The Unscented Particle Filter

Rudolph van der Merwe
(OGI)

Nando de Freitas
(UC Berkeley)

Arnaud Doucet
(Cambridge University)

Eric Wan
(OGI)

Outline

- **Optimal Estimation & Filtering**
- **Optimal Recursive Bayesian Solution**
- **Practical Solutions**
  - Gaussian approximations (EKF, UKF)
  - Sequential Monte Carlo methods (Particle Filters)
- **The Unscented Particle Filter**
  - The Unscented Transformation and UKF
  - Applications of UT/UKF to Particle Filters
- **Experimental Results**
- **Conclusions**

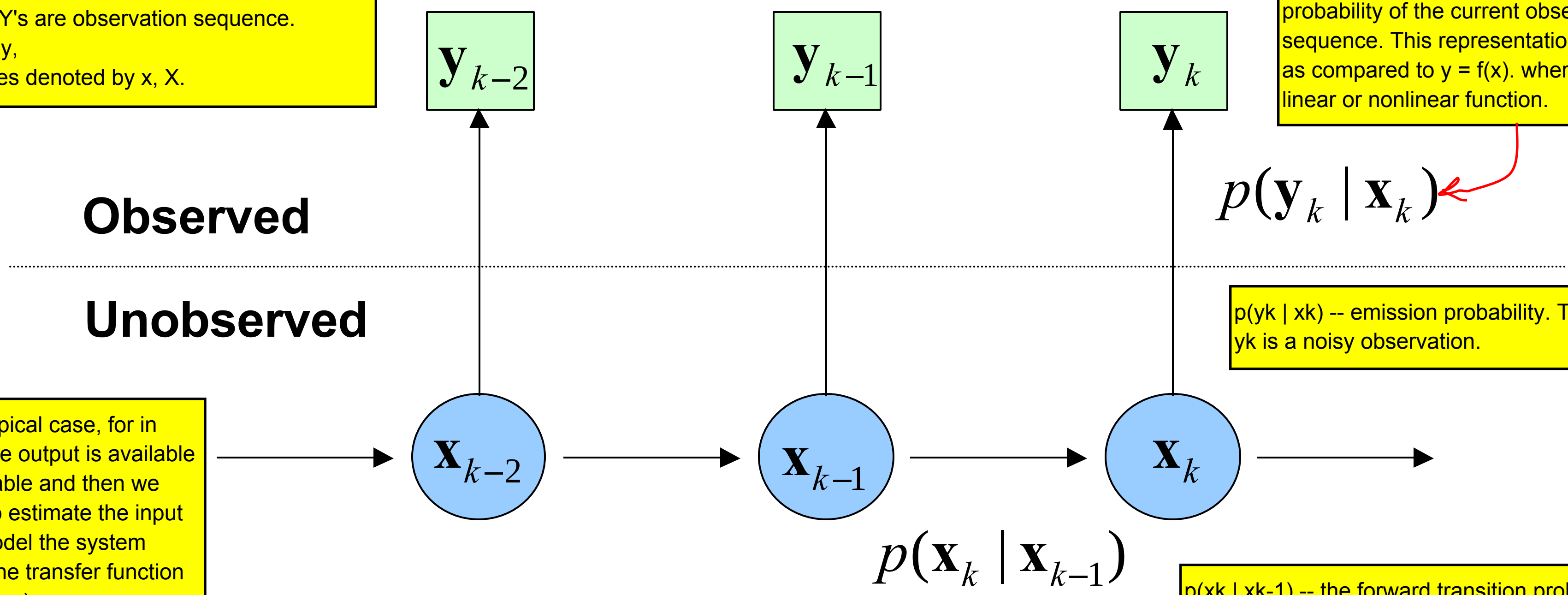
Filtering

- **General problem statement**

In Speech : y implies utterences, or speech data.
x implies articulator positions, vocal tract parameters.

small y's are instantaneous observations
capital Y's are observation sequence.
Similarly,
for states denoted by x, X.

This is a nice equation - representing the probability of the current observation given the sequence. This representation is more appropriate as compared to y = f(x). where in f(x), would be linear or nonlinear function.



p(yk | xk) -- emission probability. This is true as we know, yk is a noisy observation.

This is a typical case, for in practice, the output is available or measurable and then we either try to estimate the input or try to model the system (estimate the transfer function of the system)

p(xk | xk-1) -- the forward transition probability.

- Filtering is the problem of sequentially estimating the states (parameters or hidden variables) of a system as a set of observations become available on-line.

Filtering

- **Solution of sequential estimation problem given by**
  - Posterior density : $p(\mathbf{X}_k \mid \mathbf{Y}_k)$

This is posterior density because we are estimating the current state after the current observation has been made. This is related to evidence and prior.

$$\mathbf{X}_k = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}$$
$$\mathbf{Y}_k = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_k\}$$

  - By recursively computing a marginal of the posterior, the *filtering density*,

$$p(\mathbf{x}_k \mid \mathbf{Y}_k)$$

one need not keep track of the complete history of the states.

Filtering

- **Given the filtering density, a number of estimates of the system state can be calculated:**

Still right now we are not sure of how p(xk|Yk) is estimated.

- Mean (optimal MMSE estimate of state)

We need to check the notation used here: why is the xk within the expectation NOT BOLD? Find Out!!

$$\hat{\mathbf{x}}_k = E[x_k \mid \mathbf{Y}_k] = \int \mathbf{x}_k p(\mathbf{x}_k \mid \mathbf{Y}_k) d\mathbf{x}_k$$

- Mode

The mode is defined as the value with the maximum density function.

- Median

The median is defined as the value in the middle of the distribution if the values are arranged in ascending or descending order.

- Confidence intervals

The degree of certainity with which the value will lie with the threshold. This threshold is called as confidence intervals

- Kurtosis, etc.

Fourth order expectation of the value is called Kurtosis. This is not the formal definition, but i am not able to paste the formula here, but i still wanted to give some idea as to what Kurtosis is, hence the mention. For more information, refer any statistics book.

State Space Formulation of System

- **General discrete-time nonlinear, non-Gaussian dynamic system**

This representation is used for writing the prior equation.

state

process noise

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{v}_{k-1})$$

known input

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{n}_k)$$

noisy observation

measurement noise

The assumption is specific to the nonlinear sate space models. for ref. www.cavs.msstate.edu/~patil/nonlinear_state_space_models.pdf, page 3 out of page 8, section 2.1 The Probablistic Model

- Assumptions : Assumptions laid within the format of discussion. These are not the assumptions from the State Space Formulations.

1) States follow a first order Markov process

assumed that the current state is dependent only on the previous state.

$$p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \cdots, \mathbf{x}_0) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$$

The assumption validity needs to be checked for the case of nonlinear system. But, to start with the assumption offers simplicity in calculations.

2) Observations independent given the states

Where A is either one observation (any one) yn or any set of previous observations. Yn. Actually here n taking any value between o to k.

$$p(\mathbf{y}_k \mid \mathbf{x}_k, A) = p(\mathbf{y}_k \mid \mathbf{x}_k)$$

Assumption # 2 is used in finding the estimate of filtering density.

Slide from : cslu.ece.ogi.edu/publications/ps/UPF_CSLU_talk.pdf
Our reference : www.cavs.msstate.edu/~patil/particle_filter/UPF_CSLU_talk.pdf

Recursive Bayesian Estimation

- **Given this state space model, how do we recursively estimate the filtering density ?**

First equation is got from the Basic Baye's Rule.

Predicte (update) stage

p(a | b) = p(b | a) p(a) / p(b)

Joint Probability Distribution of current observation, observation sequence til previous observations given the current state.
Similarly, the denominator consists of Joint Probability Distribution of current observation and the previous observation sequence.

p(Yk | xk)
= p(yk, yk-1, yk-2, ... y0 | xk)
= club yk-1, yk-2, y0 as Yk-1
= p(yk, Yk-1 | xk)

is P(a,b) = P(a|b) P(b) ?

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \frac{p(\mathbf{Y}_k|\mathbf{x}_k)p(\mathbf{x}_k)}{p(\mathbf{Y}_k)}$$

$$= \frac{p(\mathbf{y}_k,\mathbf{Y}_{k-1}|\mathbf{x}_k)p(\mathbf{x}_k)}{p(\mathbf{y}_k,\mathbf{Y}_{k-1})}$$

$$= \frac{p(\mathbf{y}_k|\mathbf{Y}_{k-1},\mathbf{x}_k)p(\mathbf{Y}_{k-1}|\mathbf{x}_k)p(\mathbf{x}_k)}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})p(\mathbf{Y}_{k-1})}$$

$$= \frac{p(\mathbf{y}_k|\mathbf{Y}_{k-1},\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Y}_{k-1})p(\mathbf{Y}_{k-1})p(\mathbf{x}_k)}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})p(\mathbf{Y}_{k-1})p(\mathbf{x}_k)}$$

$$= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Y}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})}$$

observations independent given the states. This assumption is used.

normalising factor or constant - depends on the likelihood function p (yk / xk) defined in the measurement model

Please also refer - /www.cavs.msstate.edu/~patil/particle_filter/ Tutorial_on_particlefilter.pdf

Recursive Bayesian Estimation

This parameter we want to estimate -

update stage :

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)\,p(\mathbf{x}_k|\mathbf{Y}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})}$$

likelihood

prior

posterior

evidence

Naveen's Deivation:: [Dec 16, 2004]
IMP. Points :
1. P(a,b) = P(a | b) P(b) = P(b | a) P(a)
2. P(a) = INTEGRATION P(a,b) db
probability of a is integration of joint probability of a and b over the variable b.

So:

P(xk | Yk-1) = INTEGRATE
P(xk | xk-1, Yk-1) P(xk-1 | yk-1) dxk-1
= INTEGRATE
P(xk | xk-1) P(xk-1 | yk-1) dxk-1

Here, using the State-Space Formulation on Page 6 of this presentation, we get the Prior equation.

Prediction stage :

- **Prior** : $p(\mathbf{x}_k|\mathbf{Y}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})\,p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})\,d\mathbf{x}_{k-1}$

transition density given by *process model*

Some Authors say this [Prior and Evidence] as a STRAY gener
Chapman-Kolmovgrov Equation. Well Naveen PROVED otherw
Audience should go ahead and try to prove them to be correct.

Equation on Evidence is the Main Point of our Understanding and basis for our development Idea. Doesn't this look similar to Linear Prediction with a different that Probability is thrown out here instead of our normal Linear Predictor Coefficients.

(Propagation of past state into future before new observation is made.)

also refer eqn (3) of IEEE xsactions on signal processing, feb. 02, pp 175

Likelihood : defined in terms of *observation model*

Called the normalising constant - also refer eqn (5) of IEEE xsactions on signal processing, feb. 02, pp 175

- **Evidence** : $p(\mathbf{y}_k|\mathbf{Y}_{k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k)\,p(\mathbf{x}_k|\mathbf{Y}_{k-1})\,d\mathbf{x}_k$

Using the same explanation used for Prior, alongwith the Second assumption on the Page 6 of this presentation, that Given the state, the observation is independent of the previous observation.
P(yk | Yk-1) = INTEGRATE P(yk | xk, Yk-1) P(xk | Yk-1) dxk = INTEGRATE P(yk | xk) P(xk | Yk-1) dxk

Practical Solutions

- **Gaussian Approximations**

  Kalman filters - Extended KF [EKF], Unscented KF [UKF], Grid-based filters, Gaussian-sum filters

HMM filters are an application of approximate grid-based methods in a fixed-interval smoothing context and have been extensively used in speech processing.
In HMM based tracking - Viterbi Algorithm is used to calculate the maximum posteriori estimate of the path through the trellis.
Another approach is Baum-Welch Algorithm.
The Viterbi and Baum-Welch Algorithms are frequently applied when the sate space is approximated to be discrete. The algorithms are optimal IFF the underlying state space is truly discrete in nature.

- **Perfect Monte Carlo Simulation**

- **Sequential Monte Carlo Methods : "Particle Filters"**

  Called by different name : Bootstrap Filtering, Condensation Algorithm, Particle Filtering, Interacting Particle Filtering, Survival of the fittest.
  It is a technique for implementing a recursive Bayesian filter by MC simulations.

  - Bayesian Importance Sampling

    The weights are chosen using the principle of importance sampling.

  - Sampling-importance resampling (SIR)

Gaussian Approximations

- **Most common approach.**
- **Assume all RV statistics are *Gaussian.***
- **Optimal recursive MMSE estimate is then given by**

$$\hat{\mathbf{x}}_k = (\text{prediction of } \mathbf{x}_k) + k_k[(\text{observation of } \mathbf{y}_k) - (\text{prediction of } \mathbf{y}_k)]$$

- **Different implementations :**

Basics of Kalman Filtering, Click here

  - *Extended Kalman Filter* (EKF) : optimal quantities approximated via first order Taylor series expansion (linearization) of process and measurement models.
  - *Unscented Kalman Filter* (UKF) : optimal quantities calculated using the Unscented Transformation (accurate to second order for any nonlinearity). Drastic improvement over EKF [Wan, van der Merwe, Nelson 2000].

for a link on use of UNSCENTED KALMAN FILTER for speech processing. CLICK here.

If the link does not work, go to www.cavs.msstate.edu/~patil/particle_filter/Ukf_final_speech.pdf

- **Problem : Gaussian approximation breaks down for most nonlinear real-world applications (multi-modal distributions, non-Gaussian noise sources, etc.)**

Perfect Monte Carlo Simulation

- **Allow for a complete representation of the posterior distribution.**
- **Map intractable integrals of optimal Bayesian solution to tractable discrete sums of weighted samples drawn from the posterior distribution.**

independent, identically distributed random variable

$$\hat{p}(\mathbf{x}_k \mid \mathbf{Y}_k) = \frac{1}{N}\sum_{i=1}^{N} d\left(\mathbf{x}_k - \mathbf{x}_k^{(i)}\right)$$

$$\mathbf{x}_k^{(i)} \xleftarrow{I.I.D.} p(\mathbf{x}_k \mid \mathbf{Y}_k)$$

- **So, any estimate of the form**

$$E\left[f(\mathbf{x}_k)\right] = \int f(\mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{Y}_k)d\mathbf{x}_k$$

may be approximated by :

$$E\left[f(\mathbf{x}_k)\right] \approx \frac{1}{N}\sum_{i=1}^{N} f\left(\mathbf{x}_k^{(i)}\right)$$

Particle Filters

- **Bayesian Importance Sampling**
  - It is often impossible to sample directly from the true posterior density.
  - However, we can rather sample from a known, easy-to-sample, *proposal* distribution,

$$q(\mathbf{x}_k \mid \mathbf{Y}_k)$$

Is similar to importance density - q (Xk | Zk)

and make use of the following substitution

$$E[f(\mathbf{x}_k)] = \int f(\mathbf{x}_k) \frac{p(\mathbf{x}_k|\mathbf{Y}_k)}{q(\mathbf{x}_k|\mathbf{Y}_k)} q(\mathbf{x}_k \mid \mathbf{Y}_k)\, d\mathbf{x}_k$$

$$= \int f(\mathbf{x}_k) \frac{p(\mathbf{Y}_k|\mathbf{x}_k)\, p(\mathbf{x}_k)}{p(\mathbf{Y}_k)\, q(\mathbf{x}_k|\mathbf{Y}_k)} q(\mathbf{x}_k \mid \mathbf{Y}_k)\, d\mathbf{x}_k$$

$$= \int f(\mathbf{x}_k) \frac{w_k(\mathbf{x}_k)}{p(\mathbf{Y}_k)} q(\mathbf{x}_k \mid \mathbf{Y}_k)\, d\mathbf{x}_k$$

$$w_k(\mathbf{x}_k) = \frac{p(\mathbf{Y}_k|\mathbf{x}_k)\, p(\mathbf{x}_k)}{q(\mathbf{x}_k|\mathbf{Y}_k)}$$

Particle Filters

$$E\left[f(\mathbf{x}_k)\right] = \frac{1}{p(\mathbf{Y}_k)} \int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q\left(\mathbf{x}_k \mid \mathbf{Y}_k\right) d\mathbf{x}_k$$

$$= \frac{\int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q\left(\mathbf{x}_k \mid \mathbf{Y}_k\right) d\mathbf{x}_k}{\int p(\mathbf{Y}_k \mid \mathbf{x}_k) p(\mathbf{x}_k) \frac{q(\mathbf{x}_k \mid \mathbf{Y}_k)}{q(\mathbf{x}_k \mid \mathbf{Y}_k)} d\mathbf{x}_k}$$

$$= \frac{\int f(\mathbf{x}_k) w_k(\mathbf{x}_k) q\left(\mathbf{x}_k \mid \mathbf{Y}_k\right) d\mathbf{x}_k}{\int w_k(\mathbf{x}_k) q\left(\mathbf{x}_k \mid \mathbf{Y}_k\right) d\mathbf{x}_k}$$

$$= \frac{E_{q(\mathbf{x}_k \mid \mathbf{Y}_k)}\left[w_k(\mathbf{x}_k) f(\mathbf{x}_k)\right]}{E_{q(\mathbf{x}_k \mid \mathbf{Y}_k)}\left[w_k(\mathbf{x}_k)\right]}$$

Particle Filters

- So, by drawing samples from $q(\mathbf{x}_k \mid \mathbf{Y}_k)$, we can approximate expectations of interest by the following:

$$E\left[f(\mathbf{x}_k)\right] \approx \frac{\frac{1}{N}\sum_{i=1}^{N} w_k(\mathbf{x}_k^{(i)}) f(\mathbf{x}_k^{(i)})}{\frac{1}{N}\sum_{i=1}^{N} w_k(\mathbf{x}_k^{(i)})}$$

$$\approx \sum_{i=1}^{N} \tilde{w}_k(\mathbf{x}_k^{(i)}) f(\mathbf{x}_k^{(i)})$$

- Where the normalized importance weights are given by

$$\tilde{w}_k(\mathbf{x}_k^{(i)}) = \frac{w_k(\mathbf{x}_k^{(i)})}{\sum_{j=1}^{N} w_k(\mathbf{x}_k^{(j)})}$$

Particle Filters

- Using the state space assumptions (1st order Markov / observational independence given state), the importance weights can be estimated recursively by [proof in De Freitas (2000)]

$$w_k = w_{k-1} \frac{p(\mathbf{y}_k \mid \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{k-1})}{q(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k)}$$

- Problem with SIS is that the variance of the importance weights increase stochastically over time [Kong et al. (1994), Doucet et al. (1999)]

- To solve this, we need to resample the particles
  - keep / multiply particles with high importance weights
  - discard particles with low importance weights

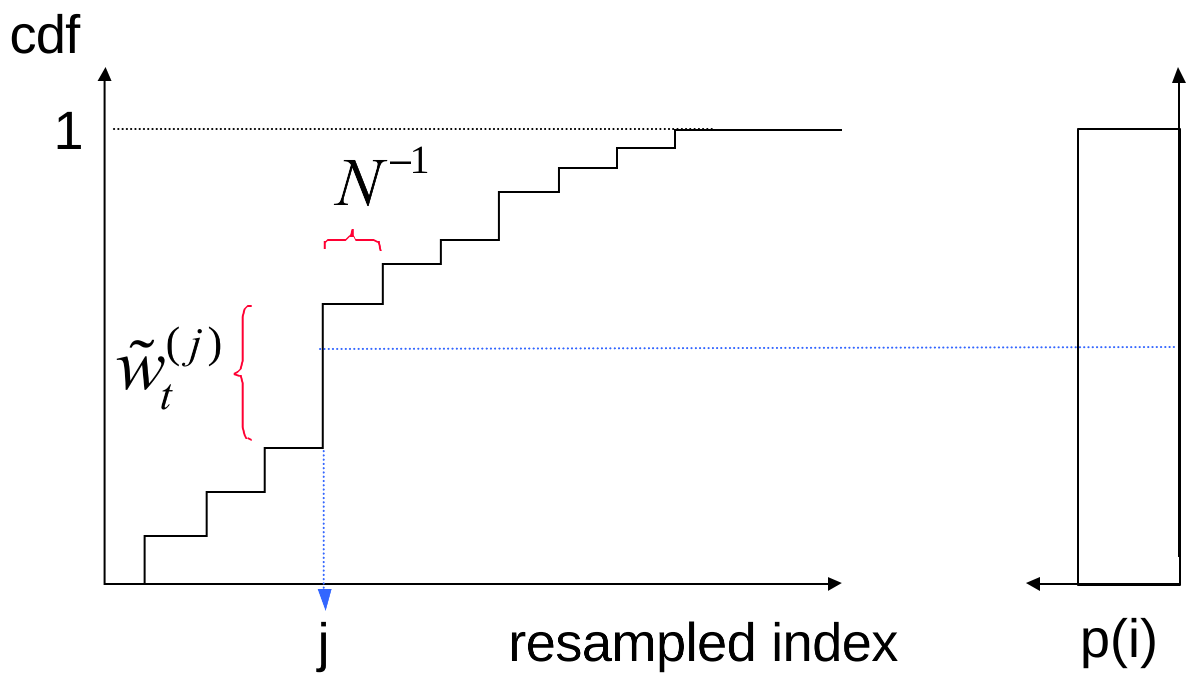
- *Sampling-importance Resampling* (SIR)

Particle Filters

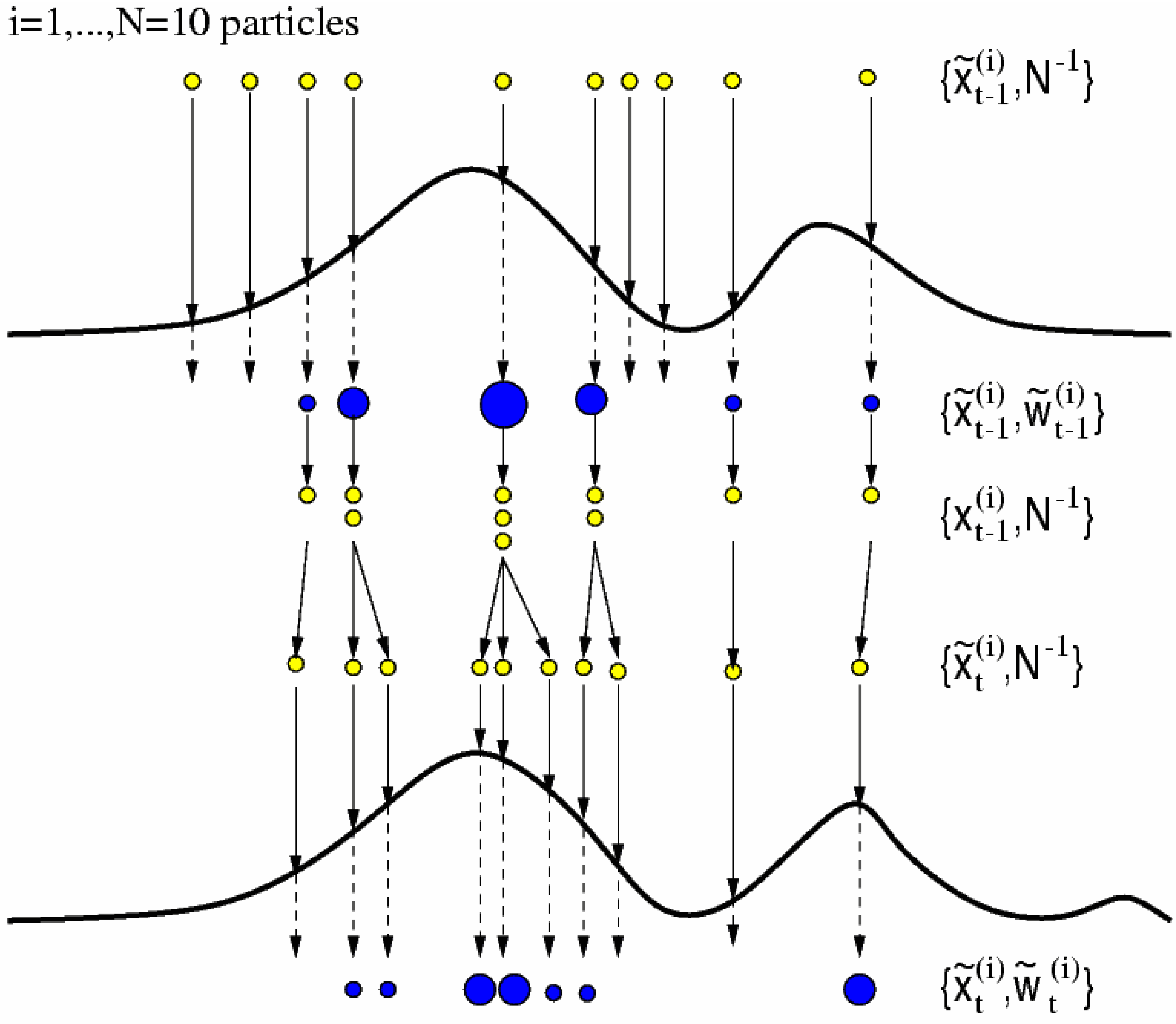
- **Sampling-importance Resampling**
  - Maps the *N* unequally weighted particles into a new set of *N* equally weighted samples.

$$\left\{\mathbf{x}_k^{(i)}, \tilde{w}_k^{(i)}\right\} \rightarrow \left\{\mathbf{x}_k^{(j)}, N^{-1}\right\}$$

  - Method proposed by Gordon, Salmond & Smith (1993) and proven mathematically by Gordon (1994).



Particle Filters



Particle Filters

- **Choice of Proposal Distribution**

$$w_k = w_{k-1} \frac{p(\mathbf{y}_k \mid \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{k-1})}{q(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k)}$$

Absolutely Critical Design Step

critical design issue for successful particle filter

- samples/particles are drawn from this distribution
- used to evaluate importance weights

- Requirements

1) Support of proposal distribution must include support of true posterior distribution, i.e. *heavy-tailed* distributions are preferable.

2) Must include most recent observations.

Particle Filters

- Most popular choice of proposal distribution does not satisfy these requirements though:

$$q\left(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k\right) = p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}\right)$$

[Isard and Blake 96, Kitagawa 96, Gordon et al. 93, Beadle and Djuric 97, Avitzour 95]

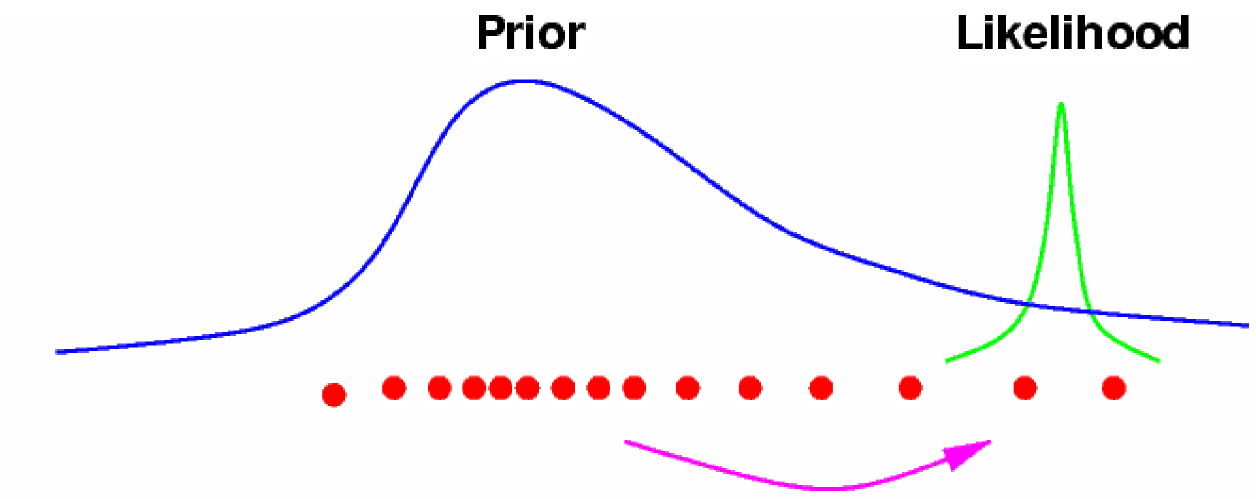
- Easy to implement :

$$w_k = w_{k-1} \frac{p\left(\mathbf{y}_k \mid \mathbf{x}_k\right) p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}\right)}{p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}\right)}$$
$$= w_{k-1} p\left(\mathbf{y}_k \mid \mathbf{x}_k\right)$$

- Does not incorporate most recent observation though !

Improving Particle Filters

- **Incorporate New Observations into Proposal**



  - Use Gaussian approximation (i.e. Kalman filter) to generate proposal by combining new observation with prior

$$q\left(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k\right) = p_G\left(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k\right)$$
$$= \mathcal{N}\left(\hat{\mathbf{x}}_k, \text{cov}[\mathbf{x}_k]\right)$$

Improving Particle Filters

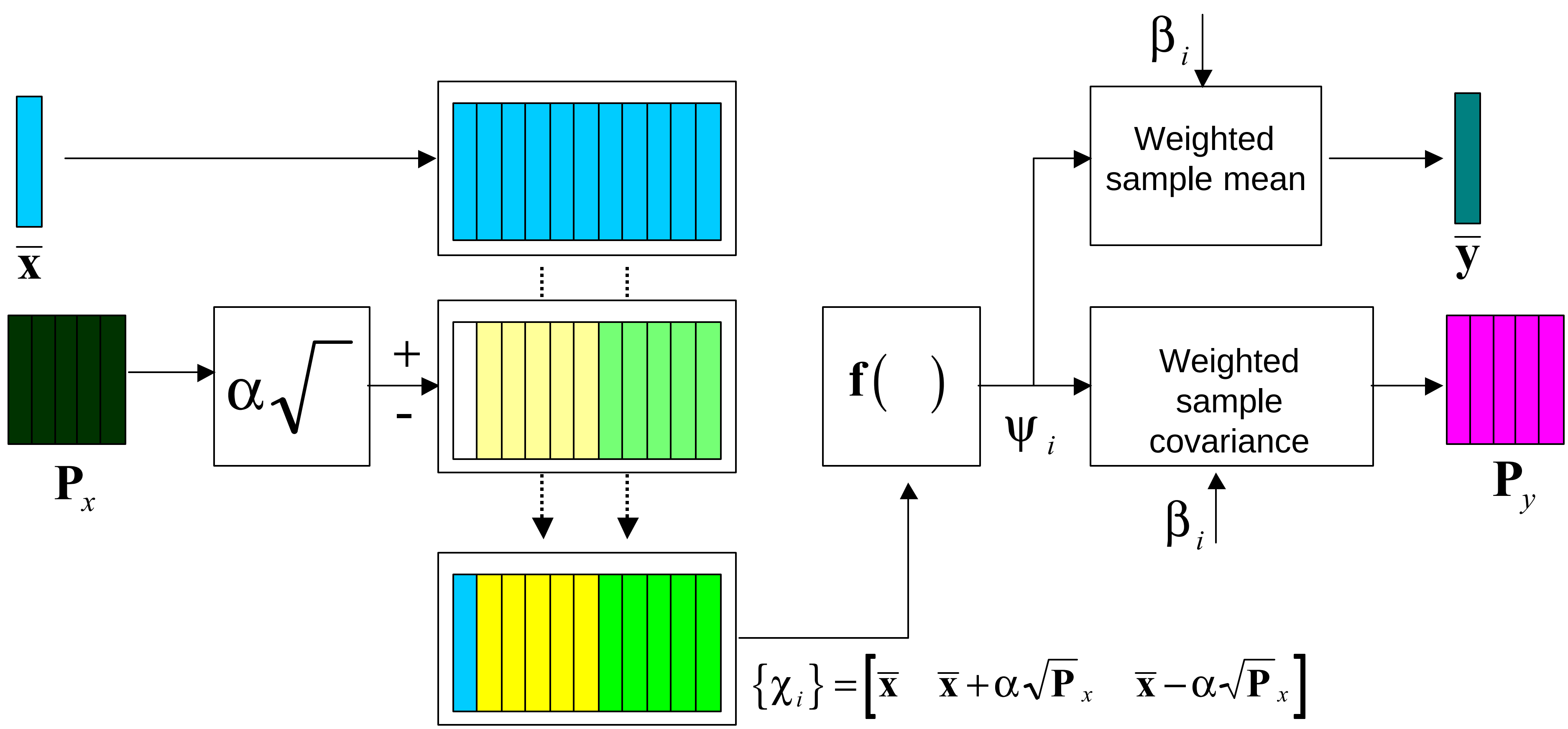
- ***Extented Kalman Filter*** **Proposal Generation**
  - De Freitas (1998), Doucet (1998), Pitt & Shephard (1999).
  - Greatly improved performance compared to standard particle filter in problems with very accurate measurements, i.e. likelihood very peaked in comparison to prior.
  - In highly nonlinear problems, the EKF tends to be very inaccurate and *underestimates* the true covariance of the state. This violates the distribution support requirement for the proposal distribution and can lead to poor performance and filter divergence.

- **We propose the use of the *Unscented Kalman Filter* for proposal generation to address these problems !**

Improving Particle Filters

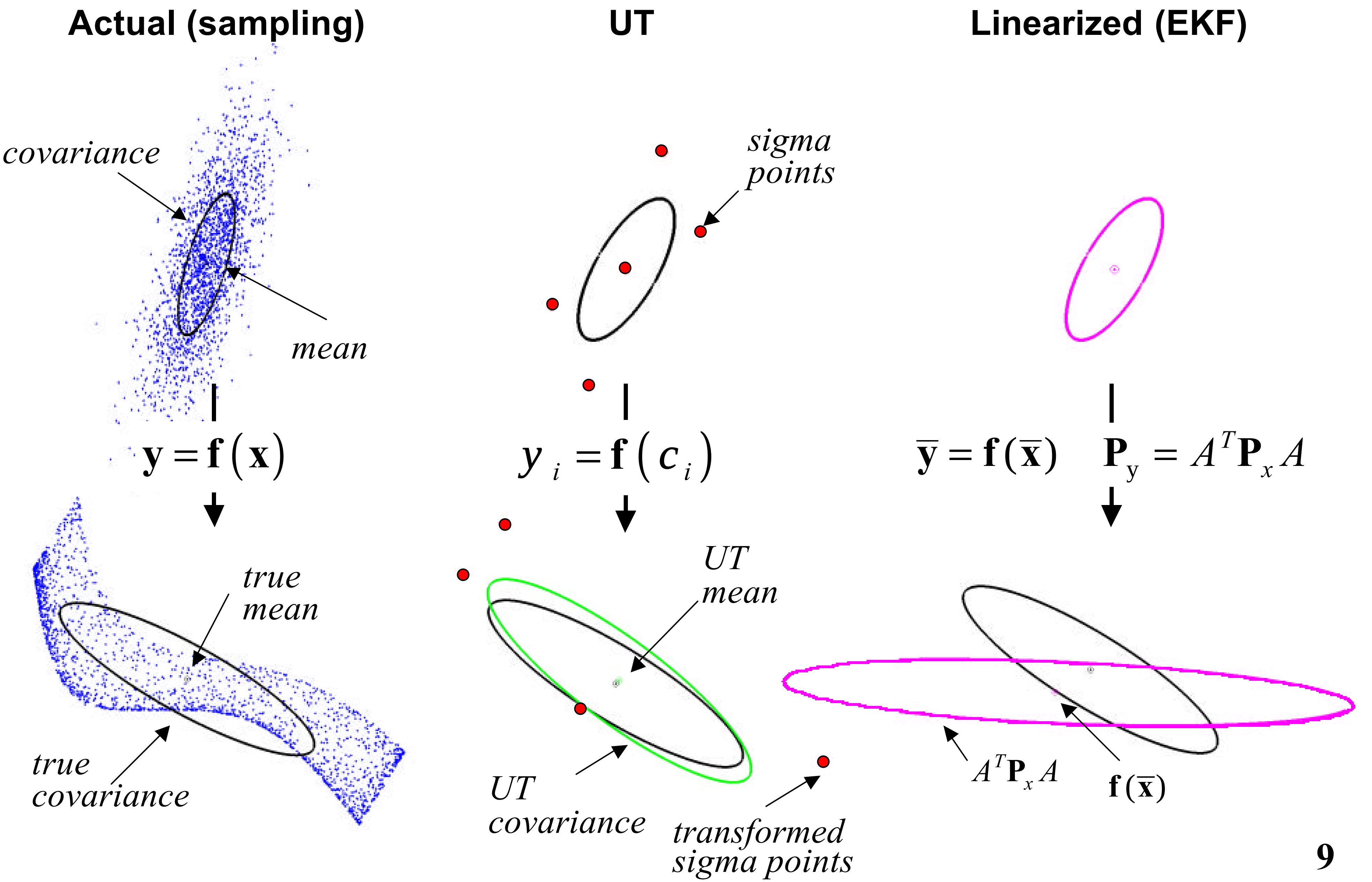
- ***Unscented Kalman Filter*** **Proposal Generation**
  - UKF is a recursive MMSE estimator based on the Unscented Transformation (UT).
  - UT : Method for calculating the statistics of a RV that undergoes a nonlinear transformation (Julier and Uhlmann 1997)
  - UT/UKF : - accurate to 3rd order for Gaussians
    - higher order errors scaled by choice of transform parameters.
  - More accurate estimates than EKF (Wan, van der Merwe, Nelson 2000)
  - Have some control over higher order moments, i.e. kurtosis, etc. → heavy tailed distributions !

Unscented Transformation

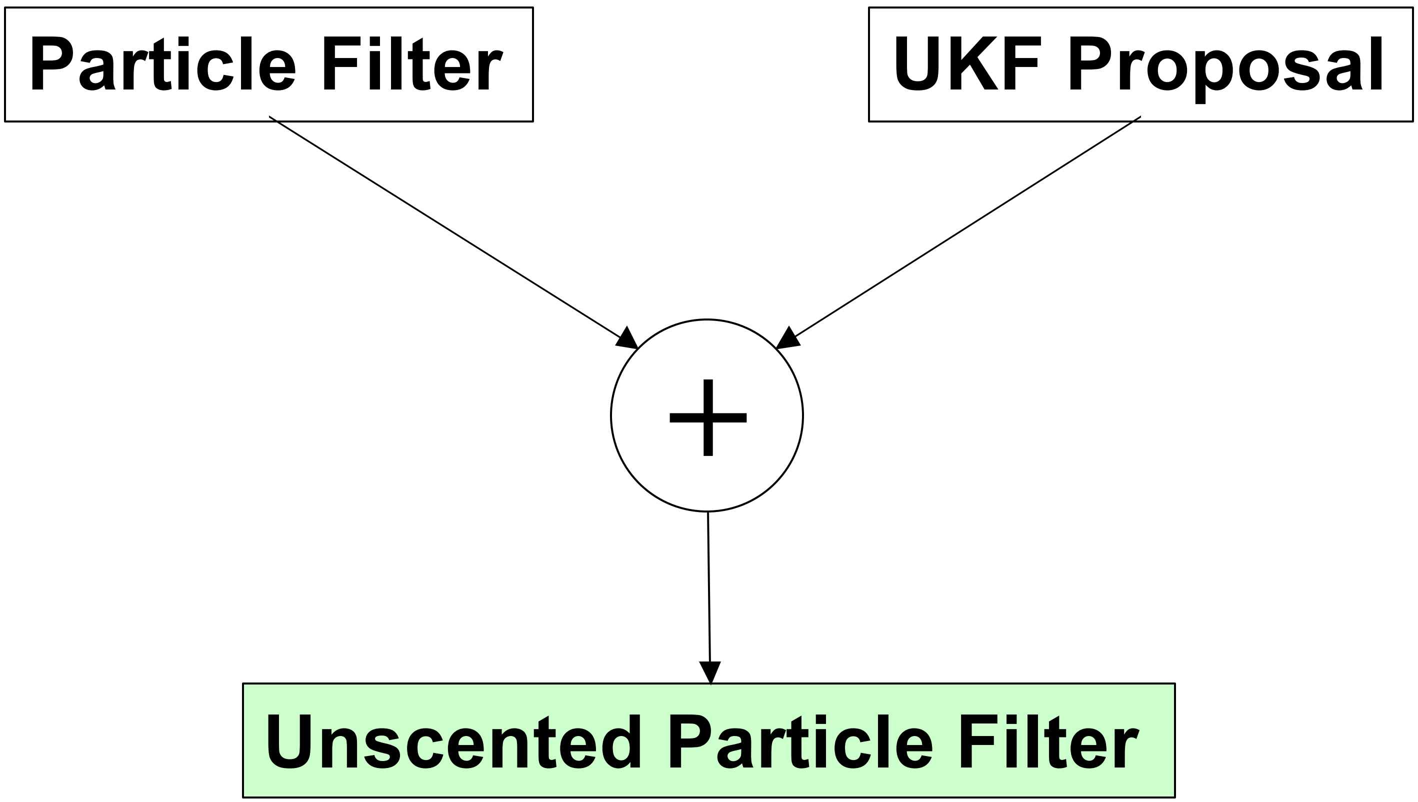


$$\{\chi_i\} = \left[\bar{\mathbf{x}} \quad \bar{\mathbf{x}} + \alpha\sqrt{\mathbf{P}_x} \quad \bar{\mathbf{x}} - \alpha\sqrt{\mathbf{P}_x}\right]$$

The Unscented Transformation



Unscented Particle Filter



Experimental Results

- **Synthetic Experiment**
  - Time-series
    - process model :

$$x_{k+1} = 1 + \sin(\omega \pi k) + \phi x_k + v_k$$

process noise (Gamma)

    - nonstationary observation model :

$$y_k = \begin{cases} \phi x_k^2 + n_k & k \le 30 \\ \phi x_k - 2 + n_k & k > 30 \end{cases}$$

measurement noise (Gaussian)

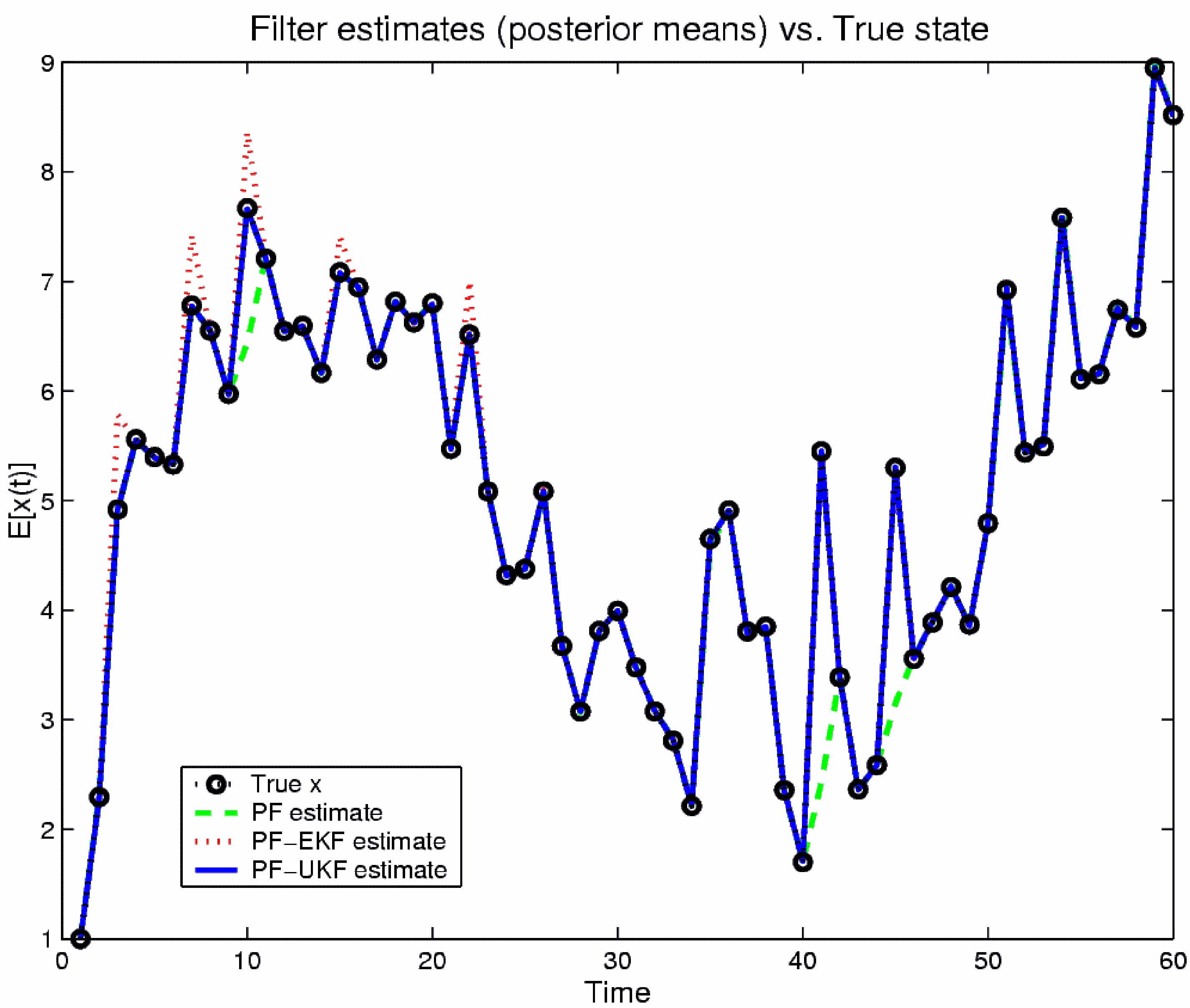
Experimental Results

- **Synthetic Experiment : (100 independent runs)**

Filter	MSE	
	mean	variance
Extended Kalman Filter (EKF)	0.374	0.015
Unscented Kalman Filter (UKF)	0.280	0.012
Particle Filter : generic	0.424	0.053
Particle Filter : EKF proposal	0.310	0.016
Unscented Particle Filter	0.070	0.006

Experimental Results

- **Synthetic Experiment**



Experimental Results

- **Pricing Financial Options**
  - Options : financial derivative that gives the holder the right (but not obligation) to do something in the future.
    - Call option : - allow holder to *buy* an underlying cash product
      - at a *specified future date* ("maturity time")
      - for a *predetermined price* ("strike price")
    - Put option : - allow holder to *sell* an underlying cash product
  - *Black Scholes* partial differential equation
    - Main industry standard for pricing options

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \tfrac{1}{2} s^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf$$

option value (f)

risk-free interest rate (r)

value of underlying cash product (S)

volatility of cash product (s)

Experimental Results

- **Pricing Financial Options**
  - Black & Scholes (1973) derived the following pricing solution:

$$C = S\mathcal{N}_c(d_1) - Xe^{-rt_m}\mathcal{N}_c(d_2)$$

$$P = -S\mathcal{N}_c(-d1) + Xe^{-rt_m}\mathcal{N}_c(-d_2)$$

$$d_1 = \frac{\ln(S/X) + (r + \frac{1}{2}s^2)t_m}{s\sqrt{t_m}}$$

$$d_2 = d_1 - s\sqrt{t_m}$$

$$\mathcal{N}_c(.) = \text{cumulative normal distribution}$$

Experimental Results

- **Pricing Financial Options**
  - State-space representation to model system for particle filters
    - Hidden states : r , s
    - Output observations: C , P
    - Known control signals: t_m , S
  - Estimate call and put prices over a 204 day period on the FTSE-100 index.
    - Performance : normalized square error for one-step-ahead predictions

$$NSE = \sqrt{\sum_k \left(\mathbf{y}_k - \hat{\mathbf{y}}_k\right)}$$

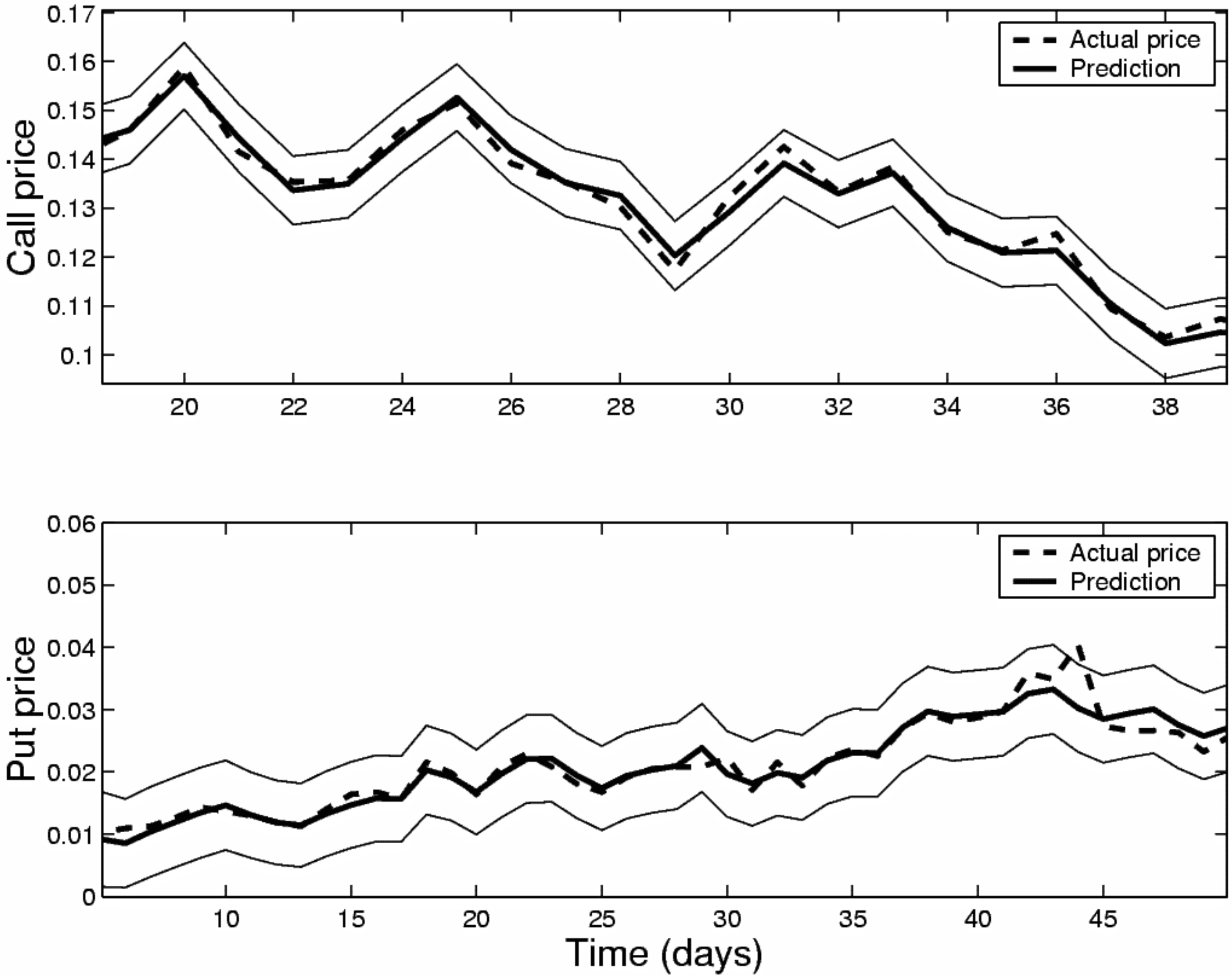
Experimental Results

- **Options Pricing Experiment : (100 independent runs)**

Option Type	Algorithm	NSE	
		mean	var
Call	Trivial	0.078	0.000
	Extended Kalman Filter (EKF)	0.037	0.000
	Unscented Kalman Filter (UKF)	0.037	0.000
	Particle Filter : generic	0.037	0.000
	Particle Filter : EKF proposal	0.092	0.508
	Unscented Particle Filter	*0.009*	*0.000*
Put	Trivial	0.035	0.000
	Extended Kalman Filter (EKF)	0.023	0.000
	Unscented Kalman Filter (UKF)	0.023	0.000
	Particle Filter : generic	0.023	0.000
	Particle Filter : EKF proposal	0.024	0.007
	Unscented Particle Filter	*0.008*	*0.000*

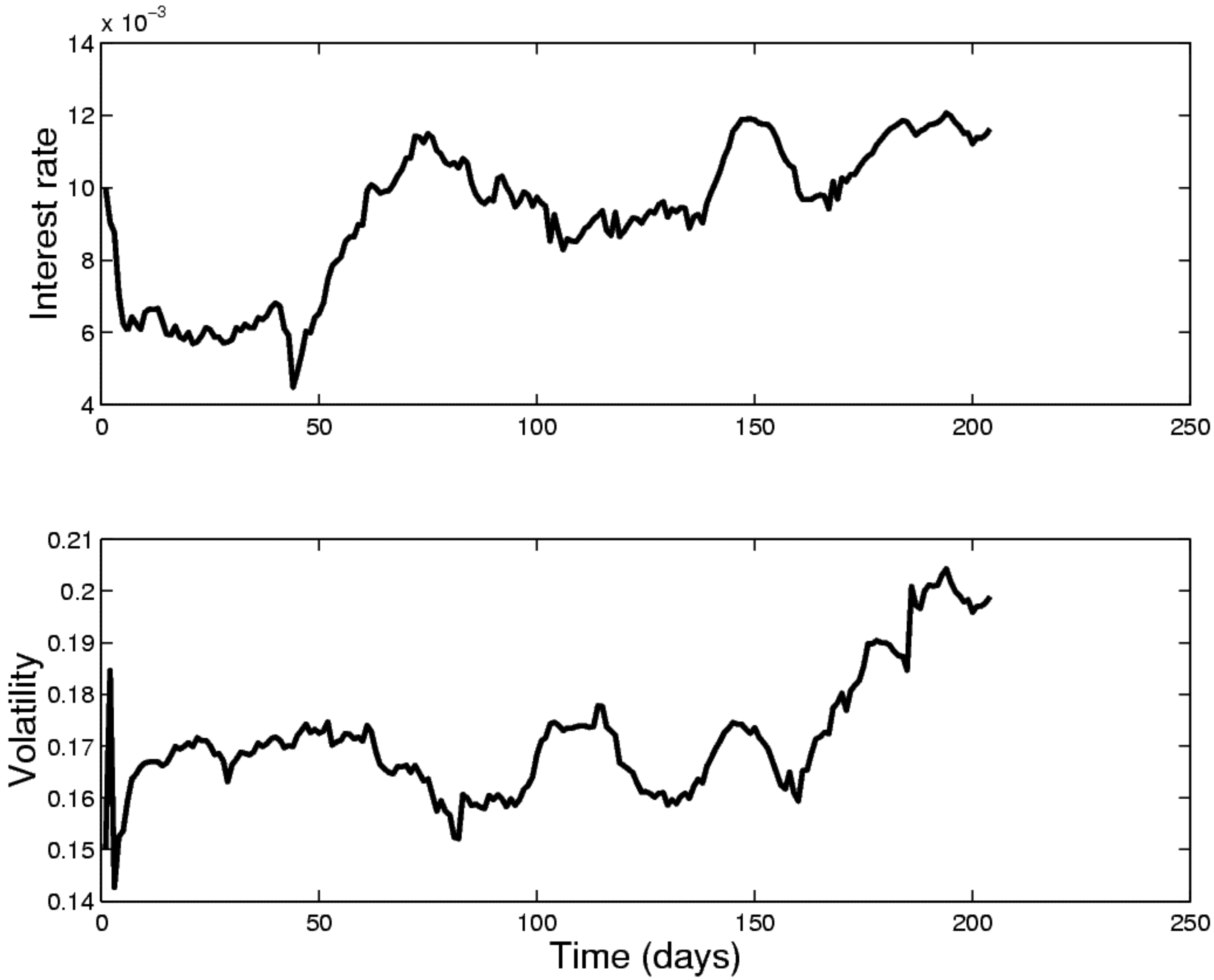
Experimental Results

- **Options Pricing Experiment : UPF one-step-ahead predictions**



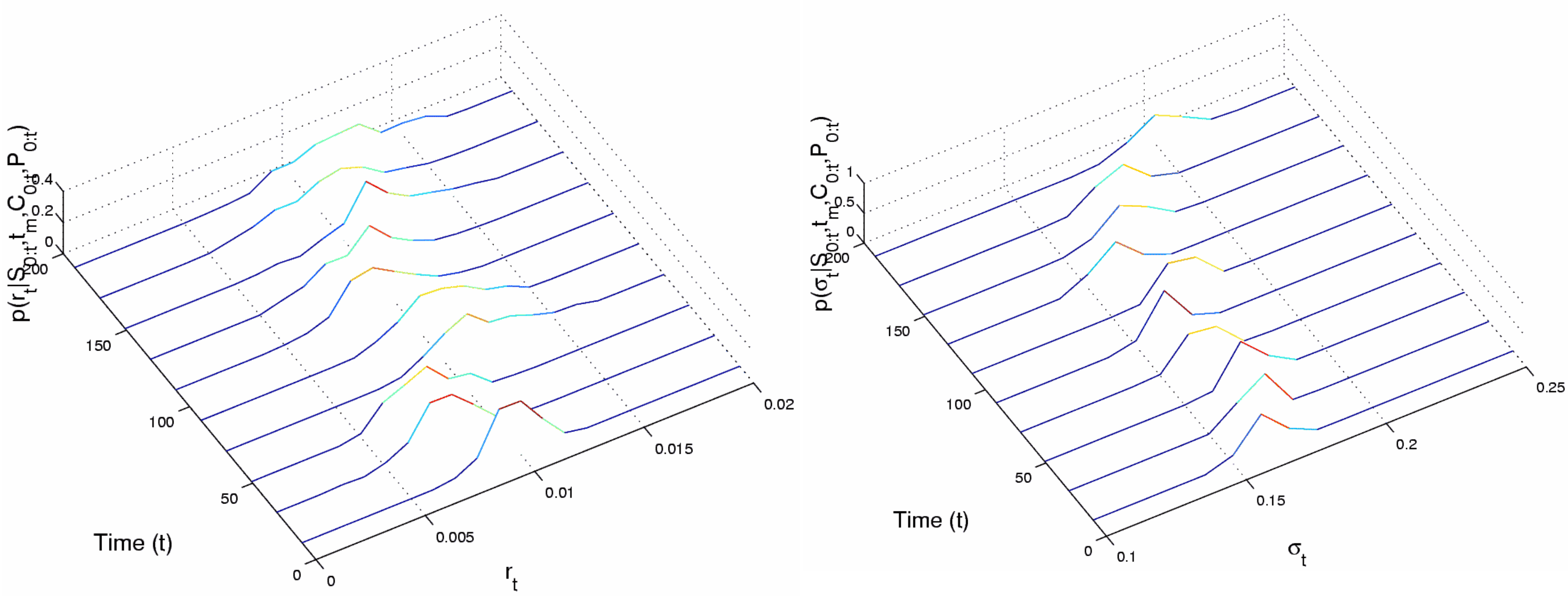
Experimental Results

- **Options Pricing Experiment : Estimated interest rate and volatility**



Experimental Results

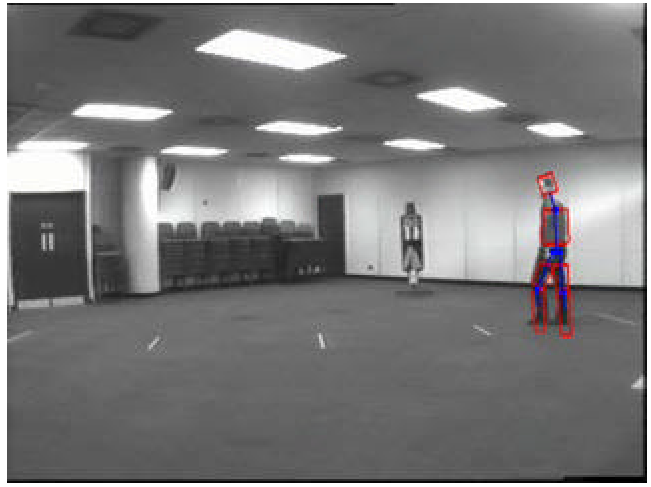
- **Options Pricing Experiment : Probability distributions of implied interest rate and volatility**



Particle Filter Demos

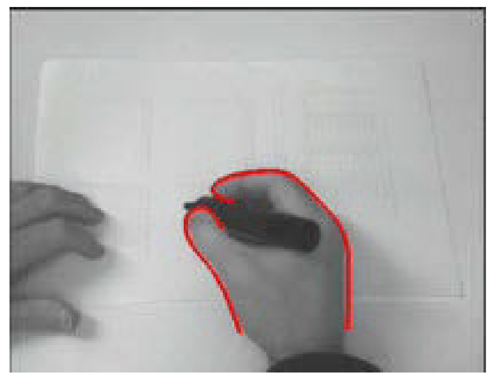
- ***Visual Dynamics Group*, Oxford. (Andrew Blake)**

Tracking agile motion



Tracking motion against camouflage

Mixed state tracking



Conclusions

- **Particle filters allow for a practical but complete representation of posterior probability distribution.**
- **Applicable to general nonlinear, non-Gaussian estimation problems where standard Gaussian approximations fail.**
- **Particle filters rely on importance sampling, so the proper choice of proposal distribution is very important:**
  - Standard approach (i.e. transition prior proposal) fails when likelihood of new data is very peaked (accurate sensors) or for heavy-tailed noise sources.
  - EKF proposal : Incorporates new observations, but can diverge due to inaccurate and inconsistent state estimates.
  - *Unscented Particle Filter* : UKF proposal
    - More consistent and accurate state estimates.
    - Larger support overlap, can have heavier tailed distributions.
    - Theory predicts and experiments prove ***significantly better*** performance.

The End