



Andrieu, C. (2016). On random- and systematic-scan samplers. *Biometrika*, 103(3), 719-726. <https://doi.org/10.1093/biomet/asw019>

Peer reviewed version

Link to published version (if available):
[10.1093/biomet/asw019](https://doi.org/10.1093/biomet/asw019)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Oxford University Press at <http://biomet.oxfordjournals.org/content/103/3.toc>.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

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

5

SUMMARY

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

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

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

and we rewrite the second term as

$$\begin{aligned} \sum_{0 \leq i < j \leq n-1} E_\pi \{ f(X_i) f(X_j) \} &= \sum_{0 \leq i < j \leq n-1} E_\pi \{ f(X_i) \Pi_{\sigma^{i:j-1}(1)} f(X_j) \} \\ &= \sum_{0 \leq i < j \leq n-1} \langle f, \Pi_{\sigma^{i:j-1}(1)} f \rangle_\pi \\ &= \sum_{0 \leq i < j \leq n-1} \langle f, \Pi_{\sigma^{i:j-i-1+i}(1)} f \rangle_\pi \\ &= \sum_{0 \leq i < n-1} \sum_{m=1}^{n-1-i} \langle f, \Pi_{\sigma^{i:m-1+i}(1)} f \rangle_\pi. \end{aligned} \tag{10}$$

Now we have

$$\begin{aligned} n^{-1} \sum_{0 < i < n-1} \sum_{m=1}^{n-1-i} \langle f, \Pi_{\sigma^{i:m-1+i}(1)} f \rangle_\pi \\ = \sum_{q=1}^k \frac{\lfloor (n-2-q)/k \rfloor}{n} \frac{1}{\lfloor (n-2-q)/k \rfloor} \sum_{p \in P_{k,q}} \sum_{m=1}^{n-1-pk-q} \langle f, \Pi_{\sigma^{0:m-1}(q)} f \rangle_\pi, \end{aligned} \tag{15}$$

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

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

20

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

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

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

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

we have

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

25 from which the result follows. □

Proof of Lemma 2. We have

$$\begin{aligned} \langle f, [I - \lambda \Xi]^{-1} f \rangle &= \langle (I - \lambda \Xi)(I - \lambda \Xi)^{-1} f, (I - \lambda \Xi)^{-1} f \rangle \\ &= \langle (I - \lambda S)(I - \lambda \Xi)^{-1} f, (I - \lambda \Xi)^{-1} f \rangle \\ &= \sup_{h \in \mathcal{H}} 2 \langle (I - \lambda \Xi)^{-1} f, h \rangle - \langle h, (I - \lambda S)^{-1} h \rangle \\ 30 \quad &= \sup_{h \in \mathcal{H}} 2 \langle f, (I - \lambda \Xi^*)^{-1} h \rangle - \langle h, (I - \lambda S)^{-1} h \rangle \\ &= \sup_{g \in \mathcal{H}} 2 \langle f, g \rangle - \langle (I - \lambda \Xi^*) g, (I - \lambda S)^{-1} (I - \lambda \Xi^*) g \rangle \\ &= \sup_{g \in \mathcal{H}} 2 \langle f, g \rangle - \langle g, (I - \lambda S) g \rangle - \langle \lambda A g, (I - \lambda S)^{-1} \lambda A g \rangle \end{aligned}$$

where we have used that $\Xi = S + A$, $\Xi^* = S - A$, that for any $g \in \mathcal{H}$, $\langle g, A g \rangle = -\langle A g, g \rangle = 0$, Lemma 5 for the self-adjoint operator $(I - \lambda S)$, $\{(I - \lambda \Xi)^{-1}\}^* = (I - \lambda \Xi^*)^{-1}$, set $g = (I - \lambda \Xi^*)^{-1} h$ and again used the property $\langle g, A g \rangle = 0$. From Lemma 5 the supremum on the 35 third line is attained for $\hat{h} = (I - \lambda S)(I - \lambda \Xi)^{-1} f$, which translates into $\hat{g} = (I - \lambda \Xi^*)^{-1} \hat{h}$ on the last line. Consequently using again Lemma 5 for the operator $I - \lambda S$ we deduce that

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

□

The following provides a useful variational representation of the quadratic form of the inverse 40 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*

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

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

45 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 $T^* = \mathfrak{S}^{-1} \circ \Delta = \mathfrak{S}^{-1} \circ T \circ \mathfrak{S}^{-1}$.

50

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. \square

BIBLIOGRAPHY

55

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