First principles calculation of configurational energy density of states in 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 16, 2019)

In this work a variant of the Wang and Landau algorithm for calculation of the configurational energy density of states is proposed. The algorithm is referred to as B_LENDER, which is an acronym for $B_L$end Each New Density Each Round and an adjective for how it was created and functions. 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 sub lattices in crystal materials. In this work the algorithm is tested with the 2d Ising model to bench mark the algorithm. This was done in the spirit of a "Fermi" problem to make a prediction of the wall time and core hours necessary for the algorithm to evaluate the partition function of a 3×3×3 supercell of disordered cubic Li_2OHCl.

INTRODUCTION

For crystalline materials with disordered sub-lattices such as the Li-ion solid state electrolyte Li_2OHCl[1–5] it is desirable to calculate from first principles methods(such as density functional theory[6]) the density of energy states $G(E_j)$. Here the energy density of states is meant to correspond to the energies of the distinct lattice configurations. With the energy density of states the partition function,

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

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), Ω corresponds to the number of possible configurations and energies in the set $\{\Sigma_i, e_i\}_\Omega$, Π to number of possible distinct energies E_j, k_B is Boltzman's constant, and T is the temperature. One method to solve this problem could be temperature dependent simulations involving the Metropolis algorithm and histogram re-weighting techniques[7, 8]. Another algorithm called the Wang and Landau algorithm[9] has been developed which is temperature independent. An issue with these algorithms 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 a method is proposed that combines the use of random sets along with the importance sampling method of the Wang and Landau algorithm that is meant to work towards the goal of highly parallel importance sampling algorithms that mesh well with high performance computing architectures. The algorithm developed in this work is referred to as the B_LENDER ($B_L$end Each New Density Each Round) algorithm. The name B_LENDER functions as an adjective as well as an acronym. This comes from how it blends the ideas of a random set and the Wang and Landau method of sampling with problity proportional to the inverse of the density of states, and also due to the nature of the algorithm iteratively blending histograms to produce a converged density of states. The Wang and Landau method does have parallel versions, including restricting random walkers to specific energy ranges or allowing the walkers to explore the entire space while periodically communicating with each other [10–12]. One of the most recent forms of Wang and Landau sampling is the replica exhcange formalism, which is a powerful way to parrelize over walkers. In principle all of the different forms of the Wang and Landau sampling currently used are based around the concept of sampling until a flat historgram is reached followed by a reduction in the modification factor of the density of states. The novel aspect of the B_LENDER algrotihm is that is formulated in a self consistent fashion and is believed to naturally evolve to a flat histogram of the visted energy levels. The B_LENDER algorithm is also natural to parallelize as it is based on a set of random walkers that each can explore the entire energy range.

In practice the calculation of the energy density of states $G(E_j)$ may is possible through uniformly randomly sampling the configuration space $\{\Sigma_i, e_i\}_\Omega$ of the Ω. The problem with this method is that if Ω is large, which it is for many problems, then the computationally effort to achieve convergence is not feasible. While the methods described previously of multi canaonical sampling and the Wang and Landau algorithm tackles this issue this work aimed to produce a Wang and Landau algorithm algorithm that is highly parallel in terms of the calculation of the energies. The algorithm designed in this work is inheriently designed for the compuation of the configurational density of states of lattice models, regardless of the calculational method for the energies. In particular the algorithm produced in this work is designed for problems were the calculation of the energy is much longer than updates in the density of states.

ALGORITHM

The B$_L$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. $G_r(E_j)^I,\ \{\Sigma_s, e_s\}^I_{\mathcal{S}}$
2. $\{\Sigma_s, e_s\}^I_{\mathcal{S}} \rightarrow \{\Sigma^{'}_s, e^{'}_s\}^I_{\mathcal{S}}$
3. $\Sigma^{'I}_s, e^{'I}_s \rightarrow \Sigma^{I+1}_s, e^{I+1}_s\ P = min[1, H(e_s)^I/H(e^{'}_s)^I]$ $else\ \Sigma^I_s, e^I_s \rightarrow \Sigma^{I+1}_s, e^{I+1}_s$
4. $G_r(E_j)^{I+1} =$

$$G_r(E_j)^I + \frac{C_o\mathcal{H}(E_j, \{e_s\}^{I+1}_{\mathcal{S}})}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}} G_r(E_j)^I =$$
$$G_r(E_j)^I (1 + \frac{C_o\mathcal{H}(E_j, \{e_s\}^{I+1}_{\mathcal{S}})}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}}) \tag{2}$$

Where $G_r(E_j)^0 \equiv [1 + \frac{C_o}{S}\mathcal{H}(E_j, \{e_s\}^0_{\mathcal{S}})]$ 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 $\{\Sigma_s, e_s\}^0_{\mathcal{S}}$ 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 $\{\Sigma_s, e_s\}^I_{\mathcal{S}}$ to produced a "perturbed" set $\{\Sigma^{'}_s, e^{'}_s\}^I_{\mathcal{S}}$, 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 $G_r(e_s)^I/G_r(e^{'}_s)^I$ then the perturbed configuration and energy $\Sigma^{'I}_s, e^{'I}_s$ goes to $\Sigma^{I+1}_s, e^{I+1}_s$, else the unperturbed configuration and energy Σ^I_s, e^I_s goes to $\Sigma^{I+1}_s, e^{I+1}_s$. This step (third) is dervied from the Wang and Landau method of sampling with probablity proportional to the inverse of the density of states. In the fourth step a histogram of the updated $\{e_s\}^{I+1}_{\mathcal{S}}$ energies is made and added (blended) in to the current density of states $G_r(E_j)^I$ by multiplying by a constant C_o(which affects the convergence properties) and the relative probability of each energy E_j in the intermediary density of states $G_r(E_j)^I$. 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 re-normalize the iterated relative density of states $G_r(E_j)^f$ at the final iteration $I = f$ as follows,

$$1.\ A = \sum_j G_r(E_j)^f$$
$$2.\ G(E_j) \approx G_r(E_j)^f \frac{\Omega}{A}\ , \tag{3}$$

to produce the properly normalized estimated value of $G(E_j)$.

An important discussion point of this algorithm (Eq 2) is the update of the relative density(step four) of states 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 the typical Wang and Landau sampling the update of the density of states is preformed by multiplication combined with a periodic reduction of the multiplication factor. 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 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 $G^M_r \equiv max[G_r(E_j)]$ the log of $\sum_j H(E_j)^I$ can be written as,

$$G^{LS}_r \equiv \log[\sum_j G_r(E_j)^I] = \log[G^M_r \frac{\sum_j G_r(E_j)^I}{G^M_r}] =$$
$$\log[G^M_r] + \log[\sum_j e^{\ln[G_r(E_j)] - \ln[G^M_r]}]\ . \tag{4}$$

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

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

In this form the alogrithm 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.

Rationale for Algorithm

The above algorithm can be rationalized by considering the adiabatic properties of a more general historgram $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$. Where $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$ is a historgram generated with $\mathcal{S}$ walkers simulated to $\mathcal{I}$ iterations using Wang and Landau importance sampling with respect to a fixed estimate of the density of states $G_r(E_j)$. Considering that the sampled energies are being generated with a probablility proportional to the exact density of states $G(E_j)$ and accepted with propablity inversely proportional to the relavtive(estimate) of the density of states $G_r(E_j)$, a historgram $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$ in equilibrium with $G(E_j)$ and $G_r(E_j)$ should follow the proportionality,

$$\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}}) \propto \frac{G(E_j)/\Omega}{G_r(E_j)/A} \; . \tag{7}$$

Where A is the sum over $G_r(E_j)$. Consider here that $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$ is generated with a fixed $G_r(E_j)$ for a certain number of iterations $\mathcal{I}$, then its sum is $\mathcal{I} \times S$. So in normalizing the proportionality in Eq 7 to $\mathcal{I} \times S$ we get,

$$\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}}) = \mathcal{S}\mathcal{I}\frac{G(E_j)}{G_r(E_j)}(\sum_j^{\Pi}\frac{G(E_j)}{G_r(E_j)})^{-1} \; . \tag{8}$$

Solving for the exact density of states gives,

$$G(E_j) = \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})G_r(E_j)\Phi \; , \tag{9}$$

where,

$$\Phi \equiv \sum_j^{\Pi}\frac{G(E_j)}{G_r(E_j)}\frac{1}{\mathcal{S}\mathcal{I}} \; . \tag{10}$$

It is important to note that Eq 9 assumes that $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$ is in equilibrium and simulated to an abritary level of precision with respect to $G(E_j)$ and $G_r(E_j)$. In general we may assume an approximately equals in Eq 9. From Eq 9 it is important to note that Φ is equivalent for all energy levels and does not affect relative probablities. In this sense we can rewrite Eq 9 as an equation for updating the relative density of states as follows,

$$G_r(E_j)^{I+1} = \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})G_r(E_j)^I\phi \; . \tag{11}$$

Where ϕ is now a function of our choice that is constant across energy levels at each iteration. Due to inaccuracies in $\mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})$ Eq 11 is not likely to be a stable way to converge $G_r(E_j)$ so a better form would be "blending" some of the old and new $G_r(E_j)$ together which gives a general form of the B$_L$ENDER algorithm,

$$G_r(E_j)^{I+1} = G_r(E_j)^I + \mathcal{H}(E_j, \{e_s\}_{\mathcal{S}\times\mathcal{I}})G_r(E_j)^I\phi \; . \tag{12}$$

It is now seen that the algorithm in Eq 2 is a form of this algorithm (Eq 12) where the histogram is generated with Wang and Landau importance sampling with $\mathcal{I} = 1$ and,

$$\phi \equiv \frac{C_o}{[\sum_j G_r(E_j)^I]^{\frac{1}{N}}} \; . \tag{13}$$

BENCH MARK WITH 2-D ISING MODEL

In this work the algorithm discussed is tested using the 2d square zero field Ising model with lattice dimension of even number[13–15]. The configurations Σ_i and energies e_i of the 2d Ising model are inherently defined by the lattice site spin variables and coupling constant J. The first test is of the effectiveness of the algorithm in calculating the density of states of the 2-d ising model. To test the accuracy of the simulations the results will be compared to the exact result solved by Beale [16]. The accuracy of the simulation will be determined by the average of errors given as,

$$\begin{aligned}\mathcal{E}(I,o) &= <|\epsilon(E_j,I,o)|>_j \\ &= \frac{1}{\Pi}\sum_{j=1}^{\Pi}\frac{|\ln(G_{ex}(E_j)) - \ln(G_r(E_j,I,o))|}{|\ln(G_{ex}(E_j))|} \; .\end{aligned} \tag{14}$$

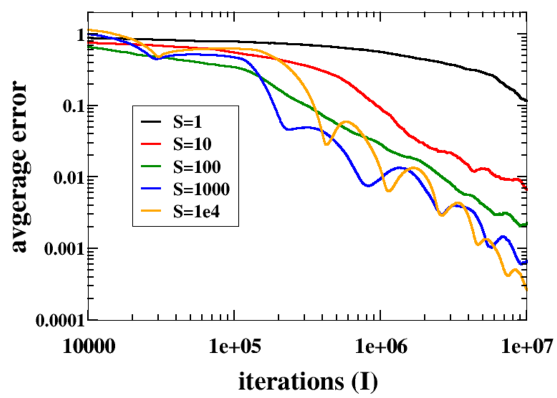
Where $G_{ex}(E_j)$ is the exact density of states, $G_r(E_j, I, o)$ is the calculated density of states at iteration number I from initial conditions and trajectory o, and $|\epsilon(E_j, I, o)|$ is the absolute value of the fractional error for a specific energy level. The primed configurations were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

This first test of the algorithm is with the 32X32 Ising model. While the ideal value of N is not known a-priori it was found in this work that a values of $N = 0.1$ was computationally efficient for the 32X32 Ising model. In Fig 1 the value of the average error calculated with Eq 14 is shown up to 1e7 iterations for $S = 1$, 10, 100, 1000, and $1e4$. The data in Fig 1 is averaged over 36 individual simulations for each value of S. The results show linear scaling from $S = 1$ to $S = 10$ and then another order of magnitude improvement from $S = 10$ to $S = 1000$, no improvement is discernable going to $S = 1e4$. The periodic flucuations in the avg error are also noted in going to larger S, it is hypothesized that these fluctions are related to the tunneling time of the walkers. The results at $S = 1000$ show that the average error is comparable to a linear speed up of the error reported for a single random walker in the orignial Wang and Landau algorithm. Defining a effective Monte Carlo step defined here as,

$$MC = \frac{S \times I}{\#E} \; , \tag{15}$$

where $\#E$ is the number of energies. For $S = 1000$, with the number of energies for the 2-d Ising model given by $n \times n$, at $I = 1e7$ gives $MC \approx 1e6$. With the value of the average error being < 0.001 for $S = 1000$ at $MC \approx 1e6$ the B$_L$ENDER algorithm is performing very well in terms of parallel speed up as compared to the reports for the orignial Wang and Landau algorithm for a single

(a)



(b)

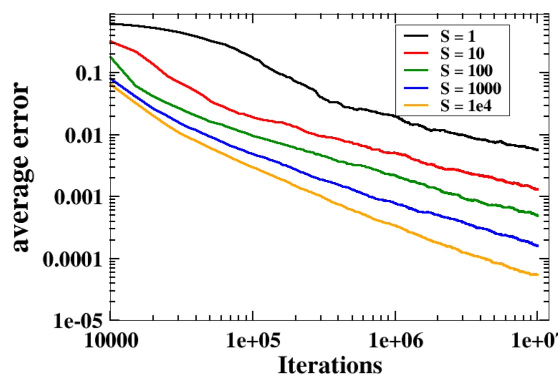


FIG. 1.

Walker. In comparison a single walker ($\mathcal{S} = 1$) simulated to $MC = 1e6$ using $N = 0.1$ has an average error of ≈ 0.002, where the result is averaged over 36 independant calculations.

Another aspect of the algorithm to consider is the dependence on the value of N. In Fig 2 the dependence on N is shown for the 32X32 Ising model and 10X10 ising model, simulated to $I = 1e6$ and $I = 1e7$ respectively, with $\mathcal{S} = 100$, and averaged over 36 independant calculations. The results show that for the larger 32X32 model the dependence on N is more pronounced and that the optimal value of N is lower than for the smaller 10X10 model.

APPLICATION TO LLTO

The purpose of developing the B$_L$ENDER algorithm was to develop and algorithm suitable for the needs of solid state density functional theory calculations of disordered lattice materials. 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

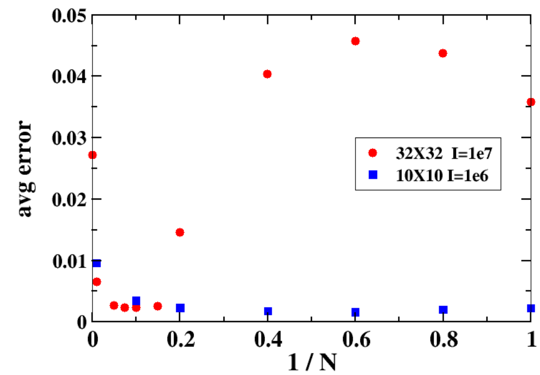


FIG. 2.

done as independant 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 is applied to the lithium and lanthanum sub lattice 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 B$_L$ENDER 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 the disordering of the lithium and lanthanum sub lattice and associated lattice distoritions.

Background on LLTO

LLTO is a complex material comprised of a variety of stoiciometries and phases but in this work the study is restricted to the tetragonal P4mmm phase of the stoiciometry $Li_{0.5}La_{0.5}TiO_3$. This phase has been characterized as being a perovsite structure.

Computational Details

The methods used in the calculation of the total energies of the lattice configurations of LLTO in this work were density functional theory using the VASP code. The PBE variant of generalized gradient approximation was used. Wave functions were represented by a planve wave basis with a 425eV cut off and self consistent cycles were converged with a energy difference of 1e-5 eV. The k-point mesh was 1X1X2 shifted by 0.5 0.5 0.5 for the 3X3X1 supercells of the LiTiO6 unit cell. The paramaters for the B$_L$ENDER algorithm were $\mathcal{S} = 10$ and $N = 1$. The value of omega was estimated as 10 times the combinatoric number of configurations of the Li La ordering into the A-site cages given as,

$$\Omega \approx 10\frac{18!}{9!9!} \, . \tag{16}$$

Results

While an exact value of Ω is not needed for the algorithm to converge experience suggests that being close as possible is computationally beneficial. Estimating that Ω is greater than the combinatoric calculation of the Li and La in the A-site cages comes from the possibility of multiple distinct lattice distoritions for each type of A-site cage configuration.

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] Z. D. Hood, H. Wang, A. S. Pandian, J. K. Keum, and C. Liang. Li_2OHCl crystalline electrolyte for stable metallic lithium anodes. *J. Am. Chem. Soc.*, 138(6):1768–1771, 2016.

[2] Y. Li, W. Zhou, S. Xin, S. Li, X. Zhu, J.and L, Z. Cui, Q. Jia, J. Zhou, Y. Zhao, and J. B. Goodenough. Fluorine-doped antiperovskite electrolyte for all-solid-state lithium-ion batteries. *Angewandte Chemie International Edition*, 55(34):9965–9968, 2016.

[3] G. Schwering, A. Hönnerscheid, L. v. Wüllen, and M. Jansen. High lithium ionic conductivity in the lithium halide hydrates $Li_{3-n}(OH_n)Cl$ $(0.83 \leq n \leq 2)$ and $Li_{3-n}(OH_n)Br$ $(1 \leq n \leq 2)$ at ambient temperatures. *ChemPhysChem*, 4(4):343–348, 2003.

[4] J. D. Howard, Z. D. Hood, and N. A. W. Holzwarth. Fundamental aspects of the structural and electrolyte properties of Li_2OHCl from simulations and experiment. *Phys. Rev. Materials*, 1:075406, Dec 2017.

[5] A. Y. Song, Y. Xiao, K. Turcheniuk, P. Upadhya, A. Ramanujapuram, J. Benson, A. Magasinski, M. Olguin, L. Meda, O. Borodin, and G. Yushin. Protons enhance conductivities in lithium halide hydroxide/lithium oxyhalide solid electrolytes by forming rotating hydroxy groups. *Adv. Energy Mater.*, page 1700971, 2018.

[6] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[7] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[8] D. P. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Physics*. Cambridge University Press, fourth edition edition, 2015.

[9] F. Wang and D. P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.*, 86:2050–2053, Mar 2001.

[10] J. Yin and D. P. Landau. Massively parallel wanglandau sampling on multiple gpus. *Comput. Phys. Commun.*, 183(8):1568 – 1573, 2012.

[11] L. Zhan. A parallel implementation of the wanglandau algorithm. *Comput. Phys. Commun.*, 179(5):339 – 344, 2008.

[12] T. Vogel, Y. W. Li, T. Wüst, and D. Landau. Generic, hierarchical framework for massively parallel wang-landau sampling. 110:210603, 05 2013.

[13] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Elsevier, 2016. ISBN 978-0-12-083180-7.

[14] L. Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65:117–149, 1944.

[15] B. M. McCoy and T. T. Wu. *The Two-Dimensional Ising Model*. Harvard University Press, 1973.

[16] P. D. Beale. Exact distribution of energies in the two-dimensional ising model. *Phys. Rev. Lett.*, 76:78–81, Jan 1996.