Analysis of the convergence of the $1/t$ and Wang-Landau algorithms in the calculation of multidimensional integrals

R. E. Belardinelli,* S. Manzi, and V. D. Pereyra
Departamento de Física, Instistuto Nacional de Física Aplicada, Universidad Nacional de San Luis, CONICET, Chacabuco 917,5700 San Luis, Argentina

(Received 29 May 2008; published 5 December 2008)

In this Brief Report, the convergence of the $1/t$ and Wang-Landau algorithms in the calculation of multidimensional numerical integrals is analyzed. Both simulation methods are applied to a wide variety of integrals without restrictions in one, two, and higher dimensions. The efficiency and accuracy of both algorithms are determined by the dynamical behavior of the errors between the exact and the calculated values of the integral. It is observed that the time dependence of the error calculated with the $1/t$ algorithm varies as $N^{-1/2}$ [with N the number of Monte Carlo (MC) trials], in quantitative agreement with the simple sampling Monte Carlo method. In contrast, the error calculated with the Wang-Landau algorithm saturates in time, evidencing the nonconvergence of this method. The sources of error for both methods are also determined.

DOI: 10.1103/PhysRevE.78.067701 PACS number(s): 02.60.Jh, 05.10.Ln, 02.70.Uu

It is well known that the Wang-Landau (WL) algorithm is one of the most refreshing variations of the Monte Carlo simulation methods [1]. Its effectiveness is based on the simplicity and versatility of the algorithm in calculating the density of states $g(E)$ with high accuracy [where $g(E)$ represents the number of all possible states or configurations for an energy level E of a given physical system]. The algorithm has been successfully used in many problems of statistical physics, biophysics, and other fields. There have been several papers in recent years dealing with improvements and sophisticated implementations of the WL iterative process in discrete and continuous systems (see references in [2]).

However, the most controversial point in the application of the WL and other variations of the algorithm is the saturation of the error between the calculated and the real $g(E)$. In fact, for long enough time, the error approaches a constant value. This problem was first evidenced in Refs. [3,4]. On the other hand, other authors [5–9] have studied the accuracy, efficiency, and convergence of the WL algorithm. Some of them [5,8,9] have demonstrated the convergence of the WL algorithm by different arguments. In particular, Zhou and Bhatt [5] have presented a mathematical analysis of the WL algorithm. They gave proof of the convergence and the sources of errors for the WL algorithm as well as strategies for improvement.

The error saturation and the algorithm convergence are certainly two contradictory results. Moreover, in Ref. [10], an analytical demonstration of the nonconvergence of the WL algorithm has been presented. The authors have deduced that in those methods in which the refinement parameter decreases faster than $1/t$ (with t the Monte Carlo time) the calculated density of states reaches a constant value for long times, and therefore the error saturates. To overcome this limitation, they introduced a modified version of the WL algorithm in which the refinement parameter is scaled down as $1/t$ instead of exponentially [11]. The $1/t$ algorithm has been successfully applied to several statistical systems [2,10–13].

*rbelar@unsl.edu.ar

The saturation of the error and consequently the nonconvergence of the WL algorithm have been demonstrated for a discrete system [10], namely, the Ising model. However, the mathematical arguments about the source of error seem to be more general, and can be extended to all algorithms that consider a refinement parameter decreasing faster than $1/t$. In all these cases, the nonconvergence of the method can be guaranteed.

In addition, for continuous systems, there are few simulations where a comparison has been made with the exact density of states [17]. The reason for that may be attributed to the nonavailability of exact results for any nontrivial system having a continuous energy spectrum.

Recently, Li *et al.* [14] reported a new application of the WL algorithm to the simplest continuous system, namely, numerical integration. The basic idea is to relate the density of states $g(E)$ to the distribution $g(y)$, where $g(y)$ represents the fraction of the integration domain ($[a,b]$ in one dimension) that lies within a certain interval $[y,y+dy]$. This idea was first proposed by Tröster and Dellago [15]. Simultaneously, Liang [16] developed a generalization of the WL algorithm to continuous systems.

In principle, the WL method of integration presents various advantages over the conventional Monte Carlo (MC) integration scheme (see Ref. [14]). Its potential appears for ill-behaved integrals and for higher-dimensional integration problems, since, in general, the random walk remains one dimensional. However, the accuracy of the WL algorithm is poorer than in the simple sampling MC method, at least for one and two dimensions, as is shown in Ref. [14].

The aim of this work is to study the convergence of the $1/t$ and Wang-Landau algorithms in the calculation of multidimensional numerical integrals. The dynamical behavior and sources of error in both algorithms are also analyzed.

The $1/t$ algorithm is adapted to the numerical calculation of multidimensional integrals. The basic idea is as follows. To evaluate the definite integral $\int_a^b y(x)dx$ it is necessary to determine the proportion of the integration domain that lies within a certain interval $[y,y+dy]$, i.e., the measure $\{x|x \in [a,b], y \leqslant y(x) \leqslant y+dy\}$. The distribution $g(y)$ can be gen-

erated by measuring this fraction. Provided that the lower bound $y_{\min}$ and the upper bound $y_{\max}$ of the integral are known, the integral can be approximated by

$$I = \int_a^b y(x)dx \approx \sum_{y_{\min}}^{y_{\max}} g(y)y. \tag{1}$$

To build the distribution $g(y)$, the interval $[y_{\max}-y_{\min}]$ is divided into $L=[y_{\max}-y_{\min}]/dy$ segments. The MC time is defined as $t=N/L$, where N is the number of Monte Carlo trials. The MC time is related to the size of the y domain. In order to compare the algorithms, all the quantities will be related to the number of MC trials, N.

The algorithm is as follows.

(i) Choose a value of x_i at random and calculate the corresponding value of y_i; then set $S(y)=0$ for all values of y, $F_0=1$, and fix F_{final} or equivalently $t_{\text{final}}=1/F_{\text{final}}$.

(ii) Choose at random a value of x_f and change y_i to y_f according to the probability given by $P(y_i \to y_f) = \min\{1, e^{[S(y_i)-S(y_f)]}\}$.

(iii) Increase $S(y) \to S(y)+F_k$.

(iv) After some fixed steps (i.e., 1000 MC times) check that all the sites y corresponding to the same F_k will be visited by the random walker at least once; then refine $F_k = F_k/2$.

(v) If $F_{k+1} \leqslant 1/t=L/N$ then make $F_{k+1}=F(t)=1/t=L/N$. Hereafter $F(t)$ is updated at each MC time.

The step (iv) is not used for the rest of the experiment. As soon as the refinement parameter takes $F=1/t$ functionality, it goes down independently of the number of times that the site y is visited by the random walker.

(vi) If $t>t_{\text{final}}$ $[F(t)<F_{\text{final}}]$ then the process stops. Otherwise, go to (ii).

Other auxiliary histograms are not necessary, with the exception of $S(y)=\ln[g(y)]$. In spite of this, as soon as the algorithm converges, the histogram corresponding to the number of visits becomes flat. The relative distribution function $g(y)$ provided by the algorithm must be appropriately normalized as in Ref. [14].

In order to compare the methods, the following integrals are calculated by using the $1/t$, Wang-Landau, and simple sampling (SS) algorithms:

$$I_{1\text{D}} = \int_{-2}^{2} (x^5 - 4x^3 + x^2 - x)\sin(4x)dx, \tag{2}$$

$$I_{2\text{D}} = \int_{-1}^{1}\int_{-1}^{1} (x_1^6 - x_1x_2^3 + x_1^2x_2 + 2x_1)\sin(4x_1+1) \times \cos(4x_2)dx_1dx_2, \tag{3}$$

$$I_{\pi} = \frac{1}{4}\int_0^1 \sqrt{1-x^2}dx = \pi, \tag{4}$$

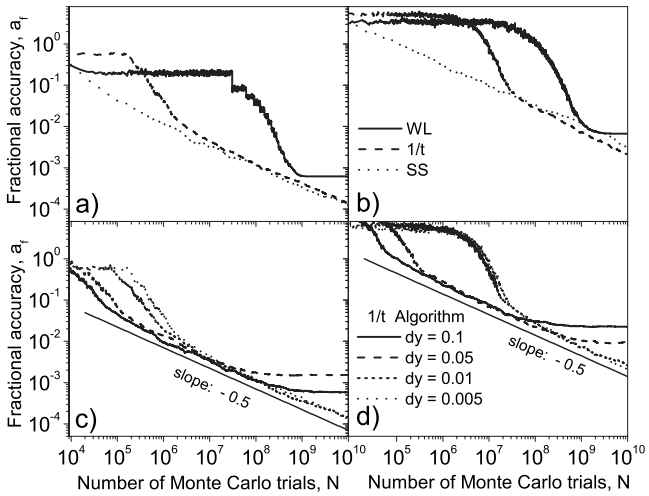

FIG. 1. Dynamical behavior of fractional accuracy a_f calculated by means of Wang-Landau, $1/t$, and simple sampling Monte Carlo integration: (a) for $I_{1\text{D}}$, with $dy=0.005$ and $L=3066$; (b) for $I_{2\text{D}}$, with $dy=0.005$ and $L=1213$. (c) and (d) show the effect of bin width on $I_{1\text{D}}$ and $I_{2\text{D}}$.

$$I_{n\text{D}} = \int_0^1\int_0^1 \cdots \int_0^1 \prod_{i=1}^{i=n} \cos(ix_i)dx_1dx_2\cdots dx_n = \prod_{i=1}^{i=n} \frac{\sin(ix_i)}{i}. \tag{5}$$

The first two integrals were introduced in Ref. [14]. These integrals have no particular physical or mathematical significance. However, they are very useful in comparing the convergence of the three MC algorithms.

The fractional accuracy is defined as

$$a_f(N) = |(I_{\text{MC}}(N) - I_{\text{exact}})/I_{\text{exact}}|, \tag{6}$$

where $I_{\text{MC}}(N)$ and I_{exact} denote the numerical estimate from the MC simulation and the exact value of the integral, respectively.

Hereafter, all the errors estimated are obtained from 100 independent simulations and the WL calculations have been made by controlling the histogram every 10 000 MC steps with a flatness criterion $p=0.9$.

In Figs. 1(a) and 1(b) the fractional accuracy a_f for $I_{1\text{D}}$ and $I_{2\text{D}}$ as a function of the number of MC trials is plotted using the three algorithms. The fractional accuracy a_f is in close agreement, following a $1/\sqrt{N}$ dependence for the $1/t$ and SS algorithms; while for the WL algorithm it saturates at $N \approx 10^9$. Therefore, for the WL algorithm, a_f does not scale as $N^{-1/2}$, as is observed.

The bin width certainly introduces a systematic error in all the algorithms that use the distribution function $g(y)$ as strategy to calculate numerical integrals. The effect of dy on the error for the $1/t$ algorithm is shown in Figs. 1(c) and 1(d) [the integrals and conditions are the same as in Figs. 1(a) and 1(b)].

The dependence of the error on the bin width will be determined by the characteristic of the function $y(x)$. In fact, one should expect that for smaller $dy \neq 0$ the saturation will occur at longer times, i.e., for the two-dimensional integral

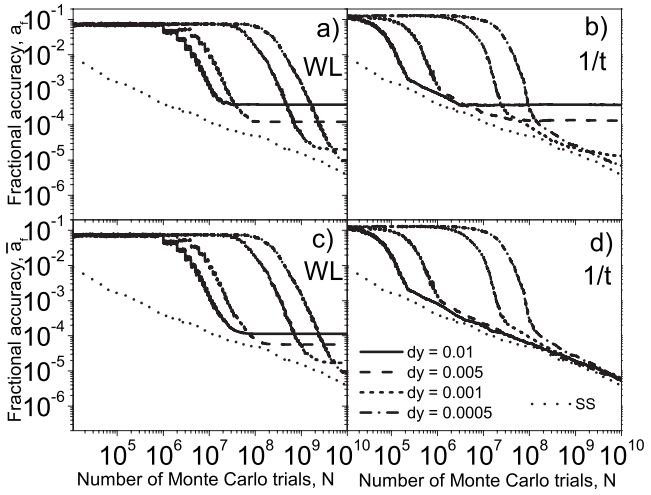

FIG. 2. Comparison of the dynamical behavior of fractional accuracies a_f and $\bar{a}_f$ for different values of bin width dy, in the estimation of the integral I_π.

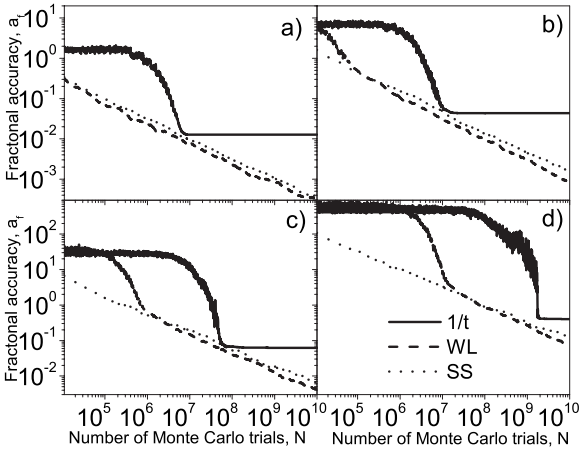

FIG. 3. Dynamical behavior of fractional accuracy a_f for n-dimensional integral $I_{n\mathrm{D}}$ with (a) $n=3$, (b) $n=4$, (c) $n=5$, and (d) $n=6$, using the three algorithms. In both cases, the WL and $1/t$ algorithms the bin width is $dy=0.05$.

shown in Fig. 1(d). However, this is not always valid, as shown in Fig. 1(c), where the error for $dy=0.05$ saturates before corresponding to $dy=0.1$.

To analyze the influence of the bin width in the behavior of the error in both WL and $1/t$ algorithms, the integral I_π is considered. Using the expression of $f(y)$ given in Eq. (4), which is a one-to-one function in the interval (0,1), one can obtain the exact value of $g(y)$ as

$$g_{\mathrm{ex}}(y) = \sqrt{1-(y+dy)^2} - \sqrt{1-y^2}. \tag{7}$$

Then, for a given value of dy, the corresponding exact value of the integral is $I_\pi(dy)_{\mathrm{ex}} = \Sigma_{y_{\min}}^{y_{\max}} g_{\mathrm{ex}}(y)y$. For example, for $dy=0.01$, the value of the integral is $I_\pi(0.01)_{\mathrm{ex}} = 3.140\,417\,032\ldots$. A new fractional accuracy related to the value of dy is defined as $\bar{a}_f(N,dy) = |I_{MC}(N) - I_\pi(dy)_{\mathrm{ex}}/I_\pi(dy)_{\mathrm{ex}}|$.

The fractional accuracy, a_f, versus the MC trials for different values of the bin width, dy, is shown in Fig. 2(a) (WL) and Fig. 2(b) $(1/t)$. In both cases the error saturates. However, the sources of the saturation obey different causes. To demonstrate this fact, the fractional accuracy, $\bar{a}_f$, for both algorithms is shown in Fig. 2(c)) (WL) and Fig. 2(d) $(1/t)$. In these cases for each value of dy, the exact value of $I_\pi(dy)_{\mathrm{ex}}$ is obtained. While for the WL results the error still saturates [Fig. 2(c)], in $1/t$ no saturation occurs [Fig. 2(d)]. Thus, for a given dy the error approaches asymptotically the exact value of the integral $I_\pi(dy)_{\mathrm{ex}}$ (different from π). This is clear evidence of the convergence of the $1/t$ algorithm in all discrete systems. The source of saturation in the $1/t$ algorithm in a continuous system may be attributed exclusively to the bin width dy.

Finally, the behavior of both algorithms in the calculation of higher-dimensional integrals is analyzed. In Figs. 3(a)–3(d), the integrals $I_{n\mathrm{D}}$ with $n=3,4,5,6$ are plotted. In all cases $dy=0.05$ and $L=40$. The error for the WL method saturates, demonstrating the nonconvergence of the algorithm in all cases. On the other hand, the $1/t$ and SS algorithms are in close agreement.

The numerical estimates of the statistical error for all the integrals obtained by using the three algorithms are given in Table I. The number of final Monte Carlo trials per run is $N_{\mathrm{final}}=10^{10}$. The values of the estimates shown in Table I confirm that the $1/t$ is more accurate than the WL algorithm for all the integrals in any dimension.

The numerical integration is an excellent demonstration to prove the convergence of the algorithms for different reasons. (i) In many cases the integrals can be solved analytically, and therefore the dynamical behavior of the error can

TABLE I. Numerical estimates of integrals calculated by using the Wang-Landau, $1/t$, and simple sampling Monte Carlo methods. Results and error estimates are obtained from 100 independent simulations.

Integral	WL algorithm	$1/t$ algorithm	Simple MC method	Exact
$I_{1\mathrm{D}}$	1.635580(285)	1.635617(27)	1.635752(23)	1.63564436296…
$I_{2\mathrm{D}}$	−0.0179671(152)	−0.0179790(51)	−0.0179841(68)	−0.01797992646…
I_π	3.1415799(35)	3.1415819(23)	3.1415920(15)	3.14159265358…
$I_{3\mathrm{D}}$	0.01801608(2989)	0.01799079(46)	0.01799666(78)	0.01799626791…
$I_{4\mathrm{D}}$	−0.00339585(1798)	−0.00340505(54)	−0.00340506(55)	0.00340490511…
$I_{5\mathrm{D}}$	0.00065747(505)	0.00065298(47)	0.00065225(49)	0.00065300923…
$I_{6\mathrm{D}}$	−0.000031919(3288)	−0.00003003(15)	−0.00003110(38)	−0.00003041015….

be easily checked. (ii) For some one-to-one functions, the exact expression for the distribution function $g(y)$ can be obtained. This is an advantage over physical systems, where only in a few cases is the exact density of states known, particularly for continuous systems. (iii) The initial and final states are not correlated with their neighborhood, namely, a given initial state can be changed to any other final state in the integration domain of $y(x)$. In this sense, if a given algorithm cannot converge appropriately in the calculation of a numerical integral, it will be more difficult to do it in those physical systems where the initial and final states are strongly correlated.

In continuous systems, as for numerical integrals, the bin width introduces a saturation of the error also in the $1/t$ algorithm. However, when the continuum is approached by a discrete lattice and the corresponding value of the distribution function can be obtained exactly, the calculation approaches asymptotically the exact value of the integral without error saturation.

The behavior of the error in the $1/t$ algorithm is in close agreement with the SS result, following the $1/\sqrt{N}$ law. Moreover, the $1/t$ algorithm can be used as a reference in the calculation of the density of states in physical systems because up to now there are no other methods that can calculate the error below the limiting curve $1/\sqrt{N}$. In contrast, the behavior of the error in the WL algorithm, for a single value of the parameters dy and p, does not follow the $1/\sqrt{N}$ at any time.

Very recently, in Ref. [18], Zhou and Su introduced a new mathematical support for the $1/t$ algorithm.

In summary, the dynamical behavior of the $1/t$ algorithm in a multidimensional numerical integration is analyzed, with the conclusion that the algorithm is always convergent in discrete systems. For a continuous model, the only source of error saturation is the grid discretization. On the other hand, further evidence of the nonconvergence of the WL algorithm is given.

We thank Dr. M. R. Gomez for reading the manuscript. This work was partially supported by the CONICET (Argentina).

[1] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[2] N. G. Fytas, A. Malakis, and I. A. Hadjiagapiou, J. Stat. Mech.: Theory Exp. (2008), P11009.
[3] Q. Yan and J. J. de Pablo, Phys. Rev. Lett. **90**, 035701 (2003).
[4] J.-S. Wang and R. H. Swendsen, J. Stat. Phys. **106**, 245 (2002).
[5] C. Zhou and R. N. Bhatt, Phys. Rev. E **72**, 025701(R) (2005).
[6] H. K. Lee, Y. Okabe, and D. P. Landau, Comput. Phys. Commun. **175**, 36 (2006).
[7] S. Sinha and S. K. Roy, e-print arXiv:0711.1031, Phys. Lett. A (to be published).
[8] D. J. Earl and M. W. Deem, J. Phys. Chem. B **109**, 6701 (2005).
[9] A. N. Morozov and S. H. Lin, Phys. Rev. E **76**, 026701 (2007).
[10] R. E. Belardinelli and V. D. Pereyra, J. Chem. Phys. **127**, 184105 (2007).
[11] R. E. Belardinelli and V. D. Pereyra, Phys. Rev. E **75**, 046701 (2007).
[12] P. Ojeda, A. Londono, N.-Y. Chen, and M. Garcia, Biophys. J. (to be published).
[13] A. Malakis, A. Nihat Berker, I. A. Hadjiagapiou, and N. G. Fytas, e-print arXiv:0809.4241.
[14] Y. W. Li, T. Wüst, D. P. Landau, and H. Q. Lin, Comput. Phys. Commun. **177**, 524 (2007).
[15] A. Tröster and C. Dellago, Phys. Rev. E **71**, 066705 (2005).
[16] F. Liang, J. Am. Stat. Assoc. **100**, 1311 (2005).
[17] K. Mukhopadhyay, N. Ghoshal, and S. K. Roy, Phys. Lett. A **372**, 3369 (2008).
[18] C. Zhou and Jia Su, Phys. Rev. E **78**, 046705 (2008).