Adaptive Monte Carlo Algorithm in $d = 1$

A. Balaž, A. Belić and A. Bogojević

Institute of Physics, P.O.B. 57, Belgrade 11001, Yugoslavia

Abstract

We develop an efficient adaptive Monte Carlo algorithm for calculation of partition functions in low dimensions, and demonstrate its use on the case of the anharmonic q^4 oscillator in $d = 1$. The trajectories needed to estimate the path integral are sampled from a correlated Gaussian centered at the average $\langle q \rangle$. The proposed scheme is a generalization of earlier work done in $d = 0$ [1].

1 Introduction

Path integrals play a central role in many areas of physics, most notably in statistical mechanics and quantum field theory where they represent the natural formalism in which to cast the theory [2, 3]. Unfortunately, except in rare cases, the only way to solve path integrals exactly is through numerical simulations. All numerical methods imply some sort of discretization in real or inverse space (space-time). This is usually done in two steps: the path integral is first approximated with a multiple integral which is then calculated using standard discretization. The most economic way for the calculation of these integrals is the Monte Carlo method which approximates the path integral with a sum of the contributions of a finite number of representative trajectories.

If we do not need to calculate the partition function, but only the averages of certain physical quantities, then the calculation can proceed using the efficient Metropolis $\mathrm{M(RT)}^2$ algorithm [4]. The calculation of the partition function is significantly more difficult however, and there still does not exist a generic efficient algorithm for its evaluation. For this reason we have initiated a series of investigations [5, 6, 7] whose goal is to develop just such a generic algorithm. In this paper we extend our previous results in $d = 0$ to calculations in $d = 1$ dimensions.

2 Formalism

Statistical mechanics in $d = 1$ is equivalent to (Euclidean) quantum mechanics. Having this in mind, we label our dynamical variables $q(\tau)$. We have investigated the anharmonic oscillator in a constant magnetic field B. The Hamiltonian of this system is

$$\mathcal{H}(q, \dot{q}) = \frac{1}{2}\,\dot{q}^2 + \frac{1}{2}\,q^2 + 16gq^4 \ . \tag{1}$$

It is well known [2, 3] that the partition function can be written as a path integral over all periodic trajectories $q(\tau) = q(\tau + \beta)$ of $\exp\left[-\int_0^\beta d\tau\,(\mathcal{H} - Bq)\right]$. In this paper we shall in fact study a related, and a bit more general object—the associated generating functional. In this case the path integral is taken over all trajectories that go from $q(0) = 0$ to $q(\beta) = a$. Using the Feynman-Kac discretization we divide the interval of propagation $(0,\ \beta)$ into N equal subintervals of length $\varepsilon = \beta/N$, and find [3]:

$$\begin{aligned} Z_N(a,\ \beta;\ g,\ B) \;&=\; \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\mathrm{d}q_1\cdots\mathrm{d}q_{N-1}}{(2\pi\varepsilon)^{N/2}}\exp\left\{-\varepsilon\sum_{n=0}^{N-1}\left[\frac{1}{2}\left(\frac{q_{n+1}-q_n}{\varepsilon}\right)^2 + \right.\right. \\ &\left.\left. +\ \frac{1}{2}\left(\frac{q_{n+1}+q_n}{2}\right)^2 + g(q_{n+1}+q_n)^4 - B\frac{q_{n+1}+q_n}{2}\right]\right\}\ . \end{aligned} \tag{2}$$

The most important point in calculating a general integral $\int \mathrm{d}^d x\ f(x)$ using the Monte Carlo method lies in the choice of the normalized distribution p in the expression

$$\int \mathrm{d}^d x\ f(x) = \int \mathrm{d}^d x\ \frac{f(x)}{p(x)}\,p(x) \equiv \left\langle \frac{f}{p}\right\rangle_p \approx \frac{1}{N_{mc}}\sum_{i=1}^{N_{mc}}\frac{f(x_i)}{p(x_i)}\ , \tag{3}$$

where N_{mc} represents the number of trajectories sampled in the Monte Carlo calculation, while x_i are the points sampled from the distribution p. The square of the error of this estimator is given [8] by the variance

$$\sigma^2_{f/p} = \frac{1}{N_{mc}-1}\left[\left\langle\left(\frac{f}{p}\right)^2\right\rangle_p - \left\langle\frac{f}{p}\right\rangle_p^2\right], \tag{4}$$

and the function p is chosen so as to minimize the error. In this paper we use the general Gaussian distribution

$$p(q_1,\cdots,q_{N-1}) = \frac{1}{\sqrt{(2\pi)^{N-1}\det A}}\exp\left\{-\frac{1}{2}\sum_{m,\ n=1}^{N-1}(q_m-\bar{q}_m)A_{mn}(q_n-\bar{q}_n)\right\}, \tag{5}$$

where A is a given symmetrical positive definite matrix, and $\bar{q}_m$ are appropriately chosen centers of the distribution. Distribution (5) is a direct generalization of the one we have used earlier in the $d = 0$ case [1]. The integrand in (2) can be well approximated within the class (5), so that the function f/p, after optimization of the Gaussian distribution, is almost constant, making the error small.

We have investigated two possible choices for $\bar{q}_m$. The mean field approximation suggests that we use $\bar{q}_m = q^{mf}(m\varepsilon)$, where $q^{mf}(\tau)$ minimizes the energy disregarding thermal fluctuations, i.e. is the solution of the equation of motion $\ddot{q} = q + 64gq^3 - B$. Another choice for $\bar{q}_m$ was motivated by our work in $d = 0$ dimension [1], where we have shown that it is more efficient to center the Gaussian at the magnetization $M = \langle q\rangle$.

This scheme was designated IGA for Improved Gaussian Approximation [5, 6, 7]. The straight forward generalization to $d = 1$ is to use $\bar{q}_m = M(m\varepsilon)$, where $M(\tau)$ is the thermal average value of $q(\tau)$. Unlike $q^{mf}(\tau)$, the magnetization $M(\tau)$ is not known beforehand. This problem was circumvented by employing an adaptive procedure as discussed in [1]. We start the simulation using the distribution p centered at $\bar{q} = q^{mf}$, and thereafter every N'_{mc} ($\ll N_{mc}$) Monte Carlo steps adjust $\bar{q}$ to the current estimate of $M = \langle q \rangle$. After a while the magnetization M settles, and a longer run is then performed to calculate the partition function with required precision.

Having made our choices for the centers of the Gaussians we still have to specify the matrix A. The simplest choice for A is to use the uncorrelated Gaussian $A_{mn} = \delta_{mn}/\sigma^2$, where the width σ is chosen to minimize the error. On the other hand, if A is chosen to cancel all the quadratic terms in the exponent of the function f/p (correlated Gaussian) then the distribution p contains the full dynamics of the linear harmonic oscillator. It is not surprising that the trajectories sampled from such a distribution have been found to be more representative than those generated from the uncorrelated Gaussian, i.e. that the corresponding errors are smaller. It should be noted that the choices of A and $\bar{q}$ are independent, since $\bar{q}$ only appears in the linear and constant terms in $\ln(f/p)$.

3 Results

The error associated with the use of the uncorrelated Gaussian grows exponentially with N, regardless of the choice of σ and $\bar{q}$. This effectively limits the numerical work to the modest values $N \lesssim 10$, making the extrapolation to $N \to \infty$ difficult. The correlated Gaussian scheme is much more efficient. In that case, the errors decrease with N, enabling us to easily work with $N \sim 100$.

The algorithm was verified by comparison with the $g = 0$ case which can be solved analytically. In the $g \neq 0$ case the estimates of the generating functional Z_N, obtained for various values of N, fit to a power series in $1/N$. In practice it is enough to stop at the quadratic term

$$Z_N = Z_{ex} + \frac{c_1}{N} + \frac{c_2}{N^2} \ . \tag{6}$$

This is shown in Fig. 1 (a) for the cases $a = \beta = g = 1$, $B = -1$ and $B = -10$. As we can see, the relative deviation of Z_N from the estimate Z_{ex} of the exact value (obtained by extrapolation to $N \to \infty$) is a smooth function. On the basis of this numerical evidence one can conclude that the continuum limit is well behaved, justifying the extrapolation procedure. All of the results for $g \neq 0$ have been obtained using this kind of extrapolation. Fig. 1 (b) shows how the errors decrease as N is increased. At all times one gets better results (smaller errors) by centering the Gaussians at the average values $M(\tau)$ rather than at mean field value $q^{mf}(\tau)$. This is in complete agreement with what has been found in

[1]. For large enough N, the errors ΔZ_N are also well approximated by a polynomial in $1/N$, as shown in Fig. 1 (b). The displayed results were obtained using $N_{mc} = 10^6$ trajectories.

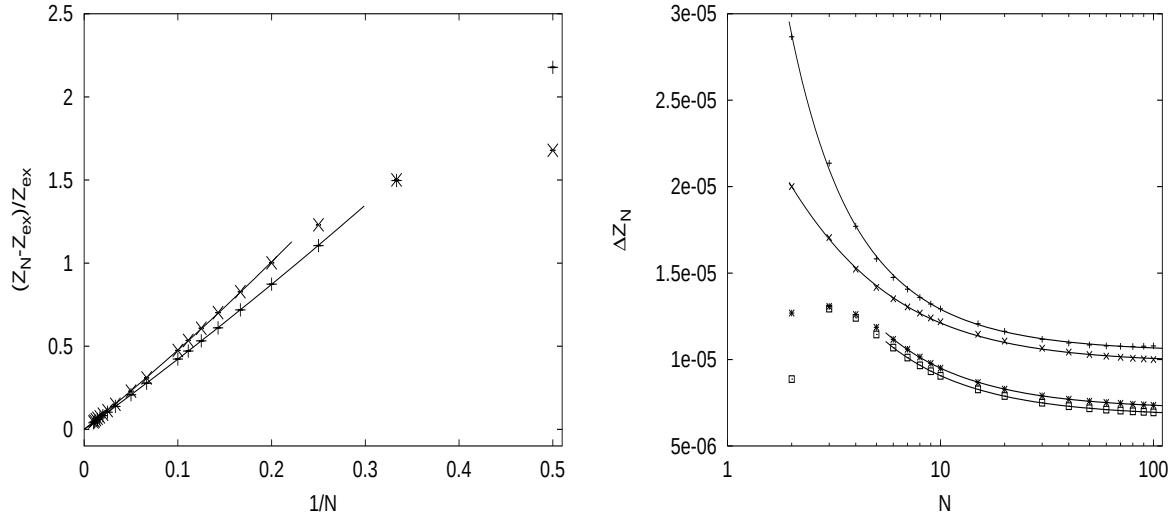

Figure 1: (a) The relative deviation of the estimates for the generating functional from the exact values for $a = \beta = g = 1$, $B = -10$ (top) and $B = -1$ (bottom). (b) The corresponding errors for (from top to bottom) $B = -1$, $\bar{q}_m = q^{mf}(m\varepsilon)$; $B = -1$, $\bar{q}_m = M(m\varepsilon)$; $B = -10$, $\bar{q}_m = q^{mf}(m\varepsilon)$; $B = -10$, $\bar{q}_m = M(m\varepsilon)$.

In this paper we have studied various strategies for generating trajectories needed to calculate the partition function in $d = 1$. We have shown that it is always more efficient to center the Gaussian distribution p on the magnetization M rather then on the mean field value. To illustrate the difference between the two, in Fig. 2 we show the trajectories used in calculating the generating functionals in the preceding cases. For $a = \beta = 1$ and $g = 0$, $B = -1$ the difference between q^{mf} and M is insignificant, while for the cases $g = 1$, $B = -1$ as well as $g = 1$, $B = -10$ there is a sizeable difference between them, leading to a substantial increase in efficiency.

We believe that there is still more room for improvement in the efficiency of the presented algorithm. The choice of the centers $\bar{q}$ is already optimal, hence further improvements may come through a full optimization of matrix A. An even more promising venue for improvement lies in the use of alternate discretization schemes. We plan to pursue both of these possibilities in future work.

The numerical simulations presented in this work were done at the Institute of Physics on a SGI Origin 2000 super-computer. We would like to acknowledge the kind help of the staff of the IPCF (Institute of Physics Computing Facilities). This work was financed in part by the Ministry of Science and Technology of the Republic of Serbia under research projects 01M01 and 01E15.

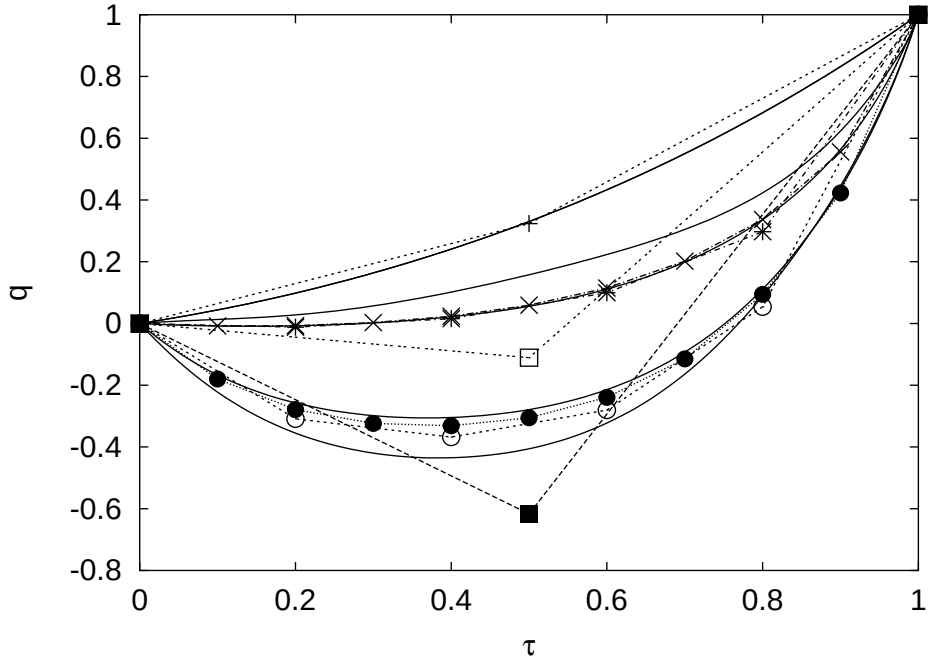

Figure 2: Comparison of the mean field solution $q^{mf}(\tau)$ with the magnetization $M(\tau)$ for $N = 2$, 5, 10, and 100 for employed sets of parameters: $a = \beta = 1$ and $B = -1$, $g = 0$ (top group), $B = -1$, $g = 1$ (middle group) and $B = -10$, $g = 1$ (bottom group).

References

[1] A. Balaž, A. Belić, A. Bogojević, Phys. Low-Dim. Struct. **9/10** (1999), 149.

[2] J. W. Negele and H. Orland, *Quantum Many-Particle Systems*, Addison-Wesley, New York, 1988.

[3] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.

[4] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, J. Chem. Phys. **21** (1953), 1087.

[5] A. Balaž, A. Belić, A. Bogojević, SFIN A2 (1998).

[6] A. Balaž, A. Belić, A. Bogojević, Facta Universitatis **1** (1998), 113.

[7] A. Balaž, A. Belić, A. Bogojević, Phys. Low-Dim. Struct. **5/6** (1999), 1.

[8] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, Vol. 1: Basics, John Wiley and Sons, New York, 1986.