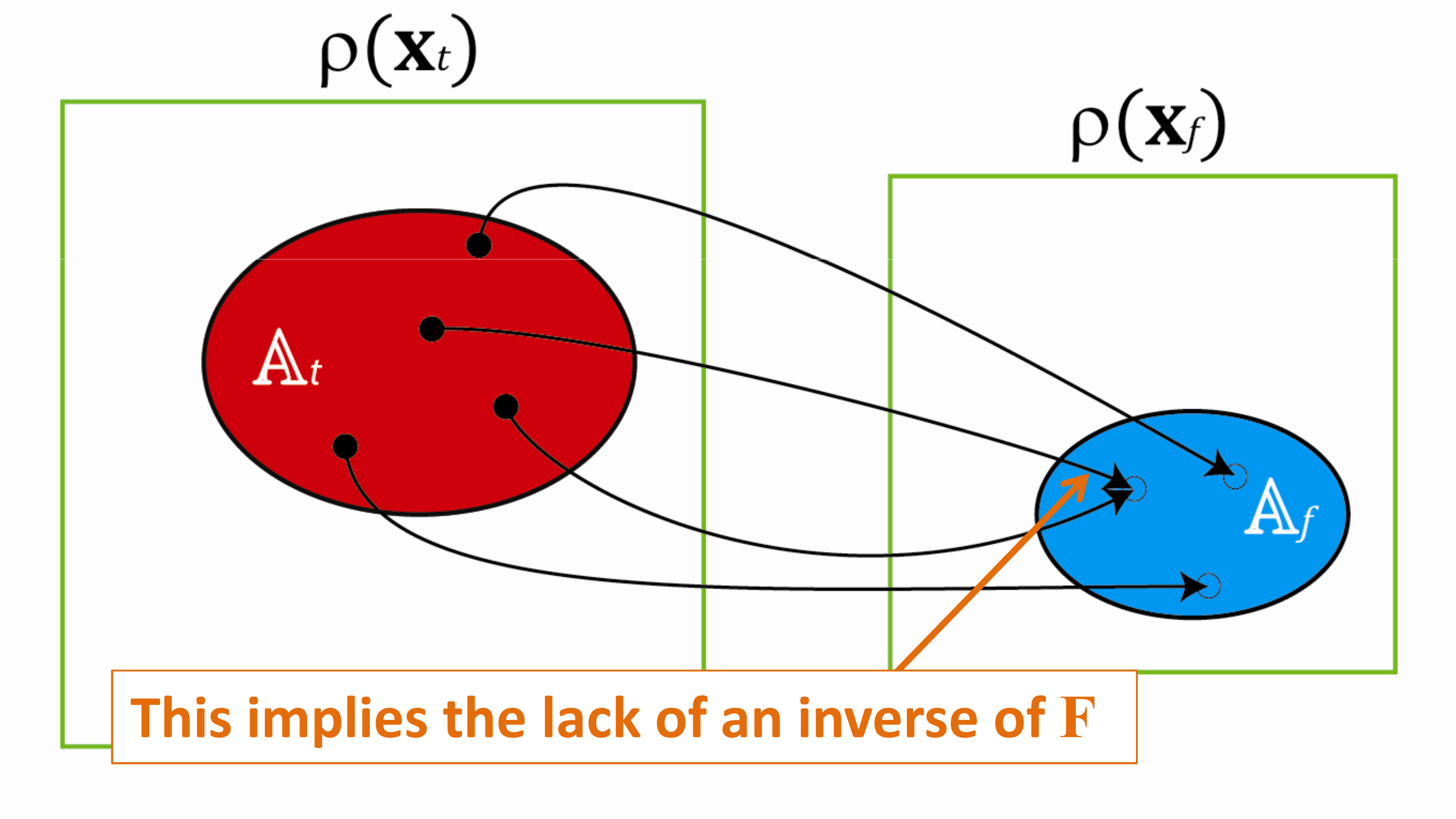
True and Forecast "Attractors"



We will label this function $\mathbf{F}\colon \mathbb{A}_t \to \mathbb{A}_f$

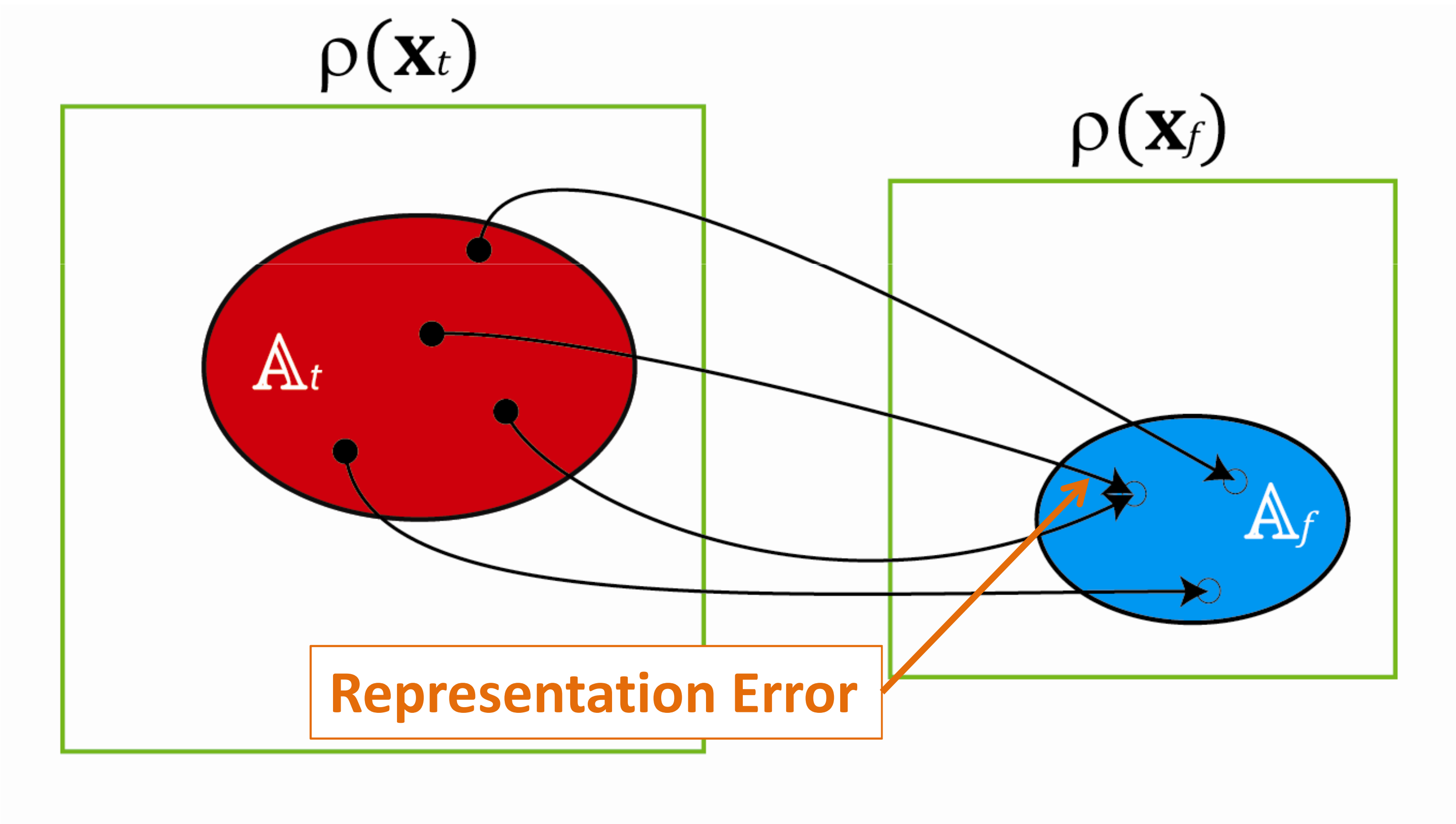
True and Forecast "Attractors"

The 'smoothing operator' $\mathbf{F}$ is interesting because it does not have an inverse.



True and Forecast "Attractors"

Not all true states can be represented by model states.



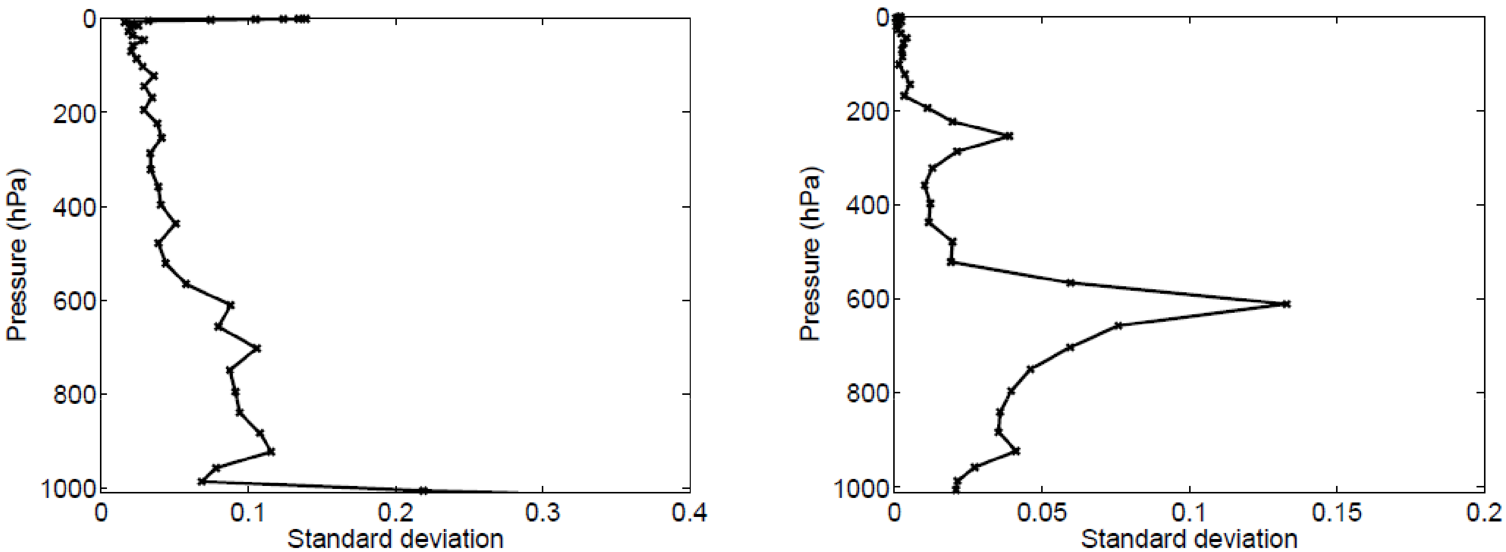
Some Examples of $\mathbf{F}$

Daley (1993), Liu & Rabier (2002) assume that the model state is a truncation in spectral space of a high resolution 'true' state and write

$$\mathbf{x}_t = \mathbf{F}_t \hat{\mathbf{x}}_t \ , \quad \mathbf{x}_f = \mathbf{F}_f \hat{\mathbf{x}}_f = \mathbf{F}_f \mathbf{T} \hat{\mathbf{x}}_t = \mathbf{F}_f \mathbf{T} \mathbf{F}_t^* \mathbf{x}_t$$

They then derive an expression for the representation error for this operator. Previously we have used this formulation to examine representation error using data from the Met Office High Resolution model.

Experimental Results



(a) Temperature (K)   (b) log(Specific humidity) (kg/kg)

Standard Deviations of Representation Errors

Waller et al, *QJ Royal Meteor Soc*, 2014, 140: 1189 – 1197

Some Examples of $\mathbf{F}$: Statistical

- One could define $\mathbf{F}$ from high and low-resolution forecasts
  - Run mesoscale model over a season
  - Run global model over the same season
  - Perform regression

Linear: $$\mathbf{x}_f = \overline{\mathbf{x}}_f + \mathbf{F}\left[\mathbf{x}_t - \overline{\mathbf{x}}_t\right], \quad \mathbf{F} = \mathbf{X}_f \mathbf{X}_t^T \left[\mathbf{X}_t \mathbf{X}_t^T\right]^{-1}$$

Quadratic: $$\mathbf{x}_f = \overline{\mathbf{x}}_f + \hat{\mathbf{F}}\left[\mathbf{x}_t - \overline{\mathbf{x}}_t\right], \quad \hat{\mathbf{F}} = \mathbf{X}_f \hat{\mathbf{X}}_t^T \left[\hat{\mathbf{X}}_t \hat{\mathbf{X}}_t^T\right]^{-1}$$

Some Examples of $\mathbf{F}$: Parameterized

- Lilly (1962) constructs a forecast model by choosing to simulate grid cell averages
  - Parameterize sub-grid scale processes through grid cell averaged fluxes of heat, momentum, etc.
- This would be approximately equivalent to

$$\mathbf{x}_f = \mathbf{S}\mathbf{x}_t$$

where $\mathbf{S}$ is a smoothing operator.

Conversion Densities

- Map: $\mathbf{x}_f = \mathbf{F}\left(\mathbf{x}_t\right)$
- Implication:
  - Given $\mathbf{x}_t$ and $\mathbf{F}$ there is one and only one forecast state $\mathbf{x}_f$
- This implies that $\rho\left(\mathbf{x}_f \middle| \mathbf{x}_t\right) = \delta\left(\mathbf{x}_f - \mathbf{F}\left(\mathbf{x}_t\right)\right)$
- Bayes' Rule states that the "priors" are related by the conversion densities

$$\rho\left(\mathbf{x}_f \middle| \mathbf{x}_t\right)\rho\left(\mathbf{x}_t \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right) = \rho\left(\mathbf{x}_t \middle| \mathbf{x}_f\right)\rho\left(\mathbf{x}_f \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right)$$

Using the Conversion Densities: Forecast Prior

- We may map the true prior density onto the forecast models attracting manifold as:

$$\rho\left(\mathbf{x}_f \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right) = \int_{-\infty}^{\infty} \rho\left(\mathbf{x}_f \middle| \mathbf{x}_t\right) \rho\left(\mathbf{x}_t \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right) d\mathbf{x}_t$$

- This density describes the distribution of forecasts $\mathbf{x}_f$ obtained from sampling from the true prior and mapping through $\mathbf{F}$.

Using the Conversion Densities: Forecast Posterior

- If there's a forecast prior then there is a forecast posterior

$$\rho\left(\mathbf{x}_f \middle| \mathbf{Y}\right) = C_N \rho\left(\mathbf{y}_N \middle| \mathbf{x}_f\right) \rho\left(\mathbf{x}_f \middle| \mathbf{y}_{N-1}, \mathbf{y}_{N-2}, \ldots\right)$$

where

$$\boxed{\rho\left(\mathbf{y}_N \middle| \mathbf{x}_f\right) = \int_{-\infty}^{\infty} \rho\left(\mathbf{y}_N \middle| \mathbf{x}_t\right) \rho\left(\mathbf{x}_t \middle| \mathbf{x}_f\right) d\mathbf{x}_t}$$

Note that we have re-defined the goal of our DA!

High Resolution – True Model

We use a Gaussian model for which

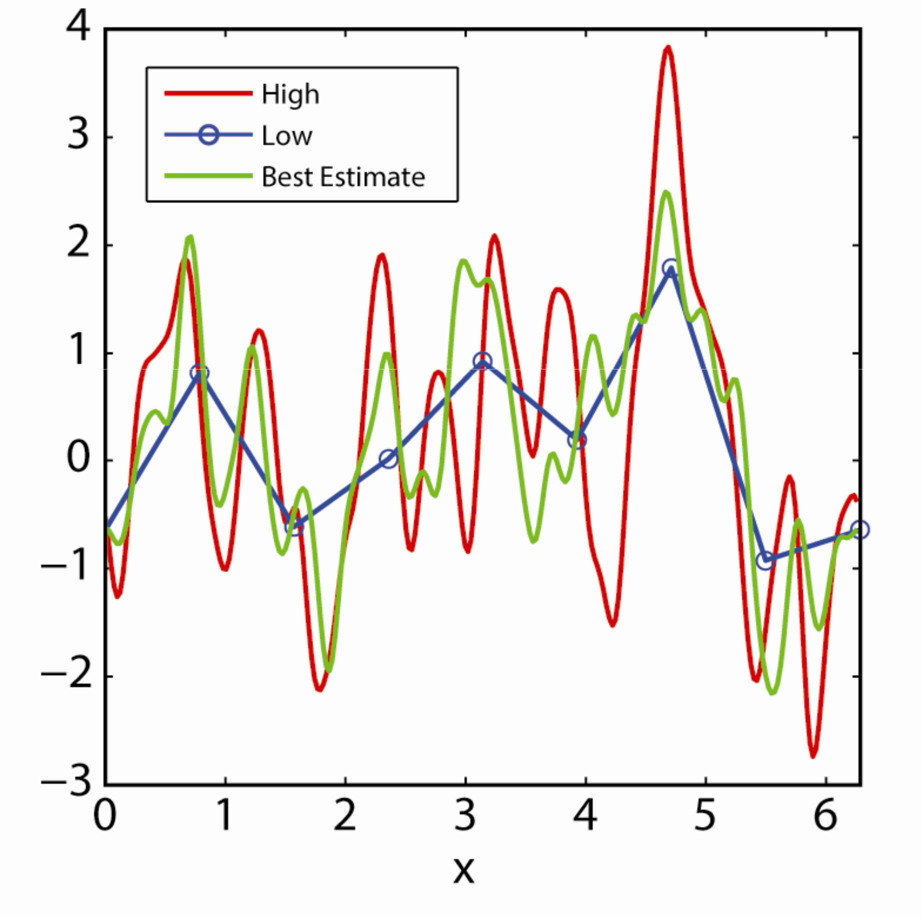
$$\mathbf{x}_t = \overline{\mathbf{x}}_t + \mathbf{Z}\boldsymbol{\eta}$$

where $\boldsymbol{\eta} \sim N(\mathbf{0},\mathbf{1})$

This model is constrained to produce states consistent with

$$\mathbf{P}_t = \mathbf{Z}\mathbf{Z}^T = \mathbf{E}\boldsymbol{\Gamma}\mathbf{E}^T$$

The length of the "true" state vector is N.



Low Resolution – Forecast Model

Relate the forecasts and the true model

$$\mathbf{x}_f = \mathbf{S}\mathbf{x}_t$$

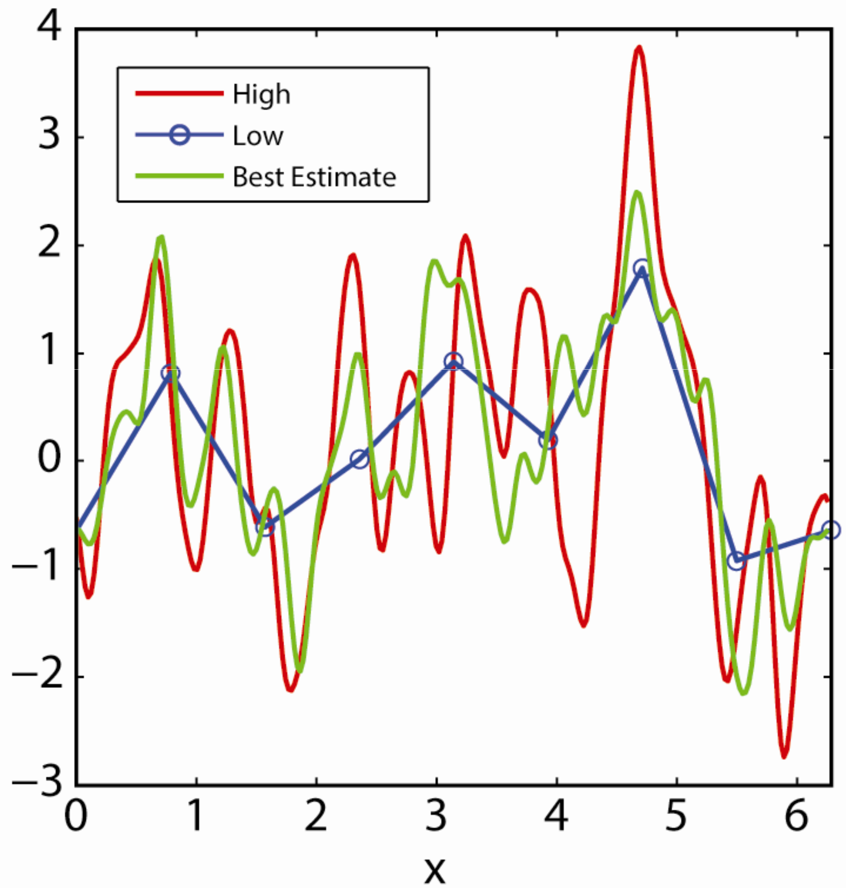
where

$$\mathbf{S} = \mathbf{E}_L \begin{bmatrix} \mathbf{D}^{1/2}\mathbf{T} & \mathbf{0} \end{bmatrix} \mathbf{E}^T$$

The low-resolution forecast is constrained such that

$$\mathbf{P}_f = \mathbf{S}\mathbf{P}_t\mathbf{S}^T$$

The length of the "forecast" state vector is M.



Mean of Conversion Density

The model estimate of the observation is

$$\mathbf{y}_f\left(\mathbf{x}_f\right) = \int_{-\infty}^{\infty} \mathbf{y}_N \rho\left(\mathbf{y}_N \middle| \mathbf{x}_f\right) d\mathbf{y}_N$$

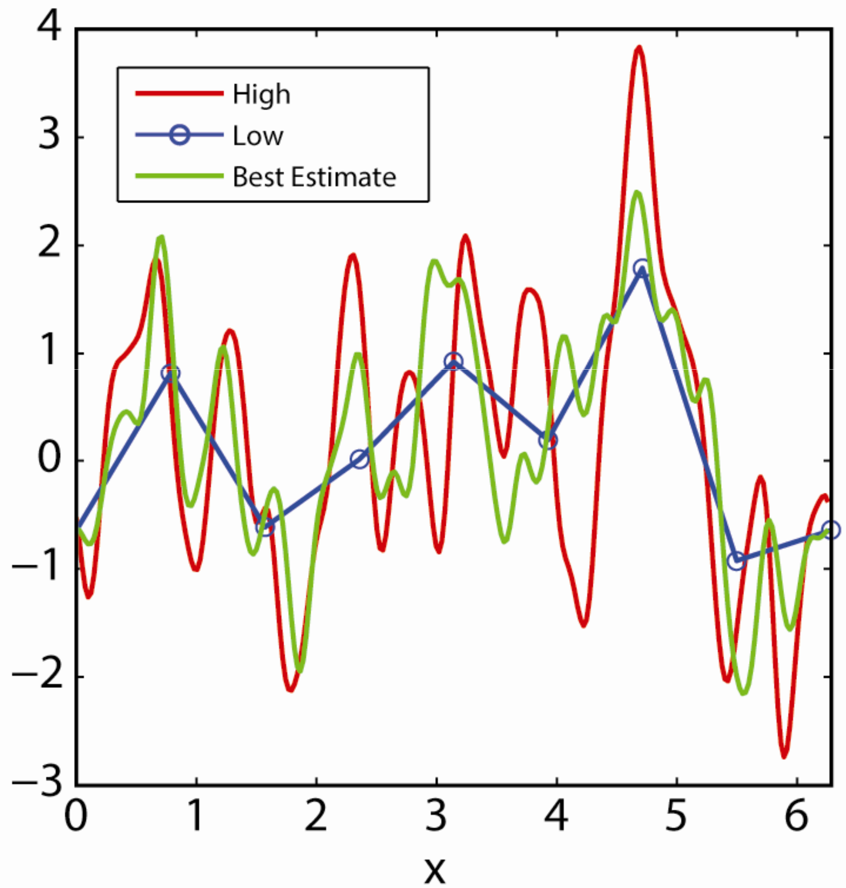
$$= \mathbf{H} \int_{-\infty}^{\infty} \mathbf{x}_t \rho\left(\mathbf{x}_t \middle| \mathbf{x}_f\right) d\mathbf{x}_t$$

Because our map is linear:

$$\mathbf{y}_f\left(\mathbf{x}_f\right) = \mathbf{H}\overline{\mathbf{x}}_t + \mathbf{H}\mathbf{G}_p\left[\mathbf{x}_f - \overline{\mathbf{x}}_f\right]$$

where

$$\mathbf{G}_p = \mathbf{Z}\left[\mathbf{S}\mathbf{Z}\right]^{\dagger} = \mathbf{S}^{\dagger}$$



Observation Error Covariance Matrix

The observation error covariance matrix for this case is

$$\mathbf{R}_f = \underbrace{\mathbf{R}_i}_{\text{Instrument}} + \underbrace{\mathbf{H}\mathbf{E}\boldsymbol{\Theta}\mathbf{E}^T\mathbf{H}^T}_{\text{Representation}}$$

where Θ is a diagonal matrix with a diagonal equal to

$$\Theta_i = \begin{cases} 0, & i = 1,\ldots,M \\ \Gamma_i, & i = M+1,\ldots,N \end{cases}$$

Observation Error Covariance Matrix

The observation error covariance matrix for this case is

$$\mathbf{R}_f = \underbrace{\mathbf{R}_i}_{\text{Instrument}} + \underbrace{\mathbf{HE\Theta E}^T\mathbf{H}^T}_{\text{Representation}}$$

where Θ is a diagonal matrix with a diagonal equal to

$$\Theta_i = \begin{cases} 0, & i = 1, \ldots, M \\ \Gamma_i, & i = M+1, \ldots, N \end{cases}$$

1. Representation error arises from truncation only!

Observation Error Covariance Matrix

The ob error covariance matrix is for this case

$$\mathbf{R}_f = \underbrace{\mathbf{R}_i}_{\text{Instrument}} + \underbrace{\mathbf{H}\mathbf{E}\boldsymbol{\Theta}\mathbf{E}^T\mathbf{H}^T}_{\text{Representation}}$$

where Θ is a diagonal matrix with a diagonal equal to

$$\Theta_i = \begin{cases} 0, & i = 1,\ldots,M \\ \Gamma_i, & i = M+1,\ldots,N \end{cases}$$

2. Representation error is the covariance matrix of all the stuff the ob sees but the model does not.

Data Assimilation for Forecast States

The data assimilation algorithm for the forecast posterior mean is

Analysis: $$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_f + \mathbf{G}\left[\mathbf{v} - \langle \mathbf{v} \rangle\right], \quad \mathbf{v} = \mathbf{y} - \mathbf{H}_f \bar{\mathbf{x}}_f$$

Gain: $$\mathbf{G} = \mathbf{P}_f \mathbf{H}_f^T \left[\mathbf{H}_f \mathbf{P}_f \mathbf{H}_f^T + \mathbf{R}_f\right]^{-1}$$

Observation Operator: $$\mathbf{H}_f \equiv \mathbf{H}\mathbf{S}^\dagger$$

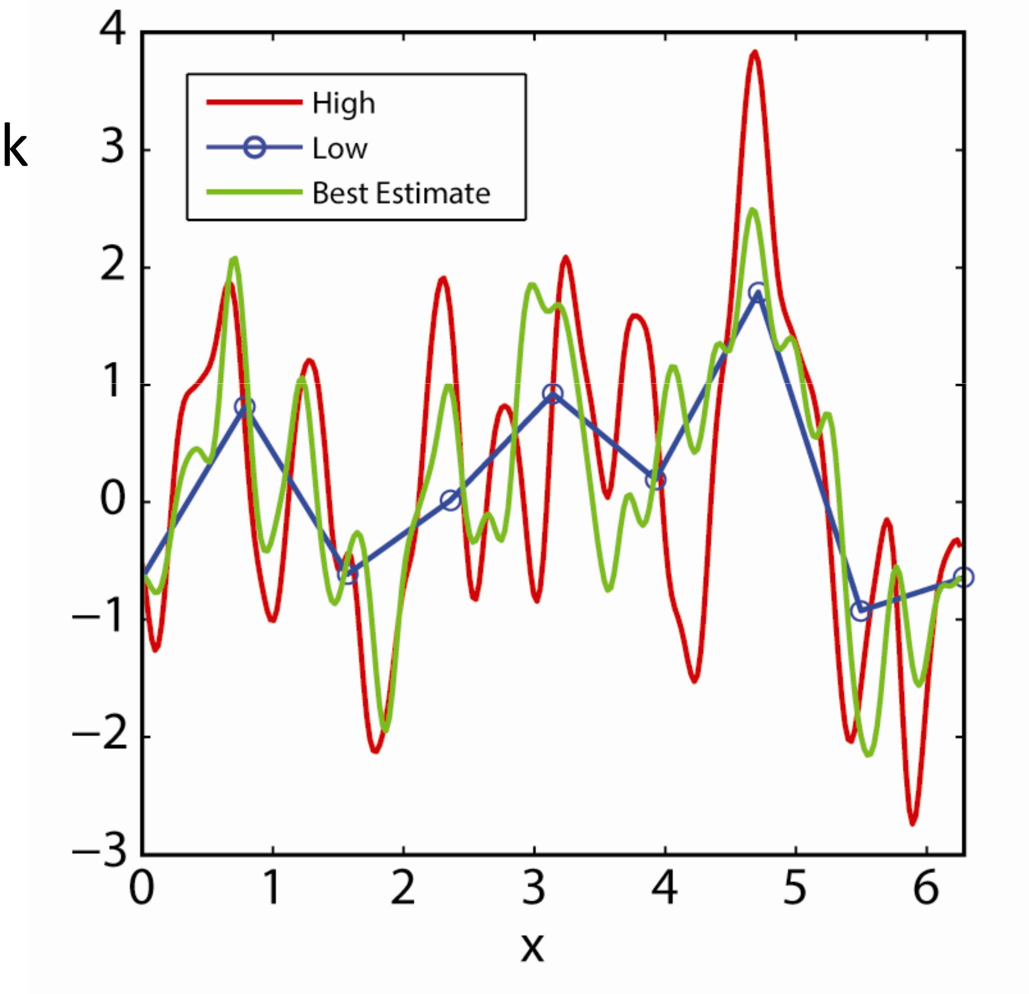
We choose this observation operator because the variance is correct up to its M eigenvalues!

$$\mathbf{H}_f \mathbf{P}_f \mathbf{H}_f^T = \mathbf{H}\mathbf{S}^\dagger \mathbf{P}_f \mathbf{S}^{\dagger T} \mathbf{H}^T = \mathbf{H}\mathbf{E}\begin{bmatrix} \boldsymbol{\Gamma}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{E}^T \mathbf{H}^T$$

We Get the Post-Processing for Free!

Post-process the forecast back to the true attractor:

$$\mathbf{x}_p\left(\mathbf{x}_f\right) = \overline{\mathbf{x}}_t + \mathbf{S}^\dagger\left[\mathbf{x}_f - \overline{\mathbf{x}}_f\right]$$

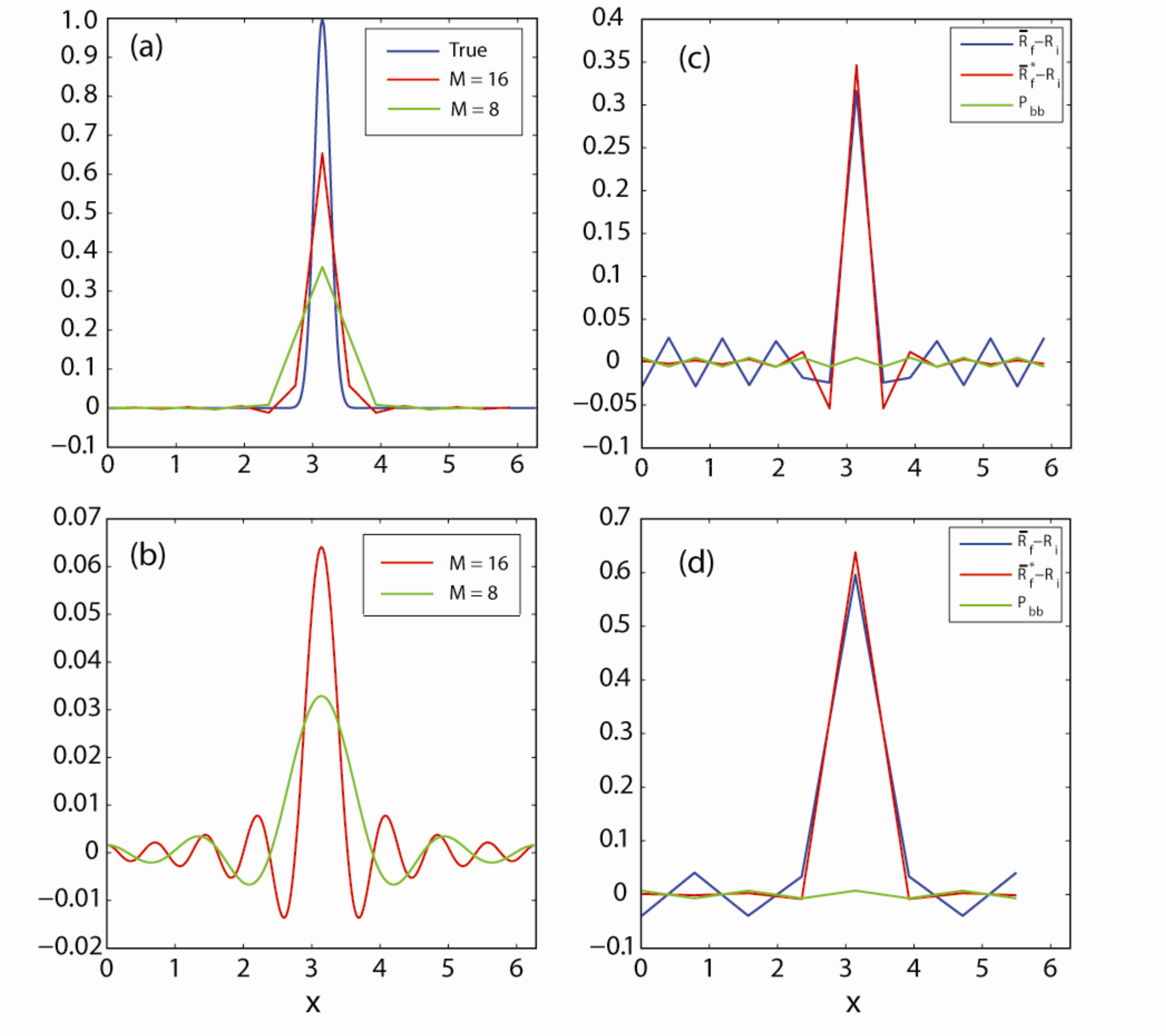


Summary

- Representation error arises from truncation.
  - The representation error covariance matrix is the covariance matrix of all the stuff the forecast model can't see.
- Representation error should be accounted for by re-defining the observation operator such that it includes a "bias correction" algorithm for the forecast.
- Algorithm:
  - Use observations of true attractor to create states on forecast attractor.
  - Use forecast model to integrate forecast states forward on forecast attractor.
  - Post-process forecast back to true attractor.

Hodyss & Nichols, *Tellus A*, 2015, 67, 24822

Pure Spectral Truncation: D = I



Gaussian Smoother: D =/ I