Particles, Points and Positions: Recent Advances in Modelling and Processing of Agile Objects

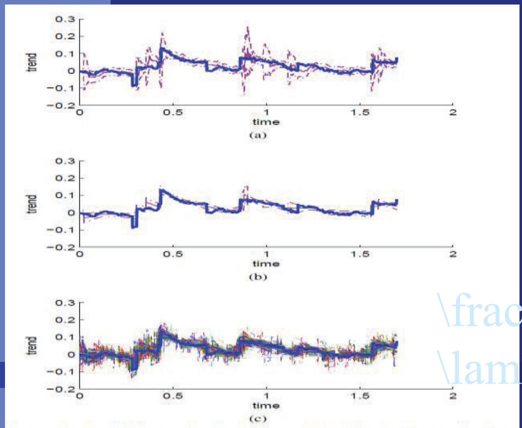
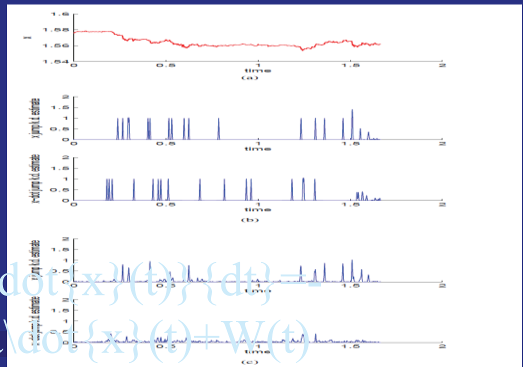
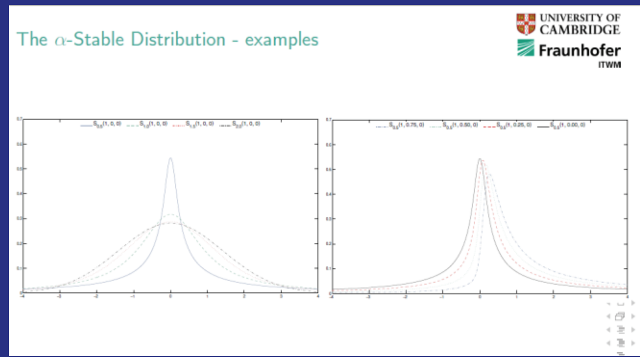


Fig. 5. Trend estimates using RBVRPF + Kalman filter (a) RBVRPF + RTS smoother (b) and RBVRPS + RTS smoother. Solid lines show means of each particle. Dashed lines show mean ±2 standard deviations.

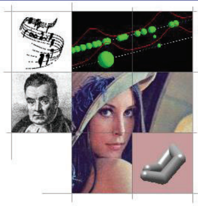
$$\frac{d\dot{x}(t)}{dt}=-\lambda\dot{x}(t)+W(t)$$



ig. 6. Foreign exchange data (a), along with filter (a) and smoother (b) ernel density estimates for the state and trend jump times.

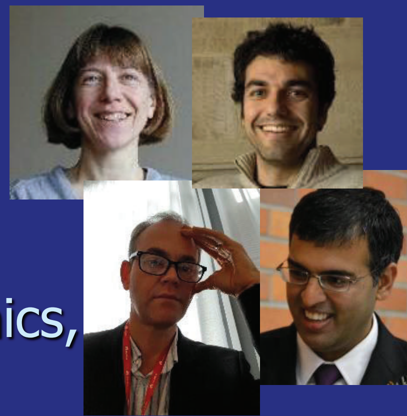


Simon Godsill
Signal Processing and Communications Lab.
University of Cambridge
www-sigproc.eng.cam.ac.uk/~sjg



Signal Processing and Communications Laboratory (SigProC)

- 5 Academic staff (Simon Godsill, Joan Lasenby, Albert Guillen-Fabregas, Ramji Venkataramanan, George Cantwell (Oct 2023))
- ~8 post-doctoral research fellows
- ~20 PhD students
- Diverse research topics, including:
  - Image and 3D data processing,
  - Computer vision and computer graphics,
  - Audio and music processing,
  - Statistical methodology (especially Bayesian methods),
  - Tracking, Sensor Fusion, Intentionality Inference
  - Information Theory and Communications



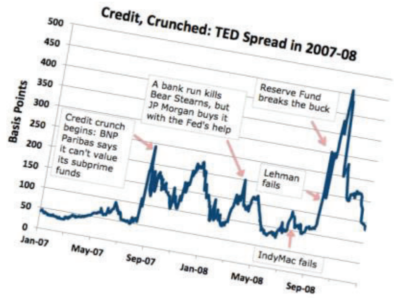
Motivation and Background

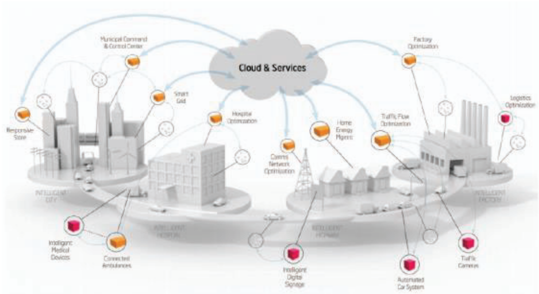
- Study of evolving spatio-temporal processes with incomplete and ambiguous measurement data
- Wish to infer in the presence of highly non-Gaussian (heavy-tailed) behaviours.
- Use powerful combinations of *continuous-time* stochastic processes models with modern Bayesian computational techniques.

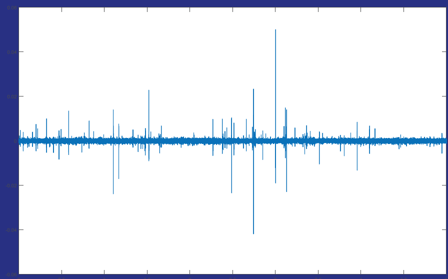
- In collaboration with (at least!):

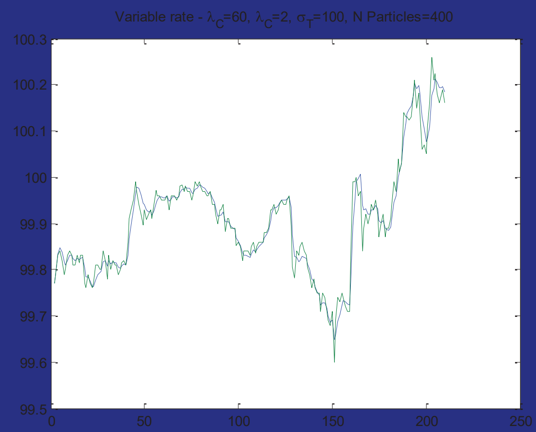
Yaman Kindap, Lily Li, Patrick Gan, Marina Riabiz, Ioannis Kontoyiannis, Marcos Tapia-Costa, Joe Johnston, Pete Bunch, Tohid Ardeshiri, Bashar Ahmad, Tatjana Lemke …

Framework: Heavy Tails and Asymmetry



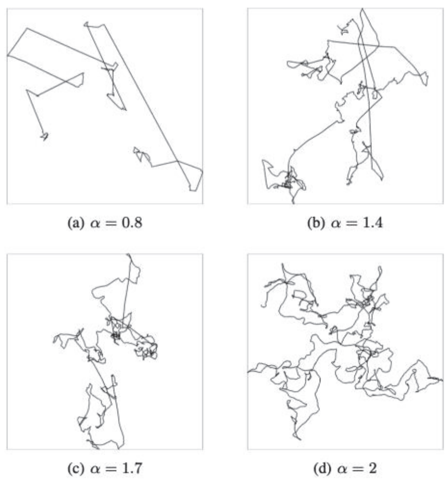


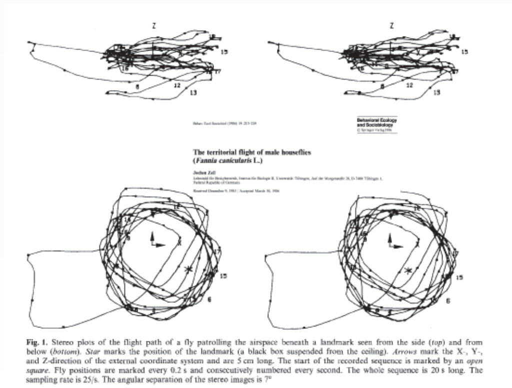


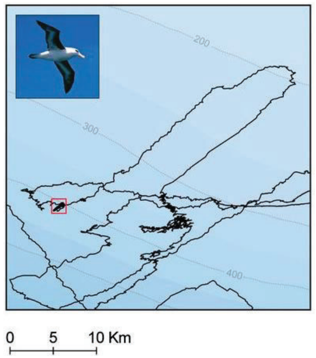


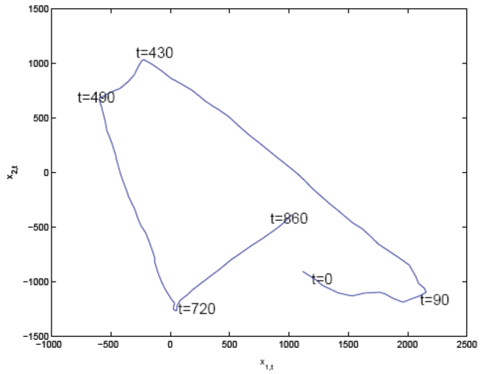
Spatio-temporal processes

- Irregular Movement (e.g. animal foraging, drones, etc.)



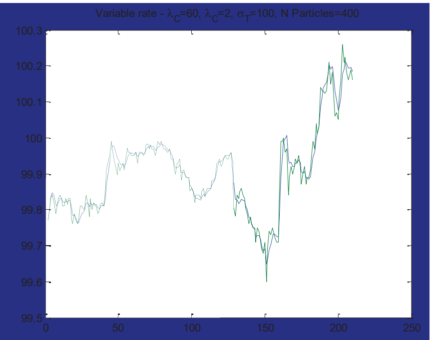






https://motchallenge.net/data/MOT17

Modelling with Jumps



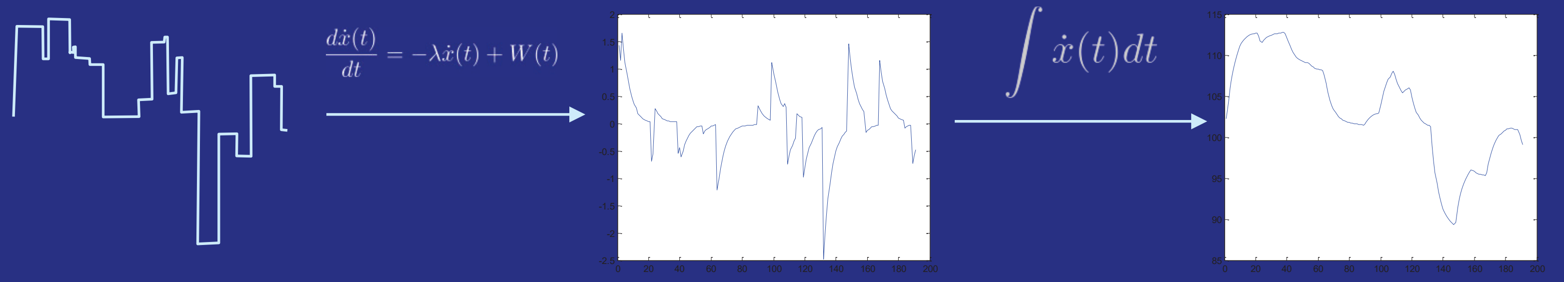
Generic linear form is:

$$dX(t) = AX(t)dt + hdW(t)$$

Example 1d case: stochastic trend model with jumps:

$$\frac{d\dot{x}(t)}{dt} = -\lambda \dot{x}(t) + W(t)$$

Now assume that $\{W(t)\}$ is not Brownian motion, but is made up of random 'jumps' at random times and random Gaussian amplitudes:

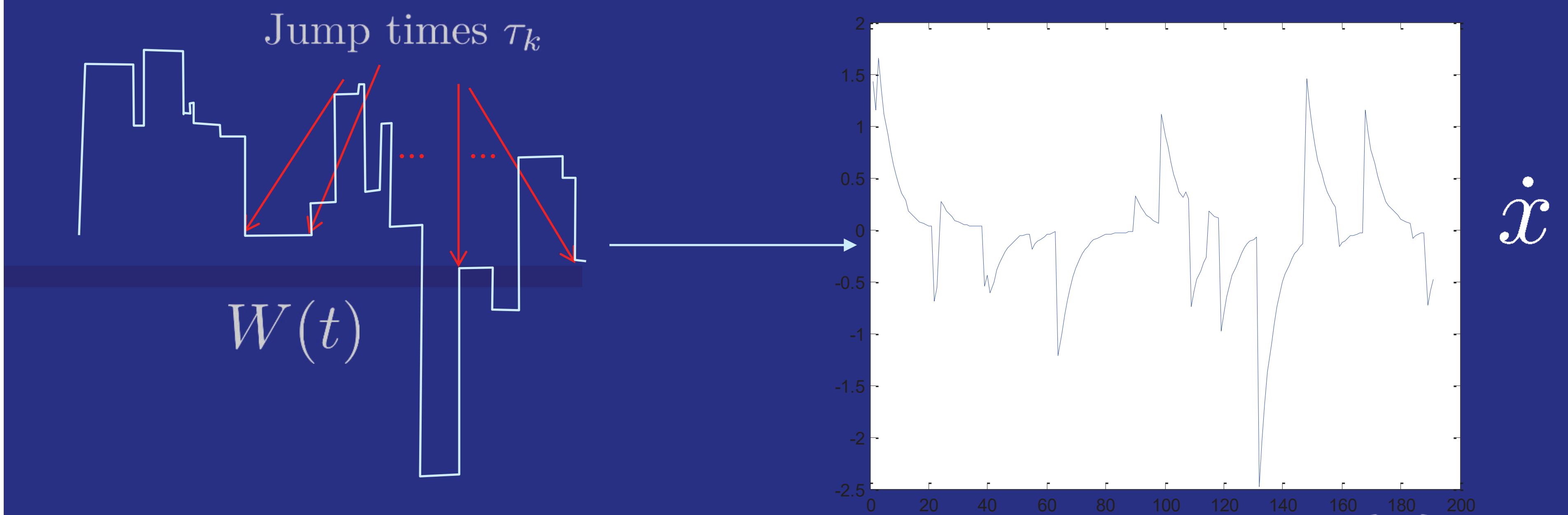


$W(t)$ $\dot{x}(t)$ $x(t)$

Observe on a discrete time 'skeleton':

$$y_k = Hx(t_k) + v_k, \quad v_k \sim \mathcal{N}(0, C_v)$$

See Christensen, Murphy and Godsill, IEEE J Sel. Top. SP, 2012

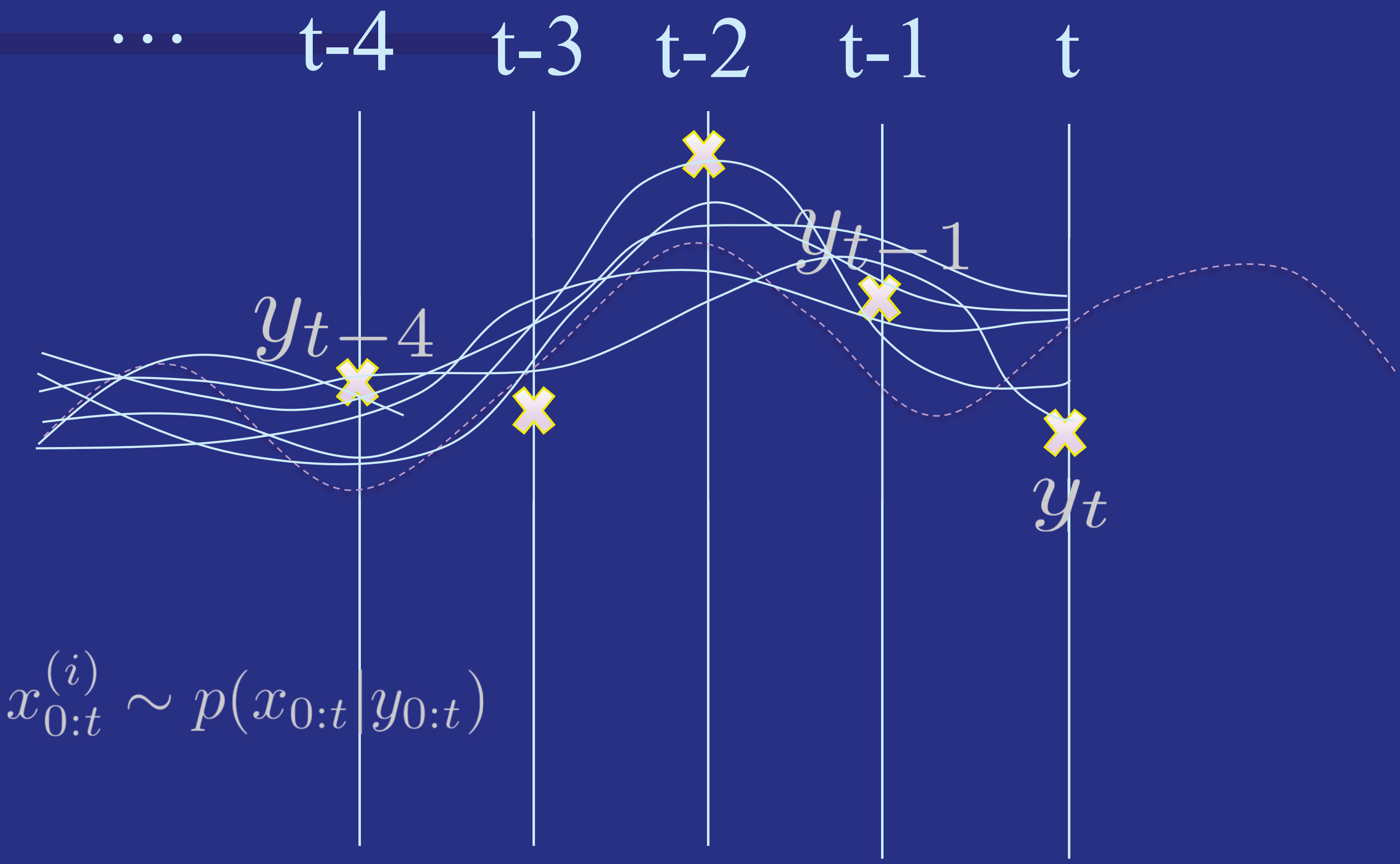


- Each particular realisation is fully characterised by its jump times $\{\tau_j\}$.
- The transition density $f(x(t)|x(s), \{\tau_j\})$ is *conditionally Gaussian*.
- Hence, under the linear/ Gaussian observation model we can compute a Gaussian likelihood using the Kalman filter, conditioned on the jump times:

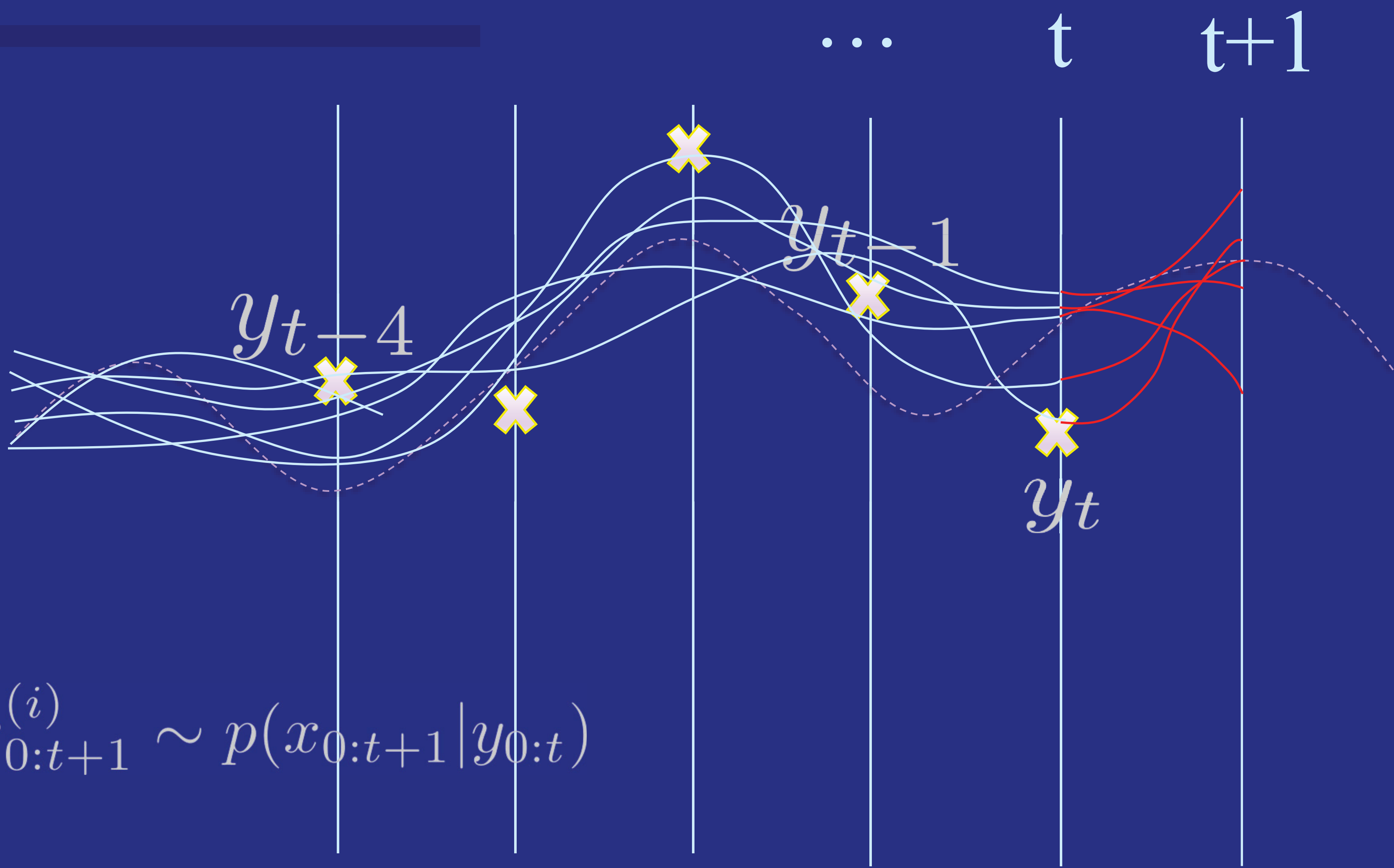
$$p(y_k|y_{1:k-1}, \{\tau_j\}_{\tau_j \le t_k}), \quad \text{Computed with KF PED}$$

- This makes the process tractable for sequential inference schemes such as Sequential Monte Carlo (particle filter)

The Particle Filter: first step. Time t: many random draws from the `path', $x_{0:t}$

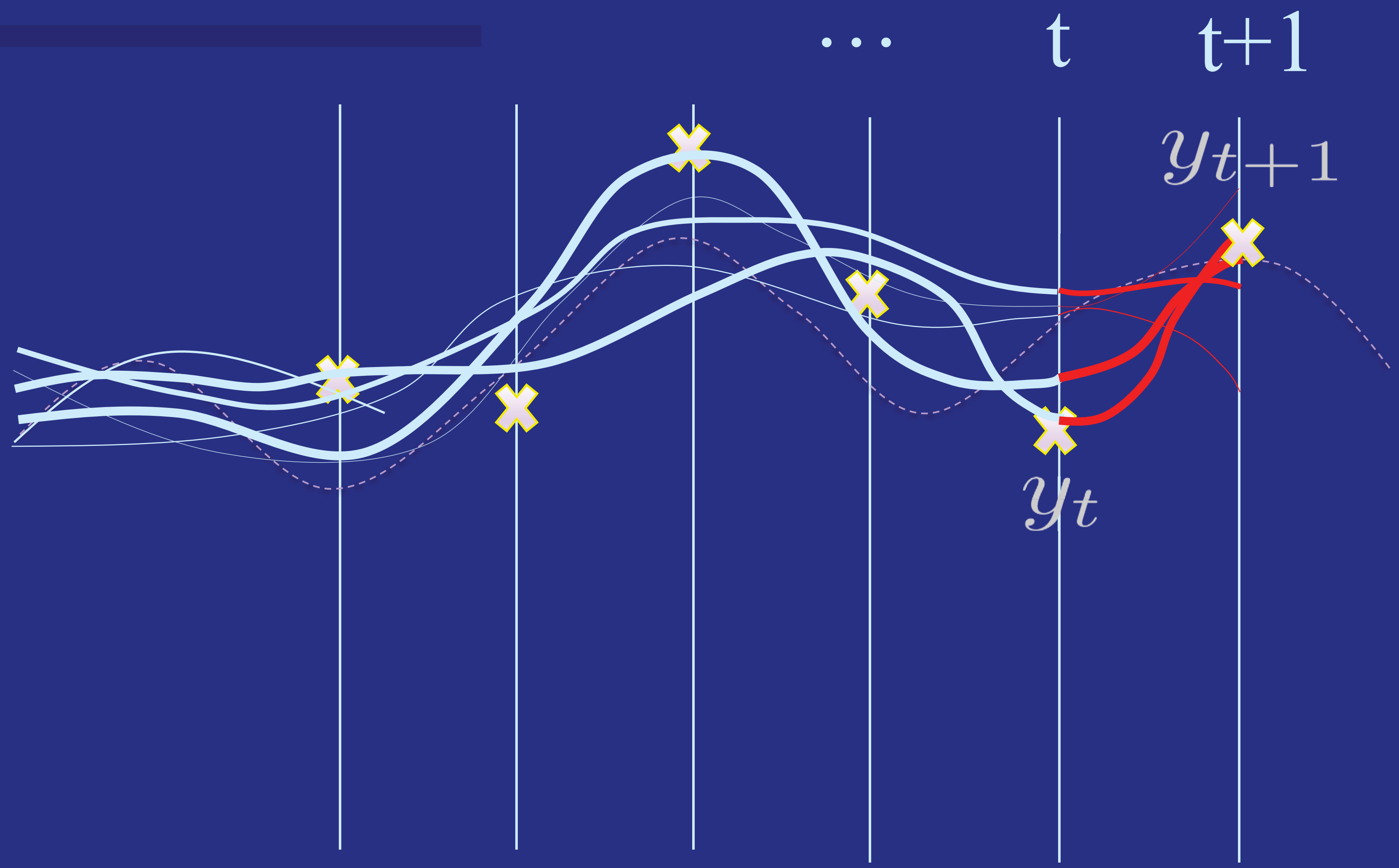


The Particle Filter: Prediction step. Extend each path randomly to time t+1 using $f(x_{t+1}|x_t)$

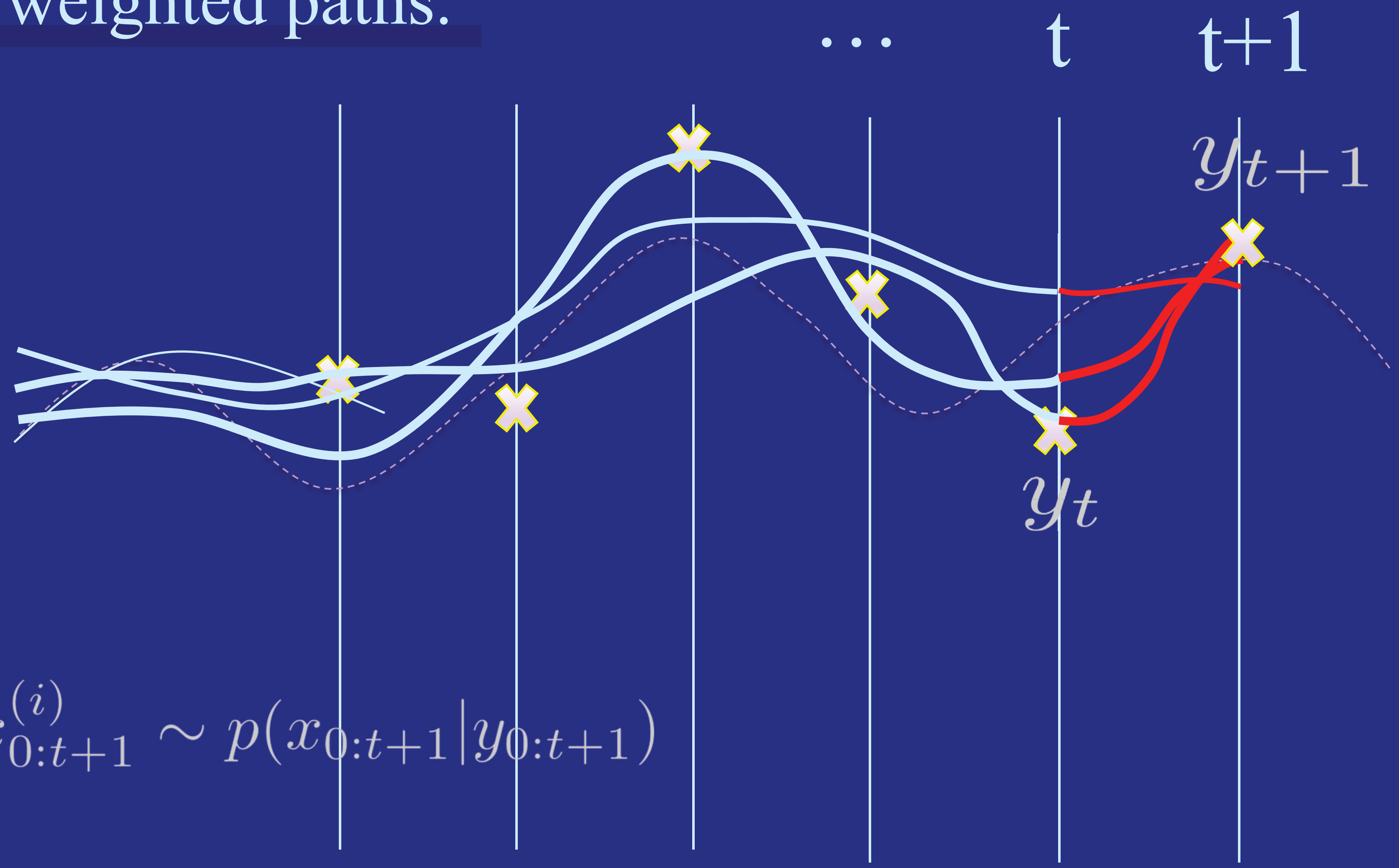


$$x_{0:t+1}^{(i)} \sim p(x_{0:t+1}|y_{0:t})$$

The Particle Filter: Update step. Compute an importance weight for each path $\propto g(y_{t+1}|x_{t+1})$



The Particle Filter: Final step. Randomly prune out low weight paths and boost the number of high weighted paths.



$x^{(i)}_{0:t+1} \sim p(x_{0:t+1}|y_{0:t+1})$

VRPF (particle) Example: tracking a manoeuvring ship through persistent clutter

Ship trajectory provided by QinetiQ Winfrith

Single sensor at (0,0)

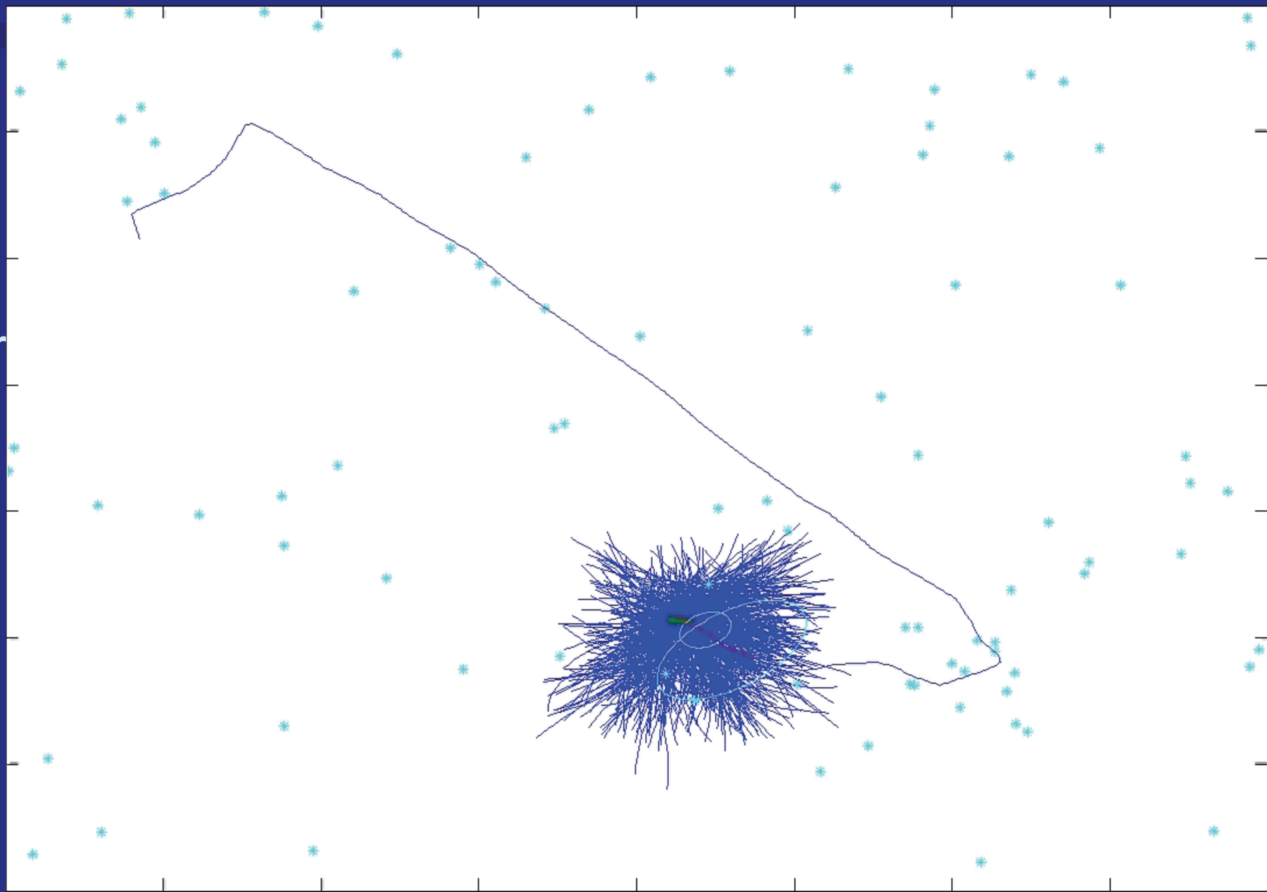
Mean number of clutter points λ_C=100

Mean number of target points = 3

Persistent clutter `hot-spot'

Poisson likelihood + VRPF

400 Particles

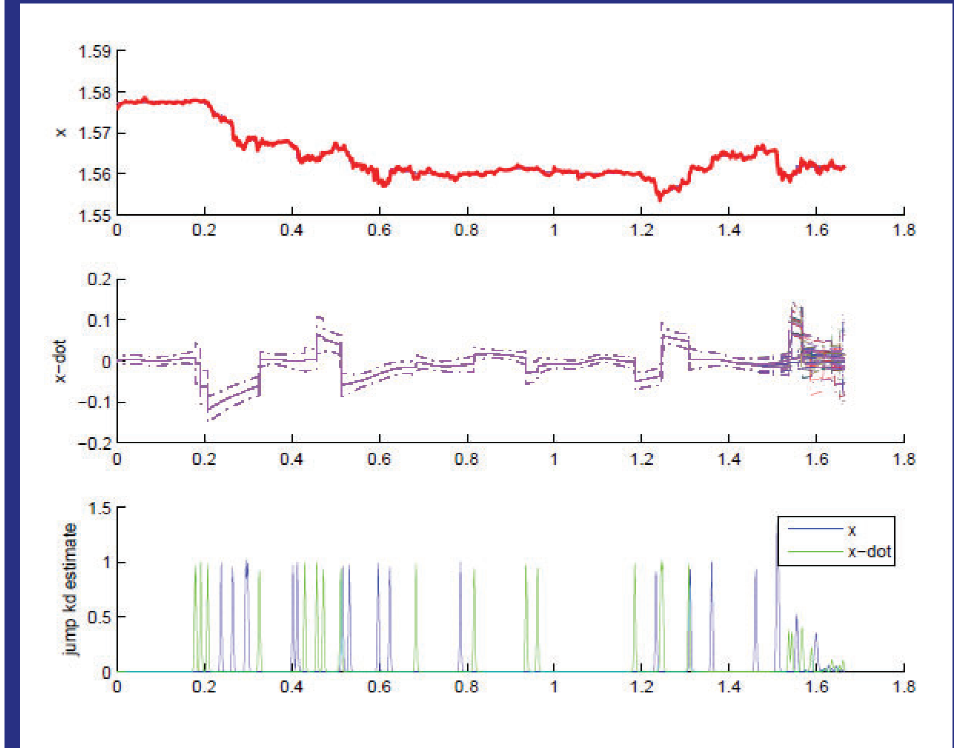


[Uses Variable Rate Particle Filter, see Godsill et al. 2006 Proc. IEEE, P. Bunch and S. Godsill (IEEE tr SP 2013a,2013b)]

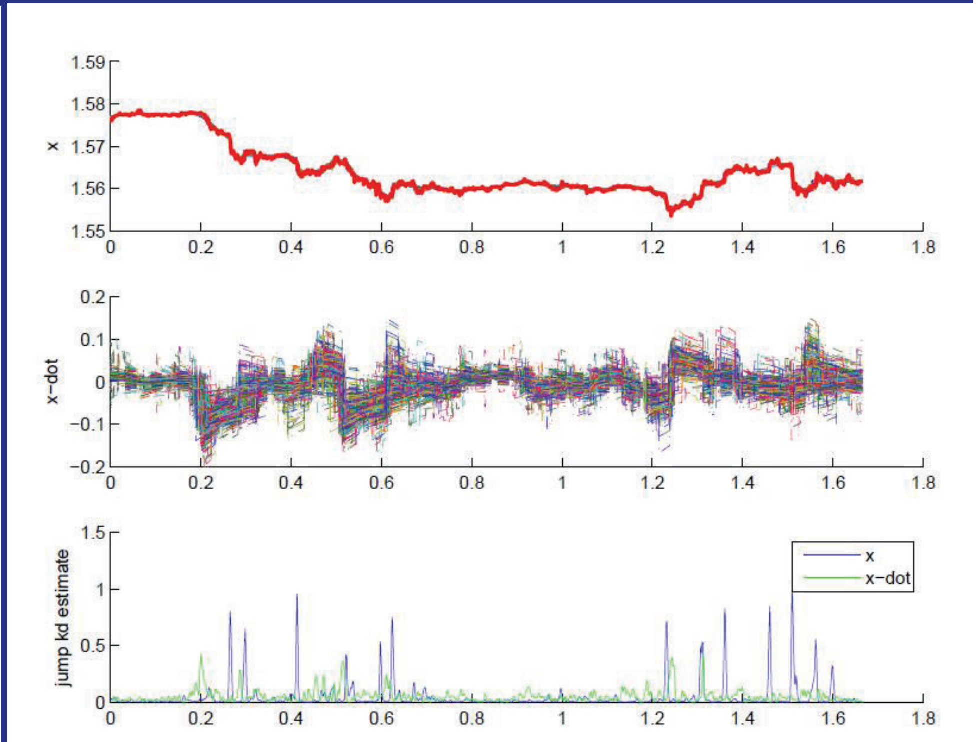
Filtering and smoothing with the stochastic trend jump model

April08 USD-GBR

Filtering only:



Forward filtering/backward sampling:



[See P. Bunch and S. Godsill (IEEE tr SP 2013a,2013b), Christensen, Murphy, Godsill (IEEE Sel. Ar. SP 2012), Sarkka, Bunch and Godsill (IFAC 2012), Godsill, Doucet, West (2004)]

The Lévy state space model

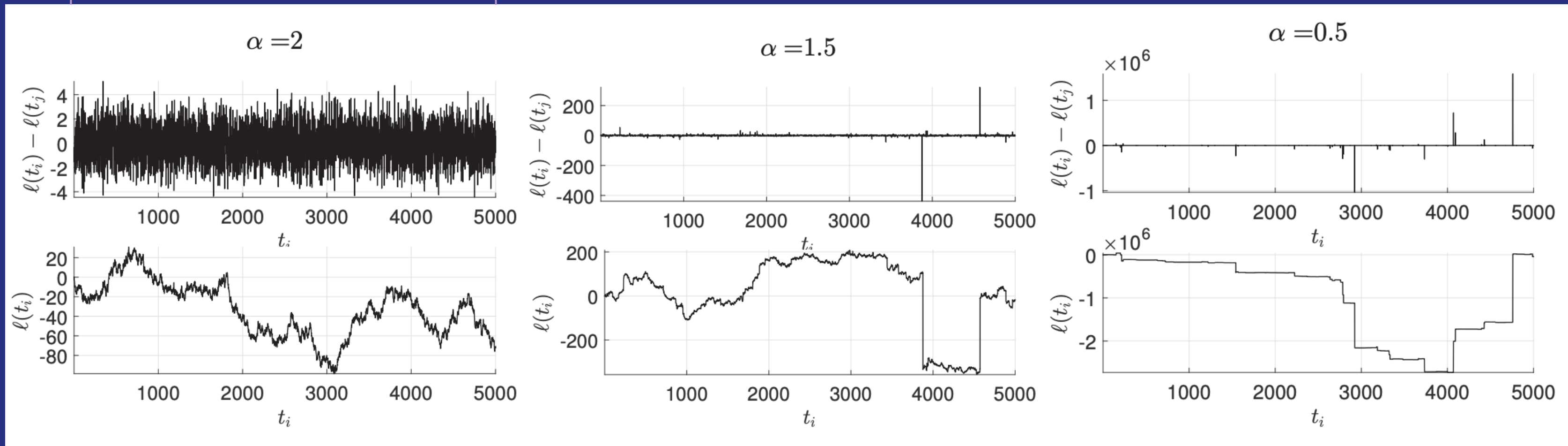
- Previously modelled jumps as a finite activity process
- Perhaps more realistic to model the jumps as an infinite collection of large/ small/tiny jumps occurring in each finite time interval
- It turns out that much more general classes of non-Gaussian processes can be obtained this way:
  - α-Stable, Student-t, variance-gamma, generalised hyperbolic, … see e.g. Cont and Tankov (2002)
- Recent work has shown that these too can be inferred within an optimal Bayesian framework using powerful representations based on Poisson processes, see e.g.:

  - [Gan and Godsill, 2020] R. Gan and **S. Godsill** (2020) α-Stable Levy State-space Models for Manoeuvring Object Tracking, in *Proc. of the International Conf. on Information Fusion*, South Africa.
  - [Riabiz et al., 2020] M. Riabiz, T. Ardeshiri, I. Kontoyiannis and **S. Godsill** (2020) Nonasymptotic Gaussian Approximation for Inference with Stable Noise,. 2018,arXiv 1802.10065. *IEEE Trans. on Information Theory, 2020*
  - [Godsill et al., 2019] **S . Godsill** and M. Riabiz and I. Kontoyiannis (2019), The Lévy State Space Model, Arxiv 1912.12524.

Lévy process models of non-Gaussianity

Gaussian

Non-Gaussian (heavy-tailed)



- Wish to model broad classes of heavy-tailed driving noise to suit application
- Adopt a generic Lévy process approach in which driving noise is modelled as *pure jump* processes in continuous time (and observed at random discrete times).
- Elegant and (fairly!) simple alternative to the standard Gaussian (Brownian motion)
- Many possible distributions: alpha-stable, Generalised Hyperbolic (inc. Student-t, normal-Gamma and normal-inverse Gaussian), normal tempered-stable, … see e.g. Cont and Tankov 2002
- These methods are ideal for irregularly sampled heterogeneous data sources as they construct the path of the process at arbitrary time points – in contrast with other non-Gaussian discrete-time models

Brownian vs non-Gaussian state space models

- We will be interested in tracking and other models that can be written in a state-space form with state evolution in continuous time:

$$dx(t) = Ax(t)dt + HdW(t)$$

and we will need to be able to characterise transition PDFs for treatment of discrete time measurements:

$$p(x(t)|x(t-\Delta))$$

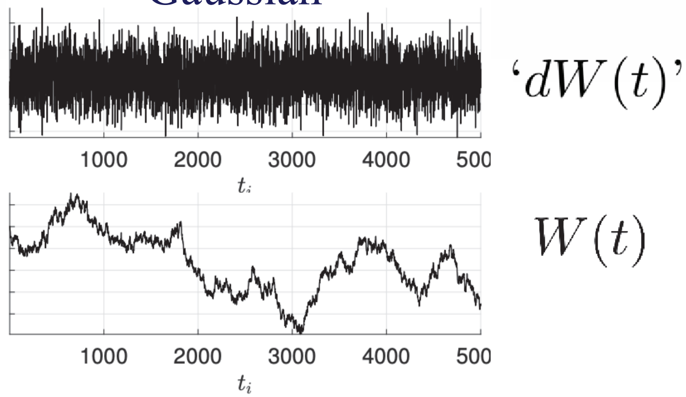
- If $\{W(t)\}$ is Gaussian (Brownian) then all calculations are simple (Kalman filter, standard stuff...)
- We are extending to a non-Gaussian $W(t)$ that moves *only* by small perturbations at random times $\{\tau_i\}_{i=1}^{\infty}$ ('jumps'), using a special conditionally Gaussian class ('Mean and Scale mixture of Gaussians') such that at any jump time τ_i:

$$dW(\tau_i) \sim \mathcal{N}(x_i\mu_W, x_i\sigma_W^2)$$

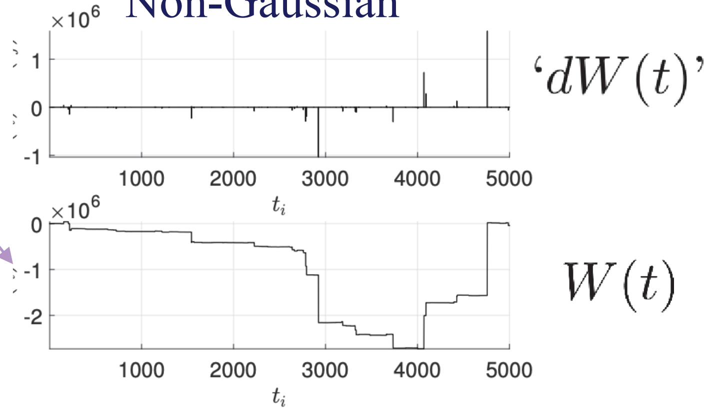
where $\{x_i\}_{i=1}^{\infty}$ are the jumps of another 'subordinator' process $X(t)$

- Because of the conditionally Gaussian structure we can implement the new models using banks of *standard achitectures* (Kalman filters etc.) inside standard particle filters
- And because of the 'discrete' structure (jumps) it is even easier to compute $p(x(t)|x(t-\Delta))$ than the pure Gaussian case for each Kalman filter

Gaussian



Non-Gaussian



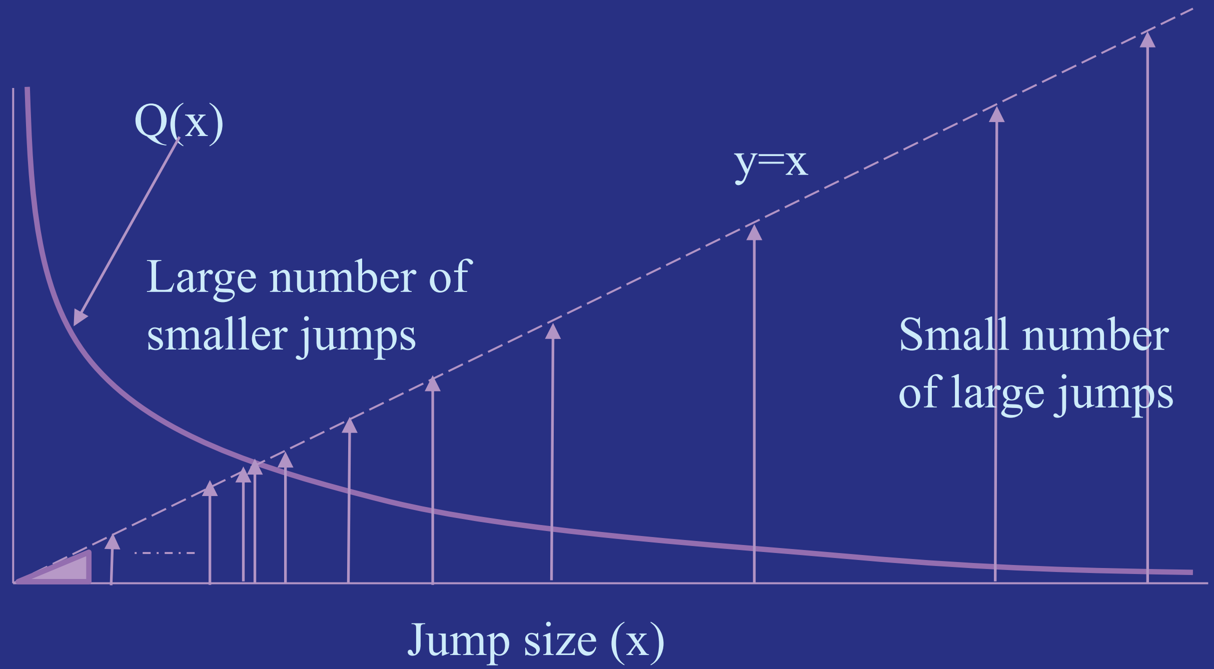
See Also:

$$(\alpha\text{-stable case}):\ dW(\tau_i) \sim \mathcal{N}(x_i\mu_W, x_i^2\sigma_W^2)$$

The Lévy State Space Model (1919) Godsill, Riabiz, Kontoyiannis https://arxiv.org/abs/1912.12524

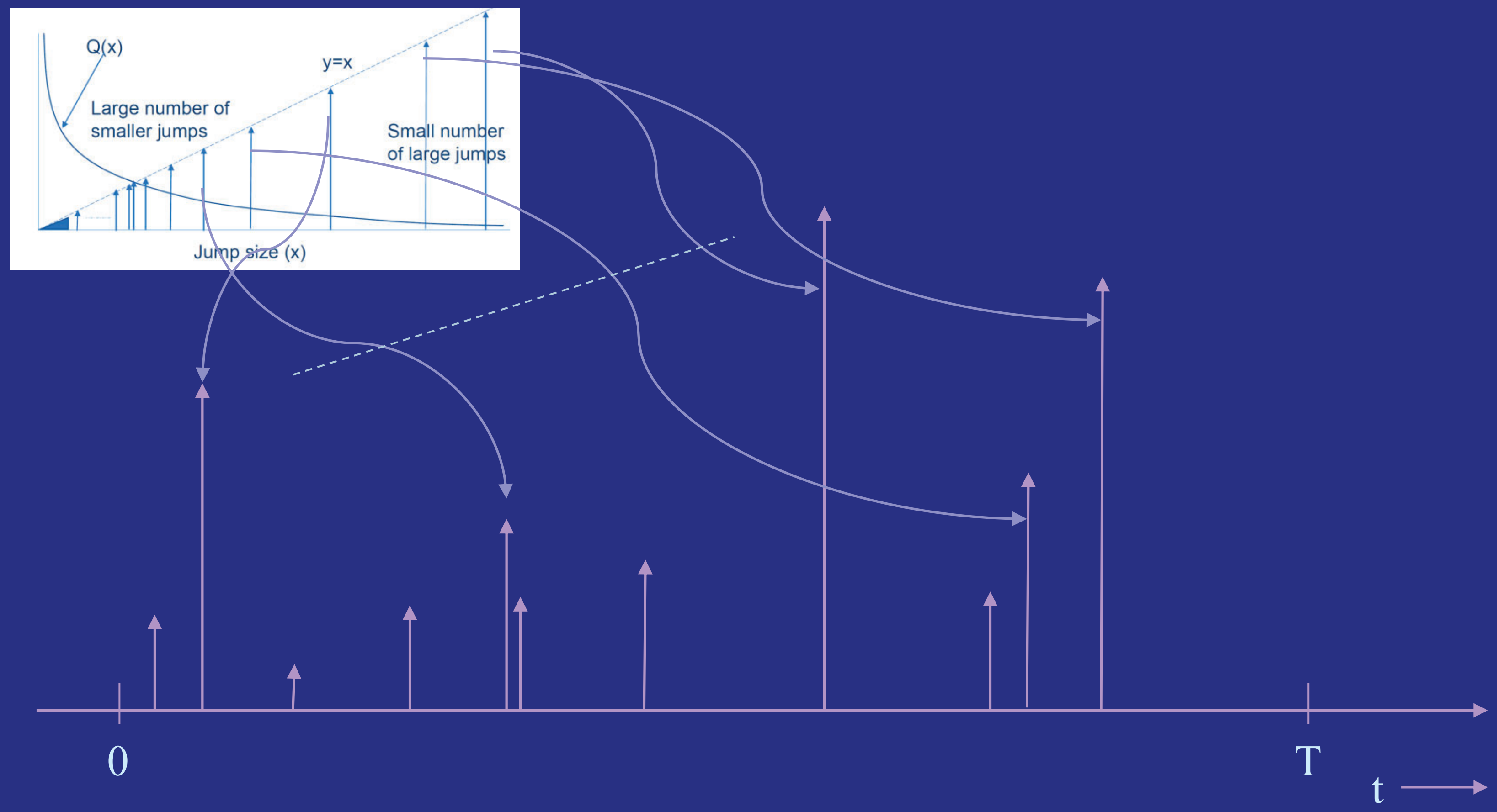
Lévy process models of non-Gausianity

- The Jumps $\{x_i\}_{i=1}^{\infty}$ are characterised by a Poisson process with non-uniform intensity function Q(x), the `Lévy Density':



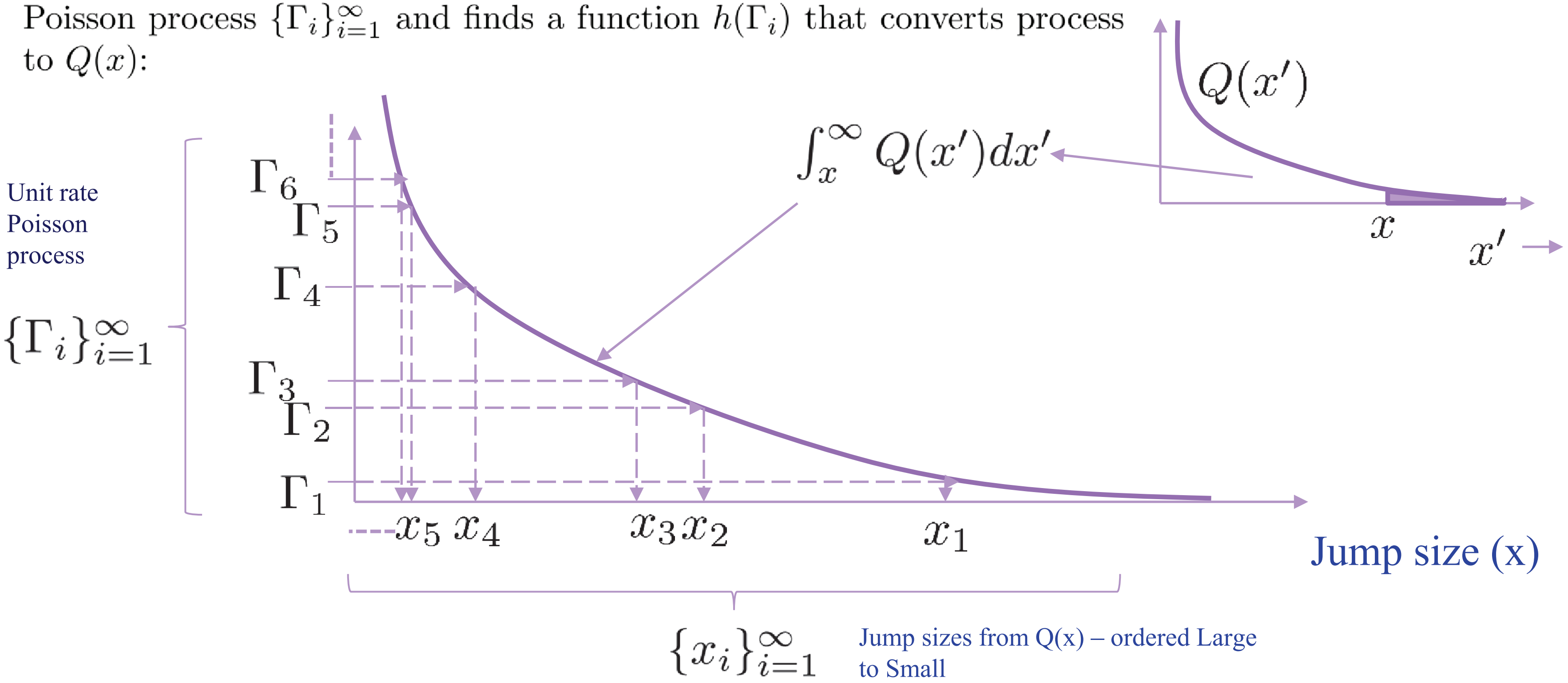
- This is a Poisson process where the average number of points in interval $(x, x+dx)$ is $Q(x)dx$

- Jumps are then uniformly randomly scattered across the time axis [0,T]:



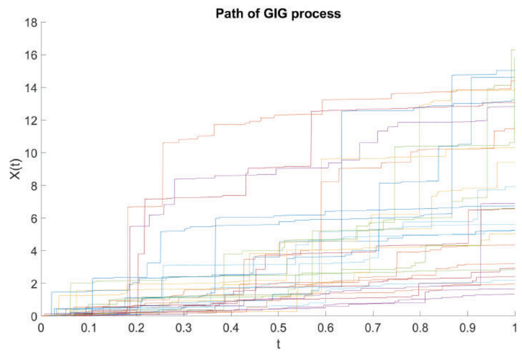
How to sample from Q(x)?

- In general $\int_0^\infty Q(x)dx \to \infty$ so can't sample directly (*Infinite Activity*)
- The classical method (Fergusson and Klaas 1970's) starts with a uniform Poisson process $\{\Gamma_i\}_{i=1}^\infty$ and finds a function $h(\Gamma_i)$ that converts process to $Q(x)$:

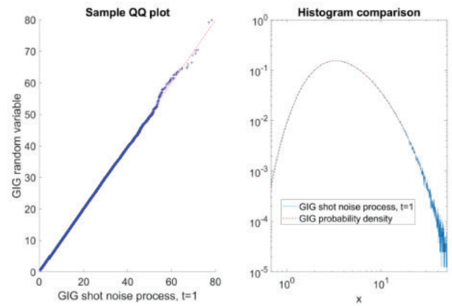


- Turns out the optimal function $h()$ is the inverse of $\int_x^\infty Q(x')dx'$.
- Can think of this as the direct analogue of sampling random variables using the *inverse CDF* method
- Also turns out that $h()$ can't be calculated for most of the processes we wish to use (e.g. the Generalised Inverse Gaussian (GIG))
- A lot of the fun then has been in developing effective alternative strategies for these cases:

  Godsill and Kindap (2021) Point process simulation of generalised inverse Gaussian processes and estimation of the Jaeger integral, Stats and Comp.



Study pdf at t=1:



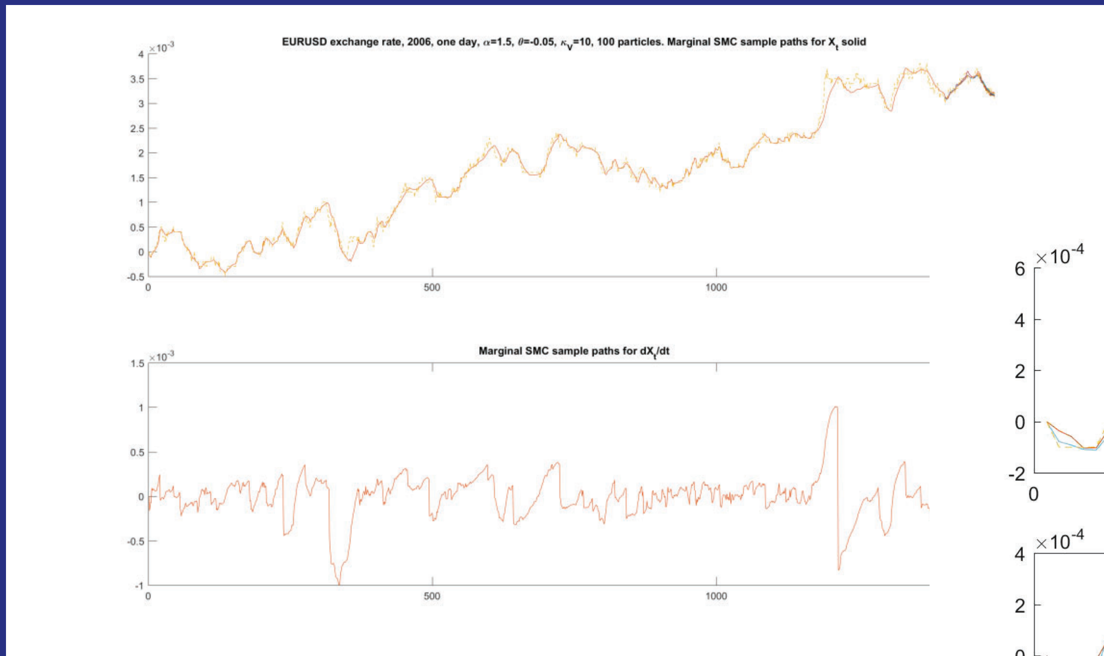
Inference Procedures

- Forward simulation of the jump sizes and times allows a random generation of paths of the continous-time process $X(t)$ over any chosen time interval
- This, coupled with the conditionally Gaussian form of the jumps, means that the process may be fitted very efficiently using simulation-based inference such as Particle Filtering and Markov chain Monte Carlo
- For implementation details see:

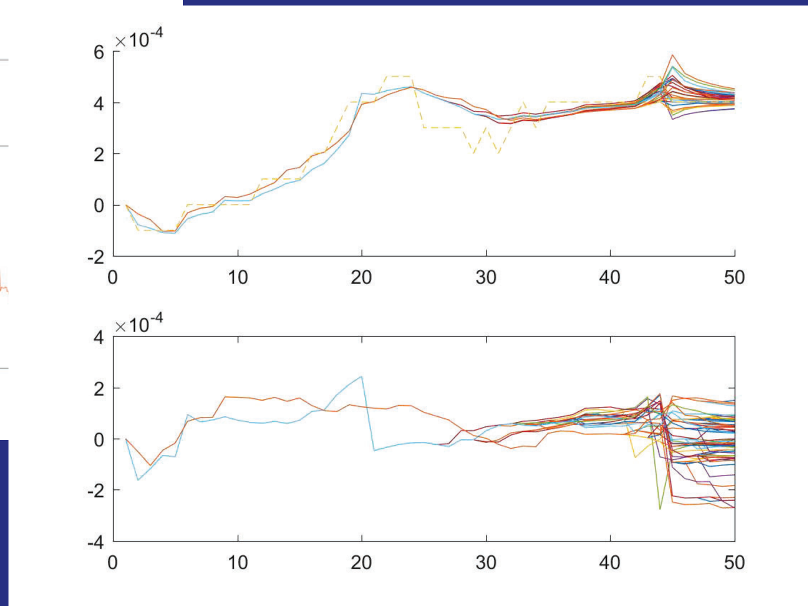
  The Lévy State Space Model. Simon Godsill, Marina Riabiz, Ioannis Kontoyiannis (Proc. Asilomar 2019)

  Inference for Variance-Gamma Driven Stochastic Systems Johnston, Kindap and Godsill (Fusion 2023 (to appear
- Implementation example for exchange rate data follows...

Application to High Frequency EuroDollar data (tick Data)



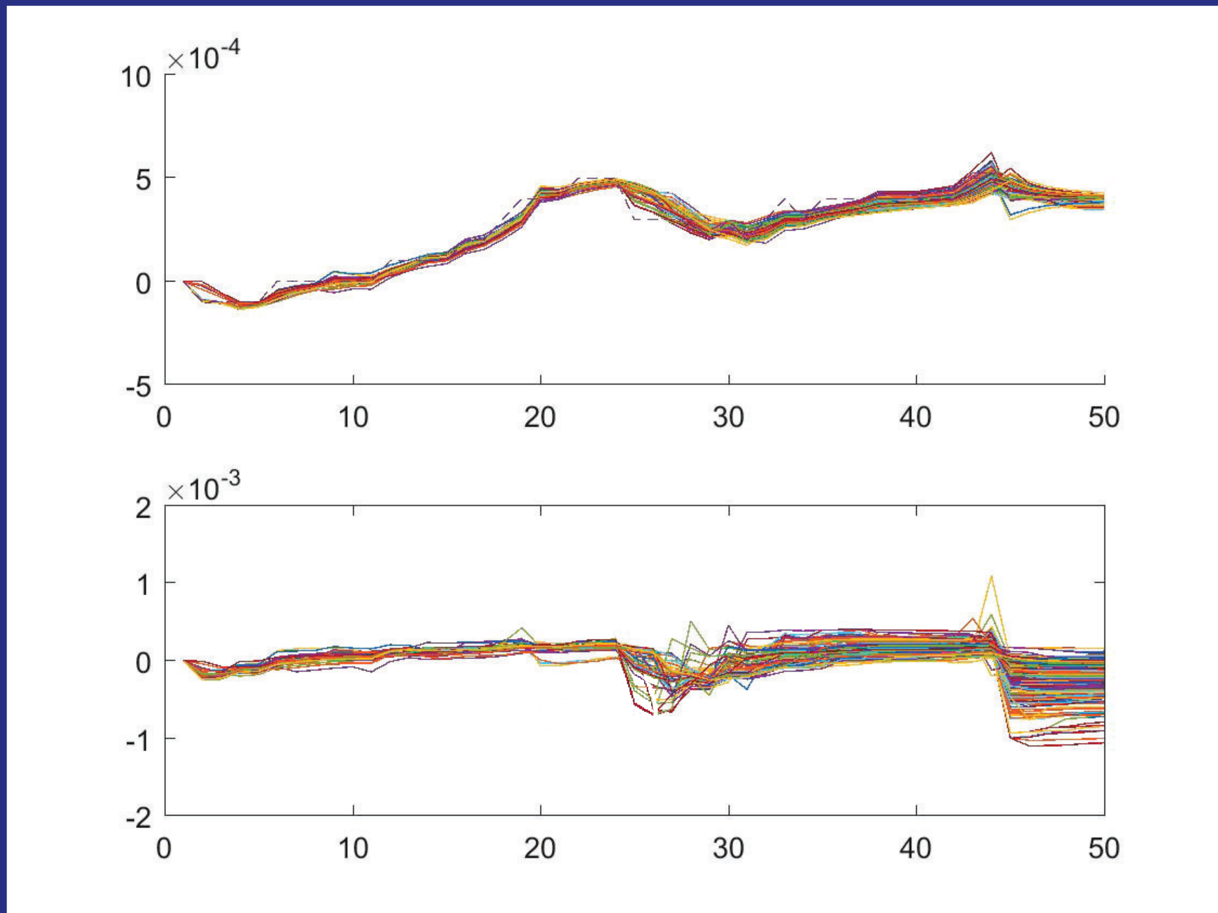
Marginal Monte Carlo filter. alpha=1.5, 100 particles



Alpha=0.8, 4000 particles:

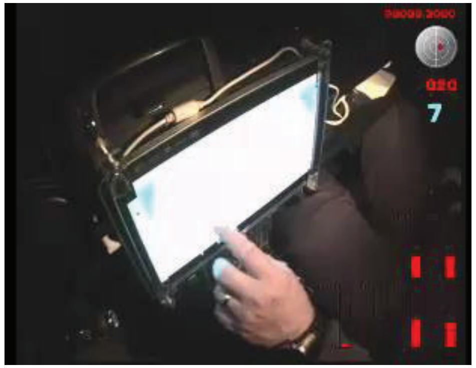
$x(t)$

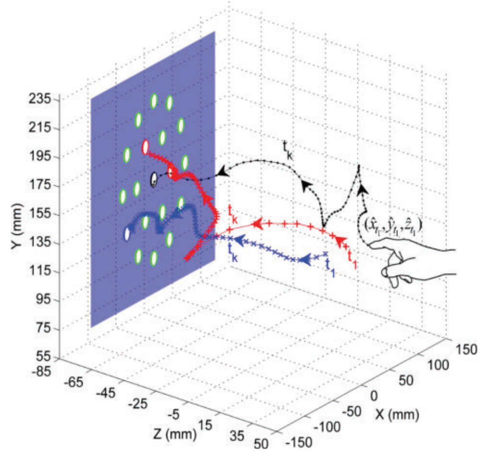
$\dot{x}(t)$

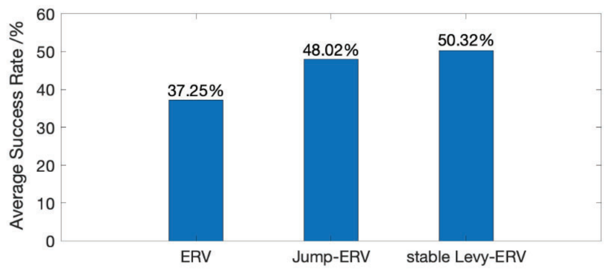


Example: Intentionality analysis for perturbed pointing task in-vehicle

Result for perturbed pointing data from automobile UI systems:





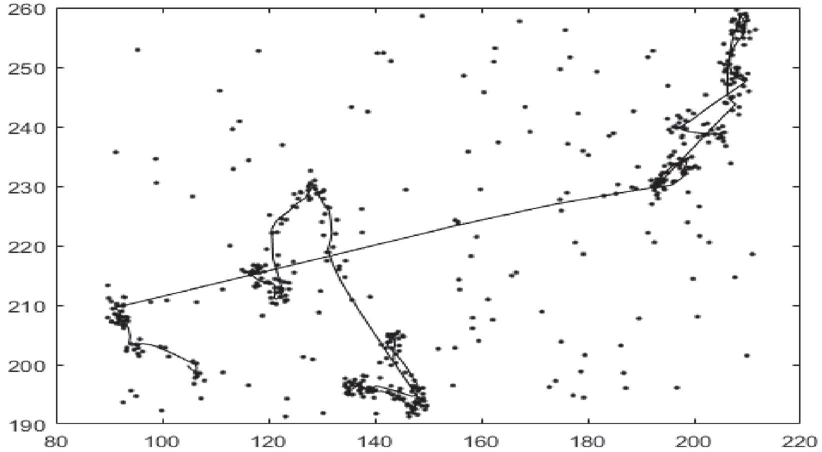


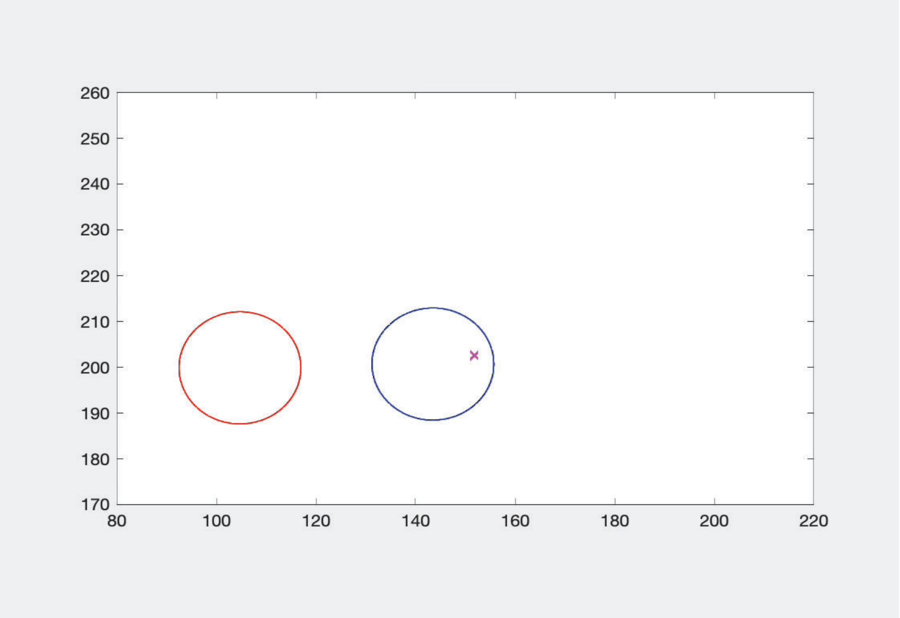
Gan, R., Ahmad, B. I., & Godsill, S. J. (2021). Levy State-Space Models for Tracking and Intent Prediction of Highly Maneuverable Objects. *IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, 57* (4), 2021-2038.

Example: α-stable Lévy State-Space Model for multiple objects in clutter

`Langevin' dynamics:

$$d\dot{X}(t) = -\lambda \dot{X}(t)dt + dW(t)$$





Black lines are ground truth; crosses are measurements;
colored lines are estimates plus 95% confidence ellipse

Conclusion

- A general framework for inference in heavy-tailed non-Gaussian stochastic processes
- Straight forward computations using conditionally Gaussian models and particle filters
- Non-parametric estimation of Q()? – non-parametric Bayes/ ML
- Applications in tracking models, multiple objects, vector Levy processes, bounds on convergence etc.
- Some recent results on Arxiv:

Generalised shot noise representations of stochastic systems driven by non-Gaussian Lévy processes Marcos Tapia Costa, Ioannis Kontoyiannis, Simon Godsill

Point process simulation of generalised hyperbolic Lévy processes. Yaman Kindap, Simon Godsill

Non-Gaussian Process Regression Yaman Kındap, Simon Godsill

A new idea – the non-Gaussian Process (NGP) model

- Here we apply the same Levy process principles to a Gaussian process (GP) model.
- We take a standard GP {W(t)} with covariance function

$$cov(W(t), W(t')) = C(t, t')$$

- We use the same class of `subordinator' jump process {X(t)} to modulate locally the covariance function (`time-change' operation):

$$cov(W(t), W(t')) = C(X(t), X(t'))$$

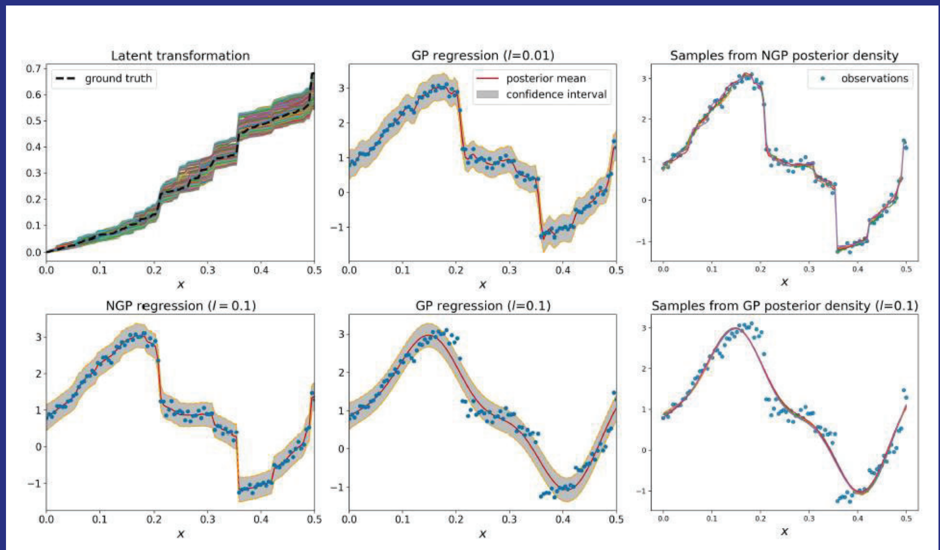
- This allows for non-Gaussian perturbations to the process, but retains once again the structure of a bank of standard GPs, each with a differently modulated covariance function. Examples…

Kindap, Godsill (2022), **Non-Gaussian Process Regression,** arXiv:2209.03117

Preliminary examples

X(t)

W(t)



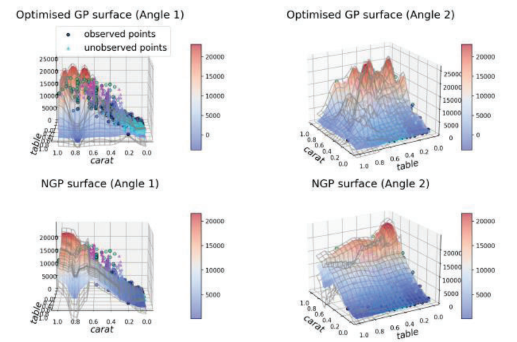


Figure 3: Regression analysis results for NGP and GP models with for the diamond price data set using a TS subordinator. The posterior means are plotted as a surface and the ±3 standard deviation surface is overlaid on the mean as a wireframe plot.

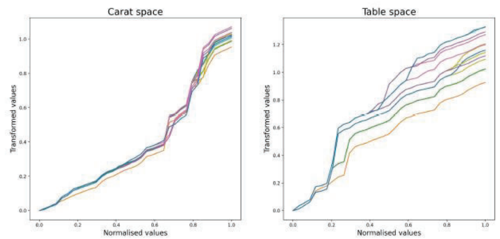


Figure 4: Posterior subordinator samples for a TS subordinator.

Spatial Tracking models using the NGP are currently under development