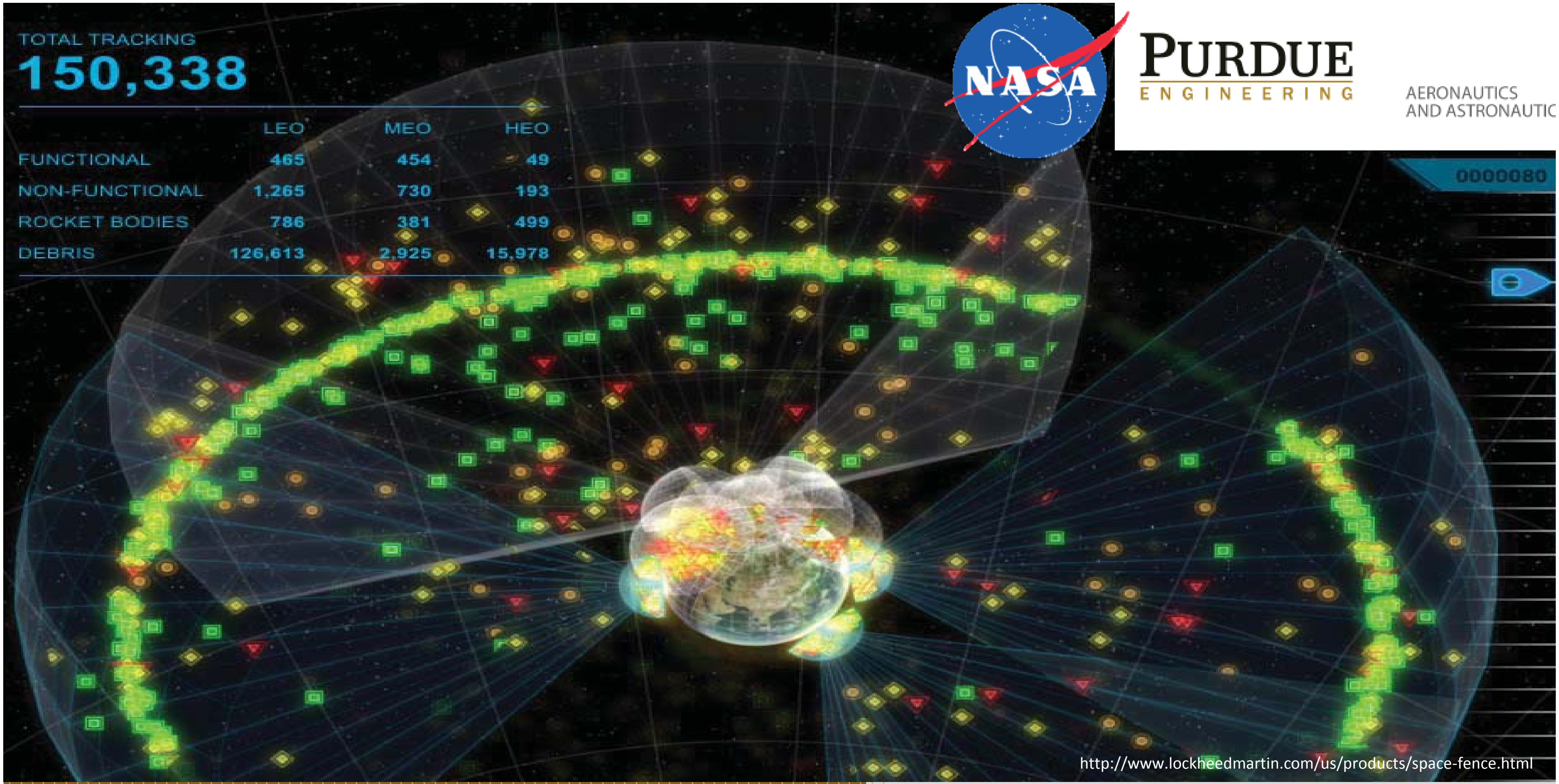
NASA

PURDUE ENGINEERING

AERONAUTICS AND ASTRONAUTIC



http://www.lockheedmartin.com/us/products/space-fence.html

INFORMATION MEASURES FOR STATISTICAL ORBIT DETERMINATION USING NONLINEAR FILTERS

Presenter: Dr. Alinda Mashiku

Collaborators:
Prof. James Garrison
Dr. Russell Carpenter

UMD Seminar Presentation
April 29th 2013

OUTLINE

OUTLINE

1. INTRODUCTION AND MOTIVATION:

RESEARCH SIGNIFICANCE

PROBLEM

In 2009, the U.S Strategic Command space objects (assets and debris) tracking update:

- ≈ 19,000 space objects (diameter >10 cm) are being tracked
- > 600,000 space objects (10cm >diameter>1 cm) are unobserved

Recent impacts have also added more objects (space debris) to track:

- Satellite Collision Russian Cosmos 2251 and the Iridium satellite in 2009
- Destruction of Chinese satellite in 2007

Artist rendering of space object collision, courtesy CNES

REQUIREMENTS

Space objects require:

- Collision avoidance mitigations
- Orbit maintenance/maneuvers (assets)
- Cataloging and identification (debris)

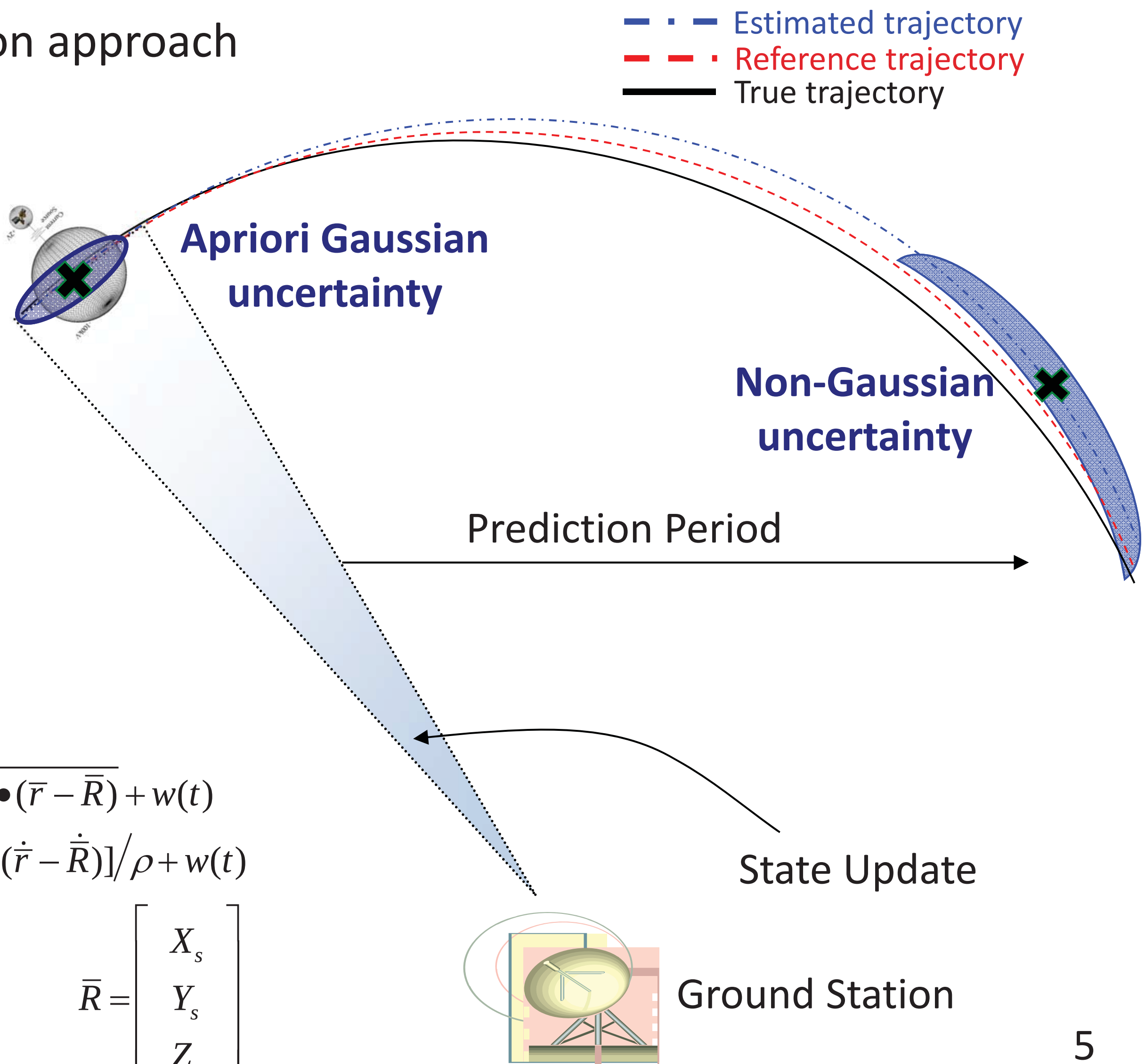
Specifics for the requirements:

- Performed within required accuracies
- *Cost effective*: Applicable to large numbers of space objects

STATISTICAL ORBIT DETERMINATION

CURRENTLY

Statistical Orbit Determination approach



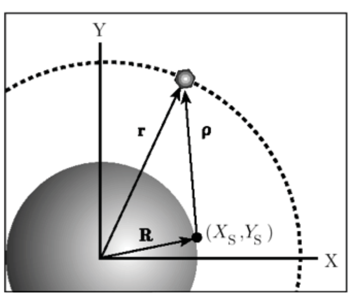
Dynamic model

$$\ddot{\bar{X}} = -\frac{\mu}{r^3}\bar{r} + \text{perturbing accelerations}$$

v(t)

$$\dot{\mathrm{X}} = \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \\ \dot{V}_x \\ \dot{V}_y \\ \dot{V}_z \end{bmatrix} + v(t) = \begin{bmatrix} V_x \\ V_y \\ V_z \\ -\mu X/r^3 \\ -\mu Y/r^3 \\ -\mu Z/r^3 \end{bmatrix} + v(t)$$

Measurement model



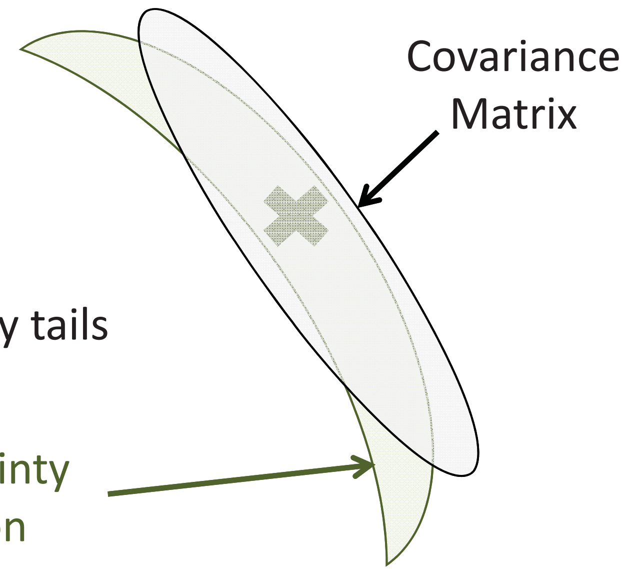
$$\rho = \sqrt{(\bar{r} - \bar{R}) \bullet (\bar{r} - \bar{R})} + w(t)$$

$$\dot{\rho} = [(\bar{r} - \bar{R}) \bullet (\dot{\bar{r}} - \dot{\bar{R}})]/\rho + w(t)$$

$$\bar{r} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \bar{R} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix}$$

1. INTRODUCTION AND MOTIVATION:

UNCERTAINTY CHARACTERIZATION



CURRENT LIMITATIONS

State estimate's uncertainty

- Covariance Matrix is not enough
- Gaussian representation ignores the information in the heavy tails

PROBLEM STATEMENT

1. Need **Full Probability Density Function (PDF) representation** for **low probability** events present in the **heavy tails** (Non-Gaussian)
2. Use a **nonlinear filter** that is capable of **full non-Gaussian PDF state estimation**
3. Need a **compressed representation** of this **PDF distribution** for real case scenario applications

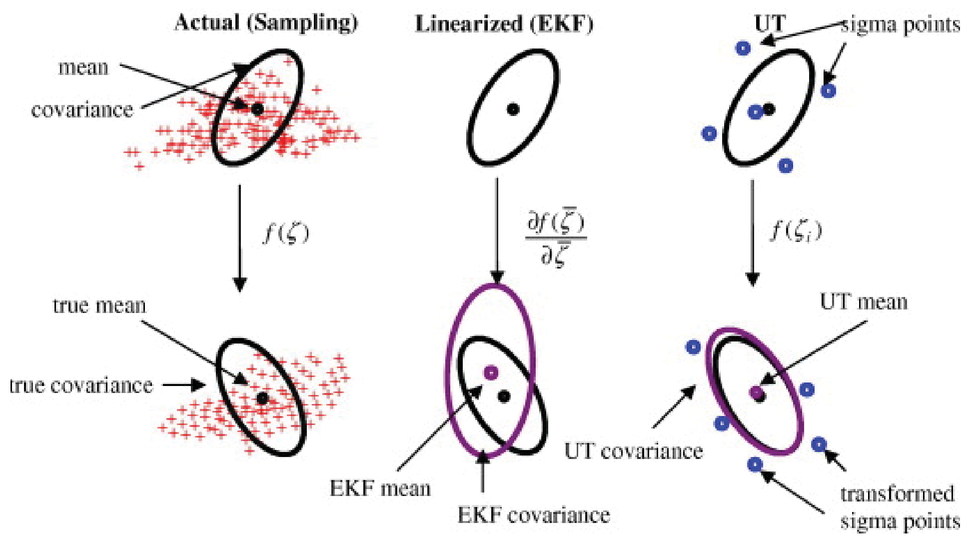
OUTLINE

1 Introduction and Motivation
2 Literature Review
- OD Current Estimation Methods (EKF and UKF)
- PDF approximation (GMM and PCE)
- Particle Filter
3 Approach
4 Research Contributions
5 Statistical Orbit Determination
6 PDF Compression
7 Simulations and Results
8 Conclusions and Future Work
9 Acknowledgements

OD CURRENT ESTIMATION METHODS

- ***Extended Kalman Filter*** (EKF) (Tapley, Shutz and Born 2004)
  - Nonlinear models assuming Gaussian uncertainties
  - Linearization about a current mean (computes Jacobian)

- ***Unscented Kalman Filter*** (UKF) (Julier and Uhlmann 1997 et al)
  - Uses a series of weighted sample points to approximate the mean and covariance



http://ars.els-cdn.com/content/image/1-s2.0-S0021999108000132-gr4.jpg

- ***Batch Weighted Least Squares*** (Gauss 1809)
  - Estimates the state x by minimizing the performance index $J(x) = \frac{1}{2}\varepsilon^T W \varepsilon$

PDF APPROXIMATION METHOD

- ***Gaussian Mixtures Model*** (GMM) (DeMars et al 2011)

$$p(x) = \sum_{i=1}^{L} w_i p_g(x; \mu_i, P_i)$$

where p_g is a Gaussian PDF

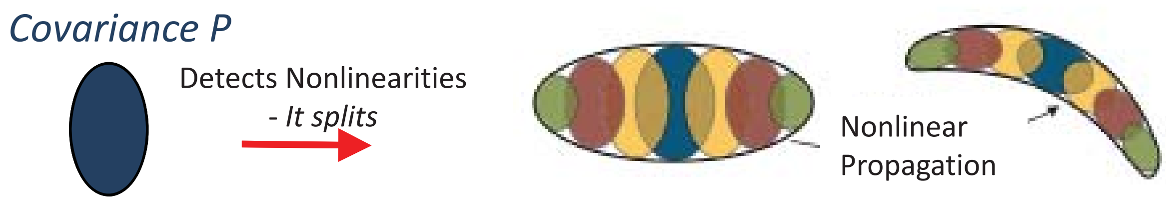


Figure Ref: J.T. Hartwood et al. Gaussian Sum filters for Space Surveillance: Theory and Simulations. *JGCD*, Vol. 34. No. 6. Nov-Dec 2011

  - Differential entropy d/dt[H(x)] is used as a measure of nonlinear detection
  - Uses UKF as a filter for each p_g

Entropy

$$H(x) = \frac{1}{2}\log|2\pi e P|$$

- ***Polynomial Chaos Expansion*** (PCE) (Jones et al 2012)
  - Solution of the Stochastic Differential Equations as linear expansions of multivariate polynomials
  - PCE estimates the coefficients c_α of the expansion

$$\hat{X}(t,\xi) = \sum_{\alpha \in \Lambda_{p,d}} c_\alpha(t)\psi_\alpha(\xi)$$

My approach:

Propose using the **Particle Filter**

- Fully captures **PDF (Probability Density Function)** distribution
- Uses a large number of samples to approach the **optimal estimate**

PARTICLE FILTER

The Particle Filter (PF) is a sequential nonlinear estimator that

- Uses random independent particles to represent the random state **x** as a PDF
- As N increases, we reach the optimal PDF representation

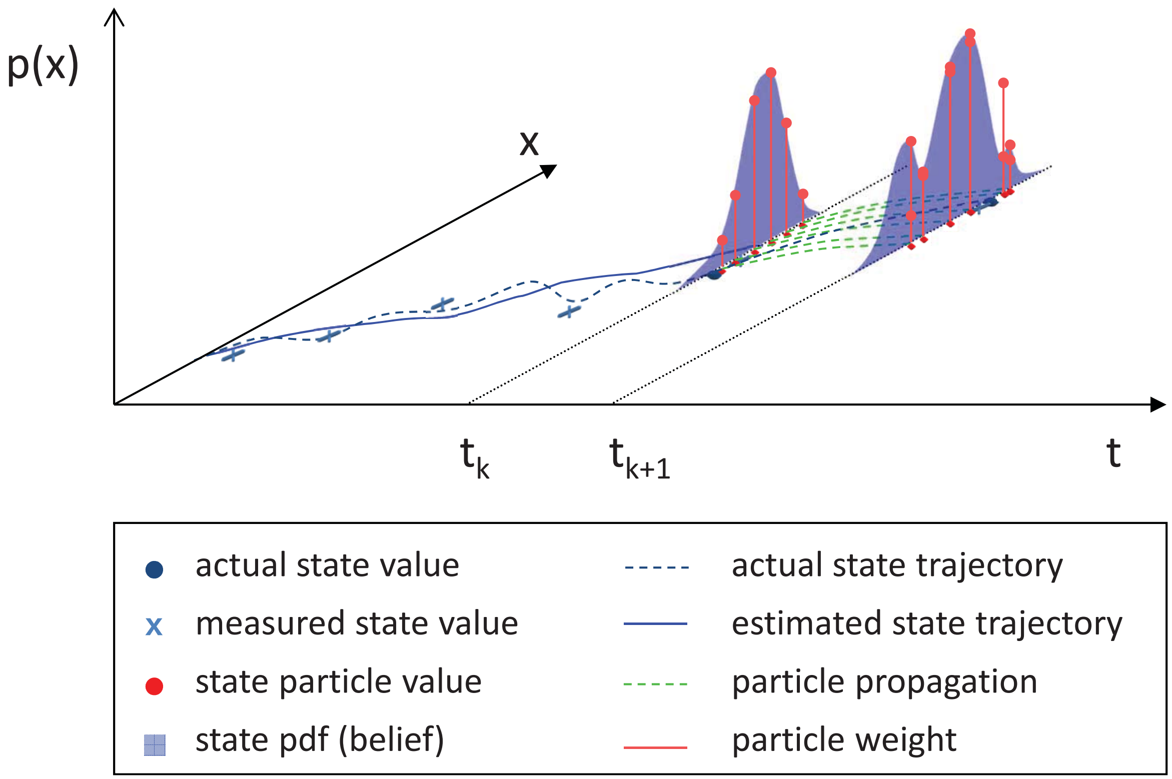
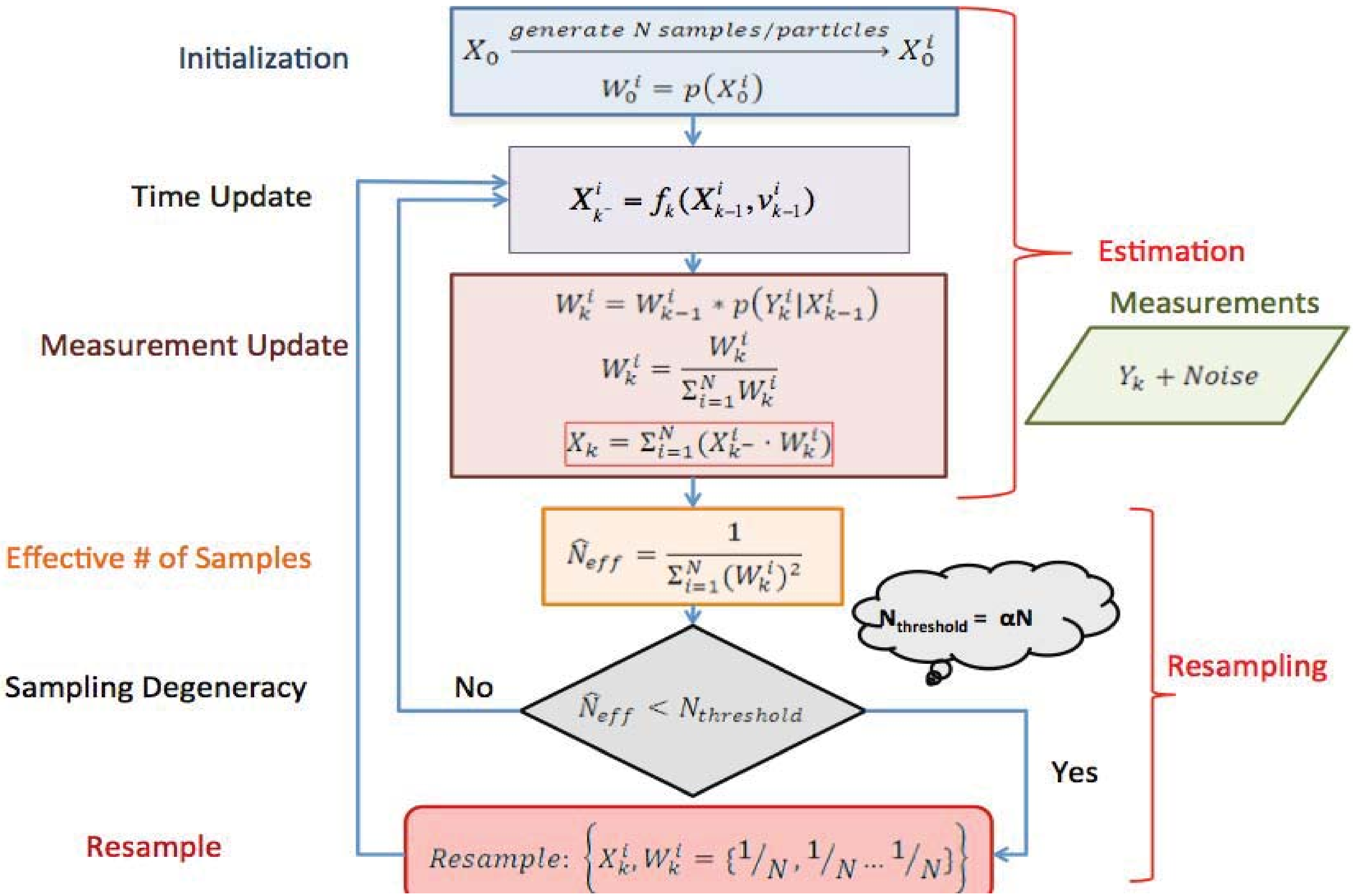


Figure Ref: Bhaskar Saha, Introduction to Particle Filter. Ames Research Center. April 16, 2008

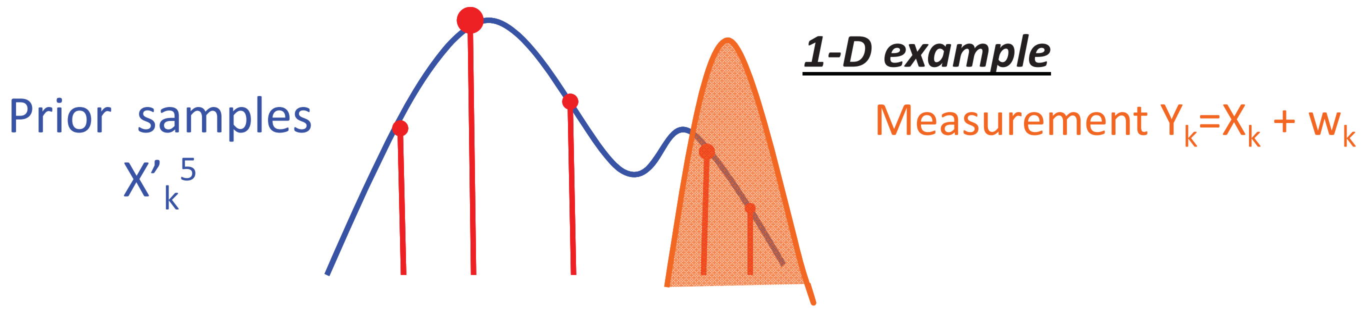
GENERIC PARTICLE FILTER ALGORITHM



PARTICLE FILTER

MAJOR DRAWBACK:

Sample Degeneracy when data outliers occur or when measurement noise is small



CIRCUMVENTION:

- **Resampling of particles** ($N_{effective\ samples} < N_{threshold}$)
- **Adaptive PF:** Soto (2005), Gang and Xiao-Jun (2008), Hwang and Speyer (2011).
- **Other PF Methods studied:** Auxiliary Particle Filter, Regularized Particle Filter, Rao-Blackwell/Marginalized Particle Filter

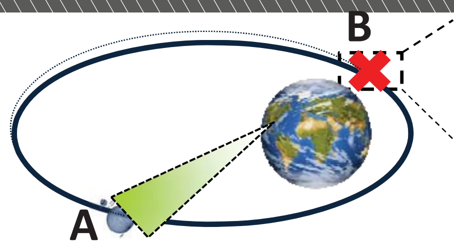
Our Approach for PF in Statistical Orbit Determination:

Used Generic Particle Filter

- Effective: sporadic measurements
- "Jitter" resampled particles by adding noise

OUTLINE

APPROACH



Multivariate Density

$p(X, Y, Z, V_X, V_Y, V_Z)$

(A) Particle Filter:

- ***Meas. Updates*** *and* ***Predictions*** *to* ***B***
- *Large number of particles -* ***compression***

(B)

Predicted State
L x N
(states$_{EC}$ x particles)

→ NFA +.. FastICA →

Uncorrelated (UC)
L x N
(states$_{UC}$ x particles)

→ Binning →

Probability Mass Fcn (PMF)
L x #bins X 2
(states$_{UC}$ x probabilities)

↓ WT/FFT + Compressions ↓

Coefficients for PMF
L x C
(states$_{UC}$ x coefficients)

↓ Inverse Transforms WT/FFT ↓

Compressed PMF
L x #bins + 3L
(states$_{UC}$ x probabilities x abscissa values)

← Discrete Sampling ←

Compressed & Uncorrelated (UC)
L x N
(states$_{UC}$ x particles)

← Inverse FastICA +..NFA ←

Reconstructed State
L x N
(states$_{ECrec}$ x particles)

Information Measures ↕ ***As Lossless as possible***

OUTLINE

1. Introduction and Motivation
2. Literature Review
3. Approach
4. Research Contributions
5. Statistical Orbit Determination
6. PDF Compression
7. Simulations and Results
8. Conclusions and Future Work
9. Acknowledgements

RESEARCH CONTRIBUTIONS

Demonstrated the use of the **PF** as a **statistical OD method** for **non-linear state estimation** and **uncertainty predictions** as a **full state PDF**

Used **nonlinear multivariate decorrelation** methods to decorrelate the multivariate state PDF

Performed **univariate uncorrelated state PDF compressions** for data allocation and transmission **cost reductions**

Demonstrated the **potential** for **cost effective ephemeris storage** for the **cataloging of space objects and debris**

OUTLINE

1 Introduction and Motivation
2 Literature Review
3 Approach
4 Research Contributions
5 Statistical Orbit Determination
6 PDF Compression
7 Simulations and Results
8 Conclusions and Future Work
9 Acknowledgements

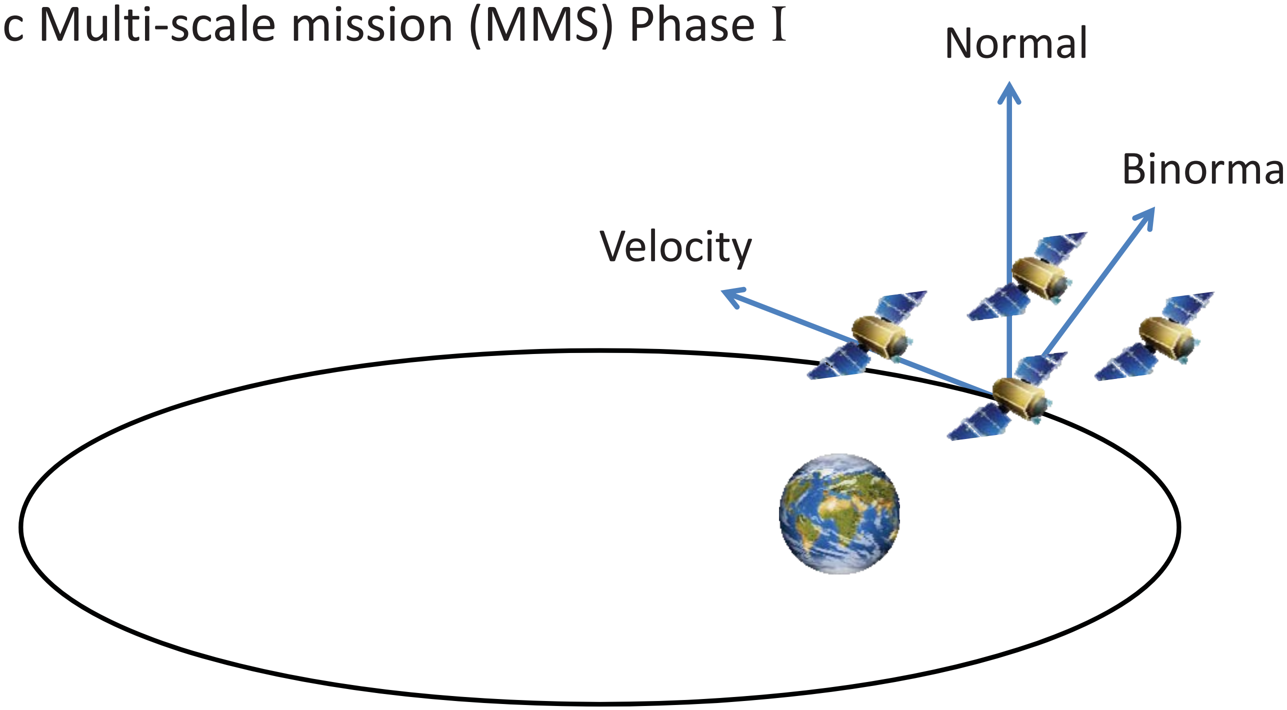
(5–6) **A** and **B**

OUTLINE

MMS ORBIT AND THE VBN FRAME

The Magnetospheric Multi-scale mission (MMS) Phase I

Orbital elements	
a	42095
e	0.8182
i	28.5°
Ω	357.857°
ω	298.225°
M_0	180°



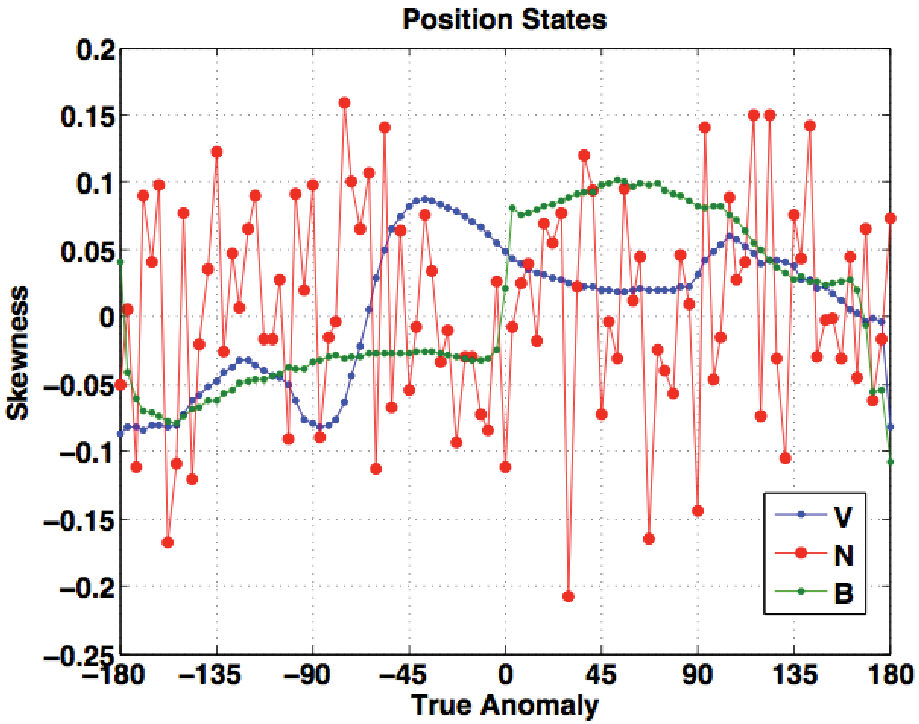
The VBN frame is the reference path orbital velocity, binormal and normal directions.

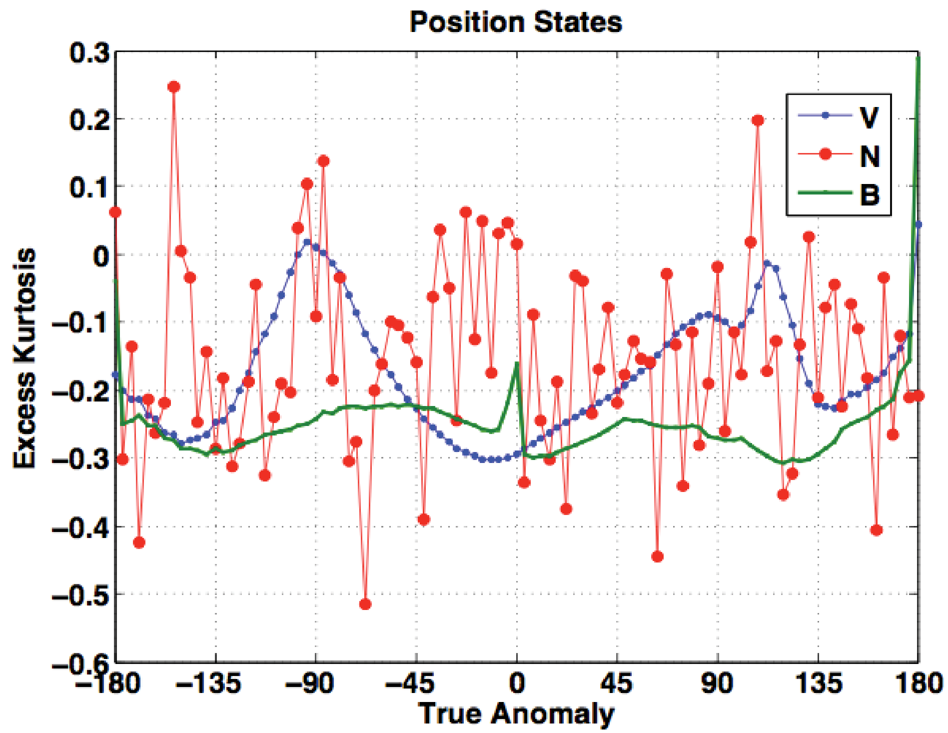
- Rotating frame
- Easy to visualize propagations

SKEWNESS AND EXCESS-KURTOSIS

Illustrating the non-Gaussian content of the PDF propagations of the MMS orbit for 1 period

- Apogee as the initial condition
- Perigee at true anomaly = 0°





Skewness: a measure of asymmetry

(3rd order moment)

$$Skewness = \frac{E\{(X-\mu)^3\}}{[E\{(X-\mu)^2\}]^{3/2}}$$

Excess-Kurtosis: a measure of peak intensity

(4th order moment)

$$Excess-Kurtosis = \frac{E\{(X-\mu)^4\}}{[E\{(X-\mu)^2\}]^2} - 3$$

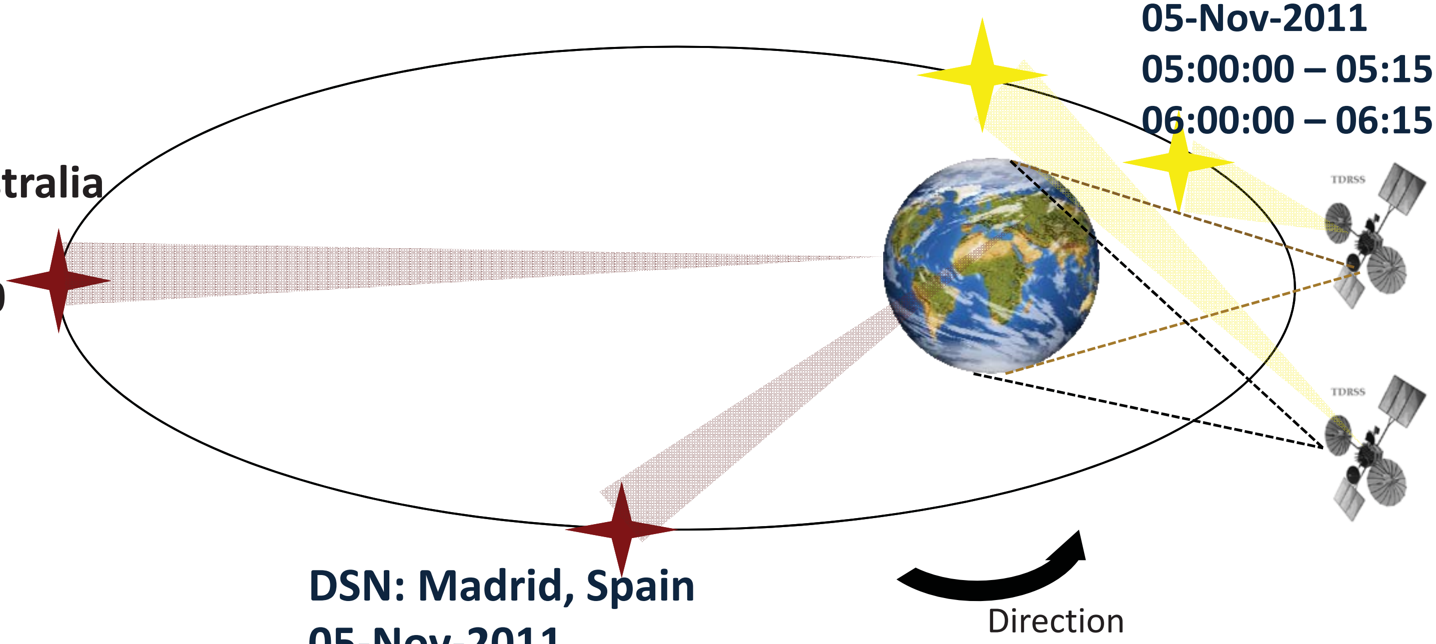
5 STATISTICAL ORBIT DETERMINATION:

SCENARIO A

Epoch: 04-Nov-2011 16:00:00
Period: 24 hrs

TDRS-1 and TDRS-3
05-Nov-2011
05:00:00 – 05:15:00
06:00:00 – 06:15:00

DSN: Canberra, Australia
04-Nov-2011
16:00:00 – 17:15:00



DSN: Madrid, Spain
05-Nov-2011
00:00:00 – 01:15:00

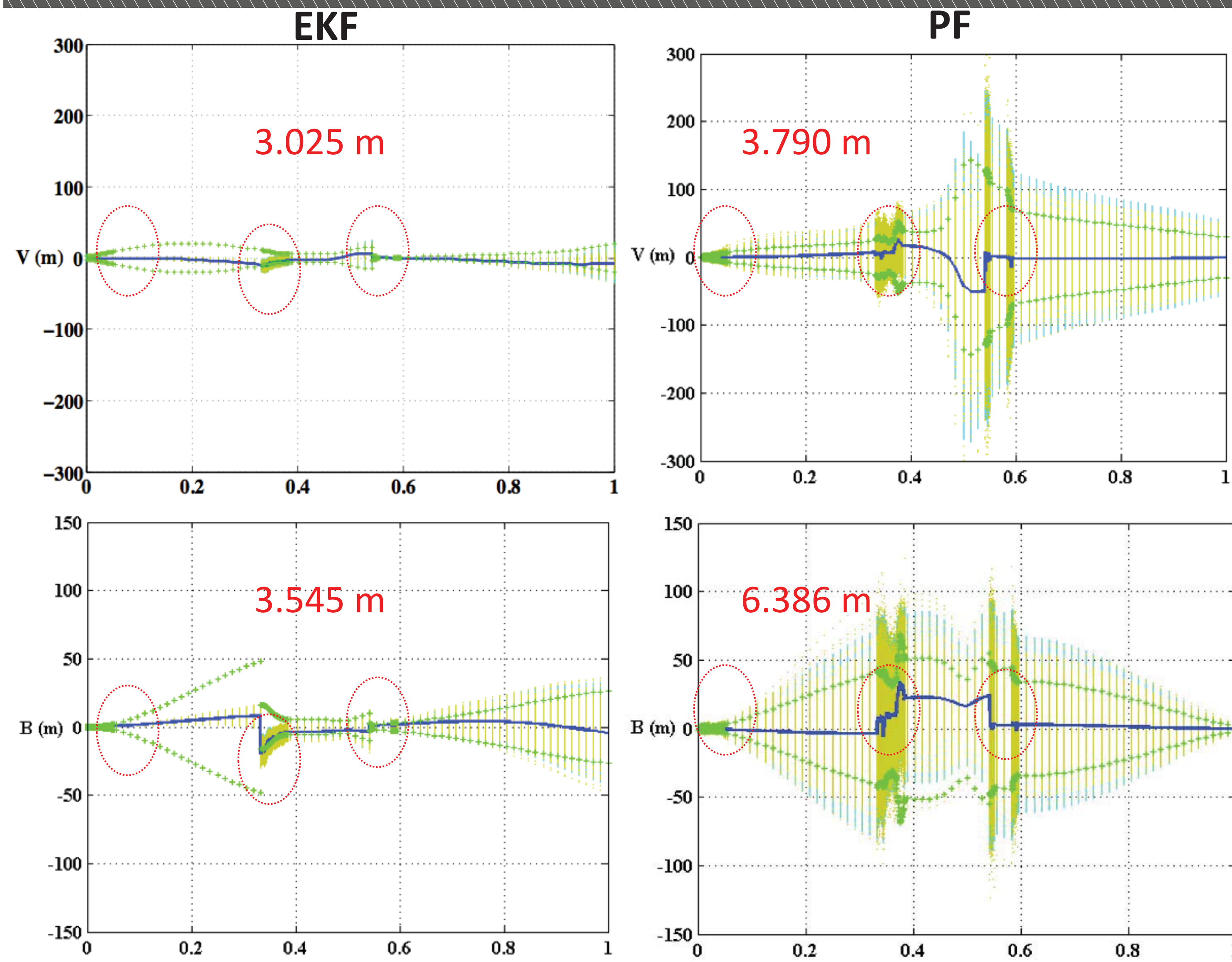
Estimation of state vector **X**: using the Extended Kalman Filter(EKF).
using the Particle Filter for **1000** particles .

CASE 1: Over 1 orbital period with 4 measurements
CASE 2: CASE 1 **AND** an additional 4 orbital periods without measurements

PF code was developed for the Orbit Determination Toolbox (ODTBX) that uses some of its plotting capabilities.

5 STATISTICAL ORBIT DETERMINATION:

CASE 1 : MEASUREMENT UPDATES



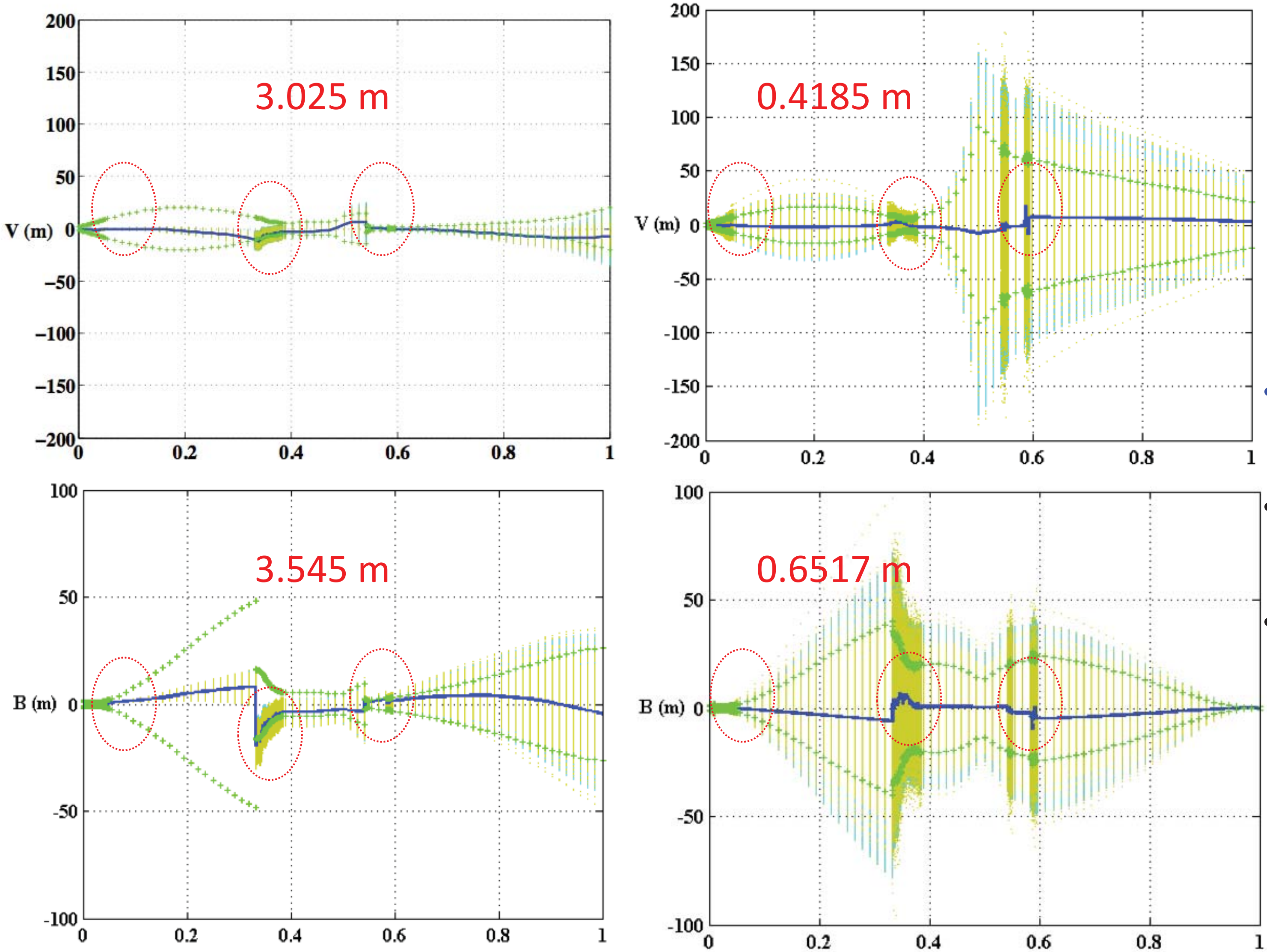
- Meas. update errors
- Ensemble mean error
- 2σ from P formal (EKF)
- 2σ from P empirical (PF)
- Ensemble errors
- 1-3 σ of ensbl. errors

- **NOTE: Resampling threshold $N_{thr} = 0.25$ N**
- **EKF outperforms PF based on mean errors**
- **PF has larger uncertainty bounds vs EKF**

Transformed Position States from XYZ_{ECI} to VNB

CASE 1 : MEASUREMENT UPDATES

EKF **PF**

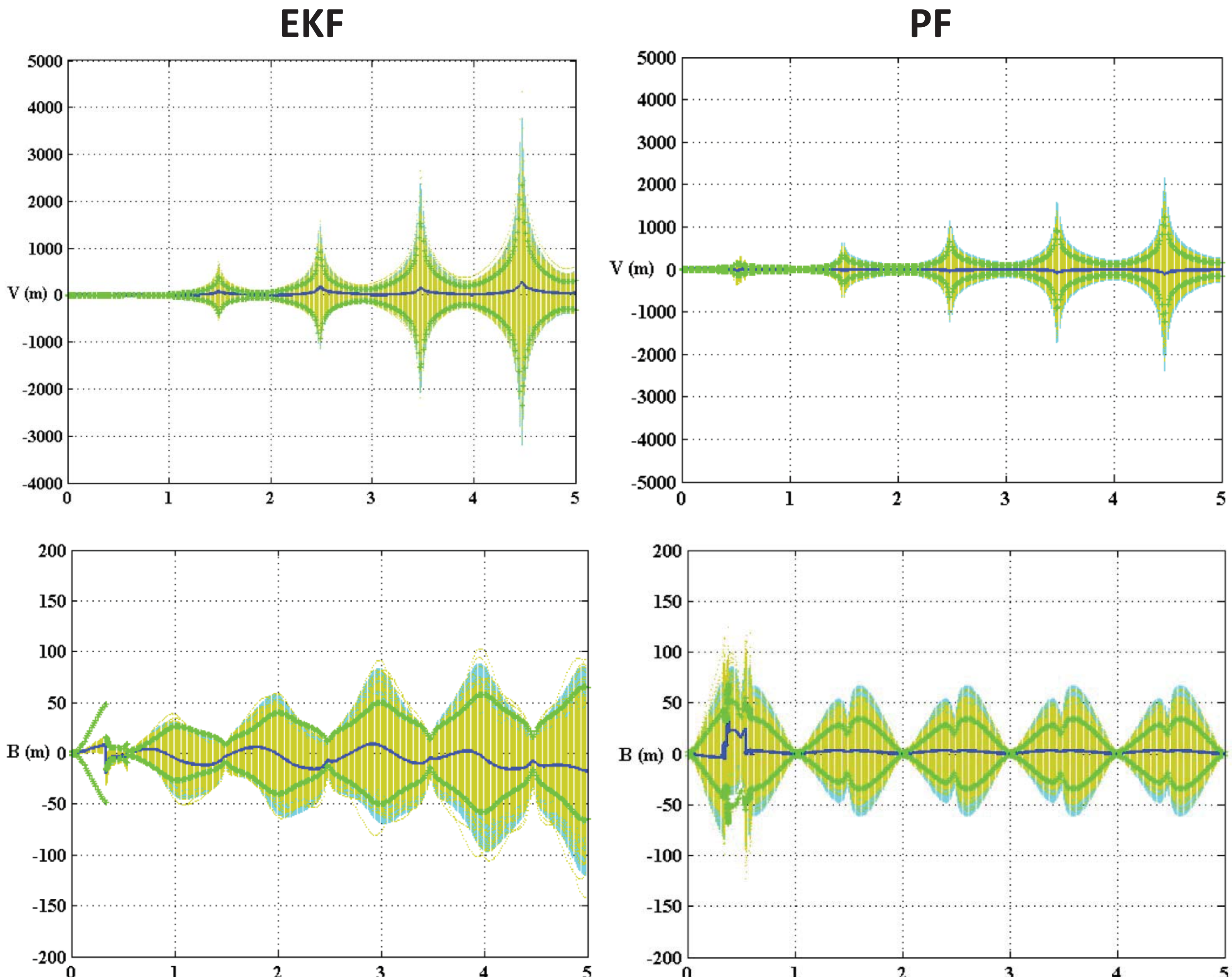


- Meas. update errors
- Ensemble mean error
- 2σ from P formal (EKF)
- 2σ from P empirical (PF)
- Ensemble errors
- 1-3 σ of ensbl. errors

- NOTE: Resampling threshold $N_{thr} = 0.75$ N
- PF outperforms EKF based on mean errors
- PF has larger uncertainty bounds vs EKF

Transformed Position States from XYZ_{ECI} to VNB

CASE 2 : PREDICTIONS



- Ensemble mean error
+ 2σ from P formal (EKF)
+ 2σ from P empirical (PF)
| Ensemble errors
| 1-3 σ of ensbl. errors

Max. Prediction errors

State	EKF	PF
V(m)	**283**	**-109**
B(m)	**-18.1**	**2.99**

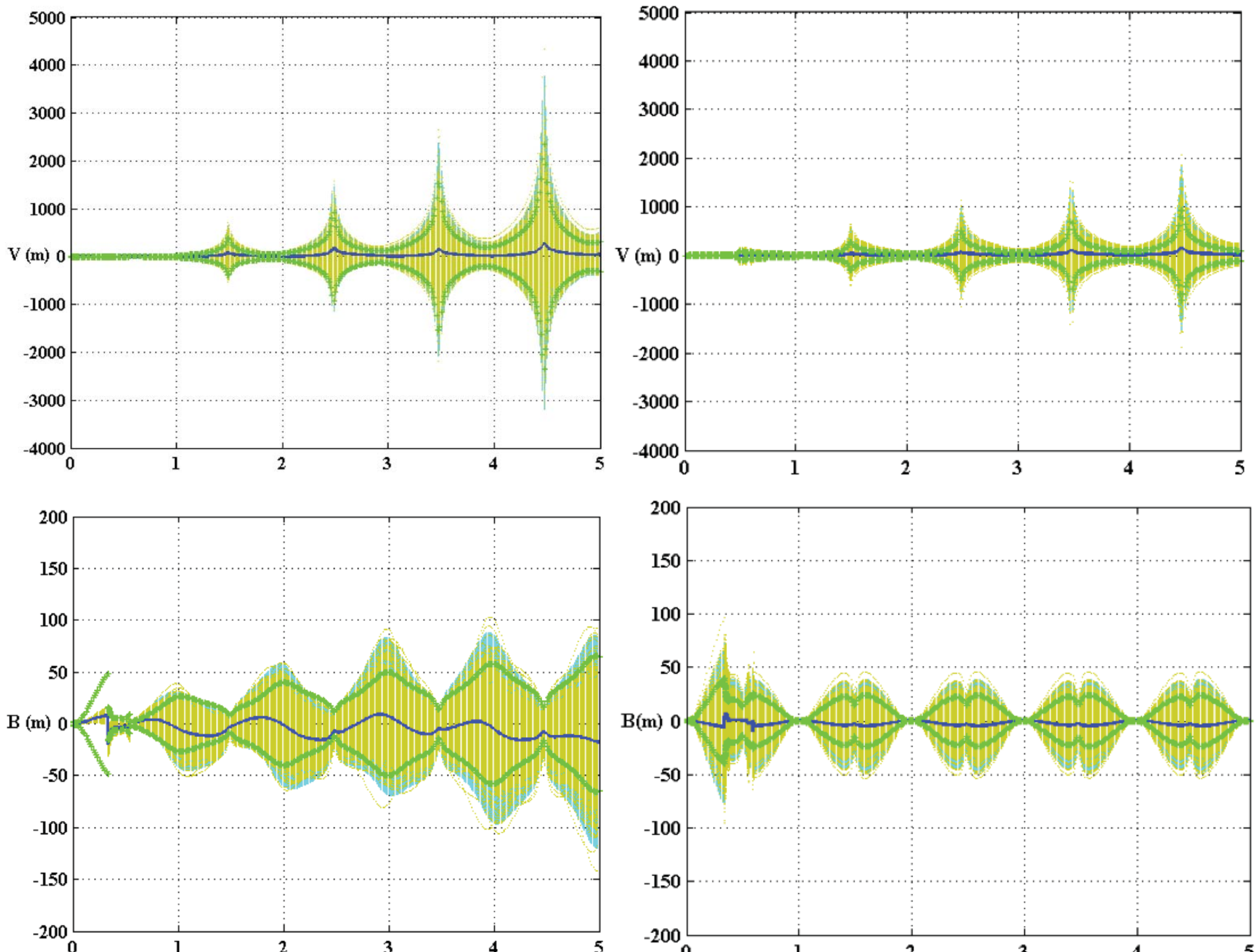
- **Post-updates $0.25N_{thr}$**
- **PF uncertainty bounds are tighter than EKF**
- **Bi-normal component has good accuracies at both apogee and perigee for the PF vs EKF.**

Transformed Position States from XYZ_{ECI} to VNB

5 STATISTICAL ORBIT DETERMINATION:

CASE 2 : PREDICTIONS

EKF **PF**



- Ensemble mean error
+ 2σ from P formal (EKF)
+ 2σ from P empirical (PF)
| Ensemble errors
| 1-3 σ of ensbl. errors

Max. Prediction errors

State	EKF	PF
V(m)	**283**	**148.5**
B(m)	**-18.1**	**-4.65**

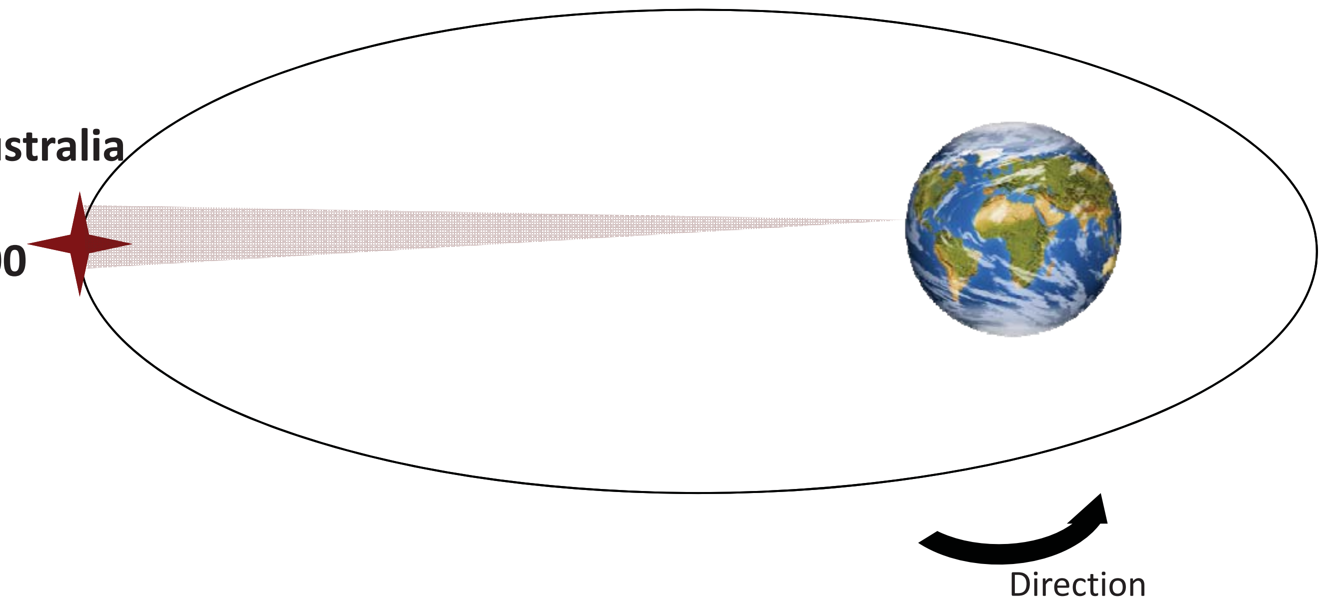
- **Post-updates $0.75N_{thr}$**
- **PF uncertainty bounds are tighter than EKF**
- **Bi-normal component has good accuracies at both apogee and perigee for the PF vs EKF.**

Transformed Position States from XYZ_{ECI} to VNB

SCENARIO B

Epoch: 04-Nov-2011 16:00:00
Period: 24 hrs

DSN: Canberra, Australia
05-Nov-2011
16:45:00 – 17:00:00



Particle Filter for **1000 particles** vs. **GMM** using **UKF**

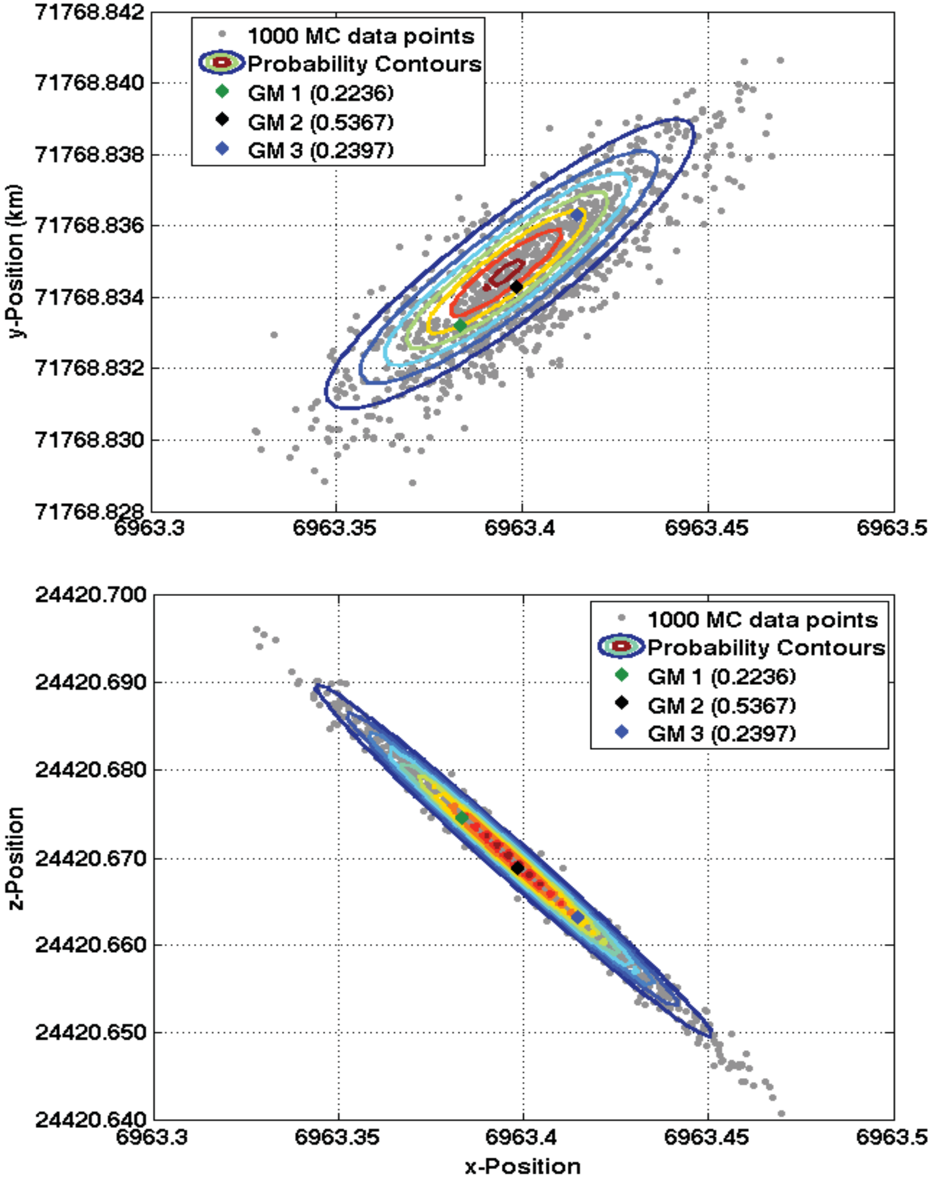
Scenario: Over 1 orbital period with 15 min measurement update at end of 1 period

View the State PDF at the Measurement Update instance: A) As planar contour plots
B) As 3-D PDF plots

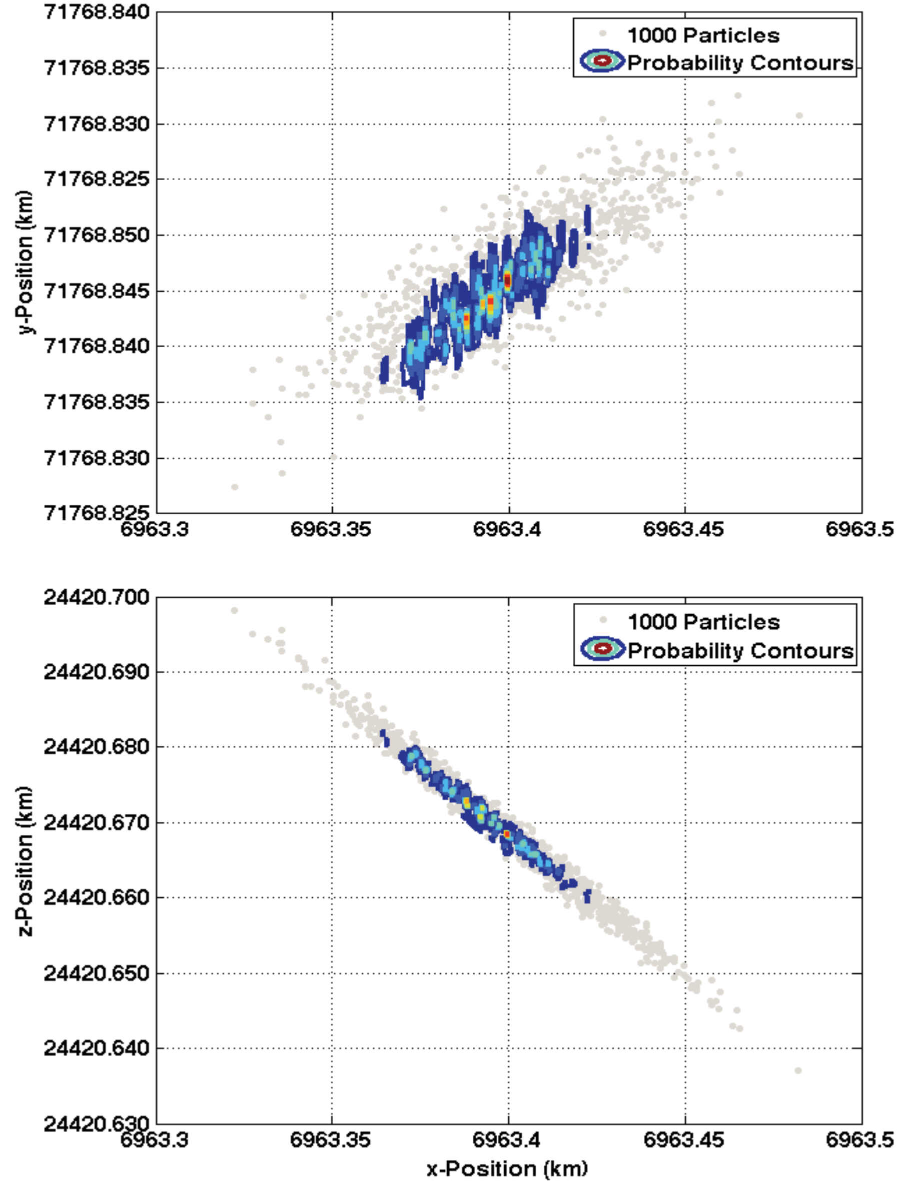
5 STATISTICAL ORBIT DETERMINATION:

MEASUREMENT UPDATED STATE PDF

GMM + UKF

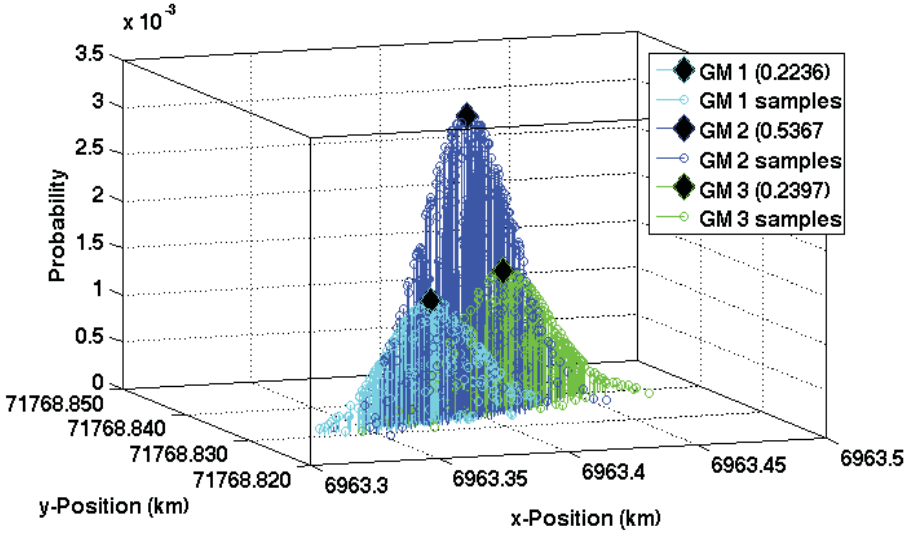


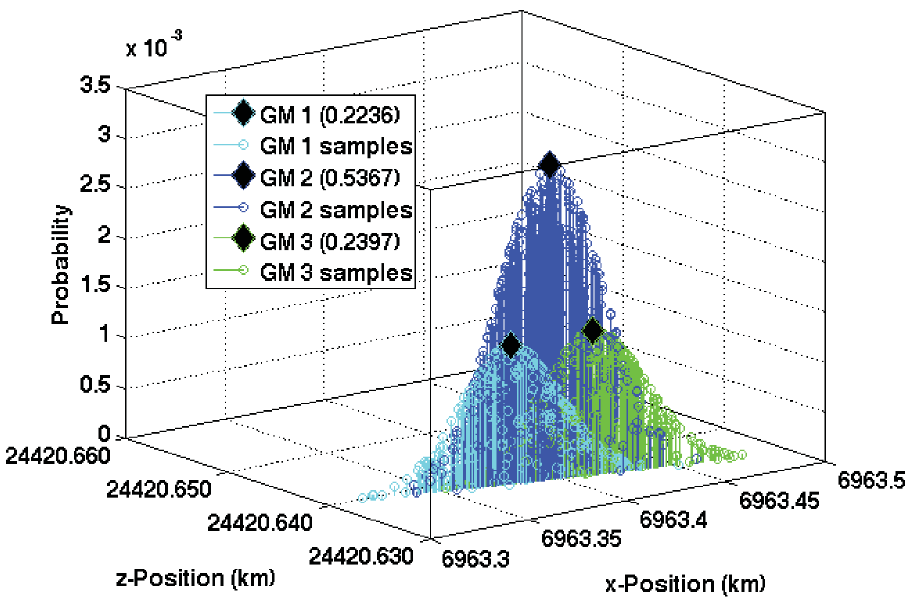
PF



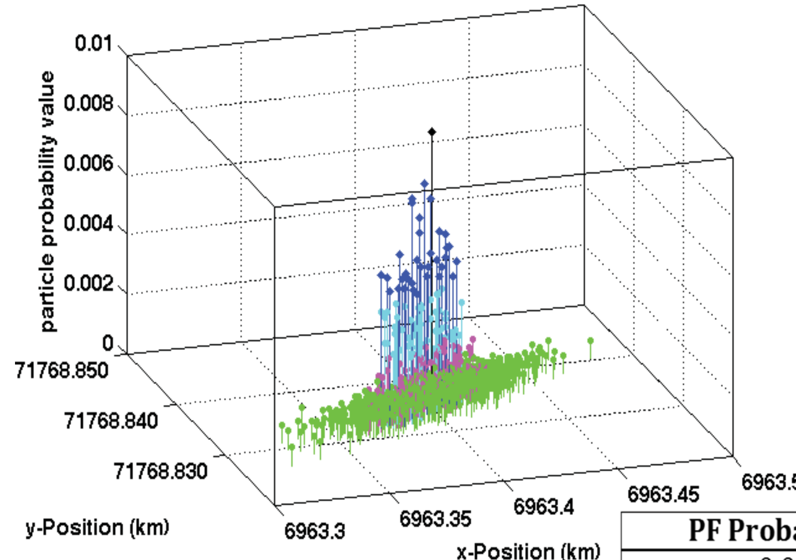
MEASUREMENT UPDATED STATE PDF

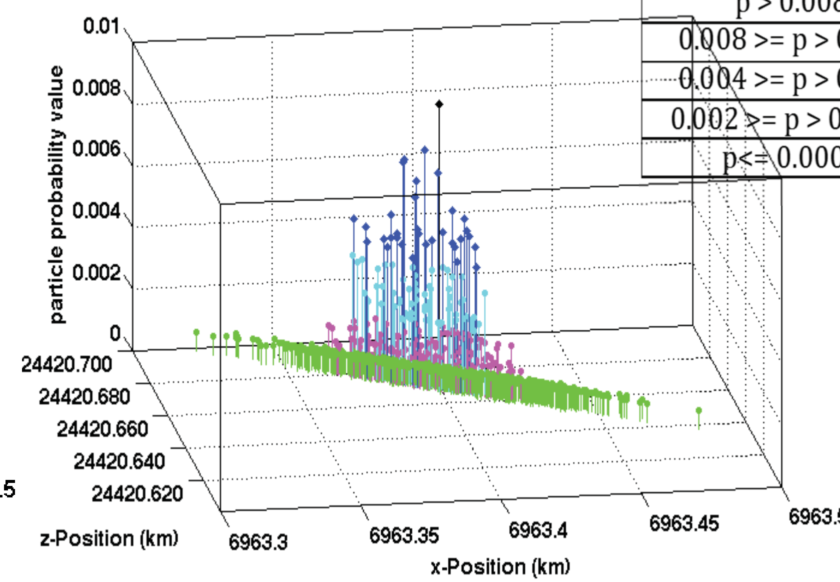
GMM + UKF





PF





PF Probability Scale	
p > 0.008	black
0.008 >= p > 0.004	dark blue
0.004 >= p > 0.002	light blue
0.002 >= p > 0.0006	magenta
p<= 0.0006	green

OUTLINE

1 Introduction and Motivation
2 Literature Review
3 Approach
4 Research Contributions
5 Statistical Orbit Determination
6 PDF Compression
  - Decorrelation Background (PCA and ICA)
  - Dimensional Reduction Example and Simulations
  - Nonlinear Factor Analysis
  - Compression Methods (FFT and WT)
  - Reconstruction Approach
  - Compression Rates
7 Simulations and Results
8 Conclusions and Future Work
9 Acknowledgements

DECORRELATION BACKGROUND

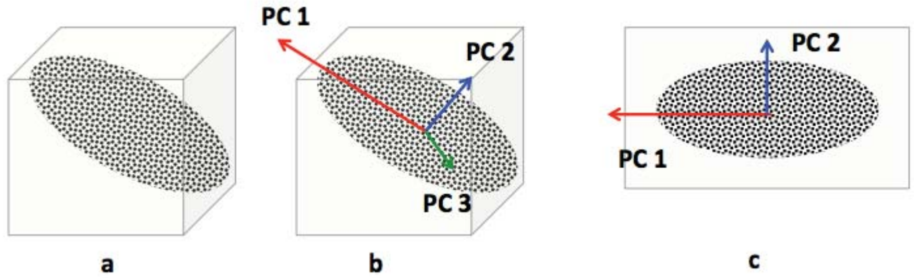
PRINCIPAL COMPONENT ANALYSIS (PCA)

$$\Sigma_x = W \Lambda W^T$$

$$s = W^T x$$

x : zero mean data

s : principal components



PCA finds the **principal components** in a new transformed coordinate system with the **greatest variance** as the **first component** and the subsequent ones follow.

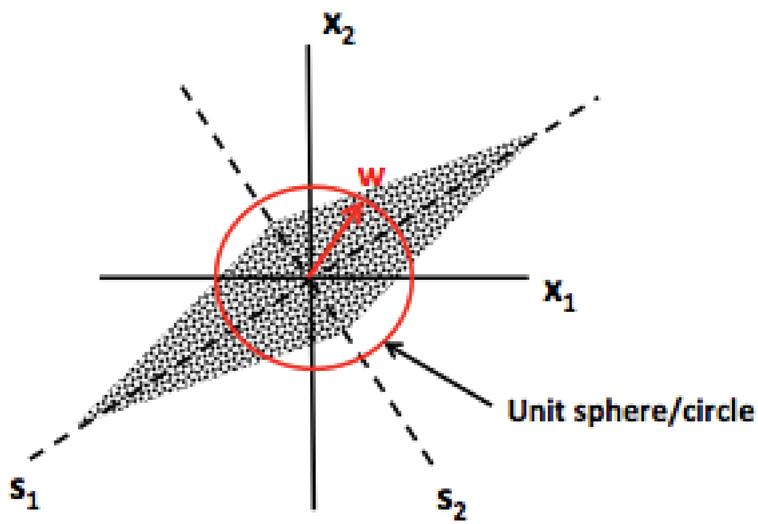
INDEPENDENT COMPONENT ANALYSIS (ICA)

$$x = Bs$$

$$s = B^T x$$

x : whitened data

s : independent states



ICA finds the **linear combinations** of the **unit sphered observations x** ($w^T x$) that are in the **direction w of the max or min Kurtosis** ***(4th order cumulant)*** B = {w1 w2}.

FastICA: uses a **fixed point iteration scheme** *(MUCH FASTER).*

$$\text{Kurtosis}(s_i) = E\{s_i^4\} - 3(E\{s_i^2\})^2 \text{ where } s_i = w^T x$$

$$w(i+1) = E\{x(w(i)^T x)^3\} - 3w(i) \text{ iterated } |w(i+1)^T w(i)| \to 1$$

6 PDF COMPRESSION:

DIMENSIONAL REDUCTION EXAMPLE

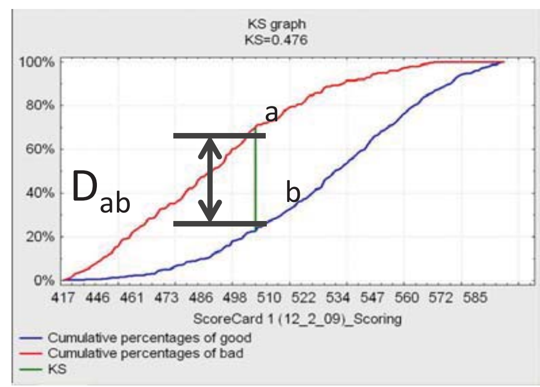
Orbit parameters: eccentricity = 0.2, P_0 ➔ $\sigma^2_{x,y,z}$= $(1000m)^2$ and $\sigma^2_{vx,vy,vz}$= $(1m/s)^2$

1 rev of propagation, Epoch of interest at perigee**,** Period = 18hrs

- Initial state dimension L = 6 ***ICA or PCA*** ⟶ 4 components
  - *Choice of 4 : Out-of-plane motion is decoupled from in-plane motion*
    - – Fundamental nonlinearity comes from Kepler's equation
  - *Demonstrate potential for augmented states not necessarily required for accurate state prediction* (Station location errors, Range biases etc)

- The reconstructed distributions are measured using:
  - The Kolmogorov-Smirnov (K-S) test
    - – Quantifies the max. distance D between the two Cumulative Distribution Functions (CDF)

$$D_{ab} = \sup_X \left| F_a(X) - G_b(X) \right|$$

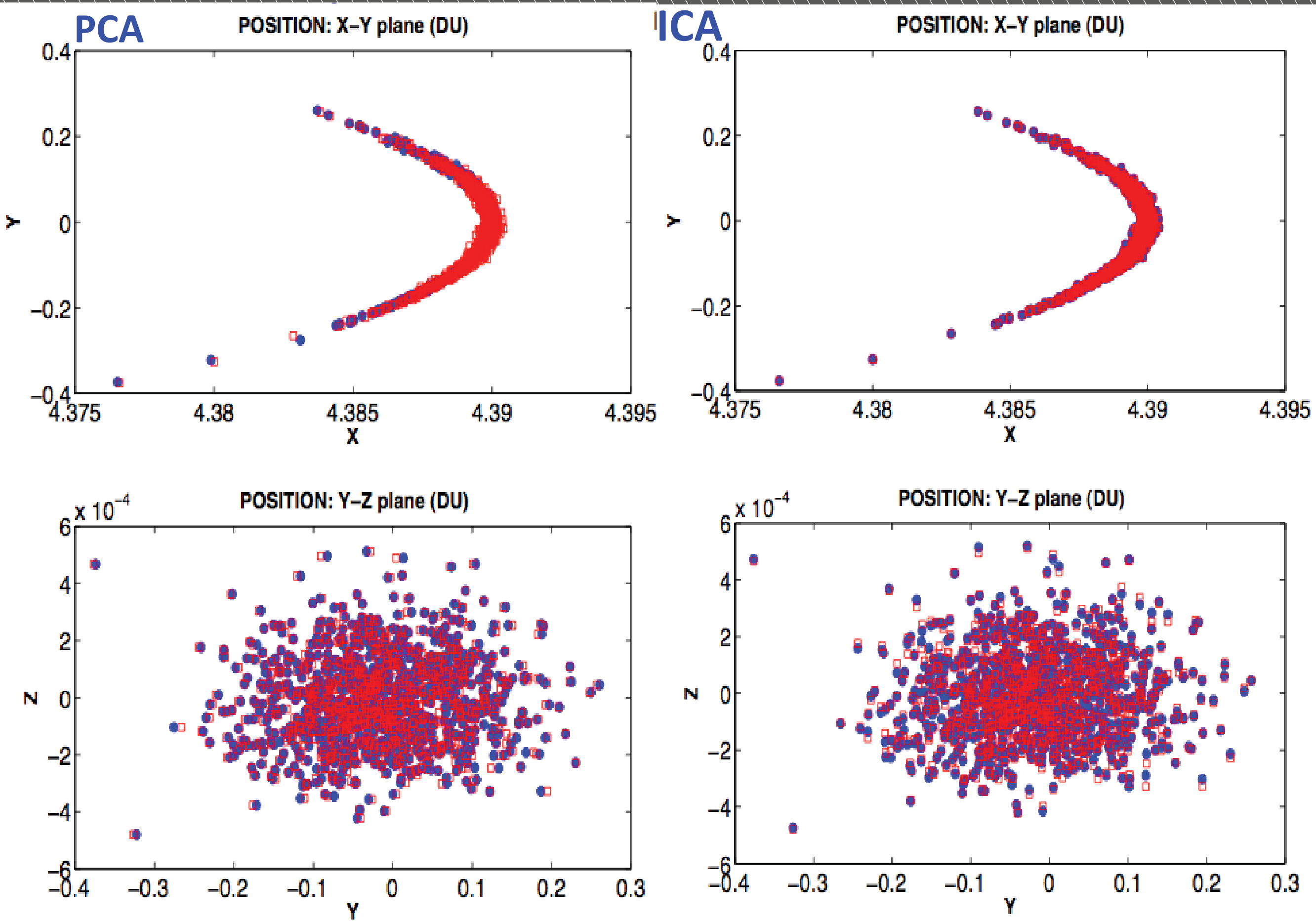


http://www.statsoft.com/Portals/0/Support/ks%20graph.JPG

- The state particles are scaled to **canonical units**
  - Ensures equal weightings for velocity components (affects eigenvalues)
  - Distance : (1DU) = 6378.145km, Velocity : (1DU/ TU) = 7.90536828km/s

6 PDF COMPRESSION:

DIMENSIONAL REDUCTION SIMULATION



Original Data

Reconstructed Data

Excess Kurtosis	PCA	ICA
1st	1.30	24.39
2nd	22.66	2.55
3rd	0.46	0.02
4th	-0.08	-0.4

K-S Statistic	PCA	ICA
X	0.0740	0.0080
Y	0.0100	0.0020
Z	0.0030	0.0150

Shortcomings of PCA/ICA:

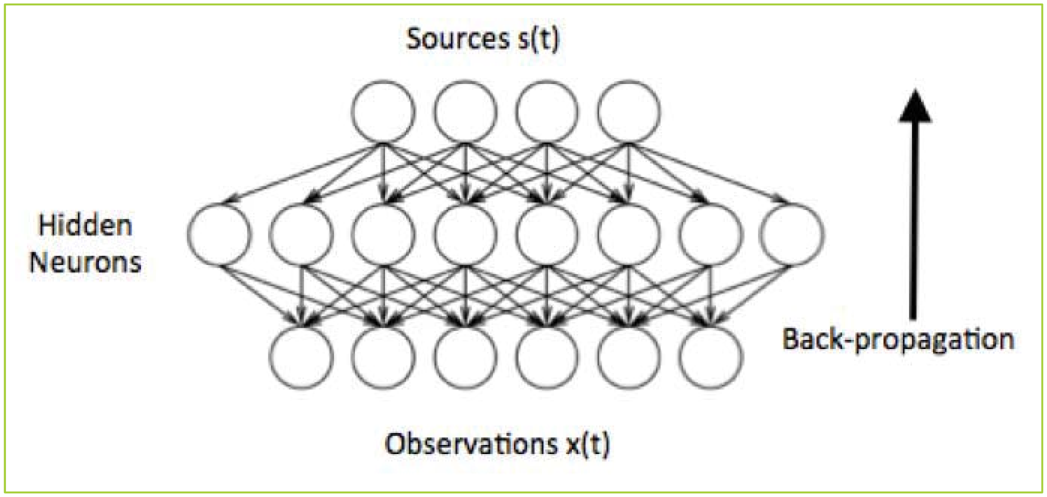
- **For high non-Gaussian behavior the components were not fully decorrelated**

Solution

- **Use Nonlinear Factor Analysis**

NONLINEAR FACTOR ANALYSIS

The Nonlinear Factor Analysis (NFA) is a nonlinear mapping of the sources **s** to the observations **x** modeled by a multilayer perceptron (MLP) network model.



$$\begin{aligned} x &= f(s,\theta)+n \\ &= \mathcal{B}\phi(\mathcal{A}s+a)+b+n \\ &= \mathcal{B}\tanh(\mathcal{A}s+a)+b+n \end{aligned}$$

Our goal is to estimate the sources **s** and the unknown variables θ =($\mathcal{A}$, $\mathcal{B}$,**a**,**b**)

$\mathcal{A} \in \Re^{H_n \times L}$, $\mathcal{B} \in \Re^{L \times H_n}$, **a** $\in \Re^{H_n}$ and **b** $\in \Re^{L}$ based on minimizing the cost function between the posterior p and the approximate q given by the Kullback-Leibler Divergence metric *D*

$$D(q(\Theta\,|\,\xi)\,\|\,p(\Theta\,|\,x,\mathcal{H})) = \int q(\Theta\,|\,\xi)\ln\frac{q(\Theta\,|\,\xi)}{p(\Theta\,|\,x,\mathcal{H})}d\Theta$$

where $\Theta = \{\bar{\theta}, \tilde{\theta}, \bar{s}, \tilde{s}\}$ *are random variables*

ξ and $\mathcal{H}$ are some prior assumptions.

Ref: Theory and MATLAB code developed by Dr. Antti Honkela at Helsinki University of Technology, Finland

NONLINEAR FACTOR ANALYSIS

The prior assumptions ξ and $\mathcal{H}$ are given as follows:

$$x = \mathcal{B}\tanh(\mathcal{A}s + a) + b + n$$

The variance is parameterized to v = log σ

$x \sim N(f(s,\theta), diag(e^{2V_n}))$

$s \sim N(0, diag(e^{2V_s}))$

Parameters

$A, B, a, b \sim N(m_\theta, diag(e^{2V_\theta}))$

(the above together: ***p***)

$$D = \int q(\Theta|\xi)\ln\frac{q(\Theta|\xi)}{p(\Theta|x,\mathcal{H})}d\Theta$$

$$= \int q(\Theta|\xi)\ln\frac{q(\Theta|\xi)}{p(\Theta,x|\mathcal{H})}d\Theta + \ln p(x|\mathcal{H})$$

≠ f(**θ**)

$$C = \int q(\Theta|\xi)\ln\frac{q(\Theta|\xi)}{p(\Theta,x|\mathcal{H})}d\Theta$$

$$C = E[\ln q(\Theta|\xi)] + E[-\ln p(\Theta,x|\mathcal{H})]$$

$$\min C = C_q + C_p = E[\ln q(\Theta|\xi)] + E[-\ln p(\Theta,x|\mathcal{H})]$$

w.r.t. $\bar{\Theta}$ *and* $\tilde{\Theta}$

$$\frac{\partial C}{\partial\bar{\Theta}} = \frac{\partial C_q}{\partial\bar{\Theta}} + \frac{\partial C_p}{\partial\bar{\Theta}} = 0 \quad and \quad \frac{\partial C}{\partial\tilde{\Theta}} = \frac{\partial C_q}{\partial\tilde{\Theta}} + \frac{\partial C_p}{\partial\tilde{\Theta}} = 0$$

where $\bar{\Theta}$ *and* $\tilde{\Theta}$ *are the mean and variance respectively*

$\bar{\Theta} = \{\bar{A}, \bar{B}, \bar{a}, \bar{b}, \bar{s}\}$ and

$\tilde{\Theta} = \{\tilde{A}, \tilde{B}, \tilde{a}, \tilde{b}, \tilde{s}\}$

are the solve-for parameters.

<u>Known</u>: x and priors for $\bar{\Theta}$ and $\tilde{\Theta}$

and

Hyperparameters:

$v_n, v_s, v_{B_j} \sim N(m_v, diag(e^{2V_v}))$

$(m_a, v_a, m_b, v_b, m_{v_n}, v_{v_n}, m_{v_{B_i}}, v_{v_{B_i}}) \sim N(0, 100^2)$ uninformative priors

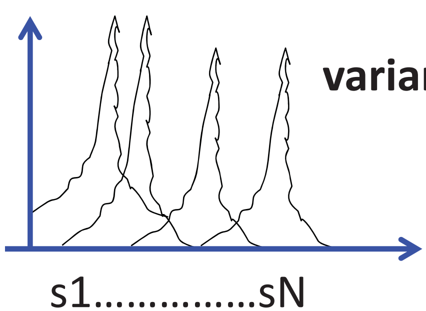
NONLINEAR FACTOR ANALYSIS

The obtained mean and variance terms for **s** and θ:

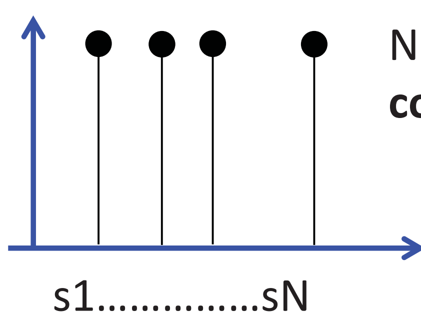
- mean **s:** L x N
- variance **s:** L x N
- mean θ: $\mathcal{A}(H_n \times L)$, $\mathcal{B}(L \times H_n)$, **a**($H_n$ x 1) and **b**(L x 1)
- variance θ: $\mathcal{A}(H_n \times L)$, $\mathcal{B}(L \times H_n)$, **a**($H_n$ x 1) and **b**(L x 1)

L : state dimension
N: No. particles
H_n :No. Neurons

The variance terms for the sources **s** are assumed to be zero (and variances θ=0)



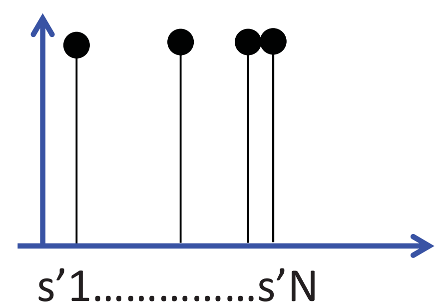
variance s(i) = **0**



Non-Gaussian and **correlated**

FastICA

Non-Gaussian and **uncorrelated**



- FastICA produces uncorrelated sources **s’**

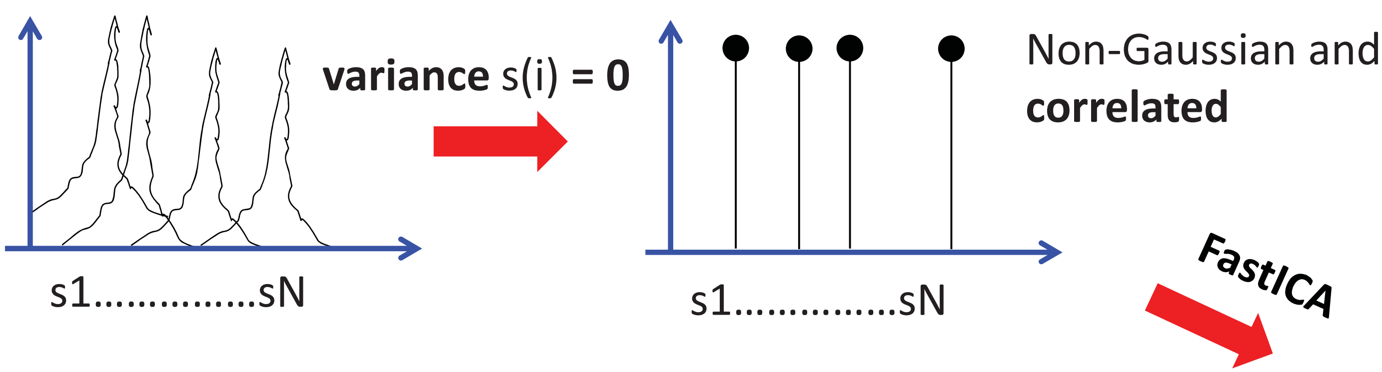
NONLINEAR FACTOR ANALYSIS

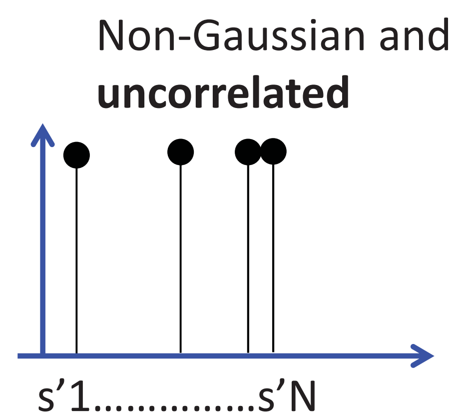
The obtained mean and variance terms for **s** and θ:

- mean **s:** L x N
- variance **s:** L x N
- mean θ: $\mathcal{A}$(H_n x L), $\mathcal{B}$(L x H_n), **a**(H_n x 1) and **b**(L x 1)
- variance θ: $\mathcal{A}$(H_n x L), $\mathcal{B}$(L x H_n), **a**(H_n x 1) and **b**(L x 1)

L : state dimension
N: No. particles
H_n :No. Neurons

The variance terms for the sources **s** are assumed to be zero (and variances θ=0)





- FastICA produces uncorrelated sources **s’**

COMPRESSION METHODS

The uncorrelated sources **s’,** are binned into a 2^7 histogram then transformed into a normalized discrete probability density function (PMF) p for compression.

FAST-FOURIER TRANSFORM

- The FFT calculates the coefficients of the discrete PDF as follows:

$$c(k) = \sum_{j=1}^{n} p(j) e^{\frac{-2\pi i}{n}(j-1)(k-1)} \text{ where } (k = 1, \ldots n)$$

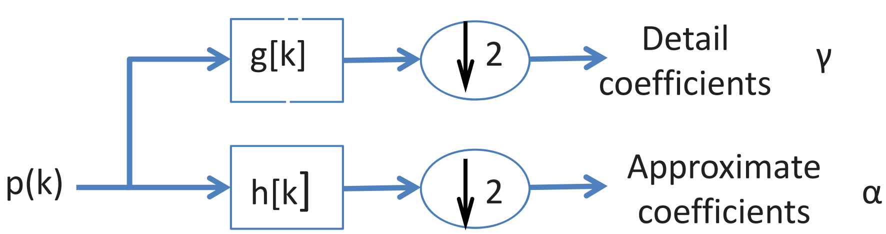
p(j) is the discrete probability value normalized from the histogram

WAVELET TRANSFORM

- Use a wavelet filter (Daubechies 2) to extract terms that represent the PMF p as a function of approximate and detail coefficients

$$\alpha_{j0}[n] = \sum_k p[k] * h[2n-k]$$

$$\gamma_j[n] = \sum_k p[k] * g[2n-k]$$



$$p(k) = \sum_{n=-\infty}^{\infty} ((\alpha_{j0}[n] * h[2n-k]) + (\gamma_j[n] * g[2n-k]))$$

$j0$ is the first level of decomposition

j are the subsequent levels of decomposition

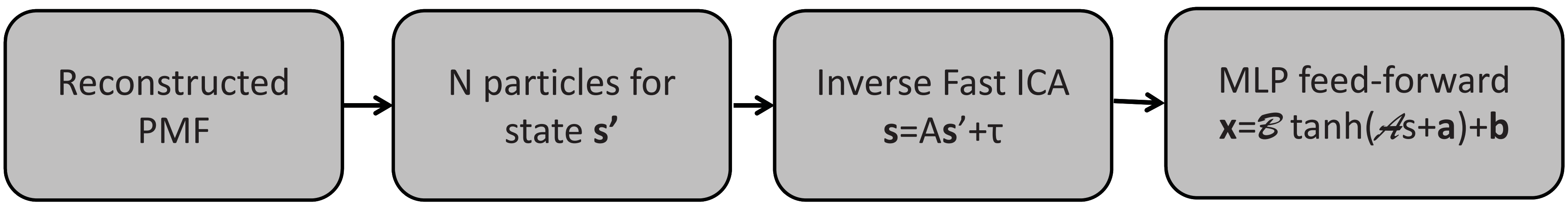
6 PDF COMPRESSION:

RECONSTRUCTION APPROACH

The uncorrelated compressed PMF representations are reconstructed by taking the inverse transforms of the FFT and the WT.

IFFT: $\hat{p}(j)=\sum_{k=1}^{m} c(k)e^{\frac{2\pi i}{n}(j-1)(k-1)}$ where $(j=1,\ldots n)$ and $(m<n)$

IWT: $\hat{p}(k)=\sum_{n=-F}^{F}((\alpha_{j0}[n]*h[-k+2n])+(\gamma_j[n]*g[-k+2n]))$ where $(k=1,\ldots n)$



- Filters (WT)
- Abscissa values
- Coefficients

- Discrete sampling
- Uniform weights

- Mixing matrix A
- Mean vector τ

- The mean variables
  Θ = (𝓐, 𝓑, **a, b**)
- Mean sources **s**

x is the reconstructed state in canonical units

OUTLINE

1 Introduction and Motivation
2 Literature Review
3 Approach
4 Research Contributions
5 Statistical Orbit Determination
6 PDF Compression
7 Simulations and Results
    - Highly non-Gaussian state
    - Decorrelation
    - Compression
    - Information Measures
    - Reconstructions
8 Conclusions and Future Work
9 Acknowledgements

HIGHLY NON-GAUSSIAN STATE

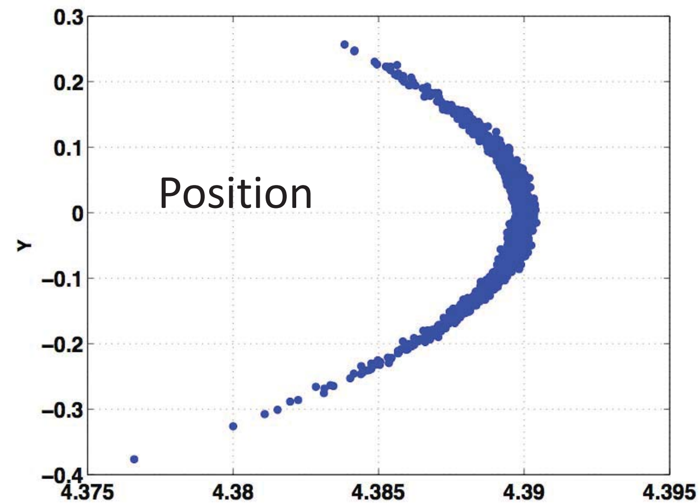
A planar eccentric orbit:

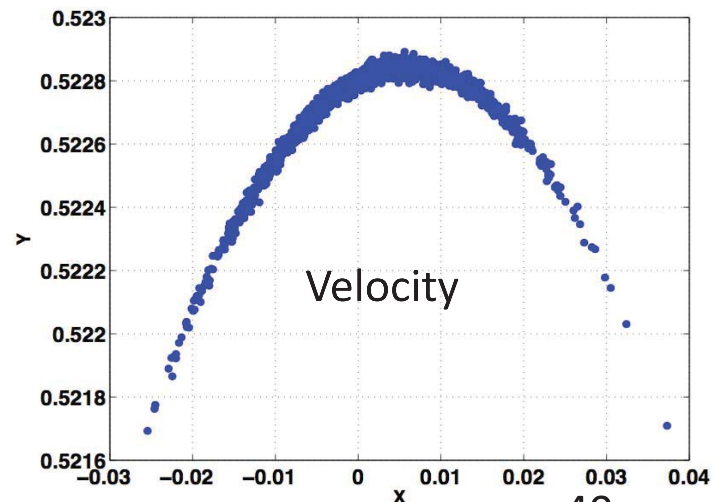
- ***P_0** variances (1 km)^2 for position and (0.001 km/s)^2 for velocity*
- *Ephemeris at true anomaly $\theta = 0^o$ after 1 period*
- *N = 2000 particles*
- *Period 18.1013 hrs.*

Initial Condition $\mathbf{X_0}$:

[28000 km, 0 km, 0 km/s, 4.1331 km/s]

Orbital Elements	Numerical Value
a	42,000 km
e	0.2
i	0 rad
Ω	0 rad
ω	0 rad
M_e	0
Period	18.1013 hrs





DECORRELATION

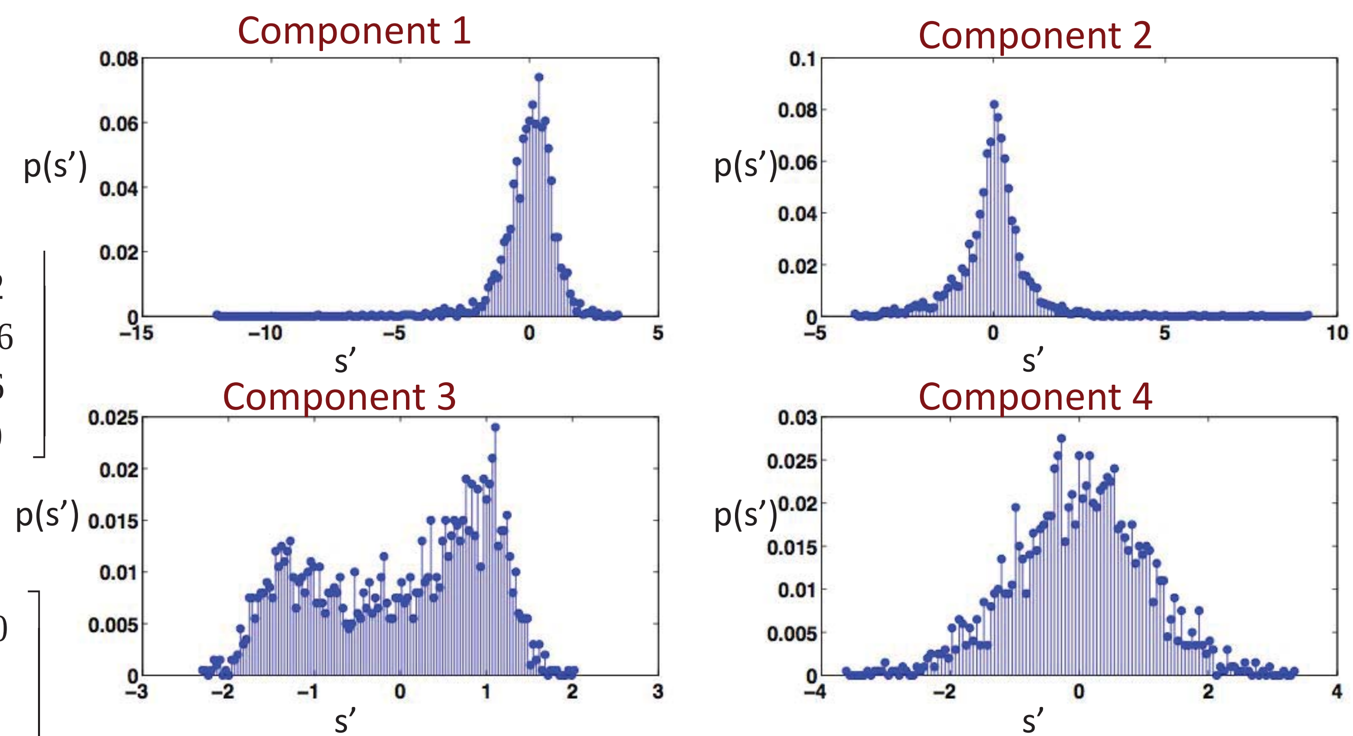
Canonical states (DU and DU/TU) → NFA + FastICA (20 neurons) → Uncorrelated components **s'** → Discrete PMF from a 2^7 binned Histogram

The rank correlation matrix:

$P_\rho = \mathbf{I}$,implies decorrelation

$$P_{\rho,NFA} = \begin{bmatrix} 1.00 & -0.62 & -0.21 & 0.02 \\ -0.62 & 1.00 & 0.21 & -0.06 \\ -0.21 & 0.21 & 1.00 & 0.06 \\ 0.02 & -0.06 & 0.06 & 1.00 \end{bmatrix}$$

$$P_{\rho,NFA+FastICA} = \begin{bmatrix} 1 & 0.0 & -0.0 & -0.0 \\ 0.0 & 1 & 0.0 & 0.0 \\ -0.0 & 0.0 & 1 & 0.0 \\ -0.0 & 0.0 & 0.0 & 1 \end{bmatrix}$$

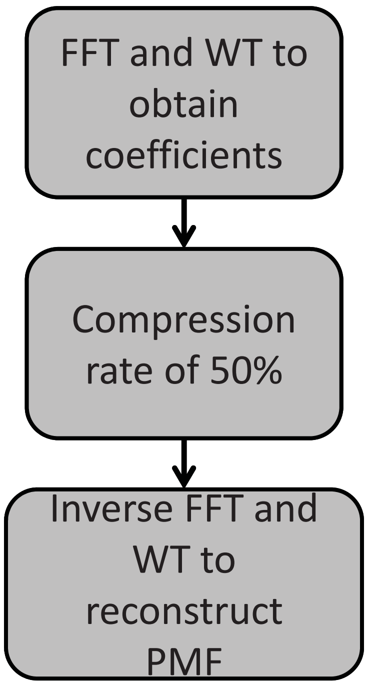


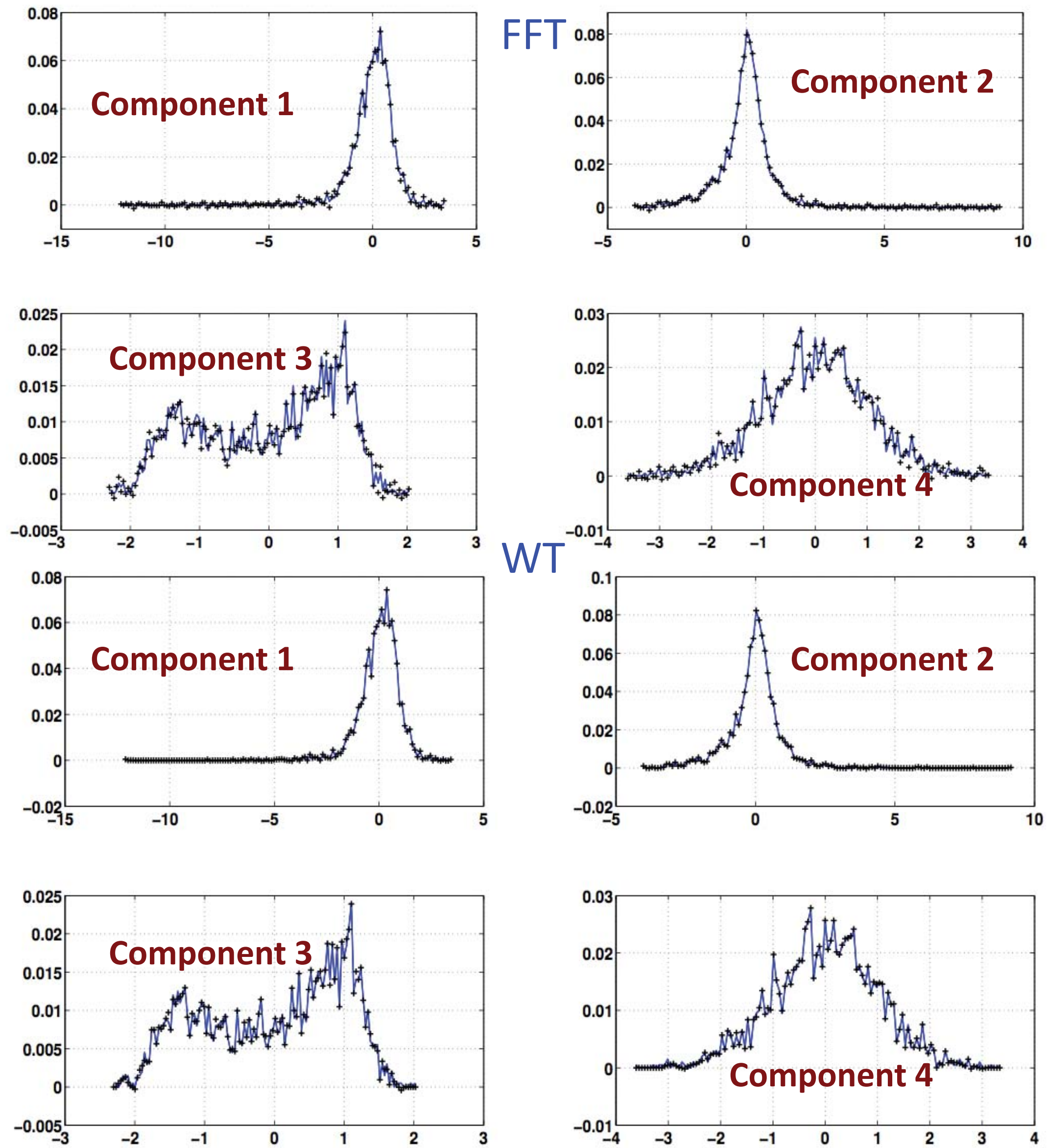
COMPRESSION

Component PDF compression and reconstruction:

Legend

- *Original*
+ *Compressed*





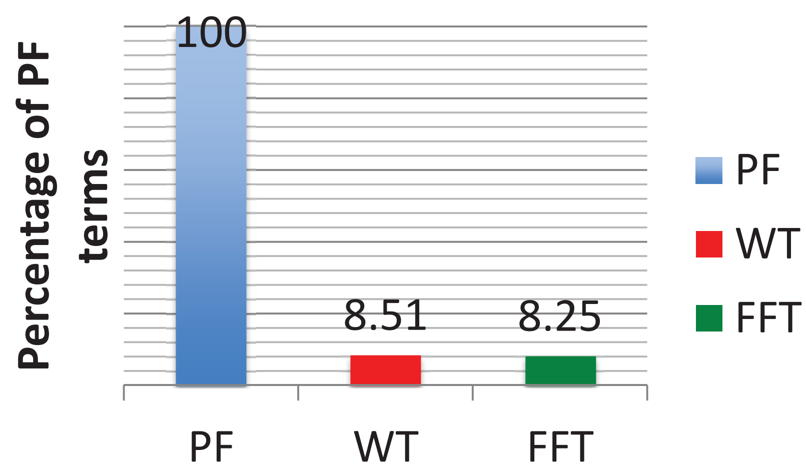
Component	FFT D_{KL} (o||c)	WT D_{KL} (o||c)
1	0.1759	0.0010
2	0.1003	0.0188
3	0.0686	0.1001
4	0.1121	0.1568

INFORMATION MEASURES

TERMS	# TERMS WT	# TERMS FFT
Coefficients **c**	273	260
Wavelet Filters	8	N/A
Abscissa PDF	12	12
Mixing Matrix **A**	16	16
Mean(**s**) = τ	4	4
$\mathcal{A}$	80	80
a	20	20
$\mathcal{B}$	80	80
b	4	4
$v_{\mathcal{A}}$	80	80
v_a	20	20
$v_{\mathcal{B}}$	80	80
v_b	4	4
TOTAL TERMS	681	660

Cost of storing number of terms

Method	# Terms	Cost (Method/PF)
PF Predictions	8000	1
WT Compressions	681	**0.0851**
FFT Compressions	660	**0.0825**



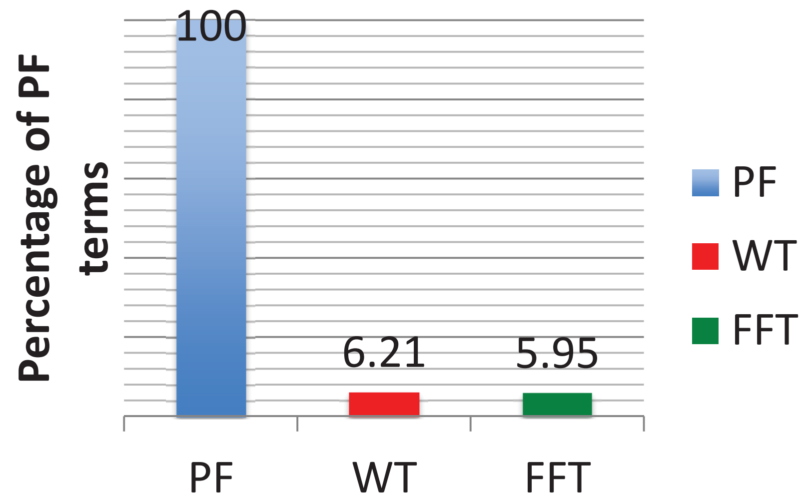
<u>However</u>: need to quantify accuracies based on the state particle reconstructions

INFORMATION MEASURES

TERMS	# TERMS WT	# TERMS FFT
Coefficients **c**	273	260
Wavelet Filters	8	N/A
Abscissa PDF	12	12
Mixing Matrix **A**	16	16
Mean(**s**) = τ	4	4
$\mathcal{A}$	80	80
a	20	20
$\mathcal{B}$	80	80
b	4	4
TOTAL TERMS	497	476

Cost of storing number of terms

Method	# Terms	Cost (Method/PF)
PF Predictions	8000	1
WT Compressions	497	**0.0621**
FFT Compressions	476	**0.0595**



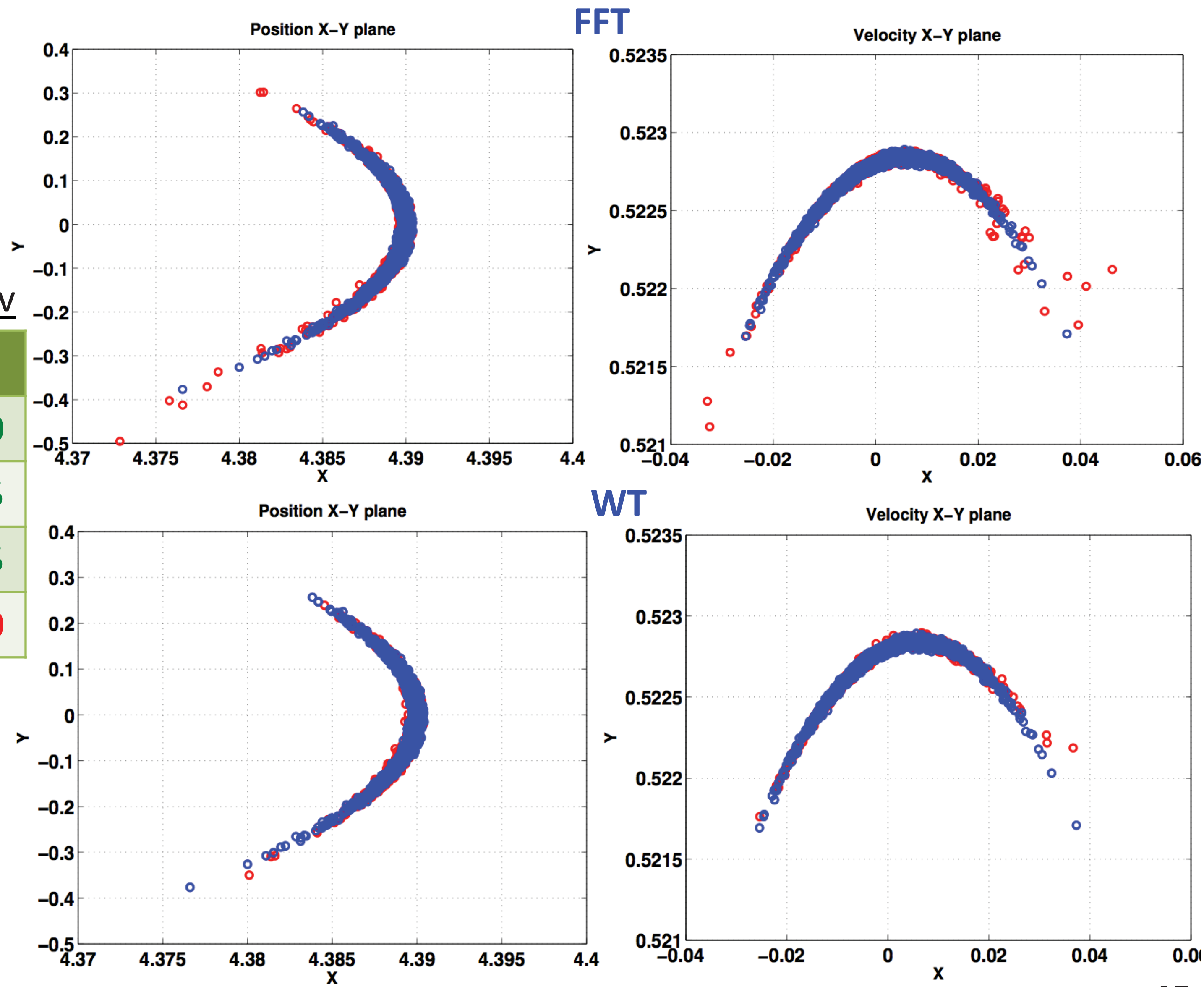
***However**: need to quantify accuracies based on the state particle reconstructions*

RECONSTRUCTIONS

Legend

o *Original*

o Reconstructed



Kolmogorov-Smirnov

K-S	FFT	WT
X	0.0680	0.0510
Y	0.0480	0.0435
V_X	0.0440	0.0425
V_Y	0.0405	0.0570

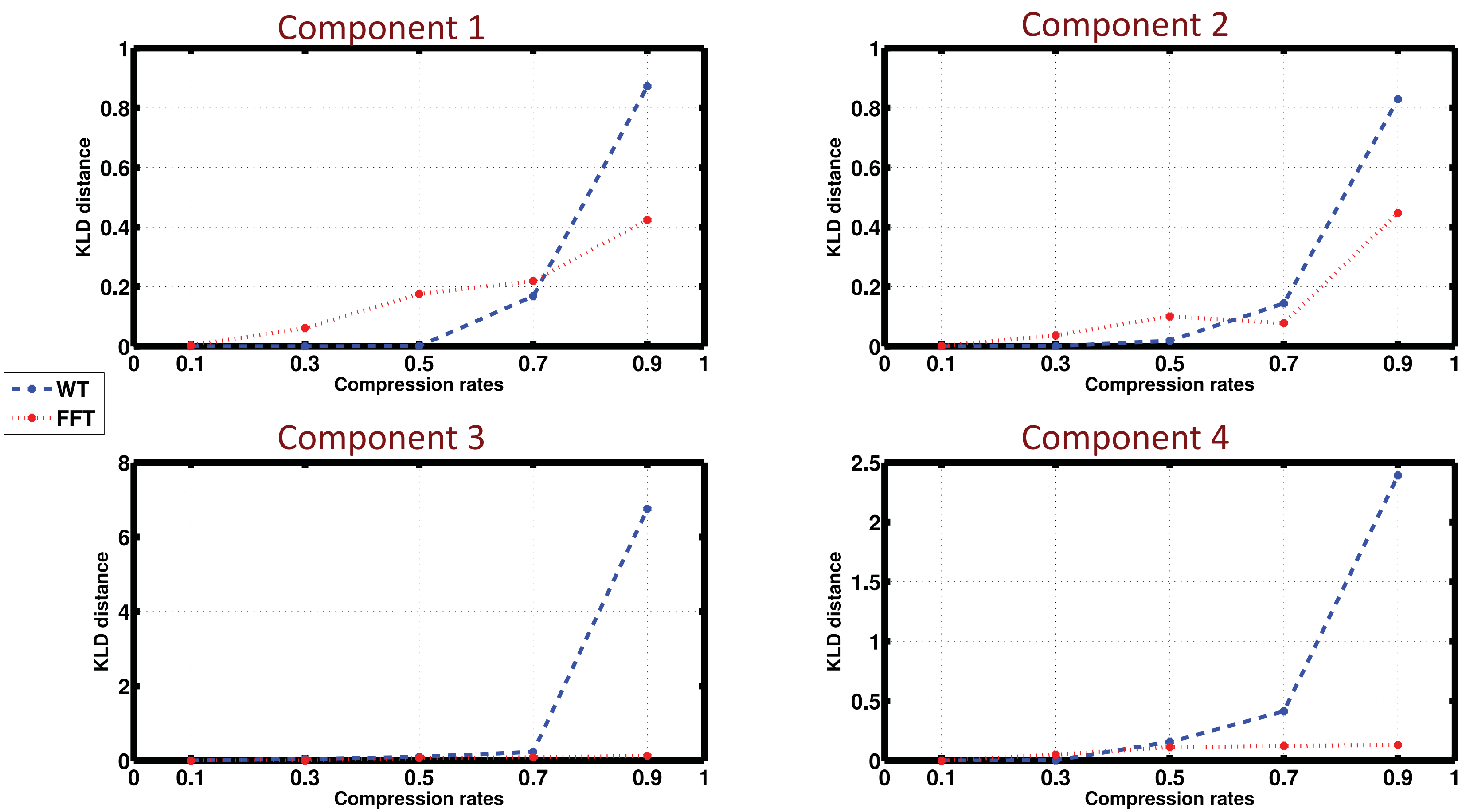
$$D_{OR} = \sup_X \left| F_O(X) - G_R(X) \right|$$

O: Original CDF *F*

R: Reconstructed CDF *G*

COMPRESSION RATES

KLD distance is calculated at different compression rates: [0.1 0.3 0.5 0.7 0.9]



- The WT coefficients has a strong bias between the approximate and detail coefficients
- WT performs better at lower compression rates compared to the FFT

OUTLINE

1 Introduction and Motivation
2 Literature Review
3 Approach
4 Research Contributions
5 Statistical Orbit Determination
6 PDF Compression
7 Simulations and Results
8 Conclusions and Future Work
9 Acknowledgements

CONCLUSIONS AND FUTURE WORK

1. Demonstrated the use of the Particle Filter:
   a) A nonlinear estimator
   b) Capable of incorporating non-Gaussian uncertainties during predictions

2. Demonstrated the use of the Independent Component Analysis (ICA) and Principal component analysis for dimensional reduction

3. Demonstrated the use of Nonlinear Factor Analysis (NFA) followed by FastICA to achieve uncorrelated states for PDF compression

4. The Wavelet Transform and the Fast-Fourier Transform demonstrated as effective methods for the compression and reconstruction of univariate PDFs.

POTENTIAL FUTURE WORK

1. Use other smooth functions to represent the PMF of the components versus WT and FFT for compression
2. Determine the limitations of NFA with respect to the number of particles (apart from computational cost)
3. Other nonlinear decorrelation methods could be implemented ex. The Nonlinear Independent Factor Analysis (NIFA) that uses Gaussian mixtures for sources

OUTLINE

ACKNOWLEDGEMENTS

- Faculty advisor Prof. J.L. Garrison
- NASA GSRP fellowship Technical Advisor Dr. J. Russell Carpenter
- Faculty committee members: Prof. Frazho and Prof. Howell
- NASA Goddard Space Flight Center and the Navigation and the Mission Design Branch (NMDB: Code 595) for the GSRP fellowship (2010-2013) Grant No. NNX10AL88H

QUESTIONS

Thank you.

CLASSICAL ESTIMATION METHODS

Extended Kalman Filter (LaViola 2003 et al)

- Nonlinear models assuming Gaussian uncertainties
- Linearization about a current mean (computes Jacobian)

$$X_k = f_k\left(X_{k-1}, w_{k-1}\right)$$

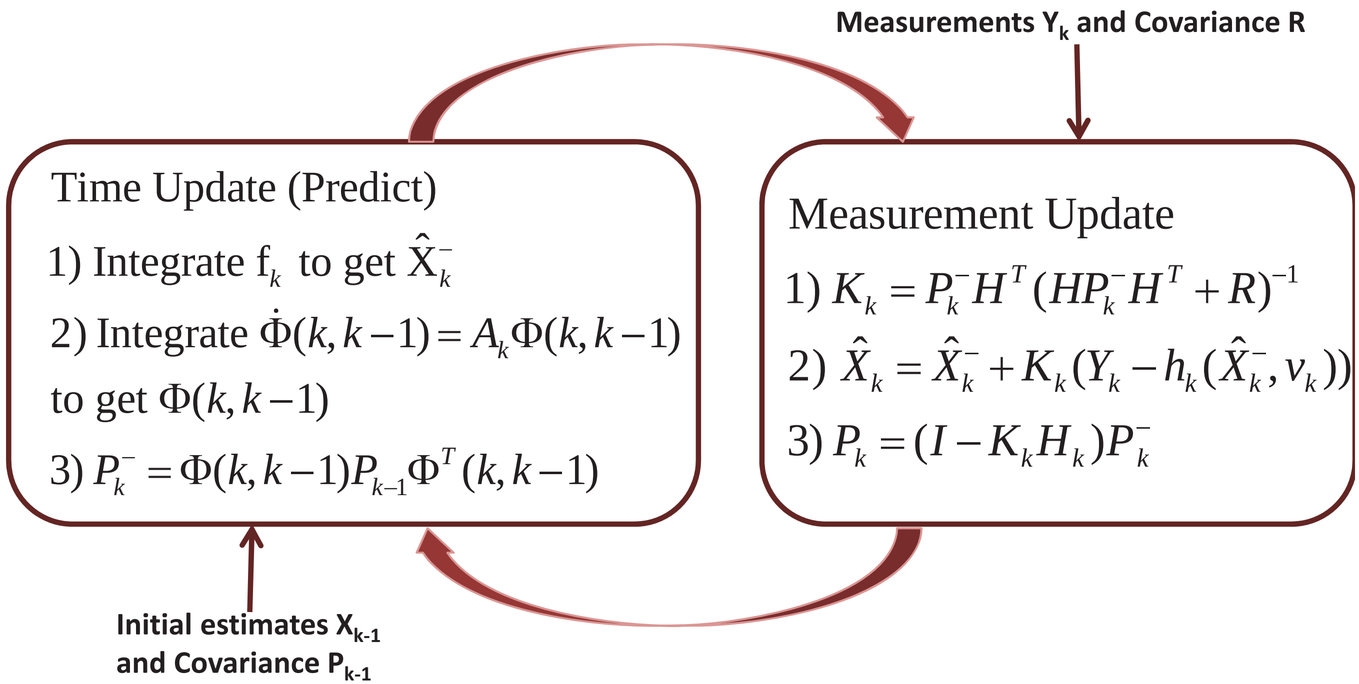
$$Y_k = h_k\left(X_k, v_k\right)$$

$$A_k = \frac{\partial f}{\partial x}\Big|_{x=x_k}$$

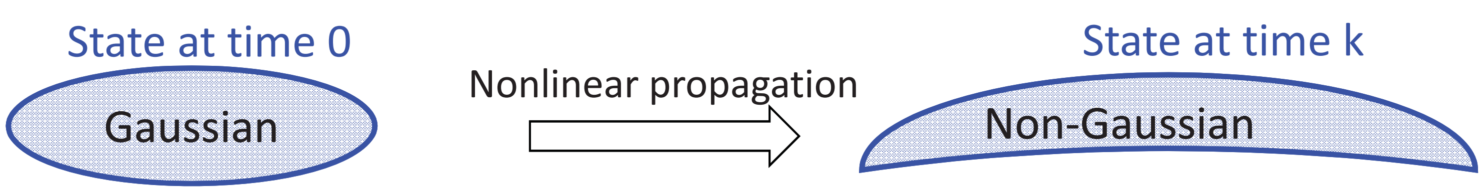
$$H_k = \frac{\partial h}{\partial x}\Big|_{x=x_k}$$

$$v_k \sim N\left(0, R\right)$$

$$w_{k-1} \sim N\left(0, Q\right)$$



ISSUE



CLASSICAL ESTIMATION METHODS (CONT.)

Unscented Kalman Filter (Julier and Uhlmann 2004 et al)

- Uses a series of weighted sample points to approximate the mean and covariance

$$\boldsymbol{X}_k^0 = x_k$$

$$\boldsymbol{X}_k^i = \left[\bar{x}_k + \left(\sqrt{(L+\lambda)P_k}\right)_i\right], i = 1,\ldots L$$

$$\boldsymbol{X}_k^i = \left[\bar{x}_k - \left(\sqrt{(L+\lambda)P_k}\right)_i\right], i = L+1,\ldots 2L$$

2L+1 Sigma points

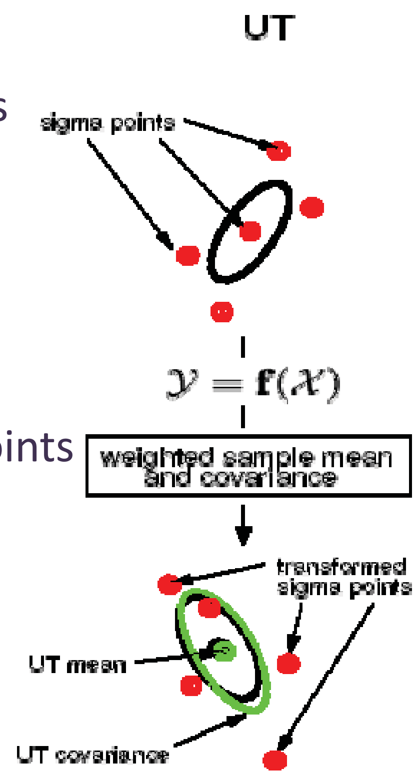
$$W_0^m = \lambda / (L+\lambda)$$

$$W_0^c = \lambda / (L+\lambda) + (1-\alpha^2+\beta)$$

$$W_i^m = W_i^c = 1 / (2(L+\lambda)) \quad \text{for } i = 1,\ldots.2L,$$

Time ***i*** sigma points and weights

where L is the dimension of the state and k is time



http://www.cslu.ogi.edu/nsel/ukf/node6.html

ISSUE

State at time t_0

Gaussian

Nonlinear propagation

State at time t_k

Non-Gaussian

PARTICLE FILTER

Based on Bayes Theory

Posterior Density

Likelihood

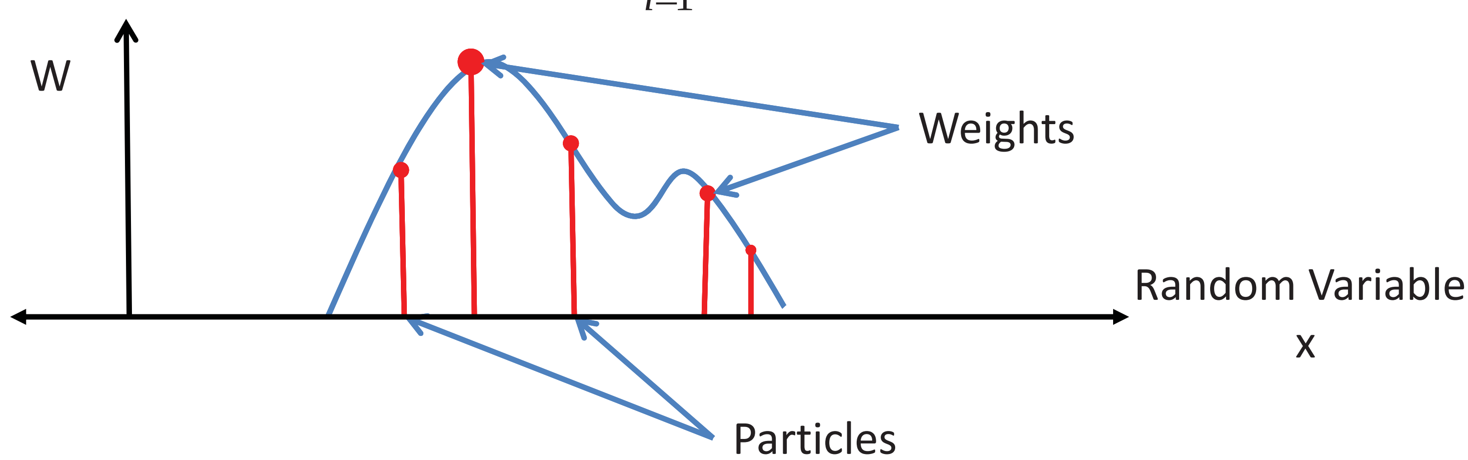
Prior Density

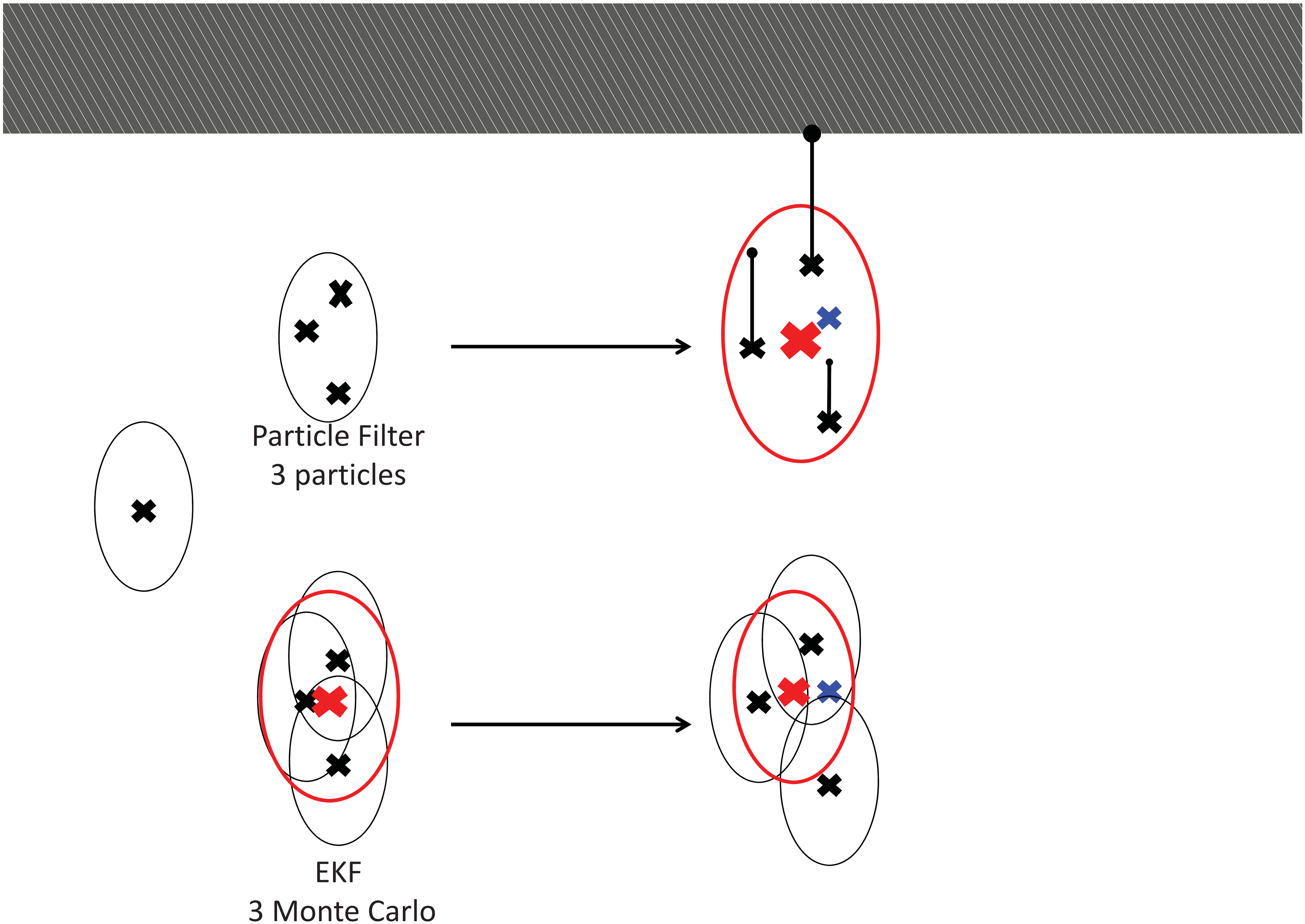
$$p(X \mid Y) = \frac{p(Y \mid X)\, p(X)}{p(Y)}$$

Normalizing Constant/ (Importance Sampling) Proposal Distribution

Estimate the Posterior Density:

$$\hat{p}(X_k \mid Y_k) \approx \sum_{i=1}^{N_p} W_k^i \times \delta(X_k - X_k^i) \qquad \text{where,} \quad \sum_{i=1}^{N_p} W^i = 1$$





Particle Filter
3 particles
EKF
3 Monte Carlo

FASTICA

$$kurt(w^T x) = E\{(w^T x)^4\} - 3[E\{(w^T x)^2\}]^2 = E\{(w^T x)^4\} - 3\|w\|^4$$

$$J(w) = E\{(w^T x)^4\} - 3\|w\|^4 + F(\|w\|^2)$$

F is a penalty term due to the constraint on w. The online learning algorithm has the form

$$w(t+1) = w(t) \pm \mu(t)[x(t)(w(t)^T x(t))^3 - 3\|w(t)\|^2 w(t) + f(\|w(t)\|^2)w(t)]$$

μ is the learning rate sequence, f is the derivative of F/2

The first two terms in the brackets are the gradient of kurt(w^Tx) and the third term is the gradient of $F(\|w\|^2)$.

The fixed points w of the learning rule takes the expectation of the learning rule and equating the change in weight to zero

$$E\{x(w^T x)^3\} - 3\|w\|^2 w + f(\|w\|^2)w = 0$$

$$w = scalar \times (E\{x(w^T x)^3\} - 3\|w\|^2 w)$$

FastICA

$$w(k) = E\{x(w(k-1)^T x)^3\} - 3w(k-1)$$

Divide w(k) by its norm (scalar $\times$ w)

$|w(k)^T w(k-1)| \to 1,$ let k = k+1

NONLINEAR FACTOR ANALYSIS

The prior assumptions ξ and $\mathcal{H}$ are given as follows:

The variance is parameterized to v = log σ

$$x \sim N(f(s,\theta), diag(e^{2V_n}))$$

$$s \sim N(0, diag(e^{2V_s}))$$

Parameters

$$A \sim N(0, I)$$

$$B_j \sim N(0, diag(e^{2V_{B_j}}))$$

$$a \sim N(m_a, diag(e^{2V_a}))$$

$$b \sim N(m_b, diag(e^{2V_b}))$$

Hyperparameters:

$$v_n \sim N(m_{v_n}, diag(e^{2V_{v_n}}))$$

$$v_s \sim N(m_{v_x}, diag(e^{2V_{v_s}}))$$

$$v_{B_j} \sim N(m_{v_{B_j}}, diag(e^{2V_{v_{B_j}}}))$$

$(m_a, v_a, m_b, v_b, m_{v_n}, v_{v_n}, m_{v_{B_j}}, v_{v_{B_j}}) \sim N(0, 100^2)$ uninformative priors

$$D = \int q(\Theta|\xi) \ln \frac{q(\Theta|\xi)}{p(\Theta|x,\mathcal{H})} d\Theta$$

$$= \int q(\Theta|\xi) \ln \frac{q(\Theta|\xi)}{p(\Theta,x|\mathcal{H})} d\Theta + \ln p(x|\mathcal{H})$$

≠ f(θ)

$$C = \int q(\Theta|\xi) \ln \frac{q(\Theta|\xi)}{p(\Theta,x|\mathcal{H})} d\Theta$$

$$C = E[\ln q(\Theta|\xi)] + E[-\ln p(\Theta,x|\mathcal{H})]$$

$$\min C = C_q + C_p = E[\ln q(\Theta|\xi)] + E[-\ln p(\Theta,x|\mathcal{H})]$$

w.r.t. $\bar{\Theta}$ *and* $\tilde{\Theta}$

$$\frac{\partial C}{\partial \bar{\Theta}} = \frac{\partial C_q}{\partial \bar{\Theta}} + \frac{\partial C_p}{\partial \bar{\Theta}} = 0 \quad and \quad \frac{\partial C}{\partial \tilde{\Theta}} = \frac{\partial C_q}{\partial \tilde{\Theta}} + \frac{\partial C_p}{\partial \tilde{\Theta}} = 0$$

where $\bar{\Theta}$ *and* $\tilde{\Theta}$ *are the mean and variance respectively*

NONLINEAR FACTOR ANALYSIS

$$C = C_q + C_p = E[\ln q(s,\theta \mid \xi)] + E[-\ln p(s,\theta,x \mid \mathrm{H})]$$

$$C_q : \sum_{i,N} E[\ln q(s_i(N))] + \sum_j E[\ln q(\theta_j)] = \sum_{all\_terms} -\tfrac{1}{2} - \tfrac{1}{2}\ln(2\pi\tilde{\Theta}_j)$$

$$q(s_i(N)) = N(s_i(N); \overline{s}_i(N), \tilde{s}_i(N)), \qquad q(\theta_j) = N(\theta_j; \overline{\theta}_j, \tilde{\theta}_j)$$

$$p(s) = N(s; 0, diag(\exp(2v_s)))$$

$$C_p : E[-\ln p(s,\theta,x \mid \mathrm{H})] = E[-\ln p(x \mid s,\theta,\mathrm{H})] + E[-\ln p(s \mid \theta,\mathrm{H})] + E[-\ln p(\theta \mid \mathrm{H})]$$

$$\theta \sim N(m, \exp(2v));$$

$$E[-\ln p(\theta \mid m, v, \mathrm{H})] = \iint -\ln N(\theta; m, \exp(2v)) q(m) q(v)\, dm\, dv$$

$$= \frac{1}{2}\ln(2\pi) + \overline{v}_\theta + [(\overline{\theta} - \overline{m}_\theta)^2 + \tilde{\theta} + \tilde{m}_\theta]\exp(2\tilde{v}_\theta - 2\overline{v}_\theta)$$

where

$$\ln p(\theta \mid m, v, \mathrm{H}) = \ln\left[\frac{1}{\sqrt{2\pi\exp(2v_\theta)}}\exp\left(-\frac{(\theta - m_\theta)^2}{2\exp(2v_\theta)}\right)\right]$$

COMPRESSION METHODS

The uncorrelated sources **s’,** are binned into a 2^7 histogram then transformed into a normalized discrete PMF p for compression.

FAST-FOURIER TRANSFORM

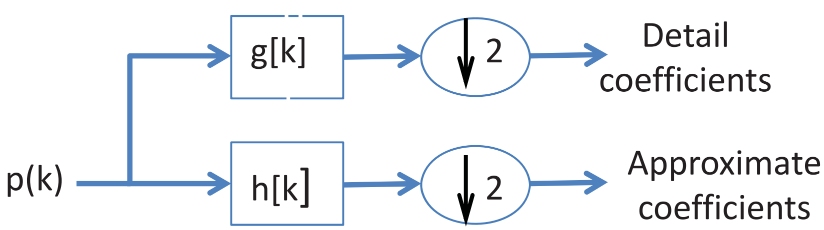
- The FFT calculates the coefficients of the discrete PMF as follows:

$$c(k)=\sum_{j=1}^{n} p(j)e^{\frac{-2\pi i}{n}(j-1)(k-1)} \text{ where } (k=1,\ldots n)$$

p(j) is the discrete probability value normalized from the histogram

WAVELET TRANSFORM

- Use a wavelet filter (Daubechies 2) to extract terms that represent the function p as a function of approximate and detail coefficients



$$p(k)=\sum_{n=-\infty}^{\infty}((\alpha_{j0}[n]\cdot h[2n-k])+(\gamma_j[n]\cdot g[2n-k]))$$

$j0$ is the first level of decomposition

j are the subsequent levels of decomposition

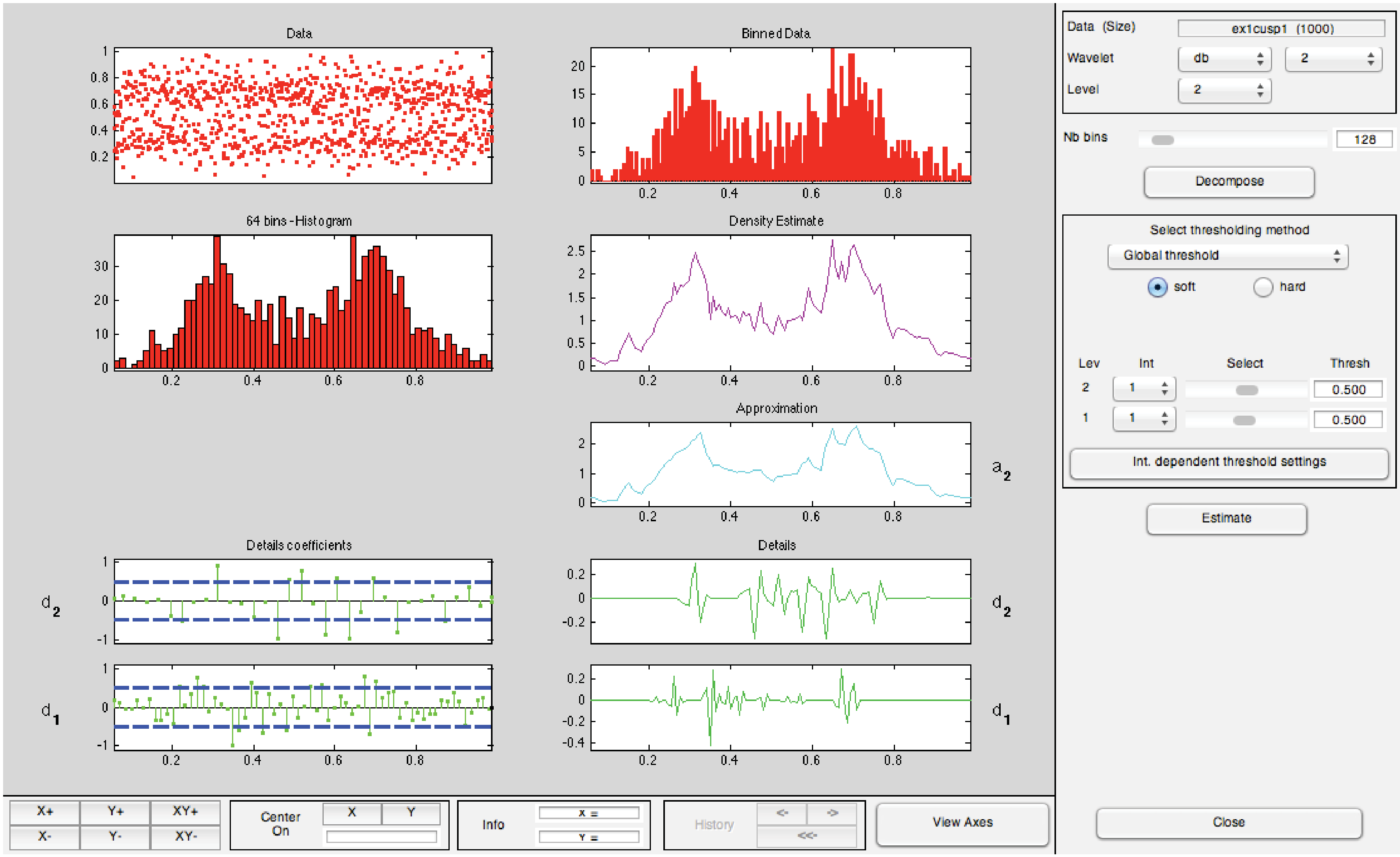
$$c[k]=(p*f)[k]=\sum_{n=-\infty}^{\infty} p[n]*f[2k-n],$$

where $f=\{h,g\},\ c[k]=\{\alpha_{j0}[k],\gamma_j[k]\}$

$$\alpha_{j0}[k]=\sum_{n} p[n]*h[2k-n]$$

$$\gamma_j[k]=\sum_{n} p[n]*g[2k-n]$$

WAVELET COMPRESSION



CONCLUSIONS

1. Demonstrated the use of the Particle Filter:
    1. A nonlinear estimator
    2. Capable of incorporating non-Gaussian uncertainties during predictions

2. Demonstrated the use of the Independent Component Analysis (ICA) and Principal component analysis for dimensional reduction

3. Demonstrated the use of Nonlinear Factor Analysis (NFA) followed by FastICA to achieve uncorrelated states for PDF compression

4. The Wavelet Transform demonstrated better compression results over lower compression rates compared to the Fast-Fourier Transform

5. The Kullback-Leibler divergence and the Kolomogorov-Smirnov tests were used as Information measures to quantify the reconstructions

6. Ephemeris compression and reconstruction were able to achieve great results for compression rates

7. The reduced number of terms can be potentially used for cataloging accurate non-Gaussian distributed ephemeris

FUTURE WORK

1. Implementation of the Particle Filter using adaptive number of samples for numeric efficiency and redundancy elimination.
2. Include a higher fidelity force model to incorporate additional relevant perturbations such as atmospheric drag, solar radiation pressure, third body perturbations (Sun and moon), J2 effects etc.
3. Develop a systematic way of determining the number of hidden neurons to use as well as the number of iterations during the training phase (possibly optimize based on constraints)
4. Other decorrelation methods that do not assume Gaussian initial conditions as the NFA does, could be implemented ex. The Nonlinear Independent Factor Analysis (NIFA) (nonlinear counterpart of the ICA)
5. Optimize the value n in the determination of the number of bins 2^n for the histogram generator
6. Use other smooth functions to represent the PDF of the components versus a PMF before compression
7. Determine the limitations of NFA, NIFA with the number of particles (apart from computational cost)