Commun. Comput. Phys.
doi: 10.4208/cicp.190313.171013a

Vol. **15**, No. 5, pp. 1352-1367
May 2014

The Lognormal Distribution and Quantum Monte Carlo Data

Mervlyn Moodley*

School of Chemistry and Physics, Quantum Research Group, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban, 4000, South Africa.

Received 19 March 2013; Accepted (in revised version) 17 October 2013

Communicated by Michel A. Van Hove

Available online 26 February 2014

Abstract. Quantum Monte Carlo data are often afflicted with distributions that resemble lognormal probability distributions and consequently their statistical analysis cannot be based on simple Gaussian assumptions. To this extent a method is introduced to estimate these distributions and thus give better estimates to errors associated with them. This method entails reconstructing the probability distribution of a set of data, with given mean and variance, that has been assumed to be lognormal prior to undergoing a blocking or renormalization transformation. In doing so, we perform a numerical evaluation of the renormalized sum of lognormal random variables. This technique is applied to a simple quantum model utilizing the single-thread Monte Carlo algorithm to estimate the ground state energy or dominant eigenvalue of a Hamiltonian matrix.

PACS: 02.70.Ss, 02.50.Cw, 02.50.Ng, 02.50.-r

Key words: Lognormal distribution, renormalization transformation, quantum Monte Carlo.

1 Introduction

Quantum Monte Carlo simulations utilizing the technique of multiplying weights together often give spurious results when one calculates expectation values of operators. Often, one is faced with a dilemma when having to choose a final estimate together with its corresponding error estimate from a set of estimators converging to the exact result. This may arise as a consequence of the estimators developing a distribution that is somewhat different from the Gaussian distribution. Correct statistical inference is based on the assumption that data under consideration adheres to a specific distribution. If this distribution is incorrect, the results obtained after statistical analysis may be invalid [1].

*Corresponding author. *Email address:* moodleym2@ukzn.ac.za (M. Moodley)

With respect to quantum Monte Carlo applications where one is interested in the statistical iteration of some operator, Hetherington [2] observed that the probability distribution of the estimators depends on the number of Monte Carlo iterations. In fact, as shown in this paper, the estimators exhibit a lognormal distribution that has been block-transformed a number of times. This attribute is inherited from the distribution of the product of weights associated with importance sampling. By the central limit theorem, the lognormal distribution should approach the Gaussian limit for a sufficiently large number of block transformations. The estimators however are sometimes not blocked sufficiently often to have reached the Gaussian limit but they do resemble the Gaussian distribution with slight deviations. It would therefore be incorrect to assume that standard statistical analysis, giving the average plus or minus one standard error to be within a 68% confidence interval, is appropriate here.

In the context of statistical and probability theory, a block or renormalization transformation as described in this work, corresponds to the renormalized sum of identically independent random variables. In this paper we consider the sum of lognormal random variables. Sums of lognormal random variables appear in many branches of science [3,4] and finance [5] but most prominently in the field of communications [6,7]. For a historic perspective of finding the distribution of sums of lognormally distributed random variables, see [8]. The difficulty in evaluating these sums of distributions analytically is due to the fact that the characteristic function of the lognormal distribution is not know in closed form and as a result approximation methods are used. Many of these approximations are based on approximating the sum of lognormal variables by another lognormal variable [6,7,9–11]. Other methods of approximation have also been introduced by Beaulieu et al. [12,13]. In the results presented in this work, the sums of lognormal distributions were evaluated simply by using the trapezoidal rule which produced excellent results without resorting to more elaborate numerical integration techniques [14].

The paper is organized as follows. In Section 2 we recall the standard statistical methods applied to a set of data if normality is assumed. In Section 3, we describe a method of calculating the number of times a set of data has undergone a block transformation by relating the cumulants of the blocked data to that of the original data. From this, a recursion relation results which relates successive blocked cumulants. Section 4 focuses on the lognormal distribution and here we construct the block transformed lognormal distribution numerically by first calculating the characteristic function and then Fourier transforming to obtain the probability distribution. A recipe is given in Section 5 to construct the probability distribution of a set of data, with given mean and variance, that has been assumed to be lognormal prior to blocking. By constructing this distribution we are able to give better estimates of the standard errors corresponding to a desired confidence interval. As an application of these methods we consider data obtained from using the single-thread Monte Carlo technique to estimate the groundstate energy (dominant eigenvalue) of a 3×3 symmetric Hamiltonian matrix. This is described in Section 6. Here we consider data with small ensemble sizes that do not ideally converge to an expected value. By constructing the probability distributions of these data, we show that the er-