

# First principles calculation of the configurational energy density of states for LLTO with new Wang and Landau algorithm variant

Jason D. Howard<sup>1</sup>

<sup>1</sup>*Materials Science Division, Argonne National Lab, Lemont, IL, 60439, USA*

(Dated: February 21, 2020)

In this work a variant of the Wang and Landau algorithm for calculation of the configurational energy density of states is proposed. The algorithm was developed for the purpose of working towards the goal of using first principles simulations, such as density functional theory, to calculate the partition function of disordered sublattices in crystal materials. The expensive calculations of first principles methods make a parallel algorithm necessary for a practical computation of the configurational energy density of states within a supercell approximation of a solid state material. The developed algorithm is natural to parallelize, is developed from a self consistent perspective, and was developed purposely for lattice based problems encountered in the study of disordered crystal sublattices. The algorithm developed in this work is tested with the 2d Ising model to bench mark the algorithm and to help provide insight for implementing the algorithm to a materials science application. Tests with the 2d Ising model revealed that the algorithm has similar behavior to the N-fold 1/t form of the Wang and Landau algorithm. The algorithm is then applied to the lithium and lanthanum sublattice of the solid state lithium ion conductor  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . This was done to help understand the disordered nature of the lithium and lanthanum. The results find, overall, that the algorithm performs very well for the 2d Ising model and that the results for  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$  are consistent with experiment while providing additional insight into the lithium and lanthanum ordering in the material.

## I. INTRODUCTION

For crystalline materials with disordered sublattices such as the lithium ion solid state electrolyte LLTO it is desirable to calculate from first principles methods(such as density functional theory[1]) the configurational energy density states  $G(E_j)$ . Here the energy density of states refers to the energies of the distinct lattice configurations. With the energy density of states the partition function,

$$Z = \sum_i^{\Omega} e^{\frac{-e_i}{k_B T}} = \sum_j^{\Pi} G(E_j) e^{\frac{-E_j}{k_B T}}, \quad (1)$$

can be determined and from it many important thermodynamics properties such as the free energy, entropy, specific heat, and ensemble averages calculated. In Eq. (1),  $\Omega$  corresponds to the number of possible configurations and energies in the set  $\{\Sigma_i, e_i\}_{\Omega}$ ,  $\Pi$  to number of possible distinct energies  $E_j$ ,  $k_B$  is Boltzmann's constant, and  $T$  is the temperature. If  $\Omega$  is small enough one of the simplest ways to solve this problem could be direct random sampling of the configuration space[2]. For most problems  $\Omega$  will be too large to allow direct random sampling of the configuration space to be practical, making an importance sampling algorithm necessary. A very well known importance sampling method to solve this problem could be temperature dependent simulations involving the Metropolis algorithm and sampling with probability proportional to  $\exp(\frac{-e_i}{k_B T})$  along with histogram re-weighting techniques[3–5]. Another more advanced method is the multi-canonical method proposed by Berg et al. [6, 7]. A variant of multicanonical sampling that samples the density of states directly known as entropic

sampling developed by Lee [8] could also be used. These algorithms require a good estimate of the density of states to be effective. Another algorithm called the Wang and Landau algorithm [9, 10] has been developed which is temperature independent, is based on a random walk in energy space with probability inversely proportional to the current estimate of the density of states, and builds up the density of states as the algorithm progresses. An issue with these algorithms(if using a single walker) in use with first principles methods such as density functional theory is the large number of iterations needed, which would require a prohibitively long wall time at the current performance power of computers. In this paper an algorithm is proposed that combines the use of random sets along with the importance sampling method of the Wang and Landau algorithm, this importance sampling is similar to the entropic sampling proposed by Lee [8]. The algorithm also used the principle of the Wang and Landau algorithm to build up an estimate of the density of states as the algorithm progresses. The proposed algorithm is meant to work towards the goal of a highly parallel importance sampling algorithm that directly calculates the density of states, meshes well with mid level high performance computing architectures(such as Argonne's BEBOP), and has a minimum of paramaters for implementation. The algorithm developed in this work is referred to as the B<sub>L</sub>ENDER (B<sub>L</sub>end Each New Density Each Round) algorithm.

The Wang and Landau method does have parallel versions, including restricting random walkers to specific energy ranges, allowing the walkers to explore the entire space while periodically communicating with each other, and methods based on a replica exchange frame work [11–14]. The B<sub>L</sub>ENDER algorithm is characterized by

allowing the walkers to explore the entire energy range and communication with each other through an update to the density of states at each iteration. In principle many of the different forms of the Wang and Landau sampling currently used are based around the concept of sampling until a flat histogram of the visited energies is reached followed by a reduction in a modification factor to the density of states and that the calculated density of states is multiplied by this factor every time an energy level is visited. One issue with these types of Wang and Landau simulations is that, being based on a flat histogram of the visited energies, the energy range must be specified *a priori*. Another issue is that the original Wang and Landau formulation for the reduction in the modification factor of the density of states has been shown to be non-convergent[15–17]. There have been advancements made in understanding how to reduce the modification factor by Belardinelli et al. [18] whom developed the  $1/t$  algorithm which is proven to be convergent, this result was verified by the work of Zhou et al. [19]. The novel aspects of the B<sub>L</sub>ENDER algorithm include, a continuous adaptation of the modification factor to the density of states using the current sum of the density of states as a regulator, using the number of configurations as a parameter in the modification factor, and using a histogram of the currently visited energies as a parameter in the modification factor. The algorithm in this work was also formulated in a self consistent fashion, is believed to be convergent based on its comparison with the  $1/t$  algorithm, and is natural to parallelize as it is based on a set of random walkers. The algorithm was developed for ease of use in the application to disordered sublattices of crystal systems.

In this work the formulated algorithm is benched marked with the 2d Ising model as a standard means of testing performance. The tests allow for a comparison to exact results and to previous benchmarks of other algorithms. The tests with the 2d Ising model also allow for insight in how to implement the algorithm to a materials science problem. The main goal in this work was to calculate with first principles methods the configurational energy density of states of the lithium ion conductor  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . There have been reports of the Wang and Landau algorithm used with first principles calculations to calculate magnetic properties of materials and order to disorder properties of alloys[20, 21].  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$  is part of a family of possible stoichiometries  $\text{Li}_{3x}\text{La}_{2/3-x}\text{TiO}_3$  of interest as solid state lithium ion conductors[22–28]. For all of the possible stoichiometries there is a tendency towards ordering of the lithium and lanthanum into lithium rich layers and lanthanum rich layers. The primary calculation of this work is that of the temperature dependent order parameter related to the lanthanum rich layer in  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . This calculation both serves to benchmark the application of the algorithm to a materials science problem with experimental knowns and to provide further insight into the physics of the material.

The rest of the article is organized as follows; section II explains the B<sub>L</sub>ENDER algorithm, section III covers the testing of the algorithm with the 2d Ising model, section IV covers the application of the algorithm to LLTO along with the computational details of the first principles methods, and section V is the conclusions.

## II. ALGORITHM

The B<sub>L</sub>ENDER algorithm proposed in this work is given as follows. It is noted that the following algorithm is in terms of producing a relative density of states  $G_r(E_j)^I$ , where  $I$  is the iteration number.

1.  $G_r(E_j)^I, \{\Sigma_s, e_s\}_{\mathcal{S}}^I$
  2.  $\{\Sigma_s, e_s\}_{\mathcal{S}}^I \rightarrow \{\Sigma'_s, e'_s\}_{\mathcal{S}}^I$
  3.  $\Sigma'_s, e'_s \rightarrow \Sigma_s^{I+1}, e_s^{I+1} \quad P = \min[1, G_r(e_s)^I / G_r(e'_s)^I]$   
     else  $\Sigma'_s, e'_s \rightarrow \Sigma_s^{I+1}, e_s^{I+1}$
  4.  $G_r(E_j)^{I+1} =$
- $$G_r(E_j)^I + \frac{C_o \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}}^{I+1})}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}} G_r(E_j)^I =$$
- $$G_r(E_j)^I (1 + \frac{C_o \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}}^{I+1})}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}}) \quad (2)$$

Where  $G_r(E_j)^0 \equiv [1 + \frac{C_o}{\mathcal{S}} \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}}^0)]$  with  $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}})$  being a histogram function that counts the number of energies  $E_j$  in the set  $\{e_s\}_{\mathcal{S}}$ . In this work  $\{\Sigma_s, e_s\}_{\mathcal{S}}^0$  is a randomly(uniformly) drawn set from the configuration space  $\{\Sigma_i, e_i\}_{\Omega}$ . In the second step a random change is applied to each element of the sampled set  $\{\Sigma_s, e_s\}_{\mathcal{S}}^I$  to produced a “perturbed” set  $\{\Sigma'_s, e'_s\}_{\mathcal{S}}^I$ , for the Ising model this could be randomly flipping a spin. In the third step a random number is drawn between zero and one for every sampled configuration, if this number is less then the ratio of the current density of states of the unperturbed to perturbed energies  $G_r(e_s)^I / G_r(e'_s)^I$  then the perturbed configuration and energy  $\Sigma'_s, e'_s$  goes to  $\Sigma_s^{I+1}, e_s^{I+1}$ , else the unperturbed configuration and energy  $\Sigma_s^I, e_s^I$  goes to  $\Sigma_s^{I+1}, e_s^{I+1}$ . This step (third) is derived from the Wang and Landau method of sampling with probability proportional to the inverse of the density of states. In the fourth step a histogram of the updated  $\{e_s\}_{\mathcal{S}}^{I+1}$  energies is made and added (blended) into the current density of states  $G_r(E_j)^I$  by multiplying by a constant  $C_o$ (which affects the convergence properties) and  $G_r(E_j)^I$  divided by the sum of the density of states to the  $1/N$  power. The  $1/N$  power is introduced as tuning parameter to affect the convergence properties and was discovered through empirical testing with the 2d Ising model. The fourth step is also shown in terms of multiplication which is discussed later. In this work it was found  $C_o = \Omega^{\frac{1}{N}}$  was

computationally efficient. After the algorithm is deemed to be complete it is necessary to renormalize the iterated relative density of states  $G_r(E_j)^f$  at the final iteration  $I = f$  as follows,

$$\begin{aligned} 1. & A = \sum_j G_r(E_j)^f \\ 2. & G_{calc}(E_j) = G_r(E_j)^f \frac{\Omega}{A}, \end{aligned} \quad (3)$$

to produce the properly normalized estimated value of  $G(E_j)$ . In principle,  $G_r(E_j)$  can also be renormalized based on information of the number of configurations in a given bin. For example if the ground state is known to have a given degeneracy then the entire density of states can be normalized such that the ground state bin has the correct degeneracy.

An important discussion point of this algorithm (Eq. 2) is the update of the relative density of states (step four) being presented as addition and multiplication. In the addition form the self consistent nature of the update is clear, in the sense that the density of states is updated by adding a piece proportional to the counts in the histogram of the random set times the relative proportion of that energy level in the current estimate of the density of states. In typical Wang and Landau sampling the update of the density of states is preformed by multiplication of the density of states every time an energy level is visited combined with a periodic reduction of the multiplication factor when a histogram of visited energies reaches a predetermined flatness criteria, the histogram of the visited energies is then reset to zero. In the multiplication form of step four of this algorithm (Eq. 2) it is seen that the dependence on one over the sum of the density of states serves to naturally reduce the multiplication factor as the simulation progresses. The multiplication form is also useful when  $\Omega$  is large and the sum of the density of states is larger than a typical floating point number. In this case the log of the density of states can be stored and the update performed through addition of logs. Taking  $G_r^M \equiv \max[G_r(E_j)]$  the log of  $\sum_j G_r(E_j)^I$  can be written as,

$$\begin{aligned} G_r^{LS} &\equiv \log[\sum_j G_r(E_j)^I] = \log[G_r^M \frac{\sum_j G_r(E_j)^I}{G_r^M}] = \\ &\log[G_r^M] + \log[\sum_j e^{\ln[G_r(E_j)] - \ln[G_r^M]}]. \end{aligned} \quad (4)$$

With  $G_r^{LS}$  from Eq. 4 the log update form of step four of the algorithm (Eq. 2) can be written as the following,

$$\begin{aligned} &\log[G_r(E_j)^I (1 + \frac{C_o \mathcal{H}(E_j, \{e_s\}_S^{I+1})}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}})] = \\ &\log[G_r(E_j)^I] + \log[1 + \mathcal{H}(E_j, \{e_s\}_S^{I+1}) e^{\ln[C_o] - \frac{1}{N} G_r^{LS}}]. \end{aligned} \quad (5)$$

In this form the algorithm can be implemented even when  $\Omega$  is large. To implement the ratio of the density of states

in step two of the algorithm,

$$e^{\ln[G_r(e_s)^I] - \ln[G_r(e'_s)^I]}, \quad (6)$$

can be used.

### III. BENCH MARK WITH 2D ISING MODEL

In this work the algorithm discussed is tested using the 2d square zero field Ising model with lattice dimension of even number [29–31] and periodic boundary conditions. The configurations  $\Sigma_i$  and energies  $e_i$  of the 2d Ising model are inherently defined by the lattice site spin variables  $\sigma_{k,l}$  which take the values  $\pm 1$  and coupling constant  $J$ . Explicitly the energy  $e_i$  for a given configuration  $\Sigma_i$  of a  $n \times n$  Ising lattice is given by,

$$e_i = -J \sum_{k,l=1}^n \sigma_{k,l}^i (\sigma_{k+1,l}^i + \sigma_{k,l+1}^i). \quad (7)$$

The first test is of the effectiveness of the algorithm in calculating the density of states of the 2d Ising model. To test the accuracy of the simulations the results will be compared to the exact result solved by Beale [32]. The accuracy of the simulation will be determined by the error defined as,

$$\begin{aligned} \mathcal{E}(I, o) &= \langle |\epsilon(E_j, I, o)| \rangle_j \\ &= \frac{1}{\Pi} \sum_{j=1}^{\Pi} \frac{|\ln[G_{ex}(E_j)] - \ln[G_{calc}(E_j, I, o)]|}{|\ln[G_{ex}(E_j)]|}. \end{aligned} \quad (8)$$

Where  $G_{ex}(E_j)$  is the exact density of states,  $G_{calc}(E_j, I, o)$  is the calculated density of states from Eq. 3 at iteration number  $I$  from initial conditions and trajectory  $o$ , and  $|\epsilon(E_j, I, o)|$  is the absolute value of the fractional error for a specific energy level. The primed configurations in this work were generated by randomly flipping one spin on the Ising lattice

The first test of the algorithm is with the  $32 \times 32$  Ising model. In a materials science problem with first principles calculations the system size is not expected to be anywhere near the size of the  $32 \times 32$  Ising model so these results are included to show the algorithm may have potential for larger system size. While the ideal value of  $N$  is not known prior to the calculation it was found in this work that a value of  $N = 0.1$  was computationally efficient for the  $32 \times 32$  Ising model. In Fig. 1(a) the value of the average error calculated with Eq. 8 is shown up to  $1e7$  iterations for  $\mathcal{S} = 1, 10, 100, 1000$ , and  $1e4$ . The data in Fig. 1(a) is averaged over 36 individual simulations for each value of  $\mathcal{S}$ . The results show linear scaling from  $\mathcal{S} = 1$  to  $\mathcal{S} = 10$  and then another order of magnitude improvement from  $\mathcal{S} = 10$  to  $\mathcal{S} = 1000$ , no significant improvement is discernible going to  $\mathcal{S} = 1e4$ . The periodic fluctuations in the average error are also noted in going to larger  $\mathcal{S}$ , the exact nature of these fluctuations are unclear at the moment.

In Fig. 1(b) are also the results for the average error of 36 independent simulations of a  $10 \times 10$  Ising model for the different number walkers  $\mathcal{S} = 1, 10, 100, 1000$ , and  $1e4$ , and with  $N = 1$ . The results show that the scaling is quite good as the number of walkers increases. This result is encouraging because the number of configurations that will be typical for a supercell approximation in a first principles calculation is not expected to exceed the large number of  $\approx 10^{30}$  configurations in the  $10 \times 10$  Ising model. This test on the smaller model indicates the algorithm is suitable for implementation in the materials science problem tackled in this paper.

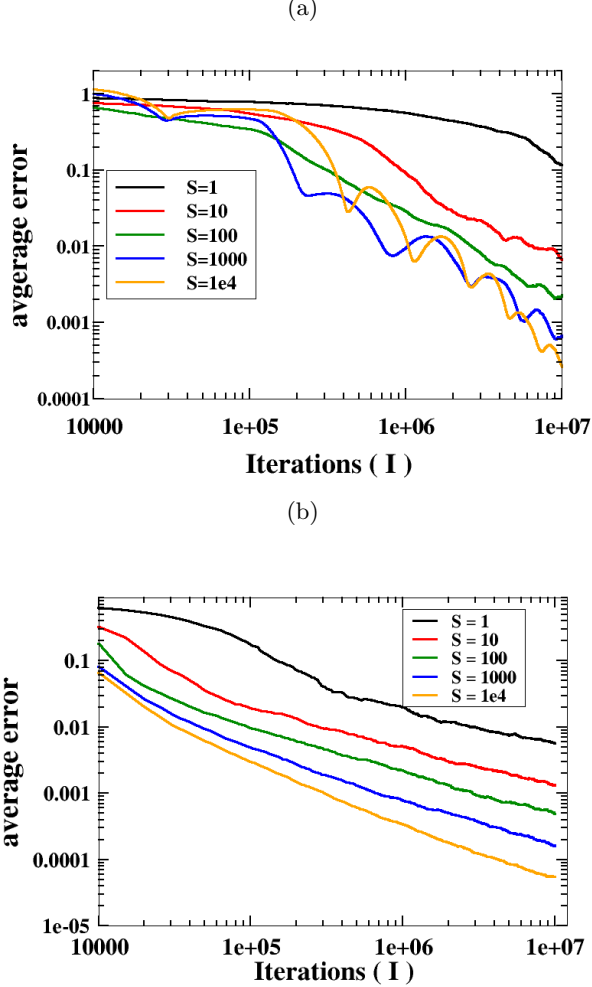


FIG. 1. (color online) Average error from 36 simulations calculated from Eq. 8 with  $\mathcal{S} = 1, 10, 100, 1000$ , and  $1e4$  for (a) the  $32 \times 32$  Ising model with  $1/N = 0.1$ , (b) the  $10 \times 10$  Ising model with  $1/N = 1$ .

Another aspect of the algorithm to consider is the dependence on the value of  $N$  and of  $C_o$ . In Fig. 2(a) and (b) the dependence on  $N$  is shown for the  $32 \times 32$  and  $10 \times 10$  Ising models, simulated to  $I = 1e7$  and  $I = 1e6$  respectively for  $\mathcal{S} = 100$  (a) and  $\mathcal{S} = 10$  (b). The results in Fig. 2(a) and (b) were averaged over 36 independent simulations. The results in Fig. 2(a) show that

for the larger  $32 \times 32$  model with  $\mathcal{S} = 100$  the dependence on  $N$  is more pronounced and that the optimal value of  $1/N$  is lower than for the results for  $32 \times 32$  model with  $\mathcal{S} = 10$  and the smaller  $10 \times 10$  model for both  $\mathcal{S} = 10$  and  $\mathcal{S} = 100$ . The more pronounced convergence dependence on  $N$  for the larger  $32 \times 32$  model at larger  $\mathcal{S}$  does pose a problem if one was to implement the the algorithm for a new system where the density of states is not known beforehand because there is no current evidence to predict what the optimal parameter would be. The tests with the  $10 \times 10$  model suggest that for a smaller system size that the convergence dependence on  $N$  is less pronounced and that  $N = 1$  is sufficient. The results presented here suggest that if one was to use the algorithm for a larger system size some method of predicting the optimal value of  $N$  would be required. These tests have used a value of  $C_o = \Omega^{1/N}$ , this value was based on tests showing this to be a computationally efficient choice that can be predicted based on knowledge of the system. In Fig. 3(a) and (b) is shown the error for the  $32 \times 32$  and  $10 \times 10$  Ising model vs  $C_o/\Omega^{1/N}$  with,  $1/N = 0.1$  and  $1/N = 1$  respectively,  $I = 1e7$  and  $I = 1e6$  respectively, and with  $\mathcal{S} = 100$  (a) and  $\mathcal{S} = 10$  (b). The results in Fig. 3(a) and (b) were averaged over 36 independent runs. The results in Fig. 3(a) and (b) show that the results are relatively insensitive to the value of  $C_o$  within several orders of magnitude of  $\Omega^{1/N}$  and that the main feature is a sudden increase in error going below some lower bound of  $C_o$  and a plateau of the error for  $C_o$  above this lower bound.

#### A. Relationship of the $B_L$ ENDER algorithm to the $N$ -fold $1/t$ algorithm

In the introduction it was mentioned that an improvement to update of the Wang and Landau method had been made by Belardinelli et al. [18], which was referred to as the  $1/t$  algorithm. It was found in this work that for the  $\mathcal{S} = 1$ ,  $N = 1$ , and  $C_o = \Omega$  that the  $B_L$ ENDER algorithm had similar behavior to the results of Belardinelli et al.. To make the comparison we first note that in the multiplication form of the  $B_L$ ENDER algorithm with  $C_o = \Omega$  and  $N = 1$  we can define,

$$f(I) \equiv \left(1 + \frac{\Omega \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}=1}^{I+1})}{[\sum_j G_r(E_j)^I]}\right) \quad (9)$$

such that the update to the density of states is given by,

$$G_r(E_j)^{I+1} = G_r(E_j)^I f(I). \quad (10)$$

This is in the original form of the Wang and Landau algorithm, the difference being how  $f(I)$  changes with time(iteration). In the work by Belardinelli et al. They defined a value  $F(t) = \log(f(t))$  and found an optimal form of  $F(t) = \frac{c}{t^p}$ , with  $t$  being the effective Monte Carlo time as defined by,

$$t = \frac{SI}{\Pi}. \quad (11)$$

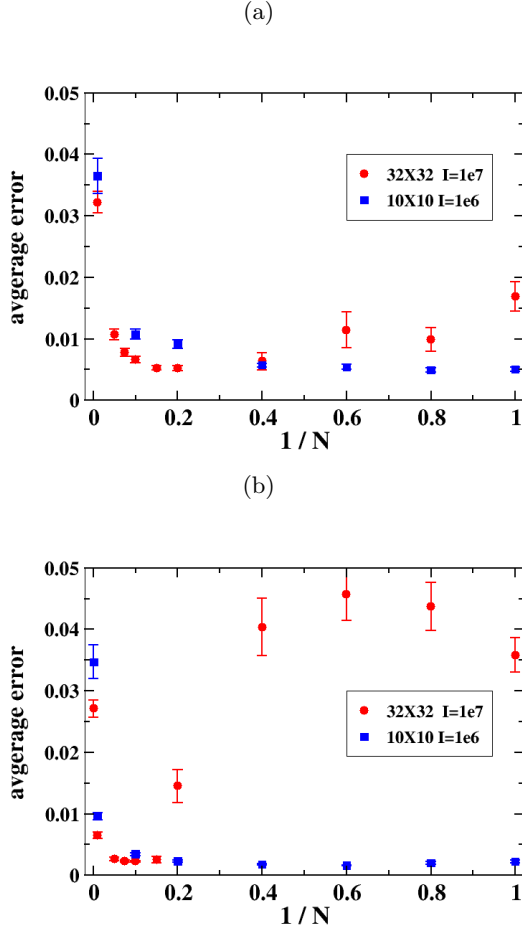


FIG. 2. (color online) Average error calculated from Eq. 8 from 36 simulations for (a)  $S = 10$  and (b)  $S = 100$  vs the value of  $1/N$  for the  $32 \times 32$  Ising model simulated to  $1e7$  iterations as red circles and the  $10 \times 10$  Ising model simulated to  $1e6$  iterations as blue squares. Errorbars show the standard deviation of the mean.

They found the optimal values to be  $c = 1$  and  $p = 1$ . They also described a more complex variant of this algorithm called the N-fold  $1/t$  that incorporated the average life time of the initial configurations. In their work they used a  $8 \times 8$  Ising model, so to compare, calculations of the average of  $F(t)$  and the error defined by Eq. 8 from 36 simulations were done for the  $B_L$ ENDER algorithm. The results for the average value of  $F(t)$  are shown in black in Fig. 4(a) and the average error in Fig. 4(b). In comparing the results to those from Fig. 5 of Belardinelli et al. [18] shown in red in Fig. 4, a striking similarity is seen in the behavior as compared to the N-fold  $1/t$  algorithm. It is clear from this comparison that there is a mathematical relationship between these two algorithms although at the moment the exact nature of this relationship is unclear. To further analyze the results of Fig. 4(a) it is noted that assuming  $F(t) = \frac{c}{t^p}$  a linear fit of  $\log F(t)$  vs  $\log(t)$  will predict the coefficients  $c$  and  $p$ . Fitting the results from the  $B_L$ ENDER algorithm in Fig.

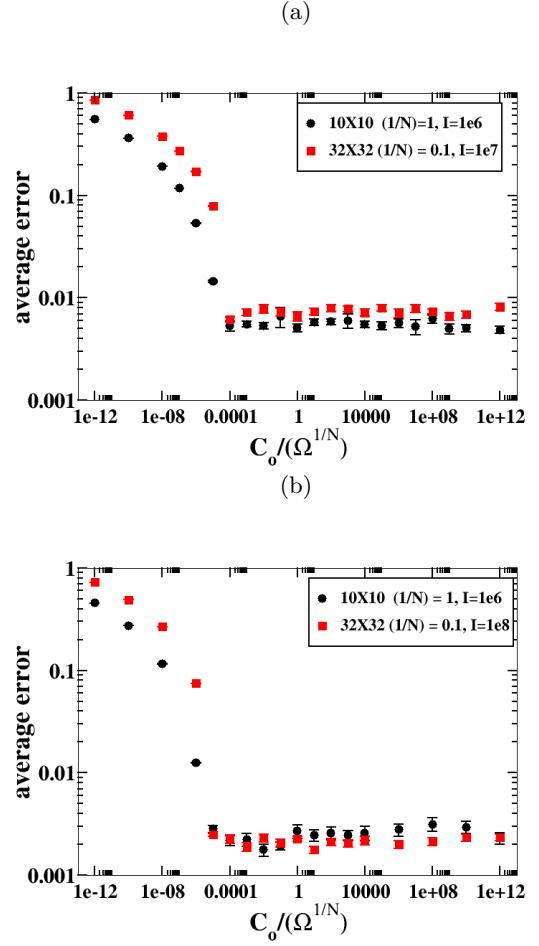


FIG. 3. (color online) Average error from 36 simulations calculated from Eq. 8 for the  $10 \times 10$  (black circles) and  $32 \times 32$  (red squares) Ising model simulated to  $1e6$  and  $1e7$  iterations and with  $1/N = 1$  and  $1/N = 0.1$  respectively for (a)  $S = 10$  and (b)  $S = 100$  vs the value of  $C_o/\Omega^{1/N}$ . Error bars show the standard deviation of the mean.

4(a) in this manner gives  $c = 1.8$  and  $p = 1$ . So the value of  $p$  is found to be in agreement with that of Belardinelli et al. while  $c$  is marginally larger. These similarities with the N-fold  $1/t$  algorithm, which is proven to be convergent, give further confidence in the  $B_L$ ENDER algorithm and suggest it is likely to be convergent as well.

#### IV. APPLICATION TO LLTO

The purpose of developing the  $B_L$ ENDER algorithm was to develop an algorithm suitable for the needs of solid state density functional theory calculations of disordered crystal sublattices. Due to the long run time of density functional theory calculations the parallel nature of  $B_L$ ENDER allows for calculations of each energy to be done as independent job submissions to a computer cluster. The results can then be processed by a script running on the head node. In this work the  $B_L$ ENDER algorithm

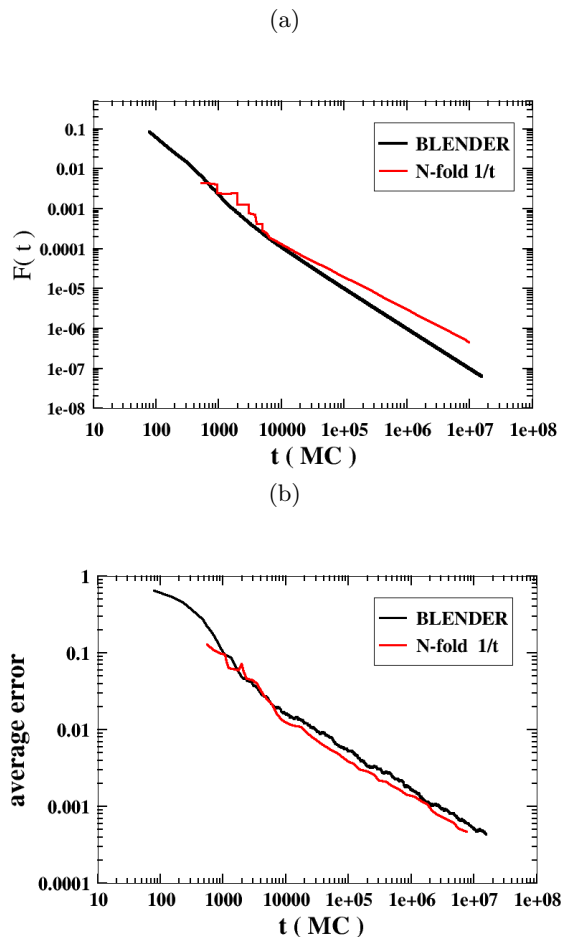


FIG. 4. (color online) In black (a) The average of  $F(t)$  and (b) the average error defined by Eq. 8 for 36 simulations of the  $8 \times 8$  2d Ising model with the BLENDER algorithm with  $S = 1$ ,  $N = 1$ , and  $C_o = \Omega$ . In red (a)  $F(t)$  results taken from the work of Belardinelli et al. [18] (b) the error defined by Eq. 8 taken from the work of Belardinelli et al.. Here  $t = \frac{t}{64}$  or equivalently one Monte-Carlo time step (MC) as defined by Eq. 11.

is applied to the lithium and lanthanum sublattice of the solid state lithium ion electrolyte  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . The goal of this study was to both, perform a calculation with BLENDER of a real material system that is fairly well understood, and also to learn something new in the process. Specifically the desired knowledge to be gained is a better understanding of the lithium and lanthanum sublattice disordering.

### A. Background on LLTO

LLTO is a complex material comprised of a variety of stoichiometries and phases but in this work the study is restricted to the reported tetragonal P4mmm phase of the stoichiometry  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$  [23, 27]. A unit cell of this structure is shown in Fig. 5. The lattice parameters

for this unit cell were taken from the experimental results from Ibarra et al. [23];  $3.8688(4)\text{\AA}$  for a- and b-axes, and  $7.7463(2)$  for c-axis. This unit cell is representative of an ordered form of  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$  where the lithium and lanthanum are separated into separate layers on the high symmetry A-sites. Where the A-site refers to the general perovskite formula unit  $\text{ABX}_3$ . The structure in Fig. 5 is actually structurally unstable and the energy can be lowered by lattice distortions which manifests as tilts in the titanium oxygen octahedra and the lithium distorting off of the high symmetry A-sites. The instability of the structure in Fig. 5 is evidenced by the imaginary phonon modes calculated by Moriwake et al. [24].

The physics of interest in this study is to understand the disordering of lithium and lanthanum between layers. It is reported for this phase that the lanthanum are mostly mixed between layers when the samples are slow cooled during synthesis and if quenched from high temperature the lanthanum ordering is reported to be completely mixed between layers [23]. Apart from the mixing of lithium and lanthanum between layers there is also configurational complexity associated with octahedral tilting. In this work the BLENDER algorithm is used to evaluate the density of states associated with local minimum corresponding to the lithium and lanthanum ordering and associated lattice distortions.

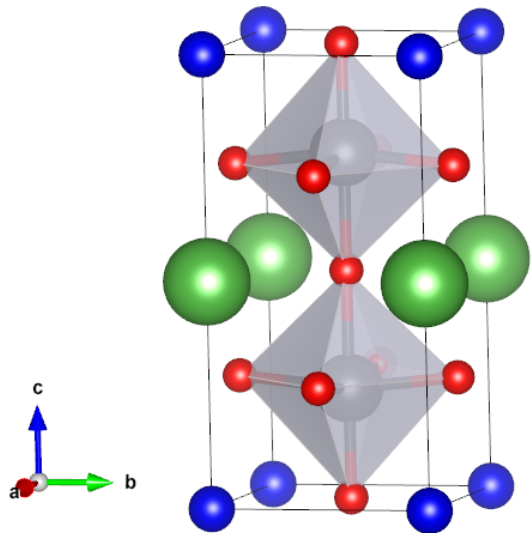


FIG. 5. (color online) 10 atom unit cell of P4/mmm  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . Where dark blue spheres are lithium, green spheres are lanthanum, red spheres are oxygen, and grey spheres inside of octahedra are titanium.

### B. Computational Details

In this work a  $3 \times 3 \times 1$  supercell of the unit cell depicted in Fig. 5 was used as an approximation to bulk

$\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$ . While not an ideal size as it is restrictive of the possible lattice configurations and to the types of domains of octahedral tilting that can form it is the largest supercell practical for performing the configurational Monte-Carlo in this work.

An important aspect of completing this study is a scheme for producing the initial and primed configurations in the iterative process of the B<sub>L</sub>ENDER algorithm. The scheme used in this study was to first generate a set of lithium and lanthanum randomly placed on the high symmetry A-sites where occupancy is restricted to one, then a small amount of noise on the order of  $\pm 0.2\text{\AA}$  was added to each lithium and lanthanum coordinate. These configurations were then relaxed to a local minimum which formed the first set of configurations in the iterative process. Then the primed configurations (step 2) were formed by swapping a random lithium and lanthanum atom and placing them back on the high symmetry A-site along with a new amount of random noise, these configurations were then relaxed to a local minimum. The random noise off the A-site served to allow for searching the distorted lattice configuration space.

The method used in the calculation of the total energies of the lattice configurations of LLTO in this work was density functional theory using the Vienna Ab-initio Simulation Package (VASP) [33–36] within the projector augmented wave formalism[37]. The Perdew-Burke-Ernzerhof (PBE) variant of generalized gradient approximation was used for the exchange and correlation functional[38]. The valence electron configurations for the data sets were;  $5p^65d^16s^2$  for La,  $2s^1$  for Li,  $3p^63d^23s^2$  for Ti, and  $2s^22p^4$  for O. The calculations also took advantage of the “soft” option for La and O. The total energy cut off for expansion of the plane waves was 250 eV. Self consistent cycles were converged with a energy difference of  $< 2.5\text{e-}5$  eV and relaxation of atomic coordinates was terminated when the difference in total energy between ionic relaxation steps was  $< 2.5\text{e-}4$  eV. A  $1\text{X}1\text{X}2$  gamma centered k-point mesh was used for the  $3\text{X}3\text{X}1$  supercells of the  $\text{LiLaTiO}_6$  unit cell. Each calculation of an energy was completed with a 36 processor broadwell node with an average wall time of 15 minutes per calculation. These cutoffs and parameters were chosen to maximize computational efficiency while retaining enough accuracy to capture important physical properties of LLTO. To test the accuracy of these methods 10 structures were calculated with fixed coordinates at these convergence criteria and more accurate PAW data sets and cutoffs. The more accurate PAW data sets included the valence electron configurations;  $5s^25p^65d^16s^2$  for La,  $1s^22s^1$  for Li,  $3p^63d^23s^2$  for Ti, and  $2s^22p^4$  for O. The cutoffs for the more accurate calculations were 450 eV for the plane wave basis, and  $2\text{X}2\text{X}3$  gamma centered k-points. The average magnitude of the error in relative energy between structures from this test was 0.05 eV.

The calculations were performed at the experimental lattice parameters  $3.8688\text{\AA}$  for a- and b-axes, and  $7.7463\text{\AA}$  for c-axis. The parameters for the B<sub>L</sub>ENDER algo-

rithm were  $S = 10$  and  $N = 1$ . The bin width used for determining  $G_r(E_j)$  was chosen to 0.02 eV. The value of  $\Omega$  was estimated as 100 times the combinatoric number of configurations of the lithium and lanthanum ordering onto the A-site given as,

$$\Omega \approx 100 \frac{18!}{9!9!}. \quad (12)$$

While an exact value of  $\Omega$  is not needed for the algorithm to converge experience from the 2d Ising model suggests that being within several orders of magnitude is sufficient. Estimating that  $\Omega$  is greater than the combinatoric calculation of the lithium and lanthanum in the A-site cages comes from the possibility of multiple distinct lattice distortions for each type of A-site cage configurations. An approximate upper bound on the number of distinct lattice distortions for each A-site cage configuration can be based on the experimental and theoretical known that lithium tend to occupy the six possible sites corresponding to local minimum near the oxygen windows connecting different A-site cages and that lanthanum tend to occupy the center of the cage [24, 28, 39, 40]. If every lithium could occupy one of these six locations within the A-site cage irrespective of the ordering of the other lithium the number of distinct lattice distortions could be estimated as  $6^9 \approx 1\text{e}7$ . So the approximation  $\Omega \approx \leq 1\text{e}7 \frac{18!}{9!9!}$  can be made. It is physically reasonable that a significant fraction of these configurations will be unstable so that the estimate of  $\Omega$  in Eq. 12 is likely to be within several orders of magnitude of  $\Omega$ .

### C. Results

Using the parameters and configurational enumeration scheme specified above a simulation was performed to 3,000 iterations for the  $3\text{X}3\text{X}1$ , 90 atom supercell. After 150 iterations the algorithm was restricted to look in the energy range less than 1.25 eV higher than the lowest energy found at that time. This was to improve computational efficiency by preventing the walkers from exploring an unnecessarily high energy range. While 3,000 iterations is not ideally converged, it was sufficient to gain further understanding of the material. In principle it would be desirable to use some type of stopping criteria to determine convergence of the simulation such as in the work of Caprica[41] whom tracked thermodynamic quantities such as the peak of specific heat to determine convergence. In this work convergence is limited by computational resources. For the 30,000 total structures calculated over the 3,000 iterations the total expense was 270,000 core hours. It is expected that the qualitative aspects of the results are well accounted for despite the limited number of iterations.

The main focus of the results is the nature of the lithium and lanthanum sublattice ordering. To accomplish this the order parameter of interest is that of the oc-

cupancy of lanthanum in the lanthanum rich layer along the c-axis. In the work by Ibarra et al. [23] they refer to this order parameter as  $La1$ , the same convention will be used in this work. This order parameter  $La1$  is defined as the number of lanthanum in the lanthanum rich layer divided by the total number that could occupy the layer. As an example the unit cell in Fig. 5 would have  $La1 = 1$ . It is important to note in this work the  $3 \times 3 \times 1$  supercell restricts the configurations along the a- and b-axes from having alternate layering of lithium and lanthanum rich layers. Ideally the calculations would be done with at least a  $4 \times 4 \times 1$  supercell but the computational effort is beyond the scope of this work. The results later will have to be interpreted taking this systematic supercell error into account.

To calculate the ensemble average of these order parameters first arithmetic averages of the order parameter at each energy level  $E_j$  are calculated from the primed configurations ( $\{\Sigma'_s\}$ ) that occurred during the simulation. The arithmetic average of a general order parameter  $O$  over all configurations with energy  $E_j$  is denoted by  $\langle O \rangle_j$ . Then with these the ensemble average is computed as,

$$\langle O \rangle = \sum_{j=1}^{\Pi} \langle O \rangle_j G_r(E_j) \frac{e^{-\frac{E_j}{k_B T}}}{Z}. \quad (13)$$

Where  $Z = \sum_{j=1}^{\Pi} G_r(E_j) \exp(-\frac{E_j}{k_B T})$ . It is noted that normalization of the relative density of states to the appropriate number of configurations is not necessary for the calculation of the ensemble average of an order parameter. If wanting to compare free energies ( $-k_B T \ln(Z)$ ) between phases it would be necessary to normalize the density of states properly to obtain an accurate calculation of the free energy.

The first main result to report is a view of the convergence of  $G_r(E_j)$  as a function of the iterations. In Fig. 6,  $G_r(E_j)$  is shown at  $I = 500, 1000$ , and  $3,000$  with the y-axis plotted on a log scale. The  $G_r(E_j)$  shown in Fig. 6 are plotted such that the lowest energy of  $G_r(E_j)$  found at the particular iteration shown is set to zero, the sharp cutoff at higher energy was the upper limit to energy range, and the plots are normalized by dividing through by the minimum of  $G_r(E_j)$  at that iteration. The main characteristic of the results by 3,000 iterations is the presence of some low energy states with a energy gap up to a more continuous spectrum of states, this gap could be due to insufficient sampling at lower energies. The lowest energy configuration is characterized as having  $La1 = 1$ , that being having alternate layers of lithium and lanthanum along the c-axis. It is not however equivalent to the unit cell shown in Fig. 5, in that the structures have distinct lattice distortions.

The next result is the arithmetic averages of the  $La1$  order parameter, which is shown in Fig. 7(a) along with the number of samples used to determine each value in Fig. 7(b). The results in Fig. 7a show an over all tendency for more mixing of lithium and lanthanum between

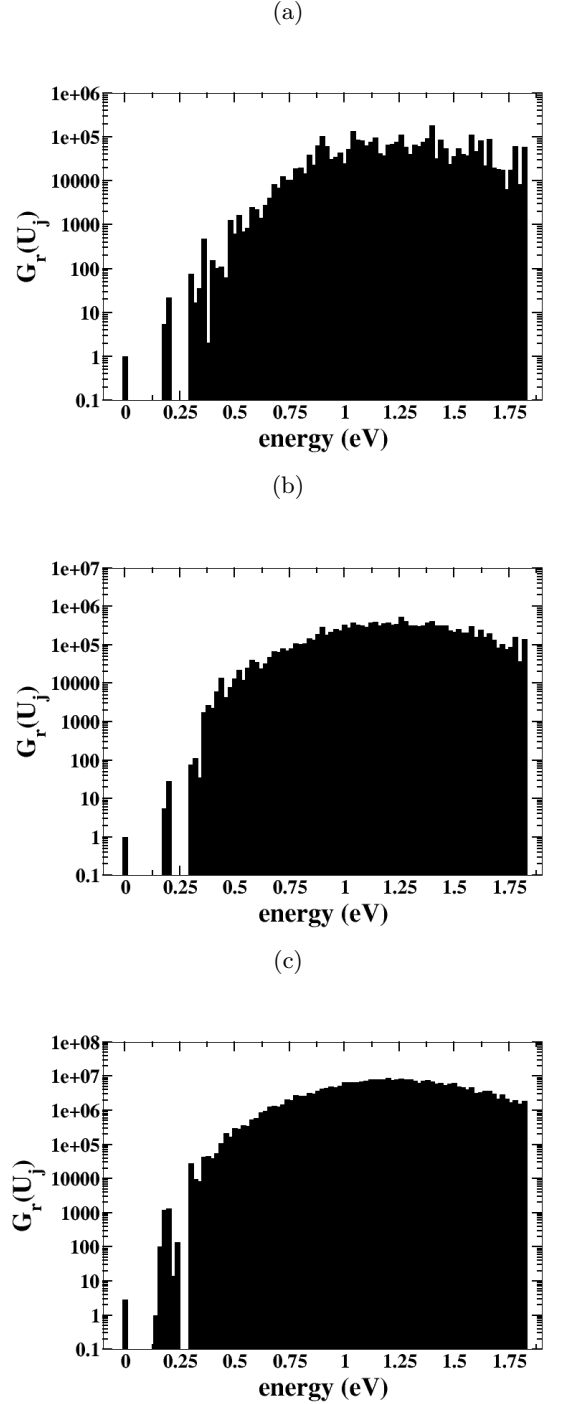


FIG. 6. Plots of  $G_r(E_j)$  at (a) 500 iterations, (b) 1,000 iterations, and (c) 3,000 iterations. The plots are normalized by dividing through by the smallest value of  $G_r(E_j)$  at that particular iteration. The plots are shown with a log scale on the y-axis.

layers for higher energies. The ensemble average of  $La1$  is shown in Fig. 8 which shows a phase transition from completely segregated lithium and lanthanum between layers to mostly mixed between layers and increased mix-



ing with increasing temperature. In Fig. 8 the values for  $La1$  range from approximately 0.675 at 500K to 0.625 at 1500K. For the  $3\times3\times1$  model the minimum possible value of  $La1$  is  $5/9 = 0.56$ . These results are in qualitative agreement with the  $La1 = 0.53$  reported by Ibarra et al. [23] and it is expected that that larger supercell models would increase the available configurational entropy and result in a reduced value of  $La1$  in the disordered phase.

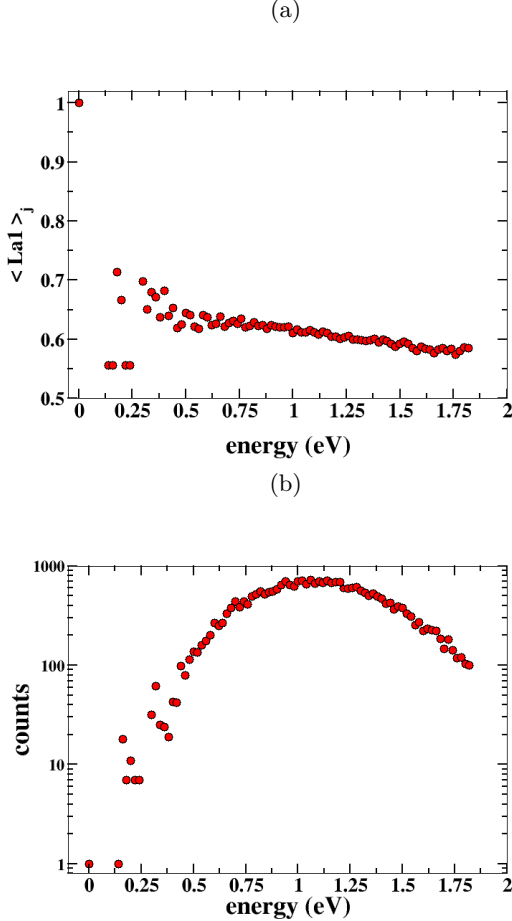


FIG. 7. (a) Arithmetic averages of  $La1$  order parameter as a function of energy. (b) Number of counts to determine each value of the  $La1$  order parameter for each energy.

Some indication of convergence of the 3000 iterations comes from an inspection of the flatness of the histogram of the energies visited during the simulation. The visited configurations and energies are defined by step 3 of Eq. 2. A histogram of the visited energies during the simulation are shown in Fig. 10. The results show that the histogram is qualitatively flat for  $\approx > 0.5$  eV. This result along with the counts for calculating the  $La1$  order parameter shown in Fig. 7(b) suggest that the results are most well converged for  $> 0.5$  eV. In this sense the statement can be made that the strongest result of this computation is the trend for more mixed configurations to have higher energy seen in the range  $> 0.5$  eV. While

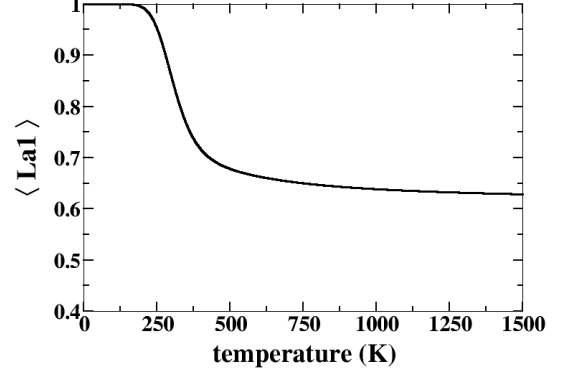


FIG. 8. Ensemble average of  $La1$  order parameter calculated with Eq 13 as function of temperature.

the first principles methods used can be considered coarse grained in terms of PAW data sets, total energy cutoffs, and k-points as observed from testing with more accurate methods it is expected that some error cancellation is present such that the overall trend seen in Fig. 7(a) is still qualitatively correct. It must be said that the ensemble average of  $La1$  is highly dependent on the low energy structures as per the exponential nature of the partition function. In this regard the observed phase transition in Fig. 8 can not be expected to be an accurate prediction of a transition temperature. The most important result of Fig 8 is the high temperature region above the phase transition showing a tendency for greater mixing of lithium and lanthanum for increasing temperature.

To gain some further insight in to the structures found during the simulation the two lowest energy structures are shown in Fig. 9(a) and (b). In Fig. 9(a) it is seen that the lowest energy structure has the lithium and lanthanum completely segregated between layers along the c-axis. In Fig. 9(b) that the second lowest energy structure has a fully occupied layer of lithium along the a-axis along with two partially occupied layers. Due to the pseudo cubic nature of the system it is expected that there would be nearly energetically identical structures consisting of completely segregated lithium and lanthanum layers along the a and b axes. This feature of ambiguity of the orientation of how the lithium and lanthanum rich layers can form are a likely driving force for the numerous domain boundaries observed in experiments for other stoichiometries synthesized by annealing from high temperature[22, 24]. Due to the  $3\times3\times1$  used in this study these structures were not possible to realize in the simulation but as evidenced by Fig. 9(b) segregation of the lithium and lanthanum into separate layers is still energetically favorable. A common feature of both Fig. 9(a) and (b) is the lithium sitting off of the high symmetry A-site near the oxygen windows separating A-site cages. This feature has been previously reported experimentally and theoretically in the literature[24, 28, 39, 40].

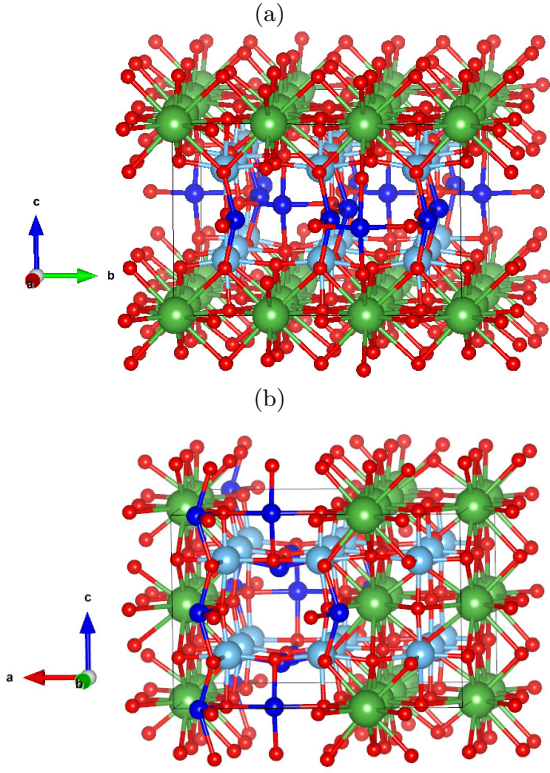


FIG. 9. (color online) (a) Lowest energy structure found during simulation. (b) Second lowest energy structure found during the simulation. Where dark blue spheres are lithium, green spheres are lanthanum, red spheres are oxygen, and light blue spheres are titanium.

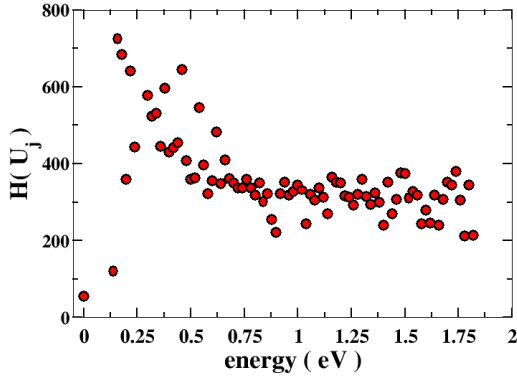


FIG. 10. Histogram of the visited energies over the 3000 iterations of the simulation. Where visited here means the accepted energies as per step 3 of Eq. 2.

## V. CONCLUSIONS

This work has presented a parallel variant of the Wang and Landau algorithm referred to as  $B_L$ ENDER (  $B_L$ end Each New Density Each Round). The algorithm was developed purposely for use with disordered crystal sublattices and is naturally parallel. It's design makes it facile to implement on a mid level high performance computer such as Argonne's BEBOP where jobs for a structural energy calculation can be independently submitted to compute nodes and managed by a script running on a head node. It was trialed using the 2d Ising model and showed good performance for the  $10 \times 10$  Ising model with a minimal number of implementation parameters. Results for the  $32 \times 32$  Ising model suggest the algorithm could have applicability to larger system sizes provided a tuning parameter is chosen appropriately, currently it is not known how to choose this parameter before hand without numerical testing. In the case of using one walker the  $B_L$ ENDER algorithm was also found to have similar behavior as to the N-fold  $1/t$  algorithm of Belardinelli et al.[18]. Knowledge gained from testing with the 2d Ising model allowed for an informed implementation to the real material science problem of studying the lithium and lanthanum sublattice disorder of  $\text{Li}_{0.5}\text{La}_{0.5}\text{TiO}_3$  using density functional theory methods. The simulations of the disordered lithium ion conductor LLTO were in qualitative agreement with experiment and provided further insight into the disordered nature of the material. It was found that lower energy structures favored segregating lithium and lanthanum into separate layers and that structures with lithium and lanthanum mixed between layers were on average higher in energy than more segregated structures. Thermodynamic analysis of the order parameter related to lithium and lanthanum intermixing between layers showed at phase transition between completely segregated to mostly mixed tending to more mixed at higher temperatures. Overall the results show that the algorithm performed well in simulating both the 2d Ising model and with first principles calculations of a real material system.

## ACKNOWLEDGMENTS

This work was supported by the Center for Electrical Energy Storage: Tailored Interfaces, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences at Argonne National Laboratory under Contract DE-AC02-06CH11357. I would like to thank the Laboratory Computing Resource Center (LCRC) faculty of Argonne National Lab for their support and maintenance of the computing resources that made this project possible.

[1] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[2] J. D. Howard and N. A. W. Holzwarth, Phys. Rev. B **99**, 014109 (2019).

- [3] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [4] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988).
- [5] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Physics*, fourth edition ed. (Cambridge University Press, 2015).
- [6] B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
- [7] B. A. Berg and T. Celik, *Phys. Rev. Lett.* **69**, 2292 (1992).
- [8] J. Lee, *Phys. Rev. Lett.* **71**, 211 (1993).
- [9] F. Wang and D. P. Landau, *Phys. Rev. Lett.* **86**, 2050 (2001).
- [10] F. Wang and D. P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
- [11] J. Yin and D. P. Landau, *Comput. Phys. Commun.* **183**, 1568 (2012).
- [12] L. Zhan, *Comput. Phys. Commun.* **179**, 339 (2008).
- [13] T. Vogel, Y. W. Li, T. Wüst, and D. Landau, *Phys. Rev. Lett.*, **110**, 210603 (2013).
- [14] T. Vogel, Y. W. Li, T. Wüst, and D. P. Landau, *Phys. Rev. E* **90**, 023302 (2014).
- [15] Q. Yan and J. J. de Pablo, *Phys. Rev. Lett.* **90**, 035701 (2003).
- [16] R. E. Belardinelli and V. D. Pereyra, *Phys. Rev. E* **75**, 046701 (2007).
- [17] R. E. Belardinelli and V. D. Pereyra, *Phys. Rev. E* **93**, 053306 (2016).
- [18] R. E. Belardinelli and V. D. Pereyra, *The Journal of Chemical Physics* **127**, 184105 (2007), <https://doi.org/10.1063/1.2803061>.
- [19] C. Zhou and J. Su, *Phys. Rev. E* **78**, 046705 (2008).
- [20] M. Eisenbach, D. M. Nicholson, A. Rusanu, and G. Brown, *Journal of Applied Physics* **109**, 07E138 (2011), <https://doi.org/10.1063/1.3562218>.
- [21] S. N. Khan and M. Eisenbach, *Phys. Rev. B* **93**, 024203 (2016).
- [22] X. Gao, C. A. J. Fisher, T. Kimura, Y. H. Ikuhara, A. Kuwabara, H. Moriwake, H. Oki, T. Tojigamori, K. Kohama, and Y. Ikuhara, *J. Mater. Chem. A* **2**, 843 (2014).
- [23] J. Ibarra, A. Vrez, C. Len, J. Santamara, L. Torres-Martinez, and J. Sanz, *Solid State Ionics* **134**, 219 (2000).
- [24] H. Moriwake, X. Gao, A. Kuwabara, C. A. Fisher, T. Kimura, Y. H. Ikuhara, K. Kohama, T. Tojigamori, and Y. Ikuhara, *Journal of Power Sources* **276**, 203 (2015).
- [25] H. Geng, A. Mei, C. Dong, Y. Lin, and C. Nan, *Journal of Alloys and Compounds* **481**, 555 (2009).
- [26] A. Belous, O. Yanchevskiy, O. V'yunov, O. Bohnke, C. Bohnke, F. Le Berre, and J.-L. Fourquet, *Chemistry of Materials* **16**, 407 (2004), <https://doi.org/10.1021/cm034820x>.
- [27] S. Stramare, V. Thangadurai, and W. Weppner, *Chemistry of Materials* **15**, 3974 (2003), <https://doi.org/10.1021/cm0300516>.
- [28] M. Catti, *Chemistry of Materials* **19**, 3963 (2007), <https://doi.org/10.1021/cm0709469>.
- [29] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Elsevier, 2016) iSBN 978-0-12-083180-7.
- [30] L. Onsager, *Phys. Rev.* **65**, 117 (1944).
- [31] B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model* (Harvard University Press, 1973).
- [32] P. D. Beale, *Phys. Rev. Lett.* **76**, 78 (1996).
- [33] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [34] G. Kresse and J. Hafner, *Phys. Rev. B* **49**, 14251 (1994).
- [35] G. Kresse and J. Furthmüller, *Computational Materials Science* **6**, 15 (1996).
- [36] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [37] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [38] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [39] X. Gao, C. A. J. Fisher, T. Kimura, Y. H. Ikuhara, H. Moriwake, A. Kuwabara, H. Oki, T. Tojigamori, R. Huang, and Y. Ikuhara, *Chemistry of Materials* **25**, 1607 (2013), <https://doi.org/10.1021/cm3041357>.
- [40] J. A. Alonso, J. Sanz, J. Santamara, C. Len, A. Vrez, and M. T. Fernández-Daz, *Angewandte Chemie International Edition* **39**, 619 (2000), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291521-3773%2820000204%2939%3A3%3C619%3A%3AAID-ANIE619%3E3.0.CO%3B2-O>.
- [41] A. A. Caparica, *Phys. Rev. E* **89**, 043301 (2014).