Applicability of Quasi-Monte Carlo for lattice systems

Andreas Ammon\textsuperscript{1,2}, Tobias Hartung\textsuperscript{1,2}, Karl Jansen\textsuperscript{2}, Hernan Leovey\textsuperscript{3}, Andreas Griewank\textsuperscript{3}, Michael Müller-Preussker\textsuperscript{1}

\textsuperscript{1}Humboldt-University Berlin, Physics Department, \textsuperscript{2}NIC, DESY, Zeuthen, \textsuperscript{3}Humboldt-University Berlin, Mathematics Department

July 29 2013
31\textsuperscript{st} International Symposium on Lattice Field Theory

corresponding publication: \texttt{arXiv:1302.6419}, (in referee process with CPC)
Outline

Motivation

The (An)Harmonic Oscillator on the Lattice

Results

Outlook/Conclusions
typical lattice problem:

\[ Z = \int \mathcal{D}x \ e^{-S[x]} ; \quad x = (x_1, \ldots, x_d) \]  

(1)

\[ \langle O \rangle = Z^{-1} \int \mathcal{D}x \ e^{-S[x]} \ O[x] \]  

(2)

- stochastic approximation through Markov chain Monte Carlo methods: Metropolis algorithm, HMC, ...
- finite Markov chain: \( x_1, \ldots, x_N \rightarrow N \) samples of \( O: O_1, \ldots, O_N \)
- \( O_i \) random variables with variance \( \sigma_O^2 \)
- estimate \( \langle O \rangle = \frac{1}{N} \sum_{i=1}^{N} O_i \) has standard error

\[ \Delta \langle O \rangle = \frac{\sigma_O}{\sqrt{N}} \]

- need 100 times more statistics to get additional digit of precision
- past improvements: reduce \( \sigma_O \) and auto-correlation
- Improved error scaling would be highly desirable!
quasi-Monte Carlo (QMC) is an approach to improve the asymptotic error behaviour
see for example F. Kuo, Ch. Schwab and I. Sloan, 2012 [KSS12]

- construction of deterministic low-discrepancy point-sets in arbitrary many dimensions
- low-discrepancy $\rightarrow$ “more uniform” (see below)
- promises $N^{-1}$ asymptotic error behaviour for integrands with certain properties (e.g. Gaussian)
- $\rightarrow$ **two** times more digits with the same number of samples!!
- applied successfully to financial problems (see bibliography)
QMC point sets are more uniform

How does an actual uniform sampling in two dimensions look like?

**Example:** 512 two-dimensional pseudo-random points

- sample 512 points
- introduce grid of $8 \times 8$ equal squares
- count number of points in each square
- count occurrence of 1, 2, ... points in a square (histogram of histogram)

$\approx$ Poisson distribution with $\lambda = \bar{n} = 8$

uneven sampling $\rightarrow$ larger stochastic error
QMC point set (2d Sobol samples):

- each square contains same number of points → delta distribution
- even coverage
- less stochastic fluctuations
- simulate effect of higher statistics with much less samples
- in this sense QMC is exactly what we want
- randomisation possible (RQMC) w/o changing properties → practical error estimation

**Figure:** 512 uniform 2d Sobol points
lattice action (see “Creutz and Freedman” [CF81]):

\[ S = a \sum_{i=1}^{d} \left( \frac{M_0}{2} \frac{(x_{i+1} - x_i)^2}{a^2} + \frac{\mu^2}{2} x_i^2 + \lambda x_i^4 \right) ; \quad x_{d+1} = x_1 \quad (p.b.) \]

- \( M_0 \) ... particle mass
- \( \mu^2 = M_0 \omega^2 \) ... frequency/spring constant
- \( a \) ... lattice spacing
- \( d \) ... number of lattice sites \( \rightarrow T = da \) ... time extent

- \( \lambda = 0 \rightarrow \) harmonic oscillator
- \( \lambda > 0 \rightarrow \) anharmonic oscillator, \( \mu^2 < 0 \rightarrow \) double well potential

**Figure:** two cases for the anharmonic potential
primary observables

\[ \langle x^2 \rangle = \langle \frac{1}{d} \sum_i x_i^2 \rangle \]  \hspace{1cm} (3)

\[ \langle x^4 \rangle = \langle \frac{1}{d} \sum_i x_i^4 \rangle \]  \hspace{1cm} (4)

\[ \langle x_k x_{k+j} \rangle = \langle \frac{1}{d} \sum_i x_i x_{i+j} \rangle \ldots \text{correlator} \]  \hspace{1cm} (5)

derived quantities

\[ E_0 = 3\lambda \langle x^4 \rangle + \mu^2 \langle x^2 \rangle + \frac{\mu^4}{16} \]  \hspace{1cm} (6)

\[ E_1 - E_0 = \text{energy gap from correlator fit} \]  \hspace{1cm} (7)

theoretically known for \( a \to 0 \), \( T = da \to \infty \) (iterative method)
Blankenbecler, DeGrand and Sugar 1980 [BDS80]
Experiment I: Harmonic Oscillator \((\lambda = 0, \mu^2 > 0)\)

Partition function can be written as multivariate Gaussian integral

\[
Z = \int \mathcal{D}x \exp \left( -\frac{1}{2}x^t C^{-1}x \right)
\]

\[
C^{-1} = \frac{2M_0}{a} \left( (1 + \frac{a^2\mu^2}{2M_0})\delta_{ij} - \frac{1}{2} (\delta_{ij+1} + \delta_{ij-1}) \right)
\]

covariance matrix: \( C = SDS^t \rightarrow D = \text{diag}(\beta_1, \ldots, \beta_d) \quad \beta_i \in \mathbb{R}^+ \)

\[
x = Sw \quad \Rightarrow \quad Z \rightarrow \int \mathcal{D}w \exp \left( -\sum_i \frac{1}{2\beta_i}w_i^2 \right)
\]

→ Sampling algorithm

- generate uniform \( z \in [0, 1]^d \) (pseudo random / QMC)
- \( w_i = \sqrt{\beta_i};\Phi^{-1}(z_i), \Phi^{-1} \ldots \) inverse standard normal CDF
  - ordering of eigenvalues \( \beta_1 > \beta_2 > \ldots > \beta_d \) when using QMC
  - like ordering of importance \( z_1 > z_2 > \ldots > z_d \)
- \( x_i = S_{ij}w_j \) (Hartley transformation, involutive: \( S = S^{-1} = S^t \))
harmonic oscillator results

parameters: $\mu^2 = 2.0$, $M_0 = 0.5$, $a = 0.5$ & $d = 100$

Figure: left: asymptotic error behaviour of MC/QMC, right: fit of QMC error $\sim N^\alpha$

- QMC at work
- trivial, but successful application to physical problem
Experiment II: Anharmonic Oscillator ($\lambda = 1$, $\mu^2 < 0$)

direct sampling not possible because of anharmonic part of the potential
→ do reweighting

$$Z = \int \mathcal{D}x \exp \left( -\frac{1}{2}x^t C^{-1}x - a\lambda \sum_i x_i^4 \right)$$

(C$^{-1}$ indefinite ($\mu^2 < 0$) → define $C_{\text{sim}}^{-1} = \frac{2M_0}{a} \left( (1 + \mu_{\text{sim}}^2 \frac{a^2}{2M_0})\delta_{ij} - \frac{1}{2} (\delta_{ij+1} + \delta_{ij-1}) \right)$

with $\mu_{\text{sim}}^2 > 0$, arbitrary insert the "productive 0"

$$Z = \int \mathcal{D}x \exp \left( -\frac{1}{2}x^t C_{\text{sim}}^{-1}x - \frac{1}{2}x^t(C^{-1} - C_{\text{sim}}^{-1})x - a\lambda \sum_i x_i^4 \right)$$

$$= \int \mathcal{D}x \ e^{-\frac{1}{2}x^t C_{\text{sim}}^{-1}x} W(x)$$

→ sampling like harmonic oscillator but with $C \rightarrow C_{\text{sim}}$ observable estimation from samples $(x^j)_{j=1,...,N}$:

$$\langle O \rangle \approx \frac{\sum_j W(x^j)O(x^j)}{\sum_j W(x^j)} \quad W(x) = e^{-\frac{1}{2}...}$$
numerical results/anharmonic oscillator

parameters: \( M_0 = 0.5 \), \( a = 0.015 \), \( \mu^2 = -16 \)

fit: \( \Delta O \sim CN^\alpha \)

<table>
<thead>
<tr>
<th></th>
<th>( O )</th>
<th>( \alpha )</th>
<th>log ( C )</th>
<th>( \chi^2/dof )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d = 100 )</td>
<td>( X^2 )</td>
<td>-0.763(8)</td>
<td>2.0(1)</td>
<td>7.9 / 6</td>
</tr>
<tr>
<td></td>
<td>( X^4 )</td>
<td>-0.758(8)</td>
<td>4.0(1)</td>
<td>13.2 / 6</td>
</tr>
<tr>
<td></td>
<td>( E_0 )</td>
<td>-0.737(9)</td>
<td>4.0(1)</td>
<td>8.3 / 6</td>
</tr>
<tr>
<td>( d = 1000 )</td>
<td>( X^2 )</td>
<td>-0.758(14)</td>
<td>2.0(2)</td>
<td>5.0 / 4</td>
</tr>
<tr>
<td></td>
<td>( X^4 )</td>
<td>-0.755(14)</td>
<td>4.0(2)</td>
<td>5.7 / 4</td>
</tr>
<tr>
<td></td>
<td>( E_0 )</td>
<td>-0.737(13)</td>
<td>4.0(2)</td>
<td>4.0 / 4</td>
</tr>
</tbody>
</table>

energy gap

- asymptotic behaviour of correlator
- non-trivial observable
- not possible to detect on present parameter setup ($T$ too small)
- changed $\mu^2 = -16 \rightarrow \mu^2 = -4$
- energy gap: $0.0015 \rightarrow 1.576$

result obtained for $d = 100$, $N = 2^5, 2^8, 2^{11}, 2^{14}$ and 400 Sobol’ sequences each:

$$\alpha = -0.735(13)$$

(Tobias Hartung, 2013, personal communication)
outlook & conclusions

- harmonic oscillator: QMC works perfectly (as expected)
- anharmonic oscillator: significantly improved error scaling $\rightarrow N^{-\frac{3}{4}}$

remaining questions:
  - Why do we observe this $N^{-\frac{3}{4}}$ behaviour??
  - further improvements by generalised choice of $C_{sim}$?
  - other, possibly non-Gaussian, sampling methods

- next step: one-dimensional spin model in cosine discretisation

$$S[\phi] = la \sum_{i} -\frac{1}{a^2} \cos(\phi_{i+1} - \phi_i)$$ (14)

study $\chi_Q$ (topological susceptibility) and $\Delta E = E_1 - E_0$ (energy gap)


