# Disk Packing in a Square: A New Global Optimization Approach

B. Addis
Dip. di Sistemi e Informatica, Università di Firenze, via di Santa Marta 3,I-50139 Firenze, Italy
b.addis@ing.unifi.it
M. Locatelli
Dip.di Informatica, Università di Torino, corso svizzera 18, I-10149 Torino, Italy locatell@di.unito.it
F. Schoen
Dip. di Sistemi e Informatica, Università di Firenze, Via di Santa Marta 3,I-50139 Firenze, Italy
schoen@ing.unifi.it

We present a new computational approach to the problem of placing $n$ identical non overlapping disks in the unit square in such a way that their radius is maximized. The problem has been studied in a large number of papers, both from a theoretical and from a computational point of view. In this paper we conjecture that the problem possesses a so-called funneling landscape, a feature which is commonly found in molecular conformation problems. Based upon this conjecture we develop a stochastic search algorithm which displays excellent numerical performance. Thanks to this algorithm we could improve over previously known putative optima in the range $n \leq 130$ in as many as 32 instances, the smallest of which is $n = 53$. First experiments in the related problem of packing equal spheres in the unit cube led us to an improvement for $n = 28$ spheres.

*Key words:* disk packing; circle packing; continuous dispersion problem; sphere packing; basin hopping; global optimization; stochastic methods
*History:*

---

## 1. Introduction

Packing $n$ equal, non overlapping, disks inside a two-dimensional figure $D$ is a well known and widely studied geometrical problem. Usual choices for the container set $D$ are regular convex figures like the unit square, an equilateral triangle or a circle. The objective of the problem is to maximize the common radius of $n$ nonoverlapping disks contained in $D$. If we denote by $\mathbf{z}^i = (x_i, y_i) \in \mathbb{R}^2$ the center of disk $i$ and by $C(\mathbf{z}^i, r)$ the disk with center $\mathbf{z}^i$ and

radius $r$, then the problem can be formalized as follows

$$\begin{aligned} r_n = \max\ & r & & \\ C(\mathbf{z}^i, r) &\subseteq D & & i = 1, \ldots, n \\ C^{int}(\mathbf{z}^i, r) \cap C^{int}(\mathbf{z}^j, r) &= \emptyset & & \forall\ i \neq j \end{aligned} \tag{1}$$

where $C^{int}$ denotes the interior of a disk. The problem can be viewed as a cutting problem where we are given a two-dimensional material with shape $D$ and we are asked to produce $n$ equal circular objects out of it in such a way that the wasted part is minimized. Notice that if the radius of the circular objects is fixed, an equivalent problem is that of choosing $D$ with the prescribed shape but "as small as possible" in order to be able to produce the $n$ circular objects (e.g. when $D$ is a square this amounts to minimizing its side length).

Although the approach we propose in the following sections could be employed with any convex figure $D$, our attention will be concentrated on the case of packing disks in the unit square, which is the most widely studied in the literature. It is useful to recall at this point that this packing problem is strictly related to the problem of scattering $n$ points in the unit square in such a way that their minimal distance is maximized. The latter problem can be formalized as the following maximin problem

$$\begin{aligned} d_n = \max \min_{i \neq j}\ & \|\mathbf{w}^i - \mathbf{w}^j\|_2 & & \\ \mathbf{w}^i &\in [0,1]^2 & & i = 1, \ldots, n \end{aligned} \tag{2}$$

where $\|\cdot\|_2$ denotes the Euclidean distance. It is well known (see e.g. [Croft et al.(1991)]) and easily proven that the following relation between the optimal values of problems (1) and (2) holds

$$r_n = \frac{d_n}{2(d_n + 1)}$$

and easy relations also allow to derive an optimal solution for any of the two problems from an optimal solution of the other one.

Two are the main directions which research in this field has followed: the first deals with finding packings with proven optimality. This way proofs have been found for optimal packings of up to 9 disks [Schaer (1965), Schaer et al. (1965), Schwartz (1970)] and for some isolated larger $n$ values ($n = 14$ [Wengerodt (1987)], $n = 16$ [Wengerodt (1983)], $n = 25$ [Wengerodt (1987)], $n = 36$ [Wengerodt et al. (1987)]). A different research direction within this stream has aimed towards finding computer-aided proofs. Computational

proofs of optimality have been found for $n$ up to 20 [de Groot et al. (1991)], $21 \leq n \leq 27$ [Nurmela et al. (1999)], $28 \leq n \leq 30$ [Markot et al. (2005)]. In [Locatelli et al. (2002)] optimality within precision $10^{-5}$ has been proven for $n$ up to 35 and for $n = 38, 39$. Computer-aided optimality proofs turn out to be quite computationally demanding. It is interesting to observe that these proofs are usually based on subdivisions of the unit square into nonoverlapping subrectangles each of which is guaranteed to contain at most one point of an optimal solution, and on the subsequent analysis of all the possible combinations of $n$ such subrectangles. As a consequence, the computational burden does not increase regularly with $n$ but has a sudden increase each time there is a need to increase the number of subrectangles (and then also the number of possible combinations) in order to guarantee that each of them contains at most one point of an optimal solution.

The difficulty of proving optimality led to the development of heuristic approaches aiming at improving best known results without giving optimality proofs for them. This represents the second main branch of research. For small $n$ values a remarkable case is $n = 10$. Starting from 1970 and over two decades many improvements have been proposed until in 1990 a computer-aided optimality proof has been given for this case [de Groot et al. (1991)]. More recently different researchers have found new best known solutions for larger $n$ values and all currently best known solutions for $n$ up to 300 are reported in the Packomania web site [Packomania] maintained and continuously updated by E. Specht. In what follows we briefly recall some approaches proposed for this problem in the last decade.

In [Nurmela et al. (1997)] the authors define an "energy" function, strictly related to problem (2)

$$E = \sum_{i \neq j} \left( \frac{\lambda}{d_{ij}^2} \right)^m \tag{3}$$

where $d_{ij}$ is the distance between points $i$ and $j$, $\lambda$ is a scaling factor, and $m$ is a positive integer. They adopt a multistart approach starting a local optimization from at least 50 randomly generated solutions. Noting that when (3) is minimized, the corresponding solutions converge to those of (2) as $m \to +\infty$, for each initial point the authors first perform a local search with a small $m$ value and once they have reached a local minimum they increase the value of $m$, repeating this scheme until $m$ becomes very large. Energy (3) should be minimized over the box $[0, 1]^{2n}$ but the authors transform the problem into an unconstrained one by an appropriate change of variable. The authors also recognize some regular patterns

of disks which are optimal or presumably optimal for small $n$ values but become nonoptimal for $n$ large enough. The best known among such patterns is the square lattice packing of $n = k^2$ points which is optimal for $k$ up to 6 but is not for $k = 7$.

In [Graham et al. (1996)] the authors consider the patterns proposed in [Nurmela et al. (1997)] and extend them with new ones. They employ a billiards simulation method which also allows them to identify threshold indices above which it is guaranteed that the identified regular patterns become nonoptimal.

In [Boll et al.(2000)] a two-phase approach is proposed. The first phase is an approximation one. During this phase each point in turn is moved along appropriately chosen directions with a step-size which is exponentially decreased during the run. The second phase is a refining one where the result of the first phase is the starting point for the billiards simulation method.

In [Casado et al.(1998)] initially the unit square is subdivided into $k \times k$ subsquares, where $k = \lceil \sqrt{n} \rceil$, and the initial solution is obtained by placing the $n$ points at the center of $n$ randomly selected distinct subsquares. Then, each point is randomly perturbed and the perturbed point may be accepted even when it is nonimproving (i.e., backtracking is allowed during the search). Starting from the results in [Casado et al.(1998)], in [Szabò (2000)] the author discusses some new regular patterns of points.

We also mention [van Dam et al.], where a different but strictly related problem is considered. Some solutions of the packing problem have a strong *shadow* effect, i.e. if we project them over the $x$ or the $y$-axis, then the number of distinct projected points is much lower than the overall number of points. Therefore, the authors consider solutions with the additional constraint that these are also well separated when projected over the $x$- and $y$-axes. They propose both an exact branch-and-bound approach which returns optimal solutions to this modified problem for $n$ up to 70, and a heuristic approach for $n$ up to 1000.

This paper belongs to the second branch of research, i.e. we propose a new heuristic method for the problem of packing disks in the unit square. In Section 2 we will introduce the concept of funneling landscape which comes from molecular conformation problems but seems to be particularly well suited also for disk packing. In Section 3 we will give the details of our approach. In Section 4 we will present some quite encouraging computational results. Finally, in Section 5 we will propose some possible directions for future research.

## 2. Funneling landscapes

Molecular conformation problems are quite closely related to disk packing. In these problems we are given a set of $n$ atoms or particles and the model of an energy function which, in its simplest versions, only depends on pairwise distances between atoms. Differently from disk packing, atoms lie in the three-dimensional space. Moreover, the optimization problem is unconstrained, i.e. atoms do not have to be packed into a container with a prescribed shape. However, when some popular energy models, like the Lennard-Jones or the Morse ones, are employed, atoms tend to gather together giving rise to regular geometrical shapes (mostly icosahedra, decahedra, face centered cubic). We might also say that atoms tend to be packed but the shape of the container is not a rigid one. The relationship between the two problems is also remarked by the energy function (3), whose definition resembles that of popular energy models and which is employed in [Nurmela et al. (1997)] to solve the packing problem (see the previous section).

The minimization of energy models turns out to be an extremely difficult task. However significant simplification derived when the funneling landscape of the energy was taken into account. A possible definition of the concept of funneling, based on "disconnectivity graphs", can be found e.g. in [Wales et al. (1997)]. Here we give an alternative definition based on neighborhoods of local minima (see also [Locatelli (2005)]). Let $\mathcal{N}$ be a neighborhood structure defined upon the set $\mathcal{X}$ of all local minima of a given objective function $f$. Then, a funnel can be defined as a maximal subset $\mathcal{Y} \subseteq \mathcal{X}$ of local minima with the following property: there exists a local minimum $\bar{X} \in \mathcal{Y}$ such that for all $X \in \mathcal{Y}$ a decreasing sequence of neighbor local minima in $\mathcal{Y}$ starting at $X$ and ending at $\bar{X}$ exists, i.e.

$$\begin{aligned} \exists\, X_0, X_1, \ldots, X_t : X_i \in \mathcal{N}(X_{i-1}) \cap \mathcal{Y} \qquad & i = 1, \ldots, t \\ f(X_i) < f(X_{i-1}) \qquad & X_0 = X, X_t = \bar{X}. \end{aligned}$$

The common final endpoint of the sequences is called *funnel bottom*. There are at least two issues that deserve special attention in the above definition. First of all, in this definition *only local minima* of the objective function are considered. It is now common practice in molecular conformation problems to restrict the search to local minima, as local searches partially eliminate the strong roughness of energy models. The second important issue is that with the above definition we can focus our attention on funnel bottoms. Indeed, these always include the global minimum of our problem. It has been observed in molecular

conformation problems that, in spite of the huge number of local minima of these problems (conjectured to grow exponentially with the number $n$ of atoms), with an appropriately chosen neighborhood structure $\mathcal{N}$ (neither too "small", nor too "large") the number of funnel bottoms becomes very small – in the easiest cases there is a single funnel bottom coinciding with the global minimum. Therefore, any efficient algorithm to detect funnel bottoms is a good candidate to solve molecular conformation problems. One such algorithm is Basin Hopping (see [Wales et al. (1997)]). In what follows we describe its monotonic variant (see [Leary (2000)])

**MBH($X$ : initial local minimum)**

**Step 1.** Let $Y \in \mathcal{N}(X)$;

**Step 2.** **if** $f(Y) < f(X)$ **then** set $X := Y$;
**else** reject $Y$;

**Step 3.** Repeat Steps 1–2 until $MaxNoImprove$ consecutive rejections have occurred;
**return** $X$;

$Y$ is usually obtained starting a local search from a random perturbation of $X$. In spite of its simplicity, this algorithm is very efficient in detecting funnel bottoms. Note that when multiple funnel bottoms exist, it is usually necessary to run MBH many times from different initial solutions in order to detect the funnel bottom corresponding to the global minimum.

It is also interesting to remark the relation between MBH and an heuristic approach for combinatorial optimization problems, namely Iterated Local Search (ILS). Both MBH and ILS perform local moves between local minima, i.e. when they generate a new local minimum they do not completely disrupt the structure of the previous local minimum (as it happens, for instance, when we restart from a completely new random solution) but they try to partially preserve it.

In the next section we will see how to adapt the MBH method to packing problems.

# 3. Monotonic Basin Hopping for packing problems

In order to apply the MBH method to our packing problem, the first step is to reformulate it as a mathematical programming problem for which efficient local search procedures exist.

It is easy to see that problem (2) can be reformulated as the following constrained problem

$$\begin{aligned} d_n = \max\ & d \\ (x_i - x_j)^2 + (y_i - y_j)^2 \geq d^2 \qquad & \forall\ i \neq j \\ x_i, y_i \in [0,1] \qquad & i = 1, \ldots, n \end{aligned} \tag{4}$$

This problem has a linear objective function, but reverse convex (see e.g. [Horst at al. (1996)]) quadratic constraints. Thus this problem is nonconvex and highly multimodal. However, with respect to general problems with nonconvex constraints, for which even finding feasible solutions may be an extremely hard task, here feasible solutions are easy to find.

In order to define MBH a neighborhood structure has to be defined. We remark that the idea of "neighbor solutions" is already present in the existing literature. Indeed, the moves of single points in [Boll et al.(2000)] and [Casado et al.(1998)] can be viewed within this framework. However, in our definition there are two major differences: the first one is that *all* points in a solution are simultaneously perturbed; then, following the MBH approach, we do not stop at the perturbed solution but we move to the local minimum reached when starting a local search from the perturbed solution.

The perturbation is defined as follows. Let $\xi_n = \frac{\alpha}{\sqrt{n}}$ (in our computation we set $\alpha = 0.5$). For each variable $z$, where $z = x_i$ or $y_i$, $i = 1, \ldots, n$, if we denote by $\hat{z}$ the current value of this variable, then the new value of the variable is uniformly sampled over the interval

$$[\max\{0, \hat{z} - \xi_n\}, \min\{1, \hat{z} + \xi_n\}] \subseteq [0,1].$$

Some comments are needed at this point. First of all we need to comment our choice of the value for $\xi_n$. It is well known (see e.g. [Croft et al.(1991)]) that

$$d_n \sim \frac{2^{1/2} 3^{-1/4}}{\sqrt{n}}. \tag{5}$$

Then, let us consider an "active" pair of points of an optimal solution (a pair of points is said to be active if the distance between the two points is equal to $d_n$, i.e. if the associated constraint in (4) is active). The maximum possible step-size allowed by our perturbation method is given by $\sqrt{2}\xi_n$ which is only slightly greater than $d_n/2$. This basically means that after the perturbation, with a very high probability, both points will still lie on the same side of a maximally separating line. This way the structure of the previous solution is still partially preserved, as it has to be for local moves. Also note that the perturbation of a point

only affects its neighbors (it is easily seen that all points at distance larger than $d_n + 2\sqrt{2}\xi_n$ from a given point will still be at a distance larger than $d_n$ after the perturbation).

Next we note that there is some asymmetry in the perturbation of a point when this is close to the border. Indeed, in this case the perturbed point will be generated with a higher probability "towards" the interior of the unit square rather than "towards" the border (just think about the perturbation of a point lying at a vertex of the unit square). Therefore, this perturbation mechanism somehow acts as an implicit compression tool, which might have a positive effect. Indeed, for molecular conformation problems it has been observed that the insertion of explicit compression tools has extremely positive effects in detecting the most challenging global minima (see [doye (2000)], [Doye et al. (2004)], [Locatelli et al. (2003)]).

Finally, we remark that our perturbation is still incomplete. Indeed, Problem (4) has a further variable which is $d$. This variable is perturbed in a different way with respect to the others. In order to generate a perturbed point within the feasible region of problem (4)the value for $d$ has to be chosen in $[0, \bar{d}]$, where $\bar{d}$ is the minimum distance between points of the perturbed solution, i.e.

$$\bar{d} = \min_{i \neq j} \sqrt{(\bar{x}_i - \bar{x}_j)^2 + (\bar{y}_i - \bar{y}_j)^2}. \tag{6}$$

It has been experimentally observed that setting $d = \bar{d}$ is not an efficient choice, while the choice $d = 0$ turns out to be quite efficient. This can be explained as follows. With $d = \bar{d}$ the perturbed solution lies on the border of the feasible region, while with $d = 0$ the perturbed solution usually lies in the interior of the feasible region, which gives more freedom to the local search procedure started from the perturbed solution.

## 3.1. Population Basin Hopping

Although we cannot exhibit a formal proof, it seems that, as the dimension of the problem increases, the number of funnels tends to increase. Following an approach experimented with success for the optimization of molecular structures [Grosso et al. (2005)], we implemented a population based version of MBH for the problem of disk packing. The idea is that of working with a set of locally optimal disk packings (a population of parents). Each solution in the set is perturbed, according to the same schema as before, and from each perturbed solution a local search is started, producing a new set of solutions (a population of children). At this point each solution of the new population (each child) is compared with each element in the population of parents looking for the closest one. If the child is better than its closest parent,

it replaces it. After all children have been compared with the population of parents and a new set of parents has been obtained, the procedure restarts. In order to implement this strategy, a measure of similarity is needed between pairs of solutions. In our first experiments, we used the ideas of [Lee et a. (2001)] and defined dissimilarity in the following way. Let $\rho$ be a threshold; given a disk packing $X = \{X_i \in \mathbb{R}^2\}_{i=1}^n$, let

$$N_{X;\rho}(i) := |\{j \in 1, \ldots, n, j \neq i : \|X_i - X_j\| \leq \rho\}|$$

i.e., $N_{X;\rho}(i)$ is the number of disks whose distance from disk $i$ does not exceed $\rho$. Let

$$H_{X;\rho}(k) := |\{i \in 1, \ldots, n : N_{X;\rho}(i) = k\}|$$

be the histogram of $N$. Then we can compare two packings $X$ and $Y$ evaluating

$$D(X, Y) = \sum_{k=0}^{n} (k+1)|H_{X;\rho}(k) - H_{Y;\rho}(k)|.$$

As a general rule it is advisable to choose a value for $\rho$ which is somewhat larger than the expected minimum distance between pairs of points; choosing a $\rho$ value which is too strict might not be a good choice. In fact in many configurations, even putative optimal ones, there are "free disks", i.e. disks that can be moved without violating any constraint. Being relatively free to move, these disks might or might not be in contact with others, so that through a dissimilarity measure two configurations which differ only for the location of free disks might erroneously be considered as different. In our computations we usually set $\rho = 1.5/\sqrt{n}$.

This measure of dissimilarity is quite easy to compute and displays a very good discriminating power; moreover it is invariant to rigid transformations of the packing configurations.

# 4. Computational experiments

In our computational experiments, for the standard version of MBH we employed $MaxNoImprove = 100$, while for the population variant (PBH) we choose 40 elements in the population and $MaxNoImprove = 30$; in PBH we used this parameter on the whole population, i.e. we choose to stop the algorithm as soon as 30 generations of children had been obtained with no improvement in the best one. For local optimization, after a few experimentation, we choose to adopt the large scale SQP method SNOPT [Gill et al. (2002)] with feasibility tolerance parameters set to $10^{-12}$.

As already mentioned in the Introduction, putative optimal configurations are maintained by E. Specht on his web site [Packomania]; there, after a new finding for $n = 97$ in January 2003, Specht comments: *"It seems to be very unlikely to find still better packings for* $n \leq 100$*, but it is possible"*. Basically, since many different methods had been tested on these instances, it seemed to be reasonable to conjecture that packings known at that time for $n$ up to 100 were indeed the optimal ones. Therefore, in order to check the robustness of our method the first aim was to show that we were able to efficiently reproduce at least all the best known packings up to $n = 100$. What we obtained was actually much more than that. Indeed, in this range we have been able to detect many new best known solutions, now reported in [Packomania] and to confirm most of the remaining. Stimulated by this success we performed numerical experiments up to $n = 130$ again obtaining many improved putative optima. In the following table we report the situation up to now for $n \in [50, 130]$ – we did not report any result for smaller values of $n$ because we could usually very easily confirm all of the known putative optima. In normal typeface we report the values of $n$ for which we could obtain the previously known putative optimum with a maximum error on the distance of $10^{-12}$. **Bold** numbers correspond to cases in which we could improve over previously known putative optima. *Italic* is reserved for those instances for which we were not able to obtain, within the prescribed accuracy, the putative optimum. We remark that in this problem an apparently small improvement in the distance can, and usually does, imply a significantly different geometry. Just for illustration we present in Figure 1 a comparison between the previous putative optimum and our improved configuration for $n = 104$.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 51 | 52 | **53** | 54 | 55 | 56 | 57 | 58 | **59** |
| 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | **68** | 69 |
| 70 | 71 | 72 | **73** | 74 | *75* | 76 | **77** | **78** | 79 |
| 80 | 81 | 82 | 83 | 84 | **85** | **86** | 87 | **88** | 89 |
| 90 | 91 | 92 | 93 | 94 | **95** | 96 | 97 | **98** | 99 |
| **100** | **101** | 102 | 103 | **104** | **105** | **106** | **107** | **108** | **109** |
| 110 | **111** | *112* | **113** | 114 | **115** | **116** | **117** | **118** | **119** |
| 120 | **121** | **122** | **123** | 124 | 125 | 126 | **127** | **128** | **129** |
| **130** | | | | | | | | | |

The results summarized in the above table have been obtained by executing 100 independent experiments with MBH for $n \leq 100$ and 10 independent runs with PBH for all values of $n$. Many of the new records and (when no new record has been detected) of the old putative optima have been found several times by both methods (see also Section 4.1). For what

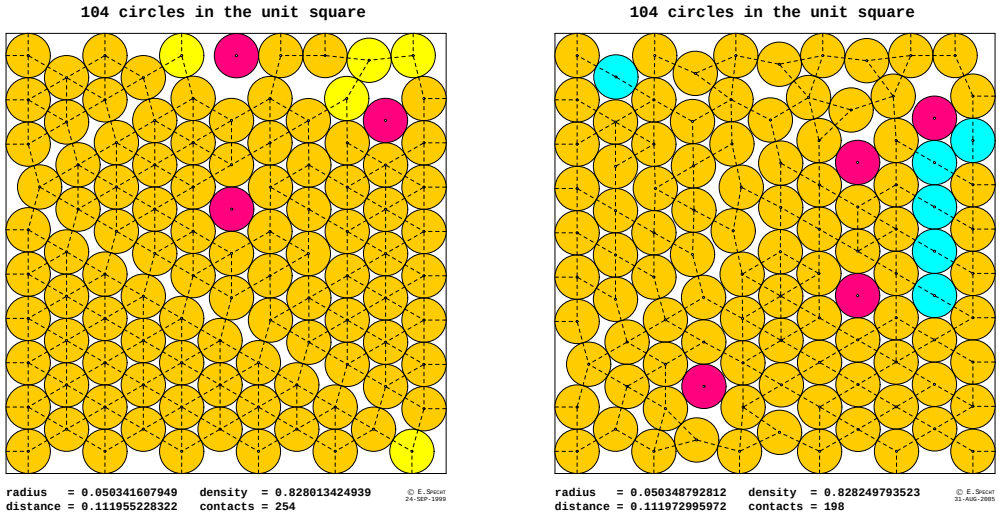


Figure 1: Putative optimal packings for $n = 104$: previously known record (left, $d = 0.111955228322$) and new putative optimum (right, $d = 0.111972995972$)

concerns failures, only two cases have occurred. $n = 112$ could not be found even after running 100 PBH independent experiments. We feel that, for larger dimensions, there will be more and more instances in which a simple algorithm like the one introduced in this paper will not be sufficient and some sort of local moves, tailored to this particular problem should be included. The case $n = 75$ is particularly significative: we could indeed find a configuration whose distance is just $10^{-10}$ smaller than that of the putative optimum deposited in the web site of Specht, but most of the times we reached a configuration whose distance is more than $10^{-7}$ smaller than that of the putative optimum. However, visual inspection of the latter configuration versus the putative optimal one does not reveal any difference in their structure. Thus it is very likely that the failure is caused by a failure in *local* optimization, which is trapped in a configuration which is slightly worse than the optimal. However, even performing various runs of perturbation and optimization from our best solution could not lead to the putative optimum. Neither we could obtain the putative optimum changing the local optimization routine, either by substituting it with other constrained optimization software, or by using the method implemented by Specht and available from his web site. Thus it seems that the region of attraction of this optimum is significantly narrow and some different strategy should be identified to deal with this phenomenon. Research is planned on this subject.

## 4.1. Further observations on the performance of MBH

¿From the point of view of efficiency we observe that solving problem (4) is equivalent to solving any problem with the following form:

$$\begin{aligned} &\max\ f(d) \\ (x_i - x_j)^2 + (y_i - y_j)^2 &\geq g(d^2) && \forall\ i \neq j \\ x_i, y_i &\in [0,1] && i = 1, \ldots, n \end{aligned} \tag{7}$$

where $f$ in a monotonic increasing function and $g$ is some continuous function satisfying

$$g(d^2) \begin{cases} = d^2 & \text{for } d \geq d^\star \\ \leq d^2 & \text{for } d < d^\star \end{cases}$$

where $d^\star$ denotes a lower bound for the optimal value of (4). Although any function $f$ and $g$ satisfying the above conditions lead to formulations which are equivalent from the theoretical point of view, considerable differences can be observed in the practical behavior (see also [Dimnaku et al. (2005)] for a discussion about the impact of equivalent formulations).

Another interesting aspect is the impact of compression. As we remarked in Section 3, the mechanism used to generate perturbed points in MBH introduces a sort of implicit compression, favoring generation of points in the interior of the square. Therefore we tested the influence on performance of the addition of a simple explicit compression mechanism: after having perturbed a solution according to the mechanism described in Section (3), we further modify this solution by multiplying all of the coordinates by a factor $r \in (0,1)$ (in particular, we tested $r = 0.8$); thus the resulting configuration shrinks to fit within the square $[0,r]^2$.

In order to test the impact of different choices, we selected four test environments (listed in Table 1) and performed for each of them 100 runs with $n$ ranging from $n = 31$ (the first $n$ value for which a computer-aided proof of optimality has not been given yet) up to $n = 100$. Note that choice C1 is equivalent to (4) and that the introduction of the *tol* value is necessary to guarantee the differentiability of the objective function at $d = 0$. Also note that at each iteration of a MBH run, the $d^\star$ value is the current record value for that run.

In Figures 2 and 3 we report some statistics on the number of successes obtained by our method. We have to warn that the notion of success is not easy to define in this problem. We decided to declare a run a success when the final record value $d^*$ is within $5 * 10^{-11}$ from the best known value. However, it turns out that in some runs which are declared as

| Choice | $f$ | $g$ | Compression |
|---|---|---|---|
| C1 | $d$ | $d^2$ | No |
| C2 | $100d$ | $d^2 - [\max\{0, d^\star - d\}]^2$ | No |
| C3 | $100\sqrt{d + tol}$ | $d^2 - [\max\{0, d^\star - d\}]^2$ | No |
| C4 | $100\sqrt{d + tol}$ | $d^2 - [\max\{0, d^\star - d\}]^2$ | Yes |

Table 1: Different tested choices ($tol = 10^{-5}$).

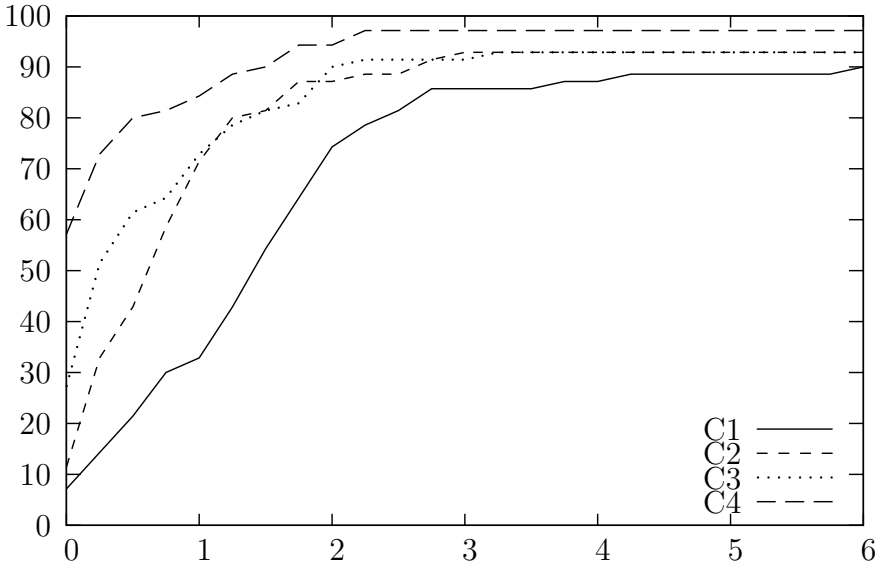


Figure 2: Performance profiles of the four tested choices (see [Dolan et al. (2002)] for detailed information on this kind of graphics). Each curve represents the percentage of problems where the number of successes is at least $2^{-x}$ times that of the best choice. In particular, test choice C4 is the best one in 57% of the problems while in 84% cases its success rate is at least 1/2 with respect to the best one.

failures according to this criterion, the "correct" configuration is actually detected but the local solver is unable to return a solution with the prescribed precision. This happened e.g. for the $n = 86$ case but even more evidently for the already discussed $n = 75$ case.

We also computed the average of the final record values and it turned out that the highest average was attained by choice C4 in 47 out of 70 cases, by choice C3 in 21 out of 70 cases, by choice C2 in 2 out of 70 cases, and by choice C1 only in 1 out of 70 cases. This shows that imposing a very high slope around $d = 0$ pushes MBH towards solutions with high function values (and the addition of compression further increases this tendency). Note also that the case $n = 76$ could only be solved with choice C4 while with all the other choices only a solution differing by more than $10^{-5}$ from the best known one could be reached.

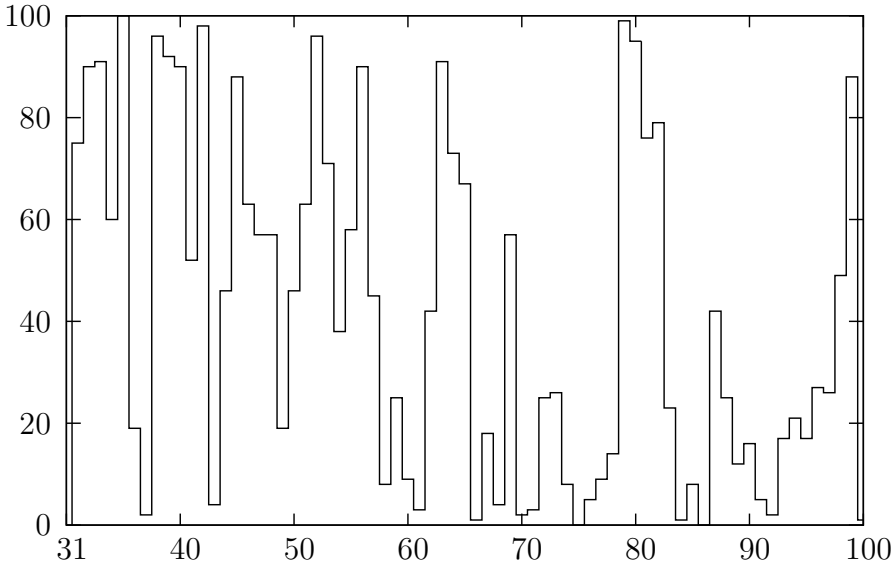


Figure 3: Percentage of successes for choice C4 on the test set ($n = 31, \ldots, 100$)

# 5. Directions for future research

Although the above computational results are already quite satisfactory, there are still many possible ways through which we might hope to improve the proposed approach.

One possible improvement is the local search procedure. We have employed the SNOPT procedure, which is a general one for constrained problems, but it is possible that alternative local search procedures which take into account the special structure of problem (4) or the equivalent ones (7) may represent an improvement, if not from the point of view of the quality of the results, at least from the point of view of computation times. We think, for example, to local optimization methods specifically designed for nonconvex problems.

It is also not strictly necessary to use reformulation (4) or the equivalent ones (7) of our problem. An interesting alternative could be to consider the unconstrained minimization of the energy function (3) as done in [Nurmela et al. (1997)].

As remarked in Section 3, the perturbation mechanism acts as an implicit compression tool. But explicit compression tools as those employed in two-phase local searches for molecular conformation problems (see [Doye et al. (2004)], [Locatelli et al. (2003)]) could also be helpful.

Finally, in molecular conformation problems, especially for large number of atoms, great benefits come from so called direct mutation. This is a nonrandom move where "bad" atoms, i.e. atoms whose contribution to the total energy is not as high as expected, are removed

from their site and inserted at a "good" site. Something similar could also be thought for the packing problem by removing e.g. a disk which "blocks" other disks, so that these can reach new positions.

The above suggestions are only some of the possible ones which could be used to improve the current approach. However, we believe that the recognition of a funneling landscape structure for the solutions of the packing problem and the consequent definition of a MBH approach for this problem represent a remarkable step in the direction of solving it, as also confirmed by the computational results. We also remark again that, although we restricted our attention to the problem of scattering points within a square (for which a wide literature exists), the same approach can be easily extended to the problem of scattering points within other convex figures, like an equilateral triangle or a disk: it is enough to substitute the bound constraints for the variables in (4) with the constraints defining our convex figure. Finally, this approach is by no means constrained to two-dimensional packing problems. We just started a few computational experiments in sphere packing in the unit cube in $\mathbb{R}^3$ and, with respect to the results published in [Gensane (2004)], we could obtain an improved configuration for packing $n = 28$ spheres, while confirming most of the results reported in the cited paper. Details on these extensions will appear elsewhere.

# Acknowledgements

This research has been partially supported by Progetto FIRB "Ottimizzazione Non Lineare su Larga Scala"

# References

[Boll et al.(2000)] Boll D.V., Donovan J., Graham R.L., Lubachevsky B.D., Improving dense packings of equal disks in a square, *The Electronic J. of Comb.*, 7, R46 (2000)

[Casado et al.(1998)] Casado L.G., Garcia I., Szabó P.G., Csendes T., Equal circles packing in square II: new results for up to 100 circles using the TAMSASS-PECS algorithm, in *New trends in equilibrium systems: Mátraháza Optimization Days*, Kluwer Academic Publishers (1998)

[Croft et al.(1991)] Croft H.G., Falconer K.J., Guy R.K., *Unsolved problems in geometry*, Springer-Verlag, New York, 1991

[van Dam et al.] van Dam E., Husslage B., den Hertog D., Melissen H., Maximin Latin hybercube designs in two dimensions, to appear in *Operations Research*

[Dimnaku et al. (2005)] Dimnaku A., Kincaid R., Trosset M.W., Approximate solutions of continuous dispersion problems, *Annals of OR*, 136, 65-80 (2005)

[Dolan et al. (2002)] Dolan E.D., Morè J.J., Benchmarking optimization software with performance profiles, *Mathematical Programming*, 91, 201–213 (2002)

[doye (2000)] Doye J.P.K., The effect of compression on the global optimization of atomic clusters, *Phys.Rev.E*, 62, 8753-8761 (2000)

[Doye et al. (2004)] Doye J.P.K., Leary R.H., Locatelli M., Schoen F., The global optimization of Morse clusters by potential transformations, *INFORMS J. Computing*, 16, 371-379 (2004)

[Gensane (2004)] Gensane T. Dense packings of equal spheres in a cube, *The Electronic Journal of Combinatorics*, 11, 1–17 (2004)

[Graham et al. (1996)] Graham R.L., Lubachevsky B.D., Repeated patterns of dense packings of equal disks in a square, *The Electronic J. of Comb.*, 3, 1-16 (1996)

[de Groot et al. (1991)] de Groot C., Peikert R., Würtz D., Monagan M., Packing circles in a square: a review and new results, *System Modelling and Optimization, Proc. 15th IFIP Conf. Zürich*, 45-54 (1991)

[Gill et al. (2002)] Gill P.E., Murray W., Saunders M. A., SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization, *SIAM J. Optim.*, 12, 979–1006, (2002)

[Grosso et al. (2005)] Grosso A., Locatelli M., Schoen F., A Population Based Approach for Hard Global Optimization Problems Based on Dissimilarity Measures, *Optimization on line*, 1056, (2005)

[Horst at al. (1996)] Horst R., Tuy, H., *Global Optimization: Deterministic Approaches*, 3rd edition, Springer, Heidelberg (1996)

[Leary (2000)] Leary R.H., Global optimization on funneling landscapes, *J. Global Optim.*, 18, 367–383, (2000)

[Lee et a. (2001)] Lee J., Lee I., Lee J., Unbiased Global Optimization of Lennard-Jones Clusters for $N \leq 201$ by Conformational Space Annealing Method, *Phys. Rev. Lett.*, 91, 080201 (2003)

[Locatelli et al. (2002)] Locatelli M., Raber U., Packing equal circles in a square: a deterministic global optimization approach, *Discrete Applied Mathematics*, 122,139-166 (2002)

[Locatelli et al. (2003)] Locatelli M., Schoen F., Efficient algorithms for large scale global optimization: Lennard-Jones clusters, *Computational Optimization and Applications*, 26, 173-190 (2003)

[Locatelli (2005)] Locatelli M., On the multilevel structure of global optimization problems, *Computational Optimization and Applications*, 30, 5-22 (2005)

[Maranas et al. (1995)] Maranas C.D., Floudas C.A., Pardalos P.M., New results in the packing of equal circles in a square, *Discrete Mathematics*, 142, 287-293 (1995)

[Markot et al. (2005)] Markot M.C., Csendes T., A new verified optimization technique for the "packing circles in a unit square" problems, to appear in *SIAM Journal on Optimization* (2005)

[Nurmela et al. (1999)] Nurmela K.J., Oestergard P.R.J., More Optimal Packings of Equal Circles in a Square, *Discrete Comput. Geom.*, 22, 439-457 (1999)

[Nurmela et al. (1997)] Nurmela K.J., Oestergard P.R.J., Packing up to 50 equal circles in a square, *Discrete Comput. Geom.*, 18, 111-120 (1997)

[Packomania] Packomania web site maintained by E.Specht, `www.packomania.com`

[Schaer (1965)] Schaer J., The densest packing of nine circles in a square, *Canadian Mathematical Bulletin*, 8, 273-277 (1965)

[Schaer et al. (1965)] Schaer J., Meir A,, On a geometric extremum problem, *Canadian Mathematical Bulletin*, 8, 21-27 (1965)

[Schwartz (1970)] Schwartz B.L., Separating points in a square, *J. Recreational Math.*, 3, 195-204 (1970)

[Szabò (2000)] Szabò P.G., Some new structures for the "Equal Circles Packing in a Square" problem, *Central European Journal of Operations Research*, 8, 79-91 (2000)

[Wales et al. (1997)] Wales D.J., Doye J.P.K., Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms, *J. Phys. Chem. A*, 101, 5111–5116, (1997)

[Wengerodt (1983)] Wengerodt G., Die dichteste Packung von 16 Kreisen in einem Quadrat, *Beiträge Algebra Geom.*,16, 173-190 (1983)

[Wengerodt (1987)] Wengerodt G., Die dichteste Packung von 14 Kreisen in einem Quadrat, *Beiträge Algebra Geom.*,25, 25-46 (1987)

[Wengerodt et al. (1987)] Wengerodt G., Kirchner K., Die dichteste Packung von 36 Kreisen in einem Quadrat, *Beiträge Algebra Geom.*,25, 147-159 (1987)

[Wengerodt (1987)] Wengerodt G., Die dichteste Packung von 25 Kreisen in einem Quadrat, *Ann.Univ.Sci.Budapest Eötvös Sect. Math.*,30, 3-15 (1987)