Biometrika (2016), **xx**, x, *pp*. 1–3

Supplementary material for 'On the random- versus systematic-scan sampler choice'

BY C. ANDRIEU

School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.

c.andrieu@bristol.ac.uk

SUMMARY

This supplement contains proofs of Proposition 1 and Lemma 2, as well as three lemmas.

Proof of Proposition 1. Let $f \in L_0^2(\mathsf{X},\pi)$. Then for $n \geq 1$,

$$E_\pi\{\textstyle\sum_{i=0}^{n-1} f(X_i) \sum_{j=0}^{n-1} f(X_j)\} = nE_\pi\left\{f^2(X_0)\right\} + 2 \sum_{0\leq i<j\leq n-1} E_\pi\{f(X_i)f(X_j)\},$$

and we rewrite the second term as

$$\begin{aligned}\sum_{0\leq i<j\leq n-1} E_\pi\{f(X_i)f(X_j)\} &= \sum_{0\leq i<j\leq n-1} E_\pi\{f(X_i)\Pi_{\sigma^{i:j-1}(1)} f(X_i)\} \\ &= \sum_{0\leq i<j\leq n-1} \left\langle f, \Pi_{\sigma^{i:j-1}(1)} f\right\rangle_\pi \\ &= \sum_{0\leq i<j\leq n-1} \left\langle f, \Pi_{\sigma^{i:j-i-1+i}(1)} f\right\rangle_\pi \\ &= \sum_{0\leq i<n-1} \sum_{m=1}^{n-1-i} \left\langle f, \Pi_{\sigma^{i:m-1+i}(1)} f\right\rangle_\pi .\end{aligned}$$

Now we have

$$\begin{aligned} n^{-1} &\sum_{0<i<n-1} \sum_{m=1}^{n-1-i} \left\langle f, \Pi_{\sigma^{i:m-1+i}(1)} f\right\rangle_\pi \\ &= \sum_{q=1}^{k} \frac{\lfloor (n-2-q)/k \rfloor}{n} \frac{1}{\lfloor (n-2-q)/k \rfloor} \sum_{p\in P_{k,q}} \sum_{m=1}^{n-1-pk-q} \left\langle f, \Pi_{\sigma^{0:m-1}(q)} f\right\rangle_\pi ,\end{aligned}$$

where $P_{k,q} = \{m \in \mathbb{N} : 0 < mk + q < n-1\}$. The term $i = 0$ is treated similarly. We conclude by letting $n \to \infty$ and using a Cesàro sum argument for each $q \in \{1,\ldots,k\}$. □

We let $\bar{\mathrm{pr}}(\cdot)$ denote the probability distribution of the Markov chain defined by $\mathcal{T}$.

LEMMA 4. *For* $m \geq 0$ *and* $A_1,\ldots,A_m \in \mathcal{X}^m$,

$$\bar{\mathrm{pr}}\left(X_0^{\sigma^0(1)} \in A_0, X_1^{\sigma^1(1)} \in A_1, \ldots, X_m^{\sigma^m(1)} \in A_m\right) = \mathrm{pr}\left(X_0 \in A_0, X_1 \in A_1, \ldots, X_m \in A_m\right) .$$

Proof. By construction for $A \in \mathcal{X}$, $\bar{\text{pr}}\left(X_0^{\sigma^0(1)} \in A\right) = \pi\big(A\big) = \text{pr}\left(X_0 \in A\right)$ and for $i \geq 1$, component $\sigma^i(1)$ is generated by kernel $\Pi_{\sigma^{i-1}(1)}$ and with

$$\mathcal{G}_i = \sigma\left\{\big(X_m^{(1)}, \ldots, X_m^{(k)}\big), 0 \leq m \leq i\right\}$$

we have

$$\bar{\text{pr}}\left(X_i^{\sigma^i(1)} \in A \mid \mathcal{G}_{i-1}\right) = \Pi_{\sigma^{i-1}(1)}\left(X_{i-1}^{\sigma^{i-1}(1)}, A\right)$$

from which the result follows. □

Proof of Lemma 2. We have

$$\begin{aligned}
\big\langle f, \big[I - \lambda \Xi\big]^{-1} f\big\rangle &= \big\langle \big(I - \lambda \Xi\big)\big(I - \lambda \Xi\big)^{-1} f, \big(I - \lambda \Xi\big)^{-1} f\big\rangle \\
&= \big\langle \big(I - \lambda S\big)\big(I - \lambda \Xi\big)^{-1} f, \big(I - \lambda \Xi\big)^{-1} f\big\rangle \\
&= \sup_{h \in \mathcal{H}} 2\big\langle \big(I - \lambda \Xi\big)^{-1} f, h\big\rangle - \big\langle h, \big(I - \lambda S\big)^{-1} h\big\rangle \\
&= \sup_{h \in \mathcal{H}} 2\big\langle f, \big(I - \lambda \Xi^*\big)^{-1} h\big\rangle - \big\langle h, \big(I - \lambda S\big)^{-1} h\big\rangle \\
&= \sup_{g \in \mathcal{H}} 2\big\langle f, g\big\rangle - \big\langle \big(I - \lambda \Xi^*\big) g, \big(I - \lambda S\big)^{-1}\big(I - \lambda \Xi^*\big) g\big\rangle \\
&= \sup_{g \in \mathcal{H}} 2\big\langle f, g\big\rangle - \big\langle g, \big(I - \lambda S\big) g\big\rangle - \big\langle \lambda A g, \big(I - \lambda S\big)^{-1} \lambda A g\big\rangle
\end{aligned}$$

where we have used that $\Xi = S + A$, $\Xi^* = S - A$, that for any $g \in \mathcal{H}$, $\big\langle g, Ag\big\rangle = -\big\langle Ag, g\big\rangle = 0$, Lemma 5 for the self-adjoint operator $\big(I - \lambda S\big)$, $\big\{\big(I - \lambda \Xi\big)^{-1}\big\}^* = \big(I - \lambda \Xi^*\big)^{-1}$, set $g = \big(I - \lambda \Xi^*\big)^{-1} h$ and again used the property $\big\langle g, Ag\big\rangle = 0$. From Lemma 5 the supremum on the third line is attained for $\hat{h} = \big(I - \lambda S\big)\big(I - \lambda \Xi\big)^{-1} f$, which translates into $\hat{g} = \big(I - \lambda \Xi^*\big)^{-1} \hat{h}$ on the last line. Consequently using again Lemma 5 for the operator $I - \lambda S$ we deduce that

$$\big\langle f, \big(I - \lambda \Xi\big)^{-1} f\big\rangle \leq \big\langle f, \big(I - \lambda S\big)^{-1} f\big\rangle - \lambda^2 \big\langle A\hat{g}, \big(I - \lambda S\big)^{-1} A\hat{g}\big\rangle.$$

□

The following provides a useful variational representation of the quadratic form of the inverse of a positive self-adjoint operators, attributed to Bellman, and used for example by Caracciolo et al. (1990).

LEMMA 5. *Let Ξ be a self-adjoint operator on a Hilbert space $\mathcal{H}$, satisfying $\langle f, \Xi f\rangle \geq 0$ for all $f \in \mathcal{H}$ and such that the inverse Ξ^{-1} exists. Then*

$$\left\langle f, \Xi^{-1} f\right\rangle = \sup_{g \in \mathcal{H}} \left(2\left\langle f, g\right\rangle - \left\langle g, \Xi g\right\rangle\right),$$

where the supremum is attained with $g = \Xi^{-1} f$.

The operator T is not self-adjoint, but one can easily determine the expression for its adjoint T^* in terms of $\mathfrak{S}$ and Δ or T. Visualizing T as a block diagonal matrix may be helpful.

LEMMA 6. *Let T, Δ and $\mathfrak{S}$ be as defined in §2 and §3. Then*

1. *the adjoint of $\mathfrak{S}$ is $\mathfrak{S}^* = \mathfrak{S}^{-1}$,*

2. $\Delta^* = \Delta$, *that is* Δ *is self-adjoint,*
3. $T = \Delta \circ \mathfrak{S}$ *and the adjoint of* T *is* $T^* = \mathfrak{S}^{-1} \circ \Delta = \mathfrak{S}^{-1} \circ T \circ \mathfrak{S}^{-1}$.

Proof. The first statement follows from

$$\langle \varphi, \mathfrak{S}\varphi \rangle = \sum_{i=1}^{k} \langle \varphi_i, \varphi_{\sigma(i)} \rangle_\pi = \sum_{j=1}^{k} \langle \varphi_{\sigma^{-1}(j)}, \varphi_j \rangle_\pi = \langle \mathfrak{S}^{-1}\varphi, \varphi \rangle \ .$$

The second statement is direct while the third statement follows from the general fact that $T^* = \mathfrak{S}^* \circ \Delta^*$ followed by an application of the first two statements of the lemma. □

BIBLIOGRAPHY

CARACCIOLO, S., PELISSETTO, A. & SOKAL, A. D. (1990). Nonlocal Monte Carlo algorithm for self-avoiding walks with fixed endpoints. *J. Stat. Phys.* **60**, 1–53.