

# Packing a multidisperse system of hard disks in a circular environment

André Müller, Johannes J. Schneider, and Elmar Schömer\*

*Department of Physics, Mathematics, and Computer Science, Johannes Gutenberg University of Mainz, Staudinger Weg 7, 55099 Mainz, Germany*

(Received 11 July 2008; revised manuscript received 27 December 2008; published 2 February 2009)

We consider the problem of finding the densest closed packing of hard disks with proposed different radii in a circular environment, such that the radius of the circumcircle is minimal. With our approach, we are able to find denser packings for various problem instances than known from the literature. Both for the dynamics of the simulation and for the optimum values of the radii of the circumcircles, we find various scaling laws.

DOI: [10.1103/PhysRevE.79.021102](https://doi.org/10.1103/PhysRevE.79.021102)

PACS number(s): 64.60.De, 64.75.Gh, 65.40.Ba, 89.75.Da

## I. INTRODUCTION

The problem of determining a spatial arrangement or even the densest packing of hard disks and spheres is often studied, as these systems are widely used in physics as simple two- and three-dimensional models for granular matter, colloidal systems, fiber-reinforced composites, and molecular crystals [1–5]. Here usually monodisperse and bidisperse systems are considered, i.e., either all disks and spheres exhibit the same radius value or one of two different values. In recent years, the focus of research has shifted to densest packings of items with shapes for which the determination of overlappings between them is slightly more difficult, such as ellipsoids [6] and spherocylinders. It has, e.g., been found that a random packing of ellipsoids with a specific aspect ratio (M&M candies) is denser than a random packing of spheres [7]. Furthermore, densest packings of particles with long-range interactions, in confinement, and under constraints, such as e.g., shear forces and repulsive walls [8], are investigated. Usually, these packing problems are considered in either two or three dimensions. But also the packing of high-dimensional hard spheres with identical radii is of interest as this problem can be mapped on finding efficient binary codes for digital communications [9,10].

For many of these more complicated packing problems, no exact solution is known, such that heuristic optimization algorithms have to be used in order to find at least quasi-optimum configurations for these NP-complete problems. A variety of specifically constructed and general purpose methods has been applied to these problems. In order to compare the quality of these heuristic methods, a set of well-defined benchmark problems has been created, such as e.g., the problem of packing disks with identical radii in a unit square: here a given number of  $N$  identical disks has to be placed in the square in a way that the radius value of these congruent disks is maximized [11]. This problem has been studied very extensively, optimality proofs were provided for small numbers of  $N$  (e.g., for  $N \leq 27$  in [12] and for  $N=36$  in [13]). For a recent overview of this problem, see, e.g., the review [14], and the references therein. Similarly, the problems of packing circles in a circle [15], circles in a triangle, squares in a

circle, and triangles in a circle have been extensively studied. Also for these problems, exact proofs for the globality of the found solutions were partially provided and exact methods for some small numbers of  $N$  were developed (see, e.g., [16]). Most of these benchmark problems consider systems with objects of the same shape and size; sometimes two or three different sizes are considered. However, we want to focus on multidisperse systems for which all object sizes differ from each other.

Multidisperse packing problems exhibit additional fascinating properties when compared to monodisperse packing problems. The effect of multidispersity on the microstructure can be dramatic [17]. When considering, e.g., particles which partially exhibit conducting properties, one finds that they are often prevented from forming a connected network as a result of the relative size and composition of the surrounding nonconducting particles [18]. A further example involves the dissolution of a crystal comprised of multidisperse disks, where the large disks restrict the solubility of the crystal in the solvent [19].

## II. CLASSIC APPROACH TO MULTIDISPERSE PACKING

Recently, a benchmark contest [20] was performed for which the task was given as follows: Consider a system of  $N$  hard disks with radii  $r_i=i$  ( $i=1,2,\dots,N$ ). In the benchmark contest, only system sizes up to  $N=50$  were considered. These disks shall be packed in a circular environment in a way that the radius  $R$  of the circumcircle is minimal. An example for 50 disks is shown in Fig. 1. If we denote the coordinates of the midpoint of disk  $i$  as  $x_i$  and  $y_i$  and assume that the midpoint of the circumcircle is located at the origin,  $R$  is given as

$$R = \max_i(\sqrt{x_i^2 + y_i^2} + r_i). \quad (1)$$

A configuration is therefore completely described by the ordered set of locations  $(\vec{v}_i) = \left[ \begin{pmatrix} x_i \\ y_i \end{pmatrix} \right]$  of the midpoints of the various disks. The disks must not overlap. In order to meet this constraint, a penalty function

$$P(i,j) = [r_i + r_j - d(i,j) + \kappa] \Theta[r_i + r_j - d(i,j)], \quad (2)$$

between each pair  $(i,j)$  of disks is introduced, with

\*{andmuell,schneidj,schoemer}@uni-mainz.de;  
www.staff.uni-mainz.de/{schneidj,schoemer}

URL:

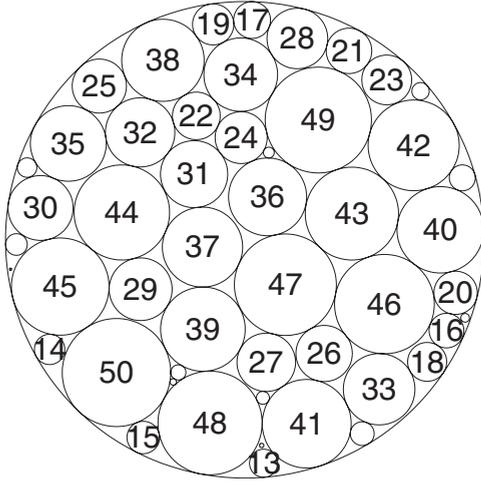


FIG. 1. Best solution found for the multidisperse packing problem with 50 disks: the circumcircle has a radius  $R = 220.600\,418\,7\dots$ . The numbers within the larger disks denote their radius values.

$$d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

being the distance between the midpoints of disks  $i$  and  $j$ ,  $\kappa$  being an offset in order to put more emphasis on small overlappings, and

$$\Theta(a) = \begin{cases} 1 & \text{if } a > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

being the Heaviside function. Thus we get the overall Hamiltonian

$$\mathcal{H} = R + \lambda \mathcal{P} = R + \lambda \frac{1}{2} \sum_{i,j} P(i,j), \quad (5)$$

with the Lagrange multiplier  $\lambda$ .  $\lambda$  has to be chosen in such a way that while  $R$  is minimized, the optimization process ends up with  $P(i,j)=0$  for all pairs of disks.

We will focus on this benchmark problem throughout this paper, as it is a very good example for a multidisperse system of hard disks, as it is a NP (non-deterministic polynomial time)-complete problem for which no exact algorithm solving this problem in polynomial time is known, and as the benchmark contest provided world record results achieved in competition between 155 groups from 32 countries having taken part in the contest, with which we can compare our results.

We apply the classic physical optimization algorithm simulated annealing [21] to this problem. Within this approach, one starts out with a randomly generated configuration and an initial high temperature. Then the system is subjected to a cooling schedule in which the temperature is gradually decreased towards zero, such that the system undergoes a transition from a high-energy unordered regime to a low-energy ordered configuration, in which it finally freezes. At each temperature step, several moves are performed which change the current configuration  $\sigma$  slightly to a tentative new configuration  $\tau$ .

Three different move routines are used: (a) The shift move randomly selects a disk  $i$  and shifts it slightly to a new position nearby its previous place. For this purpose, two small uniformly distributed random numbers  $\xi_x$  and  $\xi_y$  are chosen from the interval  $[-\zeta_{\text{shift}}; +\zeta_{\text{shift}}]$  and added to the coordinates of the midpoint of the  $i$ th disk, such that the tentative new configuration  $\tau$  is given by

$$\tau = \left[ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \dots, \begin{pmatrix} x_i + \xi_x \\ y_i + \xi_y \end{pmatrix}, \dots, \begin{pmatrix} x_N \\ y_N \end{pmatrix} \right]. \quad (6)$$

As this shift move should only allow for very small quiverings of the disks, we choose  $\zeta_{\text{shift}}=1$ , such that the relative alignment of the disks in the configuration usually remains unchanged. (b) The jump move works the same way as the shift move, but it shifts the disk not only locally, but to a large extent, such that one disk can jump nonlocally to an entirely new location. Here we choose  $\zeta_{\text{jump}}=1000$ , such that this jump move lets one disk jump over large distances and can thus lead to a new arrangement of the disks. Here one might argue why it is necessary to work with two different move routines, working the same way, but on different length scales, especially as the jump move includes the shift move if  $\xi_x$  and  $\xi_y$  are chosen very small. However, during our tests, we found that it is necessary on the one hand to put a lot of emphasis on the small length scale in order to let all the disks quiver slightly and on the other hand, to allow also for larger jumps, even if the probability that such a move leads to an improvement if the configuration is already rather dense, might be very small. We will show how the results change with changing the move parameters in Sec. IV. (c) The third move routine, the swap move, randomly selects a pair of disks  $(i,j)$  with neighboring radius values, i.e.,  $r_i = r_j \pm 1$ , and exchanges their locations, such that the tentative new configuration  $\tau$  is given as

$$\tau = \left[ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \dots, \begin{pmatrix} x_{i+1} \\ y_{i+1} \end{pmatrix}, \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \dots, \begin{pmatrix} x_N \\ y_N \end{pmatrix} \right]. \quad (7)$$

These moves have in common that most of the configuration remains unchanged, as only one or two disks are displaced when the move is accepted. Thus, all these moves obey the local search paradigm.

Each of these moves is accepted or rejected according to the Metropolis acceptance criterion [22] with the probability

$$p(\sigma \rightarrow \tau) = \begin{cases} 1 & \text{if } \Delta\mathcal{H} \leq 0 \\ \exp[-\Delta\mathcal{H}/(k_B T)] & \text{otherwise,} \end{cases} \quad (8)$$

with  $\Delta\mathcal{H} = \mathcal{H}(\tau) - \mathcal{H}(\sigma)$  being the energy difference between the current configuration  $\sigma$  and the tentative new configuration  $\tau$ , with the temperature  $T$ , and with the Boltzmann constant  $k_B$ , which is set to 1 in computer simulations, such that the temperature is measured in the same units as the energy. From the point of view of optimization, the temperature is simply a control parameter governing the transition from a quasirandom walk regime, in which nearly all moves are accepted, to a greedy regime, in which each move leading to a deterioration is rejected. In case of acceptance, one sets  $\sigma := \tau$ , otherwise one stays in the current configuration  $\sigma$ . Then one proceeds with the next move altering  $\sigma$ . We also

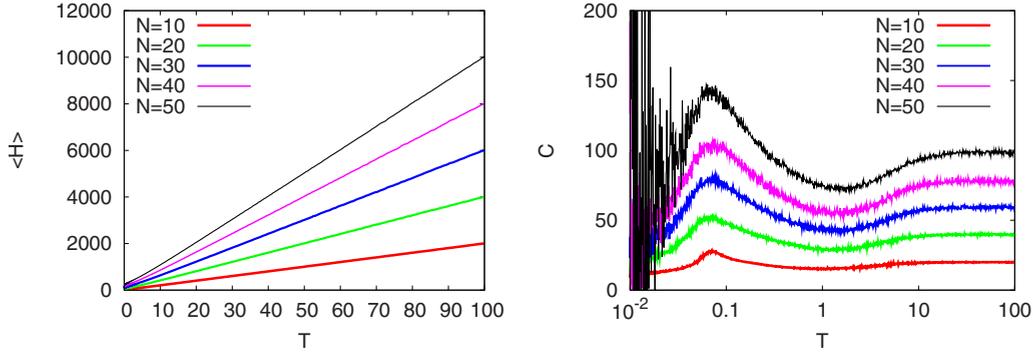


FIG. 2. (Color online) Linear decrease of the mean energy  $\langle \mathcal{H} \rangle$  with decreasing temperature  $T$  at high temperatures (left) and specific heat  $C$  vs temperature  $T$  (right) for various system sizes.

tested more modern physical optimization algorithms such as parallel tempering [23–25] and multicanonical algorithms [26,27], but they yielded worse results than simulated annealing, a result which is in accordance with the extensive investigations for the traveling salesman problem in [28].

### III. DYNAMICS OF THE OPTIMIZATION PROCESS

At first we want to have a look at the dynamics of the optimization process. We start out with the results for a parameter set we found to be optimal. In the next section, we will describe how the results change in dependence on the parameter values.

We decrease the temperature  $T$  via an exponential cooling schedule, i.e.,  $T_{\text{new}} = f \times T_{\text{old}}$  and choose the cooling factor  $f = 0.99$ , the initial temperature  $T_{\text{initial}} = 100$ , and the final temperature  $T_{\text{final}} = 10^{-4}$ . At the end, we add a Greedy step at  $T = 0$ , such that we have overall 1377 temperature steps. In each temperature step, we perform 20 000 measurements, the first 5000 of which are discarded, as the system needs some time to equilibrate at the new temperature value. In order to get independent measurements, we perform 40 sweeps before a new measurement is taken. The shift move and the jump move are called with probability  $\frac{1}{4}$  each; the swap move is called with probability  $\frac{1}{2}$ . For penalizing the overlaps, we use  $\kappa = 0$  and  $\lambda = 1$ . We start out with a random configuration. The various disks are placed randomly within a circle whose midpoint is located at the origin and whose radius is so large that hardly any overlaps occur. Then the cooling process is started.

Please note that we show results only for the system sizes  $N = 10, 20, 30, 40$ , and  $50$ , but we also checked the results for several intermediate values of  $N$ .

We find that the mean energy  $\langle \mathcal{H} \rangle$  decreases linearly with decreasing temperature  $T$  at high temperatures. When fitting the data shown in the left graphic of Fig. 2, we get exactly

$$\langle \mathcal{H} \rangle = 2NT. \quad (9)$$

This result is also found for all other system sizes we investigated and does not change if starting at higher initial temperatures or changing the amount of calculation time per temperature step, at least if this amount does not become too small. We also considered a corresponding multidisperse

packing problem of hard spheres in three dimensions, where we got the result  $\langle \mathcal{H} \rangle = 3NT$  [29]. We can generally write

$$\langle \mathcal{H} \rangle = DNk_B T, \quad (10)$$

with  $D$  being the dimension of the considered problem. This result can also be easily derived analytically in the limit of large values of  $R$  under the assumption that the particles perform their random walks independently of each other, only restricted in their movements by the barrier of the circumcircle, whose radius can be enlarged only with a probability proportional to  $\exp[-1/(k_B T)]$ . Such a linear dependence of the energy of the system on the temperature is also found for other physical systems, in which the particles move independently of each other but are restricted in their movements by a barrier, like the particles of an ideal gas. There, each dimension provides one translational degree of freedom, contributing  $\frac{1}{2}k_B T$  per particle to the overall energy, such that the energy decreases linearly with decreasing temperature. Please note that the different sizes of the disks of our multidisperse system have no influence on the linear decrease of the mean energy at high temperatures. Only the number of particles determines the prefactor.

But then there comes a point at which the various disks become aware of each other, such that this linear decrease cannot continue anymore. Instead, at low temperatures, a transition towards a freezing process takes place, until the system finally gets stuck at the bottom of some local valley in the energy landscape. This freezing transition can be best viewed when considering the specific heat

$$C = \frac{\partial \langle \mathcal{H} \rangle}{\partial T}. \quad (11)$$

For calculating the specific heat, we use the numerically more stable formula

$$C = \frac{\text{Var}(\mathcal{H})}{k_B T^2} = \frac{\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2}{k_B T^2}, \quad (12)$$

which is equivalent to formula (11) in thermal equilibrium. Looking at the right graphic in Fig. 2, we find that the specific heat is virtually constant at high temperatures, resolving the linear decrease of the mean energy  $\langle \mathcal{H} \rangle$  in this temperature range as follows:

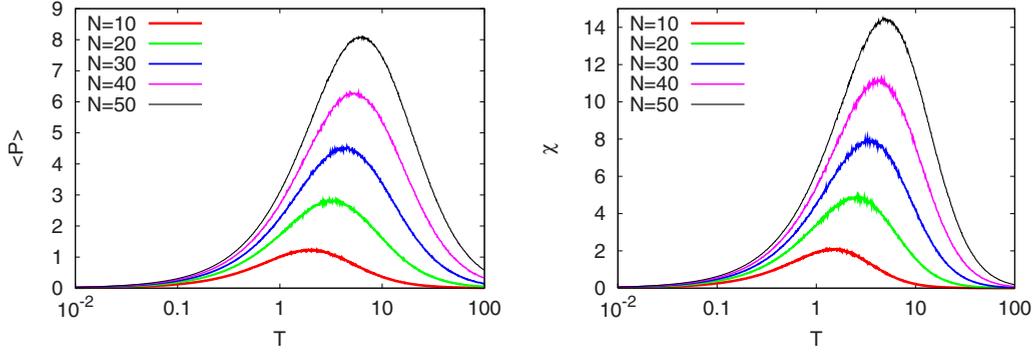


FIG. 3. (Color online) Expectation value  $\langle \mathcal{P} \rangle$  of the sum of the penalties (left) and corresponding susceptibility  $\chi$  vs temperature  $T$  (right) for various system sizes.

$$C = DNk_B. \quad (13)$$

When the temperature decreases beyond  $T \sim N$ ,  $C$  slightly decreases towards an intermediate minimum. This decrease becomes more pronounced when increasing  $N$ . The minimum value is given by

$$C_{\text{min,intermediate}} = (1.45 \pm 0.05)N. \quad (14)$$

The temperature at which this intermediate minimum lies increases slightly with increasing  $N$  from  $T \sim 1$  for  $N=10$  to  $T \sim 1.75$  for  $N=50$ .

Lowering the temperature even further, the specific heat exhibits a peak indicating the freezing transition of the system. The freezing temperature, at which the peak of the specific heat lies, is nearly system size independent, and lies at  $T_f \sim 0.07$  for small system sizes and  $T_f \sim 0.073$  for  $N \geq 30$ . The height of the peak increases linearly with the system size as follows:

$$C(T_f) \sim 2.6N. \quad (15)$$

Only the height of the peak for  $N=50$  is slightly larger.

Afterwards, the specific heat decreases again. Due to rounding errors occurring while calculating the variance, strong oscillations occur at small temperatures, for which the factor  $1/(k_B T^2)$  becomes very large. Partially, the specific heat increases  $\propto 1/T^2$  at very low temperatures.

Next we have a look at the development of the expectation value of the penalty function  $\mathcal{P} = \sum_{i < j} P(i, j)$ , which is shown in the left graphic of Fig. 3. At first sight, this graphic looks highly unusual, as one expects penalty functions to decrease from large values at the beginning to hopefully 0 at the end of the optimization run. But as already mentioned, hardly any overlaps can be formed for high temperatures, as the size of the circumcircle is so large that the various disks can be located rather distant from each other. Thus, only when decreasing  $T$  and thus decreasing  $R$ , the various disks can get closer together and the sizes of the overlaps first increase, until the penalties are sensed by the system and the overlaps decrease again.

The larger the system size, the larger is the sum of the overlaps which can occur in the system, thus the peak is higher for larger values of  $N$ . Furthermore, the peak is lo-

ated at higher temperatures for larger system sizes. If denoting the temperature at which  $\langle \mathcal{P} \rangle$  exhibits its peak as  $T_p$ , we find the power law

$$T_p \sim 0.39N^{0.71} \quad (16)$$

for the temperature at which the maximum of the peak lies and the linear dependence

$$\langle \mathcal{P} \rangle(T_p) \sim 0.17N - 0.55 \quad (17)$$

for the height of the peak, at least for the system sizes we consider.

The Hamiltonian of our packing problem in Eq. (5) is composed of two addends such as the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 - HN\mathcal{M} = - \sum_{\langle i, j \rangle} J_{ij} S_i S_j - H \sum_i S_i \quad (18)$$

of the Ising model with  $J_{ij}$  being the interaction between the Ising spins  $S_i$  and  $S_j = \pm 1$ ,  $H$  being the magnetic field,  $N$  being the number of spins, and  $\mathcal{M}$  being the magnetization of the system. There, the term  $-HN\mathcal{M}$  is called the Zeeman term, which enables the coupling of the spins to the magnetic field. We may interpret this term as a penalty function forcing the various spins to align along the direction of the magnetic field. Thus, we may draw the analogy that  $\mathcal{P}$  corresponds to the magnetization  $\mathcal{M}$  and the Lagrange multiplier  $\lambda$ , which governs the strength of the penalty function, to the magnetic field  $H$ . For the magnetic system, the susceptibility is defined as

$$\chi = \frac{\partial \langle \mathcal{M} \rangle}{\partial H} = \frac{N}{k_B T} \text{Var}(\mathcal{M}). \quad (19)$$

Analogously, we define a susceptibility for our packing problem as

$$\chi = \frac{\partial \langle \mathcal{P} \rangle}{\partial \lambda} = \frac{\text{Var}(\mathcal{P})}{k_B T}. \quad (20)$$

The results for this susceptibility are shown in the right graphic of Fig. 3. We find basically the same behavior as for  $\langle \mathcal{P} \rangle$ . If denoting the temperature at which the susceptibility exhibits its peak as critical temperature  $T_c$ , we get the power law

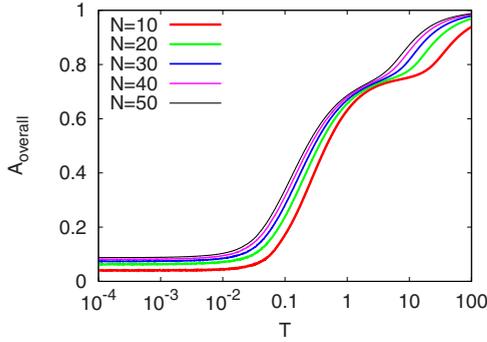


FIG. 4. (Color online) Overall acceptance rate  $A$  vs temperature  $T$  for various system sizes.

$$T_c \sim 0.31N^{0.71} \quad (21)$$

for the dependency of the critical temperature on the system size and the linear law

$$\chi(T_c) \sim 0.31N - 1.2 \quad (22)$$

for the height of the peak. Please note that  $T_c$  is much larger than  $T_f$ . Furthermore, note again that some of these results might change if working with larger system sizes.

Finally, we want to have a look at the velocity of the dynamics. For this purpose, we measure the total acceptance rate of all moves for each temperature, which is shown in Fig. 4. We find that the decrease of the acceptance rate is obviously driven by at least two different influences, such that we have a look at the partial acceptance rates for the various moves (shift move, jump move, and swap move), which are shown in [30]. We find for all partial acceptance rates that they are the larger, the larger the system size is and that they decrease sigmoidally. The acceptance rate for the jump move decreases first, then the acceptance rate of the swap move, and slightly afterwards the acceptance rate of the shift move. Generally, we find that the acceptance rates do not vanish for small temperatures. This can be easily explained: the system is able to perform trivial moves with  $\Delta\mathcal{H}=0$ . Especially the smallest disks allow for such trivial moves, as they can often be shifted or exchanged or even put to a new location by the jump move without generating overlaps. Even at our last finite temperature  $10^{-4}$ , the acceptance rate is still decreasing. At this low temperature, the acceptance rates depend on the system size for  $N \geq 20$  via the power laws

$$A_{\text{shift}}(T=10^{-4}) \sim 0.118N^{0.15} \quad (23)$$

for the shift move,

$$A_{\text{jump}}(T=10^{-4}) \sim 3 \times 10^{-9}N^{2.75} \quad (24)$$

for the jump move, and

$$A_{\text{swap}}(T=10^{-4}) \sim 2.6 \times 10^{-3}N^{0.85} \quad (25)$$

for the swap move. For smaller system sizes, there are deviations from these curves, as there are no small disks in the system anymore which can quiver slightly. Instead, all disks are packed in the densest way and the system is frozen in one configuration.

#### IV. DEPENDENCE OF THE DYNAMICS ON SIMULATION PARAMETERS

As already mentioned in Sec. II, one might argue whether the move set and the simulation parameters we used are optimal. At first sight, one might focus on the question why two move routines for moving one disk to a new location, but working on different length scales are used. At first, we only used one move routine instead of having both a shift move and a jump move and tried to find out what the optimum maximum value  $\zeta$  for moving a disk in an  $x$  and  $y$  direction was. The results for various system sizes are shown in Fig. 5. Here we called the shift/jump move and the swap move with equal probability. We found that a value of  $\zeta \sim 1$  led to the best results within a constant amount of calculation time. We worked with the same cooling schedule as described in Sec. III, but only with 44 000 sweeps per temperature step. The results are averaged over 200 optimization runs each.

However, when having a look at the observables mentioned in Sec. III, we found that the results change significantly while changing  $\zeta$ . Furthermore, the results depend on the calculation time. For the results shown in Fig. 6, the same amount of calculation time was used as for the results shown in Sec. III. The curves for the expectation value of the energy gradually approach the limiting curve for large  $\zeta$  with increasing  $\zeta$  [30]. This curve does not change anymore if increasing  $\zeta$  beyond the value  $\zeta=1000$ . The same holds true if comparing the results for  $\langle \mathcal{P} \rangle$  and for the susceptibility  $\chi$  with those in Fig. 3. The specific heat changes its form completely while increasing  $\zeta$ . For  $\zeta=100$ , one even finds a double peak structure. When watching the intermediate solutions during an optimization run performed with  $\zeta \sim 100$ , one might even be able to detect a slight clustering and ordering effect. Again, here also the results do not change anymore if increasing the value of  $\zeta$  from  $\zeta=1000$  to larger values or using more calculation time. Thus, we find that the value  $\zeta=1000$  is sufficient to get correct results for the thermodynamic observables.

In order to overcome the problem that the best results are achieved with  $\zeta \sim 1$ , but the thermodynamic behavior is only correct for large values of  $\zeta$ , we introduced both a shift move and a jump move, which are called with probability  $\frac{1}{4}$  each. We fixed  $\zeta_{\text{shift}}=1$ . The results for our observables do not change when using  $\zeta_{\text{jump}} \geq 1000$ . As shown in [30], the quality of the achieved optimization results does not depend much on the value of  $\zeta_{\text{jump}}$ . Thus, we used  $\zeta_{\text{jump}}=1000$  from then on.

When looking at locally minimum solutions, we often had the impression that they could be improved by the application of more complex moves. Therefore, we also tried further moves, which change a configuration to a larger extent but do not destroy too much of the previous ordering, such as

- (1) randomly selecting two disks with arbitrary radii and exchanging them;
- (2) exchanging three randomly selected disks with arbitrary radii;
- (3) selecting a circular area and turning all disks whose midpoints are within this area around the midpoint of the area by a randomly chosen angle;
- (4) selecting a circular area and mirroring all disks

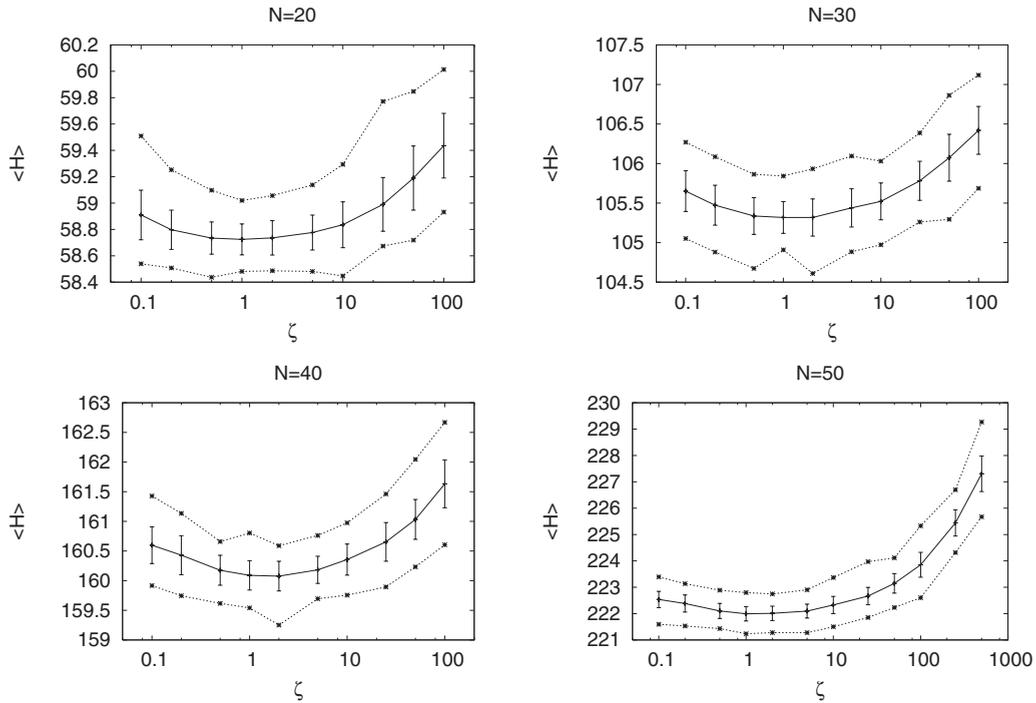


FIG. 5. Minimum, mean (with error bars), and maximum quality of results achieved for various system sizes  $N$  with the maximum amount of  $\zeta$  shifting a disk in the  $x$  and  $y$  direction: the results are averaged over 200 optimization runs each.

whose midpoints are within this area across a randomly selected line going through the midpoint of the area;

(5) selecting a circular area and performing both the turning and mirroring as above;

(6) selecting two circular areas of the same size and exchanging them in a way that the disks whose midpoints are

inside one of these circular areas are moved to the other area in the way that their midpoints are relative placed to the midpoint of the new area as they were formerly placed relatively to the midpoint of the old area, and vice versa;

(7) performing the circular area exchange move as above and combining it with the turning move;

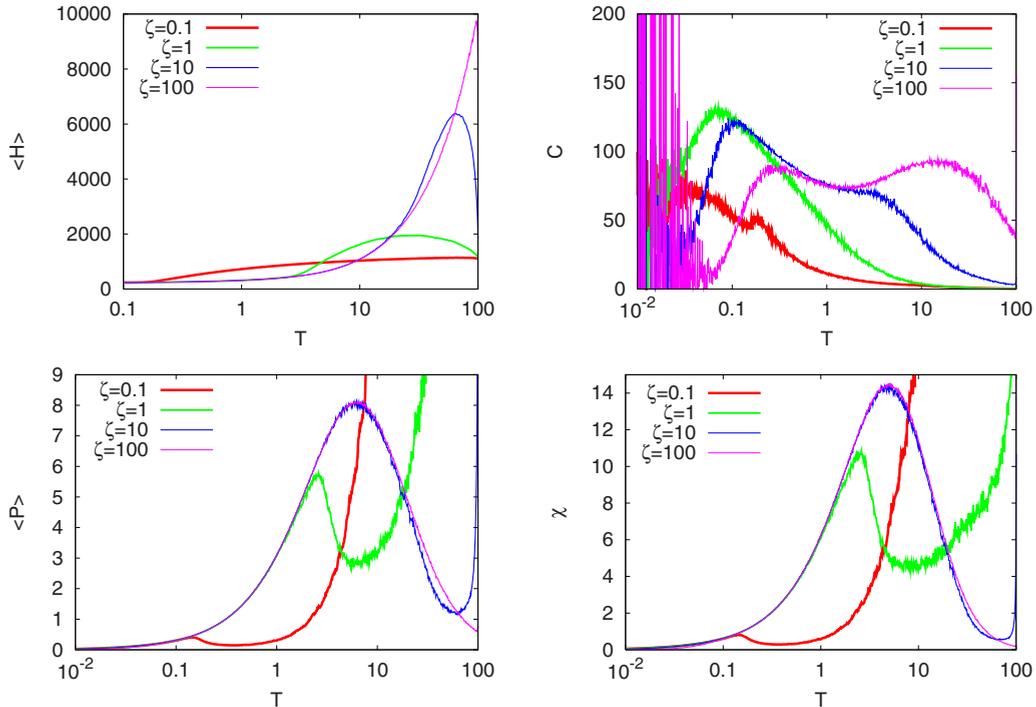


FIG. 6. (Color online) Results for the expectation value  $\langle \mathcal{H} \rangle$  of the energy, the specific heat  $C$ , the expectation value  $\langle \mathcal{P} \rangle$  of the penalty sum, and the susceptibility  $\chi$  for various values of  $\zeta$  and a system size of  $N=50$  disks.

(8) performing the circular area exchange move as above and combining it with the mirroring move;

(9) performing the circular area exchange move as above and combining it with the turning-and-mirroring move

Furthermore, we tried ruin and recreate-type moves with constructive elements [31], which first delete some of the disks (either entirely at random or within a circular area), determine the empty spaces within the configuration, and then iteratively reinsert the deleted disks in these empty spaces in an optimum way. However, within the amounts of the calculation times we used, no further improvements could be achieved with using these more complex moves, such that we prefer to stay with our local search approach here.

## V. RECTIFICATION OF NONFEASIBLE SOLUTIONS

However, using this simulated annealing approach with only a small value for the Lagrange multiplier  $\lambda$  and a vanishing offset  $\kappa$ , one often ends up with solutions, in which very small overlaps with a size up to 0.01 remain. Of course, such solutions are not feasible. In order to rectify these solutions, we use a method called contact simulator as an afterburner after the main routine which works with simulated annealing as described above. This afterburner tries to resolve overlappings by first identifying pairs  $(i, j)$  of overlapping disks and measuring the amount of these overlaps. The penalty function as defined in Eq. (2), again with vanishing offset  $\kappa$ , such that it reads

$$P(i, j) = \begin{cases} r_i + r_j - d(i, j) & \text{if disks } i \text{ and } j \text{ overlap} \\ 0 & \text{otherwise,} \end{cases} \quad (26)$$

can be used for this purpose. The disks are shifted to new positions according to

$$\vec{v}_i^{\text{shifted}} = \vec{v}_i - \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\vec{v}_j - \vec{v}_i}{|\vec{v}_j - \vec{v}_i|} \frac{P(i, j)}{2}, \quad (27)$$

with  $\vec{v}_i$  denoting again the location of disk  $i$ .

As the packings are very dense and as there is often more than one overlap, one will find that after such a shifting some overlappings are not resolved completely and even some new overlappings might occur. Thus, this procedure has to be iterated, until the convergence condition

$$\sum_{i, j} P(i, j) < \varepsilon \quad (28)$$

is met for some small number  $\varepsilon > 0$ . Depending on the original locations of the disks and on the value of  $\varepsilon$ , which we, e.g., choose as  $\varepsilon = 10^{-5}$ , 20–100 iterations are sufficient to reach convergence. By resolving these overlappings, the radius of the circumcircle will increase if a disk touching the circumcircle is shifted to the outside. When this happens, often one also finds some new holes (usually long stretched between some larger disks and looking like concave lenses) in the interior of the configuration, which is a clear indicator

that the solution is worsened by the contact simulator. In order to reduce this effect, we then measure how much the various disks lie outside the previous circumcircle with radius  $R$  and measure the quantity

$$C(i) = \begin{cases} |\vec{v}_i^{\text{shifted}}| + r_i - R & \text{if } |\vec{v}_i^{\text{shifted}}| + r_i > R \\ 0 & \text{otherwise.} \end{cases} \quad (29)$$

Then we determine the new location of a disk as

$$\vec{v}_i^{\text{new}} = \vec{v}_i^{\text{shifted}} \left( 1 - \frac{\alpha C(i)}{|\vec{v}_i^{\text{shifted}}|} \right), \quad (30)$$

with  $\alpha \in ]0; 0.5[$  being the strength of the pressure pushing the disks inside. Then we determine the new value of  $R$  and set  $\vec{v}_i := \vec{v}_i^{\text{new}}$  for the next iteration.

Furthermore, we also try to improve our results in this afterburner by trying moves such as exchanging the locations of disks with neighboring radii and putting the smallest disks close to the circumcircle, as they are less of an obstacle there for getting an excellent solution than they can be in the interior of the configuration. Finally, we end up with improved solutions, which usually exhibit a reduction of the radius of the circumcircle by an amount of  $\sim 10^{-4} - 0.4$  and a reduction of the sum of the overlaps to  $\mathcal{O}(10^{-8})$ , which is the computational limit due to the finite numerical precision.

Now one might argue whether such an afterburner is necessary, as the failure is already created by the small values we use for the parameters  $\lambda$  and  $\kappa$ , namely,  $\lambda = 1$  and  $\kappa = 0$ . Surely, when working in another area of the parameter space  $(\lambda, \kappa)$  with larger values for both  $\lambda$  and  $\kappa$ , one will always end up with feasible solutions. However, we found that we get the best results if performing many runs in parallel just in this area in which a significant percentage of the solutions exhibits some small overlappings. Then we can either choose the best feasible solution or first run the afterburner on the nonfeasible solutions rectifying them and then select the best configuration. We also experienced for other problems, such as the vehicle routing problem [32], that it is advantageous to work in an area of the parameter space in which a significant fraction, but only up to half of the configurations exhibit small violations of the imposed constraints, whereas the majority of the solutions are feasible.

## VI. INVESTIGATION OF QUASIOPTIMUM SOLUTIONS

As simulated annealing is a heuristic algorithm, there is no guarantee that it reaches the global optimum of the proposed optimization problem but it usually ends up at one of the many very good solutions which are qualitatively only slightly worse than the global optimum. A further advantage of simulated annealing in combination with the local search approach, according to which a configuration is only changed slightly and in a random way, lies in the fact that the configuration space is sampled in an unbiased way, such that ending up in one quasioptimum solution is as likely as ending up in another quasioptimum solution. Thus, simulated annealing is also a useful tool for investigating the valley structure of the energy landscape of the optimization problem, with local minimum configurations lying at the bottoms

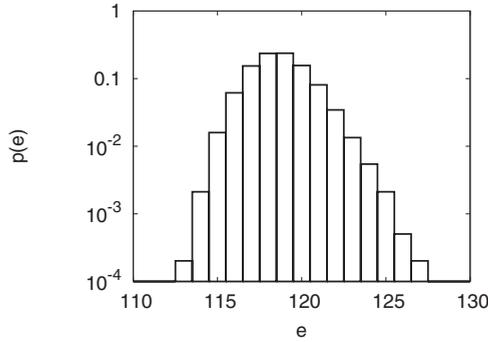


FIG. 7. Probability distribution of the number  $e$  of edges denoting the neighborhood relations between the various disks.

of these valleys. Now we want to investigate these quasioptimum solutions.

As we generally concentrated on the problem instance containing  $N=50$  disks, feeling that if we could produce very good solutions for this largest problem size considered in the benchmark contest [20] then we would have the algorithmic techniques necessary to also find good solutions for smaller system sizes, we produced most solutions for this largest problem instance and thus want to focus on it throughout this section. But we want to stress the point that we achieved analogous results for smaller system sizes, of course based on a smaller set of quasioptimum results for these smaller instances [30]. The circumcircle of the best solution for  $N=50$ , which we were able to find and which is shown in Fig. 1, has a radius value 220.600 418 7.... Additionally to this solution, we found a huge number of quasioptimum solutions being qualitatively only slightly worse than the global optimum, being locally minimal, such that they cannot be improved any further in the  $T=0$  limit of simulated annealing, but differing strongly in the arrangement of the disks. In our investigations, we use the  $Z=9923$  quasioptimum configurations with energy values  $\mathcal{H} < 222$  that we found by using simulated annealing.

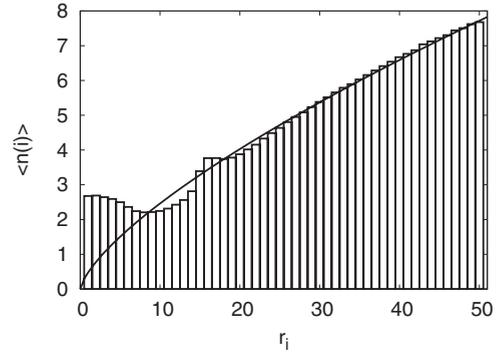


FIG. 8. Average number  $\langle n(i) \rangle$  of neighbors of disk  $i$  with radius  $r_i=i$ : the fit curve for large  $r_i$  is given by the power law  $\langle n(i) \rangle = 0.48r_i^{0.71}$ .

### A. Neighborhood relations between different disks

First of all, we want to investigate some general properties of these quasioptimum configurations and focus on the neighborhood relations between the various disks. We consider two disks  $i$  and  $j$  to be neighbors of each other if their corresponding cells in a Voronoi diagram [33,34] are adjacent to each other. As we deal with disks of different sizes, the boundaries in the Voronoi diagram are not straight lines (as it would be the case if constructing Voronoi diagrams of sets of points or sets of disks with identical radii) but hyperbolic curves, as the radii  $r_i$  of the various disks differ. (Please note that the results given below remain nearly unchanged if using a power diagram with straight lines as boundaries of the cells. The reason for this is that the disks stick so close together that the hyperbolic boundaries between adjacent disks within such an arrangement of disks are only slightly bent.) By connecting the midpoints of disks in adjacent cells, we get a graph with edges, which comprise the Delaunay triangulation. Let  $\eta^\sigma$  be an edge matrix with

$$\eta^\sigma(i,j) = \begin{cases} 1 & \text{if the cells of disks } i \text{ and } j \text{ are adjacent to each other in configuration } \sigma \\ 0 & \text{otherwise.} \end{cases} \quad (31)$$

Then the overall number  $e(\sigma)$  of edges in the configuration is simply given by

$$e(\sigma) = \frac{1}{2} \sum_{i,j} \eta^\sigma(i,j). \quad (32)$$

The minimum number of edges we find in the considered configurations is 113 and the maximum is 127. The distribution of the number of edges, which is shown in Fig. 7, exhibits a sharp peak around its mean value  $\langle e \rangle = 118.72$ .

Second, we have a look at the average number of neighbors of each disk, depending on the disk radius  $r_i$ . The aver-

age number of neighbors of disk  $i$  is given as

$$\langle n(i) \rangle = \frac{1}{Z} \sum_{\sigma} n^\sigma(i) = \frac{1}{Z} \sum_{\sigma} \sum_j \eta^\sigma(i,j). \quad (33)$$

As Fig. 8 shows, the curve for the average number  $\langle n(i) \rangle$  of neighbors to disk  $i$  with radius  $r_i=i$  exhibits a rich behavior: for the smallest disks, this number lies between 2 and 3, as these disks can be placed rather randomly, either in the interior, i.e., in one of the holes between the large disks, such that the number of their neighboring disks is 3, or in the exterior, i.e., close to the circumcircle, which is touched by

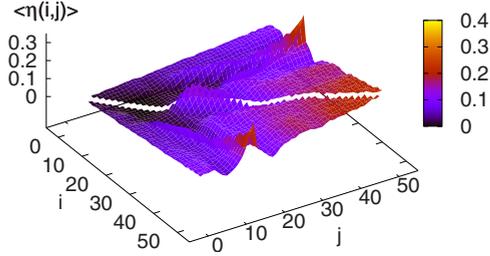


FIG. 9. (Color online) Probability that the disks  $i$  and  $j$  with radii  $r_i=i$  and  $r_j=j$  are neighboring each other.

two larger neighboring disks, such that the number of neighboring disks is 2. The second case, that small disks are put close to the circumcircle, becomes obviously more likely when increasing the disk radius  $r_i$  from 1 to 9. For  $r_i=9$ , we get  $\langle n(9) \rangle = 2.21$ . Afterwards, the number of neighbors increases with an intermediate maximum at  $r_i=17$ . For the large disks, the average number of neighbors increases with the power law

$$\langle n(i) \rangle \propto r_i^\nu, \quad (34)$$

with the exponent  $\nu = 0.71 \pm 0.02$ . Please note that if trying a linear fit, one neglects the bending of the curve and the best linear fit  $1.4 + 0.13r_i$  exhibits a larger fit error than a fit with a power law. Furthermore, note that the exponent 0.71 keeps cropping up in this problem, especially in the power law dependence of the critical temperature  $T_c$  on the system size  $N$ , but also the number of disks adjacent to the circumcircle depends on the system size via the power law  $\propto N^{0.71}$  [30]. Thus, we suspect that this value is in reality  $1/\sqrt{2}$ , though we are not yet able to prove it. For  $r_i=50$ , we get  $\langle n(50) \rangle = 7.68$ .

Here the question arises which disks are adjacent to each other in the various solutions and whether two disks which are neighboring each other in one solution, also tend to neighbor each other in further solutions, i.e., whether there are common structures in the neighborhood relationship between the disks. Such common structures, which are also called backbones and which are based on neighborhood relations, are found in many other NP-complete optimization problems, such as, e.g., for the traveling salesman problem [28,35], in which such a structure is an edge connecting two points in all solutions. Similarly, such common edges turn up in vehicle routing problems [32] and nondegenerate production planning problems [36]. When working with spin glass models, such as the Sherrington-Kirkpatrick (SK) model [37], groups of spins can be identified which are parallel or antiparallel to each other in all solutions [38]. Similarly to the definition of a backbone in [32], we define an averaged edge matrix

$$\langle \eta(i,j) \rangle = \frac{1}{Z} \sum_{\sigma} \eta^{\sigma}(i,j), \quad (35)$$

from which we can derive the backbones from those entries with  $\langle \eta(i,j) \rangle = 1$ . When looking at Fig. 9, we find that some pairs of neighboring disks are much more likely than others. Especially, there is a tendency that large disks are adjacent to

large disks. Furthermore, disks with radii  $r_i \sim 15$  tend to stick with the largest disks. Please note that we find the same behavior that disks with radii of roughly  $\sim 0.3N$  tend to stick with the largest disks also for all other system sizes  $N \geq 20$ . When looking at various quasioptimum configurations, one usually finds that these disks lie in a hole limited by two of the largest disks and the circumcircle. This result can easily be explained analytically: consider the case of two disks with radius  $N$  touching each other and touching the circumcircle with radius  $R$ . The maximum  $r_{\max}$  for the radius of a disk to be placed between the two disks and the circumcircle without generating an overlap is given by

$$r_{\max} = \frac{R}{N+4R} (2R - N - 2\sqrt{R^2 - 2RN}). \quad (36)$$

In the limit  $R \rightarrow \infty$ , we get  $r_{\max} = N/4$  and thus a result close to our simulation result for a similar scenario.

Summarizing, we find some tendencies for pairs of disks liking to stick close together. However, backbones, i.e., structures which are common to all quasioptimum solutions, cannot be identified.

## B. Properties of the metric induced by the Delaunay triangulation

Nevertheless, here we want to proceed further and have a look at the spatial arrangement of the disks not only from the point of view of nearest neighborhood, but also of next nearest neighborhood, and so on. The edges of the Delaunay triangulation induce a neighborhood metric  $\Delta$  with  $\Delta(i,j) = 1$  if the Voronoi cells of disks  $i$  and  $j$  are adjacent to each other. If disk  $j$  is adjacent to disk  $i$  and if disk  $k$  is a nearest neighbor of disk  $j$ , but not of disk  $i$ , then disk  $k$  is a next nearest neighbor of disk  $i$  and we set  $\Delta(i,k) = 2$ . When completely exploring the Delaunay graph, we can assign an integer distance value to each pair of disks  $(i,j)$ , counting the minimum number of edges which have to be used to get from disk  $i$  to disk  $j$  on this graph. Please note that the Delaunay triangulation always provides a fully connected graph, such that we can assign finite distance values to all pairs of disks. We performed this analysis on all our quasioptimum configurations, leading to the results shown in Fig. 10. Here we find that the largest disks have on average the shortest distance to each other. This result becomes even more pronounced when looking at the maximum distances occurring in our quasioptimum solutions. We find that the largest disks exhibit only a maximum distance of 5 among each other, medium sized disks a distance of 6, small disks a distance of 7, and the smallest disks a distance of 8. Obviously, it is advantageous not to have a bunch of small disks comprised in the interior of the configuration. When again having a look at disks with radii  $\sim 15$ , we also find here that they tend to stay close to the largest disks. Please note that the minimum distance  $\Delta_{\min}$  between each pair of disks is 1, as for each pair, at least one quasioptimum configuration exists, in which these disks are adjacent to each other.

The same result for the average distances can be found when we consider the Euclidean distances between the disks, as defined in Eq. (3). When taking the average distance

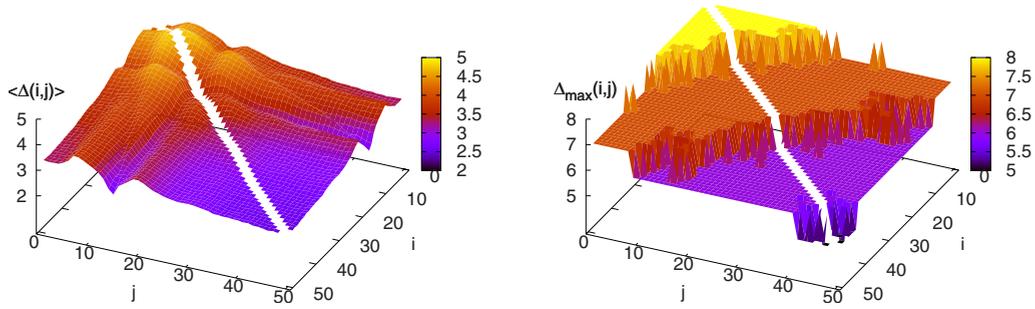


FIG. 10. (Color online) Mean and maximum distance between disks  $i$  and  $j$  with radii  $r_i=i$  and  $r_j=j$  according to the metric  $\Delta$  induced by the neighborhood relations.

$\langle d(i,j) \rangle$  between the midpoints of the disks in our quasioptimum configurations and subtracting the sum of their radii  $r_i$  and  $r_j$ , we get a similar graphic in Fig. 11 compared to the left graphic in Fig. 10, which is based on the  $\Delta$  metric.

### C. Overlaps between different solutions

At this point, we may ask how much different solutions have generally in common. For answering this question, we calculate the overlap between pairs of configurations. The unnormalized overlap  $q_{\sigma\tau}$  between the configurations  $\sigma$  and  $\tau$  is given as

$$q_{\sigma\tau} = \frac{1}{2} \sum_{i,j} \eta^\sigma(i,j) \eta^\tau(i,j). \quad (37)$$

One might want to normalize this overlap to

$$\tilde{q}_{\sigma\tau} = \frac{q_{\sigma\tau}}{\min\{e(\sigma), e(\tau)\}}, \quad (38)$$

but please note that the minimum number of edges varies slightly around its mean value  $\langle \min\{e(\sigma), e(\tau)\} \rangle = 117.79$ . Thus, we prefer to work with unnormalized overlaps, just as in the work of Kirkpatrick and Toulouse on the random-link traveling salesman problem in [39], but keep in mind that we have overlap numbers between 0 and 1 like there and in many other problems (see e.g. [40]). We are also aware of the fact that our overlaps are not algebraic numbers, as they are restricted to the interval  $[0;1]$ , again like in the work of Kirkpatrick and Toulouse, whereas when investigating con-

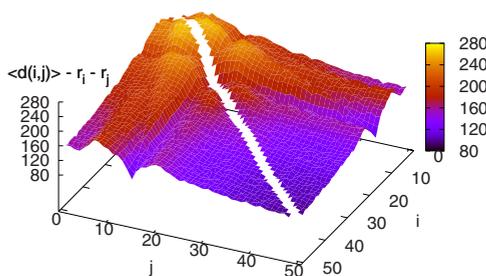


FIG. 11. (Color online) Average Euclidean distance between disks  $i$  and  $j$  with radii  $r_i=i$  and  $r_j=j$ :  $\langle d(i,j) \rangle$  denotes the average distance of the midpoints of the two corresponding disks, from which the sum of their radii is subtracted.

figurations for spin glass models, one gets overlaps in the range  $[-1;1]$ .

Investigating these overlap values, we create a histogram of the overlap values which occur between the  $9923 \times 9922/2 = 49\,228\,003$  pairs  $(\sigma, \tau)$  of configurations with  $\tau \neq \sigma$ . The probability with which the overlap values occur between our quasioptimum configurations is shown in Fig. 12. We find that there is no configuration, which is completely different from any other configuration, as  $p(q=0) = 0$ . The largest overlap value is 125, which is rather close to the maximum number of edges occurring in a configuration, which was 127, as can be seen in Fig. 7. The mean value of the overlaps is  $\langle q \rangle = 15.5754 \pm 5 \times 10^{-4}$ . The maximum of the distribution lies at  $q=15$ , which 5 601 883 pairs of configurations have as an overlap value. The distribution of the overlap values exhibits three peaks, one at  $115 \leq q \leq 125$ , a smaller one at  $83 \leq q \leq 104$ , and the largest one at  $1 \leq q \leq 41$ . Overlap values  $q=85$ ,  $q=88$ , and in the ranges  $42 \leq q \leq 82$  and  $105 \leq q \leq 114$  do not occur. If using not all configurations with  $R < 222$  but selecting only those 378 solutions with  $R < 221.2$ , only the left peak of the distribution remains, but it becomes narrower: the maximum overlap found is 24, the maximum of the peak is formed by 50 pairs of configurations with an overlap of 13.

## VII. OPTIMUM VALUES FOR THE CIRCUMCIRCLE

From the point of view of optimization, the focus lies on getting the best possible result. Here both the radius  $R$  of the

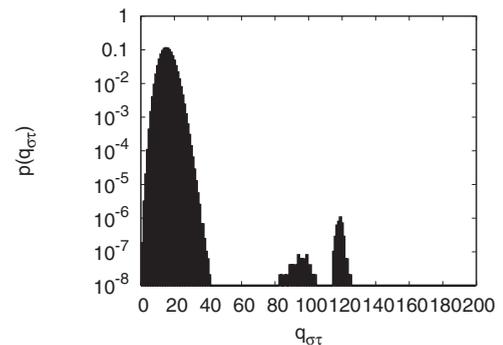


FIG. 12. Probability with which a specific value of our unnormalized overlap values  $q_{\sigma\tau}$  is found for a pair of configurations  $(\sigma, \tau)$  with  $\tau \neq \sigma$ .

circumcircle and the spatial arrangement of the disks in the corresponding configurations are of interest. We were able to match all world records for small system sizes and even to beat all existing world records for the larger system sizes considered in the benchmark contest. In this paper, we show exemplarily the world record configuration for  $N=50$  in Fig. 1. All other world record configurations which were found during the benchmark competition and by our optimization runs can be found online [41]. All these world record configurations were created with simulated annealing or its deterministic variant [42], which is also called threshold accepting [43]. As there is no exact algorithm for this problem, no solution can be proven to be optimal. In a first step, we can only refer to the best solutions found so far by the competition between 155 groups from 32 countries having taken part in the benchmark contest, having applied their various optimization techniques to this problem (for the algorithms used by the winning group, see their paper [44]), and having submitted a total of 27 490 solutions, and compare our results with the contest records listed on the contest web page [20].

Table I shows our results in comparison to the best results found by these groups during the benchmark competition, which are provided by the contest organizers Sam Byrd, Jean-Charles Meyrignac, and Al Zimmermann [20]. When having a closer look at the radii of the circumcircles for the various system sizes  $N$ , one finds that the radius  $R(N)$  increases with the system size via the power law

$$R(N) = sN^{1.5}. \quad (39)$$

The prefactor  $s$  decreases slightly with increasing  $N$  from  $s \sim 0.642$  for  $N=25$  to  $s \sim 0.624$  for  $N=50$ . If we insert this relation between  $R$  and  $N$  in Eq. (36) for the maximum value of the radius of a small disk to be placed between two disks with radius  $N$  and the circumcircle, we get  $r_{\max} \sim 0.307N$  for  $N=50$  and thus reproduce our simulation result that disks with radii  $\sim 0.3N$  tend to stick between two of the largest disks and the circumcircle, as they are best suited for filling the hole there. The ratio  $r_{\max}/N$  decreases only slightly with increasing  $N$ , such that this result also holds for the other larger instances we consider.

When fitting the data points  $(R(N))$  for  $25 \leq N \leq 50$  to a fit curve, the best fit that can be made assumes a slight logarithmic dependence of the prefactor  $s$  on the system size  $N$  as

$$s(N) \sim 0.72 + 2.5 \times 10^{-2} \ln(1/N). \quad (40)$$

Of course, we are aware of the fact that this is only a small effect and that this law cannot hold for much larger  $N$ .

The power law scaling function is to be expected for optimum results, as we can easily derive this power law for lower and upper bounds to the value of the radius of the circumcircle. The size  $R^2\pi$  of the disk, within which the various disks lie, has to be at least as large as the sum of the disk sizes, i.e.,

TABLE I. Comparison of the best results found for  $N \leq 50$  during the benchmark contest, in which 155 groups from 32 countries took part, with our results for  $N \leq 50$ .

$N$	Contest record	Our value	Record
5	9.0013977	9.0013977	matched
6	11.0570404	11.0570404	matched
7	13.4621107	13.4621107	matched
8	16.2217467	16.2217467	matched
9	19.2331939	19.2331939	matched
10	22.0001930	22.0001930	matched
11	24.9606343	24.9606343	matched
12	28.3713894	28.3713894	matched
13	31.5458670	31.5458670	matched
14	35.0956471	35.0956471	matched
15	38.8379955	38.8379955	matched
16	42.4581164	42.4581164	matched
17	46.2913421	46.2913421	matched
18	50.1197626	50.1197626	matched
19	54.2402936	54.2402936	matched
20	58.4005675	58.4005675	matched
21	62.5588771	62.5588771	matched
22	66.7602862	66.7602862	matched
23	71.1994616	71.1994616	matched
24	75.7527041	75.7491426	beaten
25	80.2858644	80.2858644	matched
26	85.0764012	84.9899391	beaten
27	89.7921816	89.7509627	beaten
28	94.5499865	94.5265365	beaten
29	99.5123179	99.4831116	beaten
30	104.5785550	104.5411690	beaten
31	109.7719469	109.6824564	beaten
32	114.8654383	114.8409490	beaten
33	120.2169571	120.0658465	beaten
34	125.4335018	125.3669392	beaten
35	131.1563546	130.9176279	beaten
36	136.5349008	136.4922446	beaten
37	142.1749805	142.0513754	beaten
38	147.8576914	147.4568453	beaten
39	153.5553012	153.3800827	beaten
40	159.4890249	159.1824078	beaten
41	165.2919097	165.0369006	beaten
42	170.9257616	170.8953274	beaten
43	177.0743401	177.0513747	beaten
44	183.1760616	183.0992248	beaten
45	189.6354391	189.2029320	beaten
46	195.9107634	195.5264351	beaten
47	202.1856117	201.7279256	beaten
48	208.6359467	208.0901593	beaten
49	214.6619520	214.2954475	beaten
50	221.0897526	220.6004187	beaten

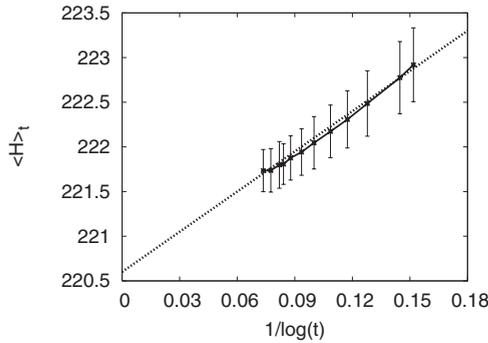


FIG. 13. Average quality  $\langle \mathcal{H} \rangle_t$  of solutions for the problem instance with  $N=50$  disks achieved after calculation time  $t$  (measured in sweeps per temperature step) vs the inverse of the logarithm of the calculation time: the data shown were produced with the same simulation parameters as in Sec. III, but with 726, 1000, 2500, 5000, 10 000, 22 000, 44 000, 88 000, 145 244, 200 000, 400 000, and 800 000 sweeps per temperature step, respectively. For each data point, we averaged the results over 200 optimization runs. A fit through all mean points gives the fit function  $220.6 + 15/\ln(t)$ .

$$R(N)^2 \pi \geq \sum_{i=1}^N r_i^2 \pi = \pi \sum_{i=1}^N i^2 = \frac{\pi}{6} N(N+1)(2N+1) > \frac{\pi}{3} N^3, \quad (41)$$

such that we have the lower bound

$$R(N) > N^{1.5}/\sqrt{3} \sim 0.577N^{1.5}. \quad (42)$$

A simple upper bound can be derived by considering only square numbers  $N=L^2$ . For these numbers, we can imagine all disks being located on the grid points of a regular square lattice with lattice constant  $2N$ . The linear dimension of this lattice is  $L \times 2N$ . Embedding this square lattice in a circle with radius  $\sqrt{2}LN$ , we get

$$R(N) < \sqrt{2} \times N^{1.5} \sim 1.414N^{1.5}. \quad (43)$$

As the series of numbers  $R(N)$  increases monotonically with  $N$ , this power law for the upper bound for the subseries of square numbers holds true for the overall series.

As already mentioned, we cannot be sure whether we were able to find globally optimum configurations for the various values of  $N$ . However, when applying simulated annealing to NP-complete spin glass problems, Grest and his co-workers found the Grest hypothesis [45] according to which the average quality  $\langle \mathcal{H} \rangle_t$  of solutions obtained after some calculation time  $t$  deviates from the true global optimum  $\mathcal{H}_0$  by

$$\langle \mathcal{H} \rangle_t - \mathcal{H}_0 \propto \frac{1}{\ln(t)^\mu}, \quad (44)$$

with the exponent  $\mu \sim 1$ . It has been found that this hypothesis is also valid for other NP-complete problems, such as the traveling salesman problem [28,46].

The Grest hypothesis also seems to be valid for this packing problem as Fig. 13 shows exemplarily for the system size  $N=50$ . We stayed with the same simulation parameters as in

Sec. III but varied the number of sweeps per temperature step, thus varying linearly the overall calculation time. Although we averaged over 200 optimization runs each, the error bars are still so large that we are unable to give a good estimate for the optimum. Thus, we can only state that we might already have reached the global optimum but we can only be sure that our best solution is close to it.

Further evidence whether we might already have reached the global optimum of a problem instance is provided on the contest web page [20] in a detailed overview of all results of the 155 groups having taken part in the competition. As a significant number of groups were able to find independently the same best known solutions for all system sizes up to  $N=23$ , we think that these small instances might already be solved optimally.

Finally, we have a look at the density  $\varrho$  of a packing, which is defined as

$$\varrho(N) = \frac{\sum_{i=1}^N r_i^2 \pi}{R^2(N) \pi}. \quad (45)$$

The density values for the various system sizes are provided in [41]. For larger system sizes, we get a monotonous increase of the density with increasing  $N$  from  $\varrho \sim 0.841$  for  $N=20$  to  $\varrho \sim 0.882$  for  $N=50$ . Please note that one finds for bidisperse systems with two types of disks with a radius ratio of less than 5 for large to small, that the density of a binary disk mixture remains constant at a packing fraction  $\varrho \sim 0.84$  [47]. However, in our problem, when increasing  $N$ , the ratio between the radii of the largest disk and the smallest disk also increases like  $N$ . This increase might be reflected in the increase of the density.

## VIII. CONCLUSION AND OUTLOOK

In this paper, we considered a multidisperse system of  $N$  hard disks with different integer radii  $r_i$  ( $1 \leq r_i \leq N$ ), for which the densest packing in a circular environment had to be found. While solving this problem with simulated annealing, we found that the cooling process exhibits some interesting properties, e.g., the critical temperature  $T_c$  scales with the system size  $N$  via the power law  $T_c \propto N^{0.71}$ . With this physical optimization approach, we were able to match all world records for the small system sizes  $5 \leq N \leq 23$  and  $N=25$  and even to beat all world record results for  $N=24$  and all larger system sizes considered in a benchmark competition which had been held recently and in which 155 groups from 32 countries took part. Besides achieving many world records, we were also able to achieve a huge number of quasioptimum solutions for this problem, which are locally minimum, and were thus able to compare these solutions for common properties and to explore the local valleys of the energy landscape. We found that the average number of neighbors the larger disks have scale with  $r_i^{0.71}$  with  $r_i$  being the radius of the disk. Furthermore, we found that disks with radii  $\sim 0.3N$  like to stay close to the largest disks. The smallest disks can be placed rather randomly. But with increasing disk radius, the probability that such a small disk is placed close to the circumcircle increases. Furthermore, we found

that the optimum radius value  $R$  of the circumsphere scales with the system size via the power law  $R \propto N^{1.5}$ , which we could also prove analytically by providing lower and upper bounds. We also found that the subspace of the quasioptimum configurations exhibit the property of ultrametricity [48].

We will continue our investigation of this problem by attempting to introduce a move set based on changes of the Delaunay graph. We already found out that the definition of the neighborhood generated by such moves for such a placement problem is not so straightforward as it is for sequencing problems, like the traveling salesman problem [49], as the deletion of an edge and insertion of another edge can result in further changes of the Delaunay graph, as the disk might be a neighbor of further disks at its location now. A further problem occurring in this approach is that the Delaunay graph does not contain the full information about the locations of the disks such that some arbitrariness remains if rebuilding a disk configuration from its Delaunay triangulation.

Furthermore, we will extend our investigations to systems of hard spheres in three and even higher dimensions. For three dimensions, in which we have already started to investigate the problem of packing a multidisperse system of hard spheres with successive integer radius values in a spherical environment, we got the result that the optimum radius of the circumsphere scales with the system size as

$$R(N) \propto N^{4/3}. \quad (46)$$

We could again verify this result analytically with lower and upper bounds [29]. Then we will move on to other particle shapes, such as ellipsoids and spherocylinders.

#### ACKNOWLEDGMENTS

This packing project has been supported by the Forschungsfonds of the Johannes Gutenberg University of Mainz, the Collaborative Research Centre Transregio 6 (SFBTR6, subproject A5) of the German Research Foundation, and the newly founded Center for Computational Research Methods in Natural Sciences at the Johannes Gutenberg University of Mainz. Furthermore, we would like to thank the John von Neumann Institute for Computing (NIC) at the Research Center Jülich, Germany, for a generous amount of computing time on the IBM Regatta under Project No. HMZ15 as well as the Center for Data Processing (ZDV) in Mainz. A.M. is very grateful to its administrator, Markus Tacke, for the introduction and his help with the computing facilities at the ZDV. J.J.S. gratefully acknowledges fruitful discussions with Uwe Krey, Ingo Morgenstern (both at University of Regensburg, Germany), and Scott Kirkpatrick (The Hebrew University of Jerusalem, Israel).

- 
- [1] H. Reiss, H. M. Ellerby, and J. A. Manzanares, *J. Phys. Chem.* **100**, 5970 (1996).
  - [2] W. B. Russel, D. A. Saville, and W. R. Schowalter, *Colloidal Dispersions* (Cambridge University Press, Cambridge, England, 1989).
  - [3] S. Torquato and F. Lado, *Proc. R. Soc. London, Ser. A* **417**, 59 (1988).
  - [4] G. Metcalfe, T. Shinbrot, J. J. McCarthy, and J. M. Ottino, *Nature (London)* **374**, 39 (1995).
  - [5] R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983).
  - [6] P. Pfliegerer and T. Schilling, *Phys. Rev. E* **75**, 020402(R) (2007).
  - [7] A. Donev, I. Cisse, D. Sachs, E. A. Variano, F. H. Stillinger, R. Connelly, S. Torquato, and P. M. Chaikin, *Science* **303**, 990 (2004).
  - [8] A. Ricci, P. Nielaba, S. Sengupta, and K. Binder, *Phys. Rev. E* **75**, 011405 (2007).
  - [9] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 2nd ed. (Springer, New York, 1993).
  - [10] N. J. A. Sloane, *Sci. Am.* **250**, 116 (1984).
  - [11] L. Moser, *Can. Math. Bull.* **3**, 78 (1960).
  - [12] R. Peikert, D. Würtz, M. Monagan, and C. de Groot, in *Packing Circles in a Square: A Review and new results*, edited by P. Kall, *System Modelling and Optimization*, Lecture Notes in Control and Information Sciences Vol. 180 (Springer, Berlin, 1992), pp. 45–54.
  - [13] K. Kirchner and C. Wengerodt, *Beitr. Algebra Geom.* **25**, 147 (1987).
  - [14] P. G. Szabó, M. C. Markót, and T. Csentes, in *Global Optimization in Geometry—Circle Packing Into the Square*, edited by C. Audet, P. Hansen, and G. Savard, *Essays and Surveys in Global Optimization* (Kluwer, Dordrecht, 2005), pp. 233–266.
  - [15] R. L. Graham, B. D. Lubachevsky, K. J. Nurmela, and P. R. J. Östergård, *Discrete Math.* **181**, 139 (1998).
  - [16] F. Fodor, *Geom. Dedic.* **74**, 139 (1999).
  - [17] O. U. Uche, F. H. Stillinger, and S. Torquato, *Physica A* **342**, 428 (2004).
  - [18] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer, New York, 2002).
  - [19] B. Mahan and R. Myers, *University Chemistry* (The Benjamin/Cummings Publishing Co., Inc, Menlo Park, CA, 1987).
  - [20] [www.recmath.org/contest/CirclePacking](http://www.recmath.org/contest/CirclePacking)
  - [21] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, *Science* **220**, 671 (1983).
  - [22] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
  - [23] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
  - [24] K. Hukushima, H. Takayama, and H. Yoshino, *J. Phys. Soc. Jpn.* **67**, 12 (1998).
  - [25] B. Coluzzi and G. Parisi, *J. Phys. A* **31**, 4349 (1998).
  - [26] B. A. Berg and T. Neuhaus, *Phys. Lett. B* **267**, 249 (1991).
  - [27] B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
  - [28] J. J. Schneider and S. Kirkpatrick, *Stochastic Optimization* (Springer, Berlin, 2006).
  - [29] A. Müller, J. J. Schneider, and E. Schömer (unpublished).
  - [30] A. Müller, diploma thesis, Johannes Gutenberg University of

- Mainz, Germany, 2008 (unpublished).
- [31] G. Schrimpf, J. Schneider, H. Stamm-Wilbrandt, and G. Dueck, *J. Comput. Phys.* **159**, 139 (2000).
- [32] J. Schneider, FGCS, *Future Gener. Comput. Syst.* **19**, 121 (2003).
- [33] J. O'Rourke, *Computational Geometry in C*, 2nd ed. (Cambridge University Press, Cambridge, 1998).
- [34] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed. (Springer, Berlin, 2008).
- [35] J. Schneider, Ch. Froschhammer, I. Morgenstern, Th. Husslein, and J. M. Singer, *Comput. Phys. Commun.* **96**, 173 (1996).
- [36] J. Schneider, J. Britze, A. Ebersbach, I. Morgenstern, and M. Puchta, *Int. J. Mod. Phys. C* **11**, 949 (2000).
- [37] D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [38] J. J. Schneider, in *Searching for Backbones—An Efficient Parallel Algorithm for Finding Groundstates in Spin Glass Models*, edited by M. Tokuyama and I. Oppenheim, 3rd International Symposium on Slow Dynamics in Complex Systems, Sendai, Japan, 3–8 November 2003, AIP Conf. Proc. No. 708 (AIP, New York, 2004), p. 426–429.
- [39] J. P. Bouchaud and P. Le Doussal, *Europhys. Lett.* **1**, 91 (1986).
- [40] F. Murtagh, *Eur. Phys. J. B* **43**, 573 (2005).
- [41] [www.staff.uni-mainz.de/schneidj/pack2Drecords](http://www.staff.uni-mainz.de/schneidj/pack2Drecords); see also EPAPS Document No. E-PLLEE8-79-131901 for the world record configurations and their densities. For more information on EPAPS, See <http://www.org/pubservs/epaps.html>
- [42] P. Moscato and J. F. Fontanari, *Phys. Lett. A* **146**, 204 (1990).
- [43] G. Dueck and T. Scheuer, *J. Comput. Phys.* **90**, 161 (1990).
- [44] B. Addis, M. Locatelli, and F. Schoen, *Oper. Res. Lett.* **36**, 37 (2008).
- [45] G. S. Grest, C. M. Soukoulis, and K. Levin, *Phys. Rev. Lett.* **56**, 1148 (1986).
- [46] J. Schneider, I. Morgenstern, and J. M. Singer, *Phys. Rev. E* **58**, 5085 (1998).
- [47] D. Bideau, A. Gervois, L. Oger, and J. P. Troadec, *J. Phys. (Paris)* **47**, 1697 (1986).
- [48] J. J. Schneider, A. Müller, and E. Schömer, *Phys. Rev. E* (to be published).
- [49] G. Stattenberger, M. Dankesreiter, F. Baumgartner, and J. J. Schneider, *J. Stat. Phys.* **129**, 623 (2007).