Accelerating the Convergence of Lattice Methods by Importance Sampling-Based Transformations

Earl Maize and John Sepikas and Jerome Spanier

Abstract Importance sampling is a powerful technique for improving the stochastic solution of quadrature problems as well as problems associated with the solution of integral equations, and a generalization of importance sampling, called weighted importance sampling, provides even more potential for error reduction. Additionally, lattice methods are particularly effective for integrating sufficiently smooth periodic functions. We will discuss the advantage of combining these ideas to transform non-periodic to periodic integrands over the unit hypercube to improve the convergence rates of lattice-based quadrature formulas. We provide a pair of examples that show that with the proper choice of importance transformation, the order in the rate of convergence of a quadrature formula can be increased significantly. This technique becomes even more effective when implemented using a family of multidimensional dyadic sequences generally called extensible lattices. Based on an extension of an idea of Sobol [17], extensible lattices are both infinite and at the same time return to lattice-based methods with the appropriate choice of sample size. The effectiveness of these sequences, both theoretically and with numerical results, is discussed. Also, there is an interesting parallel with low discrepancy sequences generated by the fractional parts of integer multiples of irrationals which may point the way to a useful construction method for extensible lattices.

Earl Maize
Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109 e-mail: earl.h.maize@jpl.nasa.gov

John Sepikas
Pasadena City College, Pasadena, CA 91105

Jerome Spanier
Beckman Laser Institute and Medical Clinic, 1002 Health Sciences Road East, University of California, Irvine CA 92612
1 Introduction

The need for estimates of multidimensional integrals is widespread. It is well known nowadays that quasi-Monte Carlo (qMC) methods can (sometimes surprisingly) provide better estimates for these purposes than classical deterministic quadrature formulas or pseudorandom Monte Carlo (MC) methods. All such qMC methods rely on the uniformity of the points selected in the integration domain. A highly desirable feature of any technique for forming such estimates is the possibility of adding sample points without recomputing the previously sampled points. For independent samples (the MC case) this is no problem but for correlated samples (the qMC case) the uniformity, as measured by the discrepancy, tends to be lowered in blocks whose size depends on the algorithm that generates the qMC sequence. Moreover, when the qMC sequence is generated by a conventional lattice rule, extending the sequence requires recomputing all of the sequence elements anew, as we will explain below. In other words, conventional lattice rule methods require deciding in advance how many qMC points are needed - an uncomfortable constraint when more points are needed. Overcoming this limitation leads to the notion of extensible lattice rules.

Let \( g \) be an \( s \)-dimensional vector of integers and form the sequence
\[
\mathbf{x}_n = \left\{ \frac{n}{N} g \right\} n = 0, 1, \ldots, N - 1 \tag{1}
\]
where the braces indicate the fractional part of each vector component. We are interested in using the \( \mathbf{x}_n \) as arguments for the approximation of an integral over the \( s \)-dimensional hypercube \( I^s \) by a sum:
\[
\theta = \int_{I^s} f(x) \, dx \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n). \tag{2}
\]

The formulas (1) and (2) define a rank-1 lattice rule, an idea that originated with Korobov [8], and has given rise to a great deal of interest in the intervening years, especially in the last 20 years.

It is well known that lattice methods are especially attractive when used to estimate integrals of smooth periodic functions. Consider the class \( E_\lambda^s(K) \) of all periodic functions \( f \) on \( I^s \) whose coefficients in the absolutely convergent Fourier series expansion
\[
f(x) = \sum_{\mathbf{h} \in \mathbb{Z}^s} c_\mathbf{h} \exp(2\pi i \mathbf{h} \cdot \mathbf{x}) \tag{3}
\]
satisfy the decay condition
\[
|c_\mathbf{h}| \leq K \frac{1}{r(\mathbf{h})^\lambda} \tag{4}
\]
with \( \lambda > 1 \) and where
\[
r(\mathbf{h}) = \max(1, |h_1|) \max(1, |h_2|) \cdots \max(1, |h_s|) \tag{5}
\]
Lattice Methods

and \( h = (h_1, \ldots, h_s) \in \mathbb{Z}^s \).

For such functions, a quick calculation shows that the error in a lattice method may be expressed as

\[
\left| \int_{\mu} f(x) dt - \frac{1}{N} \sum_{n=1}^{N} f(x_n) \right| = \sum_{\mathbf{h} \cdot \mathbf{g} \equiv 0 \text{ (mod } N)} c_{\mathbf{h}} \cdot \mathbf{g} \equiv 0 \text{ (mod } N) \leq K \sum_{\mathbf{h} \cdot \mathbf{g} \equiv 0 \text{ (mod } N)} r(\mathbf{h})^{-\lambda}. \tag{6}
\]

where the prime on the summation indicates that the sum is taken over all \( \mathbf{h} \in \mathbb{Z}^s \) except for \( \mathbf{h} = (0, \ldots, 0) \). The infinite sum appearing in (7) is a recurring figure of merit for lattice methods which we denote by

\[
P_{\lambda}(\mathbf{g}, N) = \sum_{\mathbf{h} \cdot \mathbf{g} \equiv 0 \text{ (mod } N)} r(\mathbf{h})^{-\lambda}. \tag{7}
\]

An excellent survey of lattice methods is in [13]. One can find there that there exist lattice methods whose errors satisfy

\[
\left| \int_{\mu} f(x) dt - \frac{1}{N} \sum_{n=1}^{N} f(x_n) \right| = O(N^{-\lambda} \left(\log N\right)^{\lambda s}). \tag{8}
\]

Note that the error expression in (9) now takes advantage of the additional smoothness of the integrand as represented by the size of \( \lambda \).

It is clear from (1) that the value of \( x_n \) depends on the choice of the integer \( N \). In his 1981 Ph.D. dissertation Maize [11] observed that certain infinite dyadic sequences (and their generalizations to other prime bases) can be used to define arguments \( x_n \) that do not depend on an \textit{a priori} choice of \( N \), as they do in the formula (1). This gives rise to the possibility of extending \( N \) - point lattice rule quadrature formulas to infinite sequences that revert to lattice rules for specific choices of \( N \). The simplest instance of these in the one dimensional case gives rise to the van der Corput sequence \( x_n = \phi_2(n) \) and produces an infinite sequence with the property that when \( N \) is any power of two, the points \( x_1, \ldots, x_N \) define a lattice. Maize pointed out that such sequences can be easily (and exactly) implemented in a computer and he showed how such sequences might be used to improve upon \( \left(\log N\right)^{\lambda s}/N \) convergence rates for periodic and sufficiently regular integrands.

Many of the ideas of Maize’s dissertation were published in the paper [21]. The idea of extensibility for lattice rules reappeared in a paper by Hickernell and Hong [4] and has subsequently been pursued further in [5], [6] and in other papers. Such infinite sequences, defined with respect to a number base \( b \) with the property that for every integer \( b \) the first \( b^m \) points form a lattice, are now called \textit{extendible lattice rules}. Such sequences, therefore, behave in much the same way as the initial segments of \((t, s)\) sequences do to form \((t, m, s)\) nets [12, 13]. The challenge is to es-
establish the existence of extensible lattices with favorable uniformity properties and exhibit constructive algorithms for their effective computation.

To capitalize on this possibility, we explore the potential advantage in converting nonperiodic to periodic integrands by applying transformations of the sort that are commonly used for other purposes in the context of pseudorandom Monte Carlo. Specifically, we will see that importance sampling transformations are useful candidates for such consideration.

These ideas will be illustrated by applying them to the evaluation of a simple three dimensional integral and a more challenging four dimensional example.

2 Extensible Lattices

The generation of extensible lattices in [11] was inspired by the notion of good direction numbers found in Sobol [17], which is itself a generalization of Korobov’s development of the theory of good lattice points [8]. The essential motivation is to find a way to preserve the desirable convergence properties of good lattice methods while at the same time maintaining an unlimited supply of sampling points.

2.1 Generation of Extensible Lattices

The general method for defining an extensible lattice is to select an increasing sequence of positive integers $N_1 < N_2 < \ldots$ and generating vectors of integers $g^{(1)}, g^{(2)}, \ldots$ such that each finite lattice sequence is nested within the next. That is,

$$\left\{ \frac{n}{N_k} g^{(k)}(n) \right\}_{n=0}^{N_k-1} \subseteq \left\{ \frac{n}{N_{k+1}} g^{(k+1)}(n) \right\}_{n=0}^{N_{k+1}-1}.$$  \hfill (10)

Figure 1 depicts such a nested lattice sequence.

One particular method for accomplishing this is to choose a prime $p$ and let $N_l = p^l$. If we then insist that the generating vectors satisfy $g^{(l+1)}(n) \equiv g^{(l)}(n) \mod (p^l)$ where the congruence is taken component-wise, it is easily seen that the inclusions (10) are satisfied.

If our only aim were to sample integrands with the fixed sample sizes $N_k$, $k = 1, 2, \ldots$, the definitions above would be sufficient. However, the practitioner may wish to choose an intermediate sample size, that is a sample size $N$ where $N_k < N < N_{k+1}$. For that we require a way to generate intermediate points in an extended lattice in a manner that distributes them uniformly over $I^I$. As it turns out, the van der Corput sequence provides an ideal mechanism for accomplishing this.

Since our immediate goal is to explore the practical application of this theory, we will from here on restrict ourselves to dyadic sequences; that is, extensible lattice sequences with $p = 2$. The reader is referred to [5] and [6] for the more general case.
We begin with the $s = 1$ case. Let $\nu^{(k)}, k = 1, 2, \ldots$ be a sequence of integers with $\nu^{(k+1)} \equiv \nu^{(k)} \pmod{2^k}$. For any integer $n$, represent $n$ in base 2 via $n = \varepsilon_1 + \varepsilon_2 2 + \cdots + \varepsilon_l 2^{l-1}$ and define the $n$th component of the sequence as

$$x^{(n)} = \left\{ \frac{\varepsilon_1}{2} \nu^{(1)} + \frac{\varepsilon_2}{4} \nu^{(2)} + \cdots + \frac{\varepsilon_l}{2^l} \nu^{(l)} \right\}. \quad (11)$$

Since $\nu^{(k+1)} \equiv \nu^{(k)} \pmod{2^k}$ it follows that (11) is equivalent to

$$x^{(n)} = \left\{ \phi_2(n) \nu^{(l)} \right\} \quad \text{where } l = \lfloor \log_2 n \rfloor + 1 \quad (12)$$

and $\phi_2(n)$ is van der Corput’s radical inverse function with base 2. Note that the choice $\nu^{(k)} = 1$ for all $k$ produces the classic van der Corput sequence.

Finally, given a sequence of $s$–dimensional vectors $\mathbf{v}^{(l)}$ whose components satisfy $v^{(l+1)}_j \equiv v^{(l)}_j \pmod{2^l}$ we may define the infinite $s$–dimensional sequence $\mathbf{x}^{(n)}$ component-wise via

$$x^{(n)}_j = \frac{\varepsilon_1}{2} v^{(1)}_j + \frac{\varepsilon_2}{4} v^{(2)}_j + \cdots + \frac{\varepsilon_l}{2^l} v^{(l)}_j \quad (13)$$

or as noted above,

$$x^{(n)}_j = \left\{ \phi_2(n) v^{(n)}_j \right\} \quad \text{where } l = \lfloor \log_2 n \rfloor + 1. \quad (14)$$

and then form the $s$–dimensional sequence via

$$\mathbf{x}^{(n)} = \left\{ \phi_2(n) \mathbf{v}^{(l)} \right\} \quad (15)$$
where the fractional parts are taken component-wise.

Recall that for a prime \( p \), the definition of a \( p \)-adic integer is an infinite sequence \( \{a^{(1)}, a^{(2)}, \ldots\} \) of integers where \( a^{(k+1)} \equiv a^{(k)} \pmod{p^k} \). Denote by \( O_p \), the set of all such sequences with the canonical representation \( a^{(k+1)} = a^{(k)} + bp^k \) where \( b \in \{0, 1, 2, \ldots, p-1\} \). With addition and multiplication defined componentwise (and reduction to canonical form as required), the set \( O_p \), called the \( p \)-adic integers, is an integral domain and can be imbedded in a field called the \( p \)-adic numbers. An ordinary integer \( n \) is represented in \( O_p \) via the sequence \( \{n, n, \ldots, n, \ldots\} \). In the context of \( p \)-adic integers, ordinary integers are called rational integers.

Returning to the definition of the generating vector, we observe that a sequence of integers \( \nu^{(k)} \), \( k = 1, 2, \ldots \) satisfying \( \nu^{(k+1)} \equiv \nu^{(k)} \pmod{2^k} \) is simply a 2-adic integer. It is quite straightforward to see (Maize[11], Hickernell, et. al. [5]) that by viewing the binary fractions as naturally imbedded within the 2-adic numbers and applying 2-adic arithmetic, (15) is represented by

\[
x^{(n)} = \{\phi_2(n)\nu\}
\]

where \( \nu \) is a vector of 2-adic integers.

### 2.2 Distribution Properties of Extensible Lattices

In the previous section, we described a method for generating extensible lattices which can be compactly expressed via equation (16). The existence of uniformly distributed extensible lattice sequences is confirmed via the following theorem (Maize[11]).

**Theorem 1.** Let \( \nu_1, \ldots, \nu_s \) be 2-adic integers and let \( \nu = (\nu_1, \ldots, \nu_s) \). The s-dimensional infinite sequence \( x^{(n)} = \{\phi_2(n)\nu\} \) is uniformly distributed (see [10]) if and only if \( \nu_1, \ldots, \nu_s \) are linearly independent over the rational integers.

The proof is accomplished via the use of Weyl’s criterion. Note the very intriguing analog between sequences generated by (16) and sequences of the form \( x^{(n)} = \{n\alpha\} \) where \( \alpha \) is an \( s \)-dimensional vector of irrational numbers. As it turns out in both cases, the equidistribution of the sequence hinges on the independence of the components of the generating vector over the rational numbers. We will expand upon this observation later.

Clearly there is an ample supply of generating vectors since the cardinality of the 2-adic integers is that of the continuum and any independent set will, in the limit, correctly integrate any function in \( E_\lambda^K \). The question then turns to the quantitative performance of these sequences. Most recently, Hickernell and Niederreiter [6] have examined the distribution properties of extensible lattices with respect to several figures of merit. Their result germane to our discussion here is summarized in the following theorem.
Theorem 2. For a given dimension $s$ and any $\varepsilon > 0$, there exist 2-adic generating vectors $\mathbf{v}$ and a constant $C(\lambda, \varepsilon, s)$ such that
\[
P_{\lambda}(\mathbf{v}^{(l)}, 2^l) \leq C(\lambda, \varepsilon, s)2^{-\lambda l}(\log 2^l)^{\lambda(s+1)}[\log \log (2^l + 1)]^{\lambda(1+\varepsilon)}
\]  
for $l = 1, 2, \ldots$.

where the figure of merit $P_{\lambda}(\mathbf{v}^{(l)}, 2^l)$ is defined in (8). Considering the very slowly increasing $\log(\log)$ term, we see that, comparing (17) with (7), the potential penalty for requiring that the lattices be nested is only slightly more than an additional factor of $\log(2^l)^{\lambda}$. 

2.3 Construction of Generating Vectors

From the previous sections, we know that uniformly distributed extensible lattices not only exist, but at least in theory, have distribution properties that are worse by only a factor slightly larger than $\log(N)^{\lambda}$ when compared with the best known results for general lattice methods. Unfortunately, these results are based on averaging techniques and none provides an explicit representation for good generating vectors. We are left in the position of having a good theoretical method but with no practical path to implementation.

One idea for the construction of generating vectors is a “bootstrap” method whereby one picks an appropriate figure of merit and an initial guess for a generating vector. One then examines the figure of merit for all possible candidate generating vectors of the form $\mathbf{v}^{(l+1)} \equiv \mathbf{v}^{(l)} \pmod{p^l}$. A potential pitfall of this method of course, is that, while at each step it does guarantee that the next component of the generating vector will be optimal with respect to the previous choices, it does not guarantee global optimality. Maize [11] numerically explored this technique for $\lambda = 2$ and $p = 2$. It was further studied by Hickernell, et al [5]. Niederreiter and Pillichshammer [14] have examined this method in the more general context of weighted Korobov spaces using several different figures of merit and have provided some very positive results regarding this process. For the figure of merit $P_{\lambda}(\mathbf{v}^{(l)}, 2^l)$ and remaining in base 2, the algorithm in [14] may be described as follows:

Step 1: Set $\mathbf{v}^{(1)} = (1, 1, \ldots, 1)$
Step 2: For $k = 2, 3, \ldots$, choose $\mathbf{v}^{(k)} = \mathbf{v}^{(k-1)} + 2^{k-1} \mathbf{e}$, where $\mathbf{e}_j = 0, 1$, so that $P_{\lambda}(\mathbf{v}^{(l)}, 2^l)$ is minimized. The following theorem appears in [14].

Theorem 3. With the algorithm above,
\[
P_{\lambda}(\mathbf{v}^{(l)}, 2^l) = \sum_{\mathbf{h} \cdot \mathbf{v}^{(l)} = 0 \pmod{2^l}} r(\mathbf{h})^{-2} \leq \left( \prod_{j=1}^{s} (1 + 2\zeta(\lambda)) - 1 \right) \min \left( l, \frac{2^l - 1}{2^{\lambda - 1} - 1} \right) \frac{1}{2^l}
\]  
where $\zeta$ is the Riemann zeta function.
We can see that there is a gap between this algorithm’s performance and the best results of Section 2.2. In particular, we have lost, except in the constant multiplier, any dependence on the smoothness of the integrand, $\lambda$. As noted in Maize [11] and Niederreiter and Pillichshammer [14], numerical evidence suggests that there is substantial room for improvement. We present some of the evidence in the next section.

### 2.4 Numerical Investigations

For the figure of merit $P_\lambda(\nu_\nu(\nu^l), 2^l)$ we have implemented the algorithm in the previous section for the choice $\lambda = 2$ and for dimensions $s = 1, 2, \ldots, 10$ and for sample sizes up to $2^{28}$. The normalized results multiplied by $N = 2^l$ are plotted in Figure 2. Examining the curves in the figure, it is clear that the error term is approaching zero more rapidly than $2^{-l}$ suggested by the best known theoretical results for the algorithm. The numerical results appear much more promising. In fact, referring to

![Figure 2 Normalized Figure of Merit $2^l \ast P_\lambda(\nu_\nu(\nu^l), 2^l)$](image-url)

Figure 3 where the same results are normalized by the convergence rates inferred from Theorem 2, it may not be too reckless to conjecture that there are generating vectors that will produce a convergence rate of
\[
\left| \int_{I} f(x) \, dt - \frac{1}{N} \sum_{n=1}^{N} f(x_n) \right| = O((\log N)^{\lambda s}/N^{\lambda}) \text{ for } f \in E^{\lambda}_{\lambda}(K) \tag{19}
\]

and \( N = 2^l \).

3 Integration of Periodic and Non-periodic Functions

Traditionally, importance sampling [3, 19] provides a standard Monte Carlo technique for reducing the variance in pseudorandom estimates of an integral and in solving integral equations. It achieves this by evaluating the integrand at points that are nonuniformly distributed and reweighting the integrand values to eliminate the bias engendered by the nonuniform sampling. The method attempts to generate sample points preferentially in subdomains more “important” in the sense of their relative contribution to estimates of the integral under consideration. Here we want to use the methodology of importance sampling, and its generalization to weighted uniform sampling [16, 18], not to redistribute the sample points, but rather to convert a nonperiodic integrand to one that is periodic and smooth. Our interest in doing so is to gain access to the higher rates of convergence promised by lattice methods when applied to smooth periodic integrands.
3.1 Theoretical Convergence Rates

Standard MC methods converge, of course, at the rate forecast by the central limit theorem. Thus, as the number $N$ of samples increases the integration error (as measured by the standard deviation, or the relative standard deviation, of the sample mean) reduces asymptotically as $O(N^{-1/2})$ for any $L_2$ integrand. For qMC methods there are various measures of the error, frequently referred to as figures of merit, such as the quantity $P_\lambda$ that we have consistently used in this paper. Other figures of merit are introduced elsewhere in the literature, and a more extensive discussion can be found, e.g., in [14].

3.2 Use of Importance Sampling to Periodicize

Given the improved convergence rates for lattice methods when applied to smooth periodic functions, it seems reasonable to investigate whether quadrature problems, especially those in high dimensions, can benefit from conversion of nonperiodic to periodic integrands. This is not a new idea; it was discussed already in [9, 22, 7]. More recently, the book [13] provides a number of other references related to this topic.

In Paul Chelson’s 1976 dissertation [2], see also [21] and [20], a primary focus was to make rigorous the possibility of applying quasirandom methods to the estimation of finite dimensional integrals and solutions of matrix and integral equations. The latter problems are infinite dimensional in the sense that the underlying sample space is infinite dimensional. Further, if that could be shown, Chelson wanted to know whether variance reduction techniques, such as importance sampling, could be useful in the qMC context. Chelson found that this is, indeed, the case and he established a generalized Koksma-Hlawka inequality for both the finite and infinite dimensional instances in which the term $V(f)$ involving the variation of the integrand is replaced by $V(f/g)$, where $g$ plays the role of an importance function. This gives rise to the possibility that the function $g$ can be chosen in such a way that $V(f/g) \ll V(f)$. Such an idea could then improve the estimate of the integral, but it does not increase the rate of convergence of the sum to the integral. In [11] Maize generalized Chelson’s results to weighted importance sampling [16], [18] and established a Koksma-Hlawka inequality in which the variation of the integrand $V(f)$ is replaced by $V((f-\theta h)/g)$ where $h$ is a positive weighting function.

These results established that these methods, so useful for MC applications, might offer similar advantages in the qMC context. Whereas importance sampling requires the selection of a suitable importance function, weighted uniform sampling offers much greater flexibility. The reader is referred to [21] for a discussion of these ideas.

For the estimation of $s$-dimensional integrals, the importance sampling formulation chooses a (usually) nonuniform distribution function $G(x)$ on $I^s$ and uses the
Lattice Methods

estimate

\[ \theta = \int f(t)dt \approx \frac{1}{N} \sum_{n=1}^{N} f(G^{-1}(x_n)) / g(G^{-1}(x_n)) \]  

(20)

in place of

\[ \int f(t)dt \approx \frac{1}{N} \sum_{n=1}^{N} f(x_n), \]  

(21)

where \( g(x) \) is the probability density function corresponding to \( G \). Instead of choosing \( G \) to minimize the variance (or variation in the qMC instance) of the estimator \( f(x)/g(x) \), we can choose \( G \) to convert \( f \) to a smooth periodic function and hopefully take advantage of higher order convergence rates. More generally according to Maize [11] (see also [21]), we can select a nonnegative weighting function \( h(t) \) whose integral over \( I_s \) is 1 and use

\[ \int f(t)dt \approx \sum_{n=1}^{N} \frac{f(G^{-1}(x_n))}{g(G^{-1}(x_n))} / \sum_{n=1}^{N} \frac{h(G^{-1}(x_n))}{g(G^{-1}(x_n))} \]  

(22)

in place of (21).

For example, a simple way to “periodize” an integrand is via the Bernstein polynomials

\[ \frac{1}{g(x)} = B_\alpha(x) = Kx^\alpha(1-x)^\alpha \]  

(23)

with a normalizing constant \( K \), which is clearly a periodic function on \( I \). With such a definition, it is a simple matter to calculate \( G^{-1}(x) \).

A potential danger of this method is that the Jacobian of the resulting transformation can produce much larger derivatives than those of \( f \), adversely affecting the error bounds. While we might obtain a better convergence rate, the implied constants multiplying the error term may have grown to the point where we have lost the advantage. An additional complication is that computing \( G^{-1}(x) \) can be quite cumbersome for practical problems. This is where the choice of the weighting function \( h \) can be used. As stated above, [11] (see also [21]) provides a bound for the error in weighted uniform sampling that is proportional to \( V((f - \theta h)/g) \). From this we can see that if \( h \) is chosen to mimic the behavior of the integrand, for example by choosing \( h \) to be a low order approximation of \( f \) that is readily integrated, it can both relieve the requirement that \( g \) closely mimic the integrand and at the same time, reduce the constant multipliers in the error term.

4 Examples

Let us suppose that we wish to evaluate the integral of the function \( f(x,y,z) = 4x^2yze^{xy} \) over the three-dimensional unit cube. This integral is easily evaluated to
\[ \theta = \int_{I^3} 4x^2yze^{xy}dxdydz = 2(3 - e). \]  
(24)

Figure 4 contrasts the performance of MC and a few qMC estimators in estimating this integral. The quadrature errors incurred from the use of a pseudorandom sequence, the Halton sequence, an \( \{ n\alpha \} \) sequence, and the 3-dimensional extensible lattice sequence based on the numerical algorithms from Section 2.4 are plotted versus sample size. One can easily see the rather slow \( \sqrt{N} \) performance of the pseudorandom technique and the much better performance of the quasi-random sequences.

We now focus on using an extensible lattice and the techniques from 3.2 to improve the performance of the estimates. We choose a very simple one-dimensional importance function for each variable based on the first order Bernstein polynomial

\[ G^{-1}(t) = 3t^2 - 2t^3 \]  
(25)

where \( t = x, y, \) or \( z \). A simple calculation shows that

\[ \frac{f(G^{-1}(x))}{g(G^{-1}(x))} = 6xyz(1-x)(1-y)(1-z)f(G^{-1}(x)). \]  
(26)

For a weighting function, we will mimic the behavior of \( f \) by approximating the exponential term with the first three terms of the appropriate Taylor series. After
performs the appropriate normalizations, we obtain

\[
h(x, y, z) = \frac{80}{11} x^2 y z \left(1 + xy + \frac{(xy)^2}{2}\right). \tag{27}
\]

Figure 5 illustrates the results of these numerical studies. We can see that using importance sampling to periodicize the integrand does, indeed, result in an improved convergence rate. In addition, the use of the weighting function with importance sampling maintains the improved convergence rate while improving the constant multiplier. This fairly modest example shows that when the integrand is sufficiently regular, the technique of weighted importance sampling, implemented with extensible lattices, can be an effective technique for reducing error in qMC computations.

**Fig. 5** qMC, Importance Sampling, and Weighted Importance Sampling with an Extensible Lattice

As a second example, let us consider the 4-dimensional integral

\[
\theta = \int_{\mathbb{R}^4} \left(1 + \|x\|^2\right)^{1/2} e^{-\|x\|^2} dx.
\]

which has been studied in [1] and [15]. A change of variables yields the equivalent integral

\[
\theta = \int_{\mathbb{R}^4} \left(1 + \sum_{j=1}^{4} (\varphi^{-1}(t_j)/2)^2\right)^{1/2} dt \tag{29}
\]
where $\varphi$ is the cumulative normal distribution function. For this example, we will consider varying the order of the Bernstein polynomials used to "periodize" the integrand. Figure 6 compares the relative errors derived from our method ($\alpha = 2, 3, 4$) with those based on the Genz-Patterson method used [1] for reference. Again in this more challenging example we see the benefits of our approach. Our second order method performs as well and the higher order methods outperform the Genz-Patterson quadrature method.

![Fig. 6 Periodization with Variable Order Bernstein Polynomials](image)

## 5 Summary and Future Work

We have described early efforts to develop infinite sequences of points in $I^n$ whose initial segments of length $2^m$ form a series of ever larger lattices; sequences now called extensible lattice sequences. Aside from their attractiveness for estimating finite dimensional integrals, in 3.2 we mentioned that extensible lattice sequences can also be used to solve infinite dimensional problems such as those characterized by matrix and integral equations. The sequences introduced in Section 2.1 seem well suited to this task and we hope to report in a subsequent publication on our efforts to explore this idea.
A second area of future work is to fill any part of the gap between the best currently known theoretical convergence rates and the often more optimistic evidence provided in various numerical experiments, including our own.

Finally, we plan to investigate the possible value of constructive methods for extensible lattices that make use of $p$-adic irrationals, as suggested by Theorem 1. We close with one final piece of evidence that this last idea may be a fruitful approach.

Recall from Theorem 1 that a sufficient condition for uniform distribution of a sequence of the form (16) is that the components of the generating vector $\mathbf{v}$ viewed as 2-adic integers must be linearly independent over the rationals. The first quadratic irrational in the 2-adic numbers is $\sqrt{17}$ whose first few terms as a 2-adic integer are given by $\sqrt{17} = \{1, 3, 7, 7, 23, 23, 23, 23, 23, 23, 279, \ldots \}$. We can select the generating vector $\mathbf{v} = (1, \sqrt{17})$ in two dimensions and generate the sequence as defined in (16). Figure 7 plots the normalized figure of merit $P_2(\mathbf{v}, 2^l)$ alongside the result from the optimum generating vector for $s = 2$. As Figure 7 clearly shows, while initially not as good as the vector found by the exhaustive search, $\mathbf{v} = (1, \sqrt{17})$ is indeed an effective generator of an extensible lattice for $s = 2$.

![Fig. 7 Normalized Figure of Merit for the Generator $\mathbf{v} = (1, \sqrt{17})$](image)

**Acknowledgements.** The first two authors wish to dedicate this paper to their friend and mentor, Dr. Jerome Spanier on the occasion of his eightieth birthday. The last author gratefully acknowledges partial support from the Laser Microbeam and Medical Program (LAMMP), an NIH
Biomedical Technology Resource Center (P41-RR01192). The authors would also like to thank the referees for helpful remarks and suggestions that improved the manuscript.

References