Systems of distant representatives in Euclidean space*

Adrian Dumitrescu[†] Minghui Jiang[‡]

March 9, 2015

Abstract

Given a finite family of sets, Hall's classical marriage theorem provides a necessary and sufficient condition for the existence of a system of distinct representatives for the sets in the family. Here we extend this result to a geometric setting: given a finite family of objects in the Euclidean space (e.g., convex bodies), we provide a sufficient condition for the existence of a system of distinct representatives for the objects that are also *distant* from each other. For a wide variety of geometric objects, this sufficient condition is also necessary in an asymptotic sense (i.e., apart from constant factors, the inequalities are the best possible). Our methods are constructive and lead to efficient algorithms for computing such representatives.

Keywords: Systems of distinct representatives, Lebesgue measure, lattice packing, lattice covering, bipartite matching, approximation algorithm.

1 Introduction

Let J be a finite index set, and $\mathcal{A} = \{A_j \mid j \in J\}$ be a finite family of finite sets. A system of distinct representatives (SDR) for $\mathcal{A}$ is an indexed set $S = \{a_j \mid j \in J\}$ of distinct elements with $a_j \in A_j$, $\forall j \in J$. Hall's classical marriage theorem [11] [15, Ch. 5] provides a necessary and sufficient condition for the existence of an SDR: $\mathcal{A}$ has an SDR if and only if

$$\forall I \subseteq J, \quad \left|\bigcup_{i \in I} A_i\right| \geq |I|.$$

Hall's Theorem has been extended to systems of multiple representatives by Halmos and Vaughan [12]. Let t be a positive integer. A t-wise system of distinct representatives (t-SDR) for $\mathcal{A}$ is an indexed family $T = \{B_j \mid j \in J\}$ of disjoint subsets with $B_j \subseteq A_j$ and $|B_j| = t$, $\forall j \in J$. The Halmos-Vaughan's Theorem [12] [23, Theorem 22.14] provides a necessary and sufficient condition for the existence of a t-SDR: $\mathcal{A}$ has a t-SDR if and only if

$$\forall I \subseteq J, \quad \left|\bigcup_{i \in I} A_i\right| \geq t|I|.$$

Systems of distinct representatives for objects in the Euclidean space appear naturally, perhaps more naturally than their combinatorial counterparts. Let $\mathcal{R}$ be a family of n subsets of a metric

*An extended abstract of this paper appeared in the Proceedings of the 29th ACM Symposium on Computational Geometry, ACM, 2013, pages 441–448.

†Department of Computer Science, University of Wisconsin–Milwaukee, USA. Email: dumitres@uwm.edu. Supported in part by NSF grant DMS-1001667.

‡Department of Computer Science, Utah State University, Logan, USA. Email: mjiang@cc.usu.edu.

space. The problem of *dispersion in* $\mathcal{R}$ is that of selecting n points, one in each subset, such that the minimum inter-point distance is maximized. This problem was introduced by Fiala et al. [10] as Systems of Distant Representatives, generalizing the classic problem Systems of Distinct Representatives. An especially interesting version of the dispersion problem, is in a geometric setting where $\mathcal{R}$ is a set of convex bodies (e.g., disks or rectangles in the plane). For instance, given a set of n disks in the plane, the dispersion problem is that of selecting n points, one from each disk, such that the minimum pairwise distance of the selected points is maximized. One can think of the selected points as representatives of the disks, that are far enough from each other.

Dispersion in disks is NP-hard, as Fiala et al. [10] showed that dispersion in unit disks is already NP-hard. As noted in [9], dispersion in disjoint unit disks is also NP-hard, and even APX-hard; i.e., unless P = NP, the problem does not admit any polynomial-time approximation scheme. On the other hand, several constant-ratio approximation algorithms have been proposed for disks and balls by Cabello [6] and by Dumitrescu and Jiang [9].

While such algorithms can compute representative points for an input set of disks, guaranteed to be far enough from each other when compared to an optimal solution, one would like to have a more accurate assessment of the quality of the solution (i.e., separation distance) as a function of the input family of disks. It is this direction that we pursue here, where we present asymptotically tight characterizations of the input family in terms of the separation distance between the representative points for the disks. Moreover, some of our results extend to much more general geometric objects in place of disks (or balls).

Indeed, while dispersion in disks is relatively well understood, almost nothing is known about dispersion in other geometric ranges. Here we tackle for the first time the dispersion problem in families of fat objects and in families of homothets of a convex body. (Our meaning of fatness is in fact quite broad and does not even require connectedness.)

Moreover, our methods are constructive and lead to efficient algorithms for finding distant or disjoint representatives. They reduce the geometric problems under consideration to combinatorial problems on bipartite graphs, which are then solved by using known algorithms for matching.

Notations and definitions. Let $[n]$ denote the set $\{1, 2, \ldots, n\}$. A *geometric object* (or *object*, for short) is a compact set in $\mathbb{R}^d$ ($d \geq 1$) with nonempty interior. Note that a geometric object is measurable, but it is not necessarily convex or even connected. A *convex body* is a convex geometric object.

For a geometric object B in $\mathbb{R}^d$, denote by $L(B)$ the Lebesgue measure of B, which is the length when $d = 1$, the area when $d = 2$, and in general the d-volume for any $d \geq 3$. For a finite family $\mathcal{F}$ of geometric objects in $\mathbb{R}^d$, denote by $L(\mathcal{F}) = L(\cup_{B \in \mathcal{F}} B)$ the Lebesgue measure of the union of the geometric objects in $\mathcal{F}$. For simplicity, we will use volume as a synonym for Lebesgue measure for any $d \geq 1$. Throughout the paper, all families of geometric objects considered are finite families, and by "disjoint objects" we mean "pairwise-disjoint objects".

For standard terminology regarding *lattice packing* and *lattice covering* the reader is referred to [18, Ch. 3].

We next state our results.

Cubes and boxes.

Theorem 1. *Let $\mathcal{F}$ be a family of n axis-parallel cubes in $\mathbb{R}^d$, and let t be a positive integer. Suppose that there exists $x > 0$ such that the following holds: for any k, $1 \leq k \leq n$, and for any subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, the volume of the union of the k cubes in $\mathcal{F}'$ is at least $2^d t k x^d$. Then* *one can choose tn points, with t points in each of the n cubes in $\mathcal{F}$, such that all pairwise distances among these points are at least x.*

For an axis-parallel box $B = I_1 \times \cdots \times I_d$ in $\mathbb{R}^d$, let the *minimum side length* be the minimum extent of the box over all d coordinate axes, i.e., $\min\{L(I_1), \ldots, L(I_d)\}$.

Theorem 2. *Let $\mathcal{F}$ be a family of n axis-parallel boxes in $\mathbb{R}^d$, and let t be a positive integer. Suppose that there exists $x > 0$ such that the minimum side length of each box in $\mathcal{F}$ is at least x and the following holds: for any k, $1 \leq k \leq n$, and for any subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, the volume of the union of the k boxes in $\mathcal{F}'$ is at least $2^d tkx^d$. Then one can choose tn points, with t points in each of the n boxes in $\mathcal{F}$, such that all pairwise distances among these points are at least x.*

We would like to highlight the condition that "the minimum side length of each box in $\mathcal{F}$ is at least x" in Theorem 2. Note that this minimum-length condition is not stated in Theorem 1, although it is used in our proofs of both theorems. The reason that we don't need to state it explicitly in Theorem 1 (and similarly for all our other theorems except Theorems 2 and 6) is because it is redundant: if the sufficient condition, that the volume of the union of the k cubes in $\mathcal{F}'$ is at least $2^d tkx^d$, is satisfied for any k, $1 \leq k \leq n$, and for any subfamily of size k, the case $k = 1$ requires that the volume of each cube is at least $2^d tx^d \geq x^d$ and so the side length of each cube is at least x.

Balls and fat objects.

Theorem 3. *Let $\mathcal{F}$ be a family of n balls in $\mathbb{R}^d$, and let t be a positive integer. Suppose that there exists $x > 0$ such that the following holds: for any k, $1 \leq k \leq n$, and for any subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, the volume of the union of the k balls in $\mathcal{F}'$ is at least $2^d d^{d/2} tkx^d$. Then one can choose tn points, with t points in each of the n balls in $\mathcal{F}$, such that all pairwise distances among these points are at least x.*

Given an orthogonal coordinate system Γ, we say that a geometric object B is α-fat with respect to Γ for some number $0 < \alpha \leq 1$ if there exist two concentric axis-parallel cubes P and Q, where P is a homothet of Q with ratio α, such that $P \subseteq B \subseteq Q$. Note that an axis-parallel cube itself is 1-fat. (Alternatively we can define fatness in terms of inscribed and circumscribed balls, but we prefer cubes which yield sharper constants in our bounds.)

Theorem 4. *Let $\mathcal{F}$ be a family of n α-fat objects in $\mathbb{R}^d$, and let t be a positive integer. Suppose that there exists $x > 0$ such that the following holds: for any k, $1 \leq k \leq n$, and for any subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, the volume of the union of the k objects in $\mathcal{F}'$ is at least $2^d \alpha^{-d} tkx^d$. Then one can choose tn points, with t points in each of the n objects in $\mathcal{F}$, such that all pairwise distances among these points are at least x.*

An asymptotically tight example. Apart from the constants appearing in the volume conditions, the bounds in all the theorems above are asymptotically the best possible. For instance, if d is fixed, an almost tight example in regard to Theorem 1 is as follows. Consider n axis-parallel cubes of side length x, whose centers are along one of the coordinate axes. Let the distance between consecutive centers be cx for some constant $0 < c < 1$. Then the union of any k cubes has volume at least $((k-1)cx + x)x^{d-1} = \Omega(kx^d)$, and the union of all n cubes has volume exactly $((n-1)cx + x)x^{d-1}$. By a standard packing argument, one can show that for a suitable $c = c(d)$, the union of the n cubes cannot hold n points with all pairwise distances at least x.

Indeed, let $\mathrm{vol}_d(r)$ denote the volume of the sphere of radius r in $\mathbb{R}^d$. It is well-known that

$$\mathrm{vol}_d(r) = \begin{cases} \dfrac{\pi^{d/2}}{(d/2)!} \cdot r^d & \text{if } d \text{ is even}, \\ \dfrac{2 \cdot (2\pi)^{(d-1)/2}}{1 \cdot 3 \cdots d} \cdot r^d & \text{if } d \text{ is odd}. \end{cases} \tag{1}$$

Requiring that all pairwise distances are at least x implies that the rectangular box obtained by extending the union of the n cubes by $x/2$ along every axis in both directions contains n pairwise-disjoint balls of radius $x/2$, namely

$$n\, \mathrm{vol}_d(x/2) \le ((n-1)cx + 2x)(2x)^{d-1}. \tag{2}$$

Consequently, for a suitable

$$c(d) = \Theta\left(d^{-1/2}\left(\frac{e\pi}{8d}\right)^{d/2}\right),$$

the inequality (2) does not hold.

Translates and homothets of a convex body.

Theorem 5. *Let $\mathcal{F}$ be a family of n homothets of a convex body C in $\mathbb{R}^d$, and let t be a positive integer. Suppose that there exists $x > 0$ such that the following holds: for any k, $1 \le k \le n$, and for any subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size k, the volume of the union of the k homothets in $\mathcal{F}'$ is at least $4^d d^{2d} t k x^d$ times the volume of C. Then one can choose tn interior-disjoint translates of xC with t translates in each of the n homothets of C in $\mathcal{F}$.*

The interior-disjoint translates of xC in the homothets of C in Theorem 5 are better viewed as "disjoint" representatives rather than "distant" representatives in the preceding theorems. Systems of "disjoint representatives" have been studied by Aharoni and Haxell [1] as matchings of hyperedges in hypergraphs; see also [16]. Knuth and Raghunathan [14] introduced a closely related concept called systems of "compatible representatives", and studied the METAFONT labeling problem in this framework. We next consider a more general algorithmic problem:

LARGEST DISJOINT REPRESENTATIVES

Instance: A convex body C and n geometric objects $R_1, \ldots, R_n$ in $\mathbb{R}^d$.

Problem: Find n interior-disjoint translates of a scaled copy λC of the convex body C, one translate in each object, such that the scale factor λ is maximized.

Note that the geometric objects R_i in the problem LARGEST DISJOINT REPRESENTATIVES are not necessarily homothets of the convex body C. The following theorem gives an approximation algorithm for this problem:

Theorem 6. *Given a lattice Λ in $\mathbb{R}^d$ that supports both a lattice covering of $\mathbb{R}^d$ by $-C$ and a lattice packing in $\mathbb{R}^d$ of $x(C-C)$, where $0 < x < 1$, there is an algorithm based on bipartite matching that approximates* LARGEST DISJOINT REPRESENTATIVES *with ratio $x - \epsilon$ for arbitrarily small $\epsilon > 0$.*

We refer to Figure 1 for two examples. When C is an axis-parallel unit square centered at the origin in the plane, the square lattice of cell length 1 supports a lattice tiling of the plane by C, which is both a lattice covering of the plane by $-C = C$ and a lattice packing in the plane of $(1/2)(C - C) = C$. Thus the approximation ratio of the algorithm in Theorem 6 for finding maximum disjoint squares in a set of n given objects is $1/2 - \epsilon$. When C is a disk of unit radius centered at the origin in the plane, the triangular lattice of cell length $\sqrt{3}$ supports both a lattice covering of $\mathbb{R}^2$ by $-C = C$ and a lattice packing in $\mathbb{R}^2$ of $(\sqrt{3}/4)(C - C) = (\sqrt{3}/2)C$. Thus the approximation ratio of the algorithm in Theorem 6 for finding maximum disjoint disks in a set of n given objects is $\sqrt{3}/4 - \epsilon$.

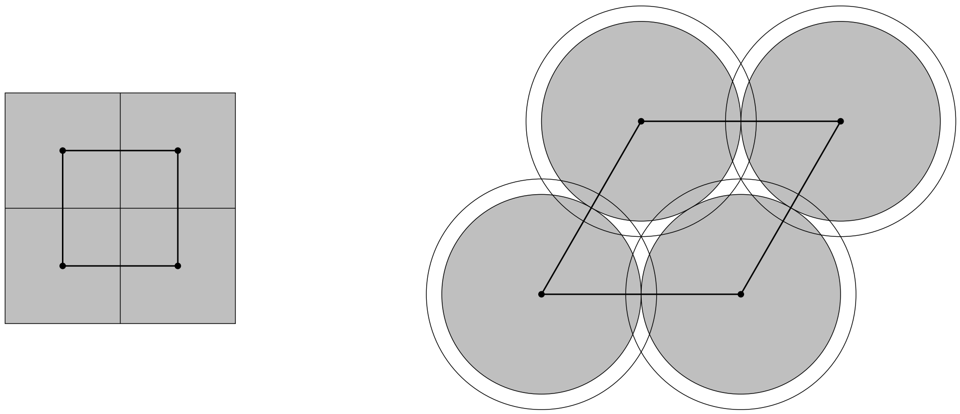

Figure 1: Left: When C is a unit square, a square lattice of cell length 1 supports both a lattice covering by C and a lattice packing of C. Right: When C is a disk of unit radius, a triangular lattice of cell length $\sqrt{3}$ supports both a lattice covering by C and a lattice packing of $(\sqrt{3}/2)C$.

Related results. We did not find any results with similar flavor in the literature with the exception of some special cases. However, we mention some recent results from the broad category of dispersion problems and also point to some classical works applying the lattice technique to the study of convex bodies.

For packing of n axis-parallel congruent squares in a given rectilinear polygon such that the side length of the squares is maximized, Baur and Fekete [2] presented a $\frac{2}{3}$-approximation algorithm, and proved that the problem is NP-hard to approximate with ratio larger than $\frac{13}{14}$. Note that this is a specialized variant of the problem Largest Disjoint Representatives we mentioned earlier, where the n objects are the same rectilinear polygon. A $\frac{2}{3}$-approximation algorithm for the problem of packing n unit disks in a rectangle without overlapping an existent set of m unit disks in the same rectangle, has been obtained by Benkert et al. [3].

Several problems on selecting independent sets of large volume and multiple questions and techniques in relation to lattices have been considered by Rado for various classes of convex bodies, in his three papers entitled "Some covering theorems" [20, 21, 22]. In particular, we employ an old technique of Rado [20, Theorem 10(iii)] in finding large sets of lattice points covered by the union of a family of geometric objects; see also [4] for a similar application of this technique. Going further back in time, the first to apply the lattice technique to the study of convex bodies was probably Minkowski [17]; see also [13, Ch. 24], [18, Ch. 1].

2 Cubes and boxes

The key to our results is recognizing and exploiting the power of lattices. We first deal with the case of cubes (squares in $\mathbb{R}^2$) that naturally fits in the framework of lattices.

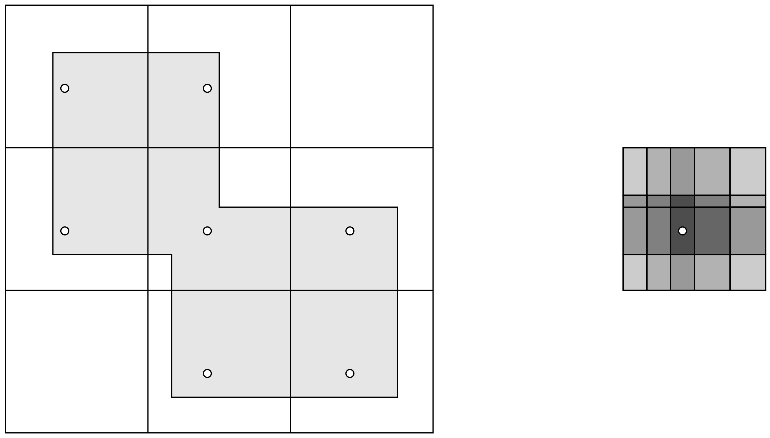

Figure 2: Illustration for the proof of Lemma 1 in $\mathbb{R}^2$.

Lemma 1. *Let Λ be a cubic lattice of cell size $x > 0$ and let $\mathcal{F}$ be a family of axis-parallel boxes of minimum side length at least x in $\mathbb{R}^d$, $d \geq 1$. If the volume of the union of these boxes is at least $2^d m x^d$, for some positive integer m, then they cover at least m lattice points.*

Proof. Fix a cell of the lattice, say τ, outside of the configuration. Translate all lattice cells partially or totally covered by the union of the boxes in $\mathcal{F}$ to τ. Since the covered volume is at least $2^d m x^d$, there exists a point ξ in τ covered at least $2^d m$ times. This means that there exist $2^d m$ distinct points p_i in $2^d m$ distinct cells σ_i of Λ respectively, $1 \leq i \leq 2^d m$, with the same relative offset as ξ in τ, that are covered by the union. See Figure 2.

Consider any index $i \in [2^d m]$. Since the point p_i in the cell σ_i is covered by some axis-parallel box with minimum side length at least x, it follows that at least one of the 2^d vertices (lattice points) of σ_i is also covered by the same box. Select one such covered lattice point for each cell σ_i arbitrarily. Observe that each lattice point can be selected at most 2^d times, by the 2^d adjacent cells in Λ, if at all. It follows that at least $2^d m / 2^d = m$ distinct lattice points are covered by the union, as required. □

Proof of Theorem 1. Let Λ be a cubic lattice of cell size x placed arbitrarily, but fixed throughout the proof. For each cube $Q \in \mathcal{F}$, let Q_Λ be the subset of lattice points covered by Q. We obtain in this way a family $\mathcal{A}$ of subsets over the set of lattice points, one subset $Q_\Lambda \in \mathcal{A}$ for each cube $Q \in \mathcal{F}$. By the volume assumption in the theorem we know that for every k and every subfamily of k cubes, the volume of the union of these k cubes is at least $2^d ktx^d$. In particular, for $k = 1$ it implies that each cube $Q \in \mathcal{F}$ has volume at least $2^d tx^d > x^d$, hence its side length is at least x. Since the volume of the union of any k cubes is at least $2^d ktx^d$, it follows from Lemma 1 that the k cubes Q cover at least tk lattice points, i.e., the union of the corresponding k subsets Q_Λ has size at least tk. It then follows by the Halmos-Vaughan Theorem [12] that $\mathcal{A}$ has a t-SDR. Equivalently, one can choose t distinct points from the subset $Q_\Lambda \in \mathcal{A}$ for each cube $Q \in \mathcal{F}$, that is, tn points in total. Moreover, since these points are lattice points, their pairwise distances are at least x, as required. □

Proof of Theorem 2. The proof is analogous to the proof of Theorem 1 with each cube replaced by a box. The condition that the minimum side length of each box in $\mathcal{F}$ is at least x is given explicitly in the theorem. □

Remarks. Lemma 1 and Theorem 2 can be further generalized, with each axis-parallel box of minimum side length at least x replaced by the Minkowski sum of an arbitrary point set and an axis-parallel cube of side length x. The above proofs give in fact a stronger separation, namely the L_∞ distance (rather than the L_2 distance) between any two representatives is at least x.

3 Balls and fat objects

To prove Theorem 3, we reduce the case of balls to that of cubes, i.e., we reduce the proof of Theorem 3 to that of Theorem 1. We start with two lemmas of independent interest that are reminiscent of the Kneser-Poulsen conjecture for disks [5]:

Lemma 2. *Let $\mathcal{F}$ be a family of n intervals on the line $I_i = [c_i - \ell_i/2, c_i + \ell_i/2]$, $1 \le i \le n$, where the ith interval I_i is centered at c_i and has length ℓ_i. Let $X(\mathcal{F}, \lambda)$ denote the corresponding family of n intervals of the same length $I'_i = [\lambda c_i - \ell_i/2, \lambda c_i + \ell_i/2]$, $1 \le i \le n$, which are centered at λc_i instead of c_i $(L(I_i) = L(I'_i), 1 \le i \le n)$. Suppose that $\lambda \ge 1$. Then $L(\mathcal{F}) \le L(X(\mathcal{F}, \lambda))$.*

Proof. Without loss of generality we can assume that the n intervals have distinct centers and that $c_1 < \cdots < c_n$. Observe that the intervals in $X(\mathcal{F}, \lambda)$ appear in the same order of the centers as in $\mathcal{F}$. We proceed by induction on n. For $n = 1$, there is nothing to prove. Let now $n \ge 2$, and assume that the lemma holds for $n-1$ intervals. We can assume without loss of generality that no interval I_i is completely contained in another interval $I_j \in \mathcal{F}$; indeed, then the inequality in the lemma follows by induction:

$$L(\mathcal{F}) = L(\mathcal{F} \setminus I_i) \le L(X(\mathcal{F} \setminus I_i, \lambda)) \le L(X(\mathcal{F}, \lambda)).$$

For each j, $1 \le j \le n$, denote by $\mathcal{F}_j$ the subfamily of j intervals I_i, $1 \le i \le j$. Then we have $L(\mathcal{F}_{n-1}) \le L(X(\mathcal{F}_{n-1}, \lambda))$ by the induction hypothesis. By the previous assumption, for each i, $1 \le i \le n-1$, the interval I_n is either completely to the right of I_i, or partially covered by I_i at the left end of I_n. In either case, the right endpoint of I_n is to the right of the right endpoints of all other intervals I_i, i.e., $(c_n + \ell_n/2) - (c_i + \ell_i/2) > 0$ for $1 \le i \le n-1$. Since $\lambda \ge 1$, and $c_i < c_n$ for $1 \le i \le n-1$, this further yields

$$(c_n + \ell_n/2) - (c_i + \ell_i/2) \le (\lambda c_n + \ell_n/2) - (\lambda c_i + \ell_i/2), \text{ for } 1 \le i \le n-1.$$

In particular,

$$\min_{1 \le i \le n-1} [(c_n + \ell_n/2) - (c_i + \ell_i/2)] \le \min_{1 \le i \le n-1} [(\lambda c_n + \ell_n/2) - (\lambda c_i + \ell_i/2)]. \tag{3}$$

We consider two cases.

1. If I'_n is disjoint from any other interval in $X(\mathcal{F}_{n-1})$, the inequality in the lemma follows by induction:

$$L(\mathcal{F}_n) \le L(\mathcal{F}_{n-1}) + L(I_n) \le L(X(\mathcal{F}_{n-1}, \lambda)) + L(I_n) = L(X(\mathcal{F}_{n-1}, \lambda)) + L(I'_n) = L(X(\mathcal{F}_n, \lambda)).$$

2. If I'_n overlaps some interval in $X(\mathcal{F}_{n-1})$, I_n must overlap the corresponding interval in $\mathcal{F}_{n-1}$, and then again by induction and using (3) we have

$$\begin{aligned} L(\mathcal{F}_n) &= L(\mathcal{F}_{n-1}) + \min_{1 \le i \le n-1} [(c_n + \ell_n/2) - (c_i + \ell_i/2)] \\ &\le L(X(\mathcal{F}_{n-1}, \lambda)) + \min_{1 \le i \le n-1} [(\lambda c_n + \ell_n/2) - (\lambda c_i + \ell_i/2)] \le L(X(\mathcal{F}_n, \lambda)). \end{aligned}$$

This completes the proof. □

Lemma 3. *Let $\mathcal{F}$ be a family of n axis-parallel boxes in $\mathbb{R}^d$ where the ith box B_i is centered at $(c_{i,1}, \dots, c_{i,d})$, $1 \le i \le n$. Let $X(\mathcal{F}, \lambda)$ denote the corresponding family of n axis-parallel boxes in $\mathbb{R}^d$ where the ith box is a translate of B_i centered at $(\lambda c_{i,1}, \dots, \lambda c_{i,d})$, $1 \le i \le n$. Suppose that $\lambda \ge 1$. Then $L(\mathcal{F}) \le L(X(\mathcal{F}, \lambda))$.*

Proof. For $0 \le j \le d$, let $X_j(\mathcal{F}, \lambda)$ denote the family of n boxes in $\mathbb{R}^d$ where the ith box is a translate of B_i centered at $(\lambda c_{i,1}, \dots, \lambda c_{i,j}, c_{i,j+1}, \dots, c_{i,d})$, $1 \le i \le n$. Then $X_0(\mathcal{F}, \lambda) = \mathcal{F}$ and $X_d(\mathcal{F}, \lambda) = X(\mathcal{F}, \lambda)$. It suffices to show that $L(X_{j-1}(\mathcal{F}, \lambda)) \le L(X_j(\mathcal{F}, \lambda))$ for all $1 \le j \le d$.

Fix $j \in [d]$. For any family $\mathcal{F}$ of axis-parallel boxes and any line H parallel to the jth axis, denote by $\mathcal{F} \Cap H$ the family of intervals $\{B \cap H \mid B \in \mathcal{F}, B \cap H \ne \emptyset\}$. For each box in $X_{j-1}(\mathcal{F}, \lambda)$ that intersects H, the corresponding translate of the box in $X_j(\mathcal{F}, \lambda)$ also intersects H. By Lemma 2 we have $L(X_{j-1}(\mathcal{F}, \lambda) \Cap H) \le L(X_j(\mathcal{F}, \lambda) \Cap H)$. Then, by integrating this inequality over all lines that are parallel to the jth axis and intersect at least one box in $X_{j-1}(\mathcal{F}, \lambda)$ we get $L(X_{j-1}(\mathcal{F}, \lambda)) \le L(X_j(\mathcal{F}, \lambda))$, as required. □

Proof of Theorem 3. For each ball $B \in \mathcal{F}$, denote by P_B the axis-parallel cube inscribed in B, and denote by Q_B the axis-parallel cube circumscribed about B. Observe that P_B and Q_B are concentric. Moreover Q_B is a homothet of P_B with ratio $\lambda = \sqrt{d}$. Let $\mathcal{P} = \{P_B \mid B \in \mathcal{F}\}$ and $\mathcal{Q} = \{Q \mid B \in \mathcal{F}\}$.

Consider any $k \in [n]$ and an arbitrary subset $\mathcal{F}_k$ of k balls in $\mathcal{F}$. Let $\mathcal{P}_k = \{P_B \mid B \in \mathcal{F}_k\}$ and $\mathcal{Q}_k = \{Q_B \mid B \in \mathcal{F}_k\}$. We clearly have $L(\mathcal{P}_k) \le L(\mathcal{F}_k) \le L(\mathcal{Q}_k)$. Consider the family $\mathcal{Q}'_k = X(\mathcal{Q}_k, \lambda)$ of axis-parallel cubes obtained from $\mathcal{Q}_k$ by spreading the cubes without changing their sizes to increase the pairwise distances of their centers by a factor of λ. By Lemma 3, we have $L(\mathcal{Q}_k) \le L(\mathcal{Q}'_k)$. Observe that $\mathcal{Q}'_k$ can be also obtained from $\mathcal{P}_k$, by uniformly scaling both the sizes and the pairwise distances of the cubes by a factor of λ. Hence $L(\mathcal{Q}'_k) = \lambda^d L(\mathcal{P}_k) = d^{d/2} L(\mathcal{P}_k)$. Putting all these together yields $L(\mathcal{F}_k) \le L(\mathcal{Q}_k) \le L(\mathcal{Q}'_k) = d^{d/2} L(\mathcal{P}_k)$.

By the volume assumption in the theorem, $L(\mathcal{F}_k) \ge 2^d d^{d/2} tkx^d$. Then, from the inequality $L(\mathcal{F}_k) \le d^{d/2} L(\mathcal{P}_k)$, we have $L(\mathcal{P}_k) \ge L(\mathcal{F}_k)/d^{d/2} \ge 2^d tkx^d$. This lower bound of $2^d tkx^d$ on $L(\mathcal{P}_k)$ holds for any subfamily $\mathcal{P}_k$ of k cubes in $\mathcal{P}$. Thus by Theorem 1, one can choose tn points, with t points in each cube in $\mathcal{P}$, such that all pairwise distances among these points are at least x. Since the cubes in $\mathcal{P}$ are contained in the respective balls in $\mathcal{F}$, it follows that one can choose tn points, with t points in each ball in $\mathcal{F}$, such that all pairwise distances among these points are at least x. □

Proof of Theorem 4. The proof of this theorem is analogous to the proof of Theorem 3, with each ball replaced by an α-fat object, where the factor λ is equal to $1/\alpha$ instead of $\sqrt{d}$. □

4 Translates and homothets of a convex body

The next two lemmas follow by affine transformations from Lemmas 1 and 3.

Lemma 4. *Let Λ be a lattice in $\mathbb{R}^d$ generated by d linearly independent vectors $\vec{u}_1, \dots, \vec{u}_d$. Let σ be the fundamental parallelepiped (cell) of Λ induced by the 2^d vectors $m_1\vec{u}_1 + \dots + m_d\vec{u}_d$, where $m_i \in \{0, 1\}$ for $1 \le i \le d$. Let $\mathcal{F}'$ be a family of parallelepipeds in $\mathbb{R}^d$ that are parallel to σ and have side lengths at least $|\vec{u}_i|$ along the vectors $\vec{u}_i$, $1 \le i \le d$. If the volume of the union of these parallelepipeds is at least $2^d m$ times the volume of σ, for some positive integer m, then they cover at least m lattice points.*

Lemma 5. *Let $\mathcal{F}$ be a family of n pairwise-parallel parallelepipeds in $\mathbb{R}^d$, where the ith parallelepipeds P_i is centered at $(c_{i,1},\ldots,c_{i,d})$, $1 \le i \le n$. Let $X(\mathcal{F},\lambda)$ denote the corresponding family of n parallelepipeds in $\mathbb{R}^d$ where the ith parallelepiped is a translate of P_i centered at $(\lambda c_{i,1},\ldots,\lambda c_{i,d})$, $1 \le i \le n$. Suppose that $\lambda \ge 1$. Then $L(\mathcal{F}) \le L(X(\mathcal{F},\lambda))$.*

Proof of Theorem 5. The volume assumption in the theorem for $k = 1$ implies that each member of $\mathcal{F}$ has volume at least $4^d d^{2d} t x^d L(C) \ge (4d^2x)^d L(C)$, and hence is a homothet of C with ratio at least $4d^2x > 2dx > x$. By a lemma of Chakerian and Stein [7], for every convex body C in $\mathbb{R}^d$, there exist two (not necessarily rectangular) parallelepipeds P and Q homothetic with ratio d such that $P \subseteq C \subseteq Q$.

Let $\mathcal{F}_k \subseteq \mathcal{F}$ be any subfamily of k homothets of C, where $1 \le k \le n$. Let $\mathcal{P}_k$ be the corresponding family of homothets of P contained in the homothets of C in $\mathcal{F}_k$, and let $\mathcal{Q}_k$ be the corresponding family of homothets of Q containing the homothets of C in $\mathcal{F}_k$; $|\mathcal{P}_k| = |\mathcal{Q}_k| = |\mathcal{F}_k| = k$. Note that $L(Q) = d^d L(P)$. Thus by a similar analysis as in the proof of Theorem 3 that uses Lemma 5 instead of Lemma 3 (for parallelepipeds instead of axis-parallel boxes), we have $L(\mathcal{F}_k) \le L(\mathcal{Q}_k) \le d^d L(\mathcal{P}_k)$.

By the volume assumption in the theorem, $L(\mathcal{F}_k) \ge 4^d d^{2d} t k x^d L(C) \ge 4^d d^{2d} t k x^d L(P)$. Then, from the inequality $L(\mathcal{F}_k) \le d^d L(\mathcal{P}_k)$, we have $L(\mathcal{P}_k) \ge L(\mathcal{F}_k)/d^d \ge 4^d d^d t k x^d L(P)$. Recall that each homothet of C in $\mathcal{F}$ is at least as large as $2dxC$, hence it contains a homothet of P at least as large as $2dxP$. Fix a lattice Λ whose cells are translates of $R = 2dxP$, with cell volume $2^d d^d x^d L(P)$. Then the homothets of P in $\mathcal{P}_k$ are at least as large as the lattice cells. Moreover, since $4^d d^d t k x^d L(P) = (2^d t k) 2^d d^d x^d L(P)$, it follows by Lemma 4 that the k parallelepipeds in $\mathcal{P}_k$ cover at least tk lattice points. Hence the k homothets of C in $\mathcal{F}_k$ cover at least tk lattice points too. This statement holds for any k and any subfamily $\mathcal{F}_k \subseteq \mathcal{F}$ of k homothets of C. Then, by the Halmos-Vaughan Theorem [12], there exist tn lattice points in Λ with t points in each of the n homothets of C in $\mathcal{F}$.

Each of these tn lattice points is contained in some homothet of C in $\mathcal{F}$. For each such lattice point p' contained in some homothet C' of C in $\mathcal{F}$, there is a translate of xC that contains p' and is contained in C' (recall that each homothet in $\mathcal{F}$ is at least as large as xC). Any two translates of xC containing two different lattice points are interior-disjoint. This is because each translate of xC containing a lattice point is contained in a translate of xQ containing the lattice point. Since xQ is a translate of $dxP = \frac{1}{2}R$, each translate of xQ containing a lattice point is in turn contained in a translate of R centered at the lattice point. The translates of R centered at different lattice points are interior-disjoint because the cells of the lattice are translates of R. Thus we obtain tn interior-disjoint translates of xC with t translates in each of the n homothets of C in $\mathcal{F}$. □

5 Approximation algorithm for Largest Disjoint Representatives

In this section we prove Theorem 6 by giving an approximation algorithm based on bipartite matching for the problem Largest Disjoint Representatives. Let Λ be a lattice that supports both a lattice covering of $\mathbb{R}^d$ by translates of $-C$ and a lattice packing in $\mathbb{R}^d$ of $x(C-C)$ for some $0 < x < 1$. Let $\epsilon > 0$ be arbitrarily small; without loss of generality, $0 < \epsilon < x$. Put $\epsilon' = \frac{\epsilon}{2x}$. The algorithm uses a binary search with a decision procedure. Given a candidate scale factor λ, the decision procedure either reports (correctly) that the scale factor λ is too large, or finds n translates using a suitable smaller scale factor $\lambda'' = \lambda' x$, where $\lambda' = \lambda(1-\epsilon')$.

Before the binary search, the algorithm first computes a range $[\lambda_{\text{low}}, \lambda_{\text{high}}]$ for the candidate scale factor λ. For each geometric object R_i, $1 \le i \le n$, compute the largest scale factor λ_i such that R_i contains a homothet C_i of C with ratio λ_i. Let $\lambda_{\text{high}} = \min\{\lambda_1,\ldots,\lambda_n\}$, and let

$\lambda_{\text{low}} = \lambda_{\text{high}}/(4d^2 n^{1/d})$. The algorithm then conducts a binary search with range $[l, h]$ initialized to $[\lambda_{\text{low}}, \lambda_{\text{high}}]$; for the current range $[l, h]$, the candidate scale factor λ is set to $(l + h)/2$. If the decision procedure reports that the candidate scale factor λ is too large, the algorithm updates the upper bound h to λ; otherwise, it updates the lower bound l to λ. The search stops when the lower bound l is at least $1 - \epsilon'$ times the upper bound h. The result of the decision procedure on the final lower bound l is returned.

The decision procedure works as follows. Construct a bipartite graph G with the n objects R_i, $1 \leq i \leq n$, as n vertices on one side, and with the lattice points in $\lambda'\Lambda$ that are covered by the union of the objects on the other side, such that an object R_i and a lattice point p are connected by an edge if and only if there exists a translate of $\lambda''C$ containing p and contained in R_i. Find a maximum matching in G. If the size of the maximum matching is less than n, report that the given scale factor λ is too large. Otherwise, return a translate of $\lambda''C$ in each object that contains the corresponding lattice point.

Correctness. The correctness of the upper bound λ_{high} of the search range is obvious; we next refer to the lower bound λ_{low}. Let $\mathcal{C}$ be the family of homothets C_i of C with ratio λ_i contained in the geometric objects R_i, $1 \leq i \leq n$. Note that for any $k \in [n]$ and for any subfamily of k homothets of C in $\mathcal{C}$, the volume of the union of the k homothets is at least the volume of each one, which is at least λ_{high}^d times the volume of C. Choose $x_0 = \lambda_{\text{high}}/(4d^2 n^{1/d})$ to satisfy the equality $4^d d^{2d} n x_0^d = \lambda_{\text{high}}^d$. With this choice, for any $k \in [n]$ and any subfamily $\mathcal{C}_k$ of k homothets from $\mathcal{C}$, the volume of the union of the homothets in $\mathcal{C}_k$ is $L(\mathcal{C}_k) \geq \lambda_{\text{high}}^d L(C) = 4^d d^{2d} n x_0^d L(C) \geq 4^d d^{2d} k x_0^d L(C)$. Then, by Theorem 5, each homothet C_i (hence each object R_i too) contains an interior-disjoint translate of $x_0 C$. Thus the lower bound $\lambda_{\text{low}} = x_0 = \lambda_{\text{high}}/(4d^2 n^{1/d})$ of the search range is also correct.

We next show the correctness of the decision procedure. Suppose that there exist n interior-disjoint translates of λC, one in each object, then there exist n disjoint (both interior and boundary) translates of $\lambda' C = \lambda(1-\epsilon')C$, one in each object. Since Λ is a lattice that supports a lattice covering of $\mathbb{R}^d$ by translates of $-C$, $\lambda'\Lambda$ is a lattice that supports a lattice covering of $\mathbb{R}^d$ by translates of $-\lambda' C$. Then each point q in $\mathbb{R}^d$ is covered by at least one translate $-\lambda' C + p$ of $-\lambda' C$ for some lattice point p in $\lambda'\Lambda$. Equivalently, each translate $\lambda' C + q$ of $\lambda' C$ covers some lattice point p in $\lambda'\Lambda$. Therefore the n disjoint translates of $\lambda' C$ cover n distinct lattice points in $\lambda'\Lambda$, one in each object. Since Λ is also a lattice that supports a lattice packing in $\mathbb{R}^d$ of translates of $x(C-C)$, $\lambda'\Lambda$ is a lattice that supports a lattice packing in $\mathbb{R}^d$ of translates of $\lambda' x(C-C) = \lambda''(C-C)$. Each translate of $\lambda''C$ that contains a lattice point of $\lambda'\Lambda$ is contained in a translate of $\lambda''(C-C)$ centered at the lattice point. Since the translates of $\lambda''(C-C)$ centered at different lattice points are interior-disjoint (by the packing assumption), it follows that the translates of $\lambda''C$ containing different lattice points in $\lambda'\Lambda$ are interior-disjoint too. In particular, the algorithm finds n interior-disjoint translates of $\lambda''C$, one in each object.

If the decision procedure reports that the given scale factor λ is too large, then there do not exist n interior-disjoint translates of λC, one in each object. Indeed, otherwise the algorithm would find n interior-disjoint translates of $\lambda''C$, one in each object, as shown in the preceding paragraph. So the decision procedure reports correctly that the given scale factor λ is too large.

Finally, the binary search and the decision procedure together yield an approximation algorithm with ratio $x(1-\epsilon')(1-\epsilon') > x(1-2\epsilon') = x - \epsilon$, as desired.

Running time. The running time of the above algorithm is not necessarily polynomial. Specifically, although polynomial-time algorithms for bipartite matching are well-known (see e.g. [8, Ch. 26] and [19]), the size of the bipartite graph G constructed by the decision procedure may not be polynomial in the "input size" of the problem, which is the total size necessary to encode the convex body C and the n geometric objects R_i, $1 \le i \le n$. For example, when the convex body C and the objects R_i are axis-parallel squares, they can be encoded efficiently by specifying the coordinates of their corners. If an object R_i is very large compared to $\lambda'' C$, then the number of lattice points p in $\lambda' \Lambda$ that are covered by some translate of $\lambda'' C$ contained in R_i could be exponential in the input size.

This issue can be addressed as follows. First, instead of taking all lattice points in $\lambda' \Lambda$ that are covered by the union of the objects as the vertices on one side of the bipartite graph, we only take, for each object R_i, a set S_i of up to n lattice points p in $\lambda' \Lambda$ that are covered by some translate of $\lambda'' C$ contained in R_i (that is, if the number of such lattice points exceeds n, we take any n of them and ignore the rest). We thus collect a set $S = S_1 \cup \cdots \cup S_n$ of at most n^2 lattice points. Next, we add edges to connect the n objects and the at most n^2 lattice points following the same rule as before, i.e., an object R_i and a lattice point $p \in S$ are connected by an edge if and only if there exists a translate of $\lambda'' C$ containing p and contained in R_i. We thus have a reduced bipartite graph G' with $O(n^2)$ vertices and $O(n^3)$ edges.

We claim that G' has a matching of size n if and only if G has a matching of size n. The direct implication is trivial since G' is a subgraph of G. Now consider the reverse implication. Suppose there is a matching of size n in G. If an edge e in the matching connects an object R_i to a lattice point $q \notin S_i$, then the size of S_i must be exactly n (since if $|S_i|$ were smaller than n, then q would be part of S_i). Thus we can replace the edge e by another edge e' that connects R_i to some lattice point $p \in S_i$ not incident to any edge in the matching. By repeating this replacement step at most n times, at most once for each object, we obtain a matching of size n in G'.

Recall that the decision procedure takes one of two different actions depending only on whether the bipartite graph G has a matching of size n. From the above claim, it follows that the correctness of the algorithm is unaffected when we replace G by the reduced bipartite graph G'.

We have shown that the running time of the algorithm can be made polynomial in the input size provided we can compute, in polynomial time, a set S_i of up to n lattice points for each object R_i. This is achievable in many natural scenarios. We give two simple examples:

1. When the geometric objects R_i are rectilinear polygons and the convex body C is an axis-parallel square (correspondingly the lattice Λ is a square lattice with $x = 1/2$), we can use the standard sweep-line technique to enumerate, for each rectilinear polygon R_i of m_i vertices, the first n lattice points satisfying the containment constraint, in lexicographical order of their x and y coordinates, with running time polynomial in both m_i and n. This yields an approximation ratio of $1/2 - \epsilon$.

2. When the geometric objects R_i and the convex body C are all disks (correspondingly, the lattice Λ is a triangular lattice with $x = \sqrt{3}/4$), we can use a priority queue to enumerate the first n lattice points contained in R_i, in increasing order of their distances to the center of R_i, in $O(n \log n)$ time. This yields an approximation ratio of $\sqrt{3}/4 - \epsilon = 0.433\ldots - \epsilon$; it is interesting to compare this ratio to the current best approximation ratio of $3/8 = 0.375$ for the related dispersion problem of finding n points in n disks, one point in each disk, to maximize the minimum inter-point distance [6].

6 Conclusion

One can compute distant (respectively disjoint) representatives as guaranteed by Theorems 1–4 (respectively Theorem 5). Translating the proofs of these theorems into algorithms requires some additional effort. Specifically, it amounts to computing systems of distinct representatives for the corresponding families of finite sets of lattice points for a fixed x (as in the approximation algorithm in Theorem 6, computing systems of distinct representatives is equivalent to computing maximum-cardinality matchings in bipartite graphs, and admits polynomial-time algorithms). Finding the largest x (within a factor of $1+\varepsilon$) that yields distinct representatives is easily conducted by a binary search. Fortunately, exactly as in the combinatorial setting of systems of *distinct* representatives, for computing systems of *distant* or *disjoint* representatives there is no need to check inequalities for an exponential number of subsets as they appear listed in the conditions of Theorems 1 through 5.

Regarding Theorem 6, the approximation algorithm requires a lattice satisfying certain properties. This leads to the algorithmic question of how to find such a lattice. Two fundamental steps are: (i) checking whether a given lattice supports a lattice packing of a given convex body; and (ii) checking whether a given lattice supports a lattice covering by a given convex body.

References

[1] R. Aharoni and P. Haxell, Hall's theorem for hypergraphs, *Journal of Graph Theory*, **35** (2000), 83–88.

[2] C. Baur and S. P. Fekete, Approximation of geometric dispersion problems, *Algorithmica*, **30** (2001), 451–470.

[3] M. Benkert, J. Gudmundsson, C. Knauer, R. van Oostrum, and A. Wolff, A polynomial-time approximation algorithm for a geometric dispersion problem, *International Journal of Computational Geometry and Applications*, **19** (2009), 267–288.

[4] S. Bereg, A. Dumitrescu, and M. Jiang, Maximum area independent set in disk intersection graphs, *International Journal of Computational Geometry & Applications*, **20** (2010), 105–118.

[5] K. Bezdek and R. Connelly, Pushing disks apart—the Kneser-Poulsen conjecture in the plane, *Journal für die reine und angewandte Mathematik*, **553** (2002), 221–236.

[6] S. Cabello, Approximation algorithms for spreading points, *Journal of Algorithms*, **62** (2007), 49–73.

[7] G. D. Chakerian and S. K. Stein, Some intersection properties of convex bodies, *Proceedings of the American Mathematical Society*, **18** (1967), 109–112.

[8] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd edition, The MIT Press, Cambridge, 2009.

[9] A. Dumitrescu and M. Jiang, Dispersion in disks, *Theory of Computing Systems*, **51** (2012), 125–142.

[10] J. Fiala, J. Kratochvíl, and A. Proskurowski, Systems of distant representatives, *Discrete Applied Mathematics*, **145** (2005), 306–316.

[11] P. Hall, On representatives of subsets, *Journal of the London Mathematical Society*, **10** (1935), 26–30.

[12] P. R. Halmos and H. E. Vaughan, The marriage problem, *American Journal of Mathematics*, **72** (1950), 214–215.

[13] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th edition, Oxford University Press, 1979.

[14] D. E. Knuth and A. Raghunathan, The problem of compatible representatives, *SIAM Journal on Discrete Mathematics*, **5** (1992), 422–427.

[15] J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 2nd edition, New York, 2001.

[16] L. Martinez and L. Montejano, Geometric variants of Hall's Theorem through Sperner's Lemma, *Electronic Notes in Discrete Mathematics*, **44** (2013), 127–132.

[17] H. Minkowski, *Geometrie der Zahlen*, Leipzig, 1896 (reprinted in 2009).

[18] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, John Wiley, New York, 1995.

[19] W. R. Pulleyblank, Matchings and extensions, in *Handbook of Combinatorics*, Vol. I; (R. L. Graham, M. Grötschel, and L. Lovász, editors), Elsevier, Amsterdam, 1995, pp. 179–232.

[20] R. Rado, Some covering theorems (I), *Proceedings of the London Mathematical Society*, **51** (1949), 232–264.

[21] R. Rado, Some covering theorems (II), *Proceedings of the London Mathematical Society*, **53** (1951), 243–267.

[22] R. Rado, Some covering theorems (III), *Journal of the London Mathematical Society*, **42** (1968), 127–130.

[23] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Springer-Verlag, Berlin, 2003.