SIAM J. NUMER. ANAL.
Vol. 62, No. 3, pp. 1021–1038

GAIN COEFFICIENTS FOR SCRAMBLED HALTON POINTS*

ART B. OWEN† AND ZEXIN PAN†

Abstract. Randomized quasi-Monte Carlo, via certain scramblings of digital nets, produces unbiased estimates of $\int_{[0,1]^d} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ with a variance that is $o(1/n)$ for any $f \in L^2[0,1]^d$. It also satisfies some nonasymptotic bounds where the variance is no larger than some $\Gamma < \infty$ times the ordinary Monte Carlo variance. For scrambled Sobol' points, this quantity Γ grows exponentially in d. For scrambled Faure points, $\Gamma \leqslant \exp(1) \doteq 2.718$ in any dimension, but those points are awkward to use for large d. This paper shows that certain scramblings of Halton sequences have gains below an explicit bound that is $O(\log d)$ but not $O((\log d)^{1-\epsilon})$ for any $\epsilon > 0$ as $d \to \infty$. For $6 \leqslant d \leqslant 10^6$, the upper bound on the gain coefficient is never larger than $3/2 + \log(d/2)$.

Key words. cubature, quadrature, randomized quasi-Monte Carlo, scrambled nets

MSC codes. 65D32, 91G60

DOI. 10.1137/23M1601882

1. Introduction.

High dimensional integrals are often computed by plain Monte Carlo (MC) sampling. In its basic form, we sample random vectors independent and identically distributed (i.i.d.) from their distribution, evaluate some quantity of interest on the sampled vectors, and average the resulting values. It is often possible to use a rich set of transformations from $\mathbb{U}[0,1]^d$ (see [4]) to generate the needed random vectors. We can then write the integral of interest as $\mu = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and approximate it via $\hat{\mu} = (1/n)\sum_{i=0}^{n-1} f(\boldsymbol{x}_i)$ for $\boldsymbol{x}_i \overset{\text{i.i.d.}}{\sim} \mathbb{U}[0,1]^d$.

In quasi-MC (QMC) sampling [5, 6, 16], deterministic points $\boldsymbol{x}_i \in [0,1]^d$ are chosen strategically to nearly minimize a measure of distance between the discrete uniform distribution on $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1}\}$ and the continuous uniform distribution on $[0,1]^d$. Such distances are known as discrepancies [2]. The most widely studied one is the star discrepancy $D_n^*(\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1})$ which is a multivariate generalization of the Kolmogorov–Smirnov distance between discrete and continuous uniform distributions. It is possible to attain $D_n^* = O(\log(n)^{d-1}/n)$. Then the Koksma–Hlawka inequality [12] yields $|\hat{\mu} - \mu| = O(n^{-1+\epsilon})$ for any $\epsilon > 0$, when f has bounded variation in the sense of Hardy and Krause, which we write as $f \in \mathrm{BVHK}$.

While $\log(n)^{d-1} = O(n^{\epsilon})$ for any $\epsilon > 0$ it is natural to question whether that is a good description for large d and modest n. Surprisingly, this expression seems reasonable for applied work. Those logarithmic powers apply for adversarially chosen integrands f that never seem to arise in practice [26] and it is challenging to construct even one such integrand requiring a power of $\log(n)$ above 1 [22], even when exploiting known weaknesses of some QMC constructions.

Some (but not all) randomized QMC (RQMC) methods provide stronger assurances that high powers of $\log(n)$ do not correspond to very bad accuracy. In RQMC, one takes QMC points $\boldsymbol{a}_0, \dots, \boldsymbol{a}_{n-1}$ and a random transformation τ such that $\boldsymbol{x}_i = \tau(\boldsymbol{a}_i) \sim \mathbb{U}[0,1]^d$ individually, while $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}$ collectively have low discrepancy. See [13] and [21, Chapter 17]. This allows us to get i.i.d. replicates $\hat{\mu}_r$ for

*Received by the editors September 15, 2023; accepted for publication (in revised form) February 13, 2024; published electronically May 2, 2024.
https://doi.org/10.1137/23M1601882
Funding: Supported by the National Science Foundation under grant DMS-2152780.
†Stanford University, Stanford, CA 94305 USA (owen@stanford.edu, zep002@stanford.edu).

$r = 1, \dots, R$ that are unbiased for μ and we can use them to estimate the RQMC sampling variance.

Some RQMC methods give unbiased estimates of μ with variance no larger than $\Gamma\sigma^2/n$ for some $\Gamma < \infty$, where σ^2/n is the variance of $\hat{\mu}$ under i.i.d. sampling. This bounds how much the powers of $\log(n)$ can make RQMC worse than plain MC which is the natural default comparison for RQMC. Also, if $f \in \mathrm{BVHK}$ then $\mathrm{var}(\hat{\mu}) = O(n^{-2+\epsilon})$ for any $\epsilon > 0$.

We take as our starting point, the nested uniform scrambling of digital nets from [18]. That method provides an estimate $\hat{\mu}$ with many desirable properties noted in [23]. It is unbiased: if $f \in L^1[0,1]^d$ then $\mathbb{E}(\hat{\mu}) = \mu$. There is a strong law of large numbers: if $f \in L^{1+\epsilon}[0,1]^d$ for some $\epsilon > 0$ then $\Pr(\lim_{n\to\infty} \hat{\mu} = \mu) = 1$. If $f \in L^2[0,1]^d$ then $\mathrm{var}(\hat{\mu}) = o(1/n)$. If f is sufficiently smooth, so that it has mixed partial derivatives with respect to each input at most once that are in $L^2[0,1]^d$, then $\mathrm{var}(\hat{\mu}) = O(n^{-3}(\log n)^{d-1})$. The property of most interest here is that if $f \in L^2[0,1]^d$ then there exists $\Gamma < \infty$ such that $\mathrm{var}(\hat{\mu}) \leqslant \Gamma\sigma^2/n$. This quantity Γ is called a "gain coefficient."

The most popular QMC points are the digital nets and sequences of Sobol' [27]. They are constructed using dyadic (base 2) representations and are designed for sample sizes $n = 2^m$. The properties described above for RQMC can be attained using either the nested uniform scrambling of [18] or the faster linear scrambling plus digital shift of [14]. Writing the original Sobol' points $\boldsymbol{a}_i = (a_{i1}, \dots, a_{id}) \in [0,1]^d$, and then writing each a_{ij} in terms of bits, the RQMC points $\boldsymbol{x}_i$ are obtained by taking their bits to be certain randomizations of the bits of a_{ij}.

The scrambled Sobol' points have a disadvantage in that the known upper bounds for Γ grow exponentially with dimension d. The bound in [24] is $\Gamma \leqslant 2^{d+t-1}$, where $t \geqslant 0$ is the quality parameter of the Sobol' points. So the upper bound on Γ is exponential in d. This has been generalized by [10, Theorem 3.5] to $\Gamma \leqslant b^{t+d-1}/(b-1)^d$ for randomizations of digital nets in base b (for prime powers b). We know that the upper bound is not strict for all digital nets. For instance, in [24] we were able to exhibit a digital net with $\Gamma = 2^{d+t-2}$. However, Corollary 5 of that paper gives a lower bound. It shows that for scrambled nets in base 2 that either $\Gamma = n$ or $\Gamma = 2^{t^*_{1:d}+d-1}$. In the second case $\Gamma \geqslant 2^{d-1}$ because the quantity $t^*_{1:d}$ is nonnegative. The case $\Gamma = n$ arises only for digital nets with a flaw in their "generator matrices." So we consider gain coefficients Γ that are exponentially large in d to be the norm. Theorem 4.5 of [10] generalized this exponential lower bound to digital nets in general bases $b \geqslant 2$.

A smaller value of Γ can be found by scrambling the digital nets of Faure [7]. While Sobol's points are constructed in base 2, Faure's points are constructed in a more general integer base $b \geqslant 2$. Scrambling the points of Faure, provides a bound of $\Gamma \leqslant [b/(b-1)]^{d-1}$ in dimension d [19]. Because his construction requires $b \geqslant d$ it follows that the maximal gain cannot exceed $\exp(1) \doteq 2.718$ in any dimension. Faure's construction requires b to be a prime number; however, it generalizes to the case where b is a power of a prime [15].

Unfortunately, the point sets of Faure do not seem to do as well in practice as those of Sobol'. This can be explained by the fact that to get nontrivial equidistribution in s-dimensional marginal projections of $[0,1]^d$ they require at least b^s points to be used. Because $b \geqslant d$, we then need to use $n \geqslant d^s$ points to gain an appreciable advantage over plain MC in averaging the s-dimensional interactions in an ANOVA decomposition of f. QMC and RQMC points typically have very uniform 1 dimensional marginal projections $\{x_{0j}, \dots, x_{n-1,j}\}$ and so the difficulties with Faure points arise when d^2 or d^3 would be an uncomfortably large value for n.

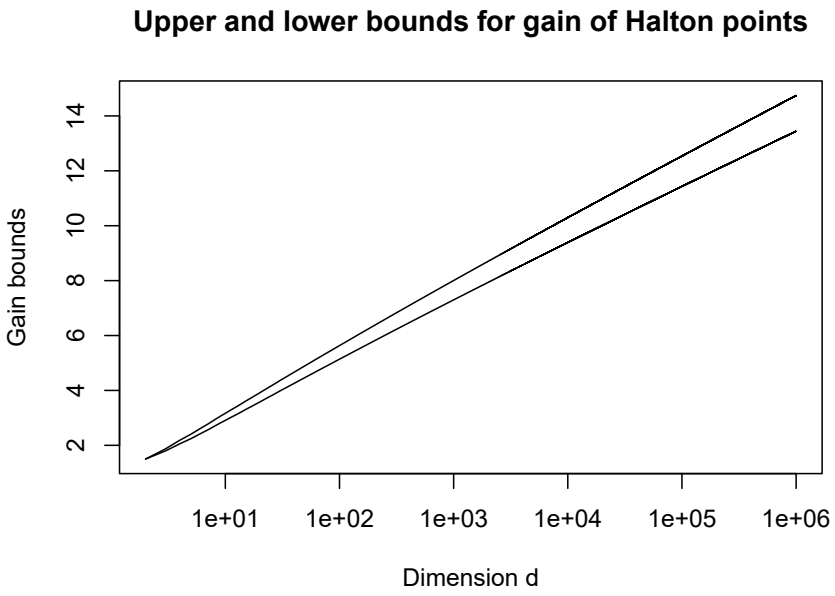

FIG. 1. *This figure shows the upper and lower bounds for* Γ_d *from* (1.1). *The horizontal axis is the dimension* d *for* $2 \leqslant d \leqslant 10^6$.

There is thus a gap. How can we get RQMC constructions that converge faster than those of Faure while having better upper bounds on Γ than those of Sobol'? This article proposes scrambling of Halton points [11] as a solution. Halton points are less commonly used than Sobol' points now, probably due to experience or beliefs that Sobol' points provide greater accuracy. Here, we show that Halton points have gain parameters that grow at most slowly with dimension. Letting Γ_d be the largest gain coefficient in d dimensions, our main theoretical results are upper and lower bounds for Γ_d. We easily find that $\Gamma_1 = 1$ and our bounds imply that

$$\frac{3}{4}\prod_{j=1}^{d}\frac{b_j+1}{b_j} \leqslant \Gamma_d \leqslant \frac{1}{2}\prod_{j=1}^{d}\frac{b_j}{b_j-1} \tag{1.1}$$

both hold for all $d \geqslant 2$, where as described below, b_j is the jth prime number. Using (1.1) we show that $\Gamma_d = O(\log d)$. We also show that Γ_d cannot be $O((\log d)^{1-\epsilon})$ for any $\epsilon > 0$. The bounds in (1.1) are shown in Figure 1. For $6 \leqslant d \leqslant 10^6$, the upper bound on Γ_d never exceeds $3/2 + \log(d/2)$, though that may fail to hold for some $d > 10^6$.

This logarithmic rate for Γ_d is much slower than the exponential rate for scrambled Sobol' points. We might then prefer to use scrambled Halton points in settings where we very much want to avoid the worst outcomes even if it means less accuracy on benign cases. Halton points are also easier to use than Faure points when d is large. If we rank the RQMC methods by worst case variances we prefer Faure to Halton to Sobol'. In high dimensional settings with nonpathological integrands we might reasonably prefer the reverse order. Then Halton, coming second both times, may be a good compromise choice.

The rest of this paper is organized as follows. Section 2 introduces some notation, defines the Halton points and introduces gain coefficients for all nonempty subsets

of $s \leqslant d$ variables and all vectors of s nonnegative integers. Section 3 gives some expressions for gain coefficients at special sample sizes n. It also shows that the gain coefficients are $O(1/n)$ from which the scrambled Halton variance is $o(1/n)$ for any integrand in $L^2[0,1]^d$. Section 4 has numerical examples to illustrate how gain coefficients vary with n. Section 5 has two theorems that identify precisely where the worst gain coefficients must lie and then establishes the upper bound in (1.1). Section 6 establishes the lower bound in (1.1). Section 7 has brief conclusions.

2. Background material.

2.1. Basic notation. We use $\mathbb{R}$ for the real numbers, $\mathbb{Z}$ for the integers, $\mathbb{N}$ for the positive integers, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $\mathbb{Z}_m = \{0, 1, \dots, m-1\}$ for $m \in \mathbb{N}$. We use $1{:}d$ to denote $\{1, 2, \dots, d\}$.

For $u \subseteq 1{:}d$, we use $|u|$ for the cardinality of u and $-u$ for the complementary set $1{:}d \setminus u$. A vector of zeros is denoted by $\mathbf{0}$. If $u = \{j_1, j_2, \dots, j_{|u|}\}$ then we use $\mathbb{N}_0^u$ to denote a copy of $\mathbb{N}_0^{|u|}$ that can be indexed by the elements of u. For example, from any $\boldsymbol{k} \in \mathbb{N}_0^{\{1,2,4\}}$ we can obtain components k_1, k_2, and k_4.

For $z \in \mathbb{R}$, we let $\lfloor z \rfloor = \max\{y \in \mathbb{Z} \mid y \leqslant z\}$. For $a \in \mathbb{N}_0$ and $b \in \mathbb{N}$ the residue of a modulo b is $a - \lfloor a/b \rfloor b$ which we denote by $(a \bmod b)$.

The expressions $\mathbf{1}_A$ and $\mathbf{1}\{A\}$ are both indicators, taking the value 1 when A holds and 0 when A does not hold. The choice of which to use is made based on readability.

2.2. Halton points. Let b_j be the jth smallest prime number for $j \in \mathbb{N}$. The base b_j digits of $i \in \mathbb{N}_0$ are denoted $a_{ij\ell}$. That is, for $i \in \mathbb{N}_0$ and $j \in \mathbb{N}$, we can write

$$i = \sum_{\ell=1}^{\infty} a_{ij\ell} b_j^{\ell-1}$$

for $a_{ij\ell} \in \mathbb{Z}_{b_j}$. This sum has only finitely many nonzero terms for any $i \in \mathbb{N}_0$. The unscrambled Halton points are $\boldsymbol{a}_i \in [0,1)^d$ for $i \in \mathbb{N}$ with

$$a_{ij} = \sum_{\ell=1}^{\infty} a_{ij\ell} b_j^{-\ell} \tag{2.1}$$

for $j \in 1{:}d$. Halton points can be defined by taking b_j to be any d relatively prime natural numbers. In practice, the first d primes are almost always used and we will work with that assumption.

Here is a brief intuitive description of why Halton points fill the unit cube nearly uniformly. For more details see [11]. For $j = 1$, as integers i alternate between even and odd, the first digit a_{i11} alternates between 0 and 1 and then the point a_{i1} alternates between being in $[0, 1/2)$ and $[1/2, 1)$ so we always have nearly half of the points in $[0, 1/2)$ and half in $[1/2, 1)$. More generally, any consecutive 2^k integers i contain all values of $\mathbb{Z}_{2^k}$ and then the corresponding a_{i1} will be balanced over $[r/2^k, (r+1)/2^k)$ for all $r \in \mathbb{Z}_{2^r}$. Still more generally, for $j \geqslant 1$ and any $b_j^{k_j}$ consecutive indices $i \in \mathbb{N}_0$, the values a_{ij} stratify over $[r/b_j^{k_j}, (r+1)/b_j^{k_j})$ for $r \in \mathbb{Z}_{b_j^{k_j}}$. For $\boldsymbol{k} \in \mathbb{N}_0^d$ we can consider the Halton strata

$$S_{\boldsymbol{r}}(\boldsymbol{k}) = \prod_{j=1}^{d} \left[\frac{r_j}{b_j^{k_j}}, \frac{r_j + 1}{b_j^{k_j}} \right) \tag{2.2}$$

with $r_j \in \mathbb{Z}_{b_j^{k_j}}$. By the Chinese remainder theorem, every consecutive batch of $\prod_{j=1}^d b_j^{k_j}$ points has exactly one member in each of the strata above. Any subsequent batch of fewer than $\prod_{j=1}^d b_j^{k_j}$ points is spread through those strata, with at most one of them in each stratum. Smaller bases b_j tend to provide better equidistribution properties than larger bases do. As a result, when using Halton points, it can be very valuable to arrange for the most important input variables to have the lowest indices. A perfect definition of variable importance would be tautological and not very helpful. In practice, one can use scientific understanding/intuition or proxy measures such as Sobol' indices [3] to order the inputs.

While Halton points are asymptotically equidistributed, it is well known that for small n and large d, the points tend to show unwanted structure. For $i < 100$, $a_{i,26} = (i \bmod 101)/101$ and $a_{i,27} = (i \bmod 103)/103$ are collinear. There have been many proposals to break up this unwanted structure by, for example, replacing $a_{ij\ell}$ in (2.1) by some permuted values $\pi(a_{ij\ell})$, where $\pi(\cdot)$ can depend on j and ℓ. There are deterministic proposals in [1], [8], and [28] and others described in [9]. There is a random permutation proposal in [17] with a study and implementation in [20] and another kind of randomization in [29].

Here we consider two randomizations. One is the nested uniform scramble [18] in base b_j applied to the jth component of $\boldsymbol{a}_i$ with all d randomizations statistically independent of each other. The other is the random linear scramble, with digital shift, from [14]. Faure and Lemieux [9] have considered the linear scramble, without a digital shift, for Halton points. They did not use random scrambles but instead did a computer search to find a scramble to recommend for general use.

2.3. Gain coefficients. Digital nets are similar to Halton points, except that they use the same base b for every component of the n points. Gain coefficients for scrambled digital nets were presented in [19]. They arise from a d-fold tensor product of a base b Haar wavelet basis for $L^2[0,1]$. For Halton points, we use instead a tensor product of Haar wavelet basis functions with the jth one defined in terms of base b_j. For nonempty $u \subseteq 1{:}d$, $\boldsymbol{k} \in \mathbb{N}_0^u$, and integer $n \geqslant 1$, define the gain coefficient

$$
\begin{aligned}
G_{u,\boldsymbol{k}}(n) &= \frac{1}{n}\prod_{j\in u}(b_j-1)^{-1}\widetilde{G}_{u,\boldsymbol{k}}(n), \quad \text{where}\\
\widetilde{G}_{u,\boldsymbol{k}}(n) &= \sum_{i=0}^{n-1}\sum_{i'=0}^{n-1}\prod_{j\in u}\Bigl(b_j\mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor=\lfloor b_j^{k_j+1}a_{i'j}\rfloor} - \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor=\lfloor b_j^{k_j}a_{i'j}\rfloor}\Bigr).
\end{aligned}
\tag{2.3}
$$

This formula is a generalization of the one in [19, Theorem 2] that uses the same base b in every dimension. These gain coefficients apply to scrambling of arbitrary point sets, though they have useful simplifications for some QMC points.

Each $f \in L^2[0,1]^d$ has variance components $\sigma^2_{u,\boldsymbol{k}}$ defined through the wavelet basis. The variance σ^2 of f satisfies

$$\sigma^2 = \sum_{u\subseteq 1:d}\sum_{\boldsymbol{k}\in\mathbb{N}_0^u}\sigma^2_{u,\boldsymbol{k}}.$$

We take $\sigma^2_{\varnothing,()} = 0$ because it corresponds to a constant term which does not contribute to the sampling variance. If we use $n \geqslant 1$ randomized Halton points then the estimate

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=0}^{n-1} f(\boldsymbol{x}_i)$$

is an unbiased estimate of $\mu = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ with variance

$$\frac{1}{n}\sum_{\varnothing\neq u\subseteq 1:d}\sum_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n)\sigma^2_{u,\boldsymbol{k}} \leqslant \frac{\Gamma_d(n)\sigma^2}{n},$$

where

$$\Gamma_d(n) = \max_{\varnothing\neq u\subseteq 1:d}\sup_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n).$$

Estimation using scrambled Halton points cannot have more than $\Gamma_d(n)$ times the variance from using plain MC points. It is then interesting to bound $\Gamma_d(n)$. We will also get a bound for

$$\Gamma_d = \sup_{n\in\mathbb{N}} \Gamma_d(n).$$

2.4. Preliminary results. Here we present some elementary results to simplify some of the derivations for gain coefficients. We begin by defining two quantities that frequently arise in our expressions. For nonempty $u \subseteq 1{:}d$ and any $v \subseteq u$, let

$$H_{u,v} = \prod_{j\in v} b_j \prod_{j\in u-v}(-1). \tag{2.4}$$

Then for $\boldsymbol{k}\in\mathbb{N}_0^u$ define

$$m_{u,v,\boldsymbol{k}} = \prod_{j\in v} b_j^{k_j+1}\prod_{j\in u-v} b_j^{k_j}. \tag{2.5}$$

By inclusion-exclusion, we may write the product in the gain formula (2.3) as

$$\begin{aligned}&\prod_{j\in u}\Bigl(b_j \mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor} - \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor}\Bigr)\\ &\quad= \sum_{v\subseteq u} H_{u,v}\prod_{j\in v}\mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor}\prod_{j\in u-v}\mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor}.\end{aligned}$$

Next, we develop an expression for the coefficient of $H_{u,v}$ in the expression above.

For a_{ij} given by (2.1) and $r\geqslant 0$,

$$\lfloor b_j^r a_{ij}\rfloor = \left\lfloor \sum_{\ell=1}^{\infty} b_j^{r-\ell}a_{ij\ell}\right\rfloor = \sum_{\ell=1}^{r} b_j^{r-\ell}a_{ij\ell}.$$

Therefore $\lfloor b_j^r a_{ij}\rfloor = \lfloor b_j^r a_{i'j}\rfloor$ if and only if

$$\sum_{\ell=1}^{r} b_j^{r-\ell}a_{ij\ell} = \sum_{\ell=1}^{r} b_j^{r-\ell}a_{i'j\ell}$$

which holds if and only $i = i' \bmod b_j^r$. Then using the Chinese remainder theorem

$$\begin{aligned}&\prod_{j\in v}\mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor}\prod_{j\in u-v}\mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor}\\ &\quad= \prod_{j\in v}\mathbf{1}\{i = i' \bmod b_j^{k_j+1}\}\prod_{j\in u-v}\mathbf{1}\{i = i' \bmod b_j^{k_j}\}\\ &\quad= \mathbf{1}\{i = i' \bmod m_{u,v,\boldsymbol{k}}\}.\end{aligned} \tag{2.6}$$

For $m, n \in \mathbb{N}$ let

$$C_{m,n} = \sum_{i=0}^{n-1} \sum_{i'=0}^{n-1} \mathbf{1}\{i = i' \bmod m\}. \tag{2.7}$$

Then the unnormalized gain coefficients from (2.3) satisfy

$$\widetilde{G}_{u,\boldsymbol{k}}(n) = \sum_{v \subseteq u} H_{u,v} C_{m_{u,v,\boldsymbol{k}},n} = \sum_{v \subseteq u} H_{u,v} C_{u,v,\boldsymbol{k}}(n), \tag{2.8}$$

where $C_{u,v,\boldsymbol{k}}(n)$ is a more readable replacement for $C_{m_{u,v,\boldsymbol{k}},n}$.

PROPOSITION 2.1. *For $m, n \in \mathbb{N}$,*

$$C_{m,n} = n + (2n - m)\lfloor n/m \rfloor - m\lfloor n/m \rfloor^2. \tag{2.9}$$

Proof. Write $n = mq + r$ for quotient $q = \lfloor n/m \rfloor \in \mathbb{N}_0$ and remainder $r \in \mathbb{Z}_m$. Then as explained below,

$$\begin{aligned} C_{m,n} &= mq^2 + (2q+1)r \\ &= mq^2 + (2q+1)(n - mq) \\ &= m\lfloor n/m \rfloor^2 + (2\lfloor n/m \rfloor + 1)(n - m\lfloor n/m \rfloor) \\ &= n + (2n - m)\lfloor n/m \rfloor - m\lfloor n/m \rfloor^2. \end{aligned}$$

The mq^2 term comes from $\sum_{i=0}^{mq-1} \sum_{i'=0}^{mq-1} \mathbf{1}\{i = i' \bmod m\}$. We get a count of qr from $\sum_{i=0}^{mq-1} \sum_{i'=mq}^{mq+r-1} \mathbf{1}\{i = i' \bmod m\}$ and another qr with the indices i and i' reversed. Finally, $\sum_{i=mq}^{mq+r-1} \sum_{i'=mq}^{mq+r-1} \mathbf{1}\{i = i' \bmod m\} = r$. ∎

We may write the fractional part $\lfloor n/m \rfloor$ arising in $C_{m,n}$ by $n/m - \varepsilon$ for some $0 \leqslant \varepsilon \leqslant 1$ for each $m = m_{u,v,\boldsymbol{k}}$. Doing this we get

$$C_{u,v,\boldsymbol{k}}(n) = \frac{n^2}{m_{u,v,\boldsymbol{k}}} + m_{u,v,\boldsymbol{k}} \varepsilon_v (1 - \varepsilon_v), \tag{2.10}$$

where $0 \leqslant \varepsilon_v \leqslant 1$, which we will use later.

3. Nonasymptotic results. Here we show some nonasymptotic properties of the gain coefficients. We also show that for scrambled Halton points $\mathrm{var}(\hat{\mu}) = o(1/n)$ when $f \in L^2[0,1]^d$.

Let $\underline{m}_{u,\boldsymbol{k}} = m_{u,\varnothing,\boldsymbol{k}}$ and $\overline{m}_{u,\boldsymbol{k}} = m_{u,u,\boldsymbol{k}}$. These are the minimal and maximal values of $m_{u,v,\boldsymbol{k}}$, respectively. We assume throughout that $u \neq \varnothing$.

PROPOSITION 3.1. *If $1 \leqslant n < \underline{m}_{u,\boldsymbol{k}}$ then*

$$G_{u,\boldsymbol{k}}(n) = 1.$$

Proof. If $n < \underline{m}_{u,\boldsymbol{k}} = m_{u,\varnothing,\boldsymbol{k}}$ then $\lfloor n/m_{u,v,\boldsymbol{k}} \rfloor = 0$ and from (2.9), $C_{u,v,\boldsymbol{k}}(n) = n$. In this case

$$\widetilde{G}_{u,\boldsymbol{k}} = \sum_{v \subseteq u} H_{u,v} C_{u,v,\boldsymbol{k}}(n) = n \sum_{v \subseteq u} \prod_{j \in v} b_j \prod_{j \in u - v} (-1) = n \prod_{j \in u} (b_j - 1).$$

Therefore $G_{u,\boldsymbol{k}} = 1$, because the gain coefficients in (2.3) are defined with a normalizing factor of $\prod_{j \in u} (b_j - 1)^{-1}/n$. ∎

PROPOSITION 3.2. *If $n = q\overline{m}_{u,\boldsymbol{k}}$ for $q \in \mathbb{N}$, then*

$$G_{u,\boldsymbol{k}}(n) = 0.$$

Proof. If $n = q\overline{m}_{u,\boldsymbol{k}}$ for $q \in \mathbb{N}$ then for all $v \subseteq u$,

$$\begin{aligned} C_{u,v,\boldsymbol{k}}(n) &= n + (2n - m_{u,v,\boldsymbol{k}})(n/m_{u,v,\boldsymbol{k}}) - m_{u,v,\boldsymbol{k}}(n/m_{u,v,\boldsymbol{k}})^2 \\ &= n^2/m_{u,v,\boldsymbol{k}}. \end{aligned}$$

Now $\tilde{G}_{u,\boldsymbol{k}}(n)$ equals

$$\begin{aligned} \sum_{v\subseteq u} H_{u,v}\frac{n^2}{m_{u,v,\boldsymbol{k}}} &= n^2 \sum_{v\subseteq u}\left[\prod_{j\in v} b_j \prod_{j\in u-v}(-1)\right]\prod_{j\in u} b_j^{-k_j}\prod_{j\in v} b_j^{-1} \\ &= \frac{n^2}{m_{u,\varnothing,\boldsymbol{k}}}\sum_{v\subseteq u}(-1)^{|u-v|} = 0 \end{aligned} \tag{3.1}$$

and so $G_{u,\boldsymbol{k}}(n) = 0$ by (2.3). ∎

A gain of zero is the expected result. For such n we have attained zero discrepancy for all of the Halton strata congruent to $\prod_{j\in u}[0, 1/b_j^{k_j+1})\prod_{j\in -u}[0,1)$. There are $\overline{m}_{u,\boldsymbol{k}}$ such strata defined by u and $\boldsymbol{k}$, and so $G_{u,\boldsymbol{k}}(n)$ cannot be zero for $n < \overline{m}_{u,\boldsymbol{k}}$. Next we show that $G_{u,\boldsymbol{k}}(n)$ cannot reattain its maximal value for any $n > \overline{m}_{u,\boldsymbol{k}}$.

PROPOSITION 3.3. *Let $n = q\overline{m}_{u,\boldsymbol{k}} + r$ for $q \in \mathbb{N}$ and $r \in \mathbb{Z}_{\overline{m}_{u,\boldsymbol{k}}} \setminus \{0\}$. Then*

$$G_{u,\boldsymbol{k}}(n) = \frac{r}{n} G_{u,\boldsymbol{k}}(r). \tag{3.2}$$

Proof. For any $i' \in \mathbb{N}$ and any $r \in \mathbb{Z}_{\overline{m}_{u,\boldsymbol{k}}}$,

$$\begin{aligned} &\sum_{i=r}^{r+q\overline{m}_{u,\boldsymbol{k}}-1} \prod_{j\in u} b_j \mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor} - \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor} \\ &= \sum_{i=r}^{r+q\overline{m}_{u,\boldsymbol{k}}-1} \sum_{v\subseteq u} H_{u,v} \prod_{j\in v} \mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor} \times \prod_{j\in u-v} \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor} \\ &= \sum_{v\subseteq u} H_{u,v} \sum_{i=r}^{r+q\overline{m}_{u,\boldsymbol{k}}-1} \mathbf{1}_{i=i' \bmod m_{u,v,\boldsymbol{k}}} \\ &= \sum_{v\subseteq u} H_{u,v} \prod_{j\in u-v} b_j \\ &= 0. \end{aligned} \tag{3.3}$$

The last step follows by the argument used in the proof of Proposition 3.2. If $r > 0$ then using (3.3) in sums over both i' and i

$$\begin{aligned} \widetilde{G}_{u,\boldsymbol{k}}(n) &= \sum_{i=0}^{q\overline{m}_{u,\boldsymbol{k}}+r-1} \sum_{i'=0}^{q\overline{m}_{u,\boldsymbol{k}}+r-1} \prod_{j\in u} b_j \mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor} - \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor} \\ &= \sum_{i=0}^{r-1}\sum_{i'=0}^{r-1} \prod_{j\in u} b_j \mathbf{1}_{\lfloor b_j^{k_j+1}a_{ij}\rfloor = \lfloor b_j^{k_j+1}a_{i'j}\rfloor} - \mathbf{1}_{\lfloor b_j^{k_j}a_{ij}\rfloor = \lfloor b_j^{k_j}a_{i'j}\rfloor} \\ &= \widetilde{G}_{u,\boldsymbol{k}}(r). \end{aligned}$$

Now (3.2) follows by the normalization in (2.3). ∎

We left the case $r = 0$ out of Proposition 3.3. We know that $G_{u,\boldsymbol{k}}(n) = 0$ in that case. However we have not chosen a convention for $G_{u,\boldsymbol{k}}(0)$. We think that $G_{u,\boldsymbol{k}}(0) = 1$ is reasonable since $n = 0$ for RQMC is the same as $n = 0$ for MC, but we have not found another need for such a convention.

COROLLARY 3.4. *If $f \in L^2[0,1]^d$ and $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}$ are points of a Halton sequence randomized with a nested uniform scramble, or a random linear scramble with digital shift, then*

$$\lim_{n\to\infty} n \cdot \operatorname{var}\left(\frac{1}{n}\sum_{i=0}^{n-1} f(\boldsymbol{x}_i)\right) = 0.$$

Proof. Let f have variance components $\sigma^2_{u,\boldsymbol{k}}$. Then

$$\begin{aligned} n \cdot \operatorname{var}(\hat{\mu}) &= \sum_{\varnothing\neq u\subseteq 1:d} \sum_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n)\sigma^2_{u,\boldsymbol{k}} \\ &= \sum_{\varnothing\neq u\subseteq 1:d} \sum_{\boldsymbol{k}\in\mathbb{N}_0^u} \frac{n \bmod \overline{m}_{u,\boldsymbol{k}}}{n} G_{u,\boldsymbol{k}}(n \bmod \overline{m}_{u,\boldsymbol{k}})\sigma^2_{u,\boldsymbol{k}} \\ &\leqslant \Gamma_d \sum_{\varnothing\neq u\subseteq 1:d} \sum_{\boldsymbol{k}\in\mathbb{N}_0^u} \frac{n \bmod \overline{m}_{u,\boldsymbol{k}}}{n} \sigma^2_{u,\boldsymbol{k}} \\ &\to 0 \end{aligned}$$

as $n \to \infty$. ∎

The next proposition shows that any values of $G_{u,\boldsymbol{k}}(n)$ reappear as values of $G_{u,\boldsymbol{k}'}(n')$, where $\boldsymbol{k}'$ is any vector in $\mathbb{N}_0^u$ no smaller than $\boldsymbol{k}$ componentwise and n' is some value $n' \geqslant n$.

PROPOSITION 3.5. *For $j \in u \subseteq 1{:}d$ and $\boldsymbol{k} \in \mathbb{N}_0^u$ define $\boldsymbol{k}'$ by $k'_j = k_j + 1$ and $k'_\ell = k_\ell$ for $\ell \in u - \{j\}$. Then*

$$G_{u,\boldsymbol{k}'}(nb_j) = G_{u,\boldsymbol{k}}(n) \tag{3.4}$$

for all $n \in \mathbb{N}$.

Proof. First

$$\widetilde{G}_{u,\boldsymbol{k}'}(b_j n) = \sum_{v\subseteq u} H_{u,v} C_{u,v,\boldsymbol{k}'}(b_j n).$$

Now

$$\begin{aligned} C_{u,v,\boldsymbol{k}'}(b_j n) &= b_j n + (2b_j n - m_{u,v,\boldsymbol{k}'})\lfloor nb_j/m_{u,v,\boldsymbol{k}'}\rfloor - m_{u,v,\boldsymbol{k}'}\lfloor b_j n/m_{u,v,\boldsymbol{k}'}\rfloor^2 \\ &= b_j n + (2b_j n - b_j m_{u,v,\boldsymbol{k}})\lfloor n/m_{u,v,\boldsymbol{k}}\rfloor - b_j m_{u,v,\boldsymbol{k}}\lfloor n/m_{u,v,\boldsymbol{k}}\rfloor^2 \\ &= b_j C_{u,v,\boldsymbol{k}}(n). \end{aligned}$$

It follows that $\widetilde{G}_{u,\boldsymbol{k}'}(b_j n) = b_j \widetilde{G}_{u,\boldsymbol{k}}$. Then (3.4) holds after normalization. ∎

COROLLARY 3.6. *For nonempty $u \subseteq 1{:}d$ and $\boldsymbol{k} \in \mathbb{N}_0^u$,*

$$G_{u,\boldsymbol{k}}\left(n\prod_{j\in u} b_j^{k_j}\right) = G_{u,\boldsymbol{0}}(n)$$

holds for all $n \geqslant 1$.

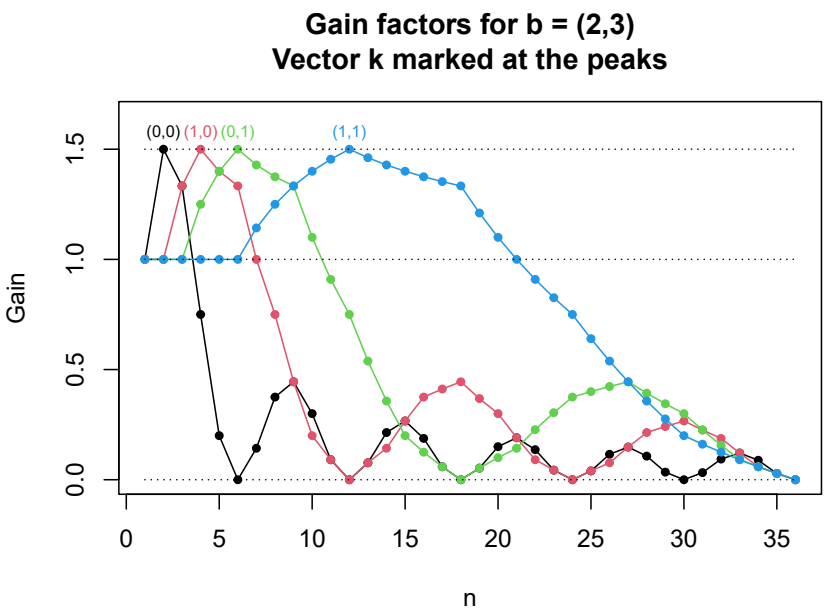

FIG. 2. *For $d=2$ with $\boldsymbol{b}=(2,3)$ this figure shows the gains for $\boldsymbol{k}\in\{(0,0),(0,1),(1,0),(1,1)\}$ versus $1\leqslant n\leqslant 36$. At $n=36$ all four of these gains are zero. The same peak value $3/2$ is attained by all four curves. In all cases, the maximum is attained at $n=2\times b_1^{k_1}\times b_2^{k_2}$. The horizontal reference lines are at gains 0, 1, and $3/2$.*

Proof. We make $\sum_{j\in u}k_j$ applications of Proposition 3.5. ∎

4. Example computations. It is straightforward to compute the gain coefficients for scrambled Halton points in some settings of interest. Figure 2 shows the gain coefficients in the smallest interesting case: $d=2$ and $\boldsymbol{b}=(2,3)$ for $1\leqslant n\leqslant 36$. We see that all $\boldsymbol{k}\in\{(0,0),(0,1),(1,0),(1,0)\}$ attain the same maximal gain factor of $3/2$. All of the curves start at gain equal to one for $n=1$. This makes sense because $n=1$ scrambled Halton point is mathematically equivalent to $n=1$ MC point. The curves are initially one for all $n\leqslant\prod_{j\in u}b_j^{k_j}$ (see Proposition 3.1) and then with some oscillation, they reach zero at $n=\prod_{j\in u}b_j^{k_j+1}$ (see Proposition 3.2). After reaching zero they keep oscillating, but they will never again (for any larger n) reattain their maximum (see Proposition 3.3). The curve for $\boldsymbol{k}$ attains its peak at $n=2\prod_{j\in u}b_j^{k_j}$. The factor $\prod_{j\in u}b_j^{k_j}$ is in line with Proposition 3.5.

Figure 3 shows gain coefficients for $d=3$ with $\boldsymbol{b}=(2,3,5)$. The values of n range from 1 to 1000. Vectors $\boldsymbol{k}$ with $\prod_{j\in u}b_j^{k_j}>1000$ have gain 1 for all n in this range. The plot shows gain curves for all other vectors $\boldsymbol{k}$. It is clear that any value of n has a maximal gain close to the overall maximum (empirically $9/8$). In this worst case sense, the scrambled Halton points do not have especially good values of n. In another sense, described next, there do exist especially good values of n.

If we anticipate that smaller values of $|u|$ and of $\prod_{j\in u}b_j^{k_j}$ correspond to more important features of the function, then values of n that are divisible by products of small powers of the b_j have an advantage. We see in Figure 2 that special values of n give gain equal to zero for some of the effects with small $\boldsymbol{k}$. From Figure 3 we can see that selecting such a special value of n will not give a meaningful penalty with regard to worst case behavior. This leaves us more free to use convenient or highly composite

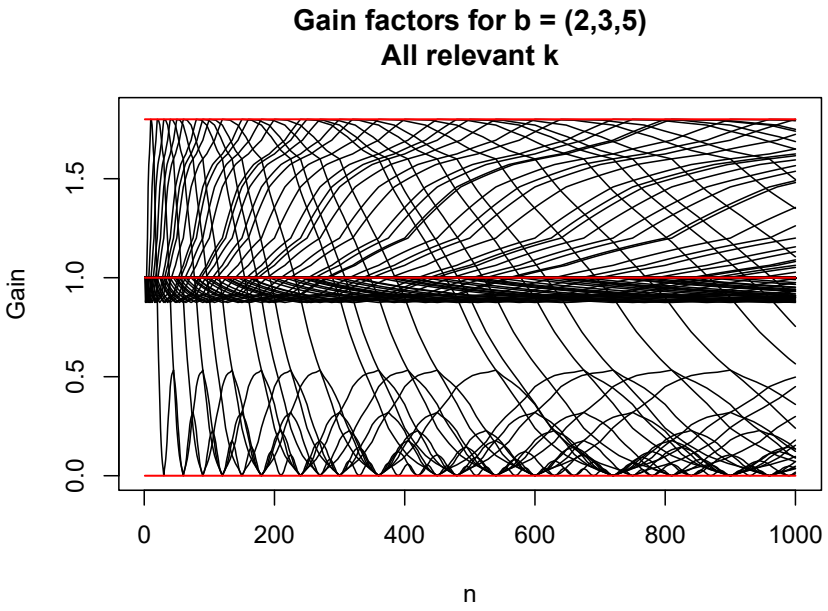

FIG. 3. *For $d = 3$ with $\boldsymbol{b} = (2,3,5)$ this figure shows gain factors $G_{u,\boldsymbol{k}}(n)$ versus n for all nonempty $u \subseteq \{1,2,3\}$ and all $\boldsymbol{k}$ with $\prod_{j\in u} b_j^{k_j} < n$. For any other $\boldsymbol{k}$ we know that $G_{u,\boldsymbol{k}}(n) = 1$ over this range for n. There are horizontal reference lines at gains 0, 1, and 9/5.*

values of n. Values of n that are powers of 10 are often popular with users. For the Halton sequence, such n are very good for the first and third input dimensions. A value like $n = 1800 = 2^3 3^2 5^2$ can be expected to give good results when the integrand depends strongly on the first three components of $\boldsymbol{x}$ in a smooth way. A user who wants n to be a power of 10 might then use bases 2 and 5 for what they think are the most and second most important input variables, respectively.

A striking feature of Figure 3 is a thick band between gains of 1 and 7/8. The latter value is $G_{1:3,\mathbf{0}}(2)$. The gains for every $\boldsymbol{k}$ decrease from 1 to 7/8 before rising to 9/5.

In Figures 2 and 3 we never see any $G_{u,\boldsymbol{k}}(n) > \max_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n)$. Theorem 5.1 in section 5 proves that this can never happen. Theorem 5.2 in section 5 shows that if $v \subsetneq u$ then $\sup_{n\geqslant 1} G_{v,\mathbf{0}}(n) \leqslant \sup_{n\geqslant 1} G_{u,\mathbf{0}}(n)$. Therefore the largest gains for d variables arise in $G_{1:d,\mathbf{0}}(n)$ and we only need to consider n from 1 to $\prod_{j=1}^d b_j$.

5. Upper bounds for gain. It is of interest to know the largest possible values of gain coefficients. Here, Theorem 5.1 shows that we only need to consider $\boldsymbol{k} = \mathbf{0}$. Then Theorem 5.2 shows that we only need to consider $u = 1{:}d$. Applying Proposition 3.3, the largest possible gain for $d \geqslant 1$ is one of $G_{1:d,\mathbf{0}}(n)$ for $1 \leqslant n \leqslant \prod_{j=1}^d b_j$.

THEOREM 5.1. *For all $1 \leqslant d < \infty$ and all nonempty $u \subseteq 1{:}d$ and all $\boldsymbol{k} \in \mathbb{N}_0^u$,*

$$\sup_{n\in\mathbb{N}} G_{u,\boldsymbol{k}}(n) = \sup_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n). \tag{5.1}$$

Proof. Let $b^* = \prod_{j\in u} b_j^{k_j}$. Corollary 3.6 shows that

$$G_{u,\boldsymbol{k}}(nb^*) = G_{u,\mathbf{0}}(n). \tag{5.2}$$

It suffices to show that for n' such that $nb^* \leqslant n' \leqslant (n+1)b^*$, $G_{u,\boldsymbol{k}}(n')$ is maximized at one of the endpoints. That is, we will show that

$$\sup_{nb^* \leqslant n' \leqslant (n+1)b^*} G_{u,\boldsymbol{k}}(n') = \max\bigl(G_{u,\boldsymbol{k}}(nb^*), G_{u,\boldsymbol{k}}((n+1)b^*)\bigr)$$

which is at most $\sup_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n)$ by (5.2). Letting $\varepsilon'_v = n'/m_{u,v,\boldsymbol{k}} - \lfloor n'/m_{u,v,\boldsymbol{k}} \rfloor$ we use (2.8) and (2.10) to write

$$\widetilde{G}_{u,\boldsymbol{k}}(n') = \sum_{v\subseteq u} H_{u,v}\left(\frac{n^2}{m_{u,v,\boldsymbol{k}}} + m_{u,v,\boldsymbol{k}}\varepsilon'_v(1-\varepsilon'_v)\right) = \sum_{v\subseteq u} H_{u,v} m_{u,v,\boldsymbol{k}}\varepsilon'_v(1-\varepsilon'_v). \tag{5.3}$$

The first part of the sum vanishes as it did in (3.1). We write $n' = nb^* + r$ for $0 \leqslant r \leqslant b^*$. Because $m_{u,v,\boldsymbol{k}} = b^* m_{u,v,\mathbf{0}}$,

$$\begin{aligned}
\varepsilon'_v &= \frac{n'}{m_{u,v,\boldsymbol{k}}} - \left\lfloor \frac{n'}{m_{u,v,\boldsymbol{k}}} \right\rfloor \\
&= \frac{n + r/b^*}{m_{u,v,\mathbf{0}}} - \left\lfloor \frac{n + r/b^*}{m_{u,v,\mathbf{0}}} \right\rfloor \\
&= \frac{r}{b^* m_{u,v,\mathbf{0}}} + \frac{n}{m_{u,v,\mathbf{0}}} - \left\lfloor \frac{n}{m_{u,v,\mathbf{0}}} \right\rfloor \\
&= \frac{r}{b^* m_{u,v,\mathbf{0}}} + \varepsilon_v,
\end{aligned}$$

where $\varepsilon_v = n/m_{u,v,\mathbf{0}} - \lfloor n/m_{u,v,\mathbf{0}} \rfloor$. Now the normalized gain coefficients $G_{u,\boldsymbol{k}}(n')$ can be expressed as

$$\begin{aligned}
G_{u,\boldsymbol{k}}(n') &= \frac{1}{n'}\prod_{j\in u}(b_j-1)^{-1}\sum_{v\subseteq u} H_{u,v} m_{u,v,\boldsymbol{k}}\varepsilon'_v(1-\varepsilon'_v) \\
&= \prod_{j\in u}(b_j-1)^{-1}\sum_{v\subseteq u} H_{u,v}\frac{b^* m_{u,v,\mathbf{0}}}{nb^*+r}\varepsilon'_v(1-\varepsilon'_v) \\
&= \prod_{j\in u}(b_j-1)^{-1}\sum_{v\subseteq u} H_{u,v}\frac{m_{u,v,\mathbf{0}}}{n+r/b^*}\Bigl(\varepsilon_v + \frac{r}{b^* m_{u,v,\mathbf{0}}}\Bigr)\Bigl(1-\varepsilon_v - \frac{r}{b^* m_{u,v,\mathbf{0}}}\Bigr) \\
&= \prod_{j\in u}(b_j-1)^{-1}\sum_{v\subseteq u} H_{u,v}\frac{m_{u,v,\mathbf{0}}}{n+x}\left(\varepsilon_v(1-\varepsilon_v) + (1-2\varepsilon_v)\frac{x}{m_{u,v,\mathbf{0}}} - \frac{x^2}{m_{u,v,\mathbf{0}}^2}\right),
\end{aligned}$$

where we have replaced r/b^* with x. Let us extend the domain of x to all real numbers in $[0,1]$. Our goal becomes to prove that $G_{u,\boldsymbol{k}}(n')$, as a function of x, is monotonic on $[0,1]$.

First notice that because $H_{u,v} = \prod_{j\in v} b_j \prod_{j\in u-v}(-1) = (-1)^{|u-v|} m_{u,v,\mathbf{0}}$,

$$\sum_{v\subseteq u} H_{u,v}\frac{m_{u,v,\mathbf{0}}}{n+x}\frac{x^2}{m_{u,v,\mathbf{0}}^2} = \frac{x^2}{n+x}\sum_{v\subseteq u}(-1)^{|u-v|} = 0.$$

This allows us to rewrite $\prod_{j\in u}(b_j-1)G_{u,\boldsymbol{k}}(n')$ as

$$\begin{aligned}
&\sum_{v\subseteq u} H_{u,v}\frac{m_{u,v,\mathbf{0}}}{n+x}\Big(\varepsilon_v(1-\varepsilon_v)+(1-2\varepsilon_v)\frac{x}{m_{u,v,\mathbf{0}}}\Big)\\
&\quad=\sum_{v\subseteq u} H_{u,v}m_{u,v,\mathbf{0}}\varepsilon_v(1-\varepsilon_v)\frac{1}{n+x}+\sum_{v\subseteq u}H_{u,v}(1-2\varepsilon_v)\frac{x}{n+x}\\
&\quad=\frac{1}{n+x}\sum_{v\subseteq u}H_{u,v}\Big(m_{u,v,\mathbf{0}}\varepsilon_v(1-\varepsilon_v)-n(1-2\varepsilon_v)\Big)+\sum_{v\subseteq u}H_{u,v}(1-2\varepsilon_v).
\end{aligned}$$

Monotonicity of $G_{u,\boldsymbol{k}}(n')$ follows from monotonicity of $1/(n+x)$ on $[0,1]$ and hence $G_{u,\boldsymbol{k}}(n')$ is maximized at nb^* or $(n+1)b^*$. ∎

THEOREM 5.2. *For all $1\leqslant d<\infty$ and all nonempty $v\subseteq u\subseteq 1{:}d$,*

$$\sup_{n\in\mathbb{N}} G_{v,\mathbf{0}}(n)\leqslant\sup_{n\in\mathbb{N}}G_{u,\mathbf{0}}(n). \tag{5.4}$$

Proof. It suffices to show the conclusion holds when $u-v$ is a single element j^* and apply induction. Denote the maximizer of $G_{v,\mathbf{0}}(n)$ as n^*. Our goal is to show that

$$\sup_{n\in\mathbb{N}} G_{v,\mathbf{0}}(n)=G_{v,\mathbf{0}}(n^*)\leqslant G_{u,\mathbf{0}}(b_{j^*}n^*)\leqslant\sup_{n\in\mathbb{N}}G_{u,\mathbf{0}}(n). \tag{5.5}$$

For any subset $w\subseteq v$, we define $w_+=w\cup\{j^*\}$. Then

$$\begin{aligned}
H_{u,w_+}&=b_{j^*}H_{v,w}, & H_{u,w}&=-H_{v,w},\\
m_{u,w_+,\mathbf{0}}&=b_{j^*}m_{v,w,\mathbf{0}}, \quad\text{and} & m_{u,w,\mathbf{0}}&=m_{v,w,\mathbf{0}}.
\end{aligned} \tag{5.6}$$

We also introduce $K(x)=x(1-x)$ to simplify some expressions. Starting with (5.3), applying identities from (5.6), and using (3.1), we get for any n divisible by b_{j^*} that

$$\begin{aligned}
\widetilde{G}_{u,\mathbf{0}}(n)&=\sum_{w\subseteq u}H_{u,w}m_{u,w,\mathbf{0}}K(\varepsilon_w)\\
&=\sum_{w\subseteq v}H_{u,w_+}m_{u,w_+,\mathbf{0}}K\Big(\frac{n}{m_{u,w_+,\mathbf{0}}}-\Big\lfloor\frac{n}{m_{u,w_+,\mathbf{0}}}\Big\rfloor\Big)\\
&\quad+\sum_{w\subseteq v}H_{u,w}m_{u,w,\mathbf{0}}K\Big(\frac{n}{m_{u,w,\mathbf{0}}}-\Big\lfloor\frac{n}{m_{u,w,\mathbf{0}}}\Big\rfloor\Big)\\
&=\sum_{w\subseteq v}b_{j^*}^2H_{v,w}m_{v,w,\mathbf{0}}K\Big(\frac{n}{b_{j^*}m_{v,w,\mathbf{0}}}-\Big\lfloor\frac{n}{b_{j^*}m_{v,w,\mathbf{0}}}\Big\rfloor\Big)\\
&\quad-\sum_{w\subseteq v}H_{v,w}m_{v,w,\mathbf{0}}K\Big(\frac{n}{m_{v,w,\mathbf{0}}}-\Big\lfloor\frac{n}{m_{v,w,\mathbf{0}}}\Big\rfloor\Big)\\
&=b_{j^*}^2\widetilde{G}_{v,\mathbf{0}}(n/b_{j^*})-\widetilde{G}_{v,\mathbf{0}}(n).
\end{aligned} \tag{5.7}$$

The corresponding normalized coefficient is

$$\begin{aligned}
G_{u,\mathbf{0}}(n)&=\frac{1}{n}\prod_{j\in u}(b_j-1)^{-1}\Big(b_{j^*}^2\widetilde{G}_{v,\mathbf{0}}(n/b_{j^*})-\widetilde{G}_{v,\mathbf{0}}(n)\Big)\\
&=\frac{b_{j^*}}{(b_{j^*}-1)n}\prod_{j\in v}(b_j-1)^{-1}\widetilde{G}_{v,\mathbf{0}}(n/b_{j^*})-\frac{1}{(b_{j^*}-1)b_{j^*}n}\prod_{j\in v}(b_j-1)^{-1}\widetilde{G}_{v,\mathbf{0}}(n)\\
&=\frac{b_{j^*}}{b_{j^*}-1}G_{v,\mathbf{0}}(n/b_{j^*})-\frac{1}{b_{j^*}-1}G_{v,\mathbf{0}}(n).
\end{aligned}$$

Now, using the fact that n^* is the maximizer of $G_{v,\mathbf{0}}(n)$

$$\begin{aligned} G_{u,\mathbf{0}}(b_{j^*}n^*) &= \frac{b_{j^*}}{b_{j^*}-1}G_{v,\mathbf{0}}(n^*) - \frac{1}{b_{j^*}-1}G_{v,\mathbf{0}}(b_{j^*}n^*) \\ &\geqslant G_{v,\mathbf{0}}(n^*). \end{aligned}$$

The theorem immediately follows from (5.5). ∎

THEOREM 5.3. *For scrambled Halton points, the gains satisfy*

$$\sup_{n\in\mathbb{N}} G_{u,\boldsymbol{k}}(n) \leqslant \prod_{j\in u-\{j_m\}} \frac{b_j}{b_j-1} \tag{5.8}$$

for all $d \geqslant 1$, all nonempty $u \subseteq 1{:}d$, and all $\boldsymbol{k} \in \mathbb{N}_0^u$, where $j_m = \arg\min_{j\in u} b_j$ and an empty product above equals 1 by convention.

Proof. According to Theorem 5.1, it suffices to prove the theorem for $\boldsymbol{k} = \mathbf{0}$. By Proposition 3.3, the largest $G_{u,\mathbf{0}}(n)$ arises for $1 \leqslant n \leqslant \overline{m}$, where $\overline{m} = \overline{m}_{u,\mathbf{0}} = m_{u,u,\mathbf{0}}$. Therefore

$$\sup_{n\in\mathbb{N}} G_{u,\boldsymbol{k}}(n) = \sup_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n) = \max_{1\leqslant n\leqslant\overline{m}} G_{u,\mathbf{0}}(n).$$

We proceed by induction on $|u|$. When u only contains a single element j, a straightforward calculation shows for $1 \leqslant n \leqslant b_j$ that

$$G_{u,\mathbf{0}}(n) = \frac{b_j - n}{b_j - 1}.$$

So $\sup_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n) = G_{u,\mathbf{0}}(1) = 1$ and the theorem is trivially true for $|u| = 1$.

Now for $|u| > 1$, we assume that (5.8) holds for $v = u \setminus \{j^*\}$ with $j^* \neq j_m$ and then prove it holds for u. From (5.7) and nonnegativity of $\widetilde{G}_{v,\mathbf{0}}(n)$,

$$\widetilde{G}_{u,\mathbf{0}}(n) \leqslant \sum_{w\subseteq v} b_{j^*}^2 H_{v,w} m_{v,w,\mathbf{0}} K\biggl(\frac{n}{b_{j^*}m_{v,w,\mathbf{0}}} - \Bigl\lfloor \frac{n}{b_{j^*}m_{v,w,\mathbf{0}}} \Bigr\rfloor\biggr).$$

Let $m_{v,w,*} = b_{j^*} m_{v,w,\mathbf{0}}$ and

$$G_{v,*}(n) = \frac{1}{n}\prod_{j\in v}(b_j-1)^{-1} \sum_{w\subseteq v} H_{v,w} m_{v,w,*} K\biggl(\frac{n}{m_{v,w,*}} - \Bigl\lfloor \frac{n}{m_{v,w,*}} \Bigr\rfloor\biggr).$$

We can proceed as in the proof of Theorem 5.1 with b^* replaced by b_{j^*} and conclude that

$$\sup_{n\in\mathbb{N}} G_{v,*}(n) = \sup_{n\in\mathbb{N}} G_{v,\mathbf{0}}(n) \leqslant \prod_{j\in v-\{j_m\}} \frac{b_j}{b_j-1}.$$

Hence

$$\begin{aligned} G_{u,\mathbf{0}}(n) &\leqslant \frac{1}{n}\prod_{j\in u}(b_j-1)^{-1} \sum_{w\subseteq v} b_{j^*}^2 H_{v,w} m_{v,w,\mathbf{0}} K\biggl(\frac{n}{b_{j^*}m_{v,w,\mathbf{0}}} - \Bigl\lfloor \frac{n}{b_{j^*}m_{v,w,\mathbf{0}}} \Bigr\rfloor\biggr) \\ &= \frac{b_{j^*}}{b_{j^*}-1} G_{v,*}(n) \\ &\leqslant \prod_{j\in u-\{j_m\}} \frac{b_j}{b_j-1} \end{aligned}$$

and the theorem follows from induction. ∎

COROLLARY 5.4. *For scrambled Halton points in dimension* $d \geqslant 1$

$$\sup_{n\in\mathbb{N}} \max_{\varnothing\neq u\subseteq 1:d} \sup_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n) \leqslant \frac{1}{2}\prod_{j=1}^{d}\frac{b_j}{b_j-1}.$$

Proof. The result holds for $d=1$. For $d\geqslant 2$,

$$\sup_{n\in\mathbb{N}} \max_{\varnothing\neq u\subseteq 1:d} \sup_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n) = \sup_{n\in\mathbb{N}} G_{1:d,\mathbf{0}}(n) \leqslant \prod_{j=2}^{d}\frac{b_j}{b_j-1} = \frac{1}{2}\prod_{j=1}^{d}\frac{b_j}{b_j-1}$$

with the inequality coming from Theorem 5.3. ∎

THEOREM 5.5. *For the scrambled Halton points*

$$\Gamma_d = \sup_{n\in\mathbb{N}} \max_{\varnothing\neq u\subseteq 1:d} \sup_{\boldsymbol{k}\in\mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n) = O(\log(d)) \tag{5.9}$$

as $d\to\infty$.

Proof. First, $\log(\Gamma_d) \leqslant \sum_{j=1}^{d}\log(b_j/(b_j-1))$, where b_j is the jth prime number. For any $j\geqslant 1$ we have $b_j > \underline{b}_j = j\log(j)$ by equation (3.12) of [25]. For any $\epsilon>0$, a Taylor expansion gives

$$\log\left(\frac{b_j}{b_j-1}\right) < \log\left(\frac{1}{1-1/\underline{b}_j}\right) < \frac{1}{\underline{b}_j} + \frac{1+\epsilon}{2\underline{b}_j^2}$$

for all $j\geqslant J_1 = J_1(\epsilon)$ for some $J_1<\infty$. Then for all large enough d, some $J_2 = J_2(\epsilon)\geqslant J_1(\epsilon)$ and some constants $c_\epsilon < c'_\epsilon < \infty$

$$\begin{aligned}\log(\Gamma_d) &< c_\epsilon + \int_{J_2-1}^{d}\frac{1}{x\log(x)}\,\mathrm{d}x + \int_{J_2-1}^{d}\frac{1+\epsilon}{2(x\log(x))^2}\,\mathrm{d}x\\ &< c'_\epsilon + \int_{J_2-1}^{d}\frac{1}{x\log(x)}\,\mathrm{d}x\\ &= \log(\log(d)) + O(1)\end{aligned}$$

as $d\to\infty$. Exponentiating this relationship establishes (5.9). ∎

6. A lower bound. Here we show that the gains cannot be $O(\log(d)^{1-\epsilon})$ for any $\epsilon>0$. First we get a bound for the gain factor of any set u that includes either $j=1$ or $j=2$. This is equivalently about whether either 2 or 3 are among the primes b_j for $j\in u$.

THEOREM 6.1. *For* $1\leqslant d<\infty$ *and* $u\subseteq 1{:}d$, *if* $u\cap\{1,2\}\neq\varnothing$ *then*

$$\sup_{n\in\mathbb{N}} G_{u,\boldsymbol{k}}(n) \geqslant \prod_{j\in u-\{j^*\}}\frac{b_j+1}{b_j}$$

for any $\boldsymbol{k}\in\mathbb{N}_0^u$, *where* j^* *is any element of* $u\cap\{1,2\}$.

Proof. According to Theorem 5.1, it suffices to prove the inequality for $\boldsymbol{k}=\mathbf{0}$. For $j^*\in u\cap\{1,2\}$, let $b_* = b_{j^*}$, $n^* = \prod_{j\in u, b_j\neq b_*} b_j$, and $V = \{v\subseteq u \mid j^*\in v\}$. Because $m_{u,v,\mathbf{0}}$ divides n^* for any $v\notin V$, $\varepsilon_v = n^*/m_{u,v,\mathbf{0}} - \lfloor n^*/m_{u,v,\mathbf{0}}\rfloor = 0$. Then (5.3) simplifies to

$$\widetilde{G}_{u,\mathbf{0}}(n^*) = \sum_{v\in V} H_{u,v} m_{u,v,\mathbf{0}} K\left(\frac{n^*}{m_{u,v,\mathbf{0}}} - \Big\lfloor \frac{n^*}{m_{u,v,\mathbf{0}}} \Big\rfloor\right)$$
$$= \sum_{v\in V} (-1)^{|u-v|} \left(\prod_{j\in v} {b_j}^2\right) K\left(\frac{1}{b_*}\prod_{j\in u-v} b_j - \left\lfloor \frac{1}{b_*}\prod_{j\in u-v} b_j \right\rfloor\right),$$

where $K(x) = x(1-x)$.

When $b_* = 2$, because $\prod_{j\in u-v} b_j$ is odd,

$$K\left(\frac{1}{b_*}\prod_{j\in u-v} b_j - \left\lfloor \frac{1}{b_*}\prod_{j\in u-v} b_j \right\rfloor\right) = K\Big(\frac{1}{2}\Big) = \frac{1}{4}.$$

When $b_* = 3$, because $\prod_{j\in u-v} b_j$ is an integer not divisible by 3,

$$K\left(\frac{1}{b_*}\prod_{j\in u-v} b_j - \left\lfloor \frac{1}{b_*}\prod_{j\in u-v} b_j \right\rfloor\right) = K\left(\frac{1}{3}\right) = K\left(\frac{2}{3}\right) = \frac{2}{9}.$$

In either case,

$$K\left(\frac{1}{b_*}\prod_{j\in u-v} b_j - \left\lfloor \frac{1}{b_*}\prod_{j\in u-v} b_j \right\rfloor\right) = \frac{b_*-1}{b_*^2}$$

and the normalized coefficient equals

$$G_{u,\mathbf{0}}(n^*) = \frac{1}{n^*}\prod_{j\in u}(b_j-1)^{-1} \sum_{v\in V}(-1)^{|u-v|}\left(\prod_{j\in v} {b_j}^2\right)\frac{b_*-1}{b_*^2}$$
$$= \prod_{j\in u, b_j\neq b_*} \frac{1}{b_j(b_j-1)} \sum_{v\in V}(-1)^{|u-v|} \prod_{j\in v, b_j\neq b_*} b_j^2$$
$$= \prod_{j\in u, b_j\neq b_*} \frac{1}{b_j(b_j-1)} \prod_{j\in u, b_j\neq b_*} (b_j^2-1)$$
$$= \prod_{j\in u, b_j\neq b_*} \frac{b_j+1}{b_j}.$$

Hence

$$\sup_{n\in\mathbb{N}} G_{u,\mathbf{0}}(n) \geqslant G_{u,\mathbf{0}}(n^*) = \prod_{j\in u, b_j\neq b^*} \frac{b_j+1}{b_j}. \qquad \square$$

For $d \geqslant 2$ we divide $\prod_{j=1}^d (b_j+1)/b_j$ by either 3/2 or 4/3 and still get a lower bound. It follows that

$$\Gamma_d \geqslant \frac{3}{4}\prod_{j=1}^{d} \frac{b_j+1}{b_j}$$

for $j \geqslant 2$, while $\Gamma_1 = 1$.

COROLLARY 6.2. *For any* $\epsilon > 0$

$$\Gamma_d = \sup_{n \geqslant 1} \max_{\varnothing \neq u \subseteq 1:d} \sup_{\boldsymbol{k} \in \mathbb{N}_0^u} G_{u,\boldsymbol{k}}(n)$$

cannot be $O((\log d)^{1-\epsilon})$.

Proof. First $1{:}d \cap \{1,2\} \neq \varnothing$, so Theorem 6.1 gives $\Gamma_{1:d} \geqslant \prod_{j=2}^{d}(b_j+1)/b_j$ (which is 1 for $d = 1$). As in the proof of Theorem 5.5 we note that if $j \geqslant 6$ then $b_j < j\log(j) + j\log(\log(j))$. Then for $0 < \epsilon' < \epsilon'' < \epsilon$ and large enough j

$$\log\left(\frac{b_j+1}{b_j}\right) \geqslant \frac{1-\epsilon'}{j\log(j)+j\log(\log(j))} \geqslant \frac{1-\epsilon''}{j\log(j)}.$$

Using an integral lower bound like the one in the proof of Theorem 6.1 we get

$$\log(\Gamma_{1:d}) \geqslant c + (1-\epsilon'')\log(\log(d))$$

for some $c \in \mathbb{R}$. After exponentiating, $\Gamma_{1:d}$ cannot be $O((\log d)^{1-\epsilon})$. Finally, $\Gamma_d = \Gamma_{1:d}$ by Theorems 5.1 and 5.2. ∎

7. Conclusions. When we score RQMC methods by their worst case variance relative to plain MC, then we find that scrambled Halton points attain a much better bound than scrambled Sobol' points do, while retaining the $o(1/n)$ variance property. This does not imply that scrambled Halton points will be generally better than scrambled Sobol' points in applications, because the integrands of interest may not be ones where scrambled Sobol' points perform poorly. It does make scrambled Halton points a useful approach for settings where never performing much worse than MC is a priority. We note that we could obtain a gain uniformly bounded in d if we were to slightly increase the values b_j in use. We do not recommend this as it would be detrimental to the equidistribution properties that QMC and RQMC are designed to produce.

Acknowledgments. We thank Nabil Kahale who asked about methods with better gain bounds than scrambled Sobol' points at MCM 2023, as well as C. D. Parada who raised the same question in an email. We thank two anonymous reviewers for their helpful comments.

REFERENCES

[1] E. BRAATEN AND G. WELLER, *An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration*, J. Comput. Phys., 33 (1979), pp. 249–258.
[2] W. CHEN, A. SRIVASTAV, AND G. TRAVAGLINI, EDS., *A Panorama of Discrepancy Theory*, Springer, Cham, Switzerland, 2014.
[3] S. DA VEIGA, F. GAMBOA, B. IOOSS, AND C. PRIEUR, *Basics and Trends in Sensitivity Analysis: Theory and Practice in R*, SIAM, Philadelphia, 2021.
[4] L. DEVROYE, *Non-uniform Random Variate Generation*, Springer, New York, 1986.
[5] J. DICK, F. Y. KUO, AND I. H. SLOAN, *High-dimensional integration: The quasi-Monte Carlo way*, Acta Numer., 22 (2013), pp. 133–288.
[6] J. DICK AND F. PILLICHSHAMMER, *Digital Sequences, Discrepancy and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010.
[7] H. FAURE, *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arith., 41 (1982), pp. 337–351.
[8] H. FAURE, *Good permutations for extreme discrepancy*, J. Number Theory, 42 (1992), pp. 47–56.
[9] H. FAURE AND C. LEMIEUX, *Generalized Halton sequences in* 2008*: A comparative study*, ACM Trans. Model. Comput. Simul., 19 (2009), pp. 15:1–15:31.

[10] T. Goda and K. Suzuki, *Improved bounds on the gain coefficients for digital nets in prime power base*, J. Complexity, 76 (2023), 101722.
[11] J. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math., 2 (1960), pp. 84–90.
[12] F. J. Hickernell, *Koksma-Hlawka Inequality*, Wiley StatsRef: Statistics Reference Online, Wiley, 2014.
[13] P. L'Ecuyer and C. Lemieux, *A survey of randomized quasi-Monte Carlo methods*, in Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications, M. Dror, P. L'Ecuyer, and F. Szidarovszki, eds., Kluwer Academic, Boston, 2002, pp. 419–474.
[14] J. Matoušek, *On the L^2–discrepancy for anchored boxes*, J. Complexity, 14 (1998), pp. 527–556.
[15] H. Niederreiter, *Point sets and sequences with small discrepancy*, Monatsh. Math., 104 (1987), pp. 273–337.
[16] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
[17] G. Ökten, M. Shah, and Y. Goncharov, *Random and deterministic digit permutations of the Halton sequence*, in Monte Carlo and Quasi-Monte Carlo Methods 2010, L. Plaskota and H. Woźniakowski, eds., Springer, Berlin, 2012, pp. 609–622.
[18] A. B. Owen, *Randomly permuted (t, m, s)-nets and (t, s)-sequences*, in Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and P. J.-S. Shiue, eds., Springer, New York, 1995, pp. 299–317.
[19] A. B. Owen, *Monte Carlo variance of scrambled net quadrature*, SIAM J. Numer. Anal., 34 (1997), pp. 1884–1910.
[20] A. B. Owen, *A Randomized Halton Algorithm in R*, Technical report, preprint, arXiv:1706.02808, 2017.
[21] A. B. Owen, *Practical Quasi-Monte Carlo*, https://artowen.su.domains/mc/practicalqmc.pdf (2023).
[22] A. B. Owen and Z. Pan, *Where are the logs?*, in Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer, Springer, Cham, Switzerland, 2022.
[23] A. B. Owen and D. Rudolf, *A strong law of large numbers for scrambled net integration*, SIAM Rev., 63 (2021), pp. 360–372.
[24] Z. Pan and A. B. Owen, *The nonzero gain coefficients of Sobol's sequences are always powers of two*, J. Complexity, 75 (2023), 101700.
[25] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math., 6 (1962), pp. 64–94.
[26] C. Schlier, *A Practitioner's View on QMC Integration*, Technical report, Universität Freiburg, Fakultät für Physik, Freiburg, Germany, 2002.
[27] I. M. Sobol', *The distribution of points in a cube and the accurate evaluation of integrals,* Zh. Vychisl. Mat. Mat. Phys., 7 (1967), pp. 784–802 (in Russian).
[28] B. Vandewoestyne and R. Cools, *Good permutations for deterministic scrambled Halton sequences in terms of L2-discrepancy*, J. Comput. Appl. Math., 189 (2006), pp. 341–361.
[29] X. Wang and F. J. Hickernell, *Randomized Halton sequences*, Math. Comput. Model., 32 (2000), pp. 887–899.