Computer Physics Communications 121–122 (1999) 40–42

Computer Physics Communications

www.elsevier.nl/locate/cpc

Hybrid simulated annealing using Tsallis statistics

Rafael Salazar [1], Raúl Toral

Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC-UIB), Campus Universitat Illes Balears, E-07071 Palma de Mallorca, Spain [2]

Abstract

We use the Hybrid Monte Carlo method to sample the Tsallis configurational probability $p(\mathbf{x})$ which depends on the parameter q. We use here a homogeneous Markov chain which satisfies the detailed balance condition for all values of q and we apply this method to Simulated Annealing optimization problems. In two examples, a test function with many minima and the discrete Random phase sine-Gordon Model, we show the dependence on the q parameter of the ground state configuration energy obtained after annealing. We conclude that the performance of the simulated annealing method is improved by using $q > 1$ values. © 1999 Published by Elsevier Science B.V. All rights reserved.

The Tsallis generalized formulation of Statistical Mechanics [2] has been successful in describing the behavior of some non-extensive systems such as the large-range Ising ferromagnet, Lévy-like anomalous diffusion, two-dimensional turbulence, etc. The key ingredient of the new statistics (in the canonical ensemble) is to assign to the $\mathbf{x} \equiv (x_1, \ldots, x_N)$ configuration with energy E, the following probability:

$$p(\mathbf{x}) = \frac{[1-(1-q)E/T]^{1/(1-q)}}{Z_q}. \tag{1}$$

The parameter q depends on the system under study. This expression is such that one can recover the Boltzmann–Gibbs classical probability in the limit $q \to 1$. The generalized partition function Z_q is determined by the condition $\sum_{\mathbf{x}} p(\mathbf{x}) = 1$. Since $p(\mathbf{x})$ must be real probabilities, one needs to take $p(\mathbf{x}) = 0$, when $E(\mathbf{x}) > T/(1-q)$ for $q < 1$, and in the case $q > 1$ whenever $E(\mathbf{x}) < T/(1-q)$.

The Tsallis generalized statistics has been already applied to the Simulated Annealing method (SA). In Ref. [3], the authors introduce a new method (which they call Generalized Simulated Annealing, GSA) which is a generalization of the precedents Boltzmann SA and Fast SA methods. This generalization was made in the context of the Monte Carlo methods by using a proposal function $g(\mathbf{x}'|\mathbf{x})$, an acceptance probability $h(\mathbf{x}'|\mathbf{x})$ and an annealing schedule $T = T(t)$, in the sense that these generalized functions depend on a q parameter in such a way that when $q \to 1$ one recovers the Boltzmann SA and when $q \to 2$ one recovers the Fast SA. However, the GSA method does not satisfy the main Monte Carlo prescription, the detailed balance condition (DBC):

$$g(\mathbf{x}'|\mathbf{x})h(\mathbf{x}'|\mathbf{x})p(\mathbf{x}) = g(\mathbf{x}|\mathbf{x}')h(\mathbf{x}|\mathbf{x}')p(\mathbf{x}'). \tag{2}$$

The validity of the GSA is not questioned because one can use always an inhomogeneous Markov chain, instead of the homogeneous one satisfying the DBC; moreover, it has been very recently shown [4] that GSA satisfies the weak ergodicity condition, which gives a mathematical foundation to the conjectured convergence of the GSA process to the minimum of the cost function. Ref. [5] introduces a SA method

[1] E-mail: rafael@hp1.uib.es.

[2] http://www.imedea.uib.es.

0010-4655/99/$ – see front matter © 1999 Published by Elsevier Science B.V. All rights reserved.
PII: S0010-4655(99)00275-1

which satisfies the DBC while sampling a probability proportional to $p(\mathbf{x})^q$.

In this work we use the Hybrid Monte Carlo (HMC) method (which satisfies DBC) to sample the Tsallis distribution $p(\mathbf{x})$. As an application, we study the performance of the SA when using values of q different from 1. In previous work [6] we showed that the use of the HMC algorithm can enhance the performance of SA methods in the case of Boltzmann Statistics for large dimensional systems. The Hybrid Monte Carlo method uses the numerical integration of the molecular dynamics equations of motion as the proposal function $g(\mathbf{x}'|\mathbf{x})$:

$$\begin{aligned}\mathbf{x}' &= \mathbf{x} + \frac{\delta t^2}{2}\mathbf{F}(\mathbf{x}) + \delta t\mathbf{p},\\ \mathbf{p}' &= \mathbf{p} + \frac{\delta t}{2}\big(\mathbf{F}(\mathbf{x}) + \mathbf{F}(\mathbf{x}')\big),\end{aligned} \tag{3}$$

where $\mathbf{p}$ are random variables distributed according to the Gaussian distribution $\exp[-\mathbf{p}^2/2T]$. We use the Metropolis solution of the DBC for the acceptance probability

$$h(\mathbf{x}'|\mathbf{x}) = \min\big(1, \exp\big[-\big(\tilde{H}(\mathbf{x}', \mathbf{p}') - \tilde{H}(\mathbf{x}, \mathbf{p})\big)/T\big]\big)$$

where we consider an effective Hamiltonian $\widetilde{H}(\mathbf{x}, \mathbf{p}) = -T \ln p(\mathbf{x}) + \mathbf{p}^2/2$, which for the Tsallis probability takes the form:

$$\begin{aligned}\widetilde{H}(\mathbf{x}, \mathbf{p}) &= -\frac{T}{1-q}\ln\big[1 - (1-q)E(\mathbf{x})/T\big] + \frac{\mathbf{p}^2}{2}\\ &\equiv V_{\text{eff}}(\mathbf{x}) + \frac{\mathbf{p}^2}{2}\end{aligned} \tag{4}$$

and the forces in Eq. (3) are:

$$F_i(\mathbf{x}) = -\frac{\partial V_{\text{eff}}(\mathbf{x})}{\partial x_i}.$$

In order to avoid that the effective potential vanishes when the temperature is annealed to zero, we assume a temperature dependence on $q(T) = 1 - (1-q_0)T/T_0$, where T_0 is a reference temperature to set the potential. As in Ref. [6], we use here an exponential annealing schedule $T(t) = T(0)\mathrm{e}^{-ct}$. To show the dependence of this method on the parameter q_0 (our previous work [6] is equivalent to using $q_0 = 1$) we use two examples. The first one is the test function:

$$E(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{N}\frac{\sin(4\pi K x_i)}{\sin(2\pi x_i)} + NK \tag{5}$$

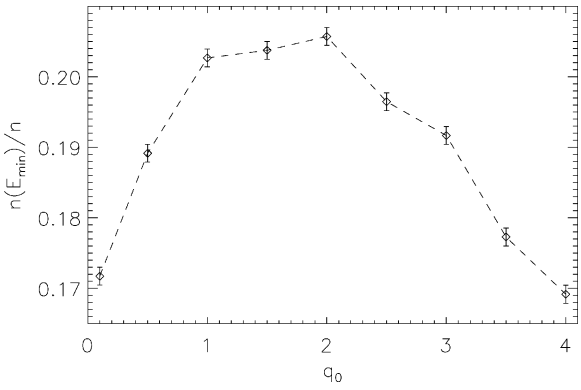

Fig. 1. Relative number of realizations which reach the lowest energy state, $E_{\min} = 0$, for the first example Eq. (5). The values shown are the result of sampling over $n = 2.5 \times 10^4$ realizations. We use here $T_0 = T(0) = 120$, and $c = 0.05$.

with $N = 30$, $K = 2$. In Fig. 1 we plot the relative number of realizations which, after annealing, reach the absolute minimum for a particular value of the annealing parameters. We can see from this plot that the highest performance is obtained in this example by using a value close to $q_0 = 2$.

The second example aims to find the ground state ($T = 0$) of the discrete Random phase sine-Gordon Model (RsGM), which has recently been the scene of considerable activity, both theoretical and numerical [7]. This model has various interpretations of interest such as a crystalline surfaces with disordered substrate and an array of flux lines in a superconducting film subject to random pinning and parallel magnetic field. The energy for this model is:

$$E(\mathbf{x}) = \frac{1}{2}\sum_{\langle i,j\rangle}(x_i - x_j)^2 - \sum_i \cos(x_i - x_j^0), \tag{6}$$

where $\langle i, j\rangle$ denotes nearest neighbors in a square lattice, and $x_j^0 \in [0, 2\pi)$ is the quenched random disorder. The results are shown in Fig. 2 for a particular value of the annealing parameters. We see again that the lowest minimum value for the energy is obtained with highest probability using $q_0 > 1$. We note that a usual approach [9] to find the ground state of the RsGM uses the Langevin molecular dynamics overdamped equations at $T = 0$. In this method, the relative number of successes is of the order of 4%. The fact that one can obtain improvements in the performance of the sampling method of the configurational

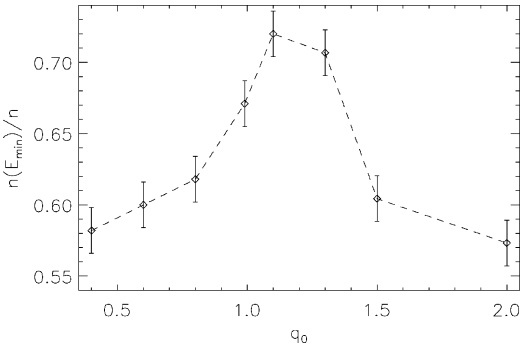

Fig. 2. Same as Fig. 1 for the ground state energy of the RsGM. We plot the relative number of realizations which reach the lowest energy $E_{\min} = -157.8810$ after averaging over $n = 200$ realizations, and using the same pinning values $\mathbf{x}^0$ in all the cases. We use here a two-dimensional lattice with $N = 32^2$ points and $T_0 = 10^5$, $T(0) = 100$, $c = 0.018$.

space, when one uses $q_0 > 1$ values, is in agreement with the results presented in [5] and is a consequence from the fact that the distribution $p(\mathbf{x})$ with $q > 1$ is broader than the one obtained using $q = 1$.

Acknowledgements

We acknowledge financial support from DGES (Spain) project numbers PB94-1167 and PB97-0141-C02-01.

References

[1] C. Tsallis, J. Stat. Phys. 52 (1988) 479.
[2] C. Tsallis, Physica A 221 (1995) 277.
[3] D.A. Stariolo, C. Tsallis, in: Annual Reviews of Computational Physics II, D. Stauffer (Ed.) (World Scientific, Singapore, 1995).
[4] H. Nishimori, J. Inoue, J. Phys. A 31 (1998) 5661.
[5] I. Andricioaei, J.E. Straub, Phys. Rev. E 53 (1996) R3055.
[6] R. Salazar, R. Toral, J. Stat. Phys. 89 (1997) 1047.
[7] D. Cule, Y. Shapir, J. Phys. A 29 (1996) 21.
[8] T. Giamarchi, P. Le Doussal, Phys. Rev. Lett. 72 (1994) 1530.
[9] A. Sánchez, A.R. Bishop, D. Cai, N. Grønbech-Jensen, F. Domíngez-Adame, Physica A 107 (1997) 326.