ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM – JOHANN RADON INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Column reduced digital nets

V. Anupindi, P. Kritzer

RICAM-Report 2024-09

Column reduced digital nets

V. Anupindi, P. Kritzer

May 17, 2024

Abstract

Digital nets provide an efficient way to generate integration nodes of quasi-Monte Carlo (QMC) rules. For certain applications, as e.g. in Uncertainty Quantification, we are interested in obtaining a speed-up in computing products of a matrix with the vectors corresponding to the nodes of a QMC rule. In the recent paper *The fast reduced QMC matrix-vector product* (J. Comput. Appl. Math. 440, 115642, 2024), a speed up was obtained by using so-called reduced lattices and row reduced digital nets. In this work, we propose a multiplication algorithm where we exploit the repetitive structure of column reduced digital nets instead of row reduced digital nets. We also provide an upper bound for the quality parameter of column reduced digital nets, which is useful for error analysis and has an advantage over the corresponding result for row reduced digital nets.

1 Introduction

1.1 The problem setting

In many applications, such as in statistics, finance, and uncertainty quantification, we would like to numerically compute

$$\int_D f(\boldsymbol{x}^\top A)\,\mathrm{d}\mu(\boldsymbol{x}), \tag{1}$$

where A is a real $s \times \tau$ matrix, by quasi-Monte Carlo (QMC) rules

$$Q_N(f) := \frac{1}{N}\sum_{k=0}^{N-1} f(\boldsymbol{x}_k^\top A), \tag{2}$$

where $\boldsymbol{x}_k = (x_k^{(1)}, \ldots, x_k^{(s)})^\top$ are column vectors corresponding to the points used in the QMC rule. Problems of this kind particularly arise in some important applications in statistics and uncertainty quantification. For instance, this approach can be used when approximating the expected value of a function with a multivariate normal random variable with some given covariance matrix, or when approximating the expected value of the solution of a PDE with random coefficients, see, e.g., [4].

Computing the vector-matrix products $\boldsymbol{x}_k^\top A$ for all $k \in \{0, \ldots, N-1\}$ takes $\mathcal{O}(Ns\,\tau)$ operations. This problem is equivalent to computing the matrix-matrix product XA, where

$$X = \left[\boldsymbol{x}_0^\top, \boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_{N-1}^\top\right]^\top$$

is the $N \times s$ matrix whose k-th row is $\boldsymbol{x}_k$. Computing XA can be infeasible in situations where s and N are both large (which happens in many applications).

In the paper [4], it is shown that when using particular types of QMC rules, the cost to evaluate $Q_N(f)$, as in (2), can be reduced to only $\mathcal{O}(\tau\, N \log N)$ operations provided that $\log N \ll s$. This reduction in computational cost is achieved by a fast matrix-matrix multiplication exploiting the fact that for specifically chosen point sets, such as (polynomial) lattice rules, the matrix X can be re-ordered to be of circulant structure.

The recent paper [1] studies an alternative method to reduce the computation time by imposing a certain structure of the points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}$. The key idea of this approach is to find situations in which the components of the points $\boldsymbol{x}_k$ have a certain repetitive structure, which then facilitates systematic fast computation of the products $\boldsymbol{x}_k^\top A$. This can be achieved by suitable modifications of (polynomial) lattice point sets using ideas from [2], but how to implement this idea for digital nets, which are more general than polynomial lattice point sets and among the most commonly used QMC node sets, is not straightforward. In [1], the authors made a first attempt and studied a reduction of the computation time for digital (t, m, s)-nets by setting certain *rows* of the generating matrices to zero (we refer to Section 1.2 for the precise definition of digital nets and their generating matrices). The basic idea in [1] is that for each of the s generating matrices $C_j^{(m)}$, $1 \le j \le s$, of the digital net, we identify a so-called reduction index $w_j \in \mathbb{Z}$ and set the last w_j rows of $C_j^{(m)}$ equal to zero. As shown in [1], this introduces a certain repetitiveness in the entries of the matrix X and speeds up the computation of the matrix-matrix product XA. We call such digital nets *row reduced* digital nets. However, for assessing the quality of reduced nets when used in QMC rules, it is more natural to study the situation where certain *columns* of the generating matrices are set to zero, since this directly corresponds to the reduced (polynomial) lattice point sets, resulting in the consideration of *column reduced* digital nets. The idea of column reduced digital nets is to set the last w_j columns of the generating matrix $C_j^{(m)}$, $1 \le j \le s$, equal to zero, instead of setting rows equal to zero. Furthermore, in the present paper, we focus on digital nets that are obtained from *digital sequences*, which implies additional structure in the generating matrices. Again, the approach of using column reduced digital nets yields a speed-up in the computation of XA, but as we will see below, it also makes it easier to assess the properties of the resulting column reduced digital nets than doing the same for row reduced digital nets. Furthermore, the error analysis for approximating (1) by (2) becomes easier. This idea was already mentioned (but not pursued) in [1], and this is what we intend to do in the present paper.

1.2 Digital nets and sequences

In this section, we give the definitions of (t, m, s)-nets and (t, s)-sequences, the digital construction method for these, and shortly outline how to assess their quality.

Let $\mathbb{F}_b$ be a finite field with b elements, where b is prime. We identify the elements of $\mathbb{F}_b$ with the set $\{0, 1, \ldots, b-1\}$. An *elementary interval* in base b and dimension s is a half-open interval of the form $\prod_{j=1}^{s} [a_j b^{-d_j}, (a_j+1) b^{-d_j})$ where the a_j, d_j are nonnegative integers with $0 \le a_j < b^{d_j}$ for $1 \le j \le s$.

In the following, we recall the definition of (t, m, s)-nets and (t, s)-sequences, which have the property that the number of points in certain elementary intervals is proportional to their sizes. This guarantees a degree of uniform distribution of the point set in $[0, 1)^s$, which is desirable when using such a point set in a QMC rule. For detailed discussions on (t, m, s)-nets and (t, s)-sequences, we refer to [6, 9].

Definition 1. For a given dimension $s \ge 1$ and nonnegative integers t, m with $0 \le t \le m$, a

(t, m, s)-*net* in base b is a point set $\mathcal{P} \subset [0,1)^s$ consisting of b^m points such that any elementary interval in base b with volume b^{t-m} contains exactly b^t points of $\mathcal{P}$.

A sequence $(\boldsymbol{x}_0, \boldsymbol{x}_1, \dots)$ of points in $[0,1)^s$ is called a (t, s)-*sequence* in base b if for all integers $m \geq t$ and $k \geq 0$, the point set consisting of the points $\boldsymbol{x}_{kb^m}, \dots, \boldsymbol{x}_{kb^m+b^m-1}$ forms a (t, m, s)-net in base b.

Note that the lower the value of t of a (t, m, s)-net or a (t, s)-sequence, the more uniformly the points are distributed in $[0,1)^s$, which is a desirable property when the point set is used as an integration node set in a QMC rule. This is the reason why t is referred to as the *quality parameter* of a net or sequence.

A (t, m, s)-net is called *strict*, if it does not fulfill the requirements of a $(t-1, m, s)$-net (for $t \geq 1$), and analogously for (t, s)-sequences. In general, any (t, m, s)-net is also a $(t+1, m, s)$-net for $t < m$.

We point out that it is, in general, a non-trivial combinatorial question of which values of t can be reached for which configurations of the other parameters. We again refer to [6, 9] for details.

A common way to generate (t, m, s)-nets and (t, s)-sequences is using the *digital method*, which was first introduced by Niederreiter in [8].

Definition 2. A *digital* (t, m, s)*-net over* $\mathbb{F}_b$ is a (t, m, s)-net $\mathcal{P} = \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_{b^m-1}\}$ where the points are constructed as follows. Let $C_1^{(m)}, \dots, C_s^{(m)}$ in $\mathbb{F}_b^{m \times m}$ be matrices over $\mathbb{F}_b$. To generate the k-th point in $\mathcal{P}$, $0 \leq k \leq b^m - 1$, we use the b-adic expansion $k = \sum_{i=0}^{m-1} k_i b^i$ with digits $k_i \in \{0, \dots, b-1\}$ which we denote by $\overrightarrow{k} = (k_0, \dots, k_{m-1})^\top$. The j-th coordinate $x_{k,j}$ of $\boldsymbol{x}_k = (x_{k,1}, \dots, x_{k,s})$ is obtained by computing

$$\overrightarrow{x}_{k,j} := C_j^{(m)} \overrightarrow{k},$$

and then setting

$$x_{k,j} := \overrightarrow{x}_{k,j} \cdot (b^{-1}, b^{-2}, \dots, b^{-m}).$$

Similarly, a *digital* (t, s)-sequence $\mathcal{S}$ over $\mathbb{F}_b$ is generated by infinite matrices $C_1, \dots, C_s$, where

$$C_j = (c_{i,r}^{(j)})_{i,r \in \mathbb{N}} \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}. \tag{3}$$

To generate the k-th point in $\mathcal{S}$, $k \geq 0$, we use the b-adic expansion $k = \sum_{i=0}^{\infty} k_i b^i$ with digits $k_i \in \{0, \dots, b-1\}$ which we denote by $\overrightarrow{k} = (k_0, k_1, \dots)^\top$. The j-th coordinate $x_{k,j}$ of $\boldsymbol{x}_k = (x_{k,1}, \dots, x_{k,s})$ is obtained by computing

$$\overrightarrow{x}_{k,j} := C_j^{(m)} \overrightarrow{k},$$

and then setting

$$x_{k,j} := \overrightarrow{x}_{k,j} \cdot (b^{-1}, b^{-2}, \dots).$$

Note that from any digital (t, s)-sequence over $\mathbb{F}_b$ with generating matrices $C_1, \dots, C_s$, we can, for $m \geq t$, derive a digital (t, m, s)-net over $\mathbb{F}_b$, simply by considering the point set generated by the left upper $m \times m$ submatrices $C_1^{(m)}, \dots, C_s^{(m)}$ of $C_1, \dots, C_s$. This is equivalent to considering the first b^m points of the (t, s)-sequence.

As pointed out above, the quality of a (t, m, s)-net or (t, s)-sequence is determined by its t-value. For digital (t, m, s)-nets and (t, s)-sequences, we can determine the t-value from rank conditions on the generating matrices, using a quantity that we shall refer to as the *linear independence parameter*.

Definition 3. For any integers $1 \le j \le s$ and $m \ge 1$, let $C_1^{(m)}, C_2^{(m)}, \ldots, C_s^{(m)}$ be $m \times m$ matrices over $\mathbb{F}_b$. Then the *linear independence parameter* $\rho_m(C_1^{(m)}, C_2^{(m)}, \ldots, C_s^{(m)})$ is defined as the largest integer such that for any choice of $d_1, \ldots, d_s \in \mathbb{N}_0$, with $d_1 + \cdots + d_s = \rho_m$, we have that

$$\begin{aligned}
&\text{the first } d_1 \text{ rows of } C_1^{(m)} \text{ together with}\\
&\text{the first } d_2 \text{ rows of } C_2^{(m)} \text{ together with}\\
&\vdots\\
&\text{the first } d_s \text{ rows of } C_s^{(m)}
\end{aligned}$$

are linearly independent over $\mathbb{F}_b$.

It is known (see, e.g., [6, 9]) that the generating matrices $C_1^{(m)}, C_2^{(m)}, \ldots, C_s^{(m)}$ of a digital (t, m, s)-net over $\mathbb{F}_b$ satisfy

$$\rho_m(C_1^{(m)}, C_2^{(m)}, \ldots, C_s^{(m)}) \ge m - t, \tag{4}$$

where we have equality if the net is a strict (t, m, s)-net. Similarly, for the generating matrices $C_1, \ldots, C_s$ of a digital (t, s)-sequence over $\mathbb{F}_b$ we must have $\rho_m(C_1^{(m)}, \ldots, C_s^{(m)}) \ge m - t$ for all $m \ge \max\{t, 1\}$, where $C_j^{(m)}$ denotes the left upper $m \times m$ submatrix of C_j for $j \in \{1, \ldots, s\}$. Hence, for digital nets and sequences, their quality can be assessed by checking linear independence conditions on the rows of the generating matrices.

2 The t-values of column reduced digital nets

Now we turn towards the primary object of our study, which is the column reduced digital nets. We note that if we take a general digital (t, m, s)-net and set some columns of its generating matrices to zero, we cannot control the quality parameter of the reduced net. However, since digital (t, s)-sequences require stronger conditions on their generating matrices, we can estimate the quality parameter of reduced digital (t, m, s)-nets derived from digital sequences by taking the nets generated by the left upper $m \times m$ submatrices of the generating matrices of the sequences.

For $m \ge t$, we consider the digital (t, m, s)-net generated by the matrices $C_1^{(m)}, \ldots, C_s^{(m)}$, derived via the above principle from a digital (t, s)-sequence with generating matrices $C_1, \ldots, C_s$, $C_j = (c_{i,r}^{(j)})$, $i, r \in \mathbb{N}$.

Let $0 = w_1 \le \cdots \le w_s \in \mathbb{N}_0$, we call these numbers the *reduction indices*, for the generating matrices $C_j^{(m)}$. We derive the corresponding reduced matrices $\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}$, with $\widetilde{C}_j^{(m)} = (\widetilde{c}_{i,r}^{(j)})$, $i, r \in \{1, 2, \ldots, m\}$, for $1 \le j \le s$, where

$$\widetilde{c}_{i,r}^{(j)} = \begin{cases} c_{i,r}^{(j)} & \text{if } r \in \{1, \ldots, m - \min(m, w_j)\},\\ 0 & \text{if } r \in \{m - \min(m, w_j) + 1, \ldots, m\}. \end{cases} \tag{5}$$

That is, the first $m - \min(m, w_j)$ columns of $\widetilde{C}_j^{(m)}$ are the same as the columns of the

matrix $C_j^{(m)}$, and we set the last $\min(m, w_j)$ columns to zero, i.e, if $w_j < m$,

$$\widetilde{C}_j^{(m)} = \begin{pmatrix} c_{1,1}^{(j)} & \dots & c_{1,(m-w_j)}^{(j)} & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots & \vdots & \\ c_{(m-w_j),1}^{(j)} & \dots & c_{(m-w_j),(m-w_j)}^{(j)} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ c_{m,1}^{(j)} & \dots & c_{m,(m-w_j)}^{(j)} & 0 & \dots & 0 \end{pmatrix}.$$

We are interested in estimating the quality parameter of the digital net generated by the $\widetilde{C}_j^{(m)}$.

Apart from the main motivation outlined in Section 1, there is another computational advantage of using column reduced digital nets. Indeed, by the general construction principle of digital point sets, the generating matrices of a digital net or sequence are multiplied over $\mathbb{F}_b$ by vectors representing the digits of the indices of the elements of the point set. By replacing the matrices $C_j^{(m)}$ by $\widetilde{C}_j^{(m)}$, we increase the sparsity of the generating matrices, which saves computation time in the generation of the point set.

Theorem 1. *Let $\mathcal{P}$ be a digital (t, m, s)-net over $\mathbb{F}_b$ with generating matrices $C_1^{(m)}, \dots, C_s^{(m)}$ derived from a digital (t, s)-sequence over $\mathbb{F}_b$, where we assume that $m \geq t$. Let $\widetilde{C}_1^{(m)}, \dots, \widetilde{C}_s^{(m)}$ be as defined in* (5) *with respect to reduction indices $0 = w_1 \leq \dots \leq w_s$ and let $\widetilde{t}$ be the minimal quality parameter of the net generated by the $\widetilde{C}_j^{(m)}$. Then,*

$$\max\{0, m - w_s - t\} \leq \rho_m\left(\widetilde{C}_1^{(m)}, \dots, \widetilde{C}_s^{(m)}\right) \leq \max\{0, m - w_s\}, \tag{6}$$

and $\widetilde{t} \leq \min\{m, w_s + t\}$.

Furthermore, if $\mathcal{P}$ is a strict digital (t, m, s)-net, it is true that

$$\rho_m\left(\widetilde{C}_1^{(m)}, \dots, \widetilde{C}_s^{(m)}\right) \leq \max\{0, m - \max\{t, w_s\}\}. \tag{7}$$

Proof. We note that we have $m \geq t$ by assumption. If $w_s \geq m$, then we trivially have $\rho_m\left(\widetilde{C}_1^{(m)}, \dots, \widetilde{C}_s^{(m)}\right) = 0$, as $\widetilde{C}_s^{(m)}$ only contains zeros, and (6) holds.

Therefore, we will assume for the rest of the proof that $w_s < m$.

We prove the second inequality in (6) first. We have

$$\widetilde{C}_s^{(m)} = \begin{pmatrix} c_{1,1}^{(s)} & \dots & c_{1,(m-w_s)}^{(s)} & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots & & \vdots \\ c_{(m-w_s),1}^{(s)} & \dots & c_{(m-w_s),(m-w_s)}^{(s)} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ c_{m,1}^{(s)} & \dots & c_{m,(m-w_s)}^{(s)} & 0 & \dots & 0 \end{pmatrix}.$$

Let D be the matrix containing the first d_1 rows of $\widetilde{C}_1^{(m)}$, the first d_2 rows of $\widetilde{C}_2^{(m)}$, etc., up to the first d_s rows of $\widetilde{C}_s^{(m)}$, where $d_1, \dots, d_s$ are nonnegative integers satisfying $d_1 + \dots + d_s = m - w_s$.

For the special choice $(d_1, \ldots, d_s) = (0, \ldots, 0, m-w_s)$, we have $\operatorname{rank}(D) = \operatorname{rank}(\widetilde{C}_s^{(m)}) \le m-w_s$. Therefore,

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) \le m - w_s.$$

Now we prove the first inequality in (6). If $m - w_s - t < 0$, the inequality is trivial. Otherwise, i.e., if $m - w_s \ge t$, we know that

$$\rho_k\left(C_1^{(k)}, \ldots, C_s^{(k)}\right) \ge k - t, \tag{8}$$

for any $k \ge t$, since our net is derived from a digital (t, s)-sequence. Here, $C_j^{(k)}$, $1 \le j \le s$, denotes the left upper $k \times k$ submatrix of C_j. In particular, we observe that for the left upper $(m - w_s) \times (m - w_s)$ submatrices of $C_1, \ldots, C_s$,

$$\rho_{(m-w_s)}(C_1^{(m-w_s)}, \ldots, C_s^{(m-w_s)}) \ge m - w_s - t.$$

We now consider arbitrary integers $d_1, \ldots, d_s \ge 0$ with $d_1 + \cdots + d_s = m - w_s - t$. Let $\boldsymbol{k}_i^{(j)}$ denote the i-th row vector of $C_j^{(m-w_s)} \in \mathbb{F}_b^{(m-w_s)\times(m-w_s)}$. We know that

$$\boldsymbol{k}_1^{(1)}, \ldots, \boldsymbol{k}_{d_1}^{(1)}, \boldsymbol{k}_1^{(2)}, \ldots, \boldsymbol{k}_{d_2}^{(2)}, \ldots, \ldots, \boldsymbol{k}_1^{(s)}, \ldots, \boldsymbol{k}_{d_s}^{(s)} \tag{9}$$

are linearly independent over $\mathbb{F}_b$. Let $\boldsymbol{c}_i^{(j)}$ denote the i-th row vector of $\widetilde{C}_i^{(m)} \in \mathbb{F}_b^{m\times m}$. We observe that for $1 \le i \le m - w_s$,

$$\boldsymbol{c}_i^{(j)} = (\boldsymbol{k}_i^{(j)}, \boldsymbol{u}_i^{(j)}) \in \mathbb{F}_b^{1\times m},$$

where the $\boldsymbol{k}_i^{(j)}$ are as above and $\boldsymbol{u}_i^{(j)} \in \mathbb{F}_b^{1\times w_s}$.

The row vectors

$$\boldsymbol{c}_1^{(1)}, \ldots, \boldsymbol{c}_{d_1}^{(1)}, \boldsymbol{c}_1^{(2)}, \ldots, \boldsymbol{c}_{d_2}^{(2)}, \ldots, \ldots, \boldsymbol{c}_1^{(s)}, \ldots, \boldsymbol{c}_{d_s}^{(s)} \tag{10}$$

are linearly independent, since otherwise the row vectors in (9), which are projections of $\boldsymbol{c}_i^{(j)}$ onto the first $m - w_s$ entries, would be linearly dependent. Therefore,

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) \ge m - w_s - t.$$

This concludes the proof of (6). Using (4) and the lower bound in (6), we obtain the upper bound for $\widetilde{t}$.

It remains to show (7).

Let D be the matrix containing the first d_1 rows of $\widetilde{C}_1^{(m)}$, the first d_2 rows of $\widetilde{C}_2^{(m)}$, etc., up to the first d_s rows of $\widetilde{C}_s^{(m)}$, where $d_1, \ldots, d_s$ are nonnegative integers. As above, for the special choice $(d_1, \ldots, d_s) = (0, \ldots, 0, m - w_s)$, we have $\operatorname{rank}(D) = \operatorname{rank}(\widetilde{C}_s^{(m)}) \le (m - w_s)$. So,

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) \le m - w_s.$$

However, since we assume that $\mathcal{P}$ is a strict digital (t, m, s)-net in this part of the proof, there must exist a choice of $(d_1, \ldots, d_s)$ with $d_1 + \cdots + d_s = m - t + 1$ such that the corresponding

rows of $C_1^{(m)}, \ldots, C_s^{(m)}$ are linearly dependent, and therefore also the corresponding rows of $\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}$ are linearly dependent. This yields

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) \leq m - t,$$

so we must have

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) \leq m - \max\{t, w_s\}.$$

□

Remark 1. For $t = 0$ and $w_s < m$ in Theorem 1, we obtain equality in (6) and therefore

$$\rho_m\left(\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}\right) = m - w_s,$$

and $\widetilde{t} = w_s$.

Remark 2. We now give an example that illustrates that the lower bound in Theorem 1 is sharp.

Assume that Q is a digital $(0, 2)$-sequence with generating matrices D_1 and D_2 (examples of Q exist, e.g., by choosing as Q a Niederreiter sequence, see [8]).

From Q, we construct a digital $(t, 2)$-sequence P, by prepending exactly t zero columns to both D_1 and D_2. That is, we construct new generating matrices C_j, $j \in \{1, 2\}$, such that

$$C_j := \left(\begin{array}{ccc|c} 0 & \ldots & 0 & \\ \vdots & \vdots & \vdots & \quad D_j \quad \\ 0 & \ldots & 0 & \\ \vdots & \vdots & \vdots & \end{array}\right).$$

It is easily checked that C_1, C_2 generate a digital $(t, 2)$-sequence; indeed, let $m \geq t$ be arbitrarily chosen but fixed. Then the matrices $C_1^{(m)}, C_2^{(m)}$ contain the matrices $D_1^{(m-t)}, D_2^{(m-t)}$ as submatrices. As D_1, D_2 generate a $(0, 2)$-sequence, for any $d_1, d_2 \in \mathbb{N}_0$ with $d_1 + d_2 = m - t$ the first d_1 rows of $D_1^{(m-t)}$ together with the first d_2 rows of $D_2^{(m-t)}$ must be linearly independent, so also the corresponding rows of $C_1^{(m)}$ and $C_2^{(m)}$ (with zeros prepended) must be linearly independent. This establishes that C_1 and C_2 generate a $(t, 2)$-sequence.

Let now $m \geq t$, and let $w_1 = 0$, and $w_2 \geq w_1$ be reduction indices such that $m - w_2 - t \geq 0$. Then $\widetilde{C}_1^{(m)} = C_1^{(m)}$, and

$$\widetilde{C}_2^{(m)} = \left(\begin{array}{ccc|c|ccc} 0 & \ldots & 0 & & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & D_2^{(m\times(m-t-w_2))} & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & & 0 & \ldots & 0 \end{array}\right),$$

where $D_2^{(m\times(m-t-w_2))}$ denotes the left upper $m \times (m - t - w_2)$ submatrix of D_2. By Theorem 1, we know that $\rho_m\left(\widetilde{C}_1^{(m)}, \widetilde{C}_2^{(m)}\right) \geq m - t - w_2$. However, $\rho_m\left(\widetilde{C}_1^{(m)}, \widetilde{C}_2^{(m)}\right) > m - t - w_2$ cannot hold since the first $m - t - w_2 + 1$ rows of $\widetilde{C}_2^{(m)}$ must be linearly dependent.

This implies that the lower bound in Theorem 1 is sharp.

Remark 3. Next, we provide an example showing that the upper bound (7) for strict digital nets in Theorem 1 is sharp.

We use the same notation as in Remark 2. We again start with the digital $(0,2)$-sequence Q. Again, we transform Q into a (t,s)-sequence, now called R, with generating matrices E_1 and E_2. For E_1, we take the first generating matrix of P from above, i.e., $E_1 = C_1$. Furthermore, we choose E_2 as

$$E_2 := \left(\begin{array}{ccc|ccc} & & & 0 & \dots & 0 & \dots \\ & D_2^{(t)} & & \vdots & \vdots & \vdots & \vdots \\ & & & 0 & \dots & 0 & \dots \\ \hline 0 & \dots & 0 & & & & \\ \vdots & \vdots & \vdots & & D_2 & & \\ 0 & \dots & 0 & & & & \\ \vdots & \vdots & \vdots & & & & \end{array}\right),$$

where $D_2^{(t)}$ is the left upper $t \times t$ submatrix of D_2. First, note that R really is a strict $(t,2)$-sequence. Indeed, if we consider the matrix $E_1^{(m)}$ for $m < t$, this matrix only contains zeros, so the quality parameter of R must be at least t. On the other hand, let $m \geq t$ and consider the matrices $E_1^{(m)}$ and $E_2^{(m)}$. Choose $d_1, d_2 \geq 0$ such that $d_1 + d_2 = m - t$, and consider the first d_1 rows of $E_1^{(m)}$ together with the first d_2 rows of $E_2^{(m)}$. We distinguish two cases.

- If $d_2 \leq t$, then it is obvious that the first d_1 rows of $E_1^{(m)}$ together with the first d_2 rows of $E_2^{(m)}$ are linearly independent, as D_1 and D_2 generate a $(0,2)$-sequence.
- If $d_2 > t$, we proceed as follows. Assume to the contrary that the first d_1 rows of $E_1^{(m)}$ together with the first d_2 rows of $E_2^{(m)}$ were not linearly independent. By the structure of E_1 and E_2, this would immediately imply that the first d_1 rows of $D_1^{(m-t)}$ together with the first $d_2 - t$ rows of $D_2^{(m-t)}$ are not linearly independent, where $d_1 + d_2 - t = m - 2t$, which would be a contradiction to the property that D_1 and D_2 generate a digital $(0,2)$-sequence.

Let now again $m \geq t$, and let $w_1 = 0$, and $w_2 \geq w_1$ be reduction indices such that $m - w_2 - t \geq 0$. Then $\widetilde{E}_1^{(m)} := E_1^{(m)}$, and

$$\widetilde{E}_2^{(m)} := \left(\begin{array}{ccc|ccc|ccc} & & & 0 & \dots & 0 & 0 & \dots & 0 \\ & D_2^{(t)} & & \vdots & & \vdots & \vdots & \vdots & \vdots \\ & & & 0 & \dots & 0 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & & & & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & D_2^{((m-t)\times(m-t-w_2))} & & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & & & & 0 & \dots & 0 \end{array}\right).$$

We again distinguish two cases.

Case 1: $\max\{t, w_2\} = w_2$. We claim that $\rho_m\left(\widetilde{E}_1^{(m)}, \widetilde{E}_2^{(m)}\right) = m - w_2$. To this end, let $d_1, d_2 \geq 0$ such that $d_1 + d_2 = m - w_2$, which implies that d_1 and d_2 are both not larger than $m - t$. Then, we consider two sub-cases.

- If $d_2 \leq t$, it is clear because of the structure of the matrices that the first d_1 rows of $\widetilde{E}_1^{(m)}$ together with the first d_2 rows of $\widetilde{E}_2^{(m)}$ are linearly independent, as D_1 and D_2 generate a $(0,2)$-sequence. This is guaranteed since we know that d_1 and d_2 are both not larger than $m-t$.

- If $d_2 > t$, we proceed as follows. Assume to the contrary that the first d_1 rows of $\widetilde{E}_1^{(m)}$ together with the first d_2 rows of $\widetilde{E}_2^{(m)}$ were not linearly independent.

  By the structure of $\widetilde{E}_1^{(m)}$ and $\widetilde{E}_2^{(m)}$, this would immediately imply that the first d_1 rows of $D_1^{(m-t)}$ together with the first $d_2 - t$ rows of $D_2^{((m-t)\times(m-t-w_2))}$ are not linearly independent, where $d_1 + d_2 - t = m - t - w_2$. Note, however, that $D_1^{(m-t)}$ contains $D_1^{(m-t-w_2)}$ as its left upper submatrix, and also $D_2^{((m-t)\times(m-t-w_2))}$ contains $D_2^{(m-t-w_2)}$ as its left upper submatrix. By the property that D_1 and D_2 generate a $(0,2)$-sequence, and by the assumption that $m - w_2 \geq t$, the first d_1 rows of $D_1^{(m-t-w_2)}$ together with the first $d_2 - t$ rows of $D_2^{(m-t-w_2)}$ must be linearly independent. The same must, however, then also hold for the corresponding rows of $D_1^{(m-t)}$ and $D_2^{((m-t)\times(m-t-w_2))}$, which yields a contradiction.

Hence we have shown that $\rho_m\left(\widetilde{E}_1^{(m)}, \widetilde{E}_2^{(m)}\right) \geq m - w_2$, and by Theorem 1 we must actually have $\rho_m\left(\widetilde{E}_1^{(m)}, \widetilde{E}_2^{(m)}\right) = m - w_2$.

Case 2: $\max\{t, w_2\} = t$. We claim that $\rho_m\left(\widetilde{E}_1^{(m)}, \widetilde{E}_2^{(m)}\right) = m - t$. To this end, let $d_1, d_2 \geq 0$ such that $d_1 + d_2 = m - t$. Also here, we distinguish two sub-cases.

- If $d_2 \leq t$, it is obvious that the first d_1 rows of $\widetilde{E}_1^{(m)}$ together with the first d_2 rows of $\widetilde{E}_2^{(m)}$ are linearly independent, as D_1 and D_2 generate a $(0,2)$-sequence. This is guaranteed since we know that d_1 and d_2 are both not larger than $m-t$.

- If $d_2 > t$, we proceed as follows. Assume to the contrary that the first d_1 rows of $\widetilde{E}_1^{(m)}$ together with the first d_2 rows of $\widetilde{E}_2^{(m)}$ were not linearly independent.

  By the structure of $\widetilde{E}_1^{(m)}$ and $\widetilde{E}_2^{(m)}$, this would immediately imply that the first d_1 rows of $D_1^{(m-t)}$ together with the first $d_2 - t$ rows of $D_2^{((m-t)\times(m-t-w_2))}$ are not linearly independent, where $d_1 + d_2 - t = m - 2t \leq m - t - w_2$. Note, however, that $D_1^{(m-t)}$ contains $D_1^{(m-t-w_2)}$ as its left upper submatrix, and also $D_2^{((m-t)\times(m-t-w_2))}$ contains $D_2^{(m-t-w_2)}$ as its left upper submatrix. By the property that D_1 and D_2 generate a $(0,2)$-sequence, by the fact that $d_1 + d_2 \leq m - t - w_2$, and by the assumption that $m - w_2 \geq t$, the first d_1 rows of $D_1^{(m-t-w_2)}$ together with the first d_2 rows of $D_2^{(m-t-w_2)}$ must be linearly independent. The same must, however, then also hold for the corresponding rows of $D_1^{(m-t)}$ and $D_2^{((m-t)\times(m-t-w_2))}$, which yields a contradiction.

In summary, we have shown that (7) is sharp for strict digital nets.

3 Projections of column reduced digital nets

Due to the important role of the t-value, one sometimes also considers a slightly refined notion of a (t, m, s)-net, which is then referred to as a $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, m, s)$-net, where $[s] := \{1, \ldots, s\}$. The latter notion means that for any $\mathfrak{u} \neq \emptyset$, $\mathfrak{u} \subseteq [s]$, the projection of the net onto those components with indices in $\mathfrak{u}$ is a $(t_{\mathfrak{u}}, m, |\mathfrak{u}|)$-net. The notion of a $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, s)$-sequence is defined analogously. Moreover, for $\mathfrak{u} \neq \emptyset$, we write $\overline{\mathfrak{u}} := \max(\mathfrak{u})$.

If we assume (which we always do in this paper) that the reduction indices satisfy $0 = w_1 \leq w_2 \leq \cdots \leq w_s$, then, for any non-empty $\mathfrak{u} \subseteq [s]$, the reduction index $w_{\overline{\mathfrak{u}}}$ is the largest among all reduction indices corresponding to $\mathfrak{u}$. This yields the following adaption of Theorem 1, which obviously can be shown in the same manner.

Corollary 1. *Let $\mathcal{P}$ be a digital $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, m, s)$-net over $\mathbb{F}_b$ with generating matrices $C_1^{(m)}, \ldots, C_s^{(m)}$, which has been derived from a digital $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, s)$-sequence, where we assume that $m \geq t$. Let $\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}$ be the reduced generating matrices with respect to reduction indices $0 = w_1 \leq \cdots \leq w_s$ and let $(\widetilde{t}_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}$ be the minimal quality parameters of the projections of the net generated by the $\widetilde{C}_j^{(m)}$. Then, for every non-empty $\mathfrak{u} \subseteq [s]$,*

$$\max\{0, m - w_{\overline{\mathfrak{u}}} - t_{\mathfrak{u}}\} \leq \rho_m((\widetilde{C}_j^{(m)})_{j\in\mathfrak{u}}) \leq \max\{0, m - w_{\overline{\mathfrak{u}}}\},$$

and $\widetilde{t}_{\mathfrak{u}} \leq \min\{m, w_{\overline{\mathfrak{u}}} + t_{\mathfrak{u}}\}$.

Furthermore, if, for a non-empty $\mathfrak{u} \subseteq [s]$, the projection of $\mathcal{P}$ onto the components in $\mathfrak{u}$ is a strict digital $(t_{\mathfrak{u}}, m, |\mathfrak{u}|)$-net, it is true that

$$\rho_m((\widetilde{C}_j^{(m)})_{j\in\mathfrak{u}}) \leq \max\{0, m - \max\{t_{\mathfrak{u}}, w_{\overline{\mathfrak{u}}}\}\}.$$

4 Applications of column reduced digital nets

4.1 A reduced matrix product algorithm

In this section, we return to the problem outlined in Section 1. Let P be a digital (t, m, s)-net over $\mathbb{F}_b$, with generating matrices $C_1^{(m)}, \ldots, C_s^{(m)}$. Let $\boldsymbol{w} = (w_j)_{j=1}^s \in \mathbb{N}_0^s$ be a sequence of reduction indices with $0 = w_1 \leq w_2 \leq \cdots \leq w_s$. Let $s^* \leq s$ be the largest index such that $w_{s^*} < m$. Let $\widetilde{C}_1^{(m)}, \ldots, \widetilde{C}_s^{(m)}$ be the reduced generating matrices corresponding to $w_1, \ldots, w_s$, and let Q be the corresponding reduced digital net. Let $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}$ be the points of Q, where we interpret $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}$ as column vectors. Let

$$X = [\boldsymbol{x}_0^\top, \boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_{N-1}^\top]^\top$$

be the $N \times s$ matrix whose k-th row is the k-th point of Q for $0 \leq k \leq N-1$.

Let $\boldsymbol{\xi}_j$ denote the j-th column of X, i.e., $X = [\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_s]$. Let $A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s]^\top$, where $\boldsymbol{a}_j \in \mathbb{R}^{1\times\tau}$ is the j-th row of A. Then we have

$$XA = [\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_s] \cdot [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s]^\top = \boldsymbol{\xi}_1\boldsymbol{a}_1 + \boldsymbol{\xi}_2\boldsymbol{a}_2 + \cdots + \boldsymbol{\xi}_s\boldsymbol{a}_s. \tag{11}$$

We will make use of a certain inherent repetitiveness of the reduced net Q, which we will illustrate by considering a reduction index $0 \leq w_j < m$ for $1 \leq j \leq s^*$, and the corresponding

generator matrix $\widetilde{C}_j^{(m)}$. The j-th components of the $N = b^m$ points of Q (i.e., the j-th column $\boldsymbol{\xi}_j$ of X) are then given by

$$\begin{aligned}\boldsymbol{\xi}_j &= \left(\left(\widetilde{C}_j^{(m)}\overrightarrow{0}\right)\cdot(b^{-1},\ldots,b^{-m}),\ldots,\left(\widetilde{C}_j^{(m)}\overrightarrow{(b^m-1)}\right)\cdot(b^{-1},\ldots,b^{-m})\right)^\top \\ &= (\underbrace{X_j,\ldots,X_j}_{b^{w_j}\text{ times}})^\top,\end{aligned}$$

where, as above, we write $\overrightarrow{k}$ to denote the vector of base b digits of length m for $k \in \{0,1\ldots,b^m-1\}$, and where

$$X_j = \left(\left(\widetilde{C}_j^{(m)}\overrightarrow{0}\right)\cdot(b^{-1},\ldots,b^{-m}),\ldots,\left(\widetilde{C}_j^{(m)}\overrightarrow{(b^{m-w_j}-1)}\right)\cdot(b^{-1},\ldots,b^{-m})\right)^\top.$$

The reason for this repetitive structure is that, for any w_j with $0 < w_j < m$, the last w_j columns of $\widetilde{C}_j^{(m)}$ are equal to zero, and thus, in the product $\widetilde{C}_j^{(m)}\overrightarrow{k}$, the last w_j entries of $\overrightarrow{k}$ become irrelevant. We will exploit this structure within Q to derive a fast matrix-matrix multiplication algorithm to compute XA.

Based on the above observations, it is possible to formulate the following algorithm to compute (11) in an efficient way. Note that for $j > s^*$ the j-th column of X consists only of zeros, so there is nothing to compute for the entries of X corresponding to these columns.

Algorithm 1 Fast reduced matrix-matrix product using column reduced digital nets

Input:
Matrix $A \in \mathbb{R}^{s\times\tau}$, integer $m \in \mathbb{N}$, prime b, reduction indices $0 = w_1 \le w_2 \le \cdots \le w_s$, corresponding generating matrices $\widetilde{C}_1^{(m)}, \dots, \widetilde{C}_s^{(m)}$ of a reduced digital net.

Set $N = b^m$ and set $P_{s^*+1} = \mathbf{0}_{1\times\tau} \in \mathbb{R}^{1\times\tau}$.
for $j = s^*$ **to** 1 **do**

- Compute X_j as

$$X_j = \left(\left(\widetilde{C}_j^{(m)}\overrightarrow{0}\right)\cdot(b^{-1},\dots,b^{-m}),\dots,\left(\widetilde{C}_j^{(m)}\overrightarrow{(b^{m-w_j}-1)}\right)\cdot(b^{-1},\dots,b^{-m})\right)^\top \in \mathbb{R}^{b^{m-w_j}\times 1}.$$

- Compute P_j as

$$P_j = \overbrace{\begin{Bmatrix} \begin{pmatrix} P_{j+1} \\ P_{j+1} \\ \vdots \\ P_{j+1} \end{pmatrix} \end{Bmatrix}}^{b^{\min(w_{j+1},m)-w_j} \text{ times}} + X_j \boldsymbol{a}_j \in \mathbb{R}^{b^{m-w_j}\times\tau},$$

  where $\boldsymbol{a}_j \in \mathbb{R}^{1\times\tau}$ denotes the j-th row of the matrix A.

end for
Set $P = P_1$.

Return: Matrix product $P = XA$.

Remark 4. The number of computations needed for Algorithm 1 is of order

$$\mathcal{O}\left(\sum_{j=1}^{s^*} b^{m-w_j}(\tau + m(m-w_j))\right).$$

Note that this algorithm also generates the points of the reduced digital net, whereas the standard multiplication or the analogous "row reduced algorithm" [1, Algorithm 4], both require pre-computed points of the digital net as input. Generating the points of a non-reduced digital net requires $\mathcal{O}(b^m s m^2)$ operations, see also [1, Algorithm 3] and the standard non-reduced matrix-matrix multiplication usually requires $\mathcal{O}(b^m s\tau)$ operations. Therefore, Algorithm 1 improves the runtime of both steps. We also point out that the number of operations necessary for Algorithm 1 is independent of s, and only depends on s^*. If the reduction indices w_j grow sufficiently fast, then s^* can be significantly lower than s.

4.2 Error analysis

In the beginning of the paper we set out the task of approximating the integral (1) by the QMC rule (2). We have shown in the previous sections how to speed up the computation of

the products $\boldsymbol{x}_k^\top A$ if we choose $\boldsymbol{x}_k$ as the points of a column reduced digital net. However, we should also keep in mind the integration error made by using a QMC rule of the form (2) using those $\boldsymbol{x}_k$.

In many applications of quasi-Monte Carlo, one considers so-called *weighted function spaces* such as weighted Sobolev or weighted Korobov spaces (see, e.g., [3, 5, 6]). The idea of studying weighted function spaces goes back to the seminal paper [11] of Sloan and Woźniakowski. The motivation for weighted spaces is that in many applications, different coordinates or different groups of coordinates may have different influence on a multivariate problem. To give a simple example, consider numerical integration of a function $f : [0,1]^s \to \mathbb{R}$, where

$$f(x_1, \dots, x_s) = \mathrm{e}^{x_1} + \frac{x_2 + \cdots + x_s}{2^s}.$$

Clearly, for large s, the first variable has much more influence on this problem than the others. In order to make such observations more precise, one introduces weights, which are nonnegative real numbers $\gamma_{\mathfrak{u}}$, one for each set $\mathfrak{u} \subseteq \{1, \dots, s\}$. Intuitively speaking, the number $\gamma_{\mathfrak{u}}$ models the influence of the variables with indices in $\mathfrak{u}$. Larger values of $\gamma_{\mathfrak{u}}$ mean more influence, smaller values less influence. Formally, we set $\gamma_\emptyset = 1$, and we write $\boldsymbol{\gamma} = \{\gamma_{\mathfrak{u}}\}_{\mathfrak{u} \subseteq \{1,\dots,s\}}$. These weights can now be used to modify the norm in a given function space, thereby modifying the set over which a suitable error measure, as for example the *worst-case error*, of a problem is considered. By making this set smaller according to the weights (in the sense that also here, certain groups of variables may have less influence than others), a problem may thus become easier to handle and even lose the curse of dimensionality, provided that suitable conditions on the weights hold. This effect also corresponds to intuition—if a problem depends on many variables, of which only some have significant influence, it is natural to expect that the problem will be easier to solve than one where all variables have the same influence.

The *weighted star discrepancy* is (via the well-known *Koksma-Hlawka inequality* or its weighted version, see, e.g., [3, 6, 9]), a measure of the worst-case quadrature error for a QMC rule with node set Q, with b^m nodes, defined as

$$D^*_{b^m,\boldsymbol{\gamma}}(Q) := \sup_{x \in (0,1]^s} \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}} \left| \Delta_{Q,\mathfrak{u}}(\boldsymbol{x}) \right|, \tag{12}$$

where

$$\Delta_{Q,\mathfrak{u}}(\boldsymbol{x}) := \frac{\#\{(y_1, \dots, y_s) \in Q \colon y_j < x_j,\, \forall j \in \mathfrak{u}\}}{b^m} - \prod_{j \in \mathfrak{u}} x_j. \tag{13}$$

Indeed, for certain weighted function classes based on Sobolev spaces of smoothness one, the weighted star discrepancy equals the worst-case quadrature error of a QMC rule with node set Q. Here, by the worst-case error, we mean the supremum of the integration error taken over the unit ball of the function class under consideration. We refer to [3, Section 5.3] for further details on the weighted Koksma-Hlawka inequality.

As shown in [10], we have

$$\begin{aligned} D^*_{b^m,\boldsymbol{\gamma}}(Q) &= \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \sup_{\boldsymbol{x} \in (0,1]^s} \gamma_{\mathfrak{u}} \left| \Delta_{Q,\mathfrak{u}}(\boldsymbol{x}) \right| \\ &= \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \gamma_{\mathfrak{u}} \sup_{\boldsymbol{x} \in (0,1]^s} \left| \Delta_{Q,\mathfrak{u}}(\boldsymbol{x}) \right|. \end{aligned}$$

In the latter expression, the suprema over $\boldsymbol{x} \in (0,1]^s$ just yield the values of the star discrepancy of the projections of Q, and thus, one can use existing discrepancy bounds for the projections

of Q. Let us proceed as follows. Assume that $\mathcal{P}$ is a digital $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, m, s)$-net over $\mathbb{F}_b$ with $m \times m$ generating matrices $C_1^{(m)}, \dots, C_s^{(m)}$ derived from a digital $((t_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}, s)$-sequence, where $m \geq t$. Let $\widetilde{\mathcal{P}}$ be the corresponding column reduced digital net based on the reduction indices $0 = w_1 \leq w_2 \leq \cdots \leq w_s$, and let $(\widetilde{t}_{\mathfrak{u}})_{\mathfrak{u}\subseteq[s]}$ be the minimal quality parameters of the projections of $\widetilde{\mathcal{P}}$.

Whenever we consider a $\mathfrak{u} \subseteq [s]$ that is not a subset of $[s^*]$, we know due to Corollary 1 that the quality parameter of the corresponding projection of $\widetilde{P}$ is m and therefore we can bound its discrepancy only trivially by 1. Whenever we have $\mathfrak{u} \subseteq [s^*]$, however, we can use existing discrepancy bounds for the corresponding net. To this end, we use the results from [7], which are, to our best knowledge, the currently best-known general upper discrepancy bounds for (t, m, s)-nets. This yields, for any non-empty set $\mathfrak{u} \subseteq [s]$,

$$\sup_{\boldsymbol{x}\in(0,1]^s} \left|\Delta_{\widetilde{\mathcal{P}},\mathfrak{u}}(\boldsymbol{x})\right| \leq \begin{cases} 1 & \text{if } \mathfrak{u} \not\subseteq [s^*], \\ (b^{\widetilde{t}_{\mathfrak{u}}}/b^m) \sum_{v=0}^{|\mathfrak{u}|-1} a_{v,b}^{(|\mathfrak{u}|)} m^v & \text{if } \mathfrak{u} \subseteq [s^*] \text{ and } |\mathfrak{u}| \geq 2, \\ b^{\widetilde{t}_{\mathfrak{u}}}/b^m & \text{if } \mathfrak{u} \subseteq [s^*] \text{ and } |\mathfrak{u}| = 1, \end{cases} \tag{14}$$

where

$$\begin{aligned} a_{v,b}^{(|\mathfrak{u}|)} &= \binom{|\mathfrak{u}|-2}{v} \left(\frac{b+2}{2}\right)^{|\mathfrak{u}|-2-v} \frac{(b-1)^v}{2^v v!} (a_{0,b}^{(2)} + |\mathfrak{u}|^2 - 4) \\ &\quad + \binom{|\mathfrak{u}|-2}{v-1} \left(\frac{b+2}{2}\right)^{|\mathfrak{u}|-1-v} \frac{(b-1)^{v-1}}{2^{v-1} v!} a_{1,b}^{(2)}, \end{aligned}$$

for $0 \leq v \leq |\mathfrak{u}| - 1$, with

$$a_{0,b}^{(2)} = \begin{cases} \frac{b+8}{4} & \text{if } b \text{ is even}, \\ \frac{b+4}{2} & \text{if } b \text{ is odd}, \end{cases} \qquad \text{and} \qquad a_{1,b}^{(2)} = \begin{cases} \frac{b^2}{4(b+1)} & \text{if } b \text{ is even}, \\ \frac{b-1}{4} & \text{if } b \text{ is odd}. \end{cases}$$

This then yields

$$D^*_{b^m,\boldsymbol{\gamma}}(\widetilde{\mathcal{P}}) \leq \max \left\{ \max_{\substack{\emptyset\neq\mathfrak{u}\subseteq[s] \\ \mathfrak{u}\not\subseteq[s^*]}} \gamma_{\mathfrak{u}}, \max_{\substack{\mathfrak{u}\subseteq[s^*] \\ |\mathfrak{u}|=1}} \gamma_{\mathfrak{u}} \frac{b^{\widetilde{t}_{\mathfrak{u}}}}{b^m}, \max_{\substack{\mathfrak{u}\subseteq[s^*] \\ |\mathfrak{u}|\geq 2}} \gamma_{\mathfrak{u}} \frac{b^{\widetilde{t}_{\mathfrak{u}}}}{b^m} \sum_{v=0}^{|\mathfrak{u}|-1} a_{v,b}^{(|\mathfrak{u}|)} m^v \right\}. \tag{15}$$

Let us analyze the three maxima in the curly brackets in (15) in greater detail. To this end, as also in [1], we restrict ourselves to product weights in the following, i.e., we assume weights $\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j$ with $\gamma_1 \geq \gamma_2 \geq \cdots > 0$.

For the first term, we proceed as in [1], namely we use that $w_j \geq m$ if $j \in \mathfrak{u} \setminus [s^*]$, and obtain for $\mathfrak{v} = \mathfrak{u} \cap [s^*]$ that

$$\gamma_{\mathfrak{u}} \leq \gamma_{\mathfrak{v}} \gamma_{\mathfrak{u}\setminus\mathfrak{v}} \frac{1}{b^m} \prod_{j\in\mathfrak{u}\setminus\mathfrak{v}} (1 + b^{w_j}) \leq \frac{1}{b^m} \prod_{j\in\mathfrak{u}} \gamma_j (1 + b^{w_j}). \tag{16}$$

For the second maximum in (15), note that we have $\mathfrak{u} = \{j\}$ for some $j \in [s^*]$, and hence $\widetilde{t}_{\mathfrak{u}} \leq \min\{m, w_j + t_{\{j\}}\}$ by Corollary 1. Consequently,

$$\max_{\substack{\mathfrak{u}\subseteq[s^*] \\ |\mathfrak{u}|=1}} \gamma_{\mathfrak{u}} \frac{b^{\widetilde{t}_{\mathfrak{u}}}}{b^m} \leq \max_{j\in[s^*]} \gamma_j \frac{b^{\min\{m, w_j + t_{\{j\}}\}}}{b^m}. \tag{17}$$

For the third maximum in (15), we again use Corollary 1, and obtain

$$\max_{\substack{\mathfrak{u}\subseteq[s^*]\\|\mathfrak{u}|\geq 2}} \gamma_{\mathfrak{u}} \frac{b^{\widetilde{t}_{\mathfrak{u}}}}{b^m} \sum_{v=0}^{|\mathfrak{u}|-1} a_{v,b}^{(|\mathfrak{u}|)} m^v \leq \max_{\substack{\mathfrak{u}\subseteq[s^*]\\|\mathfrak{u}|\geq 2}} \gamma_{\mathfrak{u}} \frac{b^{\min\{m,w_{\overline{\mathfrak{u}}}+t_{\mathfrak{u}}\}}}{b^m} \sum_{v=0}^{|\mathfrak{u}|-1} a_{v,b}^{(|\mathfrak{u}|)} m^v. \tag{18}$$

Using these estimates in (15), we obtain

$$\begin{aligned} &D^*_{b^m,\boldsymbol{\gamma}}(\widetilde{\mathcal{P}}) \\ &\leq \quad \max\left\{ \max_{\substack{\emptyset\neq\mathfrak{u}\subseteq[s]\\ \mathfrak{u}\not\subseteq[s^*]}} \frac{1}{b^m}\prod_{j\in\mathfrak{u}}\gamma_j(1+b^{w_j}), \max_{j\in[s^*]}\gamma_j\frac{b^{w_j+t_{\{j\}}}}{b^m}, \max_{\substack{\mathfrak{u}\subseteq[s^*]\\|\mathfrak{u}|\geq 2}}\gamma_{\mathfrak{u}}\frac{b^{\min\{m,w_{\overline{\mathfrak{u}}}+t_{\mathfrak{u}}\}}}{b^m}\sum_{v=0}^{|\mathfrak{u}|-1}a_{v,b}^{(|\mathfrak{u}|)}m^v \right\}. \end{aligned} \tag{19}$$

Remark 5. A few remarks on (19) are in order. Note that only the first term in the curly brackets in (19) depends on s. The two remaining terms depend on s^*, which can be independent of s if the reduction indices w_j increase sufficiently fast. However, let us give a few further details on these observations.

We may want that the first term

$$\frac{1}{b^m}\prod_{j\in\mathfrak{u}}\gamma_j(1+b^{w_j}) \leq \frac{1}{b^m}\prod_{j=1}^{s}\gamma_j(1+b^{w_j})$$

be bounded by κ/b^m for some constant $\kappa>0$ independent of s. Let $j_0\in\mathbb{N}$ be minimal such that $\gamma_j\leq 1$ for all $j>j_0$. Then we impose $\prod_{j=1}^{s}\gamma_j(1+b^{w_j})\leq\gamma_1^{j_0}\prod_{j=1}^{s}(1+\gamma_j b^{w_j})\leq\kappa$. Hence it is sufficient to choose $\kappa>\gamma_1^{j_0}$ and for all $j\in[s]$,

$$w_j := \min\left(\left\lfloor \log_b\left(\frac{\left(\frac{\kappa}{\gamma_1^{j_0}}\right)^{1/s}-1}{\gamma_j}\right)\right\rfloor, m\right). \tag{20}$$

The choice of the w_j in (20) depends on s. For sufficiently fast decaying weights γ_j, it is possible to choose the w_j such that they no longer depend on s. Indeed, suppose, e.g., that $\gamma_j=j^{-2}$. Then we could choose the w_j such that, for some $\tau\in(1,2)$,

$$w_j \leq \min\left(\left\lfloor \log_b\left(j^{2-\tau}\right)\right\rfloor, m\right). \tag{21}$$

This then yields

$$\prod_{j=1}^{s}(1+\gamma_j b^{w_j}) \leq \exp\left(\sum_{j=1}^{s}\log(1+\gamma_j b^{w_j})\right) \leq \exp\left(\sum_{j=1}^{s}\gamma_j b^{w_j}\right) \leq \exp(\zeta(\tau)),$$

where $\zeta(\cdot)$ is the Riemann zeta function. This gives a dimension-independent bound on the term $\prod_{j=1}^{s}\gamma_j(1+b^{w_j})$ from above, and hence a dimension-independent bound for all of $D^*_{b^m,\boldsymbol{\gamma}}(\widetilde{\mathcal{P}})$.

Regarding the second term in (19), this term only depends on one-dimensional projections of $\widetilde{\mathcal{P}}$. In particular, if we choose the w_j as in (21), this expression should be easy to bound

from above. This is even more so if the t-values of the one-dimensional projections of the non-reduced net $\mathcal{P}$ are low, which may often be the case (in fact, the t-values of one-dimensional projections might even be zero in many examples). Thus we can bound the second term by an expression of the form κ^*/b^m, which only depends on s^* but not on s.

Regarding the third term in (19), it crucially depends on the weights $\boldsymbol{\gamma}$ and their interplay with the quality parameters of the projections of $\mathcal{P}$, $t_{\mathfrak{u}}$. In particular, small quality parameters, in combination with sufficiently fast decaying weights and a suitable choice of the reduction indices w_j, should yield tighter error bounds. Indeed, we could proceed similarly to [7, Corollary 1], and bound the third term in (19) by a term of the form

$$\max_{\substack{\mathfrak{u}\subseteq[s^*]\\|\mathfrak{u}|\geq 2}} \gamma_{\mathfrak{u}} \frac{1}{b^m}\left(c_{|\mathfrak{u}|}m^{|\mathfrak{u}|-1}+\mathcal{O}(m^{|\mathfrak{u}|-2})\right),$$

where $c_{|\mathfrak{u}|}$ depends on $b, t_{\mathfrak{u}}$ and $|\mathfrak{u}|$, but not on m. Note that also the third term only depends on s^* and not on s, so for sufficiently fast increasing reduction indices w_j, the dimension s does not matter. In summary, we obtain

$$D^*_{b^m,\boldsymbol{\gamma}}(\widetilde{\mathcal{P}}) \leq \max\left\{\frac{\kappa}{b^m}, \frac{\kappa^*}{b^m}, \max_{\substack{\mathfrak{u}\subseteq[s^*]\\|\mathfrak{u}|\geq 2}} \gamma_{\mathfrak{u}} \frac{1}{b^m}\left(c_{|\mathfrak{u}|}m^{|\mathfrak{u}|-1}+\mathcal{O}(m^{|\mathfrak{u}|-2})\right)\right\}.$$

Remark 6. Note that our new result yields an advantage over the corresponding result for row reduced nets in [1]. In that paper, one needs to work with the quality parameters of the projections of the reduced net, which are, in general, not known. In the present paper, we benefit from the combination of the column reduction and the fact that the nets considered here are derived from digital sequences, which guarantees additional structure. Usually, it is computationally involved to determine the t-value of a digital net or sequence from the generating matrices, since many linear independence conditions need to be checked. Here, however, we can use Theorem 1 and Corollary 1, which relate the t-values of $\mathcal{P}$ to those of $\widetilde{\mathcal{P}}$, and thus give us an advantage. In particular, if $\mathcal{P}$ is obtained from, say, a Sobol' or a Niederreiter sequence, it should be possible to have t-values that are guaranteed to be reasonably low.

5 Numerical experiments

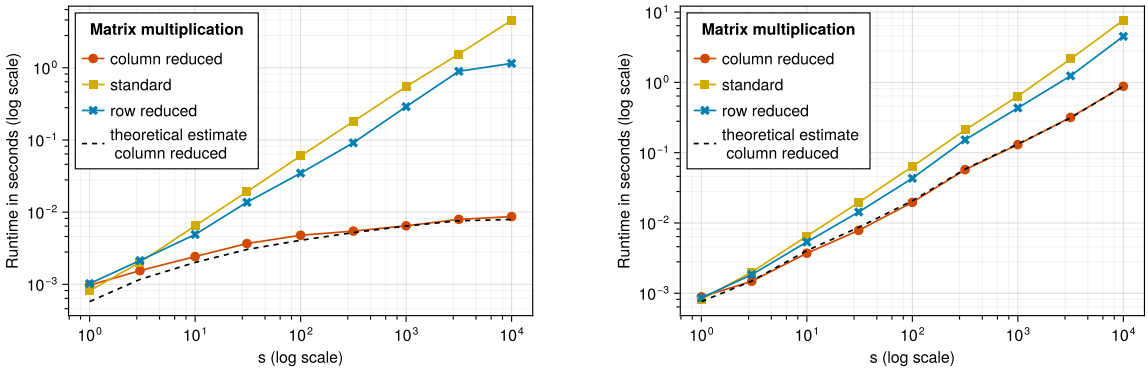
In this section, we test the computational performance of column reduced digital nets for matrix products XA, where A is an $s \times \tau$ matrix, as detailed in Section 4.1. We implemented Algorithm 1 in the Julia programming language (Version 1.9.3).[1] In the following plots, we compare the runtime of Algorithm 1 to the standard matrix multiplication and also the matrix multiplication using the points from row reduced digital nets as given in [1, Algorithm 4]. We remark that the reported runtimes are also affected by technical implementation details such as memory efficiency, a detailed discussion of which is out of scope here.

For the generating matrices $C_1^{(m)}, \ldots, C_s^{(m)}$, we used random matrices in $\mathbb{F}_b^{m\times m}$, since the matrix product computation itself does not depend on the entries of the matrix, i.e, we get similar relations of runtimes if we use generating matrices of specific digital sequences like Sobol' or Niederreiter sequences.

[1] Source code available at https://github.com/Vishnupriya-Anupindi/ReducedDigitalNets.jl

In Figure 1 we see, for fixed $b = 2, m = 12$, and $\tau = 20$, how the runtime changes as we vary s. We compare this for two different choices of reduction indices w_j. We see that in this case, using column reduced digital nets in Algorithm 1 performs better than the use of row reduced digital nets in [1, Algorithm 4] and also the standard matrix multiplication.

As the reduction indices w_j increase more slowly (as in Figure 1b), the difference in performance between the standard multiplication and Algorithm 1 reduces. We can see this also theoretically by inserting the weights in Remark 4.



(a) $w_j = \min(\lfloor \log_2(j) \rfloor, m)$ (b) $w_j = \min(\lfloor \log_2(j^{1/2}) \rfloor, m)$

Figure 1: $m = 12, \tau = 20$, varying w_j

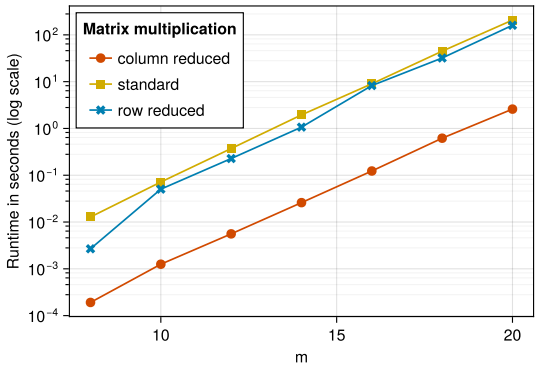


Figure 2: $s = 800, \tau = 20$, varying m

In Figure 2, we study the behavior for fixed $b = 2, s = 800$, and $\tau = 20$ as m increases. Note that we use the logarithmic scale for the time but not for m. We observe that also in this case Algorithm 1 seems to perform better than the row reduced case.

Overall, the numerical tests for the runtime using column reduced digital nets fit our theoretical estimate for the runtime as given in Remark 4 and comparison with the row reduced algorithm reveals that the column reduced algorithm could yield a better performance. Additionally to this practical advantage, column reduced matrices also have a theoretical advantage over row reduced matrices, as pointed out in Remark 6.

6 Conclusion

Column reduced digital nets have applications in the field of quasi-Monte Carlo methods. We can speed up the matrix-matrix multiplication in the quasi-Monte Carlo method by exploiting the repetitive structure of the points of a column reduced digital net. The bounds for the quality parameter (t-value) of column reduced digital nets have not been studied before.

In our research, we provide an algorithm for the matrix-matrix product using column reduced digital nets, which is faster than the standard matrix multiplication algorithm. In addition, we provide bounds for the t-value for column reduced digital nets. This is very essential for the error analysis of our method and has an advantage over the corresponding result for the row reduced nets in [1].

For future work, one could consider relaxing the conditions we impose on the t-value of the underlying digital sequence. One could also explore in-depth the interplay between column and row reduced digital nets.

Acknowledgments

The authors acknowledge the support of the Austrian Science Fund (FWF) Project P34808.

References

[1] J. Dick, A. Ebert, L. Herrmann, P. Kritzer, M. Longo. The fast reduced QMC matrix-vector product. J. Comput. Appl. Math. 440, 115642, 2024.

[2] J. Dick, P. Kritzer, G. Leobacher, F. Pillichshammer. A reduced fast component-by-component construction of lattice points for integration in weighted spaces with fast decreasing weights. J. Comput. Appl. Math. 276, 1–15, 2015.

[3] J. Dick, P. Kritzer, F. Pillichshammer. *Lattice Rules—Numerical Integration, Approximation, and Discrepancy*. Springer, Cham, 2022.

[4] J. Dick, F.Y. Kuo, Q.T. Le Gia, C. Schwab. Fast QMC matrix-vector multiplication. SIAM J. Sci. Comput. 37, A1436–A1450, 2015.

[5] J. Dick, F.Y. Kuo, I.H. Sloan. High-dimensional integration—the quasi-Monte Carlo way. Acta Numer. 22, 133–288, 2013.

[6] J. Dick, F. Pillichshammer. *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration.* Cambridge University Press, Cambridge, 2010.

[7] H. Faure, P. Kritzer. New star discrepancy bounds for (t, m, s)-nets and (t, s)-sequences. Monatsh. Math. 172, 55–75, 2013.

[8] H. Niederreiter. Low-discrepancy point sets obtained by digital constructions over finite fields, Czechoslovak Math. J. 42, 143–166, 1992.

[9] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics, 63. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1992.

[10] F. Pillichshammer. Tractability properties of the weighted star discrepancy of regular grids. J. Complexity 46, 103–112, 2018.

[11] I.H. Sloan, H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? J. Complexity 14, 1–33, 1998.

Authors' address:

Vishnupriya Anupindi
Peter Kritzer
Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstr. 69, 4040 Linz, Austria.
vishnupriya.anupindi@ricam.oeaw.ac.at
peter.kritzer@ricam.oeaw.ac.at