Limits of Order Types *

Xavier Goaoc, Alfredo Hubard, Rémi de Joannis de Verclos, Jean-Sébastien Sereni, Jan Volec

Limits of Order Types*

Xavier Goaoc† Alfredo Hubard‡ Rémi de Joannis de Verclos§
Jean-Sébastien Sereni¶ Jan Volec‖

November 6, 2018

Abstract

We apply ideas from the theory of limits of dense combinatorial structures to study order types, which are combinatorial encodings of finite point sets. Using flag algebras we obtain new numerical results on the Erdős problem of finding the minimal density of 5- or 6-tuples in convex position in an arbitrary point set, and also an inequality expressing the difficulty of sampling order types uniformly. Next we establish results on the analytic representation of limits of order types by planar measures. Our main result is a rigidity theorem: we show that if sampling two measures induce the same probability distribution on order types, then these measures are projectively equivalent provided the support of at least one of them has non-empty interior. We also show that some condition on the Hausdorff dimension of the support is necessary to obtain projective rigidity and we construct limits of order types that cannot be represented by a planar measure. Returning to combinatorial geometry we relate the regularity of this analytic representation to the aforementioned problem of Erdős on the density of k-tuples in convex position, for large k.

Keywords: Limits of structures, flag algebra, geometric measure theory, Erdős-Szekeres theorem, Sylvester's problem.

1 Introduction

The theory of *dense graph limits*, developed over the last decade by Borgs, Chayes, Lovász, Razborov, Sós, Szegedy, Vesztergombi and others, studies sequences of large graphs using a combination of equivalent formalisms: algebraic (as positive homomorphisms from certain graph algebras into $\mathbf{R}$), analytic (as measurable, symmetric functions from $[0,1]^2$ to $[0,1]$ called *graphons*) and discrete probabilistic (as families $\{D_n\}$ of probability distributions over n-vertex graphs satisfying certain relations). These viewpoints are complementary: while the algebraic formalism allows effective computations *via* semi-definite methods [30], the analytic viewpoint offers powerful methods (norm equivalence, completeness) to treat in a unified setting a diversity of graph problems such as pseudorandom graphs or property testing [25].

In this article, we combine ideas from dense graph limits with order types, which are combinatorial structures arising in geometry. The order type of a point set encodes the respective positions of its elements, and suffices to determine many of its properties, for instance its convex hull, its triangulations, or which graphs admit crossing-free straight line drawings with vertices supported on that point set. Order types have received continued attention in discrete and computational geometry since the 1980s

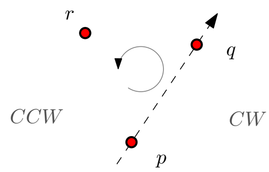
*This work was partially supported by ANR blanc PRESAGE (ANR–11-BS02–003).
†Université de Lorraine, CNRS, INRIA, LORIA, F-54000, Nancy, France. Email: xavier.goaoc@loria.fr. This author was partially supported by Institut Universitaire de France.
‡Université Paris Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE, UPEM, F-77454, Marne-la-Vallée, France. Email: alfredo.hubard@u-pem.fr.
§Radboud University Nijmegen, Netherlands. Email: r.deverclos@math.ru.nl.
¶Centre National de la Rercherche Scientifique, ICube (CSTB), Strasbourg, France. Email: sereni@kam.mff.cuni.cz.
‖Department of Mathematics, Emory University, Atlanta, USA. Email: jan@ucw.cz. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 800607. Previous affiliation: Department of Mathematics and Statistics, McGill University, Montreal, Canada, where this author was supported by CRM-ISM fellowship.

and are known to be rather intricate objects, difficult to axiomatise [33]. While order types are well defined in a variety of contexts (arbitrary dimensions, abstractly *via* the theory of oriented matroids) all point sets considered in this work are finite subsets of the Euclidean plane with no aligned triple, unless otherwise specified.

1.1 Order types and their limits

Let us first define our main objects of study.

Order types. Define the *orientation* of a triangle pqr in the plane to be *clockwise* (CW) if r lies to the right of the line pq oriented from p to q and *counter-clockwise* (CCW) if r lies to the left of that oriented line. (So the orientation of qpr is different from that of pqr.) We say that two planar point sets P and Q *have the same order type* if there exists a bijection $f\colon P \to Q$ that preserves orientations: for any triple of pairwise distinct points $p, q, r \in P$ the triangles pqr and $f(p)f(q)f(r)$ have the same orientation. The relation of having the same order type is easily checked to be an equivalence relation; the equivalence class, for this relation, of a finite point set P is called the *order type* of P. A point set P with order type ω is called a *realization* of ω.



When convenient, we extend to order types any notion that can be defined on a set of points and does not depend on a particular choice of realization. For instance we define the *size* of an order type ω to be the cardinality $|\omega|$ of any of its realization. We adopt the convention that there is exactly one order type of each of the sizes 0, 1 and 2. We let $\mathcal{O}$ be the set of order types and $\mathcal{O}_n$ the set of order types of size n.

Convergent sequences and limits of order types. We define the *density* $p(\omega, \omega')$ of an order type ω in another order type ω' as the probability that $|\omega|$ random points chosen uniformly from a point set realizing ω' have order type ω. (Observe that this probability depends solely on the order types and not on the choice of realization.) We say that a sequence $(\omega_n)_{n\in\mathbf{N}}$ of order types *converges* if the size $|\omega_n|$ goes to infinity as n goes to infinity, and for any fixed order type ω the sequence $(p(\omega, \omega_n))_n$ of densities converges. The *limit* of a convergent sequence of order types $(\omega_n)_{n\in\mathbf{N}}$ is the map

$$\left\{ \begin{array}{ccl} \mathcal{O} & \to & [0,1] \\ \omega & \mapsto & \lim_{n\to\infty} p(\omega, \omega_n) \end{array} \right.$$

A standard compactness argument reveals that limits of order types abound. Indeed, for each element ω_n in a sequence of order types, the map $\omega \in \mathcal{O} \mapsto p(\omega, \omega_n)$ can be seen as a point in $[0,1]^{\mathbf{N}}$, which is compact by Tychonoff's theorem. Any sequence of order types with sizes going to infinity therefore contains a convergent subsequence.

1.2 Problems and results

We explore the application of the theory of limits of dense graphs to order types in two directions. On one hand, the algebraic description of limits as positive homomorphisms of flag algebras makes these limits amenable to semi-definite programming methods. We implemented this approach for order types and obtained numerical results. On the other hand, the fact that measures generally define limits of order types (Corollary 8) unveils stimulating problems and interesting questions, of a more structural nature, on the relation between measures and limits of order types.

Flag algebras of order types. The starting motivation for our work is a question raised by Erdős and Guy [18] in 1973 (see also [17]): "what is the minimum number $\mathrm{conv}_k(n)$ of convex k-gons in a set of n points in the plane?". This falls within the scope of a general (and more conceptual than precise) question of Sylvester: "what is the probability that four points at random are in convex position?". Making sense of Sylvester's question implies defining a distribution on 4-tuples of points, and from the beginning of the 20th century several variants using distributions coming from the theory of convex sets were investigated (*e.g.* uniform or gaussian distributions on compact convex sets). For more background on this, the reader is referred to the survey by Ábrego, Fernández-Merchant & Salazar [4] and to the book by Brass, Moser & Pach [11, Section 8]. We also point out that Sylvester's question is actually related to several important conjectures in convex geometry [12, Chapter 3]. By a standard double-counting argument, one sees that $\mathrm{conv}_k(n+1) \geqslant \frac{n+1}{n+1-k}\mathrm{conv}_k(n)$, so the limiting density

$$c_k = \lim_{n\to\infty} \mathrm{conv}_k(n)/\binom{n}{k}$$

is well defined and equal to the supremum of this ratio for $n \in \mathbf{N}$. We apply the framework of flag algebras to order types and use the semi-definite method to obtain lower bounds on c_k for $k \in \{4,5,6\}$. As it turns out, the literature around the computation of c_4 is vast: not only does c_4 correspond to the last open case of a relaxation of Sylvester's conjecture to all open sets of finite area, but as discovered by Scheinerman and Wilf [32], its value is determined by the asymptotic behaviour of the rectilinear crossing number of the complete graph, which has been extensively investigated. For this particular case, the best lower bound we could obtain is $c_4 \geqslant 0.37843917$, which falls short of the currently best known lower bound, namely $277/729 > 0.3799$. This better lower bound is obtained by plugging results of Aichholzer *et al.* [6] and Ábrego *et al.* [1, 3] into an expression of the rectilinear crossing number found independently by Lovász *et al.* [27] and by Ábrego and Fernández-Merchant [2]. The best currently known upper bound on c_4 is $c_4 < 0.380473$, and is due to Fabila-Monroy and López [20]. Nonetheless, our method allows us to strongly improve the known lower bounds on c_5 and on c_6.

Proposition 1. *We have* $c_5 \geqslant 0.0608516$ *and* $c_6 \geqslant 0.0018311$.

To the best of our knowledge, prior to this work the only known lower bounds on any constant c_k with $k \geqslant 5$ followed from a general and important result of Erdős and Szekeres [19], *via* a simple double counting argument. One indeed sees that $c_5 \geqslant \frac{5!4!}{9!} > 0.00793$ using that nine points in the plane must contain a convex pentagon (a result attributed to Makai by Erdős and Szekeres [19], the first published proof being by Kalbfleisch, Kalbfleisch & Stanton [23]). Similarly, as Szekeres and Peters [38] proved (using a computer-search) that the Erdős-Szekeres conjecture [19] is true for convex hexagons, one can use that seventeen points in the plane must contain a convex hexagon to infer that $c_6 \geqslant \frac{11!6!}{17!} > 0.0000808$. The best upper bounds that we are aware of on these numbers are $c_5 \leqslant 0.0625$ and $c_6 \leqslant 0.005822$. We point out that Ábrego (personal communication) conjectured that $c_5 = 0.0625$. We again refer the interested reader to the survey and the book cited above [4, 11].

We prove Proposition 1 by a reformulation of limits of order types as positive homomorphisms from a so-called flag algebra of order types into $\mathbf{R}$ (see Proposition 9); this point of view allows a semidefinite programming formulation of the search for inequalities satisfied by limits of order types. Specifically, we argue that for any limit of order types ℓ

$$\ell(\diamond_5) \geqslant 0.0608516 \qquad \text{and} \qquad \ell(\diamond_6) \geqslant 0.0018311,$$

where $\diamond_k$ is the order type of k points in convex position for any positive integer k.

On a related topic, we can mention a recent application of flag algebras by Balogh, Lidický, and Salazar [9] to the (non rectilinear) crossing number of the complete graph. Their techniques differ from ours in that they use rotational systems instead of order types, and they use results about the crossing number of the complete graph for small numbers.

We now turn to another aspect of our work using flag algebras of order types. Probabilistic constructions often present extremal combinatorial properties that are beyond our imagination, a textbook example being the lower bound on Ramsey numbers for graphs devised by Erdős [16] in 1947. Sampling order types of a given size *uniformly* is of much interest to test conjectures and search for extremal examples (see *e.g.* [11, p. 326]). However the uniform distribution on order types seems out of reach as suggested by the lack of closed formulas for counting them, or heuristics to generate them, but we know of no formal justification of the hardness of approximation of this distribution. As it turns out, limits of order types can also be defined as families of probability distributions on order types with certain internal consistencies (see Proposition 7) and we exploit this interpretation to provide negative results on certain sampling strategies. Our second result obtained by the semidefinite method of flag algebras indeed shows that a broad class of random generation methods must exhibit some bias.

Proposition 2. *For any limit of order types* ℓ *there exist two order types* ω_1 *and* ω_2 *of size* 6 *such that* $\ell(\omega_1) > 1.8208\ \ell(\omega_2)$.

This unavoidable bias holds, in particular, for the random generation of order types by independent sampling of points from *any* finite Borel measure over $\mathbf{R}^2$, see Corollary 8.

Representing limits by measures. Given a finite Borel measure μ over $\mathbf{R}^2$ and an order type $\omega \in \mathcal{O}$, let $\ell_\mu(\omega)$ be the probability that $|\omega|$ random points sampled independently from μ have order type ω. It turns out (Corollary 8) that ℓ_μ is a limit of order types if and only if every line is negligible for μ. Going in the other direction, it is natural to wonder if geometric measures could serve as analytic representations of limits of order types, like graphons for limits of dense graphs. This raises several questions, here are some that we investigate: does every limit of order types enjoy such a realization? For those which do, what does the set of measures realizing them look like? In analogy with the study of realization spaces of order types [8, 14, 24, 33] we investigate how the weak topology on probability measures relates to the topology on limits.

We provide an explicit construction of a limit of order types that cannot be represented by any measure in the plane. For $t \in (0,1)$ let $\odot_t$ be the probability distribution over $\mathbf{R}^2$ supported on two concentric circles with radii 1 and t, respectively, where each of the two circles has $\odot_t$-measure $1/2$, distributed proportionally to the length on this circle. Because every line is negligible for $\odot_t$, Corollary 8 ensures that $\ell_{\odot_t}$ is a limit of order types. We define $\ell_\odot$ to be the limit of an arbitrary convergent sub-sequence of $(\ell_{\odot_{1/n}})_{n\in\mathbf{N}}$.

Proposition 3. *There exists no probability measure μ such that $\ell_\mu = \ell_\odot$.*

The proof of Proposition 3 reveals that the sequence $(\ell_{\odot_{1/n}})_{n\in\mathbf{N}}$ is in fact convergent, so $\ell_\odot$ is indeed an explicit example. Actually, an explicit description (for instance based on the reformulation of $\ell_\odot$ used in the proof of Lemma 3.6) is possible, if tedious.

Let us now turn our attention to a family of limits that has a particularly nice realization space. Firstly, observe that nonsingular affine transforms preserve order types. Thus, if μ is a measure that does not charge any line, f is a nonsingular affine transform, and $\mu' = \mu \circ f^{-1}$ is the push-forward of μ by f, then $\ell_{\mu'} = \ell_\mu$. More generally, one can take f to be the restriction to $\mathbf{R}^2$ of an "adequate" projective map (we spell out the meaning of "adequate" in Section 3.1). Our main contribution in this direction is a projective rigidity result: we show that under natural measure theoretic conditions on μ, the realization space of μ is essentially a point.

Theorem 4. *Let μ_1 and μ_2 be two compactly supported measures of $\mathbf{R}^2$ that charge no line and whose supports have non-empty interiors. If $\ell_{\mu_1} = \ell_{\mu_2}$, then there exists a projective transformation f such that $\mu_2 \circ f = \mu_1$.*

The converse of the statement of Theorem 4 is not true: it can fail for instance if the line mapped to infinity intersects the interior of $\operatorname{supp}\mu$. The approach we use to establish it actually provides a necessary and sufficient condition, expressed in terms of "spherical transforms" as defined in Section 3.1.
The third author conjectured [13] that the condition on μ can be weakened.

Conjecture 5. *If μ is a measure that charges no line and has support of Hausdorff dimension strictly greater than 1, then every measure that realizes ℓ_μ is projectively equivalent to μ.*

It is easy to see that the conditions in Conjecture 5 are necessary. Recalling that $\diamond_k$ is the order type of k points in convex position, the limit of order types that for every k gives probability 1 to $\diamond_k$ can be realized by measures of Hausdorff dimension in $(0,1)$. In particular, these are pairwise projectively inequivalent. More interestingly, we build a limit of order types ℓ_E presenting projective flexibility, and which is interesting from the combinatorial geometry perspective: indeed, ℓ_E is a nearly optimal lower bound for the aforementioned problem of Erdős, witnessing that $\log c_k = \Theta(-k^2)$.

Theorem 6. *There exists a limit of order types ℓ_E that can be realized by a measure of Hausdorff dimension s for any $s \in (0,1)$ for which $\ell_E(\diamond_k) = 2^{-\frac{k^2}{8}+O(k\log k)}$. Furthermore, ℓ_E cannot be realized by a measure that is absolutely continuous with respect to the Lebesgue measure.*

As in the rigidity result, the relation to the Hausdorff dimension seems fundamental: we observe that regular measures present a very different asymptotic behavior with respect to the density of $\diamond_k$.

2 Flag algebras

2.1 Reformulation of limits

Order types can be understood as equivalence classes of chirotopes under the action of permutations (see below). As such, they form an example of a *model* in the language of Razborov [30], and the theory of limits of order types is a special case of Razborov's work. This section provides a self-contained presentation of the probabilistic and algebraic reformulations of limits of order types.

2.1.1 Probabilistic characterization

Let $(\omega_n)_{n\in\mathbf{N}}$ be a sequence of order types converging to a limit ℓ. Let $\omega \in \mathcal{O}$, let $k \geqslant |\omega|$ and let n_0 be large enough so that $|\omega_n| \geqslant k$ for every $n \geqslant n_0$. A simple conditioning argument yields that for any $n \geqslant n_0$,

$$p(\omega, \omega_n) = \sum_{\omega' \in \mathcal{O}_k} p(\omega, \omega') p(\omega', \omega_n). \tag{1}$$

Indeed, the probability that a random sample realizes ω is the same if we sample uniformly $|\omega|$ points from a realization of ω_n, or if we first sample k points uniformly from that realization and next uniformly select a subset of $|\omega|$ of these k points. It follows that any limit ℓ of order types satisfies the following *conditioning identities*:

$$\forall \omega \in \mathcal{O}, \forall k \geqslant |\omega|, \quad \ell(\omega) = \sum_{\omega' \in \mathcal{O}_k} p(\omega, \omega')\ell(\omega'). \tag{2}$$

As one notices, Equation (1) yields that the conditioning identities are equivalent to the (seemingly weaker) condition

$$\forall \omega \in \mathcal{O}, \quad \ell(\omega) = \sum_{\omega' \in \mathcal{O}_{|\omega|+1}} p(\omega, \omega')\ell(\omega'). \tag{3}$$

However, a stronger condition than Equation (2) is needed to characterize limits of order types, as we illustrate in Example 2.1. The *split probability* $p\,(\omega', \omega''; \omega)$, where ω', ω'' and ω are order types, is the probability that a random partition of a point set realizing ω into two classes of sizes $|\omega'|$ and $|\omega''|$, chosen uniformly among all such partitions, produces two sets with respective order types ω' and ω''. (In particular $p\,(\omega', \omega''; \omega) = 0$ if $|\omega| \neq |\omega_1| + |\omega_2|$.) We provide a detailed proof of the following proposition, but the reader familiar with the topic already sees the corresponding result for dense graph limits [26, Theorem 2.2].

Proposition 7. *A function $\ell \colon \mathcal{O} \to \mathbf{R}$ is a limit of order types if and only if*

$$\forall \omega', \omega'' \in \mathcal{O}, \qquad \ell(\omega')\ell(\omega'') = \sum_{\omega \in \mathcal{O}_{|\omega'|+|\omega''|}} p(\omega', \omega''; \omega)\ell(\omega), \tag{4}$$

and for every $n \in \mathbf{N}$, the restriction $\ell_{|\mathcal{O}_n}$ is a probability distribution on $\mathcal{O}_n$.

Before establishing Proposition 7, let us first point out that the product condition (4) implies the conditioning condition (2), as is seen by taking for ω'' the (unique) order type of size 1.

Proof of Proposition 7. We start by establishing the direct implication. The fact that $\ell_{|\mathcal{O}_n}$ is a probability distribution on $\mathcal{O}_n$ follows from the definition of a limit. As for Equation (4), fix two order types ω' and ω'' in $\mathcal{O}$. Let ω_n be a random order type sampled from $\ell_{|\mathcal{O}_n}$ where $n \geqslant |\omega'| + |\omega''|$. Let

$$\alpha_n = p(\omega', \omega_n)p(\omega'', \omega_n) \qquad \text{and} \qquad \beta_n = \sum_{\omega \in \mathcal{O}_{|\omega'|+|\omega''|}} p(\omega', \omega''; \omega)p(\omega, \omega_n).$$

Now, fix some point set P with order type ω_n. By the definition, the value $p(\omega, \omega_n)$ is the probability that $|\omega|$ points sampled uniformly from P have order type ω. Now, on one hand, α_n equals the probability that two independent events both happens: (i) that a set P' of $|\omega'|$ random points chosen uniformly from P has order type ω', and (ii) that another set P'' of $|\omega''|$ random points chosen uniformly from P has order type ω''. On the other hand, observe that β_n equals the probability that (i) and (ii) happen *and* that P' and P'' are disjoint. The difference $|\alpha_n - \beta_n|$ is therefore bounded from above by the probability that P' and P'' intersect. Bounding from above the probability that P' and P'' have an intersection of one or more elements by the expected size of $P' \cap P''$ yields that

$$\left| p(\omega', \omega_n) p(\omega'', \omega_n) - \sum_{\omega \in \mathcal{O}_{|\omega'|+|\omega''|}} p(\omega', \omega''; \omega) p(\omega, \omega_n) \right| \leqslant \mathbf{E}\left(|P' \cap P''|\right) = \frac{|\omega'||\omega''|}{|\omega_n|}. \tag{5}$$

Taking $n \to \infty$ in (5) we see that ℓ satisfies Equation (4).

Conversely, suppose that ℓ satisfies both conditions. For every integer n, we pick a random order type r_n of size n^2 according to $\ell_{|\mathcal{O}_{n^2}}$. We assert that

$$\forall \omega \in \mathcal{O}, \qquad \mathbf{P}\left(\lim_{n \to \infty} p(\omega, r_n) \neq \ell(\omega) \right) = 0.$$

Since the number of order types is countable, we conclude that the random sequence $(r_n)_{n \in \mathbf{N}}$ is convergent and has limit ℓ with probability 1.

It remains to prove the assertion. Fix some $\omega \in \mathcal{O}$ and assume that $n \geqslant |\omega|$. Then

$$\mathbf{E}[p(\omega, r_n)] = \sum_{\omega' \in \mathcal{O}_{n^2}} \mathbf{P}(r_n = \omega') p(\omega, \omega') = \sum_{\omega' \in \mathcal{O}_{n^2}} \ell(\omega') p(\omega, \omega') = \ell(\omega),$$

and

$$\mathbf{Var}(p(\omega, r_n)) = \mathbf{E}[p(\omega, r_n)^2] - \mathbf{E}[p(\omega, r_n)]^2 = \sum_{\omega' \in \mathcal{O}_{n^2}} p(\omega, \omega')^2 \ell(\omega') - \ell(\omega)^2.$$

By Equation (4),

$$\mathbf{Var}(p(\omega, r_n)) = \sum_{\omega' \in \mathcal{O}_{n^2}} p(\omega, \omega')^2 \ell(\omega') - \sum_{\omega'' \in \mathcal{O}_{2|\omega|}} p(\omega, \omega; \omega'') \ell(\omega'').$$

Since $2|\omega| \leqslant |\omega|^2 \leqslant n^2 = |r_n|$, Equation (2) yields that $\ell(\omega'') = \sum_{\omega' \in \mathcal{O}_{n^2}} p(\omega'', \omega') \ell(\omega')$. So we deduce that

$$\mathbf{Var}(p(\omega, r_n)) = \sum_{\omega' \in \mathcal{O}_{n^2}} \left(p(\omega, \omega')^2 - \sum_{\omega'' \in \mathcal{O}_{2|\omega|}} p(\omega, \omega; \omega'') p(\omega'', \omega') \right) \ell(\omega').$$

We observe that Inequality (5) obtained above is valid in general, and therefore

$$\mathbf{Var}(p(\omega, r_n)) \leqslant \sum_{\omega' \in \mathcal{O}_{n^2}} \frac{|\omega|^2}{n^2} \ell(\omega') = \frac{|\omega|^2}{n^2}.$$

By Chebyshev's inequality it thus follows that for any positive ε,

$$\mathbf{P}(|p(\omega, r_n) - \ell(\omega)| > \varepsilon) \leqslant \frac{|\omega|^2}{\varepsilon^2 n^2}.$$

Therefore, for any ω and $\varepsilon > 0$, the sum $\sum_{n \geqslant 1} \mathbf{P}(|p(\omega, r_n) - \ell(\omega)| > \varepsilon)$ is finite and hence the Borel-Cantelli lemma implies that with probability 1, only finitely many of the events $\{|p(\omega, r_n) - \ell(\omega)| > \varepsilon\}_n$ happen. Consequently, it holds with probability 1 that $\lim_{n \to \infty} p(\omega, r_n) = \ell(\omega)$. $\square$

Proposition 7 provides many examples of limits of order types.

Corollary 8. *Let μ be a finite Borel measure over $\mathbf{R}^2$. The map $\ell_\mu \colon \omega \in \mathcal{O} \mapsto p(\omega, \mu)$ is a limit of order types if and only every line is negligible for μ.*

Proof. Assume that ℓ_μ is a limit of order types and let $(\omega_n)_{n \in \mathbf{N}}$ be a sequence converging to μ. Let $\therefore$ be the order type of size 3. We have

$$p(\therefore, \mu) = \ell_\mu(\therefore) = \lim_{n \to \infty} p(\therefore, \omega_n) = 1,$$

so three random points chosen independently from $\frac{1}{\mu(\mathbf{R}^2)}\mu$ are aligned with probability 0, and consequently every line is negligible for μ.

Conversely, assume that μ is a measure for which every line is negligible. For every integer $n \geqslant 3$ the restriction of ℓ_μ to $\mathcal{O}_n$ is a probability distribution. Moreover, for every order type $\omega \in \mathcal{O}$ and every integer $m > |\omega|$, we have

$$\ell_\mu(\omega) = \sum_{\omega' \in \mathcal{O}_m} p(\omega, \omega')\ell_\mu(\omega')$$

by sampling m points from μ and sub-sampling $|\omega|$ points among them uniformly. Proposition 7 then implies that ℓ_μ is a limit of order types. $\square$

We conclude this section with a simple example of a function ℓ that is not a limit of order types, and yet satisfies the conditioning identities and restricts on every $\mathcal{O}_n$ to a probability distribution.

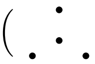
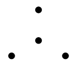
Example 2.1. Let $\ell_\circ$ be the limit where convex order types have probability 1. Let $\ell_\odot$ be the limit defined before Proposition 3 on page 4. Set ℓ to be $\frac{1}{2}(\ell_\circ + \ell_\odot)$.

First, note that if ℓ_1 and ℓ_2 are two limits of order types and ℓ_3 is any convex combination of ℓ_1 and ℓ_2, then the conditioning identities for ℓ_1 and ℓ_2 ensure that:

$$\forall \omega \in \mathcal{O}, \forall k > |\omega|, \quad \ell_3(\omega) = \sum_{\omega' \in \mathcal{O}_k} p(\omega, \omega')\ell_3(\omega').$$

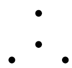
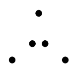
In particular, our function ℓ satisfies the conditioning identities. Similarly, it is immediate to check that for every n, the restriction of ℓ to $\mathcal{O}_n$ is a probability distribution on $\mathcal{O}_n$.

We show that ℓ does not satisfy (4) by proving that

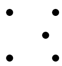
ℓ() $\ell(\bullet) = \frac{1}{32}$ and $\sum_{|\omega|=5} p$( $, \bullet\, ; \omega)\, \ell(\omega) = \frac{3}{64}$.

To establish these equalities, we compute certain values of $\ell_\odot$, which essentially boils down to computing two probabilities: (i) that the number of points on the outer circle equals that of the convex hull of the order type considered, and (ii) that the convex hull of these "outer points" contains the centre of the circle.

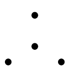
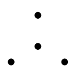
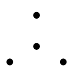
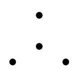
To compute (ii) requires to compute the probability that k points chosen uniformly on a circle have the circle's center in their convex hull. Let us condition on the k lines formed by the points and the center of the circle (the lines being almost surely pairwise distinct). Each point is selected uniformly from the two possible, antipodal, positions on the corresponding line, and the initial choice of the lines does not matter. So the probability reformulates as: given k pairs of antipodal points on the circle, how many of the 2^k k-gons formed by one point from each pair contain the center? For $k = 3$ the answer is 2 out of 8 configurations, while for $k = 4$ the answer is 8 out of 16 configurations. We thus infer that

$\ell_\odot$() $= 4\frac{1}{2^4}\frac{1}{4} = \frac{1}{16}$, $\ell_\odot$() $= \binom{5}{2}\frac{1}{2^5}\frac{1}{4} = \frac{5}{64}$,

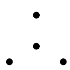
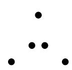
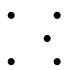
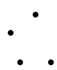
and

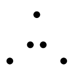
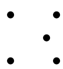
$\ell_\odot$() $= 5\frac{1}{2^5}\frac{1}{2} = \frac{5}{64}$.

It follows, on one hand, that

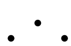
ℓ() $\ell(\bullet) = \ell$() $= \frac{1}{2}\Big(\ell_\circ$() $+ \ell_\odot$()$\Big) = \frac{1}{2}\left(0 + \frac{1}{16}\right) = \frac{1}{32}$.

and on the other hand, that

$\sum_{|\omega|=5} p$( $, \bullet\, ; \omega)\, \ell(\omega) = \frac{4}{5}\ell$() $+ \frac{2}{5}\ell$() $+ 0\ell$()

$= \frac{2}{5}\ell_\odot$() $+ \frac{1}{5}\ell_\odot$()

$$= \frac{2}{5} \times \frac{5}{64} + \frac{1}{5} \times \frac{5}{64}$$

$= \frac{3}{64} \neq \ell$() $\ell(\bullet)$.

2.1.2 Algebraic characterization

Recall that by (2) every limit ℓ of order types satisfies

$$\forall \omega \in \mathcal{O}, \forall k \geqslant |\omega|, \quad \ell(\omega) = \sum_{\omega' \in \mathcal{O}_k} p(\omega, \omega')\ell(\omega').$$

Now, let $\mathbf{R}\mathcal{O}$ be the set of all finite formal linear combinations of elements of $\mathcal{O}$ with real coefficients and consider the quotient vector space

$$\mathcal{A} = \mathbf{R}\mathcal{O}/\mathfrak{K} \qquad \text{where} \qquad \mathfrak{K} = \mathrm{vect}\Big\{\omega - \sum_{\omega' \in \mathcal{O}_{|\omega|+1}} p(\omega, \omega')\omega' : \omega \in \mathcal{O}\Big\}.$$

We define a product on $\mathcal{O}$ as follows:

$$\forall \omega_1, \omega_2 \in \mathcal{O}, \qquad \omega_1 \times \omega_2 = \sum_{\omega \in \mathcal{O}_{|\omega_1|+|\omega_2|}} p(\omega_1, \omega_2; \omega)\omega, \tag{6}$$

and we extend it linearly to $\mathbf{R}\mathcal{O}$. This extension is compatible with the quotient by $\mathfrak{K}$ and therefore turns $\mathcal{A}$ into an algebra [30, Lemma 2.4].

An algebra homomorphism from $\mathcal{A}$ to $\mathbf{R}$ is *positive* if it maps every element of $\mathcal{O}$ to a non-negative real, and we define $\mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ to be the set of positive algebra homomorphisms from $\mathcal{A}$ to $\mathbf{R}$. Observe that any algebra homomorphism sends $\cdot$, the order-type of size one, to the real 1 as $\cdot$ is the neutral element for the product of order types.

Proposition 9 ([30, Theorem 3.3b])**.** *A map $f\colon \mathcal{O} \to \mathbf{R}$ is a limit of order types if and only if its linear extension is compatible with the quotient by $\mathfrak{K}$ and defines a positive homomorphism from $\mathcal{A}$ to $\mathbf{R}$.*

We equip $\mathcal{A}$ with a partial order $\geqslant$, and write that $a_1 \geqslant a_2$ with $a_1, a_2 \in \mathcal{A}$ if the image of $a_1 - a_2$ under every positive homomorphism is non-negative. The algebra $\mathcal{A}$ allows us to compute effectively with density relations that hold for *every* limit ℓ.

Example 2.2. Let $\cdot$ be the order type on one point, ⸬ and ⁛ the two order types of size four and ⁘ , ⁙ and ⸪ the three order types of size five, seen as elements of $\mathcal{A}$. It follows from the definitions and from Equation (2) that

$$⸬ = ⁘ + \frac{3}{5}\, ⁙ + \frac{1}{5}\, ⸪ \qquad \text{and} \qquad ⁘ + ⁙ + ⸪ = \cdot \tag{7}$$

Since for any limit of order types ℓ we have $\ell(\cdot) = 1$, the above implies that

$$\ell(\diamond_4) = \ell\left(\frac{1}{5}\, ⁘ + \frac{1}{5}\, ⁙ + \frac{1}{5}\, ⸪\right) + \ell\left(\frac{4}{5}\, ⁘ + \frac{2}{5}\, ⁙\right) \geqslant \frac{1}{5}\ell(\cdot) = \frac{1}{5}$$

Using again Equation (2), and the non-negativity of ⸪ we obtain

$$\frac{2}{5}\, ⁘ \geqslant ⸬ - \frac{3}{5}\left(⁘ + ⁙ + ⸪ \right) = ⸬ - \frac{3}{5}\cdot$$

and $\ell(\diamond_5) \geqslant \frac{5}{2}\ell(\diamond_4) - \frac{3}{2}$ for any limit of order types ℓ.

2.1.3 Semi-definite method

Proposition 9 allows to search for inequalities over $\mathcal{A}$ by semidefinite programming. Let us give an intuition of how this works on an example. Here, we use the comprehensive list of all the order types of size up to 11, which was made available by Aichholzer [1] based on his work with Aurenhammer and Krasser [5] on the enumeration of order types. Throughout this paper, all non-trivial facts we use without reference on order types of small size can be traced back to that resource.

[1] http://www.ist.tugraz.at/aichholzer/research/rp/triangulations/ordertypes/

Example 2.3. A simple (mechanical) examination of the 6405 order types of size 8 reveals that $p(\diamond_4, \omega) \geqslant 19/70$ for any $\omega \in \mathcal{O}_8$. With Identity (2) this implies that  $\geqslant 19/70 \cdot$ or equivalently that $c_4 \geqslant 19/70 > 0.2714$. Observe that for any $C \in \mathcal{A}$ and any (linear extension of a) limit of order types ℓ we have $\ell(C \times C) = \ell(C)^2 \geqslant 0$ by Proposition 9. We thus have at our command an infinite source of inequalities to consider to try and improve the above bounds. For instance, a tedious but elementary computation yields that

$$\left(\frac{6}{25}\ - \frac{11}{125}\ \right)^2 + \frac{298819}{1093750} \sum_{\omega \in \mathcal{O}_8} \omega = \sum_{\omega \in \mathcal{O}_8} a_\omega \omega,$$

where $a_\omega \leqslant p(\diamond_4, \omega)$ for every $\omega \in \mathcal{O}_8$. This implies that $\ell(\diamond_4) \geqslant 298819/1093750 > 0.2732$ for any limit of order types ℓ. The search for interesting combinations of such inequalities can be done by semidefinite programming.

2.2 Improving the semidefinite method *via* rooting and averaging

The effectiveness of the semidefinite method for limits of graphs was greatly enhanced by considering partially labelled graphs. We unfold here a similar machinery, using some blend of order types and chirotopes.

Partially labelled point sets, flags, σ-flags and $\mathcal{A}^\sigma$. A point set *partially labelled* by a finite set $\mathcal{Z}$ (the *labels*) is a finite point set P together with some injective map $L\colon \mathcal{Z} \to P$. It is written $(P, \mathcal{Z}, L)$ when we need to make explicit the set of labels and the label map. Two partially labelled point sets $(P, \mathcal{Z}, L)$ and $(P', \mathcal{Z}, L')$ *have the same flag* if there exists a bijection $\phi\colon P \to P'$ that preserves both the orientation and the labelling, the latter meaning that $\phi(L(i)) = L'(i)$ for each $i \in \mathcal{Z}$. The relation of having the same flag is an equivalence relation, and *a flag* is an equivalence class for this relation. Again, we call any partially labelled point set a *realization* of its equivalence class, and the *size* $|\tau|$ of a flag τ is the cardinality of any of its realizations.

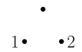
A flag where all the points are labelled, *i.e.* where $|P| = |\mathcal{Z}|$ in some realization $(P, \mathcal{Z}, L)$, is a *$\mathcal{Z}$-chirotope*. (When $\mathcal{Z} = [k] = \{1, 2, \ldots, k\}$ a $\mathcal{Z}$-chirotope coincides with the classical notion of chirotope.) Note that chirotopes correspond to types in the flag algebra terminology. Discarding the non-labelled part of a flag τ with label set $\mathcal{Z}$ yields some $\mathcal{Z}$-chirotope σ called the *root* of τ. A flag with root σ is a *σ-flag* and by $\mathcal{X}^\sigma$ we mean the set of σ-flags. The *unlabelling* $\tau^\emptyset$ of a flag τ with realization $(P, \mathcal{Z}, L)$ is the order type of P.

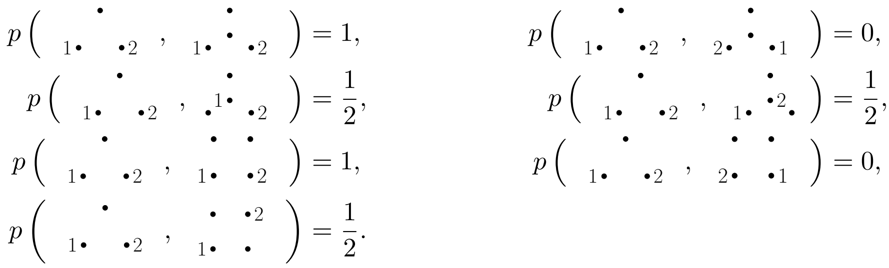
Let $\mathcal{Z}$ be a set of labels and σ a $\mathcal{Z}$-chirotope. We define densities and split probabilities for σ-flags like for order types. Namely, let τ, τ' and τ'' be σ-flags respectively realized by $(P, \mathcal{Z}, L)$, $(P', \mathcal{Z}, L')$ and $(P'', \mathcal{Z}, L'')$. The *density* of τ in τ' is the probability that for a random subset S of size $|P| - |\mathcal{Z}|$, chosen uniformly in $P' \setminus L'(\mathcal{Z})$, the partially labelled set $(S \cup L'(\mathcal{Z}), \mathcal{Z}, L')$ has flag τ. The *split probability* $p(\tau, \tau'; \tau'')$ is the probability that for a random subset S of size $|P| - |\mathcal{Z}|$, chosen uniformly in $P'' \setminus L''(\mathcal{Z})$, the partially labelled sets $(S \cup L''(\mathcal{Z}), \mathcal{Z}, L'')$ and $(P'' \setminus S, \mathcal{Z}, L'')$ have respective flags τ and τ'.

We can finally define an algebra of σ-flags as for order types. We endow the quotient vector space

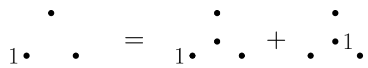
$$\mathcal{A}^\sigma = \mathbf{R}\mathcal{X}^\sigma / \mathfrak{K}^\sigma \qquad \text{where} \qquad \mathfrak{K}^\sigma = \operatorname{vect}\Big\{\omega - \sum_{\omega' \in \mathcal{X}^\sigma_{|\omega|+1}} p(\omega, \omega')\omega' : \omega \in \mathcal{X}^\sigma\Big\}$$

with the linear extension of the product defined on $\mathcal{X}^\sigma$ by $\tau \times \tau' = \sum_{\tau'' \in \mathcal{X}^\sigma_{|\tau|+|\tau'|-|\sigma|}} p(\tau, \tau'; \tau'')\tau''$.

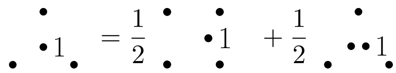
Example 2.4. Here are a few examples to illustrate the notions we just introduced. Letting σ be the unique $\mathcal{Z}$-chirotope with $|\mathcal{Z}| = 2$, there are exactly seven σ-flags on four points, three with a convex hull of size 4 and four with a convex hull of size 3. The densities of  into each of these seven σ-flags with 4 points indicated below.



Considering the quotient algebra, one sees that



and



Rooted homomorphisms and averaging. The interest of using the algebras $\mathcal{A}^\sigma$ to study $\mathcal{A}$ relies on three tools which we now introduce. We first define an *embedding* of a $\mathcal{Z}$-chirotope in an order type ω to be a σ-flag with root σ and unlabelling ω. We use *random embeddings* with the following distribution in mind: fix some point set realizing ω, consider the set I of injections $f\colon \mathcal{Z} \to P$ such that $(P, \mathcal{Z}, f)$ is a σ-flag, assume that $I \neq \varnothing$, choose some injection f_r from I uniformly at random, and consider the flag of $(P, \mathcal{Z}, f_r)$. We call this the *labelling* distribution on the embeddings of σ in ω.

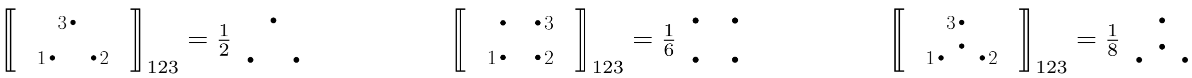
Next, we associate to any convergent sequence of order types $(\omega_n)_{n\in\mathbf{N}}$ with limit ℓ, and for every $\mathcal{Z}$-chirotope σ such that $\ell(\sigma^\varnothing) > 0$, a probability distribution on $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$. For every $n \in \mathbf{N}$, the labelling distribution on embeddings of σ in ω_n defines a probability distribution $\mathbf{P}^\sigma_\mathbf{n}$ on mappings from $\mathcal{A}^\sigma$ to $\mathbf{R}$; specifically, for each embedding θ_n of σ in ω_n we consider the map

$$f_{\theta_n} : \left\{ \begin{array}{ccc} \mathcal{A}^\sigma & \to & [0,1] \\ \tau & \mapsto & p(\tau, \theta_n) \end{array} \right.$$

and assign to it the same probability, under $\mathbf{P}^\sigma_\mathbf{n}$, as the probability of θ_n under the labelling distribution. As $\ell(\sigma^\varnothing)$ is positive, the fact that $p(\omega, \omega_n)$ converges as $n \to \infty$ for every $\omega \in \mathcal{O}$ implies the weak convergence of the sequence $(\mathbf{P}^\sigma_\mathbf{n})_{n\in\mathbf{N}}$ to a Borel probability measure $\mathbf{P}^\sigma_\ell$ on $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$ [30, Theorems 3.12 and 3.13]. Moreover, as $\ell(\sigma^\emptyset)$ is positive, the homomorphism induced by ℓ determines the probability distribution $\mathbf{P}^\sigma_\ell$ [30, Theorem 3.5].

We finally define, for every $\mathcal{Z}$-chirotope σ, an *averaging* (or downward) operator $[\![\cdot]\!]_\sigma \colon \mathcal{A}^\sigma \to \mathcal{A}$ as the linear operator defined on the elements of $\tau \in \mathcal{X}^\sigma$ by $[\![\tau]\!]_\sigma = p^\sigma_\tau \cdot \tau^\emptyset$, where p^σ_τ is the probability that a random embedding of σ to $\tau^\emptyset$ (for the labelling distribution) equals τ.

Example 2.5. Here are a few examples of σ-flags, where $\sigma = 123$ is the CCW chirotope of size 3:



For every $\mathcal{Z}$-chirotope σ and every limit of order types ℓ, we have the following important identity [30, Lemma 3.11]:

$$\forall \tau \in \mathcal{A}^\sigma, \quad \ell\left([\![\tau]\!]_\sigma\right) = \ell\left([\![\sigma]\!]_\sigma\right) \int_{\phi^\sigma \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})} \phi^\sigma(\tau)\, \mathrm{d}\mathbf{P}^\sigma_\ell, \tag{8}$$

which represents the fact that one can sample $[\![\tau]\!]_\sigma$ by first picking a copy of σ at random, and then, conditioning on the choice of σ, extend it to a copy of τ. Equation (8) in particular implies that $\ell\left([\![C^\sigma]\!]_\sigma\right) \geqslant 0$ for any $C^\sigma \in \mathcal{A}^\sigma$ such that $\phi^\sigma(C^\sigma) \geqslant 0$ almost surely for $\phi^\sigma \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$, relatively to $\mathbf{P}^\sigma_\ell$. It follows that for every limit of order types ℓ and every $\mathcal{Z}$-chirotope σ,

$$\forall C^\sigma \in \mathcal{A}^\sigma, \quad \ell\left(\left[\!\!\left[(C^\sigma)^2\right]\!\!\right]_\sigma\right) \geqslant 0. \tag{9}$$

2.3 The semidefinite method for order types

The operator $[\![\cdot]\!]_\sigma$ is linear, so for every $\phi \in \mathrm{Hom}^+(\mathcal{A},\mathbf{R})$, every $A_1^\sigma, A_2^\sigma, \dots, A_I^\sigma \in \mathcal{A}^\sigma$, and every non-negative reals $z_1, z_2, \dots, z_I$, we have

$$\phi\left(\left[\!\!\left[\sum_{i=1}^{I} z_i \cdot (A_i^\sigma)^2\right]\!\!\right]_\sigma\right) \geqslant 0.$$

Every real (symmetric) positive semidefinite matrix M of size $n \times n$ can be written as $\sum_{i=1}^n \lambda_i u_i^T u_i$ where $\lambda_1, \dots, \lambda_n$ are non-negative real numbers and $u_1, \dots, u_n$ orthonormal vectors of $\mathbf{R}^n$. It follows that for every finite set of flags $S \subseteq \mathcal{O}^\sigma$ and for every real (symmetric) positive semidefinite matrix M of size $|S| \times |S|$, we have $\phi\left([\![v_S^T M v_S]\!]_\sigma\right) \geqslant 0$, where v_S is the vector in $(\mathcal{A}^\sigma)^{|S|}$ whose ith coordinate equals the ith element of S (for some given order). This recasts the search for a good "positive" quadratic combination as a semidefinite programming problem.

Let N be an integer, $f = \sum_{\omega\in\mathcal{O}_N} f_\omega \omega$ some target function, and $\sigma_1, \dots, \sigma_k$ a finite list of chirotopes so that $|\sigma_i| \equiv N \mod 2$. For each $i \in [k]$, let v_i be the $|\mathcal{X}^{\sigma_i}_{(N+|\sigma_i|)/2}|$-dimensional vector with ith coordinate equal to the ith element of $\mathcal{X}^{\sigma_i}_{(N+|\sigma_i|)/2}$. We look for a real b as large as possible subject to the constraint that there are k real (symmetric) positive semidefinite matrices $M_1, M_2, \dots, M_k$, where M_i has size $|v_i| \times |v_i|$, so that

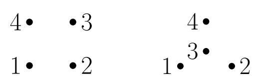
$$\forall \omega \in \mathcal{O}_N, \quad f_\omega \geqslant a_\omega \qquad \text{where} \qquad \sum_{\omega\in\mathcal{O}_N} a_\omega \omega = \sum_{i\in[k]} [\![v_i^T M_i v_i]\!]_{\sigma_i} + b \sum_{\omega\in\mathcal{O}_N} \omega. \tag{10}$$

The values of the real numbers a_ω are determined by b, the entries of the matrices $M_1, M_2, \dots, M_k$, the splitting probabilities $p(\tau', \tau''; \tau)$, where $\tau', \tau'' \in \mathcal{X}^{\sigma_i}_{(N+|\sigma_i|)/2}$ and $\tau \in \mathcal{X}^{\sigma_i}_N$, and the probabilities $p^{\sigma_i}_\tau$, where $\tau \in \mathcal{O}^{\sigma_i}_N$. Moreover, finding the maximum value of b and the entries of the matrices M_i can be formulated as a semidefinite program.

Effective semidefinite programming for flags of order types. In order to use a semidefinite programming software for finding a solution of programs in the form of (10), it is enough to generate the sets $\mathcal{O}_N$ and $\mathcal{X}^{\sigma_i}_N$, the split probabilities $p(\tau', \tau''; \tau)$, where $\tau', \tau'' \in \mathcal{X}^{\sigma_i}_{(N+|\sigma_i|)/2}$ and $\tau \in \mathcal{X}^{\sigma_i}_N$, and the probabilities $p^{\sigma_i}_\tau$, where $\tau \in \mathcal{O}^{\sigma_i}_N$.

We generated the sets and the values by brute force up to $N = 8$. The only non-trivial algorithmic step is deciding whether two order types, represented by point sets, are equal. This can be done by computing some canonical ordering of the points that turn two point sets with the same order type into point sequences with the same chirotope. A solution taking $O(n^2)$ time was proposed by Aloupis et al. [7]; the method that we implemented takes time $O(n^2 \log n)$ and seems to be folklore (we learned it from Pocchiola and Pilaud). For solving the semidefinite program itself, we used a library called CSDP [10]. The input data for CSDP was generated using the mathematical software SAGE [35].

Setting up the semidefinite programs. In the rest of this section we work with $N = 8$ and use chirotopes labelled $\sigma_1, \sigma_2, \dots, \sigma_{24}$ with σ_1 being the empty chirotope, σ_2 the only chirotope of size two, σ_3 and σ_4 the two chirotopes of size 4 depicted on the left, and $\sigma_5, \dots, \sigma_{24}$ a fixed set of 20 chirotopes of size 6 so that $\mathcal{O}_6 = \{\sigma_5^\emptyset, \dots, \sigma_{24}^\emptyset\}$. Note that since $|\mathcal{O}_6| = 20$, what follows does not depend on the choices made in labelling $\sigma_5, \dots, \sigma_{24}$. The vectors $v_1, \dots, v_{24}$ described in the previous paragraph for this choice of N and σ_i's have respective lengths 2, 44, 468, 393, 122, 112, 114, 101, 101, 103, 106, 103, 103, 120, 102, 108, 94, 90, 91, 91, 95, 95, 92 and 104.



Computations proving Proposition 1 and Proposition 2. We solved two semidefinite programs with the above choice of parameters for $f = \sum_{\omega\in\mathcal{O}_8} p(\diamond_5, \omega)$ and $f = \sum_{\omega\in\mathcal{O}_8} p(\diamond_6, \omega)$ and obtained real symmetric positive semidefinite matrices $M_1, \dots, M_{24}$ and $M'_1, \dots, M'_{24}$ with rational entries so that

$$\sum_{\omega\in\mathcal{O}_8} p(\diamond_5, \omega)\omega \geqslant \sum_{i\in[24]} [\![v_i^T M_i v_i]\!]_{\sigma_i} + \frac{15715211616602583691}{258254417031933722624} \sum_{\omega\in\mathcal{O}_8} \omega,$$

and

$$\sum_{\omega\in\mathcal{O}_8} p(\diamond_6,\omega)\omega \geqslant \sum_{i\in[24]} [\![v_i^T M_i' v_i]\!]_{\sigma_i} + \frac{67557324685725989}{36893488147419103232} \sum_{\omega\in\mathcal{O}_8} \omega.$$

The lower bounds on c_5 and c_6 then follow from Equation (2).

Assume (without loss of generality) that $\mathcal{O}_6 = \{\omega_{6,1}, \omega_{6,2}, \dots, \omega_{6,20}\}$. Solving two semidefinite programs, we obtained real symmetric positive semidefinite matrices $M_1, \dots, M_{24}$ and $M_1', \dots, M_{24}'$ as well as non-negative rational values $d_1, \dots, d_{20}$ and $d_1', \dots, d_{20}'$ so that

$$\sum_{j\in[20]} d_j \left(\omega_{6,j} - \frac{1}{32}\sum_{\omega\in\mathcal{O}_8}\omega\right) + \sum_{i\in[24]} [\![v_i^T M_i v_i]\!]_{\sigma_i} < 0$$

and

$$\sum_{j\in[20]} d_j' \left(-\omega_{6,j} + \frac{1}{18}\sum_{\omega\in\mathcal{O}_8}\omega\right) + \sum_{i\in[24]} [\![v_i^T M_i' v_i]\!]_{\sigma_i} < 0.$$

They respectively imply that there is no $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ such that, $\ell(\omega) \geqslant 1/32$ for every $\omega \in \mathcal{O}_6$, or such that $\ell(\omega) \leqslant 1/18$ for every $\omega \in \mathcal{O}_6$. Altogether this proves Proposition 2 with an imbalance bound of $32/18 > 1.77$. The better bound of Proposition 2 is obtained by a refinement of this approach where the order types with minimum and maximum probability are prescribed; this requires solving over 700 semidefinite programs.

The numerical values of the entries of all the matrices $M_1, \dots, M_{24}$ and coefficients $d_1, \dots, d_{20}$ mentioned above can be downloaded from the web page http://honza.ucw.cz/proj/ordertypes/. In fact, the matrices $M_1, \dots, M_{24}$ are not stored directly, but as an appropriate non-negative sum of squares, which makes the verification of positive semidefiniteness trivial. To make an independent verification of our computations easier, we created sage scripts called `verify_prop*.sage`, available from the same web page.

3 Representation of limits by measures

Corollary 8 asserts that every probability (or finite Borel measure) over $\mathbf{R}^2$ that charges no line defines a limit of order types. Going in the other direction, we say that a measure μ *realizes* a limit of order types ℓ if $\ell_\mu = \ell$. We examine here two questions: does every limit of order types enjoy such a realization and, for those that do, what does the set of measures realizing them look like? We answer the first question negatively in Section 3.2. We then give partial answers to the second question in Section 3.3 and Section 3.4. Every measure of $\mathbf{R}^2$ or on $\mathbf{S}^2$ that we consider is defined on the Borel σ-algebra.

3.1 Spherical geometry

If μ is a measure that charges no line and $f\colon \operatorname{supp}\mu \to \mathbf{R}^2$ is an injective map that preserves orientations, then $\mu' = \mu \circ f^{-1}$ is another measure that realizes the same limit as μ. A map that preserves orientations must preserve alignments, and therefore coincides locally with a projective map. Note, however, that if we fix two points $p, q \in \mathbf{R}^2$ and take a third point r "to infinity" in directions $\pm\vec{u}$, the orientation of the triple (p, q, r) is different for $+\vec{u}$ and for $-\vec{u}$. The appropriate geometric setting is therefore not projective, but spherical.

The spherical model. For convenience, we write $\{z > 0\}$ for $\{\,(x,y,z) \in \mathbf{R}^3 \mid z > 0\,\}$ and, similarly, $\{z < 0\}$ and $\{z = 0\}$ stand for $\{\,(x,y,z) \in \mathbf{R}^3 \mid z < 0\,\}$ and for $\{\,(x,y,z) \in \mathbf{R}^3 \mid z = 0\,\}$, respectively. Consider the standard identification of $\mathbf{R}^2$ with the open upper half of the unit sphere in $\mathbf{R}^3$ given by the bijection:

$$\iota\colon \left\{ \begin{array}{ccc} \mathbf{R}^2 & \to & \mathbf{S}^2 \cap \{z > 0\} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \frac{1}{\sqrt{x^2+y^2+1}} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \end{array} \right. .$$

The map ι sends the lines of $\mathbf{R}^2$ to the intersection of the half-space $z > 0$ with the great circles of $\mathbf{S}^2$, that is circles cut out by planes through the center of that sphere, other than the circle contained in the plane $z = 0$. We can define the orientation of a triple (p, q, r) of points of $\mathbf{S}^2$ as the sign of $\det(p, q, r)$, so that it coincides, through ι, with the planar notion. Naturally, ι transports planar measures that charge no line into spherical measures, supported on $\mathbf{S}^2 \cap \{z > 0\}$, that charge no great circle.

Spherical transforms. Let $X, Y \subset \mathbf{R}^2$. A function $f \colon X \to Y$ is a *spherical transform* if there exist a direct affine transform g and a rotation $h \in SO(\mathbf{R}^3)$ such that $h(\iota(X)) \subseteq \{z > 0\}$ and $f = (g \circ \iota^{-1} \circ h \circ \iota)_{|X}$. In words, spherical transforms are compositions of direct affine transforms with a lifting of $\mathbf{R}^2$ to the upper half of $\mathbf{S}^2$, a rotation that keeps the image of the lifting within that upper half-sphere, then a backprojection from that upper sphere to $\mathbf{R}^2$. Notice that f is defined over all of $\mathbf{R}^2$ if h preserves the great circle $\mathbf{S}^2 \cap \{z = 0\}$, and on the open halfplane $\iota^{-1}\left(\{z > 0\} \cap h^{-1}(\{z > 0\})\right)$ otherwise.

Lemma 3.1. *If μ is a measure in $\mathbf{R}^2$ that charges no line and f is a spherical transform defined over* $\operatorname{supp}\mu$*, then $\ell_\mu = \ell_{\mu \circ f^{-1}}$.*

Proof. The statement follows from the observation that f preserves orientations over $\operatorname{supp}\mu$. To see this, let $p, q, r \in \operatorname{supp}\mu$ and let us decompose $f = (g \circ \iota^{-1} \circ h \circ \iota)_{|\operatorname{supp}\mu}$ where g is a direct affine transform and $h \in SO(\mathbf{R}^3)$. Translations preserve orientation. Writing orientations as determinants reveals that direct linear transforms also preserve orientations. Altogether, the direct affine transform g thus also preserves orientations. It follows that

$$\begin{aligned}
[f(p), f(q), f(r)] &= [g \circ \iota^{-1} \circ h \circ \iota(p), g \circ \iota^{-1} \circ h \circ \iota(q), g \circ \iota^{-1} \circ h \circ \iota(r)] \\
&= [\iota^{-1} \circ h \circ \iota(p), \iota^{-1} \circ h \circ \iota(q), \iota^{-1} \circ h \circ \iota(r)] \\
&= \operatorname{sign}\det\begin{pmatrix} h \circ \iota(p) & h \circ \iota(q) & h \circ \iota(r) \end{pmatrix} \\
&= \operatorname{sign}\left(\det(h)\det\begin{pmatrix} \iota(p) & \iota(q) & \iota(r) \end{pmatrix}\right) \\
&= \operatorname{sign}\left(\det(h)\right)[p, q, r],
\end{aligned}$$

and f preserves orientations because h is direct. □

Lemma 3.2. *Let C be a closed convex set and L a line in $\mathbf{R}^2$. If $d(C, L) > 0$ and C contains no ray parallel to L, then there exists a spherical transform f defined over C such that $f(C)$ is compact.*

Proof. Let X be the open halfpane bounded by L that contains C. The image $\iota(L)$ coincides with the intersection of a great circle Γ with $\{z > 0\}$. Moreover, $\iota(X)$ is contained in one of the hemispheres bounded by Γ. Let $h \in SO(\mathbf{R}^3)$ be a rotation that maps Γ to $\mathbf{S}^2 \cap \{z = 0\}$ and maps the interior of $\iota(X)$ to (a subset of) the region $\{z > 0\}$. We set $f = (\iota^{-1} \circ h \circ \iota)_{|X}$.

We argue that if $f(C)$ is unbounded, then $d(C, L) = 0$ or C contains a ray parallel to L. Let $(p_n)_{n \in \mathbf{N}} \subset C$ such that $(f(p_n))_{n \in \mathbf{N}}$ is unbounded. Since $\mathbf{S}^2$ is compact, up to taking a subsequence we can assume that $\{h \circ \iota(p_n)\}$ converges. Its limit u must belong to $\{z = 0\}$, as otherwise $\iota^{-1}(u) \in \mathbf{R}^2$, which would contradict the assumption that $\{f(p_n)\}$ is unbounded. This means that $(\iota(p_n))$ converges to a limit $u' = h^{-1}(u) \in \Gamma$. If $u' \in \iota(L)$, then $d(C, L) = 0$. If $u' \notin \iota(L)$, then $u' \in \Gamma \cap \{z = 0\}$ and we can write $u' = (u_x : u_y : 0)$ where $\tilde{u}' = (u_x, u_y)$ is a unit direction vector for L to which the vectors $\left(\frac{p_1 p_n}{\|p_1 p_n\|}\right)_{n \in \mathbf{N}}$ converge. For any positive real number $t > 0$ and for any large enough integer n we have $\|p_1 p_n\| > t$, so the segment $p_1 p_n$ contains a point q_n with $\|p_1 q_n\| = t$. The sequence $\{q_n\}$ is contained in C by convexity, and converges to $p_1 + t\tilde{u}$. Since C is closed, this limit is also in C. The ray $p_1 + \mathbf{R}^+ \tilde{u}'$ is thus in C, which contradicts the assumption that C contains no ray parallel to L. We deduce that $f(C)$ is bounded and thus, altogether, the assumptions on L and C guarantee that $f(C)$ is compact. □

3.2 A limit that is not representable

For convenience, let us repeat the definition of $\ell_\odot$ from the introduction. For $t \in (0, 1)$ let $\odot_t$ be the probability distribution over $\mathbf{R}^2$ supported on two concentric circles, with respective radii 1 and t where each of the two circles has $\odot_t$-measure $1/2$, distributed proportionally to the length on that circle. We define $\ell_{\odot_t}$ to be the limit of order types associated to $\odot_t$. We define $\ell_\odot$ to be the limit of an arbitrary convergent sub-sequence of $(\ell_{\odot_{1/n}})_{n\in\mathbf{N}}$.

We use the following two facts about measures, limits of order types, and sequences of point sets. Let μ be a measure of the plane, and P_n a set of n points of the plane sampled i.i.d. according to μ. The *empirical measure* associated to P_n is the measure μ_{P_n} defined on all μ-measurable subsets A of $\mathbf{R}^2$ by

$$\mu_{P_n}(A) = \frac{1}{n}|P_n \cap A|.$$

Lemma 3.3. *Let μ be a measure that does not charge lines and for $n \geqslant 1$, let P_n be a set of points sampled independently from μ, with $|P_n| \underset{n\to\infty}{\to} \infty$. The empirical measure of P_n almost-surely weakly converges to μ.*

The next lemma follows from ideas similar to those used in the proof of Proposition 7.

Lemma 3.4. *Let μ be a measure that does not charge lines. If $(P_n)_{n\in\mathbf{N}}$ is a sequence of point sets with $|P_n| = n^2$ and whose empirical measures converge to μ as n goes to infinity, then the order type of P_n converges to ℓ_μ as n goes to infinity.*

3.2.1 Reduction to compact support

Define the *peeling depth* of an order type ω as the largest integer k such that $P_k \neq \emptyset$ in a sequence $\{P_i\}_{i\in\mathbf{N}}$ where P_1 realizes ω and $P_{i+1} = P_i \setminus \mathrm{conv}(P_i)$ for each $i \geqslant 1$. For instance, k points in convex position have peeling depth 1, and any k points chosen from the support of $\odot_t$ have peeling depth at most 2.

Lemma 3.5. *Fix $k \in \mathbf{N}$ and let ℓ be a limit of order types such that $\ell(\omega) = 0$ for any order type ω of peeling depth greater than k. If there exists a probability measure μ realizing ℓ, then there exists one with compact support.*

Proof. Let μ be a probability measure that realizes ℓ and let $C = \mathrm{conv}\,\mathrm{supp}\,\mu$. If an order type ω can be realized by points in the support of μ, then $\ell_\mu(\omega) > 0$. Indeed, for a positive and sufficiently small ε, any perturbation of the points of amplitude at most ε also realizes ω. It follows that the support of μ cannot contain any finite subset with peeling depth $k+1$.

We first argue that $C \neq \mathbf{R}^2$. We proceed by contradiction, so suppose $C = \mathbf{R}^2$. A theorem of Steinitz [36] asserts that for any subset $A \subseteq \mathbf{R}^2$, any point in the interior of $\mathrm{conv}\,A$ belongs to the interior of the convex hull of at most four points of A. For any point $p \in \mathbf{R}^2$, there is thus a quadruple $Q(p) \subset \mathrm{supp}\,\mu$ such that p belongs to the interior of $\mathrm{conv}\,Q(p)$. We now define $Y_1 = \{p_1\}$ where p_1 is an arbitrary point of $\mathrm{supp}\,\mu$, and $Y_{i+1} = \cup_{p\in Y_i} Q(p)$ for each $i \geqslant 1$. The set $\cup_{i=1}^{k+1} Y_i$ is contained in the support of μ and has peeling depth at least $k+1$, a contradiction. Consequently, C cannot be $\mathbf{R}^2$.

We can also rule out the case where C is a halfplane. Suppose the contrary and let L be the line bounding this halfplane. We can follow the argument showing that $C \neq \mathbf{R}^2$, except if at some stage, some point of Y_i belongs to L. Assume that this is the case for $p \in Y_i$, and note that we can assume that $i \geqslant 2$. There exists $p' \in Y_{i-1}$ such that $p \in Q(p')$; remark that the peeling depth of $\cup_{j=1}^{i} Y_j$ remains at least i if we replace p in Y_i by any point q in $\mathrm{supp}\,\mu$ such that p' belongs to the interior of $\mathrm{conv}\,Q(p') \setminus \{p\} \cup \{q\}$. There exists $\varepsilon_1 > 0$ such that for any $z \in B(p, \varepsilon_1)$, the point p' belongs to the interior of $\mathrm{conv}\,Q(p') \setminus \{p\} \cup \{z\}$. Set $\tau = \mu(B(p, \varepsilon_1))$. Since $\mu(L) = 0$, there exists $\delta > 0$ such that the set of points within distance δ of L has measure at most $\varepsilon/2$. The part of $B(p, \varepsilon_1)$ at distance at least δ from L therefore has measure at least $\varepsilon_1/2$. A dichotomy argument produces a point q in that part that is in the support of μ. We can replace p by q in Y_i and continue with the construction. It follows that C cannot be a halfplane.

For any point $p \in \partial C$ there is a line supporting C and containing p. Thus, since C is neither $\mathbf{R}^2$ nor a halfplane, C is contained in the intersection of two non-parallel, closed, halfplanes H_1 and H_2. Let L be a line disjoint from $H_1 \cap H_2$ that is parallel to neither ∂H_1 nor ∂H_2. Note that Lemma 3.2 applies, so there is a spherical transform f defined over $\mathrm{supp}\,\mu$. Lemma 3.1 ensures that the pushback of μ by f is a measure with compact support that defines the same limit of order types as μ. $\square$

Corollary 10. *If $\ell_\odot$ is realizable by a measure, then it is realizable by one with compact support.*

Proof. For any $t \in (0,1)$, if an order type ω has peeling depth at least 3, then $\ell_{\odot_t}(\omega) = 0$. It follows that $\ell_\odot$ vanishes on every order type of peeling depth three or more and therefore Lemma 3.5 yields the conclusion. $\square$

3.2.2 Excluding realizations with compact support

Assume by contradiction that there exists a measure μ that realizes $\ell_\odot$; in particular, μ charges no line. Let P_n be a set of $N = n^2$ random points sampled independently from μ. By Lemma 3.3, the empirical measures of $(P_n)_n$ weakly converge to μ and by Lemma 3.4, their order types converge to ℓ. We write $C_n = P_n \cap \partial \operatorname{conv} P_n$ for the extreme points of P_n and $I_n = P_n \setminus C_n$ for the remaining points.

Lemma 3.6. *With high probability, $0.49N \leqslant |C_n| \leqslant 0.51N$ and at least $f(N)$ edges of $\operatorname{conv} P_n$ form, with another vertex of C_n, a triangle that contains I_n, where $\lim_{N\to\infty} f(N) = +\infty$.*

Before we prove Lemma 3.6, let us see how to use it by establishing Proposition 3.

Proof of Proposition 3. Suppose, contrary to the statement, that there exists a measure μ that realizes $\ell_\odot$. In particular, μ charges no line. By Corollary 10, we can assume μ to have compact support. Let α be the length of the boundary of $\operatorname{conv}\operatorname{supp}\mu$. Note that $\alpha < \infty$ since μ has compact support. Since $\operatorname{conv} P_n$ is contained in $\operatorname{conv}\operatorname{supp}\mu$, the boundary of $\operatorname{conv} P_n$ also has length at most α (this follows for instance from the Cauchy-Crofton formula [31, Chapter 3, § 2]). Lemma 3.6 implies that there is an edge of $\operatorname{conv} P_n$ of length at most $\varepsilon = O\left(\frac{\alpha}{f(N)}\right)$ that forms, together with another vertex of C_n, a triangle containing I_n. Let L_n be the line supporting one of the edges of this triangle that is not on $\partial \operatorname{conv} P_n$. With high probability, at least $0.49|P_n|$ points of P_n lie within distance at most ε of L_n. The set of lines intersecting $\operatorname{conv}\operatorname{supp}\mu$ is compact, so the sequence $(L_n)_n$ contains a subsequence $\left(L_{i(n)}\right)_n$ converging to a line L. Since the empirical measures of $\left(P_{i(n)}\right)_n$ converge to μ, we deduce that $\mu(L) \geqslant 0.49$, a contradiction. $\square$

It remains to prove Lemma 3.6.

Proof of Lemma 3.6. Let us give another model of random point sets whose order-type is distributed like P_n. Let $\mathfrak{C}_t$ be the circle of radius t in $\mathbf{R}^2$ centered at the origin. We let $P_n^{1/2}$ be a set of $N = n^2$ random points sampled independently from $\odot_{\frac{1}{2}}$ and for $t \in \left(0, \frac{1}{2}\right)$ we let P_n^t be the point set obtained from $P_n^{1/2}$ by scaling the points on $\mathfrak{C}_{1/2}$ from the origin by a factor $2t$. Observe that P_n^t is distributed like N points sampled independently from $\odot_t$. Moreover, almost surely, there exists (a random variable) $t^* > 0$ such that the order types of P_n^t are identical for $t \in (0, t^*)$. As $t \to 0$, the sequence $(\ell_{\odot_t})_t$ converges to $\ell_\odot = \ell_\mu$, so the order type of $P_n^{t^*}$ is distributed like P_n.

Let us define $C_n^{t^*} = P_n^{t^*} \cap \partial \operatorname{conv} P_n^{t^*}$ and $I_n^{t^*} = P_n^{t^*} \setminus C_n^{t^*}$. We assert that with high probability

(a) $0.49N \leqslant |P_n^{t^*} \cap \mathfrak{C}_1| \leqslant 0.51N$; and

(b) $C_n^{t^*} = P_n^{t^*} \cap \mathfrak{C}_1$.

Since the order types of P_n and $P_n^{t^*}$ are identically distributed, this claim implies in particular the first part of the statement.

Assertion (a) follows from Hoeffding's inequality because the repartition of the points of $P_n^{t^*}$ between $\mathfrak{C}_1$ and $\mathfrak{C}_{t^*}$ follows a binomial law with parameter $1/2$. To prove Assertion (b), let us bound the probability that there exists a point from $\mathfrak{C}_{t^*}$ on $\partial \operatorname{conv} P_n^{t^*}$. If such a point exists, then there is a line through that point that cuts an arc of $\mathfrak{C}_1$ with no point of $P_n^{t^*}$. Any line through a point of $\mathfrak{C}_{t^*}$ cuts an arc of $\mathfrak{C}_1$ with length at least a fraction $\alpha(t^*)$ of the length of $\mathfrak{C}_1$. We can now cover $\mathfrak{C}_1$ by smaller intervals of length $\alpha(t^*)/2$, and we observe that $P_n^{t^*}$ must miss at least one of these smaller intervals. For t^* small enough, 8 smaller intervals suffice, and $C_n^{t^*} = P_n^{t^*} \cap \mathfrak{C}_1$ with probability least $1 - 8\left(\frac{15}{16}\right)^N$. This proves our assertion.

Let us now turn our attention to the second part of the statement of the lemma. We write $C_n^{t^*} = \{p_1, p_2, \dots\}$ where the points are labelled in clockwise order along $\partial \operatorname{conv} P_n^{t^*}$, starting in an arbitrary place. We label the points cyclically, setting $p_{i+|C_n^{t^*}|} = p_i$. Almost surely, for every i the antipodal of p_i

is not in $C_n^{t^*}$ so there exists a unique edge p_jp_{j+1} such that $p_ip_jp_{j+1}$ contains $I_n^{t^*}$. We set $j(i) = j$. Our task is to bound from below the number of distinct edges $p_{j(i)}p_{j(i)+1}$.

Let us associate to every point p_i the random variables $s_i = |i - j(i)|$. We assert that with high probability

$$\frac{|C_n^{t^*}|}{2} - N^{\frac{2}{3}} \leqslant \min_i s_i \leqslant \max_i s_i \leqslant \frac{|C_n^{t^*}|}{2} + N^{\frac{2}{3}}. \tag{11}$$

Observe that (11) implies that $j(i)$ is different for any two indices $i \in \{0, 2.1N^{2/3}, 4.1N^{2/3}, \dots, (k + 0.1)N^{2/3}\}$ where $k \in \mathbf{N}$ such that $(k + 0.1)N^{2/3} \leqslant 0.49N < (k + 1.1)N^{2/3}$. It follows that with high probability $j(i)$ takes at least $\Omega(N^{1/3})$ different values, proving the second statement.

It remains to establish (11). The random variable s_i counts the number of points on one of the halfcircles of $\mathfrak{C}_1$ bounded by p_i and its antipodal. Consider a different labeling $C_n^{t^*} = \{q_1, q_2, \dots\}$ where q_1 is an arbitrary point of $C_n^{t^*}$ and where the lines through the origin and $q_1, q_2, \dots$ are in clockwise order. Each s_i counts the number of $q_j \neq p_i$ on the clockwise half-circle from p_i to its antipodal. Observe that the position of the lines spanned by the points q_i and their antipodals are irrelevant; in fact, fixing the lines arbitrarily and picking each q_i uniformly among the two candidate points on its assigned line leads to the same distribution of the variables s_i. A purely combinatorial, equivalent, description is thus the following. Let $W = w_1w_2 \dots w_{2k}$ be a random circular word on $\{0, 1\}^{2k}$, where $k = |C_n^{t^*}|$, $w_{i+k} = 1 - w_i$ and the letters $w_1, w_2, \dots, w_k$ are chosen independently and uniformly in $\{0, 1\}$. The random variable s_i has the same distribution as $w_i + w_{i+1} + \cdots + w_{i+k-1}$. Hoeffding's inequality and a union bound then yield that with high probability, $\max_i s_i$ is at most $\frac{k}{2} + o(k)$ (and, by symmetry, that $\min_i s_i$ is at least $\frac{k}{2} - o(k)$). $\square$

3.3 The rigidity theorem

In this section, we prove Theorem 4, which states that probability distributions giving rise to the same limit of order types are projectively equivalent, provided that their support have non-empty interiors. We proceed by adapting the notion of kernels to three-variable functions. This notion generalizes the notion of geometric limit firstly by allowing probability measures of any topological space J instead of probability measures on $\mathbf{R}^2$ and secondly by using any real-valued function of J^3 instead of the chirotope of the plane. Our main tool[2] on kernels is Theorem 13, which asserts that two kernels have same sampling distribution if and only if they are isomorphic. We introduce the notions related to kernels and state Theorem 13 in Section 3.3.1, but defer all proofs to Section 3.3.3, after we have examined kernels defined from geometric probabilities and order types and deduced Theorem 4 (Section 3.3.2).

3.3.1 Rigidity for kernels (overview)

A *kernel* is a triple (J, μ, W) where (J, μ) is a probability space and W is a measurable map from J^3 to $\mathbf{R}$ such that $\sup_{J^3} |W| < \infty$. For the sake of readability, W can stand for a whole kernel (J, μ, W) when there is no ambiguity about the set and the measure involved. We are, in particular, interested in kernels based on chirotopes, that is, where $J = \mathbf{R}^2$, μ is a measure that charges no line, and W is a map to $\{-1, 0, 1\}$ given by the orientation.

Our first ingredient is a rigidity theorem for kernels (Theorem 13), which asserts that if two kernels have the same density functions (the analogue, for general kernels, of the limit of order types ℓ_μ associated with a measure μ), then they are essentially equal. We present here the main notions and results, and postpone the proofs to Section 3.3.3.

Density function. A *step function* over a probability space (J, μ) is a kernel (J, μ, U) for which there is a partition $(P_1, \dots, P_m)$ of J such that the function U is constant on $P_i \times P_j \times P_k$ for every $(i, j, k) \in [m]^3$. Since the set $U^{-1}(\{x\})$ is measurable for every $x \in \mathbf{R}$, we may always assume that P_i is measurable for $i \in [m]$.

[2] This result and its proof are direct adaptations of similar results on two-dimensional kernels and graphons; the presentation reaches the proof of Proposition 12 as directly as possible, and we refer to the monograph of Lovász [25] for (a lot) more details on two-dimensional kernels.

Let (J, μ, W) be a kernel and n a positive integer. Any n-tuple $S = (x_1, \ldots, x_n) \in J^n$ defines a function

$$\mathbf{H}(W, S)\colon \begin{cases} [n]^3 & \to \quad \mathbf{R} \\ (i,j,k) & \mapsto \quad W(x_i, x_j, x_k). \end{cases}$$

We observe that if μ_n^u is the counting measure on $[n]$, then $([n], \mu_n^u, \mathbf{H}(W,S))$ is a step function for the partition of $[n]$ into singletons. We let $\mathbf{H}(W,n)$ be the random step function $\mathbf{H}(W,S)$, where S is a μ^n-random tuple. We define $\mathcal{H}_n$ to be $\{H\colon [n]^3 \to \mathbf{R}\}$. For every $H \in \mathcal{H}_n$, the *density* $t(H,W)$ of H in W is the probability that $\mathbf{H}(W,n) = H$. We set $\mathcal{H} = \bigcup_{n \geqslant 1} \mathcal{H}_n$.

Norms and distance. A kernel (J, μ, W) has associated L_1- and L_∞-norms respectively defined by

$$\|W\|_1 = \int_{J^3} |W(x,y,z)| \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(z) \qquad \text{and} \qquad \|W\|_\infty = \sup_{J^3} |W|\,.$$

The L_1- and L_∞-norms define distances only between kernels with the same underlying probability space. To define a distance between any two kernels (J_1, μ_1, W_1) and (J_2, μ_2, W_2), we first map them both to a common probability space by measure-preserving[3] maps (like the Gromov-Hausdorff distance between metric spaces). Specifically, we define

$$\mathrm{d}(W_1, W_2) = \inf \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1$$

where the infimum is taken over all choices of a probability space (J, μ) and a pair of measure-preserving maps $\phi_i\colon J \to J_i$ for $i \in \{1,2\}$, and where W^ϕ is defined by $W^\phi(x,y,z) = W(\phi(x), \phi(y), \phi(z))$. This infimum does not change if we assume that μ is a coupling measure on $J = J_1 \times J_2$ and that ϕ_i is the natural projection of J on J_i for $i \in \{1,2\}$. Further, a theorem by Janson [21, Theorem 2] ensures that, up to a measure-preserving map, we can also assume that J_i is $[0,1]$ and μ_i is the Lebesgue measure for $i \in \{1,2\}$.

Rigidity of kernels. We now turn our attention to rigidity results for kernels. Unless otherwise specified, all topological hypothesis are to be understood in $\mathbf{R}^2$ endowed with the usual topology; this is in particular the case when we consider a measure with "non-empty interior". We start with the following.

Theorem 11. *If (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are two kernels such that $t(H, W_1) = t(H, W_2)$ for every function $H \in \mathcal{H}$, then $\mathrm{d}(W_1, W_2) = 0$.*

Theorem 11 makes no assumption on the measure, but the equality of W_1 and W_2 is only in measure, up to mapping to a common probability space. Stronger rigidity properties are possible under some regularity assumption on the kernels. Define two points x and x' in J to be *twins for* (J, μ, W) if $W(x,y,z) = W(x',y,z)$ for μ^2-almost every pair $(y,z) \in J^2$. The kernel (J, μ, W) is *twin-free* if it admits no twins.

Proposition 12. *Let (J_1, μ_1, W_1) and (J_2, μ_2, W_2) be two twin-free kernels. If $t(H, W_1) = t(H, W_2)$ for every function $H \in \mathcal{H}$, then there exist two sets $N_1 \subseteq J_1$ and $N_2 \subseteq J_2$ with $\mu_1(N_1) = 0 = \mu_2(N_2)$ and a bijection $\rho\colon J_1 \setminus N_1 \to J_2 \setminus N_2$ such that ρ and ρ^{-1} are measure preserving, and W_1 and W_2^ρ are equal μ_1^3-almost everywhere.*

For even more regular kernels, there are in fact distances on J_1 and J_2 such that a measure-preserving isometry exists. Specifically, the *neighborhood pseudo-distance* of a kernel (J, μ, W) is the function $d_W : J \times J \to \mathbf{R}^+$ given by

$$\begin{aligned} \forall (x,x') \in J^2, \qquad d_W(x,x') &= \|W(x,\cdot,\cdot) - W(x',\cdot,\cdot)\|_1 \\ &= \int_{J^2} |W(x,y,z) - W(x',y,z)| \,\mathrm{d}\mu(y)\,\mathrm{d}\mu(z). \end{aligned} \tag{12}$$

It is straightforward to check that d_W is reflexive and satisfies the triangular inequality, so it is a distance if and only if it is separated, *i.e.*, if (J, W, μ) is twin-free.

[3] If (J_1, μ_1) and (J_2, μ_2) are two probability spaces, a map $\rho\colon J_1 \to J_2$ is *measure preserving* if for every μ_2-measurable set $A' \subseteq J_2$, the set $A = \rho^{-1}(A')$ is μ_1-measurable and $\mu_1(A) = \mu_2(A')$.

A kernel (J, μ, W) is *pure* if (i) the map d_W is a distance on J, (ii) the metric space (J, d_W) is complete and separable, and (iii) the measure μ has *full support*, that is, every non-empty open set of (J, d_W) has positive μ-measure.

Theorem 13. *Let (J_1, μ_1, W_1) and (J_2, μ_2, W_2) be two pure kernels. If $t(H, W_1) = t(H, W_2)$ for every function $H \in \mathcal{H}$, then there exists an isometry $\rho\colon (J_1, d_{W_1}) \to (J_2, d_{W_2})$ such that ρ and ρ^{-1} are measure preserving, and W_1 and W_2^ρ are equal μ_1^3-almost everywhere.*

3.3.2 Kernels based on chirotopes

We now deduce Theorem 4 from Theorem 13 by studying kernels coming from chirotopes. We let χ be the map that sends a triple of points in $\mathbf{R}^2$ to their orientation. We consider kernels of the form $(\operatorname{supp}\mu, \mu, \chi_{|\operatorname{supp}\mu})$, and let $d_{\chi,\mu}$ be the associated neighborhood pseudo distance.[4]

Purity. Given two distinct points a and b in $\mathbf{R}^2$, we let (ab) be the line through a and b.

Lemma 3.7. *Let μ be a measure on $\mathbf{R}^2$. If $\operatorname{supp}\mu$ has non-empty interior, then the kernel $(\operatorname{supp}\mu, \mu, \chi)$ is twin-free.*

Proof. Let B be an open ball contained in $\operatorname{supp}\mu$, and consider two distinct points a and b in $\operatorname{supp}\mu$. There exists a line separating a from b and intersecting B (for otherwise B would have to be contained in the line (ab)). Let x and y be two distinct points in $(ab) \cap B$. By continuity of the determinant expressing orientations, there exists a positive real ϵ such that any line secant to both $B(x, \epsilon)$ and $B(y, \epsilon)$ also separates a from b. This implies that for any $x' \in B(x, \epsilon)$ and $y' \in B(y, \epsilon)$, we have $\chi(x', y', a) \neq \chi(x', y', b)$. Since $\mu^2\left(B(x, \epsilon) \times B(y, \epsilon)\right)$ is positive, it follows that a and b are not twins. ☐

Topologies. By Lemma 3.7, if the support of a measure μ has non-empty interior, then the kernel $(\operatorname{supp}\mu, \mu, \chi)$ is twin-free and $d_{\chi,\mu}$ is a distance. To show that this kernel is pure, it remains to control the topology induced by $d_{\chi,\mu}$ on $\operatorname{supp}\mu$.

Lemma 3.8. *Let μ be a compactly supported measure that charges no line. If $\operatorname{supp}\mu$ has non-empty interior, then the topologies induced on $\operatorname{supp}\mu$ by $d_{\chi,\mu}$ and $\|.\|_2$ are equivalent.*

Proof. We first remark that $d_{\chi,\mu}$ is a distance thanks to Lemma 3.7. We next argue that $d_{\chi,\mu}$ is continuous toward $\|.\|_2$. To do so, we consider a sequence (x_n) of points in $\mathbf{R}^2$ converging to some point x in the sense that $\lim_{n\to\infty} \|x_n - x\|_2 = 0$, and we prove that we also have $\lim_{n\to\infty} d_{\chi,\mu}(x_n, x) = 0$. For every pair $(y, z) \in \operatorname{supp}\mu^2$ such that x, y and z are not aligned, the set of points $x' \in \mathbf{R}^2$ satisfying the equation $\chi(x, y, z) = \chi(x', y, z)$ is an open half-space containing x. Consequently, $\chi(x_n, y, z)$ is eventually equal to $\chi(x, y, z)$. It follows that the sequence $f_n(y, z) = \chi(x, y, z) - \chi(x_n, y, z)$ tends to 0 for μ^2-almost every pair $(y, z) \in \operatorname{supp}\mu^2$. Indeed, this last statement is true if x, y and z are not aligned and the set of pairs (y, z) aligned to x is a μ^2-nullset because μ does not charge lines. We conclude by an application of the dominated convergence theorem that the sequence $d_{\chi,\mu}(x_n, x) = \int_{\operatorname{supp}\mu^2} f_n(y, z)\,\mathrm{d}\mu(y, z)$ tends to 0.

To conclude, it suffices to prove that if $d_{\chi,\mu}(x_n, x)$ tends to 0 for some sequence $(x_n)_{n\in\mathbf{N}} \in \operatorname{supp}\mu^{\mathbf{N}}$ and $x \in \operatorname{supp}\mu$ then $\|x_n - x\|_2$ tends to 0 when n goes to infinity. We use a classical compactness argument. Assume for the sake of contradiction that there is a sequence $(d_{\chi,\mu}(x_n, x))_{n\in\mathbf{N}}$ that tends to 0 while $(\|x_n - x\|_2)_{n\in\mathbf{N}}$ does not. Since $\operatorname{supp}\mu$ is compact, it is possible to extract a subsequence $(x_{\phi(n)})_{n\in\mathbf{N}}$ that converges with respect to $\|.\|_2$ to a limit $y \in \operatorname{supp}\mu$ different from x. Then, as seen above, $d_{\chi,\mu}(x_{\phi(n)}, y) \to 0$ and $d_{\chi,\mu}(x, y) = 0$. Since $d_{\chi,\mu}$ is a distance, it follows that $x = y$, which yields a contradiction. ☐

Corollary 14. *If μ is a measure that charges no line and such that $\operatorname{supp}\mu$ is compact and has non-empty interior in $\mathbf{R}^2$, then the kernel $(\operatorname{supp}\mu, \mu, \chi)$ is pure.*

Proof. By Lemma 3.7, we already know that $(\operatorname{supp}\mu, \mu, \chi)$ is twin-free, and therefore that $d_{\chi,\mu}$ is a distance. By Lemma 3.8, this distance defines the same topology on $\operatorname{supp}\mu$ as the Euclidean distance, so $(\operatorname{supp}\mu, d_{\chi,\mu})$ is complete and separable. Moreover, any open ball for $d_{\chi,\mu}$ contains an open ball for $\|\cdot\|_2$ so μ has full support. ☐

[4]We depart here from the notation introduced in Equation (12) because we manipulate several kernels based on the function χ, so we need to lift the ambiguity.

Alignments. To study how the map $\rho\colon \operatorname{supp}\mu_1 \to \operatorname{supp}\mu_2$ provided by Theorem 13 transports alignments, we reformulate alignments in topological terms. Let us define the sets

$$T^+ = \{(a,b,c) \in (\mathbf{R}^2)^3 \colon \chi(a,b,c) = 1\} \quad \text{and} \quad T^- = \{(a,b,c) \in (\mathbf{R}^2)^3 \colon \chi(a,b,c) = -1\}.$$

We define $\mathrm{cl}_{\|\|_2}(X)$ and $\partial_{\|\|_2} X$ to be the topological closure and the boundary of a set X for the topology induced by the Euclidean distance, respectively. From the expression of the orientation of a triple of points as a determinant, it comes

$$\begin{aligned}\{(a,b,c) \in (\mathbf{R}^2)^3 \colon \chi(a,b,c) = 0\} &= \partial_{\|\|_2} T^+ = \partial_{\|\|_2} T^- \\ &= \mathrm{cl}_{\|\|_2}(T^+) \setminus T^+ = \mathrm{cl}_{\|\|_2}(T^-) \setminus T^- = \mathrm{cl}_{\|\|_2}(T^+) \cap \mathrm{cl}_{\|\|_2}(T^-).\end{aligned} \tag{13}$$

We use a local version of this characterization of alignment, where we restrict T^+ or T^- to the support of a measure μ. Note that $(\operatorname{supp}\mu)^3$ may contain isolated aligned triples or other pathologies, so we have to take some care; we again rely on the assumption that $\operatorname{supp}\mu$ has non-empty interior.

Lemma 3.9. *Let X be a subset of $\mathbf{R}^2$ that contains an open ball B. A triple $t = (x,y,z) \in X^3$ with $x \in B$ is aligned if and only if it is in $\mathrm{cl}_{\|\|_2}(T^+ \cap X^3) \cap \mathrm{cl}_{\|\|_2}(T^- \cap X^3)$.*

Proof. If t is aligned, then there exist x' and x'' both in B arbitrarily close to x, and separated by the line (yz). It follow that exactly one of (x', y, z) and (x'', y, z) belongs to $T^+ \cap X^3$, and the other belongs to $T^- \cap X^3$. This implies that $t \in \mathrm{cl}_{\|\|_2}(T^+ \cap X^3) \cap \mathrm{cl}_{\|\|_2}(T^- \cap X^3)$. The other direction follows from the fact that

$$\mathrm{cl}_{\|\|_2}(T^+ \cap X^3) \cap \mathrm{cl}_{\|\|_2}(T^- \cap X^3) \subseteq \mathrm{cl}_{\|\|_2}(T^+) \cap \mathrm{cl}_{\|\|_2}(T^-)$$

and Equation (13). □

Proof of Theorem 4. We can now prove that if μ_1 and μ_2 are two compactly supported measures of $\mathbf{R}^2$ that charge no line, whose supports have non-empty interiors, and such that $\ell_{\mu_1} = \ell_{\mu_2}$, then there exists a projective transformation f such that $\mu_2 \circ f = \mu_1$.

Proof of Theorem 4. By Corollary 14, the kernels $(\operatorname{supp}\mu_i, \mu_i, \chi)$ are pure. Thus, by Theorem 13, there exists a measure-preserving isometry $\rho\colon (\operatorname{supp}\mu_1, d_{\chi,\mu_1}) \to (\operatorname{supp}\mu_2, d_{\chi,\mu_2})$ such that χ and χ^ρ are equal μ_1^3-almost everywhere. Moreover, Lemma 3.8 ensures that ρ is an homeomorphism from $(\operatorname{supp}\mu_1, \|\cdot\|_2)$ to $(\operatorname{supp}\mu_2, \|\cdot\|_2)$. It remains to analyze how ρ transports alignments.

First, let us remark that $\rho^3(T^+ \cap \operatorname{supp}\mu_1^3)$ and $T^- \cap \operatorname{supp}\mu_2^3$ are disjoint. To prove this, we note that the intersection of these two sets is open in $\operatorname{supp}\mu_2^3$ and has μ_2^3-measure 0: indeed, we see that χ and χ^ρ disagree on every element of $(T^+ \cap \operatorname{supp}\mu_1^3) \cap (\rho^{-1})^3(T^- \cap \operatorname{supp}\mu_2^3)$, and therefore

$$\mu_1^3((T^+ \cap \operatorname{supp}\mu_1^3) \cap (\rho^{-1})^3(T^- \cap \operatorname{supp}\mu_2^3)) = 0,$$

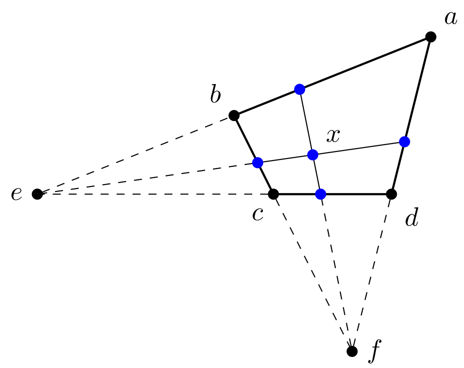
which implies that $\mu_2^3(\rho^3(T^+ \cap \operatorname{supp}\mu_1^3) \cap (T^- \cap \operatorname{supp}\mu_2^3)) = 0$. The conclusion follows because μ_2^3 has full support on $\operatorname{supp}\mu_2^3$. A symmetric argument yields that $\rho^3(T^- \cap \operatorname{supp}\mu_1^3)$ and $T^+ \cap \operatorname{supp}\mu_2^3$ are also disjoint.

Now, consider an aligned triple $t = (x,y,z) \in \operatorname{supp}\mu_1^3$ where x belongs to some open ball B_1 contained in $\operatorname{supp}\mu_1$. From Lemma 3.9, it comes that $t \in \mathrm{cl}_{\|\|_2}(T^- \cap \operatorname{supp}\mu_1^3) \cap \mathrm{cl}_{\|\|_2}(T^+ \cap \operatorname{supp}\mu_1^3)$. Thus, $\rho(t) \in \mathrm{cl}_{\|\|_2}(\rho^3(T^-) \cap \operatorname{supp}\mu_2^3) \cap \mathrm{cl}_{\|\|_2}(\rho^3(T^+) \cap \operatorname{supp}\mu_2^3)$. The disjointedness properties established in the previous paragraph imply that

$$\rho^3\left(T^- \cap \operatorname{supp}\mu_1^3\right) \subseteq \mathbf{R}^2 \setminus T^+ = \mathrm{cl}_{\|\|_2}\left(T^-\right) \quad \text{and} \quad \rho^3\left(T^+ \cap \operatorname{supp}\mu_1^3\right) \subseteq \mathbf{R}^2 \setminus T^- = \mathrm{cl}_{\|\|_2}\left(T^+\right),$$

so $\rho(t) \in \mathrm{cl}_{\|\|_2}\left(T^-\right) \cap \mathrm{cl}_{\|\|_2}\left(T^+\right)$. Since $\rho(x)$ belongs to $\rho(B_1)$, which is an open ball contained in $\operatorname{supp}\mu_2$, Lemma 3.9 ensures that $\rho(t)$ is aligned.

So ρ preserves alignments on every triple $(x,y,z) \in \operatorname{supp}\mu_1^3$ with $x \in B_1$. Recall that $B_2 = \rho(B_1)$ is an open, convex set contained in $\operatorname{supp}\mu_2$. We show that ρ preserves alignments in general by a density argument. Let $Q = abcd$ be a convex quadrilateral contained in B_1 and such that $(ab) \cap (cd)$ is a unique point e, contained in B_1, and $(ad)\cap(bc)$ is a unique point f, contained in B_1. We now subdivide Q, as depicted on the right, by defining $x=(ac)\cap(bd)$, then introducing the intersections of (ex) with (ad) and (bc), and the intersections of (fx) with (ab) and (cd). This produces four new convex quadrilaterals tiling Q. Remark that for any of these new quadrilaterals, the lines supporting opposite sides also intersect in e and f. We apply this subdivision procedure recursively to any new quadrilateral produced, and let A be the (infinite) set of vertices of all the quadrilaterals obtained this way. The set A is dense in our initial quadrilateral Q; indeed, the projective transformation that maps a to $(0,0)$, b to $(1,0)$, c to $(1,1)$ and d to $(0,1)$ maps A to the grid $\left\{\, (\frac{i}{2^k},\frac{j}{2^k}) \;\middle|\; k\in\mathbf{N} \text{ and } 0\leqslant i,j\leqslant 2^k \,\right\}$. (Note that under this transform, e and f are mapped to the line at infinity.)



Now, observe that $(\rho(a)\rho(b))$ and $(\rho(c)\rho(d))$ intersect in a unique point, namely $\rho(e)\in\operatorname{supp}\mu_2$ (note that $\rho(\operatorname{supp}\mu_1)$ cannot be contained in a line because ρ is measure preserving and μ_2 does not charge lines). Similarly, $(\rho(a)\rho(d))$ and $(\rho(b)\rho(c))$ intersect in a unique point, which is $\rho(f)\in\operatorname{supp}\mu_2$. We perform a similar recursive construction in B_2, starting from the quadrilateral $\rho(a)\rho(b)\rho(c)\rho(d)$, and let A' be the set of vertices obtained.

Let p be the unique projective transformation that maps a to $\rho(a)$, b to $\rho(b)$, c to $\rho(c)$ and d to $\rho(d)$. In particular, p preserves the colinearity of points in $\mathbf{R}^2$. This implies that if $x\in A$, then $\rho(x)=p(x)$ as all points in A and in A' are defined from Q and from $(\rho(a),\rho(b),\rho(c),\rho(d))=(p(a),p(b),p(c),p(d))$ using only colinearity properties. Hence $p_{|Q}=\rho_{|Q}$ by continuity.

It remains to show that $p(y)=\rho(y)$ for every $y\in\operatorname{supp}\mu_1$. Let $y\in\operatorname{supp}\mu_1$ and consider two lines L_1 and L_2 that both intersect the interior of Q and such that $L_1\cap L_2=\{y\}$. For $i\in\{1,2\}$, let x_i and z_i be two distinct points in $L_i\cap Q$; in particular $p(x_i)=\rho(x_i)$ and $p(z_i)=\rho(z_i)$. By definition, p preserves colinearity and since $x_1,x_2\in B_1$, both triples $\rho^3(x_i,y,z_i)$ are also aligned. It follows that

$$\rho(y)=(\rho(x_1)\rho(z_1))\cap(\rho(x_2)\rho(z_2))=(p(x_1)p(z_1))\cap(p(x_2)p(z_2))=p(y),$$

which completes the proof. $\square$

3.3.3 Rigidity for kernels (proofs)

We now prove the rigidity results stated in Section 3.3.1.

Step functions. We first argue that any kernel can be approximated by a step function.

Lemma 3.10. *Let (J,μ,W) be a kernel. For every positive real number ε, there exists a step function V on (J,μ) such that $\|W-V\|_1\leqslant\varepsilon$.*

Proof. It is well known in measure theory that for every function W and every positive ε, there exists a function $V\colon J^3\to\mathbf{R}$ that is measurable, has finite image (*i.e.*, $V(J^3)$ is finite) and satisfies $\|W-V\|_\infty\leqslant\varepsilon/2$. Let us write $V=\sum_{i=1}^k a_i\mathbf{1}_{A_i}$, where a_i is a nonzero real number and $\mathbf{1}_{A_i}$ the indicator function of a measurable set A_i for each $i\in[k]$.

Let $\mathcal{B}$ be the set of boxes of the form $P_1\times P_2\times P_3$, where P_1, P_2 and P_3 are measurable subsets of J and let X be the set of finite unions of elements of $\mathcal{B}$. For every $i\in[k]$, there is a set $B_i\in X$ such that $\mu(A_i\triangle B_i)\leqslant\frac{\varepsilon}{2k|a_i|}$. (A proof of this fact is presented in Appendix, Lemma A.1.) The function $U=\sum_{i=1}^k a_i\mathbf{1}_{B_i}$ is a step function as a linear combination of step functions. Moreover,

$$\|U-V\|_1\leqslant\sum_{i=1}^k|a_i|\cdot\|\mathbf{1}_{A_i}-\mathbf{1}_{B_i}\|_1=\sum_{i=1}^k|a_i|\cdot\mu(A_i\triangle B_i)\leqslant\frac{\varepsilon}{2}.$$

Consequently, $\|W-U\|_1\leqslant\|W-V\|_1+\|V-U\|_1\leqslant\varepsilon$, which finishes the proof. $\square$

Induced step functions. Let μ_n^u be the counting measure on $[n]$. As reported earlier (in Subsection 3.3.1), we know that $([n], \mu_n^u, \mathbf{H}(W,S))$ is a step function for the partition of $[n]$ into singletons. We now bound the distance between this step function and (J, μ, W). Thanks to Lemma 3.10, it is sufficient to deal with the case where (J, μ, W) is also a step function. We use the following (standard) terminology. An *atom* of a measure μ is a μ-measurable set A with $\mu(A) > 0$ and such that every μ-measurable subset of A has measure either $\mu(A)$ or 0. A measure is *atom-free* if it admits no atom.

Lemma 3.11. *Let (J, μ, W) be a step function and let $(P_1, \ldots, P_m)$ be a partition of J such that W is constant on $P_i \times P_j \times P_k$ for any $(i,j,k) \in [m]^3$. For every integer n and every tuple $S = (x_1, \ldots, x_n) \in J^n$,*

$$\mathrm{d}(W, \mathbf{H}(W,S)) \leqslant 6 \sum_{i=1}^{m} \left| \mu(P_i) - \frac{|\{j \in [n] \colon x_j \in P_i\}|}{n} \right| \cdot \|W\|_\infty .$$

Proof. We first eliminate any atom the measure μ may have by defining[5] an atom-free measurable space (J_0, μ_0) and a measure-preserving map $\phi_1 \colon J_0 \to J$. Now, let us write

$$Q_i = \{j \in [n] \colon x_j \in P_i\} \qquad \text{and} \qquad n_i = |Q_i| .$$

For $i \in [m]$, we set $a_i = \min\{\mu(P_i), \mu_n^u(Q_i)\}$. Since $\mu_0(\phi_1^{-1}(P_i)) = \mu(P_i) \geqslant a_i$, there is a subset $R_i \subseteq \phi_1^{-1}(P_i)$ such that $\mu_0(R_i) = a_i$. The set $J_0' = J_0 \setminus \bigcup_{i=1}^m R_i$ has measure $1 - \sum_{i=1}^m a_i = \sum_{i=1}^m (\mu_n^u(Q_i) - a_i)$. A now classical theorem of measure theory, due to Sikorski [34][6], assures that we can partition J_0' into m measurable subsets $T_1, \ldots, T_m$ such that $\mu_0(T_i) = \mu_n^u(Q_i) - a_i$ for every $i \in [m]$. (A proof that such a partition exists is presented in Appendix, Lemma A.2.)

We now construct a measure-preserving function ϕ_2 such that $\phi_2(R_i \cup T_i) = Q_i$ for each $i \in [m]$. Recall that $(Q_1, \ldots, Q_m)$ is a partition of $[n]$. We know that $\mu_0(R_i \cup T_i) = \mu_n^u(Q_i) = n_i/n$ for each $i \in [m]$, so there is a partition $(B_j^i)_{j \in Q_i}$ of $R_i \cup T_i$ into n_i parts each of μ_0-measure $1/n$, which are indexed by the elements of Q_i. If $j \in Q_i$, then we set $\phi_2(y) = j$ for every $y \in B_j^i$. Doing this for each $i \in [m]$ naturally defines a function ϕ_2 from J_0 to $[n]$. (Indeed, for every $x \in J_0$ there is a unique pair $(i,j) \in [m] \times [n]$ such that $x \in B_j^i$, and we set $\phi_2(x) = j$.) For convenience, we note that for each $j \in [n]$, there is exactly one index $i \in [m]$ such that B_j^i is defined: we call $i(j)$ this index.

The function ϕ_2 satisfies that $\phi_2(R_i \cup T_i) = Q_i$ for $i \in [m]$. Moreover, it is measure preserving since $\mu_0(\phi_2^{-1}(A)) = \sum_{j \in A} \mu_0 \left(B_j^{i(j)} \right) = |A|/n = \mu_n^u(A)$ for every $A \subseteq [n]$. To deduce our assertion, it suffices to prove that

$$\left\| W^{\phi_1} - \mathbf{H}(W,S)^{\phi_2} \right\|_1 \leqslant 6 \sum_{i=1}^{m} |\mu(P_i) - \mu_n^u(Q_i)| \cdot \|W\|_\infty .$$

If a belongs to R_i for some $i \in [m]$, then by construction $\phi_1(a) \in P_i$ and moreover $\phi_2(a) \in Q_i$, that is, $x_{\phi_2(a)} \in P_i$. Consequently, for every $(i,j,k) \in [m]^3$ the functions W^{ϕ_1} and $\mathbf{H}(W,S)^{\phi_2}$ are equal (and constant) on $R_i \times R_j \times R_k$.

It follows that W^{ϕ_1} and $\mathbf{H}(W,S)^{\phi_2}$ are equal everywhere except on the set $X = (J_0' \times J_0 \times J_0) \cup (J_0 \times J_0' \times J_0) \cup (J_0 \times J_0 \times J_0')$, which has μ_0^3-measure at most $3\mu_0(J_0') = 3 \sum_{i=1}^m (\mu_n^u(Q_i) - a_i) \leqslant 3 \sum_{i=1}^m |\mu(P_i) - \mu_n^u(Q_i)|$. Consequently,

$$\begin{aligned}
\left\| W^{\phi_1} - \mathbf{H}(W,S)^{\phi_2} \right\|_1 &\leqslant \int_X |W^{\phi_1} - \mathbf{H}(W,S)^{\phi_2}| \, \mathrm{d}\mu_0^3 \\
&\leqslant \int_X |W^{\phi_1}| \, \mathrm{d}\mu_0^3 + \int_X |\mathbf{H}(W,S)^{\phi_2}| \, \mathrm{d}\mu_0^3 \\
&\leqslant 2 \times \|W\|_\infty \times \mu_0^3(X) \\
&\leqslant 6 \|W\|_\infty \sum_{i=1}^{m} |\mu(P_i) - \mu_n^u(Q_i)| .
\end{aligned}$$

The statement follows. ☐

[5] If μ is atom-free then we may take $(J_0, \mu_0) = (J, \mu)$ and ϕ_1 equal to the identity; otherwise, we set $J_0 = J \times [0,1]$, we let μ_0 be the product measure of μ and the Lebesgue measure on $[0,1]$, and we let ϕ_1 be the projection of J_0 on its first coordinate.

[6] This result is often, and apparently wrongly, attributed to Sierpiński.

Random step functions. To control how well $\mathbf{H}(W,n)$ approximates W, we use the following lemma.

Lemma 3.12. *For every kernel (J,μ,W) and every integer n,*

$$|\mathbf{E}\left[\|\mathbf{H}(W,n)\|_1\right] - \|W\|_1| \leqslant \frac{3}{n}\|W\|_\infty .$$

Proof. Let $S=(x_1,\dots,x_n)$ be a μ^n-random tuple of J^n. The expected value E of $\|\mathbf{H}(W,S)\|_1$ is equal to

$$\int_{J^n}\frac{1}{n^3}\sum_{1\leqslant i,j,k\leqslant n}|W(x_i,x_j,x_k)|\,\mathrm{d}\mu^n(S) = \frac{1}{n^3}\sum_{1\leqslant i,j,k\leqslant n} I_{i,j,k},$$

where $I_{i,j,k}$ is the integral $\int_{J^n}|W(x_i,x_j,x_k)|\,\mathrm{d}\mu^n(S)$. First note that $0\leqslant I_{i,j,k}\leqslant\|W\|_\infty$ for every $(i,j,k)\in[n]^3$. If moreover i, j and k are pairwise different, then $I_{i,j,k}=\|W\|_1$. The number of triples of $[n]^3$ with pairwise different elements is greater than n^3-3n^2. It follows that $(n^3-3n^2)\|W\|_1\leqslant n^3E\leqslant(n^3-3n^2)\|W\|_1+3n^2\|W\|_\infty$, and further $-\frac{3}{n}\|W\|_1\leqslant E-\|W\|_1\leqslant\frac{3}{n}\|W\|_\infty$, which yields the conclusion since $\|W\|_1\leqslant\|W\|_\infty$. □

We can now prove that random step functions are good approximations relatively to our distance.

Lemma 3.13. *Let (J,μ,W) be a kernel. For every positive integer n, we consider the random kernel $([n],\mu_n^u,\mathbf{H}(W,n))$. The random sequence $(\mathrm{d}(W,\mathbf{H}(W,n)))_n$ almost surely tends to 0 as n goes to infinity.*

Proof. Let us first assume that W is a step function. We fix a partition $(P_1,\dots,P_m)$ of J such that W is constant on $P_i\times P_j\times P_k$ for any $(i,j,k)\in[m]^3$ and apply Lemma 3.11 to obtain, for any n-tuple S, the upper bound:

$$\mathrm{d}(W,\mathbf{H}(W,S))\leqslant 6\sum_{i=1}^{m}\left|\mu(P_i)-\frac{|\{j\in[n]\colon x_j\in P_i\}|}{n}\right|\cdot\|W\|_\infty . \tag{14}$$

Let us now consider a μ^n-random tuple S and set $Q_i=\{j\in[n]\colon x_j\in P_i\}$ and $n_i=|Q_i|$. For each $i\in[m]$, the parameter n_i follows a binomial law of parameter $\mu(P_i)$. Fix $\varepsilon>0$. By Hoeffding's inequality, the probability that $\left|\mu(P_i)-\frac{n_i}{n}\right|\geqslant\varepsilon$ is at most $2e^{-2\varepsilon^2n}$. Thus the union bound yields that $\mathrm{d}(W,\mathbf{H}(W,S))\leqslant 6m\varepsilon\|W\|_\infty$ with probability at least $1-2me^{-2\varepsilon^2n}$. Hence, $\mathrm{d}(W,\mathbf{H}(W,n))$ goes almost surely to 0 when n goes to infinity.

Let us now consider the general case. By Lemma 3.10, there is a step function V on (J,μ) such that $\|W-V\|_1\leqslant\varepsilon$. For every positive integer n, one can couple the random variables $\mathbf{H}(V,n)$ and $\mathbf{H}(W,n)$ such that

$$\mathbf{E}(\|\mathbf{H}(V,n)-\mathbf{H}(W,n)\|_1)\leqslant\frac{3}{n}\|V-W\|_\infty . \tag{15}$$

Indeed, let $S=(x_1,\dots,x_n)$ be a single μ^n-random tuple of J^n, and note that $\mathbf{H}(V,S)-\mathbf{H}(W,S)=\mathbf{H}(V-W,S)$. Lemma 3.12 applied to $V-W$ yields Equation (15). Since V is a step function, we know from the first part of the proof that when n goes to infinity, the random sequence $(\mathrm{d}(V,\mathbf{H}(V,n)))_n$ almost surely goes to 0. Let N be large enough to ensure that whenever $n\geqslant N$, both $\mathbf{E}(\mathrm{d}(V,\mathbf{H}(V,n)))\leqslant\varepsilon$ and $\frac{3}{n}\|V-W\|_\infty\leqslant\varepsilon$ hold. It follows that for every $n\geqslant N$,

$$\begin{aligned}
\mathbf{E}(\mathrm{d}(W,\mathbf{H}(W,n))) &\leqslant \mathbf{E}(\mathrm{d}(W,V)+\mathrm{d}(V,\mathbf{H}(V,n))+\mathrm{d}(\mathbf{H}(V,n),\mathbf{H}(W,n)))\\
&\leqslant \mathrm{d}(W,V)+\mathbf{E}(\mathrm{d}(V,\mathbf{H}(V,n)))+\mathbf{E}(\|\mathbf{H}(V,n)-\mathbf{H}(W,n)\|_1)\\
&\leqslant \|W-V\|_1+\mathbf{E}(\mathrm{d}(V,\mathbf{H}(V,n)))+\frac{3}{n}\|V-W\|_\infty\\
&\leqslant 3\varepsilon.
\end{aligned}$$

The statement follows. □

Proof of Theorem 11. We can now prove the first rigidity result that we stated: if (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are two kernels such that $t(H, W_1) = t(H, W_2)$ for every function $H \in \mathcal{H}$, then $\mathrm{d}(W_1, W_2) = 0$.

Proof of Theorem 11. The assumption implies that for every positive integer n, the two random variables $\mathbf{H}(W_1, n)$ and $\mathbf{H}(W_2, n)$ have the same distribution. Consequently, if H_n is a random function equivalent to both $\mathbf{H}(W_1, n)$ and $\mathbf{H}(W_2, n)$, then by Lemma 3.13 both $\mathrm{d}(W_1, H_n)$ and $\mathrm{d}(W_2, H_n)$ almost surely tend to 0. The triangular inequality $\mathrm{d}(W_1, W_2) \leqslant \mathrm{d}(W_1, H_n) + \mathrm{d}(H_n, W_2)$ thus ensures that $\mathrm{d}(W_1, W_2)$ equals 0. $\square$

Kernel isomorphisms We next show that there is an isomorphism between (almost all) the sets associated to these kernels that preserves (almost everywhere) the characteristic of the kernels, that is, their measures and their functions (Proposition 12). We start with an analogue of a classical result on kernels, which provides a weaker conclusion.

Lemma 3.14. *If (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are two kernels such that $\mathrm{d}(W_1, W_2) = 0$, then there exist a probability space (J, μ) and for each $i \in \{1, 2\}$ a measure-preserving map $\phi_i \colon J \to J_i$ such that*

- *for every $i \in \{1, 2\}$ and every μ-measurable set $A \subseteq J$, the set $\phi_i(A)$ is μ_i-measurable; and*
- *$W_2^{\phi_2}(x, y, z) = W_1^{\phi_1}(x, y, z)$ for μ^3-almost every $(x, y, z) \in J^3$.*

Proof. We only sketch the argument, as it is a straightforward extension of results already published, as referenced in what follows. We may first assume that $J_1 = [0, 1] = J_2$. A proof of this fact for functions of two variable was given by Janson [22, Proof of Theorem 7.1]. His argument extends directly to functions of three variables and, for that matter, to any finite number of variables. (We point out that although Janson assume a kernel to be symmetric in its two variables, in the proof of Theorem 7.1 (and, more specifically, in the proof of Lemma 7.3), this assumption is used only to ensure that the obtained function is again symmetric (and thus a kernel in Janson's sense), which is not needed here.)

Now a direct adaptation of the proof of Theorem 8.13 in the book by Lovász [25, p. 136] gives the lemma. We just give an outline of the argument. We aim to show the statement of the theorem for $J = J_1 \times J_2$, and ϕ_i being the projection of J on J_i for $i \in \{1, 2\}$. We know that $\mathrm{d}(W_1, W_2) = \inf_\mu \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^\mu$, where μ ranges over all coupling measures of $J = J_1 \times J_2$. It suffices to show that this last infimum is in fact a minimum to deduce the statement. As we assumed that $J_1 = [0, 1] = J_2$, the space of coupling measures is compact in the weak topology. Consequently, it is enough to show that the function $\mu \mapsto \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^\mu$ is *lower semicontinuous*, *i.e.*, if $(\mu_n)_n$ weakly converges to μ then

$$\liminf_n \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^{\mu_n} \geqslant \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^\mu .$$

This last inequality is Inequality (8.21) on p. 137 of *loc. cit.* and the proof follows as in the book. $\square$

The twin-free case (proof of Proposition 12). We can now prove a stronger rigidity: given two twin-free kernels (J_1, μ_1, W_1) and (J_2, μ_2, W_2), if $t(H, W_1) = t(H, W_2)$ for every function $H \in \mathcal{H}$, then there exist two sets $N_1 \subseteq J_1$ and $N_2 \subseteq J_2$ with $\mu_1(N_1) = 0 = \mu_2(N_2)$ and an invertible and measure-preserving map $\rho \colon J_1 \setminus N_1 \to J_2 \setminus N_2$ such that

1. ρ^{-1} is measure preserving; and
2. W_1 and W_2^ρ are equal μ_1^3-almost everywhere.

Proof of Proposition 12. Theorem 11 ensures that Lemma 3.14 applies: let (J, μ) and ϕ_1, ϕ_2 be the probability space and the applications given by this lemma, respectively. For $i \in \{1, 2\}$, we set $\tilde{J}_i = \phi_i(J)$ and $N_i = J_i \setminus \tilde{J}_i$. So $\mu_i(N_i) = 0$ because ϕ_i is measure preserving and $\phi_i^{-1}(\tilde{J}_i) = J$.

We start by showing that if two elements of J have the same image by ϕ_1, then they must have the same image by ϕ_2; that is, $\phi_2(\phi_1^{-1}(\{x\}))$ is a singleton for every $x \in \tilde{J}_1$. Indeed, suppose that $\phi_1^{-1}(\{x\})$ contains two distinct elements a and b. Then $W_1^{\phi_1}(a, y, z) = W_1^{\phi_1}(b, y, z)$ for every $(y, z) \in J^2$. Furthermore, we know that $W_1^{\phi_1}(a, y, z) = W_2^{\phi_2}(a, y, z)$ and $W_1^{\phi_1}(b, y, z) = W_2^{\phi_2}(b, y, z)$ for μ^2-almost every pair $(y,z) \in J^2$. Consequently, $W_2^{\phi_2}(a,y,z) = W_2^{\phi_2}(b,y,z)$ for μ^2-almost every pair $(y,z) \in J^2$. This implies that $W_2(\phi_2(a),y',z') = W_2(\phi_2(b),y',z')$ for μ_2^2-almost every pair $(y',z') \in \tilde{J}_2^2$, because $\mu(J) = \mu(\phi_2^{-1}(\phi_2(J))) = \mu_2(\phi_2(J))$. (The last equality follows from the fact that ϕ_2 is measure preserving.) Since W_2 is twin-free, we deduce that $\phi_2(a) = \phi_2(b)$. Therefore, we can define a map $\rho\colon \tilde{J}_1 \to \tilde{J}_2$ by setting $\rho(x)$ to be the unique element of $\phi_2(\phi_1^{-1}(\{x\}))$.

Now, observe that there exists a subset N of $\tilde{J}_1^3$ with $\mu_1^3(N) = 0$ such that

- $\tilde{J}_1^3 \setminus N \subseteq \phi_1(J)^3$;
- $(x,y,z) \mapsto (\rho(x),\rho(y),\rho(z))$ is defined everywhere on $\tilde{J}_1^3 \setminus N$; and
- $W_2^{\phi_2}(x,y,z) = W_1^{\phi_1}(x,y,z)$ whenever $(\phi_1(x),\phi_1(y),\phi_1(z)) \notin N$.

For $(x_1,y_1,z_1) \in \tilde{J}_1^3 \setminus N$, let $(x,y,z) \in J^3$ such that $\phi_1(x) = x_1$, $\phi_1(y) = y_1$ and $\phi_1(z) = z_1$. Then $\rho(x_1) = \phi_2(x)$, $\rho(y_1) = \phi_2(y)$ and $\rho(z_1) = \phi_2(z)$. Therefore $W_2^{\rho}(x_1,y_1,z_1) = W_2^{\phi_2}(x,y,z) = W_1^{\phi_1}(x,y,z) = W_1(x_1,y_1,z_1)$, which yields 2.

Let us show that ρ is measure preserving. We observe that if $A \subseteq \tilde{J}_2$, then $\rho^{-1}(A) = \phi_1(\phi_2^{-1}(A))$. Indeed, by the definitions

$$
\begin{aligned}
\rho^{-1}(A) &= \left\{\, x \in \tilde{J}_1 \;\middle|\; \phi_2(\phi_1^{-1}(\{x\})) \subseteq A \,\right\} \\
&= \left\{\, x \in \tilde{J}_1 \;\middle|\; \phi_1^{-1}(\{x\}) \subseteq \phi_2^{-1}(A) \,\right\} \\
&\subseteq \phi_1(\phi_2^{-1}(A)).
\end{aligned}
$$

Conversely, if $x \in \phi_1(\phi_2^{-1}(A))$, then there exists $y \in J$ such that $\phi_1(y) = x$ and $\phi_2(y) \in A$. Since $\phi_2(\phi_1^{-1}(\{x\}))$ is a singleton, we deduce that $\phi_2(\phi_1^{-1}(\{x\})) = \{\phi_2(y)\}$, which is contained in A. This proves the observation.

As a result, it suffices to prove that $\mu_2(A) = \mu_1(\phi_1(\phi_2^{-1}(A)))$ to infer that $\mu_1(\rho^{-1}(A)) = \mu_2(A)$. As ϕ_1 and ϕ_2 are measure preserving, it is enough to prove that $\phi_1^{-1}(\phi_1(\phi_2^{-1}(A))) = \phi_2^{-1}(A)$. By definition, the set on the right side is always contained in the set on the left side. For the converse inclusion, fix $x \in \phi_1^{-1}(\phi_1(\phi_2^{-1}(A)))$ and let us show that $\phi_2(x) \in A$. There exists $y \in J$ such that $\phi_1(y) = \phi_1(x)$ and $\phi_2(y) \in A$. In particular, $\phi_2(y) \in \phi_2(\phi_1^{-1}(\phi_1(\{x\})))$ and hence $\phi_2(\phi_1^{-1}(\phi_1(\{x\}))) = \{\phi_2(y)\}$. However, $x \in \phi_1^{-1}(\phi_1(\{x\}))$ and thus $\phi_2(x) = \phi_2(y) \in A$.

To see that ρ is invertible, we first define $\rho'\colon \tilde{J}_2 \to \tilde{J}_1$. To this end, one shows similarly as before that $\phi_1(\phi_2^{-1}(\{x\}))$ is a singleton for every $x \in \tilde{J}_2$. So $\rho'(x)$ can be defined as the unique element of $\phi_1(\phi_2^{-1})(\{x\})$. Now by symmetry of the roles played by ϕ_1 and ϕ_2, one sees similarly as before that ρ' is measure preserving. It remains to prove that $\rho'(\rho(x)) = x$ for every $x \in \tilde{J}_1$. Fix $x \in \tilde{J}_1$. As $\phi_2(\phi_1^{-1}(\{x\})) = \{\rho(x)\}$, there exists $y \in J$ such that $\phi_1(y) = x$ and $\phi_2(y) = \rho(x)$. Consequently, $x \in \phi_1(\phi_2^{-1}(\{\rho(x)\}))$, which is equal to $\{\rho'(\rho(x))\}$. This concludes the proof. $\square$

The case of pure kernels (proof of Theorem 13). We finally establish our strongest rigidity theorem for kernels: if (J_1,μ_1,W_1) and (J_2,μ_2,W_2) are two pure kernels and $t(H,W_1) = t(H,W_2)$ for every function $H \in \mathcal{H}$, then there exists an isometry $\rho\colon (J_1,d_{W_1}) \to (J_2,d_{W_2})$ such that W_1 and W_2^{ρ} are equal μ_1^3-almost everywhere.

Proof of Theorem 13. Let $\rho\colon J_1 \setminus N_1 \to J_2 \setminus N_2$ be the map given by Proposition 12 applied to (J_1,μ_1,W_1) and (J_2,μ_2,W_2). We first prove that we may restrict ρ to a set D_1 with μ_1-measure one such that if we fix any $x \in D_1$, then $W_1(x,y,z)$ equals $W_2^{\rho}(x,y,z)$ for μ_1^2-almost every pair $(y,z) \in J_1^2$. For $x \in J_1$, we define $I(x) \subset J_1^2$ to be the set of pairs (y,z) such that $W_1(x,y,z) \neq W_2^{\rho}(x,y,z)$. Further, let A be the set composed of each $x \in J_1$ such that $\mu_1^2(I(x)) > 0$.

We assert that $\mu_1(A) = 0$. To prove this, we set $A_\varepsilon = \left\{\, x \in J_1 \;\middle|\; \mu_1^2(I(x)) > \varepsilon \,\right\}$ and we notice that $A = \bigcup_n A_{\varepsilon_n}$ where the union is taken over a decreasing sequence $(\varepsilon_n)_{n\in\mathbf{N}}$ that tends to 0. As the union is countable, it suffices to prove that $\mu_1(A_\varepsilon) = 0$ for every $\varepsilon > 0$ to conclude that $\mu_1(A) = 0$. Fixing $\varepsilon > 0$, it follows from the definitions that $W_1(x,y,z) \neq W_2^{\rho}(x,y,z)$ for every triple in $\{\, (x,y,z) \mid x \in A_\varepsilon \text{ and } (y,z) \in I(x) \,\}$, which is a set of μ_1^3-measure at least $\varepsilon \cdot \mu_1(A_\varepsilon)$. Because of Property 2 of Proposition 12, the previous statement implies that $\varepsilon \cdot \mu_1(A_\varepsilon) = 0$, hence $\mu_1(A_\varepsilon) = 0$.

We define $D_1 = J_1 \setminus N_1 \setminus A$ and $D_2 = \rho(D_1) = J_2 \setminus N_2 \setminus \rho(A)$. We know that $\mu_1(D_1) = 1$ and the equality $\mu_2(D_2) = 1$ follows from the fact that ρ^{-1} is measure preserving. The restriction $\rho_{|D_1}\colon D_1 \to D_2$ of ρ to D_1

is an isometry between the metric spaces (D_1, d_{W_1}) and (D_2, d_{W_2}). Indeed, fixing $(x, x') \in D_1^2$, we know from the construction of D_1 that for μ_1^2-almost every pair $(y, z) \in J_1^2$ we have $W_1(x, y, z) = W_2^\rho(x, y, z)$ and $W_1(x', y, z) = W_2^\rho(x', y, z)$. So in particular $W_1(x, y, z) - W_1(x', y, z) = W_2^\rho(x, y, z) - W_2^\rho(x', y, z)$. Consequently,

$$\begin{aligned} d_{W_1}(x, x') &= \int_{J_1^2} |W_1(x, y, z) - W_1(x', y, z)| \,\mathrm{d}\mu_1^2(y, z) \\ &= \int_{J_1^2} |W_2^\rho(x, y, z) - W_2^\rho(x', y, z)| \,\mathrm{d}\mu_1^2(y, z) \\ &= \int_{J_2^2} |W_2(\rho(x), y', z') - W_2(\rho(x'), y', z')| \,\mathrm{d}\mu_2^2(y', z') \\ &= d_{W_2}(\rho(x), \rho(x')). \end{aligned}$$

This proves that $\rho_{|D_1}$ is an isometry.

Now we assume that μ_1 has full support and (J_2, d_{W_2}) is complete. In this case, $\rho_{|D_1}$ extends by continuity to an injective map $\tilde{\rho}$ on J_1. To prove this, it suffices to show that $\rho_{|D_1}$ is absolutely continuous and D_1 is dense in (J_1, d_{W_1}). The absolute continuity follows from the fact that $\rho_{|D_1}$ is an isometry. The set D_1 is dense in J_1 because every open set included in $J_1 \setminus D_1$ is an open nullset, and hence is empty as μ_1 has full support. By continuity of d_{W_1} and d_{W_2} towards themselves the extension $\tilde{\rho}$ is an isometry.

To prove the second item, it suffices to apply the previous proof to the inverse $(\rho_{|D_1})^{-1}$ of $\rho_{|D_1}$, where the roles played by (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are inverted. □

3.4 Flexibility of realizations and the probability that k random points are in convex position

For limits of order types that can be realized by at least one measure whose support has non-empty interior, the rigidity theorem completely describes the space of realizations: they are the orbit of that one realization under spherical transformations. In general, the situation can be radically different as the following easy example shows.

Example 3.1. Let $\ell_\diamond$ be the limit of a sequence of sets of points in convex position. This limit assigns the probability 1 to each $\diamond_k$, the order type of k points in convex position, and 0 to the rest. Every measure with convex support realizes $\ell_\diamond$; this allows arbitrarily disconnected support, as wells as support of any Hausdorff dimension between 0 and 1 (consider a Cantor set on $[0, 1]$ with that dimension and map the interval to the circle).

The limit $\ell_\diamond$ is exceptionally simple, it obviously maximizes $\ell(\diamond_k)$. One may wonder if this variety of realizations is also exceptional. We construct a different limit with similar realization properties that plays a role in combinatorial geometry as it gets close to minimizing $\ell(\diamond_k)$ for k large enough.

Theorem 15. *There exists a limit ℓ_E of order types such that for every $t \in (0, 1)$, the limit ℓ_E can be realized by a measure with a support of Hausdorff dimension t. Moreover, there is no measure μ that realizes ℓ_E and is, on an open set of positive μ-measure, absolutely continuous to the Lebesgue measure or to the length measure on a C^2 curve of positive length.*

Theorem 15 will follow from Lemma 3.15 for the first statement, and Proposition 17 and Lemma 3.16 for the second statement.

3.4.1 Definition of ℓ_E

It is convenient to give two presentations of ℓ_E, one geometric and the other combinatorial.

Let us start with the combinatorial definition of ℓ_E. Consider the space $E = \{0, 1\}^{\mathbf{N}}$ equipped with the coin-tossing measure. For $u, v \in E$, let $u \wedge v$ be the longest common prefix of u and v and let $\prec_{lex}$ be the lexicographic order on E. We define ℓ_E as a chirotope χ on E. Specifically, let $u, v, w \in E$ and, without loss of generality, suppose that $u \prec_{lex} v \prec_{lex} w$. We set $\chi(u, v, w) = 1$ if $|u \wedge v| < |v \wedge w|$ and $\chi(u, v, w) = -1$ otherwise. For any order type ω of size k, we let $\ell_E(\omega)$ be the probability that the

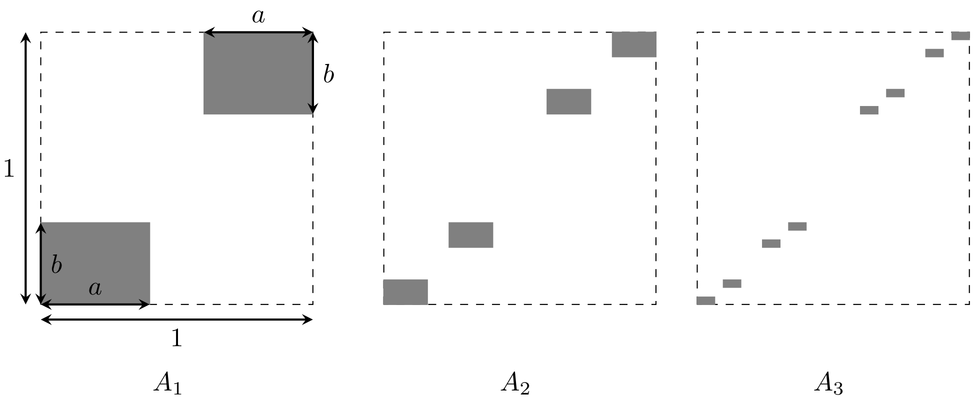


Figure 1: Definition of ℓ_E.

restriction of χ to k random elements of E chosen independently from the coin-tossing distribution equals, after unlabelling, ω. The fact that ℓ_E is a limit of order types easily follows from the geometric viewpoint.

Let us now give a geometric presentation of ℓ_E; refer to Figure 1. As is usual, let $\{0,1\}^*$ be the collection of all finite binary words. Fix some parameters a and b such that $0 < b < a < \frac{1}{2}$, and define the rectangles $R = [0,1]^2, R_0 = [0,a] \times [0,b]$ and $R_1 = [1-a,1] \times [1-b,1]$. For each $i \in \{0,1\}$, let φ_i be the affine transform fixing (i,i) and mapping R to R_i. To any word $w = i_1 i_2 \dots i_n \in \{0,1\}^*$ we associate the set $R_w = \varphi_{i_n} \circ \varphi_{i_{n-1}} \circ \dots \circ \varphi_{i_1}(R)$ and let $\mu^{a,b}$ be the probability measure such that $\mu^{a,b}(R_w) = \frac{1}{2^{|w|}}$ for every $w \in \{0,1\}^*$. We notice that $R_w \subset R_v$ if and only if v is a prefix of w. Letting $A_n = \bigcup_{w \in \{0,1\}^n} R_w$ for $n \geqslant 1$, the support of $\mu^{a,b}$ is $A = \bigcap_{n \geqslant 1} A_n$.

Lemma 3.15. *If $(a,b) \in \left(0, \frac{1}{2}\right)^2$ and $b \leqslant (1-2a)(1-2b)a$, then $\ell_{\mu^{a,b}} = \ell_E$. In particular, ℓ_E is a limit of order types.*

Proof. The measure $\mu^{a,b}$ is the image of the coin-tossing probability on $\{0,1\}^{\mathbf{N}}$ by the function $\Psi_{a,b}$ that assigns to $w \in \{0,1\}^{\mathbf{N}}$ the unique point in $\bigcap_{w_v} R_{w_v}$, where the intersection is taken over all prefixes w_v of w.

We shall prove that every point in $A \cap R_1$ lies above any line spanned by two points in $A \cap R_0$ provided that

$$b \leqslant (1-2a)(1-2b)a.$$

Since A is stable by the symmetry of center $(\frac{1}{2}, \frac{1}{2})$, it would then follow that every point in $A \cap R_0$ lies below any line spanned by two points in $A \cap R_1$. Let us show how this property of A allows us to conclude the proof. Indeed, this property yields that $\ell_{\mu^{a,b}}$ is fully determined. Let $\Psi_{a,b}(u)$, $\Psi_{a,b}(v)$ and $\Psi_{a,b}(w)$ be three pairwise distinct points in A with $u, v, w \in \{0,1\}^{\mathbf{N}}$ and assume that $u \prec_{\text{lex}} v \prec_{\text{lex}} w$. If $|u \wedge v| < |v \wedge w|$, set $p = u \wedge v$. Since $u \prec_{\text{lex}} v$, the word $p.0$ is a prefix of u and $p.1$ is a prefix of v, and therefore of w. It follows that $\Psi_{a,b}(u) \in R_{p.0}$ and $\{\Psi_{a,b}(v), \Psi_{a,b}(w)\} \subset R_{p.1}$. Moreover, the abscissa of $\Psi_{a,b}(v)$ is smaller than that of $\Psi_{a,b}(w)$ because $v \prec_{\text{lex}} w$. Consequently, $\chi(\Psi_{a,b}(u), \Psi_{a,b}(v), \Psi_{a,b}(w)) = 1 = \chi_E(u,v,w)$. The proof that $\chi(\Psi_{a,b}(u), \Psi_{a,b}(v), \Psi_{a,b}(w)) = \chi_E(u,v,w)$ when $|u \wedge v| > |v \wedge w|$ is similar and we omit it.

It remains to prove that A indeed fulfills the property announced. For two distinct points $x, y \in A$, let $\alpha(x,y)$ be the angle between the line $h(x,y)$ and the abscissa axis (so $\alpha(x,y)$ is defined modulo π). If $x \in R_0$ and $y \in R_1$, then

$$1 - 2b \leqslant \tan\alpha(x,y) \leqslant \frac{1}{1-2a},$$

since the minimum is obtained when $x_m = (0,b)$ and $y_m = (1, 1-b)$, while the maximum is obtained when $x_M = (a,0)$ and $y_M = (1-a,1)$ (see Figure 2, left.) The application of a function φ_i with $i \in \{1,2\}$ acts as follows

$$\tan\alpha(\varphi_i(x), \varphi_i(y)) = \frac{b}{a}\tan\alpha(x,y).$$

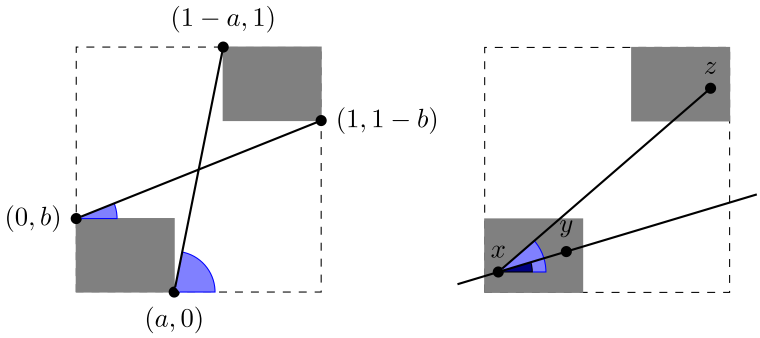


Figure 2: Left: Minimal and maximal value for $\alpha(x,y)$ when $x \in R_0$ and $y \in R_1$. Right: $z \in R_1$ is above $h(x,y)$ if $\alpha(x,y) < \alpha(x,z)$.

Since $\frac{b}{a} < 1$, it further holds that $\tan\alpha(\varphi_i(x), \varphi_i(y)) < \tan\alpha(x,y)$. By iterating this property, it eventually follows that $\tan\alpha(x,y) \leqslant \frac{1}{1-2a}$ for every $x, y \in A$ such that $x \neq y$ and x has smaller abscissa than y.

Let $x, y \in A \cap R_0$ and $z \in A \cap R_1$ such that x has smaller abscissa than y. Note that z lies above the line $h(x,y)$ if and only if $\alpha(x,y) \leqslant \alpha(x,z)$, where the values of the angles are taken in $(-\frac{\pi}{2}, \frac{\pi}{2})$ (see Figure 2, right), which in turn is equivalent to $\tan\alpha(x,y) \leqslant \tan\alpha(x,z)$. Since $\tan\alpha(x,y) \leqslant \frac{b}{a}\frac{1}{1-2a}$ and $\tan\alpha(x,z) \geqslant 1-2b$, it suffices that $\frac{1}{1-2a}\frac{b}{a} \leqslant 1-2b$, *i.e.*, $b \leqslant (1-2a)(1-2b)a$. This proves the announced property, thereby completing the demonstration. $\square$

3.4.2 The Erdős-Sylvester problem

As mentioned already in the introduction, the limit of order types ℓ_E was constructed to attack the Erdős-Sylvester problem of determining c_k. In fact, both upper and lower bounds on c_k go back to the classical results of Erdős and Szekeres who proved that for every integer k, there exists $s(k)$ such that any planar point set in general position and of size at least $s(k)$ contains k points in convex position. Erdős and Szekeres also gave a construction showing that $s(k) \geqslant 2^{k-2}+1$ and they conjectured this value to be tight. This bound was nearly achieved in a recent breakthrough of Suk [37] who improved the upper bound on $s(k)$ to $2^{k+6k^{2/3}\log k}$. The next proposition is folklore.

Proposition 16. *For every integer $k \geqslant 4$ and every limit of order types ℓ,*

$$\ell(\diamond_k) \geqslant 2^{-k^2+o(k^2)}.$$

Proof. Let $(\omega_n)_{n\in\mathbf{N}}$ be a convergent sequence of order types with limit ℓ. The Erdős-Szekeres theorem ensures that $p(\diamond_k, \omega) \geqslant \frac{1}{\binom{s(k)}{k}}$ for any order type ω of size $s(k)$. Let $k \geqslant 4$ and fix some n_0 such that $|\omega_n| \geqslant s(k)$ whenever $n \geqslant n_0$. It then follows that

$$\forall n \geqslant n_0, \qquad p(\diamond_k, \omega_n) = \sum_{\omega \in \mathcal{O}_{s(k)}} p(\diamond_k, \omega) p(\omega, \omega_n) \geqslant \frac{1}{\binom{s(k)}{k}} \sum_{\omega \in \mathcal{O}_{s(k)}} p(\omega, \omega_n) = \frac{1}{\binom{s(k)}{k}}$$

The first identity is a standard conditional probability argument: instead of taking a random k-element subset of a realization of ω_n, we first take an $s(k)$-element subset, consider their order type ω, then take a random k-element subset of a realization of ω and estimate the probability that it has order type $\diamond_k$ conditioned on the order type of ω. The last identity simply expresses that the sum for all order types ω of size $s(k)$ of the density $p(\omega, \omega_n)$ is 1. We derive that $p(\diamond_k, \omega_n) \geqslant \frac{1}{\binom{s(k)}{k}} \geqslant \frac{1}{2^{k^2+o(k^2)}}$ for any $n \geqslant n_0$, so $\ell(\diamond_k) = \lim_{n\to\infty} p(\diamond_k, \omega_n) \geqslant \frac{1}{2^{k^2+o(k^2)}}$. $\square$

Our example ℓ_E matches the order of growth in the exponent of Proposition 16. We don't know of any previous result in this direction.

Proposition 17. $\ell_E(\diamond_k) = 2^{-\frac{k^2}{8}+O(k\log k)}$.

Proof. Let μ be a measure that does not charge lines. Let X and S be two sets of random points chosen independently from μ of respective sizes k and $\lfloor k/2 \rfloor$. By the discussion above,

$$\ell_\mu(\diamond_k) = \mathbf{P}_\mu(X \text{ is in convex position}) \leqslant \binom{k}{\lfloor \frac{k}{2} \rfloor} \mathbf{P}_\mu(S \text{ is a cup or a cap}).$$

Moreover, the construction of $\mu^{a,b}$ implies that

$$\mathbf{P}_{\mu^{a,b}}(S \text{ is a cup or a cap}) = 2\,\mathbf{P}_{\mu^{a,b}}(S \text{ is a cup}).$$

The condition that $b \leqslant (1-2a)(1-2b)a$ ensures that every s-cup containing more than one point in $\phi_0(R)$ contains at most one point in $\phi_1(R)$. It follows that

$$\begin{aligned}\mathbf{P}_{\mu^{a,b}}(S \text{ is a cup}) = &\mathbf{P}_{\mu^{a,b}}(S \text{ is a cup and } S \subset \phi_0(R)) + \mathbf{P}_{\mu^{a,b}}(S \text{ is a cup and } S \subset \phi_1(R)) \\ &+ \mathbf{P}_{\mu^{a,b}}(S \cap \phi_0(R) \text{ is a cup and } |S \cap \phi_0(R)| = |S| - 1)\end{aligned}$$

Observe that

$$\mathbf{P}_{\mu^{a,b}}(S \text{ is a cup} \mid S \subset \phi_0(R)) = \mathbf{P}_{\mu^{a,b}}(S \text{ is a cup} \mid S \subset \phi_1(R)) = \mathbf{P}_{\mu^{a,b}}(S \text{ is a cup}).$$

Altogether, defining $f(s)$ to be the probability that s random points chosen independently from $\mu^{a,b}$ form a cup, we have

$$f(s) = \frac{2}{2^s} f(s) + \frac{s}{2^s} f(s-1), \quad \text{that is,} \quad f(s) = \frac{s}{2^s - 2} f(s-1).$$

Notice that $f(3) = \frac{1}{2}$. This can be seen directly from the combinatorial description of ℓ_E as follows. Consider three different sequences $u, v, w \in \{0,1\}^n$, assuming that $u \prec_{lex} v \prec_{lex} w$. The orientation of (u, v, w) depends only on the first entry of v where u and w differ, and this entry of v is uniformly distributed in $\{0, 1\}$. Altogether,

$$f(s) = \prod_{i=3}^{s} \frac{i}{2^i - 2} = 2^{-\frac{s(s+1)}{2}+3} \prod_{i=3}^{s} \frac{i}{1 - 2^{-i+1}}$$

so $f(s) = 2^{-\frac{s^2}{2}+O(s\log s)}$ and $\ell_E(\diamond_k) = 2^{-\frac{k^2}{8}+O(k\log k)}$. ☐

Until recently it was suspected that every realization of the order types of Erdős-Szekeres needed to be very spread out, in the sense that the quotient of the diameter and the minimal distance between two points on every realization had to be exponentially large. A construction of the Erdős-Szekeres example in a grid of size polynomial in n was recently achieved [15].

To some extent our next result vindicates the original intuition on realizations of the order types of Erdős-Szekeres. We show that the fast decay of $\ell_E(\diamond_k)$ exhibited in Proposition 17 drastically restricts its space of realizations.

Lemma 3.16. *Let μ be a finite measure over $\mathbf{R}^2$ for which lines are negligible and U an open set of positive μ-measure.*

(i) If μ is absolutely continuous, on U, to the Lebesgue measure then $p(\diamond_k, \mu) \geqslant 4^{-k\log k + O(k)}$.

(ii) If μ is absolutely continuous, on U, to the length measure on a C^2 curve then $p(\diamond_k, \mu) \geqslant 2^{-O(k)}$.

Proof. By the Radon-Nikodym theorem [28, 29], if a measure μ is absolutely continuous to a measure λ on a measurable set X then there exists an absolutely continuous function $\frac{d\mu}{d\lambda}$ such that $\mu(A) = \int_A \frac{d\mu}{d\lambda} d\lambda$ for every measurable set $A \subseteq X$.

Let λ_2 be the Lebesgue measure over $\mathbf{R}^2$. Since $\mu(U) > 0$ and μ is absolutely continuous to λ_2 on U, the function $\frac{d\mu}{d\lambda_2}$ is nonnegative, nonzero, and continuous on U. In particular, U contains some square A

on which $\frac{d\mu}{d\lambda_2}$ is bounded from below by some positive constant c. The probability that k random points chosen independently from μ are all in A is $\left(\frac{\mu(A)}{\mu(\mathbf{R}^2)}\right)^k$. Conditioning on this event, we deduce that

$$p(\diamond_k, \mu) \geqslant \left(\frac{\mu(A)}{\mu(\mathbf{R}^2)}\right)^k p(\diamond_k, \mu_{|A}).$$

Let $1_{\diamond_k}$ be the indicator function, over A^k, of k-tuples of points in convex position. We have

$$\begin{aligned} p(\diamond_k, \mu_{|A}) &= \int_{A^k} 1_{\diamond_k} d\mu(x_1) d\mu(x_2) \dots d\mu(x_k) \\ &= \int_{A^k} 1_{\diamond_k} \frac{d\mu}{d\lambda_2}(x_1) \frac{d\mu}{d\lambda_2}(x_2) \dots \frac{d\mu}{d\lambda_2}(x_k) d\lambda_2(x_1) d\lambda_2(x_2) \dots d\lambda_2(x_k) \\ &\geqslant c^k \int_{A^k} 1_{\diamond_k} d\lambda_2(x_1) d\lambda_2(x_2) \dots d\lambda_2(x_k) = c^k p(\diamond_k, \lambda_{2|A}) \end{aligned}$$

Valtr [39] proved that $p(\diamond_k, \lambda_{2|A}) = \frac{1}{k!^2}\binom{2k-1}{k-1}^2$, so altogether

$$p(\diamond_k, \mu) \geqslant \left(\frac{\mu(A)}{\mu(\mathbf{R}^2)}\right)^k \frac{c^k}{k!^2} \binom{2k-1}{k-1}^2 = \Omega\left(4^{-k \log k + O(k)}\right),$$

which proves the first statement.

For the 1-dimensional case, let $d\lambda_1$ be the 1-dimensional Lebesgue measure and let Γ be a C^2 curve such that μ is absolutely continuous to $\lambda_{1|\Gamma}$ on some open set U. Since $\frac{d\mu}{d\lambda_1}_{|\Gamma}$ is continuous, nonnegative and nonzero, there is an open set $U' \subseteq U$ such that $\frac{d\mu}{d\lambda_1}$ is positive on $\Gamma \cap U'$.

Since Γ is C^2, its curvature is a continuous function. As μ does not charge lines, this curvature is nonzero and there exists an open subset U'' of U such that $U'' \cap \Gamma$ is non-empty and Γ has positive curvature on U''. Up to passing to a smaller neighborhood, we can find an arc γ of our curve that has positive length and that is entirely on its convex hull. Hence, any k points on γ are in convex position. Moreover, $\frac{d\mu}{d\lambda_1}_{|\Gamma}$ is positive on U' and therefore on γ, so $\mu(\gamma)$ is positive. It follows that

$$\forall k, \quad p(\diamond_k, \mu) \geqslant \left(\frac{\mu(\gamma)}{\mu(\mathbf{R}^2)}\right)^k,$$

which proves the second statement. □

References

[1] B. M. Ábrego, M. Cetina, S. Fernández-Merchant, J. Leaños, and G. Salazar. On $\leqslant k$-edges, crossings, and halving lines of geometric drawings of K_n. *Discrete Comput. Geom.*, 48(1):192–215, 2012.

[2] B. M. Ábrego and S. Fernández-Merchant. A lower bound for the rectilinear crossing number. *Graphs Combin.*, 21(3):293–300, 2005.

[3] B. M. Ábrego, S. Fernández-Merchant, J. Leaños, and G. Salazar. A central approach to bound the number of crossings in a generalized configuration. In *The IV Latin-American Algorithms, Graphs, and Optimization Symposium*, volume 30 of *Electron. Notes Discrete Math.*, pages 273–278. Elsevier Sci. B. V., Amsterdam, 2008.

[4] B. M. Ábrego, S. Fernández-Merchant, and G. Salazar. The rectilinear crossing number of K_n: closing in (or are we?). In *Thirty essays on geometric graph theory*, pages 5–18. Springer, New York, 2013.

[5] O. Aichholzer, F. Aurenhammer, and H. Krasser. Enumerating Order Types for Small Point Sets with Applications. In *Proc.* 17^{th} *Ann. ACM Symp. Computational Geometry*, pages 11–18, Medford, Massachusetts, USA, 2001.

[6] O. Aichholzer, J. García, D. Orden, and P. Ramos. New lower bounds for the number of ($\leqslant k$)-edges and the rectilinear crossing number of K_n. *Discrete Comput. Geom.*, 38(1):1–14, 2007.

[7] G. Aloupis, J. Iacono, S. Langerman, O. Özkan, and S. Wuhrer. The complexity of order type isomorphism. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 405–415, 2014.

[8] F. Ardila, F. Rincón, and L. Williams. Positively oriented matroids are realizable. *J. Eur. Math. Soc. (JEMS)*, 19(3):815–833, 2017.

[9] J. Balogh, B. Lidický, and G. Salazar. Closing in on Hill's conjecture. *ArXiv e-prints*, Nov. 2017.

[10] B. Borchers. CSDP, A C library for semidefinite programming. *Optimization Methods and Software*, 11(1-4):613–623, 1999.

[11] P. Brass, W. Moser, and J. Pach. *Research Problems in Discrete Geometry.* Springer, 2005.

[12] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.-H. Vritsiou. *Geometry of isotropic convex bodies*, volume 196 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2014.

[13] R. de Joannis de Verclos. *Applications of limits of combinatorial structures in graph theory and in geometry.* PhD thesis, Université Grenoble-Alpes, 2017.

[14] M. G. Dobbins, A. Holmsen, and A. Hubard. Realization spaces of arrangements of convex bodies. *Discrete Comput. Geom.*, 58(1):1–29, 2017.

[15] F. Duque, R. Fabila-Monroy, and C. Hidalgo-Toscano. Point sets with small integer coordinates and no large convex polygons. *Discrete Computational Geometry*, 59, 2018.

[16] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[17] P. Erdős. Some old and new problems in combinatorial geometry. In *Convexity and graph theory (Jerusalem, 1981)*, volume 87 of *North-Holland Math. Stud.*, pages 129–136. North-Holland, Amsterdam, 1984.

[18] P. Erdős and R. K. Guy. Crossing number problems. *Amer. Math. Monthly*, 80:52–58, 1973.

[19] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[20] R. Fabila-Monroy and J. López. Computational search of small point sets with small rectilinear crossing number. *Journal of Graph Algorithms and Applications*, 18(3):393–399, 2014.

[21] S. Janson. Standard representation of multivariate functions on a general probability space. *Electron. Commun. Probab.*, 14:343–346, 2009.

[22] S. Janson. *Graphons, cut norm and distance, couplings and rearrangements*, volume 4 of *New York Journal of Mathematics. NYJM Monographs.* State University of New York, University at Albany, Albany, NY, 2013.

[23] J. D. Kalbfleisch, J. G. Kalbfleisch, and R. G. Stanton. A combinatorial problem on convex n-gons. In *Proc. Louisiana Conf. on Combinatorics, Graph Theory and Computing (Louisiana State Univ., Baton Rouge, La., 1970)*, pages 180–188. Louisiana State Univ., Baton Rouge, La., 1970.

[24] E. Katz and S. Payne. Realization spaces for tropical fans. In *Combinatorial aspects of commutative algebra and algebraic geometry*, volume 6 of *Abel Symp.*, pages 73–88. Springer, Berlin, 2011.

[25] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 2012.

[26] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[27] L. Lovász, K. Vesztergombi, U. Wagner, and E. Welzl. Convex quadrilaterals and k-sets. In *Towards a theory of geometric graphs*, volume 342 of *Contemp. Math.*, pages 139–148. Amer. Math. Soc., Providence, RI, 2004.

[28] O. Nikodym. Sur une généralisation des intégrales de M. J. Radon. *Fundamenta Mathematicae*, 15:131–179, 1930.

[29] J. Radon. Theorie und anwendungen der absolut additiven mengenfunktionen. *Sitzungsber. Acad. Wissen. Wien*, 122:1295–1438, 1913.

[30] A. A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.

[31] L. A. Santaló. *Introduction to integral geometry*, volume 1198. Hermann, 1953.

[32] E. R. Scheinerman and H. S. Wilf. The rectilinear crossing number of a complete graph and Sylvester's "four point problem" of geometric probability. *Amer. Math. Monthly*, 101(10):939–943, 1994.

[33] P. W. Shor. Stretchability of pseudolines is NP-hard. volume 4 of *Applied geometry and discrete mathematics*, pages 531–554. Amer. Math. Soc., 1991.

[34] R. Sikorski. *Funkcje rzeczywiste. Vol. I.* Monografie Matematyczne. Tom XXXV. Państwowe Wydawnictwo Naukowe, Warszawa, 1958.

[35] W. A. Stein et al. *Sage Mathematics Software (Version 6.1)*. The Sage Development Team, 2013. http://www.sagemath.org.

[36] E. Steinitz. Bedingt konvergente Reihen und konvexe Systeme. i-ii-iii. *J. Reine Angew. Math.*, 143 (1913), 128–175, 144 (1914), 1–40, 146 (1916), 1-52.

[37] A. Suk. On the Erdős-Szekeres convex polygon problem. *J. Amer. Math. Soc.*, 30:1047–1053, 2017.

[38] G. Szekeres and L. Peters. Computer solution to the 17-point Erdős-Szekeres problem. *ANZIAM J.*, 48(2):151–164, 2006.

[39] P. Valtr. Probability that n random points are in convex position. *Discrete & Computational Geometry*, 13(1):637–643, 1995.

A Some elementary properties from measure theory

In the proof of Lemma 3.10 we use the following property coming from measure theory.

Lemma A.1. *If (J, μ) is a probability space and X is a non-empty family of measurable subsets of J such that*

- *X generates the σ-algebra of measurable sets; and*
- *X is stable under finite unions and complementary operations,*

then for every $\varepsilon > 0$ and every measurable set A there is $B \in X$ such that $\mu(A \triangle B) \leqslant \varepsilon$, where $A \triangle B$ stands for the symmetric difference of A and B, that is $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Proof. One can prove the statement above by showing that the family of sets A for which there is indeed such an element $B \in X$ contains X and is stable under taking complements and countable unions.

The result is true if $A \in X$, since it then suffices to take $B = A$. Assuming that A satisfies that $\mu(A \triangle B) \leqslant \varepsilon$ for some $B \in X$, the complement $\bar{A}$ of A (in J) also satisfies $\mu(\bar{A} \triangle \bar{B}) = \mu(A \triangle B) \leqslant \varepsilon$ and $\bar{B}$ belongs to X since X is stable under taking complements. Let $(A_i)_{i \in \mathbf{N}}$ be a countable family of measurable subsets satisfying the property and let us prove that the property holds for the set $A = \bigcup_{i \in \mathbf{N}} A_i$. For each $i \in \mathbf{N}$, we know that there is a set $B_i \in X$ such that $\mu(A_i \triangle B_i) \leqslant \varepsilon/2^i$. Taking $S_k = \bigcup_{i=1}^{k} B_i$ for $k \in \mathbf{N}$ and $S_\infty = \bigcup_{i \in \mathbf{N}} B_i$, we have $\mu(A \triangle S_\infty) \leqslant \sum_{i \in \mathbf{N}} \mu(A_i \triangle B_i) \leqslant 2\varepsilon$. Note that $(S_k)_{k \in \mathbf{N}}$ is an increasing sequence of sets of X whose union is S_∞. Since μ is a probability measure, $\mu(S_\infty)$ is finite. It follows that the real number sequence $(\mu(S_k))_{k \in \mathbf{N}}$ tends to $\mu(S_\infty)$ as k tends to infinity. Let $k \in \mathbf{N}$ be an index such that $\mu(S_\infty) - \mu(S_k) \leqslant \varepsilon$. Then $\mu(S_\infty \triangle S_k) = \mu(S_\infty) - \mu(S_k) \leqslant \varepsilon$ because $S_k \subseteq S_\infty$. It thus follows that $\mu(A \triangle S_k) \leqslant \mu(A \triangle S_\infty) + \mu(S_\infty \triangle S_k) \leqslant 3\varepsilon$. ◻

We focus on Borel measures of $\mathbf{R}^n$. In this case, the atoms are precisely the singletons with positive measure.

Lemma A.2 (Sikorski [34])**.** *Let μ be an atom-free measure on a set J and let $A \subseteq J$ be a measurable set with finite μ-measure. Then for every every non-negative number $x \leqslant \mu(A)$, there is a measurable subset $B \subseteq A$ with $\mu(B) = x$.*

Proof. If y is a real number and $B_1, B_2 \subseteq J$ are measurable sets satisfying $\mu(B_1) \leqslant y \leqslant \mu(B_2)$, we define

$$\alpha(B_1, B_2, y) = \sup \{ \mu(X) \mid \mu(X) \leqslant y, B_1 \subseteq X \subseteq B_2 \}$$

where only measurable sets X are considered.

We first prove that for every measurable set B and real number y such that $y \leqslant \mu(B) < \infty$, there exists a measurable set $C \subseteq B$ satisfying $\mu(C) = \alpha(C, B, y)$. To see this, we fix a sequence $(\varepsilon_n)_{n\in\mathbf{N}}$ of positive numbers tending to 0. Next, we define an increasing sequence of sets $(C_n)_{n\in\mathbf{N}}$ satisfying $\mu(C_n) \leqslant y$ for every $n \in \mathbf{N}$ as follows. Set $C_0 = \varnothing$ and observe that by the definition of α, for each $i \geqslant 1$ there exists a measurable set C_i such that $C_{i-1} \subseteq C_i \subseteq B$ and $\mu(C_i) \geqslant \alpha(C_{i-1}, B, y) - \varepsilon_i$. Set $C = \bigcup_{n\in\mathbf{N}} C_n$ and note that $\alpha(C, B, y) \leqslant \alpha(C_n, B, y)$ since $C_n \subseteq C$ for every $n \in \mathbf{N}$. It follows that

$$\forall n \geqslant 1, \quad \mu(C_n) \geqslant \alpha(C_{n-1}, B, y) - \varepsilon_n \geqslant \alpha(C, B, y) - \varepsilon_n.$$

Letting n tends to infinity yields that $\mu(C) \geqslant \alpha(C, B, y)$. This upper bound on the supremum $\alpha(C, B, y)$ is in particular reached by C, so $\mu(C) = \alpha(C, B, y)$. This proves the property stated.

Since $x \leqslant \mu(A) < \infty$, we thus know that there exists a measurable set $C_1 \subseteq A$ such that $\mu(C_1) = \alpha(C_1, A, x) \leqslant x$. Further, as $\mu(A) - x \leqslant \mu(A \setminus C_1) < \infty$, we also know that there exists a measurable set $C_2 \subseteq A \setminus C_1$ such that $\mu(C_2) = \alpha(C_2, A \setminus C_1, \mu(A) - x) \leqslant \mu(A) - x$.

As it turns out, the set $S = A \setminus (C_1 \cup C_2)$ is an atom unless it has measure 0. Indeed, suppose on the contrary that S has a measurable subset T with $0 < \mu(T) < \mu(S)$. Then $\mu(C_1 \cup T) > x$ since $\alpha(C_1, A, x) = \mu(C_1)$. Similarly, $\mu(C_2 \cup (S \setminus T)) > \mu(A) - x$ since $\alpha(C_2, A \setminus C_1, \mu(A) - x) = \mu(C_2)$. Consequently, on the one hand $\mu(A) < \mu(C_2 \cup (S \setminus T)) + \mu(C_1 \cup T)$ while on the other hand A is the disjoint union of $(C_1 \cup T)$ and $(C_2 \cup (S \setminus T))$, which is a contradiction.

Therefore, since μ is atom-free, it follows that $\mu(S) = 0$, *i.e*, $\mu(A) = \mu(C_1) + \mu(C_2)$, which implies that $\mu(C_1) = x$. □