Optimizing the Packing of Cylinders into a Rectangular Container: A Nonlinear Approach

E. G. Birgin [*] J. M. Martínez [†] D. P. Ronconi [‡]

February 13, 2003.

Abstract

The container loading problem has important industrial and commercial applications. An increase in the number of items in a container leads to a decrease in cost. For this reason the related optimization problem is of economic importance. In this work, a procedure based on a nonlinear decision problem to solve the cylinder packing problem with identical diameters is presented. This formulation is based on the fact that the centers of the cylinders have to be inside the rectangular box defined by the base of the container (a radius far from the frontier) and far from each other at least one diameter. With this basic premise the procedure tries to find the maximum number of cylinder centers that satisfy these restrictions. The continuous nature of the problem is one of the reasons that motivated this study. A comparative study with other methods of the literature is presented and better results are achieved.

Key words: Cylinder packing, rectangular container, circular container, nonlinear programming, bound-constrained minimization, convex-constrained minimization.

[*] Department of Computer Science IME-USP, University of São Paulo, Rua do Matão 1010, Cidade Universitária, 05508-090, São Paulo SP, Brazil. This author was supported by PRONEX-Optimization 76.79.1008-00, FAPESP (Grants 01-04597-4 and 02-00094-0) and CNPq (Grant 300151/00-4). Corresponding author. FAX: +55(11)3091-6134. e-mail: egbirgin@ime.usp.br

[†] Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil. This author was supported by PRONEX-Optimization 76.79.1008-00, FAPESP (Grant 01-04597-4), CNPq and FAEP-UNICAMP. e-mail: martinez@ime.unicamp.br

[‡] Department of Production Engineering, EP-USP, University of São Paulo, Av. Prof. Almeida Prado, 128, Cidade Universitária, 05508-090, São Paulo SP, Brazil. This author was supported by FAPESP (Grants 00-01715-3 and 01-02972-2). e-mail: dronconi@usp.br

1 Introduction

The container loading problem is a three-dimensional problem that consists of arranging items of different sizes inside large objects, in such a way that an objective function (in general, the volume loaded) is optimized. This problem has important industrial and commercial applications. The goods are packed into standard containers for ease of handling so they can be transported by ship, trucks or rail car with minimum damage. An increase in the number of items in a container leads to a decrease in cost. For this reason the related optimization problem is of economic importance.

In the present work we are particularly concerned with the densest packing of identical cylinders inside a container. All cylinders are considered to have the same height. Consequently, the problem can be solved as a two-dimensional problem where the solution method has to present the position of the circular bases of the cylinders inside the rectangular container floor without overlapping.

In [11] some algorithms for packing identical circles with single geometric patterns are proposed. A practical application can be found in [12] where the authors present a container loading software for a paper manufacturer. They use an heuristic method based on several patterns to deal with the two-dimensional circle packing problem. The paper [10] is concerned with the relationship between the way circles are packed into a rectangular box and the efficiency of the resultant palletisation. In [7] algorithms based on Simulated Annealing are presented. Furthermore, the same authors [8] propose a new upper bound for this problem. In [14] several heuristic approaches are discussed for the problem of packing cylinders of different diameters inside a container. A mixed integer nonlinear formulation is proposed, but the authors warn that the associated computational effort is excessive.

In this work, we present a procedure based on a nonlinear decision problem to solve the cylinder packing problem with identical diameters. This formulation is based on the fact that the centers of the cylinders have to be inside the rectangular box defined by the base of the container (a radius far from the frontier) and far from each other at least one diameter. The computational procedure tries to find the maximum number of cylinder centers that satisfy these constraints. A comparative study with other methods of the literature is presented and better results are achieved.

This paper is organized as follows. Section 2 describes the nonlinear decision problem. Section 3 is devoted to the procedure which uses the decision problem to pack as many circles as possible. In Section 4 numerical results and a comparison with other methods are presented. Section 5 briefly describes some extensions for packing circles into circles. The last section contains final remarks.

2 Decision problem

The decision problem treated in this section is:

> *Given k circles of radio r and a rectangular box of dimension $d_1 \times d_2$, whether is it possible to locate all the circles into the box or not.*

We introduce a nonlinear model for this problem. Finding the answer for the decision problem will depend on finding the global minimizer of a nonconvex and nonlinear optimization problem. We also describe a solver to find first order stationary points (very likely, local minimizers) of the introduced model and a strategy to enhance the probability of finding global minimizers.

2.1 Nonlinear formulation

We wish to place k circles of radius r into the rectangular box $[0, d_1] \times [0, d_2]$ in such a way that the intersection between any pair of circles i and j, $i \neq j$, is at most one point, i.e., the circles are not overlapped. Therefore, given k, r, d_1 and d_2, the goal is to determine $p^1, \ldots, p^k \in [r, d_1 - r] \times [r, d_2 - r]$ solving the problem:

$$
\begin{array}{ll}
\text{Minimize} & \displaystyle\sum_{i \neq j} \max(0, (2r)^2 - \|p^i - p^j\|_2^2)^2 \\
\text{subject to} & \\
& r \leq p_1^i \leq d_1 - r, \ \text{and} \\
& r \leq p_2^i \leq d_2 - r, \ \text{for } i = 1, \ldots, k.
\end{array}
\tag{1}
$$

The points $p^1, \ldots, p^k$ are the centers of the desired circles. p_1^i and p_2^i represent the abscissa and the ordinate of $p^i \in I\!R^2, i = 1, \ldots, k$. If the objective function value at the global minimizer of this problem is zero then the answer of the decision problem is *YES*, otherwise, the answer is *NO*.

Observe that (1) is a continuous optimization problem where the objective function has continuous first derivatives but discontinuous second derivatives. This motivates the method used in the following section.

2.2 Regularized Hessians

The problem considered in Section 2.1 has the following general form:

$$
\text{Minimize } \Phi(x) \text{ s.t. } x \in \Omega \subset I\!R^n, \tag{2}
$$

where

$$
\Phi(x) = f(x) + \frac{1}{2} \sum_{i=1}^{m} \max\{0, g_i(x)\}^2,
$$

and Ω is closed and convex. In our application $f(x) \equiv 0$, but we would like to consider the formulation (2) having in mind further applications. The factor $\frac{1}{2}$ in the second term of the objective function merely simplifies derivatives.

Assuming that g_i has continuous second derivatives, it is easy to see that Φ has continuous first (but not second) derivatives. The second derivatives of Φ are, in general, discontinuous at the points where $g_i(x) = 0$. This is an disadvantage for minimization algorithms based on quadratic models, like Newton's method, which enjoys good convergence properties. We aim to overcome this disadvantage by means of a perturbation of the Hessian matrix of Φ in the points where $g_i(x) \leq 0$ for some i. Unlike the original one, the perturbed Hessian will be continuous.

Consider the associated problem

$$\text{Minimize } \psi(x,z) \text{ s.t. } x \in \Omega \subset I\!R^n, \tag{3}$$

where

$$\psi(x,z) = f(x) + \frac{1}{2}\sum_{i=1}^{m} [\ g_i(x) + z_i^2\ ]^{\ 2}.$$

In the following lemma we prove that problems (2) and (3) are equivalent. Problem (3) has continuous second derivatives but depends on the additional variables $z_1, \ldots, z_m$. We will see that discontinuities of the second derivatives of (2) correspond to singularities of the Hessian of the objective function of (3).

Lemma 2.1 *The point $\bar{x} \in \Omega$ is a global minimizer of (2) if, and only if, there exists $\bar{z} \in I\!R^m$ such that $(\bar{x}, \bar{z})$ is a global minimizer of (3). Moreover, $\Phi(\bar{x}) = \psi(\bar{x}, \bar{z})$.*

Proof. See Appendix.

This equivalence motivates us to study Newton-like minimization methods for solving (3). Problem (3) has continuous second derivatives, but it has m additional variables. Computing the gradient of ψ, we get

$$\nabla\psi(x,z) = \begin{pmatrix} \nabla f(x) + \sum_{i=1}^m [g_i(x) + z_i^2] \nabla g_i(x) \\ 2[g_i(x) + z_1^2] z_1 \\ \vdots \\ 2[g_m(x) + z_m^2] z_m \end{pmatrix},$$

and computing the Hessian matrix, we get

$$\nabla^2\psi(x,z) = \left(\begin{array}{c|ccc} \begin{array}{c}\nabla^2 f(x) + \sum_{i=1}^m \nabla g_i(x) \nabla g_i(x)^T \\ + \sum_{i=1}^m [g_i(x) + z_i^2] \nabla^2 g_i(x)\end{array} & 2z_1 \nabla g_1(x) & \ldots & 2z_m \nabla g_m(x) \\ \hline 2z_1 \nabla g_1(x)^T & 6z_1^2 + 2g_1(x) & & 0 \\ \vdots & & \ddots & \\ 2z_m \nabla g_m(x)^T & 0 & & 6z_m^2 + 2g_m(x) \end{array}\right).$$

We shall call "good pairs" (x,z) to those pairs such that $z_i^2 = -g_i(x)^2$ when $g_i(x) \le 0$ and $z_i = 0$ when $g_i(x) > 0$. In other words, $z_i^2 = \max\{0, -g_i(x)^2\}$. The denomination is justified because, as it can be seen in the proof of the lemma above, whenever (x,z) is a feasible point of (3), we can obtain a "good pair" where the functional value is not larger. If (x,z) is an iterate of a minimization method for (3), its replacement by the corresponding good pair always represents an improvement.

Therefore, it is interesting to compute the gradient and the Hessian of ψ at good pairs. In that case, assuming without loss of generality that $g_i(x) < 0$, $i = 1, \ldots, p$, $g_i(x) \ge 0$, $i = p+1, \ldots, m$, we obtain

$$\nabla\psi(x,z) = \begin{pmatrix} \nabla f(x) + \sum_{i=p+1}^m g_i(x)\nabla g_i(x) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and

$$\nabla^2\psi(x,z) =$$

$$\left(\begin{array}{c|ccc|ccc}
\begin{array}{c}\nabla^2 f(x) + \displaystyle\sum_{i=1}^m \nabla g_i(x)\nabla g_i(x)^T \\ + \displaystyle\sum_{i=p+1}^m g_i(x)\nabla^2 g_i(x)\end{array} & 2\sqrt{-g_1(x)}\nabla g_1(x) & \ldots & 2\sqrt{-g_p(x)}\nabla g_p(x) & & 0 & \\
\hline
2\sqrt{-g_1(x)}\nabla g_1(x)^T & -4g_1(x) & & 0 & & & \\
\vdots & & \ddots & & & 0 & \\
2\sqrt{-g_p(x)}\nabla g_p(x)^T & 0 & & -4g_p(x) & & & \\
\hline
 & & & & 2g_{p+1}(x) & & 0 \\
0 & & 0 & & & \ddots & \\
 & & & & 0 & & 2g_m(x)
\end{array}\right).$$

Although $\nabla^2\psi(x,z)$ is continuous, it is a singular matrix for all the points x such that $g_i(x) = 0$ for some i. So, this Hessian is singular exactly at the points where the Hessian of Φ may not exist. This confirms our previous claim on the relation between discontinuities and singularities of the Hessians of (2) and (3).

It is interesting to interprete Newtonian iterations associated to ψ in terms of the variables x only. We do this in the following theorem.

Theorem 2.1 *Assume that (x,z) is a good pair. Assume that $\overline{\Delta x} \in I\!R^n$ satisfies*

$$[\nabla^2 f(x) + \sum_{i=p+1}^m [\nabla g_i(x)\nabla g_i(x)^T + g_i(x)\nabla^2 g_i(x)]\overline{\Delta x} = -[\nabla f(x) + \sum_{i=p+1}^m g_i(x)\nabla g_i(x)]. \qquad (4)$$

Then, there exists $\overline{\Delta z} \in I\!R^m$ such that

$$\nabla^2\psi(x,z)\begin{pmatrix} \overline{\Delta x} \\ \overline{\Delta z} \end{pmatrix} = -\nabla\psi(x,z). \qquad (5)$$

Proof. See Appendix.

Remarks.

(i) The theorem above shows that, essentially, a Newtonian iteration for the minimization of $\psi(x,z)$ followed by a restoration $z_i^2 \leftarrow \max\{0, -g_i(x)^2\}$ is equivalent to a Newton iteration for minimizing $\Phi(x)$ provided that we define

$$\nabla^2 \max\{0, g_i(x)\}^2 = \nabla^2[g_i(x)^2], \text{ if } g_i(x) = 0.$$

In this way, we defined the Hessian at the points where it does not exist.

(ii) The singularity of $\nabla^2\psi(x,z)$ corresponds to the discontinuity of $\nabla^2\Phi(x)$.

These observations motivate the study of perturbations of the Newtonian system that eliminate the essential singularity. Using appropriate perturbations of singular matrices in order to improve conditioning and solvability is a common device in numerical linear algebra. Accordingly, we define the regularized Hessian $\nabla^2\psi(x,z,\varepsilon)$ for all good (x,z) and $\varepsilon > 0$ as follows:

$$\nabla^2\psi(x,z,\varepsilon) =$$

$$\left(\begin{array}{c|ccc|ccc}
\begin{array}{c}\nabla^2 f(x) + \sum_{i=1}^m \nabla g_i(x)\nabla g_i(x)^T \\ + \sum_{i=p+1}^m g_i(x)\nabla^2 g_i(x)\end{array} & 2\sqrt{-g_1(x)}\nabla g_1(x) & \dots & 2\sqrt{-g_p(x)}\nabla g_p(x) & & 0 & \\
\hline
2\sqrt{-g_1(x)}\nabla g_1(x)^T & -4g_1(x)+2\varepsilon & & 0 & & & \\
\vdots & & \ddots & & & 0 & \\
2\sqrt{-g_p(x)}\nabla g_p(x)^T & 0 & & -4g_p(x)+2\varepsilon & & & \\
\hline
 & & & & 2g_{p+1}(x)+2\varepsilon & & 0 \\
0 & & 0 & & & \ddots & \\
 & & & & 0 & & 2g_m(x)+2\varepsilon
\end{array}\right).$$

The following theorem, in analogy to Theorem 2.1, relates perturbed Newtonian iterations corresponding to the minimization of ψ to perturbed (regularized) iterations related to the minimization of Φ.

Theorem 2.2. *Assume that (x,z) is a good pair. Assume that Δx satisfies*

$$\begin{aligned}&\{\nabla^2 f(x) + \sum_{i=1}^p \frac{\varepsilon}{\varepsilon - 2g_i(x)} \nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^m [\nabla g_i(x)\nabla g_i(x)^T + g_i(x)\nabla^2 g_i(x)]\}\Delta x = \\ &-[\nabla f(x) + \sum_{i=p+1}^m g_i(x)\nabla g_i(x)].\end{aligned} \tag{6}$$

Then, there exists $\Delta z \in I\!R^m$ such that

$$\nabla^2\psi(x,z,\varepsilon)\begin{pmatrix}\Delta x \\ \Delta z\end{pmatrix} = -\nabla\psi(x,z). \tag{7}$$

Proof. See Appendix.

Remarks.

(i) Unlike Theorem 2.1, in Theorem 2.2 we see that the Δz-part of the solution of (7) is uniquely determined. This is due to the regularizing perturbation. Defining, as before,

$$\nabla^2 \max\{0, g_i(x)\}^2 = \nabla^2[g_i(x)^2]$$

when $g_i(x) = 0$, the system (6) can be written as

$$\{\nabla^2[f(x) + \frac{1}{2}\sum_{i=1}^{m} \max\{0, g_i(x)\}^2] + \sum_{i=1}^{p} \frac{\varepsilon}{\varepsilon - 2g_i(x)} \nabla g_i(x) \nabla g_i(x)^T\}\Delta x =$$
$$-\nabla[f(x) + \frac{1}{2}\sum_{i=1}^{m} \max\{0, g_i(x)\}^2]$$

or, equivalently,

$$[\nabla^2\Phi(x) + \sum_{i=1}^{p} \frac{\varepsilon}{\varepsilon - 2g_i(x)} \nabla g_i(x) \nabla g_i(x)^T]\Delta x = -\nabla\Phi(x). \tag{8}$$

(ii) Observe that the perturbation related to g_i, $i \leq p$, of $\nabla^2\Phi(x)$ tends to $\nabla g_i(x)\nabla g_i(x)^T$ when $g_i(x) \to 0$ and tends to 0 as $g_i(x)$ tends to $-\infty$. The perturbation matrix is positive semidefinite, therefore it adds stability to the system. Finally the iteration (8) do not exhibit discontinuities on the boundaries $g_i(x) = 0$.

The reasoning above leads us to define the Regularized Hessian of Φ as

$$\nabla^2\Phi(x, \varepsilon) = \nabla^2\Phi(x) + \sum_{i=1}^{p} \frac{\varepsilon}{\varepsilon - 2g_i(x)} \nabla g_i(x) \nabla g_i(x)^T. \tag{9}$$

Since the perturbation is positive semidefinite, the perturbed Hessian is positive semidefinite provided that $\nabla^2\Phi(x)$ is. This is an advantage for minimization algorithms based on quadratic models.

For solving

$$\text{Minimize } \Phi(x) \text{ s.t. } x \in \Omega$$

iterative methods are used. At each iteration k a Hessian approximation B_k is usually needed. In this work, based on the considerations above, we will use the regularized Hessian approximations given by formula (9). Experiments confirm that this is more stable than merely using $\nabla^2\Phi(x)$.

2.3 Nonlinear optimizer

The method briefly described in this section deals with the minimization of a smooth function $\varphi : I\!R^n \to I\!R$ with bounds on the variables. The feasible set is defined by $\{x \in I\!R^n \mid \ell \leq x \leq u\}$ and the problem is:

$$\text{Minimize } \varphi(x) \text{ subject to } l \leq x \leq u. \tag{10}$$

GENCAN [4] is a recently introduced active-set method for smooth box-constrained minimization. For a description of basic techniques of continuous optimization and active-set methods see, for example, [9] and [16] (pp. 326–330). GENCAN adopts the leaving-face criterion of [3], that employs the spectral projected gradients defined in [5, 6]. For the internal-to-the-face minimization it uses a general algorithm with a line search that combines backtracking and extrapolation. In the present form, GENCAN uses, for the direction chosen at each step inside the faces, a truncated-Newton approach. This means that the search vector is an approximate minimizer of the quadratic approximation of the function in the current face. Conjugate gradients are used to find this direction. The method is fully described in [4] where extensive numerical experiments assess its reliability.

In this application, the truncated-Newton approach uses as Hessian approximations the regularized Hessian approximations defined in the previous section. GENCAN with the regularized Hessian turned out to be much more efficient than the same method using the discontinuous Hessian. Satisfactory results were obtaining with $\varepsilon = 0.01$.

2.4 Solving the decision problem

The method described in the previous subsection finds first-order stationary points (very likely, local minimizers) of problem (10). To enhance the probability of finding a global minimizer, the method is started from many randomly generated initial guesses. With probability 1, eventually, the method will start from an initial guess in the basin of convergence of a global minimizer. Nevertheless, this framework has two drawbacks: (i) we do not want to wait an infinite time; and (ii) unless we have an *a priori* knowledge of the optimal cost of problem (1) at the global minimizer, we will not be able to distinguish the global minimizer from other stationary points.

In practice, we run the method starting from N different initial guesses. If a solution with optimal cost equal to zero is found then the answer for the decision problem is *YES*, else the answer is "we do not know" and we assume it is *NO*. The fluxogram in Figure 1 shows this strategy.

3 Packing as many circles as possible

In the previous section we described a strategy to answer the question of whether is it possible packing k circles into the box or not. In order to pack as many circles as possible, we act as follows.

We start trying to pack just one circle and ask to the decision problem whether this is possible or not. While the answer is *YES* we try once more. If the answer is *NO*, we stop. The diagram of Figure 2 sketches this strategy. In the figure, n is the number of circles we were able to pack. Note that it is not necessary to start trying to pack just one circle if it is known that the answer for the decision problem with, say, $k = \bar{k}$ is *YES*. In such a case, it is enough to start from $\bar{k} + 1$.

4 Numerical experiments

All the experiments were run on a Sun SparcStation 20 with the following main characteristics: 128Mbytes of RAM, 70MHz, 204.7 mips, 44.4 Mflops. Codes are in Fortran77 and the compiler option adopted was "-O4".

Looking for a global minimizer of the decision problems (1) (see Figure 1) we set $N = 50,000$, which means that we run the local-minimization solver from 50,000 different initial guesses or until we find a global minimizer (detected by its null functional cost).

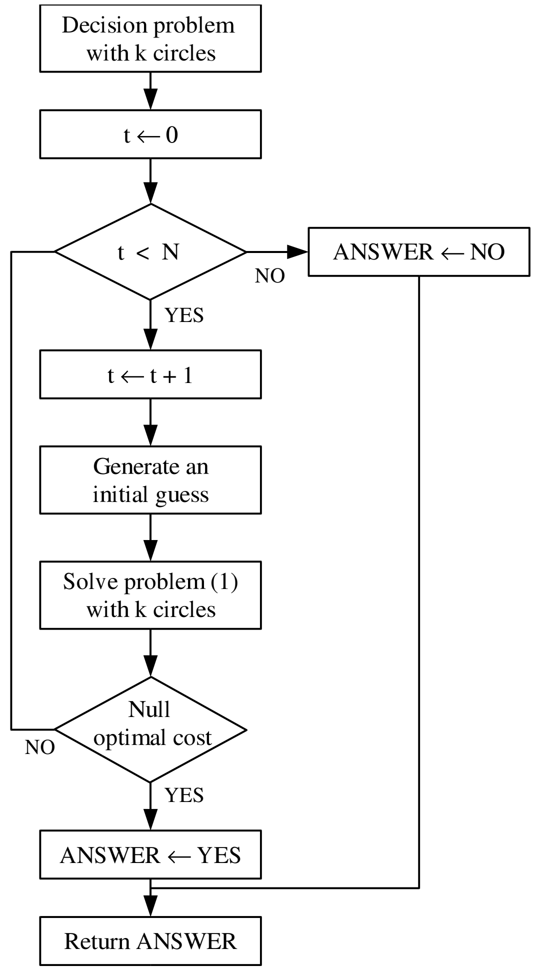


Figure 1: Solving the decision problem.

For each trial t (see Figure 1) the initial guess was generated as follows. Random points $p^i \in I\!R^2$ were generated inside the feasible region of problem (1). The random numbers where generated in the order $p_1^1, p_2^1, p_1^2, p_2^2, p_1^3, p_2^3, \ldots$ and the Schrage's random number generator [18] (double precision version) with seed t was used for a machine-independent generation of random numbers.

The first set of problems consists of 3 problems taken from [11]. The description of the problems, the number of packed cylinders presented in [11] and by the new method (called GENPACK from now on), and some figures which help to have an idea of the computational effort of GENPACK, are shown in Table 1. In the table, NDPS means "number of decision problems solved" and Time is the CPU time in seconds used to solve all these decision problems. The other columns are self-explanatory. Observe that, for the third problem, we were able to pack more circles than [11]. Figure 3 shows the solution. Note that no clear pattern is detected

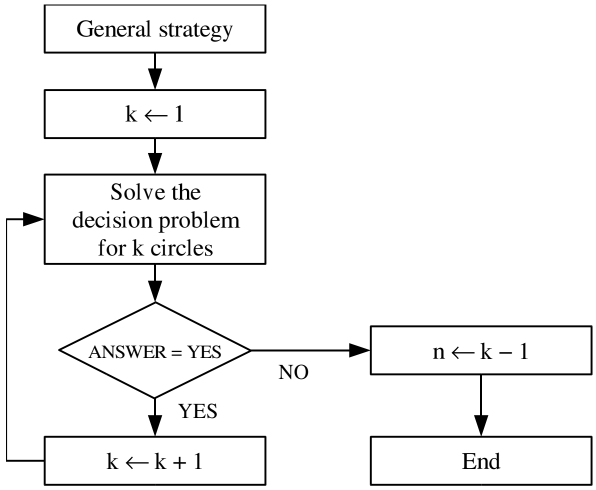


Figure 2: General scheme for the cylinder packing problem.

and that this kind of packing was not considered at all in [11].

Problem				Number of packed circles		GENPACK figures	
Name	Box dimensions		Circle radius	In [11]	By GENPACK	NDPS	Time
1	1200 ×	800	102	22	22	5	0.32
2	1200 ×	800	101	23	23	32	3.51
3	471 ×	196	14	124	**126**	289	1824.85

Table 1: Performance of GENPACK for the problems presented in [11].

The second set of problems was taken from [7, 8] and is described in Table 2. As mentioned in [7, 8], the test instances were generated such as to be representative of three possible relations between the rectangular box and the circle area: in problems 1.x, 2.x, 3.x, 4.x and 5.x, the circle area is approximately 1%, 2.5%, 5%, 10% and 15% of the rectangular box, respectively. In terms of the box shape, problems x.1, x.2, x.3 are long-rectangular, problems x.4, x.5, x.6 are short-rectangular and problems x.7, x.8, x.9 are squares.

As it can be shown in Table 2, many results for problems in subsets 1.x and 2.x were improved. Both subsets seems to be the harder ones, as mentioned in [7]. The reason for which no problems from subsets 3.x, 4.x and 5.x were improved is because, probably, since these are easier problems, the number of cylinders packed in [7] are the optimal ones. Figure 4 shows the packings found for the 6 improved problems of this second set.

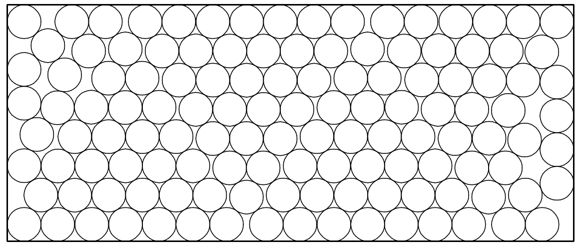


Figure 3: 126 circles of radius 14 packed in a box of dimension 471×196. In [11] a suboptimal solution with 124 circles was reported.

5 An extension: packing circles into circles

The methodology can be extended for packing identical circles into a circle of radius Δ, instead of a rectangular box. The decision problem (1) becomes

$$\begin{array}{ll} \text{Minimize} & \displaystyle\sum_{i \neq j} \max(0, (2r)^2 - \|p^i - p^j\|_2^2)^2 \\ \text{subject to} & \\ & \|p^i\|_2^2 \leq (\Delta - r)^2, \ \text{for } i = 1, \ldots, k. \end{array} \tag{11}$$

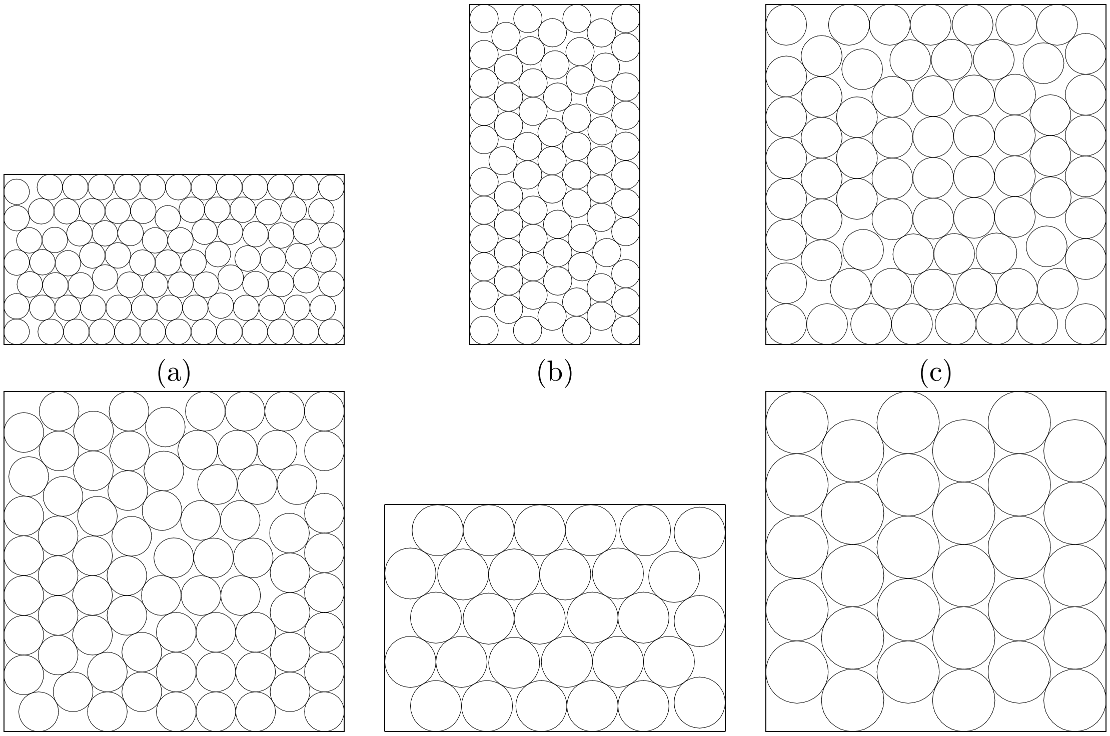
Problem (11) is not a bound-constrained problem, but a convex-constrained one. In such case, the bound-constrained solver GENCAN is not applicable any more and another solver, like SPG [5, 6] can be used. See also [1, 2] for other applications of SPG.

In [15] a similar problem is treated. The radio Δ of the circular box and the number l of circles that must be packed are fixed, and the problem is to determine the maximum radius of the identical circles to be packed. Optimal radius r_l^* for this problem with $l = 1, \ldots, 65$ and the *centers* of the l circles inside a circular box of unity radius are known (see, for example, [15]). For websites of this and related problems see [19, 17, 13] and the references therein.

If we consider circular boxes of fixed radius $1 + r_l^*$ and circular objects of radius r_l^*, we know that at least l circles can be packed, i.e., the answer of the decision problem (11) with $\Delta = 1 + r_l^*$, $r = r_l^*$, and $k = l$ is *YES*. In other words, a solution of the problem of maximizing the radius of the objects to be packed provides a lower bound for the problem of maximizing the number of packed objects. Both problems are not equivalent as 7 circles of radius r_6^* can be packed into a circular box of radius $1 + r_6^*$ (see [15]).

Table 3 shows the main characteristics of 3 problems taken from [15] and the performance of GENPACK (combined with SPG for solving the nonlinear decision subproblems (11)). Figure 5 illustrates the packings found.

The decision problems (1) and (11), for packing circles into rectangles and circles into circles, respectively, can also be extended to considerer k non-identical circles with radius $r^1, r^2, \ldots, r^k$



(a) (b) (c) (d) (e) (f)

Figure 4: Problems (a), (b), (c), (d), (e) and (f) correspond to problems 1.1, 1.3, 1.8, 1.9, 2.5 and 2.9 from [7], respectively. In problems 1.1, 1.3, 1.8, 1.9 and 2.9 one more circle was packed. In problem 2.5 two more circles were packed.

as follows:

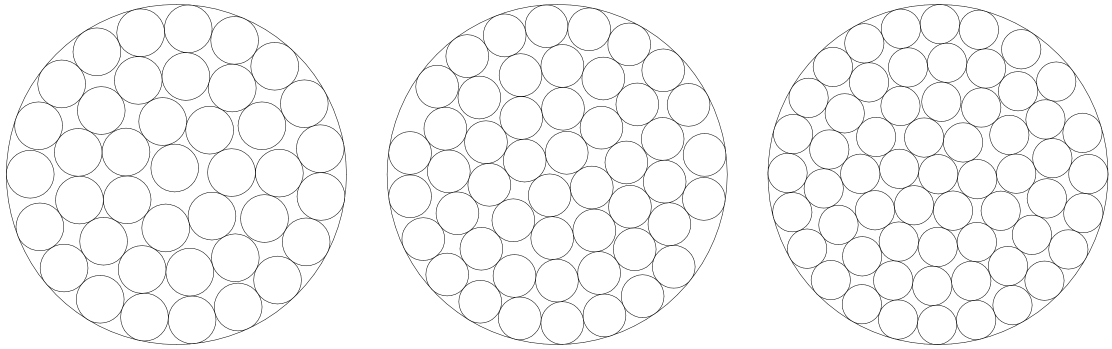
$$\begin{array}{ll} \text{Minimize} & \sum_{i \neq j} \max(0, (r^i + r^j)^2 - \|p^i - p^j\|_2^2)^2 \\ \text{subject to} & \\ & r^i \leq p_1^i \leq d_1 - r^i, \text{ and} \\ & r^i \leq p_2^i \leq d_2 - r^i, \text{ for } i = 1, \ldots, k; \end{array} \tag{12}$$

and

$$\begin{array}{ll} \text{Minimize} & \sum_{i \neq j} \max(0, (r^i + r^j)^2 - \|p^i - p^j\|_2^2)^2 \\ \text{subject to} & \\ & \|p^i\|_2^2 \leq (\Delta - r^i)^2, \text{ for } i = 1, \ldots, k. \end{array} \tag{13}$$

6 Final remarks

This work presented a methodology, based on a nonlinear decision problem, to solve the problem of packing identical circles into a rectangular box. The numerical results show that this is a



(a) (b) (c)

Figure 5: Some examples of packing circles into circles.

promising approach. It should be noted that, in principle, the problem is not concerned with the geometrical complexity of the solution. If an application imposes restrictions to the distribution of the circles, these constraints may be taken into account by means of an adequate nonlinear programming reformulation.

A common criticism against piecewise defined models in practical optimization is the lack of second derivatives in the boundary that separates the regions where the objective function "changes" its analytical definition. In this paper we proposed a way of overcoming this difficulty by means of a regularization of the Hessian. Unlike the original one, the regularized Hessian is continuous. This ensures more stability of the iterations.

The introduced model was also extended for packing identical circles into circles and non-identical circles into rectangles and circles. In the later cases, the scheme on Figure 1 for solving the decision problem is applicable. On the other hand, it is not clear which decision problems should be solved in order to maximize the used area. Some heuristics approach may be developed. This requires further research.

Acknowledgement

The authors are indebted to two anonymous referees whose comments helped a lot to improve this paper.

Problem				Number of packed circles		GENPACK figures	
Name	Box dimensions		Circle radius	In [7]	By GENPACK	NDPS	Time
1.1	160 × 80		6	90	**91**	61	734.05
1.2	100 × 200		8	84	84	4304	5791.83
1.3	120 × 240		10	73	**74**	2723	4065.91
1.4	100 × 80		5	86	86	27987	37108.39
1.5	120 × 80		6	68	68	16	23.35
1.6	120 × 100		6	87	87	864	2273.14
1.7	80 × 80		5	68	68	16794	12336.67
1.8	100 × 100		6	70	**71**	80	225.57
1.9	120 × 120		7	73	**74**	5	18.53
2.1	160 × 80		10	32	32	2	0.67
2.2	100 × 200		13	29	29	11	2.51
2.3	120 × 240		15	32	32	1	0.42
2.4	100 × 80		8	32	32	2016	232.45
2.5	120 × 80		9	28	**30**	25	3.12
2.6	120 × 100		10	30	30	12	1.42
2.7	80 × 80		7	32	32	10	7.91
2.8	100 × 100		9	30	30	1	0.18
2.9	120 × 120		11	29	**30**	33	11.21
3.1	160 × 80		14	15	15	2	0.02
3.2	100 × 200		18	15	15	26	0.41
3.3	120 × 240		21	15	15	1	0.01
3.4	100 × 80		11	16	16	30	0.75
3.5	120 × 80		12	15	15	1	0.01
3.6	120 × 100		14	14	14	17	0.30
3.7	80 × 80		10	16	16	2	0.03
3.8	100 × 100		13	13	13	1	0.02
3.9	120 × 120		15	16	16	2	0.04
4.1	160 × 80		20	8	8	1	0.04
4.2	100 × 200		25	8	8	1	0.11
4.3	120 × 240		30	8	8	1	0.11
4.4	100 × 80		16	6	6	1	0.00
4.5	120 × 80		17	7	7	1	0.00
4.6	120 × 100		20	6	6	1	0.01
4.7	80 × 80		14	6	6	1	0.00
4.8	100 × 100		18	6	6	1	0.00
4.9	120 × 120		21	6	6	1	0.00
5.1	160 × 80		25	3	3	1	0.00
5.2	100 × 200		31	3	3	1	0.00
5.3	120 × 240		37	3	3	1	0.00
5.4	100 × 80		19	4	4	1	0.00
5.5	120 × 80		21	4	4	1	0.00
5.6	120 × 100		24	4	4	1	0.00
5.7	80 × 80		17	4	4	1	0.00
5.8	100 × 100		22	4	4	1	0.00
5.9	120 × 120		26	4	4	1	0.00

Table 2: Performance of GENPACK for the problems presented in [7].

Problem			Number of packed circles		GENPACK figures	
Name	Circular box radius	Circles radius	In [14]	By GENPACK	NDPS	Time
1	1.1632960610	0.1632960610	40	40	1	5.23
2	1.1439363515	0.1439363515	50	50	332	456.11
3	1.1307835795	0.1307835795	60	60	155	363.62

Table 3: Performance of GENPACK for packing circles into circles.

References

[1] E. G. Birgin, R. Biloti, M. Tygel, L.T. Santos, Restricted optimization: a clue to a fast and accurate implementation of the Common Reflection Surface stack method, Journal of Applied Geophysics 42 (1999) 143–155.

[2] E. G. Birgin, I. Chambouleyron, J. M. Martínez, Estimation of the optical constants and the thickness of thin films using unconstrained optimization, Journal of Computational Physics 151 (1999) 862–880.

[3] E. G. Birgin, J. M. Martínez, A box constrained optimization algorithm with negative curvature directions and spectral projected gradients, Computing [Suppl] 15 (2001) 49–60.

[4] E. G. Birgin, J. M. Martínez, Large-scale active-set box-constrained optimization method with spectral projected gradients, Computational Optimization and Applications 23 (2002) 101–125.

[5] E. G. Birgin, J. M. Martínez, M. Raydan, Nonmonotone spectral projected gradient methods on convex sets, SIAM Journal on Optimization 10 (2000) 1196–1211.

[6] E. G. Birgin, J. M. Martínez, M. Raydan, SPG: Software for convex-constrained optimization, ACM Transactions on Mathematical Software 27 (2001) 340–349.

[7] M. H. Correia, J. F. Oliveira, J. S. Ferreira, Cylinder Packing by Simulated Annealing, Pesquisa Operacional 20 (2000) 269–284.

[8] M. H. Correia, J. F. Oliveira, J. S. Ferreira, A New Upper Bound for the Cylinder Packing Problem, International Transactions in Operational Research 8 (2001) 571–583.

[9] J. E. Dennis Jr., R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewoods Cliffs, 1983.

[10] K. A. Dowsland, Optimising the Palletisation of Cylinders in Cases, OR Spectrum 13 (1991) 204–212.

[11] H. Isermann, Heuristiken zur Lösung des zweidimensionalen Packproblem für Rundgefäße, OR Spectrum 54 (1991) 213–223.

[12] H. J. Fraser, J. A. George, Integrated Container Loading Software for Pulp and Paper Industry, European Journal of Operational Research 77 (1994) 466–474.

[13] E. Friedman, http://www.stetson.edu/˜efriedma/packing.html.

[14] J. A. George, J. M. George, B. W. Lamar, Packing Different-sized Circles into a Rectangular Container, European Journal of Operational Research 84 (1995) 693–712.

[15] R. L. Graham, B. D. Lubachevsky, K.J. Nurmela, P. R. J. Östergard, Dense Packing of Congruent Circles in a Circle, Discrete Mathematics 181 (1998) 139–154.

[16] D. G. Luenberger, Linear and Nonlinear Programming, Addison-Wesley, California, London, Amsterdam, Ontario, Sydney, 1984.

[17] R. Peikert, http://www.cg.inf.ethz.ch/˜peikert/personal/CirclePackings/.

[18] L. Schrage, A more portable Fortran random number generator, ACM Transactions on Mathematical Software 5 (1979) 132–138.

[19] P. G. Szabó, http://www.inf.u-szeged.hu/˜pszabo/Pack.html.

Appendix

Proof of Lemma 2.1. Assume that $x \in \Omega$. Define

$$z_i = \begin{cases} \sqrt{-g_i(x)}, & \text{if } g_i(x) \leq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then, $\Phi(x) = \psi(x, z)$. Conversely, assume that $x \in \Omega$, $z \in I\!R^m$. Replacing z by z' where

$$z_i' = \begin{cases} \sqrt{-g_i(x)}, & \text{if } g_i(x) \leq 0, \\ 0, & \text{otherwise,} \end{cases}$$

we obtain that

$$\Phi(x) = \psi(x, z') \leq \psi(x, z).$$

Therefore, the thesis is proved. □

Proof of Theorem 2.1. Let us analyze the set of solutions of the Newtonian linear system (5). Those solutions should satisfy

$$\nabla^2 f(x) + \sum_{i=1}^{m} \nabla g_i(x) \nabla g_i(x)^T + \sum_{i=p+1}^{m} g_i(x) \nabla^2 g_i(x)] \Delta x + 2 \sum_{i=1}^{p} \sqrt{-g_i(x)} \nabla g_i(x) \Delta z_i =$$
$$-[\nabla f(x) + \sum_{i=p+1}^{m} g_i(x) \nabla g_i(x)],$$

$$\sqrt{-g_i(x)} \nabla g_i(x)^T \Delta x - 2 g_i(x) \Delta z_i = 0, \; i = 1, \ldots, p,$$

and

$$g_i(x) \Delta z_i = 0, \; i = p+1, \ldots, m.$$

So, regrouping the first block and dividing the second block by $\sqrt{-g_i(x)} > 0$, we see that the solutions of (5) are the pairs $(\Delta x, \Delta z)$ that satisfy

$$\{\nabla^2 f(x) + \sum_{i=p+1}^{m} [\nabla g_i(x) \nabla g_i(x)^T + g_i(x) \nabla^2 g_i(x)]\} \Delta x + \sum_{i=1}^{p} \nabla g_i(x) [\nabla g_i(x)^T \Delta x + 2\sqrt{-g_i(x)} \Delta z_i] =$$
$$-[\nabla f(x) + \sum_{i=1}^{p+1} g_i(x) \nabla g_i(x)],$$

$$\nabla g_i(x)^T \Delta x + 2\sqrt{-g_i(x)}\Delta z_i = 0,\ i = 1, \dots, p,$$

and

$$g_i(x)\Delta z_i = 0,\ i = p+1, \dots, m.$$

So, the solutions of (5) are the pairs $(\Delta x, \Delta z)$ that satisfy

$$\{\nabla^2 f(x) + \sum_{i=p+1}^{m} [\nabla g_i(x)\nabla g_i(x)^T + g_i(x)\nabla^2 g_i(x)]\}\Delta x = -[\nabla f(x) + \sum_{i=1}^{p+1} g_i(x)\nabla g_i(x)], \tag{14}$$

$$\nabla g_i(x)^T \Delta x + 2\sqrt{-g_i(x)}\Delta z_i = 0,\ i = 1, \dots, p, \tag{15}$$

and

$$g_i(x)\Delta z_i = 0,\ i = p+1, \dots, m. \tag{16}$$

Since $\overline{\Delta x}$ satisfies the first block of the above equations, it turns out that the theorem is proved if we are able to compute the solution with Δz satisfying the second and the third block. This can be trivially done, for example, defining

$$\overline{\Delta z}_i = \frac{-\nabla g_i(x)^T \overline{\Delta x}}{2\sqrt{-g_i(x)}},\ i = 1, \dots, p,$$

$$\Delta z_i = 0, i = p+1, \dots, m.$$

So, the theorem is proved. □

Remark. Observe that the choice $\Delta z_i = 0$ for $i > p$ is not unique in the case that $g_i(x) = 0$. This is the case in which the Hessian of $\psi(x, z)$ is singular. So, the theorem proves that the existence of the Newtonian direction for ψ depends only on the existence of a solution of the first block of equations (14). When the Newtonian system of ψ is singular and $\overline{\Delta x}$ exists, there are infinitely many solutions for each i such that $g_i(x) = 0$, and these solutions are due to the freedom in the choice of Δz_i.

Proof of Theorem 2.2. As in Theorem 2.1, let us analyze the set of solutions of system (7). Those solutions should satisfy

$$[\nabla^2 f(x) + \sum_{i=1}^{m} \nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^{m} g_i(x)\nabla^2 g_i(x)]\Delta x + 2\sum_{i=1}^{p} \sqrt{-g_i(x)}\nabla g_i(x)\Delta z_i =$$
$$-[\nabla f(x) + \sum_{i=p+1}^{m} g_i(x)\nabla g_i(x)],$$

$$\sqrt{-g_i(x)}\nabla g_i(x)^T \Delta x + [\varepsilon - 2g_i(x)]\Delta z_i = 0,\ i = 1, \dots, p,$$
$$[g_i(x) + \varepsilon]\Delta z_i = 0,\ i = p+1, \dots, m.$$

Therefore, $\Delta z_i = 0$ for all $i = p+1, \dots, m$. Moreover, the second block of equations is equivalent to

$$\Delta z_i = \frac{\sqrt{-g_i(x)}\nabla g_i(x)^T \Delta x}{2g_i(x) - \varepsilon},\ i = 1, \dots, p.$$

Replacing Δz_i in the first block, we obtain:

$$[\nabla^2 f(x) + \sum_{i=1}^{m} \nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^{m} g_i(x)\nabla^2 g_i(x)]\Delta x + 2\sum_{i=1}^{p} \frac{(-g_i(x)\nabla g_i(x)\nabla g_i(x)^T)\Delta x}{2g_i(x) - \varepsilon} =$$
$$-[\nabla f(x) + \sum_{i=p+1}^{m} g_i(x)\nabla g_i(x)].$$

Therefore,

$$\{\nabla^2 f(x) + \sum_{i=1}^{p}[1 - \frac{2g_i(x)}{2g_i(x) - \varepsilon}]\nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^{m} \nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^{m} g_i(x)\nabla^2 g_i(x)\}\Delta x =$$
$$-[\nabla f(x) + \sum_{i=p+1}^{m} g_i(x)\nabla g_i(x)].$$

But

$$1 - \frac{2g_i(x)}{2g_i(x) - \varepsilon} = \frac{\varepsilon}{\varepsilon - 2g_i(x)},$$

therefore this block of equations is equivalent to

$$[\nabla^2 f(x) + \sum_{i=1}^{p} \frac{\varepsilon}{\varepsilon - 2g_i(x)}\nabla g_i(x)\nabla g_i(x)^T + \sum_{i=p+1}^{m} \nabla g_i(x)\nabla g_i(x)^T + g_i(x)\nabla^2 g_i(x)]\Delta x =$$
$$-[\nabla f(x) + \sum_{i=p+1}^{m} g_i(x)\nabla g_i(x)].$$

This completes the proof. $\square$