ADAPTIVE SIMPLE MONTE CARLO

FRED J. HICKERNELL AND YUEWEI LIU

ABSTRACT. We attempt a

1. INTRODUCTION

An *integration lattice*, L, is a subset of $\mathbb{R}^d$ and a superset of $\mathbb{Z}^d$ that has no accumulation point and that is closed under addition and subtraction modulo 1, i.e.,

$$\boldsymbol{x}, \boldsymbol{t} \in L \implies \boldsymbol{x} \pm \boldsymbol{t} \in L.$$

A *shifted* integration lattice, $L_{\boldsymbol{\Delta}}$ is defined as $L_{\boldsymbol{\Delta}} = \{\boldsymbol{x} + \boldsymbol{\Delta} \pmod 1 : \boldsymbol{x} \in L\}$ for some integration lattice L and some fixed $\boldsymbol{\Delta} \in [0,1)^d$. The shifted lattice is actually a coset of the lattice. The node set of an integration lattice is defined as those lattice points falling inside the unit cube, i.e., $P = L \cap [0,1)^d$. Likewise, the node set of a shifted integration lattice is $P_{\boldsymbol{\Delta}} = L_{\boldsymbol{\Delta}} \cap [0,1)^d$. Equivalently, the nodesets of the integration lattice and shifted integration lattice may be defined respectively by the conditions

$$\boldsymbol{x}, \boldsymbol{t} \in P \implies \boldsymbol{x} \pm \boldsymbol{t} \pmod 1 \in P, \qquad P_{\boldsymbol{\Delta}} = \{\boldsymbol{x} + \boldsymbol{\Delta} \pmod 1 : \boldsymbol{x} \in P\}.$$

The dual lattice, $L^\perp$, corresponding to an integration lattice, L, is a subset of $\mathbb{Z}^d$ satisfying

$$\boldsymbol{k} \in L^\perp \iff \boldsymbol{k}^T \boldsymbol{x} \in \mathbb{Z} \quad \forall \boldsymbol{x} \in L.$$

Equivalently, one may replace $\boldsymbol{x} \in L$ by $\boldsymbol{x}$ in the node set of L.

Let $\mathcal{F}$ denote the Banach space of real-valued functions defined on $[0,1)^d$ with absolutely convergent Fourier series, i.e.,

$$\mathcal{F} = \left\{ f \in \mathcal{L}_\infty[0,1)^d : f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{x}}, \text{ where } \left\| \hat{f} \right\|_{\ell_1} < \infty \right\}.$$

with the norm defined as $\|f\|_{\mathcal{F}} = \left\| \hat{f} \right\|_{\ell_1}$. Note that the Fourier coefficients may be defined as

$$\hat{f}(\boldsymbol{k}) = \int_{[0,1)^d} f(\boldsymbol{x}) e^{-2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}.$$

These Fourier coefficients, defined as linear functionals, $f \mapsto \hat{f}(\boldsymbol{k})$, are bounded under the norm. The problem is how to numerically estimate the integrals of functions in $\mathcal{F}$. Note that the integral is just one of the Fourier coefficients and is also a bounded linear functional on $\mathcal{F}$,

$$\mu : f \mapsto \int_{[0,1)^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \qquad \mu(f) = \hat{f}(\boldsymbol{0})$$

The problem of interest is how to numerically estimate the integrals of functions in $\mathcal{F}$. The integral is just one of the Fourier coefficients and thus also a bounded linear functional on $\mathcal{F}$:

$$\mu : f \mapsto \int_{[0,1)^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \qquad \mu(f) = \hat{f}(\mathbf{0}).$$

Integrals are approximated by *sample averages* over node sets of integration lattices. For any set $P \in [0,1)^d$, this is defined as

$$A(\cdot, P) : f \mapsto \frac{1}{|P|} \sum_{\boldsymbol{x} \in P} f(\boldsymbol{x}),$$

where $|P|$ denotes the cardinality of P. Note that $A(\cdot, P)$ is also a bounded linear functional on $\mathcal{F}$. If $P_{\boldsymbol{\Delta}}$ is the node set of a shifted integration lattice, then the sample average may be written in terms of the Fourier coefficients as

$$\begin{aligned}
A(f, P_{\boldsymbol{\Delta}}) &= \frac{1}{|P_{\boldsymbol{\Delta}}|} \sum_{\boldsymbol{x} \in P_{\boldsymbol{\Delta}}} f(\boldsymbol{x}) = \frac{1}{|P|} \sum_{\boldsymbol{x} \in P} \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T(\boldsymbol{x}+\boldsymbol{\Delta})} \\
&= \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{\Delta}} \left[\frac{1}{|P|} \sum_{\boldsymbol{x} \in P} e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{x}} \right] \\
&= \sum_{\boldsymbol{k} \in L^{\perp}} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{\Delta}} = T_{\boldsymbol{\Delta}}(S_{L^{\perp}}(f)),
\end{aligned}$$

where $T_{\boldsymbol{x}} : f \mapsto f(\boldsymbol{x})$ is the evaluation functional defined on $\mathcal{F}$, and $S_{\mathcal{K}} : \mathcal{F} \to \mathcal{F}$ is a bounded linear operator that strips out all but the specified terms of the Fourier series of a function:

$$(S_{\mathcal{K}} f)(\boldsymbol{x}) \; = \sum_{\boldsymbol{k} \in \mathcal{K}} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{x}},$$

where $\mathcal{K} \subseteq \mathbb{Z}^d$.

The error of a quasi-Monte Carlo method based on the design P is also a bounded linear functional, $\mathrm{err}(f, P) = \mu(f) - A(f; P)$. For shifted lattice rules the error may be written in terms of the Fourier coefficients as

$$\begin{aligned}
\mathrm{err}(f, P) = \mu(f) - A(f; P) &= \hat{f}(\mathbf{0}) - \sum_{\boldsymbol{k} \in L^{\perp}} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{\Delta}} \\
&= - \sum_{\boldsymbol{k} \in L^{\perp\prime}} \hat{f}(\boldsymbol{k}) e^{2\pi\sqrt{-1}\boldsymbol{k}^T\boldsymbol{\Delta}} = -T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f)),
\end{aligned}$$

where $L^{\perp\prime} = L^{\perp} \setminus \{\mathbf{0}\}$.

2. Error Estimates

Suppose that L is an integration lattice and $\tilde{L}$ is is a sub-lattice of L. Because lattices are subgroups of $\mathbb{Z}^d$, then any shifted lattice $L_{\boldsymbol{\Delta}}$ can be written as a union of $\tilde{L}_{\boldsymbol{\Delta}}$ and a finite number of shifted counterparts (cosets) of $\tilde{L}$, i.e.,

$$L_{\boldsymbol{\Delta}} = \tilde{L}_{\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_0} \cup \cdots \cup \tilde{L}_{\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_{M-1}}, \qquad \widetilde{\boldsymbol{\Delta}}_i \in P = L \cap [0,1)^d, \quad \widetilde{\boldsymbol{\Delta}}_0 = \mathbf{0},$$

where P is the nodeset of L. Let $P_{\boldsymbol{\Delta}}$ denote the nodeset of $L_{\boldsymbol{\Delta}}$, and $\widetilde{P}_i$ denote the nodeset of $\tilde{L}_{\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_i}$. It follows that $P_{\boldsymbol{\Delta}} = \widetilde{P}_0 \cup \cdots \cup \widetilde{P}_{M-1}$. Moreover, the dual lattice, $\tilde{L}^{\perp}$ is the union of shifted copies of $L^{\perp}$, i.e.,

$$\tilde{L}^{\perp} = L^{\perp}_{\boldsymbol{\nu}_0} \cup \cdots \cup L^{\perp}_{\boldsymbol{\nu}_{M-1}}, \qquad \boldsymbol{\nu}_i \in \tilde{L}^{\perp}, \quad \boldsymbol{\nu}_0 = \mathbf{0}.$$

Furthermore, this implies that

$$\tilde{L}^{\perp\prime} = L^{\perp\prime} \cup L^{\perp}_{\boldsymbol{\nu}_1} \cup \cdots \cup L^{\perp}_{\boldsymbol{\nu}_{M-1}}, \qquad \boldsymbol{\nu}_i \in \tilde{L}^{\perp}, \quad \boldsymbol{\nu}_0 = \mathbf{0}.$$

Note that

$$\boldsymbol{k}^T \widetilde{\boldsymbol{\Delta}}_i \in \mathbb{Z} \qquad \forall \boldsymbol{k} \in L^{\perp}, \tag{1}$$

by definition of the dual lattice. Letting $1_{\mathcal{A}}$ denote the characteristic function, note also that

$$\begin{aligned} 1_{L^{\perp}}(\boldsymbol{k}) &= \frac{1}{|P|} \sum_{\boldsymbol{x} \in P} e^{2\pi\sqrt{-1}\boldsymbol{k}^T \boldsymbol{x}} = \frac{1}{|\widetilde{P}|M} \sum_{\boldsymbol{x} \in \widetilde{P}} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}\boldsymbol{k}^T(\boldsymbol{x}+\widetilde{\boldsymbol{\Delta}}_i)} \\ &= \frac{1}{|\widetilde{P}|} \sum_{\boldsymbol{x} \in \widetilde{P}} e^{2\pi\sqrt{-1}\boldsymbol{k}^T \boldsymbol{x}} \frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}\boldsymbol{k}^T \widetilde{\boldsymbol{\Delta}}_i} = 1_{\tilde{L}^{\perp}}(\boldsymbol{k}) \frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}\boldsymbol{k}^T \widetilde{\boldsymbol{\Delta}}_i}, \end{aligned}$$

which implies that

$$\frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}\boldsymbol{k}^T \widetilde{\boldsymbol{\Delta}}_i} = 1_{\tilde{L}^{\perp} \setminus L^{\perp}}(\boldsymbol{k}) \tag{2}$$

One possible upper bound for $|\mathrm{err}(f, P_{\boldsymbol{\Delta}})| = |T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f))|$ would be

$$\widehat{\mathrm{err}}_0(f, P) = \frac{\mathfrak{C}}{M-1} \left| A(f, P_{\boldsymbol{\Delta}}) - A(f, \widetilde{P}_0) \right|$$

where $C > 0$ is a fudge factor. That is, an upper bound on the error of numerical integration using the nodeset, $P_{\boldsymbol{\Delta}}$, with the larger number of points is proportional to the difference between the integral approximation using the nodeset $P_{\boldsymbol{\Delta}}$ and the approximation using a subset of this nodeset, namely, $\widetilde{P}_0$. This expression can be written in terms of the Fourier coefficients of the integrand

$$\begin{aligned} \widehat{\mathrm{err}}_0(f, P_{\boldsymbol{\Delta}}) &= \frac{\mathfrak{C}}{M-1} \left| A(f, P_{\boldsymbol{\Delta}}) - A(f, \widetilde{P}_0) \right| = \frac{\mathfrak{C}}{M-1} \left| \mathrm{err}(f, P_{\boldsymbol{\Delta}}) - \mathrm{err}(f, \widetilde{P}_0) \right| \\ &= \frac{\mathfrak{C}}{M-1} \left| T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f)) - T_{\boldsymbol{\Delta}}(S_{\tilde{L}^{\perp\prime}}(f)) \right| \\ &= \frac{\mathfrak{C}}{M-1} \left| T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f) - S_{\tilde{L}^{\perp\prime}}(f)) \right| = \frac{\mathfrak{C}}{M-1} \left| T_{\boldsymbol{\Delta}}(S_{\tilde{L}^{\perp\prime} \setminus L^{\perp\prime}}(f)) \right| \\ &= \mathfrak{C} \left| \frac{1}{M-1} \sum_{i=1}^{M-1} T_{\boldsymbol{\Delta}}(S_{L^{\perp}_{\boldsymbol{\nu}_i}}(f)) \right|. \end{aligned}$$

This error bound works well provided that the average of the $T_{\boldsymbol{\Delta}}(S_{L^{\perp}_{\boldsymbol{\nu}_i}}(f))$ mimics $T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f))$. However, a potential problem may arise if there is cancellation among the $T_{\boldsymbol{\Delta}}(S_{L^{\perp}_{\boldsymbol{\nu}_i}}(f))$, causing their average to severely underestimate $T_{\boldsymbol{\Delta}}(S_{L^{\perp\prime}}(f))$.

Another potential upper bound that mitigates against some of the cancellation takes the form of the quasi-standard error:

$$[\widehat{\mathrm{err}}(f, P_{\boldsymbol{\Delta}})]^2 = \frac{\mathfrak{C}^2}{M(M-1)} \sum_{i=0}^{M-1} [A(f, \widetilde{P}_i) - A(f, P_{\boldsymbol{\Delta}})]^2$$

which can also be expressed as

$$
\begin{aligned}
[\widehat{\operatorname{err}}(f, P_{\boldsymbol{\Delta}})]^2 &= \frac{\mathfrak{C}^2}{M(M-1)} \sum_{i=0}^{M-1} [\operatorname{err}(f, \widetilde{P}_i) - \operatorname{err}(f, P_{\boldsymbol{\Delta}})]^2 \\
&= \frac{\mathfrak{C}^2}{M(M-1)} \sum_{i=0}^{M-1} \left\{ [\operatorname{err}(f, \widetilde{P}_i)]^2 - 2 \operatorname{err}(f, \widetilde{P}_i) \operatorname{err}(f, P_{\boldsymbol{\Delta}}) + [\operatorname{err}(f, P_{\boldsymbol{\Delta}})]^2 \right\} \\
&= \frac{\mathfrak{C}^2}{M-1} \left\{ \frac{1}{M} \sum_{i=0}^{M-1} [\operatorname{err}(f, \widetilde{P}_i)]^2 - [\operatorname{err}(f, P_{\boldsymbol{\Delta}})]^2 \right\}
\end{aligned}
$$

The first term in the right hand side expression above may be simplified by applying the properties of the above lattices, including (1) and (2):

$$
\begin{aligned}
\frac{1}{M} \sum_{i=0}^{M-1} [\operatorname{err}(f, \widetilde{P}_i)]^2 &= \frac{1}{M} \sum_{i=0}^{M-1} [T_{\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_i}(S_{\tilde{L}^{\perp\prime}}(f))]^2 \\
&= \sum_{\boldsymbol{k},\boldsymbol{l} \in \tilde{L}^{\perp\prime}} \left[\hat{f}(\boldsymbol{k}) \hat{f}^*(\boldsymbol{l}) \frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}(\boldsymbol{k}-\boldsymbol{l})^T(\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_i)} \right] \\
&= \sum_{\boldsymbol{k} \in \tilde{L}^{\perp\prime}} \sum_{\boldsymbol{m} \in L^{\perp\prime}} \left[\hat{f}(\boldsymbol{k}) \hat{f}^*(\boldsymbol{m}) \frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}(\boldsymbol{k}-\boldsymbol{m})^T(\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_i)} \right] \\
&\qquad + \sum_{\boldsymbol{k} \in \tilde{L}^{\perp\prime}} \sum_{j=1}^{M-1} \sum_{\boldsymbol{m} \in L^{\perp}_{\boldsymbol{\nu}_j}} \left[\hat{f}(\boldsymbol{k}) \hat{f}^*(\boldsymbol{m}) \frac{1}{M} \sum_{i=0}^{M-1} e^{2\pi\sqrt{-1}(\boldsymbol{k}-\boldsymbol{m})^T(\boldsymbol{\Delta}+\widetilde{\boldsymbol{\Delta}}_i)} \right] \\
&= \sum_{\boldsymbol{k},\boldsymbol{m} \in L^{\perp\prime}} \hat{f}(\boldsymbol{k}) \hat{f}^*(\boldsymbol{m}) e^{2\pi\sqrt{-1}(\boldsymbol{k}-\boldsymbol{m})^T \boldsymbol{\Delta}} \\
&\qquad + \sum_{j=1}^{M-1} \sum_{\boldsymbol{k},\boldsymbol{m} \in L^{\perp}_{\boldsymbol{\nu}_j}} \hat{f}(\boldsymbol{k}) \hat{f}^*(\boldsymbol{m}) e^{2\pi\sqrt{-1}(\boldsymbol{k}-\boldsymbol{m})^T \boldsymbol{\Delta}} \\
&= [\operatorname{err}(f, P_{\boldsymbol{\Delta}})]^2 + \sum_{j=1}^{M-1} \left| T_{\boldsymbol{\Delta}}(S_{L^{\perp}_{\boldsymbol{\nu}_j}}(f)) \right|^2 .
\end{aligned}
$$

This implies that

$$
\widehat{\operatorname{err}}(f, P_{\boldsymbol{\Delta}}) = \mathfrak{C} \sqrt{\frac{1}{M-1} \sum_{j=1}^{M-1} \left| T_{\boldsymbol{\Delta}}(S_{L^{\perp}_{\boldsymbol{\nu}_j}}(f)) \right|^2 }. \tag{3}
$$

References

Room E1-208, Department of Applied Mathematics, Illinois Institute of Technology, 10 W. 32ND St., Chicago, IL 60616

School of Mathematics and Statistics, Lanzhou University, Lanzhou City, Gansu, China 730000