# Supplementary material for 'MCMC for continuous-time discrete-state systems'

**Vinayak Rao**
Gatsby Computational Neuroscience Unit
University College London
vrao@gatsby.ucl.ac.uk

**Yee Whye Teh**
Gatsby Computational Neuroscience Unit
University College London
ywteh@gatsby.ucl.ac.uk

**Proposition 1** The path $(V, L, W)$ returned by algorithm 1 corresponds to a sample from the semi-Markov process parametrized by $(\pi_0, A)$.

*Proof.* Without any loss of generality, assume that the system has just entered state $s \in \mathcal{S}$ at time $0$.

Suppose that $t$ is the time of $n$th candidate jump, so that there were $n-1$ rejected transitions on the interval $[0, t]$. Let these occur at times $(w_1, w_2, \ldots, w_{n-1})$, with $t = w_n$. Recalling that these were generated from the hazard function $B_s(t)$, and letting $[n-1]$ represent the set of integers $\{1, \cdots, n-1\}$, we have:

$$
\begin{aligned}
&P((w_1, \ \ldots w_n), \{v_i = s, l_i = (w_i - w_0)\ \forall i \in [n-1]\}, v_n = s', l_n = 0 | w_0, v_0 = s) \\
&= \left(\prod_{k=1}^{n-1} B_s(l_k) \exp\left(-\int_{l_{k-1}}^{l_k} B_s(\tau)\mathrm{d}\tau\right)\left(1 - \frac{A_s(l_k)}{B_s(l_k)}\right)\right) \qquad (1)\\
&\quad \left(B_s(l_{n-1} + \Delta w_{n-1}) \exp\left(-\int_{l_{n-1}}^{l_{n-1}+\Delta w_{n-1}} B_s(\tau)\mathrm{d}\tau\right)\left(\frac{A_{ss'}(l_{n-1} + \Delta w_{n-1})}{B_s(l_{n-1} + \Delta w_{n-1})}\right)\right) \\
&= \exp\left(-\int_0^{l_{n-1}+\Delta w_{n-1}} B_s(\tau)\mathrm{d}\tau\right)\left(\prod_{k=1}^{n-1}(B_s(l_k) - A_s(l_k))\right) A_{ss'}(l_{n-1} + \Delta w_{n-1}) \qquad (2)
\end{aligned}
$$

Integrating out $w_1$ to $w_{n-1}$ (and thus $l_1$ to $l_{n-1}$), we have

$$
\begin{aligned}
&P(w_n = t, \ \{v_i = s\ \forall i \in [n-1]\}, v_n = s', l_n = 0 | w_0 = 0, v_0 = s) \qquad (3)\\
&= \exp\left(-\int_0^t B_s(\tau)\mathrm{d}\tau\right) A_{ss'}(w_n) \\
&\quad \left(\int_0^t \int_{l_1}^t \cdots \int_{l_{n-2}}^t \prod_{k=1}^{n-1} (B_s(l_k) - A_s(l_k)\mathrm{d}l_k)\right) \\
&= A_{ss'}(t) \exp\left(-\int_0^t B_s(\tau)\mathrm{d}\tau\right) \frac{1}{(n-1)!}\left(\int_0^t \mathrm{d}\tau\,(B_s(\tau) - A_s(\tau))\right)^{n-1} \qquad (4)
\end{aligned}
$$

The expression above gives the probability of transitioning from state $s$ to $s'$ after a wait of $t$ time units, with $n-1$ rejected candidate jumps. Summing out $n-1$, we get the transition probability. Thus,

$$
\begin{aligned}
&P(s_{next} = s', t_{next} = t | s_{curr} = s, t_{curr} = 0) \\
&= A_{ss'}(t) \exp\left(-\int_0^t B_s(\tau)\mathrm{d}\tau\right) \sum_{n-1=0}^{\infty} \frac{1}{(n-1)!}\left(\int_0^t \mathrm{d}\tau\,(B_s(\tau) - A_s(\tau))\right)^{n-1} \\
&= A_{ss'}(t) \exp\left(-\int_0^t A_s(\tau)\mathrm{d}\tau\right) \qquad (5)
\end{aligned}
$$

This is the desired result. $\square$

**Proposition 2** Conditioned on a trajectory $(S, T)$ of the sMJP, the thinned events $\tilde{W}$ are distributed as a Poisson process with intensity $B(t) - A(t)$.

*Proof.* We will consider the interval of time $[t_i, t_{i+1}]$, so that the sMJP entered state $s_i$ at time $t_i$, and remained there until time $t_{i+1}$, when it transitioned to state $s_{i+1}$. Exploiting the independence properties of the sMJP and the Poisson process, we only need to consider resampling thinned events on this interval. Call this set of thinned events $\tilde{W} \equiv \{\tilde{w}_1, \cdots, \tilde{w}_{n-1}\} \in [t_i, t_{i+1}]$, and call the corresponding set of labels $\tilde{V} \equiv \{\tilde{v}_1, \cdots, \tilde{v}_{n-1}\}$ and $\tilde{L} \equiv \{\tilde{l}_1, \cdots, \tilde{l}_{n-1}\}$ (to avoid notational clutter, we do not indicate that $\tilde{W}$ and $\tilde{L}$ are actually restrictions to $[t_i, t_{i+1}]$). Observe that each element of $\tilde{v}_j \in \tilde{V}$ equals $s_i$, while each element $\tilde{l}_j \in \tilde{L}$ equals $\tilde{w}_j - t_i$. We write this as $\tilde{V} = s_i$ and $\tilde{L} = \tilde{W} - t_i$. Then, by Bayes rule, we have

$$\begin{aligned}
&P(\tilde{W}, \tilde{V} = s_i, \tilde{L} = \tilde{W} - t_i | s_i, t_i, s_{i+1}, t_{i+1}) \qquad (6)\\
&\quad = \frac{P(\tilde{W}, \tilde{V} = s_i, \tilde{L} = \tilde{W} - t_i, v_n = s_{i+1}, w_n = t_{i+1}, l_n = 0 | v_0 = s_i, w_0 = t_i, l_0 = 0)}{P(s_{i+1}, t_{i+1} | s_i, t_i)}\\
&\quad = \frac{\exp\left(-\int_{t_i}^{t_{i+1}} B(\tau)\mathrm{d}\tau\right)\left(\prod_{k=1}^{n-1}\left(B(\tilde{w}_k) - A(\tilde{w}_k)\right)\right) A_{s_i s_{i+1}}(t_{i+1} - t_i)}{A_{s_i s_{i+1}}(t_{i+1} - t_i)\exp\left(-\int_{t_i}^{t_{i+1}} A(\tau)\mathrm{d}\tau\right)}\\
&\quad = \exp\left(-\int_{t_i}^{t_{i+1}} B(\tau) - A(\tau)\mathrm{d}\tau\right)\left(\prod_{k=1}^{n-1}\left(B(v_k) - A(v_k)\right)\right)
\end{aligned}$$

This is just the density of a Poisson process on $(t_i, t_{i+1})$ with intensity $(B(t) - A(t))$, which is what we set out to prove. $\square$