Slice Sampling for General Completely Random Measures – Supplement

Peiyuan Zhu **Alexandre Bouchard-Côté** **Trevor Campbell**
Department of Statistics
University of British Columbia
Vancouver, BC V6T 1Z4

We now show that the proposed slice sampler defines a valid Markov chain Monte Carlo algorithm (Theorem 0.2). In particular, (1) the exact posterior π is the invariant distribution of the Markov chain, and (2) that a law of large numbers holds: for any measurable function Φ and initial state S_0, the sequence of states $S_1, S_2, \dots$ produced by the slice sampler satisfies

$$\frac{1}{T}\sum_{t=1}^{T}\Phi(S_t) \overset{a.s.}{\to} \mathbb{E}_\pi\left[\Phi(S)\right].$$

We start with some basic notation. Let $\mathcal{S}$ be a set endowed with a σ-algebra $\mathcal{B}$, and let π be a target probability distribution on $\mathcal{S}$. A *Markov kernel* $\kappa : \mathcal{S}\times\mathcal{B}\to[0,1]$ satisfies two properties: (1) for each $B\in\mathcal{B}$, $\kappa(\cdot,B):\mathcal{S}\to[0,1]$ is a measurable function, and (2) for each $s\in\mathcal{S}$, $\kappa(s,\cdot)$ is a probability measure. $\kappa(s,B)$ can be thought of as the probability of transitioning to any state $s'\in B\subseteq\mathcal{S}$ in a single jump starting from a particular state $s\in\mathcal{S}$. Given two Markov kernels κ_1, κ_2, define the *composition* $\kappa_1\circ\kappa_2$ of the kernels—another Markov kernel—via

$$(\kappa_1\circ\kappa_2)(s,B) = \int \kappa_1(s',B)\kappa_2(s,ds').$$

As with a single kernel, the composition $(\kappa_1\circ\kappa_2)(s,B)$ can be thought of as the probability of transitioning to any state $s'\in B\subseteq\mathcal{S}$ after two jumps—first via κ_2, then via κ_1—starting from a particular state $s\in\mathcal{S}$.

One of the key conditions for a kernel κ to create a Markov chain Monte Carlo scheme for a target distribution π is π-*invariance*: if one samples $s\sim\pi$, and then simulates a transition $s'\sim\kappa(s,\cdot)$, we require that $s'\sim\pi$. In other words, for any measurable set B,

$$\int \kappa(s,B)\pi(ds) = \pi(B).$$

We use the following results in Lemma 0.1 to analyze the π-invariance of the proposed slice sampler for the posterior distribution π.

Lemma 0.1. *Let $(\kappa_j)_{j=1}^\infty$ be Markov kernels, and suppose $\mathcal{S}$ can be written as a countable partition $\mathcal{S} = \bigcup_j B_j$, $i\neq j \implies B_i\cap B_j=\emptyset$ of sets of nonzero measure $\pi(B_j)>0$.*

1. *If the κ_j are all π-invariant, and*

$$\kappa(s,B) = \lim_{J\to\infty}\left(\kappa_J\circ\cdots\circ\kappa_1\right)(s,B)$$

exists pointwise for $s\in\mathcal{S}$ and $B\in\mathcal{B}$, then κ is a π-invariant Markov kernel.

2. *If each κ_j is π_j-invariant, where*

$$\pi_j(B) = \frac{\pi(B\cap B_j)}{\pi(B_j)},$$

then

$$\kappa(s,B) = \sum_{j=1}^\infty \mathbb{1}\left[s\in B_j\right]\kappa_j(s,B)$$

is π-invariant.

Proof. For 1,

$$\begin{aligned}&\int \kappa(s,B)\pi(ds)\\ &= \int \lim_{J\to\infty}\left(\kappa_J\circ\cdots\circ\kappa_1\right)(s,B)\pi(ds)\\ &= \lim_{J\to\infty}\int\left(\kappa_J\circ\cdots\circ\kappa_1\right)(s,B)\pi(ds)\\ &= \lim_{J\to\infty}\pi(B) = \pi(B),\end{aligned}$$

where we use the fact that the finite composition of π-invariant kernels is π-invariant e.g. by [1, p. 49], and Lebesgue dominated convergence to swap the limit and integral. For 2,

$$\begin{aligned}
&\int \kappa(s,B)\pi(ds) \\
&=\sum_{j=1}^{\infty}\int \mathbb{1}\left[s\in B_j\right]\kappa_j(s,B)\pi(ds) \\
&=\sum_{j=1}^{\infty}\pi(B_j)\int \kappa_j(s,B)\frac{\mathbb{1}\left[s\in B_j\right]\pi(ds)}{\pi(B_j)} \\
&=\sum_{j=1}^{\infty}\pi(B_j)\pi_j(B)=\sum_{j=1}^{\infty}\pi(B_j\cap B)=\pi(B),
\end{aligned}$$

where we again use Lebesgue dominated convergence to swap the infinite series and integral. □

Each iteration of the slice sampler can be written as the kernel composition

$$\kappa=\kappa_{\Gamma,V}^{\text{exp}}\circ\kappa_{\Gamma,V}\circ\kappa_X\circ\kappa_\psi\circ\kappa_U.$$

The kernels $\kappa_X,\kappa_\psi,\kappa_U$ are the full conditional (i.e., Gibbs) kernels for variables X,ψ,U; the kernel $\kappa_{\Gamma,V}$ (substep 1 in the main text) is the composition of the full conditional of Γ_k,V_k for all $k\in\mathbb{N}$; standard results [1, p. 79] guarantee that each of these is π-invariant, and so their composition is π-invariant by Lemma 0.1. Note that although all of these kernels involve theoretically simulating infinitely many values, in practice this is unnecessary: truncation by U makes simulating X_{nk} and ψ_k for $k>K$ unnecessary, and we will see that the final kernel $\kappa_{\Gamma,V}^{\text{exp}}$ overwrites changes to Γ_k,V_k for $k\geq K_{\text{prev}}$, implying that the full conditional step only needs to be run for $k<K_{\text{prev}}$.

The only remaining kernel is $\kappa_{\Gamma,V}^{\text{exp}}$, which corresponds to substep 2 in the main text. This kernel samples $(\Gamma_k,V_k)_{k=K_{\text{prev}}}^{\infty}$ from their full conditional. Denote κ_j^{exp} to be the kernel that samples $(\Gamma_k,V_k)_{k=j}^{\infty}$ from their full conditional; then

$$\kappa_{\Gamma,V}^{\text{exp}}=\sum_{j=0}^{\infty}\mathbb{1}\left[K_{\text{prev}}=j\right]\kappa_j^{\text{exp}}.$$

By Lemma 0.1, we just need to show that each κ_j^{exp} is π_j-invariant, where π_j is the posterior conditioned on $K_{\text{prev}}=j$, which follows from the fact that π_j is a Gibbs kernel.

We have now shown that the Markov kernel created by the slice sampler in the main text is π-invariant. We now complete the final result in Theorem 0.2.

Theorem 0.2. *If $f>0$ and $h>0$, then for any measurable function Φ and any initial random state S_0, the sequence of states $S_1,S_2,\ldots$ produced by κ satisfies*

$$\frac{1}{T}\sum_{t=1}^{T}\Phi(S_t)\overset{a.s.}{\to}\mathbb{E}_\pi\left[\Phi(S)\right].$$

Proof. We first establish φ-irreducibility: let us set φ to the posterior distribution, let $s=(v,\gamma,x,\psi,u)$ denote an initial state, and B, a target set of configurations with positive posterior probability. It may not be possible to go from s to B in one application of κ as the current configuration of the matrix x constrains what values u can take. However this obstacle disappears by considering paths obtained by two applications of κ and visiting an intermediate state where every entry in the matrix x is set to zero. To formalize this, let $B_0=\{(v,\gamma,x,\psi,u):x_{nk}=0\ \forall n,k\}$. Then

$$\begin{aligned}
\kappa^2(s,B)&=\int\kappa(s,\mathrm{d}s')\kappa(s',B) \\
&\geq\int\mu(\mathrm{d}s')\kappa(s',B)
\end{aligned}$$

where $\mu(A)=\kappa(s,A\cap B_0)$. Using the fact that ξ is monotonically decreasing, our assumption that f and h are strictly positive, we obtain from the full conditional of X derived in the paper that μ is a strictly positive measure on B_0. Moreover, using again the same assumptions, straightforward checks on each full conditional derived in the paper shows that provided $s\in B_0$, the function $\kappa(s',B)$ is positive.

Having established φ-irreducibility, Harris recurrence follows from [2, Cor. 13] since κ is a deterministic alternation of Gibbs kernels. Therefore the law of large number follows by [3, Thm. 17.0.1, 17.1.6]. □

References

[1] Charles J Geyer. Markov chain Monte Carlo Lecture Notes, 1998.

[2] Gareth O. Roberts and Jeffrey S. Rosenthal. Harris recurrence of Metropolis-within-Gibbs and trans-dimensional Markov chains. *The Annals of Applied Probability*, 16(4): 2123–2139, 2006. ISSN 1050-5164.

[3] Sean P. Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability*. Communications and Control Engineering. Springer-Verlag, London, 1993. ISBN 978-1-4471-3269-1.