ARTICLE IN PRESS

Computer Physics Communications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Quasi-Monte Carlo methods for lattice systems: A first look☆

K. Jansen [a,b], H. Leovey [c], A. Ammon [a,d,*], A. Griewank [c], M. Müller-Preussker [d]

[a] NIC, DESY Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany

[b] Department of Physics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus

[c] Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

[d] Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany

ARTICLE INFO

Article history:
Received 11 March 2013
Received in revised form
5 July 2013
Accepted 12 October 2013
Available online xxxx

Keywords:
Quasi-Monte Carlo
Lattice field theory
Quantum mechanics
Anharmonic oscillator
Path integral
Error scaling
Reweighting

ABSTRACT

We investigate the applicability of quasi-Monte Carlo methods to Euclidean lattice systems for quantum mechanics in order to improve the asymptotic error behavior of observables for such theories. In most cases the error of an observable calculated by averaging over random observations generated from an ordinary Markov chain Monte Carlo simulation behaves like $N^{-1/2}$, where N is the number of observations. By means of quasi-Monte Carlo methods it is possible to improve this behavior for certain problems to N^{-1}, or even further if the problems are regular enough. We adapted and applied this approach to simple systems like the quantum harmonic and anharmonic oscillator and verified an improved error scaling.

Program summary

Program title: qar-0.1

Catalogue identifier: AERJ_v1_0

Program summary URL: http://cpc.cs.qub.ac.uk/summaries/AERJ_v1_0.html

Program obtainable from: CPC Program Library, Queen's University, Belfast, N. Ireland

Licensing provisions: GNU General Public Licence version 3

No. of lines in distributed program, including test data, etc.: 67759

No. of bytes in distributed program, including test data, etc.: 2165365

Distribution format: tar.gz

Programming language: C and C++.

Computer: PC.

Operating system: Tested on GNU/Linux, should be portable to other operating systems with minimal efforts.

Has the code been vectorized or parallelized?: No

RAM: The memory usage directly scales with the number of samples and dimensions: Bytes used = "number of samples" × "number of dimensions" × 8 Bytes (double precision).

Classification: 4.13, 11.5, 23.

External routines: FFTW 3 library (http://www.fftw.org)

Nature of problem:
Certain physical models formulated as a quantum field theory through the Feynman path integral, such as quantum chromodynamics, require a non-perturbative treatment of the path integral. The only known approach that achieves this is the lattice regularization. In this formulation the path integral is discretized to a finite, but very high dimensional integral. So far only Monte Carlo, and especially Markov chain-Monte Carlo methods like the Metropolis or the hybrid Monte Carlo algorithm have been used to calculate approximate solutions of the path integral. These algorithms often lead to the undesired effect

☆ This paper and its associated computer program are available via the Computer Physics Communication homepage on ScienceDirect (http://www.sciencedirect.com/science/journal/00104655).

* Corresponding author at: NIC, DESY Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany. Tel.: +49 33762 7 7192; fax: +49 33762 7 7413.
E-mail address: Andreas.Ammon@desy.de (A. Ammon).

0010-4655/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cpc.2013.10.011

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

of autocorrelation in the samples of observables and suffer in any case from the slow asymptotic error behavior proportional to $N^{-1/2}$, if N is the number of samples.

Solution method:
This program applies the quasi-Monte Carlo approach and the reweighting technique (respectively the weighted uniform sampling method) to generate uncorrelated samples of observables of the anharmonic oscillator with an improved asymptotic error behavior.

Unusual features:
The application of the quasi-Monte Carlo approach is quite revolutionary in the field of lattice field theories.

Running time:
The running time depends directly on the number of samples N and dimensions d. On modern computers a run with up to $N = 2^{16} = 65536$ (including 9 replica runs) and $d = 100$ should not take much longer than one minute.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Markov chain-Monte Carlo (Mc-MC) techniques are commonly the method of choice for the numerical evaluation of partition functions in statistical physics or path integrals in Euclidean time for models in high energy physics. The reason is that they are based on importance sampling and hence select the integration points automatically according to the corresponding weight in the integrand. Many algorithms have been developed to implement a Mc-MC, starting from the Metropolis algorithm, heat-bath and over-relaxation to cluster and hybrid Monte Carlo algorithms, see e.g. Refs. [1,2]. In this way, simulations of demanding 4-dimensional quantum field theories became possible and, in fact, were carried out very successfully to e.g. compute the low-lying hadron spectrum [3] or deriving bounds for the Higgs boson mass [4].

The drawback of Mc-MC is that it estimates the desired quantity stochastically and hence the results are affected by a statistical error which sometimes needs very long and computer time extensive samplings. Quantitatively, this sampling error behaves as $N^{-1/2}$ for a (thermalized) sample size of N. This error scaling behavior is often a real stumbling block in such Mc-MC simulations. If we consider lattice quantum chromodynamics as a typical system for Mc-MC calculations in high energy physics, then due to this error scaling and the very high computational demand of these simulations, it is often impossible to significantly decrease the error to the targeted precision. It would therefore be very desirable to have Monte Carlo methods available that possibly show a better error scaling.

Such methods in fact exist in the form of quasi-Monte Carlo (QMC) techniques [5,6], where it is known that the error scaling can be improved to an N^{-1}, or even more if the problem exhibits enough structure (smoothness, periodicity). QMC methods have been analyzed theoretically very thoroughly and comprehensively, and many successful applications in mathematical finance based on high-dimensional Gaussian integrals have been studied in the last two decades [7,8]. On the other hand, to the best of our knowledge, QMC methods were never tested successfully in high-dimensional models that are relevant for high energy physics.

In this paper we therefore want to perform a very first step towards the very challenging goal of applying QMC methods to generic field theories by looking at the non-trivial case of the anharmonic quantum mechanical oscillator discretized on a finite Euclidean time lattice and evaluated in the corresponding path integral formulation [9]. For the case of the anharmonic oscillator the system is no longer Gaussian and a successful application of QMC methods would be a first non-trivial test. Of course, even if such a test is successful, there would remain a long way to go before eventually addressing 4-dimensional quantum field theories, but a proof of concept would certainly open the promising road to attack field theories in the future.

We will start our discussion with a description of the harmonic oscillator in its time-discretized path integral formulation. Here the problem is fully Gaussian and the application of adequate QMC methods should lead to an improved error scaling, an expectation that we will see to be fulfilled. Nevertheless, the harmonic oscillator example can serve well to explain how QMC methods work and how an improved error scaling behavior is realized.

We will then proceed to look at the anharmonic oscillator and we will demonstrate that also in this case QMC leads to an improved error scaling, although the obtained rate of convergence is still not optimal and leaves space for improvements. We consider this, nevertheless, to be a very promising non-trivial result which bears the potential that also other models in quantum mechanics, e.g. the topological quantum mechanical action of Ref. [10] and even field theories, can be evaluated by QMC methods. For a first account of our studies, we refer to the proceedings contribution of Ref. [11].

2. Quantum mechanical harmonic and anharmonic oscillator

In this section we will discuss the basic steps for the quantization of the theory in the path integral approach and the discretization on a time lattice. The first step is the construction of the Lagrangian (resp. the action) of the corresponding classical mechanical system for a given path $x(t)$ of a particle with mass M_0. For a numerically stable evaluation of the path integral it is essential to pass on to Euclidean time. In this case the Lagrangian L and the action S is given by

$$L(x,t) = \frac{M_0}{2}\left(\frac{dx}{dt}\right)^2 + V(x) \tag{1}$$

$$S(x) = \int_0^T L(x,t)\, dt. \tag{2}$$

Depending on the scenario (harmonic or anharmonic oscillator) the potential $V(x)$ consists of two parts,

$$V(x) = \underbrace{\frac{\mu^2}{2}x^2}_{\text{harmonic part}} + \underbrace{\lambda x^4}_{\text{anharmonic part}}, \tag{3}$$

such that the parameter λ controls the anharmonic part of the theory. It should also be mentioned that in the anharmonic case the parameter μ^2 can take on negative values, leading then to a double well potential.

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

The next step is to discretize time into equidistant time slices with a spacing of a. The path is then only defined on the time slices:

$$t \to t_i = (i-1)\cdot a \quad i = 1 \ldots d \tag{4}$$

$$x(t) \to x_i = x(t_i). \tag{5}$$

On the lattice the derivative with respect to the time appearing in (1) (first term) will be replaced by the forward finite difference $\nabla x_i = \frac{1}{a}(x_{i+1} - x_i)$. The choice of the lattice derivative is not unique and requires special care, particularly if one considers more complicated models like lattice QCD. But in [9] it was shown that the lattice derivative chosen here permits a well defined continuum limit. Putting all the ingredients together, we can write the lattice action for the (an)harmonic oscillator as

$$S^{\text{latt}}(\mathbf{x}) = a\sum_{i=1}^{d}\left(\frac{M_0}{2}(\nabla x_i)^2 + V(x_i)\right). \tag{6}$$

For the path a cyclic boundary condition $x_{d+1} = x_1$ can be assumed. In the following the superscript "latt" will be dropped, as we will only refer to the lattice action from now on. The expectation value of an observable O of the quantized theory expressed in terms of the path integral reads as follows:

$$\langle O(\mathbf{x})\rangle = \frac{\int_{\mathbb{R}^d} O(\mathbf{x}) e^{-S(\mathbf{x})} dx_1 \ldots dx_d}{\int_{\mathbb{R}^d} e^{-S(\mathbf{x})} dx_1 \ldots dx_d}. \tag{7}$$

This expression is suitable for a numerical evaluation of certain quantities of the underlying theory. Up to now only Monte Carlo methods are known to give reliable results for dimensions $d \gg 10$. One type of such methods, often used in physics, is the Markov chain-Monte Carlo approach mostly applying the weight $\propto e^{-S(\mathbf{x})}$ for sampling paths $\{x_i\}$ (so-called "importance sampling"). In the next sections, we will provide a summary of the mathematical results for QMC methods (and their randomizations), and particularly recapitulate in a rather mathematical language the strict error scaling bounds for this methods. The reader more interested in the results may move directly to Section 7.

3. Direct Monte Carlo and quasi-Monte Carlo methods

We provide in Sections 3–7 some mathematical background of QMC methods. The reader who is more interested in our results, may proceed directly to Section 8. In many practical applications one is interested in calculating quotients of the form (7) where the action $S(.)$ and the observables $O(.)$ are usually smooth functions in high dimensions. In some special situations where one would like to deal with integrands of moderately high dimensions, one possible approach is to consider estimators $\hat{I}_1, \hat{I}_2$ for the integrals I_1, I_2 in the numerator and in the denominator of (7) separately, and then take $\hat{I}_1/\hat{I}_2$ as an estimation of $\langle O(\mathbf{x})\rangle$. Another possible approach one can consider is given by the so-called weighted uniform sampling (WUS) estimator, analyzed in [12]. In the latter case, one takes a joint estimator for the total quantity $\langle O(\mathbf{x})\rangle$, using a single direct sampling method. We will show some characteristics of the WUS estimator in Section 7, and we will refer from now on to the latter two approaches as *direct* sampling methods for estimating (7). In many interesting examples, we encounter the case were the action $S(.)$ and the observable $O(.)$ lead to integrals I_1, I_2 of Gaussian type. Then the integrals I_1, I_2 can be written in the form

$$I_i = \frac{1}{(2\pi)^{d/2}\sqrt{\det(C)}}\int_{\mathbb{R}^d} g_i(\mathbf{x}) e^{-\frac{1}{2}\mathbf{x}^\top C^{-1}\mathbf{x}} d\mathbf{x},$$
$$\mathbf{x} = (x_1, \ldots, x_d),\ i = 1, 2,$$

where C denotes the covariance matrix of the Gaussian density function. A transformation to the unit cube in $\mathbb{R}^d$ can be applied such that the corresponding integrals take the form

$$I = \int_{[0,1]^d} g(A\Phi^{-1}(\boldsymbol{z}))d\boldsymbol{z} = \int_{[0,1]^d} f(\boldsymbol{z})d\boldsymbol{z} = I_{[0,1]^d}(f),$$
$$\boldsymbol{z} = (z_1, \ldots, z_d). \tag{8}$$

Here $AA^\top = C$ is some symmetric factorization of the covariance matrix, and $\Phi^{-1}(\boldsymbol{z}) := (\Phi^{-1}(z_1), \ldots, \Phi^{-1}(z_d))^\top$, where $\Phi^{-1}(\cdot)$ represents the inverse of the normal cumulative distribution function $\Phi(\cdot)$.

In the classical direct Monte-Carlo (MC) approach one tries to estimate (8) by generating samples pseudo-randomly. One starts with a finite sequence of independent identically distributed (i.i.d.) samples $P_N = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N\}$, where the points $\boldsymbol{z}_j$, $1 \le j \le N$, have been generated from the uniform distribution in $[0, 1]^d$. Then, the quadrature rule is fixed by taking the average of the function evaluations for f

$$Q_N := \frac{1}{N}\sum_{j=1}^{N} f(\boldsymbol{z}_j)$$

as an approximation of the desired integral $\int_{[0,1]^d} f(\boldsymbol{z})\ d\boldsymbol{z}$. The resulting estimator $\hat{Q}_N$ is unbiased. The integration error can be approximated via the central limit theorem, given that f belongs to $L_2([0, 1]^d)$. The variance of the estimator $\hat{Q}_N$ is given by

$$\frac{\sigma^2}{N} = \frac{1}{N}\left(\int_{[0,1]^d} f^2(\boldsymbol{z})\ d\boldsymbol{z} - \left(\int_{[0,1]^d} f(\boldsymbol{z})\ d\boldsymbol{z}\right)^2\right).$$

As measured by its standard deviation from zero the integration error associated with the MC approach is then of order $O(N^{-1/2})$. The quality of the MC samples relies on the selected pseudo-random number generators of uniform samples. Here we use the *Mersenne Twister* generator from Matsumoto and Nishimura (see [13]). MC is in general a very reliable tool in high-dimensional integration, but the order of convergence is in fact rather poor.

In contrast, QMC methods generate deterministically point sets that are more regularly distributed than the pseudo-random points from MC (see [5,14,15,6]). Typical examples of QMC are shifted lattice rules and low-discrepancy sequences. In order to give a short introduction to the subject, we define now the classical notion of discrepancy of a finite sequence of points P_N in $[0, 1)^d$. Given $P_N = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N\}$ a set of points in $[0, 1)^d$, and a nonempty family $\mathbb{I}$ of Lebesgue-measurable sets in $[0, 1)^d$, we define the classical discrepancy function by

$$D(\mathbb{I}; P_N) := \sup_{B\in\mathbb{I}}\left|\frac{\sum_{i=1}^{N} c_B(\boldsymbol{z}_i)}{N} - \lambda_d(B)\right|,$$

where c_B is the characteristic function of B, and λ_d is the Lebesgue measure in $\mathbb{R}^d$. This allows us to define the so-called *star discrepancy* as follows.

Definition 3.1. We define the *star discrepancy* $D^\star(P_N)$ of the point set P_N by $D^\star(P_N) := D(\mathbb{I}; P_N)$, where $\mathbb{I}$ is the family of all subintervals of the form $\prod_{i=1}^{d}[0, u_i)$, with $u_i \ge 0$, $1 \le i \le d$.

The *star discrepancy* can be considered as a measure of the worst difference between the uniform distribution and the sampled distribution in $[0, 1)^d$ attributed to the point set P_N. The usual way to analyze QMC as a deterministic method is by choosing a class of integrand functions F, and a measure of discrepancy $D(P_N)$ for the point sets P_N. Then, the deterministic integration error is usually given in the form

$$\left|Q_N - \int_{[0,1]^d} f(\boldsymbol{z})\ d\boldsymbol{z}\right| \le D(P_N)V(f),$$

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

where $V(f)$ measures a particular variation of the function $f \in F$. A classical particular error bound in this form is the famous Koksma–Hlawka inequality, where $D(P_N)$ is taken to be the *star discrepancy* of the point set P_N, and $V(f)$ is the variation in the sense of Hardy and Krause of f.

In the context of QMC, a sequence of points $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots$ in $[0,1)^d$ is called a low-discrepancy sequence if

$$D^\star(\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N\}) = O(N^{-1}(\log(N))^d). \tag{9}$$

An important part of QMC constructions satisfying this asymptotic bound are known under the name of (t, d)-sequences and will be discussed in more detail in Section 4. For moderate values of N, the influence of the logarithmic term in (9) usually cannot be ignored (see 5.7 and 5.8 in [16]), because the term $N^{-1}(\log(N))^d$ grows until $N > 2^d$. This normally prevents a straightforward use of low-discrepancy sequences in practical situations with very large dimensions. The latter situations are typically very high-dimensional integration problems where all variables and interactions between variables are equally important.

The practical success of QMC sequences in very high-dimensional integration problems with enough smoothness usually relies in an appropriate combination with an *effective-dimension* reduction transformation. We will discuss the analysis of effective-dimensions in more detail in Section 6. Well investigated settings for the integration error analysis of functions exhibiting a concentration of importance in few variables or groups of few variables are the so-called *weighted reproducing kernel Hilbert spaces* (see [14]), which will be considered briefly in the following section.

3.1. Quasi-Monte Carlo errors and complexity

For error analysis of QMC methods, there are certain reproducing kernel Hilbert spaces $\mathbb{F}_d$ of functions $f : [0,1]^d \to \mathbb{R}$ that are particularly useful (see [17]). Let us denote now with $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ the inner product and norm in $\mathbb{F}_d$. A reproducing kernel is a function $K : [0,1]^d \times [0,1]^d \to \mathbb{R}$ satisfying the properties

1. $K(\cdot, \boldsymbol{y}) \in \mathbb{F}_d$ for each $\boldsymbol{y} \in [0,1]^d$
2. $\langle f, K(\cdot, \boldsymbol{y}) \rangle = f(\boldsymbol{y})$ for each $\boldsymbol{y} \in [0,1]^d$ and $f \in \mathbb{F}_d$.

If the integral

$$I(f) = \int_{[0,1]^d} f(\boldsymbol{z}) d\boldsymbol{z}$$

is a continuous functional on the space $\mathbb{F}_d$, then the worst case quadrature error

$$e_N(\mathbb{F}_d) := \sup_{f \in \mathbb{F}_d, \|f\| \le 1} |I(f) - Q_N(f)|$$

for point sets $P_N = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N\}$ and QMC algorithms for the space $\mathbb{F}_d$ can be given by

$$e_N(\mathbb{F}_d) = \sup_{\|f\| \le 1} |\langle f, h_N \rangle| = \|h_N\|$$

for some $h_N \in \mathbb{F}_d$ due to Riesz' representation theorem. In this case, the *representer* h_N of the quadrature error is given explicitly in terms of the kernel by

$$h_N(\boldsymbol{z}) = \int_{[0,1]^d} K(\boldsymbol{z}, \boldsymbol{y}) d\boldsymbol{y} - \frac{1}{N} \sum_{i=1}^{N} K(\boldsymbol{z}, \boldsymbol{z}_i), \quad \forall \boldsymbol{z} \in [0,1]^d.$$

Tensor product reproducing kernel Hilbert spaces are of particular interest, since the multivariate kernel results as the product of the underlying univariate kernels. In QMC error analysis, the weighted (anchored) tensor product Sobolev space introduced in [18] is often considered:

$$\mathbb{F}_d = \bigotimes_{i=1}^{d} W_2^1([0,1]),$$

also denoted as $\mathbb{F}_d = W_{2,\mathrm{mix}}^{(1,\ldots,1)}([0,1]^d)$, where $W_2^1([0,1])$ is the Sobolev space of absolutely continuous functions on $[0,1]$ with first order derivatives in $L_2([0,1])$. The weighted norm $\|f\|_\gamma^2 = \langle f, f \rangle_\gamma$ results from the inner product

$$\langle f, g \rangle_\gamma = \sum_{u \subseteq \{1,\ldots,d\}} \prod_{j \in u} \gamma_j^{-1} \int_{[0,1]^{|u|}} \frac{\partial^{|u|}}{\partial \boldsymbol{z}_u} f(\boldsymbol{z}_u, \boldsymbol{1}) \times \frac{\partial^{|u|}}{\partial \boldsymbol{z}_u} g(\boldsymbol{z}_u, \boldsymbol{1}) d\boldsymbol{z}_u, \tag{10}$$

where for $u \subseteq \{1, \ldots, d\}$ we denote by $|u|$ its cardinality, and $(\boldsymbol{z}_u, \boldsymbol{1})$ denotes the vector containing the coordinates of $\boldsymbol{z}$ with indices in u, and the other coordinates set equal to 1.

In this case the reproducing kernel is given by

$$K_{d,\gamma}(\boldsymbol{z}, \boldsymbol{y}) = \prod_{j=1}^{d} (1 + \gamma_j [1 - \max(z_j, y_j)]) \quad \text{for } \boldsymbol{z}, \boldsymbol{y} \in [0,1]^d.$$

The weighted tensor product Sobolev space allows for explicit QMC constructions deriving error estimates of the form

$$e_N(\mathbb{F}_d) \le C(\delta) N^{-1+\delta} \quad \text{for } \delta \in \left(0, \frac{1}{2}\right], \tag{11}$$

where the constant $C(\delta)$ is independent on the dimension d, if the sequence of weights (γ_j) satisfies the condition (see [19])

$$\sum_{j=1}^{\infty} \gamma_j^{\frac{1}{2(1-\delta)}} < \infty.$$

Traditional unweighted function spaces considered for integration suffer from the curse of dimensionality. Their weighted variants describe a setting where the variables or group of variables may vary in importance, corresponding to an anisotropic problem. Many integration problems in practice start with an isotropic setting but can be modified to an anisotropic one using a proper transformation. The concentration of importance in few variables or groups of few variables gives a partial explanation of why some very high-dimensional spaces become tractable for QMC.

Explicit QMC constructions satisfying (11) are for example *shifted lattice rules* for weighted spaces [19]. The rate (11) can be also obtained for Niederreiter and Sobol' sequences (see [20]). The idea of "weighting" the norm of the spaces to obtain tractable results can be applied in fact to more general function spaces than smooth function spaces of tensor product form, and many integration examples can be found in [14]. In our numerical experiments, we have so far used QMC algorithms based on a particular type of low-discrepancy sequence. Numerical experiments with shifted lattice rules will be carried out in the near future, following new techniques for fixing adequate weights introduced in [21].

4. Low-discrepancy (t, d)-sequences

The most well known type of low-discrepancy sequences are the so-called (t, d)-sequences. To introduce how (t, m, d)-nets and (t, d)-sequences are defined, we consider first *elementary intervals* in a integer base $b \ge 2$. Let E be any sub-interval of $[0,1)^d$ of the form $E = \prod_{i=1}^{d} [a_i b^{-c_i}, (a_i + 1) b^{-c_i})$ with $a_i, c_i \in \mathbb{N}$, $c_i \ge 0$, $0 \le a_i < b^{-c_i}$ for $1 \le i \le d$. An interval of this form is called an elementary interval in base b.

Definition 4.1. Let $0 \le t \le m$ be integers. A (t, m, d)-net in base b is a point set P_N of $N = b^m$ points in $[0,1)^d$ such that every elementary interval E in base b with $\lambda_d(E) = \frac{b^t}{b^m}$ contains exactly b^t points.

Definition 4.2. Let $t \geq 0$ be an integer. A sequence $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots$ of points in $[0,1)^d$ is a (t,d)-sequence in base b if for all integers $k \geq 0$ and $m > t$, the point set consisting of $N = b^m$ points $\boldsymbol{z}_i$ with $kb^m \leq i < (k+1)b^m$ is a (t,m,d)-net in base b.

The parameter t is called the *quality parameter* of the (t,d)-sequences. In [22, Theorem 4.17], it is shown that (t,d)-sequences are in fact low-discrepancy sequences. We reproduce this result in the following theorem.

Theorem 4.3. *The star-discrepancy $D^\star$ of the first N terms P_N of a (t,d)-sequence in base b, satisfies*

$$ND^\star(P_N) \leq C(d,b)b^t(\log(N))^d + O(b^t(\log(N))^{d-1}),$$

where the implied constants depend only on b and d. If either $d = 2$ or $b = 2$, $d = 3, 4$, we have

$$C(d,b) = \frac{1}{d}\left(\frac{b-1}{2\log(b)}\right)^d,$$

and otherwise

$$C(d,b) = \frac{1}{d!}\frac{b-1}{2\lfloor b/2\rfloor}\left(\frac{\lfloor b/2\rfloor}{\log(b)}\right)^d.$$

Explicit constructions of (t,d)-sequences are available. Examples are the generalized Faure, Sobol', Niederreiter and Niederreiter–Xing sequences. All these examples fall into the category of constructions called *digital sequences*, see [15]. To complete this section, we will describe briefly Sobol' sequences and give references for their practical implementations. Sobol' sequences are among the most widely used and recommended QMC sequences by simulation practitioners (see [7] for successful applications of Sobol' sequences in finance), and they are the QMC sequences selected for our numerical experiments.

4.1. Sobol' sequences and implementations

The pioneering work of Sobol' [23] introduced the first known construction of (t,d)-sequences, and they can be viewed now as a special case of the so-called generalized Niederreiter sequences in base $b = 2$ (see Chapter 8 in [15] and references therein). The basic construction of Sobol' sequences can be described as follows:

1. Let $p_1, \ldots, p_d$ be primitive polynomials of degree $\deg(p_i) =: e_i$ over the field $\mathbb{F}_2[x]$,
$$p_i(x) = x^{e_i} + a_{1,i}x^{e_i-1} + a_{2,i}x^{e_i-2} + \cdots + a_{e_i-1}x + 1 \quad \text{for } 1 \leq i \leq d,$$
sorted according to their degree in increasing order.
2. Let $1 \leq m_{1,i}, \ldots, m_{e_i,i}$ be odd natural numbers with $m_{k,i} < 2^k$ for $1 \leq k \leq e_i$, $1 \leq i \leq d$. Then define for $k > e_i$ recursively
$$m_{k,i} = 2a_{1,i}m_{k-1,i} \oplus \cdots \oplus 2^{e_i-1}a_{e_i-1}m_{k-e_i+1,i} \oplus 2^{e_i}m_{k-e_i,i} \oplus m_{k-e_i,i} \quad \text{for } 1 \leq i \leq d,$$
where the operator $\oplus$ is the bit-by-bit exclusive-or operator.
3. Define the *direction numbers* $v_{k,i}$ by
$$v_{k,i} = \frac{m_{k,i}}{2^k} \quad \text{for } k \geq 1,\ 1 \leq i \leq d.$$
4. Consider a natural number n with binary expansion $n = n_0 + n_1 2 + \cdots + n_{r-1}2^{r-1}$, and define
$$z_{n,i} = n_0 v_{1,i} \oplus n_1 v_{2,i} \oplus \cdots \oplus n_{r-1}v_{r,i} \quad \text{for } 1 \leq i \leq d.$$
5. Finally consider the sequence of points $\boldsymbol{z}_0, \boldsymbol{z}_1, \ldots$ in $[0,1)^d$ defined by
$$\boldsymbol{z}_n = (z_{n,1}, \ldots, z_{n,d}).$$

A sequence of points $\boldsymbol{z}_0, \boldsymbol{z}_1, \ldots$ in $[0,1)^d$ as defined above is called a Sobol' sequence. The quality parameter of Sobol' sequences is given by

$$t = \sum_{i=1}^{d}(e_i - 1).$$

The direction numbers v defined above determine the quality of low dimensional projections of the points in the Sobol' sequences. Sobol' [23] introduced an additional uniformity condition called *Property A* in order to give a criterion for selection of initial m numbers in the recurrence stated above. Efficient implementations of Sobol' sequences are based on *Gray code*. A classical reference for practical implementation is [24]. Joe and Kuo [25,26] give an alternative for selection of direction numbers based in a *weighted* approach, focused in a setting where the importance of variables decay as their dimension number increase. Moreover, one can find a three-pages-note with a short and simple description on implementation of Sobol' sequences at the website http://web.maths.unsw.edu.au/~fkuo/sobol/index.html. New developments of Sobol' sequences and comparison between available implementations can be found in [27].

5. Randomized QMC

There are some advantages in retaining the probabilistic properties of the sampling. There are practical hybrid methods permitting us to combine the good features of MC and QMC. Randomization is an important tool for QMC if we are interested in a practical error estimate of our sample quadrature Q_N to the desired integral. One goal is to randomize the deterministic point set P_N generated by QMC in such a way that the estimator $\hat{Q}_N$ preserves unbiasedness. Another important goal is to preserve the better equidistribution properties of the deterministic construction.

The simplest form of randomization applied to *digital sequences* seems to be the technique called *digital b-ary shifting*. In this case, we add a random shift $\Delta \in [0,1)^d$ to each point of the deterministic set $P_N = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N\}$ using operations over the selected ring $\mathbb{F}_b$. The application of this randomization preserves in particular the t value of any projection of the point set (see [5] and references therein). The resulting estimator is unbiased.

The second randomization method we consider is the one introduced by Owen [28] in 1995. He considered (t,m,d)-nets and (t,d)-sequences in base b and applied a randomization procedure based on permutations of the digits of the values of the coordinates of points in these nets and sequences. This can be interpreted as a random scrambling of the points of the given sequence in such a way that the net structure remains unaffected. We do not discuss Owen's randomization procedure in detail here, from now on called *Owen's scrambling*. The main results of this randomization procedure can be stated in the following proposition.

Proposition 5.1 (*Equidistribution*)**.** *A randomized (t,m,d)-net in base b using Owen's scrambling is again a (t,m,d)-net in base b with probability* 1. *A randomized (t,d)-sequence in base b using Owen's scrambling is again a (t,d)-sequence in base b with probability* 1.

Proposition 5.2 (*Uniformity*)**.** *Let $\tilde{\boldsymbol{z}}_i$ be the randomized version of a point $\boldsymbol{z}_i$ originally belonging to a (t,m,d)-net in base b or a (t,d)-sequence in base b, using Owen's scrambling. Then $\tilde{\boldsymbol{z}}_i$ has the uniform distribution in $[0,1)^d$, that is, for any Lebesgue measurable set $G \subseteq [0,1)^d$, $P(\tilde{\boldsymbol{z}}_i \in G) = \lambda_d(G)$, with λ_d the d-dimensional Lebesgue measure.*

The last two propositions state that after *Owen's scrambling* of *digital sequences* we retain unaffected the low-discrepancy

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

properties of the constructions, and that after this randomization procedure we obtain random samples uniformly distributed in $[0,1)^s$.

The basic results about the variance of the randomized QMC estimator $\hat{Q}_N$ after applying *Owen's scrambling* to (t,m,d)-nets in base b (or of (t,d)-sequences in base b) can be found in [29]. We summarize these results in the following theorem.

Theorem 5.3. *Let $\tilde{z}_i$, $1 \le i \le N$, be the points of a scrambled (t,m,d)-net in base b, and let f be a function on $[0,1)^d$ with integral I and variance $\sigma^2 = \int (f-I)^2 d\boldsymbol{z} < \infty$. Let $\hat{Q}_N = N^{-1}\sum_{i=1}^N f(\tilde{z}_i)$, where $N = b^m$. Then for the variance $V(\hat{Q}_N)$ of the randomized QMC estimator it holds that*

$$V(\hat{Q}_N) = o(N^{-1}), \quad \text{as } N \to \infty, \quad \text{and}$$
$$V(\hat{Q}_N) \le \frac{b^t}{N}\left(\frac{b+1}{b-1}\right)^d \sigma^2.$$

For $t = 0$ we have

$$V(\hat{Q}_N) \le \frac{1}{N}\left(\frac{b}{b-1}\right)^{d-1} \sigma^2.$$

The above theorem says that the variance of the randomized QMC estimator $\hat{Q}_N$ using scrambled $(0,m,d)$-nets is always smaller than a small multiple of the variance of the corresponding MC estimator. If the integrand at hand is smooth enough, using *Owen's scrambling* it can be shown that one can obtain an improved asymptotic error estimate of order $O(N^{-\frac{3}{2}-\frac{1}{d}+\delta})$, for any $\delta > 0$, see [30]. Improved scrambling techniques have been developed in [31,32].

6. Effective dimensions and sensitivity indices

In many practical applications, one encounters functions for which the total variance is concentrated in a small part of its ANOVA terms. The notion of *effective dimension* of a function was first introduced in [33] to describe the contribution of a group of variables to the total variance.

6.1. ANOVA decomposition

Using the *ANOVA (Analysis of Variance) decomposition* we decompose a function into a sum of simpler functions, see [34]. Let $D = \{1,\dots,d\}$. For any subset $\mathbf{i} \subseteq D$, let $|\mathbf{i}|$ denote its cardinality and $(D-\mathbf{i})$ be its complementary set in D. Let $\boldsymbol{z}_\mathbf{i} = (z_j : j \in \mathbf{i})$ be the $|\mathbf{i}|$-dimensional vector containing the coordinates of $\boldsymbol{z}$ with indices in $\mathbf{i}$. Now assume that f is a square integrable function. Then we can write f as the sum of 2^d ANOVA terms:

$$f(\boldsymbol{z}) = \sum_{\mathbf{i}\subseteq D} f^\mathbf{i}(\boldsymbol{z}),$$

where the ANOVA terms $f^\mathbf{i}(\mathbf{x})$ are defined recursively by

$$f^\mathbf{i}(\boldsymbol{z}) = \int_{[0,1]^{d-|\mathbf{i}|}} f(\boldsymbol{z}_\mathbf{i}, \boldsymbol{z}_{D-\mathbf{i}})d\boldsymbol{z}_{D-\mathbf{i}} - \sum_{\mathbf{j}\subsetneq\mathbf{i}} f^\mathbf{j}(\boldsymbol{z}),$$

and $f^\varnothing = I(f)$. The sum of the right-hand side is over strict subsets $\mathbf{j} \neq \mathbf{i}$, and we use the convention $\int_{[0,1]^0} f(\boldsymbol{z})d\boldsymbol{z}_\varnothing = f(\boldsymbol{z})$. The ANOVA terms enjoy the following interesting properties:

1. $\int_0^1 f^\mathbf{i}(\boldsymbol{z})dz_j = 0$ for $j \in \mathbf{i}$.
2. The decomposition is orthogonal, in that $\int_{[0,1]^d} f^\mathbf{i}(\boldsymbol{z}) f^\mathbf{j}(\boldsymbol{z})d\boldsymbol{z} = 0$ whenever $\mathbf{i} \neq \mathbf{j}$.
3. Let $\sigma^2(f) = \int_{[0,1]^d} f(\boldsymbol{z})^2\, d\boldsymbol{z} - (I(f))^2$ be the variance of f, then we have:
$$\sigma^2(f) = \sum_{\mathbf{i}\subseteq D} \sigma_\mathbf{i}^2(f), \quad \text{where } \sigma_\mathbf{i}^2(f) = \int_{[0,1]^d} f^\mathbf{i}(\boldsymbol{z})^2\, d\boldsymbol{z}$$
for $|\mathbf{i}| > 0$ is the variance of $f^\mathbf{i}$ and $\sigma_\varnothing^2(f) = 0$.

Definition 6.1.

1. f is said to have effective dimension in the superposition sense d_s with proportion p, for $0 < p < 1$, if d_s is the smallest integer that satisfies
$$\sum_{|\mathbf{i}|\le d_s} \sigma_\mathbf{i}^2(f) \ge p\sigma^2(f).$$
2. f is said to have effective dimension in the truncation sense d_t with proportion p, for $0 < p < 1$, if d_t is the smallest integer that satisfies
$$\sum_{\mathbf{i}\subseteq\{1,\dots,d_t\}} \sigma_\mathbf{i}^2(f) \ge p\sigma^2(f).$$

One can estimate the effective dimension in truncation sense based on the algorithm proposed by Wang and Fang [35]. They show that the following equality holds:

$$\int_{[0,1]^{2d-|u|}} f(\boldsymbol{z})f(\boldsymbol{z}_u, \mathbf{y}_{D-u})d\boldsymbol{z}d\mathbf{y}_{D-u} = \sum_{\mathbf{i}\subseteq u} \sigma_\mathbf{i}^2(f) + f^\varnothing.$$

Thus, for estimating the effective dimension in truncation sense, we need to estimate the following tree type of integrals

$$\int_{[0,1]^d} f(\boldsymbol{z})d\boldsymbol{z}, \qquad \int_{[0,1]^d} f^2(\boldsymbol{z})d\boldsymbol{z}, \qquad \int_{[0,1]^{2d-|u|}} f(\boldsymbol{z})f(\boldsymbol{z}_u, \mathbf{y}_{D-u})d\boldsymbol{z}d\mathbf{y}_{D-u}, \tag{12}$$

for $u = \{1,\dots,l\}$, $l = 1,2,\dots$, using MC or QMC, until the proportion of variance defining the effective dimension is reached. In many applications, the proportion value is usually taken as $p = 0.99$.

Given any nonempty family T of subsets of D, we can consider the function defined by the corresponding ANOVA terms $f_T(\boldsymbol{z}) := \sum_{\mathbf{i}\in T} f^\mathbf{i}(\boldsymbol{z})$. For example, given a fixed proportion value p we can consider the sets $T = \{\mathbf{i} : \mathbf{i} \subseteq \{1,\dots,d_t\}\}$ or $T = \{\mathbf{i} : |\mathbf{i}| \le d_s\}$ to define the *effective part* f_T of the function f in truncation or superposition sense respectively. The integration error for f of a QMC algorithm Q_N can be bounded then by

$$|I(f) - Q_N(f)| \le |I(f_T) - Q_N(f_T)| + |I(f-f_T) - Q_N(f-f_T)|. \tag{13}$$

If f_T is the effective part of a function exhibiting low-effective dimension in superposition or truncation sense, then the second error term in the right-hand side of (13) represents the integration error over the *rest function* $f - f_T$ having a relatively small variance. For many practical applications, the second error term in (13) is believed to be so small that it can be neglected (see [36]).

If the truncation effective dimension is small, then few variables are important for sampling. If the superposition effective dimension is small, say d_s equals 2, 3 or maybe 4, then some QMC sequences and their randomizations are also expected to outperform MC, because they can exhibit much better equidistributed low-dimensional projections than MC (see [36,27]).
The ordering of the variables of the integrand is important for achieving a reduction of the effective dimension in the truncation sense d_t, and usually affects the performance of QMC and their randomizations in practice. Sensitivity indices usually help to order the variables in a convenient way for integration with QMC.

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

6.2. Derivative based sensitivities

As pointed out by Sobol' and Kucherenko in [37], very often derivative-based measures of sensitivities can successfully be used for detecting non-essential variables. Small values of first-order derivatives of a function implies small values of one-dimensional total Sobol' sensitivity indices. Let $\sigma^2_{\mathbf{i}}(f)$ denote the partial variance corresponding to the ANOVA term $f^{\mathbf{i}}$. Define

$$\sigma^2_{\{j\}}(f)^{\text{tot}} = \sum_{\mathbf{i}\subset D: j\in\mathbf{i}} \sigma^2_{\mathbf{i}}(f),$$

then it is shown in [34,37] that

$$\sigma^2_{\{j\}}(f)^{\text{tot}} = \frac{1}{2}\int_{[0,1]^d}\int_0^1 [f(\boldsymbol{z}) - f(z_1,\ldots,z_{j-1}, z'_j, z_{j+1},\ldots,z_n)]^2 d\boldsymbol{z}dz'_j,$$

from which one obtain the following two results:

1. if $c < |\frac{\partial f}{\partial z_j}| < C$, then
$$\frac{c^2}{12} \le \sigma^2_{\{j\}}(f)^{\text{tot}} \le \frac{C^2}{12},$$
2. and if $\frac{\partial f}{\partial z_j} \in L_2([0,1]^d)$, then
$$\sigma^2_{\{j\}}(f)^{\text{tot}} \le \frac{1}{\pi^2}\int_{[0,1]^d}\left(\frac{\partial f}{\partial z_j}(\mathbf{x})\right)^2 d\boldsymbol{z}. \tag{14}$$

As a consequence of the bounds stated above, the total variance corresponding to non-essential variables of a function can be bounded using first-order derivatives information. In a wide variety of problems in practice, the gradient of a scalar function can be efficiently computed through algorithmic differentiation (see [38]), at a cost at most 4 times of that for evaluating the original function. Thus, a cheap method for estimating derivative-based sensitivities, and an upper bound on the effective dimension in the truncation sense (as stated in the following simple proposition), may be available using algorithmic differentiation. The variance is, clearly, invariant to a permutation of the variables. This allowed us to consider the following definition and proposition.

Definition 6.2. Given an bijection (permutation) $\pi : \{1,\ldots,d\} \to \{1,\ldots,d\}$, f is said to have π-effective dimension in the truncation sense d_t with proportion p, for $0 < p < 1$, if d_t is the smallest integer that satisfies

$$\sum_{\mathbf{i}\subset\{\pi^{-1}(1),\ldots,\pi^{-1}(d_t)\}} \sigma^2_{\mathbf{i}}(f) \ge p\sigma^2(f).$$

Proposition 6.3. *Let $f \in L_2([0,1]^d)$ such that $\frac{\partial f}{\partial z_j} \in L_2([0,1]^d) \forall 1 \le j \le n$. Consider the derivative-based sensitivities*

$$v_i := \frac{1}{\pi^2}\int_{[0,1]^d}\left(\frac{\partial f}{\partial z_j}(\boldsymbol{z})\right)^2 d\boldsymbol{z},$$

and consider any permutation $\pi^* : \{1,\ldots,d\} \to \{1,\ldots,d\}$ such that

$$v_{(\pi^*)^{-1}(k)} \ge v_{(\pi^*)^{-1}(k+1)}, \quad \forall\, 1 \le k \le d-1,$$

(a non-increasing ordering of the sensitivities v_i, resulting in what is called by the authors a Diff-decay-ordering π^*).
Let $0 < p < 1$ be a fixed proportion parameter. If there exists an integer m such that

$$\sum_{j=m+1}^{d} v_{(\pi^*)^{-1}(j)} \le (1-p)\sigma^2(f) \tag{15}$$

then it follows that the π^*-effective dimension in the truncation sense with proportion p is at most m.

Proof. Let d_t denote the π^*-effective dimension in the truncation sense with proportion p. Consider m satisfying (15) and define $T_m = \{\mathbf{i} : \mathbf{i} \subset \{(\pi^*)^{-1}(1),\ldots,(\pi^*)^{-1}(m)\}\}$ and $f_{T_m} = \sum_{\mathbf{i}\in T_m} f^{\mathbf{i}}(\boldsymbol{z})$. It follows from the L_2 orthogonality of ANOVA decomposition and (14) that

$$\begin{aligned}\sigma^2(f) - \sigma^2(f_{T_m}) = \sigma^2(f - f_{T_m}) &= \sum_{\{\mathbf{i}:\mathbf{i}\subset D\wedge\mathbf{i}\notin T_m\}} \sigma^2_{\mathbf{i}}(f) \\ &\le \sum_{j=m+1}^{d}\sum_{\{\mathbf{i}\subset D:(\pi^*)^{-1}(j)\in\mathbf{i}\}} \sigma^2_{\mathbf{i}}(f) \\ &= \sum_{j=m+1}^{d} \sigma^2_{\{(\pi^*)^{-1}(j)\}}(f)^{\text{tot}} \\ &\le \sum_{j=m+1}^{d} v_{(\pi^*)^{-1}(j)} \le (1-p)\sigma^2(f).\end{aligned}$$

It follows $\sigma^2(f_{T_m}) \ge p\sigma^2(f)$ and thus $d_t \le m$, which is what was required to be proved. □

7. Weighted uniform sampling

In this section we discuss the method we used to approximate observables as they are defined in equation (7). Before the WUS method can be applied to this expression, it is necessary to perform a transformation of the variables x_i to the d-dimensional unit cube, $[0,1]^d$. In the cases we will consider in Section 8, this transformation will always be of the form

$$x_i = \sum_j A_{ij}\Phi^{-1}(z_j), \tag{16}$$

with A being a positive definite matrix and Φ^{-1} the inverse of the PDF of the standard normal distribution. After the transformation, equation (7) reads as

$$\langle O\rangle = \frac{\int_{[0,1]^d} O(A\Phi^{-1}(\boldsymbol{z}))W(\boldsymbol{z})dz_1\ldots dz_d}{\int_{[0,1]^d} W(\boldsymbol{z})dz_1\ldots dz_d} \tag{17}$$
$$W(\boldsymbol{z}) = \exp\left[-S(A\Phi^{-1}(\boldsymbol{z})) + \frac{1}{2}\sum_i(\Phi^{-1}(z_i))^2\right].$$

Now, in the WUS method points $\boldsymbol{z}_j$, $1 \le j \le N$, are generated from a uniform distribution in $[0,1]^d$. Using these points, a quotient of integrals of the form

$$Q(f_1,f_2) := \frac{\int_{[0,1]^d} f_1(\boldsymbol{z})d\boldsymbol{z}}{\int_{[0,1]^d} f_2(\boldsymbol{z})d\boldsymbol{z}}$$

can then be approximated by taking the rule

$$Q_N(f_1,f_2) := \frac{\sum_{j=1}^{N} f_1(\boldsymbol{z}_j)}{\sum_{j=1}^{N} f_2(\boldsymbol{z}_j)}, \tag{18}$$

where the functions f_i could be of very general, in particular non-Gaussian, nature. For our example these functions can be read off from Eq. (17): $f_1 = O(A\Phi^{-1}(\boldsymbol{z}))W(\boldsymbol{z})$ and $f_2 = W(\boldsymbol{z})$. For the case that $W(\boldsymbol{z})$ is really a function of $\boldsymbol{z}$ (and not just a constant), this way of evaluating integrals over certain weight functions W is known as the reweighting technique in field theory or statistical physics. A crucial element of the WUS (reweighting) method is that

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

the sampling points have a large enough overlap with the weight functions f_i considered. The resulting WUS estimator $\hat{Q}_N(f_1, f_2)$ from (18) has been analyzed in [12] and applications have been investigated for example in [39,40]. The bias and the root mean square error (RMSE) of this estimator satisfy

$$\mathrm{Bias}(\hat{Q}_N(f_1,f_2)) = \frac{Q(f_1,f_2)\,\mathrm{var}(f_2)}{N} - \frac{\mathrm{cov}(f_1,f_2)}{N} + O\left(N^{-\frac{3}{2}}\right)$$

$$\mathrm{RMSE}(\hat{Q}_N(f_1,f_2)) = \frac{\sqrt{\mathrm{var}(f_1) + (Q(f_1,f_2))^2\mathrm{var}(f_2) - 2Q(f_1,f_2)\,\mathrm{cov}(f_1,f_2)}}{\sqrt{N}} + O\left(N^{-\frac{3}{4}}\right).$$

The bias of the estimator in this case is asymptotically negligible compared with the RMSE.

A deterministic version of the WUS estimator has been considered in [39]. In particular, it follows from Theorem 4.2 in [39] that if the integrands f_1, f_2 in (17) are of bounded variation in the sense of Hardy and Krause, then by the use of a low-discrepancy sequence $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots$ instead of i.i.d. uniform random samples we obtain the integration error asymptotic

$$|Q_N(f_1,f_2) - Q(f_1,f_2)| = O(N^{-1}(\log(N))^d).$$

Similar results for the bias and RMSE of $\hat{Q}_N(f_1, f_2)$ considering randomized QMC sequences instead of i.i.d. random samples are not known to the authors. Nevertheless, the numerical results in Section 8 (e.g. estimated ground state energy vs. theoretical values given in [41]) seem to indicate that in our examples the bias under scrambled Sobol' sequences is very small and has no practical relevance.

One clear disadvantage of WUS against Mc-MC or Importance Sampling for problems with large regions of relative low values of the integrands is that with WUS we sample over the entire unit cube $[0, 1]^d$ uniformly, thus the method is dependent on how we transformed the problem to the unit cube. In contrast, Mc-MC Importance Sampling based techniques for models in high-energy or statistical physics usually focus on characteristic or important regions of the integrands aiming to sample directly from the underlying distribution of the problem, using in this way only the most relevant sample points.

8. Numerical experiments

We consider for our numerical tests the *quantum mechanical harmonic and anharmonic oscillator* in the *path integral approach* as described in Section 2. For definiteness we repeat here the expression for the action of the system:

$$S(\mathbf{x}) = \frac{a}{2}\sum_{i=1}^{d}\left(\frac{M_0}{a^2}(x_{i+1}-x_i)^2 + \mu^2 x_i^2 + 2\lambda x_i^4\right). \tag{19}$$

We investigate the two observable functions

$$O_1(\mathbf{x}) = \frac{1}{d}\sum_{i=1}^{d} x_i^2, \qquad O_2(\mathbf{x}) = \frac{1}{d}\sum_{i=1}^{d} x_i^4,$$

using the notation $\langle X^2\rangle, \langle X^4\rangle$ for $\langle O_1(\mathbf{x})\rangle, \langle O_2(\mathbf{x})\rangle$ in our tests. In addition, we will look at the ground state energy E_0 which, by virtue of the virial theorem, is related to O_1 and O_2 by $E_0 = \mu^2 O_1 + 3\lambda O_2 + \frac{\mu^4}{16}$.

Furthermore, we provide a program (obtainable from the CPC Program Library) for the QMC simulation of the (an)harmonic oscillator together with this article, and give a description of the usage of the program in the Appendix.

8.1. Harmonic oscillator

For the harmonic oscillator we can apply immediately the direct sampling approach described in Sections 3 and 7 for calculating estimates of observables $O(.)$ by setting

$$f_1 = O(A\Phi^{-1}(\boldsymbol{z})), \qquad f_2 = 1$$

in (18). The matrix A is a square root of C, the covariance matrix of the variables x_i, appearing in the action if it is expressed as a bi-linear form: $S(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T C^{-1}\mathbf{x}$, written explicitly as

$$C_{ij}^{-1} = \frac{2M_0}{a}\left[u\delta_{ij} - \frac{1}{2}\left(\delta_{i+1\,j} + \delta_{i\,j+1}\right)\right], \quad u = 1 + \frac{a^2\mu^2}{2M_0}. \tag{20}$$

Different factorizations, namely Cholesky and PCA (principal component analysis) have been tried out. The PCA based factorization turned out to perform better in our tests, which is the reason why we will only show results for this method. Note that, independently of which factorization we have chosen for C, for the case of the harmonic oscillator we sample directly from a Gaussian distribution and the considered observables functions $O(.)$ are just multivariate polynomials of low degree. Thus, the effective dimension in the superposition sense of the resulting non-constant integrand $f_1 = O(A\Phi^{-1}(\boldsymbol{z}))$ is upper bounded by the highest degree of the polynomials defining the observables $O(.)$ (this is true because the ANOVA decomposition is known to retain a *minimal* representation [42]). Therefore the problem has intrinsic low-effective dimension in the superposition sense, and it is expected that (randomized) QMC outperforms MC in this case. The PCA factorization seems to achieve further improvements since it can reduce, in addition, the effective dimension in the truncation sense. This is usually the case for Gaussian integrands considered in mathematical finance, involving a covariance matrix with rapid decaying eigenvalues (see [7,36]). The PCA factorization can be explicitly obtained for circulant Toeplitz matrices and the matrix–vector products can be efficiently computed by means of the fast Fourier transform. Given that the covariance matrix C is circulant Toeplitz, we have that $C = G\Lambda G^T$, with $G := \mathrm{Re}(F) + \mathrm{Im}(F)$,

$$(F)_{kl} = \frac{1}{\sqrt{d}}e^{-\frac{2\pi i}{d}kl} \tag{21}$$

being the Fourier matrix and Λ the diagonal matrix of positive eigenvalues (Lemma 4 in [43]). Thus $A = G\Lambda^{\frac{1}{2}}$ is a factorization of C, and in this case one can follow a recipe for generating normals with randomized QMC based on the discrete Fourier transform and using fast Fourier transform (FFT) techniques as described in [43]:

1. Generate a randomized QMC point $\tilde{\boldsymbol{z}}$.
2. Compute $\tilde{\mathbf{y}} = \Phi^{-1}(\tilde{\boldsymbol{z}})$.
3. Compute $\tilde{\mathbf{w}} = (\sqrt{\beta_1}\tilde{y}_{\pi^{-1}(1)}, \ldots, \sqrt{\beta_d}\tilde{y}_{\pi^{-1}(d)})$, where
$$\beta_j = \left[\frac{2M_0}{a}\left(u - \cos(2\pi j/d)\right)\right]^{-1}, \quad 1 \le j \le d \tag{22}$$
are the eigenvalues in the diagonal matrix Λ, and $\pi(.)$ is a fixed permutation of the variables.
4. Compute $\tilde{\mathbf{v}} = \mathrm{FFT}(\tilde{\mathbf{w}})$.
5. Take $\tilde{\mathbf{x}} = \mathrm{Re}(\tilde{\mathbf{v}}) + \mathrm{Im}(\tilde{\mathbf{v}})$ as the resulting point sample.

It is (strongly) recommended to fix first the permutation $\pi(.)$ such that $(\beta_{\pi(j)})_{j=1}^d$ are in non-increasing order, and this permutation was taken in our experiments. If this permutation of variables does not lead to satisfactory results, the analysis described in Section 6.2 can be carried out to investigate if a possible different permutation leads to more effective dimension reduction and better results.

In the ordinary Mc-MC approximation, we used the Mersenne Twister [13] pseudo-random number generator. We note in

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

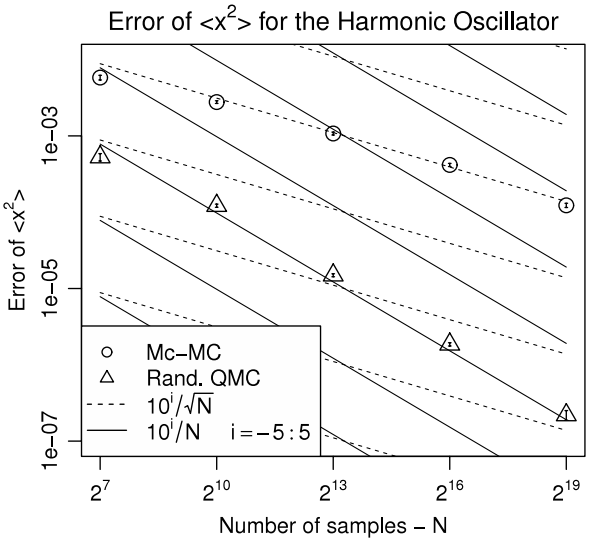

Fig. 1. The error of $\langle X^2\rangle$ in dependence on the number of samples N. The parameters here were chosen as $\lambda = 0$ (harmonic oscillator), $d = 51$, $M_0 = 0.5$ and $\mu^2 = 2.0$. The error of the error was obtained by repeating the numerical experiment 30 times, see also the text.

passing that the Mc-MC samples were generated in exactly the same way as described above for randomized QMC with the only difference that in step one Mc-MC points were generated according to the Gaussian measure of the harmonic oscillator. This corresponds to a heatbath algorithm, where all variables x_i are updated at the same time and with a reweighting procedure in the anharmonic case. For the QMC tests, we use the Sobol' sequences described in Section 4.1 from [25], with the random scrambling technique proposed by Matoušek [31]. The error of $\langle X^2\rangle$ was obtained by scrambling 10 times the QMC sequence and making 10 runs of an Mc-MC simulation (with different seeds). This procedure is repeated 30 times in both cases to obtain the error of the error. From the results shown in Fig. 1, we can see a scaling of the errors $N^{-1/2}$ for Mc-MC and N^{-1} for randomized QMC, for large N. Although this example is trivial, it was our first successful application of the QMC approach in a physical lattice model and motivated us to pass on to more complicated models.

8.2. Anharmonic oscillator

The WUS (reweighting) approach was also used for the problem of the anharmonic oscillator to estimate $\langle X^4\rangle$, $\langle X^2\rangle$ and the ground state energy of the system (E_0). With the anharmonic term in action, the probability distribution function (PDF) of the variables x_i is of non-Gaussian nature and hence becomes very complicated. This makes it very hard to generate the samples directly from the PDF of the anharmonic oscillator. Instead of this, we consider the WUS method with samples originated from an importance density (see (4.6) in [39]) of Gaussian form, leaving the anharmonic term and a fraction of the harmonic term as part of the functions f_1 and f_2 in (18). This change is in part necessary because now we choose $\mu^2 < 0$ in our test cases, and this choice breaks down the positive definiteness of the matrix C from the harmonic oscillator. Thus, we select a new covariance matrix $C^\star$ for the Gaussian samples, but we keep the sampling strategy for the x_i essentially unchanged as compared to the harmonic oscillator. The resulting weight functions in (18) are given by

$$f_1(\boldsymbol{z}) = O(A^\star \Phi^{-1}(\boldsymbol{z}))f_2(\boldsymbol{z}),$$
$$f_2(\boldsymbol{z}) = e^{-\sum_i \left(a\left(\frac{\mu^2-\mu^2_{\rm sim}}{2}\right)(A_i^\star \Phi^{-1}(\boldsymbol{z}))^2 + a\lambda (A_i^\star \Phi^{-1}(\boldsymbol{z}))^4\right)}, \tag{23}$$

where $A_i^\star$ stands for the i-row of the factorization matrix $A^\star$ of $C^\star$, and $\mu^2_{\rm sim} \geq 0$. As can be seen from the resulting weight function f_2, we have chosen the simple strategy of calibrating the diagonal of the new covariance matrix $C^\star$ by the use of a parameter $\mu^2_{\rm sim} \geq 0$. Besides the requirement of positivity on $C^\star$, one is free in the choice of the parameter $\mu_{\rm sim}$. We choose to follow the spirit of importance sampling by tuning $\mu_{\rm sim}$ to a value that reduces the fluctuations of the weights f_1 and f_2 as much as possible. The samples based on the tuned parameter $\mu_{\rm sim}$ lead us to observable averages with less variances and therefore smaller errors. Nevertheless, it seems quite difficult to find an optimal criterion for the selection of the functions f_1, f_2 and the parameter $\mu_{\rm sim}$ which leads to the best possible error behavior. At the moment we have to determine these quantities empirically and leave systematic investigations to the future.

Further, it is important to note that the PCA factorization during the generation of the Gaussian samples plays a mayor role for an efficient reduction of the effective dimension (see [33]) of the problem. For the parameters listed below, we estimated the effective dimensions in truncation sense d_t of the functions (23) to be close to 20 (for a 99% variance concentration), for estimating the integrals described in (12) with the dimension of the original system up to $d = 1000$. Thus, we observe a drastic reduction of dimensionality.

On the other hand, we found that the effective dimensions in the truncation sense of the functions (23) depend very strongly on the parameter $T = da$, i.e. the physical time extent of the system, and seems not sensible to the real dimension d. We found that for small T-values, say $T < 0.2$, the selected parameters $\mu_{\rm sim}$ and PCA lead to an effective dimension reduction, with d_t close to 4. In this case randomized QMC exhibited an N^{-1} error scaling. The situation changed by increasing the T-values. For $T = 1.5$ the effective dimensions d_t significantly increased to be close to 20. The exhibited error scaling was $N^{-\alpha}$ with $\alpha \approx 0.75$ for this case. Tests with values of $T \geq 5$ indicate that the simulations become more and more difficult in the sense that one needs more and more samples to achieve the same accuracy of an observable as compared to estimates at $T = 1.5$. Thus, in such situations the overlap of the sampling points with the functions f_i in (18) seem to be too small to reduce the fluctuations sufficiently. It seems that for this problem there exists some kind of *transition range* for the observed error scaling using randomized QMC in dependence of the time extent T, starting with a convergence rate N^{-1} for T-values less than 0.2 and decreasing to the poor convergence rate $N^{-1/2}$ for T-values higher than 5. However, we are presently exploring a more general approach for selecting a good T-dependent matrix $C^\star$ (resp. $A^\star$ in (23)) in the sampling procedure to improve the situation for larger values of T. This question and the relation to the corresponding effective dimension deserves a detailed study, in particular when more realistic models are considered. However, such an investigation, although being very interesting, goes beyond the scope of the present paper.

Nevertheless, for our numerical experiments, the parameters were set to $M_0 = 0.5, \lambda = 1.0, \mu^2 = -16$. In the two tests the lattice spacing a was adjusted such that T was kept fixed. The tuned value of $\mu^2_{\rm sim}$ generally depends on all physical parameters of the system and in particular on a. Thus, we have to adjust also $\mu_{\rm sim}$ when the lattice spacing a is changed. In particular, we set $a = 0.015$ and $\mu^2_{\rm sim} = 0.176$ for $d = 100$, whereas for $d = 1000$ $a = 0.0015$ and $\mu^2_{\rm sim} = 0.2$ was chosen. The error analysis of $\langle X^2\rangle$ and $\langle X^4\rangle$ was carried through in the same way as described for the harmonic oscillator test case discussed in the last Section 8.1. We show in Fig. 2 the error of $\langle X^2\rangle$ and E_0 as a function of the number of samples. In addition, we represent by the dashed line in Fig. 2 a fit to the data for the computed errors using the

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

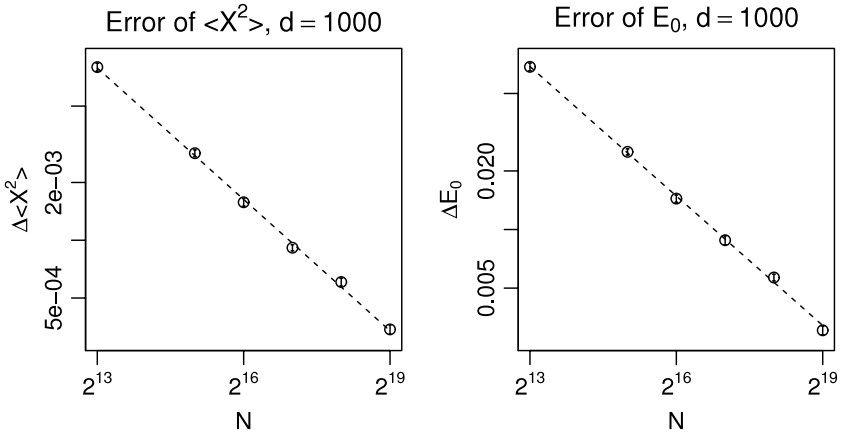

Fig. 2. The error of the observables $\langle X^2\rangle$ and E_0 as a function of the number of samples in a double logarithmic graph. The error of the error was obtained by performing 30 repetitions of the experiment with parameters chosen as $\lambda = 1.0$, $\mu^2 = -16$, $a = 0.0015$ and $d = 1000$. For the sample generation randomly scrambled Sobol' (Rand. QMC) was used with 2^{13}, 2^{15}, 2^{16}, 2^{17}, 2^{18} and 2^{19} points. The dashed line shows the fit to the data points using a fit function $\log(\Delta\langle O\rangle) \sim \log(C) + \alpha \log(N)$. The fitted exponents are $\alpha = -0.758(14)$ for $\langle X^2\rangle$ and $\alpha = -0.737(13)$ for E_0, see also Table 1.

Table 1
The results for fit parameters of the error scaling for the observables considered, i.e. X^2, X^4 and E_0, where the used fit function takes the form error $= CN^{\alpha}$, see Eq. (24). We also provide the χ^2 values as well as the number of degrees of freedom in the fit (dof).

	O	α	$\log C$	χ^2/dof
$d = 100$	X^2	−0.763(8)	2.0(1)	7.9/6
	X^4	−0.758(8)	4.0(1)	13.2/6
	E_0	−0.737(9)	4.0(1)	8.3/6
$d = 1000$	X^2	−0.758(14)	2.0(2)	5.0/4
	X^4	−0.755(14)	4.0(2)	5.7/4
	E_0	−0.737(13)	4.0(2)	4.0/4

formula

$$\log(\text{Error}(\langle O\rangle)) = \log C + \alpha \log N; \quad O = \{x^2, x^4, E_0\}, \tag{24}$$

with C and α left as free parameters. From this analysis we can obtain a quantitative determination of the exponent of the error scaling. The results for the fit parameters are listed in Table 1.

As can be inferred from Table 1, in the case of the anharmonic oscillator the error scaling exponent is only $\alpha \approx 0.76$. However, this still constitutes a much improved error scaling compared to a Mc-MC simulation with a corresponding large gain in the number of required samples to reach a desired accuracy. Moreover, the value of α is consistent for all observables considered here and independent of the dimension of the problem d, a finding which is clearly encouraging.

We finally mention that the resulting estimates of the ground state energy for $T = 1.5$ match, to at least two significant digits, the theoretical value $E_0 = 3.863$ calculated in [41], namely $\hat{E}_0 = 3.857 \pm 0.004$ for $d = 100$ and $\hat{E}_0 = 3.862 \pm 0.004$ for $d = 1000$.

9. Concluding remarks

In this article we have performed a first application of QMC methods to Euclidean lattice models. The goal was to see, whether QMC algorithms provide also in the case of non-Gaussian systems an improved error scaling behavior with respect to Markov-chain Monte Carlo methods. As a prototype system, we have considered the quantum mechanical oscillator discretized on a Euclidean time lattice, both in its harmonic (Gaussian) form as well as adding a non-Gaussian quartic term (anharmonic oscillator). For the harmonic oscillator we found a large-N (N being the number of sample points) improved error behavior, i.e. $\sim N^{-1}$ for (randomized) QMC and $\sim N^{-1/2}$ for Mc-MC.

The main result of our investigation is that also for the anharmonic oscillator, which is a non-Gaussian problem, the QMC approach leads to a significant improvement of the error scaling N^{α} with $\alpha \approx -0.76$, see Table 1 for the exact values of α for different observables and different physical situations.

Further, we found that the accessible range of $T = 1.5$ values gives already estimates of the ground state energy, compatible (within errors) with the theoretical prediction (valid in the limit $T \to \infty$ and $a \to 0$). For the case that the improved error scaling and the mild dependence on the lattice spacing a found here will also be present in more elaborate models, QMC methods have the potential to become very valuable in the future. On the other hand we observed that the applicability of the WUS (reweighting) approach seems to be limited by the physical time extent $T = da$ of the system. For values of $T \leq 1.5$ the error falls below the percent level within the investigated number of samples. For increasing values of T the error is continuously growing and at $T = 5$ we can only obtain a relative error of ≈15% for 2^{19} samples. For larger T values we expect the error to become even larger leading eventually for very large T to a situation where a meaningful evaluation of the considered quantities is no longer possible. This behavior and the relation to the effective dimension of the problem clearly needs an understanding and a dedicated investigation in the future—in particular, when more realistic models are considered.

It is clear that the quantum mechanical systems considered here are rather simple models and that there is still a long way to go if generic quantum field theories, especially gauge theories, are to be studied. Nevertheless, it is very reassuring that we find an improved error scaling behavior in the case of a quartic potential and hence a non-Gaussian system. This promising result is certainly a strong motivation for studying further QMC methods in lattice field theories and statistical mechanics.

Acknowledgments

The presented research cooperation has been supported by the Center of Computational Sciences Adlershof (CCSA), Berlin, Germany. The authors wish to express their gratitude to Alan Genz (Washington State University) and Frances Kuo (University of New South Wales, Sydney) for inspiring comments and conversations, which helped to develop the work in this article. Frances Kuo collaborated with us during her visit to the Humboldt-University Berlin in 2011. A.N., K.J. and M.M.-P. acknowledge financial support by the DFG-funded corroborative research center SFB/TR9. K. J. was supported in part by the Cyprus Research Promotion Founda-

tion under contract ΠΡΟΣΕΛΚΥΣΗ/ΕΜΠΕΙΡΟΣ/0311/16. We would like to thank specially the reviewers for their suggestions and comments which help us to improve the article.

Appendix. C++ program for the QMC simulation of the (an)harmonic oscillator

We provide a C++ program for the QMC simulation of the (an)harmonic oscillator at the following URL (needs to be provided by CCP). In the following we give a short description for the usage of the program. The program implements the suggested algorithm for the evaluation of the path integral of the harmonic and anharmonic oscillator in the QMC approach, except that the present implementation applies random digital shifts instead of random scramblings to generate different sobol sequences. Eventually, both methods should lead to very similar results.

Copyright information may be obtained from the file "README" in the package. The package is equipped with a standard "Makefile" and a cmake input file "CMakeList.txt".

A.1. Prerequisites

The only external dependency is the FFTW library version 3 which can be obtained from http://www.fftw.org. The FFTW library offers an efficient implementation of the Hartley transform.
If the library is installed in a non-standard path of your PC, you can adjust the variable "FFTWDIR" in the very beginning of the make file or via the environment variable "FFTWDIR" when using "CMakeLists.txt".

A.2. Building

Simply run 'make' when using "Makefile" or use 'ccmake <path to package source>' in an empty directory followed by a 'make'.

A.3. Parameters

Having built the executable 'qmc_quartic_reweight' you can run the program, preferably in a new empty directory, and may pass the following parameters:

Parameter	Meaning
-N <Integer>	Number of dimensions
-k <Integer>	Number of samples per estimation
-c <Integer>	Number of configurations written out to a file
-S <Integer>	Max. time separation for correlator
-a <Float>	Lattice spacing a
-M <Float>	Particle mass M_0
-m <Float>	μ^2_{sim}
-u <Float>	μ^2
-l <Float>	λ
-Q <Path to file>	File containing directions numbers

Files with direction numbers can be obtained from Frances Kuo's page
http://web.maths.unsw.edu.au/~fkuo/sobol/index.html. The program produces 10 estimations, each with the given number of samples. Output is written to files of the form <prefix>_s1_N<# dimensions>_a<a>_M0<M0>_musq <mu^2_sim>_l<lambda>_J0.000000.csv. The result of each estimation is stored in the file with the prefix "obs_macro". The first column contains the estimated value of $\langle x^2 \rangle$ and the second column contains $\langle x^4 \rangle$. Successive runs of the program with the same parameters will append 10 more estimates. Correspondingly, 30 runs should suffice to produce the statistics we used in this work.

References

[1] Christof Gattringer, Christian B. Lang, Quantum chromodynamics on the lattice, Lecture Notes in Phys. 788 (2010) 1–211.
[2] I. Montvay, G. Münster, Quantum Fields on A Lattice, in: Cambridge Monographs on Mathematical Physics, Cambridge University Press, 1994.
[3] Zoltan Fodor, Christian Hoelbling, Light hadron masses from lattice QCD, Rev. Modern Phys. 84 (2012) 449.
[4] John Bulava, Philipp Gerhold, Karl Jansen, Jim Kallarackal, Bastian Knippschild, et al., Higgs–Yukawa model in chirally-invariant lattice field theory, Adv. High Energy Phys. 2013 (2013) 875612.
[5] Pierre L'Ecuyer, Christiane Lemieux, Recent advances in randomized quasi-Monte Carlo methods, in: Moshe Dror, Pierre L'Ecuyer, Ferenc Szidarovszky (Eds.), Modeling Uncertainty, in: International Series in Operations Research & Management Science, vol. 46, Springer, US, 2005, pp. 419–474.
[6] F. Kuo, Ch. Schwab, I. Sloan, Quasi-Monte Carlo methods for high-dimensional integration: the standard (weighted Hilbert space) setting and beyond, ANZIAM J. 53 (01) (2012).
[7] Paul Glasserman, Monte Carlo Methods in Financial Engineering, in: Applications of Mathematics (New York), vol. 53, Springer-Verlag, New York, 2004, Stochastic Modelling and Applied Probability.
[8] P. Jäckel, Monte Carlo Methods in Finance, in: The Wiley Finance Series, Wiley, 2002.
[9] M. Creutz, B.A. Freedman, A statistical approach to quantum mechanics, Ann. Phys. 132 (1981) 427–462. Available online in Michael Creutz's publication list under http://thy.phy.bnl.gov/creutz/mypubs/pubs.html, visited June 17, 2013.
[10] W. Bietenholz, U. Gerber, M. Pepe, U.-J. Wiese, Topological lattice actions, J. High Energy Phys. 1012 (2010) 020.
[11] K. Jansen, H. Leovey, A. Nube, A. Griewank, M. Mueller-Preussker, A first look at quasi-Monte Carlo for lattice field theory problems, 2012. Available online under http://arxiv.org/pdf/1211.4388.pdf.
[12] M.J.D. Powell, J. Swann, Weighted uniform sampling—a Monte Carlo technique for reducing variance, J. Inst. Math. Appl. 2 (1966) 228–236.
[13] Makoto Matsumoto, Takuji Nishimura, Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Trans. Model. Comput. Simul. 8 (1) (1998) 3–30.
[14] Erich Novak, Henryk Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals, in: EMS Tracts in Mathematics, vol. 12, European Mathematical Society (EMS), Zürich, 2010.
[15] Josef Dick, Friedrich Pillichshammer, Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, New York, NY, USA, 2010.
[16] Russel E. Caflisch, Monte Carlo and quasi-Monte Carlo methods, Acta Numer. 7 (1998) 1–49. 0.
[17] Fred J. Hickernell, A generalized discrepancy and quadrature error bound, Math. Comp. 67 (1998) 299–322.
[18] Ian H. Sloan, Henryk Wozniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? J. Complexity 14 (1997) 1–33.
[19] F.Y. Kuo, Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces, J. Complexity 19 (3) (2003) 301–320.
[20] Xiaoqun Wang, Strong tractability of multivariate integration using quasi-Monte Carlo algorithms, Math. Comp. 72 (242) (2003) 823–838.
[21] A. Griewank, L. Lehmann, H. Leovey, M. Zilberman, Automatic evaluations of cross-derivatives, Math. Comp. 83 (2014) 251–274.
[22] Harald Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, in: CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
[23] I.M. Sobol', The distribution of points in a cube and the approximate evaluation of integrals, USSR Comput. Math. Math. Phys. 7 (4) (1967) 86–112.
[24] Paul Bratley, Bennett L. Fox, Algorithm 659: implementing Sobol's quasirandom sequence generator, ACM Trans. Math. Software 14 (1) (1988) 88–100.
[25] Stephen Joe, Frances Y. Kuo, Remark on algorithm 659: implementing Sobol's quasirandom sequence generator, ACM Trans. Math. Software 29 (1) (2003) 49–57.
[26] Stephen Joe, Frances Y. Kuo, Constructing Sobol sequences with better two-dimensional projections, SIAM J. Sci. Comput. 30 (5) (2008) 2635–2654.
[27] Ilya M. Sobol', Danil Asotsky, Alexander Kreinin, Sergei Kucherenko, Construction and comparison of high-dimensional Sobol' generators, Wilmott 2011 (56) (2011) 64–79.
[28] A.B. Owen, Randomly permuted (t, m, s)-nets and (t, s)-sequences, in: H. Niederreiter, P.J.-S. Shiue (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, in: Lecture Notes in Statistics, vol. 106, Springer-Verlag, 1995, pp. 299–317.

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011

[29] Art B. Owen, Monte Carlo variance of scrambled net quadrature, SIAM J. Numer. Anal. 34 (5) (1997) 1884–1910.
[30] Art B. Owen, Local antithetic sampling with scrambled nets, Ann. Statist. 36 (2008) 2319–2343.
[31] J. Matoušěk, On the L_2-discrepancy for anchored boxes, J. Complexity 14 (1998) 527–556.
[32] Shu Tezuka, Henri Faure, I-binomial scrambling of digital nets and sequences, J. Complexity 19 (6) (2003) 744–757.
[33] R.E. Caflisch, W. Morokoff, A. Owen, Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension, J. Comput. Finance 1 (1) (1997) 27–46.
[34] I.M. Sobol', Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul. 55 (1–3) (2001) 271–280. The Second IMACS Seminar on Monte Carlo Methods (Varna, 1999).
[35] Xiaoqun Wang, Kai-Tai Fang, The effective dimension and quasi-Monte Carlo integration, J. Complexity 19 (2) (2003) 101–124.
[36] X. Wang, I. Sloan, Why are high-dimensional finance problems often of low effective dimension? SIAM J. Sci. Comput. 27 (1) (2005) 159–183.
[37] I.M. Sobol', S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, Math. Comput. Simul. 79 (10) (2009) 3009–3017.
[38] Andreas Griewank, Andrea Walther, Evaluating Derivatives, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. Principles and Techniques of Algorithmic Differentiation.
[39] Jerome Spanier, Earl H. Maize, Quasi-random methods for estimating integrals using relatively small samples, SIAM Rev. 36 (1) (1994) 18–44.
[40] Russel E. Caflisch, Bradley Moskowitz, Modified Monte Carlo methods using quasi-random sequences, in: Lecture Notes in Statistics, vol. 106, Springer-Verlag, 1995, pp. 1–16.
[41] Richard Blankenbecler, Thomas A. DeGrand, R.L. Sugar, Moment method for eigenvalues and expectation values, Phys. Rev. D21 (1980) 1055.
[42] Frances Y. Kuo, Ian H. Sloan, Grzegorz W. Wasilkowski, Henryk Woźniakowski, On decompositions of multivariate functions, Math. Comp. 79 (270) (2010) 953–966.
[43] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl, I.H. Sloan, Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications, J. Comput. Phys. 230 (10) (2011) 3668–3694.

Please cite this article in press as: K. Jansen, et al., Quasi-Monte Carlo methods for lattice systems: A first look, Computer Physics Communications (2013), http://dx.doi.org/10.1016/j.cpc.2013.10.011