INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 5, Number 2, Pages 167–189

AN ALGORITHM-DRIVEN APPROACH TO ERROR ANALYSIS FOR MULTIDIMENSIONAL INTEGRATION

FRED J. HICKERNELL AND JOSEF DICK

Abstract. Most error analyses for numerical integration algorithms specify the space of integrands and then determine the convergence rate for a particular algorithm or the optimal algorithm. This article takes a different perspective of specifying the convergence rate and then finding the largest space of integrands for which the algorithm gives that desired rate. Both worst-case and randomized error analyses are provided.

Key Words. Digital nets, integration lattices, randomized, worst-case.

1. Introduction

Multi-dimensional integrals of the form

$$I(f) := \int_{\mathcal{X}_s} f(\mathbf{x})\, d\rho(\mathbf{x}), \qquad \mathcal{X}_s \subseteq \mathbb{R}^s, \tag{1}$$

arise in a number of applications. Here f is some known integrand and ρ is a given probability measure, i.e., $I(1) = 1$. For example, if $f(\mathbf{x})$ is the discounted payoff of an exotic option, and $\mathbf{x} = (x_1, \ldots, x_s)$ dictates the changes in the prices of the underlying assets that determine the payoff, then the fair price of that option is the average discounted payoff, $I(f)$, where $\mathcal{X}_s = \mathbb{R}^s$ and ρ is a multivariate normal distribution.

Error analysis of numerical integration rules typically yields error bounds and asymptotic rates of convergence for a specified Banach space of integrands. This article proposes a different approach to analyzing numerical integration rules, namely by specifying the convergence rate and the algorithm and then finding the largest space of integrands for which the algorithm gives that desired rate.

The integration rules considered here take the form of a simple average of integrand values:

$$Q_0(f) := 0, \qquad Q_n(f) := n^{-1} \sum_{i=0}^{n-1} f(\mathbf{x}_i) \quad \text{for } n > 0, \tag{2}$$

where $\{\mathbf{x}_i\}$ is the design or set of nodes where the integrand is evaluated. The nodes may be deterministic or random, but they are assumed to be independent of the integrand, making (2) a linear rule. Adaptive rules are not considered. The design is assumed to be an infinite sequence of which one uses the first n points. In

Received by the editors December 21, 2006 and, in revised form, February 21, 2007.
2000 *Mathematics Subject Classification.* 65D30.
This work was completed in part while the authors were associated with Hong Kong Baptist University and partially supported by Hong Kong Research Grants Council grants RGC/HKBU/2030/99P and RGC/HKBU/2020/02P and by Hong Kong Baptist University grant FRG/00-01/II-62.

practice one may want to consider Q_n for some increasing sequence of non-negative integers, $\mathcal{N} = \{0, n_0, n_1, \ldots\}$. A typical example is $n_m = 2^m$. Among the familiar rules of the form (2) that are considered here are simple Monte Carlo rules [4], rules based on low discrepancy sequences, such as integration lattices [11, 14, 20] or digital (t, m, s)-nets [14, 15]. Smolyak rules are similar, with the difference that one replaces n^{-1} with more general weights $a_{i,n}$, hence the approach proposed in this paper could also be applied to such rules.

Researchers have expended considerable effort to understand the strengths and weaknesses of various numerical integration rules. This is typically done by fixing a Banach space of integrands, $\mathcal{F}$, with a norm $\|\cdot\|_{\mathcal{F}}$, and computing the worst possible error of a particular rule for integrands of norm no greater than unity:

$$\begin{aligned} \operatorname{err}(f, Q_n) &:= I(f) - Q_n(f), \\ e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n) &:= \sup_{\|f\|_{\mathcal{F}} \le 1} |\operatorname{err}(f, Q_n)|, \quad n \in \mathcal{N}. \end{aligned} \tag{3}$$

The quantity $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n)$ is called the worst-case error of Q_n. Then one attempts to determine the asymptotic rate of convergence of this quantity, i.e., to show that

$$C_L(s) g(n) \le \inf_{\substack{n' \in \mathcal{N} \\ n' \le n}} e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_{n'}) \le C_U(s) g(n), \quad n = 1, 2, \ldots, \tag{4}$$

for some function $g(n)$ that tends to zero as $n \to \infty$. Typically $g(n)$ is a negative power of n or a negative power of n times some power of $\log n$. When (4) holds, one may say that $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n) \asymp g(n)$. If only an upper bound is known, then one may say that $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n) = \mathcal{O}(g(n))$. It is also of interest to know how the error depends on the dimension, s, i.e., whether $C_U(s)$ and $C_L(s)$ can be made independent of s, or polynomial in s. This corresponds to the problems of strong tractability or tractability, respectively, provided that $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_0) = 1$ and $g(n)$ decays polynomially in n^{-1}.

Knowing that a numerical integration rule has a particular convergence rate is not the full story. One would also like to know the convergence rate of the best possible rule. The worst-case difficulty of an integration problem can be defined as the error of the best possible rule:

$$e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, n) := \inf_{Q_n} e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n), \quad n \in \mathcal{N}. \tag{5}$$

If $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n) \asymp e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, n)$, then the rule Q_n is optimal. In other words, an optimal integration rule has the same convergence rate as the best rule, but their errors may differ by a constant factor (which again may depend on s).

When the numerical integration rule used is a randomized one, then it makes sense to compute the randomized error. Let $\mathcal{Q}_n$ denote the sample space of random rules Q_n, where now n denotes the average number of function evaluations used. Let μ be a probability measure on this sample space, and let rms_{Q_n} denote the root mean square using this measure μ. The randomized error for a given $\mathcal{Q}_n$ and μ is defined as

$$\begin{aligned} \operatorname{rmse}(f, \mathcal{Q}_n, \mu) &:= \operatorname*{rms}_{Q_n} |\operatorname{err}(f, Q_n)|, \\ e^{\mathrm{ra}}(\|\cdot\|_{\mathcal{F}}, \mathcal{Q}_n, \mu) &:= \sup_{\|f\|_{\mathcal{F}} \le 1} \operatorname{rmse}(f, \mathcal{Q}_n, \mu), \quad n \in \mathcal{N}. \end{aligned} \tag{6}$$

Although the norm-based approach described above is quite useful, it has a certain drawback that this article attempts to address, namely, the space of integrands $\mathcal{F}$ is fixed in advance. Once the space of integrands and the accompanying norm

are specified, all the analysis that follows depends essentially on these choices. If a given integrand f does not lie in $\mathcal{F}$, this does not necessarily imply that the rule Q_n performs badly for that f. Moreover, if f lies in $\mathcal{F}$, the error might decay much faster than the worst case. In mathematical terms, for deterministic algorithms

$$f \notin \mathcal{F} \;\not\Rightarrow\; \sup_{n\in\mathcal{N}} |\operatorname{err}(f,Q_n)| \,/ e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_n) = \infty,$$
$$f \in \mathcal{F} \;\not\Rightarrow\; \lim_{n\to\infty} \sup_{\substack{n'\in\mathcal{N}\\ n'\le n}} |\operatorname{err}(f,Q_{n'})| \,/ e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_{n'}) > 0.$$

An analogous situation holds for random algorithms.

The main idea of this article is to *construct spaces of integrands that depend on the integration rules being used* and make these spaces as big as possible given a specified convergence rate (as seen in Section 7 this approach can avoid the problems some of the classical approaches have). More specifically, we introduce a function $g : \mathcal{N} \times \mathbb{N} \to (0,1]$ with $g(0,s) = g(1,s) = 1$ for all s that specifies the desired convergence rate as a function of the number of sample points and the dimension. The spaces we consider are spanned by a sequence of orthonormal basis functions $\{\psi_{\boldsymbol{\nu}}(\mathbf{x})\}$, i.e. the functions in our spaces can be written as $f(\mathbf{x}) = \sum_{\boldsymbol{\nu}} F(\boldsymbol{\nu})\psi_{\boldsymbol{\nu}}(\mathbf{x})$ (below we will denote the span of $\{\psi_{\boldsymbol{\nu}}\}$ by Ψ). The task now is, for a given sequence of quadrature rules with increasing number of points n (in the following $\mathcal{N}$ will denote a subset of the set of natural numbers such that for each $n \in \mathcal{N}$ we have a quadrature rule using n points), to define a norm which makes this space as large as possible under the constraint that the worst-case error is bounded above by $g(n)$. More precisely, we prove the following main theorem:

Theorem 1. *For sample sizes $n \in \mathcal{N}$, dimensions $s \in \mathbb{N}$, and a desired convergence rate, $g(n,s)$, which is monotonically decreasing in n and non-decreasing in s, define*

$$\omega_{wo}(\boldsymbol{\nu},n,s) := \frac{|\operatorname{err}(\psi_{\boldsymbol{\nu}},Q_n)|}{g(n,s)} \quad \text{and} \quad \omega_{ra}(\boldsymbol{\nu},n,s) := \frac{\operatorname{rmse}(\psi_{\boldsymbol{\nu}},\mathcal{Q}_n,\mu)}{g(n,s)}. \tag{7}$$

Define the Banach spaces $\mathcal{F}_{wo,s}$ and $\mathcal{F}_{ra,s}$ as follows:

$$\|f\|_{\mathcal{F}_{wo,s}} := \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})\omega_{wo}(\boldsymbol{\nu},n,s)|, \quad \mathcal{F}_{wo,s} = \{f \in \Psi : \|f\|_{\mathcal{F}_{wo,s}} < \infty\},$$
$$\|f\|^2_{\mathcal{F}_{ra,s}} := \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})\omega_{ra}(\boldsymbol{\nu},n,s)|^2, \quad \mathcal{F}_{ra,s} = \{f \in \Psi : \|f\|_{\mathcal{F}_{ra,s}} < \infty\},$$

where the $F(\boldsymbol{\nu})$ are the $\psi_{\boldsymbol{\nu}}$-series coefficients of f. Then under conditions (9) *and* (11) *below the randomized and worst-case errors of $\mathcal{F}_{wo,s}$ and $\mathcal{F}_{ra,s}$, respectively, have the desired convergence rate, namely,*

$$e^{wo}(\|\cdot\|_{wo,s}, Q_n) \le g(n,s) \quad \text{and} \quad e^{ra}(\|\cdot\|_{ra,s}, \mathcal{Q}_n) \le g(n,s)$$

for all $n \in \mathcal{N}$ and $s \in \mathbb{N}$.

Moreover, these are the largest possible spaces with the desired convergence rates in the following sense: if $\operatorname{span}\{\psi_{\boldsymbol{\nu}}\}$ is a dense subset in the space of absolutely continuous functions $\mathcal{C}(\mathcal{X}_s)$, then for any $f \in \mathcal{C}(\mathcal{X}_s) \setminus \mathcal{F}_{wo,s}$ there exists a $\tilde{f} \in \mathcal{C}(\mathcal{X}_s) \setminus \mathcal{F}_{wo,s}$ with $|\tilde{F}(\boldsymbol{\nu})| = |F(\boldsymbol{\nu})|$ such that $\sup_{n\in\mathcal{N}} |\operatorname{err}(\tilde{f},Q_n)|/g(n,s) = \infty$. Further, if $\operatorname{span}\{\psi_{\boldsymbol{\nu}}\}$ is a dense subset in $\mathcal{L}_2(\mathcal{X}_s)$, then for all $f \in \mathcal{L}_2(\mathcal{X}_s) \setminus \mathcal{F}_{ra,s}$ it follows that $\sup_{n\in\mathcal{N}} \operatorname{rmse}(f,\mathcal{Q}_n,\mu)/g(n) = \infty$.

The norms in the theorem above now depend on the quadrature rule through $\operatorname{err}(\psi_{\boldsymbol{\nu}},Q_n)$, the worst-case error of the basis function $\psi_{\boldsymbol{\nu}}$ using the quadrature rule Q_n, and $\operatorname{rmse}(\psi_{\boldsymbol{\nu}},\mathcal{Q}_n,\mu)$, the root mean square error of the basis function $\psi_{\boldsymbol{\nu}}$ using

the random quadrature rules from the sample space $\mathcal{Q}_n$. Furthermore, these norms are also shown to be as weak as possible.

The following section provides some necessary preliminaries. Section 3 describes the method for constructing spaces of integrands for the randomized error analysis, and Section 4 gives an analogous treatment for worst-case error analysis. The proof of the main theorem of this article is given in Section 3 for the randomized error and Section 4 for the worst-case error. This theorem is applied to both lattice integration rules and digital net integration rules in Sections 5 and 6. This article concludes with some examples and a discussion.

2. Preliminaries

2.1. Basis Functions and Quadrature Rules. The integrands considered in this article are defined on some Cartesian product domain, $\mathcal{X}_s$, that is an measureable subset of $\mathbb{R}^s$. The integrands are represented as infinite series:

$$f(\mathbf{x}) = \sum_{\boldsymbol{\nu}} F(\boldsymbol{\nu})\psi_{\boldsymbol{\nu}}(\mathbf{x}), \quad \mathbf{x} \in \mathcal{X}_s \subseteq \mathbb{R}^s, \qquad \|f\|_2^2 = \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})|^2 < \infty, \tag{8}$$

where $\mathbf{x} := (x_1, x_2, \ldots, x_s)$, $\{\psi_{\boldsymbol{\nu}}(\mathbf{x})\}$ is a sequence of complex-valued, integrable, *orthonormal* basis functions with index $\boldsymbol{\nu}$ that has a countable range. Orthonormality is meant with respect to the $\mathcal{L}_2$ inner product on $\mathcal{X}_s$, i.e., $I(\bar{\psi}_{\boldsymbol{\nu}}\psi_{\boldsymbol{\lambda}})$ is unity for $\boldsymbol{\nu} = \boldsymbol{\lambda}$ and zero otherwise. This means that the series coefficients may be written as $F(\boldsymbol{\nu}) = I(f\bar{\psi}_{\boldsymbol{\nu}})$. The integrands are assumed to be square integrable (their series coefficients are square summable), to permit randomized error analysis of simple Monte Carlo methods and other random algorithms. For worst-case error analysis it is also assumed that the series are absolutely summable for any $\mathbf{x} \in \mathcal{X}_s$.

The dependence of the index $\boldsymbol{\nu}$, the basis function $\psi_{\boldsymbol{\nu}}(\mathbf{x})$, and the set of possible $\boldsymbol{\nu}$ on the dimension s are not explicitly stated for simplicity of notation. The probability measure $\rho(\mathbf{x})$ used to define the integral in (1) is assumed to have independent marginals and so the basis functions are assumed to be of tensor product form. As s increases the set of possible $\boldsymbol{\nu}$ increases monotonically. For example, for trigonometric polynomials $\psi_{\mathrm{tr},\boldsymbol{\nu}}(\mathbf{x}) := e^{2\pi\imath\boldsymbol{\nu}^T\mathbf{x}}$ for $\boldsymbol{\nu} \in \mathbb{Z}^s$ where $\imath := \sqrt{-1}$. The wavenumbers $(1, 2)$ and $(1, 2, 0)$ are taken to be equivalent, although one corresponds to $s = 2$ and the other to $s = 3$. Three examples of basis functions defined on the unit cube are discussed in Sections 5 and 6: trigonometric polynomials, Walsh functions and Haar wavelets.

It is further assumed that the function one has unit norm and corresponds to $\boldsymbol{\nu} = 0$, i.e.,

$$\psi_{\mathbf{0}}(\mathbf{x}) = 1, \qquad \|\psi_{\mathbf{0}}\|_{\mathcal{F}} = \|1\|_{\mathcal{F}} = 1, \tag{9}$$

which implies that $I(\psi_{\boldsymbol{\nu}}) = 0$ for all $\boldsymbol{\nu} \neq \mathbf{0}$. Thus error of the zero algorithm is simply the error in integrating constants, which is unity:

$$e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}}, Q_0) = e^{\mathrm{ra}}(\|\cdot\|_{\mathcal{F}}, \mathcal{Q}_0, \mu) = 1. \tag{10}$$

Quadrature rules of the form (2) include Monte Carlo and quasi-Monte Carlo integration rules, such as lattice rules and digital net rules. These rules are considered in detail in Sections 5 and 6. The performance of these rules depends on how well the design, $\{\mathbf{x}_i\}$ is chosen. As mentioned earlier, it is assumed that designs considered here are extensible in $n \in \mathcal{N}$, that is, the design for n_{m+1} points is a superset of the design for n_m points.

The quadrature rules Q_n depend implicitly on s, but it is assumed that the designs are extensible in dimension s. This means that for problems of dimension

$s+1$ one takes $\mathbf{x}_i = (x_{i1}, \ldots, x_{is})$ from the s-dimensional design and appends the value $x_{i,s+1}$ to form $\mathbf{x}_i$ for the $s+1$-dimensional design. Designs that are extensible in the dimension make sense because the probability measure $\rho(\mathbf{x})$ has independent marginals.

For the randomized error analysis it is assumed that the integration rule and the basis functions match in the sense that the errors for different basis functions are uncorrelated:

$$E_{Q_n \in \mathcal{Q}_n}\{[\mathrm{err}(\psi_{\boldsymbol{\nu}}, \mathcal{Q}_n, \mu)][\mathrm{err}(\bar{\psi}_{\boldsymbol{\lambda}}, \mathcal{Q}_n, \mu)]\} = 0 \quad \text{for } \boldsymbol{\nu} \neq \boldsymbol{\lambda}. \tag{11}$$

This is a nontrivial condition, however, Lemmas 6, 9, and 10 below gives examples of integration rules and basis functions that satisfy it. Simple Monte Carlo rules satisfy this condition for any orthonormal functions. Randomly shifted lattice rules satisfy this condition for trigonometric polynomials. Digital nets satisfy this condition for Walsh functions and Haar wavelets.

2.2. Monotonic Convergence Rates. In the subsequent sections the goal is to construct as large a space of functions as possible so that integrands in the unit ball have worst-case or randomized errors that decay like some prescribed function $g(n, s)$. It is assumed that the function $g(n, s)$ decreases monotonically to 0 in n and is non-decreasing in s, i.e. we assume

$$1 = g(0, s) = g(1, s) \geq g(n, s) \geq g(n', s) > 0 \quad \text{for } 1 \leq n \leq n' < \infty \text{ and all } s. \tag{12}$$

Functions such as $g(n, s) = n^{-\alpha}$ for $\alpha > 0$ are suitable, but functions such as $g(n, s) = n^{-\alpha}(\log n)^s$ are not suitable, although they often arise in the error analysis for multivariate integration rules. The problem with $n^{-\alpha}(\log n)^s$ is that it *does not decrease monotonically with* n.

One solution to this problem is to use $g(n, s) = n^{-\alpha+\epsilon}$, but then ϵ becomes a free parameter in the exponent. An alternative is the function

$$g(n; \alpha, \delta) := \begin{cases} e^{-\alpha[\log n + 1 - (1+\delta^{-1}\log n)^{\delta}]}, & 0 < \delta < 1,\ n \geq 1, \\ n^{-\alpha}, & \delta = 0,\ n \geq 1, \\ 1, & n = 0. \end{cases} \tag{13}$$

There is still the undetermined parameter δ, but this does not affect the power of n. This s-independent function g is used in the examples in Sections 5 and 6.

Lemma 2. *For any fixed $\alpha > 0$ and δ with $0 \leq \delta < 1$ the function $g(n, s) = g(n; \alpha, \delta)$ defined in* (13) *satisfies* (12)*, and furthermore,*

$$n^{-\alpha} = g(n; \alpha, 0) \ll n^{-\alpha}(\log n)^s \ll g(n; \alpha, \delta) \ll g(n; \alpha - \epsilon, 0) = n^{-\alpha+\epsilon} \tag{14}$$
$$\text{as } n \to \infty, \quad \text{for } 0 < \delta < 1,\ \epsilon > 0,\ s \geq 0.$$

Proof. The condition $g(0; \alpha, \delta) = g(1; \alpha, \delta) = 1$ follows from the definition. Consider the quantity

$$\tilde{g}(n) = \log\left(\frac{g(n; \alpha, \delta)}{n^{-\tilde{\alpha}}(\log n)^s}\right) = (\tilde{\alpha} - \alpha)\log n - \alpha + \alpha(1 + \delta^{-1}\log n)^{\delta} - s\log(\log n),$$

defined for all $n \geq 1$, and its first derivative with respect to n:

$$\tilde{g}'(n) = \frac{1}{n}\left[(\tilde{\alpha} - \alpha) + \alpha(1 + \delta^{-1}\log n)^{\delta-1} - \frac{s}{\log n}\right].$$

If $\tilde{\alpha} = s = 0$, then $\tilde{g}'(n) < 0$ since $\delta < 1$, and $g(n, s) = g(n; \alpha, \delta)$ decreases monotonically in n for $n \geq 1$. If $\tilde{\alpha} = \alpha$ and $s \geq 0$, then $\lim_{n\to\infty} \tilde{g}(n) = +\infty$ since

$\delta > 0$, which proves the left side of (14). If $\tilde{\alpha} < \alpha$ and $s = 0$, then $\lim_{n\to\infty} \tilde{g}(n) = -\infty$ since $\delta < 1$, which proves the right side of (14). □

The notion of *strong tractability* as studied in [16] and elsewhere is important for high dimensional problems. Briefly, integration is strongly tractable if there exists an algorithm whose worst-case or randomized error relative to the error for the zero algorithm is bounded above by $C_U g(n;\alpha,\delta)$ for some $C_U > 0$ independent of the dimension of the problem. For spaces considered here, the error of the zero algorithm is unity and can be ignored. Therefore, the notion of strong tractability in the context of this article corresponds to determining how large the series coefficients $F(\boldsymbol{\nu})$ can be while still maintaining the pre-determined worst-case or randomized convergence rate $g(n;\alpha,\delta)$.

3. Randomized Error Analysis

Condition (11) is crucial for the randomized error analysis. Under this assumption the root mean square quadrature error for a specific integrand is

$$\operatorname{rmse}(f,\mathcal{Q}_n,\mu) = \operatorname*{rms}_{Q_n\in\mathcal{Q}_n} |\operatorname{err}(f,Q_n)| = \sqrt{\sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})\operatorname{rmse}(\psi_{\boldsymbol{\nu}},\mathcal{Q}_n,\mu)|^2}.$$

It follows that

$$\sup_{n\in\mathcal{N}} \frac{\operatorname{rmse}(f,\mathcal{Q}_n,\mu)}{g(n,s)} \le 1 \iff \sup_{n\in\mathcal{N}} \sqrt{\sum_{\boldsymbol{\nu}} \left|\frac{F(\boldsymbol{\nu})\operatorname{rmse}(\psi_{\boldsymbol{\nu}},\mathcal{Q}_n,\mu)}{g(n,s)}\right|^2} \le 1$$

$$\iff \sup_{n\in\mathcal{N}} \sqrt{\sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})|^2 \omega_{\mathrm{ra}}^2(\boldsymbol{\nu},n,s)} \le 1, \tag{15}$$

where $\omega_{\mathrm{ra}}(\boldsymbol{\nu},n,s)$ was defined in (7).

This equivalence shows how to define the norm that gives the largest space of integrands with a randomized error decaying as $g(n,s)$ as given in Theorem 1. The proof of the part of Theorem 1 dealing with the randomized setting follows from (15). The root mean square integration error for the integrand f converges as fast as $g(n,s)$ if and only if the norm $\|f\|_{\mathrm{ra},s}$ is finite.

The norm $\|f\|_{\mathrm{ra},s}$ defined in Theorem 1 is not easy to calculate in practice because there is a supremum over an infinite sum parameterized by n. Since equivalent norms give the same function space, it is desirable to find equivalent norms with a simpler form. For convenience sake it may even be desirable to find a slightly stronger norm that yields a slightly smaller space of integrands for which the randomized error still has the same convergence rate. The forms of these simpler norms depend on how $\omega_{\mathrm{ra}}(\boldsymbol{\nu},n,s)$ depends on $\boldsymbol{\nu}$, n, and s.

For the rest of this section we will assume that $\{\psi_{\boldsymbol{\nu}}\}$ forms a complete orthonormal system in $\mathcal{L}_2(\mathcal{X}_s)$. The following corollary summarizes now two extreme cases.

Corollary 3. *Let a convergence rate g be given. If $\sup_{n\in\mathcal{N}} \omega_{ra}(\boldsymbol{\nu},n,s) = C_{U,s} < \infty$, then $\mathcal{F}_{ra,s} = \mathcal{L}_2(\mathcal{X}_s)$ and $e^{ra}(\|\cdot\|_{\mathcal{L}_2(\mathcal{X}_s)},\mathcal{Q}_n,\mu) = \mathcal{O}(g(n,s))$. On the other hand, if for a given s one has $\sup_{n\in\mathcal{N}} \omega_{ra}(\boldsymbol{\nu},n,s) = \infty$ for all $\boldsymbol{\nu} \ne \mathbf{0}$, then $\mathcal{F}_{ra,s}$ is the set of constant functions.*

Proof. Under the first assumption the formula for $\|f\|_{\mathrm{ra},s}^2$ can be bounded by

$$\|f\|_{\mathrm{ra},s}^2 \le \sup_{\substack{n\in\mathcal{N}\\ \boldsymbol{\nu}}} \omega_{\mathrm{ra}}^2(\boldsymbol{\nu},n,s) \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})|^2 = C_U^2 \|f\|_{\mathcal{L}_2(\mathcal{X}_s)}^2,$$

which implies that the norm $\|\cdot\|_{\mathrm{ra},s}$ is no stronger than the norm $\|\cdot\|_{\mathcal{L}_2(\mathcal{X}_s)}$. The conclusion follows.

Under the second assumption, suppose that there exists a nonconstant $f \in \mathcal{F}_{\mathrm{ra},s}$, i.e., for some $\boldsymbol{\nu}_0 \neq \mathbf{0}$, $F(\boldsymbol{\nu}_0) \neq 0$. Then by the definition in Theorem 1

$$\|f\|^2_{\mathrm{ra},s} = \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})|^2 \omega^2_{\mathrm{ra}}(\boldsymbol{\nu}, n, s) \geq \sup_{n\in\mathcal{N}} |F(\boldsymbol{\nu}_0)|^2 \omega^2_{\mathrm{ra}}(\boldsymbol{\nu}_0, n, s) = \infty.$$

This leads to a contradiction, so all $f \in \mathcal{F}_{\mathrm{ra},s}$ are constant functions. □

Simple Monte Carlo integration fits the assumptions of Corollary 3. Since the basis functions are orthonormal, condition (11) holds automatically, and the root mean square error of integrating a basis function by simple Monte Carlo is $\mathrm{rmse}(\psi_{\boldsymbol{\nu}}, \mathcal{Q}_n, \mu) = n^{-1/2}$. For $g(n,s) = n^{-1/2}$ one has $\omega_{\mathrm{ra}}(\boldsymbol{\nu}, n, s) = 1$. The first part of this corollary holds, and so $e^{\mathrm{ra}}(\|\cdot\|_{\mathcal{L}_2(\mathcal{X}_s)}, \mathrm{MC}, n) = \mathcal{O}(n^{-1/2})$. For $g(n,s) = \mathcal{O}(n^{-1/2-\epsilon})$ with $\epsilon > 0$ the second part of this corollary holds. This implies that $e^{\mathrm{ra}}(\|\cdot\|_{\mathcal{L}_2(\mathcal{X}_s)}, \mathrm{MC}, n) \asymp n^{-1/2}$, and the convergence rate for simple Monte Carlo integration can only be improved by shrinking the space of integrands to constants.

Corollary 3 is not useful for more sophisticated rules, such as randomly shifted integration lattices, because

$$\sup_{\substack{n\in\mathcal{N}\\ \boldsymbol{\nu}}} \omega_{\mathrm{ra}}(\boldsymbol{\nu}, n, s) = \infty$$

in this case. However, we note that as n increases, an increasing number of basis functions are integrated exactly. This suggests another corollary that can be used to simplify the norm $\|\cdot\|_{\mathrm{ra},s}$.

Corollary 4. *Let $\tilde{\omega}_{ra}(\boldsymbol{\nu}, n, s)$ be a function satisfying $\tilde{\omega}_{ra}(\mathbf{0}, n, s) = 1$,*

$$C_{U,s} := \sup_{\substack{n\in\mathcal{N}\\ \boldsymbol{\nu}}} \frac{\omega_{ra}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{ra}(\boldsymbol{\nu}, n, s)} < \infty, \qquad \tilde{C}^2_{U,s} := \sup_{\substack{n\in\mathcal{N}\\ \boldsymbol{\nu}}} \sum_{\substack{n'\in\mathcal{N}\\ n'\geq n}} \frac{\tilde{\omega}^2_{ra}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}^2_{ra}(\boldsymbol{\nu}, n', s)} < \infty. \tag{16}$$

For each $\boldsymbol{\nu}$ let $N(\boldsymbol{\nu}) \in \mathcal{N}$ be defined such that $\mathrm{rmse}(\psi_{\boldsymbol{\nu}}, \mathcal{Q}_n, \mu) = 0$ for all $n \in \mathcal{N}$ with $n > N(\boldsymbol{\nu})$. Define the norms $\|\cdot\|_{ra,s,k}$ for $k = 1, 2$ by

$$\|f\|^2_{ra,s,1} := \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n} |F(\boldsymbol{\nu})|^2 \tilde{\omega}^2_{ra}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s),$$

$$\|f\|^2_{ra,s,2} := \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})|^2 \tilde{\omega}^2_{ra}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s),$$

and the spaces $\mathcal{F}_{ra,s,k} = \{f \in \mathcal{L}_2(\mathcal{X}_s) : \|f\|_{ra,s,k} < \infty\}$, $k = 1, 2$. It then follows that $e^{ra}(\|\cdot\|_{ra,s,k}, \mathcal{Q}_0, \mu) = 1$ for $k = 1, 2$. Furthermore, $(C_{U,s}\tilde{C}_{U,s})^{-1} \|f\|_{ra,s} \leq \|f\|_{ra,s,1} \leq \|f\|_{ra,s,2}$ for any $f \in \mathcal{L}_2(\mathcal{X}_s)$, and so $e^{ra}(\|\cdot\|_{ra,s,k}, \mathcal{Q}_n, \mu) = \mathcal{O}(g(n,s))$ for $k = 1, 2$. On the other hand, if

$$C_{L,s} := \inf_{n'\in\mathcal{N}} \sup_{\substack{n\in\mathcal{N}\\ n\leq n'}} \inf_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} \frac{\omega_{ra}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{ra}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s)} > 0, \tag{17}$$

then $\|\cdot\|_{ra,1}$ is equivalent to $\|\cdot\|_{ra}$ and $\mathcal{F}_{ra,s,1} = \mathcal{F}_{ra,s}$.

Before proving Corollary 4 some explanation is given of the assumptions. The first part of condition (16) simply allows one to choose weights $\tilde{\omega}_{\mathrm{ra}}$ that are simpler in form and not significantly smaller than ω_{ra}. Their precise form vary from case to

case. The second part of this condition is a technical one. It is satisfied, for example, by choosing $\tilde{\omega}_{\mathrm{ra}} = \mathcal{O}(n^{\alpha})$ and any $\mathcal{N} = \{0, b, b^2, \ldots\}$ with $b > 1$. Essentially, one must ensure that $\inf n_{m+1}/n_m > 1$.

Proof. The condition that $\tilde{\omega}_{\mathrm{ra}}(\mathbf{0}, n, s) = 1$ implies that $\|1\|_{\mathrm{ra},s,k} = 1$, and so, as observed in (10), $e^{\mathrm{ra}}(\|\cdot\|_{\mathrm{ra},s,k}, \mathcal{Q}_0, \mu) = 1$ for $k = 1, 2$. Also note that $\|f\|_{\mathrm{ra},s,1} \le \|f\|_{\mathrm{ra},s,2}$ by definition.

The sum with respect to $\boldsymbol{\nu}$ in the definition of the norm $\|\cdot\|_{\mathrm{ra},s}$ can be bounded in terms of sums over the sets $\{\boldsymbol{\nu} : N(\boldsymbol{\nu}) = n'\}$ as follows:

$$
\begin{aligned}
\|f\|_{\mathrm{ra}}^2 &= \sup_{n\in\mathcal{N}} \sum_{n'\ge n} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s) \frac{\omega_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)} \\
&\le C_{U,s}^2 \sup_{n\in\mathcal{N}} \sum_{n'\ge n} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s) \\
&\le C_{U,s}^2 \sup_{n\in\mathcal{N}} \sum_{n'\ge n} \left[\frac{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n', s)} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n', s) \right] \\
&\le C_{U,s}^2 \sup_{n\in\mathcal{N}} \left(\left\{ \sum_{n'\ge n} \frac{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n', s)} \right\} \left\{ \sup_{n'\ge n} \left[\sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, n', s) \right] \right\} \right) \\
&\le C_{U,s}^2 \tilde{C}_{U,s}^2 \, \|f\|_{\mathrm{ra},s,1}^2 .
\end{aligned}
$$

This proves the first part of the corollary.

To prove the second part of the corollary a lower bound on the norm $\|\cdot\|_{\mathrm{ra},s}$ is derived with the help of assumption (17):

$$
\begin{aligned}
\|f\|_{\mathrm{ra},s}^2 &\ge \sup_{\substack{n,n'\in\mathcal{N} \\ n'\ge n}} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \omega_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s) \\
&\ge \sup_{\substack{n,n'\in\mathcal{N} \\ n'\ge n}} \left[\left\{ \inf_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} \frac{\omega_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s)} \right\} \right. \\
&\qquad \left. \times \left\{ \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s) \right\} \right] \\
&\ge \sup_{n'\in\mathcal{N}} \left[\left\{ \sup_{\substack{n\in\mathcal{N} \\ n\le n'}} \inf_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} \frac{\omega_{\mathrm{ra}}^2(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s)} \right\} \right. \\
&\qquad \left. \left\{ \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n'} |F(\boldsymbol{\nu})|^2 \, \tilde{\omega}_{\mathrm{ra}}^2(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s) \right\} \right] \\
&\ge C_{L,s}^2 \, \|f\|_{\mathrm{ra},sm1}^2 .
\end{aligned}
$$

□

Whereas $\|\cdot\|_{\mathrm{ra},s}$ requires one to compute the supremum over an infinite sequence of n of infinite series in $\boldsymbol{\nu}$, $\|\cdot\|_{\mathrm{ra},s,1}$ requires only the supremum over an infinite sequence of n of finite series in $\boldsymbol{\nu}$, and $\|\cdot\|_{\mathrm{ra},s,2}$ requires only an infinite series in $\boldsymbol{\nu}$.

4. Worst-Case Error Analysis

Now the worst-case numerical integration error is considered. For this analysis it is not enough that the integrands be square integrable. It is assumed that the series defining $f(\mathbf{x})$ is absolutely summable for all $\mathbf{x}$.

A specific rule, Q_n, applied to a particular integrand, f, has absolute error

$$\operatorname{err}(f, Q_n) = \left| \sum_{\boldsymbol{\nu}} F(\boldsymbol{\nu}) \operatorname{err}(\psi_{\boldsymbol{\nu}}, Q_n) \right| \leq \sum_{\boldsymbol{\nu}} \left| F(\boldsymbol{\nu}) \operatorname{err}(\psi_{\boldsymbol{\nu}}, Q_n) \right|,$$

with equality holding if the phase of $F(\boldsymbol{\nu})$ is chosen to make $F(\boldsymbol{\nu}) \operatorname{err}(\psi_{\boldsymbol{\nu}}, Q_n)$ non-negative (non-positive) for every $\boldsymbol{\nu}$. It follows that

$$\sup_{n \in \mathcal{N}} \left| \operatorname{err}(f, Q_n) \right| / g(n, s) \leq 1 \Longleftarrow \sup_{n \in \mathcal{N}} \sum_{\boldsymbol{\nu}} \left| F(\boldsymbol{\nu}) \operatorname{err}(\psi_{\boldsymbol{\nu}}, Q_n) / g(n, s) \right| \leq 1, \tag{18}$$

with equivalence holding if the phases of $F(\boldsymbol{\nu})$ are chosen appropriately.

This statement suggests the definition of the norm that gives the largest space of integrands with a worst-case error decaying as $g(n, s)$ as defined in Theorem 1. The definition of $\|f\|_{\mathrm{wo},s}$ corresponds to the right side of (18). The implication in (18) proves the first part of this theorem. To obtain the second part it is noted that (18) becomes an equivalence if the phases of $F(\boldsymbol{\nu})$ are chosen correctly. In other words, there may be some integrands, f, left out of $\mathcal{F}_{\mathrm{wo},s}$ for which Q_n performs well, but for each such integrand there is a related one, $\tilde{f}$ whose series coefficients have the same magnitudes as those of f for which Q_n performs poorly.

Corollary 4 has an analog in the worst-case setting. The proof is similar and is omitted. Again, for convenience, we will assume that $\{\psi_{\boldsymbol{\nu}}\}$ forms a complete orthornormal system in $\mathcal{L}_2(\mathcal{X}_s)$.

Corollary 5. *Let $\tilde{\omega}_{wo}(\boldsymbol{\nu}, n, s)$ satisfy $\tilde{\omega}_{wo}(\mathbf{0}, n, s) = 1$,*

$$C_{U,s} := \sup_{\substack{n \in \mathcal{N} \\ \boldsymbol{\nu}}} \frac{\omega_{wo}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{wo}(\boldsymbol{\nu}, n, s)} < \infty, \qquad \tilde{C}_{U,s} := \sup_{\substack{n \in \mathcal{N} \\ \boldsymbol{\nu}}} \sum_{\substack{n' \in \mathcal{N} \\ n' \geq n}} \frac{\tilde{\omega}_{wo}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{wo}(\boldsymbol{\nu}, n', s)} < \infty.$$

For each $\boldsymbol{\nu}$ let $N(\boldsymbol{\nu}) \in \mathcal{N}$ be defined such that $\operatorname{err}(\psi_{\boldsymbol{\nu}}, Q_n, \mu) = 0$ for all $n \in \mathcal{N}$ with $n > N(\boldsymbol{\nu})$. Define the norms $\|\cdot\|_{wo,s,k}$ for $k = 1, 2$ by

$$\|f\|_{wo,s,1} := \sup_{n \in \mathcal{N}} \sum_{\boldsymbol{\nu}: N(\boldsymbol{\nu}) = n} \left| F(\boldsymbol{\nu}) \right| \tilde{\omega}_{wo}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s),$$

$$\|f\|_{wo,s,2} := \sum_{\boldsymbol{\nu}} \left| F(\boldsymbol{\nu}) \right| \tilde{\omega}_{wo}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s),$$

and the spaces $\mathcal{F}_{wo,s,k} = \{f \in \Psi : \|f\|_{wo,s,k} < \infty\}$, $k = 1, 2$. It then follows that $e^{wo}(\|\cdot\|_{wo,s,k}, Q_0) = 1$ for $k = 1, 2$. Furthermore, $(C_{U,s}\tilde{C}_{U,s})^{-1} \|f\|_{wo,s} \leq \|f\|_{wo,s,1} \leq \|f\|_{wo,s,2}$ for any $f \in \Psi$, and so $e^{wo}(\|\cdot\|_{wo,s,k}, Q_n, \mu) = \mathcal{O}(g(n, s))$ for $k = 1, 2$. On the other hand, if

$$C_{L,s} := \inf_{n' \in \mathcal{N}} \sup_{\substack{n \in \mathcal{N} \\ n \leq n'}} \inf_{\boldsymbol{\nu}: N(\boldsymbol{\nu}) = n'} \frac{\omega_{wo}(\boldsymbol{\nu}, n, s)}{\tilde{\omega}_{wo}(\boldsymbol{\nu}, N(\boldsymbol{\nu}), s)} > 0, \tag{19}$$

then $\|\cdot\|_{wo,1}$ is equivalent to $\|\cdot\|_{wo}$ and $\mathcal{F}_{wo,s,1} = \mathcal{F}_{wo,s}$.

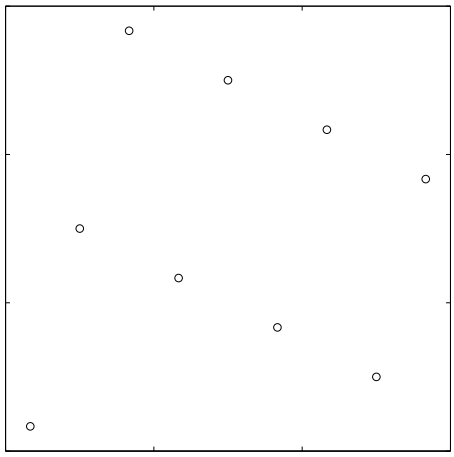

FIGURE 1. Plot of a two-dimensional nodeset of a shifted integration lattice that is also shifted digital net.

5. Extensible Lattice Rules

In this section and the next, Theorem 1 and the corollaries of the previous two sections are applied to particular quadrature rules. First extensible lattice rules are considered. The suitable domain of integration is the half-open, s-dimensional unit cube, i.e., $\mathcal{X}_s = [0,1)^s$, and the basis functions are multivariate trigonometric polynomials $\psi_{\mathrm{tr},\boldsymbol{\nu}}(\mathbf{x}) := e^{2\pi\imath\boldsymbol{\nu}^T\mathbf{x}}$ for $\boldsymbol{\nu} \in \mathbb{Z}^s$.

Integration lattices are a kind of design that places the sample points more evenly than i.i.d. random points. An integration lattice, L, is a superset of $\mathbb{Z}^s$ that is closed under addition and subtraction [11, 14, 20]. To obtain a design in the unit cube one takes the nodeset of a shifted lattice, i.e., $\{\mathbf{x}_i = \mathbf{y} + \boldsymbol{\Delta} \mod 1 : \mathbf{y} \in L \cap [0,1)^s\}$, where $\boldsymbol{\Delta}$ is some point in $[0,1)^s$. The shift $\boldsymbol{\Delta}$ can be fixed deterministically or chosen randomly uniformly on $[0,1)^s$ [2].

One commonly used integration lattice is the shifted rank-1 lattice. In this case the nodeset may be written as $\{\mathbf{x}_i = i\mathbf{h}/n + \boldsymbol{\Delta} \mod 1 : i = 0, \ldots, n-1\}$. The nodeset in Figure 1 shows the lattice rule with $n = 9$, $\mathbf{h} = (1, 8)$ and $\boldsymbol{\Delta} = (1/18, 1/18)$.

The dual lattice is defined as $L^\perp := \{\boldsymbol{\nu} \in \mathbb{Z}^s : \boldsymbol{\nu}^T\mathbf{y} \in \mathbb{Z}^s \ \forall \mathbf{y} \in L\}$. Let L_n denote a lattice whose nodeset has n points, and let $L_n^\perp$ denote the corresponding dual lattice. The lattice rules considered here are all extensible [8, 9]. This means that for $\mathcal{N} = \{0, n_0, n_1, \ldots\}$, n_{m+1} must be an integer multiple of n_m. Furthermore,

$$\mathbb{Z}^s = L_0 \subset L_{n_0} \subset L_{n_1} \subset L_{n_2} \subset \cdots \text{ and } \mathbb{Z}^s = L_0^\perp \supset L_{n_0}^\perp \supset L_{n_1}^\perp \supset L_{n_2}^\perp \supset \cdots .$$

Shifted lattice rules do a perfect job of integrating trigonometric polynomials whose wave numbers are not in the dual lattice. This is made precise in the following lemma.

Lemma 6. *For shifted lattice rules applied to trigonometric basis functions one has the following absolute and root mean square errors for $n \geq 1$:*

$$|\mathrm{err}(\psi_{tr,\boldsymbol{\nu}}, \mathit{lat}, n)| = \mathrm{rmse}(\psi_{tr,\boldsymbol{\nu}}, \mathit{ra\text{-}lat}, n) = \begin{cases} 1 & \mathit{for}\ \mathbf{0} \neq \boldsymbol{\nu} \in L_n^\perp, \\ 0 & \mathit{for}\ \boldsymbol{\nu} = \mathbf{0}\ \mathit{or}\ \boldsymbol{\nu} \notin L_n^\perp, \end{cases} \tag{20}$$

For a randomly shifted lattice rule and trigonometric basis functions condition (11) *holds.*

Proof. The proofs of (20) may be found in [14, 20, 7]. The proof of (11) may be found in [7]. □

5.1. Weakest Norms for Shifted Extensible Lattice Rules. For randomly or deterministically shifted extensible lattice rules applied to series of trigonometric polynomials one may define $\mathcal{C}_{\rm tr}([0,1)^s) := \{f : \sum_{\boldsymbol{\nu}} |F(\boldsymbol{\nu})| < \infty\}$. By Lemma 6 the norms in Theorem 1 may be written as

$$\|f\|_{\text{lat},s} = \max\left(|F(\mathbf{0})|, \sup_{0\neq n\in\mathcal{N}} \sum_{\mathbf{0}\neq\boldsymbol{\nu}\in L_n^\perp} \frac{|F(\boldsymbol{\nu})|}{g(n,s)}\right),$$

$$\|f\|^2_{\text{ra-lat},s} = \max\left(|F(\mathbf{0})|^2, \sup_{0\neq n\in\mathcal{N}} \sum_{\mathbf{0}\neq\boldsymbol{\nu}\in L_n^\perp} \left|\frac{F(\boldsymbol{\nu})}{g(n,s)}\right|^2\right).$$

The choices $\tilde{\omega}_{\rm wo}(\boldsymbol{\nu},n,s) = \tilde{\omega}_{\rm ra}(\boldsymbol{\nu},n,s) = 1/g(n,s)$ satisfy the hypotheses of Corollary 4 and 5, provided that $g(n,s)$ decays algebraically with n. Since one has an extensible lattice n_{m+1}/n_m is bounded below by two. It is natural to define $N(\mathbf{0}) = 0$ and for $\boldsymbol{\nu} \neq \mathbf{0}$ to define $N(\boldsymbol{\nu})$ as the largest $n \in \mathcal{N}$ such that $\boldsymbol{\nu} \in L_n^\perp$. It follows that

$$\|f\|_{\text{lat},s,1} = \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n} \frac{|F(\boldsymbol{\nu})|}{g(N(\boldsymbol{\nu}),s)}, \quad \|f\|^2_{\text{ra-lat},s,1} = \sup_{n\in\mathcal{N}} \sum_{\boldsymbol{\nu}:N(\boldsymbol{\nu})=n} \left|\frac{F(\boldsymbol{\nu})}{g(N(\boldsymbol{\nu}),s)}\right|^2.$$

are equivalent norms to $\|f\|_{\text{lat},s}$ and $\|f\|_{\text{ra-lat},s}$, respectively, since $C_{L,s} = 1$ in (17) and (19) for this case. Slightly stronger, but simpler to compute, norms are

$$\|f\|_{\text{lat},s,2} = \sum_{\boldsymbol{\nu}} \frac{|F(\boldsymbol{\nu})|}{g(N(\boldsymbol{\nu}),s)}, \quad \|f\|^2_{\text{ra-lat},s,2} = \sum_{\boldsymbol{\nu}} \left|\frac{F(\boldsymbol{\nu})}{g(N(\boldsymbol{\nu}),s)}\right|^2.$$

An important question is how small $N(\boldsymbol{\nu})$ can be made by a good choice of integration lattice. A smaller $N(\boldsymbol{\nu})$ implies a larger $g(N(\boldsymbol{\nu}),s)$ and therefore a larger set of integrands that can be integrated accurately by shifted lattice rules. Consider, for example, the extensible rank-1 lattice with the generating vector

$$\mathbf{h} = (1, 17797)^T \quad \text{with} \quad \mathcal{N} = \{0,1,2,4,\ldots\} \tag{21}$$

proposed in [8]. The values of $N(\nu_1,\nu_2)$ are plotted in Figure 2.

5.2. Bounds on $N(\boldsymbol{\nu})$. Good theoretical upper bounds on $N(\boldsymbol{\nu})$ are difficult to obtain, however, one may obtain lower bounds on $g(N(\boldsymbol{\nu}),s)$ as a function of $\bar{\nu}_1 \cdots \bar{\nu}_s$, where the notation

$$\bar{\nu} := \begin{cases} 1, & \nu = 0, \\ |\nu|, & \nu \neq 0, \end{cases}$$

is commonly used in the study of lattice rules. Below is an example of a new lower bound, which is an extension of a result in [9].

Theorem 7. *For $g(n;\alpha,\delta)$ defined in* (13) *there exist extensible rank-1 lattice rules with $g(N(\boldsymbol{\nu});\alpha,\delta) \geq c(\boldsymbol{\gamma},\alpha,\delta)[\overline{\nu_1/\gamma_1} \cdots \overline{\nu_s/\gamma_s}]^{-\alpha}$ for any $s,\alpha \geq 1$ and $\boldsymbol{\gamma} = (\gamma_1,\gamma_2,\ldots)$ satisfying $\sum_{j=1}^\infty \gamma_j^\delta < \infty$, where $c(\boldsymbol{\gamma},\alpha,\delta)$ is independent of the dimension s.*

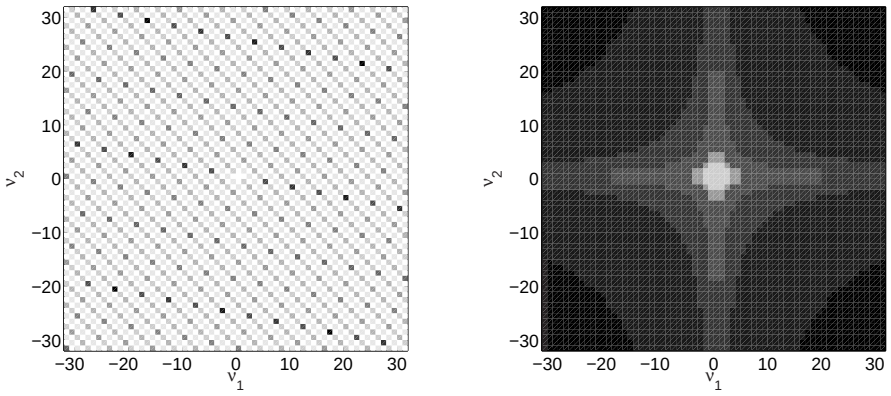

FIGURE 2. Plots of $N(\nu_1, \nu_2)$ (left) and the upper bound $N_U(\nu_1, \nu_2)$ (right) for the rank-1 integration lattice defined in (21). Lighter shades of gray denote smaller values.

The proof of this theorem relies on the lemma below. This lemma and its proof are similar to [9, Lemma 3].

Lemma 8. *Given a fixed* $\boldsymbol{\gamma} \in [0,\infty)^\infty$ *and* δ *with* $0 < \delta < 1$, *let* $S(\boldsymbol{\gamma}, m) = \prod_{j=1}^\infty [1 + \gamma_j m]$, *for* $m \ge 0$. *If* $\sum_{j=1}^\infty \gamma_j^\delta < \infty$, *then it follows that* $b^{-m} S(\boldsymbol{\gamma}, m) \le \tilde{C}(\boldsymbol{\gamma}, d) g(b^m; 1, \delta)$ *for* $m \ge 0$.

Proof. Let $\sigma_s = \sum_{j=s+1}^\infty \gamma_j^\delta$, $s = 0, 1, \ldots$. Note that by increasing s one may make σ_s arbitrarily small. We can assume without loss of generality that all $\sigma_s > 0$. Then applying some relatively elementary inequalities yields:

$$\begin{aligned}
\log[S(\boldsymbol{\gamma}, m)] &= \sum_{j=1}^\infty \log[1 + \gamma_j m] \\
&\le \sum_{j=1}^s \log[1 + \sigma_s^{-1} + \gamma_j m] + \sum_{j=s+1}^\infty \log[1 + \gamma_j m] \\
&= s \log(1 + \sigma_s^{-1}) + \sum_{j=1}^s \log[1 + \gamma_j m/(1 + \sigma_s^{-1})] \\
&\qquad + \sum_{j=s+1}^\infty \log[1 + \gamma_j m] \\
&\le s \log(1 + \sigma_s^{-1}) + \frac{\sigma_s^\delta m^\delta}{\delta} \sum_{j=1}^s \gamma_j^\delta + \frac{m^\delta}{\delta} \sum_{j=s+1}^\infty \gamma_j^\delta \\
&\le s \log(1 + \sigma_s^{-1}) + \frac{(\sigma_0 \sigma_s^\delta + \sigma_s) m^\delta}{\delta}.
\end{aligned}$$

From this upper bound it follows that

$$\log\left[\frac{b^{-m} S(\boldsymbol{\gamma}, m)}{g(b^m; 1, \delta)}\right] \le s \log(1 + \sigma_s^{-1}) + \frac{(\sigma_0 \sigma_s^\delta + \sigma_s) m^\delta}{\delta} + 1 - (1 + \delta^{-1} m \log b)^\delta.$$

Choosing s large enough makes σ_s small enough so that this quantity bounded above for all $m \ge 0$, which completes the proof. □

Proof of Theorem 7. This proof follows the arguments used to prove [9, Theorem 2] but replaces [9, Lemma 3] by Lemma 8. Just the outline of the proof is given here. For any $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots)$ define the following quantity, which is often used in the analysis of lattice rules:

$$R_\alpha(n) = \sum_{\boldsymbol{\nu} \in L_n^\perp \cap (-n/2, n/2]^s} [\overline{\nu_1/\gamma_1} \cdots \overline{\nu_s/\gamma_s}]^{-\alpha}.$$

Using arguments similar to those in [9] it can be shown that if $\sum_{j=1}^\infty \gamma_j^\delta < \infty$ and $\alpha > 1$, then there exist extensible lattice rules satisfying $R_\alpha(n) \le g(n; \alpha, \delta)/c(\boldsymbol{\gamma}, \alpha, \delta)$ for some $c(\boldsymbol{\gamma}, \alpha, \delta)$ independent of n and s. This implies that

$$[\overline{\nu_1/\gamma_1} \cdots \overline{\nu_s/\gamma_s}]^{-\alpha} \le R_\alpha(N(\boldsymbol{\nu})) \le g(N(\boldsymbol{\nu}); \alpha, d)/c(\boldsymbol{\gamma}, \alpha, d),$$

which completes the proof. □

The lower bound in Theorem 7 is rather conservative, although the author is not aware of substantially better lower bounds. Ignoring for a moment $\boldsymbol{\gamma}$ and the effect of δ, the bound in Theorem 7 in essence replaces $N(\boldsymbol{\nu})$ by

$$N_U(\boldsymbol{\nu}) := \max_{\overline{\lambda_1} \cdots \overline{\lambda_s} = \overline{\nu_1} \cdots \overline{\nu_s}} N(\boldsymbol{\lambda}).$$

This upper bound, plotted in Figure 2, is much larger than the actual values of $N(\boldsymbol{\nu})$.

6. Extensible Digital Net Rules

For prime bases b, a linearly scrambled digital sequence is defined by [14, 17, 12, 13, 10]:

$$\mathbf{x}_i = (0.x_{i11}x_{i12}\cdots, 0.x_{i21}x_{i22}\cdots, \ldots) \text{ base } b, \quad i = \cdots i_2 i_1 \text{ base } b,\ i = 0, 1, \ldots$$

$$\begin{pmatrix} x_{ij1} \\ x_{ij2} \\ \vdots \end{pmatrix} = \mathbf{L}_j \mathbf{C}_j \begin{pmatrix} i_1 \\ i_2 \\ \vdots \end{pmatrix} + \mathbf{e}_j \mod b, \quad j = 1, \ldots, s$$

where the $\mathbf{C}_j$ are the prescribed $\infty \times \infty$ generator matrices, the $\mathbf{L}_j$ are lower triangular $\infty \times \infty$ scrambling matrices, and the $\mathbf{e}_j$ are $\infty \times 1$ digital shifts. The scrambling matrices and shifts may be chosen deterministically or randomly.

For example, the design in Figure 1 corresponds to the first nine points of digitally-shifted Faure sequence [5], i.e.,

$$b = 3, \quad \mathbf{C}_1 = \mathbf{L}_1 = \mathbf{L}_2 = \mathbf{I}, \quad \mathbf{C}_2 = \begin{pmatrix} 1 & 1 & 1 & \cdots \\ 0 & 1 & 2 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \mathbf{e}_1 = \mathbf{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}.$$

Choosing the first $n = b^m$ points of a digital sequence gives a digital net, so for digital net rules one has $\mathcal{N} = \{0, 1, b, b^2, \ldots\}$. This article considers numerical integration rules based on nets because nets generally have better uniformity than an arbitrary piece of a digital sequence. The quality of a net is measured by the quality of the generator matrices, $\mathbf{C}_j$.

For any non-negative vector $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_s)$ let $\|\boldsymbol{\ell}\|_1$ denote the sum of the elements of $\boldsymbol{\ell}$. Let $U(\boldsymbol{\ell}) \subseteq \{1, \ldots, s\}$ denote the set of all j for which $\ell_j \ne 0$, and let $|U(\boldsymbol{\ell})|$ denote the cardinality of $U(\boldsymbol{\ell})$, i.e., the number of nonzero components of $\boldsymbol{\ell}$. Let $\mathbf{C}(\boldsymbol{\ell})$ be the $\|\boldsymbol{\ell}\|_1 \times \infty$ matrix formed by the first ℓ_1 rows of $\mathbf{C}_1$ followed by the first ℓ_2 rows of $\mathbf{C}_2$, etc. For any non-negative integer m, let $\mathbf{C}(\boldsymbol{\ell}, m)$ denote

the $\|\boldsymbol{\ell}\|_1 \times m$ matrix formed by the first m columns of $\mathbf{C}(\boldsymbol{\ell})$. Let $\operatorname{null}(\cdot)$ denote the null space of a matrix, and $\operatorname{def}(\cdot)$ denote the defect of a matrix, i.e., the dimension of its null space.

For a fixed $u \subseteq \{1, \dots, s\}$ the parameter T_u is the smallest number T for which any $\boldsymbol{\ell}$ with $\|\boldsymbol{\ell}\|_1 \le m - T$ and $U(\boldsymbol{\ell}) \subseteq u$ yields $\operatorname{def}(\mathbf{C}^T(\boldsymbol{\ell}, m)) = 0$, i.e., $\mathbf{C}(\boldsymbol{\ell}, m)$ has linear independent rows. In other words,

$$T_u := \min\{T : \operatorname{def}(\mathbf{C}^T(\boldsymbol{\ell}, m)) = 0 \quad \forall m \ge T, \forall \boldsymbol{\ell} \text{ with } U(\boldsymbol{\ell}) \subseteq u \ \&\ \|\boldsymbol{\ell}\|_1 \le m - T\}. \tag{22}$$

Smaller values of T_u correspond to better nets. By convention $T_\emptyset = 0$. It follows from the definition that $0 = T_\emptyset \le T_v \le T_u$ for $v \subseteq u$, and T_u provides an upper bound on $\operatorname{def}(\mathbf{C}^T(\boldsymbol{\ell}, m))$ for all m:

$$\max(0, \|\boldsymbol{\ell}\|_1 - m) \le \operatorname{def}(\mathbf{C}^T(\boldsymbol{\ell}, m)) \le \max(0, \|\boldsymbol{\ell}\|_1 - m + T_{U(\boldsymbol{\ell})}). \tag{23}$$

6.1. Walsh Functions and Haar Wavelets. Just as trigonometric polynomials are well-suited to the analysis of lattice rules, Walsh functions in base b are well-suited to the analysis of digital net rules. For any $\boldsymbol{\nu} \in \mathbb{Z}_+^s$, the set of non-negative s-dimensional integer vectors, let $\nu_{j\ell}$ denote the ℓth b-ary digit of the jth component of $\boldsymbol{\nu}$ represented in base b. Moreover, for any $\mathbf{x} \in [0,1)^s$, let $x_{\cdot j\ell}$ denote the ℓth b-ary digit of the jth component of $\mathbf{x}$. Multivariate Walsh functions are defined as

$$\psi_{\mathrm{Wa},\boldsymbol{\nu}}(\mathbf{x}) := \exp\left(\frac{2\pi\imath}{b} \sum_{j,\ell=1}^{s,\infty} \nu_{j\ell} x_{\cdot j\ell}\right), \quad \mathbf{x} \in [0,1)^s, \ \boldsymbol{\nu} \in \mathbb{Z}_+^s.$$

Before analyzing the error of scrambled digital net rules applied to Walsh functions it is helpful to define some further notation. For any $\nu \in \mathbb{Z}_+$ define the function

$$\lg(\nu) := \begin{cases} \lfloor \log_b(\nu) \rfloor + 1 & \text{for } \nu > 0 \\ 0 & \text{for } \nu = 0, \end{cases}$$

where $\lfloor \cdot \rfloor$ denotes the largest integer part of a real number or vector, i.e., the base b representation of ν has $\lg(\nu)$ digits if one ignores leading zeros. For any set $u \subseteq \{1, \dots, s\}$ let $\mathbf{1}_u$ denote the vector whose jth component is 1 for $j \in u$ and 0 otherwise. Let $\delta(\mathbf{x}, \mathbf{y}, \boldsymbol{\ell}) = 1$ if the first ℓ_j digits of x_j and y_j are the same for all $j = 1, \dots, s$ and let $\delta(\mathbf{x}, \mathbf{y}, \boldsymbol{\ell}) = 0$ otherwise. Let $\operatorname{digit}(\boldsymbol{\nu}, \boldsymbol{\ell})$ denote the $\|\boldsymbol{\ell}\|_1$-vector which is formed by the first $\boldsymbol{\ell}$ digits of $\boldsymbol{\nu}$, more precisely, let $\operatorname{digit}(\boldsymbol{\nu}, \boldsymbol{\ell}) = (\nu_{1,1}, \dots, \nu_{1,\ell_1}, \dots, \nu_{s,1}, \dots, \nu_{s,\ell_s})^T$. Let $\tilde{\mathbf{C}}(\boldsymbol{\ell}, m)$ be the analog of $\mathbf{C}(\boldsymbol{\ell}, m)$ formed from the scrambled generator matrices, $\tilde{\mathbf{C}}_j := \mathbf{L}_j \mathbf{C}_j$, and let $\mathbf{e}(\boldsymbol{\ell})$ denote the analog of $\mathbf{C}(\boldsymbol{\ell})$ formed from the digital shifts $\mathbf{e}_j$. The following lemma provides bounds on the errors of deterministic and randomly scrambled digital sequence rules applied to Walsh functions.

Lemma 9. *For scrambled net rules applied to Walsh functions one has the following absolute errors for $m = 0, 1, \dots$:*

$$|\operatorname{err}(\psi_{Wa,\boldsymbol{\nu}}, net, b^m)| = \begin{cases} 1 & \text{for } \boldsymbol{\nu} \ne \mathbf{0} \text{ and } \operatorname{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu})) \in \operatorname{null}(\tilde{\mathbf{C}}^T(\lg(\boldsymbol{\nu}), m)), \\ 0 & \text{for } \boldsymbol{\nu} = \mathbf{0} \text{ or } \operatorname{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu})) \notin \operatorname{null}(\tilde{\mathbf{C}}^T(\lg(\boldsymbol{\nu}), m)). \end{cases} \tag{24a}$$

Walsh functions are integrated exactly if $m \geq \|\lg(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}$. *Moreover, the root mean square errors for randomly scrambled net rules applied to Walsh functions are*

$$\text{rmse}(\psi_{Wa,\boldsymbol{\nu}}, \textit{ra-net}, b^m) \leq 3^{|U(\boldsymbol{\nu})|/2} b^{(T_{U(\boldsymbol{\nu})}-m)/2}, \tag{24b}$$

$$\text{rmse}(\psi_{Wa,\boldsymbol{\nu}}, \textit{ra-net}, b^m) = 0 \quad \textit{for } m \geq \|\lg(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}. \tag{24c}$$

Finally, for randomly scrambled net rules and Walsh functions condition (11) *is satisfied.*

Proof. The integrals of Walsh functions vanish except in the case of $\boldsymbol{\nu} = \mathbf{0}$, i.e., $I(\psi_{\text{Wa},\boldsymbol{\nu}}) = \delta_{\boldsymbol{\nu},\mathbf{0}}$. Moreover, the digital net rule with $n = b^m$ points applied to a Walsh function is

$$\begin{aligned} &Q_n(\psi_{\text{Wa},\boldsymbol{\nu}}; \text{net}) \\ &\quad = \frac{1}{b^m} \sum_{i=0}^{b^m-1} \exp\left\{ \frac{2\pi\imath}{b} \,\text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \left[\tilde{\mathbf{C}}(\lg(\boldsymbol{\nu}), m) \,\text{digit}(i, m) + \mathbf{e}(\lg(\boldsymbol{\nu})) \right] \right\} \\ &\qquad\qquad = \begin{cases} 0 & \text{for } \text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \tilde{\mathbf{C}}(\lg(\boldsymbol{\nu}), m) \neq \mathbf{0} \mod b, \\ \exp\left\{ \frac{2\pi\imath}{b} \,\text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \mathbf{e}(\lg(\boldsymbol{\nu})) \right\} & \text{otherwise.} \end{cases} \end{aligned}$$

This implies (24a). If $m \geq \|\lg(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}$ then it follows from the definition in (22) that $\text{def}(\tilde{\mathbf{C}}^T(\lg(\boldsymbol{\nu}), m)) = 0$, and so $|\text{err}(\psi_{\text{Wa},\boldsymbol{\nu}}, \text{net}, b^m)| = 0$. This completes the proof of the sentence after (24a) and statement (24c).

For randomly scrambled net rules it can be shown that

$$\begin{aligned} \psi_{\text{Wa},\boldsymbol{\nu}}(\mathbf{x}_i) \bar{\psi}_{\text{Wa},\boldsymbol{\lambda}}(\mathbf{x}_k) = \exp\Big\{ \frac{2\pi\imath}{b} \Big[&\text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \tilde{\mathbf{C}}(\lg(\boldsymbol{\nu}), m) \,\text{digit}(i, m) \\ &- \text{digit}(\boldsymbol{\lambda}, \lg(\boldsymbol{\lambda}))^T \tilde{\mathbf{C}}(\lg(\boldsymbol{\lambda}), m) \,\text{digit}(k, m) \\ &+ \text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \mathbf{e}(\lg(\boldsymbol{\nu})) - \text{digit}(\boldsymbol{\lambda}, \lg(\boldsymbol{\lambda}))^T \mathbf{e}(\lg(\boldsymbol{\lambda})) \Big] \Big\}. \end{aligned}$$

If $\boldsymbol{\nu} \neq \boldsymbol{\lambda}$, then the expectation of the factor corresponding to the digital shift vanishes:

$$E\left[\exp\left\{ \frac{2\pi\imath}{b} \left[\text{digit}(\boldsymbol{\nu}, \lg(\boldsymbol{\nu}))^T \mathbf{e}(\lg(\boldsymbol{\nu})) - \text{digit}(\boldsymbol{\lambda}, \lg(\boldsymbol{\lambda}))^T \mathbf{e}(\lg(\boldsymbol{\lambda})) \right] \right\} \right] = 0,$$

proving (11). If $\boldsymbol{\nu} = \boldsymbol{\lambda}$, then following a line of argument similar to that in [18, Theorem 1] yields

$$\begin{aligned} E\left[\psi_{\text{Wa},\boldsymbol{\nu}}(\mathbf{x}_i) \bar{\psi}_{\text{Wa},\boldsymbol{\nu}}(\mathbf{x}_k) \right] &= E\left[\exp\left\{ \frac{2\pi\imath}{b} \sum_{j,\ell=1}^{s,\infty} \nu_{jl}(x_{ij\ell} - x_{kj\ell}) \right\} \right] \\ &= \prod_{j \in U(\boldsymbol{\nu})} \frac{b\delta(x_{ij}, x_{kj}, \lg(\nu_j)) - \delta(x_{ij}, x_{kj}, \lg(\nu_j) - 1)}{b - 1} \\ &= \delta(\mathbf{x}_i, \mathbf{x}_k, \lg(\boldsymbol{\nu}) - \mathbf{1}_{U(\boldsymbol{\nu})}) \prod_{j \in U(\boldsymbol{\nu})} \frac{b\delta(x_{ij}, x_{kj}, \lg(\nu_j)) - 1}{b - 1} \\ &= \sum_{v \subseteq U(\boldsymbol{\nu})} \frac{(-1)^{|v|} b^{|U(\boldsymbol{\nu})|-|v|} \delta(\mathbf{x}_i, \mathbf{x}_k, \lg(\boldsymbol{\nu}) - \mathbf{1}_v)}{(b-1)^{U(\boldsymbol{\nu})}}. \end{aligned} \tag{25}$$

The sum of $\delta(\mathbf{x}_i, \mathbf{x}_k, \boldsymbol{\ell})$ over k depends on the defect of $\mathbf{C}(\boldsymbol{\ell}, m)$ as follows:

$$\frac{1}{b^m} \sum_{k=0}^{b^m-1} \delta(\mathbf{x}_i, \mathbf{x}_k, \boldsymbol{\ell}) = b^{\operatorname{def}(\mathbf{C}(\boldsymbol{\ell},m))} \quad \forall i = 0, \dots, b^m - 1.$$

Because $\mathbf{C}(\boldsymbol{\ell}, m)$ and $\mathbf{C}^T(\boldsymbol{\ell}, m)$ have the same column rank, their defects are related by

$$\begin{aligned} \max(0, m - \|\boldsymbol{\ell}\|_1) \le \operatorname{def}(\mathbf{C}(\boldsymbol{\ell}, m)) &= m - \|\boldsymbol{\ell}\|_1 + \operatorname{def}(\mathbf{C}^T(\boldsymbol{\ell}, m)) \\ &\le \max(m - \|\boldsymbol{\ell}\|_1, T_{U(\boldsymbol{\ell})}), \end{aligned} \tag{26}$$

applying (23). Then (25) implies that the expected value of the square absolute value of the quadrature rule for a Walsh function is

$$\begin{aligned} &E[Q(\psi_{\mathrm{Wa},\boldsymbol{\nu}}; \mathrm{net}) Q(\bar{\psi}_{\mathrm{Wa},\boldsymbol{\nu}}; \mathrm{net})] \\ &\quad = \left(\frac{b}{b-1}\right)^{|U(\boldsymbol{\nu})|} \sum_{v \subseteq U(\boldsymbol{\nu})} (-1)^{|v|} \left[b^{\operatorname{def}(\mathbf{C}(\lg(\boldsymbol{\nu}) - \mathbf{1}_v, m)) - |v| - m} - b^{-\|\lg(\boldsymbol{\nu})\|_1} \right], \end{aligned} \tag{27}$$

where the second term in the sum facilitates the derivation, but makes no contribution because $\sum_{v \subseteq U(\boldsymbol{\nu})} (-1)^{|v|} = 0$.

By (26) it follows that the term contained in the square brackets in (27) is nonnegative. For all sets $v \subseteq U(\boldsymbol{\nu})$ with $m - \|\lg(\boldsymbol{\nu})\|_1 + |v| \ge T_{U(\boldsymbol{\nu})}$ it follows from (26) that $\operatorname{def}(\mathbf{C}(\lg(\boldsymbol{\nu}) - \mathbf{1}_v, m)) = m - \|\lg(\boldsymbol{\nu})\|_1 + |v|$, so the term in the square brackets in (27) vanishes. For all sets $v \subseteq U(\boldsymbol{\nu})$ with $m - \|\lg(\boldsymbol{\nu})\|_1 + |v| < T_{U(\boldsymbol{\nu})}$ it follows from (26) that $\operatorname{def}(\mathbf{C}((\lg(\boldsymbol{\nu}) - \mathbf{1}_v, m)) \le T_{U(\boldsymbol{\nu})}$. Together with the binomial theorem this implies that

$$\begin{aligned} E[Q(\psi_{\mathrm{Wa},\boldsymbol{\nu}}; \mathrm{net}) Q(\bar{\psi}_{\mathrm{Wa},\boldsymbol{\nu}}; \mathrm{net})] &\le \left(\frac{b}{b-1}\right)^{|U(\boldsymbol{\nu})|} \sum_{v \subseteq U(\boldsymbol{\nu})} b^{T_{U(\lg(\boldsymbol{\nu}) - \mathbf{1}_v)} - |v| - m} \\ &\le \left(\frac{b+1}{b-1}\right)^{|U(\boldsymbol{\nu})|} b^{T_{U(\boldsymbol{\nu})} - m} \le 3^{|U(\boldsymbol{\nu})|} b^{T_{U(\boldsymbol{\nu})} - m}, \end{aligned}$$

since $b \ge 2$. This establishes (24b). □

As an alternative to Walsh functions, Haar wavelets are another set of basis functions used to analyze digital net rules. Like Walsh functions they are piecewise constant, but whereas Walsh functions have their support as the whole unit cube, Haar wavelets have their support as a box of volume $b^{|U(\boldsymbol{\nu})| - \|\lg(\boldsymbol{\nu})\|_1}$. Multivariate Haar wavelets may be defined as orthonormal linear combinations of Walsh

functions, as follows:

$$\psi_{\mathrm{Ha},\boldsymbol{\nu}}(\mathbf{x})$$

$$:= \frac{1}{b^{(\|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})|)/2}} \sum_{\substack{\boldsymbol{\lambda}:\mathrm{lg}(\boldsymbol{\lambda})=\mathrm{lg}(\boldsymbol{\nu}) \\ \lambda_{j\,\mathrm{lg}(\lambda_j)}=\nu_{j\,\mathrm{lg}(\nu_j)}}} \exp\left(\frac{-2\pi\imath}{b}\sum_{j,\ell=1}^{s,\infty}\lambda_{j\ell}\nu_{j\ell}\right)\psi_{\mathrm{Wa},\boldsymbol{\lambda}}(\mathbf{x}) \tag{28}$$

$$= \frac{1}{b^{(\|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})|)/2}} \sum_{\substack{\boldsymbol{\lambda}:\mathrm{lg}(\boldsymbol{\lambda})=\mathrm{lg}(\boldsymbol{\nu}) \\ \lambda_{j\,\mathrm{lg}(\lambda_j)}=\nu_{j\,\mathrm{lg}(\nu_j)}}} \exp\left(\frac{2\pi\imath}{b}\sum_{j,\ell=1}^{s,\infty}\lambda_{j\ell}(x_{\cdot j\ell}-\nu_{j\ell})\right)$$

$$= b^{(\|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})|)/2} \exp\left(\frac{2\pi\imath}{b}\sum_{j=1}^{s}\nu_{j\,\mathrm{lg}(\nu_j)}x_{\cdot j\,\mathrm{lg}(\nu_j)}\right)\delta(\mathbf{x},\mathbf{z}_{\boldsymbol{\nu}},\mathrm{lg}(\boldsymbol{\nu})-\mathbf{1}_{U(\boldsymbol{\nu})}), \tag{29}$$

where $\mathbf{z}_{\boldsymbol{\nu}} := (z_{\nu_1},\dots,z_{\nu_s})^T$ is defined by $\nu_j = \nu_{j\,\mathrm{lg}(\nu_j)} + z_{\nu_j} b^{\mathrm{lg}(\nu)-1}$, i.e., $\mathbf{z}_{\boldsymbol{\nu}} \in [0,1)^s$ is comprised of all but the most significant digits of $\boldsymbol{\nu}$.

Lemma 10. *For scrambled net rules applied to Haar wavelets one has the following absolute errors for $n = b^m$, $m = 0,1,\dots$:*

$$|\mathrm{err}(\psi_{Ha,\boldsymbol{\nu}}, net, b^m)| \le b^{(\|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})|)/2 + T_{U(\boldsymbol{\nu})} - m} \quad \text{for } m < \|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})| + T_{U(\boldsymbol{\nu})}, \tag{30a}$$

$$|\mathrm{err}(\psi_{Ha,\boldsymbol{\nu}}, net, b^m)| \le b^{(|U(\boldsymbol{\nu})| - \|\mathrm{lg}(\boldsymbol{\nu})\|_1)/2} \quad \text{for } \|\mathrm{lg}(\boldsymbol{\nu})\|_1 - |U(\boldsymbol{\nu})| + T_{U(\boldsymbol{\nu})} \le m < \|\mathrm{lg}(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}, \tag{30b}$$

$$|\mathrm{err}(\psi_{Ha,\boldsymbol{\nu}}, net, b^m)| = 0 \qquad \text{for } m \ge \|\mathrm{lg}(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}. \tag{30c}$$

Moreover, the root mean square errors for randomly scrambled net rules applied to Haar wavelets are

$$\mathrm{rmse}(\psi_{Ha,\boldsymbol{\nu}}, ra\text{-}net, b^m) \le 3^{|U(\boldsymbol{\nu})|/2} b^{(T_{U(\boldsymbol{\nu})}-m)/2}, \tag{30d}$$

$$\mathrm{rmse}(\psi_{Ha,\boldsymbol{\nu}}, ra\text{-}net, b^m) = 0 \quad \text{for } m \ge \|\mathrm{lg}(\boldsymbol{\nu})\|_1 + T_{U(\boldsymbol{\nu})}. \tag{30e}$$

Finally, for randomly scrambled net rules and Haar wavelets condition (11) *is satisfied.*

Proof. For small m the worst that may happen is that a proportion of $b^{\mathrm{def}(\mathbf{C}(\mathrm{lg}(\boldsymbol{\nu})-\mathbf{1}_{U(\boldsymbol{\nu})},m))-m}$ of the sample points fall inside the support of the wavelet. This implies that the worst possible numerical integration error is this proportion times the magnitude of the wavelet, i.e.,

$$\mathrm{err}(\psi_{\mathrm{Ha},\boldsymbol{\nu}}, \text{ra-net}, b^m) \le b^{(\|\mathrm{lg}(\boldsymbol{\nu})\|_1 - U(\boldsymbol{\nu}))/2 + \mathrm{def}(\mathbf{C}(\mathrm{lg}(\boldsymbol{\nu})-\mathbf{1}_{U(\boldsymbol{\nu})},m))-m}.$$

The bound on $\mathrm{def}(\mathbf{C}(\boldsymbol{\ell},m))$ in (26) derived in the proof of the previous lemma implies error bounds (30a) and (30b). Haar wavelets are integrated exactly under the analogous condition as for Walsh functions in Lemma 9, which implies (30c). Since Haar wavelets are orthonormal transformations of Walsh functions, (30d) and (30e) follow from Lemma 9. □

6.2. Error Analysis for Scrambled Digital Net Rules. For series of Walsh functions $\sup_{\mathbf{x}\in[0,1)^s}|\psi_{\mathrm{Wa},\boldsymbol{\nu}}(\mathbf{x})|=1$ for all $\mathbf{x}$, so it makes sense to define the space of possible functions for worst-case analysis as

$$\mathcal{C}_{\mathrm{Wa}}([0,1)^s):=\{f:\sum_{\boldsymbol{\nu}}|F(\boldsymbol{\nu})|<\infty\},$$

analogously to the case of trigonometric polynomials. On the other hand, Haar wavelets have support that is in general only a subset of the unit cube, in contrast to trigonometric polynomials and Walsh functions. For a given $\boldsymbol{\ell}\in\mathbb{Z}_+^s$ the Haar wavelets in the set $\{\psi_{\mathrm{Ha},\boldsymbol{\nu}}(\mathbf{x}):\lg(\boldsymbol{\nu})=\boldsymbol{\ell}\}$ have the same magnitude and mutually disjoint support. Therefore, one may define

$$\mathcal{C}_{\mathrm{Ha}}([0,1)^s):=\left\{f:\sum_{\boldsymbol{\ell}}\left[\sup_{\boldsymbol{\nu}:\lg(\boldsymbol{\nu})=\boldsymbol{\ell}}b^{(\|\boldsymbol{\ell}\|_1-|U(\boldsymbol{\ell})|)/2}\,|F(\boldsymbol{\nu})|\right]<\infty\right\}.$$

For both the deterministic and randomly scrambled net rules applied to multivariate Walsh or Haar series it is rather tedious to obtain exact formulas for $|\mathrm{err}(\psi_{\boldsymbol{\nu}},\mathrm{net},b^m)|$ and $\mathrm{rmse}(\psi_{\boldsymbol{\nu}},\mathcal{Q}_n,\mu)$, for all $\boldsymbol{\nu}$. Instead, the upper bounds in Lemmas 9–10 are used with Corollary 4 to obtain norms that are probably slightly stronger than the one defined in Theorem 1. Recall that $\mathcal{N}=\{0,1,b,b^2,\dots\}$, and note that the upper bound on the number of points needed to integrate $\psi_{\mathrm{x},\boldsymbol{\nu}}$ exactly for $\mathrm{x}=\mathrm{Wa},\mathrm{Ha}$ is $N(\boldsymbol{\nu})=b^{\|\lg(\boldsymbol{\nu})\|_1+T_{U(\boldsymbol{\nu})}}$ for $\boldsymbol{\nu}\neq\mathbf{0}$ and $N(\mathbf{0})=0$. It is convenient to choose desired convergence rate as a simple power of n, $g(n,s)=g(n,s;\alpha,0)=\min(1,n^{-\alpha})$ for both the worst-case and randomized settings, avoiding the need for $\log n$ terms as is the case for lattice rules.

The weights $\tilde{\omega}_{\mathrm{wo}}(\boldsymbol{\nu},n,s)$ and $\tilde{\omega}_{\mathrm{ra}}(\boldsymbol{\nu},n,s)$ differ from each other for digital net rules, as opposed to lattice rules. Moreover, in the worst-case setting they differ for Walsh functions and Haar wavelets. Specifically, they are

$$\tilde{\omega}_{\mathrm{wo,Wa}}(\boldsymbol{\nu},b^m,s)=b^{\alpha m},$$
$$\tilde{\omega}_{\mathrm{wo,Ha}}(\boldsymbol{\nu},b^m,s)=b^{(|U(\boldsymbol{\nu})|-\|\lg(\boldsymbol{\nu})\|_1)/2+\alpha m+\max(0,\|\lg(\boldsymbol{\nu})\|_1-|U(\boldsymbol{\nu})|+T_{U(\boldsymbol{\nu})}-m)},$$
$$\tilde{\omega}_{\mathrm{ra,x}}(\boldsymbol{\nu},b^m,s)=3^{|U(\boldsymbol{\nu})|/2}b^{T_{U(\boldsymbol{\nu})}/2+m(\alpha-1/2)},\quad \mathrm{x}=\mathrm{Wa},\mathrm{Ha}.$$

It follows that

$$\|f\|_{\mathrm{net,Wa},1}:=\sup_{m=0,1,\dots}\sum_{\boldsymbol{\nu}:\|\lg(\boldsymbol{\nu})\|_1+T_{U(\boldsymbol{\nu})}=m}b^{\alpha[T_{U(\boldsymbol{\nu})}+\|\lg(\boldsymbol{\nu})\|_1]}\,|F(\boldsymbol{\nu})|\,,$$
$$\|f\|_{\mathrm{net,Wa},2}:=\sum_{\boldsymbol{\nu}}b^{\alpha[T_{U(\boldsymbol{\nu})}+\|\lg(\boldsymbol{\nu})\|_1]}\,|F(\boldsymbol{\nu})|\,,$$
$$\|f\|_{\mathrm{net,Ha},1}:=\sup_{m=0,1,\dots}\sum_{\boldsymbol{\nu}:\|\lg(\boldsymbol{\nu})\|_1+T_{U(\boldsymbol{\nu})}=m}b^{\alpha T_{U(\boldsymbol{\nu})}+|U(\boldsymbol{\nu})|/2+(\alpha-1/2)\|\lg(\boldsymbol{\nu})\|_1}\,|F(\boldsymbol{\nu})|\,,$$
$$\|f\|_{\mathrm{net,Ha},2}:=\sum_{\boldsymbol{\nu}}b^{\alpha T_{U(\boldsymbol{\nu})}+|U(\boldsymbol{\nu})|/2+(\alpha-1/2)\|\lg(\boldsymbol{\nu})\|_1}\,|F(\boldsymbol{\nu})|\,,$$
$$\|f\|^2_{\mathrm{ra\text{-}net,x},1}:=\sup_{m=0,1,\dots}\sum_{\substack{\boldsymbol{\nu}:\|\lg(\boldsymbol{\nu})\|_1+T_{U(\boldsymbol{\nu})}=m\\ \boldsymbol{\nu}\neq\mathbf{0}}}3^{|U(\boldsymbol{\nu})|}b^{2\alpha T_{U(\boldsymbol{\nu})}+(2\alpha-1)\|\lg(\boldsymbol{\nu})\|_1}\,|F(\boldsymbol{\nu})|^2\,,$$
$$\|f\|^2_{\mathrm{ra\text{-}net,x},2}:=\sum_{\boldsymbol{\nu}}3^{|U(\boldsymbol{\nu})|}b^{2\alpha T_{U(\boldsymbol{\nu})}+(2\alpha-1)\|\lg(\boldsymbol{\nu})\|_1}\,|F(\boldsymbol{\nu})|^2$$

for $\mathrm{x}=\mathrm{Wa},\mathrm{Ha}$. To ensure that the norms for the randomized setting only admit square integrable functions it is assumed that $\alpha\geq 1/2$.

The quality factor for nets, T_u, plays an important role in the definition of these norms. The values of $T_{\{1,\ldots,s\}}$ increase as s increases. From [14] one has the following upper bounds on b^{T_u}:

$$2^{T_u} \le \prod_{j\in u} [C_{\mathrm{Sob}} j \log_2(j+1)\log_2(\log_2(j+3))] \qquad \text{for Sobol' sequences } (b=2),$$

$$b^{T_u} \le \prod_{j\in u} [2j \log_b(j+b)\log_2(\log_2(j+3))] \qquad \text{for Niederreiter sequences.}$$

Unfortunately, even for $|u|$ fixed, T_u may become large if $\max(u)$ becomes large.

7. Some examples

This section provides a couple of examples to illustrate how the algorithm-driven approach can explain certain behaviours of the integration error not explained using traditional error analysis. For simplicity we consider univariate functions defined on $[0,1]$.

Consider the three periodic functions plotted in Figure 3:

$$f_1(x) := 3 - |2x-1|\,, \quad f_2(x) = 1 + [4x(1-x)]^2, \quad f_3(x) = [4x(1-x)]^4.$$

To integrate these functions one might use the rectangle rule, $Q_{\mathrm{rect},n}(f) := [f(0) + f(1/n) + \cdots + f(1-1/n)]/n$, which is the one-dimensional special case of a lattice rule. Let $\mathcal{F}_{\mathrm{per}}^{(r)}$ denote the space of periodic functions with $r-1$ absolutely continuous, periodic derivatives and absolutely integrable rth derivatives. Define the accompanying norm as $\|f\|_{\mathcal{F}_{\mathrm{per}}^{(r)}} := |f(0)| + \int_0^1 \left|f^{(r)}(x)\right| dx$. The worst-case error of this rectangle rule is $e^{\mathrm{wo}}(\|\cdot\|_{\mathcal{F}_{\mathrm{per}}^{(r)}}, \mathcal{Q}_{\mathrm{rect},n}) \asymp n^{-r}$. Because $f_r \in \mathcal{F}_{\mathrm{per}}^{(2r-1)} \setminus \mathcal{F}_{\mathrm{per}}^{(2r)}$ for $r = 1,2,3$, one might expect that the rectangle rule has the fastest convergence rate for f_3 followed by f_2, and then f_1. While it is true that the convergence rate for f_3 is faster than that for f_2, the convergence rate for f_1 is the fastest, as seen in Figure 3 for $n = 0,1,2,4,\ldots$.

The reason for this nonintuitive behaviour is that f_1 has nonzero trigonometric Fourier coefficients, $F_1(\nu)$, for only zero and odd ν, and the rectangle rule with an even number of points integrates these odd wavenumber Fourier terms exactly. Specifically, $N(0) = 0$ and $N(\nu) = 1$ for odd ν, so $g(N(\nu);\alpha,\delta) = 1$ for all ν with $F_1(\nu) \neq 0$. Thus, the norms $\|f_1\|_{\text{ra-lat,tr},k}$ and $\|f_1\|_{\mathrm{lat,tr},k}$ for $k = 1,2$ are uniformly bounded in α. This yields a convergence rate $g(n;\alpha,\delta)$ for arbitrarily large α. On the other hand, f_2 and f_3 have nonzero Fourier coefficients for all ν and so, their convergence rates are algebraic in n^{-1}.

Consider three more functions, that are plotted in Figure 4:

$$f_4(x) := 2 + \frac{4x\ 1_{[0,3/4)}(x)}{3} + 2(1-x)\ 1_{[3/4,1]}(x),$$

$$f_5(x) := 1 + \frac{3x\ 1_{[0,2/3)}(x)}{2} + \frac{3(1-x)\ 1_{[2/3,1]}(x)}{2},$$

$$f_6(x) := \frac{3x\ 1_{[0,2/3)}(x)}{2} + 3(1-x)\ 1_{[2/3,1]}(x).$$

The first two of these are discontinuous and the third one is continuous. Consider now the sample space of stratified sampling numerical integration rules, $\mathcal{Q}_{\mathrm{strat},n}(f)$, whose elements are $Q_{\mathrm{strat},n}(f) := [f(0+\Delta_0) + f(1/n+\Delta_1) + \cdots + f(1-1/n+\Delta_{n-1})]/n$, where the Δ_i are independent and identically distributed (i.i.d.) uniform random variables on $[0,1/n)$. This is a special case of a randomly scrambled digital

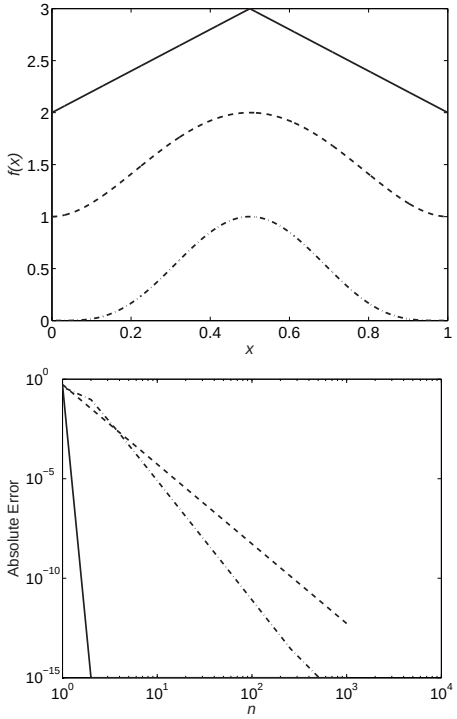

FIGURE 3. Plots of the functions f_1 (solid), f_2 (dashed), and f_3 (dot-dashed) and the absolute integration errors using a rectangle rule.

net rule. Let $\mathcal{F}^{(r)}$ denote the space of integrands with $r-1$ absolutely continuous derivatives and square integrable rth derivative with $\|f\|^2_{\mathcal{F}^{(r)}} := |f(0)|^2 + \cdots + \left|f^{(r-1)}(0)\right|^2 + \int_0^1 \left|f^{(r)}(x)\right|^2 \, dx$. Again the convergence rate of the integration error improves as the smoothness increases, $e^{\mathrm{ra}}(\|\cdot\|_{\mathcal{F}^{(r)}}, \mathcal{Q}_{\mathrm{strat},n}) \asymp n^{-r-1/2}$ for $r = 0, 1$. Because $f_6 \in \mathcal{F}^{(1)}$, but $f_5 \notin \mathcal{F}^{(1)}$, it is no suprise that f_5 is integrated much worse than f_6, as seen in Figure 4. However, $f_4 \notin \mathcal{F}^{(1)}$, but the convergence rates for f_4 and f_6 are the same.

For base 2 net rules the Walsh and Haar bases $\psi_\nu(x)$ may have discontinuities at points where x is a multiple of $2^{-\lg(\nu)}$. Discontinuities of the integrands at such points do not affect the decay rate of the series coefficients and consequently do not influence the performance of the net integration rules. However, discontinuities at other points do affect the decay rate of the Haar or Walsh series coefficients and the performance of the net integration rule. Since the discontinuity of f_4 is at $x = 3/4$ it does not influence the convergence rate of the net integration rule. However, since the discontinuity of f_5 is at $x = 2/3$ it does influence the convergence rate of the net integration rule.

8. Discussion

The algorithm-driven approach proposed here requires finding an appropriate basis to match the numerical integration rule. There must be a simple formula for $\mathrm{err}(\psi_{\boldsymbol{\nu}}, Q_n)$ for deterministic rules and $E\{[\mathrm{err}(\psi_{\boldsymbol{\nu}}, \mathcal{Q}_n, \mu)]^2\}$ for random rules, and condition (11) must be satisfied. Several examples of matching rules and bases have

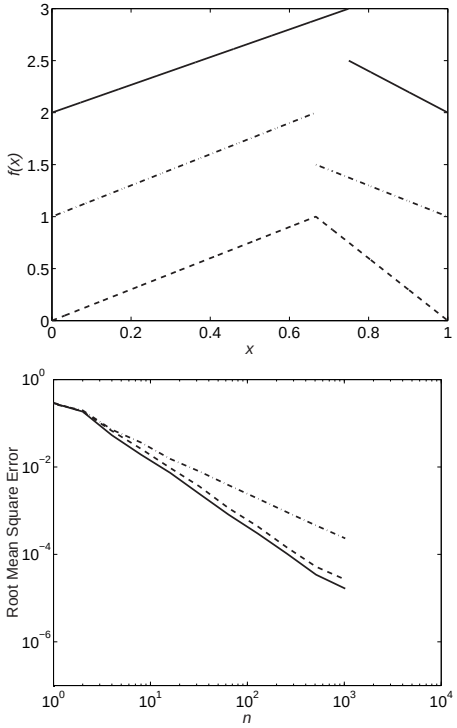

FIGURE 4. Plots of the functions f_4 (solid), f_5 (dashed), and f_6 (dot-dashed) and the root mean square integration errors using a stratified sampling rule.

been given here. However, it is not yet known what basis is appropriate for, say, random-start Halton sequences [21].

In spite of the effort involved, the algorithm-driven approach does have a number of benefits. It gives a more complete picture than the usual norm-based approach of under what conditions an algorithm performs well and under what conditions it does not. Integrands for which the numerical integration rule works well that are missed by the norm-driven approach are often picked up by the algorithm-driven approach.

One may safely assume that the weights $\omega_{\mathrm{ra}}(\boldsymbol{\nu}, n, s)$ in Theorem 1 for the randomized case are no larger than the corresponding weights $\omega_{\mathrm{wo}}(\boldsymbol{\nu}, n, s)$ for the worst case. For lattice rules applied to trigonometric Fourier series they are in fact the same. Even in this situation the norm for the randomized case is weaker because it is an $\mathcal{L}_2$-norm of the series coefficients rather than an $\mathcal{L}_1$-norm for the worst case. However, for digital net rules applied to Walsh function and Haar wavelet series the randomized case norm is even weaker because the weights $\omega_{\mathrm{ra}}(\boldsymbol{\nu}, n, s)$ are asymptotically smaller than $\omega_{\mathrm{wo}}(\boldsymbol{\nu}, n, s)$ as $n \to \infty$.

There are several problems that deserve further study. To understand what is meant by the norms defined in the algorithm-driven approach one may wish to relate the series coefficients to other properties of the integrand, such as smoothness. The analysis in this article describes the largest spaces of integrands for which a numerical integration rule yields a specified convergence rate, however it does not address the question of whether the algorithm is optimal for these spaces. For

the case of Haar wavelets integrated by net rules the approach of [6] might help. Owen and his collaborators have introduced the concept of effective dimension [1, 19]. It would be interesting to understand what the typical effective dimension for integrands in the spaces defined by the algorithm-driven approach is.

Acknowledgements

The authors would like to thank Stefan Heinrich, Henryk Woźniakowski and the referees for their comments and suggestions.

References

[1] R. E. Caflisch, W. Morokoff and A. Owen, Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension, J. Comput. Finance 1 (1997) 27–46.
[2] R. Cranley and T. N. L. Patterson, Randomization of number theoretic methods for multiple integration, SIAM J. Numer. Anal. 13 (1976) 904–914.
[3] P. J. Davis and P. Rabinowitz, Methods of Numerical Integration, Academic Press, Orlando, Florida, 1984.
[4] M. Evans and T. Swartz, Approximating Integrals via Monte Carlo and Deterministic Methods, no. 20 in Oxford Statistical Science Series, Oxford University Press, Oxford, 2000.
[5] H. Faure, Discrépance de suites associées à un système de numération (en dimension s), Acta Arith. 41 (1982) 337–351.
[6] S. Heinrich, F. J. Hickernell and R. X. Yue, Optimal quadrature for Haar wavelet spaces, Math. Comp. 73 (2004) 259–277.
[7] F. J. Hickernell, Lattice rules: How well do they measure up?, in : P. Hellekalek, G. Larcher (Eds.), Random and Quasi-Random Point Sets, Vol. 138 of Lecture Notes in Statistics, Springer-Verlag, New York, 1998, pp. 109–166.
[8] F. J. Hickernell, H. S. Hong, P. L'Écuyer and C. Lemieux, Extensible lattice sequences for quasi-Monte Carlo quadrature, SIAM J. Sci. Comput. 22 (2000) 1117–1138.
[9] F. J. Hickernell and H. Niederreiter, The existence of good extensible rank-1 lattices, J. Complexity 19 (2003) 286–300.
[10] H. S. Hong and F. J. Hickernell, Implementing scrambled digital nets, ACM Trans. Math. Softw. 29 (2003) 95–109.
[11] L. K. Hua and Y. Wang, Applications of Number Theory to Numerical Analysis, Springer-Verlag and Science Press, Berlin and Beijing, 1981.
[12] G. Larcher, Digital point sets: Analysis and applications, in: P. Hellekalek, G. Larcher (Eds.), Random and Quasi-Random Point Sets, Vol. 138 of Lecture Notes in Statistics, Springer-Verlag, New York, 1998, pp. 167–222.
[13] J. Matoušek, On the L_2-discrepancy for anchored boxes, J. Complexity 14 (1998) 527–556.
[14] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.
[15] H. Niederreiter and C. Xing, Rational Points on Curves over Finite Fields: Theory and Applications, no. 285 in London Mathematical Society Lecture Notes Series, Cambridge University Press, Cambridge, 2001.
[16] E. Novak and H. Woźniakowski, When are integration and discrepancy tractable?, in: Foundations of Computational Mathematics, 2001, pp. 211–266.
[17] A. B. Owen, Randomly permuted (t, m, s)-nets and (t, s)-sequences, in: H. Niederreiter, P. J.-S. Shiue (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Vol. 106 of Lecture Notes in Statistics, Springer-Verlag, New York, 1995, pp. 299–317.
[18] A. B. Owen, Monte Carlo variance of scrambled net quadrature, SIAM J. Numer. Anal. 34 (1997) 1884–1910.
[19] A. B. Owen, The dimension distribution and quadrature test functions, Statist. Sinica 13 (2003) 1–18.
[20] I. H. Sloan and S. Joe, Lattice Methods for Multiple Integration, Oxford University Press, Oxford, 1994.
[21] X. Wang and F. J. Hickernell, Randomized Halton sequences, Math. Comput. Modelling 32 (2000) 887–899.

Department of Applied Mathematics, Illinois Institute of Technology, Room 208, Bldg. E1, 10 W. 32nd Street, Chicago, IL 60521
E-mail: hickernell@iit.edu
URL: http://www.iit.edu/~hickerne

School of Mathematics and Statistics, UNSW, Sydney 2052, Australia
E-mail: Josef.Dick@unsw.edu.au
URL: http://www2.maths.unsw.edu.au/Contacts/profile.php?logname=josi