

# A scalable method for *ab initio* computation of free energies in nanoscale systems

M. Eisenbach\*, C.-G. Zhou†, D. M. Nicholson\*, G. Brown‡, J. Larkin§, and T. C. Schulthess¶

\*Oak Ridge National Laboratory, Oak Ridge TN 37831-6164

†Quantitative Research, J. P. Morgan Chase and Company, 12th Floor, 277 Park Ave., New York, New York 10017, USA

‡Department of Physics, Florida State University, Tallahassee, Florida 32306

§Cray Incorporated, Oak Ridge TN 37831-6008

¶Institute for Theoretical Physics and Swiss National Supercomputer Center, ETH Zurich, CH-8093 Zurich

**Abstract**—Calculating the thermodynamics of nanoscale systems presents challenges in the simultaneous treatment of the electronic structure, which determines the interactions between atoms, and the statistical fluctuations that become ever more important at shorter length scales. Here we present a highly scalable method that combines *ab initio* electronic structure techniques, we use the Locally Self-Consistent Multiple Scattering (LSMS) technique, with the Wang-Landau (WL) algorithm to compute free energies and other thermodynamic properties of nanoscale systems. The combined WL-LSMS code is targeted to the study of nanomagnetic systems that have anywhere from about one hundred to a few thousand atoms. The code scales very well on the Cray XT5 system at ORNL, sustaining 1.03 Petaflop/s in double precision on 147,464 cores.

## CONTENTS

<b>I</b>	<b>Introduction</b>	1
<b>II</b>	<b>The hybrid Wang-Landau / LSMS method to compute free energies</b>	2
II-A	The Wang-Landau method . . . . .	2
II-B	The Locally Self-consistent Multiple Scattering (LSMS) method . . . . .	3
II-C	Hybrid WL-LSMS algorithm for <i>ab initio</i> statistical physics . . . . .	4
<b>III</b>	<b>Simulation results: estimating the transition temperature for iron</b>	5
<b>IV</b>	<b>Performance of the WL-LSMS method at scale</b>	6
<b>V</b>	<b>Summary and outlook</b>	7
<b>VI</b>	<b>Acknowledgment</b>	8

## I. INTRODUCTION

Many sophisticated implementations of Kohn-Sham based density functional theory (2; 4) have evolved over the past decades (6), and with the spread of teraflop/s scale computers, large *ab initio* electronic structure calculations have become a matter of routine and have reached scales where they allow the direct study of nano-structures (3). These studies typically are concerned with the atomic and magnetic structure and dynamics, as well as optical or transport properties.

From a thermodynamic point of view, practically all large-scale *ab initio* simulations are concerned with the computation

of the internal energy of a system, which consists of the potential and kinetic energies that the atoms of the system possess by virtue of their respective positions and velocities. However, in order to explore most physical properties of a nanoscale system at finite temperatures, it will be necessary to compute the free energy, which takes into account entropic and thermal effects of the system and its environment. While potential and kinetic energies are straightforward to compute in atomistic simulations that yield positions and velocities, free energies remain a challenge since they require higher-order information.

In principle, the free energy surface of a nano-scale system can be extracted from long-time-scale molecular or spin dynamics simulations. If one traces the evolution of a system over long enough time scales, such that all of phase space is explored by the simulation, thermodynamic potentials can be computed. But even with the largest computers conceivable today, such simulations will at best be possible with empirical models, and studies that are based on *ab initio* electronic structure methods will remain out of reach even on exaflop/s scale computer systems. For systems with corrugated energy surfaces, molecular or spin dynamics simulations tend to be "stuck" in local energy minima and unrealistically long simulations would be required to sample large enough parts of phase space.

Several techniques have been designed to specifically deal with the problem of exploring complex free energy landscapes. Of particular interest are *meta dynamics* (5) and the *Wang-Landau* Monte Carlo method (11). In *meta dynamics* a set of dynamic equations are solved in the space of collective variables that are of interest for the free energy calculations. In the *Wang-Landau* method a random walk is performed explicitly in the space of collective variables. In both methods the simulated system evolves in such a way that avoids areas of phase space that have already been visited. In a certain sense, both methods provide an intelligent way to overcome the time-scale dilemma with massively parallel computers. Molecular and spin dynamics simulation techniques are serial in nature, and will therefore be very difficult to improve drastically in the coming decade, where computer performance will evolve primarily through enhancements in concurrency. *Meta dynamics* and the *Wang-Landau* method can be straightforwardly parallelized to thousands of concurrent replicas (in

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

the case of meta dynamics) or walkers (in the case of Wang-Landau). Since the individual replicas or walkers represent simulations that already today are scaled to hundreds or even thousands of processing cores, it is easy to see how these simulations will scale to millions or even tens of millions of processing elements. Studies of nanoscale systems based on these methods are therefore ideally suited for exa-scale computer systems. But scalability of the Wang-Landau method is only one of its advantages. When the underlying energy functional is temperature independent, as is often the case and in particular for the magnetic systems studies here, the Wang-Landau method will allow one to compute the thermodynamic quantities at all temperature in a single simulation, a property that reduces the total cost of the computation dramatically compared to standard Metropolis Monte Carlo or molecular and spin dynamics.

In the current paper we present the first attempt to combine the Wang-Landau method with a first-principles electronic structure technique, the Locally Self-Consistent Multiple Scattering (LSMS) (13) method, to compute from first principles the free energy of a magnetic system. We show that already on a "small" petascale computer system with 150,000 cores, the free energy of elemental Fe can be computed with sufficient accuracy, in order to determine the ferromagnetic transition temperature. In fact, the simulations we present here represent the first *ab initio* study of this kind that do not have to resort to a mean field approximation or empirical models.

In the next section, we introduce the hybrid WL-LSMS method and give a detailed discussion of the parallel implementation. In section III we present a simulation study of the magnetism in Fe, and we show that even in these preliminary simulations length scales are reached that are representative of the relevant scale in studies of magnetic nanoparticles. In section IV we discuss the performance of the WL-LSMS code at scale on the Cray XT system Jaguar at Oak Ridge National Laboratory. The code shows near optimal scaling and sustains a double precision performance in excess of one petaflop/s on 150,000 cores. We close the discussion in section V with a summary and outlook.

## II. THE HYBRID WANG-LANDAU / LSMS METHOD TO COMPUTE FREE ENERGIES

The method we devise here is targeted to systems with non-collinear atomic moments. In typical magnetic materials, the atomic moments tend to be collinear and the orientation of the magnetization varies at length scales that are much longer than the interatomic distance. Only above the magnetic ordering temperatures, where the system shows no long-range order, do non-collinearities at the atomic scale appear. However, in nano-particles the magnetic structure in the surface region can be complex and non-collinear, even at temperatures where the moments in the core region of the particle are collinear. In small particles, where the surface region contains a significant fraction of the particle volume, the surface induced disorder leads to an enhancement of the entropy and to interesting

magnetic properties of the nanoparticles that deviate significantly from the corresponding bulk material. Understanding these nanoscale effects will require simulations of the kind we develop here, since the atomic-scale magnetic structure in nanoparticles cannot be accessed experimentally.

In the hybrid Wang-Landau/LSMS method we use an extended version of the Wang-Landau algorithm (15) that is capable of effectively treating systems with continuous degrees of freedom and thermodynamic potentials of several collective variables. This method has been used already to study temperature dependent free energy barriers for the magnetic switching in realistic models of FePt nanoparticles (14), a system that in the past decade has been intensively studied for its potential applications in magnetic data storage technologies. The aim of the Wang-Landau/LSMS method that we develop here is to replace the empirical model of the nanoparticles in the simulations with parameter-free *ab initio* electronic structure calculations. The solution of the electronic structure problem is based on Kohn-Sham Density Functional Theory (KS-DFT) and we are using the LSMS method to solve the Kohn-Sham equations. This method was developed specifically to study magnetic systems with non-collinear moments on massively parallel machines and in 1998 was the first scientific application to sustain a teraflop/s on a 1500 processor Cray T3E.

### A. The Wang-Landau method

All thermodynamic potentials can be derived from the partition function

$$Z(T) = \int e^{-E(\mathbf{X})/(k_B T)} d\mathbf{X} \quad (1)$$

where  $E(\mathbf{X})$  is the internal energy of the system with the phase space described by the variable  $\mathbf{X}$  in some high dimensional space consisting of all the microscopic degrees of freedom of the system (atomic positions, velocities, and/or magnetic moments). In importance sampling Monte Carlo simulations one performs a random walk through phase space that is biased in such a way that the walker spends most of the time where the integrand in equation 1 is largest, that is, where the energy  $E(\mathbf{X})$  is small. In the Metropolis algorithm, this is accomplished by accepting a proposed move from point  $\mathbf{X}_i$  to  $\mathbf{X}_{i+1}$  with the probability

$$\min[1, \exp(-k_B T [E(\mathbf{X}_{i+1}) - E(\mathbf{X}_i)])]. \quad (2)$$

This results in a very efficient computation of the partition function at a particular temperature when the energy function is not too complex. In many cases that are common in nanoscience and biology, the energy function can have many local minima that are separated by large energy barriers. With the conventional Metropolis algorithm, the random walk can be trapped for very long time around local minima, and sampling representative parts of phase space around other local minima of the energy function can become exceedingly difficult.

The partition function in equation 1 can be rewritten in the form

$$Z(T) = \int g(E) e^{-E/(k_B T)} dE \quad (3)$$

where the density of states is defined as

$$g(E) = \int \delta(E - E(\mathbf{X})) dX \quad (4)$$

and  $\delta(E)$  is the Dirac  $\delta$ -function.

Flat histogram methods, such as the Wang-Landau algorithm, use the density of states,  $g(E)$ , for importance sampling. The analog of the Metropolis formula in this case is to accept the new configuration with probability

$$\min[1, g(E_i)/g(E_{i+1})]. \quad (5)$$

The effect is to create an equal probability of visiting each energy level in the system. In other words, a histogram of where the random walk is at the end of each Monte Carlo move would be essentially flat (save for statistical  $1/\sqrt{(N)}$  noise).

The main obstacle of flat-histogram methods is that  $g(E)$  is not known. Instead, an estimate of the density of states  $\tilde{g}(E)$  must be constructed self-consistently as the Monte Carlo estimate is generated. The Wang-Landau algorithm accomplishes this as follows. It begins with a prior estimate of the density of states,  $\tilde{g}_0(E)$ , which might be just a constant. Assuming that a Monte Carlo move to a new configuration with Energy  $E_{i+1}$  is accepted according to the criterion of equation 5, the density of states is updated with

$$\ln[\tilde{g}(E_{i+1})] \leftarrow \ln[\tilde{g}(E_{i+1})] + \ln f \quad (6)$$

where  $f$  is the modification factor that is initially set to  $\ln f = 1$ . For every accepted move, a histogram  $H(E)$  is updated, which records where the random walk has been. The estimate  $\tilde{g}(E)$  is considered converged, when

$$\min[H(E)] \leq A \text{ mean}[H(E)], \quad (7)$$

where the flatness parameter  $0 < A < 1$  controls the accuracy of the estimated  $\tilde{g}(E)$ , with increasing accuracy as  $A$  approaches unity. When this criterion is met, the modification factor is reduced such that  $\ln(f) \leftarrow \ln(f)/2$  and the histogram reset. For each iteration in  $f$ , a new visit histogram  $H(E)$  is calculated, and the process is repeated until  $\ln f \leq 1 \times 10^{-6}$  or smaller, and the estimated  $\tilde{g}(E)$  is considered the converged density of states.

Since the systems we set out to study here have continuous degrees of freedom,  $g(E)$  is a function of a continuous variable as well. Hence, when the random walk arrives at a particular configuration with energy  $E_i$ , the estimated density of state  $\tilde{g}(E)$  is updated by (15)

$$\tilde{g}(E) \rightarrow \tilde{g}(E) \times f^{k((E-E_i)/\delta)} \quad (8)$$

where  $k(x) \geq 0$  is a continuous function with compact support. In particular we choose an Epanechnikov kernel,  $k(x) = \max[0, 1 - x^2]$ , and the width  $\delta$  to be 2% of the

difference between the minimal (ferromagnetic) and maximal (anti-ferromagnetic) energies of the system.

### B. The Locally Self-consistent Multiple Scattering (LSMS) method

The LSMS method (13) calculates with almost perfect weak scaling the local spin density approximation to the diagonal part of the electron Green's function. The electron/spin density and energy are easily determined once the Green's function is known. The LSMS takes domain decomposition to its natural limit, one atom per processor. This approach was based on two assumptions about the growth of supercomputing, first, that the number of available processors would grow to keep pace with the scientific need to model larger and larger systems and second, that local memory would grow at a slower pace that could potentially limit advances in the level of treatment if, for example, the same memory were to be shared by multiple atoms. Linear scaling with system size is achieved in the LSMS by using several unique properties of the real space multiple scattering approach to the Green's function (8): 1) the Green's function is *nearsighted* as discussed at length by (9), therefore, each domain, i.e. atom requires only information from nearby atoms in order to calculate the local value of the Green's function; 2) the Green's function is analytic, therefore, the required integral over electron energy levels can be analytically continued onto a contour in the complex plane where the imaginary part of the energy further restricts its range; and 3) to generate the local electron/spin density an atom needs only a small amount of information, phase shifts (or single site scattering matrices,  $t$ ), from those atoms within the range of the Green's function. On an individual processor the bulk of the calculation is done by ZGEMM in the evaluation of the local sub-block of the inverse of the real space KKR matrix, which is constructed from the  $t$ 's of all atoms within the range of the Green's function. The very compact nature of the information that needs to be passed between processors and the high efficiency of ZGEMM are responsible for the superior performance of the LSMS code.

The structure of the code is based on the local interaction zone, LIZ, defined by the atoms encompassed by the range of the Greens function. Each processor finds local solutions to a PDE (the Schrodinger equation) from which are obtained the  $t$ 's which describe the solutions on the boundary of the domain, Fig. 1. Every processor has a list of the addresses of all processors/atoms in its LIZ. All the  $t$ 's that a processor requires in order to construct the local electron/spin density are distributed to it from its LIZ and its  $t$ 's are sent to those processors that have it in their lists, all via asynchronously posted receives and sends. Reconstructing the Kohn-Sham potential in the self-consistency cycle requires the long-range electrostatic contributions, these require distributing the single site moments of the density, a very small amount of data. For system sizes up to  $10^5$  atoms the total communication time is a small fraction of the total computation time. In our pursuit of the statistical physics of nano structures the small amount of communication within a single statistical instance of the

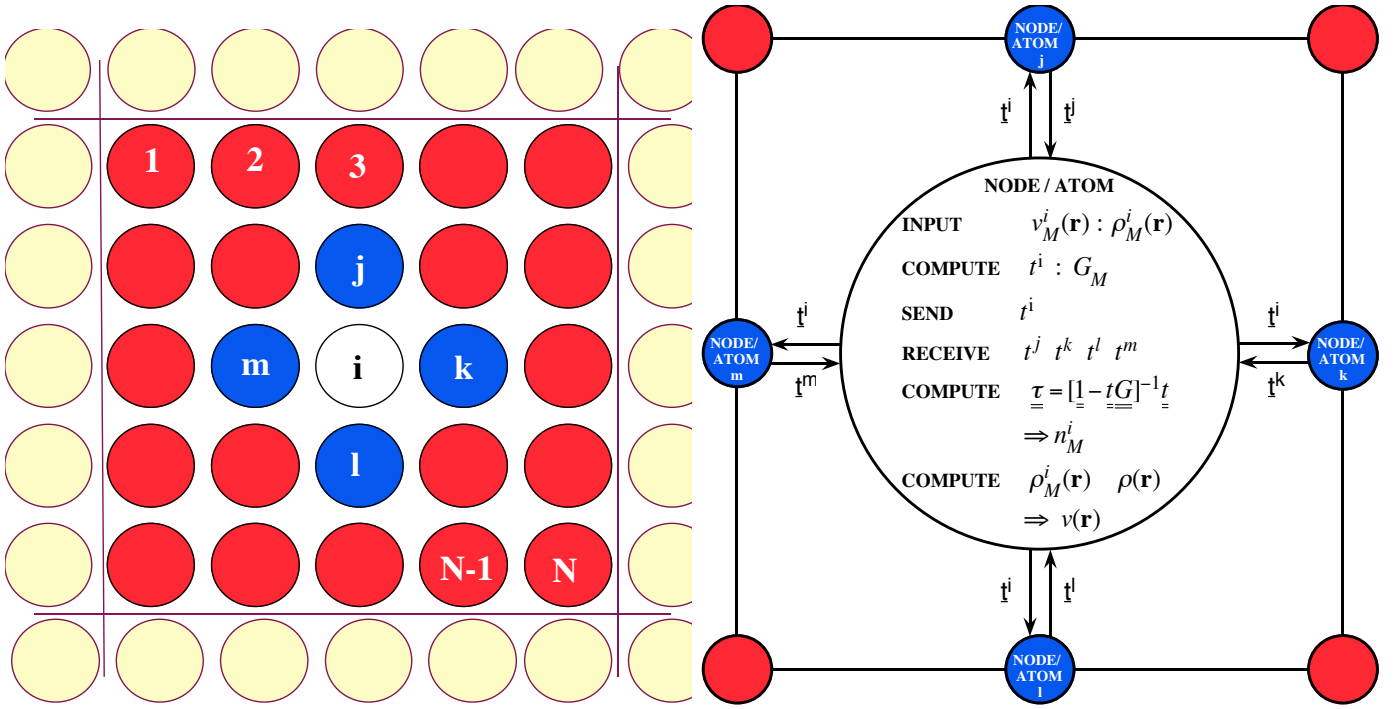


Fig. 1. Schematic, Left: LIZ centered at processor/atom  $i$ ; Right: message passing and computation.

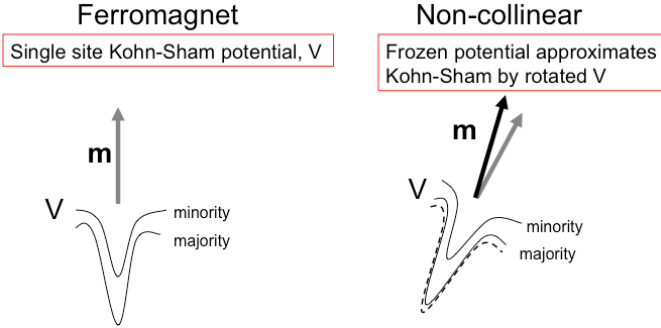


Fig. 2. Schematic of the relationship between ferromagnetic collinear Kohn-Sham potential and moment and the frozen potential rotated potential. The black arrow on the right illustrates that the moment associated with the rotated potential need not align with the rotated potential. The dashed line on the right indicates that the rotated potential will differ by small amounts from the actual Kohn-Sham potential.

LSMS will be negligible.

The frozen potential approximation (7), also referred to as the force theorem or the magnetic force theorem, allows us to bypass self consistent determination of the Kohn-Sham potential. The theorem applies to situations where the exact Kohn-Sham potential is known for one external field and a reliable transformation is available for converting the Kohn-Sham potential to that appropriate to a new external field. In the magnetic case of interest here the known potential is that of ferromagnetic Fe and the transformation corresponds to applied local magnetic fields that result in noncollinear magnetic moments, Fig. 2. In the LSMS the applied local fields

simply rotate the exchange potential on an atomic site (10). The Kohn-Sham exchange potential is represented as vectors in the Pauli matrices so it is natural to speak of the direction of the exchange potential. In this case the energy difference is given by the change in the eigenvalues sum between the ferromagnet and the rotated states which are overwhelmingly noncollinear. This approximation is valid to second order in the difference between the rotated potential and the variation of the sum of the exchange correlation and electrostatic energy with respect to the spin density matrix, i.e. the difference between input and output potentials. An alternative viewpoint is that the typically small difference in potentials represents a fluctuating, but not random, external field in which the statistical problem is being solved. One might also be concerned that the spin moments associated with the rotated potential will not point in the direction of the local rotation, however, this is not an issue. We are in fact searching over the energy landscape by making steps in the Kohn-Sham potential which has the same validity as searching over configurations defined in terms of the spin moments.  $T_c$  is not affected by this transformation and we can easily transform back to the moments which are uniquely related to the Kohn-Sham potential to determine the magnetization as a function of  $T$  in a joint density of states calculation.

### C. Hybrid WL-LSMS algorithm for *ab initio* statistical physics

Parallelization of the Wang-Landau method described above has to proceed differently from conventional Monte-Carlo methods, due to the fact that the acceptance probability for trial moves depends on the density of states  $g(E)$ , the property

to be calculated. Therefore the Wang-Landau method is not trivially parallel. In the algorithm presented here, we exploit the observation that the underlying *ab initio* LSMS energy calculations require many orders of magnitude more time (tens of seconds) than the Wang-Landau part of the algorithm. A single core executes the Wang-Landau algorithm (Alg. 1) and submits the energy calculation to a set of cores associated with an LSMS instance. (see Fig. 3) The Wang-Landau process maintains a set of moment directions for each LSMS instance and generates a new trial move for a given instance by randomly picking one moment in its set and generating a new random direction on a sphere for it. The LSMS instances in turn wait for spin configurations from the Wang-Landau driver and calculate the corresponding energies that are sent back to the Wang-Landau process. The Wang-Landau driver expects to receive the energies from the *ab initio* calculation that might arrive in a order that differs from the one in which they were submitted. While this destroys the determinism of the pseudo-random-number sequence used in the Wang-Landau driver, this has no negative effect on the convergence of the method. The energies thus obtained are used to determine the acceptance of the new random spin configuration and also to update the density of states and the histogram. Once the histogram indicates that the whole energy range has been visited it is cleared and the modification factor  $\gamma$  is reduced (note that  $\gamma = \ln f$ ). This process is repeated until the density of states is converged, which happens simultaneously with  $\gamma$  approaching zero. We expect this strategy to parallelize the Wang-Landau algorithm to scale to larger numbers of walkers, as long as the time for the Wang-Landau process to process one result from each walker is less then the time for one LSMS energy calculation.

---

**Algorithm 1** Wang-Landau/LSMS algorithm

---

- 1: initialize logarithmic density of states  $\ln g(E) \leftarrow 0$ , histogram  $h(E) \leftarrow 0$ , modification factor  $\gamma \leftarrow 1$ , and the set of magnetic moment directions for the  $M$  walkers  $\{\hat{e}\}_{1\dots M}$
  - 2: **repeat**
  - 3:   submit new random momet directions  $\{\hat{e}^{\text{new}}\}$  to idle LSMS instances
  - 4:   recieve new energy  $E_n^{\text{new}}$  from walker  $n$
  - 5:   accept new set of directions  $\{\hat{e}^{\text{new}}\}_n$  with probability  $\min[1, g(E_n^{\text{old}})/g(E_n^{\text{new}})]$
  - 6:   **if** Move accepted **then**
  - 7:      $\{\hat{e}^{\text{old}}\}_n \leftarrow \{\hat{e}^{\text{new}}\}_n$
  - 8:   **end if**
  - 9:   update density of states  $\ln g(E_n) \leftarrow \ln g(E_n) + \gamma$  and histogram  $h(E_n) \leftarrow h(E_n) + 1$
  - 10:   **if**  $h(E)$  flat **then**
  - 11:      $\gamma \leftarrow \gamma/2$ ,  $h(E) \leftarrow 0$
  - 12:   **end if**
  - 13: **until**  $g(E)$  converged, *i.e.*  $\gamma \approx 0$
- 

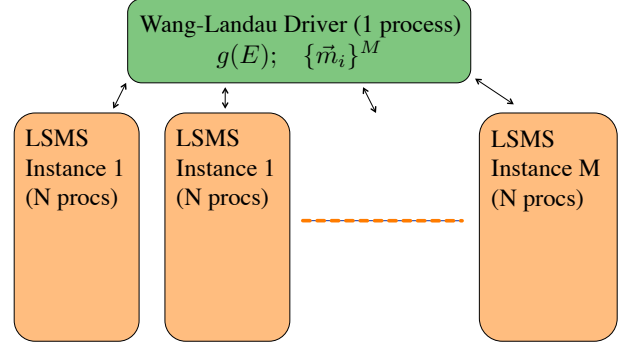


Fig. 3. Parallelization strategy of the combined Wang-Landau/LSMS algorithm. The Wang-Landau process (Alg. 1) generates random spin configurations for  $M$  walkers and updates a single density of states  $g(E)$ . The energies for these  $N$  atom systems are calculated by independent LSMS processes (Fig. 1). This results in two levels of communication, between the Wang-Landau driver and the LSMS instances, and the internal communication inside the individual LSMS instances spanning  $N$  processes each.

### III. SIMULATION RESULTS: ESTIMATING THE TRANSITION TEMPERATURE FOR IRON

In order to illustrate and quantitatively validate the WL-LSMS method, we present a calculation of the magnetic transition temperature (Curie temperature) of iron. Utilizing the methods and algorithm described above we consider a periodically repeated cells of 16 and 250 iron atoms, respectively, and converge the Wang-Landau density of states  $g(E)$  for changes in the magnetization direction on the individual iron sites. For the underlying LSMS calculation of these iron cells, the atoms were placed on a body-centered cubic lattice with a lattice parameter of  $5.42a_0$ , corresponding to the experimental value, and the local interaction zone has a radius of  $11.5a_0$ , including 65 atoms. The self-consistently converged potential for the ferromagnetic ground state was used for all the individual frozen-potential energy calculations in the combined Wang-Landau/LSMS algorithm.

TABLE I  
CONFIGURATION DETAILS AND TOTAL NUMBER OF WANG-LANDAU STEPS AS WELL AS TOTAL CPU-CORE HOURS REQUIRED TO CONVERGE THE DENSITY OF STATES OF THE 16 AND 250 ATOM SYSTEMS, RESPECTIVELY. THESE CALCULATIONS WERE PERFORMED ON A CRAY XT5 SYSTEMS WITH QUAD-CORE AMD OPTERON PROCESSORS RUNNING AT 2.3 GHZ.

atoms	WL walkers	cores	WL steps	total CPU-core hours
16	200	3,208	23,200	12,300
250	400	100,008	590,000	4,885,720

In table I we summarize the configurations and magnitudes necessary to converge the density of states of the 16 and 250 atom systems, respectively. The resulting density of states is shown in Fig. 4. While the computational resources needed for a system with several hundred atoms are considerable, the remaining calculations to compute any desired temperature dependent thermodynamic properties is marginal.

With this density of states the partition function that describes the thermodynamics of the system can be calculated

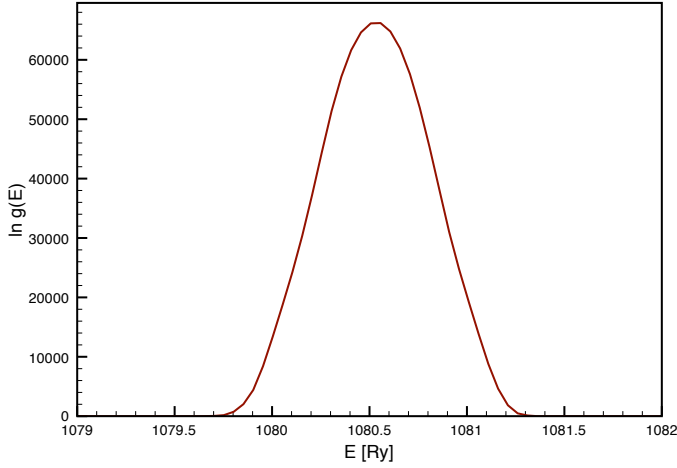
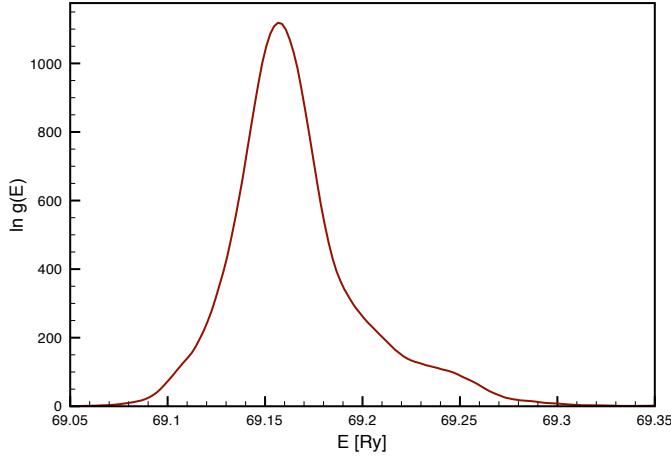


Fig. 4. The unnormalized logarithmic Wang-Landau density of states  $\ln g(E)$  for a periodic system of 16 (upper panel) and 250 (lower panel) iron atoms, respectively.

for any temperature.

$$Z'(T) = g_0 \int g(E) e^{-E/(k_B T)} dE = g_0 Z \quad (9)$$

Note that in this equation the partition function  $Z$  contains an unknown normalization factor  $g_0$  that will have consequences in calculating the free energy  $F$  (shown in Fig. 5) and entropy  $S = -\frac{\partial F}{\partial T}$  for the system.

$$F = -k_B T \ln Z' = -k_B T \ln Z - k_B T \ln g_0 \quad (10)$$

A phase transition in the system can be identified by jump in the internal energy  $U$  or a divergence the specific heat  $c = \frac{\partial U}{\partial T}$ . As the internal energy is related to the free energy and entropy by equation 11 the dependence on the unknown normalization factor  $g_0$  is canceled and our method yields absolute values for these quantities that are presented in Fig. 6.

$$U = k_B T^2 \frac{\partial \ln Z}{\partial T} = F + TS = F' + TS' \quad (11)$$

These calculations yield a transition temperature of 670K and 980K, for the 16 atom and the 250 system, respectively. The 16 site result is smaller as expected, since finite

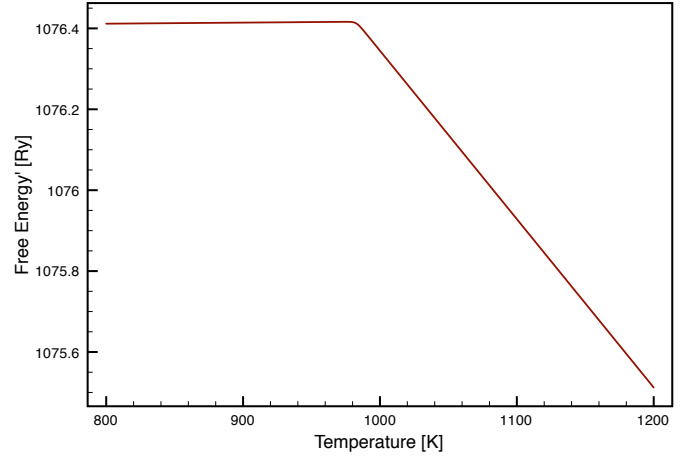


Fig. 5. Free Energy  $F'$  for a system of 250 iron atoms. The actual quantity plotted ( $F' = F + k_B T \ln g_0$ ) contains a contribution due to the unknown normalization factor  $g_0$ .

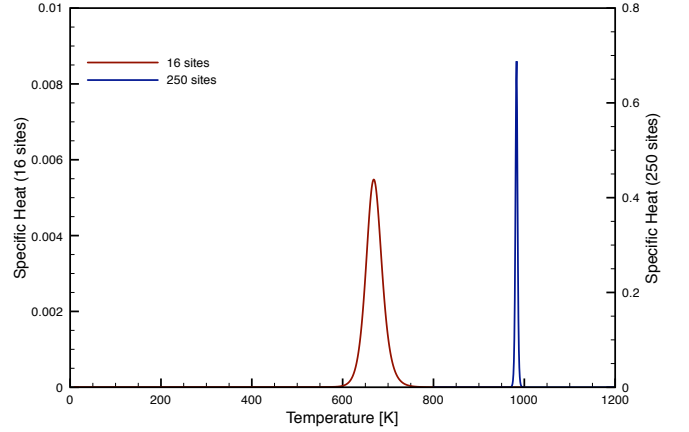


Fig. 6. Specific heat  $c$  for periodic 16 and 250 iron atom system calculated using the moments of the density of states  $g(E)$  as described in equation 16. A respective transition temperature of 670K and 980K can be read of these graphs.

size effects will lead to a larger reduction of the transition temperature for smaller cells. The result of the 250 atom calculation, however, is in remarkably good agreement with the experimentally known Curie temperature of bulk iron of 1050K. Calculations with 128 and 432 atom cells are currently under way and an estimate the true transition temperature predicted by the WL-LSMS method using the finite size scaling techniques of (1) will be published in a physics journal.

Finally we note, that for practical calculations the thermodynamic quantities can be expressed in terms of the moments

$$I_n(T) = \int E^n g(E) e^{-E/(k_B T)} dE \quad (12)$$

of the density of states  $g$ :

$$Z' = I_0 \quad (13)$$

$$F' = -k_B T \ln I_0 \quad (14)$$

$$U = I_1/I_0 \quad (15)$$

$$c = \frac{1}{k_B T^2} \left( \frac{I_2}{I_0} - \frac{I_1^2}{I_0^2} \right) \quad (16)$$

#### IV. PERFORMANCE OF THE WL-LSMS METHOD AT SCALE

The study of phase transitions in the previous section were intended to demonstrate the ability of the WL-LSMS method to predict (known) thermodynamic quantities from first principles with atomistic simulations. Precisely because the method is based on an *ab initio* electronic structure method, and not on an empirical model, it can be applied without changes to the code to any atomistic transition metal system. Of particular interest are small magnetic nanoparticles that have anywhere from around one hundred to a few thousand atoms, where the magnetic properties deviate most from bulk systems and studies of empirical models are only of limited value. Consequently it is important that the performance of the code is assessed for setups that are representative of these magnetic nanoparticles.

Hence the configuration we use for the performance analysis is as follows. The system consists of a 1024 Fe atom setup, and we study the scaling properties of the code as a function of number of walkers used in the WL simulation. Every individual LSMS calculation per walker can be distributed onto 1024 cores. On the Cray XT5 *jaguarpf* system at ORNL's National Center for Computational Sciences (NCCS), we can thus scale these calculations to up to 144 parallel WL walkers, leaving us with a total of 147,464 cores. In these performance analysis runs, each walker executes 20 WL steps, which is far fewer than a real simulations. Since the setup time of the calculations remains the same if the runs were longer, the performance numbers and scaling properties we report in this section are conservative estimates of the real numbers one would measure during production runs.

In figure 7 we show how the time to solution scales if we increase the number of WL walkers, and thus the total number of samples taken in the WL-LSMS simulation. The result shown thus represent a weak scaling plot, and the scaling behavior of the WL-LSMS method looks close to optimal, at least up to a machine with 150 thousand processors. We find a similarly optimal strong scaling behavior, if we fix the number of samples taken for every run and increase the number of walkers from 10 to 144. With the available size of machines today, we are still far from saturating the method in terms of scaling behavior.

The sustained floating point performance of the runs that correspond to the results presented in figure 7 are shown in Table II. In order to estimate the executed floating point operations of the benchmark runs, we have instrumented the WL-LSMS code with PAPI calls. The average number of floating point operations per second reported in table II is

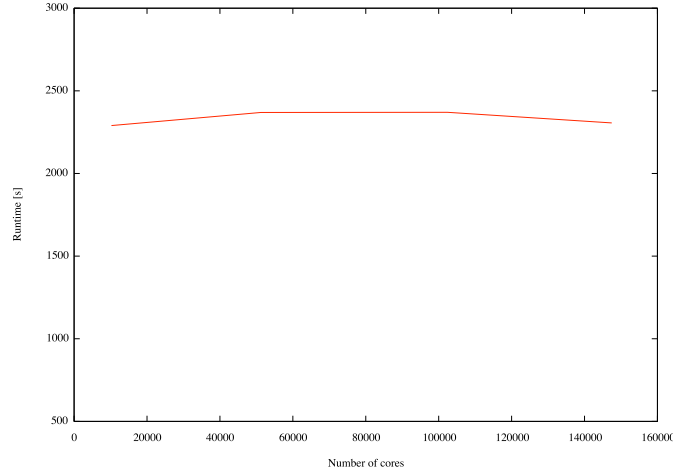


Fig. 7. Runtime of Wang-Landau-LSMS for a periodic 1024 iron atom cell demonstrating weak scaling over the number of walker instances. Each walker performed 20 Wang-Landau steps (energy calculations) and the number of walkers ranges from 10 (10248 cores) to 144 walkers (147464 cores).

TABLE II  
SUSTAINED PERFORMANCE OF THE WL-LSMS CODE RUNNING ON THE CRAY XT5 SYSTEM AT NCC, CONSISTING OF AMD QUAD-CORE OPTERON PROCESSORS THAT RUN AT 2.3 GHz AND HAVE A THEORETICAL PEAK PERFORMANCE PER CORE OF 9.2 GFLOP/S.

num. of WL walkers	num. of cores	TFlop/s	percentage of peak
10	10,248	72.546	76.9%
50	51,208	350.558	74.4%
100	102,408	698.124	74.4%
144	147,464	1028.75	75.8%

computed from the total number of retired floating point operations as reported by PAPI\_FP\_OPS events divided by the measured execution time of the runs. The run time has been measured with PAPI calls to PAPI\_get\_real\_usec as well as calls to the C time routine. For the largest runs with 144 parallel Wang-Landau walkers of 1024 atoms each and 20 steps per walker, the measured sustained performances was 1.028 petaflop/s, which on the 147,464 AMD Opteron cores running at 2.3 GHz corresponds to a fraction of 75.8% of the theoretical peak performance.

#### V. SUMMARY AND OUTLOOK

We have developed a new and highly scalable method to compute free energies of atomistic nano-scale systems. It combines the Wang-Landau algorithm, which performs a random walk in energy space, with Kohn-Sham Density Functional Theory based methods to compute the energy of the nanostructure. We have used the method in conjunction with the locally self-consistent multiple scattering method to study a magnetic system with non-collinear moments. With the specific example of bulk Fe, which we discuss here, we show that the method can be used to compute thermodynamic potentials and predict the Curie temperature without the use of adjustable parameters. The WL-LSMS code is applicable as is to any magnetic nano-particle with anywhere between

tens to a few thousand atoms, and is particularly suited for the study of transition metal alloys.

The WL-LSMS code uses a hybrid parallelization scheme. At the top level, the code parallelizes over concurrent random walkers, where we use a master-slave scheme, with a master that accumulates the density of states of the system, and the slaves that execute the random walks, each running its own instance of the LSMS method. The second parallelization level is the LSMS portion of the code, where domain decomposition is used with one atom per processing core. In typical production runs, the WL method would use a hundred to a few thousand concurrent walkers, and the LSMS portion would be parallelized over up to a few thousand processing cores. The method hence scales to hundred thousand or millions of processing cores.

Performance runs were carried out on the Cray XT5 system at ORNL. On this supercomputer, the WL-LSMS code sustains a performance of about 75% of peak. The scaling to up to 150 thousand processing cores proves to be nearly ideal. On 147,464 cores, the code sustains a double precision floating point performance of 1,029 TFlop/s.

In future work, we plan on using this code to study the temperature dependent free energy barrier for magnetic switching of FePt nanoparticles, a system that has received much attention in magnetic data storage applications. Since our approach is based on first principles electronic structure calculations, the WL-LSMS code will allow us to study directly the influence of the surface region on the magnetic properties, in particular, how the properties of the particles can be engineered with different coatings and surface treatments.

On the methods side, we plan to extend the parallelization model of the Wang-Landau method to include multiple masters. We will have two goals in mind. First, for cases where the energy evaluation are very fast, i.e. much shorter than the present case where we use the LSMS code, we will try to distribute the work of the master, in order to scale to large numbers of walkers without running into limitations of Amdahl's law. Second, we plan to make the WL method resilient to the loss of processing nodes.

Finally, the approach of using classical variables to connect first-principles calculations with Wang-Landau Monte Carlo techniques to calculate thermodynamics properties is not limited to LSMS or magnetic systems. It would be interesting, for example, to compare the Wang-Landau approach to the Metropolis-sampling results found for VASP calculations of molten Li (12).

## VI. ACKNOWLEDGMENT

We acknowledge use of resources of the National Center for Computational Sciences and the Center for Nanophase Materials Sciences, which are sponsored by the respective facilities divisions of the offices of Advanced Scientific Computing Research and Basic Energy Sciences (BES) of the U.S. Department of Energy. The present research was supported by the Laboratory Directed Research and Development fund of

ORNL, as well as the division of scientific user facilities of BES.

## REFERENCES

- [1] BINDER, K., AND LANDAU, D. P. *Phys. Rev. B* 30 (1984), 1477.
- [2] HOHENBERG, P., AND KOHN, W. *Phys. Rev.* 136 (1964), 864.
- [3] KENT, P. R. C. Computational challenges of large-scale, long-time, first-principles molecular dynamics. *J. of Physics: Conf. Ser.* 125 (2008), 012054.
- [4] KOHN, W., AND SHAM, L. *Phys. Rev.* 140 (1965), 1133.
- [5] LAIO, A., AND PARRINELLO, M. Escaping free-energy minima. *PNAS* 99 (2002), 12562–12566.
- [6] MARTIN, R. M. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge, 2004.
- [7] NICHOLSON, D. M. C., AND BROWN, R. H. *Phys. Rev. B* 67 (2003), 0164011.
- [8] NICHOLSON, D. M. C., STOCKS, G. M., AND Y WANG, E. A. *Phys. Rev. B* 50 (1994), 14686.
- [9] PRODAN, E., AND KOHN, W. *Proc. Natl. Acad. Sci.* 102 (2005), 11635.
- [10] UJFALUSSY, B., WANG, X., AND D. M. C. NICHOLSON, E. A. *J. Appl. Phys.* 85 (1999), 4824.
- [11] WANG, F., AND LANDAU, D. P. *Phys. Rev. Lett.* 86 (2001), 2050.
- [12] WANG, S., MITCHELL, S. J., AND RIKVOLD, P. A. Ab initio monte carlo simulations for nanoscopic lithium systems at different temperatures. *Comput. Mat. Sci.* 29 (2004), 145.
- [13] WANG, Y., STOCKS, G. M., AND W. A. SHELTON, E. A. *Phys. Rev. Lett.* 75 (1995), 2867.
- [14] ZHOU, C., SCHULTHESS, T. C., AND MRYASOV, O. Magnetic anisotropy of fept nanoparticles: temperature-dependent free energy barrier for switching. 2950–2952.
- [15] ZHOU, C., SCHULTHESS, T. C., TORBRGGE, S., AND LANDAU, D. P. Wang-landau algorithm for continuous models and joint density of states. *Phys. Rev. Lett.* 96 (2006), 120201.