Geometric separator theorems

Warren D. Smith*& Nicholas C. Wormald†
wds@research.NJ.NEC.COM, nick@ms.unimelb.edu.au

May 4, 1998

Abstract — **We have a 4-step method for producing "separator theorems" about geometrical objects en masse. Examples: I: Given N disjoint iso-oriented squares in the plane, there exists a rectangle with $\leq 2N/3$ squares inside, $\leq 2N/3$ squares outside, and $\leq (4+o(1))\sqrt{N}$ partly in & out. II: More generally, given N iso-oriented d-cubes in d-space, no more than κ of which cover a point, there exists a box with $\leq 2N/3$ cubes inside, $\leq 2N/3$ cubes outside, and $O(d\kappa^{1/d}N^{1-1/d})$ partly in & out. Here "2/3" is best possible, and the dependence of the bound on κ, d, and N is best possible up to the absolute constant factor in the O. III: We have more general versions where the "d-cubes" and/or the separating "box" could instead be any of a wide variety of convex objects with bounded diameter/width ratios, and where the disjointness condition can be replaced with a variety of other conditions: (a) no more than κ objects cover a point, or (b) no subcollection's measure exceeds the union of its measures by more than a factor κ, or (c) no more than κ objects, among objects whose linear dimensions vary by a factor at most λ, cover a point.**

For tons of applications of these theorems, see the companion paper "Applications of geometric separator theorems." For example, we get the first subexponential algorithms for optimal traveling salesman tour and rectilinear Steiner tree of N points in d-space, d fixed.

1 Introduction

BY COMBINING theorem components from columns A, B, C, and D of a "chinese menu," one can get a huge number of possible theorems. The present paper covers only "A" and "B" (leaving C and D to the companion paper[1]):

A. Theorems about covering d dimensional objects by smaller versions of themselves. Simplest example: a d-box may be divided into two smaller d-boxes.

B. Separator theorems about geometrical objects. Simple example (related to the example in 'A' above): Given N interior-disjoint squares in the plane, there exists a rectangle (both the squares and the rectangle have sides oriented parallel to the coordinate axes) such that $\leq 2N/3$ squares's interiors are entirely inside it, $\leq 2N/3$ are entirely outside, and $\leq (4+o(1))\sqrt{N}$ are partly inside and partly outside. In this theorem "2/3" is best possible.

We now discuss these in more detail.

A. Covering bodies with few smaller versions of themselves (§2). There are a lot of choices because we may consider different kinds of "bodies," e.g. iso-oriented (or not) boxes, cubes, simplices, spheres, L_p balls, general convex bodies. Additional freedom comes from the fact that there are three interesting versions of "versions:" scaled translates, scaled copies with both rotations and translations allowed, and volume reducing affine transformations.

B. Separator theorems about geometrical objects (§3). The example result with squares (§3.1) is, roughly speaking, proven by a $(3+1)$-step argument:

1. "Sup trick" allows us to bound the number of squares outside the separating rectangle.

2. "Pigeonhole principle" allows us to bound the number of squares inside the separating rectangle. (This depends on the covering results from 'A' – as step 0, pick one!)

3. Randomizing argument allows us to bound the number of squares partly inside and partly outside.

Each of the steps in this argument is highly generalizable, and consequently we get a large number of variations of the basic theorem, including:

1. The squares could be other things: cubes, simplices, octahedra, general convex body of CV aspect ratio (definition 2) bounded by B. In the latter case, "$(4+o(1))\sqrt{N}$" changes to "$(4+o(1))B\sqrt{N}$."

2. The separator could be things other than an iso-oriented rectangular box: a (d-)cube, simplex, octahedron, sphere, or general convex body of bounded aspect ratio. (The constants may change.)

3. The objects need not be interior-disjoint; it will suffice if at most a constant number κ of them cover any point ("κ-thick"). In some variants, even weaker conditions ("κ-overloaded," "(λ,κ)-thick;" see §1.2.3) will suffice. In such cases the bound "$4B\sqrt{N}$" becomes "$4B\sqrt{N\kappa}$."

*NECI, 4 Independence Way, Princeton NJ 08540 USA.
†Dept of Mathematics and Statistics, University of Melbourne, Parkville VIC 3052, Australia. Research supported by the ARC.

[1] A preliminary combined full version of both papers (> 50 pages) is available by `ftp ftp.nj.nec.com`, login as anonymous, password = your email address, `cd pub/wds`, `get geomsep.ps`, `quit`.

4. We need not stay in the plane; it is also possible to go to higher dimensions d. The $O(B\sqrt{N\kappa})$ bound is replaced[2] by $O(B\kappa^{1/d}N^{1-1/d})$ with an implied constant depending on d and the specifics of the theorem. In some versions the 1/3-2/3 split weakens so that the "1/3" is replaced by a decreasing function of d.

5. Instead of splitting the *objects* 1/3-2/3 at worst, we can instead split some arbitrary measure on the plane, which need not have anything to do with the objects.

6. We also have linear time algorithmic versions of our theorems (some with weaker constants).

In contrast, a separator theorem proved by Miller and Thurston [25] depends on special properties of the "inversive group" and thus works only for separating *spheres* with *spheres*. Their proof also cannot be made to work for κ-overloaded and (λ,κ)-thick spheres, and unlike us they have no optimality proof for their separator theorem. The present paper makes it clear that the Miller-Thurston result is a special case of a much more general phenomenon. But the [25] result makes up for its specialness with better performance – that is, better constants (with better dependence on the dimension d) – as we show by determining those constants for the first time[3].

We also (§3.3) have miscellaneous other separator theorems not derived from our main 4-step procedure. For example we have a separator theorem for iso-oriented d-boxes *not* required to have bounded aspect ratios.

1.1 Illustration of our game plan in action

Separator theorems are mainly useful because they allow one to design efficient "divide and conquer" algorithms. The present two papers show how to start with a covering argument A, get a separator theorem B about geometrical objects, get a separator theorem C about geometrical graphs, and wind up with an algorithmic application D. This paper covers A,B; the companion paper C,D. Here is a complete short story describing the whole process in one particular instance: 2D RSMTs.

A. [Covering argument] A 4×3 box may be tiled by two 2×3 boxes.

B1. [Sup trick] Find the largest 4×3 box such that it, and any iso-oriented box congruent to it, have $\geq N/3$ of the N square interiors entirely outside it.

B2. [Pigeonhole principle] One of the two 2×3 sub-boxes must contain bits of at least $N/3$ squares.

B3. [Randomizing argument] The boundary of a randomly diagonally translated iso-oriented 3×4 box containing the 2×3 box will in expectation cross $O(\sqrt{N})$ squares and have $\leq 2N/3$ squares inside or outside.

C. [Geometrical graph] The "diamond property" says that the squares whose diagonals are the E edges of a Rectilinear Steiner Minimum Tree (RSMT) of N sites in the plane, must be interior disjoint. Hence by turning your head 45°, from B3 there exists a rectangle with $\leq 2E/3$ edges inside and outside and whose boundary crosses $O(\sqrt{E})$ edges.

D. [Algorithmic application] By "searching over separators" [26] we may (exhaustively) guess the separating rectangle R and the combinatorial "boundary conditions" describing how the RSMT crosses R (all this with no knowledge of the RSMT!) and then divide and conquer the sites. The resulting time recurrence is $T(N)\leq N^{O(\sqrt{N})}[T(2N/3)+T(N/3)]$, so, solving, we get an algorithm for finding the RSMT of N sites in the plane in time $N^{O(\sqrt{N})}$.

1.2 Notation and basic definitions

WLOG, means "without loss of generality." Theorem statements ending with "♣:" this means "proof omitted here; see the full paper for the proof[4]." Logarithms: $\lg x=\log_2 x$, $\ln x=\log_e x$, $e\approx 2.71828$, and $\log x$ intentionally leaves the base of the logarithm unspecified, although using e will work in all our uses. For definitions of geometrical graphs, see §6.1. We sometimes denote the boundary of B by ∂B.

Definition 1 *An "iso-oriented d-box" is a cartesian product of d 1-dimensional compact real intervals.*

1.2.1 Aspect ratio

Roughly speaking, the aspect ratio of an object is the ratio of its longest dimension to its shortest. More precisely

Definition 2 *The "DW aspect ratio" of a compact set in $\mathbf{R}^d$ is the ratio of its Diameter to its Width. The "CV aspect ratio" is the dth root of the ratio of the d-volume of the smallest enclosing iso-oriented d-cube, to the set's own d-volume.*

Thus the CV aspect ratio of an iso-cube is 1, and everything else has a CV aspect ratio >1. The DW aspect ratio of a ball is 1, and everything else has a DW aspect ratio ≥ 1. (If the dimension d is fixed, then all reasonable definitions of aspect ratio are equivalent to within constant factors.)

1.2.2 Volumes of various d-dimensional objects

Definition 3 *Define $\bullet_d=\pi^{d/2}/(d/2)!$ to be[5] the d-volume of a d-ball of radius 1 [37] and $\circ_d=d\bullet_d$ to be the $(d-1)$-surface of a d-ball of radius 1, and*

$$\triangle_d=\frac{\sqrt{1+d}}{d!}(1+\frac{1}{d})^{d/2} \qquad (1)$$

to be the d-volume of a d-dimensional regular simplex inscribed in a d-ball of radius 1.

[2] In our papers, all O's are valid uniformly as *all* the quantities inside them vary independently over their full allowed ranges – in this case, as $\kappa\geq 1$, $B\geq 1$, $d\geq 2$, and $N\geq 2$ vary. In the few cases in which we want to disallow variation we will indicate so in the text, or write, e.g., $O_d(N)$, meaning that d is to be regarded as fixed while N varies unboundedly, and the implied constant factor depends on d.

[3] Much of this analysis turns out to duplicate work by Spielman and Teng [29]. But we can go further by getting better balance ratios than Miller et al.'s $(d+1):1$; e.g. we can get $2:1$ using ellipsoids whose principal axes vary by at most a constant factor.

[4] Only particularly concise or important proofs are included herein.

[5] If d is odd, the factorial function will be applied to half-integral argument, which is no problem if you know that $(-1/2)!=\sqrt{\pi}$, $0!=1$, and $n!=n\cdot(n-1)!$.

Definition 4 *An "L_1-ball" of radius r, centered at the origin in d-space, is the convex hull of the $2d$ points with coordinates that are permutations of $(\pm r, 0, 0, ..., 0)$. (If $d = 3$, this is a regular octahedron.) Its d-volume is $(2r)^2/d!$.*

1.2.3 Generalizations of the notion of "disjoint" sets

Definition 5 *A collection of point sets are "κ-thick" if no point is common to more than κ sets. (E.g., 1-thickness is equivalent to disjointness.)*

Definition 6 *For $\lambda > 1$ define "(λ, κ)-thick:" no point is covered by more than κ of the objects whose linear dimensions vary by a factor at most λ.*

Definition 7 *A collection of compact measurable sets are "κ-overloaded" if for any subcollection of the objects, the sum of their individual measures is at most a constant κ times the measure of their union.*

Note that (λ, κ)-thick and κ-overloaded sets can be infinitely thick. See figure 1.

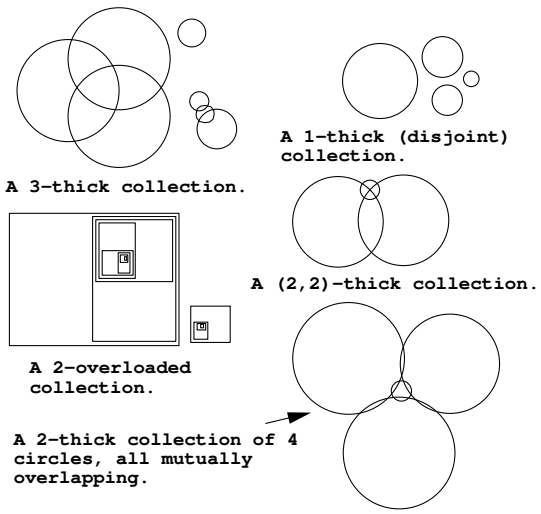

Figure 1: Some collections of objects.

2 Covering convex bodies with smaller copies

Definition 8 *A "convex d-body" is a compact convex set in $\mathbf{R}^d$, $1 \le d < \infty$, that has interior points. It is "strictly convex" if its boundary contains no line segment; equivalently, if every tangent hyperplane has exactly one point of contact with the body.*

Definition 9 *A "smooth point" on the surface of a convex body is one with a unique tangent hyperplane. A convex body will be called "smooth" if every point on its surface is smooth.*

2.1 Smaller scaled translated copies

Definition 10 *A "homothet" of a set is a scaled and translated copy of it. For an "s-homothet," the scaling factor is s, and if s is left undefined or unspecified, we will take it to mean "for some s, $0 < s < 1$."*

2.1.1 Balls, d-Simplices, Regular d-octahedra (L_1-balls), and d-Cubes

Theorem 11
For a d-ball, d-simplex, regular d-octahedron, and d-cube (or by an affine transformation, any parallelipiped) respectively: It is impossible to cover these by respectively d, d, $2d-1$, 2^d-1 s-homothets, but it is possible if we use 1 more than these. The minimum possible scale factors s are then $\sqrt{1-d^{-2}}$, $d/(d+1)$, $(d-1)/d$, $1/2$, respectively. ♣

The ball part of the theorem was previously known when $d \le 3$ [24].

2.1.2 General smooth convex d-bodies

Theorem 12 *It is impossible to cover any smooth convex d-body with d s-homothets. But every smooth convex body may be covered by $d+1$.* ♣

2.1.3 Any convex body, $d = 2$

Theorem 13 (Lassak [22]) *Any convex 2-body may be covered by 4 homothets (4 is best possible for a square), each scaled by $s = 1/\sqrt{2}$. This value of s is best possible.*

2.1.4 Hadwiger hypothesis

H.Hadwiger [4] conjectured in 1957 that the d-cube was the worst convex body, i.e. that $< 2^d$ smaller self-homothets would suffice to cover any convex d-body *except* for bodies affinely equivalent to a d-cube, which as we've seen require exactly 2^d. This remains open.

For a formulation as an illumination problem see Boltyanskii [3] [4].

Another equivalent formulation arises from applying Minkowski's [5] "polar map." The number of copies needed is the same as the number of open halfspaces through the origin needed to cover the surface of the polar body, and in such a way that each (closed) face of the polar body is entirely contained in at least one of the halfspaces. Thus, the polar equivalent to Hadwiger's conjecture is: "The surface of any convex d-body B containing the origin may be covered by 2^d open halfspaces through the origin in such a way that every (closed) face of B is entirely contained in at least one of the halfspaces."

A related geometric separator conjecture, which would yield new post-office [9] algorithms, is

Conjecture 14 *For any convex d-polytope with F faces, there exists a hyperplane with at least $\lfloor c_d F \rfloor$ faces entirely to each side of it. (Further conjecture: $c_d = 2^{-O(d)}$ suffices.)*

Lassak [23] showed any convex d-body could be covered by

$$(d+1)d^{d-1} - (d-1)(d-2)^{d-1} \sim (1-e^{-2})d^d \qquad (2)$$

s-homothets.

Schramm [35] showed that for any *strictly* convex d-body invariant under a group of reflections acting *irreducibly*[6], $d+1$ s-homothets suffice.

[6] That is, in such a way that only the body's centroid, or the entire body – but no other subset of the body – is preserved by every group action.

Rogers [34] showed that for centrally symmetric convex bodies,

$$(1+s^{-1})^d\,(\ln d+\ln\ln d+5)d \tag{3}$$

s-homothets suffice. This bound with $s=1$ is valid for s-homothets with s infinitesimally below 1.

Rogers and Zong [34] also found that for general convex bodies, $\binom{2d}{d}(\ln d+\ln\ln d+5)d$ s-homothets suffice. (This is better than Lassak's bound (EQ 2) when $d\geq 6$.) Also, a convex body K may be covered by $2^d(\ln d+\ln\ln d+5)d$ translated copies of its negation $-K$. Both these and (EQ 3) were consequences of the following theorem

Theorem 15 (Rogers and Zong's homothet covering theorem) *To cover a convex d-body K with translated copies of a convex d-body H, it suffices to use*

$$\frac{\mathrm{vol}(K-H)}{\mathrm{vol}(H)}\zeta_d(H) \tag{4}$$

copies, where $\zeta_d(H)$ is the minimal covering density of d-space by translates of H, and where "$K-H$" denotes the set $\{\vec{x}-\vec{y}\mid \vec{x}\in K,\ \vec{y}\in H\}$.

The consequences mentioned then follow from Rogers's earlier [32] bound $\zeta_d(H)\leq(\ln d+\ln\ln d+5)d$, and the inequalities

$$2^d\leq\frac{\mathrm{vol}(K-K)}{\mathrm{vol}(K)}\leq\binom{2d}{d} \tag{5}$$

which are respectively a special case of the "Brunn-Minkowski theorem" (EQ 6) and the upper bound from [33] [7]. These are tight respectively for a centro-symmetric d-body and a d-simplex.

The Brunn-Minkowski inequality [5] states that if A and B are convex d-bodies,

$$\mathrm{vol}(A)^{1/d}+\mathrm{vol}(B)^{1/d}\leq\mathrm{vol}(A+B)^{1/d}. \tag{6}$$

A reverse form of this inequality due to V.D.Milman [26] [29] [30] states that there exists an absolute constant C such that for each A, B there exists a determinant-1 affine transformation u such that

$$\mathrm{vol}(A+uB)^{1/d}\leq[\mathrm{vol}(A)^{1/d}+\mathrm{vol}(uB)^{1/d}]C_d,\quad C_d\leq C. \tag{7}$$

The minimal values of C_d seem to be totally unknown, aside from the fact that $1\leq C_d\leq C<\infty$. However,

Theorem 16 *For all $d\geq 1$, $C_d<\sqrt{2\pi d/e}$.* ♣

Consequently (see also [21])

Theorem 17 (Ultra general homothet covering theorem) *There exists an absolute constant C so that: For any convex d-bodies A and B, there exists a determinant-1 affine transformation u such that A may be covered by*

$$(\ln d+\ln\ln d+5)dC_d^d\,\frac{[\mathrm{vol}(A)^{1/d}+\mathrm{vol}(B)^{1/d}]^d}{\mathrm{vol}(B)} \tag{8}$$

translates of uB, where either $C_d\geq 1$ is defined by theorem 16, or $C_d=C$ (either is valid).

2.2 Rotations allowed?

Definition 18 *A "rotothet" of a set is a scaled and possibly rotated and translated, copy of it. For an "s-rotothet," the scaling factor is s, and if s is left undefined or unspecified, we will take it to mean "for some s, $0<s<1$."*

How many s-rotothets of a convex d-body are required to cover it, if the copies may each be independently rotated and translated?

The answer can be smaller than in the case where rotations are not allowed (although not for a regular simplex or a ball, in which cases $d+1$ are always required). For example[7], a 45-45-90 right triangle is tiled by only 2 copies of itself, each scaled by $1/\sqrt{2}$.

In d-space for all sufficiently large d, at least $1.203^{\sqrt{d}}$ s-rotothets are needed to cover the convex d-bodies constructed by Kahn and Kalai [20] [28].

When $d=2$, we have

Theorem 19 *Any 2D convex body may be covered by 3 smaller rotothets — and 3 are necessary for the disk and for the equilateral triangle.*

Proof. Let the body be called Q. We may assume Q is not a body of constant width, because [8] then 3 s-homothets, $s=0.9101$, would suffice. In the below we'll implicitly use the fact that a convex body's surface is continuous and (in 2D) "differentiable to the left" (and right). It is also differentiable almost everywhere[8].

So assume Q has width equal to its diameter (which is the maximum possible width) only over some fraction γ, $0\leq\gamma<1$, of the possible rotation angles.

Rotate the object randomly within a set, of measure μ, $0<\mu<1-\gamma$, in which we get East-West widths bounded above by some bound below the diameter of Q. Translate 2 copies, shrunk a sufficiently tiny amount, a sufficiently tiny amount North and South. Now only 2 small bits will be left uncovered on the East and West (unless, e.g. the westmost point is non-unique, which won't happen with probability 1 because of the random rotation). Now the maximum distance between any pair of points in these 2 small bits is less than the diameter of the object. Hence these two small regions may both be covered by a third rotothet. □

2.2.1 Simplices divisible into two congruent scaled versions of themselves

Conjecture 20 (The chinese egg simplex conjecture) *The only d-simplex divisible into two congruent scaled versions of itself, in any dimension $d\geq 2$, is the 45-45-90 right triangle in 2D.*

Theorem 21 *Conjecture 20 is true in 2D. It is also true in 3D if attention is restricted to simplices in which the cutting plane is a plane of mirror symmetry.* ♣

[7] Conjecturally, this example is unique – see §2.2.1.

[8] In fact, convex d-bodies are 2-time differentiable almost everywhere, as was proved by A.D.Aleksandrov in 1936. Some results of this kind are surveyed in [18].

The full paper also gives a plausibility argument for conjecture 20 for $d = 3$ and all $d \geq 4$. The latter is based on counting the number of equations constraining the choice of possible cutting hyperplanes, and seeing that it is less than the number of parameters defining such a hyperplane. In any fixed dimension it should be possible to settle the conjecture by a finite case analysis.

2.3 Affinities allowed? Answer: 2?

Definition 22 *An "affinity" of a set is an affine (i.e. general linear) transformation of it. For an "s-affinity," the volume scaling factor (|determinant|) is s^d, and if s is left undefined or unspecified, we will take it to mean "for some s, $0 < s < 1$."*

How many s-affinities (where the affine transformations are chosen independently for each copy) are needed to cover a convex d-body?

We can tile d-boxes and d-simplices with only 2 affine versions of themselves, each with $s^d = 1/2$, i.e. half volume. A unit volume d-ball can be covered by two d-ellipsoids, each of volume

$$s^d = (\frac{d}{d+1})^{(d+1)/2}(\frac{d}{d-1})^{(d-1)/2} < \exp(\frac{-1}{2d}), \qquad (9)$$

and this is tight[9]. These facts suggest

Conjecture 23 (Two affines?) *Two s-affinities of a convex d-body always suffice to cover it[10]!*

Theorem 24 *Conjecture 23 is true for smooth convex d-bodies (they may even be allowed to have 1 non-smooth point), and also for general convex bodies when $d = 2$.* ♣

3 Separator theorems about geometrical objects

3.1 A separator theorem for d-cubes

Theorem 25 (Cube separator theorem) *Let there be a set $\mathcal{S}$ of N iso-oriented d-cubes in a euclidean d-space, whose interiors are κ-thick, or more generally κ-overloaded (definition 7) or (λ, κ)-thick (definition 6). Then there exists an iso-oriented d-box (with maximal sidelength ≤ 2 times the minimal sidelength) with at most $2N/3$ cube interiors entirely inside it, at most $2N/3$ cube interiors entirely outside (where the restrictions that "cardinality $\leq 2N/3$" may be generalized to be "weight $\leq 2/3$" where "weight" is defined via any of a wide class of mass-1 measures on d-space, which needn't bear any relation to $\mathcal{S}$), and moreover:*
(a) Let $\epsilon > 0$. If the interiors are κ-thick, the number of cube interiors partly inside and partly outside the box is

$$\leq c_d(\epsilon)\,\kappa^{1/d}\,N^{1-1/d} + \left(\frac{2}{\epsilon}+2\right)^d \kappa \qquad (10)$$

where

$$c_d(\epsilon) = \left(\frac{[1+2^d(H_{2d}-H_d-\frac{1}{2}+O(\epsilon d))]\,(2d)!}{d!}\right)^{1/d} \qquad (11)$$

provided ϵd is sufficiently small, where $H_m = \sum_{j=1}^m j^{-1}$ is the mth harmonic number.
(b) Let $K > 0$. If the interiors are κ-overloaded, the number of cubes with side $< K$ times the maximal sidelength of the box and whose interiors are partly inside and partly outside the box is $O\left(Kd\kappa^{1/d}N^{1-1/d}\right)$.
(c) If the interiors are (λ, κ)-thick, then given any $K > 0$, the number of cubes which have side less than K times the maximal sidelength of the box and whose interiors are partly inside and partly outside the box is

$$[\frac{\lambda}{1-\lambda^{-(d-1)}} + \kappa^{1/d}O(Kd)]N^{1-1/d}. \qquad (12)$$

Remark. Take κ fixed. When $N \to \infty$ with d fixed, best results in (a) are obtained if ϵ is of order $N^{-\frac{d-1}{(d+1)d}}$, in which case (EQ 10) is $\leq c_d(0)\kappa^{1/d}N^{1-1/d}(1+o(1))$. Note that $H_{2d} - H_d < \ln 2$, with equality in the limit as $d \to \infty$.
Remark. When $d = 2$ and $d = 3$ the main term in the right hand side of (EQ 10) (ignoring ϵ's) becomes respectively $4\sqrt{\kappa N}$ and $(232\kappa)^{1/3}N^{2/3}$. When $d \to \infty$, it is asymptotic to

$$\frac{8}{e}d\kappa^{1/d}N^{1-1/d}. \qquad (13)$$

Proof sketch. We will begin by finding a maximum d-volume box B_0 of specified shape with at least $1/3$ of the weight outside it and any homothet of it. Next we'll divide this box into two smaller boxes. By the pigeonhole principle, at least one of the sub-boxes must contain $1/3$ of the weight inside it. Finally, we will argue that the boundary of a box containing the good sub-box but contained in B_0, chosen at random from a certain probability distribution, will in expectation cross few cubes. This will prove (a). Proving (b) and (c) requires a few extra tricks.
(Almost complete) proof. Part (a) is treated first, so assume the cube interiors are κ-thick. Let B be a brick (iso-oriented d-box) with side length ratios $d+1{:}d+2{:}\ldots{:}2d$ (not necessarily in that order) having maximal d-volume subject to the constraint that every iso-oriented brick congruent to it has at least $N/3$ cube interiors (or at least $1/3$ of some arbitrary[11] measure on d-space) entirely outside.

WLOG (by scaling) B_0's dimensions are in fact exactly $(d+1) \times (d+2) \times \ldots \times (2d)$. Note that an infinitesimally

[9] This fact underlies the "ellipsoid algorithm" for convex programming, cf. lemma 3.1.34 of [17]. If the principal axes of the ellipsoid are $L_1 \leq L_2 \leq \ldots \leq L_d$, then we can slice it with a hyperplane through its center and perpendicular to the L_d axis, and then cover each of the two resulting hemiellipsoids with ellipsoids with axes $dL_i/\sqrt{d^2-1}$ for $1 \leq i \leq d-1$, and $dL_d/(d+1)$. Let the DW aspect ratio of the original ellipsoid be $A = L_d/L_1$. Notice that if $A \geq \sqrt{(d+1)/(d-1)}$, then the two smaller ellipsoids will have DW aspect ratios $\leq A$, while if $A \leq \sqrt{(d+1)/(d-1)}$, then the two smaller ellipsoids will have DW aspect ratios $\leq \sqrt{(d+1)/(d-1)}$ also.

[10] One might further conjecture that the scaling factor on the left hand side of (EQ 9) is the worst possible, i.e. the ball is the worst convex object.

[11] The measure defining the notion of "weight" has to be such that this maximum can exist, or at least the corresponding supremum, in which case B's size is chosen arbitrarily close to the supremum.

expanded version of B_0 will have $\leq 1/3$ of the weight outside it[12].

Cut B_0 in half by bisecting its largest dimension to get two subbricks B_i and B_{ii} with side lengths $d \times (d+1) \times \cdots \times (2d-1)$. By the *pigeonhole principle* at least one of these contains or intersects at least $N/3$ cubes of $\mathcal{S}$ ($\geq 1/3$ weight).

WLOG that subbrick is B_i. Also WLOG (by a translation) B_i's "lower left" corner (i.e. the one with minimal coordinates in every direction) lies at the origin.

Note that if every sidelength of B_i were expanded by adding 1, the result would be a brick congruent to B_0. So consider in fact the family $\mathcal{F}$ of translates of B_0 whose min-coordinate corners are $(-t, -t, -t, \ldots, -t)$ for $0 < t < 1$.

Every member of $\mathcal{F}$ contains B_i and hence contains or intersects at least $N/3$ cubes of $\mathcal{S}$. Also, every member of $\mathcal{F}$, like B_0, is contained in a brick congruent to B_0 and hence (by the definition of B_0) must have at least $N/3$ cube interiors entirely outside it. So in order to complete the proof, we need only to show that some member (in fact, a random member) of $\mathcal{F}$ cuts sufficiently few cubes of $\mathcal{S}$. See figure 2 for a picture in 2 dimensions.

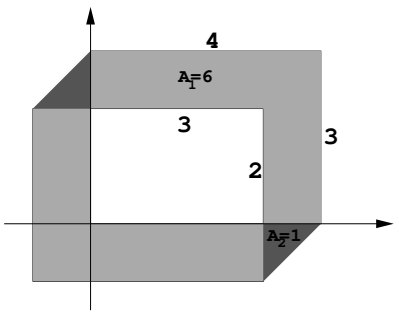

Figure 2: Sliding boxes in $\mathcal{F}$.

Let $A_0 = (2d)!/d!$ be the total d-volume (i.e. area, in the case $d = 2$) of B_0. Let the d-volume swept out by the surfaces of the boxes in $\mathcal{F}$ as t goes from 0 to 1 be $A_1 + A_2$, where A_k is the portion swept through k times by the box surface. (Notice that $k = 1$ and $k = 2$ are possible, but $k \geq 3$ is not, due to convexity.) Thus in figure 2, A_2 is represented by the two triangular areas.

By considering the volume of the annulus swept out by the brick's surface,

$$A_1 + 2A_2 = 1 \cdot S_0 \tag{14}$$

where $S_0 = 2A_0 \sum_{j=d+1}^{2d} j^{-1} = 2(H_{2d} - H_d)A_0$ is the surface $(d-1)$-area of B_0. Meanwhile

$$A_1 + A_2 = \frac{S_0 + A_0}{2} \tag{15}$$

is the volume swept out by the "top" (i.e. with higher coordinate values) faces of the brick, plus the relevant volume inside the bottommost brick in $\mathcal{F}$ (i.e. $A_0/2$ since the "hole" in the middle is half the volume). Solving, we find $A_1 = A_0$ and $A_2 = (S_0 - A_0)/2$.

[12]For the precise details of how to repalce "infinitesimally" with "ϵ's and δ's," see the full paper.

Armed with the above d-volume formulae, we're ready to proceed. If $\epsilon > 0$, we may discard all cubes of sidelength greater than ϵd which intersect the annulus described above, because (due to κ-thickness) the number of cubes we thereby ignore will be $\leq (2/\epsilon + 2)^d \kappa$, i.e., a constant independent of N, which we've added to the right hand side of (EQ 10) to compensate.

Give the space inside A_k monetary value k^d dollars per unit d-volume ($k \in \{1, 2\}$). Let t_i be the probability that a random member of $\mathcal{F}$ cuts the ith d-cube. If this cube has side length s_i then $t_i = ks_i$, and the space it occupies is worth $k^d s_i^d = t_i^d$ dollars, if it is wholly inside A_k. It is easily checked that cubes which are partly in A_1 and partly in A_2 also have value at least t_i^d dollars. Cubes in $\mathcal{S}$ which are partly outside $A_1 \cup A_2$ will also have this property if we give value 2^d per unit volume to all the space outside $A_1 \cup A_2$ which can be covered by cubes of side length ϵd which intersect $A_1 \cup A_2$. Note this space includes part of the "hole," but nevertheless its volume is $O(\epsilon d A_0)$.

If C is the expected number of cubes of $\mathcal{S}$ cut by a random member of $\mathcal{F}$, then

$$C = \sum_{i=1}^{N} t_i. \tag{16}$$

But any point in space can only be covered by at most κ cubes in $\mathcal{S}$, which gives the total-dollars constraint

$$\sum_{i=1}^{N} t_i^d \leq (A_1 + 2^d A_2 + O(2^d \epsilon d A_0))\kappa. \tag{17}$$

Maximizing C subject to this constraint, Hölder's inequality implies that the maximum occurs when all t_i are equal, and we finally get

$$\begin{aligned} C &\leq N^{1-1/d}(A_1 + 2^d A_2 + O(2^d \epsilon d A_0))^{1/d} \kappa^{1/d} \\ &= N^{1-1/d} A_0^{1/d} \kappa^{1/d} \\ &\quad \times (1 + 2^d [H_{2d} - H_d - 1/2 + O(\epsilon d)])^{1/d}. \end{aligned} \tag{18}$$

This is equivalent to (EQ 10). Since some member of $\mathcal{F}$ must cut the expected number (or fewer) cubes of $\mathcal{S}$, we are done. (The asymptotic behavior in (EQ 13) comes from Stirling's formula $x!^{1/x} \sim x/e$.) Part (a) is proven.

For part (b), assume $\mathcal{S}$ is κ-overloaded. Follow the proof of part (a), but write $t_i = 2s_i$ for an upper bound on the probability that the ith d-cube is cut, and instead of giving different parts of space different values, note only that the total volume which can be covered by cubes with sidelength at most K times any of the sidelengths of the d-box is $O(Kd)^d$. By the κ-overloaded property, the total volume of all these cubes is at most $\kappa \cdot O(Kd)^d$. With this upper bound on $\sum s_i^d$, maximising $\sum 2s_i$ yields the required bound.

For part (c), assume $\mathcal{S}$ is (λ, κ)-thick. Again, follow the proof of part (a), but this time discard the cubes of sidelength $\geq 2Kd$ (not ϵd as before). The expected number of intersections C, which must be maximised, is (EQ 16) as before, but the constraint (EQ 17) does not necessarily hold since there is only a bounded number of overlaps

on cubes which are roughly the same size. Write the summation (EQ 16) as $S_1 + S_2$ where S_1 is the contribution from all $t_i < (A\kappa/N)^{1/d}$ and S_2 is the rest. Immediately, $S_1 < A^{1/d}N^{1-1/d}$. To bound S_2, note by the definition of (λ,κ)-thickness that for all x,

$$\sum_{x \le t_i \le \min\{\lambda x, 2Kd\}} t_i^d \le A\kappa \tag{19}$$

for some constant $A = O(Kd)^d$ related to the volume of the annulus, cf. (EQ 17).

For $j \ge 0$ let b_j denote the number of i for which $\lambda^j \le t_i N^{1/d}(A\kappa)^{-1/d} \le \min\{\lambda^{j+1}, 2Kd\}$. By the constraint above, $\lambda^{dj} b_j/N \le 1$ for each j, i.e.

$$b_j \le \lambda^{-dj} N. \tag{20}$$

Hence

$$S_2 < \sum_{j\ge 0} \lambda^{j+1} b_j N^{-1/d} \le \frac{\lambda}{1-\lambda^{-(d-1)}} N^{1-1/d} \tag{21}$$

if $d \ge 2$. Hence

$$C = S_1 + S_2 \le \left(\frac{\lambda}{1-\lambda^{-(d-1)}} + (A\kappa)^{1/d}\right) N^{1-1/d}, \tag{22}$$

as required for (c). □

Remarks:

(i) Theorem 25 and its proof also hold inside a "cubical[13] d-torus" (iso-oriented d-box with opposite faces identified to give "periodic boundary conditions"). ♣

(ii) If, instead of iso-oriented d-cubes, $\mathcal{S}$ consists of any other assortment of objects with the property that their "CV aspect ratio" (definition 2) is bounded above by τ, then the theorem still holds, but with κ in (EQ 10) replaced by $\kappa\tau^d$.

(iii) It is also possible to demand that the separator shape be something other than an iso-oriented d-box. Let us recall the proof technique. We began by finding a maximal d-volume separator shape B_0 with at least a constant fraction p of the weight outside it. We then subdivided this shape into 2 smaller sets – more generally we may cover, rather[14] than subdivide, the shape by $m \ge 2$ sets. By the pigeonhole principle, at least one of these sets, call it Q, must contain $\ge (1-p)/m$ of the weight. (The best value of p to pick is $1/(m+1)$; we are going to prove a $(m/[m+1], O(N^{1-1/d}))$ separation result.) We then argued that a random separator shape congruent to B_0 and containing Q would, in expectation, cross no more than $cN^{1-1/d}\kappa^{1/d}$ objects from $\mathcal{S}$. Here the value of c will depend upon the separator shape, the shape and size of the covering objects, and the probability distribution you use.

The full paper gives a very general separator theorem and a table of important specific separator theorems. In the present condensed paper we'll content ourselves with the following ultra simplified and weakened version.

[13] Or rectangular – provided the torus has bounded ratio of maximum to minimum sidelength. See the full paper for the precise statement.

[14] And this is the point of the large number of "covering by smaller copies" theorems in §2.

Theorem 26 (Highly general separator theorem) *Let there be a κ-thick set of N convex objects in d-space, each with DW aspect ratio $\le B_1$. Let Q be a convex shape of DW aspect ratio $\le B_2$ that may be covered by m smaller copies of itself, each copy having s^d times smaller d-volume, and let $0 < \epsilon < s/2$. Then there exists a convex body T equivalent to Q such that $\le mN/(m+1)$ of the objects are completely inside T, $\le mN/(m+1)$ are completely outside, and at most*

$$O_Q(1+\epsilon)B_1B_2^2(\kappa \ln\frac{1}{s})^{1/d}N^{1-1/d} + O_Q(1+\frac{1}{\epsilon})^d\kappa \tag{23}$$

are partly inside and partly outside. ♣

In theorem 26, "smaller copies" and "equivalent" are allowed to mean either homothets, or rotothets, or even volume-reducing affinities, provided these are affinities which preserve the validity of the bound on aspect ratio, e.g. as in footnote 9.

3.2 A separator theorem for d-boxes

Theorem 27 (Separator hyperplane for κ-thick iso-oriented d-boxes.) *Given N iso-oriented d-boxes whose interiors are κ-thick: There exists a hyperplane, orthogonal to a coordinate axis, such that at least*

$$\lfloor (N+1-\kappa)/(2d) \rfloor \tag{24}$$

of the d-box interiors lie to each side of the hyperplane.

Proof[15].

When $d = 1$, the result is easy. So suppose $d \ge 2$.

Find the leftmost hyperplane with at least $\lfloor (N+1-\kappa)/(2d) \rfloor$ boxes entirely to its left. (Or any hyperplane between this and its right-facing alter ego.) If there are $\lfloor (N+1-\kappa)/(2d) \rfloor$ boxes to its right, we are done. Otherwise, we have a $(d-1)$-dimensional problem inside the hyperplane, and by induction on d, by solving this problem we will get a hyperplane with at least $\lfloor (M+1-\kappa)/(2d-2) \rfloor$ boxes to each side of it, where $M+1 > N+1-2\lfloor (N+1-\kappa)/(2d) \rfloor$, i.e. $M+1 > (1-1/d)(N+1)+\kappa/d$. Hence, the number of boxes on each side will be $\ge \lfloor (1-1/d)(N+1-\kappa)/(2d-2) \rfloor = \lfloor (N+1-\kappa)/(2d) \rfloor$. □

The full paper also gives some other variants of this result and shows

Theorem 28 (Algorithmic version) *The hyperplane of theorem 27 may be found in $O(Nd)$ steps. Also, if the $2d$ lists of the lower and upper endpoints of the intervals defining the boxes's ith coordinates are pre-*sorted*, then the* best *such hyperplane (according to a wide variety of optimality measures) may be found in $O(Nd)$ steps.* ♣

[15] This generalizes [1] to work for arbitrary κ.
Examples in [1] show that (EQ 24) is best possible for an infinite number of N for each $d \ge 1$, when $\kappa = 1$. We also have d-dimensional examples with $\kappa = 1$, $N = 2md$ in which no iso-oriented hyperplane cuts off more than m boxes, and with $\kappa = 2$, $N = 2 + 2d$ in which no iso-oriented hyperplane cuts off more than 1 box.

3.3 Some miscellaneous separation results

Definition 29 *A "Rado point" of a unit mass measure μ in d-space is a point such that any hyperplane through it will split the the measure so that $\leq d/(d+1)$ mass will lie in either open halfspace. A "tight Rado point" is a Rado point such that there exists a hyperplane through it for which this inequality is tight.*

R.Rado [31] showed that any measure in d-space has a Rado point, indeed a tight one. The 1:d balance guarantee can't be improved [11] for *general* measures, but:

Theorem 30 ($1/e$ split for convex bodies) *Any hyperplanar cut through a convex d-body's center of mass cuts off $> e^{-1} \approx .367879$ of the volume.*

Proof. Call the body B. WLOG let its center of mass be $\vec{0}$ and the cut halfspace be $x \geq 0$. We may WLOG apply "Steiner reflection symmetrizations" to B about any "mirror hyperplane" M containing the x-axis. (See §9, pages 76-77, of [5].) Thus WLOG B is a convex body of revolution with axial line being the x-axis. Let its maximal-x point be $x_{\max}$. Then the "worst" convex B (minimizing the volume fraction cut off) is uniquely a cone with apex $x_{\max}$ and with base at $x = -x_{\max}/(d+1)$. The result follows. □

Theorem 31 *Given $\binom{k+d}{d} - 1$ measures in d-space, there exists an algebraic surface of degree k which bisects them all (that is, once the surface is removed, two sets – where the degree-k polynomial is > 0 and < 0 – remain, and each set contains $\leq 1/2$ of the mass of each measure) simultaneously.*

Remark. Particular cases of this theorem include [16] "Given k measures on the real line. Then there exist $\leq k$ points such that if the line is cut at those points, and the odd numbered pieces are collected together and called S, evens are T, then S and T each have $\leq 1/2$ of each of the k measures, simultaneously." and "Given 5 measures in the plane, there exists a conic curve which bisects them all simultaneously[16]."

Proof. To prove the latter special case, map the plane into 5-space via

$$(x, y) \;\rightarrow\; (x, y, x^2, y^2, xy) \tag{25}$$

then apply the "ham sandwich theorem [39]." The general case is similar. □

3.4 Strengthenings of Miller & Thurston's sphere separator theorem

Theorem 32 (Strengthened Miller Thurston sphere separator theorem I) *Given N balls in d-space, whose interiors are κ-thick, there exists a sphere S such that $\leq N/(d+2)$ of the balls lie entirely inside S, $\leq N/(d+2)$ of the balls lie entirely outside S, and $\leq c_d \kappa^{1/d} N^{1-1/d}$ of the balls are cut by S, where $c_1 = 1$, $c_2 = 2$, $c_3 < 1.317$, $c_4 < 1.014$, $c_5 < 0.840$, and more generally if $d \geq 2$,*

$$c_d \leq \frac{1}{2d}\left(\frac{\mathsf{O}_{d+1}}{\bullet_d}\right)^{1+1/d} = \left(\frac{\pi}{2d}\right)^{1/2} \cdot \left[1 + \frac{1}{3d} + O(d^{-3/2})\right]. \tag{26}$$

(See definition 3.) This also works with any of a wide class of weight functions where instead of demanding that $\leq N/(d+2)$ of the balls lie entirely inside and outside, we demand $\leq 1/(d+2)$ of the weight lie inside or outside. ♣

This strengthens the theorem which was the central achievement in [25], but for an inexplicable reason the results there were stated in weak forms involving "big-O" notation with undetermined constants (which in fact could depend on d in a totally unspecified way)[17]. Most of the present theorem was shown independently by [29]. The proof depends on mapping the balls to spherical caps on the surface of a sphere in $d+1$ dimensions by a reverse "stereographic projection." A topological argument shows that the sphere may be chosen so that its center is a "Rado point" of representative points in the caps (or more generally, of some arbitrary weight measure). Then a probabilistic argument considering cap areas shows that a random hyperplanar cut through this sphere's center will (in expectation) cut few caps; by the Rado point property it will divide the weight in a 1:d balance at worst; this cut maps back to a spherical cut in $\mathbf{R}^d$.

Theorem 33 (Additional strengthening of Miller by constant factors) *The bound $c_d \kappa^{1/d} N^{1-1/d}$ of the previous theorem may be decreased as follows.*

1. *If $\rho(d, \kappa) \in (0, 1]$ is the maximal possible average euclidean radius achievable by a κ-thick set of N spherical caps on a unit sphere in $(d+1)$-space,* expressed as a fraction *of the euclidean radius of a spherical cap of area $\mathsf{O}_d \kappa / N$, then we may multiply c_d by*

$$\rho(d, \kappa)\,(1 + o(1)). \tag{27}$$

   For all sufficiently large d, $0.5 \leq \rho(d, 1) < 0.660185901$ if N is large enough.

2. *For the unweighted case only, we may (also) multiply c_d by*

$$\frac{1 + (d+1)^{1-1/d}}{(4+2d)^{1-1/d}}\,(1 + o(1)), \tag{28}$$

   which when d is large is approximately $1/2$.

Here the "$o(1)$" terms apply in the limit where $\kappa_d N^{d-1} \rightarrow \infty$ with d fixed.

Remark. When $d = 2$, we had $c_2 = 2$. Applying item 1 only yields $c_2' < 1.90463$, while applying both yields $c_2''' < 1.83973$.

[16] That is, once the curve is removed, each piece of the plane that remains (there hereby are two, where the quadratic form is > 0 and < 0) has $\leq 1/2$ of the mass of each measure.

[17] The Miller-Thurston theorem cannot be made to hold for separating, e.g., squares, at least not by means of any simple trick such as circumscribing circles about the squares. This is because configurations of disjoint squares exist whose circumballs are infinitely thick.

Remark. When $d = 3$, we had $c_3 \approx 1.317$. Applying item 1 alone yields (conjecturally) $c_3' \approx 0.91820$; applying item 2 alone yields $c_3'' \approx 0.99853$; and applying both yields (conjecturally) $c_3''' \approx 0.91685$.

Proof sketch. The two methods of improving the constants in the previous theorem are respectively: known bounds on the wasted area in packings (or κ-thick configurations) of spherical caps may be employed to reduce the bound by ~34-50%, and in the unweighted case we get another ~50% off by intentionally using a "tight" Rado cut, which exactly has a 1:d balance. □

The full paper gives an extensive discussion of known and conjectured results re $\rho(d,\kappa)$, e.g. $\rho(2,1) = \pi/\sqrt{12}$, $\lim_{\kappa\to\infty}\rho(d,\kappa) = 1$, $0.5 \le \lim_{d\to\infty}\rho(d,\kappa)^{1/d} < 0.660185901$, and conjecturally $\rho(3,1) \le 0.77413$.

3.5 Two thirds is best possible

Theorem 34 *For each $N > 0$, there exists a set of N disjoint disks in the plane, such that any circle C, cutting $O(\sqrt{N})$ of the disks, and with at most fN disks entirely inside and at most fN entirely outside, necessarily has $f \ge 2/3 - o(1)$.*

Proof. Make 3 small-diameter clusters of $N/3$ disks located near the vertices of an equilateral triangle. Each cluster will be an "exponential spiral" of disks whose radius increases by a constant factor every turn of the spiral, and with a constant number of disks per turn, and such that each disk touches its neighbors in the spiral ordering. For example, one could use the "loxodromic progression" of circles (pictured in figure 9 page 114 of [11]) each of radius $g + \sqrt{g} \approx 2.89005$ (where $g = (1+\sqrt{5})/2 \approx 1.61803$ is the "golden ratio") in which any 4 consecutive disks are mutually tangent.

In order to get $f < 2/3$, C would have to cut at least 1 of the 3 clusters into two parts, each of cardinality of order N. If C's radius is large compared to the diameter of a cluster, then this is impossible without cutting order N disks. If C's radius is comparable to or smaller than the diameter of a cluster, then both the other clusters will lie outside C. □

Remarks.

(i) The same 2D example, but with d-balls instead of 2-balls, shows that $f < 2/3$ is also impossible for sphere separators for balls in d-space, for any $d \ge 2$.

(ii) This also works in a large number of other scenarios, for example if the separator is a square, ellipse of bounded eccentricity, or equilateral triangle, rather than a circle. Or if the objects are squares rather than discs. In all these cases also, 2/3 is best possible[18].

[18] One nice proof for squares arises from clusters that are subsets of the well known "golden spiral" tiling of the plane by squares, where each successive square's sidelength is g times larger. This is best made by starting with a golden rectangle ($g \times 1$), dividing it into a square and a golden rectangle, and continuing on. (Also see figure 2.4.9 of [19].)

3.6 The dependency on d, κ, and N is best possible up to a factor ~ 4.5

In theorem 25, the number of d-cubes partly inside and partly outside the separating box was

$$O(d\kappa^{1/d}N^{1-1/d}) \tag{29}$$

as $N \to \infty$.

Theorem 35 (Construction showing best possible) *For κ-thick objects, this bound (EQ 29) is optimal up to the implied constant factor in the O, assuming we are living in toroidal d-space as in remark (i) after theorem 25.*

Proof sketch. When $d = 1$, the theorem is trivial, so assume $d \ge 2$.

Consider placing iso-oriented unit cubes so that their lower-coordinate corners lie at the points of the lattice consisting of the integer linear combinations of the rows of the following $d \times d$ Toeplitz[19] matrix M:

```
 1  0  0  ...  0  0  X
-X  1  0  ...  0  0  0
 0 -X  1  0  ...  0  0
 0  0 -X  1  0  ...  0
             ...
 0  0  0  ...  0 -X  1 .
```

We claim ♣ that this is a tiling of d-space by cubes of side 1 and cubes of side X, (for any desired X, $0 < X < 1$), such that each side-X cube is adjacent to $2d$ side-1 cubes (1 per face). Each side-1 cube is adjacent to $2d$ side-X cubes (1 per face) as well as to some side 1-cubes.

This tiling is of independent interest. When $d = 2$ it was known to the ancients (cf. figure 2.4.2g of [19]), but already when $d = 3$ it's apparently new.

If X is irrational, then every point of every hyperplane (except for the union of a measure-0 set with a set of L_∞ diameter $\le 1 + X$) lies inside a cubical tile. Thus in an extremely strong sense, there are no good cut hyperplanes. Also, no two cubes share a (mutually complete) $(d-1)$-face[20], or for that matter, apparently any k-face for any $k > 0$.

If $X = p/q$ is rational we have a tiling of a cubical d-torus of side $q+p$ by cubes of sides 1 and X, such that every point of every hyperplane (except for the union of a measure-0 set with a set of L_∞ diameter $\le 1 + X$) lies inside a cubical tile.

Tile a d-torus universe with a roughly $N^{1/d} \times N^{1/d} \times \ldots \times N^{1/d}$ grid where each cube has side 1 or side α, where α is arbitrarily close to 1, as above.

Now, consider a separating d-box. This box cannot be a 1/3-2/3 separator unless its "inside" and "outside" both have d-volumes between $.33N$ and $.67N$.

Now, the box's surface area S (by an isoperimetric theorem for d-boxes which is readily proven by Steiner symmetrization) must then obey $S \ge 2d(.33N)^{(d-1)/d}$. Then,

[19] A matrix is "Toeplitz" if it is constant on diagonals.

[20] The small cubes do share their entire face with part of the face of a large neighbor, though.

for any d-box with surface area S, we claim at least $(1-o(1))S$ of our cubes intersect its surface[21].

To conclude, we've shown the existence of an example with $\kappa = 1$ such that any 1/3-2/3 separating box must cross at least

$$2d\kappa^{1/d}(.33N)^{1-1/d} \tag{30}$$

of our cubes. The same result for any $\kappa > 1$ arises by superimposing κ copies of the example with $\kappa = 1$. □

This is within an asymptotic factor of 4.5 of the upper bound (EQ 13) of theorem 25.

3.7 Algorithmic versions

For this subsection, assume the input consists of iso-oriented d-cubes with disjoint interiors. Extensions to the κ-thick case are quite easy. In $O(N^{2d+1})$ steps, one could consider all possible inequivalent rectangle shapes and thus find the best rectangle separator by brute force. However, this approach is inefficient.

A better approach is based on the idea of a "separating d-annulus." This is two concentric d-boxes of bounded CV aspect ratios and with a ratio of linear dimensions bounded below by some constant greater than 1, such that at least a constant fraction of the objects's boundaries lie inside the inner box, and at least a constant fraction lie outside the outer box.

If we can find a separating annulus, then it immediately follows from the assumption that the interiors of the objects are disjoint – by a randomizing argument similar to the one in the proof of theorem 25 involving "$\mathcal{F}$" – that a random d-box containing the inner box and contained in the outer box, will cut an expected number of $O(N^{1-1/d})$ of the objects. Since a nonnegative random variable lies at or below twice its expectation value with probability $\geq 1/2$, we may, then, simply guess a box and then confirm in $O(Nd)$ steps that it works (with an expected number of 2 guesses being required before succeeding)[22].

We know of two efficient ways to find a separating d-annulus; one is deterministic and the other is randomized.

3.7.1 Deterministic method

For simplicity, we will describe the method in 2D assuming the objects we are separating are iso-oriented squares. We will find a separating d-annulus made of two concentric d-cubes with sidelength ratios 1 : 3. Unfortunately, this method does not achieve the optimal constant "2/3" in the split in theorem 25.

1. We assume WLOG that the N input squares are in general position, in particular the coordinates of all their corners are distinct. The reason we may assume this is because when we input them we could preshrink them by random factors selected from the range $(1-\epsilon, 1)$, where $\epsilon > 0$ is infinitesimal.

2. For any $\epsilon > 0$, we can find[23] in linear time numbers x_1, x_2, x_3 such that the four intervals $(-\infty, x_1]$, $(x_1, x_2]$, $(x_2, x_3]$, (x_3, ∞) each contain 1/4 (plus or minus ϵ) of the x-coordinates of the left hand sides of the squares.

3. Choose whichever of the two finite intervals I is shorter – WLOG $I = (x_2, x_3]$. Let W denote the set of squares whose left sides fall in $(x_2, x_3]$ and now do the same trick with y-coordinates of tops of squares in W – giving intervals $(-\infty, y_1]$, $(y_1, y_2]$, $(y_2, y_3]$, (y_3, ∞) – let J be the shorter of the two finite ones. Then $1/16 \pm 2\epsilon$ of all squares have their top left corners in the rectangle $R = I \times J$. Choose whichever of I and J is longer, let the inner square S_0 of the annulus just contain R, and let the outer square S_1 have the same center but with side *almost* three times longer.

4. Conclusion: Immediately $R \subseteq S_0$ contains bits of at least $N/16$ squares. Also, S_1 does not reach across the other of the two inner intervals in the long direction (e.g. if the long direction of R is x-direction, it does not reach x-coordinate x_1), and so at least $N/4$ of the squares are not wholly inside S_1.

The full paper shows how to generalize this algorithm to d dimensions while guaranteeing that at least $1/(2^{d+2}-2^d-2)$ of the weight lies on each side of the separating box.

3.7.2 Randomized method

The randomized method is as follows. Pick a random subset q of the N objects, and find their optimal separator (or just find any separating d-annulus) by brute force in (e.g., if the objects are iso-oriented d-cubes) $O(|q|)^{d+1}$ steps. (Anyhow, some function of $|q|$ and d steps, more generally.)

Now we claim that, if $|q|$ is a sufficiently large constant – it will suffice if $|q| = \Omega(\epsilon^{-2}d^3\ln(d/\epsilon))$ – then the resulting separator will in fact *be* a separator with a split balance only $1+\epsilon$ times worse than best possible, with constant success probability. Success may be verified in $O(Nd)$ steps.

This may be proven using VC dimension techniques ♣ depending on

Lemma 36 *The VC dimension of doubly convex d-annuli (that is, convex d-bodies with a convex d-body "hole" removed) is $\leq 2d+2$.*

Proof sketch. The following $(2d+3)$-point set can't be shattered: $d+1$ points forming the vertices of a d-simplex, another $d+1$ forming an enclosed d-simplex, and a final point enclosed by both simplices. □

3.7.3 Comparison

The approach of §3.7.1 ran in $O(dN)$ time but produced a possibly exponentially poorly balanced – $\Omega(2^d)$:1 – split. The randomized approach (§3.7.2) produces a well balanced split – $O(1)$:1 at worst, and within $(1+\epsilon)$ of the optimal balance, with high probability – but consumes $(d/\epsilon)^{O(d)}N$ (expected) runtime. It maximally generalizes (to separating other things than cubes with other things than boxes)

[21] Worries about overcounting cubes which intersect more than one face of the box may be avoided by realizing that they are asymptotically negligible if every sidelength of the d-box is $\gg 1$. On the other hand, if any sidelength were only $O(1)$, then the box's surface area would be far larger than our bound on S, easily overcoming any $2d$ overcounting factor if $N = d^{\Omega(d)}$.

[22] The full paper also shows how to find the *best* separating box concentric with the inner annulus box deterministically in $O(N \log N)$ time and such a box, at most a constant factor worse than optimal, deterministically in $O(N)$ time.

[23] By using a linear time selection algorithm [2] [15].

trivially, but the deterministic method is not so easily generalized.

References

[1] F. d'Amore & P.G.Franciosa: Separating sets of hyperrectangles, Int'l. J. Comput. Geom. Applic. 3,2 (1993) 155-165; and see missing figure: 3,3 (1993) 345.

[2] M.Blum, R.W.Floyd, V.Pratt, R.L.Rivest, R.E.Tarjan: Time bounds for selection, JCSS 7 (1973) 448-461.

[3] V.G.Boltyanskii: Problem of illumination of the boundary of a convex body (in Russian), Izvestiya Moldovskog Filiala Akad. Nauk.SSSR 10, 76 (1960) 79-86.

[4] V.G.Boltjansky & I.Ts.Gohberg: Results and problems in combinatorial geometry, Cambridge Univ. Press 1985.

[5] T.Bonnesen & W.Fenchel: Theory of convex bodies, BCS associates, Moscow Idaho USA 1987.

[6] K.Borsuk: Drei Sätze über die n-dimensionalen euklidische Sphäre, Fundamenta Math. 20 (1933) 177-190.

[7] G.D.Chakerian: Inequalities for the difference body of a convex body, Proc. Amer. Math. Soc. 18 (1967) 879-884.

[8] G.D.Chakerian & G.T.Sallee: An intersection theorem for sets of constant width, Duke Math. J. 36 (1969) 165-170.

[9] Kenneth L. Clarkson: A probabilistic algorithm for the post office problem, STOC 17 (1985) 175-184.

[10] J.H.Conway & N.J.A.Sloane: Sphere packings, lattices, and groups, Springer 1988.

[11] H.S.M. Coxeter: Loxodromic sequences of tangent spheres, Aequationes Math. 1 (1968) 104-121.

[12] L. Fejes Toth: Lagerungen in der Ebene, auf der Kugel, und im Raum, (2nd ed.) Springer 1972 (GMW #65).

[13] L.Few: Multiple packing of spheres, a survey, pages 88-93 in W.Fenchel (ed.): Proc. Colloq. Convexity (Copenhagen 1965) Kobenhavn Univ. Mat. Inst. 1967.

[14] L.Few & P.Kanagasabapathy: The double packing of spheres, J. London Math. Soc. 44 (1969) 141-146.

[15] R.W.Floyd & R.Rivest: Expected time bounds for selection, Commun. ACM 18,3 (March 1975) 165-173.

[16] C.H.Goldberg & D.B.West: Bisection of circle colorings, SIAM J. Algeb. & Discr. Meth. 6 (1985) 93-106.

[17] M.Grötschel, L.Lovasz, A.Schrivjer: Geometric algorithms and combinatorial optimization, Springer 1990.

[18] P.Gruber (ed.): Handbook of convex geometry, Elsevier 1993.

[19] Branko Grunbaum, G.C. Shephard: Tilings and patterns, W.H. Freeman, 1989.

[20] J.Kahn & G.Kalai: A counterexample to Borsuk's conjecture, Bull. AMS 29 (1993) 60-62.

[21] H.Konig & V.D.Milman: On the covering numbers of convex bodies, pp. 82-95 in Springer Lecture Notes in Math. #1267 (1987).

[22] M.Lassak: Covering a plane convex body by 4 homothetical copies with the smallest positive ratio, Geometriae Dedicata 21 (1986) 157-167.

[23] M.Lassak: Covering the boundary of a convex set by tiles, Proc. Amer. Math. Soc. 104 (1988) 269-272.

[24] J. B. M. Melissen: Packing and Coverings with Circles, PhD thesis, 1997 (Universiteit Utrecht). To be published as book by Kluwer.

[25] Gary L. Miller, Shang-Hua Teng, William Thurston, & Stephen A. Vavasis: Separators for sphere-packings and nearest neighbor graphs, J. ACM, 44, 1 (Jan 1997) 1-29.

[26] V.D.Milman: Inegalite de Brunn-Minkowski inverse et applications a la theorie locale des espaces normes [An inverse form of the Brunn-Minkowski inequality with applications to the local theory of normed spaces], C.R.Acad.Sci.Paris I Math. 302,1 (1986) 25-28. MR87f:52018.

[27] D.J.Muder: A new bound on the local density of sphere packings, D&CG 10 (1993) 351-375.

[28] A.Nilli: On Borsuk's problem, Contemp. Math. 178 (1994) 209-210.

[29] Gilles Pisier: The volume of convex bodies and Banach space geometry, Cambridge University Press 1989.

[30] Gilles Pisier: J.Reine & Angew. Math. 393 (1989) 115-131. MR90a:46038.

[31] R.Rado: A theorem on general measure, J.London Math. Soc. 21 (1946) 291-300.

[32] C.A.Rogers: A note on coverings, Mathematika 4,7 (1957) 1-6.

[33] C.A.Rogers & G.C.Shephard: The difference body of a convex body, Archiv der Math. 8 (1957) 220-233.

[34] C.A.Rogers & C.Zong: Covering convex bodies by translates of convex bodies, Mathematika 44 (1997) 215-218.

[35] Oded Schramm: Illuminating sets of constant width, Mathematika 35 (1988) 180-189.

[36] W.D.Smith: Studies in computational geometry motivated by mesh generation, (485 pp.), PhD. thesis, Princeton University applied math dept., 1988. Available from University Microfilms International, order number 9002713.

[37] D.M.Y.Sommerville: An introduction to the geometry of N dimensions, Methuen, London, 1929.

[38] Daniel A. Spielman and Shang-Hua Teng: Disk Packings and Planar Separators, 12th Annual ACM Symposium on Computational Geometry (1996) 349-358.

[39] A.H.Stone & J.W.Tukey: Generalized 'sandwich' theorems, Duke Math. J. 9 (1942) 356-359.