Learning Relational Kalman Filtering

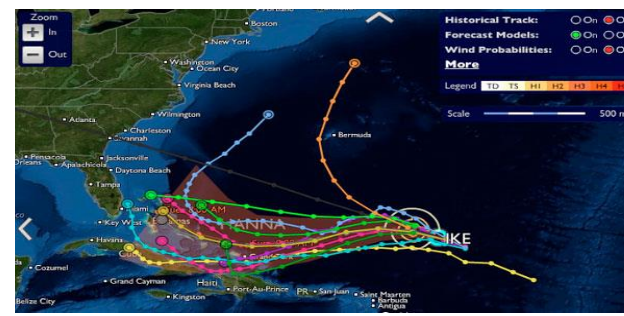
Jaesik Choi*
Ulsan National Institute of Science and Technology
***Speaker**

Eyal Amir, Tiangfang Xu and Albert J. Valocchi
University of Illinois at Urbana-Champaign

Kalman Filter

- Kalman Filter is an algorithm which produces estimates of unknown variables given a series of measurements (w/ noise) over time.

- Numerous applications in
  - Robot localization
  - Econometrics (time series)
  - Military: rocket and missile guidance
  - Autopilot
  - Weather forecasting
  - Speech enhancement
  - …



Kalman Filtering: an example

- Input statements
  - **John**’s house price was **$0.39M at 2014**.
  - Each year, **John**’s house price **increases 5%**.
  - **John**’s house price is around the sold price.
  - **John**’s house is sold sporadically.
- Question: what is the price of John’s house each year?

Kalman Filtering: an example

- Input statements
  - **John**'s house price was **$0.39M at 2014**.
  - Each year, **John**'s house price **increases 5%**.
  - **John**'s house price is around the sold price.
  - **John**'s house is sold sporadically.

- Question: what is the price of John's house each year?

Transition Model

$$x^{15'}_{John} = 1.05x^{14'}_{John} + \varepsilon_{tras}$$

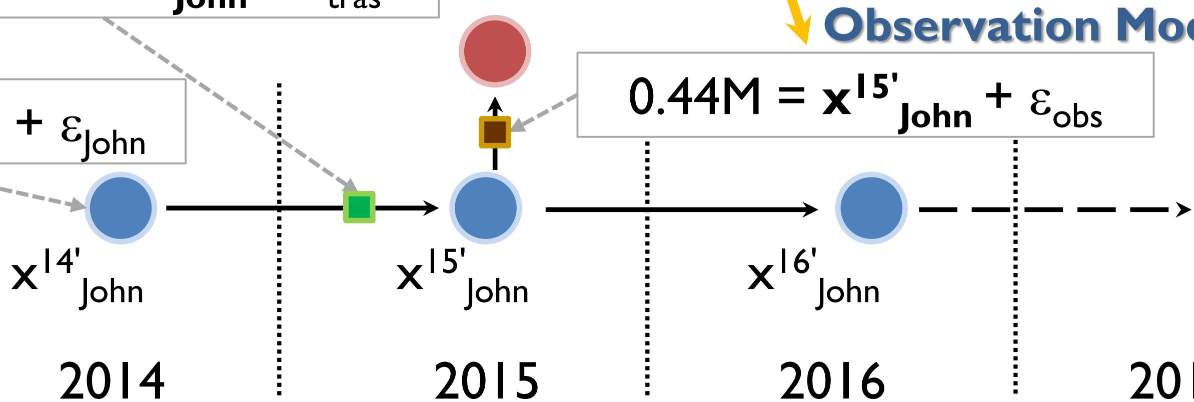
$$x^{14'}_{John} = 0.39M + \varepsilon_{John}$$

Sold at $0.44M

$\varepsilon \sim N(0,\sigma^2)$

Observation Model

$$0.44M = x^{15'}_{John} + \varepsilon_{obs}$$



Kalman Filtering: an example

- Input statements
  - **John**'s house price was **\$0.39M at 2014**.
  - Each year, **John**'s house price **increases 5%**.
  - **John**'s house price is around the sold price.
  - **John**'s house is sold sporadically.

- Question: what is the price of John's house each year?

Transition Model

$$x^{15'}_{John} = 1.05x^{14'}_{John} + \varepsilon_{tras}$$

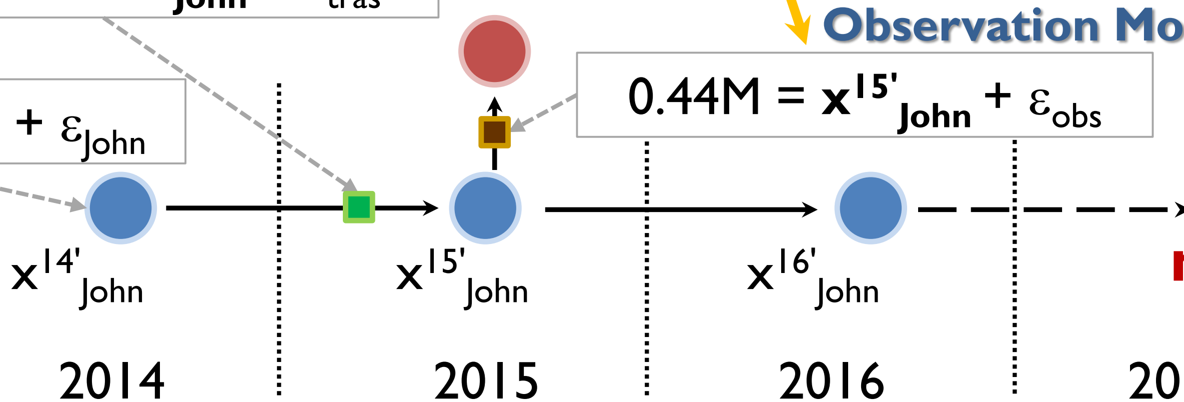
Sold at \$0.44M

$\varepsilon \sim N(0,\sigma^2)$

Observation Model

$$0.44M = x^{15'}_{John} + \varepsilon_{obs}$$

$$x^{14'}_{John} = 0.39M + \varepsilon_{John}$$



$O(n^3)$

n: # of rvs

Relational Kalman Filtering (RFK):

[Choi Guzman, Amir, IJCAI-11] & [Ahmadi, Kersting, Sanner, IJCAI-11]

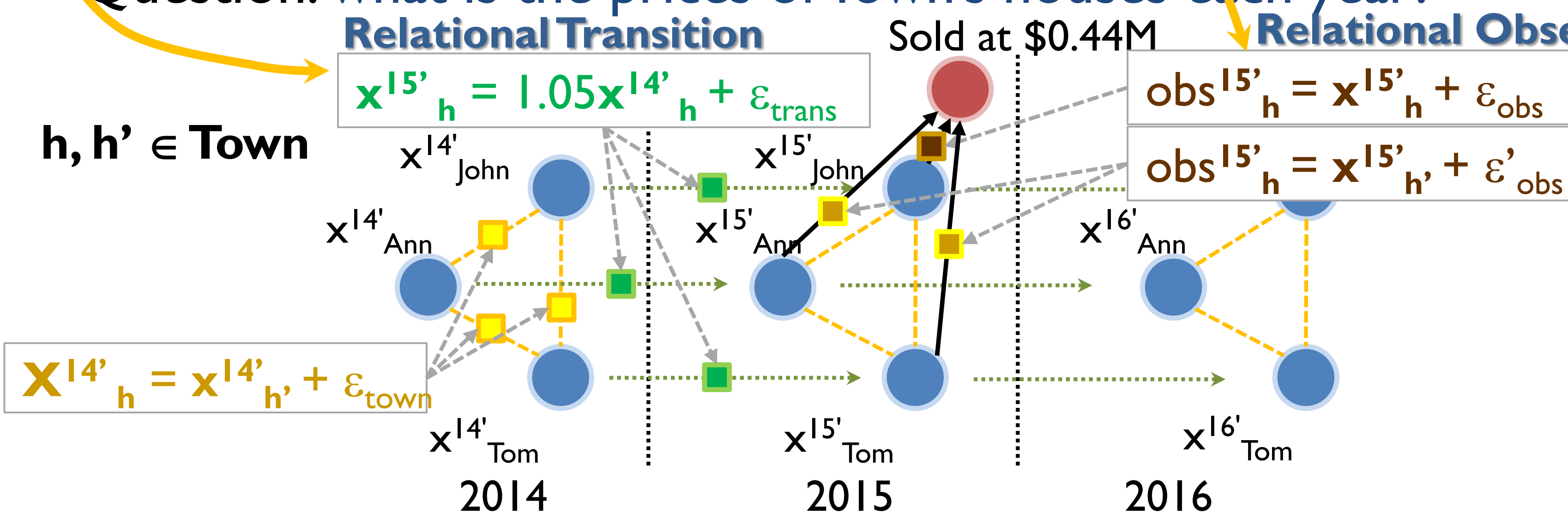
- Input statements
  - **Town is a set of houses.**
  - **Town's** houses have initial prices **at 2014**.
  - Each year, **Town**'s house prices **increase 5%**.
  - **Town**'s house prices are around sold prices.
  - **Town**'s houses are sold sporadically.
- Question: what is the prices of Town's houses each year?

Relational Kalman Filtering:

[Choi Guzman, Amir, IJCAI-11] & [Ahmadi, Kersting, Sanner, IJCAI-11]

- Input statements
  - **Town is a set of houses.**
  - **Town's** houses have initial prices **at 2014**.
  - Each year, **Town**'s house prices **increase 5%**.
  - **Town**'s house prices are around sold prices.
  - **Town**'s houses are sold sporadically.

- Question: what is the prices of Town's houses each year?

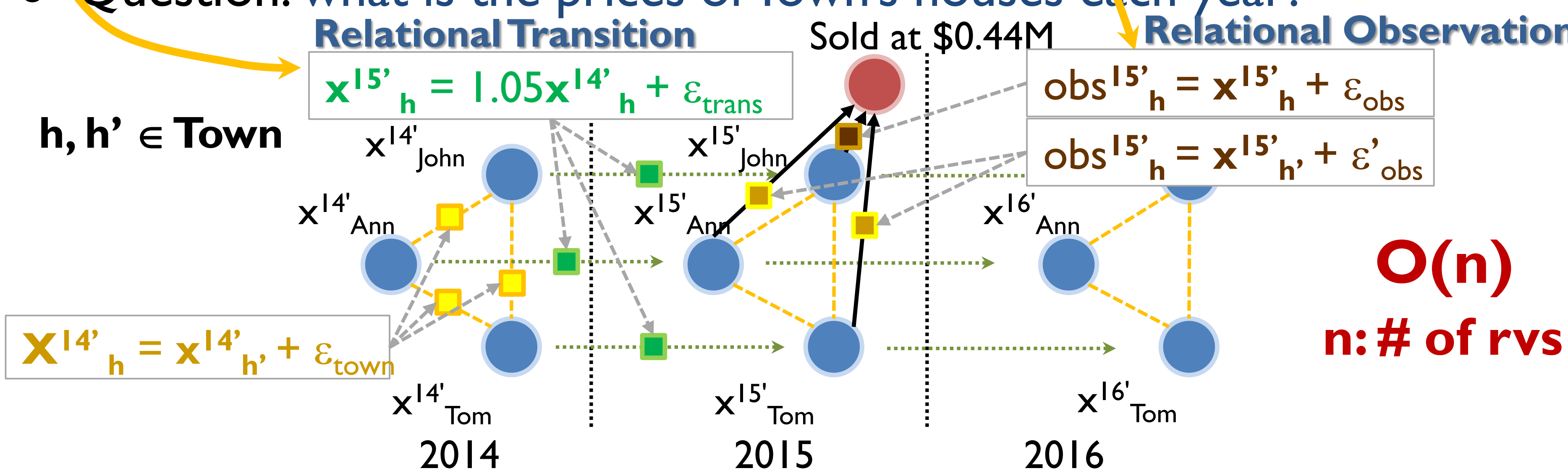


Relational Kalman Filtering:

[Choi Guzman, Amir, IJCAI-11] & [Ahmadi, Kersting, Sanner, IJCAI-11]

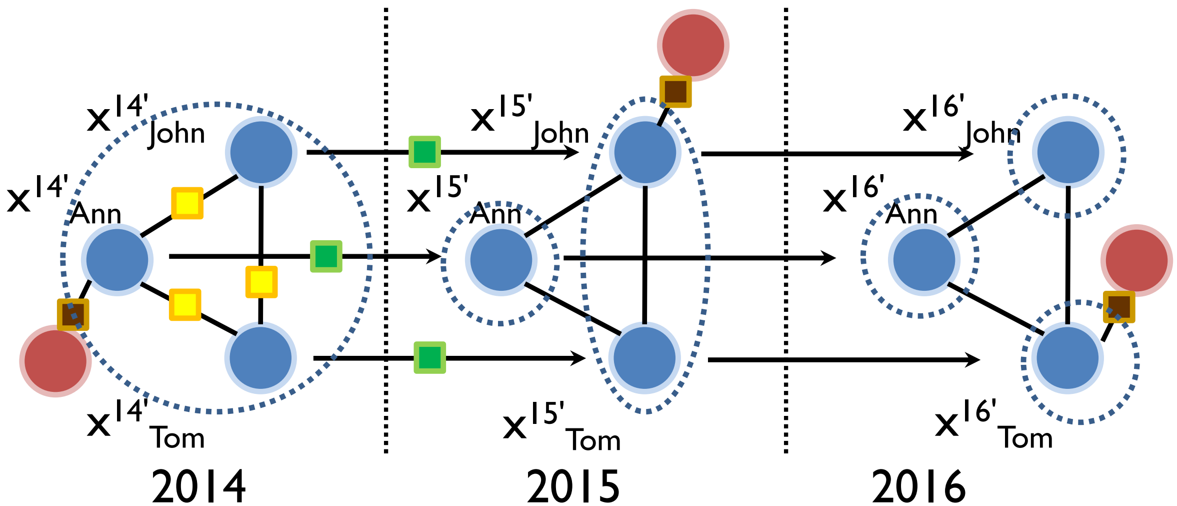
- Input statements
  - **Town is a set of houses.**
  - **Town's** houses have initial prices **at 2014**.
  - Each year, **Town**'s house prices **increase 5%**.
  - **Town**'s house prices are around sold prices.
  - **Town**'s houses are sold sporadically.

- Question: what is the prices of Town's houses each year?

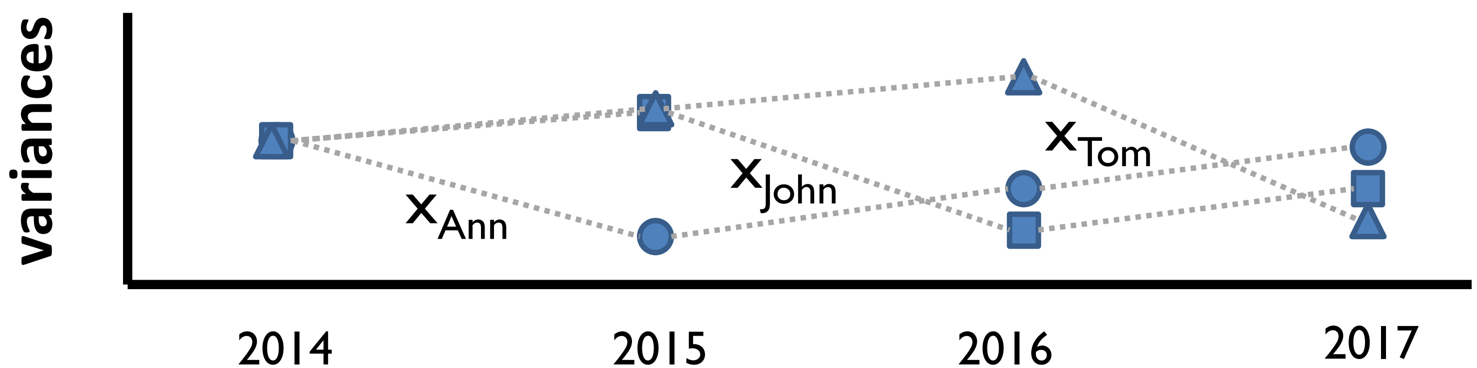


Current Issue:

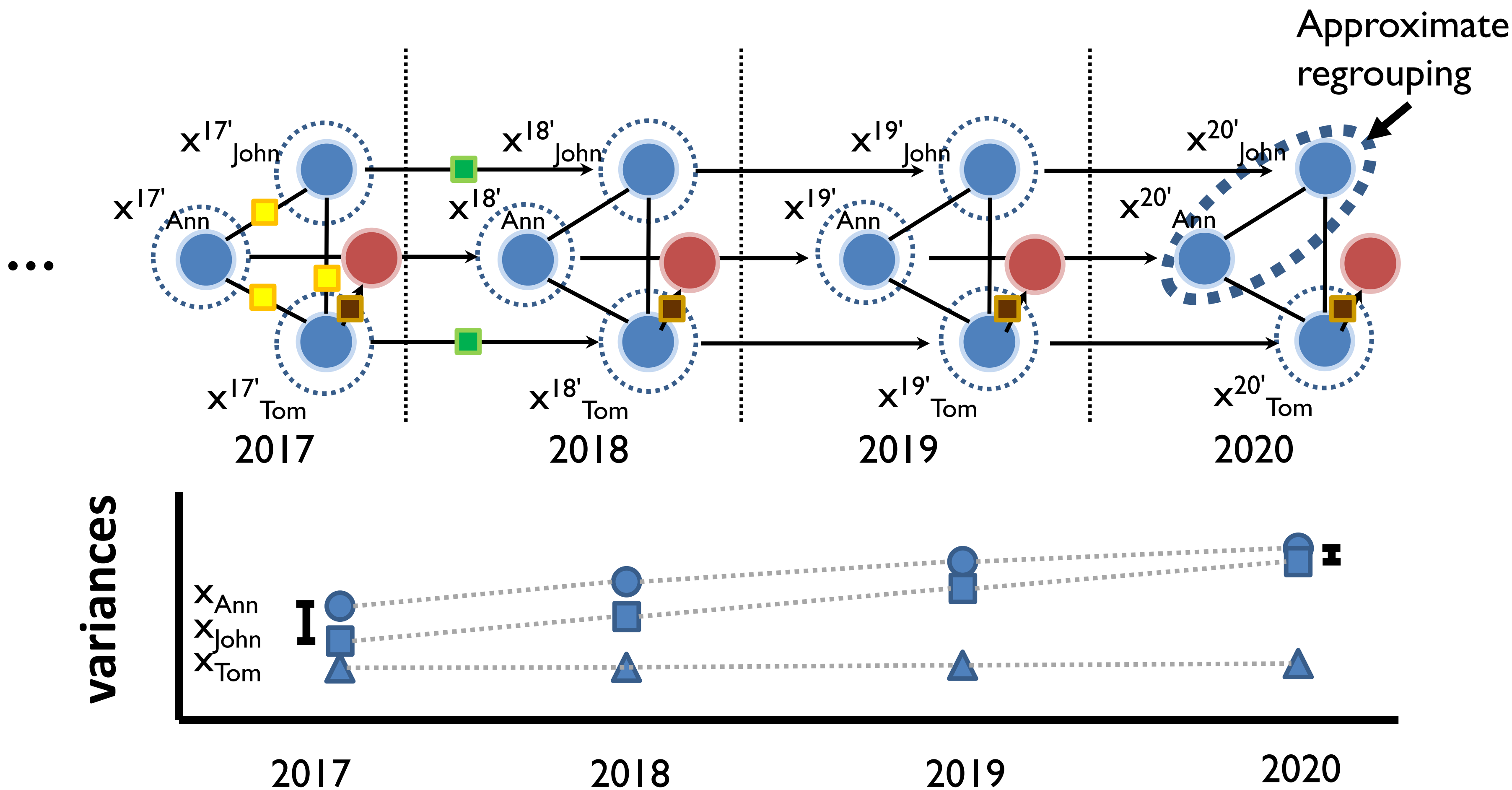
Sparse Observations → Model Degenerations



Relational O(n) → Ground O(n³)



Main Finding:
Relational Obs Prevent RFK from Degenerating!



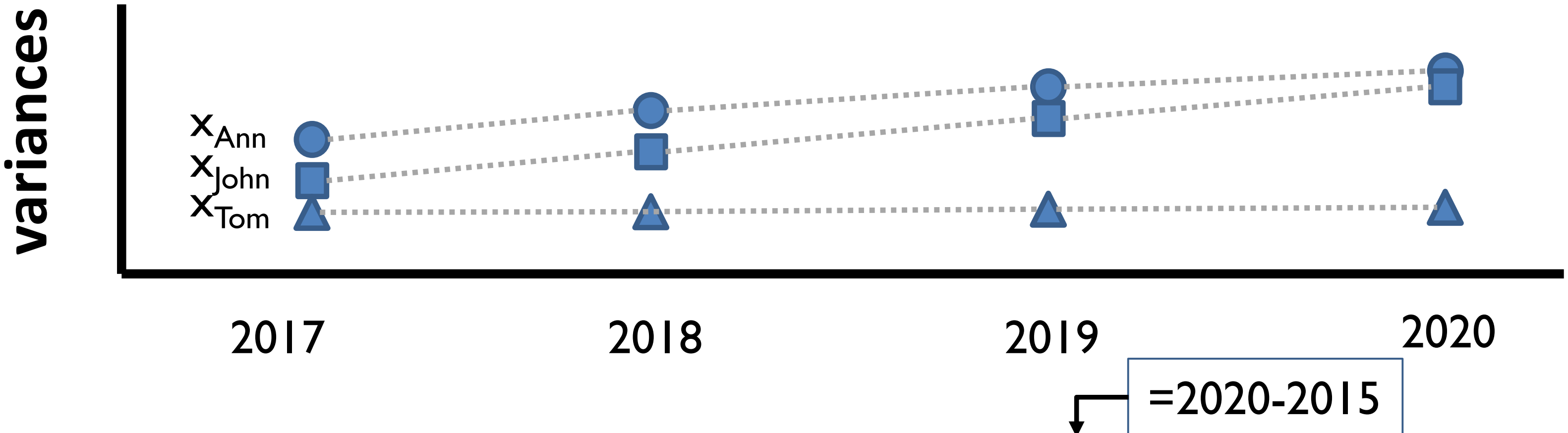
Main Theoretical Result

Theorem: *For two rvs (X and X') in a set (atom) A of RKF*

(1) X and X' have ***no obs for the previous k steps,***
(2) At least one obs is made to the other rvs *in A each time step*

Then, for c>1, the following holds,

$$|Var(X) - Var(X')| \leq O(c^{-k}).$$



$$Var(X_{John}^{20}) - Var(X_{Ann}^{20}) \leq O(c^{-5}).$$

When conditions (1) and (2) are satisfied,

We can recover a relational model out of a degenerated model!

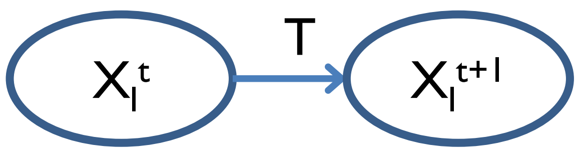
Parameter Learning for RKF Models

Parameter Learning Problem:

Input:
- (Relational) Sets of random variables
- A sequence of observations

G_T, G_O: Gaussian Noise

Transition Models

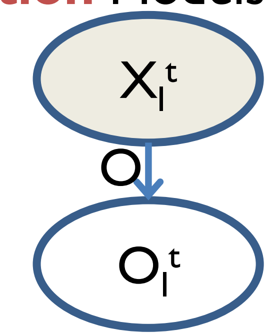


$X_1^{t+1} = X_1^t + \mathbf{B_T} U_1^t + \mathbf{G_T}$

Observation Models

Hidden

Obs



$O_1^t = \mathbf{H_O} X_1^t + \mathbf{G_O}$

$U_t(x)$: user input for x at time t

B_T, H_O : coefficients

Output: ○ Relational Parameters for RKFs ($\mathbf{B_T}, \mathbf{G_T}, \mathbf{H_O}, \mathbf{G_O}$)

Parameter Learning for RKF Models

Proposition:* *Maximum Likelihood Estimates (MLEs) of RKF models* (B_T, G_T, H_O, G_O) *are empirical means of MLEs of the KF.

In case of, the covariance matrix (e.g., G_T, and G_O)

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

The MLE of KF

$$\frac{\sum_i b_{ii}}{n} = b$$

$$\frac{\sum_{ij(i\neq j)} b_{ij}}{n(n-1)} = b'$$

$$\begin{bmatrix} b & b' & \cdots & b' \\ b' & b & \cdots & b' \\ \vdots & \vdots & \ddots & \vdots \\ b' & b' & \cdots & b \end{bmatrix}$$

The MLE of RKF

(1) Learn Ground KF

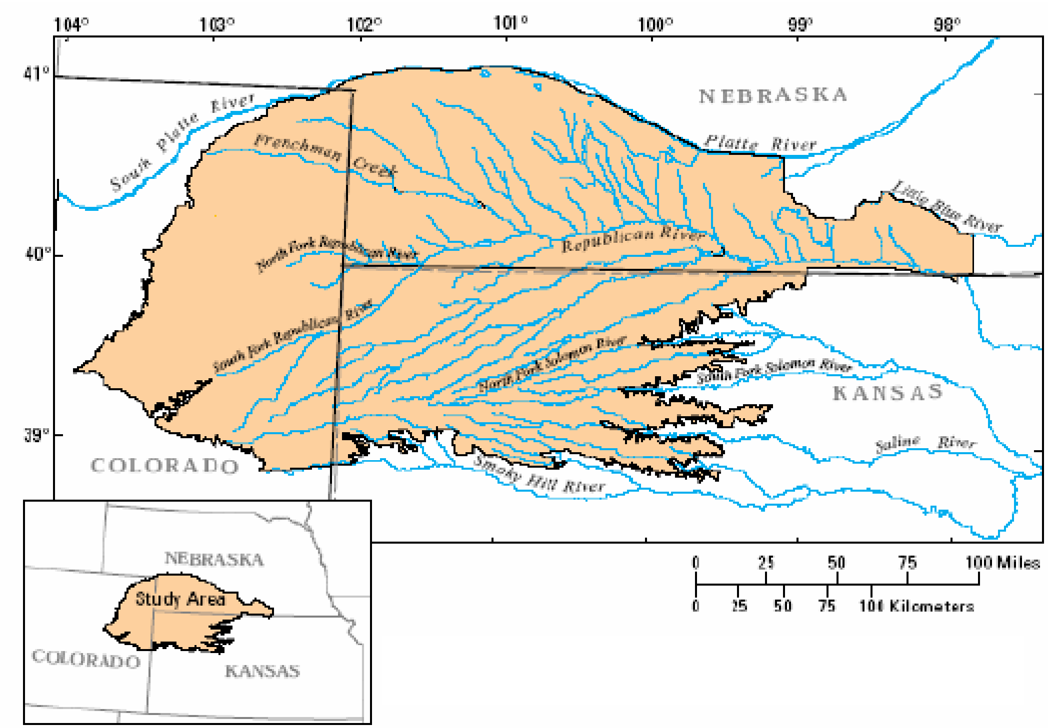
[Ghahramani and Hinton, 1996]
[K. Murphy, 1998]

(2) BlockAverage Operation

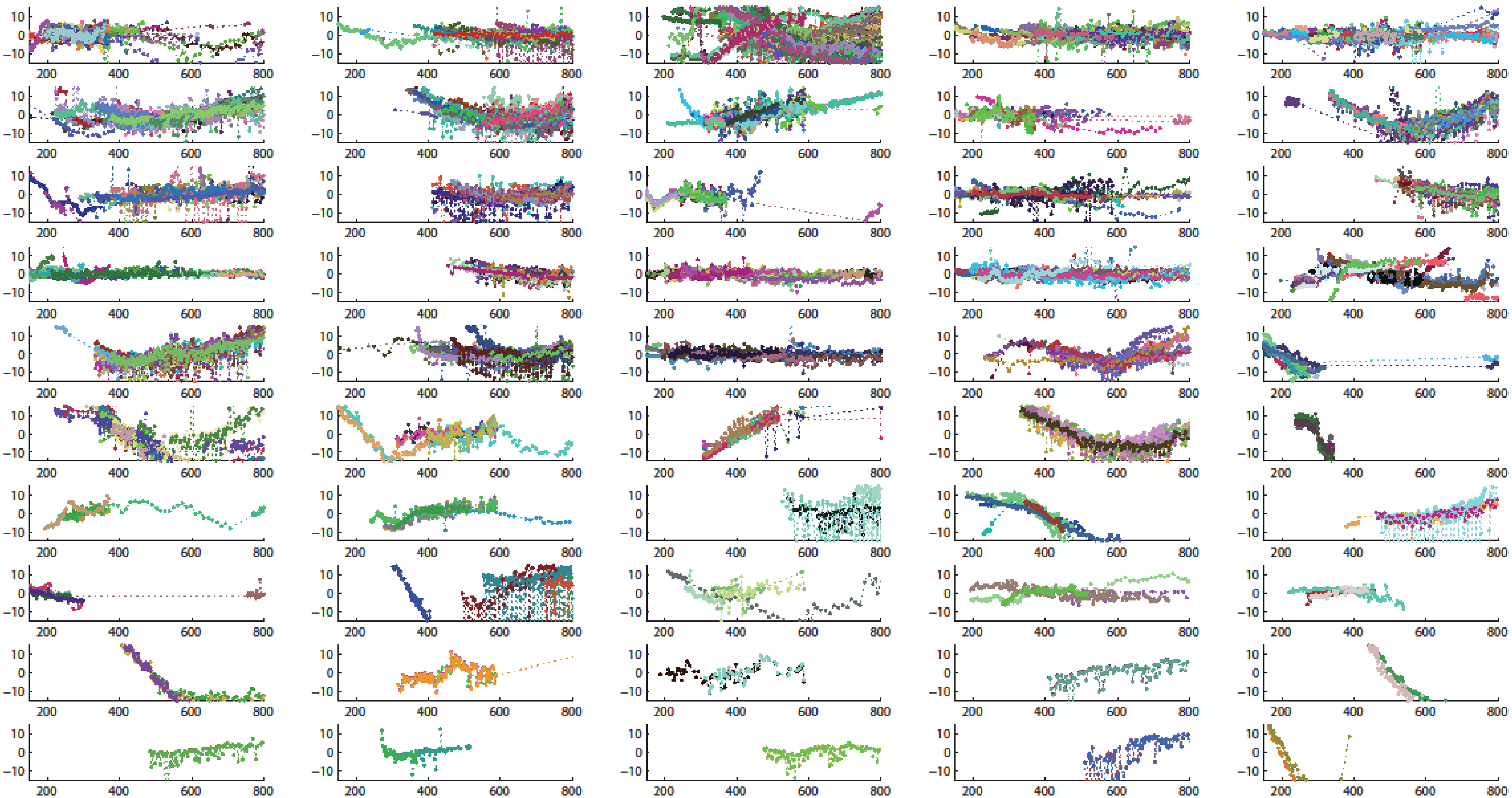
(3) Derive RKF

Experiments (Groundwater Models)

- Dataset: RRCA (Republican River Compact Administration)
  - The model has measures (water levels) for 3078 water wells.
  - The measures span from 1918 to 2007 (about 900 months).
  - It has over 300,000 measurements.

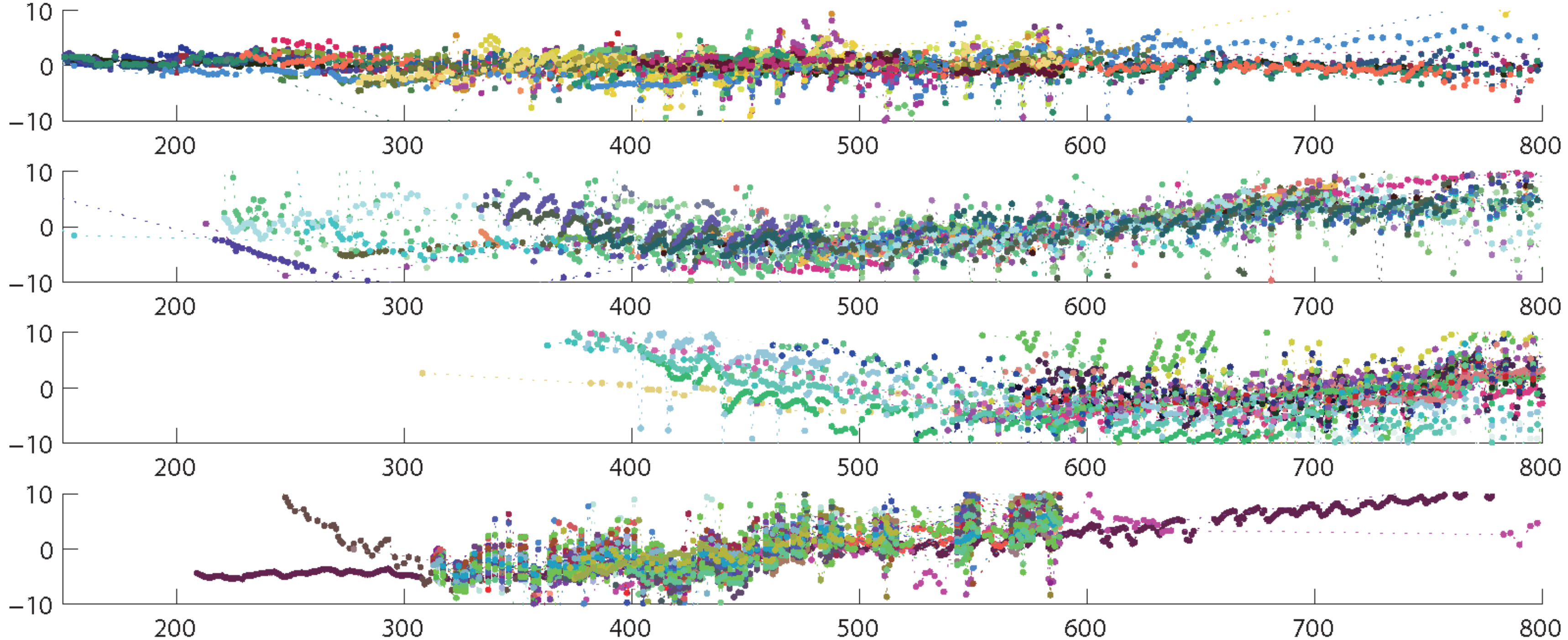


Relational Information (Clustering Wells)
by Spectral Clustering [Ng, Jordan, Weiss, 2001]



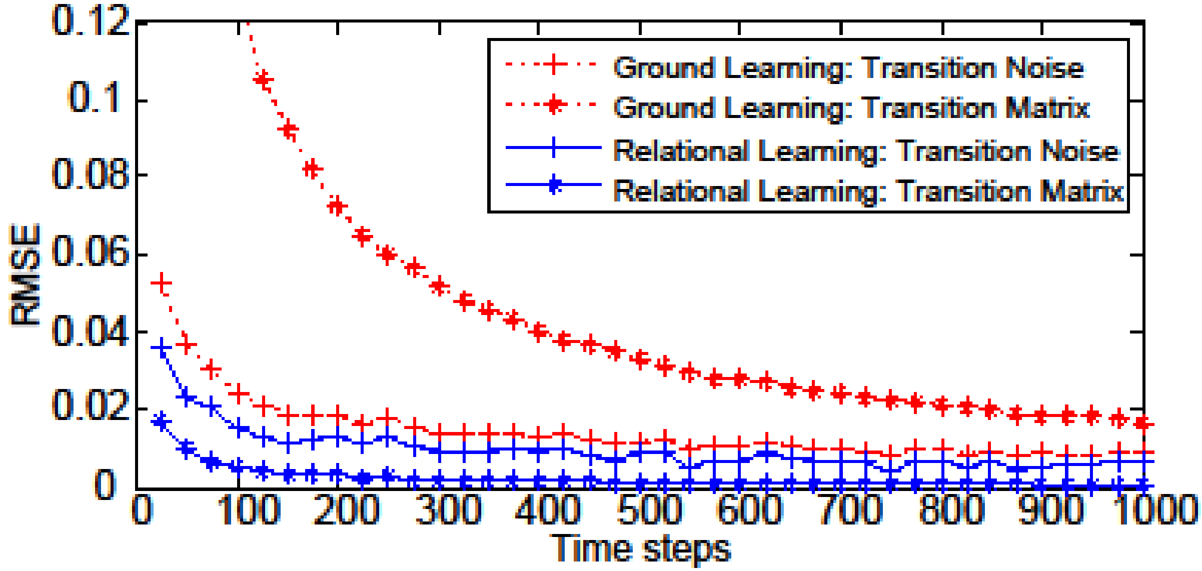
Relational Information (Clustering Wells)

by Spectral Clustering [Ng, Jordan, Weiss, 2001]



Learning and Prediction with RKF

- Parameter Learning in simulation



- Prediction accuracy on the RRCA model

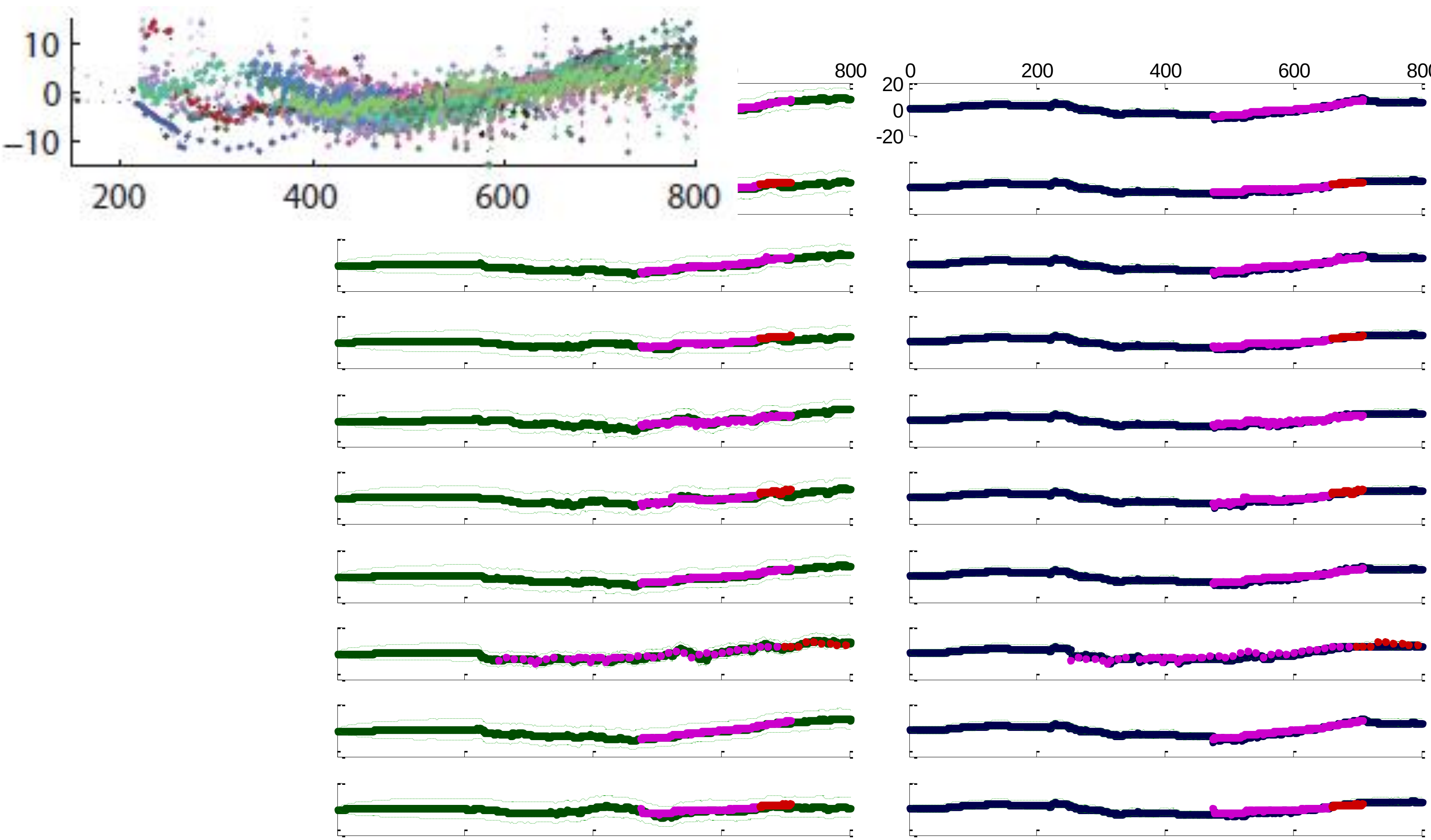
	Vanilla KF	Relational KF
RMSE (Root Mean Square Error)	5.10	**4.36**
Negative Log of Probability -log(P(data|pred))	4.91	**3.88**

Conclusions

- We show that **relational obs may prevent RKFs from degenerating**
- We present **the first parameter learning algorithm for relational continuous models**
- **S/W download soon will be available at http://pail.unist.ac.kr/LRKF/**

Thank you!

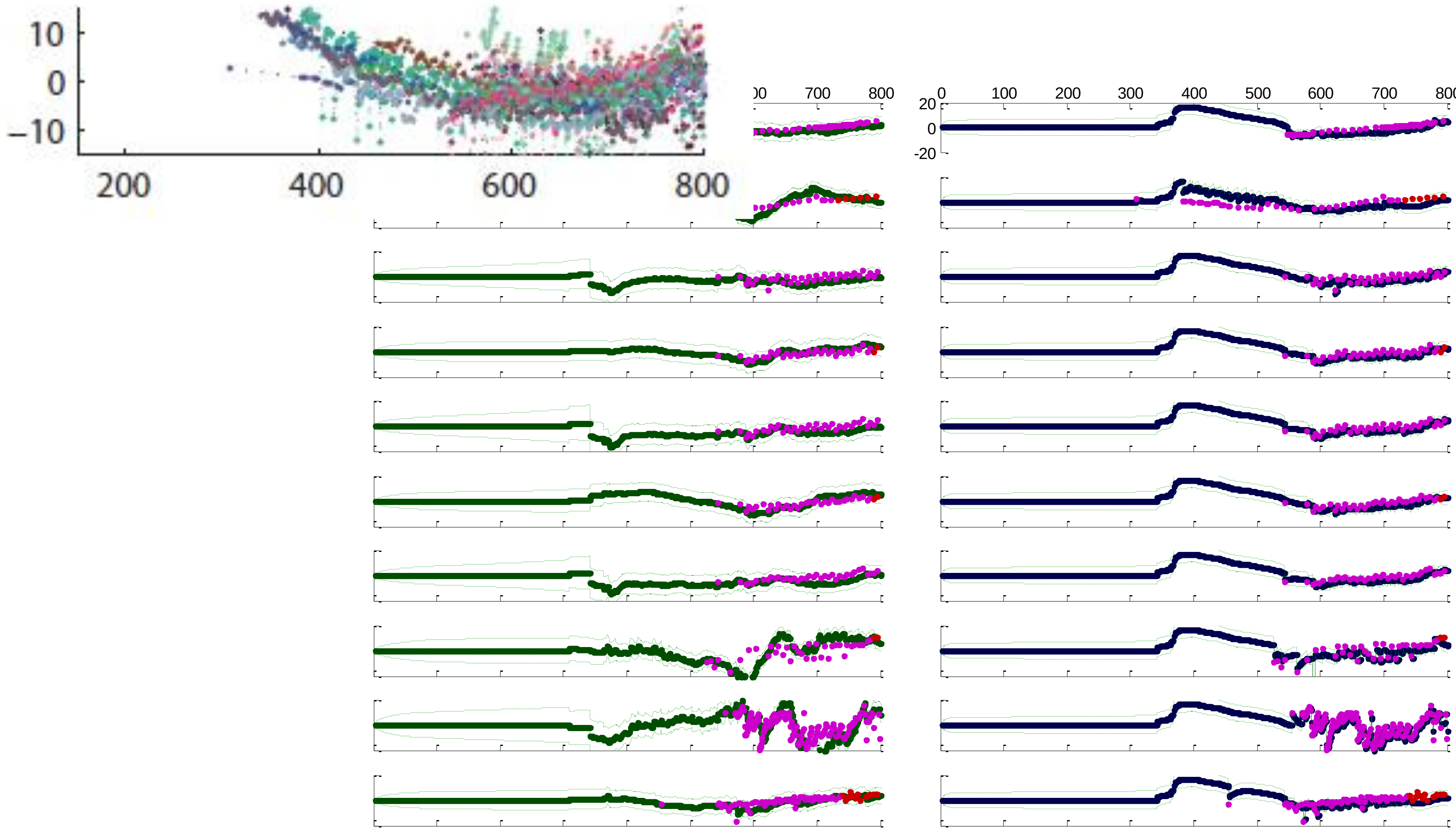
State Estimation: Vanilla KF vs Relational KF



Vanilla KF

Relational KF

State Estimation: Vanilla KF vs Relational KF



Vanilla KF   Relational KF

Dense Observations → No Degeneration