SPHERE PACKING BOUNDS VIA RESCALING

HENRY COHN AND ANDREW SALMON

Abstract. We study the relationship between local and global density for sphere packings, and in particular the convergence of packing densities in large, compact regions to the Euclidean limit. We axiomatize key properties of sphere packing bounds by the concept of a packing bound function, and we study the special case of sandwich functions, which give a framework for inequalities given by the Lovász sandwich theorem. We show that every packing bound function tends to a Euclidean limit on rectifiable sets, generalizing the work of Borodachov, Hardin, and Saff. Linear and semidefinite programming bounds yield packing bound functions, and we develop a Lasserre hierarchy that converges to the optimal sphere packing density.

Contents
1. Introduction 1
2. Packing bound functions 4
2.1. Packing and covering as packing bound functions 6
2.2. Consequences of uniform continuity 7
3. Packing bound functions via graphs 9
4. Asymptotics of packing bound functions on rectifiable sets 16
5. The Euclidean limits of $\vartheta'$, $\vartheta$, and $\vartheta^+$ 21
6. The Lasserre hierarchy for sphere packing 30
6.1. Review of bounds for compact spaces 31
6.2. The Lasserre hierarchy as sandwich functions 34
6.3. The Euclidean limit of the Lasserre hierarchy 35
Acknowledgements 41
References 41

1. Introduction

Packing problems in Euclidean space involve a balance between local and global behavior: the constraint that the bodies being packed cannot overlap is purely local, but dense local configurations do not always extend to dense global packings. For example, the regular dodecahedron is the smallest possible Voronoi cell for a three-dimensional sphere packing [26], but regular dodecahedra do not tile space, and in fact rhombic dodecahedra yield the best overall packing density [25, 27]. This tension between local and global optimality is called geometrical frustration in physics (see, for example, [40]), and it is not well understood mathematically.
Aside from the case of packing convex, centrally symmetric bodies in at most two dimensions [19, Section 25], it is not known how to reduce the general packing problem to considering a bounded number of bodies, and there is little reason to believe such a reduction is always possible, especially in high dimensions. Instead, the global packing density can be obtained as a limit of the packing densities in bounded but increasingly large regions [23], and this limiting process seems to be essential. For packing of spheres and covering by bodies of bounded diameter, the first appearance of this sort of limit appears to be in a paper by Kolmogorov and Tikhomirov [33, Theorem IX], which proves a limit theorem for Jordan-measurable sets (see [32] for an English translation).

Borodachov, Hardin, and Saff [7] analyzed this limit in broad generality. To formulate their result, we need some notation. Given a bounded subset $C$ of $\mathbb{R}^d$, let $\text{pack}(C)$ be the largest possible size of a subset $X$ of $C$ such that all points in $X$ are at distance at least 2 from each other. In other words, $\text{pack}(C)$ is the maximum number of unit spheres that can be centered at points of $C$, if their interiors are not allowed to overlap. Borodachov, Hardin, and Saff work with subsets that may not be full-dimensional, such as the surface of a sphere in $\mathbb{R}^d$. Recall that a Borel subset of $\mathbb{R}^d$ is called $n$-rectifiable if it is the image of a bounded Borel subset of $\mathbb{R}^n$ under a Lipschitz function from $\mathbb{R}^n$ to $\mathbb{R}^d$. Let $\mathcal{H}_n$ denote $n$-dimensional Hausdorff measure on $\mathbb{R}^d$, let $\mathcal{L}_d$ denote $d$-dimensional Lebesgue measure, let $B^n_r(x)$ be the open ball of radius $r$ in $\mathbb{R}^d$ centered at $x$, and let $rC$ be $C$ dilated by a factor of $r$.

The following theorem is a special case of Theorem 2.2 in [7], after rescaling:

**Theorem 1.1** (Borodachov, Hardin, and Saff [7]). Let $1 \leq n \leq d$, and let $C$ be a compact, $n$-rectifiable subset of $\mathbb{R}^d$ with $\mathcal{H}_n(C) > 0$. Then the limit

$$\lim_{r \to \infty} \frac{\text{pack}(rC)}{\mathcal{H}_n(rC)} \mathcal{H}_n(B^n_1)$$

exists and equals the sphere packing density in $\mathbb{R}^n$.

In this paper, we begin by extending Theorem 1.1 to a class of functions we call packing bound functions. The function $\text{pack}$ defined above will be a packing bound function, as will various upper bounds for pack, including linear and semidefinite programming bounds. Our extension thus analyzes how these sphere packing bounds behave when applied to increasingly large regions of space. For comparison, Hardin, Saff, and Vlasiuk [29] analyze conditions on short-range interactions between particles that suffice to obtain fairly general asymptotics. Our approach differs conceptually in studying not just optimization problems over particle configurations, but also related quantities such as bounds.

To define packing bound functions, we use the following notation. Let $d(x, y) = \|x - y\|$ be the standard $\ell^2$ metric on $\mathbb{R}^d$ (the two uses of $d$ are not ambiguous in practice), and let $\overline{C}$ denote the closure of $C$. For two subsets $C$ and $C'$ of $\mathbb{R}^d$, the distance $d(C, C')$ is the infimum over all distances $d(x, x')$ such that $x \in C$ and $x' \in C'$ (with $d(C, C') = \infty$ if $C$ or $C'$ is empty), and $C(\varepsilon) = \{x \in \mathbb{R}^d : d(x, C) < \varepsilon\}$ is the $\varepsilon$-neighborhood of $C$. We say a function $\psi : C \to C'$ is distance-increasing if $d(\psi(x), \psi(y)) \geq d(x, y)$ for all $x, y \in C$. Let $\mathcal{B}_d$ be the set of all bounded Borel subsets of $\mathbb{R}^d$, and let $\mathcal{B} = \bigcup_d \mathcal{B}_d$.

**Definition 1.2.** A packing bound function is a map $A$ from $\mathcal{B}$ to $[0, \infty)$ such that the following axioms hold for all elements $C$ and $C'$ of $\mathcal{B}$ and $\varepsilon > 0$: 


(1) (Sphere bound) If $C$ is a nonempty set contained in the interior of a ball of radius 1, then $A(C) = 1$.

(2) (Lipschitz inequality) If there exists a distance-increasing function $\psi: C \to C'$, then $A(C) \leq A(C')$.

(3) (Union axiom) If $C$ and $C'$ are subsets of the same ambient space $\mathbb{R}^d$ and $d(C, C') \geq 2$, then $A(C \cup C') = A(C) + A(C')$.

(4) (Mesh axiom) If $C \subseteq C'(\varepsilon)$, then $A(1 + \varepsilon C) \leq A(C')$.

One example is the function pack. For another, let $\text{cov}(C)$ be the size of the smallest covering of $C$ by open balls of radius 1, with sphere centers not necessarily in $C$. In other words, for a subset $C$ of $\mathbb{R}^d$, $\text{cov}(C)$ is the infimum of $|X|$ over subsets $X \subseteq \mathbb{R}^d$ such that $C \subseteq X(1)$. We will show that both pack and cov are packing bound functions, and that they are the extreme packing bound functions. More precisely, under the partial ordering $A_1 \leq A_2$ defined by $A_1(C) \leq A_2(C)$ for all $C$, every packing bound function $A$ satisfies $\text{pack} \leq A \leq \text{cov}$. The inequality $\text{pack} \leq A$ explains the name “packing bound function.”

Between pack and cov, we construct a range of packing bound functions, most notably the linear programming bound. Each of these functions can be viewed in two ways, as an upper bound for packings or as a lower bound for coverings, and each packing bound function has consistent large-scale limiting behavior:

**Theorem 1.3.** Let $1 \leq n \leq d$, let $C$ be a compact, $n$-rectifiable subset of $\mathbb{R}^d$ with $\mathcal{H}_n(C) > 0$, and let $A$ be any packing bound function. Then the limit

$$
\lim_{r \to \infty} \frac{A(rC)}{\mathcal{H}_n(rC)}
$$

exists and depends only on $A$ and $n$, rather than the choice of $C$ or $d$.

In fact, we will prove a somewhat more general result in Theorem 4.12, with the same hypotheses as Theorem 2.2 in [7]. For the special case of the function cov, we learned after completing this work that convergence was independently proved at the same time by Anderson, Reznikov, Vlasiuk, and White [3].

Because pack is the smallest packing bound function, the limit

$$
\lim_{r \to \infty} \frac{A(rC)}{\mathcal{H}_n(rC)} \mathcal{H}_n(B^n_1)
$$

is always an upper bound for the sphere packing density in $\mathbb{R}^n$. One example is the linear programming bound of Cohn and Elkies [13], which we obtain as a limit of a Delsarte linear program for bounded regions. Our formulation of this limit appears to be new, although some of the techniques are related to work of Cohn, de Courcy-Ireland, and Zhao [16, 12].

Semidefinite programming hierarchies are an important generalization of linear programming bounds, which give upper bounds for the size of the largest packing in any compact set by the work of de Laat and Vallentin [35]. Many of the best bounds known for spherical codes [5] and sphere packing [13, 11] come from linear and semidefinite programming relaxations along these lines. To formulate these semidefinite programming hierarchies as packing bound functions, we will define a discrete Lasserre hierarchy. Unlike the topological Lasserre hierarchy from [35], the

---

1This duality is less interesting than it might sound, because any number less than 1 is a trivial lower bound for covering, and any number greater than 1 is a trivial upper bound for packing.
discrete hierarchy uses only the graph structure of distances in the space $C \subseteq \mathbb{R}^d$ and not its topology. This construction yields a packing bound function, and thus abstractly, there must exist a Euclidean limit. We show that this Euclidean limit agrees with the limit of the topological approach from [35], and we formulate it as an optimization problem on Euclidean space. We also prove that as $t \to \infty$, the $t$-th level of the hierarchy approaches the optimal density of sphere packing in Euclidean space.

Although 3-point bounds for spherical codes have been known since the work of Bachoc and Vallentin [5], obtaining 3-point bounds or other semidefinite programming bounds that refine the linear programming bound in $\mathbb{R}^d$ has been a longstanding open problem. Formulating these refinements was one of the main motivations of the present work and will be used in subsequent work with David de Laat [15] to give new upper bounds on sphere packing in all dimensions up through 12 in which the exact answer is not known (i.e., not 1, 2, 3, or 8 dimensions).

2. Packing bound functions

In this section, $A$ will denote an arbitrary packing bound function, and $C$ and $C'$ will be elements of $\mathcal{B}$. We begin by examining some of the consequences of Definition 1.2. The sphere bound, Lipschitz, and union axioms are quite natural, while the mesh axiom is a little more subtle. Intuitively, it says that a packing can be modified slightly so that the sphere centers are forced to live on a mesh if that mesh is sufficiently fine. Later, we will give some consequences of this axiom to packings when $C'$ is a dense subset as well as packings on a countable nested union. These applications use the following special case of the mesh axiom, which we refer to as the continuity property:

**Proposition 2.1** (Uniform continuity). For every $\varepsilon > 0$, there exists some $\delta > 0$ such that for all $C$ and $C'$, if $C \subseteq C'(\delta)$, then $A(C) \leq A((1 + \varepsilon)C')$.

Specifically, it follows from the mesh axiom that we can take $\delta = \varepsilon/(1 + \varepsilon)$, but the specific choice of $\delta$ is not important in many applications.

Another important special case of the defining properties is monotonicity, which we obtain by applying the Lipschitz inequality to the identity function:

**Proposition 2.2** (Monotonicity). If $C \subseteq C'$, then $A(C) \leq A(C')$.

It follows that every packing bound function is an upper bound for pack:

**Corollary 2.3.** For every $C$, $\text{pack}(C) \leq A(C)$.

**Proof.** If $X$ is a subset of $C$ such that all points in $X$ are at distance at least two from each other, then $A(X) = |X|$ by the union and sphere bound properties, and $A(X) \leq A(C)$ by monotonicity. \qed

Invariance of packing bound functions under isometry also follows immediately from the Lipschitz inequality:

**Proposition 2.4** (Isometry invariance). If $C$ and $C'$ are isometric, then $A(C) = A(C')$.

We also observe a useful and basic inequality, which we call the union bound:

**Proposition 2.5** (Union bound). If $C$ and $C'$ are subsets of the same ambient space, then $A(C \cup C') \leq A(C) + A(C')$. 
Remark 2.8. Euclidean bound

We say that a packing bound function satisfies the Euclidean bound when

\[ \lim_{x \to 0} \frac{\mathcal{L}_n(C(\varepsilon))}{\mathcal{L}_n(B^n)} = 1/2. \]

Because neighborhoods \( C(\varepsilon) \) of bounded Borel sets \( C \) are Jordan-measurable,\(^2\) it further suffices to prove the Euclidean bound for Jordan-measurable sets. Specifically,

\[ A(C) \geq A\left( \frac{C(\varepsilon)}{1 + \varepsilon} \right) \]

by the mesh axiom, and

\[ \lim_{\varepsilon \to 0} \frac{\mathcal{L}_n(C(\varepsilon))}{\mathcal{L}_n(C)} = \mathcal{L}_n(C) \]

when \( C \) is compact, by Theorem 3.2.39 in [18]. Moreover, one can show that the Euclidean bound for \( C \) is automatically satisfied if \( C \) is capable of tiling Euclidean space with overlap of measure 0.

Proof. First, by monotonicity we reduce to the case when \( C \) and \( C' \) are disjoint. Now consider the union of copies of \( C \) and \( C' \) placed very far apart (further than 2 plus the diameter of \( C \cup C' \)), which we call \( D \). Then the map from \( C \cup C' \) to \( D \) sending \( C \) to the copy of \( C \) in \( D \) and \( C' \) to the copy of \( C' \) in \( D \) is distance-increasing, and therefore \( A(C \cup C') \leq A(D) = A(C) + A(C') \) by the Lipschitz and union properties.

Proposition 2.6 (Density). If \( \varepsilon > 0 \) and \( C' \subseteq C \) is dense, then

\[ A(C) \leq A((1 + \varepsilon)C'). \]

Proof. Every dense subset \( C' \) satisfies \( C \subseteq C'(\delta) \) for each \( \delta > 0 \). By continuity, we can choose \( \delta > 0 \) so that \( C \subseteq C'(\delta) \) implies \( A(C) \leq A((1 + \varepsilon)C') \).

Consider the \( n \)-cube \( I^n = [0,1]^n \). By the union bound, \( A(2I^n) \leq 2^n A(I^n) \), and so \( A(2^k I^n)/2^{kn} \) is a weakly decreasing sequence in \( k \). On the other hand, it is bounded below by \( 1/2^n \) because \( A(2^k I^n) \geq \text{pack}(2^k I^n) \geq (1 + 2^{k-1})^n \) when \( k \geq 1 \) from using the subset \( X = 2^k I^n \cap 2Z^n \). Therefore, \( A(2^k I^n)/2^{kn} \) must converge to some positive number.

Definition 2.7. For a packing bound function \( A \), let \( \delta_{A,n} = \lim_{k \to \infty} A(2^k I^n)/2^{kn} \). We say that a packing bound function satisfies the Euclidean bound if every bounded Borel subset \( C \subseteq \mathbb{R}^n \) satisfies \( A(C) \geq \delta_{A,n} \mathcal{L}_n(C) \).

Because \( A \geq \text{pack} \), the quantity \( \delta_{A,n} \mathcal{L}_n(B^n) \) is always an upper bound for the sphere packing density \( \delta_{\text{pack},n} \mathcal{L}_n(B^n) \) in \( \mathbb{R}^n \). Similarly, the inequality \( A \leq \text{cov} \) from Proposition 2.10 will imply that \( \delta_{A,n} \mathcal{L}_n(B^n) \) is a lower bound for the sphere covering density \( \delta_{\text{cov},n} \mathcal{L}_n(B^n) \) in \( \mathbb{R}^n \).

Remark 2.8. By Proposition 2.6, it suffices to prove the Euclidean bound when \( C \) is compact. Specifically, if the Euclidean bound holds for \( \overline{C} \), then for every \( \varepsilon > 0 \),

\[ A(C) \geq A\left( \frac{\overline{C}}{1 + \varepsilon} \right) \geq \delta_{\overline{C},n} \mathcal{L}_n\left( \frac{\overline{C}}{1 + \varepsilon} \right) \geq \delta_{A,n} \frac{\mathcal{L}_n(C)}{(1 + \varepsilon)^n}. \]

Because neighborhoods \( \overline{C}(\varepsilon) \) of bounded Borel sets \( C \) are Jordan-measurable,\(^2\) it further suffices to prove the Euclidean bound for Jordan-measurable sets. Specifically,

\[ A(C) \geq A\left( \frac{\overline{C}(\varepsilon)}{1 + \varepsilon} \right) \]

by the mesh axiom, and

\[ \lim_{\varepsilon \to 0} \frac{\mathcal{L}_n(C(\varepsilon))}{\mathcal{L}_n(C)} = \mathcal{L}_n(C) \]

when \( C \) is compact, by Theorem 3.2.39 in [18]. Moreover, one can show that the Euclidean bound for \( C \) is automatically satisfied if \( C \) is capable of tiling Euclidean space with overlap of measure 0.

\(^2\)To see why, notice that this amounts to the claim that the boundary of \( C(\varepsilon) \) has Lebesgue measure zero. The function \( x \mapsto d(x,C) \) is a Lipschitz function, and so it is differentiable almost everywhere by Rademacher’s theorem. Thus, it suffices to show that the set of points \( x \) where \( x \mapsto d(x,C) \) is differentiable and \( d(x,C) = \varepsilon \) has measure zero. For such points, \( \lim_{\varepsilon \to 0} \frac{\mathcal{L}_n(C(\varepsilon) \cap B^n(x))}{\mathcal{L}_n(B^n(x))} = 1/2 \), and the result follows by the Lebesgue density theorem.
2.1. Packing and covering as packing bound functions. We say a packing in $C$ is a subset $X$ of $C$ such that all points in $X$ are at distance at least 2 from each other. Recall that $\text{pack}(C)$ is the size of the largest packing in $C$.

**Proposition 2.9.** The function $\text{pack}$ is a packing bound function that satisfies the Euclidean bound.

**Proof.** Among the axioms, the sphere bound, Lipschitz inequality, and union axiom are immediate. For the mesh axiom, suppose $C \subseteq C'(\varepsilon)$, and let $X$ be a packing in $\frac{1}{1+\varepsilon}C$. Then $(1+\varepsilon)X$ is a subset of $C$ with all distances at least $2+2\varepsilon$. Each $x \in X$ is within $\varepsilon$ of some point $y(x)$ in $C'$ because $C \subseteq C'(\varepsilon)$, and so $Y = \{y(x) : x \in X\}$ is a packing in $C'$. It now follows that $A(C') \geq |Y|$ by Corollary 2.3, and thus $A(C') \geq A(\frac{1}{1+\varepsilon}C)$, as desired.

All that remains is to prove the Euclidean bound, for which we use an averaging argument. Let $C \subseteq \mathbb{R}^n$ be a Borel set with $C \subseteq [-r, r]^n$, and choose a packing $X_k$ in $2^kI^n$ for each $k$ such that

$$\lim_{k \to \infty} \frac{|X_k|}{2^kn} = \delta_{\text{pack},n}.$$

We will obtain a lower bound for $\text{pack}(C)$ by averaging over intersections of translates of $C$ with $X_k$. To do so, consider the cube $R = [-r, 2^k+r]^n$ containing $2^kI^n$. Because $C \subseteq [-r, r]^n$,

$$\int_R \#(X_k \cap (C+t)) \, dt = \int_{\mathbb{R}^n} \#(X_k \cap (C+t)) \, dt = |X_k|L_d(C).$$

It follows that for some $t \in R$,

$$\#(X_k \cap (C+t)) \geq \frac{|X_k|L_d(C)}{L_d(R)} = \frac{|X_k|L_d(C)}{(2r + 2^k)^n},$$

and therefore

$$\text{pack}(C) \geq \#(X_k \cap (C + t)) \geq \frac{|X_k|}{(2r + 2^k)^n} L_n(C).$$

Taking the limit as $k \to \infty$ completes the proof. \hfill \qed

Recall that $\text{cov}(C)$ is the smallest covering of $C$ by open balls of radius 1, with sphere centers not necessarily in $C$. That is, for $C$ in the ambient space $\mathbb{R}^d$, $\text{cov}(C)$ is the infimum of $|X|$ over subsets $X \subseteq \mathbb{R}^d$ such that $C \subseteq X(1)$. This infimum is independent of the ambient space chosen: if $C$ is contained in a proper subspace $\mathbb{R}^{d'}$ of $\mathbb{R}^d$, then the points in $X$ can be orthogonally projected to $\mathbb{R}^{d'}$. We call such sets $X$ coverings of $C$ with open balls of radius 1.

**Proposition 2.10.** The function $\text{cov}$ is a packing bound function. It is the maximal packing bound function among all packing bound functions, and it does not satisfy the Euclidean bound.

**Proof.** First, we check the sphere bound, union axiom, and mesh axiom. The function $\text{cov}$ trivially satisfies the sphere bound. For the union axiom, if $C$ and $C'$ are separated by distance at least 2, then there is no open ball of radius 1 that intersects both $C$ and $C'$, and so covering $C \cup C'$ is the same as covering $C$ and $C'$ separately. For the mesh axiom, suppose $C \subseteq C'(\varepsilon)$, and let $X$ be a covering of $C'$ with open balls of radius 1. Then $(1+\varepsilon)$-balls centered at $X$ cover $C'(\varepsilon)$, and so
these \((1 + \varepsilon)\)-balls cover \(C\). Therefore \(\frac{1}{1 + \varepsilon} X\) gives a covering of \(\frac{1}{1 + \varepsilon} C\) by balls of radius 1.

The Lipschitz inequality requires a little more argument. Let \(\psi : C \to C'\) be a distance-increasing map, with \(C' \subseteq \mathbb{R}^d\) and \(C \subseteq \mathbb{R}^d\), and suppose \(X'\) is a covering of \(C'\) with open balls of radius 1. We would like to obtain a covering of \(C\) of the same size. To do so, note that the function \(\psi\) is injective, and its inverse \(\psi^{-1}\) on \(\psi(C)\) is a Lipschitz function with Lipschitz constant 1. By the Kirszbraun theorem, we can extend \(\psi^{-1}\) to a Lipschitz function \(\phi : \mathbb{R}^d' \to \mathbb{R}^d\), again with Lipschitz constant 1.

If \(x\) is a point in \(C\), then \(d(\psi(x), x') < 1\) for some \(x' \in X'\), from which it follows that \(d(x, \phi(x')) < 1\). Thus, \(\{\phi(x') : x' \in X'\}\) is a covering of \(C\), as desired.

To see that \(\text{cov}\) is the largest packing bound function, suppose \(A\) is some other packing bound function, and write \(C \subseteq \bigcup_{i=1}^{N} B_i\) where the sets \(B_i\) are open balls of radius 1. Then \(A(C) \leq \sum_{i=1}^{N} A(B_i) = N\) by monotonicity and the union bound.

The Euclidean bound fails for \(d > 1\), because \(L_d(B_1^d)\delta_{\text{cov}, d} > 1\), while \(\text{cov}(B_1^d) = 1\). Note that the inequality \(L_d(B_1^d)\delta_{\text{cov}, d} > 1\) simply says that the sphere covering density in \(\mathbb{R}^d\) is strictly greater than 1, which holds because spheres cannot tile space when \(d > 1\). (If the covering density were 1, then a compactness argument would show that closed unit balls cover space with only measure-zero overlap. However, a point in space can be on the boundary of at most two such balls, which does not yield a covering locally.)

2.2. Consequences of uniform continuity. The idea of Proposition 2.1 can be restated as follows. If \(C\) is an arbitrary set, we can choose any sufficiently fine mesh \(C'\) and stipulate that our sphere centers must live on \(C'\) as long as we dilate the mesh by a small amount. In the remainder of this section, we give some consequences of this continuity property.

**Proposition 2.11** (Nestedunion). Let \(C_1 \subseteq C_2 \subseteq \ldots\) be Borel sets such that \(\bigcup_i C_i\) is bounded. Then for each \(\varepsilon > 0\), there exists a \(j\) such that

\[
A\left(\bigcup_i C_i\right) \leq A((1 + \varepsilon)C_j).
\]

**Proof.** Choose \(\delta\) as in Proposition 2.1. Then

\[
\bigcup_i C_i \subseteq \left(\bigcup_i C_i\right)(\delta) \subseteq \bigcup_j C_j(\delta).
\]

Because \(\bigcup_i C_i\) is compact, it must be contained in \(C_j(\delta)\) for some \(j\), which completes the proof. \(\square\)

We will use the normalized Hausdorff measure, by which we mean

\[
\mathcal{H}_n(C) = \lim_{\varepsilon \to 0^+} \mathcal{H}_{n, \varepsilon}(C),
\]

where

\[
\mathcal{H}_{n, \varepsilon}(C) = \frac{L_n(B_1^d)}{2^n} \inf \left\{ \sum_{i \in I} \text{diam}(C_i)^n : C \subseteq \bigcup_{i \in I} C_i \text{ with diam}(C_i) < \varepsilon \right\}.
\]
Recall that the \emph{n-dimensional Minkowski content} $\mathcal{M}_n(C)$ of a set $C \subseteq \mathbb{R}^d$ with $d \geq n$ is defined by

$$\mathcal{M}_n(C) = \lim_{\varepsilon \to 0^+} \frac{\mathcal{L}_d(C(\varepsilon))}{\mathcal{L}_{d-n}(B_\varepsilon^{d-n})},$$

when this limit exists (we take $\mathcal{L}_0(B_0^d) = 1$ when $n = d$). The \emph{upper} or \emph{lower} Minkowski content, denoted $\overline{\mathcal{M}}_n$ or $\underline{\mathcal{M}}_n$, is given by the lim sup or lim inf, respectively. For a compact $n$-rectifiable set $C$, \cite[Theorem 3.2.39]{HN} tells us that $\mathcal{H}_n(C) = \mathcal{M}_n(C)$.

**Proposition 2.12.** Let $C \subseteq \mathbb{R}^d$ be a bounded Borel set, and suppose $\overline{\mathcal{M}}_n(C) < \infty$. Then

$$\limsup_{r \to \infty} \frac{A(rC)}{r^n} < \infty.$$ 

**Proof.** Because $\overline{\mathcal{M}}_n(C) < \infty$, there is some constant $M$ such that for $\delta$ sufficiently small,

$$\mathcal{L}_d(C(\delta)) \leq M\delta^{d-n}.$$ 

Consider the collection $\mathcal{F}$ of balls of radius $\delta$ at each point of $C$. By the Vitali covering lemma \cite[Theorem 2.1]{HN}, there is a subset $\mathcal{F}'$ consisting of disjoint balls such that $\bigcup_{B \in \mathcal{F}'} 5B$ contains $\bigcup_{B \in \mathcal{F}} B = C(\delta)$. Because the balls in $\mathcal{F}'$ are disjoint,

$$|\mathcal{F}'| \leq \frac{\mathcal{L}_d(C(\delta))}{\mathcal{L}_d(B_1^d)} \leq \frac{M\delta^{d-n}}{\mathcal{L}_d(B_1^d)} \delta^n \leq M' \delta^n,$$

where $M' = M/\mathcal{L}_d(B_1^d)$. Now we let $r = 1/(6\delta)$. Because $\bigcup_{B \in \mathcal{F}'} 5B$ covers $rC$ and the spheres have radius $r\delta < 1$, monotonicity and the union and sphere bounds imply that

$$A(rC) \leq \sum_{B \in \mathcal{F}'} A(5rB) \leq M'/\delta^n.$$ 

In particular,

$$A(rC)/r^n \leq M'/r^\delta = M'6^n,$$

and we conclude that $\limsup_{r \to \infty} A(rC)/r^n < \infty$ because this bound holds whenever $r = 1/(6\delta)$ with $\delta$ sufficiently small. \hfill $\square$

We isolate a geometric lemma in preparation for Proposition 2.14.

**Lemma 2.13.** If $C'$ is any compact subset of $C \subseteq \mathbb{R}^d$ with $\mathcal{M}_n(C)$ and $\mathcal{M}_n(C')$ defined and $\infty > \mathcal{M}_n(C') > \mathcal{M}_n(C) - \varepsilon$, then there is a $\delta^*$ such that for all $\delta \leq \delta^*$, there is a packing $\mathcal{B}$ of disjoint balls of radius $\delta/5$ in $C(\delta) \setminus C'(\delta)$ containing at most $M\varepsilon/\delta^n$ balls for some universal constant $M$ depending on the dimensions $n$ and $d$, such that $\bigcup_{B \in \mathcal{B}} 5B$ is a covering of $C \setminus C'(\delta)$.

**Proof.** Since the Minkowski contents of $C$ and $C'$ are within $\varepsilon$, for sufficiently small $\delta^* \geq \delta > 0$,

$$\mathcal{L}_d(C(\delta) \setminus C'(\delta)) \leq \varepsilon\delta^{d-n} \cdot 2\mathcal{L}_{d-n}(B_1^{d-n}).$$

For every $\delta$, consider the collection of balls $\mathcal{F}$ of radius exactly $\delta/5$ at each point of $C \setminus C'(\delta)$. By the Vitali covering lemma, there are disjoint balls $\mathcal{B}$ of radius $\delta/5$ such that $5B$ contains the union of the balls in $\mathcal{F}$. These balls are a sphere packing and are contained in $C(\delta) \setminus C'(\delta)$, and therefore

$$\mathcal{L}_d(B_{\delta/5}^d)|\mathcal{B}| \leq \mathcal{L}_d(C(\delta) \setminus C'(\delta)) \leq \varepsilon\delta^{d-n} \cdot 2\mathcal{L}_{d-n}(B_1^{d-n}).$$

It follows that $|\mathcal{B}| \leq M\varepsilon/\delta^n$, where $M$ depends only on $n$ and $d$. \hfill $\square$
In the process of working with an arbitrary rectifiable set, we will need to pass from a set to a subset of the same Minkowski content that cuts out a bad subset. We need to show that this removal does not affect the first order asymptotics of packing. This result will also be used in working with compact sets without appealing to the Euclidean bound. The following proposition is an analogue of the “regularity lemma” in the Riesz energy setting [8, Lemma 8.6.9].

**Proposition 2.14.** For each \( n \), bounded Borel set \( C \subseteq \mathbb{R}^d \) with \( \mathcal{M}_n(C) \) defined, and \( k \geq 1 \), there is an \( \varepsilon^* > 0 \) such that for any \( \varepsilon \leq \varepsilon^* \), if \( C' \) is any compact subset of \( C \) with \( \mathcal{M}_n(C') \) defined and \( \infty > \mathcal{M}_n(C') > \mathcal{M}_n(C) - \varepsilon \), then

\[
\left( \frac{k}{k+1} \right)^n \limsup_{r \to \infty} \frac{A(rC)}{r^n} \leq \limsup_{r \to \infty} \frac{A(rC')}{r^n} + M\varepsilon k^n
\]

for a universal constant \( M \) depending on \( n \) and \( d \). The same statement also holds with \( \limsup \) replaced by \( \liminf \).

**Proof.** Choose \( \delta^* \) and \( M \) as in Lemma 2.13, and let \( \delta = 1/(kr) \leq \delta^* \) for sufficiently large \( r \). By Lemma 2.13, there is a collection \( 5\mathcal{B} \) of at most \( M\varepsilon/\delta^n \) open balls of radius \( \delta \) that cover \( C \setminus C'(\delta) \).

Because \( C \subseteq (C' \cup (C \setminus C'(\delta)))\)(\( \delta \)), scaling by a factor of \( r \) shows that

\[
A\left( \frac{1}{1 + r\delta} rC \right) \leq A(rC') + A(rC \setminus rC'(r\delta)).
\]

Now since \( 1/(1 + r\delta) = k/(k+1) \),

\[
\limsup_{r \to \infty} \frac{A\left( \frac{k}{k+1} rC \right)}{r^n} \leq \limsup_{r \to \infty} \frac{A(rC')}{r^n} + \limsup_{r \to \infty} \frac{A(rC \setminus rC'(r\delta))}{r^n}.
\]

Finally, because \( rC \setminus rC'(r\delta) \) is covered with at most \( M\varepsilon/\delta^n \) open balls of radius \( \delta r = 1/k \leq 1 \), the last term on the right side can be bounded by \( M\varepsilon k^n \), while the left side is

\[
\left( \frac{k}{k+1} \right)^n \limsup_{r \to \infty} \frac{A(rC)}{r^n},
\]

which is what we wanted to show. For the \( \liminf \) version, we instead use the inequality

\[
\liminf_{r \to \infty} \frac{A\left( \frac{k}{k+1} rC \right)}{r^n} \leq \liminf_{r \to \infty} \frac{A(rC')}{r^n} + \limsup_{r \to \infty} \frac{A(rC \setminus rC'(r\delta))}{r^n}
\]

and conclude in the same way. \( \square \)

### 3. Packing bound functions via graphs

Consider the category \( \mathcal{G} \) whose objects are (loopless, undirected) graphs \( G = (V,E) \) with finite chromatic number, and whose morphisms are graph homomorphisms. There is a natural operation of join, where the join of two graphs \( G \) and \( H \), denoted \( G * H \), is the graph with vertex set the disjoint union of the vertex sets \( V(G) \) and \( V(H) \) and edge set all edges in \( G \) and \( H \) together with all pairs \( (x,y) \) with \( x \in V(G) \) and \( y \in V(H) \). For simplicity, we denote the edge containing vertices \( x \) and \( y \) by \( xy \), and the empty graph by \( \emptyset \). The complement \( \overline{G} \) of a graph \( G \) has the same vertex set and the complementary edge set.
By passing to complements, we can consider a related category which shows the connection to packing problems in discrete geometry more clearly. This category $\mathbb{T}$ has as its objects graphs with finite clique covering number, meaning there is a finite collection of cliques $C_1, \ldots, C_n$ such that $V \subseteq \bigcup_{i=1}^{n} C_i$. Morphisms from $G$ to $G'$ are maps $f: V \to V'$ with the following two properties:

1. If $(f(x), f(y)) \in E'$, then $(x, y) \in E$. That is, the preimage of any edge is an edge.
2. The preimage of each $x' \in V'$ is a clique, possibly empty.

There is a natural operation of disjoint union graphs, which we denote $\sqcup$. This is not a coproduct on the category. There is an isomorphism of categories $\mathbb{G} \to \mathbb{T}$ sending a graph to its complement, which sends disjoint unions to joins and vice versa.

In geometric terms, we think of graphs in $\mathbb{T}$ as follows: the vertices are points in a bounded Borel subset of $\mathbb{R}^d$, and the edges indicate which points are at distance less than 2 from each other. Equivalently, edges in $G$ indicate which pairs of points are at distance at least 2, i.e., far enough away that they do not interact. We will use this category to construct sandwich functions (named after the Lovász sandwich theorem [31]), which describe the packing bound functions that depend only on the underlying graph structure, rather than the additional information supplied by the metric.

A sandwich function is a function $\Psi$ from objects of $\mathbb{G}$ to nonnegative real numbers with the following properties:

1. (Sphere bound) $\Psi(\emptyset) = 0$ and $\Psi(G) = 1$ if $G$ is a single vertex.
2. (Functoriality) If $G \to G'$ is a morphism, then $\Psi(G) \leq \Psi(G')$. That is, $\Psi$ is a functor from $\mathbb{G}$ to the poset of nonnegative real numbers.
3. (Additivity of join) For all $G$ and $G'$, $\Psi(G * G') = \Psi(G) + \Psi(G')$.

**Theorem 3.1.** Every sandwich function $\Psi$ gives a packing bound function $A$ on bounded subsets of Euclidean spaces as follows. Given a bounded Borel subset $C$ of $\mathbb{R}^n$, let $G(C)$ be the graph whose vertex set is $C$ and whose edge set is pairs $x, y \in C$ such that $|x - y| \geq 2$. Then $A(C) = \Phi(G(C))$.

**Proof.** To prove that $A$ is a packing bound function, we check the axioms:

1. (Sphere bound) If $C$ is empty, then so is $G(C)$, and therefore $A(C) = \Phi(G(C)) = 0$. If $C$ is contained in an open ball of radius 1, then $G(C)$ has no edges and thus admits a morphism in $\mathbb{G}$ to a single point, so $A(C) = \Phi(G(C)) \leq 1$. On the other hand, a single point admits a morphism to any nonempty graph by inclusion, and so $\Phi(G(C)) \geq 1$ if $C \neq \emptyset$.
2. (Lipschitz) If $f: C \to C'$ is a map satisfying $d(x, y) \leq d(f(x), f(y))$, then it induces a morphism $f: G(C) \to G(C')$ on graphs.
3. (Union) If $C$ and $C'$ are separated by distance at least 2, then $G(C \cup C') = G(C) * G(C')$, and $A(C \cup C') = \Phi(G(C) * G(C')) = \Phi(G(C)) + \Phi(G(C')) = A(C) + A(C')$.
4. (Mesh) If $C \subseteq C'(\varepsilon)$, then there exists a map $f: C \to C'$ sending each point in $C$ to a point in $C'$ such that $|x - f(x)| < \varepsilon$. We claim the composite map $\tilde{f}: \frac{1}{1+\varepsilon}C \to C \to C'$ defined by $f(x) = f((1 + \varepsilon)x)$ yields a $\mathbb{G}$ morphism $\tilde{f}: G(\frac{1}{1+\varepsilon}C) \to G(C')$. To see why, note that if $x, y \in \frac{1}{1+\varepsilon}C$ with $|x - y| \geq 2$,
then \(|(1 + \varepsilon)x - (1 + \varepsilon)y| \geq 2(1 + \varepsilon)\) and thus \(|\tilde{f}(x) - \tilde{f}(y)| = |f((1 + \varepsilon)x) - f((1 + \varepsilon)y)|
\geq |(1 + \varepsilon)x - (1 + \varepsilon)y|
- |f((1 + \varepsilon)x) - (1 + \varepsilon)x| - |f((1 + \varepsilon)y) - (1 + \varepsilon)y|
\geq 2(1 + \varepsilon) - \varepsilon - \varepsilon = 2\). □

**Remark 3.2.** Many packing bound functions, including packing and the discrete Lasserre and \(k\)-point bound hierarchies, come from sandwich functions on graphs using this theorem. However, sphere covering does not, as we will see below.

**Proposition 3.3.** The clique number \(\omega\) and chromatic number \(\chi\) are sandwich functions, and every sandwich function \(\Phi\) satisfies the following Lovász sandwich theorem:

\[
\omega(G) \leq \Phi(G) \leq \chi(G).
\]

**Proof.** To check that \(\omega\) is a sandwich function, note that the image of a clique under a graph homomorphism remains a clique, while the union of two cliques in \(G_1\) and \(G_2\) becomes a clique in their join. For chromatic number, the preimage of an independent set under a graph homomorphism is independent, so an \(n\)-coloring pulls back to an \(n\)-coloring under a graph homomorphism. Moreover, the only way to color the join of two graphs is to color each part separately.

Now suppose \(\Phi\) is a sandwich function. Any clique is the iterated join of one-point graphs, and its inclusion into the whole graph is a graph homomorphism, thus showing that if \(K_n \to G\) is the inclusion of a clique then \(n = \Phi(K_n) \leq \Phi(G)\). By considering all such cliques, this shows \(\omega(G) \leq \Phi(G)\).

On the other hand, \(n\)-colorings of \(G\) correspond to graph homomorphisms \(G \to K_n\), with the color sets being the preimages of the vertices of \(K_n\). Thus, if \(G\) has an \(n\)-coloring, then \(\Phi(G) \leq \Phi(K_n) = n\). By considering all possible colorings, this shows \(\Phi(G) \leq \chi(G)\).

**Corollary 3.4.** The packing bound function \(\text{cov}\) does not come from a sandwich function.

**Proof.** Let \(C\) be the vertices of an equilateral triangle of side length strictly between \(\sqrt{3}\) and 2. This set cannot be covered by a single sphere, but its underlying graph \(G(C)\) is independent. Thus, \(\chi(G(C))\) is strictly smaller than the covering number, which would be impossible if covering came from a sandwich function.

Given a sandwich function \(\Phi\), we often identify \(\Phi\) with the corresponding packing bound function notationally. For example, we write \(\delta_{\Phi,n}\) in Definition 2.7.

**Remark 3.5.** The quantity \(\delta_{\chi,n}/\mathcal{L}_n(B_1^\varepsilon)\) appeared in [33, Theorem IX] as a limit of their notion of \(\varepsilon\)-covering of Jordan-measurable sets by bodies of diameter at most \(2\varepsilon\) (see [33, Definition 1]) as \(\varepsilon \to 0^+\).

It would be interesting to compare \(\delta_{\chi,n}\), the “clique cover” or “diameter cover” density of Euclidean space \(\mathbb{R}^n\), with the density of the best sphere covering \(\delta_{\text{cov},n}\). Do these quantities coincide? This question was posed by Lenz and Heppes and is open even in the plane [9, Section 1.3, Conjecture 4].

The original Lovász sandwich theorem was proved for a graph invariant \(\vartheta\) and used to give an upper bound for Shannon capacity. We will see that both of these quantities, as well as some generalizations, are sandwich functions.
Shannon capacity is defined in terms of the strong graph product. Recall that the strong graph product \( G \cdot H \) is the graph whose vertex set is \( V(G) \times V(H) \) and whose edge set consists of distinct pairs \((u_1, v_1), (u_2, v_2) \in V(G) \times V(H)\) such that \( u_1 u_2 \in E(G) \) or \( u_1 = u_2 \) and likewise \( v_1 v_2 \in E(G) \) or \( v_1 = v_2 \). Let \( G^n \) denote the \( n \)-fold strong product of \( G \) with itself. Shannon capacity, denoted \( \Theta(G) \), is \( \lim_{n \to \infty} \alpha(G^n)^{1/n} \), which exists for graphs whose complements have finite chromatic number; here \( \alpha(G) \) denotes the independence number of \( G \). Shannon capacity is not a sandwich function, because it is not additive under join, as shown by Alon [1].

We now define the Lovász theta number \( \vartheta \) and its variants \( \vartheta' \) and \( \vartheta^+ \), first for finite graphs [38, Sections 4.2 and 4.4], and then in general. For finite graphs \( G \), consider positive semidefinite \( V(G) \times V(G) \) matrices \( M \) (i.e., \( |V(G)| \times |V(G)| \) matrices indexed by \( V(G) \)) such that \( \text{Tr}(M) = 1 \).

**Definition 3.6.** For \( G \) finite, we define

1. \( \vartheta^+(G) \) to be the maximum of \( \sum_{i,j} M_{ij} \) over all positive semidefinite \( V(G) \times V(G) \) matrices \( M \) with trace 1 such that \( M_{ij} \leq 0 \) for \( ij \in E(G) \),
2. \( \vartheta(G) \) to be the maximum of \( \sum_{i,j} M_{ij} \) over all positive semidefinite \( V(G) \times V(G) \) matrices \( M \) with trace 1 such that \( M_{ij} = 0 \) for \( ij \in E(G) \), and
3. \( \vartheta'(G) \) to be the maximum of \( \sum_{i,j} M_{ij} \) over all positive semidefinite \( V(G) \times V(G) \) matrices \( M \) with trace 1 such that \( M_{ij} = 0 \) for \( ij \in E(G) \) and \( M_{ij} \geq 0 \) for all \( i, j \).

For \( G \) infinite, we define \( \vartheta(G) = \sup_{H \subseteq G} \vartheta(H) \), where the supremum ranges over all induced finite subgraphs of \( G \), and similarly for \( \vartheta' \) and \( \vartheta^+ \).

Because we can extend a matrix by 0, \( \vartheta(G) = \sup_{H \subseteq G} \vartheta(H) \) holds for finite graphs \( G \), showing that our definition is internally consistent. Moreover, there is no harm in requiring that our matrices \( M \) are symmetric, because we can average \( M \) with its transpose. We will assume symmetry as part of the definition of positive semidefiniteness.

It is known [38, Equation 69] that

\[
\alpha(G) \leq \vartheta'(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi(G).
\]

Moreover, Lovász showed that \( \Theta(G) \leq \vartheta(G) \).

For later use, it will be helpful to formulate the dual semidefinite programs defining \( \vartheta \), \( \vartheta' \), and \( \vartheta^+ \). We will state the duals in terms of positive semidefinite kernels on \( G \), that is, functions \( K : V(G) \times V(G) \to \mathbb{R} \) such that \( K(x, y) = K(y, x) \) and \( \sum_{x, y \in V(G)} w_x w_y K(x, y) \geq 0 \) for all choices of weights \( w : V(G) \to \mathbb{R} \). This is of course equivalent to positive semidefinite \( V(G) \times V(G) \) matrices, but the language of kernels will be convenient.

**Proposition 3.7.** Let \( G \) be a finite graph. Then

1. \( \vartheta'(G) \) is the minimum of \( t \) over all positive semidefinite kernels \( K \) on \( G \) such that \( K(x, x) = t - 1 \) for all \( x \in V(G) \) and \( K(x, y) \leq -1 \) for \( xy \notin E(G) \),
2. \( \vartheta(G) \) is the minimum of \( t \) over all feasible \( K \) and \( t \) for \( \vartheta'(G) \) that additionally satisfy \( K(x, y) = -1 \) for \( xy \notin E(G) \), and
3. \( \vartheta^+(G) \) is the minimum of \( t \) over all feasible \( K \) and \( t \) for \( \vartheta'(G) \) that additionally satisfy \( K(x, y) \geq -1 \) for \( xy \in E(G) \).

For a proof, see [38, Sections 4.2 and 4.4].

We now are ready to prove that these give examples of sandwich functions.
Theorem 3.8. The functions $G \mapsto \vartheta(G)$, $G \mapsto \vartheta'(G)$, and $G \mapsto \vartheta^+(G)$ are sandwich functions.

To aid with the proof, we have the following general lemma.

Lemma 3.9. Suppose $\Phi(G) = \sup_{H \subseteq G} \Phi(H)$, where the supremum is over all induced finite subgraphs. If $\Phi$ satisfies the sandwich function axioms for all finite graphs, then $\Phi$ satisfies them for all graphs in $G$.

Proof. For a graph homomorphism $f : G \to H$,

$$\Phi(G) = \sup_{F \subseteq G} \Phi(F) \leq \sup_{F \subseteq G} \Phi(f(F)) \leq \sup_{F' \subseteq H} \Phi(F') = \Phi(H),$$

where the suprema are over induced finite subgraphs $F$ and $F'$ of $G$ and $H$, respectively.

Every induced finite subgraph $F$ of the join $G * H$ is contained in the join $F_G * F_H$ of some induced finite subgraphs $F_G \subseteq G$ and $F_H \subseteq H$. Thus, $\Phi(F) \leq \Phi(F_G) + \Phi(F_H)$ and so $\Phi(G * H) \leq \Phi(G) + \Phi(H)$. Conversely, for any induced finite subgraphs $F_G \subseteq G$ and $F_H \subseteq H$, $\Phi(F_G) + \Phi(F_H) = \Phi(F_G * F_H) \leq \Phi(G * H)$, so taking the supremum over such finite subgraphs gives the other inequality. Therefore $\Phi(G * H) = \Phi(G) + \Phi(H)$. □

Proof of Theorem 3.8. By the lemma, we reduce to working only with finite graphs. First, we will prove functoriality. Let $f : \overline{G} \to \overline{H}$ be a graph homomorphism. For a $V(\overline{G}) \times V(\overline{G})$ matrix $M$, we define a $V(\overline{H}) \times V(\overline{H})$ matrix $f_* M$ by

$$(f_* M)_{xy} = \sum_{i \in f^{-1}(x)} \sum_{j \in f^{-1}(y)} M_{ij}.$$  

We claim that if $M$ is positive semidefinite, then $f_* M$ is positive semidefinite. We want to show that for arbitrary weights $w_x$ for $x \in V(H)$,

$$\sum_{x,y \in V(H)} w_x w_y (f_* M)_{xy} \geq 0.$$  

That holds because the sum equals

$$\sum_{i,j \in V(G)} w_{f(i)} w_{f(j)} M_{ij},$$

which is nonnegative by the positive semidefiniteness of $M$.

To prove functoriality for $\vartheta^+$, we check as follows that if $M$ is feasible for $\vartheta^+(G)$, then $(\text{Tr}(f_* M))^{-1} f_* M$ is feasible for $\vartheta^+(H)$ and $0 < \text{Tr}(f_* M) \leq 1$. The preimage of any point is a clique in $G$, so $(f_* M)_{xx} \leq \sum_{i \in f^{-1}(x)} M_{ii}$ and hence $\text{Tr}(f_* M) \leq \text{Tr}(M) = 1$. Since $f_* M$ is nontrivial and positive semidefinite, its trace must be strictly positive. If $xy \in E(H)$, and if $f(i) = x$ and $f(j) = y$, then $ij \in E(G)$, so $M_{ij} \leq 0$. Taking the sum over the preimages shows that $(f_* M)_{xy} \leq 0.$
This shows that the matrix is indeed feasible. Now
\[
\vartheta^+(G) = \max_M \sum_{i,j \in V(G)} M_{ij}
\]
\[
= \max_M \sum_{x,y \in V(H)} (f_* M)_{xy}
\]
\[
\leq \max_M \sum_{x,y \in V(H)} (\text{Tr}(f_* M))^{-1} (f_* M)_{xy} \leq \vartheta^+(H),
\]
as desired.

The semidefinite program defining \(\vartheta(G)\) is slightly more stringent, requiring \(M_{ij} = 0\) rather than just \(M_{ij} \leq 0\) for \(ij \in E(G)\). Suppose \(M\) is feasible for \(\vartheta(G)\). As a result \(\text{Tr}(f_* M) = 1\), so we can ignore the normalizing factor \((\text{Tr}(f_* M))^{-1}\). By the same reasoning as before, \((f_* M)_{xy} = \sum_{i \in f^{-1}(x), j \in f^{-1}(y)} M_{ij} = 0\) for \(xy \in E(H)\), since the sum over \(i\) and \(j\) ranges over pairs such that \(ij \in E(G)\). This shows that \(f_* M\) must be feasible for \(\vartheta(H)\), and so this construction shows that \(\vartheta(G) \leq \vartheta(H)\).

Finally, suppose \(M\) is feasible for \(\vartheta'(G)\). This imposes nonnegativity on its entries, and so \(f_* M\) also has nonnegative entries and is feasible for \(\vartheta'(H)\). This construction therefore shows that \(\vartheta'(G) \leq \vartheta'(H)\), and functoriality is proved for finite graphs in all cases.

What remains is to show the additivity under join, or from the perspective of the category \(\mathbb{C}\), additivity under taking disjoint unions of graphs. We break up this statement into two inequalities, which we prove separately.

First, we prove subadditivity, i.e., \(\Psi(G \sqcup H) \leq \Psi(G) + \Psi(H)\) when \(\Psi\) is \(\vartheta\), \(\vartheta'\), or \(\vartheta^+\). We prove the inequality using the dual formulation in terms of kernels \(K\). Let \(K_G\) and \(K_H\) be positive semidefinite kernels that certify the values \(\Psi(G)\) and \(\Psi(H)\) in Proposition 3.7, with \(K_G(x, x) = t_G - 1\) and \(K_H(y, y) = t_H - 1\).

First, we let \(1_{S \times T}(x, y)\) be the function which is 1 for \((x, y) \in S \times T\) and 0 otherwise. We claim that for each \(\alpha > 0\), the kernel
\[
L_\alpha = \alpha^2 1_{V(G) \times V(G)} + \alpha^2 1_{V(H) \times V(H)} - 1_{V(G) \times V(H)} - 1_{V(H) \times V(G)}
\]
on \(G \sqcup H\) is positive semidefinite. To see why, note that for weights \(w_x\),
\[
\sum_{x_1, x_2 \in V(G)} w_{x_1} w_{x_2} \alpha^2 + \sum_{y_1, y_2 \in V(H)} w_{y_1} w_{y_2} \alpha^2 - 2 \sum_{x \in V(G)} w_x w_y
\]
\[
= \left( \alpha \sum_{x \in V(G)} w_x - \alpha^{-1} \sum_{y \in V(H)} w_y \right)^2 \geq 0.
\]

We combine the kernels \(K_G\) and \(K_H\) by setting
\[
K = (\alpha^2 + 1)K_G + (\alpha^{-2} + 1)K_H + L_\alpha,
\]
which is again positive semidefinite on \(G \sqcup H\). In particular, we take \(\alpha^2 = t_H/t_G\). Then for \(x \in V(G)\),
\[
K(x, x) = (\alpha^2 + 1)t_G - 1 = t_G + t_H - 1,
\]
and for \(y \in V(H)\),
\[
K(y, y) = (\alpha^{-2} + 1)t_H - 1 = t_G + t_H - 1,
\]
so \(K(z, z) = t_G + t_H - 1\) for all \(z\). All that remains is to show that \(K\) is feasible for \(\Psi(G \sqcup H)\).

If \(K_G\) is feasible for \(\vartheta'(G)\) and \(K_H\) is feasible for \(\vartheta'(H)\), then \(K\) satisfies \(K(z_1, z_2) \leq -1\) for \(z_1, z_2 \notin E(G) \cup E(H)\) and is thus feasible for \(\vartheta'(G \sqcup H)\). By direct computation, we can check that if \(K_G(x, y) = -1\) for \(xy \notin E(G)\) and \(K_H(x, y) = -1\) for \(xy \notin E(H)\), then \(K(z_1, z_2) = -1\) for \(z_1, z_2 \notin E(G) \cup E(H)\). This shows that feasibility for \(\vartheta(G)\) and \(\vartheta(H)\) implies feasibility of \(K\) for \(\vartheta(G \sqcup H)\). Finally, if \(K_G(x, y) \geq -1\) for \(xy \in E(G)\), then \((\alpha^2 + 1)K_G(x, y) \geq -\alpha^2 - 1\), so \(K(x, y) \geq -1\) for \(x, y \in V(G)\) and \(xy \in E(G)\). Assuming feasibility of \(K_H\) for \(\vartheta^+(H)\) as well shows that \(K\) is feasible for \(\vartheta^+(G \sqcup H)\).

We now prove the other direction of the inequality. Let \(M(G)\) and \(M(H)\) be symmetric positive semidefinite \(V(G) \times V(G)\) and \(V(H) \times V(H)\) matrices, respectively, used to establish lower bounds for \(\Psi(G)\) and \(\Psi(H)\). Define a \((V(G)) \times V(H))\) matrix \(\sigma(G, H)\) so that

\[
\sigma(G, H)_{xy} = \left( \sum_{i \in V(G)} M(G)_{ix} \right) \left( \sum_{j \in V(H)} M(H)_{jy} \right),
\]

and similarly define \(\sigma(H, G)\) as the \((V(H)) \times V(G))\) transpose matrix. We define the total mass of the matrices \(M(G)\) and \(M(H)\) by \(\Sigma(G) = \sum_{ij \in V(G) \times V(G)} M(G)_{ij}\) and \(\Sigma(H) = \sum_{ij \in V(H) \times V(H)} M(H)_{ij}\). These are objectives for our semidefinite program that we may assume to be positive. Define a \((V(G)) \times V(H))\) \((V(G)) \times V(H))\) matrix \(M(G \sqcup H)\) by

\[
M(G \sqcup H)_{xy} = \begin{cases} 
\Sigma(G) M(G)_{xy} / \Sigma(G) + \Sigma(H) & \text{for } (x, y) \in (V(G) \times V(G)), \\
\Sigma(H) M(H)_{xy} / \Sigma(H) + \Sigma(G) & \text{for } (x, y) \in (V(H) \times V(H)), \\
\sigma(G, H)_{xy} / \Sigma(G) + \Sigma(H) & \text{for } (x, y) \in (V(G) \times V(H)), \\
\sigma(H, G)_{xy} / \Sigma(G) + \Sigma(H) & \text{for } (x, y) \in (V(H) \times V(G)).
\end{cases}
\]

We claim, first, that \(M(G \sqcup H)\) is positive semidefinite. Let \(w_x\) be arbitrary weights for \(x \in V(G) \cup V(H)\). Then we want to show that

\[
(\Sigma(G) + \Sigma(H)) \sum_{x,y} w_x w_y M(G \sqcup H)_{xy} \geq 0.
\]

The left side is equal to

\[
\Sigma(G) \sum_{x,y \in V(G)} w_x w_y M(G)_{xy} + \Sigma(H) \sum_{x,y \in V(H)} w_x w_y M(H)_{xy} + 2 \sum_{i,x \in V(G)} w_x w_y M(G)_{iz} M(H)_{zy},
\]

By Cauchy-Schwarz inequality,

\[
\Sigma(G) \sum_{x,y \in V(G)} w_x w_y M(G)_{xy} \geq \left( \sum_{i,x \in V(G)} w_x M(G)_{ix} \right)^2 ;
\]
specifically, this is the Cauchy-Schwartz inequality 
\[ \langle w, 1 \rangle^2 \leq \langle 1, 1 \rangle \langle w, w \rangle \]
for the inner product on \( \mathbb{R}^{V(G)} \) defined by \( \langle a, b \rangle = \sum_{x,y} a_x b_y M_{xy} \). Similarly,

\[
\Sigma(H) \sum_{x,y \in V(H)} w_x w_y M(H)_{xy} \geq \left( \sum_{j,y \in V(H)} w_y M(H)_{jy} \right)^2.
\]

Applying these inequalities shows that the left side is at least
\[
\left( \sum_{i,x \in V(G)} w_x M(G)_{ix} + \sum_{j,y \in V(H)} w_y M(H)_{jy} \right)^2,
\]
which is nonnegative.

Next, suppose that \( M(G) \) and \( M(H) \) have trace 1. Then \( M(G \sqcup H) \) has trace 1 by direct computation. Additionally,

\[
\sum_{x,y \in V(G) \cup V(H)} M(G \sqcup H)_{xy} = \Sigma(G) + \Sigma(H)
\]
by a straightforward computation.

If \( xy \) is an edge in the disjoint union, it is an edge in one of \( G \) or \( H \), and \( M(G \sqcup H)_{xy} \) has the same sign as \( M(G)_{xy} \) or \( M(H)_{xy} \), depending on whether the edge is in \( G \) or \( H \). Thus, if \( M(G) \) and \( M(H) \) are feasible for \( \vartheta^+(G) \) and \( \vartheta^+(H) \), respectively, then \( M(G \sqcup H) \) is feasible for \( \vartheta^+(G \sqcup H) \). If \( M(G) \) and \( M(H) \) are feasible for \( \vartheta \), then \( M(G \sqcup H) \) is feasible for \( \vartheta(G \sqcup H) \). Finally, if \( M(G) \) and \( M(H) \) have nonnegative entries, then \( M(G \sqcup H) \) has nonnegative entries, so \( M(G \sqcup H) \) is feasible for \( \vartheta'(G \sqcup H) \) in this case. Thus, this construction completes the proof of additivity under disjoint union. \qed

4. Asymptotics of packing bound functions on rectifiable sets

Our main result is a convergence theorem for all packing bound functions and all \((\mathcal{H}_n, n)\)-rectifiable sets \( C \) satisfying \( \mathcal{M}_n(C) = \mathcal{H}_n(C) \). This recovers a statement similar to best packing on rectifiable sets [7], which is in fact a little stronger than Theorem 1.3. The structure of the proof will mimic in several places the proof of the Poppy Seed Bagel Theorem [28]. First, we will show the results for Jordan-measurable subsets of \( \mathbb{R}^n \). As in Section 2, \( A \) will denote an arbitrary packing bound function, \( C \) and \( C' \) will be elements of \( \mathcal{B} \), and \( I^n \) will be the unit cube in \( \mathbb{R}^n \).

**Proposition 4.1.** If \( C \) is a product of intervals, then

\[
\lim_{r \to \infty} A(rC)/r^n = \delta_{A,n} \mathcal{L}_n(C).
\]

**Proof.** This follows from the union bounds together with the following geometric fact: for each \( \varepsilon > 0 \), there exist \( R \) and \( \varepsilon' > 0 \) such that for all \( r > R \), \( rC \) contains disjoint copies of the unit cube occupying at least a \( 1 - \varepsilon \) fraction of the volume of \( rC \) and separated by distance at least \( \varepsilon' \), and is contained in a cover by copies of the unit cube of total volume fraction at most \( 1 + \varepsilon \).
Choosing $k$ such that $A(2^k I^n)2^{-kn}$ is within $\varepsilon$ the limit $\delta_{A,n}$ and such that $2^k \varepsilon' \geq 2$, and applying both union bounds, we find that

$$(1 - \varepsilon)^2 \delta_{A,n} \mathcal{L}_n(C) \leq (1 - \varepsilon) \frac{A(2^k I^n)}{2^{kn}} \mathcal{L}_n(C)$$

$$\leq A(2^{k+1} I^n) \frac{r_{n+1}}{2^{kn} r^n}$$

$$\leq (1 + \varepsilon) (2^{k+1} I^n) \frac{r_{n+1}}{2^{kn} r^n} \mathcal{L}_n(C)$$

$$\leq (1 + \varepsilon)^2 \delta_{A,n} \mathcal{L}_n(C). \quad \square$$

**Proposition 4.2.** If $C$ is Jordan-measurable, then $\lim_{r \to \infty} A(rC)/r^n = \delta_{A,n} \mathcal{L}_n(C)$.

**Proof.** Let $\{D_i\}$ be a finite covering of $C$ by small cubes such that $\sum_i \mathcal{L}_n(D_i) < \mathcal{L}_n(C) + \varepsilon$. Then

$$\limsup_{r \to \infty} \frac{A(rC)}{r^n} \leq \frac{1}{\varepsilon} \sum_i \lim_{r \to \infty} \frac{A(rD_i)}{r^n}$$

$$\leq \delta_{A,n} (\mathcal{L}_n(C) + \varepsilon)$$

Let $\{C_i\}$ be a finite collection of disjoint, closed cubes contained in $C$ such that $\sum_i \mathcal{L}_n(C_i) > \mathcal{L}_n(C) - \varepsilon$. Then $\delta := \min_{i,j} d(C_i, C_j) > 0$. For $r > 2/\delta$,

$$A(rC) \geq A\left(r \bigcup_i C_i\right) = \sum_i A(rC_i).$$

Dividing by $r^n$ and passing to the limit infimum as $r \to \infty$, we conclude that

$$\liminf_{r \to \infty} \frac{A(rC)}{r^n} \geq \delta_{A,n} (\mathcal{L}_n(C) - \varepsilon),$$

which completes the proof. \square

Next we prove the result for all compact sets.

**Lemma 4.3.** For every compact set $D$ in $\mathbb{R}^n$,

$$\limsup_{r \to \infty} \frac{A(rD)}{r^n} \leq \delta_{A,n} \mathcal{L}_n(D).$$

**Proof.** For every $\varepsilon$, there is a Jordan-measurable set $C$ containing $D$ such that $\mathcal{L}_n(C \setminus D) < \varepsilon$ (see Remark 2.8). By the Lipschitz inequality, $A(rC) \geq A(rD)$, and we conclude the result by dividing by $r^n$ and taking the limit sup. \square

For the other direction, we first prove a lemma on compact Jordan-measurable sets:

**Lemma 4.4.** For each compact Jordan-measurable $B \subseteq \mathbb{R}^n$ and $\varepsilon > 0$ fixed, there is an $\varepsilon' > 0$ such that for any compact subset $C \subseteq B$ with $\mathcal{L}_n(C) \geq \mathcal{L}_n(B) - \varepsilon'$,

$$(1 - \varepsilon) \liminf_{r \to \infty} \frac{A(rB)}{r^n} \leq \liminf_{r \to \infty} \frac{A(rC)}{r^n} + \varepsilon.$$ 

**Proof.** This is a consequence of Proposition 2.14, using the fact that Minkowski content is equal to Lebesgue measure for compact subsets of $\mathbb{R}^n$ (see [18, Theorem 3.2.39]). \square
We will need the Besicovitch covering lemma (as in [8, Theorem 8.6.10]), which we state as follows:

**Lemma 4.5.** Let $\mu$ be a Borel measure on $\mathbb{R}^p$, let $A \subseteq \mathbb{R}^p$ be a set of finite $\mu$-measure, and let $\mathcal{F}$ be a set of balls such that for all $x \in A$, the infimum of $r$ over the set of balls $B(x, r)$ in $\mathcal{F}$ is 0 (that is, $\mathcal{F}$ contains balls of arbitrarily small radius centered at all points of $A$). Then there is a countable subcollection of $\mathcal{F}$ that are pairwise disjoint and cover $\mu$-almost all $A$.

We are now prepared to prove the complementary inequality to Lemma 4.3 for all compact sets. This argument adapts the proof of [8, Theorem 8.6.11].

**Theorem 4.6.** Let $A$ be a packing bound function. For any compact set $C$ in $\mathbb{R}^n$,

$$
\lim_{r \to \infty} \frac{A(rC)}{r^n} = \delta_{A,n} \mathcal{L}_n(C).
$$

**Proof.** Let $\varepsilon > 0$ be fixed. It suffices to prove that

$$
\liminf_{r \to \infty} \frac{A(rC)}{r^n} \geq (1 - \varepsilon) \delta_{A,n} \mathcal{L}_n(C) - \varepsilon.
$$

Define the set

$$
C^*: = \left\{ x \in C : \lim_{r \to 0^+} \frac{\mathcal{L}_n(B(x, r) \cap C)}{\mathcal{L}_n(B(x, r))} = 1 \right\}.
$$

This set satisfies $\mathcal{L}_n(C^*) = \mathcal{L}_n(C)$ by the Lebesgue density theorem.

Now to apply Besicovitch covering lemma, we let the set $\mathcal{F}$ consist of all closed balls $B(x, r)$ around points $x \in C^*$ such that $r < 1$ and

$$
\frac{\mathcal{L}_n(B(x, r) \cap C^*)}{\mathcal{L}_n(B(x, r))} \geq 1 - \varepsilon',
$$

where $\varepsilon' > 0$ can be taken to be arbitrarily small. By the Besicovitch covering lemma, we can choose a countable disjoint subcollection $B_i$ of closed balls whose union covers almost all of $C^*$ and hence almost all of $C$. Define $C_i = C \cap B_i$.

We can choose $N$ so that

$$
\mathcal{L}_n \left( \bigcup_{i=1}^N B_i \right) \geq (1 - \varepsilon) \mathcal{L}_n(C).
$$

Because $B_i$ is in $\mathcal{F}$ and $\bigcup_{i=1}^N B_i \subseteq C(1)$,

$$
\mathcal{L}_n \left( \bigcup_{i=1}^N C_i \right) \geq \left( 1 - \frac{\varepsilon'}{\mathcal{L}_n(C(1))} \right) \mathcal{L}_n \left( \bigcup_{i=1}^N B_i \right) \geq \mathcal{L}_n \left( \bigcup_{i=1}^N B_i \right) - \varepsilon'.
$$

Then since the balls are compact sets, they are separated from each other, and we may find a $\delta > 0$ such that $d(B_i, B_j) \geq \delta$ for all distinct $i, j \leq N$. Now Lemma 4.4 tells us that if $\varepsilon'$ is small enough relative to $\varepsilon$, then

$$
\liminf_{r \to \infty} \frac{A(r \bigcup_{i=1}^N C_i)}{r^n} \geq (1 - \varepsilon) \liminf_{r \to \infty} \frac{A(r \bigcup_{i=1}^N B_i)}{r^n} - \varepsilon.
$$

For $r \geq 2\delta^{-1}$, we can apply the union axiom in the definition of a packing bound function to obtain

$$
A \left( r \bigcup_{i=1}^N B_i \right) = \sum_{i=1}^N A(rB_i).
$$
Combining these inequalities yields
\[
\liminf_{r \to \infty} \frac{A(rC)}{r^n} \geq \liminf_{r \to \infty} \frac{A\left(r \bigcup_{i=1}^{N} C_i\right)}{r^n} \\
\geq (1 - \varepsilon) \liminf_{r \to \infty} \frac{A\left(r \bigcup_{i=1}^{N} B_i\right)}{r^n} - \varepsilon \\
= (1 - \varepsilon) \liminf_{r \to \infty} \sum_{i=1}^{N} \frac{A(rB_i)}{r^n} - \varepsilon \\
= (1 - \varepsilon) \sum_{i=1}^{N} \delta_{A,n} \mathcal{L}_n(rB_i) - \varepsilon,
\]
where the last equality holds because \(B_i\) is Jordan-measurable. Finally, we obtain a lower bound of
\[
(1 - \varepsilon) \sum_{i=1}^{N} \delta_{A,n} \mathcal{L}_n(rB_i) - \varepsilon \geq (1 - \varepsilon)^2 \delta_{A,n} \mathcal{L}_n(C) - \varepsilon,
\]
as desired. \(\square\)

**Remark 4.7.** The above theorem can be proved much more easily if the Euclidean bound is available for \(A\).

Proposition 2.6 automatically gives an extension to arbitrary subsets of \(\mathbb{R}^n\) that uses Minkowski content instead of Lebesgue measure.

**Corollary 4.8.** For an arbitrary bounded Borel subset \(C \subseteq \mathbb{R}^n\),
\[
\lim_{r \to \infty} \frac{A(rC)}{r^n} = \delta_{A,n} \mathcal{M}_n(C).
\]

Note that because \(C \subseteq \mathbb{R}^n\), the Minkowski content \(\mathcal{M}_n(C)\) always exists.

**Proof.** Let \(\overline{C}\) be the closure of \(C\), which is compact and \(n\)-rectifiable and hence satisfies \(\mathcal{M}_n(C) = \mathcal{M}_n(\overline{C}) = \mathcal{L}_n(\overline{C})\). By density, this implies for any \(\varepsilon > 0\),
\[
A(r\overline{C}) \leq A(r(1 + \varepsilon)C) \leq A(r(1 + \varepsilon)\overline{C})
\]
and after dividing by \(r^n\) the left and right sides both converge to \(\delta_{A,n} \mathcal{L}_n(\overline{C})(1 + O(\varepsilon))\) as \(r \to \infty\). \(\square\)

Now, we wish to extend the result to \(n\)-rectifiable sets, in fact, to a slightly more general setting. We need some notions from geometric measure theory.

**Definition 4.9.** For a measure \(\mu\), a \((\mu, n)\)-rectifiable set is a bounded Borel subset \(E\) of \(\mathbb{R}^d\) such that there are Lipschitz maps \(\psi_i : \mathbb{R}^n \to \mathbb{R}^d\) and bounded Borel subsets \(E_i\) of \(\mathbb{R}^n\) for which \(\mu(E \setminus \bigcup_i \psi_i(E_i)) = 0\).

We have the following lemma, which is [18, Lemma 3.2.18]:

**Lemma 4.10.** Let \(C\) be an \((\mathcal{H}_n, n)\)-rectifiable set. Then for every \(\varepsilon > 0\), there are compact subsets \(C_1, C_2, \ldots \subseteq \mathbb{R}^n\) and bi-Lipschitz maps \(\psi_i : C_i \to C\) with Lipschitz constant \(1 + \varepsilon\) (in both directions) such that the sets \(\psi_i(C_i)\) are disjoint and
\[
\mathcal{H}_n\left(C \setminus \bigcup_{i=1}^{\infty} \psi_i(C_i)\right) = 0.
\]
We will also need [8, Lemma 8.7.2]:

**Lemma 4.11.** If $C$ is a compact $(\mathcal{H}_n, n)$-rectifiable set with $\mathcal{M}_n(C) = \mathcal{H}_n(C)$, then every compact subset $K$ of $C$ is $(\mathcal{H}_n, n)$-rectifiable and satisfies $\mathcal{M}_n(K) = \mathcal{H}_n(K)$.

Now we can prove our main theorem:

**Theorem 4.12.** Let $C \subseteq \mathbb{R}^d$ be an $(\mathcal{H}_n, n)$-rectifiable set with closure $\overline{C}$ satisfying the property that $\mathcal{M}_n(C) = \mathcal{H}_n(\overline{C}) < \infty$, and let $A$ be a packing bound function. Then

$$\lim_{r \to \infty} \frac{A(rC)}{r^n} = \delta_{A,n} \mathcal{M}_n(C).$$

In particular, the above theorem holds for all compact smooth $n$-manifolds or compact subsets of smooth $n$-manifolds embedded in $\mathbb{R}^d$ for some $d$. Since compact $n$-rectifiable sets are also $(\mathcal{H}_n, n)$-rectifiable and satisfy $\mathcal{M}_n(C) = \mathcal{H}_n(C)$, Theorem 4.12 implies Theorem 1.3.

**Proof.** By using the density property of $A$, we see that $A(rC) \leq A(r(1 + \varepsilon)C) \leq A(r(1 + \varepsilon)\overline{C})$, which reduces the goal to proving the statement for $\overline{C}$. From now on we assume $C$ is compact.

We first prove the $\geq$ direction. In the notation of Lemma 4.10, choose $N$ so that $\sum_{i=1}^{N} \mathcal{H}_n(\psi_i(C_i)) \geq (1 - \varepsilon)\mathcal{H}_n(C)$. By compactness, there is a $\delta > 0$ such that $d(\psi_i(C_i), \psi_j(C_j)) \geq \delta$ for all distinct $i, j \leq N$. Then for $r \geq 2\delta^{-1}$,

$$A(rC) \geq \sum_{i=1}^{N} A(r\psi_i(C_i)) \geq \sum_{i=1}^{N} A((1 + \varepsilon)^{-1}rC_i),$$

where the last inequality follows from the Lipschitz property. Thus,

$$\frac{A(rC)}{r^n \mathcal{H}_n(C)} \geq (1 - \varepsilon) \frac{\sum_{i=1}^{N} A((1 + \varepsilon)^{-1}rC_i)}{\sum_{i=1}^{N} r^n \mathcal{H}_n(\psi_i(C_i))}.$$ 

Furthermore, $\mathcal{H}_n(\psi_i(C_i)) \leq (1 + \varepsilon)^n \mathcal{H}_n(C_i)$ because $\psi_i$ has Lipschitz constant $1 + \varepsilon$ (this bound follows directly from the definition of $\mathcal{H}_n$; see, for example, [2, Proposition 2.49(iv)]), and thus $\mathcal{H}_n(\psi_i(C_i)) \leq (1 + \varepsilon)^{2n} \mathcal{H}_n((1 + \varepsilon)^{-1}C_i)$. We conclude that

$$\lim_{r \to \infty} \inf \frac{A(rC)}{r^n \mathcal{H}_n(C)} \geq (1 - \varepsilon)(1 + \varepsilon)^{-2n} \delta_{A,n},$$

and the conclusion follows by letting $\varepsilon$ tend to 0.

Now we wish to prove the other direction,

$$\limsup_{r \to \infty} A(rC)r^{-n} \leq \delta_{A,n} \mathcal{M}_n(C).$$

We note that $\bigcup_i \psi_i(C_i)$ is a compact subset of $C$, and since $\mathcal{M}_n(C) = \mathcal{H}_n(C)$ and the complement of $\bigcup_i \psi_i(C_i)$ has measure 0 under $\mathcal{H}_n$, it follows that $\mathcal{M}_n(C) = \mathcal{H}_n(\bigcup_i \psi_i(C_i)) = \mathcal{M}_n(\bigcup_i \psi_i(C_i))$ by Lemma 4.11. So by Proposition 2.14,

$$\limsup_{r \to \infty} A(rC)r^{-n} = \limsup_{r \to \infty} \frac{A(r\bigcup_i \psi_i(C_i))}{r^n},$$

and it suffices to prove the corresponding bound for $\bigcup_i \psi_i(C_i)$. 

Now by density,
\[ A\left( r \bigcup_i \psi_i(C_i) \right) \leq A\left( r(1 + \varepsilon) \bigcup_{i=1}^\infty \psi_i(C_i) \right). \]

By the nested union and Lipschitz properties,
\[ A\left( r(1 + \varepsilon) \bigcup_{i=1}^\infty \psi_i(C_i) \right) \leq \lim_{N \to \infty} A\left( r(1 + \varepsilon)^2 \bigcup_{i=1}^N \psi_i(C_i) \right) \leq \sum_{i=1}^\infty A(r(1 + \varepsilon)^3 C_i). \]

Dividing by \( r^n \) and taking the limit as \( r \to \infty \), we find that
\[ \limsup_{r \to \infty} \frac{A(rC)}{r^n} \leq (1 + \varepsilon)^3 \delta_{A,n} \sum_{i=1}^\infty \mathcal{H}_n(C_i) \leq (1 + \varepsilon)^4 \delta_{A,n} \mathcal{H}_n(C). \]

Letting \( \varepsilon \to 0 \) completes the result. \( \square \)

**Remark 4.13.** We do not know whether Theorem 4.12 generalizes to sets \( C \) with non-integral Hausdorff dimension \( n \). The limit of \( A(rC)/r^n \) at least makes sense, and the value \( \delta_{A,n} \) can also sometimes be generalized for non-integral \( n \). For the linear programming bound, for example, the radial Fourier transform is used, which can be written in terms of a Bessel function \( J_{\nu} \) with \( \nu = \frac{d}{2} - 1 \). By taking \( \nu \) to be any real number greater than \( -\frac{1}{2} \), we can formally write down a linear program for non-integral dimensional Euclidean space, but it is not clear whether it has any relationship to the limit of \( A(rC)/r^n \). If there is indeed a generalization in this direction, that would give geometric meaning to the versions of the linear program when the dimension is not integral. Could this program even be sharp in some non-integral dimensions?

5. **The Euclidean limits of \( \vartheta' \), \( \vartheta \), and \( \vartheta^+ \)**

In this section, we consider the Euclidean limits of the packing bound functions corresponding to the sandwich functions \( \vartheta' \), \( \vartheta \), and \( \vartheta^+ \). It turns out that \( \delta_{\vartheta,n} \mathcal{L}_n(B_1^n) \) is the Cohn-Elkies linear programming bound [13] for the sphere packing density in \( \mathbb{R}^n \), which is the best bound known for large \( n \). In addition, the Euclidean limits of \( \vartheta \) and \( \vartheta^+ \) can be computed exactly, and they are equal to \( \mathcal{L}_n(B_1^n)^{-1} \), where \( \mathcal{L}_n(B_1^n) \) is the volume of an \( n \)-ball of radius 1. In other words,
\[ \delta_{\vartheta,n} = \delta_{\vartheta^+,n} = \pi^{-\frac{n}{2}} \frac{n}{2} + 1, \]
and the resulting density bounds for the sphere packing problem are trivial:
\[ \delta_{\vartheta,n} \mathcal{L}_n(B_1^n) = \delta_{\vartheta^+,n} \mathcal{L}_n(B_1^n) = 1. \]

Despite the weakness of the corresponding sphere packing bounds, the Euclidean limits \( \delta_{\vartheta,n} \) and \( \delta_{\vartheta^+,n} \) involve interesting mathematics. These ideas originated in Siegel’s proof of Minkowski's theorem via Poisson summation (see, for example, Section 2.11.4 in [17]). Minkowski’s theorem can be interpreted as saying that the
maximum lattice packing density of a convex, origin-symmetric body is 1, which
Siegel proved by applying Poisson summation to the convolution of the indicator
function of the convex body with itself. It is natural to ask whether another auxiliary
function could prove a better bound than 1. Siegel showed that the answer is no
under certain hypotheses [41], which amounts to computing the limits \( \delta_{\varphi,n} \) and
\( \delta_{\varphi^+,n} \), and Gorbachev [22] rediscovered this theorem with a different proof.

In this section and the next, we will first discuss how the Delsarte problem gives
a packing bound function, and then we will generalize the results to the Lasserre
hierarchy from [35]. We will denote the packing bound functions corresponding to
\( \varphi \), \( \varphi' \), and \( \varphi^+ \) under Theorems 3.1 and 3.8 by \( \varphi \), \( \varphi' \), and \( \varphi^+ \) again; this is an abuse of
notation, but it will not cause any actual ambiguity. We will also define topological
variants \( \varphi_{\text{top}} \), \( \varphi'_{\text{top}} \), and \( \varphi^+_{\text{top}} \) below, which will impose continuity.

First, we need a few definitions. By a finite signed measure, we mean a signed
Borel measure \( \mu \) on \( \mathbb{R}^d \) such that \(-\infty < \mu(A) < \infty \) for every Borel set \( A \) (bounded or
not). If \( \mu \) and \( \mu_1, \mu_2, \ldots \) are finite signed measures, we say \( \mu_n \) is weak*- convergent to
\( \mu \) if every bounded, continuous function \( f: \mathbb{R}^d \to \mathbb{R} \) satisfies

\[
\int f \, d\mu_n \to \int f \, d\mu
\]
as \( n \to \infty \).

A positive semidefinite kernel on a bounded Borel set \( C \) is a function \( K: C \times C \to \mathbb{R} \) such that \( K(x, y) = K(y, x) \) and for every finite subset \( S \) of \( C \) and function
\( w: S \to \mathbb{R} \),

\[
\sum_{x,y \in S} K(x, y)w(x)w(y) \geq 0.
\]

For example, \( K(x, y) := f(x)f(y) \) is positive semidefinite for any function \( f : C \to \mathbb{R} \),
and we denote this kernel by \( f \otimes f \). If \( K \) is continuous, then the inequality defining
positive semidefiniteness is equivalent to

\[
\int \int_{C \times C} K(x, y) \, d\mu(x) \, d\mu(y) \geq 0
\]
for all finite signed measures \( \mu \); here finite sets \( S \) correspond to \( \mu \) with finite
support (i.e., linear combinations of delta functions supported at points), and finitely
supported signed measures are weak*- dense among all finite signed measures. A
finite signed measure \( \nu \) is positive semidefinite on \( C \times C \) if

\[
\int \int_{C \times C} K(x, y) \, d\nu(x, y) \geq 0
\]
for all continuous, positive semidefinite kernels \( K \) on \( C \). For simplicity, we often will
call \( \nu \) positive semidefinite on \( C \), rather than \( C \times C \). For finitely supported positive
semidefinite measures \( \nu \), \( \int_{C \times C} K(x, y) \, d\nu(x, y) \geq 0 \) for all positive semidefinite
kernels \( K \), even if \( K \) is not continuous.

In this language, we can interpret the construction of packing bound functions
from Theorem 3.1, Definition 3.6, and Theorem 3.8 as follows. Let \( \Delta(C) \) denote the
diagonal \( \{(x, x) : x \in C\} \) in \( C \times C \).

**Proposition 5.1.** For each bounded Borel subset \( C \) of \( \mathbb{R}^d \),

1. \( \varphi^+(C) \) is the supremum of \( \nu(C \times C) \) over finitely supported positive semi-
definite signed measures \( \nu \) on \( C \) such that \( \nu(\Delta(C)) = 1 \) and \( \nu \leq 0 \) on the
set of pairs \( (x, y) \) such that \( 0 < |x - y| < 2 \),
(2) $\vartheta(C)$ is the supremum of $\nu(C \times C)$ over finitely supported positive semidefinite signed measures $\nu$ on $C$ such that $\nu(\Delta(C)) = 1$ and $\nu = 0$ for pairs $(x, y)$ such that $0 < |x - y| < 2$, and

(3) $\vartheta'(C)$ is the supremum of $\nu(C \times C)$ over finitely supported positive semidefinite signed measures $\nu$ on $C$ such that $\nu(\Delta(C)) = 1, \nu = 0$ for pairs $(x, y)$ such that $0 < |x - y| < 2$, and $\nu \geq 0$ everywhere.

Proof. This follows immediately by considering matrices as finitely supported signed measures on $C \times C$. Every measure $\nu$ as above gives a feasible solution for the corresponding sandwich function $\vartheta^+, \vartheta$, or $\vartheta'$ over some induced finite subgraph. Conversely, any feasible solution over an induced finite subgraph gives a feasible $\nu$. Taking the supremum over feasible solutions on both sides yields the result.

Similarly, the dual semidefinite programs are as follows:

(1) $\vartheta'(C)^*$ is the infimum of $t$ over all positive semidefinite kernels $K$ on finite subsets $S$ of $C$ such that $K(x, x) = t - 1$ for all $x \in S$ and $K(x, y) \leq -1$ when $|x - y| \geq 2$,

(2) $\vartheta(C)^*$ is the infimum of $t$ over all positive semidefinite kernels $K$ on finite subsets $S$ of $C$ such that $K(x, x) = t - 1$ for all $x \in S$ and $K(x, y) = -1$ when $|x - y| \geq 2$, and

(3) $\vartheta^+(C)^*$ is the infimum of $t$ over all positive semidefinite kernels $K$ on finite subsets $S$ of $C$ such that $K(x, x) = t - 1$ for all $x \in S$, $K(x, y) \geq -1$ for all $x, y \in S$, and $K(x, y) = -1$ when $|x - y| \geq 2$.

It follows from Proposition 3.7 that $\vartheta'(C) = \vartheta'(C)^*, \vartheta(C) = \vartheta(C)^*$, and $\vartheta^+(C) = \vartheta^+(C)^*$, and that we can take $S = C$ in the dual programs when $C$ is finite. The description in terms of finitely supported signed measures suggests that we should consider the analogous problems using arbitrary finite signed measures instead. Such measures have the advantage that one can average feasible solutions over the action of a compact Lie group of isometries and arrive at invariant semidefinite programs for packing problems [4]. We will refer to these as the topological analogues of the discrete $\vartheta^+, \vartheta$, and $\vartheta'$, and denote them by $\vartheta^{+, \top}, \vartheta^{\top},$ and $\vartheta'^{\top},$ respectively. In other words,

(1) $\vartheta^{+, \top}(C)$ is the supremum of $\nu(C \times C)$ over finite, positive semidefinite signed measures $\nu$ on $C$ such that $\nu(\Delta(C)) = 1$ and $\nu \leq 0$ on the set of pairs $(x, y)$ such that $0 < |x - y| < 2$,

(2) $\vartheta^{\top}(C)$ is the supremum of $\nu(C \times C)$ over finite, positive semidefinite signed measures $\nu$ on $C$ such that $\nu(\Delta(C)) = 1$ and $\nu = 0$ for pairs $(x, y)$ such that $0 < |x - y| < 2$, and

(3) $\vartheta'^{\top}(C)$ is the supremum of $\nu(C \times C)$ over finite, positive semidefinite signed measures $\nu$ on $C$ such that $\nu(\Delta(C)) = 1, \nu = 0$ for pairs $(x, y)$ such that $0 < |x - y| < 2$, and $\nu \geq 0$ everywhere.

The topological analogue of the dual semidefinite programs for $\vartheta^+, \vartheta$, and $\vartheta'$ uses continuous kernels. We define

(1) $\vartheta^{+, \top}(C)^*$ to be the infimum of $t$ over all continuous, positive semidefinite kernels $K$ on $C$ such that $K(x, x) \leq t - 1$ for all $x$ and $K(x, y) \leq -1$ when $|x - y| \geq 2$,

(2) $\vartheta^{\top}(C)^*$ to be the infimum of $t$ over all continuous, positive semidefinite kernels $K$ on $C$ such that $K(x, x) \leq t - 1$ for all $x$ and $K(x, y) = -1$ when $|x - y| \geq 2$, and
(3) $\vartheta^{+,\text{top}}(C)^*$ to be the infimum of $t$ over all continuous, positive semidefinite kernels $K$ on $C$ such that such that $K(x, x) \leq t - 1$ for all $x$, $K(x, y) = -1$ when $|x - y| \geq 2$, and $K(x, y) \geq -1$ everywhere.

Note that here we require just $K(x, x) \leq t - 1$, rather than $K(x, x) = t - 1$ as in the finite case (Proposition 3.7). That change makes no difference in the finite case, but it will be convenient in Proposition 5.3.

The following proposition is the topological analogue of Proposition 3.7.

**Proposition 5.2** (Weak duality). For each bounded Borel set $C$,

$$\vartheta^{t,\text{top}}(C) \leq \vartheta^{t,\text{top}}(C)^*, \quad \vartheta^{\text{top}}(C) \leq \vartheta^{\text{top}}(C)^*, \quad \text{and} \quad \vartheta^{+,\text{top}}(C) \leq \vartheta^{+,\text{top}}(C)^*.$$  

**Proof.** Let $A$ be $\vartheta^{t,\text{top}}$, $\vartheta^{\text{top}}$, or $\vartheta^{+,\text{top}}$. In each case, the hypotheses on $\nu$, $t$, and $K$ in the definitions of $A(C)$ and $A(C)^*$ show that

$$\nu(C \times C) = \iint_{C \times C} d\nu(x, y) \leq \iint_{C \times C} (1 + K(x, y)) d\nu(x, y) \leq t \nu(\Delta(C)) = t,$$

as desired. \hfill \Box

The following proposition clarifies the relationship between the topological and discrete invariants.

**Proposition 5.3.** Let $C \subseteq \mathbb{R}^d$ be compact, and let $C(\varepsilon)$ denote an $\varepsilon$-neighborhood around $C$. Then

$$\vartheta^{t,\text{top}}(C) = \lim_{\varepsilon \to 0^+} \vartheta'(C(\varepsilon)).$$

In addition, for any $\varepsilon > 0$,

$$\vartheta'(C) \leq \vartheta^{t,\text{top}}(C) \leq \vartheta^{t,\text{top}}(C)^* \leq \vartheta'(C(\varepsilon)),$$

$$\vartheta^{+}(C) \leq \vartheta^{+,\text{top}}(C) \leq \vartheta^{+,\text{top}}(C)^* \leq \vartheta^{+}(C(\varepsilon)),$$

and

$$\vartheta(C) \leq \vartheta^{\text{top}}(C) \leq \vartheta^{\text{top}}(C)^* \leq \vartheta(C(\varepsilon)).$$

We do not know whether there is any compact set $C \subseteq \mathbb{R}^d$ for which $\vartheta'(C) \neq \vartheta^{t,\text{top}}(C)$.

**Corollary 5.4** (Strong duality). For each bounded Borel set $C$,

$$\vartheta^{t,\text{top}}(C) = \vartheta^{t,\text{top}}(C)^*.$$  

This corollary follows directly from Proposition 5.3, and will be generalized to the Lasserre hierarchy in Theorem 6.10. We do not know whether strong duality holds for $\theta$ or $\theta^+$, but they are less significant as packing bounds.

The proof of Proposition 5.3 will use the following lemma, which allows us to reduce a problem on an infinite topological space to a problem on a finite simplicial complex.

**Lemma 5.5.** For every compact subset $C \subseteq \mathbb{R}^d$ and $\varepsilon > 0$, there is a finite, pure simplicial $d$-complex $Y$ embedded in $\mathbb{R}^d$ such that $C \subseteq Y \subseteq C(\varepsilon)$ and every simplex in $Y$ has diameter less than $\varepsilon$.

Here, saying $Y$ is finite means the complex is built from finitely many simplices in $\mathbb{R}^d$, not that its geometric realization is a finite set.
Proof. Given a triangulation of $\mathbb{R}^d$ with simplices of diameter less than $\varepsilon$, let $Y$ consist of the simplices that are contained in $C(\varepsilon)$. To obtain such a triangulation, we can start with a decomposition of a cube of diameter less than $\varepsilon$ into $d!$ simplices whose vertices are vertices of the cube, and then extend the decomposition to $\mathbb{R}^d$ via reflection across the faces of the cube.

Proof of Proposition 5.3. We first prove that
\[
\limsup_{\varepsilon \to 0^+} \vartheta'(C(\varepsilon)) \leq \vartheta'^{\text{top}}(C).
\]
Choose a sequence $\varepsilon_n \to 0$ and $\varepsilon_n > 0$ and feasible solutions $\nu_n$ for $\vartheta'(C(\varepsilon_n))$. Since all these measures are nonnegative, their supports are bounded, and their total measures are bounded by $\vartheta'(C(\max_n \varepsilon_n))$, there is a weak-* convergent subsequence. By passing to such a subsequence, we can assume that the entire sequence is weak-* convergent.

Let $\nu$ be the weak-* limit of $\nu_n$; then $\nu$ is a nonnegative measure because each $\nu_n$ is nonnegative. It must be supported on $\bigcap_n C(\varepsilon_n)$, which is equal to $C$ by the assumption that $C \subseteq \mathbb{R}^d$ is compact. It must be positive semidefinite because every continuous function $f$ on $C$ can be extended to a continuous function $f'$ on $\mathbb{R}^d$ with compact support, and $\nu(f \otimes f) = \nu(f' \otimes f')$, which is the limit of the sequence $\nu_n(f' \otimes f')$, each term of which is nonnegative. Finally, $\nu$ must be supported on $\{(x, y) : x = y$ or $|x - y| \geq 2\}$, because $\{(x, y) : 0 < |x - y| < 2\}$ is an open set and thus
\[
\nu(\{(x, y) : 0 < |x - y| < 2\}) \leq \liminf_{n \to \infty} \nu_n(\{(x, y) \in C \times C : 0 < |x - y| < 2\}) = \liminf_{n \to \infty} 0 = 0.
\]
Therefore $\nu$ is a feasible solution for $\vartheta'^{\text{top}}(C)$. We also note that the objective $\nu(C \times C) = \nu(\mathbb{R}^d \times \mathbb{R}^d)$ is a continuous functional for the weak-* topology (because it is the integral of the constant function 1). By taking $\nu_n$ to satisfy $\nu_n(\mathbb{R}^d \times \mathbb{R}^d) \geq \vartheta'(C(\varepsilon_n)) - \delta$ for very small $\delta > 0$, we can guarantee that
\[
\limsup_{\varepsilon \to 0^+} \vartheta'(C(\varepsilon)) - \delta \leq \vartheta'^{\text{top}}(C).
\]
The desired inequality follows.

Let $A = \vartheta, \vartheta^+, \text{ or } \vartheta'$. It remains to prove that
\[
A(C) \leq A^{\text{top}}(C) \leq A^{\text{top}}(C)^* \leq A(C(\varepsilon))
\]
for each $\varepsilon > 0$. The first inequality follows immediately from the definitions, and the second by weak duality in Proposition 5.3, so we just need to show the last inequality. In fact, we will prove an upper bound of $A((1 + \varepsilon)C(\varepsilon))$. That bound looks a little weaker, but $(1 + \varepsilon)C(\varepsilon)$ is contained in a neighborhood of $C(\varepsilon)$ of radius $\varepsilon \sup_{x \in C(\varepsilon)} |x| = \varepsilon^2 + \varepsilon \sup_{x \in C} |x|$, and therefore $A((1 + \varepsilon)C(\varepsilon)) \leq A(C(\delta))$, where $\delta = \varepsilon(1 + \varepsilon + \sup_{x \in C} |x|)$. Thus, we can obtain the desired upper bound simply by decreasing $\varepsilon$, to make $\delta$ as small as we want. To prove the upper bound of $A((1 + \varepsilon)C(\varepsilon))$, we will use the dual semidefinite program given in Proposition 3.7, and we will relate it to $A^{\text{top}}(C)^*$ by discretizing space using a suitable simplicial complex.

Let $Y$ be a finite, pure simplicial $d$-complex such that $C \subseteq Y \subseteq C(\varepsilon)$ and all simplices in $Y$ have diameter less than $\varepsilon$, as in Lemma 5.5, and let $Y_0$ be the set of vertices of simplices in $Y$. Each point $x$ in $Y$ can be written using barycentric
coordinates as \( x = \sum_i a_i x_i \) with \( \{x_i\}_{i=0}^d \subseteq Y_0 \) defining a top-dimensional simplex in \( Y \), \( \sum_i a_i = 1 \), and \( a_i \geq 0 \). Moreover, such an expression is unique in the sense that two such expressions for \( x \) yield simplices that intersect in a face containing \( x \), and \( a_i = 0 \) for vertices \( x_i \) not contained in the closure of that face.

Now given a positive semidefinite kernel \( K \) on \( Y_0 \), define a kernel \( L \) on all of \( Y \) by setting

\[
L(x, y) = L \left( \sum_{i=0}^d a_i x_i, \sum_{j=0}^d b_j y_j \right) = \sum_{i,j=0}^d a_i b_j K(x_i, y_j)
\]

given barycentric coordinates \( x = \sum_i a_i x_i \) and \( y = \sum_j b_j y_j \) as above. The uniqueness of these coordinates show that \( L \) is well defined and continuous. To show that \( L \) is positive semidefinite, we must show that

\[
\sum_{m,n} w_m w_n L(x_m, x_n) \geq 0
\]

for each finite set of points \( x_n \) in \( Y \) with weights \( w_n \). If we use barycentric coordinates \( x_n = \sum_i a_{i,n} x_{i,n} \), then

\[
\sum_{m,n} w_m w_n L(x_m, x_n) = \sum_{m,n,i,j} w_m w_n a_{i,m} a_{j,n} K(x_{i,m}, x_{j,n}),
\]

which is indeed nonnegative because \( K \) is positive semidefinite. Thus, \( L \) is a continuous, positive semidefinite kernel on \( Y \). The remaining argument will split into cases, depending on whether we are analyzing \( \vartheta' \), \( \vartheta \), or \( \vartheta^+ \).

First, suppose \( K \) is a feasible solution for \( \vartheta'(Y_0)^* \), with \( K(x, x) = t - 1 \) for all \( x \). Then \( L(x, y) \) is a convex combination of values \( K(x_i, y_j) \) with \( \max(\{|x-x_i|, |y-y_j|\}) \leq \varepsilon \) and therefore \( L(x, y) \leq -1 \) for \( |x-y| \geq 2 + 2\varepsilon \). Furthermore, \( K \) takes its maximum value on the diagonal, because it is a positive semidefinite kernel, and thus \( L(x, x) \leq t - 1 \). This construction shows that

\[
\vartheta'^{\text{top}} \left( \frac{Y}{1 + \varepsilon} \right)^* \leq \vartheta'(Y_0)^* = \vartheta'(Y_0) \leq \vartheta'(C(\varepsilon)),
\]

where the rescaling by a factor of \( 1 + \varepsilon \) takes care of the \( \varepsilon \) in the inequality \( |x-y| \geq 2 + 2\varepsilon \) above, and rescaling space implies that

\[
\vartheta'^{\text{top}}(Y)^* \leq \vartheta'((1 + \varepsilon)C(\varepsilon)).
\]

By restricting \( L \) to \( C \), we conclude that \( \vartheta'^{\text{top}}(C)^* \leq \vartheta'((1 + \varepsilon)C(\varepsilon)) \), which concludes the proof.

Suppose instead that \( K \) is a feasible solution for \( \vartheta(Y_0)^* \). Then by the same reasoning, \( L(x, y) = -1 \) for \( |x - y| \geq 2 + 2\varepsilon \), so \( \vartheta^{\text{top}}(C)^* \leq \vartheta((1 + \varepsilon)C(\varepsilon)) \).

Finally, suppose \( K \) is a feasible solution for \( \vartheta^+(Y_0)^* \). Then \( L(x, y) \geq -1 \) everywhere since it is a convex combination of \( K(x_i, y_j) \geq -1 \), so \( \vartheta^{+\text{top}}(C)^* \leq \vartheta((1 + \varepsilon)C(\varepsilon)) \).

\[\Box\]

Our strategy for computing the Euclidean limits will be to compute the corresponding Euclidean limits of the topological analogues. Even though we do not know whether they are packing bound functions, their limits are still well defined. The following lemma shows that the Euclidean limits of the discrete and topological versions of a sandwich function are the same.
Lemma 5.6. Let \( A \) be a packing bound function, and let \( A^{\text{top}} \) be a real-valued function defined on compact subsets of Euclidean space and satisfying
\[
A(C) \leq A^{\text{top}}(C) \leq A^{\text{top}}(C)^* \leq A(C(\varepsilon))
\]
whenever \( \varepsilon > 0 \). Letting \( I^n = [0,1]^n \) be the unit cube in \( \mathbb{R}^n \) as usual,
\[
\lim_{r \to \infty} \frac{A^{\text{top}}(rI^n)}{r^n} = \lim_{r \to \infty} \frac{A^{\text{top}}(rI^n)^*}{r^n} = \delta_{A,n}.
\]

Proof. This follows immediately by noting that
\[
A(rI^n) \leq A^{\text{top}}(rI^n) \leq A^{\text{top}}(rI^n)^* \leq A((r+1)I^n).
\]

We are ready for the main result of this section, an explicit description of the Euclidean limits of \( \vartheta' \), \( \vartheta \), and \( \vartheta^+ \).

Theorem 5.7. The quantity \( \mathcal{L}_n(B_1^n)\delta_{\vartheta',n} \) is the linear programming bound for the sphere packing density in \( \mathbb{R}^n \), while \( \mathcal{L}_n(B_1^n)\delta_{\vartheta,n} = \mathcal{L}_n(B_1^n)\delta_{\vartheta^+,n} = 1 \).

The key step in the proof of the theorem is the following lemma. Recall that a function \( f : \mathbb{R}^n \to \mathbb{R} \) is positive semidefinite if the kernel \( K \) defined by \( K(x,y) = f(x-y) \) is positive semidefinite. In particular, \( f \) must be an even function.

Lemma 5.8. The Euclidean limits of \( \vartheta' \), \( \vartheta \), and \( \vartheta^+ \) are characterized as follows:

1. \( \delta_{\vartheta',n} \) is the infimum of \( f(0)/\hat{f}(0) \) over all continuous, integrable, positive semidefinite functions \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( f(x) \leq 0 \) for \( |x| \geq 2 \) and \( \hat{f}(0) > 0 \),
2. \( \delta_{\vartheta,n} \) is the infimum of \( f(0)/\hat{f}(0) \) over all feasible solutions to \( \delta_{\vartheta',n} \) that additionally satisfy \( f(x) = 0 \) for \( |x| \geq 2 \), and
3. \( \delta_{\vartheta^+,n} \) is the infimum of \( f(0)/\hat{f}(0) \) over all feasible solutions to \( \delta_{\vartheta,n} \) that additionally satisfy \( f(x) \geq 0 \) for \( |x| \leq 2 \).

Note that without loss of generality, we can assume these functions \( f \) are radial functions (by averaging over all rotations), because the constraints and objective functions are radially symmetric. Furthermore, we can assume that \( f \) has compact support; this is automatic for \( \vartheta \) and \( \vartheta^+ \), and can be proved as follows for \( \vartheta' \) by mollifying a feasible function \( f \). The convolution \( 1_{B_{R/2}^n}(0)*1_{B_{R/2}^n}(0) \) is a continuous function supported in \( B_R^2(0) \), and it is positive semidefinite since its Fourier transform is \( \left( \mathbb{1}_{B_{R/2}^n}(0) \right)^2 \). If we normalize it by setting \( g_R := \left( 1_{B_{R/2}^n}(0)*1_{B_{R/2}^n}(0) \right)/\mathcal{L}_n(B_R^n(0)) \) so that \( g_R(0) = 1 \), then \( g_R \) converges pointwise to \( 1 \) everywhere as \( R \to \infty \). In particular, the product \( f_R := f \cdot g_R \) is continuous and supported in \( B_R^n(0) \), it is positive semidefinite by the Schur product theorem (Theorem 7.5.3 in [30]), and \( \lim_{R \to \infty} \hat{f}_R(0) = \hat{f}(0) \) by dominated convergence. By replacing \( f \) with \( f_R \), we can come arbitrarily close to the ratio \( f(0)/\hat{f}(0) \) using compactly supported functions.

Proof. We will apply Lemma 5.6 and Proposition 5.3 throughout. Let \( K \) be a continuous, positive semidefinite kernel on \( rI^n \subseteq \mathbb{R}^n \), such that \( K \) is a feasible solution for one of the dual topological bounds. We extend it by \( 0 \) to give a positive semidefinite kernel on \( \mathbb{R}^n \). Define \( \tilde{f} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) by the formula
\[
\tilde{f}(x,y) = \frac{1}{r^{2n}} \int_{\mathbb{R}^n} K(x+z,y+z) + 1_{rI^n \times rI^n}(x+z,y+z) \, dz.
\]
Then \( \tilde{f} \) is translation invariant (i.e., \( \tilde{f}(x+t, y+t) = \tilde{f}(x, y) \)) and compactly supported, and it is a continuous function because translating integrable functions is continuous under the \( L^1 \) norm. Thus, the function \( f: \mathbb{R}^n \to \mathbb{R} \) given by \( \tilde{f}(x, y) = f(x - y) \) is well defined and a continuous, compactly supported function. The function \( \tilde{f} \) is a positive semidefinite kernel because it is an integral of such kernels, and so \( f \) is positive semidefinite by definition. We have

\[
\hat{f}(0) = r^{-2n} \int_{\mathbb{R}^n \times \mathbb{R}^n} K(x, y) \, dx \, dy + 1 \geq 1,
\]

and if \( K(x) = t - 1 \) for all \( x \), then

\[
f(0) = t \cdot \frac{r}{n}.
\]

Now suppose \( K \) is feasible for \( \vartheta^{r, \text{top}}(rI^n)^* \). Then \( \tilde{f}(x, y) \leq 0 \) for \( |x - y| \geq 2 \), because the integrand defining \( \tilde{f}(x, y) \) cannot be positive in this case. Thus, the infimum over \( f \) satisfying the conditions in the lemma statement is bounded above by \( r^{-n} \vartheta^{r, \text{top}}(rI^n)^* \).

Finally, suppose \( K \) is feasible for \( \vartheta^{+, \text{top}}(rI^n)^* \). Then \( \tilde{f}(x, y) = 0 \) for \( |x - y| \leq 2 \), because the integrand defining \( \tilde{f}(x, y) \) is nonnegative. Therefore, we conclude one inequality for each of the statements in the lemma.

For the other direction, suppose \( f: \mathbb{R}^n \to \mathbb{R} \) is continuous, positive semidefinite, and integrable, with \( \hat{f}(0) > 0 \). As noted after the lemma statement, we can also assume that \( \text{supp}(f) \subseteq B^n_R(0) \) for some \( R \).

It is not hard to show that

\[
\lim_{r \to \infty} \frac{\int_{rI^n \times rI^n} f(x-y) \, dx \, dy}{r^n} = \hat{f}(0).
\]

Fix \( \varepsilon > 0 \), and let \( r \) be large enough that

\[
(5.1) \quad \int_{rI^n \times rI^n} f(x-y) \, dx \, dy \leq r^n (1 + \varepsilon) \hat{f}(0).
\]

and \( r > 2R \). Let \( J^n = [R, r-R]_n \subseteq rI^n \), and define \( K: J^n \times J^n \to \mathbb{R} \) by

\[
K(x, y) = f(x-y) - \frac{1}{r^n} \hat{f}(0).
\]

We claim that \( K \) is positive semidefinite. Let \( S \) be a finite subset of \( J^n \), and let \( w_x \) be real weights for \( x \in S \). It suffices to show that

\[
(5.2) \quad \sum_{x, y \in S} w_x w_y K(x, y) \geq 0.
\]

To do so, we define a signed measure \( \nu \) by

\[
\sum_{x \in S} w_x \delta_x - \frac{\sum_{x \in S} w_x}{r^n} \mu_r,
\]

where \( \delta_x \) is a unit mass placed at the point \( x \) and \( \mu_r \) is the Lebesgue measure on \( rI^n \). Positive semidefiniteness of \( f \) implies

\[
\int\int f(x-y) \, d\nu(x) \, d\nu(y) \geq 0,
\]
and the left side of this inequality is equal to
\[
\sum_{x,y \in S} w_x w_y f(x - y) + \left(\frac{\sum_{x \in S} w_x}{r^{2n}}\right)^2 \int_{r I^n \times r I^n} f(x - y) \, dx \, dy \\
- \frac{2}{\rho^n} \left(\sum_{x \in S} w_x\right) \sum_{x \in S} w_x \int_{r I^n} f(x - y) \, dy.
\]
(5.3)

Because \( S \subseteq [R, r - R]^n \) and \( \text{supp}(f) \subseteq B_R^n(0) \),
\[
\int_{r I^n} f(x - y) \, dy = \hat{f}(0)
\]
for each \( x \in S \). If we substitute this identity and (5.1) into (5.3), we conclude that
\[
\sum_{x,y \in S} w_x w_y f(x - y) - \left(\sum_{x \in S} w_x\right)^2 \frac{1 - \varepsilon}{\rho^n} \hat{f}(0) \geq 0,
\]
which is equivalent to (5.2). Thus, \( K \) is positive semidefinite.

Because \( \hat{f}(0) > 0 \), we can rescale \( f \) so that
\[
\hat{f}(0) = \frac{r^n}{1 - \varepsilon}.
\]
Then \( K(x, y) = f(x - y) - 1 \) is positive semidefinite on \( J^n \), and in particular \( K(x, x) = f(0) - 1 \).

Now suppose that \( f \) satisfies \( f(x) \leq 0 \) for \( |x| \geq 2 \), as in the case of \( \vartheta' \). Then \( K(x, y) \leq -1 \) for \( |x - y| \geq 2 \), so \( K \) is feasible for \( \vartheta', \text{top}(J^n)^* \). The bound we get on \( \vartheta', \text{top}(J^n)^* \) is
\[
f(0) = \frac{r^n}{1 - \varepsilon} \frac{f(0)}{\hat{f}(0)}.
\]
Dividing by \( (r - 2R)^n \) and letting \( r \to \infty \) and then \( \varepsilon \to 0 \) allows us to conclude that
\[
\lim_{r \to \infty} \frac{\vartheta', \text{top}((r - 2R)I^n)^*}{(r - 2R)^n} \leq \frac{f(0)}{\hat{f}(0)}
\]
and the left side is equal to \( \delta_{\vartheta', n} \). This completes the proof for the case of \( \vartheta' \).

Suppose instead that \( f(x) = 0 \) for \( |x| \geq 2 \), as in the case of \( \vartheta \). Then \( K(x, y) = -1 \) for \( |x - y| \geq 2 \), and so \( K \) is feasible for \( \vartheta^{\text{top}}(J^n)^* \). By the same reasoning as above, we conclude that
\[
\delta_{\vartheta, n} \leq \frac{f(0)}{\hat{f}(0)}.
\]
Finally, suppose that \( f(x) \geq 0 \) for \( |x| \leq 2 \), as in the case of \( \vartheta^+ \). Then \( K(x, y) \geq -1 \) for \( |x - y| \leq 2 \), and so \( K \) is feasible for \( \vartheta^+, \text{top}(J^n)^* \). We conclude that
\[
\delta_{\vartheta^+, n} \leq \frac{f(0)}{\hat{f}(0)},
\]
which completes the proof. \( \square \)

Proof of Theorem 5.7. The lemma directly gives the statement for \( \delta_{\vartheta', n} \) and the linear programming bound. All that remains is to compute the optimal solutions to the following problems:
(1) Minimize \( f(0) / \hat{f}(0) \) over all continuous, positive semidefinite functions \( f \) such that \( f(x) = 0 \) for \( |x| \geq 2 \) and \( \hat{f}(0) > 0 \).

(2) Minimize \( f(0) / \hat{f}(0) \) over all continuous, positive semidefinite functions \( f \) such that \( f(x) \geq 0 \) for \( |x| \leq 2 \), \( f(x) = 0 \) for \( |x| \geq 2 \), and \( \hat{f}(0) > 0 \).

These problems were solved by Gorbachev [22], and Siegel had given the same answer to essentially the same problems in [41]. For the convenience of the reader, we will sketch Gorbachev’s proof.

The optimal function \( f \) is the convolution
\[
 f = 1_{B^n_1(0)} * 1_{B^n_1(0)}
\]
and the objective is \( f(0) / \hat{f}(0) = L_n(B^n_1) - 1 \). The proof of optimality is a consequence of the quadrature formula of Ben Ghanem and Frappier [6], which incidentally also shows that the Levenshtein bound is optimal among certain band-limited solutions to the linear programming bound problem [21, 11]. Specifically, we use the \( p = 0 \) case of Lemma 4 in [6], under the minimal hypotheses established by Grozev and Rahman [24] (which are not stated explicitly for this identity in [6] but follow from the same proof). For any continuous, radial function \( f : \mathbb{R}^n \to \mathbb{R} \) supported on \( B^n_2(0) \) whose Fourier transform is integrable, the formula says that
\[
 f(0) = \frac{1}{L_n(B^n_{r/2})} \hat{f}(0) + \sum_{m=1}^{\infty} \alpha_m \hat{f}
\left( \frac{\lambda_m}{\pi r} \right),
\]
where the node points \( \lambda_m \) are the positive roots of the Bessel function \( J_{n/2} \) and the coefficients \( \alpha_m \) are certain explicit quantities with \( \alpha_m > 0 \). Any feasible solution \( f \) to problems (1) or (2) above satisfies these hypotheses with \( r = 2 \) and has \( \hat{f} \geq 0 \) (see, for example, Corollary 1.26 in Chapter I of [43] for why \( \hat{f} \) is integrable). Since the infinite sum is nonnegative, we conclude that
\[
 \frac{f(0)}{\hat{f}(0)} \geq L_n(B^n_1)^{-1}.
\]

Remark 5.9. In the one-dimensional case,
\[
 1 = L_1(B^1_1) \delta_{\text{pack},1} \leq L_1(B^1_1) \delta_{\theta^{(1)},1} \leq L_1(B^1_1) \delta_{\text{cov},1} = 1.
\]

In other words, the fact that sphere covering and sphere packing have the same density in \( \mathbb{R}^1 \) implies that the linear programming bound must be sharp in that case. Of course this fact can be proved directly by exhibiting a closed-form auxiliary function, but it is interesting to prove it without the need to construct any explicit auxiliary function.

6. The Lasserre hierarchy for sphere packing

The Lasserre hierarchy [36] is an important family of semidefinite relaxations of the independence number; this hierarchy starts with \( \theta' \) and extends it to successively stronger bounds, which converge to the exact independence number. Based on foundations laid by Laurent [37] and by de Laat and Vallentin [35], in this section we show that the Lasserre hierarchy consists of sandwich functions, and we write their Euclidean limits as optimization problems. The net result is a generalization of the linear programming bound to a hierarchy of bounds that converge to the exact sphere packing density.
6.1. Review of bounds for compact spaces. A topological packing graph is a graph and a Hausdorff topological space such that every finite clique is contained in an open clique. Equivalently, each vertex and each edge is contained in an open clique. Every graph is a topological packing graph with the discrete topology, and every compact topological packing graph has finite independence number (and furthermore finite clique covering number, which is an even stronger assertion). The edges in a topological packing graph indicate pairs of vertices that are too close together to appear in the same packing, and independent sets correspond to packings.

The Lasserre hierarchy on compact topological packing graphs was introduced by de Laat and Vallentin [35]. We begin by reviewing this hierarchy.

Definition 6.1. We set the following notation:

1. For a topological space $V$, let $C(V)$ be the set of continuous functions from $V$ to $\mathbb{R}$, let $C_0(V) \subseteq C(V)$ consist of the functions that vanish at infinity, and let $C_c(V) \subseteq C(V)$ consist of those with compact support.

2. For a locally compact Hausdorff space $V$, let $\mathcal{M}_+(V)$ denote the set of Radon measures on $V$ (i.e., regular Borel measures), and let $\mathcal{M}_-(V)$ denote the set of finite, regular signed Borel measures on $V$. By the Riesz representation theorem, $\mathcal{M}_+(V) = C_0(V)^*$ under the pairing given by integration (see, for example [10, Theorem 7.3.6]).

3. Given a compact topological packing graph $G$, let $I_{t,G}$ be the set of independent sets in $V(G)$ of size exactly equal to $t$, with the topology given as a subset of the quotient of $V(G)$ under the map $(v_1, \ldots, v_t) \mapsto \{v_1, \ldots, v_t\}$. In particular, if $t > \alpha(G)$ then this set is empty, and $I_{\infty,G} = \{\emptyset\}$.

4. Given a compact topological packing graph $G$, let $I_{t,G}$ be the set of independent sets of size at most $t$ in $G$, with the topology given by the disjoint union of $I_{k,G}$ for $k = 0, \ldots, t$.

We often write $I_t$ instead of $I_{t,G}$ when the context is clear. We can think of $I_t$ as a moduli space of packings on $G$ by independent sets of size at most $t$, and we will often use the covariant functoriality of these moduli spaces for morphisms in the category $\mathcal{G}$. When $G$ is a compact topological packing graph, the space $I_t$ is compact by Lemma 1 in [35], and we will make frequent use of the duality between $C(I_t)$ and $\mathcal{M}_\pm(I_t)$. A key role will be played by an operator

$$A_t: C(I_t \times I_t) \to C(I_{2t}).$$

For $f \in C(I_t \times I_t)$ and $S \subseteq I_{2t}$, the function $A_t f$ is defined by

$$A_t f(S) = \sum_{J \cup J' = S} f(J, J').$$

We define the operator $A_t^*: \mathcal{M}_\pm(I_{2t}) \to \mathcal{M}_\pm(I_t \times I_t)$ to be the adjoint of $A_t$. Specifically, for $\mu \in \mathcal{M}_\pm(I_{2t})$, the signed measure $A_t^* \mu$ is characterized by

$$\left(\int_{I_t \times I_t} f(J, J') dA_t^* \mu(J, J') = \int_{I_{2t}} \sum_{J \cup J' = S} f(J, J') d\mu(S) \right)$$

for all $f \in C(I_t \times I_t)$. 

There are corresponding notions of positive semidefiniteness for kernels and measures on $I_t$. A kernel $K: I_t \times I_t \to \mathbb{R}$ is positive semidefinite if for every finite subset $S \subseteq I_t$, the matrix $(K(J, J'))_{J, J' \in S}$ is positive semidefinite. Equivalently, $K(J, J') = K(J', J)$ and for every finite subset $S \subseteq I_t$ and any weights $w_J \in \mathbb{R}$ for $J \in S$, 

$$
\sum_{J, J' \in S} w_J w_{J'} K(J, J') \geq 0.
$$

A signed measure $\mu$ on $I_t \times I_t$ is positive semidefinite if every continuous, positive semidefinite kernel $K: I_t \times I_t \to \mathbb{R}$ satisfies 

$$
\iint_{I_t \times I_t} K(J, J') \, d\mu(J, J') \geq 0.
$$

Note that for compact topological packing graphs, $I_t$ is compact. In more general cases we use only kernels with compact support. By Mercer’s theorem [42, Theorem 3.11.9], an equivalent condition is that for every continuous function $f: I_t \to \mathbb{R}$, 

$$
\mu(f \otimes f) := \iint_{I_t \times I_t} f(J) f(J') \, d\mu(J, J') \geq 0.
$$

When $G$ is a finite graph (which must have the discrete topology, because it must be Hausdorff), a symmetric measure $\mu \in \mathcal{M}_+(I_t \times I_t)$ is positive semidefinite if and only if the function $(J, J') \mapsto \mu(\{J\} \times \{J'\})$ is a positive semidefinite kernel, because the cone of positive semidefinite matrices is self-dual.

**Definition 6.2.** For a compact topological packing graph $G$, we define $\lambda^\top_{t}(G)$ to be the supremum of $\lambda(I_{t-1})$ over all $\lambda \in \mathcal{M}_+(I_{2t})$ such that $A^*_t \lambda$ is positive semidefinite as a measure on $I_t \times I_t$ and $\lambda(\emptyset) = 1$. We define $\lambda^\top_{t'}(G)$ the same way, with the additional requirement that $\lambda$ be a positive measure.

**Remark 6.3.** In [35], de Laat and Vallentin use the notation $\lambda_t(G)$ to refer to what we call the hierarchy $\lambda^\top_t(G)$. However, the notation we use here is more consistent with the convention of adding the ‘$'$ for optimization over positive measures, as in the relationship between $\lambda'$ and $\lambda^\top'$. The “top” indicates the dependence on the topology of $G$, along the lines of $\lambda'$ and $\lambda^\top'$.

**Proposition 6.4.** For every compact topological packing graph $G$, $\lambda^\top_t(G)$ and $\lambda^\top_{t'}(G)$ are nonincreasing in $t$ and satisfy 

$$
\lambda^\top_{2t}(G) \leq \lambda^\top_{t'}(G) \leq \lambda^\top_t(G).
$$

**Proof.** To see that these quantities are nonincreasing, consider restricting a measure $\lambda$ on $I_{2(t+1)}$ to $I_{2t}$. If $\lambda$ is feasible for $\lambda^\top_{t+1}(G)$, resp. $\lambda^\top_{t+1}(G)$, then its restriction is feasible for $\lambda^\top_t(G)$, resp. $\lambda^\top_{t'}(G)$. (Note that every continuous, positive semidefinite kernel on $I_t \times I_t$ extends by zero to such a kernel on $I_{t+1} \times I_{t+1}$.)

For the remaining inequality, $\lambda^\top_{t'}(G) \leq \lambda^\top_t(G)$ follows immediately from the definitions, and thus it suffices to prove that $\lambda^\top_{2t}(G) \leq \lambda^\top_t(G)$. Let $\lambda \in \mathcal{M}_+(I_{2t})$ be feasible for $\lambda^\top_{2t}$, and consider its restriction to $I_{2t}$. We will show that this restriction is feasible for $\lambda^\top_t(G)$. Since $A^*_t \lambda$ is positive semidefinite, the restriction of $A^*_t \lambda$ to the diagonal copy of $I_{2t}$ in $I_{2t} \times I_{2t}$ must be nonnegative. However, this restriction is just the restriction of $\lambda$ to $I_{2t}$, which is therefore feasible for $\lambda^\top_{t'}(G)$. \hfill $\square$
We will use the following convergence property, where $\alpha(G)$ is the independence number of $G$ (i.e., the size of the largest packing):

**Proposition 6.5** (de Laat and Vallentin [35]). Let $G$ be a compact topological packing graph. Then

$$\text{las}_{\alpha(G)}^\text{top}(G) = \text{las}_{\alpha(G)}^\text{top}(G) = \alpha(G).$$

**Remark 6.6.** The paper [35, Theorem 2] only gives the statement for $\text{las}_{\alpha(G)}^\text{top}(G)$, but their proof extends word for word to $\text{las}_{\alpha(G)}^\text{top}(G)$.

The empty set is an isolated point in $I_t$, and for some purposes it is useful to omit it. There is an equivalent formulation of $\text{las}_t$ without $\{\emptyset\}$ using the Schur complement. For $\mu$ a measure on $I_t \setminus \{\emptyset\}$, let $(A_t^\emptyset)^* \mu$ be the restriction of $A_t^* \mu$ to $I_t \setminus \{\emptyset\} \times I_t \setminus \{\emptyset\}$, and define the external tensor product $\mu \otimes \nu$ of two measures as the measure taking value $\mu(A)\nu(B)$ on $A \times B$. The following proposition follows from taking the Schur complement:

**Proposition 6.7.** The supremum of $\mu(I_{=1})$ over measures $\mu \in \mathcal{M}_+(I_{2t} \setminus \{\emptyset\})$ such that $(A_t^\emptyset)^* \mu - \mu \otimes \mu$ is positive semidefinite is $\text{las}_{t}^\text{top}(G)$.

Here $\mu \otimes \mu$ refers to a measure on $I_t \setminus \{\emptyset\} \times I_t \setminus \{\emptyset\}$, so that it is comparable with $(A_t^\emptyset)^* \mu$.

**Proof.** Define a new measure $\overline{\mu}$ on $I_{2t}$ that is equal to $\mu$ on $I_{2t} \setminus \{\emptyset\}$ and puts a mass of 1 on the point $\emptyset$. The version of the Schur complement in [35, Lemma 9] shows that if $(A_t^\emptyset)^* \mu - \mu \otimes \mu$ is positive semidefinite, then $A_t^* \overline{\mu}$ is a positive semidefinite measure and so $\overline{\mu}$ is a feasible measure for $\text{las}_{t}^\text{top}(G)$. \hfill $\square$

**Remark 6.8.** One can consider a weakened hierarchy in which the positive semidefiniteness condition above is weakened to the conditions that $(A_t^\emptyset)^* \mu$ is positive semidefinite and that

$$2\mu(I_{=2}) + \mu(I_{=1}) = (A_t^\emptyset)^* \mu(I_{=1} \times I_{=1}) \geq \mu(I_{=1})^2.$$  

This weakening was discussed briefly in [34, equation (3) in Section 2.1]. In particular, for the case $t = 1$, this weakening of the hierarchy is equivalent to $\text{las}_{1}^\text{top}$. Furthermore, de Laat and Vallentin [35] show that $\text{las}_{1}^\text{top}$ is actually equivalent to $\text{las}_{1}^\text{top}^*$.

The convex dual of the optimization problem defining $\text{las}_{t}^\text{top}(G)$ is as follows:

**Definition 6.9.** For a compact topological packing graph $G$, we define $\text{las}_{t}^\text{top}(G)^*$ to be the infimum of $K(\emptyset, \emptyset)$ over continuous kernels $K : I_t \times I_t \to \mathbb{R}$ such that

1. $K$ is positive semidefinite,
2. $A_t K(\{x\}) \leq -1$, and
3. $A_t K(S) \leq 0$ for $S \in I_t$ with $|S| \geq 2$.

By [35, Theorem 1], strong duality holds:

**Theorem 6.10** (Strong duality). Let $G$ be a compact topological packing graph. Then

$$\text{las}_{t}^\text{top}(G) = \text{las}_{t}^\text{top}(G)^*.$$  

Note that the proof of this theorem in the published version of [35] contains a minor gap, which is filled in the arXiv version.
6.2. The Lasserre hierarchy as sandwich functions.

**Definition 6.11.** For a finite graph $H$ and for $A = \text{las}_t$ or $\text{las}'_t$, define $\text{las}_t(G) = \text{las}^\text{top}_t(G)$ and $\text{las}'_t(G) = \text{las}'^\text{top}_t(G)$.

For a discrete graph $G$, possibly infinite but with finite clique covering number, we define $\text{las}_t(G)$, resp. $\text{las}'_t(G)$, to be the supremum over all induced finite subgraphs $H \subseteq G$ of $\text{las}_t(H)$, resp. $\text{las}'_t(H)$.

This definition coincides on finite graphs with $\text{las}'^\text{top}_t(G)$. We do not know whether they agree on all compact topological packing graphs, even for the case $t = 1$ (which is $\vartheta'$). Using Theorem 6.13 below, for compact subsets of $\mathbb{R}^n$, this question can be reformulated purely in terms of $\text{las}'_t$.

**Theorem 6.12.** The functions $G \mapsto \text{las}_t(G)$ and $G \mapsto \text{las}'_t(G)$ are sandwich functions.

**Proof of Theorem 6.12.** Recall that Lemma 3.9 shows that we only need to check the statement on finite graphs. The case when $G$ is a point is clear.

Consider a graph homomorphism from $G$ to $H$, where $G$ and $H$ are finite graphs, and let $f: V(G) \to V(H)$ be the underlying map on vertex sets. Let us also use $f$ for the induced map on independent sets $f: I_{=k,G} \to I_{=k,H}$ for each $k$. If $\mu$ is a measure on $I_{t,G}$, we define the measure $f_*\mu$ on $I_{t,H}$ via

$$f_*\mu(A) = \mu(f^{-1}(A)).$$

First, we show that if $\mu$ is feasible for $\text{las}_t(G)$, then $f_*\mu$ is feasible for $\text{las}_t(H)$ with the same objective function value. If $A^*_t\mu$ is positive semidefinite on $I_{t,G} \times I_{t,G}$, then $A^*_tf_*\mu$ is positive semidefinite on $I_{t,H} \times I_{t,H}$. To see why, we will use (6.1); for finite graphs, it says that

$$A^*_t\mu(J \times J') = \begin{cases} \mu(J \cup J') & \text{if } J \cup J' \in I_{2t,G}, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Let $w_J$ be a set of real weights for $J \in I_{t,H}$. If we set $w_K = w_J$ for $K \in I_{t,G}$ with $f(K) = J$ and $w_K = 0$ otherwise, then

$$\sum_{J, J' \in I_{t,H}} w_Jw_{J'}f_*\mu(J \cup J')$$

$$= \sum_{J, J' \in I_{t,H}} w_Jw_{J'}\mu(K \cup K' : K \in f^{-1}(J), K' \in f^{-1}(J'))$$

$$= \sum_{K, K' \in I_{t,G}} w_Kw_{K'}\mu(K \cup K') \geq 0.$$

This shows that $\text{las}_t(G) \leq \text{las}_t(H)$. Furthermore, if $\mu$ is nonnegative, then $f_*\mu$ is nonnegative, and thus $\text{las}'_t(G) \leq \text{las}'_t(H)$ as well.

All the remains to check is the union axiom. Let $G$ and $H$ be disjoint graphs, and consider $G \sqcup H$. We first want to show that $\text{las}_t(G \sqcup H) \geq \text{las}_t(G) + \text{las}_t(H)$ and similarly for $\text{las}'_t(G \sqcup H)$. Let $\mu_G$ and $\mu_H$ be measures on $I_{2t,G}$ and $I_{2t,H}$. To define a measure $\mu$ on $I_{2t,G\sqcup H}$, it suffices to define the measure of a singleton $\{S\}$ with $S \in I_{2t,G\sqcup H}$. We define this by

$$\mu\{S\} = \mu_G\{S \cap V(G)\}\mu_H\{S \cap V(H)\}.$$
If \( \mu_G(\{\emptyset\}) = \mu_H(\{\emptyset\}) = 1 \), then \( \mu(\{\emptyset\}) = 1 \) immediately and \( \mu(I_1) = \mu_G(I_{1,G}) + \mu_H(I_{1,H}) \). Thus, the objective is additive. Moreover, if \( \mu_G \) and \( \mu_H \) are both nonnegative, then \( \mu \) is as well. It suffices to show that if \( \lambda^*_I \mu_G \) and \( \lambda^*_I \mu_H \) are positive semidefinite, then \( \lambda^*_I \mu \) is. Because \( G \cup H \) is finite and the cone of positive semidefinite matrices is self-dual, \( \lambda^*_I \mu \) is positive semidefinite if and only if the matrix

\[
(\mu(\{J \cup J\}'))_{J,J' \in I_{1,G \cup H}}
\]

is positive semidefinite, where we set \( \mu(\{S\}) = 0 \) if \( S \notin I_{2,G \cup H} \). This matrix is the tensor product of the corresponding matrices for \( \mu_G \) and \( \mu_H \), and the desired conclusion follows.

Finally, we want to show that \( \lambda_s(G \cup H) \leq \lambda_s(G) + \lambda_s(H) \) and similarly for \( \lambda_s' \). Let \( \mu \) be a measure on \( I_{2,G \cup H} \) and consider its restriction \( \mu_G \) to \( I_{2,G} \) and restriction \( \mu_H \) to \( I_{2,H} \). We see that

\[
\mu_G(I_{1,G}) + \mu_H(I_{1,H}) = \mu(I_{1,G \cup H}).
\]

If \( \mu \) is nonnegative, then \( \mu_G \) and \( \mu_H \) are, and if \( \lambda^*_I \mu \) is positive semidefinite, then \( \lambda^*_I \mu_G \) and \( \lambda^*_I \mu_H \) are (they correspond to submatrices of the positive semidefinite matrix for \( G \cup H \)). Finally,

\[
\mu(\{\emptyset\}) = \mu_G(\{\emptyset\}) = \mu_H(\{\emptyset\}).
\]

By starting with a feasible point for \( \lambda_s(G \cup H) \) that is arbitrarily close to the optimum, we conclude that \( \lambda_s'(G \cup H) \leq \lambda_s'(G) + \lambda_s'(H) \), and similarly \( \lambda_s'(G \cup H) \leq \lambda_s'(G) + \lambda_s'(H) \). □

Because these functions are packing bounds functions, we can conclude that a Euclidean limit must exist. In the next section, we formulate the Euclidean limit of \( \lambda_s \) and \( \lambda_s^{'top} \) in terms of a semidefinite program.

### 6.3. The Euclidean limit of the Lasserre hierarchy.

For any subset \( C \subseteq \mathbb{R}^n \), consider the graph with vertex set \( C \) and edges consisting of pairs \( (x, y) \) with \( 0 < d(x, y) < 2 \). If \( C \) is bounded, we can obtain \( \lambda_s'(C) \) by treating the graph as a discrete graph. If \( C \) is compact, then the graph is a compact topological packing graph and we can obtain \( \lambda_s^{'top}(C) \). For any \( C \), we obtain a topological packing graph, and thus the collection \( I_{t,C} \) of independent sets of size at most \( t \) has a topology, but it will not be compact in general. For example, \( I_{t,\mathbb{R}^n} \) admits an action of \( \mathbb{R}^n \) by translation.

Much like the case of \( \vartheta'(C) \), we can interpret \( \lambda_s'(C) \) as a supremum over finitely supported measures coming from finite subgraphs of \( C \). From this point of view, the inequality \( \lambda_s'(C) \leq \lambda_s^{'top}(C) \) is immediate when \( C \) is compact. The following theorem gives a more precise relationship.

**Theorem 6.13.** Let \( C \) be a compact subset of \( \mathbb{R}^d \), and let \( \overline{C(\varepsilon)} \) be the closure of \( C(\varepsilon) \), the \( \varepsilon \)-neighborhood of \( C \). Then

\[
\lambda_s^{'top}(C) = \lim_{\varepsilon \to 0^+} \lambda_s'(C(\varepsilon)) = \lim_{\varepsilon \to 0^+} \lambda_s^{'top}(\overline{C(\varepsilon)}).
\]

**Proof.** First, we will show that

\[
\lim_{\varepsilon \to 0^+} \lambda_s^{'top}(\overline{C(\varepsilon)}) = \lambda_s^{'top}(C).
\]

We will prove both inequalities separately. One inequality is immediate by inclusion of \( C \subseteq C(\varepsilon) \). For the other direction, since \( \overline{C(\varepsilon)} \) is compact, there is a measure
achieving the optimal value of \( \text{las}^{t,\text{top}}_i(C(\varepsilon)) \) for any \( \varepsilon \). Take \( \mu_n \) achieving this for \( \text{las}^{t,\text{top}}_i(C(\varepsilon_n)) \) for a sequence \( \varepsilon_n \to 0 \). This is a sequence of weak*-bounded measures, and by the Banach-Alaoglu theorem there is a convergent subsequence in the weak* topology, which we denote \( \mu_n \to \mu \). We claim that \( \mu \) is feasible for \( \text{las}^{t,\text{top}}_i(C) \). The support of \( \mu \) must be contained in the intersection, and it must be supported on independent sets by the corresponding property for \( \mu_n \). Moreover, it must be positive since \( \mu_n \) are positive measures. To prove positive semidefiniteness for \( A^*_t \mu \), it suffices to check that \( \int f(x)f(y)\,dA^*_t \mu(x,y) \geq 0 \) for every continuous \( f: I_t, C \to \mathbb{R} \), because Mercer’s theorem [42, Theorem 3.11.9] allows us to write any positive semidefinite kernel as an infinite sum of terms of the form \( f(x)f(y) \). Let \( f_\varepsilon \) be a compactly supported, continuous extension of \( f \) to \( I_{t,C(\varepsilon)} \). Using \( f_\varepsilon \), we can write the integral as a limit:

\[
\int f(x)f(y)\,dA^*_t \mu(x,y) = \lim_{n \to \infty} \int f_\varepsilon(x)f_\varepsilon(y)\,dA^*_t \mu_n(x,y).
\]

Therefore \( A^*_t \mu \) is positive semidefinite, and \( \mu \) is feasible for \( \text{las}^{t,\text{top}}_i(C) \). Since the objective is a continuous functional for the weak*-topology, \( \mu \) gives a lower bound of \( \lim_{\varepsilon \to 0^+} \text{las}^{t,\text{top}}_i(C(\varepsilon)) \) for \( \text{las}^{t,\text{top}}_i(C) \).

As a consequence of this argument, note that for compact sets \( C \),

\[
\lim_{\varepsilon \to 0^+} \text{las}^{t,\text{top}}_i((1+\varepsilon)C) = \text{las}^{t,\text{top}}_i(C),
\]

because for every \( \varepsilon > 0 \), there is some \( \varepsilon' > 0 \) such that \((1+\varepsilon')C \) embeds into \( C(\varepsilon) \).

Now, to complete the proof it suffices to prove that

\[
\text{las}^{t,\text{top}}_i(C) \leq \lim_{\varepsilon \to 0^+} \text{las}^{t,\text{top}}_i(C(\varepsilon)).
\]

Let \( Y \) be a geometric simplicial complex such that \( C \subseteq Y \subseteq C(\varepsilon) \) and every simplex has diameter at most \( \varepsilon \), constructed as in Lemma 5.5. We use \( Y_0 \) to denote the vertices of \( Y \) and \( Y_d \) to denote its top-dimensional simplices. For any positive semidefinite kernel on \( I_{t,Y_0} \), we will produce a positive semidefinite kernel on the whole \( I_{t,Y} \) by viewing the simplicial complex as a union of finite elements and by taking convex combinations and extending the kernel by linearity. This is the same idea as in the proof of Proposition 5.3, but with more cumbersome notation because of the use of \( I_t \). Using this approach, we will obtain feasible kernels for \( \text{las}^{t,\text{top}}_i(C)^* \) from those for \( \text{las}^{t,\text{top}}_i(C(\varepsilon))^* \), specifically those using the subset \( Y_0 \).

Let \( K \) be a positive semidefinite kernel on \( I_{t,Y_0} \). Then we can define a positive semidefinite kernel \( L \) on \( I_{t,Y} \) as follows. Given any point in \( Y \), we can randomly round it to a vertex in \( Y_0 \) by using barycentric coordinates: if the point is \( y = \lambda_0 y_0 + \cdots + \lambda_d y_d \) with \( \lambda_i \geq 0 \), \( \sum_i \lambda_i = 1 \), and \( \{y_0, \ldots, y_d\} \) a top-dimensional simplex, then we round \( y \) to \( y_i \) with probability \( \lambda_i \). This process is well defined, because the only way \( y \) can be in several top-dimensional simplices is if all the weights not coming from their intersection vanish. Similarly, we can round an independent set \( J \) by rounding each point in it independently. We will denote the rounded version of \( J \) by the random variable \( r(J) \). One subtlety is that \( r(J) \) may not be an independent set, because two points at distance less than \( 2 + 2\varepsilon \) may round to points at distance less than 2 (recall that the simplices have diameter at most \( \varepsilon \)). That will not be a problem, since we can extend \( K \) by zero to obtain a positive semidefinite kernel on arbitrary sets of size at most \( t \), not just independent sets.
Using this notion of rounding, we define \( L(J, J') \) as the expected value

\[
\mathbb{E}K(r(J), r(J'))
\]

of \( K(r(J), r(J')) \) when we round each point in \( J \cup J' \) independently. Then \( L \) is a continuous function on \( \mathcal{I}_{t,Y} \times \mathcal{I}_{t,Y} \). To show that it is a positive semidefinite kernel, we must show that for all weights \( w_J \in \mathbb{R} \) for \( J \in \mathcal{I}_{t,Y} \) that vanish for all but finitely many \( J \),

\[
\sum_{J,J' \in \mathcal{I}_{t,Y}} w_J w_{J'} L(J, J') \geq 0.
\]

To prove this inequality, consider randomly rounding the points of \( Y \) independently. Then

\[
\sum_{J,J' \in \mathcal{I}_{t,Y}} w_J w_{J'} L(J, J') = \mathbb{E} \sum_{J,J' \in \mathcal{I}_{t,Y}} w_J w_{J'} K(r(J), r(J')),
\]

which is nonnegative because \( K \) is a positive semidefinite kernel.

All the remains to check is the conditions on \( A_tL \) for a feasible kernel. If \( A_t K(\{x\}) \leq -1 \) for all \( x \in Y_0 \), then \( A_t L(\{x\}) \leq -1 \) for all \( x \in Y \), because \( A_t L(\{x\}) = \mathbb{E} A_t K(r(\{x\})) \). However, the case of \( A_t L(S) \) with \( |S| \geq 2 \) is slightly more subtle. The issue is that \( r(S) \) may not be an independent set even if \( S \) is. However, if the minimal distance between points in \( S \) is at least \( 2 + 2 \varepsilon \), then \( r(S) \) must always be an independent set. In that case,

\[
A_t L(S) = \sum_{J,J' \subseteq S, J \cup J' = S} L(J, J') = \mathbb{E} \sum_{J,J' \subseteq r(S), J \cup J' = r(S)} K(J, J') \leq 0
\]

if \( A_t K(S) \leq 0 \). In other words, we obtain a feasible kernel for \((1 + \varepsilon)^{-1} C\), rather than \( C \).

After rescaling space by a factor of \( 1 + \varepsilon \), we conclude that

\[
las_t^{\text{top}}(C) \leq las_t^t((1 + \varepsilon) Y_0) \leq las_t^t((1 + \varepsilon) C(\varepsilon)).
\]

For any \( \varepsilon' \), we can choose \( \varepsilon \) so that \((1 + \varepsilon) C(\varepsilon) \subseteq C(\varepsilon') \) and letting \( \varepsilon' \to 0 \) proves the inequality, and hence the result. \( \square \)

**Corollary 6.14.** Let \( I^n \) be the unit cube in \( \mathbb{R}^n \). Then

\[
\lim_{r \to \infty} \frac{\text{las}_t^{\text{top}}(r I^n)}{r^n} = \lim_{r \to \infty} \frac{\text{las}_t(r I^n)}{r^n}.
\]

**Proof.** This corollary follows from Theorem 6.13 together with the fact that \( I^n(\varepsilon) \) embeds into \((1 + 2\varepsilon) I^n\). \( \square \)

We are now ready to give the Euclidean limit of the Lasserre hierarchy. Let \( \mathbb{R}^n \) act on Radon measures on \( \mathcal{I}_{2t,\mathbb{R}^n} \) by translation, and consider translation-invariant measures. Any translation-invariant measure \( \mu \) will restrict to \( \mathcal{I}_{1,\mathbb{R}^n} \) as some multiple of the Lebesgue measure, and we define \( \tilde{\mu}(0) \) to be this multiple. That is,

\[
\mu|_{\mathcal{I}_{1,\mathbb{R}^n}}(C) = \tilde{\mu}(0) \mathcal{L}_n(C)
\]

for Borel sets \( C \subseteq \mathbb{R}^n \).

**Definition 6.15.** Let \( \text{las}_t^t(\mathbb{R}^n) \) be the supremum of \( \tilde{\mu}(0) \) over all translation-invariant Radon measures \( \mu \) on \( \mathcal{I}_{2t,\mathbb{R}^n} \) such that

1. \( A_t^t \mu \) is positive semidefinite as a measure on \( \mathcal{I}_t \times \mathcal{I}_t \),
2. \( \mu \) is nonnegative, and
(3) \( \mu(\emptyset) = 1 \).

We call such a measure \( \mu \) a correlation measure of order \( 2t \) with center density \( \tilde{\mu}(0) \).

Let \( \mathcal{P} \) be a periodic packing in \( \mathbb{R}^n \), i.e., the union of finitely many translates of a lattice \( \Lambda \) such that no two points of \( \mathcal{P} \) are closer than distance \( 2 \) apart, and let \( D \) be a fundamental parallelootope for \( \Lambda \). To obtain a correlation measure from \( \mathcal{P} \), we define \( \mu_\mathcal{P} \) to be the Radon measure on \( I_{2t,\mathbb{R}^n} \) characterized by

\[
\int_{I_{2t,\mathbb{R}^n}} f \, d\mu_\mathcal{P} = \frac{1}{\mathcal{L}_n(D)} \int_{D} \sum_{S \subseteq \mathcal{P} + v, |S| \leq 2t} f(S) \, dv
\]

for compactly supported, continuous functions \( f : I_{2t,\mathbb{R}^n} \to \mathbb{R} \). In other words, \( \mu_\mathcal{P} \mid_{I_{-k}} \) is essential the correlation function of order \( k \) for \( \mathcal{P} \). The purpose of averaging over \( v \in D \) is to make \( \mu_\mathcal{P} \) translation-invariant. The center density \( \tilde{\mu}_\mathcal{P}(0) \) is the usual center density of the sphere packing \( \mathcal{P} \), i.e., \( N/\mathcal{L}_n(D) \) if \( \mathcal{P} \) consists of \( N \) translates of \( \Lambda \).

To show that \( \mu_\mathcal{P} \) is a correlation measure, all that remains is to prove that \( A_t^* \mu_\mathcal{P} \) is positive semidefinite. Let \( K : I_t \times I_t \to \mathbb{R} \) be a continuous, positive semidefinite kernel with compact support. Then

\[
\int_{I_{2t,\mathbb{R}^n} \times I_{2t,\mathbb{R}^n}} K \, dA_t^* \mu_\mathcal{P} = \int_{I_{2t,\mathbb{R}^n}} A_t K \, d\mu_\mathcal{P}
\]

\[
= \frac{1}{\mathcal{L}_n(D)} \int_{D} \sum_{S \subseteq \mathcal{P} + v, |S| \leq 2t} \sum_{J,J' \in I_t, j,j' = S} K(J,J') \, dv
\]

\[
= \frac{1}{\mathcal{L}_n(D)} \int_{D} \sum_{J,J' \in I_t, J,J' \subseteq \mathcal{P} + v} K(J,J') \, dv,
\]

and

\[
\sum_{J,J' \in I_t, J,J' \subseteq \mathcal{P} + v} K(J,J') \geq 0
\]

because \( K \) is positive semidefinite. (Note that this is a finite sum, because \( K \) has compact support.)

The main result of this section is that the quantity \( \text{las}_t'(\mathbb{R}^n) \) is the Euclidean limit of the packing bound function \( \text{las}_t' \).

**Theorem 6.16.** For each \( n \) and \( t \),

\[ \text{las}_t'(\mathbb{R}^n) = \delta_{\text{las}_t'}, \]

**Proof.** Because of Corollary 6.14, it suffices to check that

\[
\lim_{r \to \infty} \frac{\text{las}_t^{*,\text{top}}(rI^n)}{r^n} = \text{las}_t'(\mathbb{R}^n).
\]

If \( \mu \) is a correlation measure of order \( 2t \) (i.e., feasible for \( \text{las}_t'(\mathbb{R}^n) \)), then restricting \( \mu \) to \( I_{2t,rI^n} \) gives a feasible measure for \( \text{las}_t^{*,\text{top}}(rI^n) \) with objective \( r^n \text{las}_t'(\mathbb{R}^n) \). This shows that

\[
\frac{\text{las}_t^{*,\text{top}}(rI^n)}{r^n} \geq \text{las}_t'(\mathbb{R}^n)
\]

for any \( r \).
To prove the other direction of the inequality, we use a tiling construction. In this case, we use \( \text{las}_t \) instead of \( \text{las}_t^{\text{top}} \). Equivalently, we restrict our attention to feasible measures with finite support. Given such a measure for \( rI^n \), we can extend it to \( \mathbb{R}^n \) as follows.

Consider a tiling \( \mathbb{R}^n = \bigcup_{i \in T} T_i \), where each tile \( T_i \) is congruent to \((r + 2)I^n\) and the tiles are translates under the cubic lattice \((r + 2)\mathbb{Z}^n \subseteq \mathbb{R}^n \). Let \( Q_i \subseteq T_i \) be the centered copy of \( rI^n \) in \( T_i \) (with \( g_i : Q_i \rightarrow rI^n \) performing this identification), so the induced subgraph of \( Q := \bigcup_{i \in T} Q_i \) is a disjoint union over \( i \in T \). Let \( \mu_{rI^n} \) be a finitely supported measure on \( I_{2rI^n} \) such that \( \mu_{rI^n}(\{\emptyset\}) = 1 \). We now define a measure \( \mu_Q \) on \( I_{2r, Q} \) as follows. For any \( S \in I_{2r, Q} \), let

\[
\mu_Q(S) = \prod_{i \in T} \mu_{rI^n}(\{g_i(S \cap Q_i)\}).
\]

This equation directly defines \( \mu_Q \) for all single-element subsets of \( I_{2r, Q} \), and for all Borel subsets as an atomic measure. Note in particular that for each \( S \), all but finitely many factors in the infinite product are 1. Furthermore, for any compact set \( C \subseteq Q \), the measure \( \mu_Q|_{I_{2r, C}} \) has finite support. We extend \( \mu_Q \) to a measure on \( I_{2r, R^n} \) by zero.

Next, we show that \( A_t^* \mu_Q \) is positive semidefinite. In other words,

\[
\int \int_{I_{2r, R^n} \times I_{2r, R^n}} K dA_t^* \mu_Q \geq 0
\]

for every compactly supported, continuous, positive semidefinite kernel \( K : I_{2r, R^n} \times I_{2r, R^n} \rightarrow \mathbb{R} \). Every compact subset of \( I_{2r, R^n} \) is contained in \( I_{2r, C} \) for some compact subset \( C \) of \( \mathbb{R}^n \), and thus only finitely many cubes \( Q_i \) play a role for any given \( K \). Because \( \mu_{rI^n} \) has finite support, what we need to check is an assertion about positive semidefinite matrices. Specifically, the relevant matrix for \( A_t^* \mu_Q \) is a tensor power of that for \( A_{t'} \mu_{rI^n} \), just as in the verification of the union axiom in the proof of Theorem 6.12, and positive semidefiniteness is therefore preserved.

All that remains is to average \( \mu_Q \) under the action of \( \mathbb{R}^n \) by translation. Because \( \mu \) is already invariant under translation by a lattice \((r + 2)\mathbb{Z}^n \), we can average over the action of the quotient torus, which is a compact group. Therefore, the average is well defined.

By construction, if \( \mu_{rI^n} \) is feasible for \( \text{las}_t'(rI^n) \), then the result \( \mu \) after averaging is a translation-invariant measure on \( I_{2r, R^n} \) that is feasible for \( \text{las}_t'(\mathbb{R}^n) \), with objective

\[
\frac{\mu_{rI^n}(I_{r=rI^n})}{(r + 2)^n}.
\]

Letting \( r \) be arbitrarily large and optimizing over all choices of \( \mu_{rI^n} \) gives the result. \( \square \)

**Corollary 6.17.** For each \( t \), \( \text{las}_t'(\mathbb{R}^n) \) is an upper bound on sphere packing, and

\[
\lim_{t \rightarrow \infty} \text{las}_t'(\mathbb{R}^n) = \delta_{\text{pack}, n}.
\]

**Proof.** For fixed \( r \), we may choose \( t \) sufficiently large so that \( \text{las}_t'(rI^n) = \alpha(rI^n) \) and \( r^{-n} \alpha(rI^n) \) gives an upper bound for \( \text{las}_t'(\mathbb{R}^n) \). As \( r \) becomes large, this bound will come arbitrarily close to the optimal sphere center density. \( \square \)

We can formulate an optimization problem dual to \( \text{las}_t'(\mathbb{R}^n) \) as follows.
Definition 6.18. Let \( \text{las}_t'({\mathbb R}^n)^* \) be the infimum of \( K(\emptyset, \emptyset) \) over all continuous kernels \( K: I_{-1,1} \times I_{-1,1} \to {\mathbb R} \) with compact support such that

1. \( K \) is positive semidefinite,
2. \( A_t K(S) \leq 0 \) whenever \( |S| > 1 \), and
3. \( \int_{I_{-1,1} \times I_{-1,1}} A_t K \, d\mathcal{L}_n \leq -1 \),

where we view Lebesgue measure \( \mathcal{L}_n \) as a measure on \( I_{-1,1} \) by identifying \( I_{-1,1} = {\mathbb R}^n \). We call such a \( K \) an auxiliary function of order \( 2t \).

Remark 6.19. One unsatisfying feature of the above optimization problem is that the kernel \( K \) cannot be made invariant under the action of \( {\mathbb R}^n \), as that would require that \( K(\emptyset, \{ x \}) \) take a constant value, contradicting the third condition in the definition. To formulate a dual problem in a way that allows for solutions invariant under the group action, one could use the Schur complement formulation discussed in Remark 6.8, at the cost of complicating the statement of the optimization problem.

Let \( \mu \) be a correlation measure and \( K \) an auxiliary function, both of order \( 2t \). Weak duality follows immediately from

\[
0 \leq \int_{I_{-1,1} \times I_{-1,1}} K \, dA_t^* \mu = \int_{I_{-1,1}^2} A_t K \, d\mu \leq K(\emptyset, \emptyset) - \hat{\mu}(0).
\]

The relationship with \( \text{las}_t'({\mathbb R}^n)^* \) is simple. Given any correlation measure \( \mu \) of order \( 2t \) for \( {\mathbb R}^n \), restricting \( \mu \) to \( I_{2t,1}^2 \) gives a feasible measure for \( \text{las}_t'({\mathbb R}^n)^* \) with objective \( r^n \hat{\mu}(0) \), and letting \( r \to \infty \) shows that

\[
\text{las}_t'({\mathbb R}^n)^* \leq \delta_{\text{las}_t', n}
\]

by Corollary 6.14 (this argument is the first part of the proof of Theorem 6.16). Conversely, suppose \( K \) is any feasible kernel for \( \text{las}_t'({\mathbb R}^n)^* \). Then extending \( r^{-n} K \) by zero gives a feasible kernel for \( \text{las}_t'({\mathbb R}^n)^* \) with objective \( r^{-n} K(\emptyset, \emptyset) \), and thus

\[
\text{las}_t'({\mathbb R}^n)^* \leq \delta_{\text{las}_t', n}.
\]

By combining these inequalities with Theorem 6.16, we obtain strong duality:

**Theorem 6.20.** For each \( n \) and \( t \),

\[
\text{las}_t'({\mathbb R}^n) = \text{las}_t'({\mathbb R}^n)^* = \delta_{\text{las}_t', n}.
\]

Note that the optimum in \( \text{las}_t'({\mathbb R}^n)^* \) will generally not be achieved unless we broaden the class of auxiliary functions.

**Remark 6.21.** The Lasserre hierarchy bound \( \text{las}_1'({\mathbb R}^n) \) is equivalent to the linear programming bound, because \( \text{las}_1' = \vartheta_1' \), as shown in [35, Theorem 3]. It is therefore sharp for \( n = 1, 8, \) and \( 24 \), and conjecturally for \( n = 2 \) (see [13, 44, 14]). For which other pairs \((n, t)\) might there be a sharp bound? It is unclear whether these sharp bounds are a peculiar phenomenon for \( t = 1 \), or whether we can expect further cases with \( t > 1 \). It is even conceivable that for each dimension \( n \), some finite value of \( t \) yields a sharp bound.

One hint that additional sharp bounds might be possible in Euclidean space comes from the case of binary codes of block length 20 and minimal distance 8. Gijswijt, Mittelmann, and Schrijver [20] obtained a sharp bound for the size of such a code using \( \text{las}_2' \) (see Section VIII of their paper for the reduction to \( \text{las}_2' \)).
This paper is a spinoff from a larger project with David de Laat, with whom we had numerous helpful discussions but who declined coauthorship. We also thank Austin Anderson, Alexander Reznikov, Oleksandr Vlasiuk, and Edward White for pointing out reference [33].

**References**

[1] N. Alon, *The Shannon capacity of a union*, Combinatorica 18 (1998), no. 3, 301–310. MR1721946 doi:10.1007/PL00009824

[2] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000. MR1857292

[3] A. Anderson, A. Reznikov, O. Vlasiuk, and E. White, *Polarization and covering on sets of low smoothness*, in preparation, 2021.

[4] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin, *Invariant semidefinite programs*, Handbook on semidefinite, conic and polynomial optimization (M. F. Anjos and J. B. Lasserre, eds.), Internat. Ser. Oper. Res. Management Sci., vol. 166, Springer, New York, 2012, pp. 219–269. arXiv:1007.2905 MR2894697 doi:10.1007/978-1-4614-0769-0_9

[5] C. Bachoc and F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. 21 (2008), no. 3, 909–924. arXiv:math/0608426 MR2393433 doi:10.1090/S0894-0347-07-00589-9

[6] R. Ben Ghanem and C. Frappier, *Explicit quadrature formulae for entire functions of exponential type*, J. Approx. Theory 92 (1998), no. 2, 267–279. MR1604935 doi:10.1006/jath.1997.3122

[7] S. V. Borodachov, D. P. Hardin, and E. B. Saff, *Asymptotics of best-packing on rectifiable sets*, Proc. Amer. Math. Soc. 135 (2007), no. 8, 2369–2380. arXiv:math-ph/0605021 MR2302558 doi:10.1090/S0002-9939-07-08975-7

[8] H. Cohn and M. de Courcy-Ireland, *The Gaussian core model in high dimensions*, Duke Math. J. 167 (2018), no. 13, 2417–2455. arXiv:1603.09684 MR3855354 doi:10.1215/00127094-2018-00018

[9] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) 185 (2017), no. 3, 1017–1033. arXiv:1603.06518 MR3664817 doi:10.4007/annals.2017.185.3.8

[10] H. Cohn and N. Elkies, *New upper bounds on sphere packings I*, Ann. of Math. (2) 157 (2003), no. 2, 689–714. arXiv:math.MG/0110009 MR1973059 doi:10.4007/annals.2003.157.689

[11] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) 185 (2017), no. 3, 1017–1033. arXiv:1603.06518 MR3664817 doi:10.4007/annals.2017.185.3.8

[12] H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, in preparation, 2021.

[13] H. Cohn and Y. Zhao, *Sphere packing bounds via spherical codes*, Duke Math. J. 163 (2014), no. 10, 1965–2002. arXiv:1212.5966 MR3229046 doi:10.1215/00127094-2738857

[14] H. Dym and H. P. McKean, *Fourier series and integrals*, Academic Press, New York-London, 1972. MR0442564

[15] H. Federer, *Geometric measure theory*, reprint of the 1969 ed., Classics in Mathematics, Springer-Verlag, 1996. MR0257325 doi:10.1007/978-3-642-62010-2

[16] L. Fejes Tóth, *Regular figures*, A Pergamon Press Book, The Macmillan Co., New York, 1964. MR0165423

[17] P. Brass, W. Moser, and J. Pach, *Research problems in discrete geometry*, Springer, New York, 2005. MR2163782 doi:10.1007/978-0-387-29929-7

[18] D. L. Cohn, *Measure theory*, second ed., Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser/Springer, New York, 2013. MR3098996 doi:10.1007/978-1-4614-6956-8

[19] L. Fejes Tóth, *Regular figures*, A Pergamon Press Book, The Macmillan Co., New York, 1964. MR0165423

[20] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) 185 (2017), no. 3, 1017–1033. arXiv:1603.06518 MR3664817 doi:10.4007/annals.2017.185.3.8
[21] D. V. Gorbachev, Extremum problems for entire functions of exponential spherical type, Math. Notes 68 (2000), no. 2, 159–166. MR1822646 doi:10.1007/BF02675341

[22] , Extremum problem for periodic functions supported in a ball, Math. Notes 69 (2001), no. 3, 313–319. MR1846833 doi:10.1023/A:1010275206760

[23] H. Groemer, Existenzsätze für Lagerungen im Euklidischen Raum, Math. Z. 81 (1963), 260–278. MR163222 doi:10.1007/BF01111546

[24] G. R. Grozev and Q. I. Rahman, A quadrature formula with zeros of Bessel functions as nodes, Math. Comp. 64 (1995), no. 210, 715–725. MR1277767 doi:10.2307/2153447

[25] T. C. Hales, A proof of the Kepler conjecture, Ann. of Math. (2) 162 (2005), no. 3, 1065–1185. MR2179728 doi:10.4007/annals.2005.162.1065

[26] T. C. Hales and S. McLaughlin, The dodecahedral conjecture, J. Amer. Math. Soc. 23 (2010), no. 2, 299–344. arXiv:math/9811079 MR2601036 doi:10.1090/S0894-0347-09-00647-X

[27] T. Hales, M. Adams, G. Bauer, T. D. Dang, J. Harrison, L. T. Hoang, V. Magron, S. McLaughlin, T. T. Nguyen, Q. T. Nguyen, T. Nipkow, S. Obua, J. Pleso, J. Rute, A. Solovyev, T. H. A. Ta, N. T. Tran, T. D. Trieu, J. Urban, K. Vu, and R. Zumkeller, A formal proof of the Kepler conjecture, Forum Math. Pi 5 (2017), e2, 29. arXiv:1501.02155 MR3659768 doi:10.1017/fmp.2017.1

[28] D. P. Hardin and E. B. Saff, Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds, Adv. Math. 193 (2005), no. 1, 174–204. arXiv:math-ph/0311024 MR2132763 doi:10.1016/j.aim.2004.05.006

[29] D. Hardin, E. B. Saff, and O. Vlasiuk, Asymptotic properties of short-range interaction functionals, preprint, 2020. arXiv:2010.11937

[30] R. A. Horn and C. R. Johnson, Matrix analysis, second ed., Cambridge University Press, Cambridge, 2013. MR2978290

[31] D. E. Knuth, The sandwich theorem, Electron. J. Combin. 1 (1994), Article 1, 49 pp. arXiv:math/9312214 MR1269161 doi:10.37236/1193

[32] A. N. Kolmogorov, Selected works. III. Information theory and the theory of algorithms, Springer Collected Works in Mathematics, Springer, Dordrecht, 2019. MR3822138 doi:10.1007/978-94-017-2973-4

[33] A. N. Kolmogorov and V. M. Tikhomirov, ε-entropy and ε-capacity of sets in function spaces, Uspehi Mat. Nauk 14 (1959), no. 2 (86), 3–86. MR0112032

[34] D. de Laat, F. C. Machado, F. M. de Oliveira Filho, and F. Vallentin, k-point semidefinite programming bounds for equiangular lines, Math. Program., to appear. arXiv:1812.06045 doi:10.1007/s10107-021-01638-x

[35] D. de Laat and F. Vallentin, A semidefinite programming hierarchy for packing problems in discrete geometry, Math. Program. 151 (2015), no. 2, Ser. B, 529–553. arXiv:1311.3789 MR3348162 doi:10.1007/s10107-014-0843-4

[36] M. Laurent, A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming, Math. Oper. Res. 28 (2003), no. 3, 470–496. MR1997246 doi:10.1287/moor.28.3.470.16391

[37] M. Laurent and F. Rendl, Semidefinite programming and integer programming, Discrete optimization (K. Aardal, G. Nemhauser, and R. Weismantel, eds.), Handbooks in Operations Research and Management Science, vol. 12, Elsevier, 2005, pp. 393–514. doi:10.1016/S0927-0507(05)12008-8

[38] P. Mattila, Geometry of sets and measures in Euclidean spaces: fractals and rectifiability, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995. MR1333890 doi:10.1017/CBO9780511562381

[39] J.-F. Sadoc and R. Mosseri, Geometrical frustration, Collection Aléa-Saclay: Monographs and Texts in Statistical Physics, Cambridge University Press, Cambridge, 1999. MR1716082 doi:10.1017/CBO9780511599934

[40] C. L. Siegel, Über Gitterpunkte in convexen Körpren und ein damit zusammenhängendes Extremalproblem, Acta Math. 65 (1935), no. 1, 307–323. MR1555407 doi:10.1007/BF02420049

[41] B. Simon, Operator theory, A Comprehensive Course in Analysis, Part 4, American Mathematical Society, Providence, RI, 2015. MR3364494 doi:10.1090/simon/004

[42] E. M. Stein and G. Weiss, Introduction to Fourier analysis on Euclidean spaces, Princeton Mathematical Series, No. 32, Princeton University Press, Princeton, NJ, 1971. MR0304972
[44] M. S. Viazovska, *The sphere packing problem in dimension 8*, Ann. of Math. (2) **185** (2017), no. 3, 991–1015. arXiv:1603.04246 MR3664816 doi:10.4007/annals.2017.185.3.7

Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142, USA

Email address: cohn@microsoft.com

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Email address: asalmon@mit.edu