Propagating Uncertainty Pertinent to the Parameters of the Dynamic Model and Likelihood

Prof Simon Maskell, University of Liverpool

March 2024

Current Projects

Current projects related to Space:

- Dstl:
  - 'Stone Soup for Astrodynamics', £121k
  - 'Survey Space Object Manoeuvre Detection', £92k
  - 'Maximisation of Sensitivity over Long Timelines for Resolving Unresolved Artefacts of Threats', £206k (with Manchester)
- PhD studentships (co-funded by Dstl and EPSRC)
  - Gemma Cook, 'Improving Conjunction Analysis using a Combination of High-Fidelity Astrodynamics Models and Advanced Numerical Bayesian methods'
  - Ben Oakes, 'Scheduling Surveillance of Space Objects'
- ESA: Robust Orbit Determination (with GMV and UC3M), £87k

Previous Projects

Previous projects have looked at:

- (Numerical) Integrators (with UCL);
- Collision Threat;
- Sensor Management and Control (with LJMU/LCO Global);
- Future Space Populations (with UCL/Roke).

Common Denominators (1/2)

Common Denominators (2/2)

ODEs and SDEs

Assume that the true motion is an ODE in continuous time:

$$\dot{x}_t = f_0(x_t) \tag{1}$$

We approximate with $f(x_t)$ and so have an error:

$$\dot{x}_t = f(x_t) + \underbrace{f_0(x_t) - f(x_t)}_{\text{error}} = f(x_t) + e_t \tag{2}$$

where e_t are a stochastic approximation to the forces we are not modelling: We should really therefore think about SDEs, not ODEs.

Discretising SDEs Exactly

If SDE is linear with Brownian noise (with covariance, σ), then:

$$\dot{x}_t = Fx_t + e_t \Leftrightarrow p(x_\Delta) = \mathcal{N}(\mu(\Delta, x_0, F), \Sigma(\Delta, x_0, F, \sigma)) \quad (3)$$

where we can calculate $\mu(\Delta, x_0, F)$ and $\Sigma(\Delta, x_0, F, \sigma)$ using Matrix exponentials (eg calculated using Van Loan), eg for (nearly) constant velocity model:

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \sigma = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{q} \end{bmatrix} \Leftrightarrow \mu = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}, \Sigma = \begin{bmatrix} \frac{\Delta^3}{3} & \frac{\Delta^2}{2} \\ \frac{\Delta^2}{2} & \Delta \end{bmatrix} \tilde{q} \quad (4)$$

Discretising Relevant SDEs

- SDEs pertinent to astrodynamics are not linear
  - Particularly if we consider manoeuvres between observations
- We can approximate them
  - Use Van Loan on approximating linear SDE
  - Use piecewise linear approximation (aka data augmentation)
- Alternative (popular) approach
  - Use numerical (ODE) integrator/propagator of choice
  - Add noise at end of interval
    - Particularly poor approximation for long intervals
    - Choosing noise size and structure is, at best, ad-hoc
- We really need to think this through a lot more carefully
- We also (probably) need a time-varying approximation fidelity
  - We probably need fast models with accurate uncertainty (for Initial Orbit Determination and in the presence of manoeuvres)
  - We probably need better models (eg for track custody)

Algorithms

Algorithms exist for handling uncertainty in the context of data association:

- Auction (aka Munkres, Hungarian, JVC);
- MHT (aka MFA, S-D assignment);
- JPDA, which can be sped up using;
  - Belief Propagation (approximate);
  - Efficient Hypothesis Management (exact).
    - Patented until June 2024.

However, we need to articulate:

- Intra-sensor data association should be relatively easy
  - Calibration errors are self-consistent
- Inter-sensor data association is ambiguous
  - Inevitable calibration errors wrt sensor biases

Judicious Definition of the State

Consider a state, X_t, where:

$$X_t = \begin{bmatrix} x_t \\ \delta_{1,t} \\ \delta_{1,t} \end{bmatrix} \tag{5}$$

Consider two sensors (for $i \in \{1, 2\}$) with additive noise $w_{i,t}$:

$$y_{i,t} = h_1 \left(x_t + \delta_{i,t} \right) + w_{i,t} \tag{6}$$

Consider large initial uncertainty with respect to x_t:

- Measuring $y_{1,t}$ causes negative correlation between x_t and $\delta_{1,t}$;
- Predictions of $y_{1,t+1}$ will have low uncertainty
  - Ensuring intra-sensor data association is easy;
- Predictions of $y_{2,t+1}$ will have high uncertainty
  - Articulating inter-sensor data association challenge.

Sales Pitch

Particle filters are:

- Adept at propagating uncertainty over time.
  - Avoids local Gaussian approximation inherent in Gaussian sum
- Readily parallelised (including resampling).
- Difficult to apply to astrodynamics

Why?

Particle Filters 101

Particle filter assumes that:

$$p\left(x_t|y_{1:t}\right) \approx \sum_{i=1}^{N} w_t^i \delta\left(x_t - x_t^i\right) \tag{7}$$

Particles are propagated through a proposal, which we choose:

$$x_t^i \sim q\left(x_t^i|x_{t-1}^i, y_t\right) \tag{8}$$

and given a new weight:

$$w_t^i = w_{t-1}^i \frac{p\left(x_t^i|x_{t-1}^i\right) p\left(y_t|x_t^i\right)}{q\left(x_t^i|x_{t-1}^i, y_t\right)} \tag{9}$$

Existing choices of Proposal

Existing choices of proposal include:

- Dynamics ('bootstrap' particle filter):
  - (Very) easy to implement;
  - Terrible with informative measurements, eg in initial orbit determination
- Extended Kalman filter:
  - Theoretically invalid but worryingly popular;
  - Works quite well if problem is locally linear;
  - Until it doesn't.
- Unscented Kalman filter:
  - Theoretically valid;
  - Works quite well if problem is locally linear.

New Proposal

In MCMC, techniques are mature that use gradient information to move samples around arbitrary pdfs:

- Langevin: one step stochastically towards uphill;
- Hamiltonian Monte-Carlo (HMC): multiple steps;
- No U-Turn-Sampler (NUTS): adaptive number of steps.

We have adapted NUTS so we can use it as a proposal in a particle filter.

- Result is reminiscent of Particle Flow
  - Out-performs at least one Particle Flow variant.
- We understand NUTS and can get it to work on lots of problems.

Residual Issue

NUTS approximates (quite well) an optimal proposal, which would be such that:

$$x_t^i \sim p\left(x_t^i | x_{t-1}^i, y_t\right) \tag{10}$$

and:

$$w_t^i = w_{t-1}^i p\left(y_t | x_{t-1}^i\right) \tag{11}$$

This is good and not good enough:

- Good because realisation of x_t^i doesn't impact w_t^i;
- Not good enough because w_t^i is still a function of x_{t-1}^i;

Which is problematic for systems that have a long memory

- eg those pertinent to astrodynamics.

Fixed Lag Filtering

We to go back in time and revisit the historic samples we had.

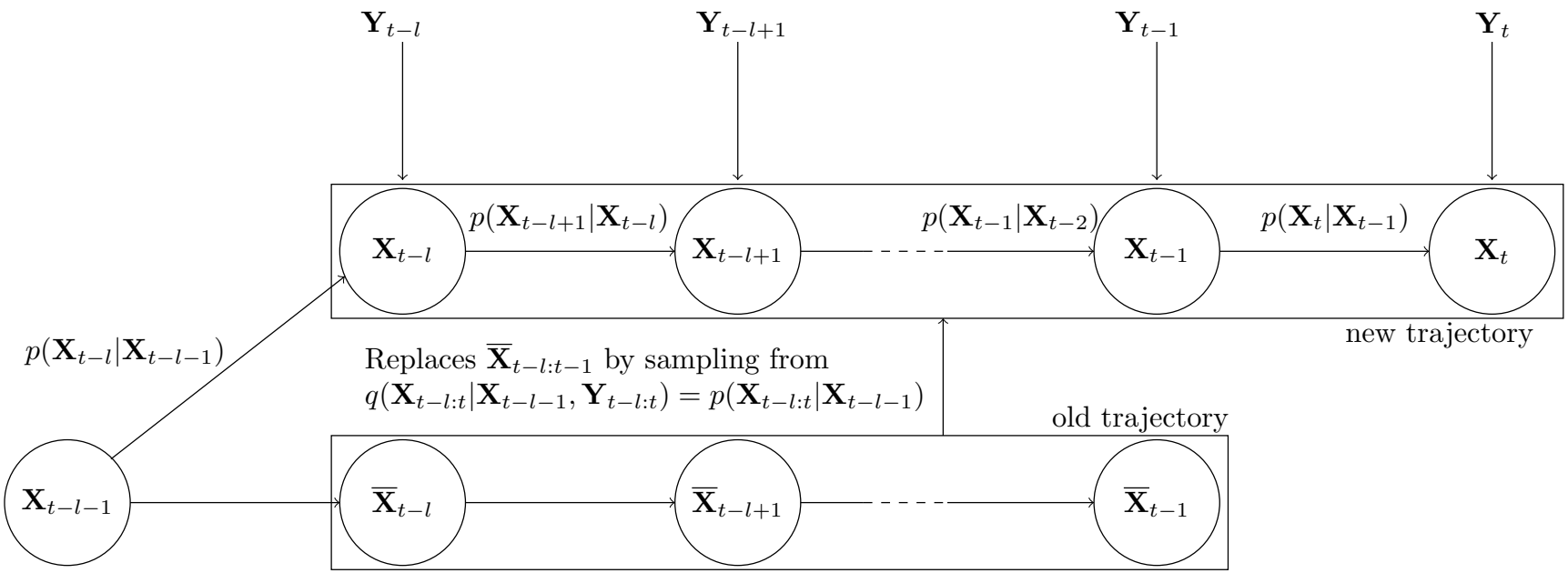
- Motivates Fixed-Lag Sequential Monte Carlo (FL-SMC);
- Aka 'Block-Sampling';
- Somewhere between a particle filter and full parameter estimation with PMCMC/SMC^2.
- Could be used adapted to perform calculations pertinent to conjunction analysis directly (eg using Markov bridges)

Maths gets a little complicated, but see:

- A Varsi, L Devlin, P Horridge and S Maskell. A General-Purpose Fixed-Lag No-U-Turn Sampler for Nonlinear Non-Gaussian State Space Models. Accepted for publication in IEEE Trans AES, 2024. `https://ieeexplore.ieee.org/document/10463105`

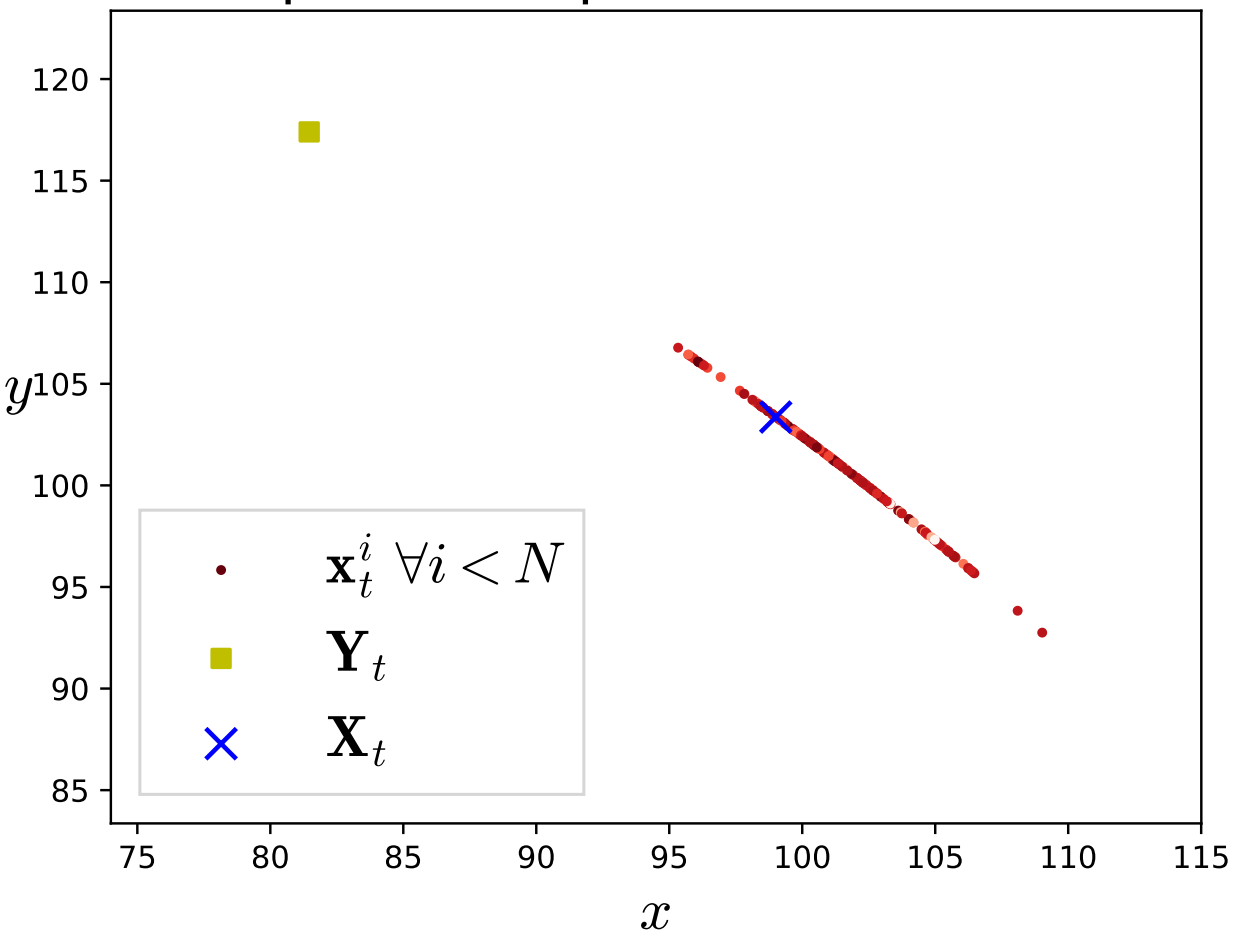
Stone Soup implementation (using GODOT) under development

Fixed-Lag SMC



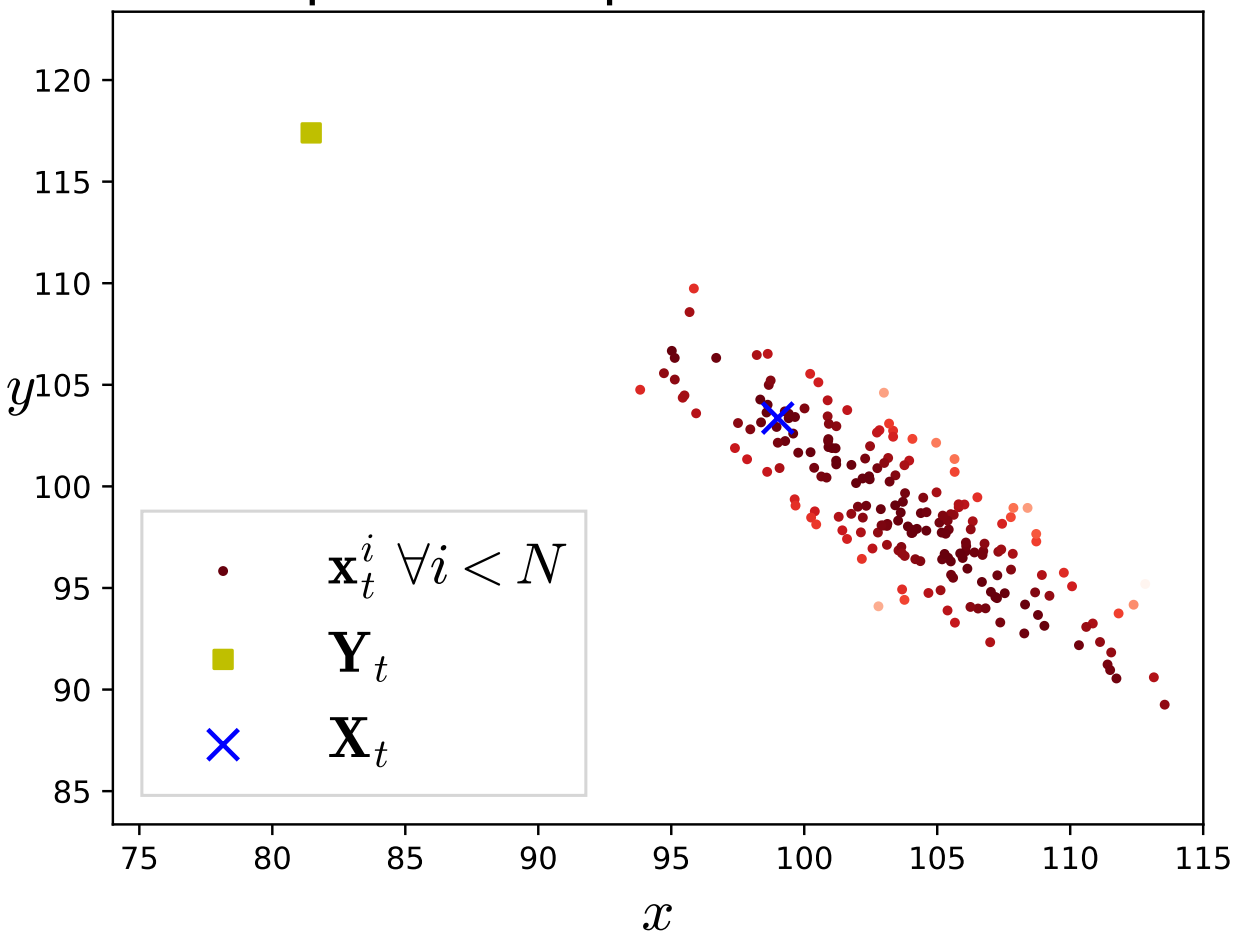
Exemplar FL-NUTS Results

STM model with nonlinear non-Gaussian posterior: pdf for FL-NUTS



Exemplar FL-UKF Results

STM model with nonlinear non-Gaussian posterior: pdf for FL-UKF



Results

Table: Nonlinear non-Gaussian posterior - 200 particles, $T = 4$.

Filter	l	**Time [s]**	$\mu_{\mathbf{x}_t}$	μ_ρ	$\frac{N_{eff}}{N}$
FL-SMC	0	0.01	52.47	2.87	0.005
FL-SMC	3	0.04	31.73	2.96	0.016
FL-EKF	0	0.03	191.03	26.37	0.131
FL-EKF	3	0.11	194.78	25.94	0.116
FL-UKF	0	0.03	56.24	6.45	0.107
FL-UKF	3	0.12	32.51	6.73	0.215
PFIPF	–	0.03	50.48	5.83	0.128
S-NUTS	–	0.03	249.08	11.68	0.443*
FL-NUTS w/o Opt	0	0.04	51.93	2.95	0.005
FL-NUTS w/o Opt	3	0.15	29.46	3.06	0.015
FL-NUTS	0	0.06	28.72	1.62	0.174
FL-NUTS	3	0.24	11.36	1.59	0.562

Conclusions

- We need better ways to solve SDEs pertinent to Astrodynamics;
- There's a pragmatic way to handle inter-sensor and intra-sensor data association ambiguity;
- The world has a new particle filter.