Simulated annealing with Tsallis weights for a system of interacting hard spheres



2002 J. Phys. A: Math. Gen. 35 6261

(http://iopscience.iop.org/0305-4470/35/30/304)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 6261–6269 PII: S0305-4470(02)33987-8

# Simulated annealing with Tsallis weights for a system of interacting hard spheres

**J Kłos**[1] **and S Kobe**[2]

[1] Instytut Fizyki, Uniwersytet im. A Mickiewicza, Coll. Physicum, ul. Umultowska 85, 61-614 Poznań, Poland
[2] Institut für Theoretische Physik, Technische Universität Dresden, D-01062 Dresden, Germany

Received 15 February 2002, in final form 20 May 2002
Published 19 July 2002
Online at stacks.iop.org/JPhysA/35/6261

**Abstract**

In this paper we consider the Monte Carlo simulated annealings based on the Tsallis and threshold acceptance probabilities. This is done by applying them to the 2D system of many particles both interacting among themselves by a dipole–dipole potential and affected by the gravitational force. Our results confirm that for large and negative values of the Tsallis parameter $q$, the generalized method leads to low-energy configurations at higher temperatures, in fewer Monte Carlo steps and after shorter CPU time. We also suggest a qualitative explanation for them in terms of the acceptance probabilities used.

PACS numbers: 02.60.Pn, 05.10.Ln, 0.5.20.−y

## 1. Introduction

Finding low-energy configurations for systems with many interacting bodies is a very hard minimization problem. Due to a huge number of local minima separated from each other by barriers in the energy function, the simplest and direct methods such as the gradient descent one are likely to fail. In order to handle such systems more powerful algorithms with the statistical mechanics background should be applied instead. Among the simulated annealings [1] is one of the most useful and efficient methods. The idea can be implemented by a Monte Carlo (MC) approach which ensures that the decrease of the temperature is in accordance with the annealing schedule. To avoid trapping in metastable states the cooling process must be sufficiently slow, and at each annealing step the system ought to be not far from thermalized. The method, although generally successful, brings about some technical difficulties, mainly caused by the annealing schedule itself. Geman and Geman [2] showed that if the temperature is lowered as the inverse logarithm of time the algorithm is sure to find a global minimum. However, this kind of cooling protocol is extremely slow and therefore of very limited practical use. Because of that some faster annealings have been worked out where arriving at the ground state(s) is no longer guaranteed, and approximate solutions are accepted [3]. Some of the latest

  

ideas devoted to the method are based on the Tsallis generalized statistics [4, 5] that is hoped to be most useful in explaining properties of nonextensive systems. The spectacular examples of successful applications of generalized simulated annealing include the solution of the travelling salesman problem and of the protein folding one [6–8]. It turned out that, with a proper choice of the Tsallis parameter $q$, it was possible to improve the efficiency of the procedure compared with that of classical simulated annealing. The optimal configurations were found in fewer annealing steps and at higher temperatures. Recently, a modified version of the generalized acceptance probability has been proposed, such that for some finite values of $q$, the Metropolis and the original Tsallis ones are recovered, and furthermore, the limit $q \to -\infty$ leads to the so-called threshold accepting. It has been shown that the optimal objective function values such as the final mean energy and the final ground state probability, reached with the use of the optimal annealing schedule, are better with decreasing $q$. Therefore, it has been suggested that in order to obtain the best optimization results, the threshold simulated annealing ought to be preferred to Tsallis statistics [9–11].

In this paper, we present the application of the Tsallis and threshold simulated annealings to a two-dimensional (2D) system of interacting hard spheres, first introduced by Pierański [12]. Steel spheres in a horizontal box were magnetized with a vertically oriented magnetic field so that they would repulse each other. By a slight tilting of the box the gravitational force could act on the spheres. Since the resulting stable configurations of the system recalled rainbows, they were called *the gravity rainbow*. Some interesting physical properties of the low-energy states of the whole class of similar ones, i.e., of the so-called *conformal crystals*, have also been found and reported. Defined as 2D structures resulting from conformal mappings of subsets of 2D periodic lattices, they have appeared in systems such as crystals, magnetic holes and natural, biological ones [13–16]. Conformal crystals have already been investigated by means of the elasticity theory and a version of the numerical gradient method [13, 17].

In this paper we compare Tsallis simulated annealing with the Metropolis and threshold ones. The comparison criteria include the speed of minimization with respect to the temperature, the number of MC steps per particle (one MCS is defined as $N$ random reconfigurations in particle positions, where $N$ denotes the number of the particles) and the CPU time needed to complete the calculations. The efficiency of the algorithms is analysed by considering the mean final energies. We also discuss the results in terms of the acceptance probabilities used, and suggest an intuitive explanation for them. Successful applications of generalized simulated annealings to magnetized particles in an external field are likely to be an indication that it might also be useful for the whole class and other systems of many interacting bodies. Moreover, a low-energy configuration obtained for a larger system is presented as an example.

## 2. Model and simulations

The generalized acceptance probability, inspired by the Tsallis statistics, is of the form

$$p_q(\Delta E/T) = \begin{cases} 1 & \text{if} \quad \Delta E \leqslant 0 \\ [1-(1-q)\Delta E/T]^{\frac{1}{(1-q)}} & \\ & \text{if} \quad \Delta E > 0 \text{ and } (1-q)\Delta E/T < 1 \\ 0 & \text{if} \quad \Delta E > 0 \text{ and } (1-q)\Delta E/T \geqslant 1 \end{cases} \tag{1}$$

where $\Delta E = E_{\text{new}} - E_{\text{old}}$ ($E_{\text{new}}$ and $E_{\text{old}}$ are the energies after and before a random reconfiguration, the Boltzmann constant $k_B$ is set to 1) and $q$ stands for the Tsallis parameter

appearing in the generalized canonical distribution of the states. It can be shown that the limit $q \to 1$ yields the traditional Metropolis algorithm [5, 18]

$$p_1(\Delta E/T) = \begin{cases} 1 & \text{if} \quad \Delta E \leqslant 0 \\ \exp(-\Delta E/T) & \text{if} \quad \Delta E > 0. \end{cases} \tag{2}$$

Since algorithm (2) obeys the detailed balance condition, it is secured that at fixed temperature $T$, irrespective of the initial state of the system, the generated Markov chain converges to the Boltzmann equilibrium distribution. The third implementation that we take into account is threshold accepting defined as

$$p_{TA}(\Delta E/T) = \begin{cases} 1 & \text{if} \quad \Delta E/T \leqslant 1 \\ 0 & \text{if} \quad \Delta E/T > 1. \end{cases} \tag{3}$$

From the point of view of the computational effort this annealing seems advantageous because, unlike the previous algorithms, there is no need to evaluate the exponential or power functions. It should also be stressed that for both the Tsallis and the threshold accepting the detailed balance condition does not hold for finite $T$. This in turn means that the convergence towards an equilibrium distribution is not *a priori* guaranteed, and the master equation should be analysed instead. However, as long as the stochastic optimization itself is the main object of the investigation (unlike statistical averaging which requires an equilibrium distribution), these two probabilities are recognized techniques [9].

We use the above algorithms for the 2D system of $N$ hard spheres confined to an open rectangular box with infinitely high walls and the bottom of width $L$. Due to the magnetic moments of the particles caused by the magnetic field, they repulse each other by the dipole–dipole potential, and additionally, are all affected by the gravity. The dimensionless total energy

$$E = \alpha \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij}^{-3} + U_{hc} + \sum_{i=1}^{N} z_i \tag{4}$$

is subject to the minimization procedure with the free parameters $\alpha$, $L$, $N$. The parameter $\alpha$ denotes the magnitude of the applied magnetic field, $U_{hc}$ is the hard core potential of the box edges and the particles and $z_i$ stands for the vertical coordinate of the $i$th particle [17].

The annealings are carried out by applying the MC algorithms within single-particle dynamics with the auxiliary condition for the particles not to leave the box and not to overlap with each other. The cooling schedule is such that the temperature is lowered exponentially

$$T = T_i \left( \frac{T_f}{T_i} \right)^{\frac{n}{N_a}} \tag{5}$$

either after $100N$ trial reconfigurations or after $10N$ accepted rearrangements in the particle positions, depending on which occurs earlier. $T_i, T_f$ are the initial and final temperatures, respectively, $N_a$ denotes the total number of annealing steps ($n = 0, 1, 2, \ldots, N_a$). Though, this kind of temperature schedule is very likely not optimal for the considered system in the case of the optimal control theory [11], we have chosen it for the comparison because of its simplicity and good performance reported earlier in the literature [6–8].

## 3. Results

Some representative results for the system with $N = 100$ are shown in figures 1–3, where in the case $q = 1$ the calculations were performed using the traditional Metropolis algorithm (2). Each curve is obtained after averaging the computational data over 50 independent runs with

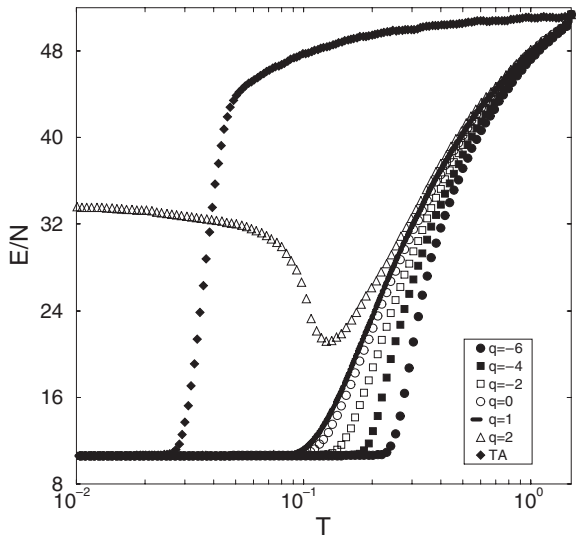


**Figure 1.** Energy per particle $E/N$ versus temperature $T$ (both in dimensionless energy units) averaged over 50 independent runs for the system with $N = 100$, $L = 32.4$, $\alpha = 15.2$. The data refer to various $q$ in equation (1) and to threshold accepting defined by equation (3). For $q = 1$ the Metropolis acceptance probability (2) is used.

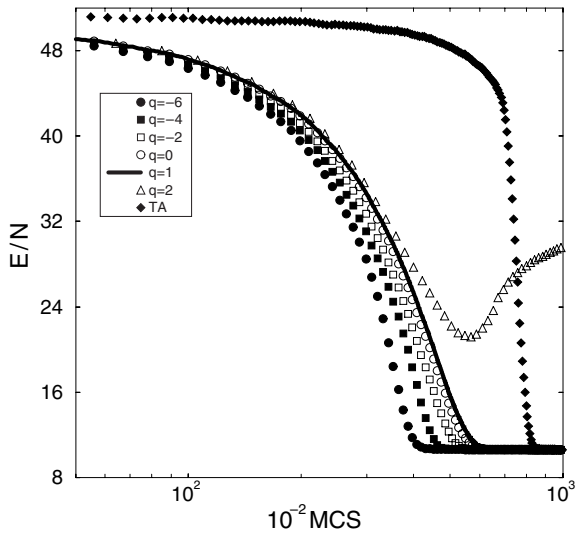


**Figure 2.** Energy per particle $E/N$ versus MCS averaged over 50 independent runs for the same system and acceptance probabilities as in figure 1.

random initial positions of the particles and identical values of $L, N, \alpha$. Using equation (1), low-energy states are approached at higher temperatures for $q < 1$, especially for large and negative values of this parameter (figure 1). This observation is in qualitative agreement with the earlier results, where a similar tendency was reported [6, 7, 19, 20]. For $q$ not far from one no major difference from the Metropolis algorithm can be seen (due to the fact that for $q \to 1$ the acceptance probability of equation (2) is recovered). The minimization carried out with

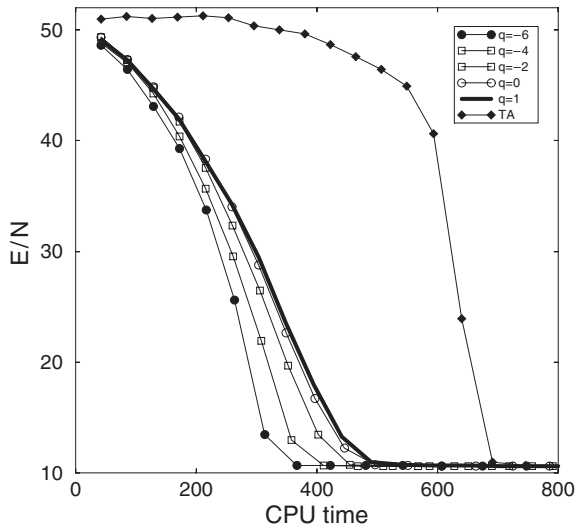


**Figure 3.** Energy per particle $E/N$ versus CPU time (in seconds) for both the same system and acceptance probabilities as in figure 1. Each time measurement was made after every 500 annealing steps.

the threshold acceptance probability of equation (3) proceeds in a different way. At higher temperatures no major energy decrease can be observed and the tendency to search for its lower values with dropping temperature is slower. The same conclusion is reached with respect to the number of MCS needed for acceptable minima (figure 2). Equation (1) leads to states with low energy after a considerably smaller number of them compared with equations (2) and (3). Since the values of $q < 1$ are expected to cause the algorithm to be more effective in local rather than global optimizations [7], we have also performed a few simulations with a huge number of cooling steps with $q > 1$. However, each time at a certain temperature the minimization either breaks down completely or still requires many more annealing steps.

For practical reasons, the comparison between the acceptance probabilities should also be supplemented by considering the CPU time as the next criterion. This is because a faster performance with the temperature need not simultaneously be the case with respect to the CPU time [7]. Depending on the considered system, the adapted annealing protocol and the algorithm itself, the minimization with respect to the former might be relatively slow, whereas that with respect to the latter fast or vice versa. Actually, the application of a certain algorithm that enables the reaching energies considerably lower than those obtained with others, after the same number of annealing steps, might cost much more CPU time. However, as indicated in figure 3, it does not take place for the system considered here. The complete minimization with the use of equation (1) with $q = -6$ consumes about 40% less CPU time than that of equation (2) or (3).

Finally, the final mean energies $E_{\text{final}}$ reached at the end of the annealings are shown in figure 4. For the exponential temperature sequence, threshold accepting turns out to be slightly less optimal than the others. The Tsallis acceptance probability reveals a tendency to be more optimal, especially for the values of $q < 0$. In fact, figure 4 suggests that for the adapted annealing schedule there exists an intermediate regime of $q \approx -4$ for which the algorithm is both very fast and optimal. Further acceleration of the minimization with decreasing $q$ is, however, likely to bring about various difficulties. Firstly, for extremely large absolute values of the parameter some roundoff errors are going to emerge [6]. Secondly, as the analysis

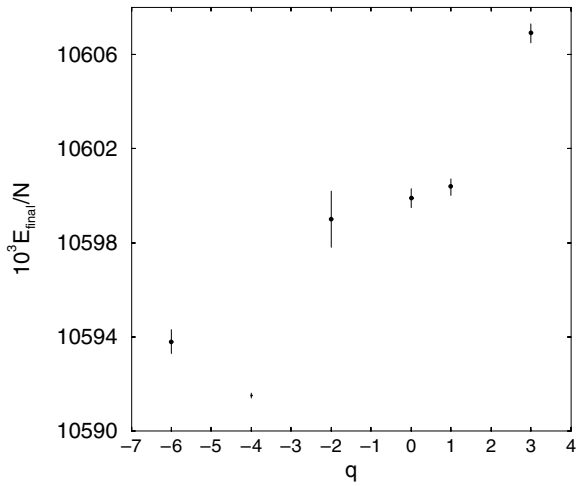


**Figure 4.** Final mean energies per particle $E_{\text{final}}/N$ versus $q$ averaged over 50 independent runs for the same system as in figure 1. The error bars are dispersions in $E_{\text{final}}/N$. $q = 3$ refers to $E_{\text{final}}/N$ found with threshold accepting.

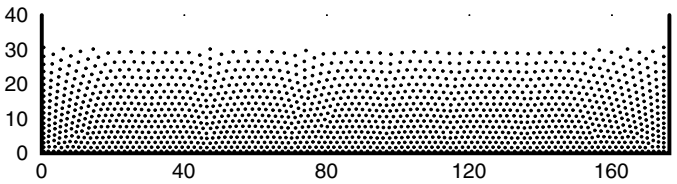


**Figure 5.** Example of the low-energy configuration found for the system with $N = 1500$, $L = 175.2$, $\alpha = 11.25$.

of the Tsallis acceptance probability indicates [9, 19], lowering $q$ is associated with gradual breaking ergodicity, i.e., with limiting the available configuration space to be explored by the random walker. This in turn might prevent it from finding the proper minima.

Apart from the tests themselves, we have also got interested in the low-energy configurations of a larger system with $N = 1500$ particles. The example structure formed by the balls at the end of the annealings is shown in figure 5. Locally, the most characteristic features of the gravity rainbow are preserved. Each particle (apart from those at the bottom, the top and on the walls) has six nearest neighbours at distances from each other depending on the position. This also causes the density distribution of the balls to be nonuniform in the 2D space. However, the global picture of the configuration looks different. There are no clear arcs with horizontal periodicity typical of the real gravity rainbow. Only some of its fragments recall the latter which provides a conformation of the earlier, purely theoretical results [17].

## 4. Remark on the acceptance probabilities

In this chapter, we present some intuitive analysis concerning the properties of the various acceptance probabilities. It should be considered as a way of giving a deeper insight into the matter rather than a rigorous mathematical proof. The faster performance of the Tsallis

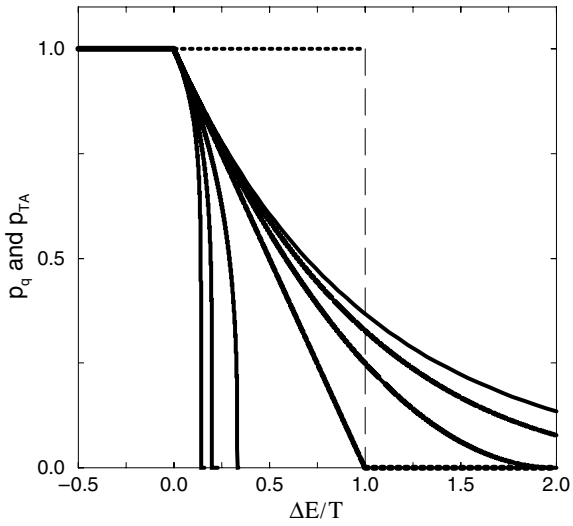


**Figure 6.** Equation (1) versus $\Delta E/T$ for various $q$. From the left to the right: $q = -6, -4, -2, 0, 0.5, 0.8, 1$. Threshold accepting of equation (3) is shown with the dotted line.

generalized annealing over the Metropolis algorithm might be a direct consequence of the property of the acceptance probabilities (figure 6) that for $q < 1$,

$$p_q(\Delta E/T) \leqslant p_1(\Delta E/T). \tag{6}$$

Actually, for $\Delta E \leqslant 0$ both of them are equal to one. For $\Delta E > 0$ the left-hand side of (6) is set to zero whenever $(1-q)\Delta E/T \geqslant 1$. For $(1-q)\Delta E/T < 1$, the relation

$$[1-(1-q)\Delta E/T]^{\frac{1}{1-q}} < \exp(-\Delta E/T) \tag{7}$$

is fulfilled, and (6) also holds. Namely, for $0 < r = 1-(1-q)\Delta E/T < 1$, inequality (7) is equivalent to

$$r - \ln(r) > 1 \tag{8}$$

which is always the case in the considered interval. Moreover, it is easy to note that if $q_1 < q_2 < 1$ then

$$p_{q_1}(\Delta E/T) \leqslant p_{q_2}(\Delta E/T) \tag{9}$$

that in turn means that the smaller the $q$ the more pronounced the inequality (6). In other words, while both acceptance probabilities always accept decreases in the energy, the Tsallis method has the tendency to accept increases in it more seldom (and therefore is faster in searching for low-energy states). The tendency is stronger for large negative $q$. Moreover, it should be stressed that this observation is of complementary character to the earlier results. The property of 'one-way transitions' of the Tsallis algorithm with $q < 1$ has previously been emphasized in [9] in terms of broken ergodicity. From the definitions of the acceptance probabilities it also follows that for $q < 0$

$$p_q(\Delta E/T) \leqslant p_{TA}(\Delta E/T) \tag{10}$$

suggesting that for the given cooling scheme, the Tsallis algorithm might be even faster in minimizing with $T$ than threshold accepting. When $q \ll 0$ the minimization goes very fast but the probability that $(1-q)\Delta E/T \geqslant 1$ becomes larger. In that case, as a consequence of

definition (1), the algorithm turns into the steepest decent method even at relatively higher $T$ increasing the probability of the random walker to get trapped in a local minimum. Therefore, depending on the system under consideration, negative values of $q$ taken to speed up the computations should not be chosen extremely large. On the other hand, the above remark does not hold if the system energy landscape contains a huge (even continuum) number of minima with practically indistinguishable depths. Then, the nonergodicity does not influence the final energy by definition, and extremely fast algorithms might be very useful (for instance to compare some final configurations with very similar energies).

## 5. Summary

In the paper we have applied the Tsallis and the threshold simulated annealings to the model 2D system of many interacting hard spheres. The former has also included the case of $q \to 1$ which yields the traditional Metropolis algorithm. In order to follow the real experiment, the interaction potential has been chosen dipole–dipole like and the gravitational force has also been taken into acount. The simulations have enabled us to inspect the performance of the methods with respect to the speed of minimization with the dropping temperature, the number of MCS and with the CPU time. The three different criteria have been supplemented by considering the mean final energies $E_{\text{final}}$, i.e., the algorithm efficiency.

On one hand, our results are in a qualitative agreement with some of those reported earlier concerning the applications of the Tsallis statistics to various systems [6, 7, 19]. Within the exponential cooling schedule, using the generalized acceptance probability with $q < 1$, it is possible to find acceptable solutions at higher temperatures and in fewer MCS. Moreover, as shown by our CPU time tests, though the computation of the generalized acceptance probability of equation (1) consumes more time than the step function given by equation (3), the former has turned out to be faster and a bit more optimal as an optimization algorithm. This observation, on the other hand, seems a bit contradictory to [9–11], where it is claimed that in order to get the best possible, optimization results with the smallest computational effort, threshold accepting should be preferred to the Tsallis probability. However, this is not the case, and the results presented in this paper are in agreement with the purely theoretical conclusions. As is pointed out, without knowledge of the optimal temperature sequence for threshold accepting (in the case of the optimal control theory), it might be better to carry out the minimization using the Tsallis or Metropolis based acceptance rules within a good annealing schedule. Our simulations indicate that this is actually the case. The exponential temperature sequence we have chosen, though very simple and very widely used in practical applications, is far from being optimal for threshold accepting applied to the considered system. Therefore, due to the lack of ergodicity in the threshold acceptance rule, trapping in local minima is a little more probable. Strictly speaking, the mean final value of the objective function and the speed of minimization strongly depend on the chosen temperature sequence [11]. This also means that the question of testing various annealing methods is rather relative. Better or worse performance of an algorithm should not be treated as absolute but as limited to the chosen cooling scheme.

Finally, we have also analysed the property of the Tsallis acceptance probability that is likely to cause the algorithm to be faster in minimizing with the temperature. For positive energy fluctuations and negative values of $q$, the tendency of the generalized function to reject such moves is better pronounced. In other words, when $T$ and $q$ are lowered, some paths in the configuration space start becoming unidirectional, meaning that the Markov chain gradually gets more and more prevented from crossing energy barriers. This, however, increases the probability of the walker being trapped in a local minimum.

## Acknowledgments

A partial support from the State Committee for Scientific Research via grant no 8 T11F 027 16 is acknowledged. Most of the calculations were carried out in Poznań Supercomputing and Networking Center (PCSS). We thank G Kamieniarz and G Musiał for helpful discussions and valuable comments.

## References

[1] Kirkpatrick S, Gelatt C D Jr and Vecchi M P 1983 *Science* **220** 671
[2] Geman S and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721
[3] Szu H and Hartley R 1987 *Phys. Lett.* A **122** 157
[4] Tsallis C 1988 *J. Stat. Phys.* **52** 479
[5] Stariolo D A and Tsallis C 1995 *Annual Reviews of Computational Physics* vol II ed D Stauffer (Singapore: World Scientific) p 343
[6] Penna T J P 1995 *Phys. Rev.* E **51** R1
[7] Andricioaei I and Straub J E 1996 *Phys. Rev.* E **53** R3055
[8] Hansmann U H E 2002 *Choas, Solitons Fractals* **13** 507
[9] Fachat A, Hoffmann K H and Franz A 2000 *Comput. Phys. Commun.* **132** 232
[10] Franz A, Hoffmann K H and Salamon P 2000 *Phys. Rev. Lett.* **86** 5219
[11] Franz A and Hoffmann K H 2002 *J. Comput. Phys.* **176** 196
[12] Pierański P 1989 *Phase Transitions in Soft Condensed Matter* ed T Rieste and D Sherrington (New York: Plenum)
[13] Rothen F, Pierański P, Rivier N and Joyet A 1993 *Eur. J. Phys.* **14** 227
[14] Skjeltorp A T 1989 *Phase Transitions in Soft Condensed Matter* ed T Rieste and D Sherrington (New York: Plenum)
[15] Kolosov V Yu 1955 *Aperiodic'94* ed G Chapuis (Singapore: World Scientific)
[16] Rothen F and Koch A J 1989 *J. Physique* **50** 663
[17] Wojciechowski K W and Kłos J 1996 *J. Phys. A: Math. Gen.* **29** 3963
[18] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 *J. Chem. Phys.* **21** 1087
[19] Kłos J and Kobe S 2000 *Acta Phys. Pol.* A **97** 943
[20] Kłos J and Kobe S 2001 *Nonextensive Statistical Mechanics and Its Applications (Springer Lecture Notes in Physics* vol 560) ed S Abe and Y Okamoto (Berlin: Springer) p 253