

# **First Year Report: Fixed Lag Particle Filtering for Target Tracking**

Pete Bunch

July 2011



**UNIVERSITY OF  
CAMBRIDGE**

---

---

# *Contents*

---

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Particle Methods</b>	<b>7</b>
2.1	Bayes of our lives - what is filtering? . . . . .	7
2.2	Keep Kalman carry on - the Kalman filter and its extensions . . . . .	8
2.3	Tough as old bootstraps - the traditional approach . . . . .	9
2.3.1	Auxiliary sampling . . . . .	10
2.3.2	Degeneracy and resampling . . . . .	11
2.3.3	Importance distributions . . . . .	11
2.4	MCMC in da house - series vs. parallel . . . . .	11
<b>3</b>	<b>Approaches to Tracking</b>	<b>13</b>
<b>4</b>	<b>Fixed Lag Estimation for Tracking</b>	<b>15</b>
4.1	Why do now what we can do later? . . . . .	15
4.2	A mathematical framework for fixed lag estimation . . . . .	16
4.2.1	Artificial conditionals . . . . .	17
4.2.2	History Proposals . . . . .	17
4.3	Fixed lag estimation for tracking . . . . .	17
4.3.1	Importance distributions . . . . .	17
<b>5</b>	<b>Fixed Lag Particle Filters</b>	<b>19</b>
5.1	Sequential Importance Sampling and Resampling . . . . .	19
5.1.1	Coping with dimensionality . . . . .	19
5.2	Markov Chain Monte Carlo . . . . .	19
5.3	Marginalised Particle Filters . . . . .	19
<b>6</b>	<b>Plan Of Future Research</b>	<b>21</b>



## *Introduction*

---

Everyone loves to track things. Yay!



# Particle Methods

## 2.1 Bayes of our lives - what is filtering?

Many tasks in signal processing, science in general, and indeed life, require us to make some estimate of an unknown quantity from indirect, incomplete, or inaccurate observations. By constructing a model to explain how these observations depend on the underlying state, we can infer something about that state. We will express this observation model in terms of a likelihood function:

$$P(Y|X) \tag{2.1}$$

where  $X$  is the state and  $Y$  the observations. This is not the whole story - in many cases we are not estimating our unknown state “from scratch”. Previous experience, prejudice, and prior knowledge can also contribute to our estimates. The likelihood and prior terms can be combined through our friend, Bayes rule (?), to calculate the posterior probability of the state, i.e. the probability of the state given the observations:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \tag{2.2}$$

We may not be able to pinpoint an exact value with certainty, but the probability of different candidates may now be compared. Mathematically, we can assign a probability distribution to the state space of  $X$ .

Often the quantity in which we are interested,  $X$ , is changing over time, and we would like to estimate its value at each point in time given only the observations received so far. In this case we will generally assume that the state is Markovian, i.e. that  $X_t|X_{t-1}$  is independent of  $X_{1:t-2}$ . This is the traditional filtering problem. By constructing a model for the evolution of the unknown state, we can now derive our prior information from our estimate at the previous time step. In discrete time, we now write:

$$P(X_t|Y_{1:t}) = \frac{\int P(Y_t|X_t)P(X_t|X_{t-1})P(X_{t-1}|Y_{1:t-1})dX_{t-1}}{P(Y_t|Y_{1:t-1})} \tag{2.3}$$

where the subscript indicates the time and ranges are notated by : in the MATLAB style.

Instead of marginalising the previous state, we may sometimes want to consider the joint state distribution over all time instances. This may similarly be expanded as:

$$P(X_{1:t}|Y_{1:t}) = \frac{P(Y_t|X_t)P(X_t|X_{t-1})P(X_{1:t-1}|Y_{1:t-1})}{P(Y_t|Y_{1:t-1})} \quad (2.4)$$

The filtering distribution may then be obtained by marginalising out the previous states.

So far, we have expressed the problem in terms of distributions, but sometimes we will need to consider recursive random variable equations. In the most general form:

$$X_t = f_t(X_{t-1}, V_t) \quad (2.5)$$

$$Y_t = g_t(X_t, W_t) \quad (2.6)$$

where  $f_t$  and  $g_t$  are known deterministic functions and  $V_t$  and  $W_t$  are random variables, known as the process and observation noise respectively. Again, we assume  $X_t$  is Markovian.

## 2.2 Keep Kalman carry on - the Kalman filter and its extensions

In a few simple cases the filtering set-up permits the derivation of closed form posterior distributions at each time instant. Most notably, the Kalman filter (KF) (?) is an analytic filter for models with continuous state and observation variables, in which both transition and observation models are linear transformations with Gaussian innovations.

$$x_t = Ax_t + v_t \quad (2.7)$$

$$y_t = Cx_t + w_t \quad (2.8)$$

where  $v_t$  and  $w_t$  are now Gaussian random variables with zero mean and covariance matrices  $Q$  and  $R$ . We use lower case variables here to emphasise that these are “nice”, continuous vectors. (As we shall see, our state variable will later be sets or lists).

Kalman’s solution for the linear-Gaussian case is given by:

$$P(x_t|y_{1:t}) = \mathcal{N}(x_t|\mu_t, \Sigma_t) \quad (2.9)$$

$$P(x_t|y_{1:t-1}) = \mathcal{N}(x_t|\hat{\mu}_t, \hat{\Sigma}_t) \quad (2.10)$$

where  $\mu_t$ ,  $\Sigma_t$ , etc. are given by the following recursions.

Time Updates:

$$\hat{\mu}_t = A\mu_{t-1} \quad (2.11)$$

$$\hat{\Sigma}_t = A\Sigma_{t-1}A^T + Q \quad (2.12)$$

Measurement Updates:

$$z_t = y_t - C\hat{\mu}_t \quad (2.13)$$

$$S_t = C\hat{\Sigma}_tC^T + R \quad (2.14)$$

$$K_t = \hat{\Sigma}_tC^TS_t^{-1} \quad (2.15)$$

$$\mu_t = \hat{\mu}_t + K_tz_t \quad (2.16)$$



$$\Sigma_t = (I - K_t C) \hat{\Sigma}_t \quad (2.17)$$

The KF is delightful because it not only provides us with a closed-form analytic solution, but the complexity of that solution does not increase as we receive additional measurements. This is a consequence of the fact that the Gaussian distribution is its own conjugate prior. Unfortunately, no other such convenient cases have been discovered (?). Analytic solutions to non-linear, non-Gaussian filtering problems generally require unacceptable conditions, such as zero process noise  $Q = 0$  (?).

Given the loveliness of the KF, the instinct when faced by an intractable non-linear filtering problem is to linearise it. This produces the Extended Kalman Filter (EKF), and is achieved by replacing the  $A$  and  $C$  matrices in equations 2.11 through 2.17 above with Jacobians:

$$A_t = \left. \frac{\partial f}{\partial x_t} \right|_{\mu_{t-1}} \quad (2.18)$$

$$C_t = \left. \frac{\partial g}{\partial x_t} \right|_{\hat{\mu}_t} \quad (2.19)$$

ADD SOMETHING ABOUT SMOOTHING.

## 2.3 Tough as old bootstraps - the traditional approach

In general we will not be so lucky as to have a problem with linear-Gaussian dynamics. In this case, a particle filter (PF) may be the best alternative. With a PF, we approximate a probability distribution with a set of (weighted) samples drawn from that distribution.

$$P(X) \approx \frac{1}{N} \sum_m W^{(m)} \delta_X(x^{(m)}) \quad (2.20)$$

where  $\delta(x)$  represents a unit probability point mass at a point  $x$ , and  $\sum_m W^{(m)} = 1$ . Such a method allows us to represent any probability distribution of arbitrary complexity, including multidimensional, multimodal, mixed distributions. As the number of particles increases, the accuracy of the approximation improves at the expense of computational complexity. Thus we have the required tool for estimation in non-linear, non-Gaussian scenarios.

We still face the problem of how to generate these samples. The conventional method for this is Sequential Importance Sampling with Resampling (SISR). For a more complete and traditional introduction to this method, see (?) or (?). Here we follow an outline similar to that used for the derivation of the auxiliary particle filter of (?).

Suppose we have a particle approximation to the joint posterior distribution from the previous frame,  $\hat{P}(X_{1:t-1}|Y_{1:t-1})$ . Each particle represents a path through time,  $X_{1:t-1}^{(m)}$ , and has an associated weight,  $W_{t-1}^{(m)}$ . Let us imagine also that the unweighted particles are samples from another distribution:

$$\mu(X_{1:t-1}|Y_{1:t-1}) \approx \frac{1}{N} \sum_m W^{(m)} \delta_X(x_{1:t}^{(m)}) \quad (2.21)$$

Thus for a given particle,

$$\hat{P}(X_{1:t-1}^{(m)}|Y_{1:t-1}) = W_t^{(m)}\mu(X_{1:t-1}^{(m)}|Y_{1:t-1}) \quad (2.22)$$

We would like to generate a new particle set which includes the current time instance. We propose a new set of extended tracks from a factored proposal distribution  $X_{1:t} \sim q(X_{1:t}|Y_{1:t}) = q(X_t|Y_t, X_{1:t-1})q(X_{1:t-1}|Y_{1:t})$ . The two factors are the proposal probabilities for the new state value,  $X_t$  and the history,  $X_{1:t-1}$ , respectively. Particles are then weighted to take account of the difference between the targeted posterior distribution and the importance distribution:

$$W_t^{(m)} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{q(X_{1:t}^{(m)}|Y_{1:t})} \quad (2.23)$$

The simplest choice for the history proposal is  $q(X_{1:t-1}|Y_{1:t}) = \mu(X_{1:t-1}|Y_{1:t-1})$ , which can be implemented by simply keeping the same set of paths as the previous particle set. This is equivalent to an ordinary IS step with no resampling, and importance weights are given by:

$$W_t^{(m)} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{q(X_{1:t}^{(m)}|Y_{1:t})} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{\mu(X_{1:t-1}^{(m)}|Y_{1:t-1})q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \approx \frac{W_{t-1}^{(m)}P(Y_t|X_t^{(m)})P(X_t^{(m)}|X_{t-1}^{(m)})}{q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \quad (2.24)$$

Alternatively, we could use  $q(X_{1:t-1}|Y_{1:t}) = \hat{P}(X_{1:t-1}|Y_{1:t-1})$  as the history proposal, i.e. sample from the weighted particle distribution which approximates the previous posterior. This is equivalent to an IS step preceeded by resampling. Importance weights are now given by:

$$W_t^{(m)} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{q(X_{1:t}^{(m)}|Y_{1:t})} \approx \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{P(X_{1:t-1}^{(m)}|Y_{1:t-1})q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \approx \frac{P(Y_t|X_t^{(m)})P(X_t^{(m)}|X_{t-1}^{(m)})}{q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \quad (2.25)$$

### 2.3.1 Auxiliary sampling

We can generalise the form of our history proposal distribution by weighting the particles from the previous posterior distribution with any arbitrary set of weights.

$$q(X_{1:t-1}|Y_{1:t}) = \frac{1}{N} \sum_m V_t^{(m)} \delta_X(x_{1:t}^{(m)}) \quad (2.26)$$

Now we have

$$\hat{P}(X_{1:t-1}^{(m)}|Y_{1:t-1}) = \frac{W_{t-1}^{(m)}}{V_t^{(m)}} q(X_{1:t-1}^{(m)}|Y_{1:t-1}) \quad (2.27)$$

giving a general form for the importance weights

$$W_t^{(m)} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{q(X_{1:t}^{(m)}|Y_{1:t})} = \frac{P(X_{1:t}^{(m)}|Y_{1:t})}{\mu(X_{1:t-1}^{(m)}|Y_{1:t-1})q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \approx \frac{W_{t-1}^{(m)}}{V_t^{(m)}} \times \frac{P(Y_t|X_t^{(m)})P(X_t^{(m)}|X_{t-1}^{(m)})}{q(X_t^{(m)}|X_{t-1}^{(m)}, Y_t)} \quad (2.28)$$

### 2.3.2 Degeneracy and resampling

REWRITE THIS ENTIRELY

Which of the two choices of history proposal should we use? The first,  $\mu(X_{1:t-1}|Y_{1:t-1})$ , is the simplest to implement, because we can simply keep the same set of paths for  $1 : t - 1$ . However, the recursive form of the importance weights means that the variance of these weights will increase over time. The result is the well-documented particle degeneracy effect, whereby all the weight coalesces in one or a few particles, and the weights of the rest tend to zero. Intuitively, this is a poor representation of the distribution - we may as well not have the zero-weight particles! The solution is to use resampling, or in the formulation posed above to use the second form of history proposal,  $\hat{P}(X_{1:t-1}|Y_{1:t-1})$ . This biases the sampling of histories towards those with high weights, encouraging those with low weights to be discarded. Because the importance weights for this type of proposal are not calculated recursively, the variance is reduced.

Particle degeneracy can be quantified ESS

Resampling strategies - systematic, multinomial, etc.

### 2.3.3 Importance distributions

It remains to choose the the importance distribution for the current state,  $q(X_t|X_{t-1}, Y_t)$ . In the original “bootstrap filter” of (?), this was set equal to the transition density,  $P(X_t|X_{t-1})$ , which leads to cancellation in the expression for importance weights. This is simple but not necessarily optimal. For example if the process noise is high, the samples of  $X_t$  will be widely spread. If, however, the observation noise is comparatively low, many or most of the samples will be far from the observation and will have a low weight. This is undesirable, as discussed in section ??.

An improved choice of - Optimal importance dist

## 2.4 MCMC in da house - series vs. parallel

S-MCMC (Godsill) approach



## *Approaches to Tracking*

---

A brief overview JPDAF MHT Brief mention of PFs



# *Fixed Lag Estimation for Tracking*

---

## 4.1 Why do now what we can do later?

In chapter 2, we looked at filtering methods, that is, estimation of a hidden state  $X_t$  given observations up until this time  $Y_{1:t}$ . For such a task, Kalman filters and particle filters are our tools of choice. However, for difficult target tracking problems, there may simply not be enough information available to make a good estimate. Consider a target with a low detection probability and high clutter. There is likely to be multiple observations with which the target could be associated in each frame, or the target could disappear entirely for a period. The result is that there is a very large number of possible routes through the observation space. A conventional particle filter will have to maintain particles on each of the possible routes until such time as they become negligibly unlikely. The correct route may at first only have a few particles on it, and when it is identified as the correct route in later frames the particle diversity may be poor. In the worst case, the correct route may at first look so unlikely (e.g. consecutive missed detections), that no particles follow it, and the track is lost.

This problem can be addressed by allowing not only a proposal of a new state  $X_t$  on the end of each path, but also the re-proposal of the preceding states as well. Thus, particles which have followed an incorrect route may be “redirected”, to follow a more probable course. This helps maintain particle diversity. If few or no particles follow the correct route at first, particles may be diverted later once the course becomes more apparent.

In an ordinary particle filter, we estimate the joint posterior,  $P(X_{1:t}|Y_{1:t})$  at each time instant. Thus, our estimates of the state distribution at previous times are updated at each step, but only by resampling. Diversity in the marginal distributions for previous states only decreases as time progresses. Conversely, in using such a system of re-proposals we allow diversity to be maintained in previous states. To limit the computational complexity, we constrain the region in which re-proposals may occur to a lagging window of  $L$  time steps.

## 4.2 A mathematical framework for fixed lag estimation

We now consider a mathematical framework for fixed lag estimation. This method was devised in (?) and (?).

As before, the target posterior distribution in which we are interested is the familiar  $P(X_{1:t}|Y_{1:t})$ . However, the proposal mechanism now becomes more complex, because we will be replacing states in an existing particle. We first propose a particle from which to take the state “history”, that is  $X_{1:t-L}$ . This may be chosen from the particle approximation from any of the previous  $L$  processing steps. However, we get more of the path than we need, because each particle at lag  $s$  is a set of states  $X_{1:t-s}$ . The final  $L-s$  states will be replaced when by a new “tip”,  $X'_{t-L+1:t}$ , drawn from an importance distribution. The complete proposal is thus

$$\{X_{1:t-L}, X'_{t-L+1:t}\} \sim q(s) \int q(X_{1:t-s}|Y_{1:t-s})q(X'_{t-L+1}|X_{1:t-s}, Y_{t-L+1:t})dX_{t-L+1:t-s} \quad (4.1)$$

where  $q(X_{1:t-s}|Y_{1:t-s})$  is a proposal distribution using the arbitrarily-weighted particles from  $\hat{P}(X_{1:t-s}|Y_{1:t-s})$ . We cannot evaluate this. In general, the integral will be intractable. If we restrict our proposals to depend only on the history, i.e. use  $q(X_{t-L+1}|X_{1:t-L}, Y_{t-L+1:t})$ , then the proposal becomes

$$\{X_{1:t-L}, X'_{t-L+1:t}\} \sim q(s)\hat{P}(X_{1:t-L}|Y_{1:t-s})q(X'_{t-L+1}|X_{1:t-s}, Y_{t-L+1:t}) \quad (4.2)$$

However, with this proposal, in order to evaluate importance weights or acceptance probabilities we still need to calculate

$$P(X_{1:t-L}|Y_{1:t-s}) = \frac{P(Y_{t-L+1:t-s}|X_{t-L}^{(m)})P(X_{1:t-L}|Y_{1:t-L})}{P(Y_{t-L+1:t-s}|Y_{1:t-L})} \quad (4.3)$$

The problematic term is:

$$P(Y_{t-L+1:t}|X_{t-L}^{(m)}) = \int P(Y_{t-L+1:t}|X_{t-L}^{(m)}, X_{t-L+1:t})P(X_{t-L+1:t}|X_{t-L}^{(m)})dX_{t-L+1:t-s} \quad (4.4)$$

In general, this will be intractable. The exception is the case where  $s = L$ , but this will in general not work well - such a requirement implies ignoring all the estimation we have done since the start of the window!

The solution to this tractability problem proposed in (?) is to augment the dimension of the target distribution to include the discarded tracks. The new target is:

$$P(X_{1:t-L}, X'_{t-L+1:t}|Y_{1:t})\rho(X_{t-L+1:t-s}|X_{1:t-L}, X'_{t-L+1:t}) \quad (4.5)$$

Once a particle approximation has been generated for this target the required posterior distribution may be obtained by marginalisation. As we are simply going to discard the old track sections  $X_{t-L+1:t-s}$ , the choice of  $\rho(\cdot)$  does not alter the distribution of the posterior. However, the variance of weights/acceptance probabilities may be affected.

With an expanded space for the target distribution there is no need to marginalise the previous states. The proposal distribution now becomes:

$$\{X_{1:t-s}, X'_{t-L+1:t}\} \sim q(s)q(X_{1:t-s}|Y_{1:t-s})q(X'_{t-L+1}|X_{1:t-s}, Y_{t-L+1:t}) \quad (4.6)$$



#### 4.2.1 Artificial conditionals

We use  $\rho(\cdot) = q(\cdot)$ . Why?

#### 4.2.2 History Proposals

“Conservative” resampling as a history proposal.

### 4.3 Fixed lag estimation for tracking

Erm... i’ve kinda said what I was going to say here in the previous section

#### 4.3.1 Importance distributions

The dimensionality of a fixed lag particle filter will generally be very large. Consider a problem where each target has a 4-dimensional kinematic state (position and velocity in  $x$  and  $y$ ) and an observation-association index. This gives us five dimensions per target per time step. If we use a lag window with length  $L = 5$  and 5 targets we will have a 125 dimensional state. If we used a basic, bootstrap approach to such a problem, the chances of even a single particle following the correct path are negligible. Instead we must exploit the strong correlations between states arising from the structure of the problem. We first factorise the proposal thus:

$$\begin{aligned} q(x_{t-L+1:t}, \lambda_{t-L+1:t} | x_{t-L}, Y_{t-L+1:t}) \\ &= q(\lambda_{t-L+1:t} | x_{t-L}, Y_{t-L+1:t}) \\ &= q(x_{t-L+1:t} | x_{t-L}, \lambda_{t-L+1:t}, Y_{t-L+1:t}) \end{aligned} \quad (4.7)$$

Thus we can sample first the association variables then the state variables. Each of these terms can be further factorised over time, as we shall see. If we were to replace each of the factors with its corresponding posterior distribution then this factorisation would recreate the “optimal” importance distribution.

#### Association proposals

The proposal for the association variables is a discrete distribution over the possible observations with which the target could be associated. Within the factorisation above, the “optimal” form of the proposal is:

$$\begin{aligned} q(\lambda_{t-L+1:t} | x_{t-L}, Y_{t-L+1:t}) \\ &\propto P(Y_{t-L+1:t} | \lambda_{t-L+1:t}, x_{t-L}) P(\lambda_{t-L+1:t} | x_{t-L}) \\ &= \prod_{\substack{k=1:L \\ tt=t-L+k}} \int P(Y_{tt} | x_{tt}, \lambda_k) P(x_{tt} | Y_{tt+1:t}, \lambda_{tt+1:t}, x_{t-L}) dx_{tt} P(\lambda_{tt}) \end{aligned} \quad (4.8)$$

This can be calculated analytically for the linear-Gaussian case. For other cases, the EKF approximations may be used. The only problematic term in the expression above is the state

distribution over  $x_{tt}$ , which will require  $k$  integrals to calculate. Although this is will be analytic with Gaussian dynamics, its complexity may become unmanageable. This term represents the probability of states given the set of observations associated with this target at later times. We can render this calculation more manageable by replacing the whole set of future observations with just one.

$$P(x_k|Y_{k+1:t}, \lambda_{k+1:t}, x_{t-L}) \approx P(x_k|Y_{k+d}, \lambda_{k+d}, x_{t-L}) \quad (4.9)$$

In general we will use  $d = 1$ , as the closest observation in time will give us the most information about  $x_{tt}$ . However, if the target is not detected at time  $tt + 1$ , then we can increase  $d$  to find the next detection of the target.

Using this approximation, for  $\lambda_{t-L+k} \neq 0$ , we have

$$q(\lambda_{t-L+1:t}|x_{t-L}, Y_{t-L+1:t}) \propto \prod_{\substack{k=1:L \\ tt=t-L+k}} \mathcal{N}(y_{tt}^{(\lambda_{tt})}|m_{tt}, S_{tt}) \quad (4.10)$$

where usually

$$S_{tt} = [I - R^{-1}C_{tt}\Sigma_{tt}C_{tt}R^{-1}]^{-1} \quad (4.11)$$

$$m_{tt} = SR^{-1}C_{tt}\Sigma_{tt}[(A^d)^T C_{tt+d}^T R_d^{-1} y_{tt+d}^{\lambda_{tt+d}} + Q_k^{-1} A^k x_{t-L}] \quad (4.12)$$

$$\Sigma_{tt} = [C_{tt}^T R^{-1} C_{tt} + (A^d)^T C_{tt+d}^T R_d^{-1} C_{tt+d} A^d + Q_k^{-1}]^{-1} \quad (4.13)$$

$$Q_d = \sum_{l=0}^{d-1} A^l Q (A^l)^T \quad (4.14)$$

$$R_d = R + C_{tt+d} Q_d (C_{tt+d})^T \quad (4.15)$$

We will need a different expression for the case when  $tt = t$ , because no future associations have yet been proposed. Similarly, if  $tt < t$  but the future associations have all been proposed as missed detections, then there are no future observations to guide us, whatever choice of  $d$  we use. In these cases we have:

$$S_{tt} = R_d \quad (4.16)$$

$$m_{tt} = C_{tt} A^k x_{t-L} \quad (4.17)$$

For full derivations, see Appendix.

So we now have a full mechanism for association proposals.

Put in some pseudo-code here.

Add a bit about adapting the MD proposal when there is clearly no valid observation.

### *Forward filtering backward sampling*

When the state space model is linear and both transition and observation densities are Gaussian, we can propose directly from the “optimal” importance distribution for the states. This is the forward-filtering-backward-sampling algorithm of (?). Starting with  $x_t$ :

$$q(x_t|x_{t-L}, \lambda_{t-L+1:t}, Y_{t-L+1:t}) = \quad (4.18)$$

## *Fixed Lag Particle Filters*

---

### 5.1 Sequential Importance Sampling and Resampling

#### 5.1.1 *Coping with dimensionality*

### 5.2 Markov Chain Monte Carlo

### 5.3 Marginalised Particle Filters



## *Plan Of Future Research*

---

radar, insects, crowds, cells, jump-diffusions, cursor tracking...



# *Bibliography*

---

- [1] Pierre Simon Laplace. Memoir on the probability of the causes of events. *Statistical Science*, 1(3):pp. 364–378, 1774.
- [2] R E Kalman. A new approach to linear filtering and prediction problems. *Journal Of Basic Engineering*, 82(Series D):35–45, 1960.
- [3] F. Daum. Nonlinear filters: beyond the kalman filter. *Aerospace and Electronic Systems Magazine, IEEE*, 20(8):57 –69, aug. 2005.
- [4] O. Cappé, S.J. Godsill, and E. Moulines. An overview of existing methods and recent advances in sequential monte carlo. *Proceedings of the IEEE*, 95(5):899–924, 2007.
- [5] A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In D Crisan and B Rozovsky, editors, *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2009.
- [6] Michael K. Pitt and Neil Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, 94(446):pp. 590–599, 1999.
- [7] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F*, 140(2):107 –113, apr 1993.
- [8] Arnaud Doucet, Mark Briers, and Stéphane Sénécal. Efficient block sampling strategies for sequential monte carlo. *Journal of Computational and Graphical Statistics*, 15(3):693–711, 2006.
- [9] M. Briers, A. Doucet, S.R. Maskell, and P.R. Horridge. Fixed-lag sequential monte carlo data association. In *Proceedings of SPIE*, volume 6236, page 62360S, 2006. Paper copy from British Library.