Generalized simulated tempering for exploring strong phase transitions

Jaegil Kim[a] and John E. Straub[b]
Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA

(Received 28 July 2010; accepted 28 September 2010; published online 15 October 2010; publisher error corrected 19 October 2010)

An extension of the simulation tempering algorithm is proposed. It is shown to be particularly suited to the exploration of first-order phase transition systems characterized by the backbending or *S*-loop in the statistical temperature or a microcanonical caloric curve. A guided Markov process in an auxiliary parameter space systematically combines a set of parametrized Tsallis-weight ensemble simulations, which are targeted to transform unstable or metastable energy states of canonical ensembles into stable ones and smoothly join ordered and disordered phases across phase transition regions via a succession of unimodal energy distributions. The inverse mapping between the sampling weight and the effective temperature enables an optimal selection of relevant Tsallis-weight parameters. A semianalytic expression for the biasing weight in parameter space is adaptively updated "on the fly" during the simulation to achieve rapid convergence. Accelerated tunneling transitions with a comprehensive sampling for phase-coexistent states are explicitly demonstrated in systems subject to strong hysteresis including Potts and Ising spin models and a 147 atom Lennard-Jones cluster. © *2010 American Institute of Physics*. [doi:10.1063/1.3503503]

I. INTRODUCTION

The molecular dynamics (MD) and Monte Carlo (MC) algorithms have become the primary tools for studying equilibrium properties of diverse physical, chemical, and biological systems.[1–3] However, simulating complex systems comprised of many particles with competing interactions often struggle to attain an adequate sampling of configurational space due to the trapping problem arising from rugged energy landscapes.[4,5] A great deal of efforts has been devoted to the development of enhanced sampling algorithms to accelerate the convergence of MC and MD simulations.[6–25]

One promising approach to achieving this goal is to combine multiple canonical ensemble simulations with an additional Markov process in temperature space. A well-known method exploiting this idea is the simulated tempering (ST)[8,9] or expanded ensemble method,[10] in which broken ergodicity at low temperatures is alleviated via a random walk in temperature space with a proper biasing weight facilitating temperature transitions.

While the ST approach has been widely applied to complex systems,[26–32] a number of challenges must be set to realize the full potential of the approach, the most demanding being the unknown dependence of the temperature weight. Since the exact weight corresponding to a relative free energy is not known *a priori*, preliminary simulations refining the weight are required. The refinement process is susceptible to local trapping in a complex energy landscape when the temperature weight for newly sampled regions is not well optimized. This limits the practical application of the ST in comparison to the temperature replica exchange method (*t*REM) or parallel tempering (PT)[6,14,15] that does not require an extra biasing weight.

Recently, several sophisticated methods have been proposed to accelerate the determination of the temperature weight.[30–34] One robust approach exploits the "average guide"[31] or "adaptive integration method,"[32] in which the temperature weight is adaptively updated by numerically integrating the derivative of the Helmholtz free energy. By determining the internal energy estimate from running averages of simulated trajectories, the weight determination is significantly accelerated. The usefulness of the average guide has been demonstrated in the combination of the ST and generalized ensemble sampling. In generalized simulated tempering (GST),[35] optimally parametrized non-Boltzmann weights combined with the weight-dependent average guide enable a systematic improvement over acceptance ratios for subensemble transitions, exploiting delocalized energy distributions.

In more recent years, considerable attention has been paid to the application of ST to biomolecular simulations[36–38] due to its better scaling property with system size and superior acceptance ratio for temperature transitions. Several comparison studies[39–41] revealed that ST with a properly chosen temperature weight exhibits faster diffusion in temperature space and a better sampling efficiency in comparison to the *t*REM. Combinations of the ST and other enhanced sampling methods have further explored this synergistic advantage.[42–45] However, the widespread use of ST to systems with complex energy landscapes and applications involving strong phase transitions, in which a large energy gap intervenes between two macroscopic phases, still remains challenging.

In this paper, the applicability of ST is further extended to diverse first-order phase transition systems by analyzing

[a]Electronic mail: jaegilkim89@gmail.com.
[b]Electronic mail: straub@bu.edu.

how conventional ST struggles to sample metastable states associated with the backbending (or *S*-loop) in the statistical temperature $T_S(E)$ (or microcanonical caloric curve)[46] characteristic to first-order phase transitions in finite size systems.[47–54] Our analysis reveals that an intrinsic instability of the canonical ensemble to a negative slope region in $T_S(E)$ is a primary cause for poor acceptance for temperature transitions in ST.

Exploiting this finding and the one-to-one correspondence between the sampling weight and the effective temperature,[55] an inverse mapping strategy is proposed to design optimally parametrized non-Boltzmann sampling weights. It is shown that a set of linear effective temperatures with a proper negative slope yields optimal Tsallis-weight ensembles,[56] avoiding an instability of the canonical ensemble for the backbending region in $T_S(E)$ and transforming unstable or metastable energy states of canonical ensembles into stables ones with successive unimodal energy distributions.

Another key ingredient of our method is an adaptive update scheme for the biasing weight in parameter space. A semianalytic expression for the biasing weight is derived by identifying the average guide[35] for a newly designed Tsallis weight. Faster determination of the biasing weight and an optimal selection of Tsallis parameters via the inverse mapping enables a comprehensive sampling for phase-coexistent states with a significant acceleration of tunneling transitions between two macroscopic phases as exemplified in Potts model.[57] Illustrative simulations for 256×256 Ising model demonstrate that our method is equally effective in exploring strong phase transition systems associated with nonbackbending behavior in $T_S(E)$.

The effectiveness of our method in treating continuum systems has been further explored in a 147 atom Lennard-Jones cluster (LJ_{147}) subject to strong hysteresis characterized by a severe backbending in $T_S(E)$.[47,48] It is shown that coupling the GST walker to independent Tsallis-weight runs via unconditional configurational swaps greatly shortens the weight determination process and facilitates tunneling transitions without disturbing detailed balance. The underlying mechanism of accelerated tunneling transitions in the GST coupled to Tsallis-weight runs is discussed in detail.

The paper is organized as follows. A general formulation of the ST and GST is outlined in Sec. II. The failure of the ST in sampling metastable, mixed states with the backbending region in $T_S(E)$ is analyzed in Sec. III. The inverse mapping strategy for designing optimal noncanonical ensemble weights is presented with a summary of simulation protocols for the GST in Sec. IV. The performance of our method is examined and quantitatively compared with Wang–Landau sampling[20] and the *t*REM for Potts and Ising models in Sec. V and the LJ_{147} cluster in Sec. VI. Our conclusions are presented in Sec. VII.

II. THEORETICAL BACKGROUND

A. Simulated tempering

The ST algorithm[8–10] combines several canonical ensemble runs with an additional Markov process on a set of reciprocal temperatures $\beta_j=[k_BT_j]^{-1}$ $(j=1,\ldots,M)$, k_B being the Boltzmann constant. Here j represents a subensemble index and M is the number of subensembles.

Two types of Markov moves are alternatively performed during the ST simulation: (1) a configurational move sampling the Gibbs–Boltzmann weight $W_{GB}=\exp\{-\beta E(\mathbf{x})\}$, $E(\mathbf{x})$ being the potential energy, which is accepted with the Metropolis criterion

$$A_{ST}^{intra}(\mathbf{x}\rightarrow\mathbf{x}';\beta)=\min[1,e^{-\beta(E'-E)}], \tag{1}$$

$E=E(\mathbf{x})$ and $E'=E(\mathbf{x}')$, and (2) an infrequent temperature transition subject to the acceptance probability

$$A_{ST}^{inter}(\beta\rightarrow\beta';\mathbf{x})=\min[1,G(\beta')/G(\beta)e^{-(\beta'-\beta)E}], \tag{2}$$

$G(\beta)$ being the unknown temperature weight biasing transitions in β space. The ultimate goal of ST is to perform a random walk in β space with a proper $G(\beta)$, providing an effective means for escaping from trapped states and overcoming undesirable broken ergodicity at low temperatures.

B. Generalized simulated tempering

The basic idea of ST is naturally extended to non-Boltzmann weights (GST),[35,61] in which a configurational space is sampled by an optimally designed generalized ensemble weight $W_{GST}=\exp\{-w(E;\lambda)\}$, $w(E;\lambda)$ being the effective potential. Here the parameter λ plays a similar role to β in ST. The dynamic range of λ should cover low and high energy regions and W_{GST} is usually selected to produce a delocalized energy distribution.

In GST, configurational moves are accepted with probability

$$A_{GST}^{intra}(\mathbf{x}\rightarrow\mathbf{x}';\lambda)=\min[1,e^{-w(E',\lambda)+w(E,\lambda)}], \tag{3}$$

and trial moves in λ space are accepted with probability

$$A_{GST}^{inter}(\lambda\rightarrow\lambda';\mathbf{x})=\min[1,G(\lambda')/G(\lambda)e^{w(E;\lambda)-w(E;\lambda')}], \tag{4}$$

$G(\lambda)=\exp\{\mathcal{G}(\lambda)\}$ being a biasing weight in λ space. The effective free energy $\mathcal{G}(\lambda)$ is a key quantity in facilitating transitions and realizing a random walk in λ space.

The probability density function (PDF) in λ space $P(\lambda)$ is proportional to $Z(\lambda)G(\lambda)/\mathcal{N}$ in GST, $Z(\lambda)=\int W_{GST}(\mathbf{x};\lambda)d\mathbf{x}$ being a partition function for the subensemble λ and $\mathcal{N}$ being a normalization constant. A random walk in λ space is accomplished by determining an ideal effective free energy,

$$\mathcal{G}_{id}(\lambda)=-\ln Z(\lambda), \tag{5}$$

which reduces to $\beta f(\beta)$ in ST, $f(\beta)$ being the Helmholtz free energy.

C. Average guide for $\partial\mathcal{G}_{id}/\partial\lambda$

The bottleneck in both ST and GST is the determination of the unknown effective free energy $\mathcal{G}(\beta)$ or $\mathcal{G}(\lambda)$. Recently, this determination has been significantly facilitated by identifying the average guide for the derivative of the effective

free energy.[31,35] The average guide for $\mu_{\rm id}(\lambda)=\partial\mathcal{G}_{\rm id}/\partial\lambda$ naturally emerges from differentiating both sides of Eq. (5) with respect to λ as

$$\mu_{\rm id}(\lambda)=\langle\mu(E;\lambda)\rangle_\lambda=\bar{\mu}(\lambda), \tag{6}$$

$\mu(E;\lambda)=\partial w(E;\lambda)/\partial\lambda$, and $\langle\mathcal{A}\rangle_\lambda$ indicates a subensemble average $\int dE P_\alpha(E)\mathcal{A}$, $P_\lambda(E)$ being the generalized probability density function (GPDF) at λ.

The most promising feature exploiting the average guide is that $\mathcal{G}(\lambda)$ is adaptively determined during the simulation by integrating $\bar{\mu}(\lambda)$, which is obtained as a subensemble average for $\mu(E;\lambda)$ over running trajectories. $\bar{\mu}(\beta)$ reduces to the internal energy $U(\beta)=\langle E\rangle_\beta$ in the conventional ST with $\mathcal{G}_{\rm id}(\beta)=\beta f(\beta)$. Once the analytic form of $w(E;\lambda)$ is known, $\bar{\mu}(\lambda)$ in GST is uniquely determined. In some cases, performing ST in temperature T rather than β is more convenient employing the average guide $\bar{\mu}(T)=-\langle E\rangle_T/T^2=-U(T)/T^2$.[31]

III. SIMULATED TEMPERING AROUND FIRST-ORDER PHASE TRANSITIONS

A. Backbending in $T_S(E)$

The weight determination is significantly accelerated with the average guide[31] or other clever methods.[30,32–34] However, ST still struggles to attain its maximum power for strong phase transitions. The situation becomes more acute in the vicinity of first-order phase transitions, in which a large energy gap involving the latent heat intervenes between $P_{\beta<\beta_c}(E)$ and $P_{\beta>\beta_c}(E)$, β_c being the transition temperature.

This difficulty is associated with an onset of a convex region ($\partial^2 S/\partial E^2>0$) in the microcanonical entropy $S(E)$, as sketched in Fig. 1(a).[47–49] The convex "intruder" in $S(E)$ gives rise to an intriguing negative slope region in the statistical temperature in Fig. 1(b),

$$T_S(E)=[\partial S/\partial E]^{-1}, \tag{7}$$

so called "backbending" or S-loop behavior.[47–54]

The backbending in $T_S(E)$ manifests an inherent instability of the canonical ensemble to phase-coexistent states, which invokes a bimodal structure in the canonical PDF $P_\beta(E)=e^{-\mathcal{F}_{\rm ST}(E;\beta)}/Z(\beta)$, $\mathcal{F}_{\rm ST}(E;\beta)=E-TS(E)$ being the Helmholtz free energy density. Noting that $\mathcal{F}_{\rm ST}(E;\beta_c)$ has three extrema at energies satisfying $T_S(E_i^*)=T_c$ $(E_1^*<E_2^*<E_3^*)$, $P_{\beta_c}(E)$ becomes double peaked at stable crossing points E_1^* and E_3^* with a local minimum at the unstable crossing point E_2^*.[54] Identifying $\beta\mathcal{F}''_{\rm ST}(E)=\beta^2 T'_S(E)$, the prime being the differentiation with respect to E, reveals that the intermediate energy states between E_u^1 and E_u^2 are intrinsically unstable for the canonical ensemble.

When the system size L is small, both free energy minima at E_1^* and E_3^* can be sampled across a free energy barrier at E_2^* [see Fig. 1(c)]. However, as L grows the backbending region becomes inaccessible due to a high free energy barrier. "Tunneling transitions" between the two macroscopic phases become rare, resulting in a localized $P_\beta(E)$ around E_1^* or E_3^*, depending on whether $\beta>\beta_c$ or not.

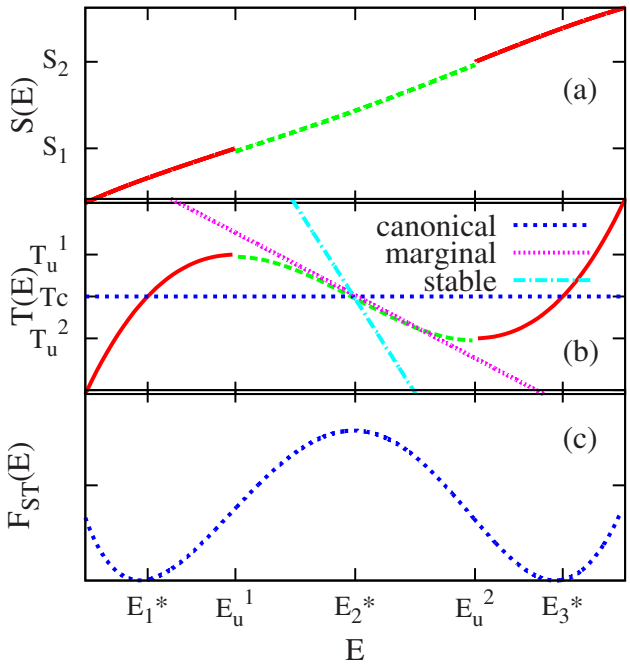

FIG. 1. (a) Convex regions (dashed line) in $S(E)$ between E_u^1 and E_u^2, (b) the backbending (S-loop) in $T_S(E)$ and effective temperatures $T(E;\lambda_\alpha)=T_c+\gamma_\alpha(E-E_2^*)$, $\gamma_\alpha=T'(E_2^*;\lambda_\alpha)$, and $T_S(E_i^*)=T_c$ $(E_1^*<E_2^*<E_3^*)$, and (c) $\mathcal{F}_{\rm ST}(E;\beta_c)=E-T_cS(E)$. In (b), unstable energy states around E_2^* in the canonical ensemble $\gamma_\alpha=0$ become stable in the generalized ensemble of Eq. (16) with $\gamma_\alpha<\gamma_S^{\rm min}=T'_S(E_2^*)$. Marginal corresponds to $\gamma_\alpha=\gamma_S^{\rm min}$.

B. Temperature transitions near β_c

To illustrate how ST fails in sampling across the backbending, consider a typical temperature transition between β_l $(>\beta_c)$ and β_h $(<\beta_c)$, close to β_c, at which $P_{\beta_l}(E)$ and $P_{\beta_h}(E)$ are sharply peaked at E_1^* and E_3^*, respectively. Assuming that $\mu_{\rm id}(\beta)=U(\beta)$ is known, the acceptance probability in Eq. (2) transforms into $\min[1,e^{h(\beta_l,\beta_h;E)}]$ where

$$h(\beta_l,\beta_h)=\int_{\beta_l}^{\beta_h}[\mu_{\rm id}(\beta)-E]d\beta=(U^*-E)(\beta_h-\beta_l), \tag{8}$$

and $U^*=U(\beta^*)$ for $\beta_h<\beta^*<\beta_l$. As $U(\beta)$ smoothly varies across the transition region, we apply the mean-value theorem for $\int_{\beta_l}^{\beta_h}\mu_{\rm id}(\beta)d\beta=\mu_{\rm id}(\beta^*)(\beta_h-\beta_l)$. By taking a linear approximation for $U(\beta)\approx E_1^*+(\Delta E^*/\Delta\beta)(\beta-\beta_l)$, $\Delta E^*=E_3^*-E_1^*$, and $\Delta\beta=\beta_h-\beta_l$, we identify $U^*\approx(E_1^*+E_3^*)/2$. For the sharply peaked $P_{\beta_l}(E)$ at E_1^*, $h(\beta_l,\beta_h;E)\approx\Delta E^*\Delta\beta/2<0$ since $\Delta\beta<0$ and $\Delta E^*>0$, implying that transitions from β_l to β_h are exponentially suppressed. The same situation occurs for the reverse transition from β_h to β_l with $h(\beta_h,\beta_l;E)\approx\Delta E^*\Delta\beta/2$ due to the sharply peaked $P_{\beta_h}(E)$ at E_3^*.

This analysis demonstrates that regardless of the current energy states, the acceptance for temperature transitions crossing β_c diminishes exponentially as

$$A_{\rm ST}^{\rm inter}(\beta\to\beta';E)\approx\min[1,e^{-|\Delta E^*(\beta'-\beta)|/2}]. \tag{9}$$

Note that the acceptance for temperature transitions in ST is always greater than the acceptance for replica exchanges near β_c in the tREM, which varies as $\min[1,e^{-|\Delta E^*(\beta'-\beta)|}]$,[58] as has been proven for the general case.[39]

IV. GENERALIZED SIMULATED TEMPERING FOR FIRST-ORDER PHASE TRANSITIONS

A. Inverse mapping

Our analysis reveals that the canonical ensemble is not an optimal distribution for ST when challenged to sample across a backbending region because of its intrinsic instability. Is it possible to construct a noncanonical ensemble, transforming unstable or metastable states of the canonical ensemble into stable ones, with a unimodal energy distribution? The Gaussian ensemble approach[59] and its parallel verstion[60] accomplish this goal by multiplying the Boltzmann factor by a Gaussian in energy.

The challenge in our study is to construct a set of generalized ensemble weights, $W_{\mathrm{GST}}(E;\lambda_\alpha)$ $(\alpha=1,\ldots,M)$, which, as the parameter λ_α varies, successively samples unstable or metastable energy states for the backbending region in $T_S(E)$. A key relationship in designing such an optimal weight is an inverse mapping between the sampling weight and its effective temperature,

$$T(E;\lambda_\alpha)=\left[\frac{\partial w(E;\lambda_\alpha)}{\partial E}\right]^{-1}. \tag{10}$$

A necessary and sufficient condition for $T(E;\lambda_\alpha)$, yielding a unimodal PDF for the negative slope region in $T_S(E)$, is derived by identifying a stability condition of an extremum E^*_α of the generalized PDF. If $P_\alpha(E)=e^{-\mathcal{F}_{\mathrm{GST}}(E;\lambda_\alpha)}/Z(\lambda_\alpha)$, $\mathcal{F}_{\mathrm{GST}}=w(E;\lambda_\alpha)-S(E)$ being a generalized free energy density, the extremum of $\mathcal{F}'_{\mathrm{GST}}(E^*_\alpha)=0$ is determined by

$$T(E^*_\alpha;\lambda_\alpha)=T_S(E^*_\alpha)=T^*_\alpha, \tag{11}$$

and its stability is identified by

$$\mathcal{F}''_{\mathrm{GST}}(E^*_\alpha)=(\gamma_S-\gamma_\alpha)/T^{*2}_\alpha, \tag{12}$$

$\gamma_S=T'_S(E^*_\alpha)$, and $\gamma_\alpha=T'(E^*_\alpha;\lambda_\alpha)$. Here the prime denotes a differentiation with respect to E.

The key finding in both Eqs. (11) and (12) is that a unimodal $P_\alpha(E)$ can arise from forming the unique crossing point E^*_α between $T_S(E)$ and $T(E;\lambda_\alpha)$ subject to $\gamma_\alpha<\gamma_S$. Indeed, a second order expansion of $P_\alpha(E)$ at E^*_α verifies

$$P_\alpha(E)\approx e^{-(E-E^*_\alpha)^2/2\sigma_\alpha}, \tag{13}$$

$\sigma_\alpha=T^{*2}_\alpha/(\gamma_S-\gamma_\alpha)$, illustrating that the "stable" in Fig. 1(b), with $\gamma_\alpha<\gamma_S$, generates a Gaussian PDF centered at E^*_α with the positive σ_α even for a negative γ_S. The GPDF approaches $\delta(E-E^*_\alpha)$, with $\gamma_\alpha\to-\infty$, corresponding to the microcanonical ensemble, and becomes locally flat around E^*_α, with $\gamma_\alpha=\gamma_S$ ["marginal" in Fig. 1(b)]. The crossing point E^*_α is unstable for $\gamma_\alpha>\gamma_S$ as in the canonical ensemble.

B. Linear effective temperature—Tsallis weight

The simplest choice for the effective temperature is a linear one

$$T(E;\lambda_\alpha)=\lambda_\alpha+\gamma_0(E-E_0), \tag{14}$$

the control parameter λ_α being the T-intercept at an arbitrary E_0 in two-dimensional (E,T) space and γ_0 being a linear slope. As illustrated in Fig. 2(a), a set of linear effective temperatures with a negative slope $\gamma_0<\gamma_S^{\min}$, $\gamma_S^{\min}$ being the minimum value of γ_S, forms successive stable crossing points with $T_S(E)$ across the backbending region, forming Gaussians centered at $\{E^*_\alpha\}$.

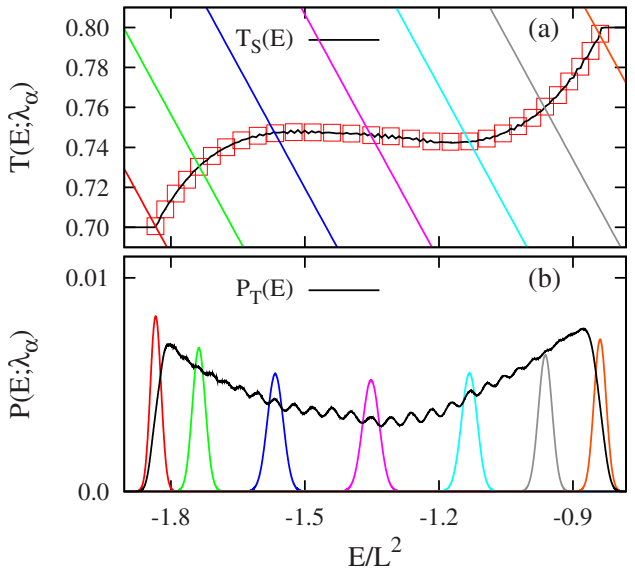

FIG. 2. (a) $T(E;\lambda_\alpha)$ (solid lines) and $[E^*_\alpha,T^*_\alpha]$ (squares) and (b) $P_\alpha(E)$ and $P_T(E)$ in GST_1 for a Potts model with $L=64$. In both (a) and (b), $\alpha=1$, 5, 10, 15, 20, 25, and 30 from left to right. $T_S(E)$ in (a) is determined by STMC and $P_T(E)$ in (b) is magnified 20 times for visualization.

Since $T_S(E)$ monotonically increases outside the transition region, the optimal γ_0 is easily guessed from the approximate $T_S(E)$ by connecting points $[\tilde{U}_\alpha=\tilde{U}(T_\alpha),T_\alpha]$, $\tilde{U}(T_\alpha)$ being an average energy derived from a short canonical run at T_α. For example, γ_0 can be safely chosen as $\gamma_L=(T_1-T_M)/(\tilde{U}_M-\tilde{U}_1)$. Once γ_0 is fixed the dynamic range of λ_α is determined between $\lambda_1=T_1$ and $\lambda_M=T_M-\gamma_0(\tilde{U}_M-\tilde{U}_1)$ with $E_0=\tilde{U}_1$. Notice that the first and Mth effective temperatures are constructed to cross $(\tilde{U}_1,T_1)$ and $(\tilde{U}_M,T_M)$, respectively. Intermediate λ_α $(\alpha=2,\ldots,M-1)$ are chosen by equally dividing λ space as

$$\lambda_\alpha=\lambda_1+(\alpha-1)\Delta\lambda, \tag{15}$$

$\Delta\lambda=(\lambda_M-\lambda_1))/(M-1)$.

The linear effective temperature of Eq. (14) produces a generic form of the Tsallis weight[56] proposed in nonextensive statistical mechanics as

$$W_{TS}\sim[\lambda_\alpha+\gamma_0(E-E_0)]^{-1/\gamma_0}=T(E;\lambda_\alpha)^{-1/\gamma_0}, \tag{16}$$

implying that the GST exploiting Eq. (14) is equivalent to the Tsallis-weight based ST.[35,61] However, note that the parametrized Tsallis weights with $\gamma_0\leq\gamma_S^{\min}$ in the GST are introduced to transform unstable energy states of the canonical ensemble into stable ones rather than simply producing more delocalized $P_\alpha(E)$ in other variants.[35,61–65]

C. Average guide $\mu(\lambda_\alpha)$

Denoting

$$\mu(E;\lambda_\alpha)=\frac{\partial w_{TS}}{\partial \lambda_\alpha}=\frac{1}{\gamma_0}\frac{1}{\lambda_\alpha+\gamma_0(E-E_0)}, \tag{17}$$

$w_{TS}=-\ln W_{TS}$, the average guide associated with Eq. (16) is obtained as

$$\bar{\mu}(\lambda_\alpha)=\frac{1}{\gamma_0}\left\langle\frac{1}{T(E;\lambda_\alpha)}\right\rangle_\alpha. \tag{18}$$

For a sufficiently small γ_0 ($\ll\gamma_S^{\min}$), yielding a sharply peaked $P_\alpha(E)$ at E_α^*, the subensemble average in Eq. (18) mostly arises from the neighborhood of E_α^*, resulting in

$$\bar{\mu}(\lambda_\alpha)\approx\bar{\mu}^{\max}(\lambda_\alpha)=\frac{1}{\gamma_0}\frac{1}{T_\alpha^*}, \tag{19}$$

$T_\alpha^*=T(E_\alpha^*;\lambda_\alpha)=\lambda_\alpha+\gamma_0(E_\alpha^*-E_0)$. Throughout this work we utilize Eq. (19) for the average guide by identifying the most probable energy set $\{E_\alpha^*\}$ in $P_\alpha(E)$.

D. Effective free energy $\mathcal{G}(\lambda_\alpha)$

Integrating the average guide $\bar{\mu}(\lambda)$ with respect to λ gives rise to the effective free energy,

$$\mathcal{G}(\lambda_\alpha)=\gamma_0^{-1}\int_{\lambda_1}^{\lambda_\alpha}1/T(E_\lambda^*;\lambda)d\lambda, \tag{20}$$

E_λ^* being the most probable energy in $P_\lambda(E)$. Interpolating $[\lambda_\alpha,T_\alpha^*]$ as $T(E_\lambda^*;\lambda)=T_\alpha^*+\eta_\alpha(\lambda-\lambda_\alpha)$ for $\lambda\in[\lambda_\alpha,\lambda_{\alpha+1}]$, $\eta_\alpha=(T_{\alpha+1}^*-T_\alpha^*)/(\lambda_{\alpha+1}-\lambda_\alpha)$, and integrating Eq. (20) yields a semianalytic effective free energy,

$$\mathcal{G}(\lambda_m)=\sum_{\alpha=1}^{m-1}\Delta\mathcal{G}_\alpha^+(\gamma_0)=\sum_{\alpha=1}^{m-1}\frac{1}{\gamma_0\eta_\alpha}\ln(T_{\alpha+1}^*/T_\alpha^*), \tag{21}$$

$\Delta\mathcal{G}_\alpha^+(\gamma_0)=\mathcal{G}(\lambda_{\alpha+1})-\mathcal{G}(\lambda_\alpha)$, which is completely determined by identifying $\{E_\alpha^*\}$.

Accordingly, subensemble transitions in GST proceed with

$$A_{\mathrm{GST}}^{\mathrm{inter}}(\alpha\to\alpha\pm1;E)=\min[1,e^{\Delta_\alpha^\pm(\gamma_0)}], \tag{22}$$

where $\Delta_\alpha^\pm(\gamma_0)=\Delta\mathcal{G}_\alpha^\pm(\gamma_0)-\Delta w_\alpha^\pm(\gamma_0)$, $\Delta\mathcal{G}_\alpha^-(\gamma_0)=-\Delta\mathcal{G}_{\alpha-1}^+(\gamma_0)$, and $\Delta w_\alpha^\pm(\gamma_0)=w_{TS}(E;\lambda_{\alpha\pm1})-w_{TS}(E;\lambda_\alpha)$. In the asymptotic limit of $\gamma_0\to0$, GST reduces to conventional ST associated with $W_{\mathrm{GB}}\sim e^{-(E-E_0)/\lambda_\alpha}$, λ_α being equally distributed temperatures between T_1 and T_M. As demonstrated in the Appendix, the analytic continuation of $\gamma_0\to0$ in Eq. (21) recovers the effective free energy $\mathcal{G}_{\mathrm{ST}}(T_\alpha)$ of Eq. (A1) in conventional ST.

E. Detailed simulation protocol

(1) Run short canonical runs at T_1 and T_M to determine approximate average energies $\tilde{U}_1$ and $\tilde{U}_M$, respectively. Select a proper γ_0 based on $\gamma_L=(T_1-T_M)/(\tilde{U}_M-\tilde{U}_1)$ and determine λ_α using Eq. (15) between $\lambda_1=T_1$ and $\lambda_M=T_M+\gamma_0(\tilde{U}_M-\tilde{U}_1)$ with $E_0=\tilde{U}_1$. The initial $\mathcal{G}_0(\lambda_\alpha)$ is guessed from $[E_\alpha^*,T_\alpha^*]$ obtained by short Tsallis-weight runs.

(2) Run GST with Eqs. (3) and (22). Update $\mathcal{G}_n(\lambda_\alpha)$ regularly by identifying $\{E_\alpha^*\}$ from the accumulated $P_\alpha(E)$. The subscript n in $\mathcal{G}_n(\lambda_\alpha)$ denotes the number of updates.

(3) Complete a production run with a frozen $\mathcal{G}_n(\lambda_\alpha)$ and compute the entropy estimate $\tilde{S}(E)$ by joining subensemble runs with the weighted histogram analysis method.[7] All canonical thermodynamic properties can be calculated at arbitrary temperature with $\tilde{S}(E)$.

V. APPLICATIONS TO DISCRETE SYSTEMS

A. Potts model

To illustrate how effectively GST performs around a typical first-order phase transition, we have examined eight state Potts model as a benchmark. The energy is determined as $E=-\Sigma_{\langle ij\rangle}\delta(S_i,S_j)$,[57] where the sum runs over the nearest-neighbor spins on the $L\times L$ square lattice and Kronecker δ takes the value 1 if $S_i=S_j$ and 0 otherwise. Here $S_i=1,2,\ldots,Q=8$ are spin variables.

To determine γ_0 and a dynamic range of λ_α, we first performed short canonical runs for 2×10^4 MC sweeps (MCSs) at $T_1=0.7$ and $T_M=0.8$, yielding $\tilde{U}_1\approx-7507.8$ and $\tilde{U}_M\approx-3438.7$. One MCS means L^2 MC trial moves for all spins. Based on $\gamma_L\approx-2.5\times10^{-5}$, several GST simulations were performed with varying γ_0 for $L=64$ as summarized in Table I. $\mathcal{G}_0(\lambda_\alpha)$ was determined from $[E_\alpha^*,T_\alpha^*]$ guessed by Tsallis-weight runs for each 2×10^4 MCS, and $\mathcal{G}_n(\lambda_\alpha)$ was updated every 10^5 MCS with subensemble transitions every MCS.

Since the effective temperatures $T(E;\lambda_\alpha)$ in Fig. 2(a) are designed to produce narrow $P_\alpha(E)$ with $\gamma_0\ll\gamma_L$, the most probable set $[E_\alpha^*,T_\alpha^*]$ quickly converges toward a locus of $T_S(E)$. Indeed, the profile of $[E_\alpha^*,T_\alpha^*]$ determined from the GST_1 in Table I for 2×10^8 MCS shows a perfect coincidence with $T_S(E)$ determined from the statistical temperature Monte Carlo (STMC) simulation, and $\{E_\alpha^*\}$ exactly correspond to the crossing points between $T_S(E)$ and $T(E;\lambda_\alpha)$ in Fig. 2(a). The backbending in $T_S(E)$ is apparent between $E_u^1\approx-1.6L^2$ and $E_u^2\approx-1.15L^2$, and unimodal energy distributions $P_\alpha(E)$ centered at $\{E_\alpha^*\}$ smoothly join ordered and disordered phases in Fig. 2(b). The superimposed energy distribution $P_T(E)=1/\Sigma N_\alpha\Sigma_\alpha N_\alpha P_\alpha(E)$, N_α being the number of sampled configurations at λ_α, is fairly uniform across the backbending region with a characteristic structure stemming from narrow peaks in $P_\alpha(E)$.

Due to the adaptive refinement of $\mathcal{G}_n(\lambda_\alpha)$, transitions in λ space do not satisfy a rigorous detailed balance at the begin-

TABLE I. Simulation details of GST and WL simulations for Potts model with $L=64$ and 128 and Ising model with $L=256$. τ_E and t_S denote mean tunneling times and total simulation times, respectively. The asterisk * in GST_n^* indicates that the GST simulation is performed with new effective temperatures, Eq. (23), exploiting $[E_\alpha^*, T_\alpha^*]$ determined from the GST_n simulation.

System	Method	L	M	γ_0	τ_E (MCS)	t_S (MCS)
Potts	GST_1	64	30	$-0.000\,1$	3.03×10^5	1.5×10^8
	GST_2	64	30	$-0.000\,05$	3.19×10^5	1.5×10^8
	GST_3	64	30	$-0.000\,01$	6.12×10^5	1.5×10^8
	GST_4	64	20	$-0.000\,05$	3.43×10^5	1.5×10^8
	GST_1^*	64	30	$-0.000\,01$	2.13×10^4	1.5×10^8
	WL	64			2.34×10^5	1.5×10^8
	GST_5	128	50	-1.87×10^{-4}	2.67×10^6	2.0×10^8
	GST_5^*	128	50	-1.87×10^{-5}	1.31×10^6	2.0×10^8
	WL	128			1.81×10^6	3.2×10^8
Ising	GST_6	256	60	$-0.000\,1$	5.32×10^5	2.5×10^7
	GST_6^*	256	60	4.5×10^{-6}	2.81×10^5	1.5×10^8
	GST_7^*	256	30	4.5×10^{-6}	2.42×10^5	1.0×10^8
	GST_8^*	256	90	4.5×10^{-6}	4.04×10^5	1.0×10^8
	WL	256			2.38×10^5	1.5×10^8

ning of the simulation. However, as the most probable energies $\{E_\alpha^*\}$ converge to crossing points between $T(E;\lambda_\alpha)$ and $T_S(E)$, $\mathcal{G}_n(\lambda_\alpha)$ becomes effectively frozen and detailed balance is recovered. Notice that energies $\{E_\alpha^*\}$ at 5×10^6 MCS already coincide with those converged at 2×10^8 MCS in Fig. 3(a). Also, $\bar{\mu}^{\max}(\lambda_\alpha)$ in Eq. (18) becomes indistinguishable from Eq. (19) after 5×10^7 MCS in Fig. 3(b). We found $|\mathcal{G}_{n>100}(\lambda_\alpha)-\mathcal{G}_\infty(\lambda_\alpha)|\le10^{-2}$ for all α, $\mathcal{G}_\infty$ being $\mathcal{G}_n$ at 2×10^8 MCS, implying that the biasing weight in parameter space is practically frozen after 10^7 MCS. A production run was performed with the frozen $\mathcal{G}_{500}(\lambda_\alpha)$.

Typical random walks in both E and λ space show numerous transitions in Figs. 4(a) and 4(b), respectively. Individual energy trajectories of each subensemble in Fig. 4(c) illustrate that localized energy distributions centered at $\{E_\alpha^*\}$ smoothly bridge between ordered and disordered phases across the backbending region. For comparison, we performed the *t*REM simulation with $M=30$, in which temperatures are evenly distributed between 0.7 and 0.8, and replica exchanges were attempted every MCS per replica. Two effectively disjointed sampling domains with no replica exchange are apparent in Fig. 4(d). We also performed the entropic version of Wang–Landau (WL) sampling for the same energy range. In WL sampling, the entropy estimate $\tilde{S}(E)$ is dynamically refined as $\tilde{S}(E)=\tilde{S}(E)+\ln f$ for a visit to E. The

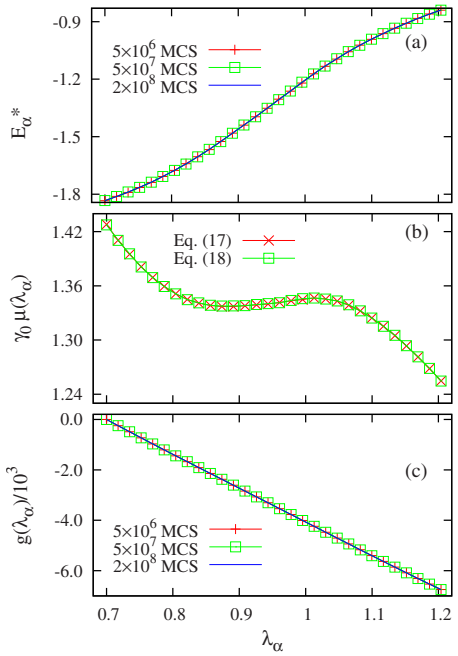

FIG. 3. (a) Profiles of $\{E_\alpha^*\}$, (b) $\bar{\mu}(\lambda_\alpha)$ and $\mu^{\max}(\lambda_\alpha)$ after 5×10^7 MCS, and (c) $\mathcal{G}_n(\lambda_\alpha)$ in the GST_1.

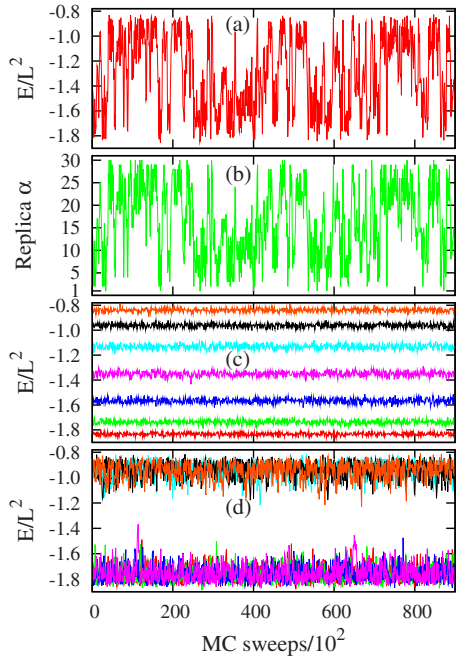

FIG. 4. Simulated trajectories in (a) energy and (b) replica space, (c) selected energy profiles for each subensemble α of the GST_1 simulation, and (d) energy trajectories of the *t*REM for a Potts model with $L=64$. $\alpha=1$, 5, 10, 15, 20, 25, and 30 from bottom to top in (c) and the same color scheme is applied to (c) and (d).

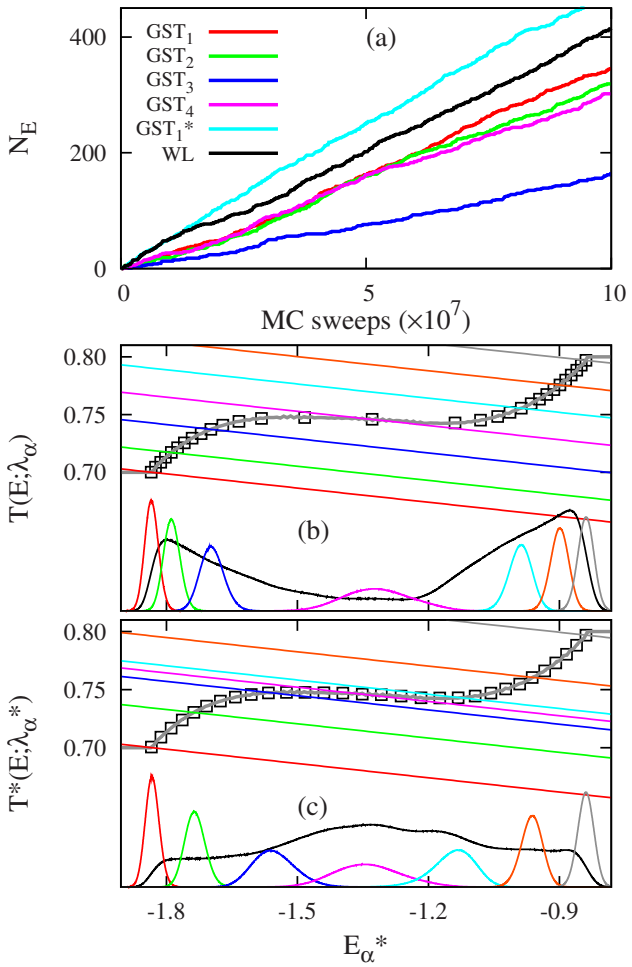

FIG. 5. (a) Accumulated tunneling transitions N_E in GST enhanced over WL in Table I, and $T(E;\lambda_\alpha)$, $[E^*_\alpha, T^*_\alpha]$, $P_\alpha(E)$, and $P_T(E)$ (b) for GST_3 and (c) GST^*_1. In (b) and (c), $\alpha=1$, 5, 10, 15, 20, 25, and 30 from left to right.

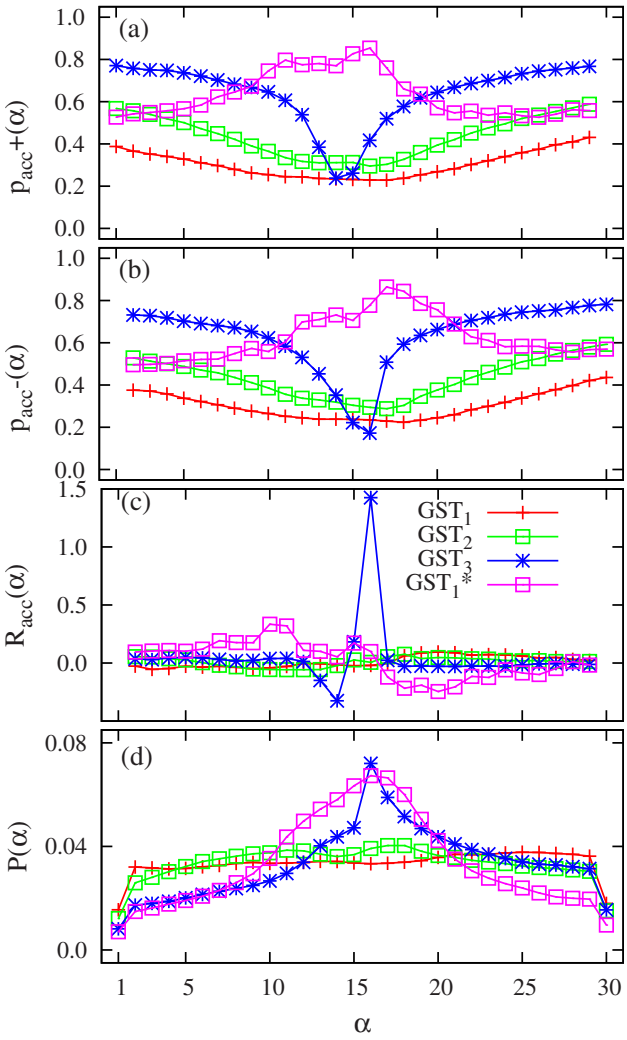

FIG. 6. Average acceptance probabilities (a) $p^+_{acc}(\alpha)$ and (b) $p^-_{acc}(\alpha)$, and (c) $R_{acc}(\alpha)=p^+_{acc}(\alpha)/p^-_{acc}(\alpha)-1$ and (d) $P(\lambda_\alpha)$ for GST simulations in Table I. The same color scheme is applied to (a), (b), and (d).

modification factor f reduces to $\sqrt{f}$ starting from 1.01 once $|H(E)-\bar{H}|/\bar{H}\leq 0.2$, $\bar{H}$ being the average histogram. $\tilde{S}(E)$ was refined down to $f_d=f-1=10^{-9}$ for 1.5×10^8 MCS.

To systematically quantify the global convergence of simulations, accumulated tunneling transitions N_E (Refs. 12, 67, and 68) were computed between $-1.80L^2$ and $-0.85L^2$, approximately corresponding to E^*_1 and E^*_M, respectively. The accumulated transitions N_E in Fig. 5(a) increase linearly with respect to MCS, implying that barrier crossing rates are almost constant throughout simulations. It is found that N_E in the GST_1 and GST_2 with $\gamma_0=-0.0001$ and $-0.000\,05$, respectively, are slightly less than N_E corresponding to the WL computed for additional 10^8 MCS. A drop of N_E in GST_3 with further increasing $\gamma_0=-0.000\,01$ is attributed to less efficient sampling for the transition region stemming from less formation of E^*_α in the backbending region with the biased $P_T(E)$ in Fig. 5(b). On the other hand, N_E remains nearly constant with much fewer subensembles of $M=20$ in GST_4 with a proper $\gamma_0=-0.000\,05$.

Tunneling is facilitated by unimodal PDFs located in the backbending region. Forming more crossing points in the transition region is crucial for an optimal performance of the GST. To investigate how a distribution of $\{E^*_\alpha\}$ affects N_E, the GST* simulations in Table I were performed with a set of new effective temperatures,

$$T^*(E;\lambda^*_\alpha)=\lambda^*_\alpha+\gamma_0(E-E_0), \tag{23}$$

$\lambda^*_\alpha=T^*_\alpha-\gamma_0(E^*_\alpha-E_0)$. Notice that $T^*(E;\lambda^*_\alpha)$ are selected to cross $[E^*_\alpha, T^*_\alpha]$ regardless of γ_0. We used the most probable data set $[E^*_\alpha, T^*_\alpha]$ of the GST_1. Due to a dense distribution of $\{E^*_\alpha\}$ in the transition region, the GST^*_1 gives rise to three times more N_E in Fig. 5(a), and a more delocalized $P_T(E)$ in Fig. 5(c), even for the same $\gamma_0=-0.000\,01$ in comparison to the GST_3.

Average acceptance probabilities $p^\pm_{acc}(\alpha)$ from α to $\alpha\pm1$ show a nonmonotonous γ_0 dependence in Figs. 6(a) and 6(b). Both $p^\pm_{acc}(\alpha)$ in GST_3 exhibit a minimum dip around $\alpha=15$. In forming a more dense distribution of $\{E^*_\alpha\}$ upon decreasing γ_0 to $-0.000\,05$, $p^\pm_{acc}(\alpha)$ increases gradually for the transition region, but declines in ordered and disordered phases due to a narrowing of $P_\alpha(E)$. As γ_0 is further decreased to -0.0001 in GST_1, narrowed $P_\alpha(E)$ cause an overall decrease in $p^\pm_{acc}(\alpha)$, but a compact distribution of $\{E^*_\alpha\}$ produces a more uniform $p^\pm_{acc}(\alpha)$. A dramatic enhancement of $p^\pm_{acc}(\alpha)$ in the transition region is observed for GST^*_1, giving rise to the highest N_E in Fig. 5(a) by maintaining a dense distribution of $\{E^*_\alpha\}$ with more delocalized $P_\alpha(E)$.

To mark a departure from a random walk in λ space we compared the ratio $R_{\rm acc}(\alpha)=p^{+}_{\rm acc}(\alpha)/p^{-}_{\rm acc}(\alpha)-1$ in Fig. 6(c). A perfect random walk corresponds to $R_{\rm acc}(\alpha)=0$. $R_{\rm acc}(\alpha)$ in GST_3 in Fig. 6(c) exhibits a large deviation from zero at $\alpha=14$ and 16, indicating a biased flux toward ordered and disordered phases in subensemble transitions. On the other hand, $R_{\rm acc}(\alpha)\approx 0$ throughout α in both GST_1 and GST_2 signifies a random walk in parameter space, yielding a nearly flat $P(\alpha)$ in Fig. 6(d). A drop of $P(\alpha)$ at $\alpha=1$ and 30 in Fig. 6(d) is due to restricted trial moves in λ_α. Of particular interest is that the profile of $R_{\rm acc}(\alpha)$ in GST_1^* is opposite to that of GST_3, implying that λ_α transitions are biased toward the phase transition region, and the peaked $P(\alpha)$ around $\alpha=15$ in GST_1^* stems from a centralized flux into the transition region, while the peak in GST_3 at $\alpha=15$ is due to trapping.

Our analysis reveals that an optimal way to maximize performance of the GST is to first determine a dense distribution of $\{E^*_\alpha\}$ for the transition region and then switch to a production run with Eq. (23) for a new $\gamma_0\approx\gamma_S^{\rm min}$. We applied this strategy to Potts model with $L=128$. A dense distribution of $\{E^*_\alpha\}$ is determined by the GST_5 simulation for 5×10^7 MCS. The parameter $\gamma_0=-1.87\times10^{-4}\ll\gamma_L=-6.2\times10^{-6}$ is set by short canonical runs at $T_1=0.7$ and $T_M=0.8$. Two production runs with different γ_0 (see Table I) were performed for each 2×10^8 MCS with the fixed $\mathcal{G}(\lambda_\alpha)$ in Eq. (21). We also performed WL sampling by refining $\tilde{S}(E)$ down to $f_d=10^{-9}$ for 3.2×10^8 MCS for the energy range between $-1.83L^2$ and $-0.85L^2$. The reduction of f is quick in the beginning, but slows down as f_d approaches zero.

Even with a much shorter time for the weight refinement (5×10^7 MCS for the GST_5^* and 3.2×10^8 MCS for the WL), N_E of the GST_5^* is higher than that of the WL simulation in Fig. 7(a). A faithful sampling for the transition region results in the uniform $P_T(E)$ via compact assignment of GPDFs in Fig. 7(b). Heat capacities C_v determined by the reweighting in Fig. 7(c) for $L=64$ collapse into a single line regardless of simulation conditions, while C_v of WL simulation exhibits a small deviation from those of the GST simulations due to a residual error in the refinement of $\tilde{S}(E)$. The corrected C_v computed through multicanonical sampling for an additional 10^8 MCS (Ref. 69), denoted as WL*, shows perfect agreement with that of the GST. On the other hand, heat capacities of both GST and WL for $L=128$ are in good agreement in Fig. 7(d).

B. Ising model

To examine how efficiently the GST algorithm explores a phase transition associated with a nonbackbending $T_S(E)$, the GST method is applied to Ising model with $L=256$. Based on $\tilde{U}_1\approx-1.75L^2$ and $\tilde{U}_{M=60}\approx-1.03L^2$ at $T_1=2.0$ and $T_M=2.6$, respectively, we first performed the GST_6 simulation for 2.5×10^7 MCS with $\gamma_0=-0.0001$ to determine a dense distribution of $\{E^*_\alpha\}$ in Fig. 8(a).

In contrast to Potts model, $T_S(E)$ traced by $[E^*_\alpha,T^*_\alpha]$ monotonically increases with a characteristic slope variation near $T_c\approx2.27$. Based on $[E^*_\alpha,T^*_\alpha]$ of GST_6, we performed the GST_6^* simulation with $\gamma_0=4.55\times10^{-6}\approx\gamma_S^{\rm min}$. In the case of a nonbackbending $T_S(E)$, γ_0 is not necessarily negative because stable crossing points can be formed as long as $\gamma_0\leqslant\gamma_S^{\rm min}$. Here we determined γ_0 as the minimum value among the linear slopes connecting $[E^*_\alpha,T^*_\alpha]$ to produce more delocalized $P_\alpha(E)$.

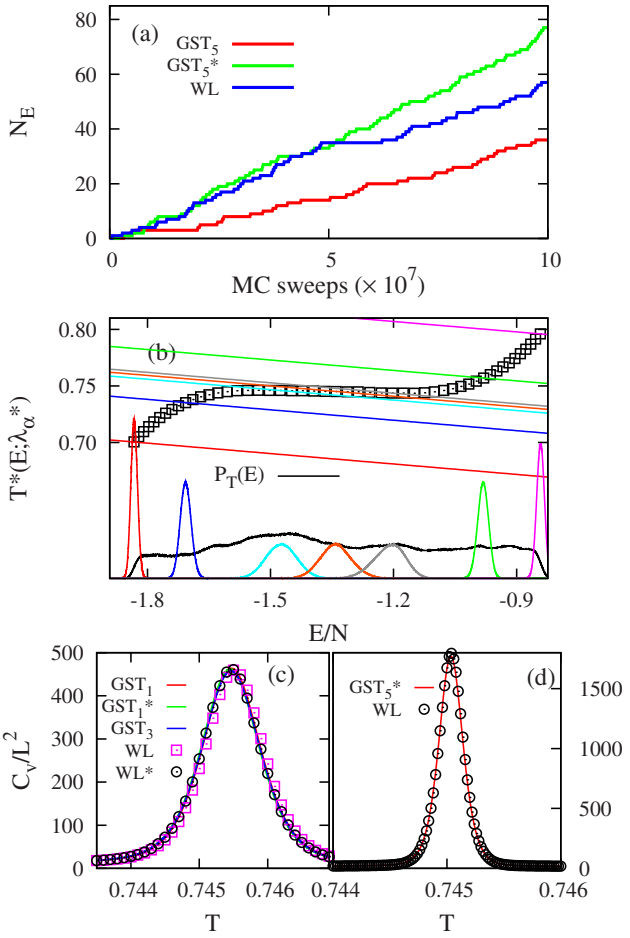

FIG. 7. (a) Enhanced tunneling transitions N_E in GST relative to WL, (b) $T^*(E;\lambda^*_\alpha)$, $[E^*_\alpha,T^*_\alpha]$, $P_\alpha(E)$, and $P_T(E)$ in GST_5^* for Potts spins with $L=128$, and C_v for Potts model with (c) $L=64$ and (d) $L=128$. $\alpha=1$, 10, 20, 25, 30, 35, 40, and 50 from left to right in (b).

As illustrated in Fig. 8(b), the GPDFs, $P_\alpha(E)$, in low and high energy regions are almost Gaussians centered at E^*_α, while $P_{30}(E)$ associated with $T^*(E;\lambda^*_{30})$, tangential to $T_S(E)$ around E^*_{30}, exhibits a much broader distribution in Fig. 8(b). A faithful sampling in the transition region with a compact distribution of $\{E^*_\alpha\}$ is marked by $P_T(E)$ peaked around $E\approx-1.42L^2$. In comparison to GST_6, N_E computed between $-1.68L^2$ and $-1.07L^2$ in Fig. 8(c) is doubled in GST_6^*, with τ_E comparable to that of the WL simulation in which $\tilde{S}(E)$ was refined for 1.5×10^8 MCS down to $f_d=1.8\times10^{-11}$.

To investigate M-dependence of N_E two additional simulations of GST_7^* and GST_8^* were performed for each 10^8 MCS with $M=30$ and 90, respectively. The compact data set $[E^*_\alpha,T^*_\alpha]$ was determined for 2.5×10^7 MCS with $\gamma_0=-0.0001$. In comparison to GST_6^*, N_E becomes more accelerated in GST_7^* with $M=30$, resulting in the essentially same τ_E as WL, while GST_9^* with $M=90$ shows a much longer τ_E for the same $\gamma_0=4.55\times10^{-6}$. This means that adding more subensembles does not always lead to an enhancement of tunneling transitions especially if sufficient energy overlaps be-

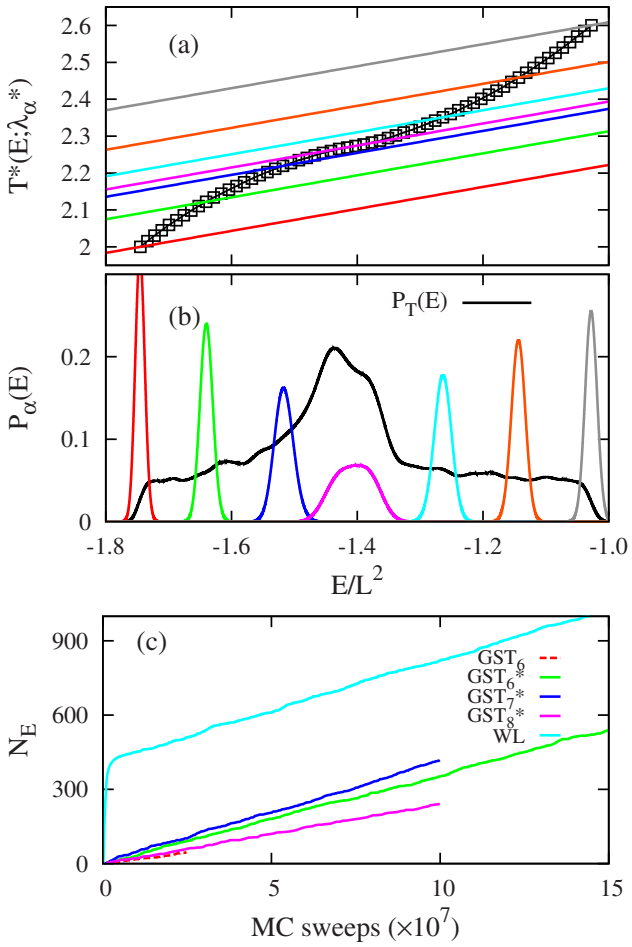

FIG. 8. (a) $T^*(E;\lambda_\alpha^*)$ and (b) $P_\alpha(E)$ in GST_6^*, and (c) N_E in GST and WL simulations for Ising model with $L=256$. $\alpha=1$, 10, 20, 30, 40, 50, and 60 from bottom to top in (a) and from left to right in (b). In (c), WL shows a very steep increase of N_E in the beginning due to a strong bias in $\tilde{S}(E)$, but starts to develop a linear increase of N_E as $f_d \to 0$. Notice that τ_E of WL was determined from a linear regime of N_E, shifted down by 500 for comparison.

tween neighboring subensembles already form in the transition region with broader GPDFs. Actually, increasing M requires more subensemble transitions for a tunneling, leading to an increase of τ_E as demonstrated in GST_8^*.

A comprehensive sampling for the phase transition region with a comparable τ_E to WL, while requiring a much smaller time for weight refinement, demonstrates that GST method with an optimal parameter set is more efficient than the WL algorithm for phase transitions with a nonbackbending $T_S(E)$.

VI. APPLICATIONS TO CONTINUUM SYSTEMS

The 147 atom Lennard-Jones cluster LJ_{147} is an ideal benchmark to test the effectiveness of our method in continuum systems. A solid-liquid transition is associated with a severe backbending in $T_S(E)$ (Refs. 47 and 48) and a strong hysteresis significantly hampers the weight refinement in various generalized ensemble methods.[68] The *t*REM or PT approach is known to be most effective as it employs a known sampling weight and a long relaxation process to the global minimum is greatly shortened via configurational swaps with low temperature replicas that typically explore the global minimum basin throughout the simulation.

Due to the adaptive refinement of $\mathcal{G}(\lambda_\alpha)$, GST experiences difficulty in effective sampling arising from strong hysteresis. To alleviate this problem, we coupled the GST walker to separate Tsallis-weight runs with the weight $W_{TS}(E;\lambda_\alpha)$, denoted by TW_α ($\alpha=1,\ldots,M$), via additional Markov moves swapping configurations between the GST walker and the TW_α. The configuration $\mathbf{x}$ at the current subensemble α in the GST walker is regularly swapped with the configuration $\mathbf{x}'$ in the TW_α. Since both configurations $\mathbf{x}$ and $\mathbf{x}'$ are sampled by the same $W_{TS}(E;\lambda_\alpha)$, these configurational swaps are unconditionally accepted without disturbing detailed balance.

The purpose of these configurational swaps is twofold. First, as in the *t*REM or PT, configurational swaps with low energy TW_α runs allow the GST walker to easily access the global minimum basin facilitating weight determination. Second, coupling the GST walker to intermediate TW_α runs around the phase transition region, constantly sampling phase-coexistent states, facilitates tunneling transitions by avoiding an intrinsic time scale for phase-coexistent states.

All GST simulations in Table II were performed with the same $\gamma_0=-0.01$ and $M=30$. τ_{ex} was varied to examine how configurational swaps between GST and the TW_α affect the weight determination and N_E. Here τ_{ex} represents a time interval for configurational swaps, which varies from one MCS to ∞ corresponding to no configurational swaps, and determines the extent of synchronization between sampling processes. The global minimum structure was taken to be the initial configuration with a spherical wall at the radius $R_c=4.0$. To minimize the extra computational burden, the total simulation time for TW_α runs was adjusted to be approximately one tenth of that in the GST by propagating TW_α runs for $(15\times M)$ MC moves during $(N\times M)$ MC moves in the GST walker. The initial guess for $\mathcal{G}_0(\lambda_\alpha)$ was determined by TW_α runs for 10^4 MCS.

The difficulty of the weight determination in the presence of strong hysteresis is clearly demonstrated in Figs. 9(a) and 9(b), in which E and $[E_\alpha^*, T_\alpha^*]$ in the GST_4 ($\tau_{ex}=\infty$) are plotted as a function of progress in the simulation. Even though the simulation was started from the global minimum,

TABLE II. Simulation details and mean tunneling times τ_E for GST and *t*REM simulations for LJ_{147}. τ_{ex} denotes a time interval for configurational swaps between the GST and TW_α runs.

Method	N	M	γ_0	τ_{ex} (MCS)	τ_E (MCS)	t_S (MCS)
GST_1	147	30	−0.01	1	4.16×10^4	2.5×10^8
GST_2	147	30	−0.01	30	4.51×10^4	2.5×10^8
GST_3	147	30	−0.01	300	6.83×10^4	2.5×10^8
GST_4	147	30	−0.01	∞	9.30×10^6	2.5×10^8
*t*REM	147	30			3.84×10^6	5.0×10^8

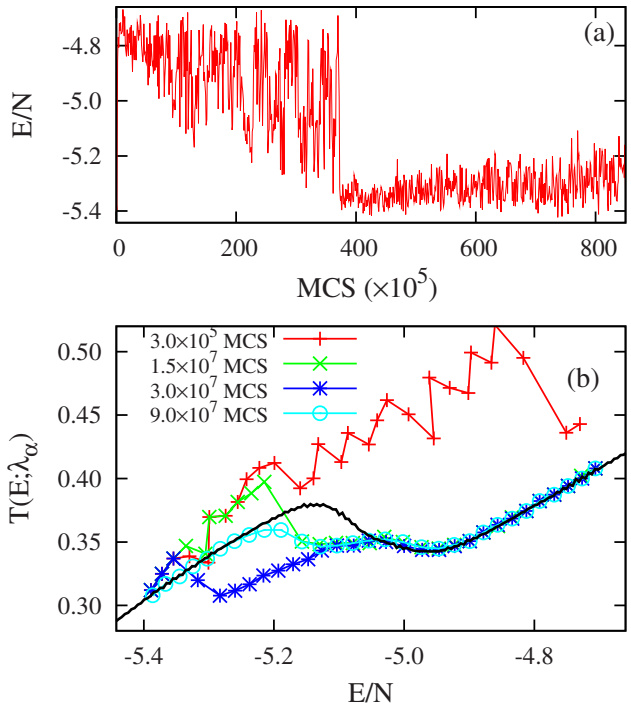

FIG. 9. (a) E and (b) $[E^*_\alpha, T^*_\alpha]$ in the GST_4 simulation for LJ_{147} in Table II. The solid line in (b) corresponds to $T_S(E)$ determined by RESTMC (Ref. 66).

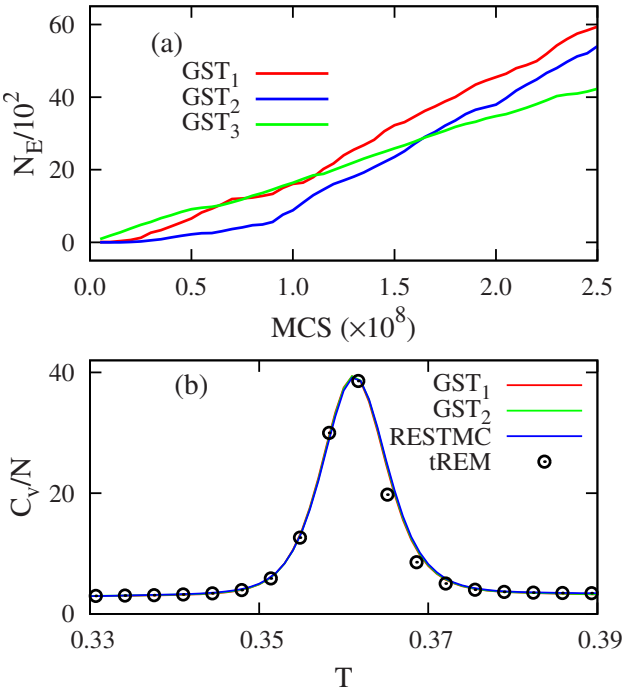

FIG. 11. (a) N_E in GST simulations with varying τ_{ex} in Table II and (b) reweighted heat capacities in GST, RESTMC, and tREM simulations for LJ_{147}.

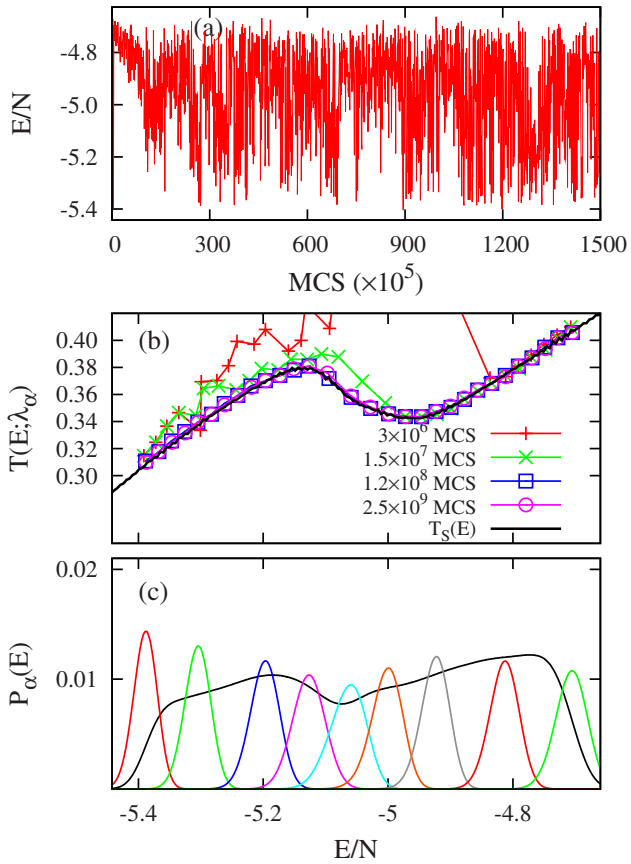

FIG. 10. (a) E, (b) $[E^*_\alpha, T^*_\alpha]$, and (c) $P_\alpha(E)$ and $P_T(E)$ in GST_1 simulation for LJ_{147} in Table II. α=1, 5, 10, 13, 15, 17, 20, 25, and 30 from left to right in (c).

the poorly estimated $\mathcal{G}_0(\lambda_\alpha)$ causes an immediate transition to liquid states in Fig. 9(a). Note that a locus of $[E^*_\alpha, T^*_\alpha]$ at 3×10^5 MCS significantly deviates from the converged $T_S(E)$ determined by replica exchange statistical temperature Monte Carlo (RESTMC) (Ref. 66) in Fig. 9(b). Once the system makes a transition to liquid states, sampling the global minimum basin is significantly hindered. Transient trapping in supercooled states is characterized by the stepwise variation of $[E^*_\alpha, T^*_\alpha]$ at 3×10^7 MCS in Fig. 9(b).

On the other hand, GST_1 with $\tau_{ex}=1$ effectively avoids trapping in metastable states and reaches the global minimum basin after 5×10^6 MCS via configurational swaps to low temperature TW_α runs, as seen in Fig. 10(a). The profiles of $[E^*_\alpha, T^*_\alpha]$ in the GST_1 are smoothly varying and converge to $T_S(E)$ at 1.2×10^8 MCS in Fig. 10(b) even for the backbending region. Frequent solid-liquid transitions are observed in Fig. 10(c) with frozen $\mathcal{G}_n(\lambda_\alpha)$ beyond 1.2×10^8 MCS across a unimodal $P_\alpha(E)$ peaked at stable crossing points $\{E^*_\alpha\}$, resulting in a uniform $P_T(E)$.

Once $\mathcal{G}_n(\lambda_\alpha)$ becomes effectively frozen in both GST_1 and GST_2 after 2.5×10^7 MCS and 10^8 MCS, respectively, N_E develops a linear relationship in Fig. 11(a). More frequent configurational swaps between the GST walker and the TW_α accelerate the weight determination with an early onset of a linear increase in N_E in the GST_1 ($\tau_{ex}=1$). To examine the effect of τ_{ex} on N_E, we also performed the GST_3 simulation ($\tau_{ex}=300$) with well-equilibrated initial configurations. A linear increase in N_E is observed in Fig. 11(a) with a slightly increased τ_E in Table II. The weak dependence of τ_E upon variation of τ_{ex} implies that the main factor accelerating N_E is coupling between the GST walker and TW_α runs. The speed-up of N_E measured by $R_\tau=\tau_{tREM}/\tau_{GST}$ is about 10^2-fold.

What is the underlying mechanism for additional accelerations of N_E in GST coupled to the TW_α? In a single replica simulation, forming mixed-phase configurations from ordered or disordered states is sequential and cumulative. τ_E is mainly restricted by an intrinsic time scale of a diffusive search for mixed-phase states (as in the 10^2 order longer τ_E in the GST_4 with $\tau_{ex}=\infty$). On the other hand, some of intermediate TW_α runs located in the phase transition regions retain phase-mixed configurations. As such, configurational swaps between the GST walker and these intermediate TW_α runs provide a nonphysical route for a tunneling.

Reweighted canonical heat capacities of various GST simulations collapse onto a single line demonstrating perfect agreement with those of RESTMC in Fig. 11(b), while the *t*REM shows a minor difference around $T_c \approx 0.36$ even for longer simulation time.

VII. CONCLUSIONS

An efficient GST method particularly suited to first-order phase transition systems has been proposed. Optimally parametrized Tsallis-weight ensemble simulations, in which unstable or metastable states of the canonical ensemble in the negative slope region of $T_S(E)$ transform into stable states forming successive unimodal energy distributions, are systematically combined via a guided Markov process in an auxiliary parameter space.

The inverse mapping between the effective temperature and the sampling weight enables a systematic selection of relevant Tsallis parameters. The resulting distributions smoothly bridge between ordered and disordered phases, allowing for a comprehensive sampling of phase-coexistent states. The determination of the weight in parameter space is greatly accelerated from the adaptive update scheme for the semianalytic effective free energy. The analytic continuation of the GST to the conventional ST in the asymptotic limit of $\gamma_0 \to 0$ is explicitly demonstrated.

The effectiveness of our method has been demonstrated in discrete spin systems with varying system size L and simulation conditions. The quantitative comparison between the GST and the WL algorithms in Potts spin systems with $L=64$ and 128 reveals that our method with optimal parameters exhibits more frequent tunneling transitions than in WL simulations. In comparison to the WL, requiring a long weight refinement with increasing L, our method enables the accurate calculation of thermodynamic averages with a shorter time for weight refinement. It is also shown that GST is equally effective in the simulation of strong phase transitions with nonbackbending $T_S(E)$, as exemplified in Ising spin system with $L=256$.

The performance of our method in continuum systems is examined in a 147 atom Lennard-Jones cluster associated with a severe backbending in $T_S(E)$. We have shown that strong hysteresis and trapping in metastable states in the weight refinement of the GST is significantly suppressed by coupling the GST walker to separate Tsallis-weight runs via unconditional configurational swaps, while maintaining a detailed balance. Faster weight determination and considerable acceleration of tunneling transitions were explicitly demonstrated through coupling of the GST walker and Tsallis-weight runs.

Finally, it should be noted that the performance of GST can be further improved with an optimal allocation scheme for λ_α. Here we used an equal λ_α allocation for simplicity. Various adaptive temperature allocation schemes are developed to maintain uniform acceptance ratios based on average energy differences of neighboring replicas in conventional *t*REM.[70–72] Since GST can locate E^*_α very quickly, the same strategy can be applied to produce a uniform acceptance by adaptively modifying a distribution of λ_α. Recently, an efficient feedback algorithm for the temperature allocation in *t*REM has been developed to maximize N_E for systems associated with a peaked C_v.[23] By analyzing a current of subensemble transitions on λ space in GST, this feedback algorithm can be implemented to optimize a distribution of λ_α more systematically.

ACKNOWLEDGMENTS

We are grateful to the National Science Foundation (Grant No. CHE-0750309) and the National Institutes of Health (Grant No. R01 GM076688) for the generous support of our research and the Center for Computational Science at Boston University for resources supporting this study.

APPENDIX: GST AS $\gamma_0 \to 0$

The effective free energy $\mathcal{G}_{ST}(T_\alpha)$ in conventional ST running at temperatures T_α is determined as

$$\mathcal{G}_{ST}(T_\alpha) = \int_{T_1}^{T_\alpha} \bar{\mu}(T) dT, \tag{A1}$$

$\bar{\mu}(T) = -U(T)/T^2$, $U(T)$ being the internal energy. Interpolating $[T_\alpha, U_\alpha]$ as $U(T) = U_\alpha + \xi_\alpha (T - T_\alpha)$ for $T \in [T_\alpha, T_{\alpha+1}]$, $\xi_\alpha = (U_{\alpha+1} - U_\alpha)/(T_{\alpha+1} - T_\alpha)$, and integrating Eq. (A1) yields

$$\mathcal{G}_{ST}(T_m) = \sum_{\alpha=1}^{m-1} \Delta \mathcal{G}^+_{\alpha,ST}, \tag{A2}$$

where

$$\Delta \mathcal{G}^+_{\alpha,ST} = -\frac{U_{\alpha+1} - U_\alpha}{T_{\alpha+1} - T_\alpha} \ln \frac{T_{\alpha+1}}{T_\alpha} + \frac{U_{\alpha+1}}{T_{\alpha+1}} - \frac{U_\alpha}{T_\alpha}. \tag{A3}$$

To investigate an asymptotic behavior of $\mathcal{G}(\lambda_\alpha)$ in GST as $\gamma_0 \to 0$, we decompose $\Delta \mathcal{G}^+_\alpha(\gamma_0)$ in Eq. (21) into

$$\Delta \mathcal{G}^0_\alpha(\gamma_0) + \Delta \mathcal{G}^d_\alpha(\gamma_0), \tag{A4}$$

where

$$\Delta \mathcal{G}^0_\alpha(\gamma_0) = -\frac{E^*_{\alpha+1} - E^*_\alpha}{T^*_{\alpha+1} - T^*_\alpha} \ln(T^*_{\alpha+1}/T^*_\alpha) \tag{A5}$$

and $\Delta \mathcal{G}^d_\alpha(\gamma_0) = \gamma_0^{-1} \ln(T^*_{\alpha+1}/T^*_\alpha)$. Here we used the identity $\eta_\alpha^{-1} = [1 - \gamma_0 (E^*_{\alpha+1} - E_\alpha)/(T^*_{\alpha+1} - T^*_\alpha)]$. Note that the divergence in $\Delta \mathcal{G}^d_\alpha(\gamma_0)$ prevents an analytic continuation of Eq. (A4) in the limit of $\gamma_0 \to 0$. However, this catastrophic situation is alleviated as $\Delta \mathcal{G}^d_\alpha(\gamma_0)$ cancels out by the same divergent term in $\Delta w^+_\alpha(\gamma_0)$ as

$$\Delta w_\alpha^+(\gamma_0) = \frac{1}{\gamma_0}\ln\frac{T(E;\lambda_{\alpha+1})}{T(E;\lambda_\alpha)} = \Delta w_\alpha^0(\gamma_0) + \Delta\mathcal{G}_\alpha^d(\gamma_0), \quad \text{(A6)}$$

where

$$\Delta w_\alpha^0 = \frac{1}{\gamma_0}\ln\frac{1+\gamma_0/T^*_{\alpha+1}(E-E^*_{\alpha+1})}{1+\gamma_0/T^*_\alpha(E-E^*_\alpha)}. \quad \text{(A7)}$$

Here we use the identity $T(E;\lambda_\alpha)=T^*_\alpha+\gamma_0(E-E^*_\alpha)$ since $T(E;\lambda_\alpha)$ crosses $[E^*_\alpha,T^*_\alpha]$.

Subtracting Eq. (A6) from Eq. (A4) results in the cancellation of $\Delta\mathcal{G}_\alpha^d(\gamma_0)$, yielding

$$\Delta_\alpha^+(\gamma_0) = \Delta\mathcal{G}_\alpha^0(\gamma_0) - \Delta w_\alpha^0(\gamma_0), \quad \text{(A8)}$$

which is always finite and well-defined for all γ_0. In the limit of $\gamma_0\to 0$, Eq. (A8) reduces to

$$\Delta_\alpha^+(0) = \Delta\mathcal{G}_\alpha^+(0) - \left(\frac{E}{T^*_{\alpha+1}} - \frac{E}{T^*_\alpha}\right), \quad \text{(A9)}$$

where

$$\Delta\mathcal{G}_\alpha^+(0) = -\frac{E^*_{\alpha+1}-E^*_\alpha}{T^*_{\alpha+1}-T^*_\alpha}\ln\frac{T^*_{\alpha+1}}{T^*_\alpha} + \frac{E^*_{\alpha+1}}{T^*_{\alpha+1}} - \frac{E^*_\alpha}{T^*_\alpha}.$$

Since $T_\alpha=T^*_\alpha$ and $U_\alpha=E^*_\alpha$ at $\gamma_0=0$, $\Delta\mathcal{G}_\alpha^+(0)$ becomes identical to $\Delta\mathcal{G}_{\alpha,\mathrm{ST}}^+$ and the second term in Eq. (A9) reduces to $\Delta w_\alpha^+(0)=w_{\mathrm{GB}}(E;T_{\alpha+1})-w_{\mathrm{GB}}(E;T_\alpha)$, illustrating that conventional ST is recovered via the analytic continuation of $\gamma_0\to 0$ in GST.

[1] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[2] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 1996).
[3] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon, Oxford, 1999).
[4] B. J. Berne and J. E. Straub, Curr. Opin. Struct. Biol. **7**, 181 (1997).
[5] A. Mitsutake, Y. Sugita, and Y. Okamoto, Biopolymers **60**, 96 (2001).
[6] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[7] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1195 (1989).
[8] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[9] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[10] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[11] J. P. Valleau, J. Chem. Phys. **99**, 4718 (1993).
[12] B. A. Berg and T. Celik, Phys. Rev. Lett. **69**, 2292 (1992); Phys. Lett. B **267**, 249 (1991).
[13] J. Lee, Phys. Rev. Lett. **71**, 211 (1993).
[14] C. J. Geyer and A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[15] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[16] I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996); J. Chem. Phys. **107**, 9117 (1997).
[17] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997); U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **56**, 2228 (1997).
[18] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[19] Y. Pak and S. Wang, J. Chem. Phys. **111**, 4359 (1999).
[20] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001).
[21] H. Fukunishi, O. Watanabe, and S. Takada, J. Chem. Phys. **116**, 9058 (2002).
[22] P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, Proc. Natl. Acad. Sci. U.S.A. **102**, 13749 (2005).
[23] S. Trebst, M. Troyer, and U. H. E. Hansmann, J. Chem. Phys. **124**, 174903 (2006).
[24] J. Kim, J. E. Straub, and T. Keyes, Phys. Rev. Lett. **97**, 050601 (2006); J. Chem. Phys. **126**, 135101 (2007); Phys. Rev. E **76**, 011913 (2007).
[25] C. Zhang and J. Ma, J. Chem. Phys. **132**, 244101 (2010).
[26] W. Kerler and P. Rehberg, Phys. Rev. E **50**, 4220 (1994).
[27] U. H. E. Hansmann and Y. Okamoto, Phys. Rev. E **54**, 5863 (1996).
[28] A. Irbäck and E. Sandelin, J. Chem. Phys. **103**, 10298 (1995).
[29] A. Irbäck, F. Sjunnesson, and S. Wallin, Proc. Natl. Acad. Sci. U.S.A. **97**, 13614 (2000).
[30] A. Mitsutake and Y. Okamoto, Chem. Phys. Lett. **332**, 131 (2000).
[31] J. Kim, Y. Fukunishi, and H. Nakamura, Chem. Phys. Lett. **392**, 34 (2004); J. Kim, Y. Fukunishi, A. Kidera, and H. Nakamura, Phys. Rev. E **69**, 021101 (2004).
[32] M. Fasnacht, R. H. Swendsen, and J. M. Rosenberg, Phys. Rev. E **69**, 056704 (2004).
[33] S. Park and V. S. Pande, Phys. Rev. E **76**, 016703 (2007).
[34] C. Zhang and J. Ma, Phys. Rev. E **76**, 036708 (2007).
[35] J. Kim, Y. Fukunishi, A. Kidera, and H. Nakamura, J. Chem. Phys. **121**, 5590 (2004).
[36] G. R. Bowman and V. S. Pande, Proteins **74**, 777 (2009).
[37] X. Huang, G. R. Bowman, and V. S. Pande, J. Chem. Phys. **128**, 205106 (2008).
[38] S. Rauscher, C. Neale, and R. Pomes, J. Chem. Theory Comput. **5**, 2640 (2009).
[39] S. Park, Phys. Rev. E **77**, 016709 (2008).
[40] C. Zhang and J. Ma, J. Chem. Phys. **129**, 134112 (2008).
[41] E. Rosta and G. Hummer, J. Chem. Phys. **132**, 034102 (2010).
[42] A. Mitsutake and Y. Okamoto, J. Chem. Phys. **130**, 214105 (2009); Phys. Rev. E **79**, 047701 (2009).
[43] A. Mitsutake, J. Chem. Phys. **131**, 094105 (2009).
[44] R. Denschlag, M. Lingenheil, P. Tavan, and G. Mathias, J. Chem. Theory Comput. **5**, 2847 (2009).
[45] L. Yang, O. Shao, and Y. Q. Gao, J. Chem. Phys. **130**, 124111 (2009).
[46] K. Huang, *Statistical Mechanics* (Wiley, New York, 1972).
[47] P. Labastie and R. L. Whetten, Phys. Rev. Lett. **65**, 1567 (1990).
[48] D. J. Wales and R. S. Berry, Phys. Rev. Lett. **73**, 2875 (1994).
[49] D. H. E. Gross, Rep. Prog. Phys. **53**, 605 (1990).
[50] M. D'Agostino, F. Gulminelli, Ph. Chomaz, M. Bruno, F. Cannata, R. Bougault, F. Gramegna, I. Iori, N. Le Neindre, G. V. Margagliotti, A. Moroni, and G. Vannini, Phys. Lett. B **473**, 219 (2000).
[51] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmuller, B. von Issendorff, and H. Haberland, Phys. Rev. Lett. **86**, 1191 (2001).
[52] C. Junghans, M. Bachmann, and W. Janke, Phys. Rev. Lett. **97**, 218103 (2006).
[53] H. Hernández-Rojas and J. M. Gomez, Phys. Rev. Lett. **100**, 258104 (2008).
[54] J. Kim, T. Keyes, and J. E. Straub, Phys. Rev. E **79**, 030902(R) (2009).
[55] J. Kim, Y. Fukunishi, and H. Nakamura, Phys. Rev. E **67**, 011105 (2003).
[56] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[57] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982); **55**, 315(E) (1983).
[58] J. Kim, T. Keyes, and J. E. Straub, J. Chem. Phys. **132**, 224107 (2010).
[59] M. S. S. Challa and J. H. Hetherington, Phys. Rev. Lett. **60**, 77 (1988).
[60] T. Neuhaus, M. P. Magiera, and U. H. E. Hansmann, Phys. Rev. E **76**, 045701(R) (2007).
[61] T. W. Whitfield, L. Bu, and J. E. Straub, Physica A **305**, 157 (2002).
[62] I. Fukuda and H. Nakamura, Phys. Rev. E **65**, 026105 (2002).
[63] S. Jang, S. Shin, and Y. Pak, Phys. Rev. Lett. **91**, 058305 (2003).
[64] S. Jang, E. Kim, and Y. Pak, J. Chem. Phys. **128**, 105102 (2008).
[65] J. Kim and J. E. Straub, J. Chem. Phys. **130**, 144114 (2009).
[66] J. Kim, T. Keyes, and J. E. Straub, J. Chem. Phys. **130**, 124112 (2009).
[67] P. Dayal, S. Trebst, S. Wessel, D. Wurtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, Phys. Rev. Lett. **92**, 097201 (2004).
[68] P. Poulain, F. Calvo, R. Antoine, M. Broyer, and Ph. Dugourd, Phys. Rev. E **73**, 056704 (2006).
[69] To fix a residual error in $\tilde{S}(E)$ refined by the WL, the multicanonical sampling (Ref. 12) was performed for additional 10^8 MCS with $\tilde{S}(E)$. A significant deviation from a uniform energy sampling was found in the resulting $P_{MU}(E)$. The corrected entropy estimate $\tilde{S}_{MU}(E)$ is obtained as $\tilde{S}(E)+\ln P_{MU}(E)$.
[70] D. A. Kofke, J. Chem. Phys. **120**, 10852 (2004); **121**, 1167 (2004).
[71] K. Y. Sanbonmatsu and A. E. Garcia, Proteins **46**, 225 (2002).
[72] N. Rathore, M. Chopra, and J. J. de Pablo, J. Chem. Phys. **122**, 024111 (2005).