A SHAPE OPTIMIZATION APPROACH TO THE PROBLEM OF COVERING A TWO-DIMENSIONAL REGION WITH MINIMUM-RADIUS IDENTICAL BALLS*

E. G. BIRGIN†, A. LAURAIN‡, R. MASSAMBONE†, AND A. G. SANTANA†

Abstract. We investigate the problem of covering a region in the plane with the union of m identical balls of minimum radius. The region to be covered may be disconnected, nonconvex, have Lipschitz boundary and in particular may have corners. Nullifying the area of the complement of the union of balls with respect to the region to be covered is considered as the constraint, while minimizing the balls' radius is the objective function. The first-order sensitivity analysis of the area to be nullified in the constraint is performed using shape optimization techniques. Bi-Lipschitz mappings are built to model small perturbations of the nonsmooth shape defined via unions and intersections; this allows us to compute the derivative of the constraint via the notion of shape derivative. The considered approach is fairly general and can be adapted to tackle other relevant nonsmooth shape optimization problems. By discretizing the integrals that appear in the formulation of the problem and its derivatives, a nonlinear programming problem is obtained. From the practical point of view, the region to be covered is modeled by an oracle that, for a given point, answers whether it belongs to the region or not. No additional information on the region is required. Numerical examples in which the nonlinear programming problem is solved with an Augmented Lagrangian approach are presented. The experiments illustrate the wide variety of regions whose covering can be addressed with the proposed approach.

Keywords: Covering problem, nonsmooth shape optimization, shape derivatives.

AMS subject classification: 49Q10, 49J52, 49Q12

1. Introduction.

In this work we consider the problem of finding the minimum radius r of m identical balls $B(x_i, r)$, $i = 1, \dots, m$, whose union covers a given arbitrary region $A \subset \mathbb{R}^d$. The covering problem has a wide variety of practical applications ranging from the configuration of a gamma ray machine radiotherapy equipment unit [26] to placing base stations [10]. The problem of covering the d-dimensional space or a bounded region with overlapping identical balls minimizing the number of balls or their radius represents a challenging problem that has been studied for more than half a century [8, 34]. An attempt of devising a formula for the area of a ball that is covered by two other identical balls in the plane was reported in 1962 in [41, pp.184,185]. The author said "*It was found that a single 'formula' could not be obtained for the area covered but an algorithm was devised which uses no less than eight formulae depending on certain geometric properties of the covering configuration.*" He further concluded that "*The impossibility of obtaining any reasonable 'formula' for the function we are trying to maximize in the relatively trivial case $m = 2$ seems to indicate the futility of the analytical approach especially when m is large. On this sad note the general analytical approach was abandoned and another method of a somewhat experimental nature [hereafter named black box maximization], using high-speed electronic computers, was adopted.*" Since then, several approaches to the problem are based on different kind of numerical optimization techniques. Although some of the techniques can be applied with small variations to arbitrary dimensions, applica-

*This work has been partially supported by FAPESP (grants 2013/07375-0, 2016/01860-1, 2018/24293-0, 2019/25258-7) and CNPq (grants 302682/2019-8, 304258/2018-0, 408175/2018-4).

†Dept. of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, Cidade Universitária, 05508-090, São Paulo, SP, Brazil. e-mails: egbirgin@ime.usp.br, rmassambone@ime.usp.br, and ags@ime.usp.br

‡Dept. of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, Cidade Universitária, 05508-090, São Paulo, SP, Brazil. e-mail: laurain@ime.usp.br

tions and the appeal of representing solutions graphically justify the attention that has been given to the cases $d = 2, 3$.

In [30] and [31] the cases in which A is an equilateral triangle and a square are considered, respectively. In both cases a two-level optimization strategy is considered. In the inner level, the radius r is fixed and a feasibility problem is solved to determine whether, with the fixed radius, there exist balls' centers $x_1, \dots, x_m \in \mathbb{R}^2$ such that the balls cover A. BFGS [29], a quasi-Newton method for smooth unconstrained minimization, is used to perform this task (presumably, minimizing the squared residual). Unfortunately, it is not explicit in [30] and [31] how the feasibility problem is modeled and how its first-order derivatives are computed. Depending on whether the balls with fixed radius cover A or not, a discrete rule is used in the outer level to update r. The method stops with a prescribed precision on the radius. In [36] the problem in which A is a region given by the union and the difference of polygons is considered. A mathematical programming model is proposed and analyzed. The proposed method is based on the computation of a feasible descent direction [42] that requires solving a linear programming problem at each iteration. In [18, 27, 28] a simulated annealing approach with an adaptive mesh is considered. Balls' centers are chosen as points in the mesh. Then, points in the mesh are assigned to the closest center using Voronoi tessellation and, as a consequence, the optimal radius for balls with the given centers to cover all points in the mesh is easily obtained. Neighbor solutions constituted by perturbations of the current centers are evaluated and accepted as in a classical local search strategy within the framework of the simulated annealing approach. The cases in which A is a rectangle, a triangle, and a square are tackled with slight variations of this approach in [18], [27], and [28], respectively. In [40, 38], arbitrary 2D and 3D regions are considered but the problem of covering the region is replaced by the problem of covering an arbitrary chosen set of points within A. Then, a specific optimization technique named hyperbolic penalization [39] is applied. In [2] the problem of covering an arbitrary region A is modeled as a nonlinear semidefinite programming problem with the help of convex algebraic geometry tools. The introduced model describes the covering problem without resorting to discretizations, but depends on some polynomials of unknown degrees with impracticable large bounds and whose coefficients are hard-to-compute. The resulting problem is solved with an Augmented Lagrangian (AL) method for nonlinear semidefinite programming. Solving the AL subproblems requires several spectral decompositions per iteration, being very time consuming; thus, only a limited number of numerical examples is exhibited.

In the present work, the covering problem is tackled from a shape optimization perspective. In a broad sense, shape optimization is the study of optimization problems where the variable is a geometric object, usually a subset of $\mathbb{R}^d$; see [12, 17, 35]. The covering problem may be naturally formulated as a nonsmooth shape optimization problem, as A may be nonsmooth, and the union of balls $B(x_i, r)$ covering A can be seen, except for degenerate cases, as a union of curvilinear polygons. To be more precise, Lipschitz domains and transformations seem to be the natural framework to model covering with a union of balls. Shape sensitivity analysis in a Lipschitz setting is well-understood – a family of Lipschitz domains is parameterized via diffeomorphisms applied to a reference shape, then the integral on the moving domain is pulled back to the reference domain, and in this way the so-called *shape derivative* [12, 17, 35] can be computed. In this paper, standard shape derivative formulae for Lipschitz domains are used to compute the sensitivity of the constraint.

The covering problem, formulated as a shape optimization problem, features an interesting class of moving nonsmooth domains that has received little attention in the literature so far, that is, moving domains defined via unions and intersections of subcomponents animated by their own independent motions. To be more specific, in this approach the variable domain is the complement of the union of balls with respect to the region to be covered, where each ball may either be dilated or be translated in an arbitrary direction. The problem then consists in minimizing the radius of the identical balls, with the constraint that the area of this variable domain vanishes. The main task is then to compute the derivative of this constraint with respect to translation and dilations of the balls. Specialized methods have been developed to compute the first-order derivative of the area of a union of balls: in two dimensions for dilations and translations in [20], and for translations in any dimension in [9], where the derivative is expressed as a linear combination of the derivatives of the distances between the centers. Nevertheless, a general methodology for the sensitivity analysis of shape funcionals depending on sets defined via unions and intersections is lacking. The main challenge we are facing in this setting is the construction of a bi-Lipschitz mapping between the reference domain and the moving domain, which also needs to coincide with the basic transformations of the subcomponents. Our main contribution is to show that stretching, moving spheres and their intersection with a fixed set may be represented by a bi-Lipschitz map, which allows us to use the known shape derivative formulae. The techniques and ideas developed in this work to build such a mapping are fairly general and can be used in two dimensions for shape functionals involving sets defined via unions and/or intersections, involving the solutions of partial differential equations, and to compute second-order derivatives. They can also be used to study the structure of first- and second-order shape derivatives, a topic that is well-understood in the smooth framework but has been less investigated in the nonsmooth case; see the pioneering work [11] and the recent contributions [14, 15, 22, 23]. Some of these techniques are nevertheless specific to two dimensions and distinct methods should be devised to treat the case of higher dimensions.

The rest of this work is organized as follows. In Section 2 we describe the shape optimization formulation of the covering problem considered in this paper, and we give the formulae for the gradient of its constraint. Section 3 is devoted to the proof of differentiability of the constraint function. We first show that, under some natural nondegeneracy conditions, the structure of the variable domain is preserved, for small translations and dilations of the balls. This is a prerequisite to perform shape sensitivity analysis and compute shape derivatives. Then we build the bi-Lipschitz mapping between the reference domain and the moving domain, and we use it to compute the derivatives. In Section 4 we describe algorithms to approximate areas and line integrals appearing in the constraint and its derivatives, and provide convergence estimates for the approximations. In Section 5, numerical experiments illustrate the applicability of the introduced approach to a variety of regions A to be covered. Conclusions and lines for future research are given in the last section.

Notation. For a given set $\omega \subset \mathbb{R}^2$, $\partial\omega$ denotes its boundary, $\overline{\omega}$ its closure, and ω^c its complement. The notation $\|\cdot\|$ is used for the Euclidean norm. The divergence of a sufficiently smooth vector field $\mathbb{R}^2 \ni (x,y) \mapsto V(x,y) = (V_1(x,y), V_2(x,y)) \in \mathbb{R}^2$ is defined by $\operatorname{div} V := \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y}$, and its Jacobian matrix is denoted DV.

2. The continuous problem.

Let A be an open bounded subset of $\mathbb{R}^2$ and $\Omega(\boldsymbol{x}, r) = \bigcup_{i=1}^m B(x_i, r)$, where $\boldsymbol{x} := \{x_i\}_{i=1}^m$ and $B(x_i, r)$ are open balls with centers $x_i \in \mathbb{R}^2$ and radii r. We consider the problem of covering A using a fixed number m of balls $B(x_i, r)$ with minimal radius r, i.e., we are looking for a vector $(\boldsymbol{x}, r) \in \mathbb{R}^{2m+1}$

such that $A \subset \Omega(\boldsymbol{x}, r)$ with minimal r. The problem can be formulated as

$$\underset{(\boldsymbol{x},r)\in\mathbb{R}^{2m+1}}{\text{Minimize}} \; r \;\text{ subject to }\; G(\boldsymbol{x}, r) = 0, \tag{2.1}$$

where

$$G(\boldsymbol{x}, r) := \mathrm{Vol}(A \setminus \Omega(\boldsymbol{x}, r)) \tag{2.2}$$

and $\mathrm{Vol}(A \setminus \Omega(\boldsymbol{x}, r))$ denotes the volume of $A \setminus \Omega(\boldsymbol{x}, r)$.

The function G can be interpreted as the composition of a so-called *shape functional* $A \setminus \Omega \mapsto \mathrm{Vol}(A \setminus \Omega)$ with a function $(\boldsymbol{x}, r) \mapsto A \setminus \Omega(\boldsymbol{x}, r)$. Under some geometric conditions detailed in the next sections, the derivative of such a function can be computed using techniques of shape calculus and in particular via the concept of *shape derivative* [12, 17, 24, 25, 35]. In the forthcoming sections we prove that

$$\nabla G(\boldsymbol{x}, r) = -\left(\int_{\partial B(x_1,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A} \nu(z)\,dz, \cdots, \int_{\partial B(x_m,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A} \nu(z)\,dz, \int_{\partial\Omega(\boldsymbol{x},r)\cap A} dz\right)^{\top}, \tag{2.3}$$

where ν is the outward unit normal vector to $\Omega(\boldsymbol{x}, r)$. Note that $\nabla G(\boldsymbol{x}, r)$ is a block vector of size $2m + 1$ since ν is a vector with two components.

Remark 2.1. The results of this section may be extended to several other relevant situations. In particular, the case of different radii r_i can be obtained immediately. Say $\Omega(\boldsymbol{x}, \boldsymbol{r})$ is now a union of balls with different radii $\boldsymbol{r} := \{r_i\}_{i=1}^m$. Then the partial derivative with respect to r_i of the function $(\boldsymbol{x}, \boldsymbol{r}) \mapsto G(\boldsymbol{x}, \boldsymbol{r})$ is $\partial_{r_i} G(\boldsymbol{x}, \boldsymbol{r}) = -\int_{\partial B(x_i,r_i)\cap\partial\Omega(\boldsymbol{x},\boldsymbol{r})\cap A} dz$.

3. Proof of differentiability of G.

In this section we prove the formula (2.3) for ∇G. Assumption 3.1 below precludes that two balls be exactly superposed, that two balls be tangent, and that more than two balls' boundaries intersect at the same point. The assumption makes the task of proving that ∇G is given by (2.3) simpler. As it will be shown in Section 3.5, there are situations in which the assumption does not hold and ∇G is still given by (2.3); while there are also situations in which the assumption does not hold and ∇G does not exist. It is not a restrictive assumption, indeed if Assumption 3.1 is not satisfied for some configuration of $\Omega(\boldsymbol{x}, r)$, then it can be satisfied using an arbitrary small perturbation of r or $\boldsymbol{x} = \{x_i\}_{i=1}^m$. In other words, the assumption excludes a null-measure set of balls' configurations in $\mathbb{R}^{2m+1}$; and, thus, supposing it holds does not represent a practical issue of concern.

Assumption 3.1. The centers $\{x_i\}_{i=1}^m$ satisfy $\|x_i - x_j\| \neq 0$ and $\|x_i - x_j\| \neq 2r$ for all $1 \leq i, j \leq m$, $i \neq j$. Also, for all $1 \leq i, j, k \leq m$ with i, j, k pairwise distinct, we have $\partial B(x_i, r) \cap \partial B(x_j, r) \cap \partial B(x_k, r) = \emptyset$.

We consider two types of perturbed sets for the optimization. First of all, $\Omega(\boldsymbol{x}, r + t\delta r) \cap A$ arises from a perturbation $r + t\delta r$ of the radius while the centers $\boldsymbol{x}$ are fixed. Second, the sets $\Omega(\boldsymbol{x} + t\delta\boldsymbol{x}, r) \cap A$ correspond to translations of $B(x_i, r)$, i.e., to perturbations of the centers $\boldsymbol{x} + t\delta\boldsymbol{x} = \{x_i + t\delta x_i\}_{i=1}^m$ with a fixed radius r. The shape sensitivity analysis of the area of these perturbed domains is achieved through integration by substitution. The integral on the perturbed domain is pulled back onto the unperturbed domain, and then the derivative with respect to t of the integrand can be computed. In order to apply integration by substitution, one needs at least a bi-Lipschitz mapping between the reference domain and the perturbed domain. In the case of the radius perturbation for instance, the reference domain would be $\Omega(\boldsymbol{x}, r) \cap A$ and the perturbed domain $\Omega(\boldsymbol{x}, r + t\delta r) \cap A$. The objective is then to build a bi-Lipschitz mapping $T_t : \overline{\Omega(\boldsymbol{x}, r) \cap A} \to \mathbb{R}^2$ such that $T_t(\Omega(\boldsymbol{x}, r) \cap A) = \Omega(\boldsymbol{x}, r + t\delta r) \cap A$ and $T_t(\partial(\Omega(\boldsymbol{x}, r) \cap A)) = \partial(\Omega(\boldsymbol{x}, r + t\delta r) \cap A)$. In the case of center perturbations we are looking for a bi-Lipschitz mapping such that $T_t(\Omega(\boldsymbol{x}, r) \cap A) = \Omega(\boldsymbol{x} + t\delta\boldsymbol{x}, r) \cap A$ and $T_t(\partial(\Omega(\boldsymbol{x}, r) \cap A)) = \partial(\Omega(\boldsymbol{x} + t\delta\boldsymbol{x}, r) \cap A)$.

The main difficulty with building T_t is that $\Omega(\boldsymbol{x}, r) \cap A$ is defined via unions of balls and intersection with A. Taken individually, the transformations of $B(x_i, r)$ are simple translations and dilations. Unfortunately, it is not possible to simply sum these simple transformations up to obtain T_t, as this would yield a discontinuous T_t. Even though the construction of T_t is rather technical, the main ideas may be summarized as follows. The boundary of $\Omega(\boldsymbol{x}, r) \cap A$ can be decomposed into a union of curves and singular points where two circles meet or where a circle meets the boundary of A. The crucial observation is that for small t, the motion of a singular point is entirely determined by the translations or dilations of the balls $B(x_i, r)$. This can be easily understood by considering the intersection between two translating or dilating circles. On the smooth parts of the boundary of $\Omega(\boldsymbol{x}, r) \cap A$ there is more freedom for building T_t, using the fact that small displacements along a smooth subset of the boundary do not modify the shape globally. Thus, the main idea of the construction is to first determine T_t at the singular points using the implicit function theorem, and then to appropriately extend T_t to the smooth parts of $\partial(\Omega(\boldsymbol{x}, r) \cap A)$, so that T_t is bi-Lipschitz and models a translation or a dilation on each $B(x_i, r)$.

3.1. Construction of a mapping corresponding to a perturbation of the radius. Theorem 3.2 guarantees that under Assumption 3.1, and for sufficiently small t, the structure of $\Omega(\boldsymbol{x}, r+t\delta r)$ is stable, in the sense that $\partial\Omega(\boldsymbol{x}, r+t\delta r)$ is composed of a constant number of connected components and arcs, and that no topological changes occur, such as splitting, merging, or holes appearing in $\Omega(\boldsymbol{x}, r + t\delta r)$. This result is necessary for building a bi-Lipschitz mapping field between $\Omega(\boldsymbol{x}, r)$ and $\Omega(\boldsymbol{x}, r + t\delta r)$ in Theorem 3.3. If topological changes were occuring for instance, the perturbation of $\Omega(\boldsymbol{x}, r)$ could not be described by a bi-Lipschitz transformation. In this case, techniques of asymptotic analysis would have to be used to study the variation of G; several examples of such singular situations are presented in Section 3.5.

THEOREM 3.2. *Suppose that Assumption 3.1 holds. Then there exists $t_0 > 0$ such that for all $t \in [0, t_0]$ we have the following decomposition*

$$\partial\Omega(\boldsymbol{x}, r + t\delta r) = \bigcup_{k=1}^{\bar{k}} \mathcal{E}_k(t) \quad \text{and} \quad \mathcal{E}_k(t) = \bigcup_{\ell=1}^{\bar{\ell}_k} \mathcal{A}_{k,\ell}(t), \tag{3.1}$$

where $\bar{k} \geq 1$ and $\bar{\ell}_k \geq 1$ are independent of t, and $\{\mathcal{E}_k(t)\}_{k=1}^{\bar{k}}$ are the connected components of $\partial\Omega(\boldsymbol{x}, r+t\delta r)$. Also, for each $k = 1, \dots, \bar{k}$ and $\ell = 1, \dots, \bar{\ell}_k$, there exists a unique index $i_{k,\ell}$, independent of t, such that $\mathcal{A}_{k,\ell}(t)$ is a subarc of $\partial B(x_{i_{k,\ell}}, r+t\delta r)$ parameterized by an angle aperture $[\theta^{\text{in}}_{k,\ell}(t), \theta^{\text{out}}_{k,\ell}(t)]$, and $t \mapsto \theta^{\text{in}}_{k,\ell}(t)$, $t \mapsto \theta^{\text{out}}_{k,\ell}(t)$ are continuous functions on $[0, t_0]$.

Proof. Let $\mathcal{I} := \{1, \dots, m\}$ and introduce $\mathcal{Z}_i := \bigcup_{j\in\mathcal{I}, j\neq i} \partial B(x_i, r) \cap \partial B(x_j, r)$. Notice that $\mathcal{Z}_i \subset \partial B(x_i, r)$, that $\mathcal{Z}_i$ may be empty, and that the cardinal $\bar{\alpha}_i$ of $\mathcal{Z}_i$ is always even due to Assumption 3.1. The points of $\mathcal{Z}_i$ can be described, in local polar coordinates with the pole x_i, by angles $\theta_{i,\alpha} \in [0, 2\pi)$, with $\alpha = 1, \dots, \bar{\alpha}_i$. The points of $\mathcal{Z}_i$ may be ordered so that the angles $\theta_{i,\alpha}$ satisfy $0 \leq \theta_{i,1} < \theta_{i,2} < \cdots < \theta_{i,\bar{\alpha}_i} < 2\pi$.

Clearly, $\partial\Omega(\boldsymbol{x}, r)$ has a finite number $\bar{k}$ of connected components $\mathcal{E}_k$. We start by showing the decomposition into arcs

$$\partial\Omega(\boldsymbol{x},r)=\bigcup_{k=1}^{\bar{k}}\mathcal{E}_k \quad \text{and } \mathcal{E}_k=\bigcup_{\ell=1}^{\bar{\ell}_k}\mathcal{A}_{k,\ell}, \tag{3.2}$$

where each arc $\mathcal{A}_{k,\ell}$ satisfies $\mathcal{A}_{k,\ell}\subset\partial B(x_{i_\ell},r)$ for some index $i_\ell\in\mathcal{I}$, and the end-points of $\mathcal{A}_{k,\ell}$ are two consecutive points of $\mathcal{Z}_{i_\ell}$, in the order determined by the angles $\{\theta_{i_\ell,\alpha}\}_{\alpha=1}^{\bar{\alpha}_{i_\ell}}$. Note that the index i_ℓ is unique thanks to Assumption 3.1.

For a given $k\in\{1,\ldots,\bar{k}\}$, the first arc $\mathcal{A}_{k,1}\subset\mathcal{E}_k$ is chosen arbitrarily. If $\mathcal{Z}_{i_1}=\emptyset$, then we have $\mathcal{E}_k=\mathcal{A}_{k,1}=\partial B(x_{i_1},r)$, i.e., $\bar{\ell}_k=1$. If $\mathcal{Z}_{i_1}\neq\emptyset$, then $\mathcal{A}_{k,1}$ may be parameterized either by the angle aperture $[\theta_{i_1,\gamma_1},\theta_{i_1,\gamma_1+1}]$ for some index $1\le\gamma_1\le\bar{\alpha}_{i_1}-1$ or by the angle aperture $[\theta_{i_1,\bar{\alpha}_{i_1}},\theta_{i_1,1}+2\pi]$, since the endpoints of $\mathcal{A}_{k,1}$ are consecutive points on $\mathcal{Z}_{i_1}$. Let us call z_ℓ^{in} and z_ℓ^{out} the initial and final points of $\mathcal{A}_{k,\ell}$, respectively, where the supra-indices “in” and “out” refer to a counterclockwise motion along the circles. Then we have $z_1^{\text{out}}\in\mathcal{Z}_{i_1}\cap\mathcal{Z}_{i_2}\neq\emptyset$ for some $i_2\neq i_1$. Defining $z_2^{\text{in}}:=z_1^{\text{out}}$, this determines automatically the next arc $\mathcal{A}_{k,2}\subset\partial B(x_{i_2},r)$ with initial point z_2^{in} and final point z_2^{out}, so that z_2^{in} and z_2^{out} are two consecutive points of $\mathcal{Z}_{i_2}$. Given z_ℓ^{out} for some $\ell\ge1$, the procedure can be iterated by setting $z_{\ell+1}^{\text{in}}:=z_\ell^{\text{out}}$. The procedure ends when ℓ is such that $z_\ell^{\text{out}}=z_1^{\text{in}}$, yielding the decomposition of $\mathcal{E}_k$ in (3.2) with $\bar{\ell}_k=\ell$. A simple example illustrating this geometric procedure is provided in Figure 1.

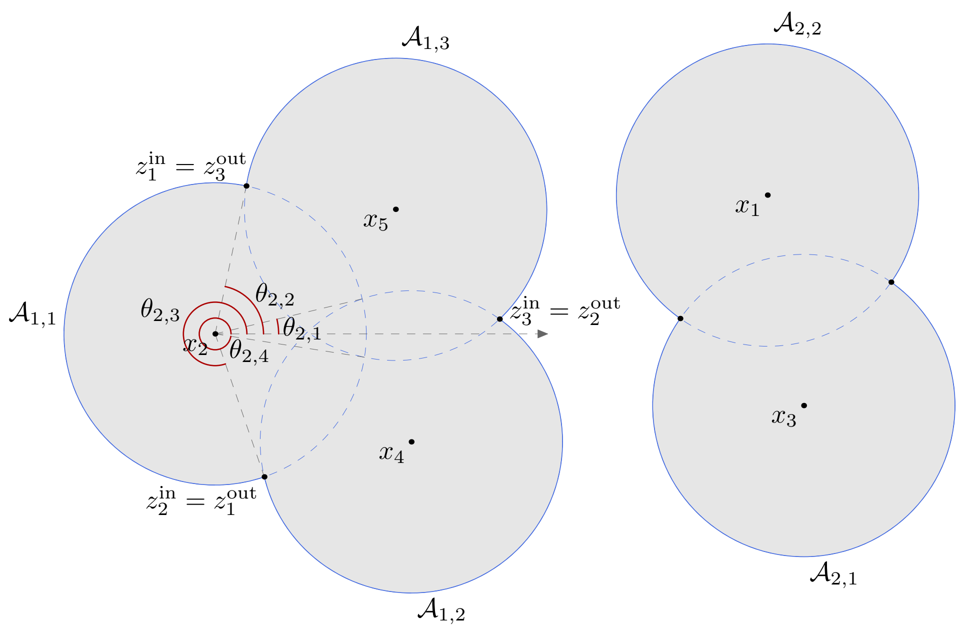


Fig. 1: An example of decomposition $\partial\Omega(\boldsymbol{x},r)=\bigcup_{k=1}^{\bar{k}}\mathcal{E}_k$ in (3.2), where $\mathcal{E}_k$ are the connected components of $\partial\Omega(\boldsymbol{x},r)$, with $\bar{k}=2$, $\mathcal{E}_k=\bigcup_{\ell=1}^{\bar{\ell}_k}\mathcal{A}_{k,\ell}$ with $\bar{\ell}_1=3$ and $\bar{\ell}_2=2$. The set $\mathcal{Z}_2:=\bigcup_{j\in\mathcal{I},j\neq2}\partial B(x_2,r)\cap\partial B(x_j,r)$ is composed of four points, hence $\bar{\alpha}_2=4$. The arc $\mathcal{A}_{1,1}$ is parameterized by the angle aperture $[\theta_{2,2},\theta_{2,3}]$, which corresponds to $i_1=2$ and $\gamma_1=2$ in the proof of Theorem 3.2.

Now that we have established the decomposition into subarcs (3.2) of the connected components of $\partial\Omega(\boldsymbol{x}, r)$, we prove that this decomposition is stable for small perturbations of the radius $r \mapsto r + t\delta r$. Let $(i, j) \in \mathcal{I}^2$, with $i \neq j$. If $\partial B(x_i, r) \cap \partial B(x_j, r) = \emptyset$, then thanks to Assumption 3.1 we also have $\partial B(x_i, r+t\delta r) \cap \partial B(x_j, r+t\delta r) = \emptyset$ for all $t \in [0, t_0]$ and $t_0 > 0$ sufficiently small. If $\partial B(x_i, r) \cap \partial B(x_j, r)$ is not empty then it is composed of exactly two points due to Assumption 3.1, i.e., $\partial B(x_i, r) \cap \partial B(x_j, r) = \{z_{ij1}, z_{ij2}\} \subset \mathcal{Z}_i$, with $z_{ij1} \neq z_{ij2}$. Using Assumption 3.1, it is clear that for all $\eta > 0$, there exists $t_0 > 0$ such that for all $t \in [0, t_0]$ we have the property $\partial B(x_i, r + t\delta r) \cap \partial B(x_j, r + t\delta r) = \{z_{ij1}(t), z_{ij2}(t)\}$, with

$$z_{ijk}(t) \in B(z_{ijk}, \eta) \text{ and } z_{ijk}(t) \to z_{ijk} \text{ as } t \to 0 \text{ for all } k \in \{1, 2\}. \tag{3.3}$$

We can also choose $\eta > 0$ sufficiently small so that

$$B(z_{i_1 j_1 k_1}, \eta) \cap B(z_{i_2 j_2 k_2}, \eta) = \emptyset \quad \text{for all } (i_1, j_1, k_1) \neq (i_2, j_2, k_2). \tag{3.4}$$

Now let us fix $\eta > 0$ and $t_0 > 0$ such that (3.3) and (3.4) are satisfied, and define

$$\mathcal{Z}_i(t) := \bigcup_{j \in \mathcal{I}, j \neq i} \partial B(x_i, r + t\delta r) \cap \partial B(x_j, r + t\delta r).$$

In view of (3.3) and (3.4), the function $p_t : \mathcal{Z}_i(t) \ni z \mapsto \operatorname{argmin}_{v \in \mathcal{Z}_i} \|z - v\| \in \mathcal{Z}_i$ defines a bijection between $\mathcal{Z}_i(t)$ and $\mathcal{Z}_i$: the injectivity of $p_t : \mathcal{Z}_i(t) \to \mathcal{Z}_i$ is due to Assumption 3.1 and the surjectivity is a consequence of (3.3). Thus, we conclude that for all $t \in [0, t_0]$ the points of $\mathcal{Z}_i(t)$ can be described, in local polar coordinates with the pole x_i, by angles $\theta_{i,\alpha}(t) \in [-\mu_0, 2\pi - \mu_0)$ for some $\mu_0 \geq 0$ independent of t, with $\alpha = 1, \ldots, \bar{\alpha}_i$, where $\bar{\alpha}_i = |\mathcal{Z}_i|$ is the cardinal of $\mathcal{Z}_i = \mathcal{Z}_i(0)$. For each $t \in [0, t_0]$, there is a bijection between the sets of angles $\{\theta_{i,\alpha}(t)\}_{\alpha=1}^{\bar{\alpha}_i}$ and $\{\theta_{i,\alpha}\}_{\alpha=1}^{\bar{\alpha}_i}$ and we have

$$-\mu_0 \leq \theta_{i,1}(t) < \theta_{i,2}(t) < \cdots < \theta_{i,\bar{\alpha}_i}(t) < 2\pi - \mu_0 \text{ for all } t \in [0, t_0]. \tag{3.5}$$

The points of $\mathcal{Z}_i(t)$ can be ordered using $\{\theta_{i,\alpha}(t)\}_{\alpha=1}^{\bar{\alpha}_i}$. Moreover, in view of (3.3) the functions $t \mapsto \theta_{i,\alpha}(t)$ are continuous on $[0, t_0]$ and we have $\theta_{i,\alpha}(0) = \theta_{i,\alpha}$ for $\alpha = 1, \ldots, \bar{\alpha}_i$.

Finally, we consider the decompositions

$$\partial\Omega(\boldsymbol{x}, r + t\delta r) = \bigcup_{k=1}^{\bar{k}(t)} \mathcal{E}_k(t) \text{ and } \mathcal{E}_k(t) = \bigcup_{\ell=1}^{\bar{\ell}_k(t)} \mathcal{A}_{k,\ell}(t),$$

where $\mathcal{E}_k(t)$ are the connected components of $\partial\Omega(\boldsymbol{x}, r + t\delta r)$. In view of the bijection between $\mathcal{Z}_i(t)$ and $\mathcal{Z}_i$, the bijection between $\{\theta_{i,\alpha}(t)\}_{\alpha=1}^{\bar{\alpha}_i}$ and $\{\theta_{i,\alpha}\}_{\alpha=1}^{\bar{\alpha}_i}$, and (3.5), we conclude that the set of subarcs of $\partial B(x_i, r)$ defined by the points of $\mathcal{Z}_i$ is also in bijection with the set of subarcs of $\partial B(x_i, r + t\delta r)$ defined by the points of $\mathcal{Z}_i(t)$. Then, employing the same procedure leading to the decompositions (3.2), we obtain that for all $t \in [0, t_0]$ we have $\bar{k}(t) = \bar{k}$ and $\bar{\ell}_k(t) = \bar{\ell}_k$ for all $k = 1, \ldots, \bar{k}$. Due to $\theta_{i,\alpha}(0) = \theta_{i,\alpha}$ and (3.5), we also have $\mathcal{A}_{k,\ell}(0) = \mathcal{A}_{k,\ell}$ and $\mathcal{A}_{k,\ell}(t) \subset \partial B(x_{i_\ell}, r + t\delta r)$ for all $t \in [0, t_0]$, where i_ℓ is the unique index such that $\mathcal{A}_{k,\ell} \subset \partial B(x_{i_\ell}, r)$. This proves the result. ∎

THEOREM 3.3. *Suppose that Assumption 3.1 holds. Then there exists $t_0 > 0$ such that for all $t \in [0, t_0]$, there exists a bi-Lipschitz mapping $T_t : \overline{\Omega(\boldsymbol{x}, r)} \to \mathbb{R}^2$ satisfying $T_t(\Omega(\boldsymbol{x}, r)) = \Omega(\boldsymbol{x}, r + t\delta r)$ and $T_t(\partial\Omega(\boldsymbol{x}, r)) = \partial\Omega(\boldsymbol{x}, r + t\delta r)$.*

Proof. First we provide a general formula for the angle $\vartheta(t)$, in local polar coordinates with the pole x_a, describing an intersection point of two circles $\partial B(x_a, r+t\delta r)$ and $\partial B(x_b, r+t\delta r)$, with $x_a, x_b \in \mathbb{R}^2, x_a \neq x_b$, and $\|x_a - x_b\| < 2r$. Introduce

$$\psi(t,\vartheta) := \|\zeta(t,\vartheta)\|^2 - (r+t\delta r)^2 \quad \text{with} \quad \zeta(t,\vartheta) := x_a - x_b + (r+t\delta r)\begin{pmatrix}\cos\vartheta \\ \sin\vartheta\end{pmatrix}.$$

Observe that $\vartheta \mapsto \zeta(t,\vartheta)$ is a parameterization of the circle $\partial B(x_a, r+t\delta r)$ in a coordinate system of center x_b, which means that the solutions of the equation $\psi(t,\vartheta) = 0$ describe the intersections between $\partial B(x_a, r+t\delta r)$ and $\partial B(x_b, r+t\delta r)$.

We compute $\partial_\vartheta \psi(0,\vartheta) = 2\langle \zeta(0,\vartheta), \partial_\vartheta \zeta(0,\vartheta)\rangle$ with

$$\zeta(0,\vartheta) = x_a - x_b + r\begin{pmatrix}\cos\vartheta \\ \sin\vartheta\end{pmatrix} \qquad \text{and} \qquad \partial_\vartheta \zeta(0,\vartheta) = r\begin{pmatrix}-\sin\vartheta \\ \cos\vartheta\end{pmatrix}. \tag{3.6}$$

Now let us select one of the two points in $\partial B(x_a, r+t\delta r) \cap \partial B(x_b, r+t\delta r)$ and let $\hat{\theta}$ be the corresponding angle in a polar coordinate system with the pole x_a. Since the conditions of Assumption 3.1 are satisfied, it is easy to see that

$$\partial_\vartheta \psi(0,\hat{\theta}) = \langle \zeta(0,\hat{\theta}), \partial_\vartheta \zeta(0,\hat{\theta})\rangle \neq 0. \tag{3.7}$$

Hence, the implicit function theorem can be applied to the function $(t,\vartheta) \mapsto \psi(t,\vartheta)$ in a neighborhood of $(0,\hat{\theta})$. This yields the existence, for t_0 sufficiently small, of a smooth function $t \mapsto \vartheta(t)$ in $[0,t_0]$ such that $\psi(t,\vartheta(t)) = 0$ in $[0,t_0]$ and $\vartheta(0) = \hat{\theta}$. We also have the derivative

$$\vartheta'(t) = -\frac{\partial_t \psi(t,\vartheta(t))}{\partial_\vartheta \psi(t,\vartheta(t))} = -\frac{\langle \zeta(t,\vartheta(t)), \partial_t \zeta(t,\vartheta(t))\rangle - (r+t\delta r)\delta r}{\langle \zeta(t,\vartheta(t)), \partial_\vartheta \zeta(t,\vartheta(t))\rangle}. \tag{3.8}$$

Now, let $\mathcal{A}$ be one of the two arcs composing the boundary of $B(x_a,r) \cup B(x_b,r)$, for instance $\mathcal{A} = \partial B(x_a,r) \cap (\overline{B(x_a,r) \cup B(x_b,r)})$, and let θ_a and θ_b be the angles parameterizing the endpoints of $\mathcal{A}$, with $\theta_a < \theta_b < \theta_a + 2\pi$ since $\mathcal{A}$ is not a circle. In view of the development above, for t_0 sufficiently small, we obtain two smooth functions $t \mapsto \theta_a(t)$ and $t \mapsto \theta_b(t)$, with $\theta_a(t) < \theta_b(t) < \theta_a(t) + 2\pi$ for all $t \in [0,t_0]$, where $\theta_a(t)$ and $\theta_b(t)$ are given by $\vartheta(t)$ with $\hat{\theta} = \theta_a$ and $\hat{\theta} = \theta_b$, respectively. The angles $\theta_a(t)$ and $\theta_b(t)$ are parameterizing the endpoints of one of the two arcs $\mathcal{A}(t)$ composing the boundary of $B(x_a, r+t\delta r) \cup B(x_b, r+t\delta r)$, with $\mathcal{A}(0) = \mathcal{A}$.

Next we define

$$\xi(t,\theta) := \alpha(t)(\theta - \theta_b) + \theta_b(t) \text{ for } (t,\theta) \in [0,t_0] \times [\theta_a,\theta_b] \qquad \text{and} \qquad \alpha(t) := \frac{\theta_b(t) - \theta_a(t)}{\theta_b - \theta_a}.$$

Then, for $\theta \in [\theta_a,\theta_b]$ we have $\xi(t,\theta) \in [\theta_a(t),\theta_b(t)]$ and $\xi(t,\theta)$ is a parameterization of $\mathcal{A}(t)$. We can parameterize a point $x \in \mathcal{A}(0)$ by

$$x = x_a + r\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}, \quad \text{and define} \quad \mathbb{T}_t(\theta) := x_a + (r+t\delta r)\begin{pmatrix}\cos\xi(t,\theta) \\ \sin\xi(t,\theta)\end{pmatrix}. \tag{3.9}$$

Writing $\xi(t,\theta) = \theta + \beta(t,\theta)$ with $\beta(t,\theta) := (\alpha(t)-1)(\theta - \theta_b(t))$, we observe that

$$\begin{pmatrix}\cos\xi(t,\theta) \\ \sin\xi(t,\theta)\end{pmatrix} = R(x_a,\beta(t,\theta))\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} = R(x_a,\beta(t,\theta))\nu,$$

where $R(x_a, \beta(t,\theta))$ is a rotation matrix of center x_a and angle $\beta(t,\theta)$, and ν is the outward unit normal vector to $\mathcal{A}$ at the point (r,θ) in polar coordinates with the pole x_a. Also, thanks to $\theta_a < \theta_b < \theta_a + 2\pi$ and $\theta \in [\theta_a, \theta_b]$, there exists a smooth bijection $\theta : \mathcal{A} \ni x \mapsto \theta(x) \in [\theta_a, \theta_b]$. Thus, using (3.9) we can define the function

$$T_t(x) := \mathbb{T}_t(\theta(x)) = x - r\nu(x) + (r + t\delta r) R(x_a, \beta(t, \theta(x)))\nu(x) \text{ for all } x \in \mathcal{A}. \tag{3.10}$$

In Figure 2 we provide an illustration of $\hat{\theta}$ and of the functions $T_t(x)$, $\xi(t,\theta)$.

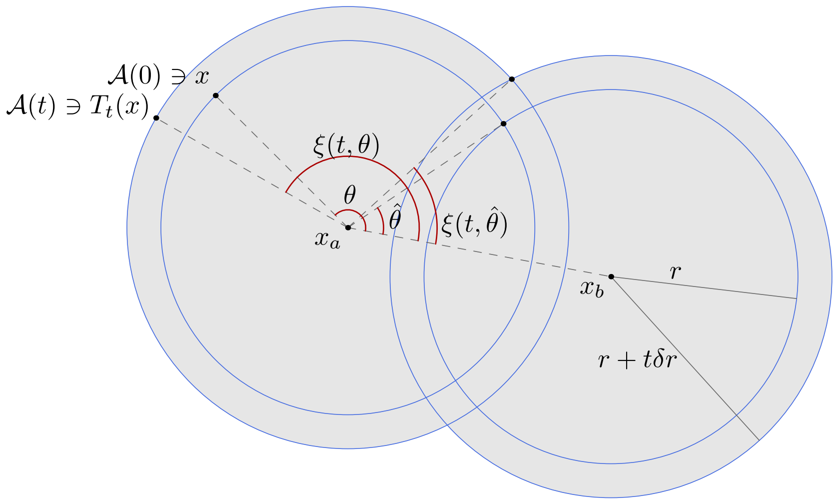


Fig. 2: Illustration of the geometric constructions in the proof of Theorem 3.3. For a given point x on the arc $\mathcal{A}(0)$, the polar coordinate $(r + t\delta r, \xi(t,\theta))$, with the pole x_a, represents the moving point $T_t(x) \in \mathcal{A}(t)$, and we have $T_0(x) = x$ and $\xi(0,\theta) = \theta \in [\theta_a, \theta_b]$. In the particular case $\theta = \hat{\theta}$, the polar coordinate $(r+t\delta r, \xi(t,\hat{\theta}))$ corresponds to an intersection point between $B(x_a, r + t\delta r)$ and $B(x_b, r + t\delta r)$, and we have $\xi(0,\hat{\theta}) = \hat{\theta} = \theta_a$.

Now we show that T_t is Lipschitz on $\mathcal{A}$. Using (3.9) we define

$$\mathbb{S}(t,\theta) := \mathbb{T}_t(\theta) - x_a - r\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = r\begin{pmatrix}\cos\xi(t,\theta) - \cos\theta\\ \sin\xi(t,\theta) - \sin\theta\end{pmatrix} + t\delta r\begin{pmatrix}\cos\xi(t,\theta)\\ \sin\xi(t,\theta)\end{pmatrix}. \tag{3.11}$$

Using $\xi(t,\theta) = \theta + \beta(t,\theta)$ we compute

$$\begin{aligned}\partial_\theta \mathbb{S}(t,\theta) &= r\begin{pmatrix}-\alpha(t)\sin\xi(t,\theta) + \sin\theta\\ \alpha(t)\cos\xi(t,\theta) - \cos\theta\end{pmatrix} + t\delta r\alpha(t)\begin{pmatrix}-\sin\xi(t,\theta)\\ \cos\xi(t,\theta)\end{pmatrix}\\ &= r\begin{pmatrix}c_1(t,\theta)\sin\theta + c_2(t,\theta)\cos\theta\\ -c_1(t,\theta)\cos\theta + c_2(t,\theta)\sin\theta\end{pmatrix} + t\delta r\alpha(t)\begin{pmatrix}-\sin\xi(t,\theta)\\ \cos\xi(t,\theta)\end{pmatrix},\end{aligned}$$

with

$$c_1(t,\theta) := 1 - \alpha(t)\cos\beta(t,\theta) \quad \text{and} \quad c_2(t,\theta) := -\alpha(t)\sin\beta(t,\theta). \tag{3.12}$$

Since $\alpha(0) = 1$ we have $\beta(0,\theta) = 0$, $c_1(0,\theta) = 0$ and $c_2(0,\theta) = 0$ for all $\theta \in [\theta_a, \theta_b]$. Thus, we obtain

$$c_1(t,\theta) = t\partial_t c_1(\xi_1,\theta) \quad \text{and} \quad c_2(t,\theta) = t\partial_t c_2(\xi_2,\theta) \tag{3.13}$$

for some $\xi_1 \in [0,t]$ and $\xi_2 \in [0,t]$. Then we compute

$$\partial_t \beta(t,\theta) = \alpha'(t)(\theta - \theta_b(t)) - (\alpha(t)-1)\theta_b'(t) \quad \text{and} \quad \alpha'(t) = \tfrac{\theta_b'(t)-\theta_a'(t)}{\theta_b - \theta_a}. \tag{3.14}$$

Using (3.6) and (3.8), we can show that for sufficiently small t_0 we have

$$|\theta_a'(t)| \le C_0 \text{ and } |\theta_b'(t)| \le C_0 \quad \text{for all } t \in [0,t_0], \tag{3.15}$$

where C_0 does not depend on t. Then, using (3.13) and (3.14) we get $|c_1(t,\theta)| \le C_1 t$ and $|c_2(t,\theta)| \le C_2 t$, where C_1 and C_2 are both independent of t and θ.

Finally, gathering (3.12), (3.13), (3.14), (3.15), and using a uniform bound on $\alpha(t)$ we obtain

$$\|\partial_\theta \mathbb{S}(t,\theta)\| \le C_3 t \quad \text{for all } t \in [0,t_0] \text{ and } \theta \in [\theta_a,\theta_b], \tag{3.16}$$

where C_3 is independent of t and θ.

Now we show that (3.16) implies the existence of a constant $C > 0$ such that $x \mapsto S(t,x) := T_t(x) - x$ is Lipschitz on $\mathcal{A}$ with Lipschitz constant Ct, i.e.,

$$\|S(t,x) - S(t,y)\| \le Ct\|x-y\|, \quad \forall (t,x,y) \in [0,t_0] \times \mathcal{A}^2. \tag{3.17}$$

Indeed if this were not the case, then there would exist a sequence $(t_n, x_n, y_n) \in [0,t_0] \times \mathcal{A}^2$ such that

$$\frac{\|S(t_n,x_n) - S(t_n,y_n)\|}{t_n\|x_n - y_n\|} \to \infty \quad \text{as } n \to +\infty. \tag{3.18}$$

Suppose that (3.18) holds. In view of (3.11) the numerator $\|S(t_n,x_n) - S(t_n,y_n)\|$ is uniformly bounded on $[0,t_0] \times \mathcal{A}^2$, thus we must have $t_n\|x_n - y_n\| \to 0$. We suppose that both $t_n \to 0$ and $\|x_n - y_n\| \to 0$, the other cases follow in a similar way. Using the compactness of $[0,t_0] \times \mathcal{A}^2$, we can extract a subsequence, still denoted (t_n, x_n, y_n) for simplicity, which converges towards $(0, x^\star, x^\star) \in [0,t_0] \times \mathcal{A}^2$. Then we write

$$\frac{\|S(t_n,x_n) - S(t_n,y_n)\|}{t_n\|x_n - y_n\|} = \underbrace{\frac{\|\mathbb{S}(t_n,\theta(x_n)) - \mathbb{S}(t_n,\theta(y_n))\|}{t_n|\theta(x_n) - \theta(y_n)|}}_{\text{bounded using (3.16) at } \theta(x^\star)} \underbrace{\frac{|\theta(x_n) - \theta(y_n)|}{\|x_n - y_n\|}}_{\text{bounded}},$$

where we have used the fact that $\|x_n - y_n\| = r \left\| \begin{pmatrix} \cos\theta(x_n) - \cos\theta(y_n) \\ \sin\theta(x_n) - \sin\theta(y_n) \end{pmatrix} \right\|$. This contradicts (3.18) which proves (3.17).

So far we have built a Lipschitz function T_t on an arc $\mathcal{A}$. We now proceed to build T_t on the entire boundary $\partial\Omega(\boldsymbol{x}, r)$. On each arc $\mathcal{A}_{k,\ell}(t) \subset \partial B(x_{i_{k,\ell}}, r + t\delta r)$ of the decomposition (3.1), T_t is built as in (3.10). Then due to (3.1) we have by construction that $T_t(\partial\Omega(\boldsymbol{x},r)) = \partial\Omega(\boldsymbol{x}, r + t\delta r)$. The continuity of T_t at the arc junctions is an immediate consequence of the definition of $\theta_a(t)$ and $\theta_b(t)$. Using the compactness of $\partial\Omega(\boldsymbol{x},r)$, the Lipschitz property (3.17) is valid on each connected component of $\partial\Omega(\boldsymbol{x},r)$. Using Kirszbraun's theorem [21] we can extend $x \mapsto S(t,x)$ to a Lipschitz function on $\Omega(\boldsymbol{x},r)$ with the same Lipschitz constant Ct.

Since $S(t,x) = T_t(x) - x$, this also defines an extension of $x \mapsto T_t(x)$ to $\Omega(\boldsymbol{x},r)$ and this shows that $x \mapsto T_t(x)$ is Lipschitz on $\Omega(\boldsymbol{x},r)$ with Lipschitz constant $1 + Ct$ for all $t \in [0,t_0]$. Since C is independent of t, we can choose t_0 sufficiently small so that $x \mapsto T_t(x)$ is invertible for all $t \in [0,t_0]$. The inverse is also Lipschitz on $\partial\Omega(\boldsymbol{x},r)$

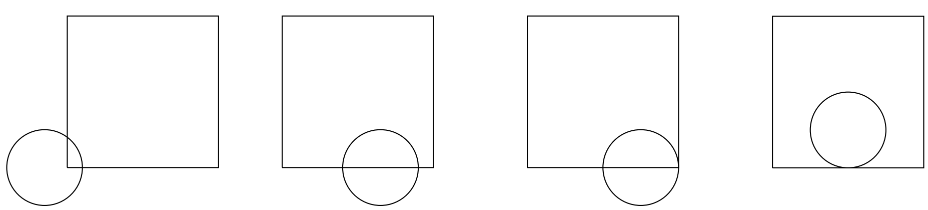
with Lipschitz constant $(1-Ct)^{-1}$ for all $t \in [0, t_0]$. This shows that $T_t : \overline{\Omega(\boldsymbol{x}, r)} \to T_t(\overline{\Omega(\boldsymbol{x}, r)})$ is bi-Lipschitz for all $t \in [0, t_0]$.

Finally, we prove that $T_t(\Omega(\boldsymbol{x}, r)) = \Omega(\boldsymbol{x}, r + t\delta r)$. Suppose first that $\partial\Omega(\boldsymbol{x}, r)$ has exactly one connected component. Since $T_t : \overline{\Omega(\boldsymbol{x}, r)} \to T_t(\overline{\Omega(\boldsymbol{x}, r)})$ is bi-Lipschitz it is also a homeomorphism, thus it maps interior points onto interior points and boundary points onto boundary points, i.e., $T_t(\Omega(\boldsymbol{x}, r))$ is the interior of $T_t(\partial\Omega(\boldsymbol{x}, r))$. According to the Jordan curve theorem [37], $\Omega(\boldsymbol{x}, r + t\delta r)$ is the interior of $\partial\Omega(\boldsymbol{x}, r + t\delta r)$. Since $\partial\Omega(\boldsymbol{x}, r + t\delta r) = T_t(\partial\Omega(\boldsymbol{x}, r))$ we also have that the interiors are the same, i.e., $T_t(\Omega(\boldsymbol{x}, r)) = \Omega(\boldsymbol{x}, r + t\delta r)$. The case where $\partial\Omega(\boldsymbol{x}, r)$ has several connected components follows in a similar way. ∎

3.2. Construction of a mapping corresponding to a perturbation of the centers. Unlike the case of the radius where the balls are dilated simultaneously, the computation of the partial derivatives of G with respect to x_i only requires the perturbation of one center x_i at a time. This can be modeled using a general setting where we build a mapping T_t between two sets $B(\hat{x}, r) \cap E$ and $B(\hat{x} + t\delta\hat{x}, r) \cap E$, where $E \subset \mathbb{R}^2$ and $B(\hat{x}, r)$ are compatible in the following sense. In what follows, a Lipschitz domain denotes an open, bounded set that is locally representable as the graph of a Lipschitz function; see [24, Def. 1] for a precise definition.

Definition 3.4. *Let ω_1, ω_2 be open subsets of $\mathbb{R}^2$. We call ω_1 and ω_2 compatible if $\omega_1 \cap \omega_2 \neq \emptyset$, ω_1 and ω_2 are Lipschitz domains, and the following conditions hold: (i) $\omega_1 \cap \omega_2$ is a Lipschitz domain; (ii) $\partial\omega_1 \cap \partial\omega_2$ is finite; (iii) $\partial\omega_1$ and $\partial\omega_2$ are locally smooth in a neighborhood of $\partial\omega_1 \cap \partial\omega_2$; (iv) $\tau_1(x) \cdot \nu_2(x) \neq 0$ for all $x \in \partial\omega_1 \cap \partial\omega_2$, where $\tau_1(x)$ is a tangent vector to $\partial\omega_1$ at x and $\nu_2(x)$ is a normal vector to $\partial\omega_2$ at x.*

Let us consider the following simple example: A is a square and $\Omega(\boldsymbol{x}, r)$ is a single ball, i.e. we have $m = 1$. Hence, the set of possible geometric configurations is three-dimensional. The sets A and $\Omega(\boldsymbol{x}, r)$ are always compatible in the sense of Definition 3.4, except when $\partial\Omega(\boldsymbol{x}, r)$ hits a corner of the square, or when $\partial\Omega(\boldsymbol{x}, r)$ and ∂A are tangent, as illustrated in Figure 3. This shows that the set of geometric configurations such that A and $\Omega(\boldsymbol{x}, r)$ are not compatible has measure zero in $\mathbb{R}^3$. Note that the examples depicted in Figure 3 are representative of the geometric configurations occurring in practice.



(a) compatible (b) compatible (c) not compatible (d) not compatible

Fig. 3: Compatibility of a ball ω_1 and a square ω_2 in the sense of Definition 3.4. In (c), condition (iii) of Definition 3.4 fails while, in (d), condition (iv) of Definition 3.4 fails.

The following result establishes the stability of the structure of $B(\hat{x}, r) \cap E$ under a small perturbation of the center $\hat{x}$ of the ball. We omit the proof of Theorem 3.5 which follows the same methodology as the proof of Theorem 3.2. Further, we build a bi-Lipschitz mapping in Theorem 3.6 between $B(\hat{x}, r) \cap E$ and $B(\hat{x} + t\delta\hat{x}, r) \cap E$;

see Figure 4.

Theorem 3.5. *Let $\hat{x}, \delta\hat{x} \in \mathbb{R}^2$, $E \subset \mathbb{R}^2$, and suppose that $B(\hat{x}, r)$ and E are compatible. Then there exists $t_0 > 0$ such that for all $t \in [0, t_0]$ we have the following decomposition*

$$\partial B(\hat{x} + t\delta\hat{x}, r) \cap E = \bigcup_{k=1}^{\bar{k}} \mathcal{A}_k(t), \tag{3.19}$$

where $\bar{k}$ is independent of t, and $\mathcal{A}_k(t)$ are subarcs of $\partial B(\hat{x} + t\delta\hat{x}, r)$ parameterized by an angle aperture $[\theta_{k,a}(t), \theta_{k,b}(t)]$, and $t \mapsto \theta_{k,a}(t)$, $t \mapsto \theta_{k,b}(t)$ are continuous functions on $[0, t_0]$.

Theorem 3.6. *Let $\hat{x}, \delta\hat{x} \in \mathbb{R}^2$, $E \subset \mathbb{R}^2$, and suppose that $B(\hat{x}, r)$ and E are compatible. Then there exists $t_0 > 0$ such that for all $t \in [0, t_0]$, there exists a bi-Lipschitz mapping $T_t : \overline{B(\hat{x}, r) \cap E} \to \mathbb{R}^2$ satisfying $T_t(B(\hat{x}, r) \cap E) = B(\hat{x} + t\delta\hat{x}, r) \cap E$ and $T_t(\partial(B(\hat{x}, r) \cap E)) = \partial(B(\hat{x} + t\delta\hat{x}, r) \cap E)$.*

Proof. We start by providing a general formula for the angle $\vartheta(t)$, in local polar coordinates with the pole $\hat{x} + t\delta\hat{x}$, describing an intersection point between the circle $\partial B(\hat{x} + t\delta\hat{x}, r)$ and ∂E. Let $z \in \partial B(\hat{x}, r) \cap \partial E$ and $\nu_E(z)$ the outward unit normal vector to E at z. Let ϕ be the oriented distance function to E, defined as $\phi(x) := d(x, E) - d(x, E^c)$, where $d(x, E)$ is the distance from x to the set E. Since we have assumed that $B(\hat{x}, r)$ and E are compatible, it follows that ∂E is locally smooth around the points $\partial B(\hat{x}, r) \cap \partial E$, hence there exists a neighborhood U_z of z such that the restriction of ϕ to U_z is smooth, $\phi(x) = 0$ and $\|\nabla\phi(x)\| = 1$ for all $x \in \partial E \cap U_z$.

Let $(r, \hat{\theta})$ denote the polar coordinates of z, with the pole $\hat{x}$. Introduce the function

$$\psi(t, \vartheta) = \phi\left(\hat{x} + t\delta\hat{x} + r\begin{pmatrix}\cos\vartheta \\ \sin\vartheta\end{pmatrix}\right).$$

We compute

$$\partial_\vartheta\psi(0, \hat{\theta}) = r\begin{pmatrix}-\sin\hat{\theta} \\ \cos\hat{\theta}\end{pmatrix} \cdot \nabla\phi\left(\hat{x} + r\begin{pmatrix}\cos\hat{\theta} \\ \sin\hat{\theta}\end{pmatrix}\right) = r\tau(z) \cdot \nabla\phi(z),$$

where $\tau(z)$ is a tangent vector to $\partial B(\hat{x}, r)$ at z. Since $B(\hat{x}, r)$ and E are compatible, $B(\hat{x}, r)$ is not tangent to ∂E and using $\|\nabla\phi(z)\| = 1$ we obtain $\tau(z) \cdot \nabla\phi(z) \neq 0$. Thus, we can apply the implicit function theorem and this yields the existence of a smooth function $[0, t_0] \ni t \mapsto \vartheta(t)$ with $\psi(t, \vartheta(t)) = 0$ and $\vartheta(0) = \hat{\theta}$. We also compute, using that $\nabla\phi(z) = \|\nabla\phi(z)\|\nu_E(z)$ since ϕ is the oriented distance function to ∂E,

$$\vartheta'(0) = -\frac{\partial_t\psi(0, \vartheta(0))}{\partial_\vartheta\psi(0, \vartheta(0))} = -\frac{\nabla\phi(z) \cdot \delta\hat{x}}{r\tau(z) \cdot \nabla\phi(z)} = -\frac{\nu_E(z) \cdot \delta\hat{x}}{r\tau(z) \cdot \nu_E(z)}. \tag{3.20}$$

Let $\mathcal{A}(t)$ be one of the arcs in the decomposition (3.19) parameterized by the angle aperture $[\theta_a(t), \theta_b(t)]$; we have dropped the index k for simplicity. The angles $\theta_a(t)$ and $\theta_b(t)$ are given by $\vartheta(t)$ with either $\hat{\theta} = \theta_a(0)$ or $\hat{\theta} = \theta_b(0)$.

Let $z_t \in \partial B(\hat{x} + t\delta\hat{x}, r) \cap \partial E$ be the point parameterized by polar coordinates $(r, \vartheta(t))$ with the pole $\hat{x} + t\delta\hat{x}$. Let $z_a \in \partial E \cap \overline{B(\hat{x}, r)}$ and $z_b \in \partial E \cap B(\hat{x}, r)^c$, both distinct from z and sufficiently close to z so that the subarc of ∂E between z_a and z_b is smooth. Let $\gamma : [0, 1] \to \mathbb{R}^2$ be a smooth parameterization of this arc satisfying $\gamma(0) = z_a$ and $\gamma(1) = z_b$. We may choose $t_0 > 0$ sufficiently small so that $z_t \in \gamma((0, 1))$

for all $t \in [0, t_0]$ and then define $\sigma(t) := \gamma^{-1}(z_t) > 0$ for all $t \in [0, t_0]$. Then define $T_t(\gamma(s)) := \gamma\left(\frac{\sigma(t)}{\sigma(0)} s\right)$ for all $0 < s < \sigma(0)$, or equivalently

$$T_t(x) := \gamma\left(\frac{\gamma^{-1}(z_t)}{\gamma^{-1}(z)}\gamma^{-1}(x)\right) \qquad \text{for all } x \in \gamma([0, \sigma(0)]). \tag{3.21}$$

Observe that $T_t(z_a) = z_a$, $T_t(z) = z_t$ and $T_t(\gamma([0, \sigma(0)])) = \gamma([0, \sigma(t)])$, which is precisely the smooth subarc of ∂E between z_a and z_t; see Figure 4 for an illustration of the construction of T_t. Then we define T_t in a similar way in neighborhoods of the other points of $\partial B(\hat{x}, r) \cap \partial E$. For all other points x of $\partial E \cap \overline{B(\hat{x}, r)}$ we set $T_t(x) = x$. Thus by construction we have $T_t(\partial E \cap \overline{B(\hat{x}, r)}) = \partial E \cap \overline{B(\hat{x} + t\delta\hat{x}, r)}$.

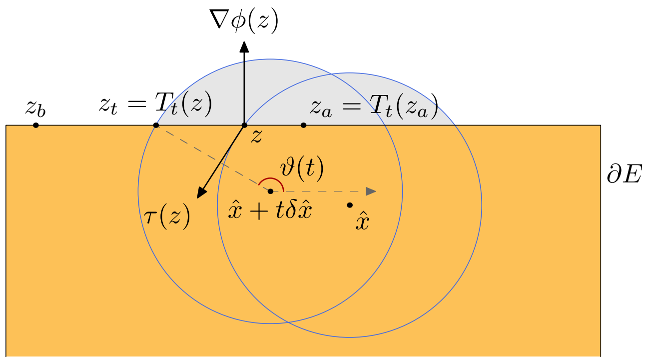


Fig. 4: Illustration of a key idea of the proof of Theorem 3.6. The point z belongs to $\partial B(\hat{x}, r) \cap \partial E$ while z_t belongs to $\partial B(\hat{x} + t\delta\hat{x}, r) \cap \partial E$. We build a transformation T_t mapping the subarc of ∂E between z_a and z to the subarc between z_a and z_t, and also mapping the arc of circle $\partial B(\hat{x}, r) \cap E$ to $\partial B(\hat{x} + t\delta\hat{x}, r) \cap E$. Note that the sets E and $B(\hat{x}, r)$ are compatible in the sense of Definition 3.4.

Let us define $S(t, \gamma(s)) := T_t(\gamma(s)) - \gamma(s)$ for all $0 < s < \sigma(0)$ and compute the derivative

$$\begin{aligned}
\partial_s[S(t, \gamma(s))] &= \frac{\sigma(t)}{\sigma(0)}\gamma'\left(\frac{\sigma(t)}{\sigma(0)} s\right) - \gamma'(s) \\
&= \gamma'\left(s + \left(\frac{\sigma(t)}{\sigma(0)} - 1\right) s\right) - \gamma'(s) + \left(\frac{\sigma(t)}{\sigma(0)} - 1\right)\gamma'\left(\frac{\sigma(t)}{\sigma(0)} s\right) \\
&= st\frac{\sigma'(\eta_0)}{\sigma(0)}\gamma''(\eta_1) + t\frac{\sigma'(\eta_0)}{\sigma(0)}\gamma'\left(\frac{\sigma(t)}{\sigma(0)} s\right),
\end{aligned}$$

with $\eta_0 \in [0, t]$ and $|\eta_1 - s| \leq st|\sigma'(\eta_0)/\sigma(0)|$. Using (3.20) and the smoothness of γ^{-1} we obtain that σ' is uniformly bounded on $[0, t_0]$. Using the smoothness of γ we obtain

$$\|\partial_s[S(t, \gamma(s))]\| \leq Ct \qquad \text{for all } t \in [0, t_0] \text{ and } 0 < s < \sigma(0),$$

for some constant C independent of t. Using a similar methodology as in the proof of Theorem 3.3, this proves that T_t is Lipschitz on $\partial E \cap \overline{B(\hat{x}, r)}$ with Lipschitz constant $1 + Ct$ for all $t \in [0, t_0]$.

The definition of T_t on the arc $\mathcal{A}(0)$ follows the same steps as in the proof of Theorem 3.3. For t_0 sufficiently small and $t \in [0, t_0]$, $\mathcal{A}(t)$ is an arc parameterized by $\theta_a(t)$ and $\theta_b(t)$, where $\theta_a(t)$ and $\theta_b(t)$ are given by $\vartheta(t)$ with $\hat{\theta} = \theta_a$ and $\hat{\theta} = \theta_b$, respectively. Then we define

$$T_t(x) := \hat{x} + t\delta\hat{x} + r\begin{pmatrix}\cos\xi(t,\theta)\\ \sin\xi(t,\theta)\end{pmatrix} \text{ with } x = \hat{x} + r\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \in \mathcal{A}(0), \tag{3.22}$$

where

$$\xi(t,\theta) := \alpha(t)(\theta-\theta_b) + \theta_b(t) \text{ for } (t,\theta)\in[0,t_0]\times[\theta_a,\theta_b] \text{ and } \alpha(t) := \tfrac{\theta_b(t)-\theta_a(t)}{\theta_b-\theta_a}. \tag{3.23}$$

The fact that T_t is Lipschitz on $\partial B(\hat{x},r)\cap\overline{E}$ with Lipschitz constant $1+Ct$, and the bi-Lipschitz extension of T_t to $\overline{B(\hat{x},r)\cap E}$ can be done as in the proof of Theorem 3.3.

We have already shown that $T_t(\partial E\cap\overline{B(\hat{x},r)}) = \partial E\cap\overline{B(\hat{x}+t\delta\hat{x},r)}$ and by construction we also have $T_t(\mathcal{A}(0)) = \mathcal{A}(t)$. This shows that $T_t(\partial(B(\hat{x},r)\cap E)) = \partial(B(\hat{x}+t\delta\hat{x},r)\cap E)$. The property $T_t(B(\hat{x},r)\cap E) = B(\hat{x}+t\delta\hat{x},r)\cap E$ is obtained in a similar way as in the proof of Theorem 3.3. ∎

3.3. Derivative of G with respect to the radius. To compute this derivative we consider a perturbation δr of the radius. The following result may be proven using Theorem 3.3 and a similar construction as in the proof of Theorem 3.6, therefore we omit its proof here. The result requires the following assumption.

Assumption 3.7. Sets $\Omega(\boldsymbol{x},r)$ and A are compatible.

Under Assumption 3.1, the set $\Omega(\boldsymbol{x},r)$ is Lipschitz, and if in addition the intersection of $\partial\Omega(\boldsymbol{x},r)$ and ∂A is empty, then Assumption 3.7 holds. Hence, in this particular case we can drop Assumption 3.7 in Theorem 3.8.

THEOREM 3.8. *Suppose that Assumption 3.1 and Assumption 3.7 hold. Then, there exists $t_0>0$ such that for all $t\in[0,t_0]$, there exists a bi-Lipschitz mapping $T_t : \overline{\Omega(\boldsymbol{x},r)\cap A} \to \mathbb{R}^2$ satisfying $T_t(\Omega(\boldsymbol{x},r)\cap A) = \Omega(\boldsymbol{x},r+t\delta r)\cap A$ and $T_t(\partial(\Omega(\boldsymbol{x},r)\cap A)) = \partial(\Omega(\boldsymbol{x},r+t\delta r)\cap A)$.*

Theorem 3.8 provides a mapping T_t that allows us to use the following integration by substitution:

$$\begin{aligned} G(\boldsymbol{x},r+t\delta r) &= \mathrm{Vol}(A\setminus\Omega(\boldsymbol{x},r+t\delta r)) = \mathrm{Vol}(A) - \mathrm{Vol}(A\cap\Omega(\boldsymbol{x},r+t\delta r)) \\ &= \mathrm{Vol}(A) - \int_{T_t(\Omega(\boldsymbol{x},r)\cap A)} dz = \mathrm{Vol}(A) - \int_{\Omega(\boldsymbol{x},r)\cap A} |\det DT_t(z)|\, dz. \end{aligned}$$

For sufficiently small t we have $|\det DT_t(z)| = \det DT_t(z)$ and $\partial_t \det DT_t(z)|_{t=0} = \operatorname{div} V(z)$, with $V := \partial_t T_t|_{t=0}$; see [12, 17, 35]. The set $\Omega(\boldsymbol{x},r)\cap A$ is Lipschitz due to Assumption 3.7, thus we may apply a divergence theorem in Lipschitz domains, see for instance [13, § 4.3, Theorem 1]. Denoting by ν the outward unit normal vector to $\Omega(\boldsymbol{x},r)$, this yields

$$\frac{d}{dt}G(r+t\delta r,\boldsymbol{x})\bigg|_{t=0} = -\int_{\Omega(\boldsymbol{x},r)\cap A} \operatorname{div} V(z)\, dz = -\int_{\partial(\Omega(\boldsymbol{x},r)\cap A)} V(z)\cdot\nu(z)\, dz. \tag{3.24}$$

The last integral in (3.24) is commonly called *boundary expression of the shape derivative* and the penultimate integral is called *volume expression*; see [12, 24, 35]. These expressions are standard for Lipschitz domains and vector fields V.

Now we compute V on $(\partial(\Omega(\boldsymbol{x},r)\cap A))\cap\partial B(x_i,r)$. In the case of $\Omega(\boldsymbol{x},r)\cap A$ we also have a decomposition into arcs similar to (3.1), and we can use (3.9) and

$\xi(0,\theta)=\theta$ to obtain

$$V = \partial_t T_t|_{t=0} = \delta r \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} + \partial_t\xi(0,\theta) r \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix} \text{ on } \partial(\Omega(\boldsymbol{x},r)\cap A)\cap\partial B(x_i,r),$$

where θ is the angle in polar coordinates with the pole x_i. Since $\nu = \binom{\cos\theta}{\sin\theta}$ on $\partial\Omega(\boldsymbol{x},r)\cap\partial B(x_i,r)$, we get $V\cdot\nu = \delta r$ on $\partial(\Omega(\boldsymbol{x},r)\cap A)\cap\partial B(x_i,r)$. We define T_t as in (3.21) or as the identity on $\partial A\cap\partial(\Omega(\boldsymbol{x},r)\cap A)$. Thus, it is easy to check that $V=\partial_t T_t|_{t=0}$ is tangent to $\partial A\cap\partial(\Omega(\boldsymbol{x},r)\cap A)$, so that $V\cdot\nu = 0$ on $\partial A\cap\partial(\Omega(\boldsymbol{x},r)\cap A)$. Gathering these results we obtain

$$\frac{d}{dt}G(r+t\delta r,\boldsymbol{x})\bigg|_{t=0} = -\int_{\partial(\Omega(\boldsymbol{x},r)\cap A)} V(z)\cdot\nu(z)\,dz = -\delta r\int_{\partial\Omega(\boldsymbol{x},r)\cap A} dz,$$

which gives the formula for the last entry of (2.3).

3.4. Derivative of G with respect to the centers. To compute this derivative we consider a perturbation $\delta\boldsymbol{x}$ such that $\delta x_i \neq 0$ for some index $i\in\mathcal{I}$ and $\delta x_j = 0$ for $j\neq i$, i.e., we consider the translation of only one ball in $\Omega(\boldsymbol{x},r)$. Introduce the notation $\Omega_{-i} := \bigcup_{j=1,j\neq i}^m B(x_j,r)$. Then we have the partition

$$\Omega(\boldsymbol{x}+t\delta\boldsymbol{x},r)\cap A = (\Omega_{-i}\cap A)\cup(B(x_i+t\delta x_i,r)\cap\Omega_{-i}^c\cap A).$$

We assume that the following condition holds.

Assumption 3.9. Sets $B(x_i,r)$ and $\Omega_{-i}^c\cap A$ are compatible.

Setting $E := \Omega_{-i}^c\cap A$, we can apply the results of Theorem 3.5 and Theorem 3.6 using Assumption 3.9. Let T_t be the bi-Lipschitz mapping given by Theorem 3.6. Then $T_t(B(x_i,r)\cap E) = B(x_i+t\delta x_i,r)\cap E$ and using an integration by substitution with the mapping T_t, we obtain

$$\begin{aligned}
G(\boldsymbol{x}+t\delta\boldsymbol{x},r) &= \mathrm{Vol}(A\setminus\Omega(\boldsymbol{x}+t\delta\boldsymbol{x},r)) = \mathrm{Vol}(A) - \mathrm{Vol}(\Omega(\boldsymbol{x}+t\delta\boldsymbol{x},r)\cap A)\\
&= \mathrm{Vol}(A) - \mathrm{Vol}(\Omega_{-i}\cap A) - \int_{B(x_i+t\delta x_i,r)\cap E} dz\\
&= \mathrm{Vol}(A) - \mathrm{Vol}(\Omega_{-i}\cap A) - \int_{T_t(B(x_i,r)\cap E)} dz\\
&= \mathrm{Vol}(A) - \mathrm{Vol}(\Omega_{-i}\cap A) - \int_{B(x_i,r)\cap E} |\det DT_t(z)|\,dz,
\end{aligned}$$

with $V := \partial_t T_t|_{t=0}$. The set $B(x_i,r)\cap E$ is Lipschitz due to Assumption 3.9, thus the divergence theorem yields

$$\frac{d}{dt}G(\boldsymbol{x}+t\delta\boldsymbol{x},r)\bigg|_{t=0} = -\int_{B(x_i,r)\cap E}\operatorname{div}V(z)\,dz = -\int_{\partial(B(x_i,r)\cap E)} V(z)\cdot\nu(z)\,dz,$$

where ν is the outward unit normal vector to $\Omega(\boldsymbol{x},r)$.

Now we compute V on $\partial(B(x_i,r)\cap E)$. Let $\mathcal{A}\subset\partial B(x_i,r)$ be an arc in the decomposition (3.19) at $t=0$, then using $\xi(0,\theta)=\theta$, (3.22) and (3.23) with $\hat{x}=x_i$ we obtain

$$V = \partial_t T_t|_{t=0} = \delta x_i + \partial_t\xi(0,\theta) r\begin{pmatrix}-\sin\xi(0,\theta)\\ \cos\xi(0,\theta)\end{pmatrix} = \delta x_i + \partial_t\xi(0,\theta) r\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix} \text{ on } \mathcal{A},$$

where θ is the angle in local polar coordinates with the pole x_i. Since $\nu = \left(\begin{smallmatrix}\cos\theta\\ \sin\theta\end{smallmatrix}\right)$ is a normal vector on $\mathcal{A}$, we get $V \cdot \nu = \delta x_i \cdot \nu$ on $\mathcal{A}$. On $\partial A \cap \partial(B(x_i, r) \cap E)$, T_t is defined by (3.21) or is the identity. Thus it is easy to check that $V = \partial_t T_t|_{t=0}$ is a tangent vector on $\partial A \cap \partial(B(x_i, r) \cap E)$, so that $V \cdot \nu = 0$ on $\partial A \cap \partial(B(x_i, r) \cap E)$. Gathering these results we obtain

$$\frac{d}{dt} G(\boldsymbol{x} + t\delta\boldsymbol{x}, r)\bigg|_{t=0} = -\int_{\partial(B(x_i,r)\cap E)} V(z) \cdot \nu(z)\, dz = -\delta x_i \cdot \int_{\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A} \nu(z)\, dz,$$

which gives the formula for the first $2m$ entries of (2.3).

3.5. Analysis of several singular cases. The theory in Sections 3.1–3.4 shows that (2.3) corresponds to the gradient of (2.2) under Assumptions 3.1, 3.7 and 3.9. From the practical point of view, the set of points $(\boldsymbol{x}, r)$ that do not satisfy these assumptions has measure zero in $\mathbb{R}^{2m+1}$; thus, it does not represent an issue. From a theoretical point of view, it is interesting to understand what may happen at these points.

Examples 3.10, 3.11, and 3.12 correspond to situations in which Assumption 3.1 does not hold. Example 3.10 corresponds to two tangent balls compactly contained in A; Example 3.11 corresponds to three balls whose boundaries intersect at a single point; and Example 3.12 corresponds to two superimposed balls, i.e., two balls whose boundaries intersect in an infinite number of points. In the first two cases, (2.3) still corresponds to the gradient of (2.2); while in the third case the gradient of (2.2) does not exist. Finally, Example 3.13 illustrates a situation in which Assumptions 3.7 and 3.9 do not hold and the gradient of (2.2) does not exist.

Example 3.10. Suppose $m = 2$, $\Omega(\boldsymbol{x} + t\delta\boldsymbol{x}, r) \subset A$ for all $t \in [0, t_0]$ and t_0 sufficiently small, and the two balls are tangent at $t = 0$, i.e., $\|x_1 - x_2\| = 2r$. Note that Assumption 3.1 is not satisfied. Two cases need to be considered to compute the gradient of G. First, if $\langle x_1 - x_2, \delta x_1 - \delta x_2\rangle \geq 0$ then it is clear that $B(x_1 + t\delta x_1, r) \cap B(x_2 + t\delta x_2, r) = \emptyset$ for all $t \in [0, t_0]$. Therefore $G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) = G(\boldsymbol{x}, r) = \mathrm{Vol}(A) - 2\pi r^2$ for all $t \in [0, t_0]$, and $\lim_{t\searrow 0}(G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) - G(\boldsymbol{x}, r))/t = 0$. Second, if $\langle x_1 - x_2, \delta x_1 - \delta x_2\rangle < 0$ then $B(x_1 + t\delta x_1, r) \cap B(x_2 + t\delta x_2, r) \neq \emptyset$ for all $t \in (0, t_0]$. Let us introduce the notation $a(t) := \mathrm{Vol}\,(B(x_1 + t\delta x_1, r) \cap B(x_2 + t\delta x_2, r))$. Using trigonometry we can show that $a(t) = 2r^2 \arccos\,(d(t)/2r) - d(t)\left(r^2 - (d(t)^2)/4\right)^{1/2}$, where $d(t) := \|x_1 + t\delta x_1 - (x_2 + t\delta x_2)\|$. It is convenient to rewrite this expression as

$$a(t) = 2r^2 \arccos\left((1 - g(t))^{1/2}\right) - 2r^2\left(g(t) + g(t)^2\right)^{1/2},$$

with $g(t) := -(2t\langle x_1 - x_2, \delta x_1 - \delta x_2\rangle + t^2\|\delta x_1 - \delta x_2\|^2)/(4r^2)$, $g(t) \geq 0$ for all $t \in [0, t_0]$ for t_0 small enough, $d(t) = 2r(1 - g(t))^{1/2}$, and $g'(0) = -\langle x_1 - x_2, \delta x_1 - \delta x_2\rangle)/(2r^2)$. After simplifications, we obtain $a'(t) = 2r^2\left(\frac{g(t)}{1-g(t)}\right)^{1/2} g'(t)$, and in particular $a'(0) = 0$. This shows that

$$\lim_{t\searrow 0} \frac{G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) - G(\boldsymbol{x}, r)}{t} = 0 \quad \text{ when } \langle x_1 - x_2, \delta x_1 - \delta x_2\rangle < 0.$$

Hence $\lim_{t\searrow 0}(G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) - G(\boldsymbol{x}, r))/t = 0$ in both cases. Proceeding in a similar way we can also show that $\lim_{t\searrow 0}(G(\boldsymbol{x}, r + t\delta r) - G(\boldsymbol{x}, r))/t = 4\pi r$. Thus $\nabla G(\boldsymbol{x}, r) = (0, \ldots, 0, 4\pi r)^\top$ in the case $\|x_1 - x_2\| = 2r$. It is easy to check that formula (2.3) also gives $\nabla G(\boldsymbol{x}, r) = (0, \ldots, 0, 4\pi r)^\top$ in this case. This indicates that, for the analyzed case, (2.3) is valid even without the satisfaction of Assumption 3.1. However, we had to use a different technique to prove that (2.3) holds, due to the fact that $G(\boldsymbol{x}+t\delta\boldsymbol{x}, r)$ takes different expressions depending on the sign of $\langle x_1 - x_2, \delta x_1 - \delta x_2\rangle$.

Example 3.11. Let $m = 3$ and x_1, x_2, x_3 be the vertices of an equilateral triangle such that the circles $\partial B(x_1, r)$, $\partial B(x_2, r)$ and $\partial B(x_3, r)$ intersect at a single point. Observe that Assumption 3.1 is not satisfied in this configuration. Then, if $\delta r < 0$ it is clear that $B(x_1, r+t\delta r) \cap B(x_2, r+t\delta r) \cap B(x_3, r+t\delta r) = \emptyset$, thus $\lim_{t\searrow 0}(G(\boldsymbol{x}, r + t\delta r) - G(\boldsymbol{x}, r))/t$ can be computed as in Section 3.3; and is equal to $\partial_r G(\boldsymbol{x}, r)$ given by (2.3). Now if $\delta r > 0$, the intersection $B(x_1, r+t\delta r) \cap B(x_2, r+t\delta r) \cap B(x_3, r+t\delta r)$ forms a well-known shape called Reuleaux triangle. An explicit calculation shows that this Reuleaux triangle is included in a ball whose area is of order $t^2\delta r^2$. Thus the first derivative of this area with respect to t at $t = 0$ is zero, hence the derivative $\lim_{t\searrow 0}(G(\boldsymbol{x}, r + t\delta r) - G(\boldsymbol{x}, r))/t$ is also equal to $\partial_r G(\boldsymbol{x}, r)$ given by (2.3) if $\delta r > 0$.

Example 3.12. Let $m = 2$, $\Omega(\boldsymbol{x}+t\delta\boldsymbol{x}, r) \subset A$ for $t \in [0, t_0]$ and t_0 sufficiently small, and the two balls are superposed at $t = 0$, i.e., $\|x_1 - x_2\| = 0$ and Assumption 3.1 is not satisfied in this configuration. Denoting $d(t) := t\|\delta x_1 - \delta x_2\|$ and $a(t) := \operatorname{Vol}\Omega(\boldsymbol{x} + t\delta\boldsymbol{x}, r)$, an explicit calculation yields $a(t) = \pi r^2 + 2r^2 \arctan\left(\frac{d(t)}{2r}\right) + rd(t)$, and consequently $a'(0) = 2r\|\delta x_1 - \delta x_2\|$.

First, expression (2.3) evaluated at $x_1 = x_2 = 0$ yields $\partial_{x_1} G(\boldsymbol{x}, r) = (0, 0)$ and $\partial_{x_2} G(\boldsymbol{x}, r) = (0, 0)$. Thus, in this case (2.3) does not give the correct value for the directional derivatives of G. Second, it is interesting to observe that, taking $\delta x_2 = 0$ to simplify, $a'(0) = 2r\|\delta x_1\|$ is equal to $\lim_{\varepsilon\searrow 0} -\partial_{x_1} G(\{x_1 + \varepsilon\delta x_1, x_2\}, r) \cdot \delta x_1$ with $\partial_{x_1} G(\{x_1 + \varepsilon\delta x_1, x_2\}, r)$ given by (2.3).

Example 3.13. Let $A = B(0, 1)$, $m = 1$, $\boldsymbol{x} = x_1 = 0$ and $r = 1$. Observe that in this example $\partial\Omega(\boldsymbol{x}, r) \cap \partial A = \partial B(0, 1)$ is not a finite set of points. Therefore Assumption 3.7 (precisely item (ii) in Definition 3.4) and Assumption 3.9 do not hold.

On the one hand, we have $G(\boldsymbol{x}, r) = 0$ and $G(\boldsymbol{x}, r+t\delta r) = 0$ for $t > 0$ and $\delta r > 0$. Thus we get in this case

$$\lim_{t\searrow 0} \frac{G(\boldsymbol{x}, r + t\delta r) - G(\boldsymbol{x}, r)}{t} = 0 \quad \text{when } \delta r > 0. \tag{3.25}$$

For $\delta r < 0$ we have $G(\boldsymbol{x}, r + t\delta r) = \pi(1 + t\delta r)^2 - \pi = \pi(2t\delta r + t^2\delta r^2)$, therefore

$$\lim_{t\searrow 0} \frac{G(\boldsymbol{x}, r + t\delta r) - G(\boldsymbol{x}, r)}{t} = 2\pi\delta r \quad \text{when } \delta r < 0. \tag{3.26}$$

This shows that G only has directional partial derivatives with respect to r at $\boldsymbol{x} = 0$. We observe that in this configuration, formula (2.3) yields the expression $\partial_r G(\boldsymbol{x}, r) = 0$ which is the same as the directional derivative (3.25). It is also interesting to observe that the other directional derivative (3.26) is equal to $\lim_{r\to 1, r<1} \partial_r G(\boldsymbol{x}, r)\delta r$ with $\partial_r G(\boldsymbol{x}, r)$ given by (2.3).

On the other hand, we have $G(\boldsymbol{x}, r) = 0$ and $G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) > 0$ for $t > 0$ and $\delta\boldsymbol{x} = \delta x_1 \neq 0$. An explicit calculation similar to the calculation in (3.12) yields

$$\lim_{t\searrow 0} \frac{G(\boldsymbol{x} + t\delta\boldsymbol{x}, r) - G(\boldsymbol{x}, r)}{t} = 2r\|\delta x_1\|. \tag{3.27}$$

However, expression (2.3) evaluated at $\boldsymbol{x} = x_1 = 0$ yields $\partial_{x_1} G(\boldsymbol{x}, r) = (0, 0)$. Thus, in this case (2.3) does not give the correct value for the directional derivatives of G.

Nevertheless, it can be checked that (3.27) is equal to $\lim_{\varepsilon \searrow 0} \partial_{x_1} G(\varepsilon \delta x_1, r) \cdot \delta x_1$ with $\partial_{x_1} G(\varepsilon \delta x_1, r)$ given by (2.3).

4. Numerical approximation of G and ∇G. In this paper we follow an optimize-then-discretize approach, i.e., we first find an expression for ∇G in the continuous setting and then discretize it. In Section 3 the gradient of G has been calculated analytically using techniques of nonsmooth shape calculus. We now show how the constraint G and its gradient ∇G may be approximated numerically. In the approximation, it is assumed that the region A to be covered is modeled by an oracle which, for a given point x, answers whether $x \in A$ or not. This is the most general way of defining a region $A \subset D$; and it reflects the fact that, from the practical point of view, A is considered to be a black-box from which no additional information is known other than the one given by the oracle. If more information about A were found available, such as an expression for its boundary, more efficient approximations could be devised.

We first prove a general result for the approximation of volumes of sets with piecewise smooth boundary using unions of square cells of a Cartesian grid. We show in Theorem 4.1 that this approximation is $O(h)$, where h is the size of the cells. In Algorithm 4.1 we implement this approach to approximate the constraint G. Then in Algorithm 4.2, a uniform discretization with step h_θ of the arcs of circles and a midpoint rule are used to approximate the integrals appearing in ∇G given by (2.3). We show in Theorem 4.2 that this approximation is of order $O(h_\theta)$.

4.1. Numerical approximation of G on a regular Cartesian grid. In this section we give a general result for the numerical approximation of volumes in a class $\mathcal{O}$ of domains with piecewise smooth boundary. In Algorithm 4.1 we implement this approach to approximate the constraint function $G(\boldsymbol{x}, r) = \mathrm{Vol}(A \setminus \Omega(\boldsymbol{x}, r))$. Suppose that $D = [0, L_D]^2$ is a square, we define $\mathcal{O}$ as the set of open and bounded subsets $\omega \subset D$ with piecewise smooth boundary $\partial\omega$, i.e.,

$$\partial\omega = \bigcup_{k=1}^{K} \overline{\Gamma_k}, \quad k = 1, \ldots, K, \tag{4.1}$$

where Γ_k is a smooth open or closed arc, $K < +\infty$, and $\overline{\Gamma_k} \cap \overline{\Gamma_j}$ is either empty or composed of one or two points, for all $j \neq k$, $j, k = 1, \ldots, K$. We observe that $\mathcal{O}$ contains non-Lipschitz domains such as domains with cracks and cusps, and also includes the sets used in the numerical experiments; see Section 5. Here $\mathrm{Per}(\partial\omega)$ denotes the perimeter of $\partial\omega$, with $\omega \in \mathcal{O}$, and χ_ω the indicator function of a set $\omega \subset \mathbb{R}^2$. In view of (4.1) and the smoothness of the Γ_k's we have $\mathrm{Per}(\partial\omega) = \sum_{k=1}^{K} \mathrm{Per}(\Gamma_k) < +\infty$.

Let the grid $\mathcal{L}$ be the set of points $z_{k,\ell} = ((k + 1/2)h, (\ell + 1/2)h)$ with $k, \ell = 0, \ldots, N-1$ and $h = L_D/N$. The point $z_{k,\ell}$ is the center of the cell $\mathcal{S}(k, \ell)$ defined by $\mathcal{S}(k,\ell) := \{(x_1, x_2) \in D \mid kh \leq x_1 \leq (k+1)h,\ \ell h \leq x_2 \leq (\ell+1)h\}$. The main idea of the proof of Theorem 4.1 is to approximate $\omega \in \mathcal{O}$ by a set ω_h that is the union of small squares of area h^2. As $h \to 0$, the symmetric difference $(\omega_h \setminus \omega) \cup (\omega \setminus \omega_h)$ behaves, roughly speaking, like a thin layer of thickness of order h concentrated on the boundary of ω. Thus, the area of the symmetric difference is of the order of the perimeter of ω times h, which allows to approximate the area of ω by the area of ω_h.

THEOREM 4.1. *Let $\omega \in \mathcal{O}$, then there exists $h_0 > 0$ such that, for all $0 < h \leq h_0$,*

$$\mathrm{Vol}(\omega) = h^2 \sum_{k,\ell=1}^{N} \chi_\omega(z_{k,\ell}) + E(h), \quad \textit{with} \quad |E(h)| \leq \sqrt{2} h\, \mathrm{Per}(\partial\omega) + \pi K h^2/2.$$

Proof. Introduce $\omega_h := \bigcup_{z_{k,\ell}\in\mathcal{L}\cap\omega} \mathcal{S}(k,\ell)$, then due to $\mathrm{Vol}(\mathcal{S}(k,\ell)) = h^2$ we have $\mathrm{Vol}(\omega_h) = h^2 \sum_{k,\ell=1}^N \chi_\omega(z_{k,\ell})$. Define $\omega_h^{\text{int}} := \{x \in \omega \mid d(x,\omega^c) \geq c_\delta h\}$ and $\omega_h^{\text{ext}} := \{x \in D \mid d(x,\omega) < c_\delta h\}$, where ω^c is the complement of ω, $c_\delta := \sqrt{2}/2 + \delta$ with $\delta > 0$, and $d(x,\omega)$ is the distance of x to the set ω. We clearly have $\omega_h^{\text{int}} \subset \omega \subset \omega_h^{\text{ext}}$. We show that, for h sufficiently small, we also have $\omega_h^{\text{int}} \subset \omega_h \subset \omega_h^{\text{ext}}$.

Suppose that $x \in \omega_h^{\text{int}}$, then $\|x - z\| \geq c_\delta h$ for all $z \in \omega^c$. There exists also $z_{k,\ell} \in \mathcal{L}$ such that $x \in \mathcal{S}(k,\ell)$. Since $z_{k,\ell}$ is the center of the cell $\mathcal{S}(k,\ell)$, we have $\|x - z_{k,\ell}\| \leq (\sqrt{2}/2)h$. Then $c_\delta h \leq \|x - z\| \leq \|x - z_{k,\ell}\| + \|z_{k,\ell} - z\|$ for all $z \in \omega^c$, which yields $\delta h = (c_\delta - \sqrt{2}/2)h \leq \|z_{k,\ell} - z\|$ for all $z \in \omega^c$. This shows that $z_{k,\ell} \in \mathcal{L}\cap\omega$ and since $x \in \mathcal{S}(k,\ell)$ this yields $x \in \omega_h$ by definition of ω_h; hence $\omega_h^{\text{int}} \subset \omega_h$.

Next we prove $\omega_h \subset \omega_h^{\text{ext}}$. Let $x \in \omega_h$, then by definition $x \in \mathcal{S}(k,\ell)$ for some $z_{k,\ell} \in \mathcal{L}\cap\omega$. Thus we have $\|x - z_{k,\ell}\| \leq (\sqrt{2}/2)h$ and $d(x,\omega) = \inf_{z\in\omega} \|x - z\| \leq \|x - z_{k,\ell}\| \leq (\sqrt{2}/2)h < c_\delta h$. This shows that $x \in \omega_h^{\text{ext}}$ and consequently $\omega_h \subset \omega_h^{\text{ext}}$. Consequently we have $\mathrm{Vol}(\omega_h^{\text{int}}) \leq \mathrm{Vol}(\omega_h) \leq \mathrm{Vol}(\omega_h^{\text{ext}})$.

Let $\Gamma_k^h := \{x \in D \mid d(x,\Gamma_k) < c_\delta h\}$ be the so-called *tubular neighborhood* of Γ_k, where $\Gamma_k \subset \partial\omega$ is one of the arcs in the decomposition (4.1). Now we prove that

$$\omega \setminus \left(\bigcup_{k=1}^K \Gamma_k^h \right) = \omega_h^{\text{int}} \subset \omega_h \subset \omega_h^{\text{ext}} \subset \omega \cup \left(\bigcup_{k=1}^K \Gamma_k^h \right). \tag{4.2}$$

We start with the rightmost inclusion. Let $x \in \omega_h^{\text{ext}} \setminus \omega$, then we have $d(x,\omega) < c_\delta h$ and consequently $d(x,\partial\omega) < c_\delta h$. Due to (4.1) this yields $d(x,\Gamma_k) < c_\delta h$ for some $k \in \{1,\dots,K\}$, and this proves $x \in \Gamma_k^h$. This proves indeed that $\omega_h^{\text{ext}} \subset \omega \cup \left(\cup_{k=1}^K \Gamma_k^h\right)$.

Now let $x \in \omega_h^{\text{int}}$, and suppose that $x \in \Gamma_k^h$ for some $k \in \{1,\dots,K\}$, then $d(x,\Gamma_k) < c_\delta h$ and consequently $d(x,\partial\omega) < c_\delta h$. This implies $d(x,\omega^c) < c_\delta h$ and then $x \notin \omega_h^{\text{int}}$, which is a contradiction. This shows that $x \in \omega \setminus \left(\cup_{k=1}^K \Gamma_k^h\right)$ and we have obtained the inclusion $\omega_h^{\text{int}} \subset \omega \setminus \left(\cup_{k=1}^K \Gamma_k^h\right)$. Conversely let $x \in \omega \setminus \left(\cup_{k=1}^K \Gamma_k^h\right)$. Suppose that $d(x,\omega^c) < c_\delta h$, then $d(x,\partial\omega) < c_\delta h$ and $d(x,\Gamma_k) < c_\delta h$ for some $k \in \{1,\dots,K\}$, which implies $x \in \Gamma_k^h$, a contradiction. This shows that $d(x,\omega^c) \geq c_\delta h$ and $x \in \omega_h^{\text{int}}$. Thus we have proved $\omega_h^{\text{int}} = \omega \setminus \left(\cup_{k=1}^K \Gamma_k^h\right)$.

Let $\mathcal{V}_k$ be the set of endpoints of the arc Γ_k, then $\mathcal{V}_k$ is included in the set of vertices of $\partial\omega$ and contains at most two vertices. For sufficiently small h, the tubular neighborhood Γ_k^h satisfies $\Gamma_k^h \subset \{x + \nu(x)\mu \mid x \in \overline{\Gamma_k}, |\mu| < c_\delta h\} \cup \bigcup_{z\in\mathcal{V}_k} \mathbb{B}(z)$, where $\mathbb{B}(z)$ is an open half-ball with center z and radius $c_\delta h$, and $\nu(x)$ is a normal vector to $\overline{\Gamma_k}$ at x. Using the results of [16, Ch. 1], there exists $h_{0,k} > 0$ independent of δ (for sufficiently small $\delta > 0$) such that

$$\mathrm{Vol}(\{x + \nu(x)\mu \mid x \in \overline{\Gamma_k}, |\mu| < c_\delta h\}) = 2c_\delta h \,\mathrm{Per}(\Gamma_k) \quad \forall h \text{ such that } 0 < h \leq h_{0,k}.$$

Since $\mathcal{V}_k$ contains at most two vertices, we obtain $\mathrm{Vol}(\Gamma_k^h) \leq 2c_\delta h\,\mathrm{Per}(\Gamma_k) + \pi(c_\delta h)^2$ for all h such that $0 < h \leq h_{0,k}$. As there is a finite number of arcs Γ_k, there exists $h_0 > 0$ such that $\sum_{k=1}^K \mathrm{Vol}(\Gamma_k^h) \leq 2c_\delta h\,\mathrm{Per}(\partial\omega) + \pi K(c_\delta h)^2$ for all h such that $0 < h \leq h_0$. From now on we suppose that $0 < h \leq h_0$. This yields

$$\mathrm{Vol}\left(\omega \cup \left(\bigcup_{k=1}^K \Gamma_k^h\right)\right) \leq \mathrm{Vol}(\omega) + \sum_{k=1}^K \mathrm{Vol}(\Gamma_k^h) \leq \mathrm{Vol}(\omega) + 2c_\delta h\,\mathrm{Per}(\partial\omega) + \pi K(c_\delta h)^2,$$

$$\mathrm{Vol}\left(\omega \setminus \left(\bigcup_{k=1}^K \Gamma_k^h\right)\right) \geq \mathrm{Vol}(\omega) - \sum_{k=1}^K \mathrm{Vol}(\Gamma_k^h) \geq \mathrm{Vol}(\omega) - 2c_\delta h\,\mathrm{Per}(\partial\omega) - \pi K(c_\delta h)^2.$$

Then, using (4.2) we obtain

$$-2c_\delta h\,\mathrm{Per}(\partial\omega) - \pi K(c_\delta h)^2 \le \mathrm{Vol}\left(\omega \setminus \left(\bigcup_{k=1}^{K} \Gamma_k^h\right)\right) - \mathrm{Vol}(\omega) = \mathrm{Vol}(\omega_h) - \mathrm{Vol}(\omega)$$

$$\le \mathrm{Vol}\left(\omega \cup \left(\bigcup_{k=1}^{K} \Gamma_k^h\right)\right) - \mathrm{Vol}(\omega) \le 2c_\delta h\,\mathrm{Per}(\partial\omega) + \pi K(c_\delta h)^2.$$

Finally this yields $|\,\mathrm{Vol}(\omega_h) - \mathrm{Vol}(\omega)| \le 2c_\delta h\,\mathrm{Per}(\partial\omega) + \pi K(c_\delta h)^2$ for all $0 < h \le h_0$. Passing to the limit $\delta \to 0$, this proves the result. ∎

Algorithm 4.1 NUMERICAL APPROXIMATION TO $\mathrm{Vol}(A \cap \Omega(\boldsymbol{x}, r))$. It considers a rectangular region D that contains A, computes a partition of D into rectangular cells with sides not larger than h, and returns the sum of the areas of the cells such that the center u of the cell satisfies $u \in A$ and u is within some ball.

Input: Region A, balls' radius r and centers $x_1, \dots, x_m$, precision $h > 0$, and bottom-left and top-right vertices $d^{\mathrm{bl}}, d^{\mathrm{tr}}$ of a rectangle $D \supseteq A$.
Output: Approximation to $\mathrm{Vol}(A \cap \Omega(\boldsymbol{x}, r))$. ($G(\boldsymbol{x}, r) = \mathrm{Vol}(A) - \mathrm{Vol}(A \cap \Omega(\boldsymbol{x}, r))$.)

```
Let n_x = ⌈(d_x^tr − d_x^bl)/h⌉, n_y = ⌈(d_y^tr − d_y^bl)/h⌉, h_x = (d_x^tr − d_x^bl)/n_x, h_y = (d_y^tr − d_y^bl)/n_y.
γ ← 0
for i = 1, ..., n_x do
    for j = 1, ..., n_y do
        Let u ← d^bl + ((i − 1/2)h_x, (j − 1/2)h_y)^T be the center of the (i, j)th cell.
        if u ∈ A and there exists k ∈ {1, ..., m} such that ‖x_k − u‖ ≤ r then
            γ ← γ + 1
return h_x h_y γ
```

4.2. Numerical approximation of ∇G. In this section we provide estimates for the numerical approximation of ∇G using Algorithm 4.2. First we observe that in (4.7), the balls $B(x_k, r)$ satisfying $\|x_i - x_k\| > 2r$ have no intersection with $\partial B(x_i, r)$, therefore we can simply ignore these. Second, if $\|x_i - x_k\| \le 2r$ there is an intersection between $\overline{B(x_i, r)}$ and $\overline{B(x_k, r)}$, so the first step of Algorithm 4.2 is to find the centers x_k satisfying $\|x_i - x_k\| \le 2r$.

Combining the results of Theorem 3.2 and Theorem 3.5 we obtain a decomposition into arcs similar to (3.1):

$$\partial\Omega(\boldsymbol{x}, r) \cap A = \bigcup_{k=1}^{\bar{k}} \mathcal{E}_k \quad \text{and} \quad \mathcal{E}_k = \bigcup_{\ell=1}^{\bar{\ell}_k} \mathcal{A}_{k,\ell}, \tag{4.3}$$

where $\bar{k} \ge 1$, $\bar{\ell}_k \ge 1$, and $\{\mathcal{E}_k\}_{k=1}^{\bar{k}}$ are the connected components of $\partial\Omega(\boldsymbol{x}, r) \cap A$. In particular we also have the decomposition

$$\partial B(x_i, r) \cap \partial\Omega(\boldsymbol{x}, r) \cap A = \bigcup_{\ell=1}^{\bar{\ell}_i} \mathcal{A}_\ell. \tag{4.4}$$

Let $\nu(z) = (\nu_1(z), \nu_2(z))$ be the outward normal vector on $\partial B(x_i, r)$, with $\nu_1(z) = \cos\theta$ and $\nu_2(z) = \sin\theta$, where θ is the angle in polar coordinates with the pole x_i. We obtain the following approximation result for Algorithm 4.2.

Theorem 4.2. *For $q=1,2$, denote by $\mathcal{G}_{i,q}$ the approximation of $(\partial G/\partial x_i)_q = \int_{\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A}\nu_q(z)\,dz$ given by Algorithm 4.2. Then we have the estimate*

$$\left|\int_{\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A}\nu_q(z)\,dz-\mathcal{G}_{i,q}\right|<h_\theta\bar{\ell}_i+\frac{2\pi\bar{\ell}_i h_\theta^2}{24r},\qquad \text{for } q=1,2, \tag{4.5}$$

where $\bar{\ell}_i$ is the number of arcs in the decomposition (4.4). Furthermore, let $\mathcal{G}_r$ be the approximation of $\partial G/\partial r=\int_{\partial\Omega(\boldsymbol{x},r)\cap A}dz$ given by Algorithm 4.2. Then we have the estimate

$$\left|\int_{\partial\Omega(\boldsymbol{x},r)\cap A}dz-\mathcal{G}_r\right|<\left(h_\theta+\frac{2\pi h_\theta^2}{24r}\right)\sum_{k=1}^{\bar{k}}\bar{\ell}_k. \tag{4.6}$$

Proof. Using (4.4) we compute

$$\int_{\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A}\nu_1(z)\,dz=\sum_{\ell=1}^{\bar{\ell}_i}\int_{\mathcal{A}_\ell}\nu_1(z)\,dz=\sum_{\ell=1}^{\bar{\ell}_i}r\int_{\theta_{\ell,a}}^{\theta_{\ell,b}}\cos(\theta)\,d\theta, \tag{4.7}$$

where $\theta_{\ell,a},\theta_{\ell,b}$ are the angles parameterizing the endpoints of the arc $\mathcal{A}_\ell$. In Algorithm 4.2 we do not compute the exact values of $\theta_{\ell,a},\theta_{\ell,b}$, therefore we cannot compute the integrals in (4.7) explicitly. Instead we use the midpoint rule with step length $\frac{h_\theta}{r}$ and check if the midpoints are in $\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A$. This corresponds to approximating the integrals $\widehat{I}_\ell:=\int_{\hat{\theta}_{\ell,a}}^{\hat{\theta}_{\ell,b}}\cos(\theta)\,d\theta$, for some $\hat{\theta}_{\ell,a},\hat{\theta}_{\ell,b}$ satisfying

$$|\hat{\theta}_{\ell,a}-\theta_{\ell,a}|\le\frac{h_\theta}{2r}\qquad\text{and}\qquad|\hat{\theta}_{\ell,b}-\theta_{\ell,b}|\le\frac{h_\theta}{2r}. \tag{4.8}$$

Let us denote I_ℓ the approximation of $\widehat{I}_\ell$ using the midpoint rule with step length $\frac{h_\theta}{r}$. We have the following estimate for this approximation:

$$\left|\int_{\hat{\theta}_{\ell,a}}^{\hat{\theta}_{\ell,b}}\cos(\theta)\,d\theta-I_\ell\right|\le\frac{(\hat{\theta}_{\ell,b}-\hat{\theta}_{\ell,a})h_\theta^2\sup_{\theta\in[\hat{\theta}_{\ell,a},\hat{\theta}_{\ell,b}]}|\cos(\theta)|}{24r^2}<\frac{2\pi h_\theta^2}{24r^2}.$$

Thus we compute

$$\left|\int_{\theta_{\ell,a}}^{\theta_{\ell,b}}\cos(\theta)\,d\theta-I_\ell\right|\le\left|\int_{\theta_{\ell,a}}^{\theta_{\ell,b}}\cos(\theta)\,d\theta-\int_{\hat{\theta}_{\ell,a}}^{\hat{\theta}_{\ell,b}}\cos(\theta)\,d\theta\right|+\left|\int_{\hat{\theta}_{\ell,a}}^{\hat{\theta}_{\ell,b}}\cos(\theta)\,d\theta-I_\ell\right|<\tfrac{h_\theta}{r}+\tfrac{2\pi h_\theta^2}{24r^2},$$

where we have used (4.8). Then using (4.7) we get

$$\left|\int_{\partial B(x_i,r)\cap\partial\Omega(\boldsymbol{x},r)\cap A}\nu_1(z)\,dz-\sum_{\ell=1}^{\bar{\ell}_i}rI_\ell\right|<h_\theta\bar{\ell}_i+\frac{2\pi\bar{\ell}_i h_\theta^2}{24r}. \tag{4.9}$$

We obtain the same estimate for $\nu_2(z)$. The estimate (4.6) is obtained in a similar way, summing over all the arcs in the decomposition (4.3). ∎

Algorithm 4.2 NUMERICAL APPROXIMATION TO $\nabla G(\boldsymbol{x}, r)$. It computes a discretization of $\partial B(x_i, r) \cap \partial\Omega(\boldsymbol{x}, r)$ for $i = 1, \dots, m$ and approximates the integrals in (2.3) using the composite middle point rule.

Input: Region A, balls' radius $r > 0$ and centers $x_1, \dots, x_m$, and precision $h > 0$.
Output: Approximation to $\nabla G(\boldsymbol{x}, r)$.

```
Let n_θ = ⌈2πr/h⌉ and h_θ = 2πr/n_θ. Set ∂G/∂r ← 0.
for i = 1, ..., m do
    Let K = {k ∈ {1, ..., m} \ {i} | ‖x_i − x_k‖ ≤ 2r}. Set ∂G/∂x_i ← 0.
    for ℓ = 1, ..., n_θ do
        θ ← (ℓ − 1/2) h_θ/r and u ← x_i + r(cos(θ), sin(θ))^T
        if u ∈ A and ‖u − x_k‖ ≥ r for all k ∈ K then
            ∂G/∂r ← ∂G/∂r + 1 and ∂G/∂x_i ← ∂G/∂x_i + (cos(θ), sin(θ))^T
return −h_θ((∂G/∂x_1)^T, ..., (∂G/∂x_m)^T, ∂G/∂r)^T
```

5. Numerical experiments. Problem (2.1), with the discretization $G_h(\boldsymbol{x}, r)$ of G computed by Algorithm 4.1, is a constrained nonlinear programming problem (with a linear objective function and a single difficult nonlinear constraint) of the form

$$\text{Minimize } f(\boldsymbol{x}, r) := r \text{ subject to } G_h(\boldsymbol{x}, r) = 0 \text{ and } r \geq 0 \tag{5.1}$$

that can be solved with an Augmented Lagrangian (AL) method [5]. In the present work we considered the safeguarded AL method Algencan [1, 5]. (See [6] for a numerical comparison with a state-of-the-art interior points method.) Algencan is based on the PHR AL function [19, 32, 33] that, for the considered problem, is defined by

$$L_\rho(\boldsymbol{x}, r, \lambda) = f(\boldsymbol{x}, r) + \frac{\rho}{2}\left[G_h(\boldsymbol{x}, r) + \frac{\lambda}{\rho}\right]^2, \tag{5.2}$$

for all $\rho > 0$, $r \geq 0$, and $\lambda \in \mathbb{R}$. Each iteration of the method consists in the approximate minimization of (5.2) subject to $r \geq 0$ followed by the update of the Lagrange multiplier λ and the penalty parameter ρ. The subproblem that consists in minimizing (5.2) subject to $r \geq 0$ is a bound-constrained minimization problem. In Algencan, bound-constrained subproblems are solved with an active-set method named Gencan [3] that uses Spectral Projected Gradient (SPG) [7] directions for "leaving faces" and a Newtonian approach "within the faces" (see [5, Ch. 9] for details). In the Newtonian approach, since second-order information is not available, Newtonian linear systems are solved with preconditioned conjugate gradients in which the Hessian-vector product is computed using an approximation to the Hessian of the AL described in [4].

The convergence theory of Algencan can be found in [5]. When applied to problem (5.1), on success, given feasibility and optimality tolerances $\varepsilon_{\text{feas}} > 0$ and $\varepsilon_{\text{opt}} > 0$, Algencan finds $(\boldsymbol{x}^\star, r^\star, \lambda^\star)$ with $r^\star > 0$ (clearly, the bound constraint $r \geq 0$ is non-active at any feasible solution) satisfying

$$\|\nabla f(\boldsymbol{x}^\star, r^\star) + \lambda^\star \nabla G_h(\boldsymbol{x}^\star, r^\star)\|_\infty \leq \varepsilon_{\text{opt}} \text{ and } \|G_h(\boldsymbol{x}^\star, r^\star)\|_\infty \leq \varepsilon_{\text{feas}}, \tag{5.3}$$

i.e., it finds a point that approximately satisfies KKT conditions for problem (5.1). In order to enhance the probability of finding an approximation to a global minimizer, we employed a simple multistart strategy. For each considered problem, Algencan was run a hundred times with an initial guess $(\boldsymbol{x}^0, r^0)$, $\boldsymbol{x}^0 = (x_1^0, \dots, x_m^0)$, such that $x_i^0 \in A \subset \mathbb{R}^2$ are random variables with uniform distribution and r^0 is a random variable with uniform distribution in $\frac{1}{m}[\frac{1}{2}, \frac{3}{2}]$. Note that $x_i \in A$ is not a constraint of the problem and that optimal solutions $(\boldsymbol{x}^\star, r^\star)$, $\boldsymbol{x}^\star = (x_1^\star, \dots, x_m^\star)$, exist such that $x_i^\star \notin A$ for some i. However, if $x_i \notin A$ and r is such that $B(x_i, r) \cap A = \emptyset$, then $\partial G/\partial x_i = 0$. Thus, if $x_i^0 \notin A$ and depending on the values of r^k along the optimization process, there exists the chance that the i-th ball stagnates in its initial configuration without contributing to the covering of A; producing in that case, with high probability, a suboptimal solution.

Algorithms 4.1 and 4.2 were implemented in Fortran 90. Algencan 3.1.1[1], which is also written in Fortran 90, was employed. All tests were conducted on a computer with a 3.4 GHz Intel Core i5 processor and 8GB 1600 MHz DDR3 RAM memory, running macOS Mojave (version 10.14.6). Code was compiled by the GFortran compiler of GCC (version 8.2.0) with the -O3 optimization directive enabled.

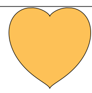
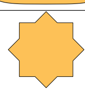
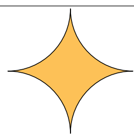
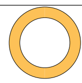
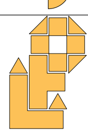
Heart		$A = \{(x,y)^T \in \mathbb{R}^2 \mid (x^2+y^2-1)^3 - x^2y^3 \le 0\}$
Soap		$A = \{(x,y)^T \in \mathbb{R}^2 \mid (2x/3)^4 + (2y)^4 \le 1\}$
Two squares		$A = A_1 \cup A_2$ $A_1 = \{(x,y)^T \in \mathbb{R}^2 \mid -0.5 \le x \le 0.5,\ -0.5 \le y \le 0.5\}$ $A_2 = \{(x,y)^T \in \mathbb{R}^2 \mid \max\{x+y, x-y, -x+y, -x-y\} \le \sqrt{2}/2\}$
Peaked star		$A = A_1 \cap (A_2 \cup A_3 \cup A_4 \cup A_5)$ $A_1 = \{(x,y)^T \in \mathbb{R}^2 \mid -0.5 \le x \le 0.5,\ -0.5 \le y \le 0.5\}$ $A_2 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x-0.5, y-0.5)^T\rVert_2 \ge 0.5\}$ $A_3 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x-0.5, y+0.5)^T\rVert_2 \ge 0.5\}$ $A_4 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x+0.5, y-0.5)^T\rVert_2 \ge 0.5\}$ $A_5 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x+0.5, y+0.5)^T\rVert_2 \ge 0.5\}$
Ring		$A = A_1 \cap A_2$ $A_1 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x,y)^T\rVert_2 \le 0.5\}$ $A_2 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x,y)^T\rVert_2 \ge 0.35\}$
Half ring		$A = A_1 \cap A_2 \cap A_3$ $A_1 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x,y)^T\rVert_2 \le 0.5\}$ $A_2 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x,y)^T\rVert_2 \ge 0.35\}$ $A_3 = \{(x,y)^T \in \mathbb{R}^2 \mid x \le 0\}$
Two half rings		$A = (A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_5 \cap A_6)$ $A_1 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x, y-0.175)^T\rVert_2 \le 0.25\}$ $A_2 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x, y-0.175)^T\rVert_2 \ge 0.10\}$ $A_3 = \{(x,y)^T \in \mathbb{R}^2 \mid x \le 0\}$ $A_4 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x, y+0.175)^T\rVert_2 \le 0.25\}$ $A_5 = \{(x,y)^T \in \mathbb{R}^2 \mid \lVert(x, y+0.175)^T\rVert_2 \ge 0.10\}$ $A_6 = \{(x,y)^T \in \mathbb{R}^2 \mid x \ge 0\}$
Disconnected		$A = \begin{array}{l} A_1(0,0,0,1,0,3) \cup A_1(0,0,1,3,0,1) \cup A_1(1.1,1.1,0,1,0,3) \cup \\ A_1(1.1,1.1,1,2,1,2) \cup A_1(2.2,3.2,0,1,0,1) \cup A_1(2.2,5.2,0,1,0,1) \cup \\ A_1(1.2,4.2,0,1,0,1) \cup A_1(3.2,4.2,0,1,0,1) \cup \\ A_2(0,3.1) \cup A_2(2.2,1.1) \cup A_3 \cup A_4 \cup A_5 \end{array}$

Table 1: Description of the considered regions A to be covered.

Table 1 shows the regions A to be covered that were considered in the numerical experiments. In the description of the "disconnected" region A,

$$\begin{array}{rcl}
A_1(\hat{x}, \hat{y}, \underline{x}, \bar{x}, \underline{y}, \bar{y}) &=& \{(x,y)^T \in \mathbb{R}^2 \mid \underline{x} \le x - \hat{x} \le \bar{x},\ \underline{y} \le y - \hat{y} \le \bar{y}\}, \\
A_2(\hat{x}, \hat{y}) &=& \{(x,y)^T \in \mathbb{R}^2 \mid y - \hat{y} \ge 0,\ y - \hat{y} \le \sqrt{3}(x - \hat{x}),\ y - \hat{y} \le -\sqrt{3}(x - \hat{x}) + \sqrt{3}\}, \\
A_3 &=& \{(x,y)^T \in \mathbb{R}^2 \mid x - 3.3 \ge 0,\ y - 5.3 \ge 0,\ (x - 3.3) + (y - 5.3) \le 1\}, \\
A_4 &=& \{(x,y)^T \in \mathbb{R}^2 \mid x - 1.1 \le 1,\ y - 5.3 \ge 0,\ -(x - 1.1) + (y - 5.3) \le 0\}, \\
A_5 &=& \{(x,y)^T \in \mathbb{R}^2 \mid x - 3.3 \ge 0,\ y - 3.1 \le 1,\ (x - 3.3) - (y - 3.1) \le 0\}.
\end{array}$$

[1]Algencan 3.1.1 is freely available at http://www.ime.usp.br/~egbirgin/tango/.

The sets A in Table 1 satisfy $A \subset D$, where D is a square of side 3 centered at the origin for the "heart", D is a rectangle with height 1 and width 3 centered at the origin for the "soap", D is a square of size $\sqrt{2}$ centered at the origin for the "two squares", D is a square of size 1 centered at the origin for the "peaked star", the "ring", the "half ring", and the "two half rings", and D is a rectangle with bottom-left corner $(0,0)$ and top-right corner $(4.3, 6.3)$ for the "disconnected" region. Taking into account the area of D, in all instances but the ones related to the "disconnected" region we considered $h = 10^{-3}$. In the "disconnected" region we considered $h = 5 \times 10^{-3}$. In Algencan, we set $\varepsilon_{\mathrm{feas}} = 0.1h$ (i.e., $\varepsilon_{\mathrm{feas}} = 5 \times 10^{-4}$ for the "disconnected" A and $\varepsilon_{\mathrm{feas}} = 10^{-4}$ in all other cases) and $\varepsilon_{\mathrm{opt}} = 10^{-1}$. The value of $\varepsilon_{\mathrm{feas}}$ is naturally related to the value of h — it would make no sense to require a tolerance much smaller than h for a constraint that is computed with precision $O(h)$.

Table 2 shows some performance metrics of the optimization procedure; while Figure 5 shows the solutions found. In the table, "trial" is the number of the initial guess (between 1 and 100) that let the optimization method find the best solution; "outit" and "innit" are the number of outer and inner iterations of the AL optimization method in that run; "Alg. 4.1" and "Alg. 4.2" are the number of calls to Algorithm 4.1 and Algorithm 4.2, i.e., the number of evaluations of G and ∇G, respectively; and "CPU time" is the CPU time in seconds. In the table and the figures, obtained radii are rounded to four decimal places. The heart-shape region A was taken from [2] where solutions for 3 and 7 balls with radii 0.8065 and 0.5524 are reported[2]. While solutions reported in [2] and here represent the same arrangement of the balls, radii obtained with the present approach are smaller. The covering of a ring with three balls is an example in which the centers of the balls are outside the region to be covered. The same phenomenon occurs with some balls in the instances with the "disconnected" region. All solutions found, except for the ones related to the "peaked star", are such that, looking with the naked eye, regions appear to be fully covered. If desired, improved solutions can be found at the expense of multiplying the effort by 100 every time h is divided by 10, since Algorithms 4.1 has time complexity $O(1/h^2)$. (As a side note, Algorithm 4.2 has time complexity $O(1/h)$.) Alternatively, better solutions could also be found by considering a dynamic multigrid approach that makes use of smaller values of h at critical places of the region to be covered. The "peaked star" case is particularly challenging because its peaks have a small area to perimeter ratio. Thus, the combination of a small but bounded-away-from zero discretization step $h > 0$ and a feasibility tolerance $\varepsilon_{\mathrm{feas}} > 0$ with the minimization of the balls' radius r is attracted by configurations with uncovered peaks.

Figure 6 shows the evolution of the optimization process in the arbitrary selected "two squares" problems with $m = 9$ balls, starting from the $70th$ initial guess (x^0, r^0) which is the one that leads to the best solution found. The top-left picture shows the initial guess. It is worth recalling that the balls' centers are randomly chosen within the region to be covered; while the initial radius is a random number in $[\frac{1}{2}, \frac{3}{2}]/m$. The picture shows that the initial radius is relatively small (with respect to the optimal one) and that the balls' center do not present any attractive feature. The initial value of the Lagrange multipliers is $\lambda^0 = 0$; and the penalty parameter ρ_0 is automatically chosen by the optimization solver in such a way that, in the augmented Lagrangian function (5.2), the term related to feasibility is one order of magnitude larger than the objective function; see [5, p.153]. This choice explains why in the first iteration the objective function (radius of the balls) is increased; while feasibility is reduced. The

[2]The values reported in [2, §5] correspond to r^2.

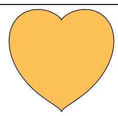
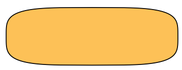
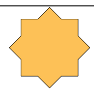
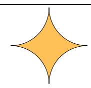
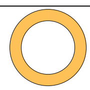
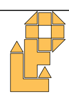
Region A	m	r^*	trial	outit	innit	Alg. 4.1	Alg. 4.2	CPU Time
	3	0.7949	100	20	155	2188	249	59.08
	7	0.5366	69	15	50	214	117	7.92
	11	0.4100	89	12	68	303	130	12.77
	15	0.3476	78	13	77	311	138	15.46
	3	0.6578	70	12	76	402	134	4.61
	7	0.4754	30	13	119	1228	185	20.11
	11	0.3564	61	13	72	261	132	6.12
	15	0.3154	69	13	80	447	140	12.77
	4	0.3810	91	11	40	222	90	2.78
	9	0.2474	70	11	45	197	94	3.18
	12	0.2064	32	10	66	346	112	6.16
	4	0.2317	82	20	136	2221	230	14.55
	5	0.1892	32	10	61	251	107	1.70
	9	0.1300	59	10	56	248	107	1.84
	3	0.4295	12	10	40	186	86	0.49
	7	0.2149	36	10	35	155	78	0.58
	11	0.1441	23	12	94	337	152	1.50
	3	0.2465	38	10	32	146	76	0.30
	7	0.1211	52	7	50	541	86	1.29
	11	0.0964	59	12	24	118	75	0.36
	3	0.2146	86	9	29	167	69	0.43
	7	0.1122	54	10	46	182	93	0.56
	11	0.0938	88	11	29	132	76	0.46
	3	1.7067	51	12	49	230	104	2.42
	7	1.1774	19	20	129	2331	215	27.14
	15	0.7820	36	20	112	1799	202	25.85

Table 2: Performance metrics of Algencan.

sequence of iterates shows that in iteration 5 the optimal arrangement has already been found; but the current radius $r^5 \approx 2.388 \times 10^{-1}$ produces a cover that leaves uncovered vertices that are visible to the naked eye. From iteration 5 to the end, increasing values of the penalty parameter produce successive iterates with increased radius and improved feasibility. The optimization process ends at iteration 11 when the required feasibility tolerance is reached.

Figure 7 shows the boxplot representation of the radii found in 100 runs of the "two squares" problem with $m \in \{4, 9, 12\}$. In the case $m = 4$, we have $r^* = r^{\min} = 0.3810$ and the median value is 0.3828, which is 4.7% larger than r^*. In the case $m = 9$, we have $r^* = r^{\min} = 0.2474$ and the median value is 0.2835, which is 14.6% larger than r^*. In the case $m = 12$, we have $r^* = r^{\min} = 0.2064$ and the median value is 0.2327, which is 12.7% larger than r^*. These quantities were computed over the runs that ended with a feasible solution, that were 100, 97, and 91, respectively. These figures, together with the small number of outliers, show that the optimization process is able to find "good quality solutions" in many cases, independently of the given initial guess.

Figure 8 and Table 3 show the results obtained by varying $h \in \{0.1, 10^{-2}, 10^{-3}, 10^{-4}\}$, with $\varepsilon_{\text{feas}} = 0.1h$ and $\varepsilon_{\text{opt}} = 0.1$, in problems "two squares" and "peaked star" with $m = 9$. The figures show that, the smaller the value of h, the higher the quality of the obtained cover. They also show that a region like the peaked star, which exhibits "small thin features", requires a smaller value of h, when compared to the two squares region, for a "reasonable" cover to be obtained. Recall that $h = 10^{-3}$ was considered in the numerical experiments shown in Figure 5 and Table 2. Figure 8 suggests that, to the naked eye, the solution obtained for the "two squares" problem with $m = 9$ considering $h = 10^{-2}$ is very similar to the one obtained with $h = 10^{-3}$. The same is true for all other problems that do not exhibit "small thin features" as the ones present in the "peaked star" problem; and due to the $O(1/h^2)$ time complexity of Algorithm 4.1, using $h = 10^{-2}$ is a hundred times faster than using $h = 10^{-3}$. This is why numerical experiments in Figure 5 and Table 2 should be understood as an illustration of the capabilities and limitations of the proposed approach; and the considered value of h must depend on the desired goal for the problem at hand. The last column in Table 3, titled "PMC" (that stands for "practical measure of time complexity") displays the total CPU time divided by the number of calls to Algorithm 4.1; and it roughly illustrates that the cost of approximating G is multiplied by 100 when h is divided by 10, as expected.

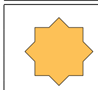
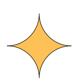
Region A	h	r^*	$\lvert G(\boldsymbol{x}^*, r^*)\rvert$	trial	outit	innit	Alg. 4.1	Alg. 4.2	CPU Time	PMC
	1e-1	0.2279	8.8e−03	7	20	133	3147	217	0.01	3e−06
	1e-2	0.2442	7.9e−04	32	9	41	223	81	0.05	1e−03
	1e-3	0.2474	5.9e−05	70	11	45	197	94	3.18	7e−02
	1e-4	0.2479	8.0e−06	85	15	83	326	150	502.64	6e+00
	1e-1	0.0762	1.0e−02	3	20	110	3894	193	0.01	9e−05
	1e-2	0.1191	1.0e−03	65	20	89	2386	175	0.19	2e−03
	1e-3	0.1300	6.9e−05	59	10	56	248	107	1.84	3e−02
	1e-4	0.1325	8.4e−06	7	11	79	317	137	224.49	3e+00

Table 3: Numerical results obtained varying $h \in \{0.1, 10^{-2}, 10^{-3}, 10^{-4}\}$, with $\varepsilon_{\text{feas}} = 0.1h$ and $\varepsilon_{\text{opt}} = 0.1$, in problems "two squares" and "peaked star" with $m = 9$. In the last column, PMC stands for "practical measurement of the complexity of Algorithm 4.1" and corresponds to the total CPU time divided by the number of calls to Algorithm 4.1.

Figure 9 corresponds to the covering of a union of two tangent unitary-diameter balls with $m = 2$ balls. This case is not covered by the theory, as the trivial solution neither satisfies Assumption 3.1 nor Assumption 3.7. Not satisfying Assumption 3.1 by having two tangent balls is in fact not an issue, since Example 3.10 shows that (2.3) still corresponds to ∇G in this case. On the other hand, not satisfying Assumption 3.7 because the intersection of the balls' border and the border of A contains infinitely many points *does* represent an issue. This is because Example 3.13 shows that, in this case, ∇G does not exist. Nevertheless, the depicted solution was found with a single run of the method, i.e., only one random initial guess. This example illustrates that a degenerate limit point does not affect the performance of the iterative optimization process that stops in finite time with a prescribed tolerance "before reaching the degenerate point that exists in the limit".

As a final illustration of the applicability of the proposed approach, Table 4 and Figure 10 show the application of the approach, with $h = 10^{-3}$, $\varepsilon_{\text{feas}} = 0.1h$, and $\varepsilon_{\text{opt}} = 0.1$, but considering 2,000 initial guesses instead of 100, to the covering of the union of three non-overlapping polygons that represent a sketch of America (a large non-convex polygon represents the continent; while two small convex polygons represent Cuba and Tierra del Fuego in the south of Argentina) [5, §13.2] with $m =$

Region A	m	r^*	trial	outit	innit	Alg. 4.1	Alg. 4.2	CPU Time
	15	0.08556	182	20	95	1844	180	2.45
	20	0.07459	1144	20	105	2806	190	3.77
	25	0.06728	1440	20	87	2086	170	2.81

Table 4: Performance metrics of Algencan applied to the problem of covering America with $m = 15, 20, 25$.

15, 20, 25 balls. In all three instances, the feasibility tolerance $\varepsilon_{\text{feas}} = 0.1h = 10^{-4}$ was reached. On the other hand, in all the three cases the method stopped because it reached the maximum of 20 outer iterations. (The same behavior can be observed in a few instances of other considered problems.) This means that the desired optimality tolerance ε_{opt} was not achieved. This could be a practical effect of reaching a solution at which ∇G is not well-defined. Solving instances with a larger number of balls or with more complex regions A faces two challenges of different nature. On the one hand, the larger the number of balls, the smaller the optimal radius; and a smaller optimal radius requires a smaller h to avoid very rough approximations. (Recall that the algorithm that approximates the constraint G has time complexity $O(1/h^2)$.) On the other hand, finding global solutions to more difficult instances (i.e., instances with more balls) might require a more elaborated *ad hoc* technique than the simple multi-start strategy adopted in the presented numerical experiments, including, for example, good quality initial guesses. Moreover, having at hand good quality initial guesses would require studying alternative nonlinear minimization methods because loosing feasibility of a potentially feasible initial guess is intrinsic to the augmented Lagrangian approach and to most of the practical nonlinear programming solvers.

6. Conclusions and future works. From the shape optimization perspective, the present work contributes to the study of shape sensitivity analysis with nonsmooth domains defined as the union of balls intersected with the domain to be covered. Studying and generalizing these techniques to three dimensions and to other types of nonsmooth domains will be a subject of future research. Regarding the covering problem, the numerical computation of the integrals defining the problem and its derivatives, as well as the availability of first-order information only, impaired the computation of precise solutions that may be required in some applications or for academic purposes. Therefore, a line for future research consists in deriving analytical expressions for second-order derivatives that would allow the application of quadratically convergent optimization methods. In some particular cases, like for example when the region A to be covered is a ball or a polygon, the objective function and its first- and second-order derivatives can be computed exactly using Voronoi diagrams.

Two related problem can also be tackled with the approach introduced in the current work. In one of them, each ball can have its own radius r_i and the goal may be minimizing the sum of the balls' perimeters, that is proportional to $\sum_{i=1}^m r_i$, or the sum of the balls' areas, that is proportional to $\sum_{i=1}^m r_i^2$. Redefining $\Omega(\boldsymbol{x}, \boldsymbol{r}) := \cup_{i=1}^m B(x_i, r_i)$ and $G(\boldsymbol{x}, \boldsymbol{r}) := \text{Vol}(A \setminus \Omega(\boldsymbol{x}, \boldsymbol{r}))$, where $\boldsymbol{r} := \{r_i\}_{i=1}^m$, expressions and algorithms to approximate $G(\boldsymbol{x}, \boldsymbol{r})$ and $\nabla G(\boldsymbol{x}, \boldsymbol{r})$ can be easily obtained with minor modifications to the introduced approach; see Remark 2.1. In the second related problem, the radius r common to all balls is fixed and the goal is to find the smallest $m \in \{1, 2, \dots\}$ and centers $x_1, \dots, x_m$ such that the balls cover a given region A. In this case, for a fixed radius r and a fixed number of balls $\bar{m}$, we define $G_{r,\bar{m}}(\boldsymbol{x}) := \mathrm{Vol}(A \setminus \Omega_{r,\bar{m}}(\boldsymbol{x}))$. The reasonable approach consists in starting with a large $\bar{m}$ and, while an $\boldsymbol{x}^*$ such that $G_{r,\bar{m}}(\boldsymbol{x}^*) = 0$ is found, reducing $\bar{m}$ by one. The feasible point $\boldsymbol{x}^*$ may be found by minimizing $F_{r,\bar{m}}(\boldsymbol{x}) := \frac{1}{2}\|G_{r,\bar{m}}(\boldsymbol{x})\|_2^2$, whose gradient is given by $\nabla F_{r,\bar{m}}(\boldsymbol{x}) = G_{r,\bar{m}}(\boldsymbol{x}) \nabla G_{r,\bar{m}}(\boldsymbol{x})$.

Acknowledgment:. The authors would like to thank the two reviewers and the associate editor, whose comments helped to significantly improve the quality of this work.

REFERENCES

[1] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt, *On augmented Lagrangian methods with general lower-level constraints*, SIAM Journal on Optimization, 18 (2008), pp. 1286–1309, https://doi.org/10.1137/060654797.

[2] E. G. Birgin, W. Gómez, G. Haeser, L. M. Mito, and D. S. Viana, *An augmented Lagrangian algorithm for nonlinear semidefinite programming applied to the covering problem*, Computational and Applied Mathematics, 39 (2020), p. article number 10, https://doi.org/10.1007/s40314-019-0991-5.

[3] E. G. Birgin and J. M. Martínez, *Large-scale active-set box-constrained optimization method with spectral projected gradients*, Computational Optimization and Applications, 23 (2002), pp. 101–125, https://doi.org/10.1023/A:1019928808826.

[4] E. G. Birgin and J. M. Martínez, *Structured minimal-memory inexact quasi-Newton method and secant preconditioners for augmented Lagrangian optimization*, Computational Optimization and Applications, 39 (2008), pp. 1–16, https://doi.org/10.1007/s10589-007-9050-z.

[5] E. G. Birgin and J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014, https://doi.org/10.1137/1.9781611973365.

[6] E. G. Birgin and J. M. Martínez, *Complexity and performance of an Augmented Lagrangian algorithm*, Optimization Methods and Software, 35 (2020), pp. 885–920, https://doi.org/10.1080/10556788.2020.1746962.

[7] E. G. Birgin, J. M. Martínez, and M. Raydan, *Nonmonotone spectral projected gradient methods on convex sets*, SIAM Journal on Optimization, 10 (2000), pp. 1196–1211, https://doi.org/10.1137/S1052623497330963.

[8] J. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag New York, New York, NY, 3rd ed., 1999, https://doi.org/10.1007/978-1-4757-6568-7.

[9] B. Csikós, *On the volume of the union of balls*, Discrete & Computational Geometry, 20 (1998), pp. 449–461, https://doi.org/10.1007/PL00009395.

[10] G. K. Das, S. Das, S. C. Nandy, and B. P. Sinha, *Efficient algorithm for placing a given number of base stations to cover a convex region*, Journal of Parallel and Distributed Computing, 66 (2006), pp. 1353–1358, https://doi.org/10.1016/j.jpdc.2006.05.004.

[11] M. C. Delfour and J.-P. Zolésio, *Structure of shape derivatives for nonsmooth domains*, Journal of Functional Analysis, 104 (1992), pp. 1–33, https://doi.org/10.1016/0022-1236(92)90087-Y.

[12] M. C. Delfour and J.-P. Zolésio, *Shapes and geometries: Metrics, analysis, differential calculus, and optimization*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2nd ed., 2011, https://doi.org/10.1137/1.9780898719826.

[13] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992, https://doi.org/10.1201/9780203747940.

[14] G. Frémiot, W. Horn, A. Laurain, M. Rao, and J. Sokołowski, *On the analysis of boundary value problems in nonsmooth domains*, Dissertationes Mathematicae, 462 (2009), pp. 1–149, https://doi.org/10.4064/dm462-0-1.

[15] P. Gangl, U. Langer, A. Laurain, H. Meftahi, and K. Sturm, *Shape optimization of an electric motor subject to nonlinear magnetostatics*, SIAM J. Sci. Comput., 37 (2015), pp. B1002–B1025, https://doi.org/10.1137/15100477X.

[16] A. Gray, *Tubes*, Springer Basel AG, Basel, Switzerland, 2nd ed., 2004, https://doi.org/10.1007/978-3-0348-7966-8.

[17] A. Henrot and M. Pierre, *Shape variation and optimization*, vol. 28 of EMS Tracts in

Mathematics, European Mathematical Society (EMS), Zürich, 2018, https://doi.org/10.4171/178, https://doi.org/10.4171/178. A geometrical analysis, English version of the French publication [MR2512810] with additions and updates.

[18] A. Heppes and J. B. M. Melissen, *Covering a rectangle with equal circles*, Periodica Mathematica Hungarica, 34 (1997), pp. 65–81, https://doi.org/10.1023/A:1004224507766.

[19] M. R. Hestenes, *Multiplier and gradient methods*, Journal of Optimization Theory and Applications, 4 (1969), pp. 303–320, https://doi.org/10.1007/BF00927673.

[20] C. Ho-Lun and H. Edelsbrunner, *Area and perimeter derivatives of a union of disks*, in Computer Science in Perspective: Essays Dedicated to Thomas Ottmann, R. Klein, H.-W. Six, and L. Wegner, eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2003, pp. 88–97, https://doi.org/10.1007/3-540-36477-3_7.

[21] M. Kirszbraun, *Über die zusammenziehende und lipschitzsche transformationen*, Fundamenta Mathematicae, 22 (1934), pp. 77–108, https://doi.org/10.4064/fm-22-1-77-108.

[22] J. Lamboley, A. Novruzi, and M. Pierre, *Estimates of first and second order shape derivatives in nonsmooth multidimensional domains and applications*, Journal of Functional Analysis, 270 (2016), pp. 2616–2652, https://doi.org/10.1016/j.jfa.2016.02.013.

[23] J. Lamboley and M. Pierre, *Structure of shape derivatives around irregular domains and applications*, Journal of Convex Analysis, 14 (2007), pp. 807–822, http://www.heldermann-verlag.de/jca/jca14/jca0646_b.pdf.

[24] A. Laurain, *Distributed and boundary expressions of first and second order shape derivatives in nonsmooth domains*, Journal de Mathématiques Pures et Appliquées, 134 (2020), pp. 328–368, https://doi.org/10.1016/j.matpur.2019.09.002.

[25] A. Laurain and K. Sturm, *Distributed shape derivative via averaged adjoint method and applications*, ESAIM: Mathematical Modelling and Numerical Analysis, 50 (2016), pp. 1241–1267, https://doi.org/10.1051/m2an/2015075.

[26] L. Liberti, N. Maculan, and Y. Zhang, *Optimal configuration of gamma ray machine radiosurgery units: the sphere covering subproblem*, Optimization Letters, 3 (2009), pp. 109–121, https://doi.org/10.1007/s11590-008-0095-4.

[27] J. B. M. Melissen, *Loosest circle coverings of an equilateral triangle*, Mathematics Magazine, 70 (1997), pp. 118–124, https://doi.org/10.1080/0025570X.1997.11996514.

[28] J. B. M. Melissen and P. C. Schuur, *Improved coverings of a square with six and eight equal circles*, The Electronic Journal of Combinatorics, 3 (1996), p. R32, https://doi.org/10.37236/1256.

[29] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag New York, New York, NY, 2nd ed., 2006, https://doi.org/10.1007/978-0-387-40065-5.

[30] K. J. Nurmela, *Conjecturally optimal coverings of an equilateral triangle with up to 36 equal circles*, Experimental Mathematics, 9 (2000), pp. 241–250, https://doi.org/10.1080/10586458.2000.10504649.

[31] K. J. Nurmela and P. R. J. Östergård, *Covering a square with up to 30 equal circles*, Tech. Report Technical Report HUT-TCS-A62, Helsinki University of Technology, Laboratory for Theoretical Computer Science, 2000, http://www.tcs.hut.fi/Publications/info/bibdb.HUT-TCS-A62.shtml.

[32] M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, in Optimization, R. Fletcher, ed., Academic Press, New York, NY, 1969, pp. 283–298, https://ci.nii.ac.jp/naid/20000915081/en/.

[33] R. T. Rockafellar, *Augmented lagrange multiplier functions and duality in nonconvex programming*, SIAM Journal on Control and Optimization, 12 (1974), pp. 268–285, https://doi.org/10.1137/0312021.

[34] C. A. Rogers, *Packing and covering*, Cambridge University Press, Cambridge, 1964, https://doi.org/10.1017/S0013091500008877.

[35] J. Sokołowski and J.-P. Zolésio, *Introduction to Shape Optimization*, Springer Series in Computational Mathematics, Springer-Verlag, Berlin, Heidelberg, 1992, https://doi.org/10.1007/978-3-642-58106-9.

[36] Y. G. Stoyan and V. M. Patsuk, *Covering a compact polygonal set by identical circles*, Computational Optimization and Applications, 46 (2010), pp. 75–92, https://doi.org/10.1007/s10589-008-9191-8.

[37] H. Tverberg, *A proof of the Jordan curve theorem*, Bulletin of the London Mathematical Society, 12 (1980), pp. 34–38, https://doi.org/10.1112/blms/12.1.34.

[38] H. M. Venceslau, D. C. Lubke, and A. E. Xavier, *Optimal covering of solid bodies by spheres via the hyperbolic smoothing technique*, Optimization Methods and Software, 30 (2015), pp. 391–403, https://doi.org/10.1080/10556788.2014.934686.

[39] A. E. Xavier, *Penalização Hiperbólica: Um novo método para resolução de problemas de*

otimização, master's thesis, COPPE/UFRJ, Rio de Janeiro, RJ, Brazil, 1982, https://www.cos.ufrj.br/index.php/pt-BR/publicacoes-pesquisa/details/15/1701.

[40] A. E. Xavier and A. A. Fernandes de Oliveira, *Optimal covering of plane domains by circles via hyperbolic smoothing*, Journal of Global Optimization, 31 (2005), pp. 493–504, https://doi.org/10.1007/s10898-004-0737-8.

[41] C. T. Zahn, *Black box maximization of circular coverage*, Journal of Research of the National Bureau of Standards – Section B Mathematics and Mathematical Physics, 66B (1962), pp. 181–216, https://doi.org/10.6028/jres.066b.020.

[42] G. Zoutendijk, *Methods of Feasible Directions: a Study in Linear and Non-Linear Programming*, Elsevier, Amsterdam, 1970, http://hdl.handle.net/2042/29608.

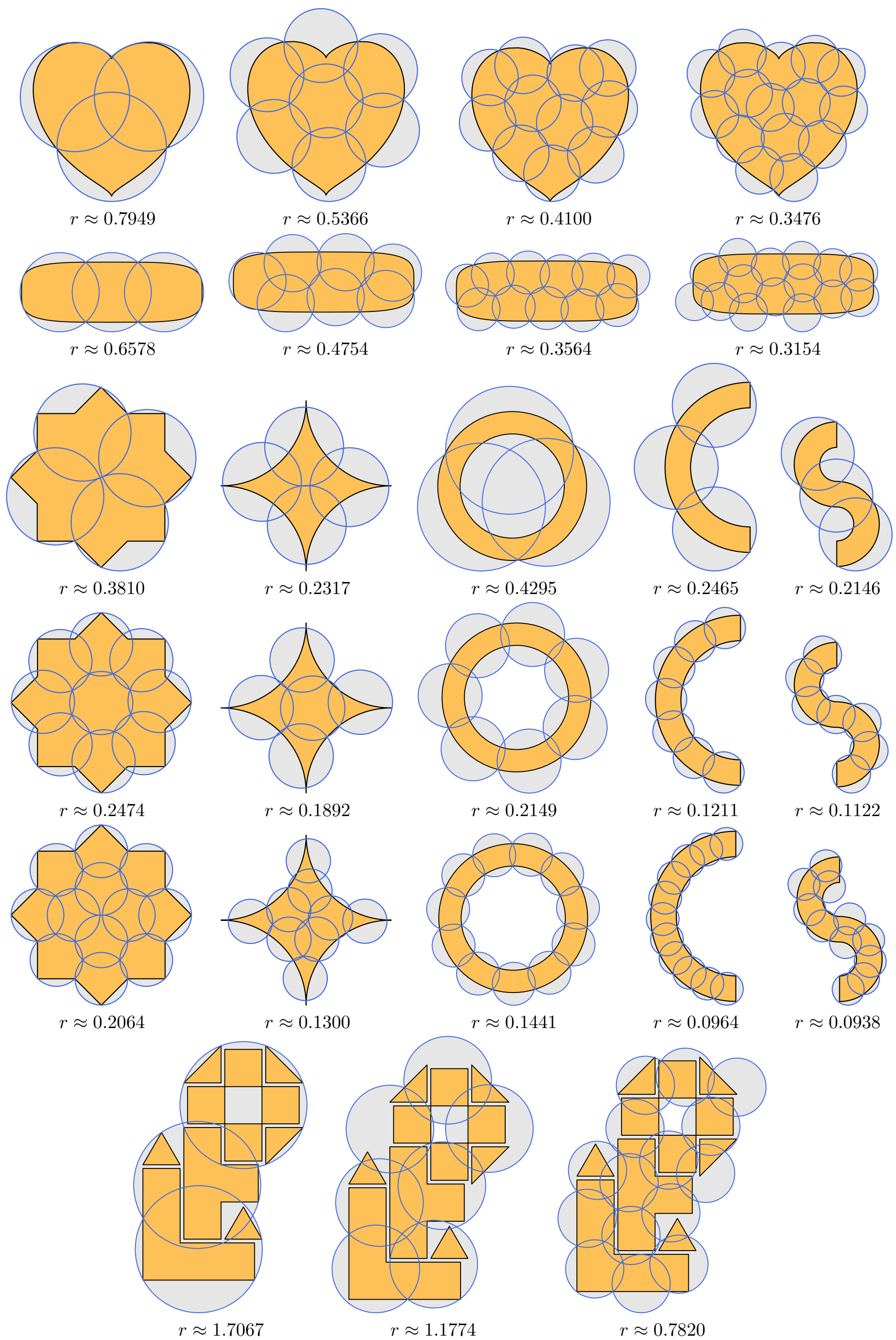


Fig. 5: Solutions found for covering regions in Table 1: heart-shape and soap-shape regions with $m = 3, 7, 11, 15$, two-squares region with $m = 4, 9, 12$, peaked star region with $m = 4, 5, 9$, ring, half-ring, and two-half-rings regions with $m = 3, 7, 11$, and disconnected region with $m = 3, 7, 15$.

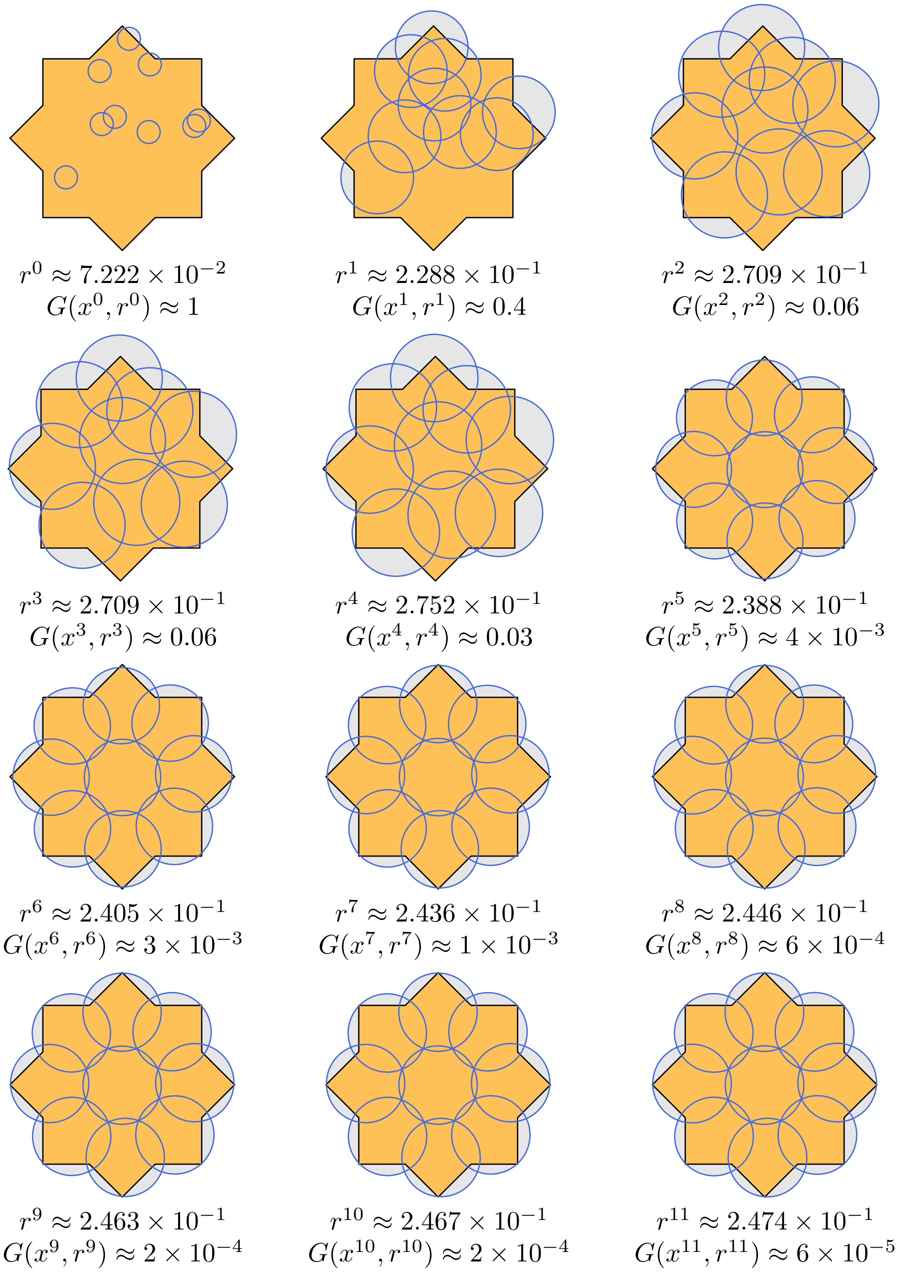


Fig. 6: Evolution of the optimization process in the “two squares” problems with $m = 9$ balls, starting from the $70th$ initial guess (x^0, r^0) which is the one that leaves to the best solution found.

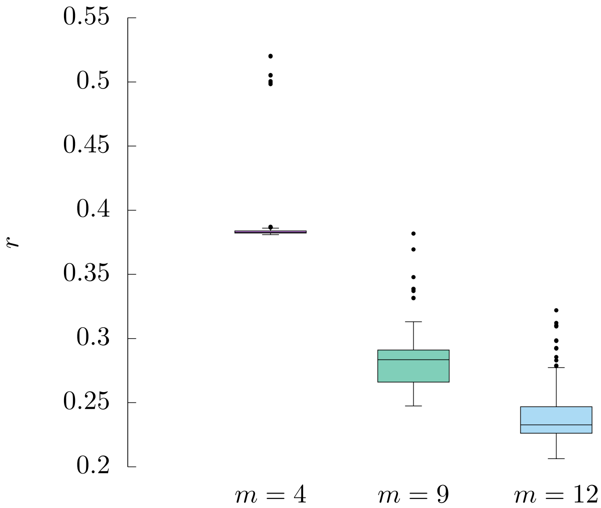
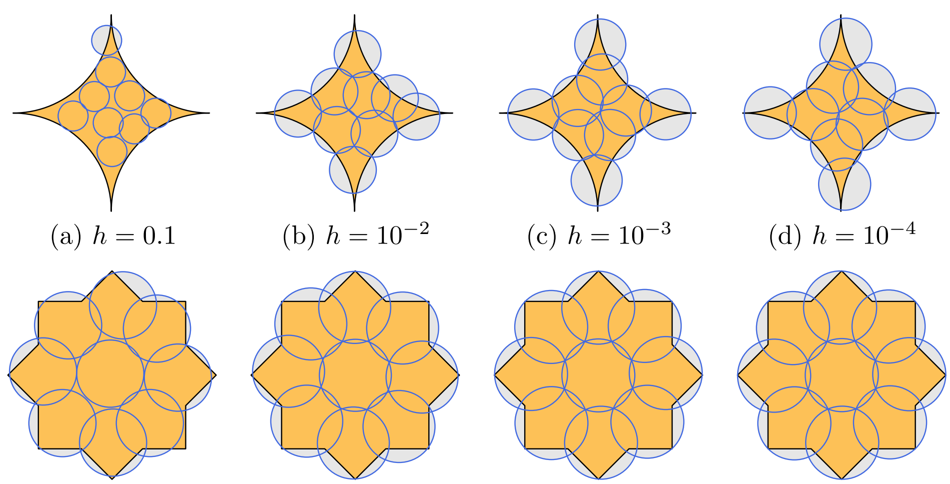


Fig. 7: Boxplot representation of the radii found in 100 runs of the "two squares" problem with $m \in \{4, 9, 12\}$.



(a) $h = 0.1$ (b) $h = 10^{-2}$ (c) $h = 10^{-3}$ (d) $h = 10^{-4}$

(e) $h = 0.1$ (f) $h = 10^{-2}$ (g) $h = 10^{-3}$ (h) $h = 10^{-4}$

Fig. 8: Solutions found varying $h \in \{0.1, 10^{-2}, 10^{-3}, 10^{-4}\}$, with $\varepsilon_{\text{feas}} = 0.1h$ and $\varepsilon_{\text{opt}} = 0.1$, in problems (a–d) "two squares" and (e–h) "peaked star" with $m = 9$.

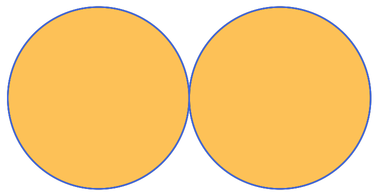
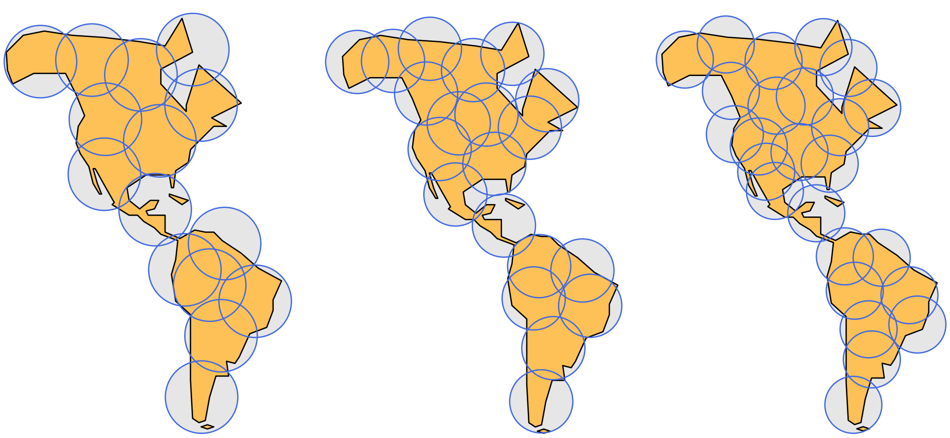


Fig. 9: An example of a degenerate case: A is the union of two tangent unitary-diameter balls to be covered by $m = 2$ balls. Even though this singular case is not covered by the theory, the solution, which is the set A itself, was found with a single run of the method.



(a) $m = 15$, $r \approx 0.08556$ (b) $m = 20$, $r \approx 0.07459$ (c) $m = 25$, $r \approx 0.06728$

Fig. 10: Solutions found for covering region America with $m = 15, 20, 25$.