Generalized ensemble method applied to study systems with strong first order transitions

E. Małołępsza\textsuperscript{1}, J. Kim\textsuperscript{2} and T. Keyes\textsuperscript{1}

\textsuperscript{1} Department of Chemistry, Boston University, Boston, MA 02215-2521
\textsuperscript{2} Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
E-mail: keyes@bu.edu

Abstract. At strong first-order phase transitions, the entropy versus energy or, at constant pressure, enthalpy, exhibits convex behavior, and the statistical temperature curve correspondingly exhibits an S-loop or back-bending. In the canonical and isothermal-isobaric ensembles, with temperature as the control variable, the probability density functions become bimodal with peaks localized outside of the S-loop region. Inside, states are unstable, and as a result simulation of equilibrium phase coexistence becomes impossible. To overcome this problem, a method was proposed by Kim, Keyes and Straub [1], where optimally designed generalized ensemble sampling was combined with replica exchange, and denoted generalized replica exchange method (gREM). This new technique uses parametrized effective sampling weights that lead to a unimodal energy distribution, transforming unstable states into stable ones. In the present study, the gREM, originally developed as a Monte Carlo algorithm, was implemented to work with molecular dynamics in an isobaric ensemble and coded into LAMMPS, a highly optimized open source molecular simulation package. The method is illustrated in a study of the very strong solid/liquid transition in water.

1. Introduction
First order phase transitions (solid↔liquid, liquid↔gas, and solid↔gas) are defined by a discontinuity in the free energy with respect to density. The phase transition starts in one phase and ends up in another, passing through coexisting region with both phases present. It involves heat absorption or release, the so called latent heat, where the temperature does not change with added heat until the phase transition is completed. Due to their ubiquitous presence, phase transitions have been extensively studied. However, computer simulation of phase transitions is a challenging task and here we focus on the source of the problem and applications of a proposed solution.

The problem in simulating first order phase transitions centers on the phase coexistence region and it is caused by unusual behaviour of entropy, $S$ [1, 2, 3, 4, 5, 6, 7]. Entropy is a function that monotonically increases with energy or enthalpy, but in the transition region its second derivative changes sign twice, and so $S$ changes curvature from concave to convex and back to concave, forming “convex intruder”. Consequently, the statistical temperature, $T_S$,

$$T_S(E) = \left[ \frac{dS}{dE} \right]^{-1}$$

also does not monotonically grow, but instead exhibits a back-bending, called an S-loop, as shown in black in Fig. 1.
Figure 1. Above: Cartoon of statistical temperature exhibiting an S-loop in first order phase transition region is shown in black. Color dashed lines represents three temperatures chosen for isothermal simulations, with blue one as transition temperature. Below: canonical probability density function (PDF) for equilibrium transition temperature indicated by blue line on the plot above, shows two peaks, each corresponding to a pure phase system, with a region in between that is not sampled.

Isothermal simulations (either as NVT or NPT) equilibrate readily in regions where only one phase is present, and the system is well sampled satisfying the ergodic hypothesis. Such ensembles correspond to red and green dots on Fig. 1, where red and green lines representing chosen temperatures intersect the statistical temperature. For both cases, below and above transition temperature, the constant temperature defining the ensemble intersects the statistical temperature only once. The intersection locates the most probable energy and minimum in free energy.

The situation changes when one would like to study the transition region and sample configurations with coexisting phases. A constant temperature in the vicinity of the transition temperature crosses the statistical temperature three times (see blue line on Fig. 1). Two filled circles correspond to minima in free energy where only one phase is present in the system, and maxima in the energy probability distribution function (PDF). The result is bimodal behavior of the PDF (see lower panel of Fig. 1). The empty circle in the middle indicates a barrier in free energy profile that might lead to broken ergodicity. Depending on the barrier height, sampling in the phase coexisting region is either inefficient or, in fact, impossible, with system in simulations stuck on either side of the barrier in minima indicated by filled blue circles.

2. The gREM

While constant-temperature ensembles are the most widely used, we would like to suggest a generalized ensemble that is characterized by a temperature function depending on energy, \( T_g(E) \) (or \( H \) in case of constant pressure simulations) \[1\]. The location of the maxima of the PDF is given by the generalized relation,

\[
T_g(E) = T_S(E).
\]  

(2)

The gREM utilizes generalized ensembles with a linear \( T_g(E) \),

\[
T_\alpha(E) = \lambda_\alpha + \eta(E - E_0),
\]  

(3)

with negative slope, \( \eta \), chosen to be large enough so that Eq. 2 has a single solution in the coexistence region, see orange lines on Fig. 2. \( E_0 \) is a reference energy. The result is an easily-simulated unimodal PDF in the transition region \[1\]. Replicas are selected with varying intercepts of the \( T_g(E) \) function, \( \lambda_\alpha \), such that the coexistence region is spanned. Replicas exchange configurations with an acceptance probability,
**Figure 2.** Cartoon of statistical temperature in black with orange lines as examples of generalized effective temperature $T_\alpha(E)$ defined in Eq. 3, each crossing the statistical temperature only once. Dashed blue line represents the tangent to the statistical temperature that defines marginal slope for effective temperature to ensure single solution of Eq. 2.

Acc, calculated to preserve detailed balance of the generalized ensemble,

$$\text{Acc} = \min[1, e^\Delta]$$  \hspace{1cm} (4)

$$\Delta = \omega_{\alpha+1}(E') - \omega_{\alpha+1}(E) + \omega_\alpha(E) - \omega_\alpha(E')$$ \hspace{1cm} (5)

where $\omega_\alpha$ is the gREM sampling weight of the $\alpha$-th replica and is calculated as follows:

$$\omega_\alpha = \int_E \left( \frac{1}{T_g(E')} \right) dE'.$$ \hspace{1cm} (6)

The original gREM method employed Monte Carlo, while here the approach was combined with the more efficient molecular dynamics method. The method was implemented by our group as a fix type module in LAMMPS [8, 9]. The required sampling is obtained from isothermal-isobaric MD by fixing the temperature at a constant “kinetic temperature”, $T_0$, and scaling the forces by the factor, $s(H) = T_0/T_g(H)$.

The new MD gREM algorithm was applied to pure water. A cubic box containing 3337 water molecules in configuration of hexagonal ice was initially equilibrated in the isobaric-isothermal ensemble at 200 K and 1 atm, and then used as the starting configuration in each generalized replica. Periodic boundary conditions were applied. The kinetic temperature was chosen as 330 K and pressure was kept as 1 atm using Nose-Hoover style non-Hamiltonian equations of motion [10, 11, 12, 13] as implemented in LAMMPS.

Water interactions were modeled by Stillinger-Weber potential [14] with parameters adjusted by Molinero [15] and denoted the mW model. In this approach each water molecule is represented as a single bead interacting via two- and three-body potential terms. Such coarse-graining reduces computational time while retaining a good reproduction of the properties of water [15].

Data gathered in all replicas were combined using the statistical temperature weighted histogram analysis method (ST-WHAM) [16] that is efficient and iteration-free method in contrast to the usual WHAM method [17].

### 3. Solid $\leftrightarrow$ liquid transition in water

Three of the most important properties of the gREM ensemble are that (1) it can successfully sample the phase coexisting region and overcome barriers, (2) though structures in this region are unstable in canonical simulations, the gREM ensemble stabilizes them and allows for equilibration [1] and (3) by using $\lambda$ to control the intersection of $T_g(H)$ with $T_S(H)$, enthalpy, not temperature, becomes the control variable for the simulation.

Characteristic structures observed around the solid $\leftrightarrow$ liquid transition in water are shown in Fig. 3. The lowest enthalpy structure corresponds to hexagonal ice occupying the entire
Figure 3. Snapshots of representative configurations observed during the solid↔liquid transition in water, shown from three points of view for easier identification of structural topology and arranged in columns. Blue dots correspond to water molecules, simulation box shown in gray. From left to right: (A) ice, (B) ice with liquid ball, (C) ice with liquid cylinder, (D) and (E) two structures with ice and liquid slabs differing in the ratio of the phases, (F) ice cylinder in liquid, (G) liquid.

Simulation box (structure denoted as A). Next, when ice starts melting, a small liquid ball (B) appears followed by a liquid cylinder (C), both surrounded by hexagonal ice. Once there is a sufficiently large amount of liquid water in the system, topology changes to liquid/ice slabs (D, E), where the interface between these two phases stays flat while the ratio of liquid to ice increases with increasing enthalpy. The next type of structures appears when there is not enough ice to form a slab and we observe instead an ice cylinder in equilibrium with bulk water (F). Finally, ice fully melts and the system consists of only bulk water (G). Topologies of the observed structures are in agreement with predictions from other studies [18].

Every change in the topology is connected to a feature in the calculated statistical temperature presented on Fig. 4. Therefore the expected S-loop in $T_S(H)$ for ice↔liquid water exhibits a complex structure in the phase coexistence region. Plateaus (or regions where $T_S(H)$ flattens) correspond to structures that belong to the same topology, while every change of topology type is in fact a transition and can be identified by an almost vertical change in $T_S(H)$.

There are two special points of interest on the $T_S(H)$ curve, namely spinodals, that define the limits of stability of both phases. The hexagonal ice in the mW model in our study can be heated up to 325.8 K before melting process starts. The experimental value of this limit equals 330±10 K [19]. In an earlier study of the mW model [20], a box of hexagonal ice was heated in canonical simulations for 100 ns until it decomposed to liquid water. The estimated limit of stability was found to be between 330 and 340 K. This value is higher compared to our result and is most likely caused the problem of overcoming the melting barrier when a liquid nucleus appears in an ice block.

The other limit of stability describes the situation where ice fully melts and changes into supercooled water at 223.4 K. The homogeneous nucleation temperature of mW water is equal to 202 K under 1 atm [15], below our estimated spinodal temperature, which is not surprising considering fact that we started the simulation from hexagonal ice, not liquid water, and as such we monitored ice melting instead of ice nucleation.
Figure 4. The statistical temperature for the hexagonal ice ↔ water transition, where each replica started from hexagonal ice, is shown as black line. The transition temperature from the Maxwell construction equals 269.1 K and is shown as a dashed blue line, while the limits of stability are shown in cyan and are equal to 223.4 K and 325.8 K. Letters A to G denote locations of structures shown on Fig. 3.

With the statistical temperature available one can apply the Maxwell construction to the inverse of $T_S(H)$ to obtain the transition temperature. In this procedure one is looking for a temperature that yields equal areas above and below it. The calculated equilibrium temperature of ice melting in mW model is equal to 269.1 K, compared to temperature of 274.6 K obtained by Molinero [15] using the phase coexistence method. Interestingly, the calculated melting temperature corresponds to a large plateau on the statistical temperature where structures consist of liquid and ice slabs with increasing amount of liquid towards higher enthalpies. This topology is identical to those used in the phase coexistence method where half of the simulation box is filled with one phase and the other half with the second phase. Location of this plateau can in fact be used to estimate the transition temperature.

In the present paper we showed preliminary results from the new MD version of an enhanced sampling algorithm, gREM, applied to the strong first order solid ↔ liquid transition in water modeled by the mW potential. We successfully calculated the statistical temperature for this transition as well as the transition temperature. We observed a series of structures with coexisting phases in several distinct topologies. Though usage of the gREM simulation method each of the observed structures was in equilibrium, in contrast to results from canonical simulations where the transition region is barely sampled.

Acknowledgments
The research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-SC0008810. We thank Profs. Bill Klein and Harvey Gould for discussions, Dr. Marek Orzechowski for help with graphics, and Dr. Christopher Knight for help with implementation.
[1] Kim J, Keyes T and Straub J E 2010 J. Chem. Phys. 132 224107
[2] Labastie P and Whetten R L 1990 Phys. Rev. Lett. 65 1567
[3] Wales D J and Berry R S 1994 Phys. Rev. Lett. 73 2875
[4] Bachmann M 2014 Thermodynamics and Statistical Mechanics of Macromolecular Systems (Cambridge University Press, Cambridge, UK)
[5] Gross D H E 2006 Physica A 365 138
[6] Junghans C, Bachmann M and Janke W 2008 J. Chem. Phys. 128 085103
[7] Behringer H and Pleimling M 2006 Phys. Rev. E 74 011108
[8] Plimpton S 1995 J. Comp. Phys. 117 1–19
[9] http://lammps.sandia.gov
[10] Martyna G L, Tobias D J and Klein M L 1994 J. Chem. Phys. 101 4177
[11] Parrinello M and Rahman A 1981 J. Appl. Phys. 52 7182
[12] Tuckerman M E, Alejandre J, Lopez-Rendon R, Jochim A L and Martyna G J 2006 J. Phys. A: Math. Gen. 39 5629
[13] Shinoda W, Shiga M and Mikami M 2004 Phys. Rev. B 69 134103
[14] Stillinger F H and Weber T A 1985 Phys. Rev. B 31 5262
[15] Molinero V and Moore E B 2009 J. Phys. Chem. B 113 4008
[16] Kim J, Keyes T and Straub J E 2011 J. Chem. Phys. 135 161103
[17] Ferrenberg A M and Swendsen R H 1988 Phys. Rev. Lett. 63 1195
[18] Schrader M, Virnau P and Binder K 2009 Phys. Rev. E 79 061104
[19] Schmeisser M, Iglev H and Laubereau A 2007 J. Phys. Chem. B 111 11271
[20] Jacobson L C, Hujo W and Molinero V 2009 J. Phys. Chem. B 113 10298