Sequential Monte Carlo Techniques for Spatiotemporal Reconstruction of Density and Temperature in the Solar Corona

Farzad Kamalabadi
University of Illinois at Urbana-Champaign and NSF

May 4, 2010 – ADA6
Joint Work with:
Mark Butala, Yuguo Chen, Russell Hewett, and Richard Frazin
1. Motivation for 3D Imaging of the Solar Corona
2. Assimilative (empirical) Models
   - from $pB$ to 3D $N_e$
   - from EUV to $T$: Differential Emission Measure Tomography
3. Dynamic Model: General Framework
4. Linear Framework
   - The ensemble Kalman filter (EnKF)
   - Covariance tapering (regularization)
   - Convergence of the EnKF
5. Numerical examples and solar data image reconstruction
Motivation: Coronal Heating

The physical processes that heat and accelerate the solar wind are not well understood.
A solar wind shock wave caused by a major solar flare or coronal mass ejection (CME) can cause a geomagnetic storm, a disturbance of Earth’s near space environment. Such storms can damage satellites, disrupt power distribution, and can harm space operations (e.g., unprotected astronauts). Empirically derived 3D $N_e$ and $T$ estimates are used to verify and improve physical models of the corona and solar wind.
Corona As Observed During A Solar Eclipse

(Photos courtesy of the HAO and Rhodes College)

(a) Solar minimum.  
(b) Solar maximum.

______________________________
http://mlso.hao.ucar.edu/eclipse80N94.html
Polarized Brightness Coronagraph Images

Data source: Mk4 coronagraph.
Sampling rate: 1 time averaged image/day.
Image resolution: 960 × 960 pixels (65% are data).
Field of view: 1.12 – 2.79 R⊙.
Pixel size: 4,350 km (at solar surface).
Figure: Artist's rendition of the dual-spacecraft STEREO with overlapping field of view (FOV) aimed at imaging 3D solar atmosphere and transient eruptions known as the coronal mass ejections.
Each pixel of a white-light, polarized brightness (pB) coronagraph image is proportional to $N_e$ integrated along the pixel’s line of sight.

The relationship is given by:

$$pB = C \int_{-\infty}^{\infty} H(s + t\theta) N_e(s + t\theta) dt$$

which approximates to:

$$\approx C \sum_{j=1}^{N} H(s + t_j\theta) x_j \Delta t_j$$

leading to:

$$y_i \approx H_i x + v_i$$
Exploitation of Solar Rotation

Solar rotation provides a unique set of line integral measurements of $N_e$ over a 14 day period (corresponding to 180° of rotation).

\[ y_1 = H_1 x + v_1 \]
\[ y_2 = H_2 x + v_2 \]

\[
\begin{pmatrix}
  y_1 \\
  \vdots \\
  y_{14}
\end{pmatrix}
= 
\begin{pmatrix}
  H_1 \\
  \vdots \\
  H_{14}
\end{pmatrix}
\begin{pmatrix}
  x \\
  \vdots \\
  v_{14}
\end{pmatrix}
\implies
y = Hx + v
\]
The (static) solar tomography inverse problem formulation:

\[ \hat{x}_{\text{static}} = \arg \min_{x \geq 0} \Psi(x) \]

The cost function:

\[ \Psi(x) = \| y - Hx \|_{W}^{2} + \sum_{i} \gamma_{i} C_{i}(x) \]  \hspace{1cm} (1)

\[ \| y - Hx \|_{W}^{2} \] : weighted residual norm,
\[ C_{i} \text{ and } \gamma_{i} \] : \( i \)-th regularization functional and regularization parameters.
Spherical Shell Cuts II (@ $r_0 = 1.23 \, R_\odot$)

Reconstructed $N_e$ on a spherical grid and comparison with EIT image.
Spherical Shell Cuts II (@ $r_0 = 1.23 \, R_\odot$)

EIT 171, 5/12/05 19:00

Ne at 1.23, same POV, Mk4
Comparison to Other $N_e$ Studies

Comparison between the reconstructed $N_e$ in various regions at $1.2 \, R_\odot$ and the corresponding $N_e$ determined from several other studies

<table>
<thead>
<tr>
<th>Region</th>
<th>Reconstructed $N_e / (10^7 , \text{cm}^{-3})$</th>
<th>Comparative $N_e / (10^7 , \text{cm}^{-3})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active region</td>
<td>19. – 38.</td>
<td>&gt;18.</td>
</tr>
<tr>
<td>Polar hole</td>
<td>0.5 – 4.4</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Equatorial hole</td>
<td>1.5 – 8.4</td>
<td>N/A</td>
</tr>
<tr>
<td>“Quiet” region</td>
<td>4. – 15.</td>
<td>16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.1</td>
</tr>
</tbody>
</table>
Differential emission measure analysis is used to determine a temperature distribution of a plasma from a set of spectra **without** attempting to determine spatial characteristics. A plasma emissivity model, e.g., CHIANTI (Young et al., 2003), is required to compute emissivity as a function of wavelength and temperature.

**Goal:** Combine the 3D imaging capabilities of SRT and the temperature resolving capabilities of DEM to estimate the 3D temperature distribution of a plasma (DEMT).

**Input:** A time series of EUV images in a set of frequency bands.

**Output:** $N_e^2(r_i, T_l)$, where $r_i$ is the location of the center of the $i$th voxel and $T_l$ is the $l$th temperature bin.
The differential emission measure is a function of the following:

$$\text{DEM}(r_i, T_l) = \int_{\Omega_{i,l}} W(r) N_e^2(r) \, dr$$

- $r_i$: Location of the center of the $i$th voxel.
- $T_l$: The $l$th temperature bin.
- $\Omega_{i,l}$: Indicator function of the $i$th voxel and the $l$th temperature bin (equal to 1 inside the $i$th voxel where the temperature falls within the $l$th bin and 0 otherwise).
- $W(r)$: The abundance of the ion in question.
- $N_e^2(r)$: The squared electron density.
The discretized equation relating the input to output is:

\[ Y_{k,j} = \sum_i A_{j,i} \sum_l Q_{k,l} \text{DEM}(r_i, T_l) + n_{k,j} \]

- **\( Y_{k,j} \)**: The measured intensity of the \( j \)th pixel in the \( k \)th spectral band.
- **\( A_{j,i} \)**: The length of the \( j \)th line of sight in the \( i \)th voxel.
- **\( Q_{k,l} \)**: The DEM kernel, constructed from a plasma emission model (CHIANTI) and the bandpass function of the \( k \)th spectral filter of the instrument measuring \( I_{k,j} \).
- **\( n_{k,j} \)**: A noise term accounting for all uncertainty in the measurement of the \( j \)th pixel in the \( k \)th spectral band.
The discretized input/output equation can then be expressed as a matrix equation \( y = Ax + n \). However, for DEMT, \( A \) is block diagonal and can be separated into two sets of equations.

**Equation 1:** Solve for the filter band emissivity (FBE) given the EUV intensity in the \( k \)th band:

\[
y_k = Wz_k + n_k, \quad 1 \leq k \leq K
\]

- \( y_k \): The measured EUV intensity in the \( k \)th spectral band stacked into a vector.
- \( W \): The line of sight weights \( A_{j,i} \) arranged into a matrix.
- \( z_k \): The FBE, i.e., the values \( \sum_l Q_{k,l} \text{DEM}(r_i, T_l) \) with fixed spectral band index \( k \) stacked into a vector.
- \( n_k \): The measurement noise (\( n_{k,j} \) in vector form).
**Equation 2:** Solve for $\text{DEM}(r_i, T_l)$ given the filter band emissivities in the $K$ spectral bands:

$$ f_i = Qd_i + v_i, \ 1 \leq i \leq I $$

- $f_i$: The FBE in the $i$th voxel, i.e., the FBE values $z_{k,i}$ with fixed voxel index $i$ stacked into a vector.
- $Q$: The DEM kernel values $Q_{k,l}$ (computed from the CHIANTI plasma emission model) arranged into a matrix.
- $d_i$: The local DEM vector, i.e., the $l$th component of $d_i$ is equal to $\text{DEM}(r_i, T_l)$ and has $L$ (the number of temperature bins) elements.
- $v_i$: Noise in the estimated FBEs - the result of uncertainties in the tomographic inversion and noise in the measurement of the EUV intensities.
• The data source is the extreme ultraviolet imager (EIT) on SOHO.
• Data are measured at wavelengths 171, 195, and 284 Å. One month of data were used starting at 10/97.
• 3D reconstructions were performed for each bandpass, effectively undoing the LOS integrations in the data.
• DEM results from these 3D reconstructions as well as reconstructions from EUVI are in Frazin, Valsquez, Kamalabadi, ”Quantitative, 3D Analysis of the Global Corona with Multi-Spacecraft Differential Emission Measure Tomography”, in press, ApJ 09.
3D Reconstruction From EIT Data @ 171 Å

1.01 Rs

1.09 Rs
3D Reconstruction From EIT Data @ 195 Å
3D Reconstruction From EIT Data @ 284 Å
Goal:

To estimate the properties of a hidden Markov random process denoted \( \{x_i, i \in \mathbb{N}\} \) at each time index \( i \) given:

- A set of statistically related (tomographic projection) measurements \( y_{1:i} = \{y_1, \ldots, y_i\} \).
- A statistical model for the time evolution of the hidden state process.

The main twist in this work is that each random vector \( x_i \) is of enormous dimension, perhaps \( > 1 \) million elements!
The general hidden Markov model (HMM):

- **Initial prior:** \( p_{x_1}(x_1) \) (2)
- **Measurement/forward model:** \( h_i(y_i|x_i) \) (3)
- **State-transition model:** \( f_i(x_{i+1}|x_i) \) (4)

\[
\text{dim}(x_i) = N \quad \text{dim}(y_i) = M
\]

**Goal:** Compute minimum mean square error (MMSE) estimates of the unknown state \( x_i \) given the measurements \( y_{1:j} \triangleq \{y_1, \ldots, y_j\} \).

\[
\hat{x}_{i|j} \triangleq \mathbb{E}[x_i|y_{1:j}] = \int x_i p(x_i|y_{1:j}) \, dx_i
\] (5)
### Filtering, Smoothing, and Prediction

| Estimate \( (\hat{x}_{i|j}) \) | Data \( (y_{1:j}) \) | Purpose |
|----------------------------------|-----------------|---------|
| Filtered \( j = i \)            |                 | Online processing – estimates are based on currently available data. |
| Predicted \( j < i \)           |                 | For forecasting the future evolution of the dynamic process. |
| Smoothed \( j > i \)            |                 | Offline processing – estimates based on all available information. |

- Smoothed and predicted estimates are typically computed by further processing of the filtered estimates \( \hat{x}_{i|i} \).

- Filtered (posterior) estimates \( \hat{x}_{i|i} \) may be recursively computed using the previous one-step prediction (prior) \( \hat{x}_{i|i-1} \) \( j \in \{i, i - 1\} \) for this talk.
Difficulties With The General Model

\[ \hat{x}_{i|i} \triangleq \mathbb{E}[x_i | y_{1:i}] = \int x_i p(x_i | y_{1:i}) \, dx_i \]

Problem: The posterior PDF \( p(x_i | y_{1:i}) \) cannot generally be found in closed form, and numerical approximation of the posterior and evaluation of the conditional mean requires \( N \) dimensional quadrature.

Quadrature requires an exponential increase in computational effort with the dimension of the state (i.e., the curse of dimensionality)!
Recall that the posterior PDF $p(x_i | y_{1:i})$ cannot generally be found in closed form.

- The particle filter (PF) circumvents the analytically intractable posterior distribution $p(x_i | y_{1:i})$ by sampling from a known proposal distribution $\pi(x_i | x_{i-1}, y_{1:i})$.

- The PF updates each particle, a proposal sample $x^l_i$ and importance weight $w^l_i$ pair, such that the weighted sample mean $\bar{x}_{i|i}$ converges to the MMSE estimate, i.e.,

$$
\bar{x}_{i|i} \triangleq \frac{\sum_{l=1}^{L} w^l_i x^l_i}{L \sum_{l=1}^{L} w^l_i} \xrightarrow{L \to \infty} \mathbb{E}[x_i | y_{1:i}]
$$

(6)

- PFs generally require a huge number of particles to adequately “cover” the probability space. They also suffer from particle impoverishment (all weight assigned to one particle after only a very few measurement updates).
The (Most Simple) Particle Filter

The particle filter (PF) is a sequential Monte Carlo method for approximating MMSE estimates.

- Importance sampling is used to avoid intractable posterior PDFs.
- Resampling (the systematic replication of particles with large importance weight and elimination of insignificantly weighted particles) must be used to mitigate particle impoverishment.
Large Dimensional Particle Filters

Standard PFs become computationally prohibitive when the state dimensions is \( N > 10 \) or \( N > 100 \) (an implication of the curse of dimensionality).

Some work in applying PFs to large dimensional problems include:

- Dimension reduction (change of basis). Only possible when the unknown process is stationary and a single fixed basis may be chosen \textit{a priori}.

- Separation of the state into nearly constant (narrow PDF) and more dynamic (broad PDF) components. (See N. Vaswani, “Particle Filtering for Large Dimensional State Spaces with Multimodal Observation Likelihoods,” IEEE Trans. Signal Processing,
Linear Additive-Noise State-Space Signal Model (Linear Gaussian Model)

Initial prior: \( \mathbb{E}[x_1] = \mu_1, \ \text{Cov}(x_1) = \Pi_1 \) \hspace{1cm} (7)

Measurement/forward model:
\( y_i = H_i x_i + v_i \) \hspace{1cm} (8)

State-transition model:
\( x_{i+1} = F_i x_i + u_i \) \hspace{1cm} (9)

- The first and second order statistics of the zero mean state \( (u_i) \) and measurement \( (v_i) \) noise are given: \( \text{Cov}(u_i) = Q_i \) and \( \text{Cov}(v_i) = R_i \).

Goal: Compute linear minimum mean square error (LMMSE) estimates of the unknown state \( x_i \) given the measurements \( y_{1:j} \).
The Kalman Filter

\[ \hat{x}_{i+1|i} = F_i \hat{x}_{i|i} \]
\[ P_{i+1|i} = F_i P_{i|i} F_i^T + Q_i \]

The Kalman Filter

\[ \begin{align*}
K_i &= P_{i|i-1} H_i^T \left( H_i P_{i|i-1} H_i^T + R_i \right)^{-1} \\
\hat{x}_{i|i} &= \hat{x}_{i|i-1} + K_i \left( y_i - H_i \hat{x}_{i|i-1} \right) \\
P_{i|i} &= P_{i|i-1} - K_i H_i P_{i|i-1}
\end{align*} \]

Measurement Update (\( \hat{x}_{i|i-1} \rightarrow \hat{x}_{i|i} \))

The \( N \times N \) matrix \( P_{i\mid j} \) is the estimator error covariance:

\[ P_{i\mid j} \triangleq \text{Cov}(x_i - \hat{x}_{i\mid j}) \]  \hspace{1cm} (10)

Almost 2 TB of computer memory is required to store \( P_{i\mid j} \) when the state dimension \( N = 10^6 \) and all operations involving \( P_{i\mid j} \) become prohibitively computationally costly!
Ensemble Kalman Filter (EnKF)

**Idea:** Update an ensemble of samples \( \{\tilde{x}^1_{i|j}, \ldots, \tilde{x}^L_{i|j}\} \) such that the sample mean and covariance approximate the KF estimate \( \hat{x}_{i|j} \) and estimator error covariance \( P_{i|j} \).

The ensemble size \( L \) is a trade-off between the estimate quality and computational effort.
• The EnKF processes an ensemble of $L$ samples of the $N$ dimensional state.

• $LN$ memory units are required to store the ensemble and the EnKF requires more memory than the KF if $L > N$.

$\Rightarrow$ A small ensemble is desirable.

• The variance (expected error) in the EnKF estimates $\tilde{x}_{i|i}$ decreases as the ensemble size $L$ is increased, i.e., the expected sample error in $\tilde{x}_{i|i}$ is controlled by the ensemble size $L$.

$\Rightarrow$ The ensemble size cannot be too small!
The primary source of sample error is due to the sample error covariance \( \tilde{P}_{i|i-1} \) which is **dominated** by sample error when the ensemble size \( L \) is small.

- **Idea:** Instead of the sample covariance \( \tilde{P}_{i|i-1} \), use the regularized covariance estimate \( C_i \circ \tilde{P}_{i|i-1} \) where \( \circ \) is the Hadamard or element-by-element matrix product.

- The proper choice of covariance taper \( C_i \) will greatly reduce the variance in \( \tilde{x}_{i|i} \) while introducing only a small bias. In other words, covariance tapering can reduce the overall (sample + bias) estimate error when the sample size \( L \) is small.

- We can compute estimates with a small ensemble size \( L \) and not pay a serious penalty (for certain problems)!
Monte Carlo Estimators

(a) Unbiased, large variance

(b) Small systematic bias, small variance

(c) Large systematic bias, small variance

- For certain problems, it is possible to greatly reduce the variance of an estimator by introducing a small bias.

- The process can reduce the mean square error

\[
\hat{x} = \arg \min_z \mathbb{E}[\|x - z\|^2] \tag{11}
\]
Convergence Of The Ensemble Kalman Filter

**Theorem:**
The EnKF estimates \( \tilde{x}_{i|i} \) converge in probability to the estimates \( \tilde{x}_{i|i}^{\infty} \) given by the localized KF (LKF), i.e.,

\[
\tilde{x}_{i|i} \overset{p.}{\rightarrow} \tilde{x}_{i|i}^{\infty}
\]  \hspace{1cm} (12)

in the limit as the ensemble size \( L \rightarrow \infty \).

- When there is no covariance tapering \( (C_i = 1_{N \times N}) \), the EnKF estimates converge to the KF estimates, i.e., \( \tilde{x}_{i|i} \overset{p.}{\rightarrow} \hat{x}_{i|i} \).

The random process \( \{x_i, i \in \mathbb{N}\} \) is a highly dynamic simulated “movie” with each frame equal to a 2D pixel image.

A low and high resolution sampling of the process are considered.

<table>
<thead>
<tr>
<th>Frame resolution</th>
<th># of frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low resolution: 32 x 32 ⇒ ( N = 1024 )</td>
<td>256</td>
</tr>
<tr>
<td>High resolution: 128 x 128 ⇒ ( N = 16,384 )</td>
<td>1024</td>
</tr>
</tbody>
</table>

Each measurement \( y_i \) is a set of parallel line integral measurements (46 at low-res and 184 at high-res) at an angle \( \theta_i \) with 1% additive white Gaussian noise.
Numerical Experiment - The Setup

- The state-transition operator is chosen as $F_i = I$, i.e, the temporal evolution of $x_i$ is modeled as a random walk.

- The state noise covariance $Q_i$ is constructed to be positive definite and has a similar spatially-localized structure as the covariance taper matrix.

- In this experiment, $Q_i$ has 9 bands and each pixel in the state noise is correlated to immediate neighbors.

- Only the diagonal of the sample error covariance is used, i.e., the covariance taper is the identity matrix $C_i = I$. 
Numerical Experiment: $32 \times 32$

$$N = 32^2 = 1024$$

$\mathbf{x}_{64}$  $\mathbf{x}_{128}$  $\mathbf{x}_{192}$  $\mathbf{x}_{256}$

$\mathbf{y}_i$  ($M = 46$)
Numerical Experiment: \(32 \times 32\)

<table>
<thead>
<tr>
<th></th>
<th>(L)KF</th>
<th>EnKF ((L = 128))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time:</td>
<td>40m</td>
<td>26s</td>
</tr>
<tr>
<td>Memory:</td>
<td>2 MB</td>
<td>0.5 MB</td>
</tr>
</tbody>
</table>
Numerical Experiment: $32 \times 32$

<table>
<thead>
<tr>
<th></th>
<th>(L)KF</th>
<th>EnKF ($L = 128$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time:</td>
<td>40m</td>
<td>26s</td>
</tr>
<tr>
<td>Memory:</td>
<td>2 MB</td>
<td>0.5 MB</td>
</tr>
</tbody>
</table>

Truth $x_{128}$

EnKF $\tilde{x}_{128|128}$

LKF $\tilde{x}_{128|128}^\infty$

KF $\hat{x}_{128|128}$
Numerical Experiment: $32 \times 32$

<table>
<thead>
<tr>
<th></th>
<th>(L)KF</th>
<th>EnKF ($L = 128$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time:</td>
<td>40m</td>
<td>26s</td>
</tr>
<tr>
<td>Memory:</td>
<td>2 MB</td>
<td>0.5 MB</td>
</tr>
</tbody>
</table>

$\mathbf{x}_{192}$

Truth

EnKF $\tilde{\mathbf{x}}_{192|192}$

LKF $\tilde{\mathbf{x}}_{192|192}^\infty$

KF $\hat{\mathbf{x}}_{192|192}$
Numerical Experiment: $32 \times 32$

<table>
<thead>
<tr>
<th></th>
<th>(L)KF</th>
<th>EnKF ($L = 128$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time:</td>
<td>40m</td>
<td>26s</td>
</tr>
<tr>
<td>Memory:</td>
<td>2 MB</td>
<td>0.5 MB</td>
</tr>
</tbody>
</table>
Numerical Experiment: $32 \times 32$

Performance comparison:

\[ e_i \triangleq \frac{\| x_i - z_i \|_2}{\| x_i \|_2} \] (13)

\( z_i \) is equal to one of (KF) \( \tilde{x}_{i|i} \), (LKF) \( \tilde{x}_{i|i}^\infty \), or (EnKF) \( \tilde{x}_{i|i} \)
Numerical Experiment: $128 \times 128$

$$N = 128^2 = 16,384$$

$\mathbf{x}_{256}$  

$\mathbf{x}_{512}$  

$\mathbf{x}_{768}$  

$\mathbf{x}_{1024}$  

$y_i$ \hspace{1cm} ($M = 184$)
Numerical Experiment: $128 \times 128$

Truth

$x_{256}$

$x_{512}$

$x_{768}$

$x_{1024}$

EnKF $\tilde{x}_{256|256}$

EnKF $\tilde{x}_{512|512}$

EnKF $\tilde{x}_{768|768}$

EnKF $\tilde{x}_{1024|1024}$
### Numerical Experiment: $128 \times 128$

<table>
<thead>
<tr>
<th>Computational requirements:</th>
<th>(L)KF (extrapolated)</th>
<th>EnKF ($L = 256$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time:</td>
<td>576d</td>
<td>6.0h</td>
</tr>
<tr>
<td>Memory:</td>
<td>512 MB</td>
<td>16 MB</td>
</tr>
</tbody>
</table>

⇒ Forget about the KF for large dimensional problems!

- We also considered an extremely low dimensional $6 \times 6$ ($N = 36$) sampling of the process.

- The minimum degeneracy PF required 2048 particles to match the same error as a 16 ensemble member EnKF.

⇒ Forget about the PF for large dimensional problems!
• We focus on October and November, 2003 (Halloween solar storm).

• Spherical reconstruction grid is $20 \times 30 \times 60$ ($N = 3.6 \times 10^4$). State of the art static reconstructions are $N = 2.9 \times 10^5$.

• Spherical reconstruction grid ranges from $1.1–2.85 \ R_\odot$ (where the solar radius $R_\odot = 7 \times 10^8$ m).

• 10 m computation for each static reconstruction and 17 m total computation for EnKF.

• Each static reconstruction is computed using a sliding window of solar measurements:

$$\begin{align*}
\hat{x}_{i}^{\text{static}} &= \arg \min_{x_i \geq 0} \| y_{(i-13):i} - H_{(i-13):i} x_i \|_2^2 + \lambda \| D x_i \|_2^2 
\end{align*}$$

(14)
4-D Reconstruction - 10/15

(a) Static reconstruction          (b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
4-D Reconstruction - 10/21

(a) Static reconstruction

(b) EnKF reconstruction
(a) Static reconstruction  (b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction    (b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  (b) EnKF reconstruction
(a) Static reconstruction

(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  

(b) EnKF reconstruction
4-D Reconstruction - 11/3

(a) Static reconstruction
(b) EnKF reconstruction
(a) Static reconstruction       (b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
4-D Reconstruction - 11/8

(a) Static reconstruction

(b) EnKF reconstruction
4-D Reconstruction - 11/10

(a) Static reconstruction
(b) EnKF reconstruction
(a) Static reconstruction  (b) EnKF reconstruction
4-D Reconstruction - 11/12

(a) Static reconstruction
(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction

(b) EnKF reconstruction
(a) Static reconstruction
(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  
(b) EnKF reconstruction
(a) Static reconstruction  (b) EnKF reconstruction
(a) Static reconstruction

(b) EnKF reconstruction
(a) Static reconstruction  (b) EnKF reconstruction
Both the static and EnKF agree with the data.

\[
y_i^{\text{Static}} = H_i \hat{x}_i^{\text{Static}} \quad \text{and} \quad y_i^{\text{EnKF}} = H_i \tilde{x}_i |_{i} \quad (15)
\]
The first column shows the EUVI-A images in the 171, 195 and 284 Å bands taken near 08:00 on 28 April 2008. The second column shows projected spherical cuts the FBE at 1.035 $R_s$. The projection angle of the spherical cut is chosen to be identical to that of the images in the first column. The third and fourth columns are similar, except they show spherical cuts taken at 1.085 and 1.135 $R_s$. The last column is a synthetic image calculated by integrating the tomographic models along the line-of-sight and thus should reasonably match the images in the first column. In all cases, we show the logarithm of the displayed quantity. For comparison purposes, the first and last columns use a common color scale for each band (row). The black streaks seen in the reconstructions near some of the active regions are artifacts caused by the Sun’s temporal variability. The polar crown filaments that are the subject of this paper are clearly visible in most of the images.
Learning Model Parameters Of Dynamical Systems

• In most physical applications, first principles physics drives the choice for the model of the dynamical system.

• In any case, the dynamic model will only be partially and imperfectly known and may depend on unobservable parameters.

• In many practical scenarios it is necessary to estimate the state of the unknown object jointly with the unknown parameters of the dynamic model.
Initial prior: \[ E[x_1] = \mu_1, \quad \text{Cov}(x_1) = \Pi_1 \] (16)

Measurement/forward model: \[ y_i = H_i x_i + v_i \] (17)

State-transition model: \[ \theta_i \sim p(\cdot|\theta_{i-1}) \] (18)
\[ x_{i+1} = F(\theta_i) x_i + G(\theta_i) u_i \] (19)

• The conditional matrix \( G(\cdot) \) allows for mixing of the state noise \( u_i \).

• Both the state transition matrix \( F(\cdot) \) and the state noise mixing matrix \( G(\cdot) \) are dependent on the hidden Markov process \( \theta_i \).

• Simple example: binary parameter \( \theta_i \) which indicates if the Sun is in a low or high state of dynamic activity.
Graphical Model Representation Of The Conditional Linear Model

- Hidden variables are represented by black dots and observed variables by open circles.

- The graph is loop free $\Rightarrow$ exact inference is possible (in principle, analogous to Baum-Welsh estimation for basic HMMs). Not computationally feasible as the time index $i$ increases because all mixture components of the parameter sequence $\theta_i$ would have to be propagated.
Conclusions

• It is now possible to construct quantitative 3D time-dependent models of the coronal density and temperature using a variety of dedicated space and ground-based measurements.

• The key to achieving this is careful formulation of the relevant statistical inference problem, in conjunction with efficient algorithmic advancements such as sequential Monte Carlo estimation techniques.

• The formulation lends itself nicely to parallelization.

• The framework is general enough to allow for inclusion of other future data products.

• The conditional linear model has been successfully used for statistical image and video processing.

• Manifold learning methods from the machine learning community may be used for the inference of such models – new algorithms are required for high dimensional problems.

• Model selection and validation are open problems – the dynamic system is unknown and training sets are not available.
The Localized Kalman Filter

The EnKF estimates $\tilde{x}_{i|i}$ converge to the LKF estimates $\tilde{x}^\infty_{i|i}$ as the ensemble size $L$ increases. The LKF may be used to quantitatively study the bias introduced by covariance tapering. Note that the LKF is as computationally expensive as the KF.

- The KF and LKF have the same initialization and time update steps.

- The KF measurement update

$$\hat{x}_{i|i} = \arg \min_{x_i} \| y_i - H_i x_i \|^2 R_i^{-1} + \| x_i - \hat{x}_{i|i-1} \|^2 P_i^{-1}_{i|i-1}$$  \hspace{1cm} (20)

- The LKF measurement update

$$\tilde{x}^\infty_{i|i} = \arg \min_{x_i} \| y_i - H_i x_i \|^2 R_i^{-1} + \| x_i - \tilde{x}^\infty_{i|i-1} \|^2 \left( C_i \circ \tilde{P}^\infty_{i|i-1} \right)^{-1}$$  \hspace{1cm} (21)
Though the LKF operates only on certain bands of the error covariance $\tilde{P}_{i|i-1}^\infty$, it is not equivalent to the banded KF (BKF).

<table>
<thead>
<tr>
<th>Method</th>
<th>Initialization</th>
</tr>
</thead>
</table>
| LKF    | $\tilde{x}_{1|0}^\infty = \mu_1$  
         | $\tilde{P}_{1|0}^\infty = \Pi_1$ |
| BKF    | $\hat{x}_{1|0}^B = \mu_1$  
         | $P_{1|0}^B = C_i \circ \Pi_1$ |
### LKF ≠ Banded Kalman Filter

<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement update</th>
</tr>
</thead>
</table>
| **LKF** | $\begin{align*} K_i^\infty &= (C_i \circ P_{i|i-1}^\infty)H_i^T[H_i(C_i \circ P_{i|i-1}^\infty)H_i^T + R_i]^{-1} \\
\tilde{x}_{i|i} &= \tilde{x}_{i|i-1} + K_i( y_i - H_i \tilde{x}_{i|i-1} ) \\
\tilde{P}_{i|i} &= \tilde{P}_{i|i-1} - K_i H_i^T \tilde{P}_{i|i-1} - \tilde{P}_{i|i-1} H_i (K_i^\infty)^T \\
&\quad + K_i (H_i \tilde{P}_{i|i-1} H_i + R_i)(K_i^\infty)^T \end{align*}$ |
| **BKF** | $\begin{align*} K_i^B &= P_{i|i-1}^B H_i^T (H_i P_{i|i-1}^B H_i^T + R_i)^{-1} \\
\hat{x}_{i|i} &= \hat{x}_{i|i-1} + K_i^B( y_i - H_i \hat{x}_{i|i-1} ) \\
P_{i|i}^B &= C_i \circ (I - K_i^B H_i) P_{i|i-1}^B \end{align*}$ |
<table>
<thead>
<tr>
<th>Method</th>
<th>Time update</th>
</tr>
</thead>
<tbody>
<tr>
<td>LKF</td>
<td>( \tilde{x}_{i+1</td>
</tr>
<tr>
<td></td>
<td>( \tilde{P}_{i+1</td>
</tr>
<tr>
<td>BKF</td>
<td>( \hat{x}_{i+1</td>
</tr>
<tr>
<td></td>
<td>( P_{i+1</td>
</tr>
</tbody>
</table>

- The BKF error covariance \( P_{i|j}^B \) is always a banded matrix.
- The LKF error covariance \( \tilde{P}_{i|j}^\infty \) is a full matrix, but only certain bands are used in the calculation of the localized Kalman gain \( \tilde{K}_i^\infty \).