

Optimal Parameters for Basin-Hopping Global Optimization Based on Tsallis Statistics

C. Shang^{1, a)} and D. J. Wales^{1, b)}

University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, UK

(Dated: 27 October 2014)

A fundamental problem associated with global optimization is the large free energy barrier for the corresponding solid-solid phase transitions for systems with multi-funnel energy landscapes. To address this issue we consider the Tsallis weight instead of the Boltzmann weight to define the acceptance ratio for basin-hopping global optimization. Benchmarks for atomic clusters show that using the optimal Tsallis weight can improve the efficiency by roughly a factor of two. We present a theory that connects the optimal parameters for the Tsallis weighting, and demonstrate that the predictions are verified for each of the test cases.

Locating the global minimum (GM) of a potential energy surface (PES) is one of the fundamental problems in theoretical chemistry. The basin-hopping algorithm (BH)^{1–4} has been widely used to search for the GM in a diverse range of systems, including atomic and molecular clusters, glass-formers, and biomolecules. The key idea in BH is to transform the PES using local minimization. After transformation, the contribution from the vibrational density of states to the entropy is replaced by the contribution from the volume of each basin of attraction, which reduces the entropy of higher energy states and thus accelerates the searching.⁵

One of the difficulties associated with global optimization is the exponential increase in the search space with system size. However, a transformed PES with a single deep funnel⁶ is relatively easy to handle.⁷ For multi-funnel PESs, it generally requires more effort to locate the GM, especially when the GM lies in a narrow funnel while the second-lowest minimum lies in a wider funnel, corresponding to higher entropy. For example, structures based on the Mackay icosahedron^{8,9} are often optimal for Lennard-Jones (LJ) clusters with the potential energy¹⁰

$$V = 4\epsilon \sum_{i < j} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (1)$$

where ϵ is the pair well depth and σ is the equilibrium pair separation. For the 55-atom cluster (LJ₅₅), the GM is a complete Mackay icosahedron.^{8,9} Hence there is only a single deep funnel on the PES. However, for LJ₃₈ and LJ₇₅, the GM is a truncated octahedron and a Marks' decahedron,¹¹ respectively. The barrier that separates the GM and the lowest energy icosahedral minimum is relatively high, and overlaps partly with minima corresponding to the liquid-like state of the cluster.^{12,13} The solid-solid transition between morphologies leads to a heat capacity feature at low temperature in each case.^{5,14} The optimal temperature for locating the GM fastest should be around this phase transition temperature, where the Boltzmann weights for the intermediate

states are small, leading to a large free energy barrier.⁵ To analyse this problem, here we use the Tsallis weight instead of the Boltzmann weight to define the acceptance ratio during the BH procedure.

The Tsallis weight was introduced as a basis for generalizing standard statistical mechanics.^{15–18} It has been used to define the acceptance ratio for some optimization problems before.^{19–21} The parameter q is used to characterize the generalization. Here we use $k = 1/(1 - q) > 0$ to simplify the equations. For the canonical ensemble, the q -exponential Tsallis weight $w(V)$ of a configuration with potential energy V relative to the lowest known minimum at V_{\min} is

$$w(V) = \begin{cases} [1 - (V - V_{\min})/(kT)]^k, & V - V_{\min} \leq kT, \\ 0, & V - V_{\min} > kT, \end{cases} \quad (2)$$

where T is the temperature used for the accept/reject step in BH. The local minima, and hence the weighting corresponding to the relative volumes of the corresponding basins of attraction, are the same as for BH with accept/reject steps based upon the Boltzmann factor $\exp(-V/k_B T)$. Below, we take $k_B = 1$ to define a convenient system of reduced units.

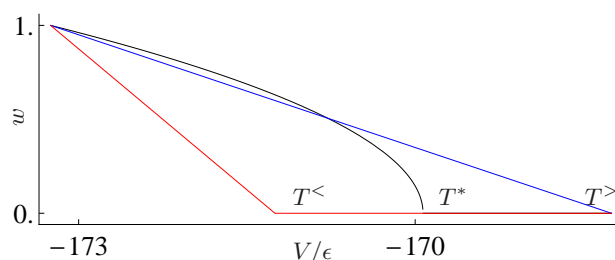


FIG. 1. Schematic illustration of relative Tsallis weights as a function of the potential energy, V , as a function of $k > 0$ at optimal, lower, and higher temperatures, T^* , $T^<$, and $T^>$, respectively. ϵ is the pair well depth from Eq. (1).

Fig. 1 illustrates the effect of choosing temperatures significantly larger or smaller than the optimal value for LJ₃₈.^{2,13} If the temperature is too small then the Tsallis weight, and hence the acceptance probability, falls to zero before we can reach local minima that enable us to cross

^{a)}cs778@cam.ac.uk

^{b)}dw34@cam.ac.uk

the barrier between the two morphologies. However, if T is too large then we explore higher energy minima well above this barrier, which is also expected to be inefficient.

We assume that the optimal temperature for BH with the Tsallis weight, T^* , should be a function of $k = 1/(1 - q)$. A suitable series expansion in k can be written as

$$kT^* = c_0 + c_1k + c_2k^2 \dots \quad (3)$$

When $k \rightarrow \infty$ ($q \rightarrow 1$), the Boltzmann weight is recovered from the Tsallis weight,¹⁵ which suggests that $T^*(k \rightarrow \infty) = T_B$ and $kT^* = c_0 + kT_B$, where T_B is the optimal temperature with the Boltzmann weight. When k is small, according to Eq. (2) we require $kT^* > \Delta V = V_{TS} - V_{\min}$, where V_{TS} is the energy of the top of the barrier separating the two funnels that contain the current lowest minimum and the GM. Hence we expect $kT^* \sim \Delta V + kT_B$. This is indeed the relation that we obtain for a variety of systems by numerical analysis.

For a basin-hopping step from V_1 to V_2 ($V_2 > V_1$), the acceptance probabilities for the Boltzmann and the Tsallis weights with optimal temperature in reduced units are

$$P_B = \exp(-\Delta V_{12}/T_B), \text{ and} \quad (4)$$

$$P_T = \left(\frac{kT^* - V_2 + V_{\min}}{kT^* - V_1 + V_{\min}} \right)^k = \left(1 - \frac{\Delta V_{12}}{V_{\max} - V_1} \right)^k, \quad (5)$$

where $\Delta V_{12} = V_2 - V_1 \geq 0$ and $V_{\max} = kT^* + V_{\min}$. When ΔV_{12} is small, we can expand P_T to first order in $\Delta V_{12}/(V_{\max} - V_1)$ as:

$$P_T = \left(1 - \frac{\Delta V_{12}}{V_{\max} - V_1} \right)^k = \exp \left[k \ln \left(1 - \frac{\Delta V_{12}}{V_{\max} - V_1} \right) \right] \quad (6)$$

$$\approx \exp \left(-\frac{k\Delta V_{12}}{V_{\max} - V_1} \right) = \exp \left(-\frac{\Delta V_{12}}{T^* + (V_{\min} - V_1)/k} \right).$$

Hence the Tsallis weight can be viewed as the Boltzmann weight with an effective temperature $T_{\text{eff}} = T^* + (V_{\min} - V_1)/k$. The possible advantages of using Tsallis weights during BH can then be understood as follows:

(1) The condition that $V - V_{\min} < kT^*$ is equivalent to $T_{\text{eff}} > 0$. The maximum energy satisfying this condition is V_{\max} in Eq. (5). The acceptance probability for any states with potential energy higher than this value is zero. Based on the relation $kT^* = \Delta V + kT_B$, we have $V_{\max} = kT_B + V_{TS}$, where V_{TS} is the energy at the top of the barrier. The region above V_{TS} that can be sampled has an energy width of kT_B , which decreases linearly with k . Compared with the Boltzmann weight, the Tsallis weight avoids sampling the area above the barrier, which is generally the liquid-like phase that is not relevant here.

(2) From Eq. (6), the effective temperature T_{eff} and P_T decrease with increasing the energy, while P_B in Eq. (4) is a constant. Based on the relation $kT^* = \Delta V + kT_B$, we find

$$\begin{cases} 1 > P_T > P_B, & T_{\text{eff}} \geq T_B, & \text{if } V_1 < V_2 \leq V_{TS}, \\ 0 < P_T < P_B, & T_{\text{eff}} \leq T_B, & \text{if } V_{TS} \leq V_1 < V_2, \\ 1 = P_T = P_B, & & \text{if } V_1 = V_2, \\ T_{\text{eff}} = T_B, & & \text{if } V_1 = V_2 = V_{TS}. \end{cases} \quad (7)$$

Compared with the Boltzmann weight at T_B , the Tsallis weight at T^* increases the sampling probability of the states beneath the top of the barrier, and decreases the probability for the states above the barrier. In particular, when $k \rightarrow 0$, the acceptance probability $P_T \rightarrow 1$ for $V_2 < V_{TS}$ and $P_T \rightarrow 0$ for $V_2 > V_{TS}$, which is equivalent to putting a limit at the top of the barrier and using an infinitely high temperature below it.

Overall, the effect of using Tsallis weights is (i) to increase the sampling probability of minima below the key transition state. This will decrease the free energy barrier between two funnels. (ii) It also decreases the sampling probability for minima above the top of the barrier and rejects moves to minima whose energies are much higher than V_{TS} . Both effects can assist escape from the wrong funnel via minima at the lower end of a liquid-like phase.

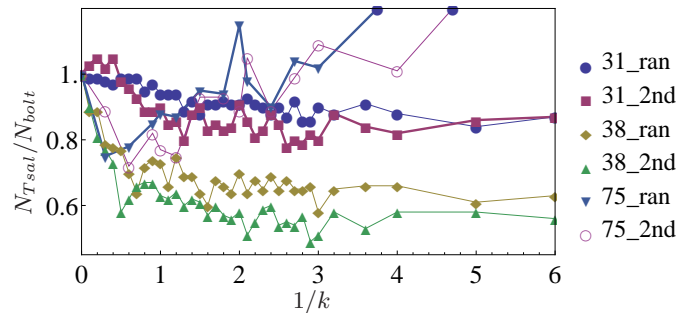


FIG. 2. The ratios of mean first encounter times in terms of step numbers between Boltzmann and Tsallis runs with different values of $1/k$ at the optimal temperatures. 31, 38 and 75 are the system sizes. “ran” corresponds to random starting points. “2nd” means that the starting points correspond to the second-lowest minimum, averaging over different random number seeds.

We first tested the efficiency for locating the GM of LJ₃₁, LJ₃₈ and LJ₇₅ clusters using Tsallis weights. For LJ₃₁, there are two competing low-energy minima with alternative Mackay and anti-Mackay overlayers.⁷ For LJ₃₈ and LJ₇₅, the second-lowest minimum lies in the funnel corresponding to incomplete Mackay icosahedral structures, and the true GM lies at the bottom of the funnel corresponding to the truncated octahedron for LJ₃₈ and to the Marks’ decahedron for LJ₇₅.^{11,22,23} Since icosahedral structures generally have higher entropy, corresponding to a wider funnel, the configuration space containing the GM is relatively narrow and deep.

The improvements obtained using Tsallis weights are shown in Fig. 2. For all clusters, test runs are considered starting from both random structures and from the second-lowest minimum. The results are averaged over 500 independent runs for LJ₃₁ and 100 independent runs for LJ₃₈ and LJ₇₅. The number of BH steps (N_{\min}) for LJ₃₁, LJ₃₈ and LJ₇₅ decrease by 16%, 42% and 25% respectively when starting from random structures and 22%, 51% and 28% when starting from the second-lowest minimum with different random number seeds. The number of energy function calls decreases by

roughly the same amount. The best efficiency is achieved when $1/k$ is around 2.5 for LJ₃₁, 3.0 for LJ₃₈ and 0.6 for LJ₇₅. For LJ₃₁ and LJ₃₈, the efficiency of Tsallis weighting decreases slightly on further decreasing k . The ratio of BH step numbers between the Boltzmann and Tsallis varies within $\pm 5\%$. For LJ₇₅, the efficiency of Tsallis weighting decreases significantly for small k and is inferior to the Boltzmann weight when $1/k > 2$.

For the Tsallis weight with $1/k = 10$, We have run LJ₃₈ from random structures at 10 temperatures scanning the range from T^* to $T^* + 10\epsilon$. The average number of BH steps is 2168, which is still lower than the 2200 steps for the Boltzmann weight with an optimal temperature. In contrast, for the Boltzmann weight, even increasing the temperature by 0.5ϵ , the number of steps increases to 4700. Hence, employing the Tsallis weight may be advantageous when investigating an unknown system, since the results are less sensitive to the temperature parameter if k is relatively small.

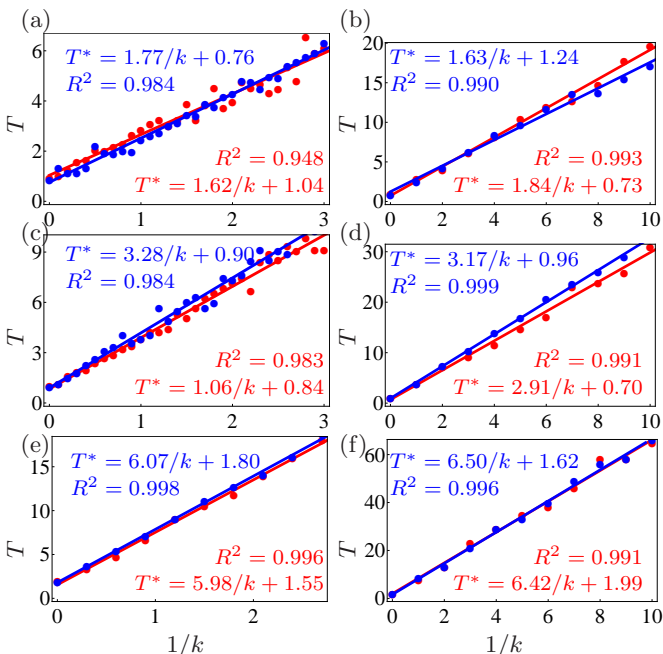


FIG. 3. The linear relationship between the optimal temperature T^* and $1/k$ for LJ₃₁ (a-b), LJ₃₈ (c-d), and LJ₇₅ (e-f). The optimal temperature T^* corresponds to the minimum number of BH steps. The blue lines and points correspond to starting from the second-lowest minimum. The red lines and points correspond to starting from random structures. For each pair of k and T values, the results are averaged over 500 independent runs for LJ₃₁ and 100 runs for LJ₃₈ and LJ₇₅.

The linear relationship obtained numerically between T^* and $1/k$ for Lennard-Jones clusters is examined in Fig. 3. The optimal temperature T^* corresponds to the minimum number of BH steps. The fit for T^* in terms of the number of energy function calls is roughly the same as Fig. 3 but a little noisier. For Boltzmann weights, when starting from random structure, the optimal temperatures are 0.95ϵ for LJ₃₁, 1.00ϵ for LJ₃₈ and 2.0ϵ for LJ₇₅. When starting from the second-lowest minimum,

the optimal temperatures are 0.87ϵ for LJ₃₁, 0.93ϵ for LJ₃₈ and 1.9ϵ for LJ₇₅. These results are consistent with the intercepts shown in Fig. 3. The slopes we obtain from Figs. 3(c-d) also exhibit good agreement with the known energy barrier for LJ₃₈ separating the second-lowest minimum from the GM, which is 3.04 .¹³ We note that the barrier here refers to the energy difference between the highest energy intermediate minimum and the second-lowest minimum along the lowest energy pathway.

Comparing the three clusters, LJ₃₈ and LJ₇₅ exhibit better linear fits and also better approximations for T_B in the high k limit. This result is reasonable because LJ₃₈ and LJ₇₅ have clearer double-funnel structures for the energy landscape. Since the approximate optimal temperature for Tsallis statistics is mainly based on the energy barrier from the alternative funnel to the funnel that contains the GM, a more pronounced double-funnel energy landscape ensures that most sampling happens within the two main funnels rather than other regions of the PES. Based on the same reasoning, starting from the second-lowest minimum generally follows the linear relationship more accurately. Starting from the second-lowest minimum also gives a larger value for ΔV , which is closer to the true value between the funnels. The ΔV observed when starting from random structures is more likely to average over other parts of the landscape.

A simple model can be used to illustrate the relationship between k and T^* . The model corresponds to a random walker²⁴ whose PES is shown in Fig. 4. A and B represent two funnels that are separated by a transition region with a width of w_T . The transition region has an average energy of V_{TS} and contains n_{TS} local minima. The walker starts from B and moves a distance of ± 1 at each step towards A. The widths of the two funnels are w_A and w_B , respectively. For simplicity, we only discuss the situation where $w_T \gg w_B \geq w_A$. Hence for relevant temperatures, the average number of steps required to move from B to A is mainly determined by w_T .

A high energy phase with a constant density of minimum, ρ_L , is connected to the transition region. The energy of minima in this phase spans the range $V_L^1 \leq V_L \leq V_L^2$, which satisfies $V_L^1 < V_{TS} \ll V_L^2$. The number of minima n_L for this high energy phase is then $n_L = \rho_L (V_L^2 - V_L^1)$. When the walker reaches the transition region, an equilibrium between this region and the high energy phase will be rapidly achieved. Hence the probability that the walker stays in the transition region with the Boltzmann weight is

$$\begin{aligned}
 P_{LT} &= \frac{n_{TS} \exp(-V_{TS}/T)}{n_{TS} \exp(-V_{TS}/T) + \sum n_i \exp(-V_i/T)} \\
 &= \frac{n_{TS} \exp(-V_{TS}/T)}{n_{TS} \exp(-V_{TS}/T) + \int_{V_L^1}^{V_L^2} \rho_L \exp(-V_L/T) dV_L} \quad (8) \\
 &= \left\{ 1 + \frac{\rho_L T [\exp(-V_L^1/T) - \exp(-V_L^2/T)]}{n_{TS} \exp(-V_{TS}/T)} \right\}^{-1}.
 \end{aligned}$$

To minimize N_s , we maximize P_{LT} and solve

$dP_{LT}/dT = 0$ to obtain:

$$T_B = V_{TS} - V_L^1 + \exp\left(\frac{V_L^1 - V_L^2}{T_B}\right)(T_B - V_{TS} + V_L^2) \quad (9)$$

$$\approx V_{TS} - V_L^1.$$

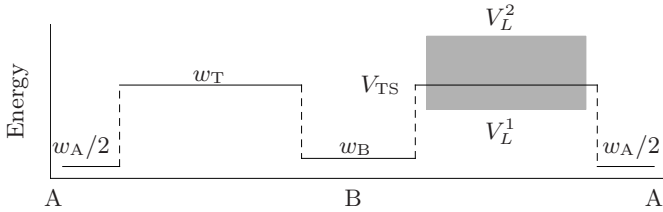


FIG. 4. The model landscape used to represent a transition between two funnels. The PES is symmetric about point “B”.

From previous work^{13,25} for LJ₃₈, we know that the energy at the top of the barrier is $V_{TS} = -170.3$, and the lower boundary of the liquid-like phase is $V_L^1 = -171.5$. Then the calculated optimal temperature according to Eq. (9) is 1.2ϵ , which is in good agreement with the optimal T_B and around the same order of magnitude as the approximate melting point.⁵

Now we replace the Boltzmann weight in Eq. (8) with the Tsallis weight. Then the probability P_{LT} is:

$$P_{LT} = \frac{n_{TS}(kT - V_{TS})^k}{n_{TS}(kT - V_{TS})^k + \sum n_i(kT - V_i)^k}$$

$$= \frac{n_{TS}(kT - V_{TS})^k}{n_{TS}(kT - V_{TS})^k + \int_{V_L^1}^{kT} \rho_L(kT - V_i)^k dV_L} \quad (10)$$

$$= \left[1 + \frac{\rho_L(kT - V_L^1)^{k+1}}{n_{TS}(k+1)(kT - V_{TS})^k} \right]^{-1}.$$

Here we only discuss the condition $kT < V_L^2$, so that the upper limit of the integral in Eq. (10) is “ kT ” rather than “ V_L^2 ”. From Eq. (10), we then calculate $dP_{LT}/dT = 0$ and obtain:

$$T^* = \frac{V_{TS}}{k} + V_{TS} - V_L^1 = \frac{V_{TS}}{k} + T_B, \quad (11)$$

which is exactly the linear relationship that we have found for Lennard-Jones clusters.

There is a possible expansion of the current theory. Consider the hopping step from V_1 to V_2 in terms of several small continuous steps. For each small step, the energy increases by dV and the acceptance probability is determined by the Boltzmann weight. Then for a given weight, if the effective temperature varies as a function of potential energy as $T_i = f(V_i)$, the acceptance probability from V_1 to V_2 can be written as:

$$p = \prod_i \exp(-dV/T_i) = \exp\left(\sum_i -dV/T_i\right) \quad (12)$$

$$= \exp\left[-\int_{V_1}^{V_2} f(V)^{-1} dV\right].$$

From Eq. (12), we can obtain the Boltzmann weighted acceptance probability when $f(V)$ is a constant, and the Tsallis weighted acceptance probability when $T_i = T - (V_i - V_{\min})/k$.

From Eq. (7) and the results represented here, if the effective temperature T_i decreases monotonically with increasing potential energy, then the optimal temperature should approximately satisfy the condition that $f(V_{TS}) = T_B$, where V_{TS} is the energy of the top of the barrier and T_B is the optimal temperature for the Boltzmann weight. The states below V_{TS} then always have an effective temperature higher than T_B , while the states above V_{TS} , which are generally liquid-like phases, always have an effective temperature lower than T_B .

In summary, we have presented a theory and numerical verification of an improved global optimisation procedure that employs Tsallis weights. Our tests include multifunnel landscapes that exhibit broken ergodicity and competing morphologies, as well as an analytical example. Hence the proposed framework should prove to be generally applicable to structure prediction throughout molecular science and soft and condensed matter systems.

ACKNOWLEDGMENTS

This work was supported by the ERC and the EPSRC.

- ¹Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 6611 (1987).
- ²D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
- ³D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
- ⁴D. Wales, *Energy Landscapes*, edited by C. U. Press, Cambridge Molecular Science (Cambridge University Press, Cambridge, U. K., 2003).
- ⁵J. P. K. Doye and D. J. Wales, Phys. Rev. Lett. **80**, 1357 (1998).
- ⁶O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
- ⁷J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **111**, 8417 (1999).
- ⁸A. L. Mackay, Acta Crystallogr. **15**, 916 (1962).
- ⁹J. A. Northby, J. Chem. Phys. **87**, 6166 (1987).
- ¹⁰J. Jones and A. Ingham, Proc. R. Soc. London, Ser. A **107**, 636 (1925).
- ¹¹L. D. Marks, Philos. Mag. A **49**, 81 (1984).
- ¹²D. J. Wales, Mol. Phys. **102**, 891 (2004).
- ¹³D. J. Wales, Mol. Phys. **100**, 3285 (2002).
- ¹⁴J. P. K. Doye, D. J. Wales, and M. A. Miller, J. Chem. Phys. **109**, 8143 (1998).
- ¹⁵C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
- ¹⁶C. Tsallis, Quim. Nova **17**, 468 (1994).
- ¹⁷C. Tsallis and D. A. Stariolo, Physica A (Amsterdam, Neth.) **233**, 395 (1996).
- ¹⁸C. Tsallis, Braz. J. Phys. **29**, 1 (1999).
- ¹⁹I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996).
- ²⁰U. H. Hansmann and Y. Okamoto, Braz. J. Phys. **29**, 187 (1999).
- ²¹H. Liu and K. D. Jordan, J. Phys. Chem. A **109**, 5203 (2005).
- ²²J. Pillardy and L. Piela, J. Phys. Chem. **99**, 11805 (1995).
- ²³J. P. K. Doye, D. J. Wales, and R. S. Berry, J. Chem. Phys. **103**, 4234 (1995).
- ²⁴B. Weesakul, Ann. Math. Stat. **32**, 765 (1961).
- ²⁵J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **110**, 6896 (1999).