Out-of-the-Box Global Optimization for Packing Problems: New Models and Improved Solutions

Timo Berthold[1,2], Dominik Kamp[3], Gioni Mexi[3], Sebastian Pokutta[2,3], Imre Pólik[4]

[1]Fair Isaac Deutschland GmbH, Germany.
[2]Institut für Mathematik, Technische Universität Berlin, Germany.
[3]Zuse Institute Berlin, Germany.
[4]Fair Isaac Corporation, USA.

Contributing authors: timoberthold@fico.com; kamp@zib.de; mexi@zib.de; pokutta@zib.de; imre@polik.net;

Abstract

Recent discoveries driven by Large Language Models (LLMs) have renewed interest in geometric packing problems. In this paper, we study several classes of such packing problems through the lens of modern global nonlinear optimization. Starting from direct nonlinear formulations, we consider packing circles in squares and fixed-perimeter rectangles, packing circles into minimum-area ellipses, packing regular polygons into regular polygons, and packing Platonic solids into Platonic solids. For ellipse packing, we derive a novel containment formulation based on the S-lemma. For polygon and Platonic-solid packing, we develop compact non-overlap formulations based on the Farkas lemma and study several natural modeling variants computationally.

Using off-the-shelf global optimization solvers, namely FICO Xpress and SCIP, we obtain numerous new incumbent solutions as well as first solutions for previously unstudied variants without any problem-specific solver modifications beyond writing down the models. Our computational results analyze the impact of various formulation choices.

Beyond the individual packing results, the paper illustrates a broader point. It provides further evidence that global nonlinear optimization has matured into an increasingly practical model-and-solve technology even for highly nonconvex problems.

1 Introduction

1.1 Motivation

The rapid progress in global optimization technology over the past decade has substantially expanded the range of nonlinear, nonconvex problems that can be solved reliably by general-purpose optimization software. State-of-the-art academic solvers such as SCIP [1] and commercial solvers such as FICO® Xpress [2], Gurobi or Hexaly combine spatial branch-and-bound, automatic linearization and convexification, sophisticated presolving and propagation, and increasingly powerful primal heuristics [3–5]. As a consequence, classes of nonlinear optimization problems that would only recently have been considered computationally far beyond the reach of generic solvers can now often be solved to proven global optimality, or at least to very high-quality primal solutions accompanied by meaningful dual bounds.

At the same time, recent developments in algorithm design based on Large Language Models (LLMs) drew attention to long-standing geometric and combinatorial problems that admit natural nonlinear optimization formulations. In particular, Google DeepMind presented the AlphaEvolve framework [6, 7], which uses LLM-generated code in an evolutionary search to produce high-quality solutions for an extensive set of mathematical problems, including variants of circle packing and hexagon packing. These developments raise a natural question: to what extent can state-of-the-art global optimization solvers match, or even surpass, such discovered algorithms when the underlying problems are formulated directly as optimization models and solved with off-the-shelf nonlinear optimization technology?[1]

Packing problems provide a particularly appealing problem class for studying this question. They have a long history in mathematics and optimization, with a track record of best-known solutions [12], and they feature highly nonconvex feasible regions. At the same time, they admit relatively intuitive nonlinear formulations, making them well suited for studying the interaction between modeling choices and solver performance. In this sense, packing problems are not only interesting in their own right, but also serve as a neat showcase for understanding the current capabilities and limitations of modern global nonlinear optimization.

1.2 Contributions

In this paper, we revisit several families of geometric packing problems from a "model-and-solve" perspective. The modeling effort differs between the problem classes. We begin with straightforward nonlinear formulations for packing circles in squares and fixed-perimeter rectangles, using only the textbook definitions of containment and non-overlap and requiring no expert knowledge. For regular polygon packing, we develop a compact non-overlap formulation based on the Farkas lemma [13]. Deriving it requires some background in optimization, but should be accessible to anyone with a graduate education in the field, and related ideas exist in the literature [14–16]. We straightforwardly extend this modeling idea to the three-dimensional setting of packing Platonic

[1] Curiously, in a study with OpenEvolve [8], a project aimed at replicating the AlphaEvolve project using open-source tools, the LLM converged to using an optimization approach [9], namely a Sequential Least Squares Programming formulation with a local solver implemented in SciPy [10, 11].

solids into Platonic solids. To the best of our knowledge, this yields the first systematic optimization study of that broader problem class beyond the special case of packing cubes into cubes. For packing circles into ellipses, we derive a novel containment model based on the S-lemma [17]. This step requires the most expertise, and we regard it as the paper's main modeling contribution. Across all classes, we did not develop any problem-specific solution algorithm or heuristic. The main requirement is formulating the problem mathematically, i.e., as an optimization model; the solution search is left to the solver.

More specifically, our contributions are threefold:

1. On the modeling side, we introduce new formulations for ellipse, polygon and Platonic-solid packing, including a practical use of the S-lemma for circle-in-ellipse containment and a Farkas-based separation framework for polygonal and polyhedral non-overlap constraints.
2. On the computational side, we obtain numerous new incumbent solutions, including for the well-studied topics of circle and cube packing, as well as first solutions for variants that appear not to have been studied before. Notably, all computational results are obtained using off-the-shelf global optimization solvers, without any problem-specific solver modifications or specialized heuristics.
3. On the methodological side, we show that formulation choices matter. For packing polygons and Platonic solids, we compare several natural model variants and observe that redundant strengthening constraints can have a significant positive computational impact, while seemingly reasonable symmetry-breaking constraints may, in fact, degrade performance.

Parts of this paper build on our earlier study in [18]. In particular, the circle-packing results in Section 4 and a subset of the polygon-packing results in Section 6 were already reported there. The present paper goes substantially beyond that earlier work by broadening the range of packing classes considered, introducing new modeling ideas for ellipses and higher-dimensional polyhedral packing, and adding a systematic empirical study of formulation variants.

The central message of this paper is not restricted to packing. Rather, our computational study provides further evidence that global nonlinear optimization is increasingly becoming an out-of-the-box technology: transparent mathematical models, combined with modern general-purpose solvers, produce highly competitive results on difficult nonconvex benchmark problems that have recently attracted attention in the context of LLM-driven discovery. See, e.g., [19] for an excellent recent study on the Heilbronn problem.

1.3 Organization of the Paper

The remainder of the paper is organized as follows. Section 2 reviews relevant background on packing problems and global nonlinear optimization. Section 3 discusses implementation aspects, including solution polishing and the computational setup. Section 4 considers packing circles with variable radii into squares and fixed-perimeter rectangles. Section 5 introduces the ellipse-packing model based on the S-lemma.

Section 6 presents the generalized polygon-packing formulation including several modeling variants and compares those computationally. Section 7 extends this model to Platonic-solid packing in three dimensions. The model can be easily extended to higher dimensions, but we restrict our computational analysis and visualizations to the three-dimensional case. Finally, Section 9 concludes with lessons learned, limitations, and directions for future work.

2 Background

This section provides an overview and the necessary background on geometric packing problems and global nonlinear optimization. The interaction between these two areas is central to this work, as packing models expose nonconvex structure that is amenable to modern global nonlinear solvers.

2.1 Packing Problems

Packing problems ask for the placement of a set of objects into a container subject to feasibility constraints (typically non-overlap and containment), while optimizing an objective such as minimizing container size, minimizing the number of bins, or maximizing packed volume. They have a long history in mathematics. The study of sphere packing dates back at least to Kepler's 1611 conjecture that a face-centered cubic arrangement yields the densest packing of equal spheres in three-dimensional space, a claim that was proven nearly four centuries later by Hales [20]. The study of packing circles in polygonal regions dates back almost 200 years [21]. Polygon packing has been studied at least since the 1950s; see [22] for an early overview. Today, the study of packing spans classical geometry (e.g., densest sphere/circle packings) and modern operations research (bin packing, strip packing, nesting). There are systematic problem typologies clarifying structural variants and modeling assumptions, enabling principled algorithm design and benchmarking; see the widely used classification in [23].

In parallel, the community established benchmark repositories to track progress and to compare methods across heterogeneous variants, including early integrated libraries such as PackLib2 [24] for multi-dimensional packing problems and 2DPackLib for two-dimensional packing [25]. Many packing and distance-constrained geometry models can be written using Euclidean norm expressions (e.g., $\|x_i - x_j\|_2 \geq r_i + r_j$ for non-overlap), yielding nonconvex nonlinear constraints with substantial symmetry. EuclidLib [26] was introduced to capture such nonconvex, Euclidean-norm-driven optimization instances across various applications, and has become a commonly used benchmarking set for global optimization research. Some of its instances have also been included in MINLPLib [27]. For a comprehensive overview of Euclidean Distance Geometry, see the survey by Liberti et al. [28].

For circle packing in particular, deterministic global optimization has long been discussed as a viable route to provable solutions and strong bounds, with comprehensive early computational evidence and application discussion in [29, 30]. The Eternity puzzle [31] is a packing problem in which triangle-based nonconvex polygon pieces must be arranged to tile a prescribed board. It drew a lot of public attention because

a £1 million prize was offered (and later awarded to two mathematicians) for a complete solution. Packing models arise throughout manufacturing and logistics: cutting stock and trim-loss reduction (paper, metal, wood, glass, textiles), palletization and container loading, warehouse storage and automated fulfillment, and layout problems in integrated circuit design. From an optimization standpoint, these applications frequently impose additional constraints beyond pure non-overlap (e.g., grouping, balancing, stability, incompatibilities), which induce mixed-integer structure even when the geometry is continuous [23].

Recent developments in solving packing problems include the work of Khajavirad [32], which provides a theoretical comparison of convexification techniques, in particular single-row linear optimization (LP) relaxations, multi-row LP relaxations, and semidefinite optimization (SDP) relaxations, for non-overlap constraints for circles and spheres modeled as nonconvex quadratically constrained quadratic problems (QCQPs), and their implications for bound quality and algorithm design. Google DeepMind introduced the AlphaEvolve framework [6, 7], which synthesizes code through LLMs to obtain new best-known solutions for a broad range of mathematical optimization tasks, including circle and hexagon packings as well as minimum-distance point configurations. Shortly after, in work that preceded the present article, we improved some of the AlphaEvolve results by using global nonlinear optimization techniques [18, 33]. The field is currently very active, with at least six groups contributing new best-known solutions to the benchmark collections curated and continuously updated at [12]. Some of these solutions have not (yet) been published elsewhere. Where no publication exists, we refer to a solution by the name of its contributors and the year listed at [12]; where an accompanying publication exists, we cite it instead.

2.2 Nonlinear Optimization

Global nonlinear optimization concerns optimization problems with nonlinear, often nonconvex objectives and constraints, where the goal is to compute solutions that are provably globally optimal. Formally, a *nonlinear optimization problem (NLP)* is defined as:

$$\begin{aligned} \min\ & f(x) \\ & g_k(x) \le 0,\ k = 1, \dots, m \\ & \ell \le x \le u, \end{aligned} \tag{2.1}$$

where $x \in \mathbb{R}^n$, $f, g_k : \mathbb{R}^n \to \mathbb{R}$ are functions built from a list of elementary functions[2] using the four arithmetic operations and composition. The variable bounds ℓ, u are in $\bar{\mathbb{R}}^n := (\mathbb{R} \cup \{\pm\infty\})^n$. The situation becomes even more challenging in the presence of integrality constraints, leading to *mixed-integer nonlinear optimization (MINLP)*.

From a complexity-theoretic perspective, NLP is fundamentally hard: continuous nonconvex quadratic optimization is NP-hard [34], and general nonlinear optimization is undecidable [35]. Algorithmically, modern global solvers extend branch-and-bound

[2]The list usually includes powers, roots, exponentials, logarithms and trigonometric functions. In addition, some solvers offer direct support for the absolute value, minimum, maximum or the sigmoid function.

and branch-and-cut frameworks from mixed-integer linear programming to nonlinear settings, a central ingredient being the construction of tight convex relaxations. The seminal work of McCormick introduced convex under- and overestimators for bilinear functions [36], forming the foundation of spatial branch-and-bound. The Reformulation-Linearization Technique (RLT) [37] and semidefinite relaxations [38] represent further systematic convexification approaches.

These methodological advances, as well as the comprehensive work on other branch-and-bound sub-algorithms such as domain propagation [39, 40], branching rules [39, 41, 42], primal heuristics [3, 5], and conflict analysis [43], are reflected in today's solver ecosystem; see [44] for a computational study. Dedicated global solvers such as BARON [45, 46] and ANTIGONE [47], and open-source frameworks such as SCIP [1, 48], implement spatial branch-and-bound and use external linear (LP), mixed-integer (MIP), and/or local nonlinear (NLP) solvers. Those works have paved the way for a second wave of global solvers in which existing commercial MIP solvers have been extended to incorporate global optimization capabilities for general NLPs and MINLPs. This was championed by FICO Xpress in 2022 [2], followed by Gurobi and Hexaly in 2023 and 2024, respectively.

3 Implementation Details

Each of the following sections will have its own set of computational results. Here we describe the common computational setup. In the paper we will list only objective values and will show a few representative or interesting solutions. There is an online supplement to the paper at https://github.com/DominikKamp/Packing, including all the models, numerical solutions and visualizations.

3.1 Solvers

We used SCIP 10.0.0 and FICO Xpress 9.8.0 to run all of our experiments. These represent the state-of-the-art in open-source and commercial global optimization, respectively. We will not report comparisons between the two solvers, as that is not the point of the study.

3.2 Multistart vs. Global Optimization

Although for most problems we could just run a global optimization solve, it was often better to first run a local optimization solve started from random starting points, and then input this solution into the global solve. Our typical solves were with a time limit of 10 000s preceded by a 5 000s multistart ramp-up, keeping the best solution across all runs. For the formulation comparisons, reduced time limits of 1 000s preceded by a 500s multistart ramp-up were used. This multistart phase is supported natively by both solvers used in this study.

3.3 Numerical Thresholds and Solution Polishing

Nonlinear optimization solvers produce solutions up to a given relative feasibility tolerance, typically 10^{-6}, meaning constraints may be violated slightly. Also, nonlinear

solvers can show a tendency to exploit feasibility tolerances to improve the objective value. Some of the improvements we have achieved are comparable to this value, so we have to be careful about feasibility. We therefore require that all reported solutions are feasible in double-precision arithmetic: every constraint, evaluated in IEEE 754 double precision, is satisfied with zero tolerance. Note that this is a statement about floating-point evaluation and not a proof of feasibility in exact real arithmetic. For the problems from the AlphaEvolve study, we use the validation code published there. For all other problems, we use our own checker, available in the online supplement.

To achieve this, we employed a three-level approach. First, all optimization solves used a tighter feasibility tolerance of $\varepsilon = 10^{-8}$. Second, we modified the constraints to make them tighter, so that the original constraints stay feasible even if the modified constraints are violated by up to ε. Third, we slightly adjust the solutions returned by the solvers, using arbitrary-precision arithmetic, to remove the remaining margins such that the double-precision check above passes. The exact nature of these changes depends on the problem and will be discussed later.

It is important to emphasize that the only reason we are doing this is to make sure that our claimed improvements are fully correct.

3.4 Computational Setup

Experiments were run single-threaded and non-exclusively on a 48-core Intel Xeon Gold 6342 CPU with 1,024 GB RAM. Models were generated in nl file format using PySCIPOpt and solved with SCIP 10.0.0 and FICO Xpress 9.8.0. For the best results reported in the online supplement we applied five solver setups: SCIP with tolerance 10^{-8}, Xpress seeds 1–3 with tolerance 10^{-8}, and Xpress with tolerance 10^{-6}.

4 Packing Circles inside a Square or a Rectangle

Among circle-packing problems, one of the most studied variants concerns packing a fixed number of unit circles into the smallest possible square; a good survey is given by Peikert [49]. For most instances of up to twenty circles, exact optimal packings are known and can be derived using algebraic geometry techniques based on polynomial systems and Gröbner basis computations; see, e.g., [50].

In this section, we investigate a variant for which it is harder to prove optimality since both the positions and radii of the circles are decision variables. Specifically, we consider packing circles of variable radii into a unit square, or, in a mild relaxation, into a rectangle of perimeter four, while maximizing the sum of their radii. Best-known solutions are mostly due to Cantrell [51], but optimality proofs are only known for some trivial cases.

4.1 Optimization Model

A key advantage of mathematical optimization approaches is that the optimization model can be changed easily. This facilitates exploring related problems by simply adding or modifying constraints, as illustrated by the two problem variants below.

Given a positive integer n, let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote the set of circles. We define the following decision variables:

$(x_i, y_i) \in \mathbb{R}^2$: coordinates of the center of circle $i \in \mathcal{N}$
$r_i \in \mathbb{R}_+$: radius of circle $i \in \mathcal{N}$
$\alpha \in \mathbb{R}_+$: width of the rectangle

We consider the problem of packing circles in a rectangle with fixed perimeter $P = 4$ and variable side lengths. We introduce a decision variable α representing the width and set the height to $H = \frac{P}{2} - \alpha = 2 - \alpha$. We can assume without loss of generality that $\alpha \leq 1$ (i.e., α is the shorter side). The optimization problem can be formulated as:

$$\max \sum_{i \in \mathcal{N}} r_i \tag{4.1a}$$

$$\text{s.t.} \quad r_i \leq x_i \leq \alpha - r_i, \qquad \text{for all } i \in \mathcal{N} \tag{4.1b}$$

$$r_i \leq y_i \leq (2 - \alpha) - r_i, \quad \text{for all } i \in \mathcal{N} \tag{4.1c}$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 \geq (r_i + r_j)^2, \qquad \text{for all } i, j \in \mathcal{N},\ i < j \tag{4.1d}$$

$$0 \leq r_i \leq \tfrac{\alpha}{2}, \qquad \text{for all } i \in \mathcal{N} \tag{4.1e}$$

$$0 < \alpha \leq 1 \tag{4.1f}$$

The objective (4.1a) is to maximize the sum of all radii. Note that this is fundamentally different from maximizing the total area covered by circles, which would be $\pi \sum_{i \in \mathcal{N}} r_i^2$; the linear objective instead tends to favor more balanced distributions of circle sizes. Constraints (4.1b) and (4.1c) ensure that each circle remains entirely within the rectangle: the center coordinates must maintain a distance of at least r_i from all rectangle boundaries. Constraints (4.1d) ensure that no two circles overlap by requiring that the Euclidean distance[3] between any pair of circle centers is at least the sum of their radii. Constraints (4.1e) provide an upper bound on each radius: no circle can have a diameter exceeding the width of the rectangle α (which is the shorter side by assumption). Finally, constraint (4.1f) bounds the width variable.

The aspect ratio determined by α is a decision variable that can be modified to maximize the sum of radii for a given number of circles. We can trivially change this formulation to packing into a unit square by fixing $\alpha = 1$. This is a crucial property of mathematical optimization modeling: the user needs to change only the model and does not have to worry about whether or how this changes the algorithm. The solvers will take care of that. This contrasts with many heuristic approaches, including LLM-generated ones, in which often a new set of heuristics needs to be developed once the model formulation changes or a new constraint is added.

The formulation has only linear and nonconvex quadratic constraints. For larger values of n the quadratic growth in the number of constraints (4.1d) will become dominant.

[3] It is advantageous to work with squared distances to avoid square roots in the formulation.

4.2 Dual Bounds

Unfortunately, model (4.1) is very weak in the sense that the dual bounds the solvers obtain for it are far from the optimal objective value. We can do better by adding a valid inequality to the problem.

Using the arithmetic-quadratic mean inequality and the fact that the sum of the areas of the circles is less than the area of the enclosing rectangle, we have

$$\sum_{i\in\mathcal{N}} r_i \le \sqrt{n}\sqrt{\sum_{i\in\mathcal{N}} r_i^2} \le \sqrt{n}\sqrt{\frac{\alpha(2-\alpha)}{\pi}} \le \sqrt{\frac{n}{\pi}}. \tag{4.2}$$

In contrast, the solvers typically start at a dual bound of $\frac{n}{2}$, corresponding to setting each r_i to $\frac{1}{2}$, and improve gradually from there. This is why we decided to add the area inequality

$$\sum_{i\in\mathcal{N}} r_i^2 \le \frac{1}{\pi} \tag{4.3}$$

to model (4.1) for all of our computational tests, and added a similar area inequality also to the other problems in this paper. This constraint not only helps to tighten the dual bound, but it also helps in finding good feasible solutions. This is because it is a constraint involving all the radii of the circles, and it is the only constraint dealing with more than two circles in the packing.

4.3 Computational Results

First, let us focus on the restricted problem, packing circles into a square. The improving solutions[4] (floored to five decimal digits) are reported in Table 1 and visualized in Figure 1. It is instructive to compare this objective value to the bound we get from inequality (4.2). For the $n = 32$ case this yields an upper bound of $\sqrt{\frac{32}{\pi}} \approx 3.191538\ldots$, which is within 8.57% of the solution. The relative tightness of this bound comes from two factors. The circles in the solutions have similar radii, so the arithmetic-quadratic mean inequality does not sacrifice much. Further, the total area of the circles is close to 1, so the area constraint is also fairly tight.

Table 1: Improving solutions for the circle-packing problem

Variant	n	Sum of radii	Previous best	Source
square	32	2.93**957**	2.93794	[6, 7]
rectangle	26	2.63**930**	2.638	[12]
rectangle	27	2.6**9015**	2.687	[12]

[4]We have found possibly even more improving solutions, but the original source [51] for the best-known packings lists only three digits after the decimal point. In almost all cases, we managed to match the best-known solutions from the literature.

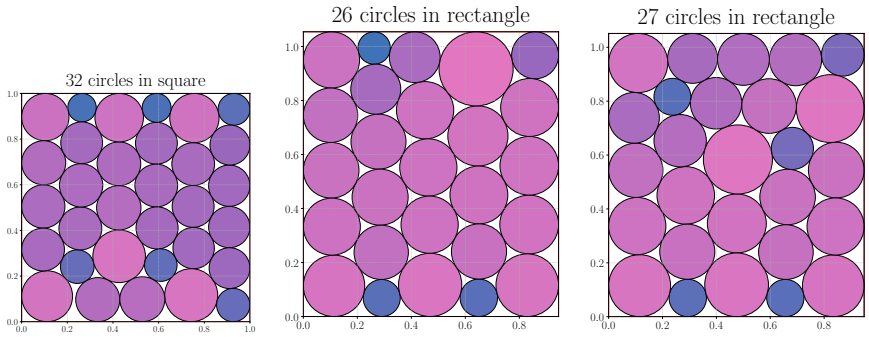


Fig. 1: Graphical representation of the new solutions for the circle-packing problem. Left: square variant ($n = 32$). Middle and right: rectangle variant ($n = 26, n = 27$). See Table 1 for details.

Now we can turn to the relaxed version, where instead of a square we are trying to pack into a rectangle of perimeter four. Obviously, all the previous solutions are still feasible for this relaxed problem, but we can do slightly better. We found the following improving solutions, again floored to five decimal digits (see Table 1). The solutions are visualized in Figure 1. Not surprisingly, the solutions all tend to use a rectangle that is almost square.

To ensure that the solutions are indeed as feasible as we can make them, we scaled the radii in the containment and overlap constraints. The circle centers are kept fixed and the radii are adjusted in two phases. First, the maximum uniform absolute radius increase δ that preserves all non-overlap and boundary constraints is computed. Then, each circle's radius is individually increased in order of increasing current radius, each time consuming the slack of its tightest remaining constraint. An absolute safety margin of 10^{-15} is applied to account for numerical rounding errors. These calculations are carried out in arbitrary-precision arithmetic.

We found it quite surprising that new solutions could be found for such a well-studied class of problems with off-the-shelf tools. In fact, it was this class of problems that we first tried from the original AlphaEvolve paper [7]. The encouraging initial results led to this much larger project.

5 Packing Circles into an Ellipse

In this section we present another interesting packing problem, where only very little progress has been made in 20 years. The problem is packing unit circles in an ellipse of minimum area. Known solutions are collected on Erich Friedman's packing website [12], the last reported solution being from 2007.

We will first present the optimization model, then see it in action. We will discuss some potential valid inequalities that can be used to strengthen the model.

5.1 Characterizing the Circle-Ellipse Containment

The main difficulty of the optimization model is to formalize mathematically that a circle is contained in an ellipse. For this we will use the S-lemma [52]:

Theorem 1 (S-lemma, Yakubovich, 1971) *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be (not necessarily homogeneous) quadratic functions and suppose that there is an $\bar{x} \in \mathbb{R}^n$ such that $g(\bar{x}) < 0$. Then the following two statements are equivalent.*

1. *There is no $x \in \mathbb{R}^n$ such that*

$$f(x) < 0, \tag{5.1a}$$
$$g(x) \leq 0. \tag{5.1b}$$

2. *There is a nonnegative multiplier $\lambda \geq 0$ such that*

$$f(x) + \lambda g(x) \geq 0, \, \forall x \in \mathbb{R}^n. \tag{5.2}$$

One could view this result as two quadratic functions seemingly behaving as if they were convex, as far as Lagrange duality is concerned. See [17] for a survey on the S-lemma and its connections to other results.

Now consider an ellipse centered at the origin, with semiaxes a and b, and a circle of radius r centered at (x_0, y_0). The circle is contained in the ellipse if and only if

$$\nexists x, y$$
$$(x - x_0)^2 + (y - y_0)^2 \leq r^2 \tag{5.3a}$$
$$b^2x^2 + a^2y^2 > a^2b^2, \tag{5.3b}$$

where the second inequality is an alternative form of the ellipse equation, avoiding the divisions. This can be written equivalently using the S-lemma:

$$\exists t \geq 0 : \forall x, y \in \mathbb{R} : a^2b^2 - b^2x^2 - a^2y^2 + t\left((x - x_0)^2 + (y - y_0)^2 - r^2\right) \geq 0. \tag{5.4}$$

This is a non-homogeneous quadratic expression without cross terms. The following lemma summarizes when this expression is nonnegative:

Lemma 2 The quadratic expression $Ax^2 + 2Bx + Cy^2 + 2Dy + E$ is nonnegative for all x, y if and only if

$$A \geq 0 \tag{5.5a}$$
$$C \geq 0 \tag{5.5b}$$
$$E \geq 0 \tag{5.5c}$$
$$AE - B^2 \geq 0 \tag{5.5d}$$
$$CE - D^2 \geq 0 \tag{5.5e}$$

$$AEC - AD^2 - CB^2 \geq 0. \tag{5.5f}$$

Proof First, a standard homogenization argument shows that $Ax^2 + 2Bx + Cy^2 + 2Dy + E$ is nonnegative for all x and y if and only if the homogeneous expression $Ax^2 + 2Bxu + Cy^2 + 2Dyu + Eu^2$ is nonnegative for all x, y and u. Since this is a homogeneous quadratic form, its nonnegativity is equivalent to the following matrix being positive semidefinite:

$$\begin{bmatrix} A & B & 0 \\ B & E & D \\ 0 & D & C \end{bmatrix}. \tag{5.6}$$

This, on the other hand, is easy to characterize by setting all principal minors to be nonnegative, yielding exactly the six inequalities[5] in the lemma. Notice that the formulas are greatly simplified because of the 0 appearing in the matrix, due to the lack of an xy term. □

We can apply this result to the expression in (5.4) to get a condition in terms of the problem data.

$$A = t - b^2 \tag{5.7a}$$
$$B = -tx_0 \tag{5.7b}$$
$$C = t - a^2 \tag{5.7c}$$
$$D = -ty_0 \tag{5.7d}$$
$$E = a^2b^2 + t\left(x_0^2 + y_0^2 - r^2\right) \tag{5.7e}$$

Now we are ready to write the complete optimization problem. We place the ellipse centered at the origin, with semiaxes a and b, which are the main decision variables. The n circles are located at $(x_i, y_i), i = 1, \ldots, n$. For each circle i we will introduce a multiplier t_i, which will certify that circle i is contained in the ellipse. The model is

$$\min \ a \cdot b \cdot \pi \tag{5.8a}$$
$$\text{s.t.} \quad (x_i - x_j)^2 + (y_i - y_j)^2 \geq 4, \quad \text{for all } i < j \tag{5.8b}$$
$$t_i - a^2 \geq 0, \quad \text{for all } i \tag{5.8c}$$
$$t_i - b^2 \geq 0, \quad \text{for all } i \tag{5.8d}$$
$$a^2b^2 + t_i(x_i^2 + y_i^2 - 1) \geq 0, \quad \text{for all } i \tag{5.8e}$$
$$(t_i - a^2)(a^2b^2 + t_i(x_i^2 + y_i^2 - 1)) - t_i^2 y_i^2 \geq 0, \quad \text{for all } i \tag{5.8f}$$
$$(t_i - b^2)(a^2b^2 + t_i(x_i^2 + y_i^2 - 1)) - t_i^2 x_i^2 \geq 0, \quad \text{for all } i \tag{5.8g}$$
$$(t_i - a^2)(t_i - b^2)(a^2b^2 + t_i(x_i^2 + y_i^2 - 1)) - t_i^2(x_i^2(t_i - a^2) + y_i^2(t_i - b^2)) \geq 0, \quad \text{for all } i \tag{5.8h}$$
$$t_i \geq 1, \quad \text{for all } i \tag{5.8i}$$
$$a, b \geq 1 \tag{5.8j}$$
$$a \geq b \tag{5.8k}$$

[5] The nonnegativity of AC follows from that of A and C and can thus be dropped.

where all indices i and j are in $1, \ldots, n$.

Constraints (5.8b) ensure that the circles do not overlap. Constraints (5.8c)–(5.8h) guarantee that the circles are all contained in the ellipse. Constraint (5.8j) is a slight strengthening of the nonnegativity of the semiaxes, to guarantee that at least a single circle can be contained in the ellipse. Constraint (5.8k) ensures that the ellipse has its longer semiaxis along the horizontal axis, which we can assume without loss of generality.

Interestingly, the variables x_i and y_i do not have a finite bound: the optimal ellipse can be elongated. This is also why a and b do not have an upper bound a priori. However, once a feasible solution with objective value z has been found, its objective value can be used to derive an upper bound on a, since we can then restrict the solution space to those ellipses that have an area smaller than z. In particular, this means that $1 \leq b \leq a \leq \frac{z}{\pi}$. These then imply finite bounds on x_i and y_i.

Overall, the model has polynomial degree 8 (coming from the a^4b^4 term in constraint (5.8h)), which is a rare phenomenon in practice.

5.2 Strengthening

First of all we can drop constraint (5.8d), because it is implied by $a \geq b$. Further, we can add a constraint that each circle center is inside the ellipse:

$$b^2x_i^2 + a^2y_i^2 \leq a^2b^2. \tag{5.9}$$

We can strengthen this further by adding some simple bounds on the centers:

$$x_i \leq a - 1 \tag{5.10a}$$
$$x_i \geq -a + 1 \tag{5.10b}$$
$$y_i \leq b - 1 \tag{5.10c}$$
$$y_i \geq -b + 1 \tag{5.10d}$$

These are simple enough not to slow down the optimization. In addition, we could add similar constraints for any point along the circles, such as $(x_i \pm 1, y_i)$ and $(x_i, y_i \pm 1)$. This would be four extra quadratic constraints per circle, which could be too much. Therefore, we did not add these constraints.

Finally, we can add a constraint about the area of the ellipse: it needs to be at least as large as the sum of the areas of the circles:

$$ab \geq n. \tag{5.11}$$

This is an objective cut, which in general does not help the optimization, but here it is beneficial in making sure that the ellipse is large enough. Using the symmetry-breaking constraint (5.8k) we can also add the constraint

$$a \geq \sqrt{n}. \tag{5.12}$$

5.3 Symmetry Breaking

There is a lot of symmetry in the formulation. We have already dealt with an obvious one, that of the containing ellipse, by making it longer along the horizontal axis. There are two other sources of symmetry.

One is the mirror symmetry of the ellipse: this allows us to convert any packing into an equivalent one simply by mirroring the location of the circles. To counter this we need to add a constraint that prefers one solution to a symmetric copy. One example is to constrain the centroid of the circles to fall in a certain quarter of the ellipse:

$$\sum_{i=1}^{n} x_i \geq 0 \tag{5.13a}$$

$$\sum_{i=1}^{n} y_i \geq 0. \tag{5.13b}$$

Unless the centroid of the solution is the origin (or lies on one of the coordinate axes), these two constraints will exclude equivalent symmetric solutions. Unfortunately, optimal configurations tend to have some kind of symmetry, so this technique is not perfect.

Another issue is the symmetry of numbering the circles. With n circles we could have $n!$ different solutions just by renumbering them. Therefore, it is common practice[6] to add a symmetry-breaking constraint. The most natural one is a canonical ordering of the circles, such as ordering their centers in increasing order of the x coordinate:

$$x_i \leq x_{i+1}, \quad 1 \leq i \leq n-1. \tag{5.14}$$

This is not perfect, because two circles can still end up with an identical x coordinate (see the solution for $n = 15$ in Figure 2), but it is simple and drastically reduces the symmetry in the problem. An added benefit of this approach is that it can be used to reduce the symmetry of the configuration itself. In particular, instead of the inequality (5.13a) we can write

$$x_{\lceil \frac{n}{2} \rceil} \geq 0, \tag{5.15}$$

which, together with the sorting constraint (5.14), ensures that at least half of the points are in the right half of the ellipse. This inequality has far fewer nonzeros than constraint (5.13a) and achieves a similar result. Note, however, that this cannot be done at the same time for the y coordinate, as we cannot assume that the points are sorted according to both coordinates.

An alternative solution is to order the points along a generic line:

$$\alpha x_i + \beta y_i \leq \alpha x_{i+1} + \beta y_{i+1}, 1 \leq i \leq n-1, \tag{5.16}$$

[6] This would not be necessary for mixed-integer *linear* problems, where exploiting symmetry has long been part of all commercial and even open-source solvers; see [53] for a comprehensive survey. On the other hand, since this kind of symmetry is much less common for nonlinear problems, and detecting it is more complicated, it has not been added to nonlinear solvers.

where α and β should be selected in a way to make it very unlikely that any of these inequalities is binding. A couple of random irrational numbers generally work, but could harm the numerics of the problem. For this application the choice of $\alpha = 2$ and $\beta = 3$ would work. This constraint introduces more nonzeros into the problem, but completely eliminates the symmetry coming from reordering the circles.

5.4 Computational Results

All solutions were post-processed and polished using the techniques in Section 3.3. In particular, to clean up these solutions we added constraints to exploit the (conjectured) structure of the solution: enforcing the apparent symmetry, rounding some values to the coordinate axes, and setting some constraints to equality (to enforce touching objects). We then re-solved the problem to see if these assumptions were correct and the same solution could still be obtained. In all cases this turned out to be true. A similar but more elaborate version of this scheme, employing Gröbner bases to obtain the coordinates in closed form, is used in [19]. In the final version of the solutions all constraints are satisfied to double floating-point precision.

5.4.1 The New Solutions

We were able to reproduce all of the known solutions[7] on Erich Friedman's packing website [12]. In addition, we found two new improving solutions. These are different configurations, not just slightly improved versions of the previous solutions.

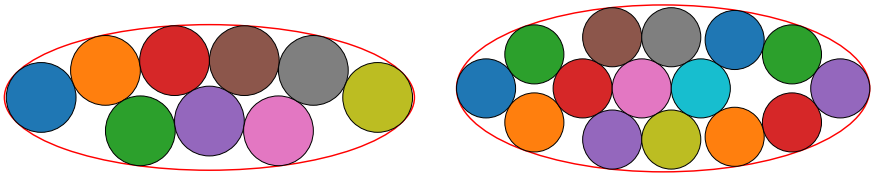


Fig. 2: New circle-packing arrangements for $n = 9$ (left, area $\pi \cdot 12.2708\ldots$, improving the previous $\pi \cdot 12.403\ldots$) and $n = 15$ (right, area $\pi \cdot 19.76329049\ldots$, improving the previous $\pi \cdot 19.824\ldots$).

The configuration for $n = 9$ in Figure 2 is symmetric about the vertical axis. The semiaxes of the ellipse are 5.8787236 and 2.087324145.

The configuration for $n = 15$ in Figure 2 is symmetric about the horizontal axis. The semiaxes of the ellipse are 6.995912684 and 2.824976723. Interestingly, this solution is not rigid: the two points with x coordinates $0.274\ldots$ do not touch the ellipse, so they are free to move. There are thus infinitely many equivalent solutions. The solution we have presented is the one where these circles are touching three other circles.

[7] We do not know how the existing solutions were found and how the circle-ellipse containment was enforced or verified for them.

6 Packing Polygons

The problem of packing n regular m-gons into a regular ℓ-gon of minimum circumradius is significantly more involved than circle packing (Section 4), as each inner element can rotate freely and non-overlap conditions must account for the specific polygonal geometry of each element. Optimization models for general polygon packing have been developed using quasi-phi-functions [14–16]; we adopt a related approach based on a variant of the Farkas lemma [13], which yields a compact formulation applicable to any regular polygon pair without shape-specific derivations.

6.1 Optimization Model

The problem is to pack n regular m-gons with unit circumradius into a regular ℓ-gon, minimizing the circumradius R of the outer polygon. Let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote the set of inner polygons, $\mathcal{M} = \{0, 1, \ldots, m-1\}$ index both the m vertices and m edges of each inner polygon, and $\mathcal{L} = \{0, 1, \ldots, \ell-1\}$ denote the ℓ edges of the outer polygon. We use the following geometric constants for regular polygons with unit circumradius:

$\rho_m = \cos(\pi/m)$*:* inradius of inner m-gons
$\rho_\ell = \cos(\pi/\ell)$*:* inradius of outer ℓ-gon
$\phi_m = 2\pi/m$*:* angular distance between adjacent vertices of m-gons
$\phi_\ell = 2\pi/\ell$*:* angular distance between adjacent vertices of ℓ-gon
$\delta_{m,j} = (j + ((m-1) \bmod 2)/2)\phi_m$*:* vertex angle for $j \in \mathcal{M}$ of m-gons
$R_{\min} = \sqrt{\frac{n \cdot m \cdot \sin(\phi_m)}{\ell \cdot \sin(\phi_\ell)}}$*:* lower bound on circumradius from area requirement

We define the following decision variables:

$R \geq 0$*:* circumradius of the scaled ℓ-gon (to be minimized)
$(x_i, y_i) \in \mathbb{R}^2$*:* coordinates of the center of inner m-gon $i \in \mathcal{N}$
$\theta_i \in [0, \phi_m]$*:* rotation angle of inner m-gon $i \in \mathcal{N}$
$a_{i,j}, b_{i,j} \in \mathbb{R}$*:* oriented inward normal components for edge $j \in \mathcal{M}$ of m-gon $i \in \mathcal{N}$
$s_{i,j} \in \mathbb{R}$*:* offset for edge $j \in \mathcal{M}$ of m-gon $i \in \mathcal{N}$
$\lambda_{i,j,k} \geq 0$*:* Farkas multiplier of pair $i, j \in \mathcal{N}$ with $i < j$ and common edge index $k \in \{1, \ldots, 2m\}$

The separation conditions are based on a version of the Farkas lemma in its strict-inequality variant. More specifically, this is the non-homogeneous version of Gordan's theorem [54], and a special case of Motzkin's transposition theorem, see, e.g., [55]: A system of strict linear inequalities $Ax > b$ is infeasible if and only if there exist nonnegative multipliers $y \geq 0$, $y \neq 0$, with $y^T A = 0$ and $y^T b \geq 0$.

Lemma 3 Two convex m-gons with half-space representations $\{a_{i,k}x + b_{i,k}y \geq s_{i,k}\}_{k \in \mathcal{M}}$ and $\{a_{j,k}x + b_{j,k}y \geq s_{j,k}\}_{k \in \mathcal{M}}$ have disjoint interiors if and only if there exist nonnegative multipliers $\lambda_1, \ldots, \lambda_{2m}$ summing to one with

$$\sum_{k \in \mathcal{M}} \lambda_{k+1} a_{i,k} + \sum_{k \in \mathcal{M}} \lambda_{k+m+1} a_{j,k} = 0, \tag{6.1a}$$

$$\sum_{k\in\mathcal{M}} \lambda_{k+1} b_{i,k} + \sum_{k\in\mathcal{M}} \lambda_{k+m+1} b_{j,k} = 0, \tag{6.1b}$$

$$\sum_{k\in\mathcal{M}} \lambda_{k+1} s_{i,k} + \sum_{k\in\mathcal{M}} \lambda_{k+m+1} s_{j,k} \geq 0, \tag{6.1c}$$

i.e., a canceling normal combination with nonnegative gap.

Proof The interiors of the two polygons are disjoint if and only if the strict linear system

$$a_{i,k}x + b_{i,k}y > s_{i,k}, \quad k \in \mathcal{M}, \tag{6.2a}$$

$$a_{j,k}x + b_{j,k}y > s_{j,k}, \quad k \in \mathcal{M} \tag{6.2b}$$

in $(x, y) \in \mathbb{R}^2$ is infeasible. By the Farkas lemma (and the theorems cited above), this system is infeasible if and only if there exist $\mu_1, \ldots, \mu_{2m} \geq 0$, not all zero, such that

$$\sum_{k\in\mathcal{M}} \mu_k a_{i,k} + \sum_{k\in\mathcal{M}} \mu_{m+k} a_{j,k} = 0, \tag{6.3a}$$

$$\sum_{k\in\mathcal{M}} \mu_k b_{i,k} + \sum_{k\in\mathcal{M}} \mu_{m+k} b_{j,k} = 0, \tag{6.3b}$$

$$\sum_{k\in\mathcal{M}} \mu_k s_{i,k} + \sum_{k\in\mathcal{M}} \mu_{m+k} s_{j,k} \geq 0. \tag{6.3c}$$

Dividing by $\sum_{p=1}^{2m} \mu_p > 0$ yields multipliers summing to one, satisfying all three conditions of the lemma. □

For each pair (i, j) with $i < j$, the $2m$ Farkas multipliers $\lambda_{i,j,k}$, one per edge of each of the two m-gons, encode this certificate. Geometrically, the partial normal sums $\sum_k \lambda_{k+1}(a_{i,k}, b_{i,k})$ and $\sum_k \lambda_{k+m+1}(a_{j,k}, b_{j,k})$ are equal and opposite, defining a normal axis with separating offset $\sum_k \lambda_{k+1} s_{i,k} + \sum_k \lambda_{k+m+1} s_{j,k}$. Conversely, given a separating axis with unit normal (u, v) and offset c, the multipliers for polygon i are obtained by solving the LP

$$\max_{\mu_k \geq 0} \quad \textstyle\sum_{k\in\mathcal{M}} \mu_k \, s_{i,k} \tag{6.4a}$$

$$\text{s.t.} \quad \textstyle\sum_{k\in\mathcal{M}} \mu_k \, (a_{i,k}, b_{i,k}) = -(u, v), \tag{6.4b}$$

whose objective does not exceed c by optimality; solving the same LP for polygon j with the negated direction gives ν_k, and combining both as $\lambda_{k+1} = \mu_k / S$ and $\lambda_{k+m+1} = \nu_k / S$ with $S := \sum_k \mu_k + \sum_k \nu_k$ yields feasible Farkas multipliers. The approach applies to any convex polygon without shape-specific derivations.

Remark 1 Since the dual-feasible bases of the LP (6.4) are defined by the vertices of the two-dimensional polygon, at most two positive adjacent Farkas multipliers per element in each pair are required to satisfy the separation conditions if possible, independent of the number of polygon vertices.

The optimization problem can be formulated as:

$$\min \; R \tag{6.5a}$$

$$R\rho_\ell + \sin(k\phi_\ell)\big(x_i + \sin(\theta_i + \delta_{m,j})\big) + \cos(k\phi_\ell)\big(y_i + \cos(\theta_i + \delta_{m,j})\big) \geq 0, \qquad \forall i \in \mathcal{N},\ j \in \mathcal{M},\ k \in \mathcal{L} \tag{6.5b}$$

$$a_{i,j} = \sin(\theta_i + j\phi_m), \qquad \forall i \in \mathcal{N},\ j \in \mathcal{M} \tag{6.5c}$$

$$b_{i,j} = \cos(\theta_i + j\phi_m), \qquad \forall i \in \mathcal{N},\ j \in \mathcal{M} \tag{6.5d}$$

$$s_{i,j} = a_{i,j}x_i + b_{i,j}y_i - \rho_m, \qquad \forall i \in \mathcal{N},\ j \in \mathcal{M} \tag{6.5e}$$

$$\sum_{k=1}^{2m} \lambda_{i,j,k} = 1, \qquad \forall i,j \in \mathcal{N},\ i < j \tag{6.5f}$$

$$\sum_{k=0}^{m-1} \lambda_{i,j,k+1}a_{i,k} + \sum_{k=0}^{m-1} \lambda_{i,j,k+m+1}a_{j,k} = 0, \qquad \forall i,j \in \mathcal{N},\ i < j \tag{6.5g}$$

$$\sum_{k=0}^{m-1} \lambda_{i,j,k+1}b_{i,k} + \sum_{k=0}^{m-1} \lambda_{i,j,k+m+1}b_{j,k} = 0, \qquad \forall i,j \in \mathcal{N},\ i < j \tag{6.5h}$$

$$\sum_{k=0}^{m-1} \lambda_{i,j,k+1}s_{i,k} + \sum_{k=0}^{m-1} \lambda_{i,j,k+m+1}s_{j,k} \geq 0, \qquad \forall i,j \in \mathcal{N},\ i < j \tag{6.5i}$$

$$0 \leq \theta_i \leq \phi_m, \qquad \forall i \in \mathcal{N} \tag{6.5j}$$

$$\lambda_{i,j,k} \geq 0, \qquad \forall i,j \in \mathcal{N},\ i < j,\ k \in \{1,\dots,2m\} \tag{6.5k}$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 \geq (2\rho_m)^2, \forall i,j \in \mathcal{N},\ i < j \tag{6.5l}$$

$$R \geq R_{\min} \tag{6.5m}$$

Constraints (6.5b) ensure that all vertices of each inner m-gon lie within the outer ℓ-gon. The vertex angles $\delta_{m,j}$ account for the positions of the inner polygon vertices relative to their inward edge normals. Each vertex must satisfy all ℓ half-space constraints defining the outer ℓ-gon.

Constraints (6.5c)–(6.5e) define the oriented half-space representation of the inner polygons. For this, each edge $j \in \mathcal{M}$ of inner polygon $i \in \mathcal{N}$ is evaluated by the oriented inward normal $(a_{i,j}, b_{i,j})$ and offset $s_{i,j}$, representing the defining half-space constraint $a_{i,j}x + b_{i,j}y \geq s_{i,j}$.

Constraints (6.5f)–(6.5i) represent the Farkas-based separation conditions as described above.

Constraint (6.5l) requires that the centers of any two inner m-gons are at least $2\rho_m$ apart, which is redundant given the Farkas separation conditions.

Constraint (6.5j) restricts the rotation angle to $[0, \phi_m]$ due to the m-fold rotational symmetry of regular m-gons, i.e., rotations beyond ϕ_m are congruent to rotations within this range. Finally, the lower bound $R_{\min}$ in (6.5m) is derived from the containment requirement that the area of the scaled outer ℓ-gon must be at least the total area of the n unit inner m-gons.

6.2 Computational Results

The resulting circumradii for all computed polygon-packing instances are listed in the online supplement[8]. We report the improving solutions in Table 2 and depict them in Figure 3.

Table 2: Improving solutions for polygon packing. References are provided if an accompanying publication is available. Otherwise the name and the year are taken from Erich Friedman's packing website.

ℓ	m	n	Circumradius	Previous	Source
4	3	12	3.13**403**	3.13802	Cantrell 2002
5	3	6	2.05**332**	2.05521	Morandi 2012
5	3	10	2.**67560**	2.71218	Cantrell 2012
5	3	11	2.**75788**	2.80900	Cantrell 2012
5	3	12	2.9**0916**	2.93400	Cantrell 2012
5	3	13	**2.97298**	3.00600	Cantrell 2012
5	3	14	3.**07677**	3.13711	Morandi 2012
5	4	9	3.2**4086**	3.28500	Hirsh 2022
5	4	10	3.3**6024**	3.38000	Cantrell 2012
5	4	11	3.**47760**	3.51400	Cantrell 2012
5	4	12	3.57**255**	3.57700	Cantrell 2012
5	4	13	3.75**495**	3.75575	Cantrell 2012
6	3	10	2.59**297**	2.59808	Friedman 2005
6	3	12	2.81**217**	2.81900	Morandi 2008
6	3	13	2.88**480**	2.88676	Friedman 2005
6	6	11	3.9**2451**	3.93010	[6, 7]
6	6	12	3.941**65**	3.94192	[6, 7]
6	6	14	4.2**6900**	4.27240	[56]
6	6	15	4.4**4728**	4.45406	[56]
6	6	16	4.5**2788**	4.53633	[56]
6	6	17	4.61**362**	4.61881	[56]
6	6	23	5.4**0001**	5.42858	[56]

[8] https://github.com/DominikKamp/Packing

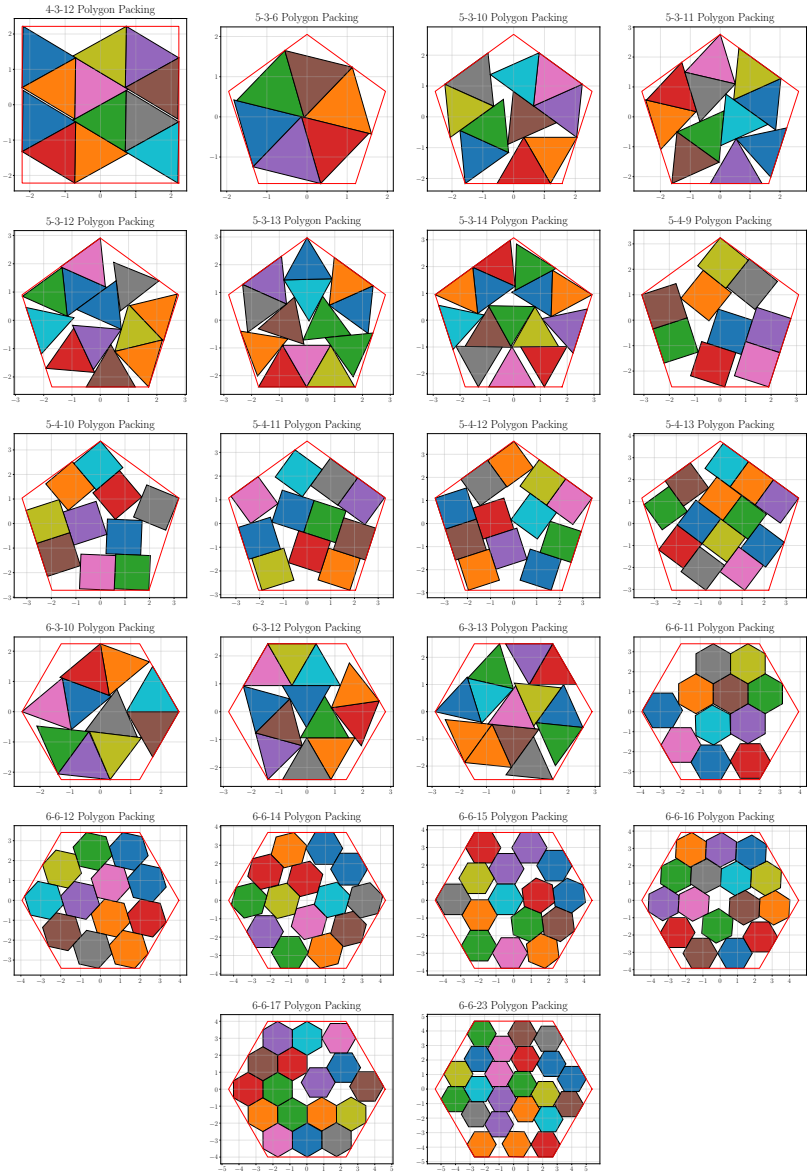


Fig. 3: Graphical representation of the solutions for the polygon-packing problem

The median optimality gap across (ℓ, m) groups ranges from 9.9% to 15.2% (average 13.0%), taking the best primal and dual bounds across all solver setups. For small n, the setups tighten the dual bound beyond the area lower bound $R_{\min}$, proving optimality for one and two inner elements and several instances with $n \leq 5$, while for $n \geq 6$, the dual bound generally remains at $R_{\min}$.

6.3 Solution Polishing

The circumradius of each inner polygon is scaled outward by a factor of $1 + \varepsilon/\rho_m$ in the containment constraints, and the inradius is increased by $\varepsilon/2$ in the separation constraints. The two different scalings ensure that the absolute geometric margin is the same (ε) between an inner element and the outer boundary as between two inner elements, so that any sufficiently precise solution with respect to solver tolerances is feasible to machine tolerance. For our runs $\varepsilon = 10^{-8}$.

After the solutions have been obtained, the rotation angles remain untouched and all center coordinates are scaled uniformly with fixed origin by a carefully chosen factor $\sigma > 0$. The scaling factor σ is computed as the minimum value that eliminates all pairwise overlaps, determined via the separating axis theorem using edge normals. Additionally, downscaling towards the origin is restricted by the containment condition for the given R to avoid unnecessary objective increase. The circumradius R is then recomputed from the scaled configuration using projections of the inner elements on the face normals of the outer element. A relative safety margin of 10^{-14} is applied to account for numerical rounding errors.

6.4 Model Variants

To investigate the impact of formulation choices on solver performance, we test five variants of the polygon-packing model. The formulation (6.5a)–(6.5m) described above, including the redundant distance constraint (6.5l), serves as the base and is denoted DIST. The remaining four variants modify or extend it as follows.

No distance constraint (NODIST)*:* The formulation NODIST removes the redundant distance constraint (6.5l) from the base formulation.

Inner separation (INNER)*:* This variant replaces the Farkas-based separation (6.5f)–(6.5i) entirely with direct separating hyperplane constraints. For each pair (i, j) with $i < j$, a separating direction angle $\alpha_{i,j} \in [0, 2\pi]$ and offset $d_{i,j} \in \mathbb{R}$ are introduced as decision variables in place of the Farkas multipliers. All vertices of polygon i must lie on the upper side of the hyperplane and all vertices of polygon j on the lower side:

$$\sin(\alpha_{i,j})\big(x_i + \sin(\theta_i + \delta_{m,k})\big) + \cos(\alpha_{i,j})\big(y_i + \cos(\theta_i + \delta_{m,k})\big) \geq d_{i,j} \tag{6.6a}$$

$$\sin(\alpha_{i,j})\big(x_j + \sin(\theta_j + \delta_{m,k})\big) + \cos(\alpha_{i,j})\big(y_j + \cos(\theta_j + \delta_{m,k})\big) \leq d_{i,j} \tag{6.6b}$$

for all $i, j \in \mathcal{N}$ with $i < j$ and $k \in \mathcal{M}$. This formulation uses fewer variables per pair (2 instead of $2m$ Farkas multipliers) but more constraints ($2m$ instead of 4 Farkas constraints including the normalization constraint).

Farkas normalization (FARKAS)*:* This variant replaces the overall normalization of Farkas multipliers (6.5f) with a lower bound for each polygon:

$$\sum_{k \in \mathcal{M}} \lambda_{i,j,k+1} \geq 1 \tag{6.7a}$$

$$\sum_{k \in \mathcal{M}} \lambda_{i,j,k+m+1} \geq 1 \tag{6.7b}$$

for all $i, j \in \mathcal{N}$ with $i < j$. The Farkas sum for each polygon needs to be positive because only the intersection of edge constraints across both full-dimensional polygons can be lower-dimensional to prove separation. The remaining Farkas constraints (6.5g)–(6.5i) are unchanged.

Symmetry breaking (SYM)*:* This variant extends DIST with elementary symmetry-breaking constraints. By permutation of variable assignments, the coordinates can always be ordered in nondecreasing order of the x coordinate as in (5.14). Furthermore, the ℓ-fold rotational and reflection symmetries of the outer polygon and of each inner polygon assignment along the outer polygon axes make it possible to restrict the centroid of all polygon centers to the cone spanned by the first half of the first angular sector:

$$\sum_{i \in \mathcal{N}} x_i \geq 0 \tag{6.8a}$$

$$-\cos(\phi_\ell/2) \sum_{i \in \mathcal{N}} x_i + \sin(\phi_\ell/2) \sum_{i \in \mathcal{N}} y_i \geq 0 \tag{6.8b}$$

Since permutation of elements does not change the coordinate sums, both symmetry reductions can be applied simultaneously.

Table 3 reports, across all instances and for each formulation, solutions found (*found*), median and mean per-instance relative difference to DIST, and counts of instances where a formulation strictly beats (*better*) or is strictly beaten by (*worse*) DIST.

Table 3: Polygon packing: overall formulation comparison across 750 instances. Solutions found, median and mean relative difference to DIST, and diff counts per formulation.

metric	DIST	NODIST	INNER	FARKAS	SYM
found	734	416	358	329	737
median	+0.000%	+0.000%	+0.575%	+0.000%	+2.178%
mean	+0.000%	+0.897%	+13.599%	+0.790%	+3.698%
better	0	4	1	1	2
worse	0	187	200	139	539

The redundant distance constraints have the strongest impact, so we use DIST as the reference formulation. The formulation NODIST, without distance constraints, is strictly worse on 187 instances versus only 4 where it is better. The alternative separation approaches INNER and FARKAS produce weaker results, while INNER additionally faces a 0.575% median (13.6% mean) quality degradation. SYM produces the worst solution quality overall: the median objective is 2.178% (3.7% mean) worse than DIST, and 539 instances are strictly worse versus only 2 where it is better. This suggests that symmetry breaking impedes performance of heuristics by unnecessarily enforcing certain orders and locations to accept a configuration.

7 Packing Platonic Solids

The problem of packing n identical Platonic solids of a certain type into a Platonic solid of the same or different type of minimum volume[9] extends the polygon-packing problem to three dimensions, where each inner solid can rotate freely and non-overlap conditions must account for the specific polyhedral geometry of each element. The Farkas-based formulation of Section 6 extends naturally to three dimensions, replacing edge normals by face normals and a single rotation angle by a three-dimensional rotational parametrization.

7.1 Optimization Model

The problem is to pack n Platonic solids of type m with unit circumradius into a Platonic solid of type ℓ, minimizing the circumradius R of the outer solid, where the types $1, \ldots, 5$ correspond to the five Platonic solids: tetrahedron, octahedron, cube, icosahedron, and dodecahedron, respectively. Let $\mathcal{N} = \{1, 2, \ldots, n\}$ denote the set of inner solids, $\mathcal{V}_m$ the set of vertices for inner solid type m, $\mathcal{F}_m$ the set of faces for inner solid type m, and $\mathcal{F}_\ell$ the set of faces for outer solid type ℓ. We use the following geometric constants for Platonic solids with unit circumradius:

$\mathbf{v}_m^j \in \mathbb{R}^3$*:* standard position for vertex $j \in \mathcal{V}_m$
$\mathbf{n}_m^f \in \mathbb{R}^3$*:* standard inward normal for face $f \in \mathcal{F}_m$
$\mathbf{n}_\ell^k \in \mathbb{R}^3$*:* standard inward normal for face $k \in \mathcal{F}_\ell$
ρ_m*:* inradius of inner solid type m
ρ_ℓ*:* inradius of outer solid type ℓ
$R_{\min} = (n \cdot V_m / V_\ell)^{1/3}$*:* lower bound on circumradius from volume requirement, where V_m, V_ℓ are the unit volumes of solid type m and ℓ

We define the following decision variables:

$R \geq 0$*:* circumradius of the scaled outer solid (to be minimized)
$(x_i, y_i, z_i) \in \mathbb{R}^3$*:* coordinates of the center of inner solid $i \in \mathcal{N}$
$\theta_i, \iota_i, \kappa_i \in [0, 2\pi]$*:* ZYX rotation angles of inner solid $i \in \mathcal{N}$
$a_{i,f}, b_{i,f}, c_{i,f} \in \mathbb{R}$*:* oriented inward normal components for face $f \in \mathcal{F}_m$ of inner solid $i \in \mathcal{N}$
$e_{i,f} \in \mathbb{R}$*:* offset for face $f \in \mathcal{F}_m$ of inner solid $i \in \mathcal{N}$
$\lambda_{i,j,k} \geq 0$*:* Farkas multiplier for pair $i, j \in \mathcal{N}$ with $i < j$ and common face index $k \in \{1, \ldots, 2|\mathcal{F}_m|\}$

The pairwise non-overlap conditions are formulated using the same Farkas certificate as for the polygon case, extended to three dimensions. The certificate carries over analogously: two convex polyhedra described by their face half-space representations have disjoint interiors if and only if there exist nonnegative multipliers for their combined face inequalities, summing to one, that yield a zero normal sum and a nonnegative offset sum. For each pair (i, j) with $i < j$, the $2|\mathcal{F}_m|$ Farkas multipliers $\lambda_{i,j,k}$, one per face of each of the two inner solids, encode this certificate; the geometric interpretation and LP-based reverse mapping of Section 6.1 carry over to three dimensions

[9] Or, equivalently, surface area, or circumradius.

analogously. The approach applies to any convex polyhedron without shape-specific derivations. Since the LP has three equality constraints (one per component of the normal equation in $\mathbb{R}^3$), LP basis theory guarantees that not more than three positive Farkas multipliers per element and pair are required.

With the 3×3 ZYX rotation matrix $\mathrm{Rot}(\theta_i, \iota_i, \kappa_i)$, the Platonic-solid packing problem can be formulated as:

$$\min \; R \tag{7.1a}$$

$$\text{s.t.} \quad R\rho_\ell + \mathbf{n}_\ell^k \cdot \left(\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} + \mathrm{Rot}(\theta_i, \iota_i, \kappa_i)\, \mathbf{v}_m^j \right) \geq 0, \quad \forall i \in \mathcal{N},\ j \in \mathcal{V}_m,\ k \in \mathcal{F}_\ell \tag{7.1b}$$

$$\begin{pmatrix} a_{i,f} \\ b_{i,f} \\ c_{i,f} \end{pmatrix} = \mathrm{Rot}(\theta_i, \iota_i, \kappa_i)\, \mathbf{n}_m^f, \quad \forall i \in \mathcal{N},\ f \in \mathcal{F}_m \tag{7.1c}$$

$$e_{i,f} = a_{i,f} x_i + b_{i,f} y_i + c_{i,f} z_i - \rho_m, \quad \forall i \in \mathcal{N},\ f \in \mathcal{F}_m \tag{7.1d}$$

$$\sum_{k=1}^{2|\mathcal{F}_m|} \lambda_{i,j,k} = 1, \quad \forall i, j \in \mathcal{N},\ i < j \tag{7.1e}$$

$$\sum_{k=0}^{|\mathcal{F}_m|-1} \lambda_{i,j,k+1} \begin{pmatrix} a_{i,k} \\ b_{i,k} \\ c_{i,k} \end{pmatrix} + \sum_{k=0}^{|\mathcal{F}_m|-1} \lambda_{i,j,k+|\mathcal{F}_m|+1} \begin{pmatrix} a_{j,k} \\ b_{j,k} \\ c_{j,k} \end{pmatrix} = \mathbf{0}, \quad \forall i, j \in \mathcal{N},\ i < j \tag{7.1f}$$

$$\sum_{k=0}^{|\mathcal{F}_m|-1} \lambda_{i,j,k+1}\, e_{i,k} + \sum_{k=0}^{|\mathcal{F}_m|-1} \lambda_{i,j,k+|\mathcal{F}_m|+1}\, e_{j,k} \geq 0, \quad \forall i, j \in \mathcal{N},\ i < j \tag{7.1g}$$

$$0 \leq \theta_i, \iota_i, \kappa_i \leq 2\pi, \quad \forall i \in \mathcal{N} \tag{7.1h}$$

$$\lambda_{i,j,k} \geq 0, \quad \forall i, j \in \mathcal{N},\ i < j,\ 1 \leq k \leq 2|\mathcal{F}_m| \tag{7.1i}$$

$$R \geq R_{\min} \tag{7.1j}$$

Constraints (7.1b) ensure that all vertices of each inner solid lie within the outer solid by requiring all oriented vertices of inner solids to satisfy all constraints of the outer solid.

Constraints (7.1c)–(7.1d) define the oriented half-space representation of each inner solid. For this, each face $f \in \mathcal{F}_m$ of inner solid i is represented by the oriented inward normal $(a_{i,f}, b_{i,f}, c_{i,f})$ and offset $e_{i,f}$, corresponding to the defining half-space constraint $a_{i,f}x + b_{i,f}y + c_{i,f}z \geq e_{i,f}$.

Constraints (7.1e)–(7.1g) represent the Farkas-based separation conditions as described above.

Constraints (7.1h) leave the rotation unrestricted since three-dimensional structures in general do not have congruence symmetries in each rotational coordinate in contrast to regular polygons in two dimensions. Finally, the lower bound $R_{\min}$ in (7.1j)

is again derived from the containment requirement that the volume of the scaled outer ℓ-solid must be at least the total volume of the n unit inner m-solids.

7.2 Computational Results

The resulting circumradii for all computed Platonic-solid packing instances are listed in the online supplement[10]. Here we list the solutions that are either improving or are new and of an interesting structure.

First, we have found an improving cube-packing solution for the case of packing 11 cubes into a larger cube. The solution is depicted in Figure 4. This is a well-researched area, so it is surprising that such a solution could be found. The numerical details of

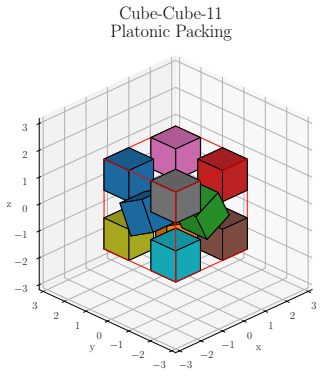


Fig. 4: Packing of 11 unit-circumradius cubes into a cube with $R = 2.89445$

the solution are summarized in Table 4.

Table 4: Improving solution for packing 11 cubes into a cube

ℓ	m	n	Circumradius	Previous	Source
3	3	11	2.**89445**	2.91210	Friedman 1998

In addition we have tested all combinations of inner and outer polyhedra with up to 30 inner ones, resulting in $5 \cdot 5 \cdot 30 = 750$ instances. For most of these problems we are the first to report any feasible solutions. A few selected configurations are depicted in Figure 5. We find it remarkable that such tight configurations could be found with a straightforward model and a general-purpose solver.

[10] https://github.com/DominikKamp/Packing

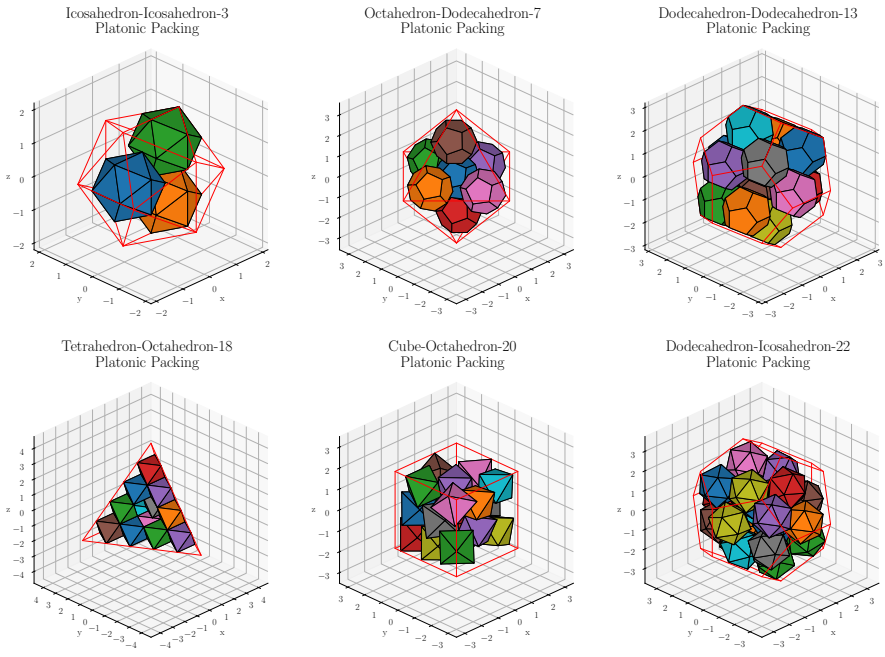


Fig. 5: Graphical representation of the solutions for selected instances of the Platonic-solid packing problem

The median optimality gaps are significantly larger than for polygons, ranging from 14.5% to 33.3% (average 26.7%). The dual bound is improved beyond $R_{\min}$ only for $n \leq 2$, indicating that further model strengthening is necessary to close the gaps.

7.3 Solution Polishing

We used a scheme analogous to the polygon-packing study, presented in Section 6.3.

7.4 Model Variants

Similar to polygon packing, we have investigated the performance of the other formulation variants. The same five formulation variants described in Section 6.4 are tested for Platonic-solid packing, with each modification extended analogously to three dimensions. Dist adds redundant center-distance constraints $(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \geq (2\rho_m)^2$; Inner replaces Farkas separation with direct inner separation hyperplanes in $\mathbb{R}^3$; Farkas uses individual normalization $\sum_k \lambda_{i,j,k+1} \geq 1$ and $\sum_k \lambda_{i,j,k+|\mathcal{F}_m|+1} \geq 1$ in place of (7.1e); and Sym extends Dist with $x_i \leq x_{i+1}$ ordering and centroid constraints into the cone spanned by the first half of the first angular sector of the first face. Table 5 reports the same metrics for the Platonic-solid packing instances.

Table 5: Platonic-solid packing: overall formulation comparison across 750 instances. Solutions found, median and mean relative difference to DIST, and diff counts per formulation.

metric	DIST	NODIST	INNER	FARKAS	SYM
found	429	263	256	208	417
median	+0.000%	+0.000%	+2.610%	+0.000%	+0.763%
mean	+0.000%	+0.836%	+34.049%	+1.263%	+1.948%
better	0	18	2	7	18
worse	0	108	160	84	264

The overall pattern resembles the polygon case, while reflecting the increased difficulty of the three-dimensional problem. DIST again serves as reference. Notably, both NODIST and SYM produce 18 instances that are strictly better than DIST. For Platonic-solid packing, INNER faces the largest quality degradation (+2.610% median, +34.0% mean) and 160 strictly worse instances. FARKAS matches DIST in median quality but is strictly worse on 84 instances versus 7 where it is better.

This is a whole new and wide-open area of research. We expect that some of the solutions we have found will soon be improved either by us or by other teams.

8 Solver Performance and Timings

Table 6 summarizes the solving behavior over the instances with a known reference, reporting the geometric mean of the time to reach an incumbent within 1%, 5% and 10% of it. The reported times count the global solve only, excluding the preceding multistart ramp-up of the Xpress runs. An incumbent within 1% of reference is found within seconds for the circle packings, and within 5% of reference within minutes for polygon and Platonic-solid packing.

Table 6: Solution time overview over the referenced instances of each problem (10 000 s time limit). The *time* columns give the geometric mean shifted by one second of the time in seconds to reach an incumbent within 1%, 5% and 10% of the reference value, when one is available. Polygon and Platonic use their distance formulations.

		time		
Problem	N	1%	5%	10%
Circle packing (square)	32	3.1	2.1	2.1
Circle packing (rectangle)	30	5.8	3.1	2.9
Polygon packing	365	1074.7	302.0	161.3
Platonic-solid packing	30	482.2	360.3	310.9

For the circles-in-ellipse problem, the best-known solution was obtained within a minute with a simple solver run. That problem did not require the elaborate computational setup of the other problems in this paper.

9 Outlook and Conclusion

In this paper we revisited a collection of geometric packing problems for which the community has been reporting best-known configurations over decades, and we obtained new best solutions for a substantial subset of them. In particular, we have improved best-known results for circle packing with variable radii in fixed-perimeter rectangles, for packing unit circles into minimum-area ellipses, for a broad range of regular polygon-packing instances, and even for the well-studied case of packing 11 cubes into a cube. For the polygon and the ellipse-packing problems, we introduced new modeling methodologies; in particular, a practical use case for the S-lemma.

Beyond improving existing benchmarks, we also argued that some of the modeling ideas we used generalize naturally to higher-dimensional settings. In particular, our Farkas-based non-overlap formulation for polygon packing generalizes to polyhedron packing in higher dimensions, which motivated us to introduce the problem class of packing Platonic solids into Platonic solids, and to report high-quality solutions for a first collection of instances which we derived from our models.

The central message, however, is not about packing. Rather, this work demonstrates that global nonlinear optimization has become an out-of-the-box technology. The improvements reported here were obtained with standard, general-purpose solvers (FICO Xpress and SCIP) on compact NLP formulations, without any problem-specific solution algorithm or heuristic: the expertise went into the models, the search was left to the solvers. Some of the authors have already argued about the importance of this point in [57].

In this sense, modern global solvers provide a competitive baseline for optimization problems involving nonlinearities, including the nonconvex geometric problems that have recently attracted attention through LLM-driven discovery. More than that, global optimization offers complementary strengths: it operates on transparent models and provides dual bounds, it generalizes when the problem to be solved changes slightly, and it does not require an expensive training/learning phase.

From a modeling perspective, we unsurprisingly found that formulation choices matter. Our systematic comparisons of five formulation variants for polygon and Platonic-solid packing (Sections 6.4 and 7.4) confirm the value of standard strengthening techniques. The addition of redundant quadratic distance constraints (Dist) emerged as the single most impactful modification, substantially increasing solution quality. These constraints are primarily thought to strengthen the dual bound, but we observed that this consequently also leads to stronger primal solutions, a behavior known from MIP [58]. Symmetry-breaking constraints (Sym) degraded solution quality. Alternative Farkas normalizations (Farkas) and direct inner separation (Inner) did not improve upon the base formulation.

There are also clear caveats. The most immediate one is the quadratic growth in model size induced by pairwise non-overlap constraints. This becomes a computational

obstacle when one moves towards instances with hundreds or thousands of objects. One simple way to mitigate that is to couple the non-overlap constraints with the symmetry-breaking constraints that sort the objects left to right, and only add the non-overlap constraints if the objects are not too far from each other. A more complicated but probably more efficient method is to separate violated non-overlap constraints and other valid inequalities dynamically. Modern global solvers are already prepared for this line of research: Xpress offers callback mechanisms for branching, cutting, and various other solver interactions, and the plugin-based design of SCIP provides analogous extensibility.

The polygon and Platonic-solid packing formulations could be further strengthened by:

1. simplifying the separation constraints combinatorially by indicating the separating vertices through integer variables, as suggested by the sparsity result discussed in connection with Lemma 3;
2. exploiting the structure of the separation constraints in a cutting-plane framework to dynamically add violated constraints only for pairs that are close to each other;
3. using density propagation, deriving lower bounds on R from area conditions on restricted element coordinates during spatial branch-and-bound.

Finally, while AlphaEvolve was a strong trigger to look at classical geometric benchmarks from a model-and-solve perspective, not every problem from that line of research [6, 7] is equally amenable to compact NLP modeling. The kissing number problem, for example, is conceptually simple to model but immediately suffers from the same combinatorial growth discussed above.

On the other hand, our earlier work [18] already demonstrated that global NLP solvers can be highly effective on distance-geometry-style instances, and we are optimistic that further problems such as the Heilbronn problem (for first results on different variants, see [19, 59]) can be attacked in the same way.

Looking ahead, we expect the scope of problems, both academic research and industrial applications, for which a model-and-solve approach is competitive to expand rapidly with continued progress in global optimization technology. The mixed-integer linear programming success story suggests a plausible trajectory: with the right combination of general yet structure-aware cuts, presolving, propagation, and primal heuristics, computational MINLP can become a black-box tool that is routinely usable at scale. We should stop looking at *global* MINLP solvers as generalizations of *local* nonlinear solvers, and instead view them as extensions of linear MIP solvers: they offer the same guarantees (primal solution, dual bound), but can handle any kind of nonlinearity. This opens the door to a much larger user base.

The recipe is, in principle, known: step back, mine the already existing rich literature on algorithms and theory for nonlinear optimization, and turn that research into robust, general-purpose solver technology. The packing case studies in this paper illustrate both that this direction is already paying off and that there remains substantial research potential. The beauty of research on general-purpose global optimization algorithms is its impact: improvement transfers across models and domains and can lead to outsized real-world impact.

Acknowledgements. We are grateful to Liding Xu (ZIB) and the FICO Xpress team, in particular Bruno Vieira and Susanne Heipcke, for their contributions during an early phase of this project. We thank Anita Hovanesian (ZIB) for her support with visualizing solutions. We are indebted to Erich Friedman for his great collection of packing problems and benchmark solutions, which sparked this work, and for incorporating our new findings on his page.

Funding disclosure. Research reported in this paper was partially supported through the Research Campus MODAL funded by the German Federal Ministry of Education and Research (fund numbers 05M14ZAM, 05M20ZBM) and the Deutsche Forschungsgemeinschaft (DFG) through the DFG Cluster of Excellence MATH+.

Competing interests. Timo Berthold and Imre Pólik are employed by Fair Isaac Corporation (FICO), which develops and sells FICO Xpress, one of the two solvers used in this study. Dominik Kamp, Gioni Mexi and Sebastian Pokutta are affiliated with Zuse Institute Berlin, which develops the SCIP Optimization Suite, the other solver used.

References

[1] Hojny, C., Besançon, M., Bestuzheva, K., Borst, S., Chmiela, A., Dionísio, J., Eifler, L., Ghannam, M., Gleixner, A., Göß, A., Hoen, A., van der Hulst, R., Kamp, D., Koch, T., Kofler, K., Lentz, J., Maher, S.J., Mexi, G., Mühmer, E., Pfetsch, M.E., Pokutta, S., Serrano, F., Shinano, Y., Turner, M., Vigerske, S., Walter, M., Weninger, D., Xu, L.: The SCIP Optimization Suite 10.0. Technical report (2025). https://doi.org/10.48550/arXiv.2511.18580

[2] Belotti, P., Berthold, T., Gally, T., Gottwald, L., Pólik, I.: Solving MINLPs to global optimality with FICO Xpress Global. Optimization, 1–19 (2025) https://doi.org/10.1080/02331934.2025.2595437

[3] Berthold, T.: Heuristic algorithms in global MINLP solvers. PhD thesis, Technische Universität Berlin (2014)

[4] Berthold, T.: A computational study of primal heuristics inside an MI(NL)P solver. Journal of Global Optimization **70**(1), 189–206 (2018) https://doi.org/10.1007/s10898-017-0600-3

[5] Berthold, T., Lodi, A., Salvagnin, D.: Primal Heuristics in Integer Programming. Cambridge University Press, Cambridge, UK (2025). https://doi.org/10.1017/9781009574792

[6] Georgiev, B., Gómez-Serrano, J., Tao, T., Wagner, A.Z.: Mathematical exploration and discovery at scale. arXiv preprint arXiv:2511.02864 (2025) https://doi.org/10.48550/arXiv.2511.02864

[7] Novikov, A., Vũ, N., Eisenberger, M., Dupont, E., Huang, P.-S., Wagner, A.Z., Shirobokov, S., Kozlovskii, B., Ruiz, F.J.R., Mehrabian, A., *et al.*: AlphaEvolve: A coding agent for scientific and algorithmic discovery. arXiv preprint arXiv:2506.13131 (2025) https://doi.org/10.48550/arXiv.2506.13131

[8] Sharma, A.: OpenEvolve: an open-source evolutionary coding agent. GitHub (2025). https://github.com/algorithmicsuperintelligence/openevolve

[9] Sharma, A.: Advanced circle packing for n=26 circles in a unit square. GitHub (2025). https://github.com/algorithmicsuperintelligence/openevolve/blob/main/examples/circle_packing/best_program.py

[10] SciPy API: Optimization and root finding (2025). https://docs.scipy.org/doc/scipy/reference/optimize.minimize-slsqp.html

[11] Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., *et al.*: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods **17**(3), 261–272 (2020) https://doi.org/10.1038/s41592-019-0686-2

[12] Friedman, E.: Packing Center — Circle Packings and Related Problems. https://erich-friedman.github.io/packing/. Online Resource

[13] Farkas, J.: Theorie der einfachen Ungleichungen. Journal für die reine und angewandte Mathematik (Crelle's Journal) **124**, 1–27 (1902) https://doi.org/10.1515/crll.1902.124.1

[14] Kallrath, J.: Cutting circles and polygons from area-minimizing rectangles. Journal of Global Optimization **43**(2-3), 299–328 (2009) https://doi.org/10.1007/s10898-007-9274-6

[15] Stoyan, Y., Pankratov, A., Romanova, T.: Quasi-phi-functions and optimal packing of ellipses. Journal of Global Optimization **65**(2), 283–307 (2016) https://doi.org/10.1007/s10898-015-0331-2

[16] Romanova, T., Bennell, J., Stoyan, Y., Pankratov, A.: Packing of concave polyhedra with continuous rotations using nonlinear optimisation. European Journal of Operational Research **268**(1), 37–53 (2018) https://doi.org/10.1016/j.ejor.2018.01.025

[17] Pólik, I., Terlaky, T.: A survey of the S-lemma. SIAM Review **49**(3), 371–418 (2007) https://doi.org/10.1137/S003614450444614X

[18] Berthold, T., Kamp, D., Mexi, G., Pokutta, S., Pólik, I.: Global optimization for combinatorial geometry problems revisited in the era of LLMs. In: Mahjoub, A.R., Kutucu, H., Küçükyavuz, S., Kaibel, V. (eds.) Combinatorial Optimization, pp. 121–135. Springer, Cham (2027). https://doi.org/10.1007/978-3-032-34519-6_10

[19] Sudermann-Merx, N.: From global optimization to exact coordinates: Heilbronn's triangle problem on the unit square using mixed-integer nonlinear optimization. arXiv preprint arXiv:2603.11107 (2026) https://doi.org/10.48550/arXiv.2603.11107

[20] Hales, T.C.: A proof of the Kepler conjecture. Annals of Mathematics **162**(3), 1065–1185 (2005) https://doi.org/10.4007/annals.2005.162.1065

[21] Bolyai, W.: Tentamen Juventutem Studiosam in Elementa Matheseos Purae, Elementaris Ac Sublimioris. Typis Collegii Reformatorum, Marosvásárhely (1832). Two volumes

[22] Fejes Tóth, L.: Regular Figures. Pergamon Press, London, UK (1964). https://doi.org/10.1016/C2013-0-01735-5

[23] Wäscher, G., Haußner, H., Schumann, H.: An improved typology of cutting and packing problems. European Journal of Operational Research **183**(3), 1109–1130 (2007) https://doi.org/10.1016/j.ejor.2005.12.047

[24] Fekete, S.P., Van der Veen, J.C.: PackLib2: An integrated library of multi-dimensional packing problems. European Journal of Operational Research **183**(3), 1131–1135 (2007) https://doi.org/10.1016/j.ejor.2006.04.023

[25] Iori, M., Lima, V.L., Martello, S., Monaci, M.: 2DPackLib: a two-dimensional cutting and packing library. Optimization Letters **16**(2), 471–480 (2022) https://doi.org/10.1007/s11590-021-01808-y

[26] Kuznetsov, A.: Convexification and global optimization of problems involving the Euclidean norm. PhD thesis, Georgia Institute of Technology (2024)

[27] MINLP Library. https://minlplib.org

[28] Liberti, L., Lavor, C., Maculan, N., Mucherino, A.: Euclidean distance geometry and applications. SIAM Review **56**(1), 3–69 (2014) https://doi.org/10.1137/120875909

[29] Locatelli, M., Raber, U.: Packing equal circles in a square: a deterministic global optimization approach. Discrete Applied Mathematics **122**(1-3), 139–166 (2002) https://doi.org/10.1016/S0166-218X(01)00359-6

[30] Castillo, I., Kampas, F.J., Pintér, J.D.: Solving circle packing problems by global optimization: numerical results and industrial applications. European Journal of Operational Research **191**(3), 786–802 (2008) https://doi.org/10.1016/j.ejor.2007.01.054

[31] Wainwright, M.: Prize specimens (2001). https://plus.maths.org/content/prize-specimens

[32] Khajavirad, A.: The circle packing problem: A theoretical comparison of various convexification techniques. Operations Research Letters **57** (2024) https://doi.org/10.1016/j.orl.2024.107197

[33] Berthold, T.: FICO Xpress Optimization Surpasses AlphaEvolve's Achievements. Blog post (2025). https://www.fico.com/blogs/best-global-optimization-solver

[34] Sahni, S.: Computationally related problems. SIAM Journal on Computing **3**(4), 262–279 (1974) https://doi.org/10.1137/0203021

[35] Matijasevič, J.V.: Enumerable sets are Diophantine. Soviet Mathematics. Doklady **11**(2), 354–358 (1970)

[36] McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems. Mathematical Programming **10**(1), 147–175 (1976) https://doi.org/10.1007/BF01580665

[37] Sherali, H.D., Adams, W.P.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM Journal on Discrete Mathematics **3**(3), 411–430 (1990) https://doi.org/10.1137/0403036

[38] Anstreicher, K.M.: Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. Journal of Global Optimization **43**(2-3), 471–484 (2009) https://doi.org/10.1007/s10898-008-9372-0

[39] Belotti, P., Lee, J., Liberti, L., Margot, F., Wächter, A.: Branching and bounds tightening techniques for non-convex MINLP. Optimization Methods & Software **24**, 597–634 (2009) https://doi.org/10.1080/10556780903087124

[40] Gleixner, A., Berthold, T., Müller, B., Weltge, S.: Three enhancements for optimization-based bound tightening. Journal of Global Optimization **67**, 731–757 (2017) https://doi.org/10.1007/s10898-016-0450-4

[41] Vigerske, S.: Decomposition in multistage stochastic programming and a constraint integer programming approach to mixed-integer nonlinear programming. PhD thesis, Humboldt-Universität zu Berlin (2013). https://doi.org/10.18452/16704

[42] Berthold, T., Geis, F.: Learning to choose branching rules for nonconvex MINLPs. arXiv preprint arXiv:2602.09996 (2026) https://doi.org/10.48550/arXiv.2602.09996

[43] Berthold, T., Witzig, J.: Conflict analysis for MINLP. INFORMS Journal on Computing **33**, 421–435 (2021) https://doi.org/10.1287/ijoc.2020.1050

[44] Berthold, T., Gleixner, A.M., Heinz, S., Vigerske, S.: Analyzing the computational impact of MIQCP solver components. Numerical Algebra, Control and Optimization **2**(4), 739–748 (2012) https://doi.org/10.3934/naco.2012.2.739

[45] Tawarmalani, M., Sahinidis, N.V.: Convexification and Global Optimization in Continuous and Mixed-integer Nonlinear Programming: Theory, Algorithms, Software, and Applications. Nonconvex Optimization and Its Applications, vol. 65. Kluwer Academic Publishers, The Netherlands (2002). https://doi.org/10.1007/978-1-4757-3532-1

[46] Tawarmalani, M., Sahinidis, N.V.: Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. Mathematical Programming **99**, 563–591 (2004) https://doi.org/10.1007/s10107-003-0467-6

[47] Misener, R., Floudas, C.A.: ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations. Journal of Global Optimization **59**(2-3), 503–526 (2014) https://doi.org/10.1007/s10898-014-0166-2

[48] Achterberg, T.: SCIP: Solving Constraint Integer Programs. Mathematical Programming Computation **1**(1), 1–41 (2009) https://doi.org/10.1007/s12532-008-0001-1

[49] Peikert, R., Würtz, D., Monagan, M., de Groot, C.: Packing circles in a square: a review and new results. In: System Modelling and Optimization: Proceedings of the 15th IFIP Conference Zurich, Switzerland, September 2–6, 1991, pp. 45–54. Springer, Berlin, Heidelberg (1992). https://doi.org/10.1007/BFb0113271

[50] Buchberger, B.: A theoretical basis for the reduction of polynomials to canonical forms. ACM SIGSAM Bulletin **10**(3), 19–29 (1976) https://doi.org/10.1145/1088216.1088219

[51] Friedman, E.: Circles in Squares (2012). https://erich-friedman.github.io/packing/cirRsqu/ Accessed 2025-12-15

[52] Yakubovich, V.A.: S-procedure in nonlinear control theory. Vestnik Leningrad University **1**, 62–77 (1971). In Russian

[53] Pfetsch, M.E., Rehn, T.: A computational comparison of symmetry handling methods for mixed integer programs. Mathematical Programming Computation **11**, 37–93 (2019) https://doi.org/10.1007/s12532-018-0140-y

[54] Gordan, P.: Über die auflösung linearer Gleichungen mit reelen Coefficienten. Mathematische Annalen **6**(1), 23–28 (1873) https://doi.org/10.1007/BF01442864

[55] Schrijver, A.: Theory of Linear and Integer Programming. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Chichester

(1986)

[56] Friedman, E.: Hexagons in Hexagons (2015). https://erich-friedman.github.io/packing/hexinhex/ Accessed 2025-12-15

[57] Berthold, T., Pólik, I.: What GenAI taught us about the future of global optimization. ORMS Today **53**(1) (2026) https://doi.org/10.1287/orms.2026.01.01

[58] Berthold, T.: Primal Heuristics for Mixed Integer Programs. Diploma thesis, Technische Universität Berlin (2006)

[59] Monji, A., Modir, A., Kocuk, B.: Solving the Heilbronn triangle problem using global optimization methods. arXiv preprint arXiv:2512.14505 (2025) https://doi.org/10.48550/arXiv.2512.14505