B_LENDER algorithm for naturally parallelizable calculation of energy density of states

Jason Howard[1]

[1] *Materials Science Division Argonne National Lab**

(Dated: November 27, 2018)

In this work an algorithm is proposed that calculates the energy density of states, this algorithm combines random sets with Wang and Landau importance sampling and is naturally parallelizable. The algorithm is referred to as B_LENDER, which is an acronym for $B_L$end Each New Density Each Round and 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 B_LENDER is tested with the 2d Ising model as an analogous system to the disorded cubic phase of the solid state lithium ion electrolyte Li_2OHCl. This was done in order to make a prediction of the wall time necessary for B_LENDER to evaluate the partition function of a 3×3×3 supercell of disordered cubic Li_2OHCl. This work also demonstrates that, during the simulation to calculate the energy density of states, linear averages of order parameters can be made for each energy level. With these averages and the energy density of states the ensemble average of the order parameter can be determined at any temperature.

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 the density of energy states $G(E_j)$. 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. One method to solve this problem could be temperature dependent simulations involving the Metropolis algorithm and histogram re-weighting techniques[6, 7]. Another algorithm called the Wang and Landau algorithm[8] 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 acronym which comes in part because how it blends the ideas of a random set and the Wang and Landau method, 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 [9–11]. B_LENDER is natural to parallelize as it is based on a set of random walkers that each can explore the entire energy range. It is hoped that due to B_LENDER being presented in the form of sets that the mathematics of set theory and real analysis can be used to analyze the algorithm and possibly further extended its capabilities from those presented in this work.

In a previous paper the notion of the master set $\{\Sigma_i, e_i\}_\Omega$ of the Ω total configurations and energies was put forth. The master set refers to the actual configurations of the defined system. In practice the calculation of the energy density of states $G(E_j)$ may only be possible through randomly sampling the configuration space. In a previous work it was shown that a properly scaled histogram of a uniformly sampled random set $\{\Sigma_s, e_s\}_\mathcal{S}$ converges to the exact density of states $G(E_j)$ as the the number of samples $\mathcal{S}$ goes to infinity. 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. This work tackles this issue and produces an algorithm that is highly parallel in terms of the calculations of the energies but also incorporates importance sampling such as in the Wang and Landau method.

The "B_LENDER" 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 $H(E_j)$.

$$
\begin{aligned}
&1.\ H(E_j)^i,\ \{\Sigma_s, e_s\}^i_{\mathcal{S}} \\
&2.\ \{\Sigma_s, e_s\}^i_{\mathcal{S}} \rightarrow \{\Sigma^{'}_s, e^{'}_s\}^i_{\mathcal{S}} \\
&3.\ H(E_j)^{Ii} = H(E_j)^i + \mathcal{H}(E_j, \{e^{'}_s\}^i_{\mathcal{S}}) \\
&4.\ \Sigma^{i'}_s, e^{i'}_s \rightarrow \Sigma^{i+1}_s, e^{i+1}_s \ P = \ min(1, H(e_s)^{Ii}/H(e^{'}_s)^{Ii}) \\
&\ else\ \Sigma^i_s, e^i_s \rightarrow \Sigma^{i+1}_s, e^{i+1}_s \\
&5.\ H(E_j)^{i+1} = H(E_j)^i + C_o \mathcal{H}(E_j, \{e_s\}^{i+1}_{\mathcal{S}}) \frac{H(E_j)^{Ii}}{\sum_j H(E_j)^{Ii}} \\
&6.\ N = \sum_j H(E_j)^{i+1} \\
&if\ H(E_j)^{i+1} \frac{\Omega}{N} < 1,\ H(E_j)^{i+1} = \frac{N}{\Omega}
\end{aligned}
\tag{2}
$$

Where $H(E_j)^0 \equiv \mathcal{H}(E_j, \{e_s\}^0_{\mathcal{S}})$ 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 histogram of the "perturbed" set is added to the current estimate of the density of states $H(E_j)^i$ to produce an intermediary density of states $H(E_j)^{Ii}$. In the fourth step a random number is drawn between zero and one for every sampled configuration, if this number is less then the ratio of the intermediary density of states $H(e_s)^{Ii}/H(e^{'}_s)^{Ii}$ 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 $\Sigma^{i+1}_s, e^{i+1}_s$. In the fifth 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 $H(E_j)^i$ 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 $H(E_j)^{Ii}$. In this work it was found a $C_o = \Omega$ was computationally efficient. The sixth step corrects for a systematic error of a energy having a density of state less than one, in principle for a discrete system a given energy should have a density of state of a least one. After the algorithm is deemed to be complete it is necessary to renormalize the iterated relative density of states $H(E_j)^f$ at the final iteration f as follows,

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

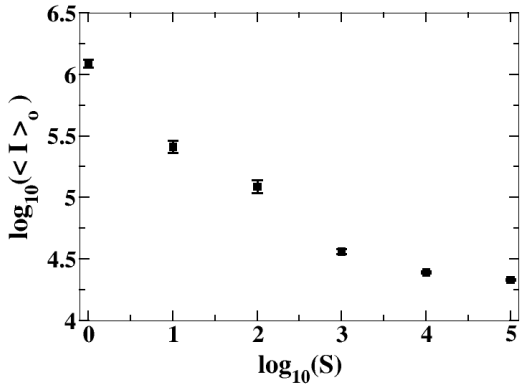
to produce the properly normalized estimated value of $G(E_j)$.

In this work the algorithm discussed is tested using the 2-d square zero field Ising model with lattice dimension of even number[12–14]. The configurations Σ_i and energies e_i of the 2-d Ising model are inherently defined by the lattice site spin variables and coupling constant J. In a previous study uniform sampling was used with first principles simulations to approximate the partition function for a $2 \times 2 \times 2$ supercell of Li_2OHCl. In going to just a $3 \times 3 \times 3$ the value of Ω would jump from $\sim 1e7$ to $\sim 1e26$. So computing the partition function for the $3 \times 3 \times 3$ system of Li_2OHCl is completely intractable from uniform sampling. In this work a $10X10$ 2d-ising model with $\Omega \approx 1.3e30$ is used as an analogous system to predict the computational effort needed for the B_LENDER algorithm to compute the partition function of a $3 \times 3 \times 3$ supercell of Li_2OHCl. The first test is a test to show the convergence of the algorithm in terms of the number of samples $\mathcal{S}$ and the number of iterations of the algorithm. To test the accuracy of the simulations the results will be compared to the exact result solved by Beale [15]. The accuracy of the simulation will be determined by the average of errors given as,

$$
\begin{aligned}
\mathcal{E}(I, o) \ &= < |\epsilon(E_j, I, o)| >_j \\
&= \frac{1}{\Pi} \sum_{j=1}^{\Pi} \frac{|\ln(G_{ex}(E_j)) - \ln(G_{bl}(E_j, I, o))|}{\ln(G_{ex}(E_j))}\ .
\end{aligned}
\tag{4}
$$

Where $G_{ex}(E_j)$ is the exact density of states, $G_{bl}(E_j, I, o)$ 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. To test the algorithm on the 10×10 Ising model, 144 individual calculations were performs at $\mathcal{S} = 1$, 10 ,100 , 1e3, 1e4, and 1e5. The number of iterations to find all the energies and the number of iterations to to reach $\mathcal{E} = 1\%$ were recorded and averaged for each value of S. These results are shown in Fig 1 (a) and (b). The plots show that the decrease in the number of iterations is not linear in including more samples $\mathcal{S}$ in the set, in the sense that there is an exponential like decay in the reduction in number of iterations for an increase in samples. Something of interest to understand is some of the basic convergence properties of the error of the simulation. To test the the convergence of the algorithm in terms of iterations 36 simulations where done for $\mathcal{S}$ =100 for 1e9 iterations. The average of the errors $< \mathcal{E}(I, o) >_o$ is shown in Fig 2 as a $\log_{10}$ $\log_{10}$ plot. The results show that the error appears to decrease approximately one order of magnitude for every 1.7 orders of magnitude increase in iterations. To further understand the errors the relative errors plots of $< |\epsilon(E_j, I, o)| >_o$ for the 144 simulations taken to $\mathcal{E} = 1\%$ and the 36 simulations taken to I =1e9 are shown in Fig 3 in black and red respectively. The results show that the relative error is larger at the less probable energy levels. The general shape of the errors seems to remain unchanged and the whole curve simply shifts when comparing the less converged results int black to the more converged results

(a) average iterations to reach $\mathcal{E} = 1\%$ vs $\mathcal{S}$



(b) average iterations to find all energies vs $\mathcal{S}$

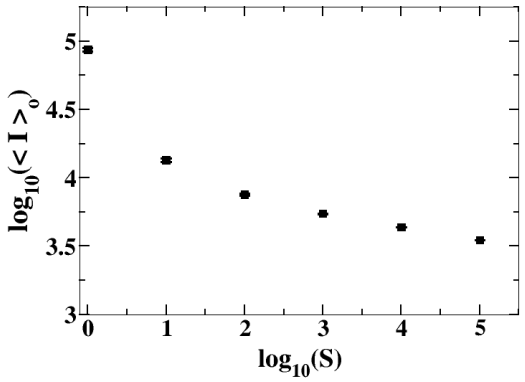


FIG. 1. A $\log_{10}$ $\log_{10}$ plot of the average iterations I to reach $\epsilon(I, o) = 1\%$ vs the set size $\mathcal{S}$ (a) and the average number of iterations I to find all of the energies vs the set size $\mathcal{S}$ (b). Error bars represent a linear propagation of the standard deviation of mean for $< I >_o$ in $\log_{10}(< I >_o)$.

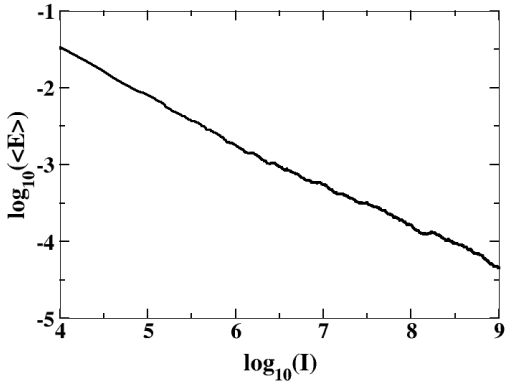


FIG. 2. $\log_{10}$ $\log_{10}$ plot of $< \mathcal{E}(I, o) >_o$ vs I. The iterations start at 1e4 because that is approximately the number of iterations when all the energies are found for a given simulation.

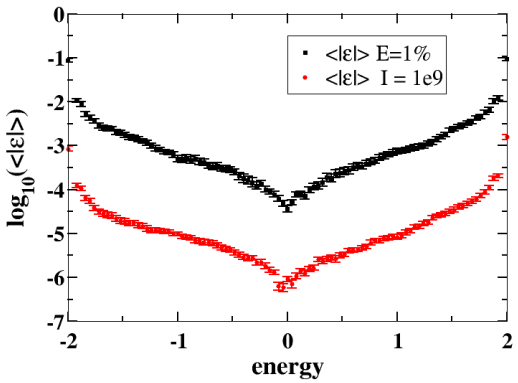


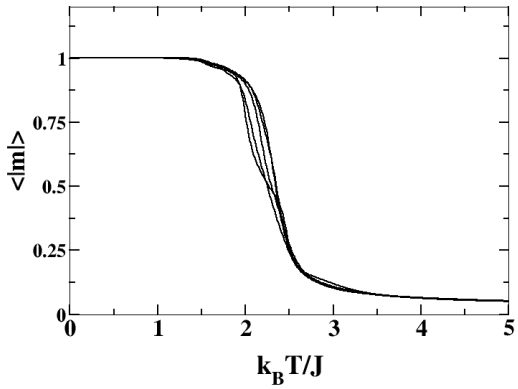
FIG. 3. plots of the $\log_{10}(< \mathcal{E}(I, o) >_o)$ vs energy per lattice site for 144 simulations taken to $\mathcal{E} = 1\%$ shown in black and 36 simulations taken to I =1e9 shown in red. Error bars are a linear propagation of the standard deviation of the mean for $< \mathcal{E}(I, o) >_o$ in the equation $\log_{10}(< \mathcal{E}(I, o) >_o)$.

red.

Now from Fig 1 (a) we could take $\mathcal{S} = 100$ as an approximate optimal value to attempt a calculation of the density of states of cubic Li_2OHCl. In willing to accept $\mathcal{E} = 10\%$ there is an approximate $10\times$ reduction in the number of iterations. So for an accuracy of $\mathcal{E} = 10\%$ that gives an approximate number of iterations $\sim 1e4$, which if the calculations of the energies took 1 minute would take 7 days. On the other hand if the level of accuracy needed was to just find all the energies which from the figure is $\sim$ 8000 iterations the calculation would take 5.5 days for 1 minute calculations of the energies. Another point to make is that for calculation of the free energy of a disordered phase, it is not completely necessary to find all energy levels, which was demonstrated in a previous work. At the current level of computing power the calculation of the energies is greater than 1 minute so it would still require too large of a computational effort to make utilizing this algorithm on the Li_2OHCl system feasible but if Moore's law continues within a decade it may be a viable option. It must be stated that the estimation on the iterations required for the Li_2OHCl system is only meant for a coarse estimate and that a more detailed analogous model would be required to improve the prediction. The benefit of the square 2-d Ising model in this work is the availability of exact solutions. Aside from the difficulties in large supercells first principles simulations the algorithm could be used with first principles on smaller systems, with model Hamiltonians, or systems defined by a model potentials.

Once the energy density of states is obtained many important thermodynamic properties can be determined but it may be considered a draw back that ensemble averages of a general order parameter $a_i \equiv a(\Sigma_i)$ are not accessible. It is shown in this work that this possible to

(a) $\mathcal{S}$ =10



(b) $\mathcal{S}$ =100

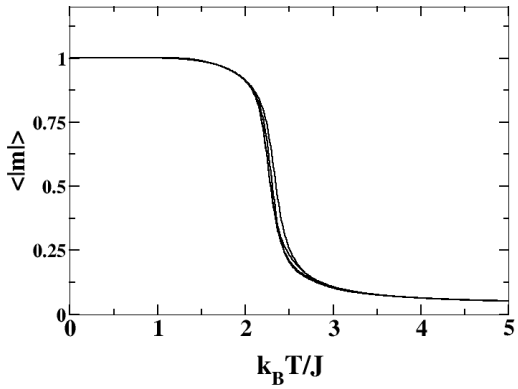


FIG. 4. The ensemble averaged absolute value of the magnetization per lattice site calculated for a 26×26 ferromagnetic Ising model using $\mathcal{S}$ =10 and $\mathcal{S}$ =100 until $\mathcal{E}(I,o) = 1\%$ shown in (a) and (b) respectively. Each plot displays five seperate calculations at each value of $\mathcal{S}$.

achieve by considering the following equation,

$$\langle a \rangle = \sum_{i=1}^{\Omega} a_i \frac{e^{-\frac{e_i}{K_b T}}}{Z} = \sum_{j=1}^{\Pi} < a_i >_j G(E_j) \frac{e^{-\frac{E_J}{K_b T}}}{Z} \ . \tag{5}$$

Where $< a_i >_j$ is the linear/simple average of a_i over all configurations with energy E_j. This equation (Eq 5) follows that given the ordered set of configurations $\{\Sigma_i\}_\Omega$ with energies $e_1 \leq e_i \leq e_\Omega$ and correspnding order paramaters $a_1 ... a_i ... a_\Omega$, the order paramaters can be paritioned into groups based on like energy. With this notion and defining $K(E_j) \equiv \sum_{l=0}^{j-1} G(E_l)$ (note that for convience: $G(E_0) \equiv 0$), Eq 5 can be written as,

$$\langle a \rangle = \sum_{j=1}^{\Pi} \sum_{k=1}^{G(E_j)} a_{K+k} \frac{e^{-\frac{E_J}{K_b T}}}{Z} = \sum_{j=1}^{\Pi} \frac{e^{-\frac{E_J}{K_b T}}}{Z} \sum_{k=1}^{G(E_j)} a_{K+k} \ . \tag{6}$$

From Eq 6 it is straight forward to define the sum over a_{K+k} in Eq 6 as,

$$\sum_{k=1}^{G(E_j)} a_{K+k} \equiv G(E_j) < a_i >_j \ , \tag{7}$$

that being the number of states with energy E_j times the average value of the order parameter a_i over those states. If an average of the values visited during the simulation is calculated then it may be possible to approximate ensemble averaged order parameters over the entire temperature range. To test this the B$_L$ENDER algorithm was applied to a 26×26 ferromagnetic Ising model with zero field with $\mathcal{S}$ = 10 and 100 to calculate the ensemble average of the absolute value of the magnetization per lattice site $\langle |m| \rangle$. In Fig 4 (a) and (b) the results of five seperate simulations are shown for $\mathcal{S}$ = 10 and 100. The results show that the known phase transition behavior of the ferromagnetic Ising model is well represented which indicates that this is a viable method for calculating ensemble averaged order parameters and that a larger value of $\mathcal{S}$ appears to reduce the variability between separate calculations. This calculation is also evidence of B$_L$ENDER being applicable to larger systems, although this is still small compared to the 256×256 model tested by Wang and Landau.

In summary the presented algorithm in this work has shown its effectiveness in accurately determining the density of states for the 2D square zero field Ising model. The simulations suggest that within a decade using B$_L$ENDER it will be feasible on a state of the art super computer to calculate the partition function of a 3×3×3 135 atom model of disordered cubic Li_2OHCl using density functional to a high level of accuracy. At the present time the algorithm should be suitable for first principles methods with smaller system sizes, with model Hamiltonians, or with models defined through model potentials. The algorithm shows promise for first principles methods in particular because the calculation of the energies can be done as individual job submissions to compute nodes managed by a script running on a head node. This work also demonstrated that if during the simulation to determine the energy density of states, linear/simple averages are made of order parameters for each energy level, then calculation of order parameters at any temperature is feasible.

* jdhoward@anl.gov

[1] Zachary D. Hood, Hui Wang, Amaresh Samuthira Pandian, Jong Kahk Keum, and Chengdu Liang. Li2ohcl crystalline electrolyte for stable metallic lithium anodes. *Journal of the American Chemical Society*, 138(6):1768–1771, 2016.

[2] Yutao Li, Weidong Zhou, Sen Xin, Shuai Li, Jinlong Zhu, Xujie L, Zhiming Cui, Quanxi Jia, Jianshi Zhou, Yusheng Zhao, and John B. Goodenough. Fluorine-doped antiperovskite electrolyte for all-solid-state lithium-ion batteries. *Angewandte Chemie International Edition*, 55(34):9965–9968, 2016.

[3] Georg Schwering, Andreas Hönnerscheid, Leo van Wüllen, and Martin Jansen. High lithium ionic conductivity in the lithium halide hydrates li3-n(ohn)cl (0.83n2) and li3-n(ohn)br (1n2) at ambient temperatures. *ChemPhysChem*, 4(4):343–348, 2003.

[4] Jason Howard, Zachary 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] Ah-Young Song, Yiran Xiao, Kostiantyn Turcheniuk, Punith Upadhya, Anirudh Ramanujapuram, Jim Benson, Alexandre Magasinski, Marco Olguin, Lamartine Meda, Oleg Borodin, and Gleb Yushin. Protons enhance conductivities in lithium halide hydroxide/lithium oxyhalide solid electrolytes by forming rotating hydroxy groups. *Advanced Energy Materials*, pages 1700971–n/a. 1700971.

[6] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[7] David P. Landau and Kurt Binder. *A Guide to Monte Carlo Simulations in Physics*. Cambridge University Press, fourth edition edition, 2015.

[8] Fugao 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.

[9] Junqi Yin and D.P. Landau. Massively parallel wang-landau sampling on multiple gpus. *Computer Physics Communications*, 183(8):1568 – 1573, 2012.

[10] Lixin Zhan. A parallel implementation of the wang-landau algorithm. *Computer Physics Communications*, 179(5):339 – 344, 2008.

[11] Thomas Vogel, Ying Wai Li, Thomas Wst, and D Landau. Generic, hierarchical framework for massively parallel wang-landau sampling. 110:210603, 05 2013.

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

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

[14] Barry M. McCoy and Tai Tsun Wu. *The Two-Dimensional Ising Model*. Harvard University Press, 1973.

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