# Variational Data Assimilation via Sparse Regularization

**Mohammad Ebtehaj**[1,2], **Efi Foufoula-Georgiou**[2], **Gilad Lerman**[1]

University of Minnesota

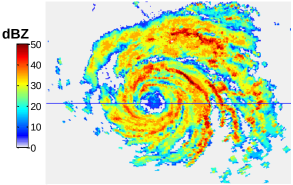
[1] School of Mathematics

[2] Department of Civil Engineering

June, 2013

# Rainfall Prior in Wavelet Domain

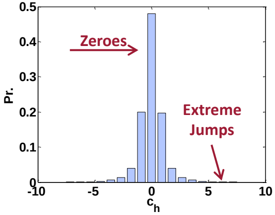
- Heavy tailed histogram with large mass around zero

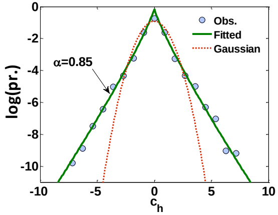



Dim:[400, 560] Range: [0, 50]

## Rainfall Prior in Wavelet Domain

- Heavy tailed histogram with large mass around zero




**Generalized Gaussian**

$$p(x) = \exp\left(-\lambda \left|x\right|^{\alpha}\right)$$

$(\alpha=2) \rightarrow$ Gaussian

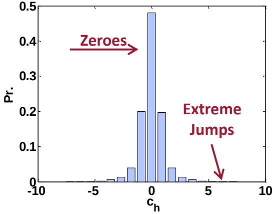
$(\alpha=1) \rightarrow$ Laplace

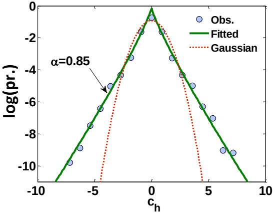
$$p(\mathbf{x}) \propto \exp\left(-\lambda \left\|\mathbf{x}\right\|_{\alpha}^{\alpha}\right)$$

where $\left\|\mathbf{x}\right\|_{\alpha}^{\alpha} = \Sigma \left|x_i\right|^{\alpha}$

# Rainfall Prior in Wavelet Domain

- Heavy tailed histogram with large mass around zero




**Generalized Gaussian**

$$p(x) = \exp\left(-\lambda \left|x\right|^{\alpha}\right)$$

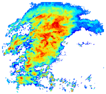
$(\alpha{=}2) \rightarrow$ Gaussian

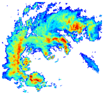
$(\alpha{=}1) \rightarrow$ Laplace

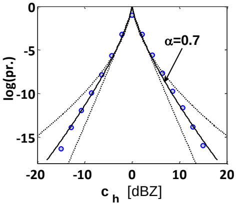
$$p(\mathbf{x}) \propto \exp\left(-\lambda \left\|\mathbf{x}\right\|_{\alpha}^{\alpha}\right)$$

where $\|\mathbf{x}\|_{\alpha}^{\alpha} = \Sigma \left|x_i\right|^{\alpha}$

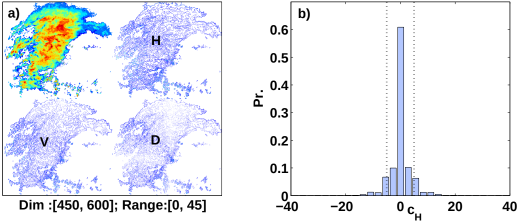
- 200 storms over TX and FL





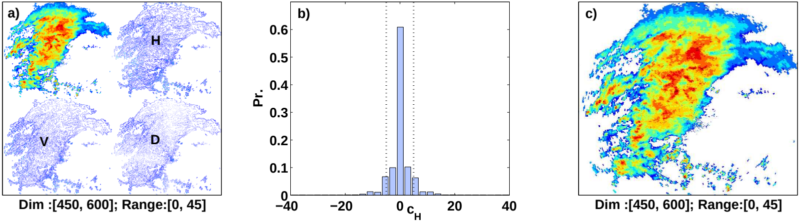


# Rainfall Sparsity




- **Sparsity** in the derivative or **wavelet domain**

# Rainfall Sparsity

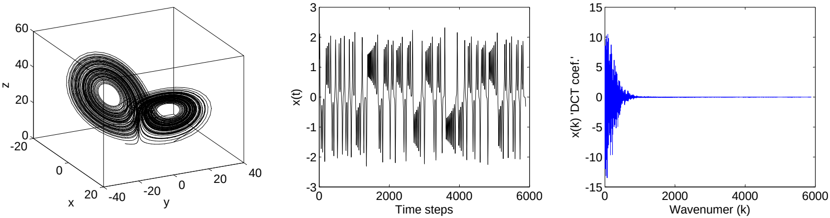



- **Sparsity** in the derivative or **wavelet domain**
- **20%** of the data contains **98%** of the total energy
- Sparsity is a strong **prior** knowledge.
- **How to incorporate sparsity?**

## Lorenz (1963)

$$\begin{aligned}\dot{x} &= \sigma\left(y - x\right)\\ \dot{y} &= x\left(\rho - z\right) - y\\ \dot{z} &= xy - \beta z\end{aligned}$$

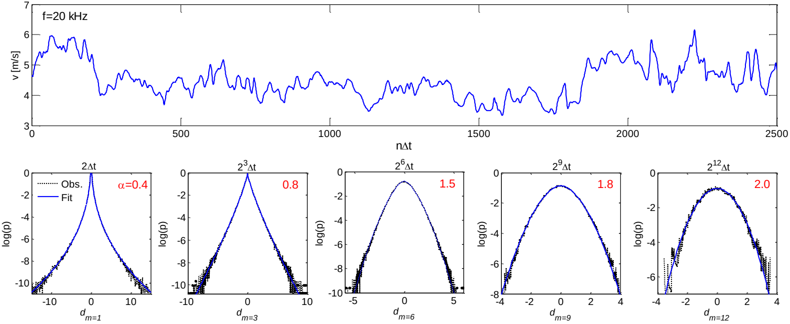
- The system is chaotic for $\sigma = 10,\ \beta = 8/3,\ \rho = 28$




- Lorenz system shows remarkable sparsity in the DCT domain
- $x(k) = \omega_k \sum_{t=1}^{N} x(t) \cos \frac{\pi(2t-1)(k-1)}{2N}$

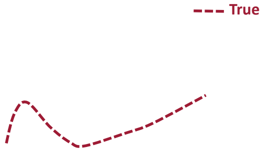
# Sparsity in Turbulent Flow

- SAFL wind tunnel
- $R_e = 4 \times 10^4$
- $f = 20$ kHz @ 105 s



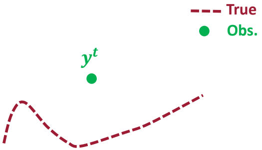
# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$



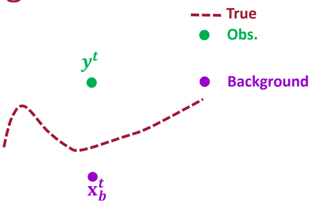
# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$
- Observation model: $\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i) + \mathbf{v}_i \in \mathbb{R}^n$



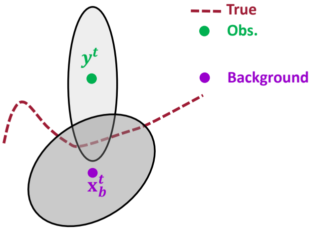
# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$
- Observation model: $\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i) + \mathbf{v}_i \in \mathbb{R}^n$
- Background state: $\mathbf{x}_0^b \in \mathbb{R}^m$



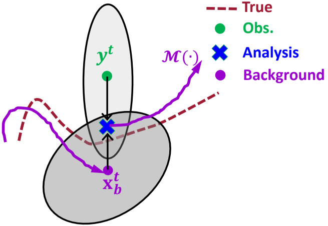
# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$
- Observation model: $\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i) + \mathbf{v}_i \in \mathbb{R}^n$
- Background state: $\mathbf{x}_0^b \in \mathbb{R}^m$
- Error: $\mathbf{v} \sim \mathcal{N}(0,\, \mathbf{R})\,, \quad \mathbf{w} \sim \mathcal{N}(0,\, \mathbf{B})$



# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$
- Observation model: $\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i) + \mathbf{v}_i \in \mathbb{R}^n$
- Background state: $\mathbf{x}_0^b \in \mathbb{R}^m$
- Error: $\mathbf{v} \sim \mathcal{N}(0,\, \mathbf{R})$, $\mathbf{w} \sim \mathcal{N}(0,\, \mathbf{B})$



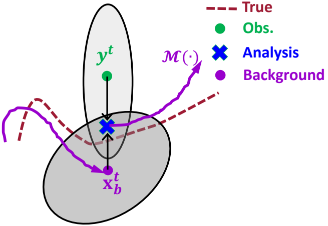

## 4D-VAR

$$\hat{\mathbf{x}}_0^a = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \sum_{i=0}^{k} \left\| \mathbf{y}_i - \mathcal{H}(\mathbf{x}_i) \right\|_{\mathbf{R}_i^{-1}}^2 + \left\| \mathbf{x}_0^b - \mathbf{x}_i \right\|_{\mathbf{B}^{-1}}^2 \right\}$$

$$\text{s.t. } \mathbf{x}_i = \mathcal{M}_{0,\,t}(\mathbf{x}_0)$$

# Variational Data Assimilation via Sparse Regularization

- The true state: $\mathbf{x}_0 \in \mathbb{R}^m$
- Observation model: $\mathbf{y}_i = \mathcal{H}(\mathbf{x}_i) + \mathbf{v}_i \in \mathbb{R}^n$
- Background state: $\mathbf{x}_0^b \in \mathbb{R}^m$
- Error: $\mathbf{v} \sim \mathcal{N}(0, \mathbf{R})$, $\mathbf{w} \sim \mathcal{N}(0, \mathbf{B})$
- $\mathbf{\Phi}$: a pre-selected **basis**




## R4D-VAR

$$\hat{\mathbf{x}}_0^a = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \sum_{i=0}^{k} \|\mathbf{y}_i - \mathcal{H}(\mathbf{x}_i)\|_{\mathbf{R}_i^{-1}}^2 + \left\|\mathbf{x}_0^b - \mathbf{x}_i\right\|_{\mathbf{B}^{-1}}^2 + \lambda \|\mathbf{\Phi}\mathbf{x}_0\|_1 \right\}$$

$$\text{s.t. } \mathbf{x}_i = \mathcal{M}_{0,t}(\mathbf{x}_0)$$

# Quadratic Programming

Assuming $\mathbf{\Phi}\mathbf{x}_0 = \mathbf{c}_0 \in \mathbb{R}^m$, then the above problem can be rewritten as,

$$\underset{\mathbf{z}_0}{\text{minimize}} \left\{ \frac{1}{2}\mathbf{c}_0^{\mathrm{T}}\mathbf{Q}\mathbf{c}_0 + \mathbf{b}^{\mathrm{T}}\mathbf{c}_0 + \lambda \left\|\mathbf{c}_0\right\|_1 \right\} \tag{1}$$

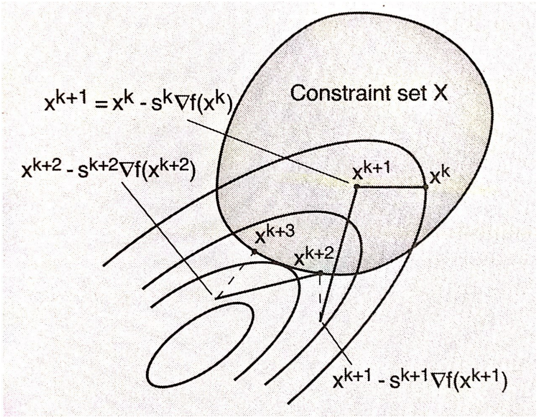
where, $\mathbf{Q} = \mathbf{\Phi}^{-\mathrm{T}} \left(\mathbf{B}^{-1} + \underline{\mathbf{H}}^{\mathrm{T}}\mathbf{R}^{-1}\underline{\mathbf{H}}\right) \mathbf{\Phi}^{-1}$ and $\mathbf{b} = -\mathbf{\Phi}^{-\mathrm{T}} \left(\mathbf{B}^{-1}\mathbf{x}_0^b + \underline{\mathbf{H}}^{\mathrm{T}}\mathbf{R}^{-1}\underline{\mathbf{y}}\right)$.

Having $\mathbf{c}_0 = \mathbf{u}_0 - \mathbf{v}_0$, where $\mathbf{u}_0 = \max\left(\mathbf{c}_0,\, 0\right) \in \mathbb{R}^m$ and $\mathbf{v}_0 = \max\left(-\mathbf{c}_0,\, 0\right) \in \mathbb{R}^m$ and then $\mathbf{w}_0 = [\mathbf{u}_0^{\mathrm{T}},\, \mathbf{v}_0^{\mathrm{T}}]^{\mathrm{T}}$, the more standard QP formulation of the problem is immediately followed as:

$$\begin{aligned} \underset{\mathbf{w}_0}{\text{minimize}} & \left\{ \frac{1}{2}\mathbf{w}_0^{\mathrm{T}} \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} \\ -\mathbf{Q} & \mathbf{Q} \end{bmatrix} \mathbf{w}_0 + \left(\lambda \mathbf{1}_{2m} + \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \end{bmatrix}\right)^{\mathrm{T}} \mathbf{w}_0 \right\} \\ \text{subject to} & \quad \mathbf{w}_0 \succcurlyeq 0. \end{aligned} \tag{2}$$

Obtaining $\hat{\mathbf{w}}_0 = [\hat{\mathbf{u}}_0^{\mathrm{T}},\, \hat{\mathbf{v}}_0^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2m}$ as the solution of (2), one can easily recover $\hat{\mathbf{c}}_0 = \hat{\mathbf{u}}_0 - \hat{\mathbf{v}}_0$ and thus the initial state of interest $\hat{\mathbf{x}}_0 = \mathbf{\Phi}^{-1}\hat{\mathbf{c}}_0$.

# Gradient Projection Method



# Advection-Diffusion Equation

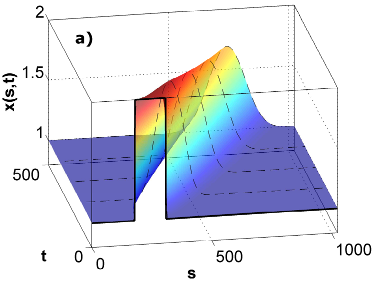
- Flat and Quadratic Top-hat (sparsity in wavelet)

$$\frac{\partial \mathbf{x}(s,\, t)}{\partial t} + a\nabla \mathbf{x}(s,\, t) = \epsilon \nabla^2 \mathbf{x}(s,\, t)$$
$$\mathbf{x}(s,\, 0) = \mathbf{x}_0(s)$$

## Advection-Diffusion Equation

- Flat and Quadratic Top-hat (sparsity in wavelet)

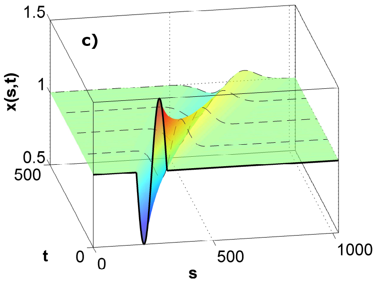
$$\frac{\partial \mathbf{x}(s,\, t)}{\partial t} + a\nabla \mathbf{x}(s,\, t) = \epsilon \nabla^2 \mathbf{x}(s,\, t)$$
$$\mathbf{x}(s,\, 0) = \mathbf{x}_0(s)$$

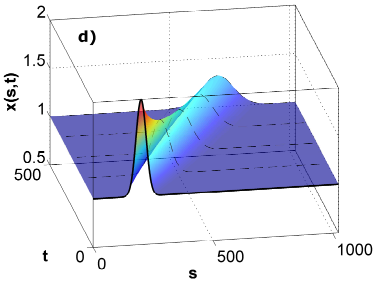


## Advection-Diffusion Equation

- Flat and Quadratic Top-hat (sparsity in wavelet)
- Window sinusoid and Squared Exponential (sparsity in DCT)

$$\frac{\partial \mathbf{x}(s,\, t)}{\partial t} + a\nabla \mathbf{x}(s,\, t) = \epsilon \nabla^2 \mathbf{x}(s,\, t)$$
$$\mathbf{x}(s,\, 0) = \mathbf{x}_0(s)$$

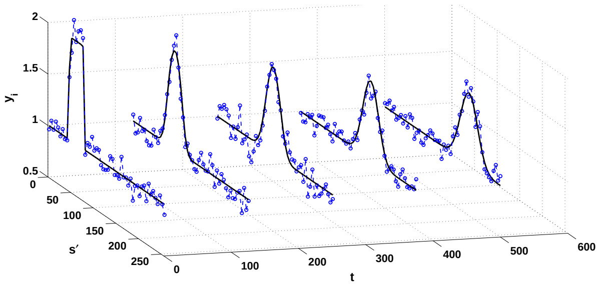


# System Equations

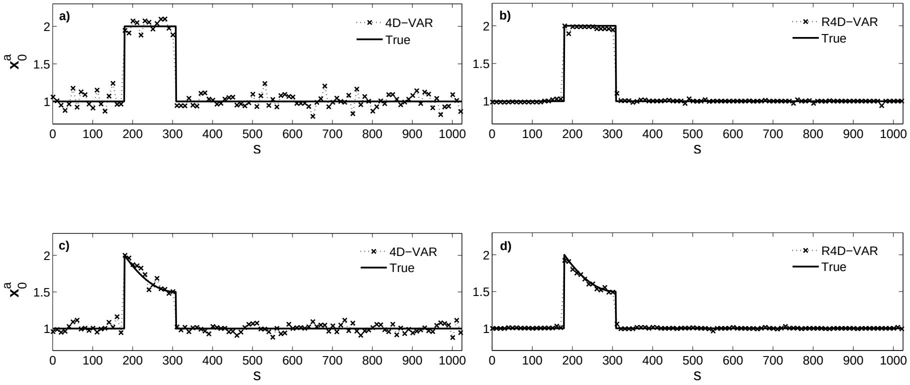
- **Model Equation:** $\mathbf{x}^i = \mathbf{M}_{0,i}\mathbf{x}^0$, where $\mathbf{M}_{0,i} = \mathbf{A}_{0,i}\mathbf{D}_{0,i}$
- **Observation Model:** $\mathbf{y}^i = \mathbf{H}\mathbf{x}^i + \mathbf{v}$ , with $\mathbf{v} \sim \mathcal{N}\left(0,\, \mathbf{R}\right)$

$$\mathbf{H} = \frac{1}{4}\begin{bmatrix} 1\,1\,1\,1 & 0\,0\,0\,0 & \cdots & 0\,0\,0\,0 \\ 0\,0\,0\,0 & 1\,1\,1\,1 & \cdots & 0\,0\,0\,0 \\ \vdots & \vdots & \vdots & \vdots \\ 0\,0\,0\,0 & 0\,0\,0\,0 & \cdots & 1\,1\,1\,1 \end{bmatrix} \in \mathbb{R}^{n\times m}$$



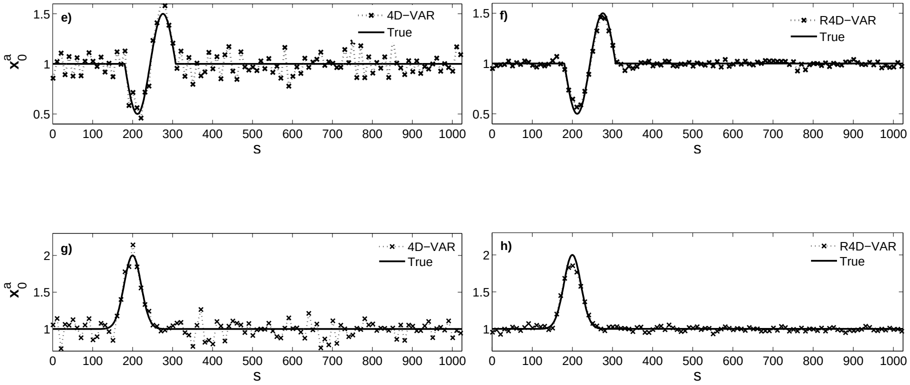
## White Background Error

- ($\mathbf{B} = \sigma_b^2 \mathbf{I}$, $\mathbf{R} = \sigma_r^2 \mathbf{I}$ ), where $\sigma_b = 0.10$ (SNR $\cong 10.5$ dB) and $\sigma_r = 0.08$ (SNR $\cong 12$ dB)



# White Background Error

- $(\mathbf{B} = \sigma_b^2\mathbf{I},\ \mathbf{R} = \sigma_r^2\mathbf{I}\ )$, where $\sigma_b = 0.10$ $(\mathrm{SNR} \cong 10.5$ dB) and $\sigma_r = 0.08$ $(\mathrm{SNR} \cong 12$ dB)



## White Background Error

- $(\mathbf{B} = \sigma_b^2 \mathbf{I},\ \mathbf{R} = \sigma_r^2 \mathbf{I}\ )$, where $\sigma_b = 0.10$ $(\mathrm{SNR} \cong 10.5$ dB) and $\sigma_r = 0.08$ $(\mathrm{SNR} \cong 12$ dB)

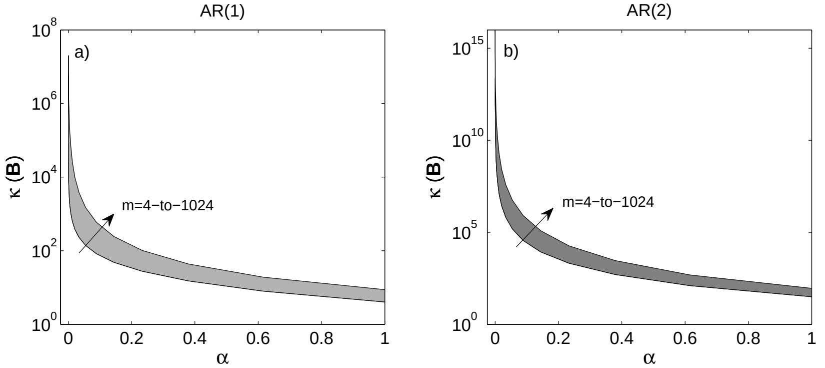
| White Background Error | | | | | | |
|---|---|---|---|---|---|---|
| | $\mathrm{MSE}_r$ | | $\mathrm{MAE}_r$ | | $\mathrm{BIAS}_r$ | |
| | R4D-Var | 4D-Var | R4D-Var | 4D-Var | R4D-Var | 4D-Var |
| **FTH** | 0.0188 | 0.0690 | 0.0099 | 0.0589 | 0.0016 | 0.0004 |
| **QTH** | 0.0152 | 0.0515 | 0.0083 | 0.0414 | 0.0030 | 0.0016 |
| **WS** | 0.0296 | 0.0959 | 0.0229 | 0.0771 | 0.0038 | 0.0022 |
| **SE** | 0.0316 | 0.0899 | 0.0235 | 0.0728 | 0.0018 | $4.26\mathrm{e}-5$ |

Table : Expected values of the $\mathrm{MSE}_r$, $\mathrm{MAE}_r$, and $\mathrm{BIAS}_r$, for 30 independent runs.

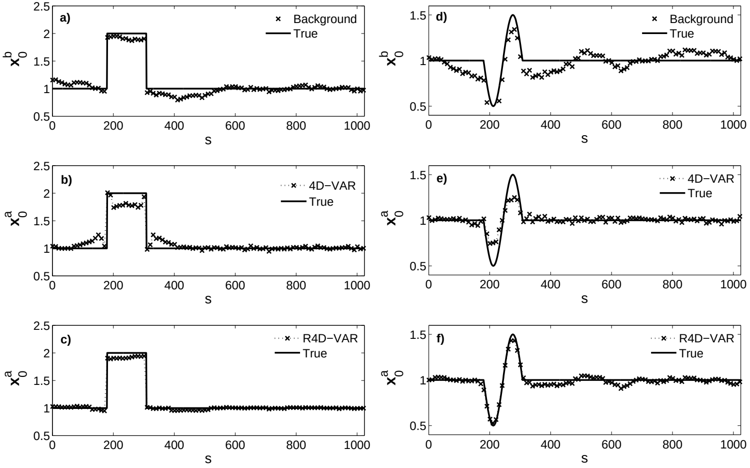
# Correlated Background Error

AR(1): $\mathbf{B}_{ij} = e^{-\alpha|i-j|}$

AR(2): $\mathbf{B}_{ij} = e^{-\alpha|i-j|}\left(1+\alpha\,|i-j|\right)$

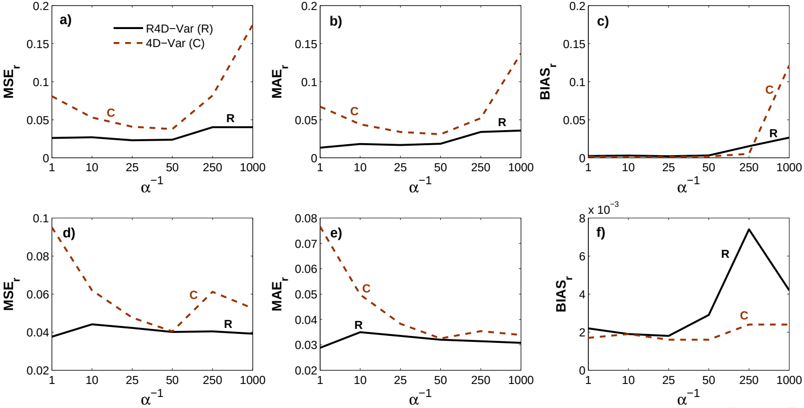


# Correlated Background Error–AR(1)



# Correlated Background Error–AR(1)

- Top panel: FTH – Bottom panel: WS
- $(\mathbf{B} = \sigma_b^2 \mathbf{C}_b,\ \mathbf{R} = \sigma_r^2 \mathbf{I}\ )$, where $\sigma_b = 0.10$ (SNR $\cong 10.5$ dB) and $\sigma_r = 0.08$ (SNR $\cong 12$ dB)



# Thank You