Referee Comment on Manuscript mr-2023-8

Monte-Carlo Analysis of Asymmetry in Three-Site Relaxation Exchange: Probing Detailed Balance

The authors present Monte Carlo simulations of a lattice gas using a dynamic model that breaks detailed balance. They determine a quantity called asymmetry parameter, which measures the breaking of detailed balance, and show that it is nonzero. They also present results for an off-lattice gas model which seems to behave in a similar manner. These findings are related to recent NMR experiments.

I strongly disagree with the main statements in the paper. In my opinion the work suffers from serious conceptual deficiencies regarding both the design of the model and the interpretation of the results, which is why I would absolutely not recommend it for publication in a regular journal. However, I understand that in this journal, the referee reports will be published alongside with the paper. Therefore, a publication might be acceptable as long as some additional technical issues have been fixed.

Let me start with the technical issues first

- The central quantity, the asymmetry parameter, is never properly defined. The only definition is found in Equation (5) which refers to the special case of a three-site exchange. The authors must add an equation defining the quantity which is actually measured in the simulations and shown in Figures 4 and 6.
- Likewise, a so-called "active site" seems to be an important ingredient either of the dynamical model or in the analysis (this does not become clear), but it is never defined. It has "different relaxation properties" but relaxation properties have not been introduced in the definition of the model before. As an "explanation", the caption of Figure 3 offers the following cryptic sentence: "If a particle cell contacts two different relaxation sites, the higher number overrides the lower number when identifying its relaxation environment." What does this mean in practice? Does the presence of an active site change the dynamics or is it just important for the analysis? And how exactly is this implemented?
- Apart from the active site element, I think I roughly understand the dynamical model of the lattice simulations, but the off-lattice simulations (Section 2.2) are not well explained at all. Simulations of hard particle models would typically be done using event-driven algorithms, where the system is propagated from one elastic collision to the next. Apparently, this was not done here, instead fixed time steps were used, which reduces the accuracy of the simulations. How exactly were the collisions implemented? For example, did the authors accurately account for the

impact parameter of each collision when calculating the new momenta of the participating particles, or did they pick them at random? What was the length of the time step? How did they handle situations when three particles collide within one time step? Such information is crucial if you report on simulation results that supposedly break the second law of thermodynamics.

- Error bars are missing throughout. They must be added in the graphs, also the numbers in the text should be given with errors, especially those for (nonzero) asymmetry parameters.
- Given the complexity of the model, the code should not just be "available upon request", it should be published together with the manuscript. This holds especially for the off-lattice code.

These issues must be fixed before the paper can be published.

Next I will summarize the conceptual deficiencies in the presentation of the paper.

Monte Carlo model :

Description of the model :

Helmholtz free energy: On page 7, it is claimed that "the particle motion is governed by the Helmholtz free energy A". However, the Helmholtz free energy is a global thermodynamic quantity and does not govern local microscopic dynamics. Probably, the authors to refer to some kind of effective coarse-grained potential here.

Dynamics and Boltzmann distribution: Same page, the authors state "The probability of a particle moving from one cell to another is given by the Boltzmann distribution $p = \exp(-\Delta A/k_B T)$". This statement does not make sense, as already apparent from the fact that the "probability" p can be larger than one, $p > 1$ for $\Delta A < 1$. It is also not consistent with the subsequent description of the algorithm, where it becomes clear that the probability of moving to a certain site also depends on the number of equivalent accessible sites etc.

Design of the model :

Internal energy: The internal energy change after moving one particle is described as $\Delta U = \mathbf{F}\,\Delta\mathbf{R}$ (page 7), where $\mathbf{F}$ is the a force acting on a particle that is constructed from the occupancy of neighboring sites. First, there is an obvious sign error there, probably a typo, it should really read $\Delta U = -\mathbf{F}\Delta\mathbf{R}$: The energy decreases if the particle follows the force. For example, in a gravitational field, if you roll downhill, your potential energy decreases. Second, and more seriously, it is easy to see that this specific force field, as it is formulated on a lattice, is

not conservative. For example, consider a system where one particle is fixed at the origin, and a second particle undergoes a cyclic motion from $(1,0) \to (2,0) \to (2,1) \to (1,1) \to (1,0)$. Then the total internal energy change after the cycle is not zero, even though the final and initial configuration are exactly the same. Therefore, this lattice force field cannot be derived from a potential.

Entropy: The probability of moving to a neighbor lattice site is associated with an entropy change, which is estimated by the sum of step lengths to unoccupied neighbor cells. This specific form of entropy is entirely heuristic and again, it cannot be derived from an effective entropy potential. One should also note that it is not necessary to include translational entropy in a proper Monte Carlo algorithm: The algorithm will automatically account for it.

Jump probability (page 8): From the previous two points, it is already clear that the quantity ΔA in the expression for p cannot be associated with a well-defined effective potential A. However, even if such a potential existed, the choice of jump probabilities seems rather arbitrary. For example, page 8 says "If $0 < p < 1$, the destination cell is chosen at random from all those with the same largest jump probability $p < 1$". This is not well motivated. Why not choose from all cells with weighted probabilities according to their jump probability?

The algorithm described here is not motivated by any microscopic considerations. With the same right, assuming that ΔA could really be derived from a global effective potential function A, one could also use a standard Metropolis algorithm, which would satisfy detailed balance by construction.

Summary: The presented Monte Carlo algorithm does not satisfy detailed balance for two reasons: First, even though the notation suggests otherwise, the underlying quantities p are not associated with a well-defined effective energy function A. Second, the jump probabilities are chosen heuristically according to some random rules which are not well-motivated. It is not surprising that these rules do not satisfy detailed balance, because imposing detailed balance usually requires special efforts.

In fact, these rules would not even guarantee global balance if A were well-defined. On the other hand, they do define some kind of stochastic Markovian dynamics, and according to the central limit theorem of finite Markov systems, the probability distribution will converge against some stationary fixed point, which however differs from the Boltzmann distribution $\mathcal{N} \exp(-\beta A)$. Furthermore, this stationary state would include persistent currents by default, because, as explained above, special efforts must be taken to remove them in such a model.

Interpretation of the results :

Thermodynamic equilibrium: The term "thermodynamic equilibrium", by definition, refers to a stationary state *without currents*. One of the central postulates of thermodynamics is that every physical closed dynamical system reaches thermodynamic equilibrium at some point. This is a postulate and might be debated. However, a system with persistent currents as described in the manuscript would not be considered to be at thermodynamic equilibrium.

Detailed balance and nonequilibrium thermodynamics: As correctly stated in the manuscript, the lack of currents is associated with microscopic detailed balance – or, putting it the other way round, breaking detailed balance normally generates currents. However, this also implies that entropy is constantly being produced, and dissipated, see, e.g., References [1-3].

Dynamical systems with broken detailed balance have been discussed in nonequilibrium thermodynamics for many decades. Physically, they are used to describe open dissipative systems, for example, living systems or active systems [1,2], which are stabilized via a steady input of energy. It is easily possible to design stochastic dynamical systems that break detailed balance, as has been done, e.g., in the present manuscript or in Refs. [3,4]. In Monte Carlo simulations, implementing such dynamics can have the advantage that a desired probability distribution function can be sampled much more efficiently [4].

Detailed balance and Monte Carlo: The Monte Carlo method has been introduced by Metropolis et al as a method to efficiently sample a desired target probability distribution. The necessary ingredient for this is to impose global balance. Detailed balance is not strictly necessary. With the exception of kinetic Monte Carlo (which has not been used here), Monte Carlo dynamics is typically not realistic.

Nevertheless, Monte Carlo is also used to study dynamical systems in a coarse-grained sense. However, it is important to note that in this type of model, you get out what you put in. If you implement Monte Carlo moves that break detailed balance, then clearly, you will find that detailed balance is broken in your system. Therefore, Monte Carlo simulations designed to model dynamics at thermal equilibrium must be set up such that the Monte Carlo moves satisfy detailed balance.

Is detailed balance always fulfilled? As stated above, the claim that closed physical dynamical systems reach thermodynamic equilibrium is a postulate. It lies at the heart of the second law of thermodynamics, but being a postulate, it could be violated in certain cases. In fact, it is violated, e.g., for integrable systems such as linear harmonic

chains. It has not been proved rigorously except for a few special cases.

On the other hand, the opposite claim that detailed balance might be broken in realistic (closed) physical system fundamentally challenges the foundations of thermodynamics. Such a claim cannot be based on Monte Carlo simulations. This is because, as explained above, Monte Carlo dynamics are inherently artificial, and it is much easier to implement dynamical models that break detailed balance than to implement models that satisfy detailed balance. The claim would have to be based on experiments, or on molecular simulations of a truly microscopic model, e.g., classical Hamiltonian dynamics or Schrödinger dynamics. In fact, there have been several claims in the past, based on atomistic simulations, that the second law might be broken in nanoscale systems. For example, spontaneous unidirectional currents through pores or the like were observed in simulations. In all of these cases, it eventually turned out that the claimed effects could be attributed to numerical artefacts of the simulations.

The central question is whether a system can thermalize, which is a valid question especially for nanoscale systems and subject of active research. Specifically, the gas diffusion case discussed in the manuscript is related to the question whether a classical ideal gas can thermalize. This is one of the few cases which has been studied very intensely and for which rigorous results exist (the H-theorem, see [5]). Ideal gases do thermalize! In the manuscript, nonideal gases with excluded volume interactions are considered, which might change the situation, but I would be very surprised if it did. This is one of the reasons why it is so important that the authors describe their simulations for the gas diffusion simulations in more detail. If they maintain the claim that detailed balance is broken in these (off-lattice) systems, they should prove it much more carefully, e.g., by systematic variation of the time step, by studying the relaxation of several quantities as a funciton of simulation time, and by a solid assessment of error bars.

References:

1 C. W. Lynn et al, PNAS 2021, 118, e2108998118.

2 F. S. Gnesotto et al, Rep. Prog. Phys. 2018, 81, 066601.

3 L. Crochik et al, Phys. Rev. E 2005, 72, 057103.

4 M. Michel et al, J. Chem. Phys. 2014, 140, 054116.

5 G. Truesdell, R.G. Muncaster, *Fundamentals of Maxwell's kinetic theory of a simple monatomic gas*, Chapter XI, in *Pure and Applied Mathematics*, Volume 83, pp. 145-172 (1980).