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

Jason D. Howard[1]

[1] *Materials Science Division, Argonne National Lab, Lemont, IL, 60439, USA**

(Dated: May 14, 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 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 algorithm developed in this work is tested with the 2d Ising model to bench mark the algorithm and to help provide insight for implementation to a materials science application. Tests with the 2d Ising model revealed that the algorithm has good performance compared to the original Wang and Landau algorithm and has similar behavior to the N-fold 1/t algorithm. The algorithm was then applied to the lithium and lanthanum sublattice of the solid state lithium ion conductor $Li_{0.5}La_{0.5}TiO_3$. 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 $Li_{0.5}La_{0.5}TiO_3$ are consistent with experiment while providing additional insight into the lithium and lanthanum ordering in the material. The primary result is that configurations with lithium and lanthanum more mixed between layers tend to be higher in energy.

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 $G(E_j)$, which will simply be referred to as the density of states. Here the density of states refers to the energies of the distinct lattice configurations. With the density of states the partition function,

$$Z = \sum_{i}^{\Omega} e^{\frac{-e_i}{k_B T}} = \sum_{j}^{\Pi} G(E_j) e^{\frac{-E_j}{k_B T}} \ , \tag{1}$$

can be determined and from it many important thermodynamic properties such as the free energy, entropy, specific heat, and ensemble averages calculated. In Eq. (1), Ω corresponds to the total number of possible configurations(Σ_i) and energies(e_i) referred to by the set $\{\Sigma_i, e_i\}_\Omega$, Π is the number of possible distinct energies E_j that appear in $\{\Sigma_i, e_i\}_\Omega$, k_B is Boltzmann's constant, and T is the temperature. If Ω is small enough one of the simplest ways to estimate $G_r(E_j)$ could be direct random sampling of the configuration space[2], which does have the advantage of being compeletly parrallel. For most problems Ω 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. The multicanonical method requires a careful choice of simulation paramaters while the entropic sampling requires 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. 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_LENDER ($B_L$end Each New Density Each Round) algorithm.

The Wang and Landau method does have parallel versions, including restricting random walkers to specific en-

* jdhoward@anl.gov

ergy 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$_L$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 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 *apriori*. Another issue is that the original Wang and Landau formulation for the reduction in the modification factor to the density of states has been shown be non-convergent[15–18]. There have been advancements made in understanding how to reduce the modification factor by Belardinelli et al. [19] whom developed the 1/t algorithm which is proven to be convergent, this result was verified by the work of Zhou et al. [18]. An issue with the 1/t algorithm as presented by Belardinelli et al. is that it requires a non-trivial preconditioning using what is referred to as the N-fold way. The N-fold way requires a careful analysis of the method to perturb the systems affect on the energy which for local models like the Ising model is possible but for first principles simulation this is not feasible. The novel aspects of the B$_L$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 is believed to be convergent based on its comparison with the 1/t algorithm and from mathematical analysis within an adiabatic assumption. The algorithm is also 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 density of states of the lithium ion conductor $Li_{0.5}La_{0.5}TiO_3$. There have been reports of the Wang and Landau algorithm used with linear scaling density functional theory methods to calculate magnetic properties of materials and order to disorder properties of alloys[20, 21]. There has also been a report of first principles calculations with replica exchange canonical Metropolis sampling for investigating ion disorder in solids by Kasamatsu et al. [22]. $Li_{0.5}La_{0.5}TiO_3$ is part of a family of possible stoichiometries $Li_{3x}La_{2/3-x}TiO_3$ of interest as solid state lithium ion conductors[23–29]. 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 $Li_{0.5}La_{0.5}TiO_3$. 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$_L$ENDER algorithm, section III covers the testing of the algorithm with the 2d Ising model including comparison with the original Wang and Landau algorithm and N-fold 1/t algorithm, section IV covers an adiabatic analysis of the alorithms convergence and the time dependence of the modification factor, section V covers the application of the algorithm to LLTO along with the computational details of the first principles methods, and section VI is the conclusions.

In this work the software package XMGRACE[30] was used in the generation of plots and VESTA[31] for the generation of structural images of LLTO.

II. ALGORITHM

The B$_L$ENDER algorithm proposed in this work is given as follows for $\mathcal{S}$ random walkers. 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. $G_r(E_j)^I,\ \{\Sigma_s, e_s\}^I_{\mathcal{S}}$
2. $\{\Sigma_s, e_s\}^I_{\mathcal{S}} \rightarrow \{\Sigma_s^{'}, e_s^{'}\}^I_{\mathcal{S}}$
3. $\Sigma_s^{'I}, e_s^{'I} \rightarrow \Sigma_s^{I+1}, e_s^{I+1}\ P = min[1, G_r(e_s)^I / G_r(e_s^{'})^I]$
   $else\ \Sigma_s^I, e_s^I \rightarrow \Sigma_s^{I+1}, e_s^{I+1}$
4. $G_r(E_j)^{I+1} =$

$$G_r(E_j)^I + \frac{C_o \mathcal{H}^I}{(A^I)^{\frac{1}{N}}} G_r(E_j)^I =$$

$$G_r(E_j)^I (1 + \frac{C_o \mathcal{H}^I}{(A^I)^{\frac{1}{N}}}) \tag{2}$$

Where $\mathcal{H}^I \equiv \mathcal{H}(E_j, \{e_s\}^I_{\mathcal{S}})$, that being an histogram function that counts the number of the currently vistied energies E_j in the set $\{e_s\}_{\mathcal{S}^I}$. The "instantaneous" histogram $\mathcal{H}^I$ is related to the total histogram of visited energies $H^I(E_j)$ by $H^I(E_j) = \sum_{i=0}^{I} \mathcal{H}^i$. The value $A^I \equiv \sum_j G_r(E_j)^I$. In this work $\{\Sigma_s, e_s\}^0_{\mathcal{S}}$ is a randomly drawn set of $\mathcal{S}$ walkers from the configuration

space $\{\Sigma_i, e_i\}_\Omega$. Also in this work the initial guess of the density of states was set to $G_r(E_j)^0 = 1 + \frac{C_o}{S}\mathcal{H}^0$, but in principle there could be other ways to intitialize the initial guess to the density of states. 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 $\Sigma_s^{'I}, e_s^{'I}$ goes to $\Sigma_s^{I+1}, e_s^{I+1}$, else the unperturbed configuration and energy Σ_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 inversely proportional to 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,

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

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 typical Wang and Landau sampling the update of the density of states is preformed by multiplication of the density of states by a factor greater than one every time an energy level is visited combined with a periodic reduction of the multiplication factor towards one 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 towards one as the simulation progresses. The multiplication form is also useful when Ω 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,

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

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,

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

In this form the algorithm can be implemented even when Ω is large. To implement the ratio of the density of states in step two of the algorithm,

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

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[32–34] and periodic boundary conditions. The configurations Σ_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 ± 1 and coupling constant J. Explicitly the energy e_i for a given configuration Σ_i of a n×n Ising lattice is given by,

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

A. Performance and properties of the B$_L$ENDER algorithm

The first test is 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 [35]. The accuracy of the simulation will be determined by the error defined as,

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

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

The first test of the algorithm is with the 32×32 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 32×32 Ising model so these results are included to show the algorithm has potential for larger system sizes. While the ideal value of N is not known prior to the calculation it was found in this work that a value of $1/N = 0.1$ was computationally efficient for the 32×32 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 $\mathcal{S}$. At $I =$ 1e7 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 $\mathcal{S} = 1000$, and at $\mathcal{S} =$ 1e4 there is marginal improvement at $I = 1e7$ but worse performance for $I <$ 1e6. 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 the results for the average error of 36 independent simulations of a 10×10 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 would be typical for a supercell approximation in a density functional theory calculation is not expected to exceed the large number of $\approx 10^{30}$ configurations in the 10×10 Ising model. Although in the work of Kahn et al. [21] they did tackel a configuration space of size of 10^{74} using linear scaling density functional theory for a 250 atom supercell. Typical density functional theory calculations scale roughly as the cube of the number of atoms, so their calculation was only feasible using the linear scaling methods.

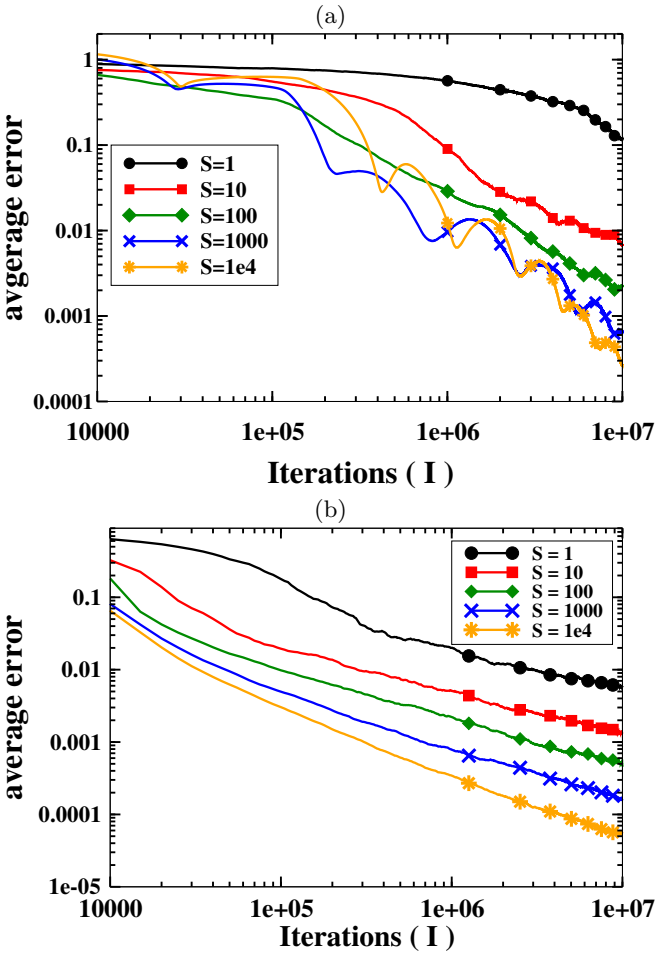

FIG. 1. (color online) Average error from 36 simulations calculated with Eq. 8 with $\mathcal{S} = 1$, 10, 100, 1000, and 1e4 for (a) the 32×32 Ising model with $1/N = 0.1$, (b) the 10×10 Ising model with $1/N = 1$.

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 32×32 and 10×10 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 32×32 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 32×32 model with $\mathcal{S} = 10$ and the smaller 10×10 model for both $\mathcal{S} = 10$ and $\mathcal{S} = 100$. The more pronounced convergence dependence on N for the larger 32×32 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 10×10 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 32×32 and 10×10 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 errors 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.

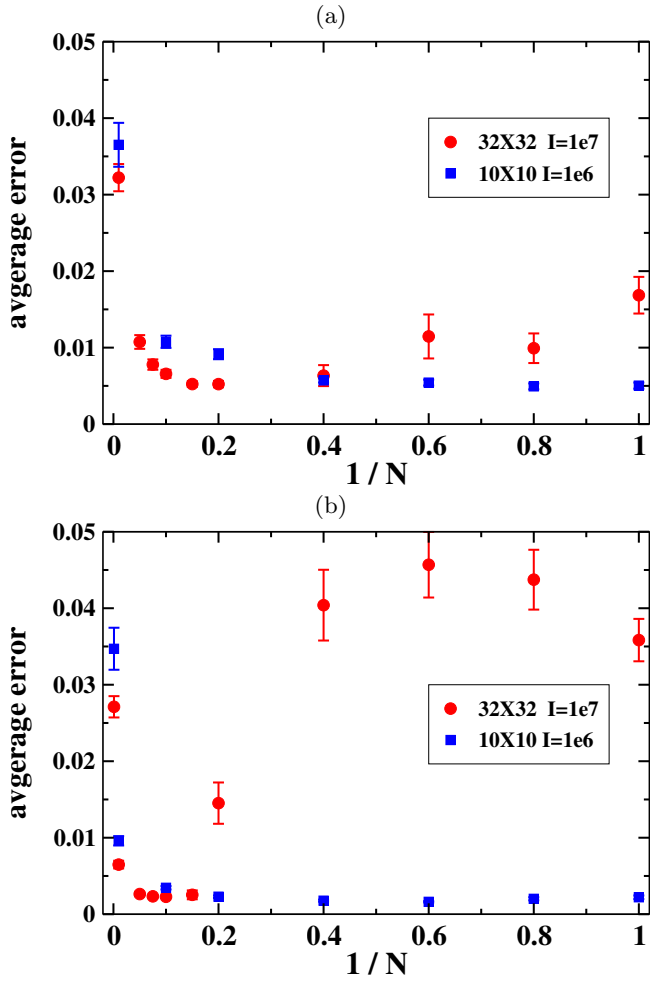

FIG. 2. (color online) Average error calculated from Eq. 8 from 36 simulations for (a) $\mathcal{S} = 10$ and (b) $\mathcal{S} = 100$ vs the value of $1/N$ for the 32×32 (red circles) and 10×10 (blue squares) Ising models simulated to 1e7 and 1e6 iterations respectively. Errorbars show the standard deviation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

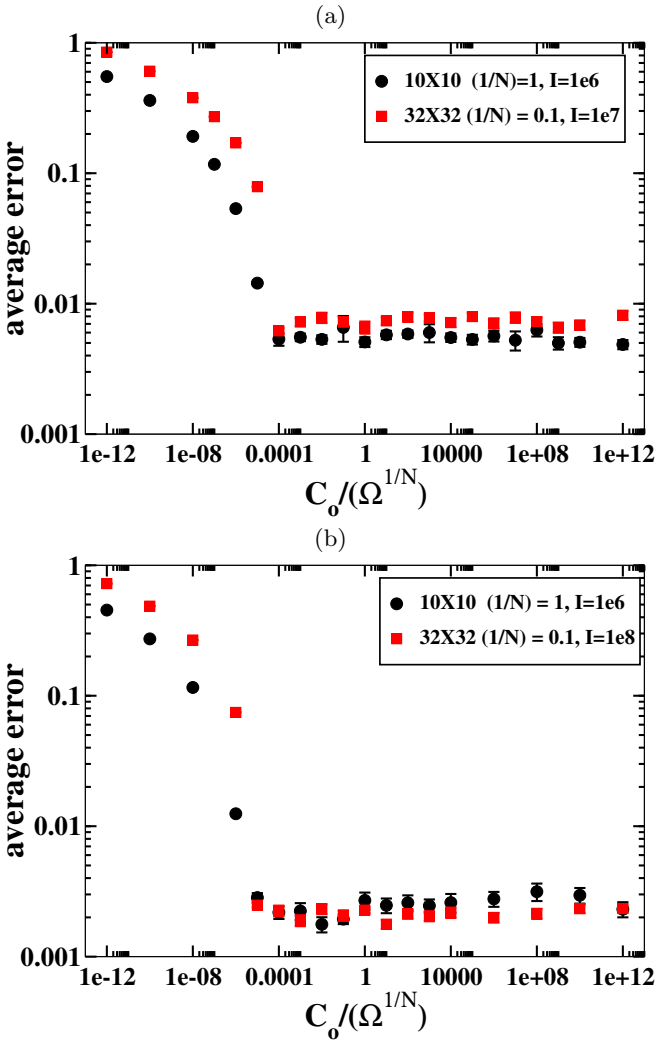

FIG. 3. (color online) Average error from 36 simulations calculated from Eq. 8 for the 10×10(black circles) and 32×32(red squares) Ising model simulated to 1e6 and 1e7 iterations and with $1/N = 1$ and $1/N = 0.1$ respectively for (a) $\mathcal{S} = 10$ and (b) $\mathcal{S} = 100$ vs the value of $C_o/\Omega^{1/N}$.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

B. Comparison with original Wang and Landau algorithm

In Wang and Landau's original work they simulated 32×32 and 50×50 2d Ising models for comparison to the exact density of states. The comparison in this section will use the 32×32 Ising model. To implement the original Wang and Landau algorithm the same parameters from their original paper will be used[9]. First we define the historgram $H(E_j)$ to be the histogram of the visited(accepted) energies during the simulation. Next define the modification factor f such that everytime a energy level is visited(accepted) then $G_r(E_j)^{I+1} = G_r(E_j)^I f$. In Wang and Landau's original work they used a flatness criteria of 80% and a reduction schedule of $f^{I+1} = \sqrt{f^I}$. The flatness criteria was described such that all energies E_j had been visited and that $min[H(E_j)]/avg[H(E_j)] > 0.8$. When the flatness criteria is meet the reduction schedule $f^{I+1} = \sqrt{f^I}$ is then implemented. The value of $f^0 = e^1$ was used. Initial configurations were generated randomly and $G_r(E_j)^0$ was set to 1 when implementing the original Wang and Landau algorithm. In the this manner the B$_L$ENDER algorithm with $1/N = 0.1$ and $\mathcal{S} = 1$ was compared to the perfomance of the original Wang and Landau algorithm. The results are shown in Fig. 4(a), which are averaged over 36 independant simulations. The comparison shows a striking difference in the error, the B$_L$ENDER algorithm tends to decrease over the entire time span while the original Wang and Landau implementation rises to a large peak early in the simulation then abrutly drops. At large number of iterations the error is nearly the same for both. To compare potential parrallel performance of the orginal Wang and Landau algorithm in a scheme similar

to the B$_L$ENDER algorithm the update,

$$G_r(E_j)^{I+1} = G_r(E_j)^I f^{\mathcal{H}(E_j,\{e_s\}_{\mathcal{S}}^{I+1})} , \qquad (9)$$

is adopted. Where $\{e_s\}_{\mathcal{S}}^{I+1}$ and $\mathcal{H}(E_j,\{e_s\}_{\mathcal{S}}^{I+1})$ are generated and function the same was an in Eq. 2. In this manner a simulation was carried out with 100 walkers ($\mathcal{S} = 100$) for the Wang and Landau alogrithm, with the same flatness criteria and reduction schedule as above, and for the B$_L$ENDER algorithm to 1e7 iterations. The results are shown in Fig 4(b), which are averaged over 36 independant simulations. The results show that both parallel implementations have siginificant improvement for equivalent number of iterations but that the original Wang and Landau still suffers the problem of rising to a large peak and then dropping suddenly. Both algorithms are shown to have similar errors at a large number of iterations.

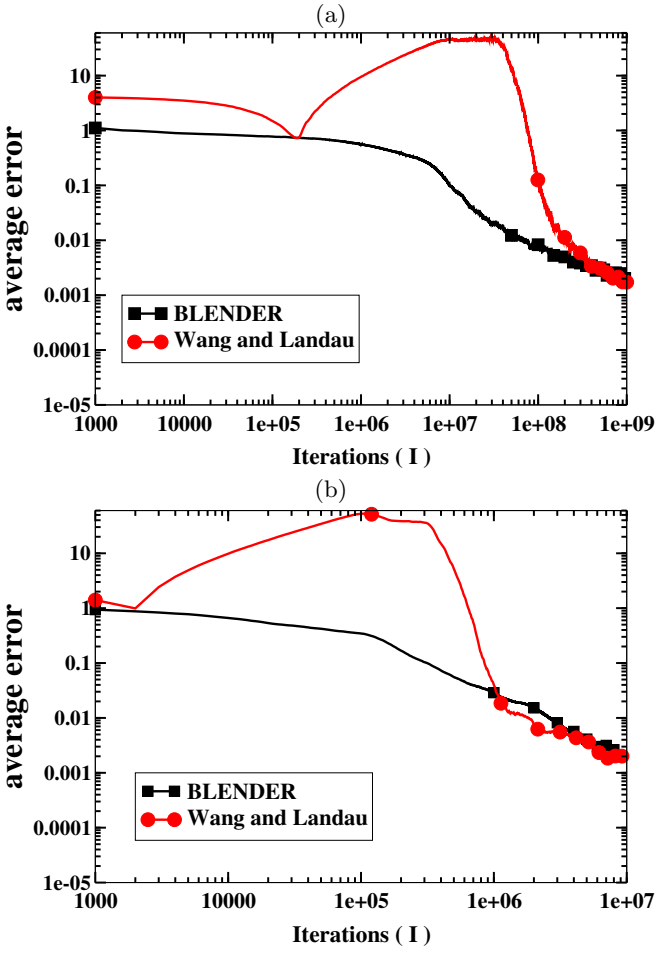

FIG. 4. (color online) Average error from 36 simulations for the orginal Wang and Landau algorithm in red with circles and the B$_L$ENDER algorithm in black with squares for (a) $\mathcal{S} = 1$ simulated to 1e9 iterations and (b) $\mathcal{S} = 100$ simulated to 1e7 iterations.

C. Relationship of the B$_L$ENDER algorithm to the N-fold 1/t algorithm

In the introduction it was mentioned that an improvement to the update to the multiplication factor in the Wang and Landau method had been made by Belardinelli et al. [19], 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 consider the generic multiplication upate of the density of states

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

For the B$_L$ENDER algorithm with $C_o = \Omega$ and $N = 1$ $f(I)$ would take the form,

$$f(I) = (1 + \frac{\Omega\mathcal{H}(E_j,\{e_s\}_{\mathcal{S}=1}^{I+1})}{[\sum_j G_r(E_j)^I]}) . \qquad (11)$$

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 defined for a single walker as,

$$t = \frac{I}{\Pi} . \qquad (12)$$

Where Π is the number of energies E_j, which is $n^2 - 1$ for the $n \times n$ Ising model. They found the optimal values to be $c = 1$ and $p = 1$. To be effective though they found this algorithm required preconditioning with the original Wang and Landau algorithm. They also described a more complex variant of this algorithm called the N-fold 1/t which incorporated the average life time of the initial configurations early in the simulation which then limited to the 1/t form. In their work they used a 8×8 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 Fig. 5(a) and the average error in Fig. 5(b). In comparing the results to those taken from Fig. 5 of Belardinelli et al. [19] shown in Fig. 5 of this work, a striking similarity is seem 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. 5(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. 5(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.

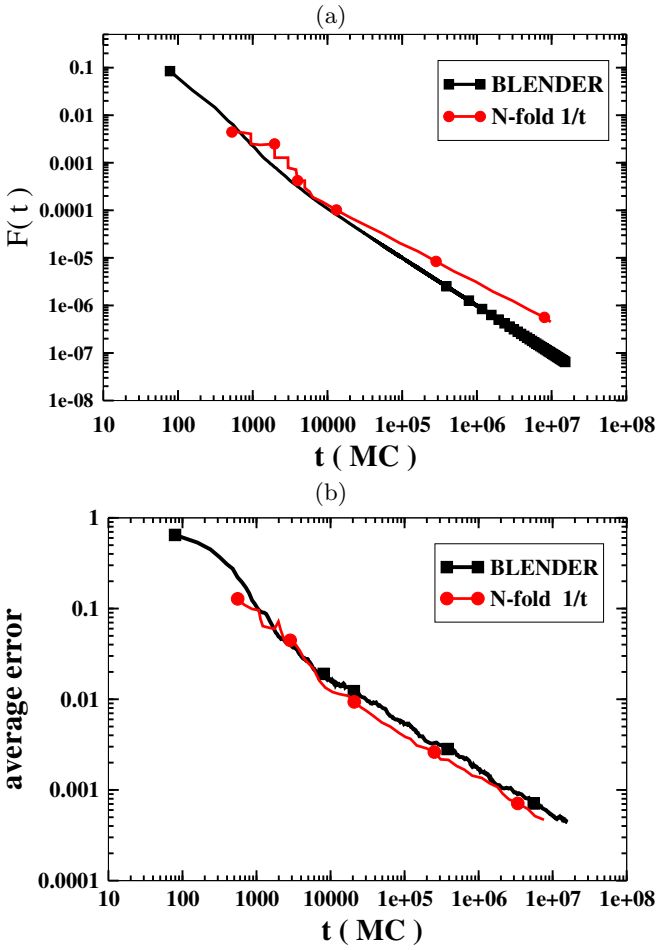

FIG. 5. (color online) In black with squares (a) The average of $F(t)$ and (b) the average error defined by Eq. 8 from 36 simulations of the 8×8 Ising model with the B$_L$ENDER algorithm with $\mathcal{S} = 1$, $N = 1$, and $C_o = \Omega$. In red with circles (a) $F(t)$ results taken from the work of Belardinelli et al. [19] (b) the error defined by Eq. 8 taken from the work of Belardinelli et al.. Here $t = \frac{I}{63}$ or equivalently one Monte Carlo time step (MC) as defined by Eq. 12.

IV. ADIABATIC ANALYSIS OF THE ALGORITHM

To better understand the nature of the algorithm we will consider the adiabatic properties of the histogram $\mathcal{H}^I \equiv \mathcal{H}(E_j, \{e_s\}^I_{\mathcal{S}})$. Considering that during a simulation the configurations are generated with probability proportional to $G(E_j)/\Omega$ and accepted inversely proportional to $G_r(E_j)^I/A^I$ the probability distribution, which we will write as $\Phi^I(E_j)$, of the $\mathcal{H}^I$ will be proportional to $[G(E_j)A^I]/[G^I_r(E_j)\Omega]$. For short we will write $G^I_r(E_j) \to G^I_r$, $G(E_j) \to G$, and $\Phi^I(E_j) \to \Phi^I$. Considering that the sum over $\mathcal{H}^I$ is constrained to $\mathcal{S}$ if we normalize Φ^I to $\mathcal{S}$ we get

$$\Phi^I = \mathcal{S}(\sum_j \frac{G}{G^I_r})^{-1}\frac{G}{G^I_r} \,. \tag{13}$$

Now an adiabatic analysis will be considered by inserting this expression for the adiabatic distribution of the histogram $\mathcal{H}^I$ into density of states update of the algorithm (step four). Doing this gives,

$$\begin{aligned} G^{I+1}_r = G^I_r + \mathcal{S}(\sum_j \frac{G}{G^I_r})^{-1}\frac{G}{G^I_r}\frac{C_o G^I_r}{(A^I)^{1/N}} = \\ G^I_r + C_o\mathcal{S}(\sum_j \frac{G}{G^I_r})^{-1}\frac{G}{(A^I)^{1/N}} \,, \end{aligned} \tag{14}$$

which will be the basis for the adiabatic analysis of the algorithm. One point to clarify is what is meant by adiabatic. A rigorous definition of adiabaticity for the algorithm can be defined that for any ϵ an I and n can be found such that,

$$\langle \sum_j |\frac{1}{n+1}\sum_{i=I}^{I+n}\mathcal{H}^i - \Phi^I|\rangle_o < \epsilon \,, \tag{15}$$

where $\langle\rangle_o$ means average over trajectories of simulation. This is a statement that the algorithm will progressively scan the energy range more and more thoroughly before a significant change to the density of states is made. From here the analysis will be continued assuming the algorithm is adiabatic.

To continued the analysis we will first determine if and under what circumstances the algorithm will converge assuming adiabacitiy. The first step is to note that,

$$\mathcal{S}C_o(\sum_j \frac{G}{G^I_r})^{-1}\frac{1}{(A^I)^{1/N}} \,, \tag{16}$$

is an iteration dependent constant(same for each bin j in the density of states), we will call this constant B^I. In this manner we will look at the progression of changes to the relative density of states,

$$\begin{aligned} G^{I+1}_r &= G^I_r + GB^I \\ G^{I+2}_r &= G^I_r + GB^I + GB^{I+1} \\ &\cdots \\ &\cdots \\ G^{I+n}_r &= G^I_r + G\sum_{i=0}^{n-1} B^{I+i} \end{aligned} \tag{17}$$

For short $\sum_{i=0}^{n-1} B^{I+i}$ will be referred to as W^n The condition for convergence can be defined such that for two bins of the density of states l and k,

$$\lim_{n\to\infty}\frac{G^I_r(E_k) + G(E_k)W^n}{G^I_r(E_l) + G(E_l)W^n} \to \frac{G(E_k)}{G(E_l)} \,. \tag{18}$$

For this to occur W^n must increase unbounded as $n \to \infty$ and not limit to zero or a constant. To make further

progress we will first rewrite B^I as ,

$$\begin{aligned} \mathcal{S}C_o(\sum_j \frac{G}{G_r^I})^{-1}\frac{1}{(A^I)^{1/N}} = \\ \mathcal{S}C_o(\sum_j \frac{G}{G_r^I})^{-1}\frac{1}{A^I(A^I)^{1/N-1}} = \\ \mathcal{S}C_o(\sum_j \frac{G}{G_r^I})^{-1}\frac{(A^I)^x}{A^I} \ , \end{aligned} \tag{19}$$

where $x = 1 - 1/N$. A sufficient condition to show convergence would be to show that the terms B^{I+i} do not go to zero as $\lim_{n\to\infty}$. In this context the lower bound of B^I will be studied. Considering that $\frac{min(G_r^I)}{\Omega} \leq (\sum_j \frac{G}{G_r^I})^{-1}$ and that $\frac{min(G_r^I)}{\Pi max(G_r^I)} \leq \frac{min(G_r^I)}{A^I}$ we can bound B^I as,

$$\frac{\mathcal{S}C_o}{\Pi\Omega}\frac{min(G_r^I)}{max(G_r^I)}(A^I)^x \leq B^I \leq \frac{\mathcal{S}C_o}{\Pi}(A^I)^x \tag{20}$$

For the case of $0 \leq x < 1$ convergence is clear because $\frac{min(G_r^I)}{max(G_r^I)}$ can not limit to zero and $(A^I)^x \geq 1$. For the case of $x < 0$ convergence is not so clear although in theory the algorithm may still be convergent based on contradiction. In Eq 20 the upper bound on B^I tells us that for $x < 0$ if the algorithm does converge the B^I will tend towards zero since A^I must grow unbounded if the algorithm converges. This means that if the algorithm does converge for $x < 0$ that the sum of the B^I terms must increase unbounded although they tend towards zero. On the other hand if the algorithm does not converge, i.e. the sum of the B^I terms tends towards zero, then A^I will not grow unbounded. This is a contradiction because if A^I does not grow unbounded then the B^I terms can't tend towards zero and the sum of them should grow unbounded and the algorithm should converge. This suggests that in theory the algorithm is convergent for $x < 0$ but it is dependent on the sum of terms that tend towards zero, which is essentially predicting slow convergence. In practice tests with the 8×8 Ising model reasonable convergence could only be achieved for "small" negative values of x.

So now considering the algorithm does converge within the adiabatic assumption for $0 \leq x < 1$ we can derive the time dependence of $F(I) = \ln[f(I)]$ where $f(I) = (1 + \frac{G}{G_r^I}B^I)$ within this analysis. First rewriting $\frac{G}{G_r^I}B^I$ as $\frac{C_o\mathcal{H}^I}{(A^I)^{1/N}}$ it is clear that $\frac{G}{G_r^I}B^I$ becomes arbitrarily small for $I \to \infty$. With $\ln(1+a) \approx a$ for small a we now have for large I,

$$\ln[F(I)] \approx \frac{GB^I}{G_r^I} \ . \tag{21}$$

Also because the algorithm is converging G_r^I is approaching $W^I G$ and $\sum_j G_r^I = A^I$ is approaching $W^I\Omega$ such that $(\sum_j \frac{G}{G_r^I})^{-1}\frac{1}{A^I} \approx \frac{W^I}{\Pi\Omega}$. With this we write in the large iteration limit $B^I \approx \frac{C_o\mathcal{S}}{\Pi\Omega}(A^I)^x$ and $G_r^I \approx G\frac{C_o\mathcal{S}}{\Pi\Omega}\sum_{i=0}^{I-1}(A^i)^x$ such that,

$$\ln[F(I)] \approx \frac{(A^I)^x}{\sum_{i=0}^{I-1}(A^i)^x} \ . \tag{22}$$

In this form it is clear that the time dependence for the case of $x = 0$ is of $1/I$ form which is what was found in comparison with the $1/t$ algorithm in the previous section. For $0 < x < 1$ the time dependence is not as clear analytically but it can be simulated by using a boot strapping procedure. Assuming we start with $G_r^0 = cG$ so that $A^0 = c\Omega$ then,

$$\begin{aligned} A^0 &= c\Omega \\ A^1 &= A^0 + \frac{C_o\mathcal{S}}{\Pi}(A^0)^x \\ A^2 &= A^1 + \frac{C_o\mathcal{S}}{\Pi}(A^1)^x \\ &..... \\ A^I &= A^{I-1} + \frac{C_o\mathcal{S}}{\Pi}(A^{I-1})^x \\ &. \end{aligned} \tag{23}$$

In this way a numerical simulation can be done to predict the value of $\ln[F(I)]$ in Eq. 22.

V. 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 the B$_L$ENDER algorithm 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 and the implementation of the algorithm is relatively simple. In this work the B$_L$ENDER algorithm is applied to the lithium and lanthanum sublattice of the solid state lithium ion electrolyte $Li_{0.5}La_{0.5}TiO_3$. The goal of this study was to both, perform a calculation with the B$_L$ENDER algorithm 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 P4/mmm phase of the stoichiometry $Li_{0.5}La_{0.5}TiO_3$[24, 28]. A unit cell of this structure is shown in Fig. 6. The lattice parameters for this unit cell were taken from the experimental results

from Ibarra et al.[24]; 3.8688(4) $\mathring{A}$ for a- and b-axes, and 7.7463(2) $\mathring{A}$ for c-axis. This unit cell is representative of an ordered form of $Li_{0.5}La_{0.5}TiO_3$ 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 ABX_3. The structure in Fig. 6 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. 6 is evidenced by the imaginary phonon modes calculated by Moriwake et al. [25].

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 [24]. In this work the B_LENDER algorithm is used to evaluate the density of states associated with local minimum corresponding to the lithium and lanthanum ordering and associated lattice distortions.

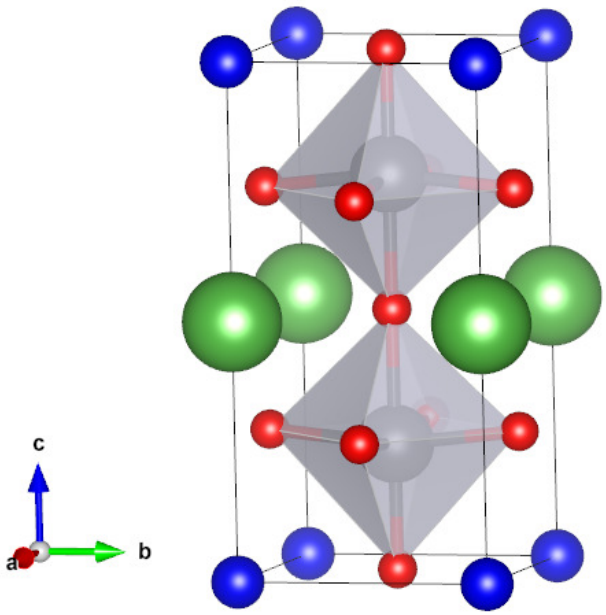

FIG. 6. (color online) 10 atom unit cell of P4/mmm $Li_{0.5}La_{0.5}TiO_3$. Where dark blue spheres are lithium, green spheres are lanthanum, red spheres are oxygen, and grey spheres inside of octahedra are titanium.

B. Computational Details

In this work a 3×3×1 90 atom supercell with periodic boundary conditions of the unit cell depicted in Fig. 6 was used as an approximation to bulk $Li_{0.5}La_{0.5}TiO_3$. 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 perturbed configurations in the iterative process of the B_LENDER 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 ± 0.2 $\mathring{A}$ 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 perturbed configurations 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 where then relaxed to a local minimum. The random noise off the A-site served to assist 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) [36–39] within the projector augmented wave formalism (PAW)[40]. The Perdew-Burke-Ernzerhof (PBE) variant of generalized gradient approximation was used for the exchange and correlation functional[41]. The valence electron configurations for the PAW data sets were; $5p^6 5d^1 6s^2$ for La, $2s^1$ for Li, $3p^6 3d^2 3s^2$ for Ti, and $2s^2 2p^4$ 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 $<$ 2.5e-5 eV and relaxation of atomic coordinates was terminated when the difference in total energy between ionic relaxation steps was $<$ 2.5e-4 eV. Electronic occupations used Guassian smearing with a width of 0.05 eV. A 1×1×2 gamma centered k-point mesh was used for the 3×3×1 supercells of the LiLaTiO6 unit cell. These cutoffs and parameters were chosen to maximize computational efficiency while retaining enough accuracy to capture important physical properties of LLTO. Each calculation of an energy was completed with a 36 processor broadwell node with an average wall time of 15 minutes per calculation. 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; $5s^2 5p^6 5d^1 6s^2$ for La, $1s^2 2s^1$ for Li, $3p^6 3d^2 3s^2$ for Ti, and $2s^2 2p^4$ for O. The cutoffs for the more accurate calculations were 450 eV for the plane wave basis, and 2×2×3 gamma centered k-points. The average magnitude in relative energy between structures from this test was 0.05 eV.

An important point to make about the calculations in this research is that they included 0 K static lattice internal energies, did not include phonon free energies, and they did not take into account relaxation of lattice parameters[42]. Ideally for fixed lattice parameters we would evaluate the parition function[2],

$$Z = \sum_{i=1}^{\Omega} e^{-\frac{(u_i + f_i(T))}{K_b T}} \ . \tag{24}$$

Where T is the temperature, k_B is Boltzmann's constant, u_i is the static lattice internal energy for configuration i, and $f_i(T)$ is the temperature dependant phonon free energy for configuration i. A configuration is defined in this work as a local minimum of the Born Oppenheimer potential energy surface. This form of the parition function poses the problem that the density of states is now temperature dependant. Take $u_i + f_i(T) \equiv e_i(T)$, let W be the number of unique $e_i(T)$, and $F_j(T)$ be the unique $e_i(T)$ then,

$$Z = \sum_{j=1}^{W} G(F_j(T)) e^{-\frac{F_j(T)}{k_b T}} \ . \tag{25}$$

This is a problem because to employ the Monte Carlo methods discussed in this work they would have to be applied at different temperatures which defeats the original purpose. This problem could be addressed by taking U_j to be the unique u_i, Π to be the number of U_j, and then considering a different form of the parition function in Eq. 24,

$$Z = \sum_{j=1}^{\Pi} < e^{-\frac{f_i(T)}{k_B T}} >_j G(U_j) e^{-\frac{U_j}{K_b T}} \ . \tag{26}$$

Here $< \exp(-f_i(T)/k_B T) >_j$ is the arithmetic average of $\exp(-f_i(T)/k_B T)$ over all configurations with static lattice internal energy U_j. In this form the static lattice density of states $G(U_j)$ could first be determined with the Monte Carlo methods described in this work and then the $< \exp(-f_i(T)/k_B T) >_j$ could then be approximated by randomly choosing some of the perturbed configurations for each U_j to evaluate the phonon density of states. With the randomly generated phonon densities of states the $< \exp(-f_i(T)/k_B T) >_j$ could be approximated for any temperature. Even in this form though the computational expense is beyond the scope of this work as phonon calculations require high cutoffs and even within the harmonic approximation are much more computationally expense than static lattice calculations[43]. So in this work the approximation,

$$Z \approx \sum_{j=1}^{\Pi} G(U_j) e^{-\frac{U_j}{K_b T}} \ , \tag{27}$$

is made. This is reasonable because the phonon free energies are expected to vary much less from configuration to configuration than the static lattice internal energies. A final comment is that relaxation of the lattice paramters could be accomplished by calculating the partition function on a grid of lattice constants, then a free energy surface could be interpolated and the lattice paramaters minimizing the free energy determined as a function of temperature. This procedure is also beyond the scope of the current research.

The calculations were performed at the experimental lattice parameters 3.8688 $\mathring{A}$ for a- and b-axes, and 7.7463 $\mathring{A}$ for c-axis. The parameters for the B_LENDER algorithm were $\mathcal{S} = 10$ and $N = 1$. The bin width used for determining $G_r(E_j)$ was chosen to be 0.02 eV. The value of Ω was estimated as 100 times the combinatoric number of configurations of the lithium and lanthanum ordering onto the A-site given as,

$$\Omega \approx 100 \frac{18!}{9!9!} \ . \tag{28}$$

While an exact value of Ω 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 Ω 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 [25, 29, 44, 45]. 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 $6^9 \approx 1e7$. So the approximation $\Omega \lessapprox 1e7 \frac{18!}{9!9!}$ can be made. It is physically reasonable that a significant fraction of these configurations will be unstable so that the estimate of Ω in Eq. 28 is likely to be within several orders of magnitude of Ω.

C. Results

Using the parameters and configurational enumeration scheme specified above a simulation was performed to 3,000 iterations for the 3×3×1, 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[46] whom tracked thermodynamic

quantities such as the peak of specific heat to determine convergence. In this work convergence is limited by computational resources. 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 occupancy of lanthanum in the lanthanum rich layer along the c-axis. In the work by Ibarra et al. [24] 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. 6 would have $La1 = 1$. It is important to note in this work the 3×3×1 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 4×4×1 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 perturbed configurations that occurred during the simulation. The arithmetic average of a general order parameter O over all configurations with energy E_j is denoted by $< O >_j$. Then with these the ensemble average is computed as,

$$\langle O \rangle = \sum_{j=1}^{\Pi} < O >_j \, G_r(E_j) \frac{e^{-\frac{E_j}{k_B T}}}{Z} \, . \qquad (29)$$

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 is a view of the convergence of $G_r(U_j)$ as a function of the iterations. Here we have switched to using U_j to highlight that the calculations only include 0 K static lattice internal energies as discussed before. In Fig. 7, $G_r(U_j)$ is shown at $I = 500$, 1000, and 3,000 with the y-axis plotted on a log scale. The $G_r(U_j)$ shown in Fig. 7 are plotted such that the lowest energy of $G_r(U_j)$ 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 $G_r(U_j)$ at that iteration. The main characteristic of the results by 3,000 iterations is the presence of some low energy states in a gap between the lowest energy state and a continous spectrum of states, this gap may due to insufficient sampling at lower energies or finite size effects.

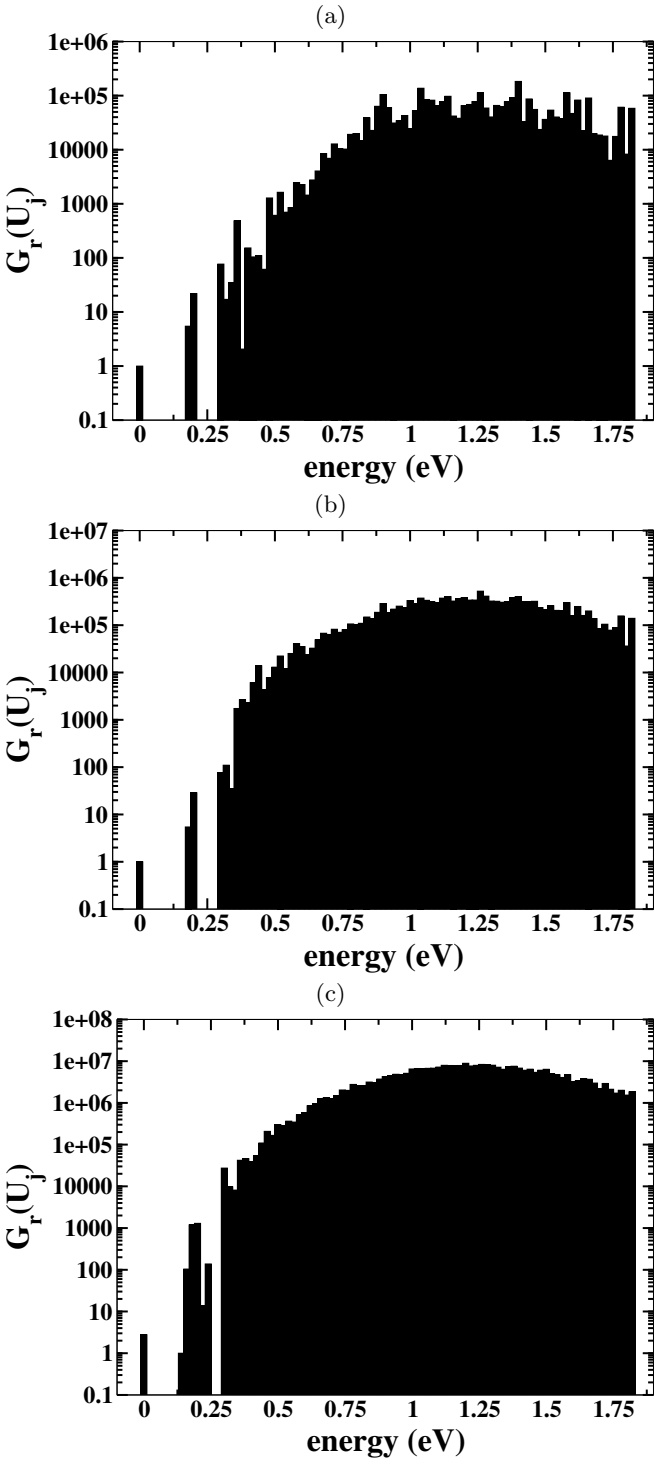

FIG. 7. Plots of $G_r(U_j)$ at (a) 500 iterations, (b) 1,000 iterations, and (c) 3,000 iterations. The plots are normalized by dividing through by the smallest value of $G_r(U_j)$ at that particular iteration. The plots are shown with a log scale on the y-axis.

It is noted, as the iterations increase, that $G_r(U_j)$ for the spectrum of higher energy states becomes noticeably smoother. The lowest energy configuration is character-

ized 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. 6, in that the structure has distinct lattice distortions.

The next result is the arithmetic averages of the $La1$ order parameter, which are shown in Fig. 8(a) along with the number of samples used to determine each value in Fig. 8(b). The results in Fig. 8(a) show an overall tendency for more mixing of lithium and lanthanum between layers for higher energies. The ensemble average of $La1$ is shown in Fig. 9, which shows a phase transition from completely segregated lithium and lanthanum between layers to mostly mixed between layers and increased mixing with increasing temperature. In Fig. 9 the values for $La1$ range from approximately 0.675 at 500 K to 0.625 at 1500 K. For the 3×3×1 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. [24] 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.

Some indication of convergence of the 3,000 iterations comes from inspecting the flatness of the histogram of visited energies 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 is shown in Fig. 11. The results show that the histogram is qualitatively flat for $\gtrapprox$ 0.5 eV. This result along with the counts for calculating the $La1$ order parameter shown in Fig. 8(b) suggest that the results are most well converged for $\gtrapprox$ 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 $\gtrapprox$ 0.5 eV. While 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, the trend seen in figure Fig. 8(a) spans an energy range much greater than the expected relative error in energies between structures. 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. 9 can not be expected to be an accurate prediction of a transition temperature. The most important result of Fig. 9 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 into the structures found during the simulation the two lowest energy structures are shown in Fig. 10(a) and (b). In Fig. 10(a) it is seen that the lowest energy structure has the lithium and lanthanum completely segregated between layers along the c-axis. In Fig. 10(b) 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

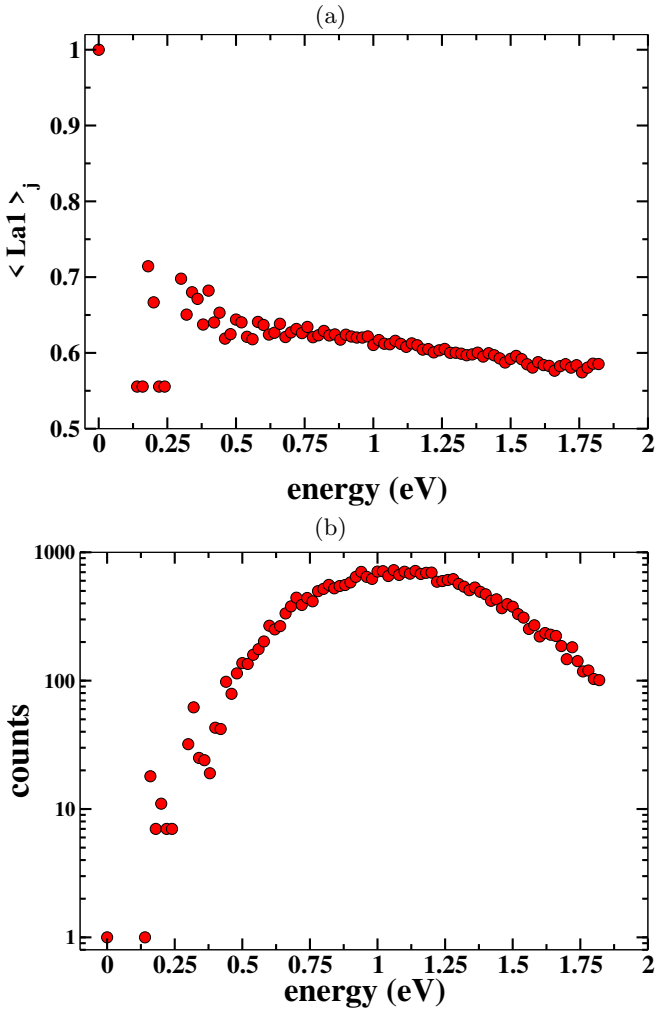

FIG. 8. (a) Arithmetic averages of the $La1$ order parameter as a function of energy. (b) Number of counts to determine each value of the $La1$ order parameter for each energy.

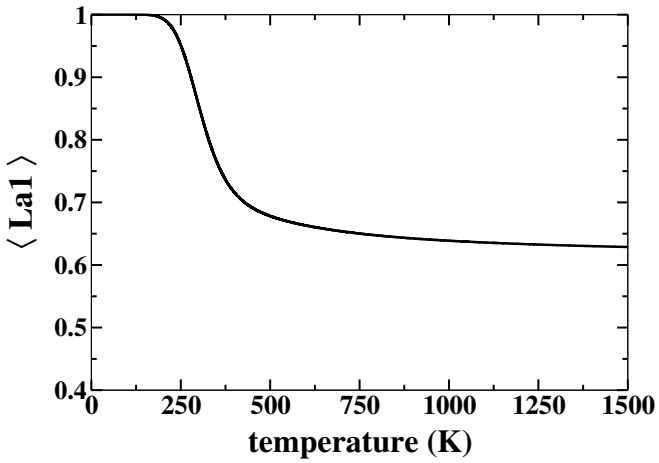

FIG. 9. Ensemble average of the $La1$ order parameter calculated with Eq. 29 as function of temperature.

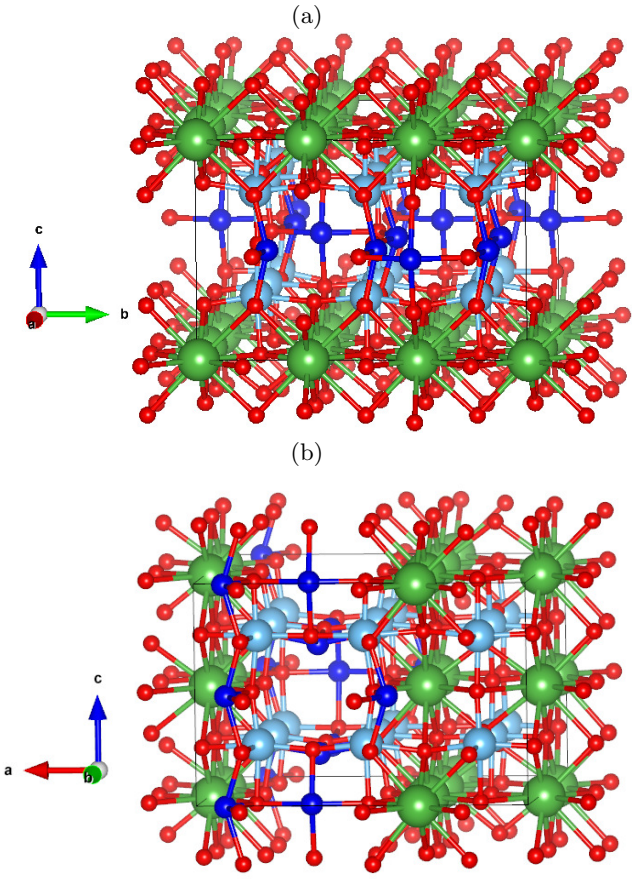

FIG. 10. (color online) (a) Lowest energy structure found during simulation. (b) Second lowest energy structure found during the simulation. Where dark blue spheres are lithium, green spheres are lanthanum, red spheres are oxygen, and light blue spheres are titanium.

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[23, 25]. Due to the 3×3×1 supercell used in this study these structures were not possible to realize in the simulation but as evidenced by Fig. 10(b) segragation of the lithium and lanthanum into separate layers is still energetically favorable. A common feature of both Fig. 10(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[25, 29, 44, 45].

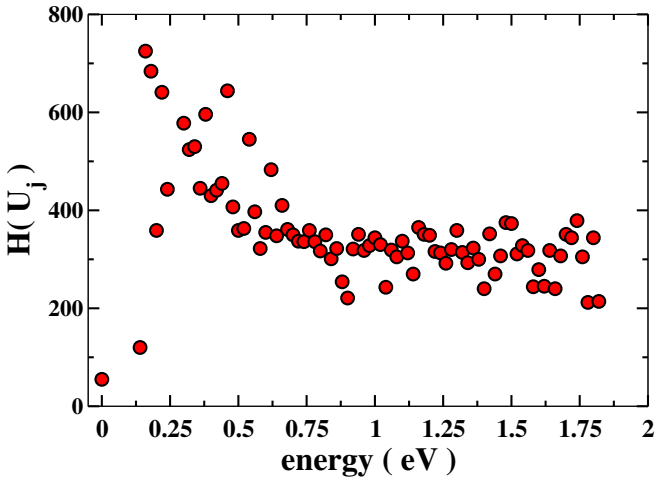

FIG. 11. Histogram of the visited energies over the 3,000 iterations of the simulation. Where visited here means the accepted energies as per step 3 of Eq. 2.

VI. CONCLUSIONS

This work has presented a parallel variant of the Wang and Landau algorithm referred to as B_LENDER($B_L$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 10×10 Ising model with a minimal number of implementation parameters. Results for the 32×32 Ising model suggest the algorithm could have applicability to larger system sizes provided a tuning parameter is chosen appropriately to maximize performance, currently it is not known how to choose this parameter before hand without numerical testing. Comparing performance with the the original Wang and Landau algorithm showed that while B_LENDER tends to decrease in error over the simulation while the error in the original Wang and Landau rises to a peak early in the simulation and then drops suddenly. Long time performance of the B_LENDER algorithm and the original Wang and Landau algorithm was found to be comparable. In the case of using one walker the B_LENDER algorithm was also found to have similar behavior as to the N-fold 1/t algorithm of Belardinelli et al.[19]. 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 $Li_{0.5}La_{0.5}TiO_3$ 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 a 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] C. Zhou and J. Su, Phys. Rev. E **78**, 046705 (2008).
[19] R. E. Belardinelli and V. D. Pereyra, The Journal of Chemical Physics **127**, 184105 (2007), https://doi.org/10.1063/1.2803061.
[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] S. Kasamatsu and O. Sugino, Journal of Physics: Condensed Matter **31**, 085901 (2019).
[23] 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).
[24] J. Ibarra, A. Vrez, C. Len, J. Santamara, L. Torres-Martnez, and J. Sanz, Solid State Ionics **134**, 219 (2000).
[25] 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).
[26] H. Geng, A. Mei, C. Dong, Y. Lin, and C. Nan, Journal of Alloys and Compounds **481**, 555 (2009).
[27] 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.
[28] S. Stramare, V. Thangadurai, and W. Weppner, Chemistry of Materials **15**, 3974 (2003), https://doi.org/10.1021/cm0300516.
[29] M. Catti, Chemistry of Materials **19**, 3963 (2007), https://doi.org/10.1021/cm0709469.
[30] XMGRACE is a two-dimensional plotting software package maintained at the webpage `http://plasma-gate.weizmann.ac.il/Grace`.
[31] K. Momma and F. Izumi, Journal of Applied Crystallography **44**, 1272 (2011).
[32] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Elsevier, 2016) iSBN 978-0-12-083180-7.
[33] L. Onsager, Phys. Rev. **65**, 117 (1944).
[34] B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model* (Harvard University Press, 1973).
[35] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).
[36] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
[37] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).
[38] G. Kresse and J. Furthmller, Computational Materials Science **6**, 15 (1996).
[39] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[40] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[41] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[42] J. D. Howard, Z. D. Hood, and N. A. W. Holzwarth, Phys. Rev. Materials **1**, 075406 (2017).

[43] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[44] 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.
[45] J. A. Alonso, J. Sanz, J. Santamara, C. Len, A. Vrez, and M. T. Fernndez-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.
[46] A. A. Caparica, Phys. Rev. E **89**, 043301 (2014).