Inference of Heartbeats in Cardiography Signals using Particle Filtering

Pete Bunch

December 2012

1 Introduction

Sometimes detecting heartbeats in a BCG or ECG signal is easy — one can simply pick the largest peaks at roughly 1 second intervals. Smoothing or differentiating may be required. At other times, more sophisticated algorithms may be necessary, because of low signal-to-noise ratios and inconsistency between heartbeats. Here, a particle filtering approach is described.

2 Preprocessing

Our BCG signals are corrupted by 50Hz mains interference, as well as a DC bias which occasionally makes large jumps. To remove these effects, the discrete-time signal is bandpass filtered between 0.2Hz and 15Hz with a 300-tap symmetric, non-causal FIR filter, and down-sampled to 30Hz. This reduces the processing burden with minimal loss of useful information.

3 Variable Rate Models

We first describe the workings of the most general continuous-time changepoint, or "variable rate" model. Consider a period of time 0 to T, between which observations, $\{y_1 \ldots y_N\}$, are made at times $\{t_1 \ldots t_N\}$. During this period, an unknown number of changepoints, K, occur at times $\{\tau_0, \tau_1 \ldots \tau_K\}$, each with associated changepoint parameters, $\{u_0, u_1 \ldots u_K\}$. The pairs $\{\tau_k, u_k\}$ are the elements of a marked point process (MPP). The latent state is a continuous-time process denoted $x(t)$. Discrete sets containing multiple values over time will be written as, e.g. $y_{1:n} = \{y_1 \ldots y_n\}$.

   The objective for inference will be to estimate the changepoint sequence. This will be denoted as $\theta = \{\tau_{0:K}, u_{0:K}\}$. At a particular time t_n, the sequence up until the current time will is $\theta_n = \{\tau_j, u_j \forall j : 0 \le \tau_j < t_n\}$. It will also be useful to define a variable for the changepoints which occur between two observation times $[t_{n_1}, t_{n_2})$, $\theta_{n_2 \backslash n_1} = \{\tau_j, u_j \forall j : t_{n_1} \le \tau_j < t_{n_2}\}$.

For notational simplicity, the following counting variables are introduced to keep track of the most recent changepoint to have occurred,

$$K(t) = \max(k : \tau_k < t) \tag{1}$$
$$K_n = K(t_n). \tag{2}$$

The changepoint sequence is assumed to be a Markov process,

$$\{\tau_k, u_k\} \sim p(\tau_k, u_k | \tau_{k-1}, u_{k-1}). \tag{3}$$

This density will be constructed such that $P(\tau_k < \tau_{k-1}) = 0$. As in (Whiteley, Johansen & Godsill 2011), a survivor function is defined as the probability that no new changepoint occurs before a given time,

$$S(\tau_k, u_k, t) = P(\tau_{k+1} > t | \tau_k, u_k)$$
$$= 1 - \int_{\tau_k}^{t} p(\xi | \tau_k, u_k) d\xi. \tag{4}$$

It is now possible to write down a prior for the changepoint sequence. This comprises a density term for each changepoint in the sequence and a survivor function term accounting for the probability that no additional changepoints occur within the time interval.

$$p(\theta_n) = S(\tau_{K_n}, u_{K_n}, t_n) p(u_0) p(\tau_0) \prod_{k=1}^{K_n} p(\tau_k, u_k | \tau_{k-1}, u_{k-1}) \tag{5}$$

The existence of such a density for a MPP is addressed in (Jacobsen 2006). The number of changepoints, K_n, is not fixed. If u_k varies over $\mathcal{U}$, and τ_k varies over the real line $\mathbb{R}$, then the space spanned by θ_n^- is the union,

$$\Theta = \bigcup_k \Theta_k \tag{6}$$
$$\Theta_k = \{k\} \times \mathbb{R}^k \times \mathcal{U}^k. \tag{7}$$

This closely resembles the variable-dimension spaces employed by reversible jump Markov chain Monte Carlo (MCMC) algorithms (Green 1995).

3.1 Conditionally Deterministic Models

Next we examine the class of variable rate models in which the state is completely specified by the changepoint sequence, with no additional random components. Such a process is commonly referred to as "piecewise-deterministic", as the latent state follows a deterministic path between changepoints. In this case, it is not necessary to discretise the state — it may be kept as a continuous variable. Thus, the state transition can be described as,

$$x(t) = f(x_{K_n}, u_{K_n}, \tau_{K_n}, t), \tau_{K_n} < t \le \tau_{K_n+1}. \tag{8}$$

By choosing $t = \tau_{K_n+1}$, this equation specifies the state at the next change-point time. Similarly, by choosing $t = t_n$, the state at the observation times may be evaluated — these latter points will be denoted $\hat{x}_n$. To complete the framework, a probabilistic measurement model must be devised for the observation process, $p(y_n|\hat{x}_n)$.

For convenience, we assume that x_0 is known in the following sections. This means that $x(t)$ may be calculated deterministically for all t given θ. This condition is easily relaxed by including x_0 as a random variable in the posterior distribution.

4 The Variable Rate Particle Filter

The variable rate particle filter (VRPF) is described in (Godsill, Vermaak, Ng & Li 2007, Godsill 2007, Whiteley et al. 2011). The objective of the algorithm is to sequentially estimate the posterior distribution of the changepoint sequence, $p(\theta_n|y_{1:n})$, at each time t_n, which is defined on the same space as the prior density (6). This distribution may be expanded using Bayes' rule,

$$p(\theta_n^-|y_{1:n}) \propto p(y_n|\theta_n, y_{1:n-1})$$
$$\times p(\theta_{n\setminus n-1}|\theta_{n-1}^-)p(\theta_{n-1}^-|y_{1:n-1}). \tag{9}$$

The transition term, $p(\theta_{n\setminus n-1}|\theta_{n-1})$, has a similar form to the changepoint prior of (5) (Jacobsen 2006), but an additional survivor function is included to account for the condition that changepoints cannot occur before t_{n-1},

$$p(\theta_{n\setminus n-1}|\theta_{n-1})$$
$$= S(\tau_{K_n}, u_{K_n}, t_n)/S(\tau_{K_{n-1}}, u_{K_{n-1}}, t_{n-1})$$
$$\times \prod_{j:t_{n-1}\leq \tau_j < t_n} p(\tau_j, u_j|\tau_{j-1}, u_{j-1}). \tag{10}$$

Practically, because changepoints will be relatively rare events in most models, it is not likely that more than one new changepoint will occur between t_{n-1} and t_n.

The target distribution of (9) cannot be calculated analytically, but may be approximated numerically, for example, by using a particle filter.

A particle filter is an algorithm for approximating a probability distribution using a set of weighted samples (or "particles") drawn from that distribution using importance sampling (IS). In this case, each particle will be a set of changepoint times and parameters.

$$\hat{p}(\theta_n|y_{1:n}) = \sum_j w_n^{(j)} \delta_{\theta_n^{(j)}}(\theta_n) \tag{11}$$

where $\delta_x(X)$ is a unit probability mass at $X = x$. The particle filter works recursively. At the nth step, a particle, $\theta_{n-1}^{(i)}$, is first resampled from those

approximating the filtering distribution at the $(n-1)$th step, using an appropriately chosen set of proposal weights, $\{v_{n-1}^{(j)}\}$ (where $\sum_j v_{n-1}^{(j)} = 1$),

$$q(\theta_{n-1}) = \sum_j v_{n-1}^{(j)} \delta_{\theta_{n-1}^{(j)}}(\theta_{n-1}). \tag{12}$$

The choice of weights determines the type of resampling used. The simplest choice, $v_{n-1}^{(j)} = 1/\aleph_F$ (where $\aleph_F$ is the number of filter particles) may be achieved by simply omitting this step all together and using the particles of $\hat{p}(\theta_{n-1}|y_{1:n-1})$. This, however, leads to degeneracy of the particle weights over time. Conventional resampling is achieved by using $v_{n-1}^{(j)} = w_{n-1}^{(j)}$. Any other choice results in an auxiliary particle filter (Pitt & Shephard 1999). For further discussion of resampling, see (Cappé, Godsill & Moulines 2007, Doucet & Johansen 2009).

Next, an extension to the changepoint sequence, $\theta_{n\backslash n-1}^{(i)}$, is sampled from an importance distribution, $q(\theta_{n\backslash n-1}|\theta_{n-1}^{(i)}, y_n)$, and concatenated with $\theta_{n-1}^{(i)}$ to create a proposal for $\theta_n^{(i)}$. Finally, the particle is weighted according to the ratio of the target and proposal densities,

$$
\begin{aligned}
w_n^{(i)} &= \frac{p(\theta_n^{-(i)}|y_{1:n})}{q(\theta_n^{(i)})}\\
&\propto \frac{p(y_n|\theta_n^{(i)}, y_{1:n-1})p(\theta_{n\backslash n-1}^{(i)}|\theta_{n-1}^{(i)})p(\theta_{n-1}^{(i)}|y_{1:n-1})}{q(\theta_{n-1}^{(i)})q(\theta_{n\backslash n-1}^{(i)}|\theta_{n-1}^{(i)}, y_n)}\\
&= \frac{w_{n-1}^{(i)}}{v_{n-1}^{(i)}} \times \frac{p(y_n|\theta_n^{(i)}, y_{1:n-1})p(\theta_{n\backslash n-1}^{(i)}|\theta_{n-1}^{(i)})}{q(\theta_{n\backslash n-1}^{(i)}|\theta_{n-1}^{(i)}, y_n)}.
\end{aligned}
\tag{13}
$$

The normalisation may be enforced by scaling the weights so that they sum to 1.

For the most basic "bootstrap" (Gordon, Salmond & Smith 1993) form of the VRPF, $\theta_{n\backslash n-1}$ may be proposed from the prior transition density (10). This can be achieved by sampling new changepoints sequentially from the transition model (3) until one falls after the current time, t_n. This final future changepoint is discarded. (This process can be thought of as sampling an entire future changepoint sequence from t_{n-1} onwards, and then marginalising those which fall after t_n.) The bootstrap proposal leads to the usual simplification of the weight formula,

$$w_n^{(i)} = \frac{w_{n-1}^{(i)}}{v_{n-1}^{(i)}} \times p(y_n|\theta_n^{(i)}, y_{1:n-1}). \tag{14}$$

It only remains to consider the likelihood term required for evaluation of the importance weights, $p(y_n|\theta_n^-, y_{1:n-1})$. The form of this term depends on the particular model under consideration.

4.1 Conditionally Deterministic Likelihoods

When the conditionally deterministic models of section 3.1 are used, the state at observation time t_n is specified by the changepoint sequence θ_n (plus the initial state, x_0), using (8). Thus, the required likelihood term is simply given by,

$$p(y_n|\theta_n, y_{1:n-1}) = p(y_n|\hat{x}_n). \tag{15}$$

This leads to a piecewise-deterministic variable rate particle filter.

In (Morelande & Gordon 2009), an improvement to this algorithm for conditionally-deterministic models was introduced in which the linear parts of the motion parameters are Rao-Blackwellised. This modification will be useful for heartbeat inference. The changepoint parameters are split into analytic and intractable parts,

$$u_k = [u_{k,a}, u_{k,i}]^T. \tag{16}$$

The particle filter is used to estimate only the nonlinear part, while the distribution of the linear part is calculated analytically for each particle, $p(u_{k,a}|u_{1:k,i}, y_{1:n})$. If the evolution of $u_{k,a}$ is linear-Gaussian, a Kalman filter is used; if it has a finite state-space, an HMM filter is used.

Provided the state depends only on the most recent changepoint, the likelihood term for the particle filter is now given by,

$$p(y_n|\theta_n, y_{1:n-1}) = \int p(y_n|\theta_n, u_{a,K_n})p(u_{a,K_n}|y_{1:n-1})du_{a,K_n}. \tag{17}$$

4.2 Improving the Variable Rate Particle Filter

The basic, bootstrap VRPF often performs poorly if changepoints are not obvious until significantly after they occur (as is often the case). For example, in a tracking scenario, if a jump occurs in the acceleration, but only the position is observed, then this change may not be obvious until a number of additional observations have arrived. In this case, the estimation may be improved by the introduction of resample-move (RM) steps (Gilks & Berzuini 2001). In an RM scheme, optional Metropolis-Hastings (MH) moves are conducted to alter the particle states after the importance sampling has taken place. For variable rate models, any one of the previous changepoints, τ_k, or associated parameters, u_k, could be adjusted. Because more observations are available than when the changepoint was first proposed, it may be possible to construct more informed proposals and so move the changepoints towards regions with higher posterior probability. It is even possible to retrospectively add or remove changepoints, using reversible jump MH moves (Green 1995). Variable rate particle filters using RM with piecewise deterministic models are described in (Whiteley et al. 2011, Gilholm & Everett 2008).

Rather than conducting the IS and MH steps separately, it is possible to combine them using the framework of SMC samplers (Del Moral, Doucet &

Jasra 2006). This was suggested in (Whiteley et al. 2011), again for piece-wise deterministic dynamics, but the extension to conditionally linear-Gaussian models is straightforward.

5 BCG Model

We model the cardiography signal as a sequence of similar, irregularly-spaced heartbeat waveforms in noise, using a variable rate framework. The heartbeat start times are represented by a sequence of changepoints, $\{\tau_k\}$, and the corresponding waveforms by a the vector sequence, $\{\mathbf{w}_k\}$. These vectors are sampled at the same rate as the signal. The sampled signal is composed of measurements, $\{y_n\}$, at times $\{t_n\}$ spaced by sampling period T_s.

The signal is modelled as,

$$x(t) = \mathbf{h}(t, \tau_{K_n})^T \mathbf{w}, \tag{18}$$

$\mathbf{h}(t, \tau_{K_n})$ is a vector of interpolation coefficients. Because of the anti-aliasing filter, we can achieve perfect interpolation using,

$$\mathbf{h}(t, \tau_{K_n})_i = \text{sinc}\left(\frac{t - \tau_{K_n} - iT_s}{T_s}\right). \tag{19}$$

However, it might be computationally pragmatic to use something more sparse. Maybe some cubic-splines.

The waveform $\mathbf{w}_k$ is expected to change only slowly over time. We use a Gaussian transition density with a tight covariance matrix,

$$p(\mathbf{w}_k|\mathbf{w}_{k-1}) = \mathcal{N}(\mathbf{w}_k|\mathbf{w}_{k-1}, Q_w). \tag{20}$$

We need a model for the evolution of τ_k. Examining the intervals between successive heartbeats, there is considerable variability. Over a short time, there appears to be a hard lower bound for the beat periods, while the upward variability is much greater. On a longer time scale, this lower bound varies slowly. Based on these observations, we adopt a shifted inverse-gamma prior for the beat periods, where the shift is itself a random variable, denoted Δ_k. The evolution of Δ_k is modelled as a gamma random walk.

Finally, we assume a Gaussian observation model,

$$y_n = \mathcal{N}(y_n|x(t_n), \sigma_y^2). \tag{21}$$

5.1 Interference

The data is intermittently interrupted by various forms of interference, including the subject rolling over, vibrations, electrical effects, etc. We need to be able to detect and account for such phenomena. The models used here distinguish between two sorts of interference: short bursts of high-amplitude noise (clutter) and longer periods of vibration of any amplitude (disturbances). Clutter occurs

on the time-scale of the observations, and disturbances on the time-scale of the heartbeats. Thus, $c_n = \{1, 0\}$ indicates whether the observation at time t_n is clutter or not, and $d_k = \{1, 0\}$ indicates whether the heartbeat from τ_k to τ_{k+1} is disturbed or not. Markovian priors are used. To avoid conflicts, $P(c_n = 1 | d_{K_n} = 1) = 0$. The resulting observation model is,

$$y_n = \begin{cases} \mathcal{N}(y_n | s_n, \sigma_y^2) & c_n = 0, d_{K_n} = 0 \\ \mathcal{C}(y_n | 0, \sigma_D) & d_{K_n} = 1 \\ \mathcal{N}(y_n | 0, \sigma_C^2) & c_n = 1 \end{cases}. \tag{22}$$

$\mathcal{C}(\cdot | x_0, \gamma)$ denotes a Cauchy distribution with mode x_0 and length scale γ.

6 Inference Algorithm

With this model we can do inference using a variable rate particle filter. In addition, because of the linear-Gaussian model assumptions, we can Rao-Blackwellise the waveform variable, $\mathbf{w}_k$, which will give us a huge dimensionality reduction. Hooray! The resulting algorithm will be a similar to those of (Morelande & Gordon 2009) and (Whiteley et al. 2011).

We want our particle filter to estimate the changepoint times and also the parameters,

$$u_k = \begin{bmatrix} \Delta_k \\ d_k \end{bmatrix}. \tag{23}$$

6.1 Interference-free

The target distribution is,

$$p(\theta_n | y_{1:n}). \tag{24}$$

A standard variable rate particle filter can now be applied, using bootstrap proposals, and the Rao-Blackwellisation of $\mathbf{w}_k$,

$$
\begin{aligned}
p(y_n | \theta_n, y_{1:n-1}) &= \int p(y_n | \theta_n, \mathbf{w}_{K_n}) p(\mathbf{w}_{K_n} | \theta_n, y_{1:n-1}) d\mathbf{w}_{K_n} \\
&= \int p(y_n | s(t_n)) p(\mathbf{w}_{K_n} | \theta_n, y_{1:n-1}) d\mathbf{w}_{K_n} \\
&= \int \mathcal{N}(y_n | \mathbf{h}(t, \tau_{K_n})^T \mathbf{w}_{K_n}, \sigma_y^2) \mathcal{N}(\mathbf{w}_{K_n} | \mathbf{m}_{n-1}, \mathbf{P}_{n-1}) d\mathbf{w}_{K_n} (25)
\end{aligned}
$$

A Kalman filter is maintained for each particle to estimate the density over $\mathbf{w}_{K_n}$. If no changepoint occurs between t_{n-1} and t_n, then,

$$
\begin{aligned}
p(\mathbf{w}_{K_n} | \theta_n, y_{1:n}) &\propto p(y_n | \mathbf{w}_{K_n}, \theta_n) p(\mathbf{w}_{K_n} | \theta_n, y_{1:n-1}) \\
&= \mathcal{N}(y_n | \mathbf{h}(t, \tau_{K_n})^T \mathbf{w}_{K_n}, \sigma_y^2) \mathcal{N}(\mathbf{w}_{K_n} | \mathbf{m}_{n-1}, \mathbf{P}_{n-1}). \tag{26}
\end{aligned}
$$

7

Alternatively, if a changepoint does occur between t_{n-1} and t_n, then,

$$p(\mathbf{w}_{K_n}|\theta_n, y_{1:n}) = p(y_n|\mathbf{w}_{K_n}, \theta_n) \int p(\mathbf{w}_{K_n}|\mathbf{w}_{K_n-1})p(\mathbf{w}_{K_n-1}|\theta_n, y_{1:n-1})d\mathbf{w}_{K_n-1}$$

$$= \mathcal{N}(y_n|a_{K_n}b(t, \tau_{K_n})^T\mathbf{w}_{K_n}, \sigma_y^2)$$

$$\times \int \mathcal{N}(\mathbf{w}_{K_n}|\mathbf{w}_{K_n-1}, Q_w)\mathcal{N}(\mathbf{w}_{K_n-1}|\mathbf{m}_{n-1}, \mathbf{P}_{n-1})d\mathbf{w}_{K_n}. \ (27)$$

So everything's Gaussian, and can be calculated in closed form. Jolly good.

6.2 With Clutter and Disturbances

If we propose disturbances from the prior, then these do not affect the derivation, except that $p(y_n|\theta_n, y_{1:n-1}) = \mathcal{C}(y_n|0, \sigma_D)$ if $d_{K_n} = 1$. However, we could propose them from the optimal proposal distribution. This is a topic for further investigation.

The target distribution for the particle filter is now,

$$p(\theta_n, c_{1:n}|y_{1:n}). \tag{28}$$

Note that if we weren't Rao-Blackwellising $\mathbf{w}$ then we could calculate $p(\lambda_n|\theta_n, y_{1:n})$ in closed form, but since we are, we can't, because it would mess up the Kalman filters. We can however, propose exactly from the optimal proposal density,

$$p(c_n|\theta_n, c_{1:n-1}, y_{1:n}) = \frac{p(y_n|\theta_n, c_{1:n}, y_{1:n-1})p(c_n|c_{n-1})}{p(y_n|\theta_n, c_{1:n-1}, y_{1:n-1})}. \tag{29}$$

If we still use a bootstrap proposal for θ_n, then the weight is given by,

$$w_n = p(y_n|\theta_n, c_{1:n-1}, y_{1:n-1})$$

$$= \sum_{c_n} p(y_n|\theta_n, c_{1:n}, y_{1:n-1}), p(c_n|c_{1:n-1}), \tag{30}$$

where $p(y_n|\theta_n, c_{1:n}, y_{1:n-1})$ is given by (25) if $c_n = 0$, and $\mathcal{N}(y_n|0, \sigma_C^2)$ if $c_n = 1$.

6.3 Initialisation

Special consideration is needed for the sampling of τ_0, and u_0, the first changepoint time and parameters, and the choice of $\mathbf{m}_0$ and $\mathbf{P}_0$.

u_0 can be sampled from a prior, $p(u_0$.

τ_0 really needs to be sampled from $p(tau_0|tau_{-1} < 0, \tau_0 > 0)$. We can do this by sampling τ_{-1} from the prior, then sampling the difference $\tau_0 - \tau_{-1}$ from a modified transition density, then sampling τ_0 uniformly between 0 and u_{-1}.

We could use a completely uninformative prior for $\mathbf{w}_0$, but this risks the filter "locking on" to a quasi periodic structure other than the actual heartbeat, e.g. from peak to peak instead of trough to trough. To encourage an accurate lock, we use a medium variance and a crude heartbeat template. This can be as simple as a vector a zeros with an appropriately space $+1$ and -1 corresponding to the major positive and negative peaks.

6.4 Fixed Lag "Smoother"

The particle filter algorithm doesn't work very well. This is because it takes many observations to observe a heartbeat, and often the right particles have been discarded before they are needed. The solution to this is to use a fixed-lag smoother. The target distribution is,

$$p(\theta_{n+L}, \lambda_{1:n+L} | y_{1:n+L}), \tag{31}$$

and we allow changes to $\theta_{n+L\setminus n}$ and $\lambda_{n+1:n+L}$. In addition, we can now move forward in batches rather than running a particle filter at every observation time. The batch size, S, will need to be smaller than the window size, L.

Such a fixed lag smoother uses the SMC sampler algorithm of (Doucet, Briers & Sénécal 2006), which makes it distinct from (although still pretty similar to) the method of (Whiteley et al. 2011).

---

for $n = 0, S, 2S, \ldots$ **do**
    for $i = 1, \ldots, N_F$ **do**
        Select a particle $\{\theta_{n+L-S}, c_{1:n+L-S}\}^{(i)}$ with probability $w_{n-S}^{(i)}$
        Sample $\theta_{n+L\setminus n}^{'(i)} \sim p(\cdot|\theta_n^{(i)})$
        for l = 1,...,L **do**
            Sample $c_{n+l}^{'(i)} \sim p(\cdot|\theta_{n+L}^{(i)}, c_{n+l-1}^{(i)}, y_{1:n})$
        end for
        $w_n^{(i)} = p(y_{n+1:n+L}|\theta_{n+L}^{'(i)}, y_{1:n})/p(y_{n+1:n+L-S}|\theta_{n+L-S}, y_{1:n})$
    end for
end for

---

7 Graphs

Clear, difficult, and very difficult signal
    Inference output on a sample of each.

References

Cappé, O., Godsill, S. & Moulines, E. (2007). An overview of existing methods and recent advances in sequential Monte Carlo, *Proceedings of the IEEE* **95**: 899–924.

Del Moral, P., Doucet, A. & Jasra, A. (2006). Sequential Monte Carlo samplers, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**: 411–436.

Doucet, A., Briers, M. & Sénécal, S. (2006). Efficient block sampling strategies for sequential Monte Carlo, *Journal of Computational and Graphical Statistics* **15**(3): 693–711.

Doucet, A. & Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later, *in* D. Crisan & B. Rozovsky (eds), *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press.

Gilholm, K. & Everett, N. O. (2008). A novel algorithm for tracking targets using manoeuvre models, *IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, pp. 143–148.

Gilks, W. R. & Berzuini, C. (2001). Following a moving target – Monte Carlo inference for dynamic Bayesian models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **63**: 127–146.

Godsill, S. (2007). Particle filters for continuous-time jump models in tracking applications, *ESAIM: Proc.* **19**: 39–52.

Godsill, S., Vermaak, J., Ng, W. & Li, J. (2007). Models and algorithms for tracking of maneuvering objects using variable rate particle filters, *Proceedings of the IEEE* **95**: 925–952.

Gordon, N., Salmond, D. & Smith, A. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *Radar and Signal Processing, IEE Proceedings F* **140**: 107–113.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika* **82**: 711.

Jacobsen, M. (2006). *Point process theory and applications: Marked point and piecewise deterministic processes*, Birkhauser.

Morelande, M. & Gordon, N. (2009). Rao-Blackwellised variable rate particle filters, *Proc. 12th Int. Conf. Information Fusion*, pp. 1 –8.

Pitt, M. K. & Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters, *Journal of the American Statistical Association* **94**: 590–599.

Whiteley, N., Johansen, A. M. & Godsill, S. (2011). Monte Carlo filtering of piecewise deterministic processes, *Journal of Computational and Graphical Statistics* **20**: 119–139.