Scaled Subgaussian Vectors are Sums of Gaussians

UC Berkeley Discrete Analysis Seminar

Jake Hofgard

April 27th, 2026

Contents

1 Introduction

In this set of notes, we present a selection of results from [Son26]. In particular, building on Cecilia Chen's talk from last week, we aim to understand how the following result can be extended to higher dimensions.

> **Theorem 1.1** (One-Dimensional Subgaussians are Sums of Gaussians)**.** *There exists a universal constant $\kappa > 0$ such that any any centered, real-valued κ-subgaussian random variable X can be written as $X = G_1 + G_2 + G_3$, where G_1, G_2, G_3 are standard Gaussians.*

In particular, we may ask whether there exists universal $q \in \mathbb{N}$ such that for any dimension $n \geq 1$ and any centered 1-subgaussian random vector $X \in \mathbb{R}^n$, there are standard Gaussian random vectors $G_1, \ldots, G_q \in \mathbb{R}^n$ that satisfy

$$X = G_1 + \ldots + G_q.$$

Surprisingly, this question has powerful implications in the realm of convex geometry. In particular, a positive answer is *equivalent* to the following conjecture due to Talagrand, which effectively posits that, given any large (with respect to the Gaussian measure), closed set in $A \subset \mathbb{R}^n$, one can build a large convex set with a constant number of Minkowksi sums of A with itself. More formally:

Conjecture 1.2. *There exists a universal $q \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ and any closed set $A \in \mathbb{R}^n$ with $\gamma_n(A) \geq \frac{2}{3}$, there exists a convex body $K \subset \mathbb{R}^n$ with $\gamma_n(K) \geq \frac{1}{2}$ and*

$$K \subset \overbrace{A + \ldots + A}^{q \text{ times}},$$

where γ_n is the standard Gaussian measure on $\mathbb{R}^n$ and $+$ denotes Minkowski summation.

Below, we will outline the proof of the equivalence between Conjecture 1.2 and the existence of a universal $q \in \mathbb{N}$ such that any centered subgaussian in $\mathbb{R}^n$ is the sum of q standard Gaussians.

Although [Son26] doesn't resolve Conjecture 1.2, the author *does* make the first substantial step towards resolving Talagrand's conjecture. Specifically, the main result that we will try to show today is as follows:

Theorem 1.3. *There exists a universal $q \in \mathbb{N}$ such that for any dimension $n \geq 1$ and $\Lambda \geq 1$ and any centered $X \in \mathbb{R}^n$ that satisfies $\|X\| \leq \Lambda$ and $\|CovX\| \leq \Lambda^2 e^{-\Lambda^2}$, there are standard Gaussian vectors $G_1, \ldots, G_q \in \mathbb{R}^n$ such that*

$$X = G_1 + \ldots + G_q.$$

Informally, as long as the covariance of X is not too large with respect to its norm, then the desired decomposition holds, with q independent of the dimension n. In fact, any random vector that satisfies this condition is itself $O(1)$-subgaussian. For 1-subgaussian random vectors, this results yields the following corollary.

Corollary 1.4. *There exists a universal $q \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ and any centered 1-subgaussian vector $X \in \mathbb{R}^n$, we can write*

$$\sqrt{\frac{\log n}{n}} X = G_1 + \ldots + G_q,$$

where $G_1, \ldots, G_q \in \mathbb{R}^n$ are standard Gaussian vectors.

Although this result doesn't give a universal decomposition of 1-subgaussians as sums of a constant number of Gaussians, thereby answering Talagrand's Conjecture 1.2, it *does* provide a convincing positive result in this direction. Combining Theorem 1.3 and the approximate Carathéodory theorem, one can in fact show that for $q_n = O((n(\log n)^{-1})^{1/3})$, Conjecture 1.2 holds with q_n in place of q. This is a significant improvement over the previously best-known bound of $q_n = O(n^{1/2})$, which follows from a combination of Gaussian isoperimetry and the Steinhaus lemma, which allows us to take K to be a Euclidean ball.

1.1 Technical Outline

Interestingly, the proof of Theorem 1.3 does not reduce to the one-dimensional case presented in Theorem 1.1. Instead, three main technical tools are needed:

(i) The closure properties of the set of random vectors that may be represented as sums of Gaussian vectors.

(ii) The theory of optimal transport maps via Caffarelli's contraction theorem, which provides conditions under which the law of a random vector is given by the pushforward of a Gaussian by a Lipschitz map.

(iii) A method for partitioning a random vector into a mixture of "simple" random vectors, each uniformly distributed on a small set of points and satisfying sharp norm and covariance bounds. An example to keep in mind is a random vector uniformly distributed on the vertices of an n-simplex, scaled to have covariance matrix proportional to the identity, which is the setting of Lemma 2.3 below.

Technique (i) was presented in some detail in last week's talk, and we present additional closure properties as needed. We will present Technique (ii), including Caffarelli's contraction theorem, without proof. The majority of this set of notes deals with Technique (iii), which involves an elegant application of [MSS15, Corollary 1.5], a key piece in the solution to the Kadison-Singer conjecture.

To prove Theorem 1.3, we decompose a random vector X into "simple" vectors using Technique (iii). Then, Techniques (i) and (ii) allows us to write each simple vector as a sum of Gaussians, from which we can conclude the theorem. Corollary 1.4 then follows by writing an arbitrary subgaussian as a sum of a constant number of random vectors satisfying the conditions of Theorem 1.3.

2 Sums of Gaussian Vectors in Higher Dimensions

Slightly simplifying the notation in [Son26], we write

$$\mathcal{G}_q^n := \{\text{random vector } X \in \mathbb{R}^n : X \stackrel{d}{=} G_1 + \ldots + G_q,\ G_1, \ldots, G_q \text{ are standard Gaussian vectors}\}.$$

Throughout, a random vector $X \in \mathbb{R}^n$ is κ-subgaussian if for any unit vector $v \in \mathbb{R}^n$,

$$\mathbb{P}(|\langle X, v\rangle| \geq t) \leq 2\exp\left(-\frac{t^2}{2\kappa^2}\right).$$

With this notation in mind, we begin by characterizing $\mathcal{G}_q^n$, establishing a handful of useful *closure properties*, in the sense of operations that preserve membership in $\mathcal{G}_q^n$ (possibly altering the parameter $q \in \mathbb{N}$).

2.1 Closure Properties of $\mathcal{G}_q^n$

We outline below as a collection of the results from [Son26, Lemma 2.1, Corollaries 2.2-3, Lemma 2.4, Corollary 2.5].

(1) **Closure Under Contractive Linear Maps.** If G is a standard Gaussian random vector in $\mathbb{R}^n$, then for any linear operator A with $\|A\| \leq 1$, there are two standard Gaussians $X, Y \in \mathbb{R}^n$ such that

$$A(G) = \frac{X+Y}{2}.$$

Additionally, if $X \in \mathcal{G}_q^n$, then for any linear map $F : \mathbb{R}^n \to \mathbb{R}^n$ with $\|F\| \leq 1$, we have that $F(X) \in \mathcal{G}_{2q}^n$.

(2) **Closure Under Positive Scaling.** If $X \in \mathcal{G}_q^n$, then for any $\tau > 0$,

$$\tau X \in \mathcal{G}^n_{(\lfloor \tau \rfloor + 2)q}.$$

(3) **Vectors of Bounded Support.** If $X \in \mathbb{R}^n$ satisfies $\|X\| \leq 1$ almost surely, then $X \in \mathcal{G}_5^n$.

(4) **Images Under Lipschitz Maps.** If $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ is C-Lipschitz and G is a standard Gaussian vector in $\mathbb{R}^n$, then there are two standard Gaussians X, Y such that

$$\Psi(G) - \mathbb{E}[\Psi(G)] = C\frac{X+Y}{2}.$$

Specifically, if $C = 1$, then $\Psi(G) - \mathbb{E}[\Psi(G)] \in \mathcal{G}_4^n$ due to Property (1).

(5) **Local-to-Global Principle.** Suppose that $q \in \mathbb{N}$, $X \in \mathbb{R}^n$ is a random vector, and μ is a discrete random variable. If, for all t in the support of μ, we have that $X \mid \{\mu = t\} \in \mathcal{G}_q^n$, then $X \in \mathcal{G}_q^n$. In fact, this result has the following corollary, which we will find useful throughout.

Corollary 2.1. *Say that $q, C \in \mathbb{N}$, $X \in \mathbb{R}^n$ is a random vector, and μ is a discrete random variable. If t is in the support of μ, let Z_t equal X conditioned on the event $\{\mu = t\}$. If $Z_t - \mathbb{E}[Z_t] \in \mathcal{G}_q^n$ and $\|\mathbb{E}Z_t\| \leq C$ for all $t \in \operatorname{supp}(\mu)$, then $X - \mathbb{E}[X] \in \mathcal{G}^n_{q+10C}$.*

Proof. We can write

$$X - \mathbb{E}[X] = \sum_{t \in \operatorname{supp}(\mu)} \mathbb{1}_{\{\mu = t\}}(Z_t - \mathbb{E}[Z_t]) + \sum_{t \in \operatorname{supp}(\mu)} \mathbb{1}_{\{\mu = t\}}\mathbb{E}[Z_t] - \mathbb{E}[X].$$

By assumption and Property (3) above, the first term belongs to $\mathcal{G}_q^n$. On the other hand, because $\|\mathbb{E}[Z_t] - \mathbb{E}[X]\| \leq 2C$, Properties (2) and (3) imply that the remaining terms belong to $\mathcal{G}^n_{10C}$, yielding the result. □

The first two properties outlined above are relatively straightforward to prove and hence omitted. Coupled with a result from [MS24] concerning convolutions of finitely-supported measures with Gaussian measures, Property (3) follows almost immediately (see Proposition A.1), while Property (4) was presented in last week's talk and is proven via a clever application of Itô's lemma. The first part of Property (5) follows directly upon applying Bayes' rule.

In a similar vein, we have the celebrated result often referred to as Caffarelli's contraction theorem, which characterizes the Brenier optimal transport map between the standard Gaussian measure and a 1-uniformly log-concave measure. In line with [Son26, Theorem 2.11], we state the result in a simplified form that is sufficient for our purposes.

Theorem 2.2 (Caffarelli's Contraction Theorem)**.** *Let μ be 1-uniformly log-concave measure on $\mathbb{R}^n$, in the sense that it admits a density proportional proportional to $\exp(-V(x))$ for some convex, twice-differentiable $V : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ with $\nabla^2 V \succeq I_n$ on $\mathrm{dom}(V)$. Then, there exists a 1-Lipschitz map $F : \mathbb{R}^n \to \mathbb{R}^n$ such that $F_*(\gamma_n) = \mu$, where γ_n is the standard Gaussian measure on $\mathbb{R}^n$.*

2.2 Decomposition into Simple Vectors

With these properties in mind, we begin our proof of Theorem 1.3. The first step is a technical lemma that concretely characterizes the "simple" vectors alluded to above.

Lemma 2.3. *Given any $C > 0$, there exists universal $q \in \mathbb{N}$, depending only on C, such that the following holds: if $d, d_0 \in \mathbb{N}$, $(e_1, \ldots, e_d)$ is the standard basis of $\mathbb{R}^d$ and $F : \mathbb{R}^d \to \mathbb{R}^{d_0}$ is a linear map with $\|F\| \le C$, then for $\widehat{Z}_F$ distributed uniformly on*

$$w_1 := F(\sqrt{\log d}\, e_1), \ldots, w_d := F(\sqrt{\log d}\, e_d),$$

we have that $\widehat{Z}_F - \mathbb{E}[\widehat{Z}_F] = \mathcal{G}_q^{d_0}$.

Proof. We reduce to the case $d = d_0$, noting that we can write any F as the composition of a linear map $G : \mathbb{R}^d \to \mathbb{R}^d$ and an orthogonal projection map, invoking closure properties (1) and (2) to obtain a sum of Gaussians in the general case that once we have shown the case $d = d_0$. We can also assume, without loss of generality, that $C \le 1$ and hence that $C = 1$ due to the same closure properties of the class $\mathcal{G}_q^n$.

Given this reduction, take

$$v_1 := \sqrt{\log d}\left(e_1 - \frac{1}{d}\mathbb{1}\right), \ldots, v_d := \sqrt{\log d}\left(e_d - \frac{1}{d}\mathbb{1}\right),$$

where $\mathbb{1} \in \mathbb{R}^d$ is the all-ones vector. Take Z uniformly distributed on the v_i, noting that $\mathbb{E}[Z] = 0$ as a result. Hence, to prove the lemma, we just need to show that $Z \in \mathcal{G}_q^d$, where q is independent of d. We may identify $\mathrm{span}\{v_1, \ldots, v_d\}$ with $\mathbb{R}^{d-1}$ and define a (potential unbounded) convex polytope for each $j \in [d]$:

$$R_j := \{z \in \mathbb{R}^{d-1} : \langle z, v_j \rangle \ge \langle z, v_k \rangle \text{ for all } k \ne j\}.$$

If $G_1 \sim \mathcal{N}(0, I_{d-1})$ is a standard Gaussian on $\mathbb{R}^{d-1}$, then by symmetry, G_1 belongs to each R_j with probability $1/d$. If G_1 lives over a probability space Ω, we can hence subdivide $\Omega = \bigcup_{j=1}^d \Omega_j$, where on Ω_j, we have that $G_1 \in R_j$. One can then show that

$$\mathbb{E}[G_1 \mid \mathbb{1}_{\Omega_j} = 1] = C_d v_j$$

for some $C_d \in (1/M, M)$, where $M > 0$ is independent of the dimension d. This fact, which we provide a complete proof of in Lemma A.2 below, is a consequence of the fact that (a) the R_j are convex cones that are symmetric with respect to the hyperplane $\{z : \langle z, v_j \rangle = 0\}$ and (b) the expectation of the maximum coordinate of a standard d-dimensional Gaussian is $\Theta(\sqrt{\log d})$.

We then take μ_j to be the measure supported on R_j with density proportional to $\exp(-\|x\|_2^2/2)$ on R_j. It is easy to see that this measure is 1-uniformly log-concave, so by Caffarelli's contraction theorem, there exists a 1-Lipschitz map $\Psi_j : \mathbb{R}^d \to \mathbb{R}^{d-1}$ such that $(\Psi_j)_* \gamma_d = \mu_j$. In particular, because G_1, conditioned on $\mathbb{1}_{\Omega_j} = 1$, has distribution μ_j it follows that $G_1 - C_d v_j$ conditioned on $\mathbb{1}_{\Omega_j} = 1$ belongs to $\mathcal{G}_2^d$ due to the property established last week: if $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ is C-Lipschitz and G is a standard Gaussian vector in $\mathbb{R}^n$, then there are two standard Gaussians X, Y such that

$$\Psi(G) - \mathbb{E}[\Psi(G)] = C\frac{X+Y}{2}.$$

To conclude, note that if $Y_1 \in \mathbb{R}^{d-1}$ is a random vector that is equal to $G_1 - C_d v_j$ on Ω_j for each $j \in [d]$, then by Property (5) above, we have that $Y_1 \in \mathcal{G}_4^d$. In particular, $G_1 - Y_1 \in \mathcal{G}_5^d$. However, $G_1 - Y_1$ has the distribution $C_d Z_F$, as we subtract off the Gaussian component of Y_1 on each Ω_j. Because $C_d \in (1/M, M)$ for some universal M, the result then follows by Property (2). □

Equipped with this notion of a simple vectors, we now provide a technical result that allows us to decompose random vectors, thereby accessing the result of Lemma 2.3. The proof of this result is outlined in the appendix below, but it follows relatively easily from the corresponding result in [MSS15, Corollary 1.5], which we also state in more detail in the appendix (see Theorem A.3 below).

Theorem 2.4. *Let $k, m, n \in \mathbb{N}$ be such that $m \geq k$, and let $v_1, \dots, v_m \in \mathbb{R}^n$ be such that*

$$\|v_i\|^2 \leq 1, \; i \in [m] \quad \text{and} \quad \frac{1}{m}\sum_{i=1}^{m} v_i v_i^\top \preceq \frac{1}{k} I_n.$$

Then, there exists a partition $\{T_1, \dots, T_s\}$ of $[m]$ for some $s \in \mathbb{N}$ such that for some universal $\tilde{C} \geq 1$,

$$\tilde{C}^{-1} k < |T_j| < \tilde{C} k \quad \text{and} \quad \left\| \frac{1}{|T_j|} \sum_{i \in T_j} v_i v_i^\top \right\| \leq \frac{\tilde{C}}{k}, \; i \in [s].$$

Although the utility of this result may not be immediately clear, it provides the missing link between the conditions in Theorem 1.3 and Lemma 2.3 in the case that X is a random vector, uniformly distributed on a finite set. Of course, once we establish Theorem 1.3 for such random vectors, a standard approximation and compactness argument yields the result for all random vectors satisfying the norm and covariance bounds of the theorem.

2.3 Proofs of Main Results

Proof of Theorem 1.3, Uniform Case. Suppose that X is a random vector distributed uniformly on a finite set of vectors $\mathrm{supp}(X) := \{v_1, \dots, v_m\} \subset \mathbb{R}^n$, satisfying the bounds $\|X\| \leq \Lambda$ almost surely and $\|\mathrm{Cov} X\| \leq \Lambda^2 e^{-\Lambda^2}$. By adding copies of the v_i if necessary, we may assume, without loss of generality, that $m \geq e^{\Lambda^2}$. Now, by the conditions on X, we must have that $\|v_i\| \leq \Lambda$ for all $i \in [m]$, and similarly, that

$$\frac{1}{m}\sum_{i=1}^{m} v_i v_i^\top \preceq \Lambda^2 e^{-\Lambda^2} I_n.$$

If we take $w_i = \Lambda^{-1} v_i$ and $k := \lfloor e^{\Lambda^2} \rfloor$, we may apply Theorem 2.4 to obtain a partition $\{T_1, \dots, T_s\}$ of $[m]$ such that for universal $\tilde{C} \geq 1$ and each $j \in [s]$,

$$\tilde{C}^{-1} e^{\Lambda^2} < |T_j| < \tilde{C} e^{\Lambda^2} \tag{2.1}$$

and

$$\left\| \sum_{i \in T_j} v_i v_i^\top \right\| \leq \frac{\tilde{C}}{k} |T_j| \Lambda^2 \leq \tilde{C}^2 \Lambda^2 \leq \tilde{C}^2 (\log |T_j| + \log \tilde{C}),$$

using Equation (2.1) to obtain the final inequality. If we take $d_j := |T_j|$, then there exists universal constant $C_0 > 0$ such that for each $j \in [s]$, we can find a linear map $F_j : \mathbb{R}^{d_j} \to \mathbb{R}^n$ with $\|F_j\| \leq C_0$ and

$$v_i = F_j(\sqrt{\log d_j} e_{s_j(i)}),$$

where $(e_1, \dots, e_{d_j})$ is the standard basis of $\mathbb{R}^d$ and $s_j : T_j \to [d_j]$ is an arbitrary bijection. In fact, this constant and map can be constructed quite explicitly, as in Lemma A.4 below.

In turn, if Z_j is the random vector distributed uniformly on $\{v_i\}_{i \in T_j}$ and μ is an independent random variable on $[j]$ with corresponding probabilities $|T_j|/m$, then

$$X = \sum_{j=1}^{s} \mathbb{1}_{\{\mu = j\}} Z_j.$$

Hence, by Lemma 2.3, there exists universal q_0 (depending only on C_0) such that

$$Z_j - \mathbb{E}[Z_j] \in \mathcal{G}^n_{q_0}.$$

Moreover, the universal bound on the norm of the linear maps F_j ensures that $\|\mathbb{E} Z_j\| \leq C_0$ for all $j \in [s]$, we have that $X \in \mathcal{G}^n_{q_0 + 10C_0}$ due to Corollary 2.1, concluding the proof. □

We have now established that any random vector $X \in \mathbb{R}^n$ with sufficiently small norm and covariance (relative to the size of $\|X\|$) can be written as the sum of q standard Gaussians, where q is a universal integer. This remarkable result then implies the main attraction: Corollary 1.4.

Proof of Corollary 1.4. Let $M > 0$ be a constant to be chosen and partition $X = Y_1 + Y_2$, where $Y_1 = \mathbb{1}_{\{\|X\| \leq M\sqrt{n}\}} X$ and $Y_2 = X - Y_1$. Because X is 1-subgaussian, a straightforward computation shows that $\|\mathrm{Cov} X\| \leq 4$ and that $\|\mathrm{Cov} Y_1\| \leq 4$ as well, for any choice of $M > 0$. Now, note that by construction,

$$\left\| \frac{1}{2\sqrt{5}M} \sqrt{\frac{\log n}{n}} Y_1 \right\| \leq \frac{\sqrt{\log n}}{2}, \quad \left\| \mathrm{Cov}\left(\frac{1}{2\sqrt{5}M} \cdot \sqrt{\frac{\log n}{n}} Y_1 \right) \right\| \leq \frac{\log n}{n}.$$

In particular, the centered random vector

$$\frac{1}{2\sqrt{5}M} \sqrt{\frac{\log n}{n}} (Y_1 - \mathbb{E}[Y_1])$$

satisfies the condition of Theorem 1.3 with $\Lambda = \sqrt{\log n}$. Hence, using Property (2), there exists universal $p \in \mathbb{N}$ such that

$$\sqrt{\frac{\log n}{n}}(Y_1 - \mathbb{E}[Y_1]) \in \mathcal{G}_p^n.$$

On the other hand, if $M > 10$, then the 1-subgaussianity of X ensures that

$$\mathbb{P}[kM\sqrt{n} < \|Y_2\| \le (k+1)M\sqrt{n}] \le \exp(-100k^2 n)$$

for each $k \in \mathbb{N}$; see [Ver18, Proposition 6.2.1]. Defining events

$$E_k := \{kM\sqrt{n} < \|Y_2\| \le (k+1)M\sqrt{n}\}$$

we can partition the underlying probability space Ω into sets $\Omega_1, \Omega_2, \ldots$ such that for each $k \in \mathbb{N}$,

$$E_k \subset \Omega_k \subset E_k \cup \{Y_2 = 0\}.$$

Conditioned on $\mathbb{1}_{\Omega_k} = 1$, note that $\frac{1}{100M}Y_2$ is equal to zero with probability at least $1 - \exp(-(k+1)^2 n)$, and its norm is certainly at most $\frac{1}{2}(k+1)\sqrt{n}$ (in fact, it is likely much smaller than this). Additionally, a straightforward computation shows that if we set Z_k to be $\frac{1}{100M}Y_2$ conditioned on $\mathbb{1}_{\Omega_k} = 1$, then

$$\|\mathbb{E}[Z_k]\| \le (k+1)M\sqrt{n}\exp(-(k+1)^2 n) \le C_0$$

for some universal $C_0 > 0$. Similarly, a coarse bound shows that

$$\|\mathrm{Cov} Z_k\| \le (k+1)^2 n \exp(-(k+1)^2 n).$$

Hence, Theorem 1.3 applies with $\Lambda = (k+1)\sqrt{n}$, yielding a universal integer q such that

$$Z_k - \mathbb{E}[Z_k] \in \mathcal{G}_q^n.$$

By Corollary 2.1, it follows that

$$\frac{1}{100M}(Y_2 - \mathbb{E}[Y_2]) \in \mathcal{G}_{q+10C_0}^n,$$

and in turn, for some universal $q' \in \mathbb{N}$,

$$\sqrt{\frac{\log n}{n}}(Y_2 - \mathbb{E}[Y_2]) \in \mathcal{G}_{q'}^n.$$

upon multiplying through by $100M\sqrt{\frac{\log n}{n}}$ and applying Property (2). Finally, because X is centered,

$$X = X - \mathbb{E}[X] = (Y_1 - \mathbb{E}[Y_1]) + (Y_2 - \mathbb{E}[Y_2]),$$

and the proof is complete after we replace q' with $2q'$. □

3 Connections to Convex Geometry

Although [Son26] does not quite resolve Conjecture 1.2, the scaling presented in Corollary 1.4 is the first substantive step towards resolving the conjecture by means of the equivalence described Section 1. In this section, we describe some of the additional geometric consequences of Theorem 1.1 and Corollary 1.4 before providing a sketch of the proof that Conjecture 1.2 is equivalent to the question of whether it is possible to describe any 1-subgaussian vector as the sum of q standard Gaussians.

3.1 Permutation Invariant Sets and Ellipsoids

In two special settings, [Son26] proves that one can nearly resolve Conjecture 1.2.

For the first setting, recall that $A \subset \mathbb{R}^n$ is permutation invariant when $(x_1, \ldots, x_n) \in A$ if and only if $(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in A$ for any permutation $\sigma \in S_n$. The following result, which crucially leverages Theorem 1.1, shows that Conjecture 1.2 is true if we impose the further assumption that A is permutation invariant.

Theorem 3.1 (Corollary 0.5 in [Son26]). *There is a universal integer $q \in \mathbb{N}$ such that if $A \subset \mathbb{R}^n$ is a permutation invariant, closed set with $\gamma_n(A) \geq \frac{2}{3}$, then there is a permutation invariant convex body $K \subset \mathbb{R}^n$ with $\gamma_n(K) \geq \frac{1}{2}$ and $K \subset A_{(q)}$, the q-fold Minkowski sum of A with itself.*

The proof of this result is very delicate, but the key invocation of Theorem 1.1 comes by constructing a permutation invariant convex set $\Omega \subset \{x_1 + \ldots + x_n = 0\}$ such that the empirical measures

$$\nu := \frac{1}{n} \sum_{i=1}^{n} \delta_{\kappa' y_i}$$

satisfy the conditions of Theorem 1.3 for all $(y_1, \ldots, y_n) \in \Omega$, where $\kappa' > 0$ is the sufficiently small constant from Theorem 1.1. By writing ν as the pushforward of a measure on $\mathbb{R}^3$ with marginals γ_1 under the map $S : (s, t, u) \mapsto s + t + u$ due to Theorem 1.1, one may define a family of measures $\mu_i(U) := n\mu(U \cap S^{-1}(\kappa' y_i))$. In turn, if $\overline{\mu} = \otimes_{i=1}^{n} \mu_i$ is the product measure defined by the μ_i, we have that $\kappa' y = (x_1, \ldots, x_n) + (x_{n+1}, \ldots, x_{2n}) + (x_{2n+1}, \ldots, x_{3n})$ for $\overline{\mu}$-almost every $(x_1, \ldots, x_{3n}) \in \mathbb{R}^{3n}$ by construction. In particular, one reduces the task to showing that with positive $\overline{\mu}$-probability (bounded below by some universal constant), each of the three sectors of the vector $(x_1, \ldots, x_{3n}) \in \mathbb{R}^{3n}$ belongs to the set in question. Although the remainder of the proof is quite technical, the key reduction is indeed provided by Theorem 1.1.

If we remove the permutation invariance required by Theorem 3.1 and instead focus on the largest ellipsoid that can be placed inside the Minkowski sum of large sets, one obtains the following result from Theorem 1.3.

Theorem 3.2 (Corollary 0.6 in [Son26]). *There exists a universal integer $q \in \mathbb{N}$ such that if $A \subset \mathbb{R}^n$ is closed with $\gamma_n(A) \geq \frac{2}{3}$, then there exists an ellipsoid $E \subset \mathbb{R}^n$ with $\gamma_n(E) \geq \frac{1}{2}$ and $\sqrt{\frac{\log n}{n}} E \subset A_{(q)}$.*

One can show that it suffices to prove the result in the setting that A is balanced. Unlike the proof of Theorem 3.1, the proof then proceeds via contradiction. In particular, if the result fails for some sufficiently large q, then one may construct a symmetric random vector X_0, supported on the complement of $\sqrt{\frac{n}{\log n}}A_{(q)}$, with $\|\mathrm{Cov}X_0\| \leq 10$. Then, the truncation $X_1 := \mathbb{1}_{\{\|X_0\|^2 \leq Mn\}} X_0$ is centered, satisfies the same covariance bound, and is such that

$$\left\| \frac{1}{10}\sqrt{\frac{\log n}{n}} X_1 \right\| \leq \sqrt{\log n}, \quad \left\| \mathrm{Cov}\left(\frac{1}{10}\sqrt{\frac{\log n}{n}} X_1 \right) \right\| \leq \frac{\log n}{n}.$$

This allows us to apply Theorem 1.3 to $\frac{1}{10}\sqrt{\frac{\log n}{n}}X_1$, which yields

$$\sqrt{\frac{\log n}{n}} X_1 = 10(G_1 + \ldots + G_q)$$

for some standard Gaussian $G_1, \ldots, G_q$. One can show that, with positive probability and for some $k_1 < q/10$, we have that $G_i \in A_{(k_1)}$ and $X_1 = X_0$ with positive probability, showing that for some $x \in \mathrm{supp}(X_0)$ and points $a_1, \ldots, a_q \in A_{(k_1)}$,

$$\sqrt{\frac{\log n}{n}} x = 10(a_1 + \ldots a_q) \in 10A_{(k_1)} \subset A_{(q)},$$

a contradiction of the fact that X_0 is supported on the complement of $\sqrt{\frac{n}{\log n}}A_{(q)}$.

It is stated in [Son26] that, upon applying the approximate Carathéodory theorem from [Ver18, Theorem 0.0.2] in conjunction with Theorem 3.2, one can show that for $q_n = O((n(\log n)^{-1})^{1/3})$, if $A \subset \mathbb{R}^n$ is closed and balanced with $\gamma_n(A) \geq \frac{2}{3}$, then there is a convex body $K \subset \mathbb{R}^n$ with

$$\gamma_n(K) \geq \frac{1}{2}, \quad K \subset A_{(q_n)}.$$

While I haven't worked out the technical details of this statement, I believe that it will follow from taking $K = \mathrm{conv}(E \cap A)$, where E is the ellipsoid from Theorem 3.2.

3.2 Talagrand's Conjecture and Sums of Gaussians

Finally, we briefly discuss the equivalence between Conjecture 1.2 and representing 1-subgaussians as sums of Gaussians. This equivalence in terms of the following quantities in [Son26]:

(1) $q_{C,n}$ is the smallest $q \in \mathbb{N}$ such that for any closed set $A \subset \mathbb{R}^n$ with $\gamma_n(A) \geq \frac{2}{3}$, $A_{(q)}$ contains a convex body K with $\gamma_n(K) \geq \frac{1}{2}$.

(2) $q'_{C,n}$ is the smallest $q \in \mathbb{N}$ such that for any balanced closed set $A \subset \mathbb{R}^n$ with $\gamma_n(A) \geq \frac{2}{3}$, $A_{(q)}$ contains a convex body K with $\gamma_n(K) \geq \frac{1}{2}$.

(3) $q_{S,n}$ is the smallest $q \in \mathbb{N}$ such that for any centered 1-subgaussian random vector $X \in \mathbb{R}^n$, we have that $X \in \mathcal{G}_q^n$.

(3) $q'_{S,n}$ is the smallest $q \in \mathbb{N}$ such that for any centered, symmetric 1-subgaussian random vector $X \in \mathbb{R}^n$, there is $Y \in \mathcal{G}_q^n$ and a random vector R with $\mathbb{P}(\|R\| \leq 1) \geq \frac{1}{2}$ such that $\mathrm{supp}(Y + R) \subseteq \mathrm{supp}(X)$.

Counterintuitively, if any one of these quantities is a dimension independent constant, then they all are.

Theorem 3.3 (Theorem 1.1 in [Son26]). *As $n \to \infty$, the following are equivalent: (i) $q_{C,n} = O(1)$, (ii) $q'_{C,n} = O(1)$, (iii) $q_{S,n} = O(1)$, and (iv) $q'_{S,n} = O(1)$.*

The proof of this result, as one might expect, is rather involved and requires some machinery from convex geometry. However, it is not difficult to provide an outline upon noting that (i) clearly implies (ii) and (iii) implies (iv). Hence, one only needs to show that (ii) implies (iii) and (iv) implies (i) to conclude that all four are equivalent.

To see that (ii) and (iii) are equivalent, one considers the empirical measures

$$V_x = \frac{1}{M} \sum_{i=1}^{M} \delta_{x_i}$$

associated with sets of points $(x_1, \ldots, x_M) \in \mathbb{R}^{nM}$. If $X \in \mathbb{R}^n$ is a fixed 1-subgaussian vector and $\varepsilon > 0$, then take

$$\mathcal{X}_{n,M,\varepsilon} := \{x \in \mathbb{R}^{nM} : W_1(V_x, X) \leq \varepsilon\}$$

and

$$\mathcal{Y}_{n,M,\varepsilon} := \{x \in \mathbb{R}^{nM} : W_1(V_x, \lambda G) \leq \varepsilon \text{ for some } \lambda \in [-1,1]\}.$$

One may then show, using (ii) and the Glivenko-Cantelli theorem concerning the convergence of empirical measures, that there exists a convex body $K_{n,M} \subset \mathbb{R}^{nM}$ such that if $X_1, \ldots, X_M$ are i.i.d. copies of X, then $(X_1, \ldots, X_M) \in K_{n,M} \cap \frac{1}{p}\mathcal{X}_{n,M}$ for some universal integer $p \in \mathbb{N}$. Moreover, (ii) implies the existence of some $g_1, \ldots, g_q \in \mathcal{Y}_{n,M}$ such that

$$W_1\left(\frac{1}{p}X, \lambda_1 G_1 + \ldots + \lambda_q G_q\right) \leq O(\varepsilon),$$

where $\lambda_1, \ldots, \lambda_q \in [-1,1]$ and G_i are standard Gaussian vectors. After taking an appropriate limit as $M \to \infty$ and $\varepsilon \to 0$ and applying Prokhorov's compactness theorem to establish convergence, one obtains

$$\frac{1}{p}X = \lambda_1 G_1 + \ldots + \lambda_q G_q,$$

and we may hence take $q_{S,n} \leq 2pq$ due to Property (2) above.

To see that (iv) implies (i), [Son26] reduces to the case that A is symmetric and has interior with Gaussian measure at least $\frac{2}{3}$. For such sets, it is possible, using some technical tools that we omit here, that there is some universal integer k_1 depending on $q'_{S,n}$ such that

$$\gamma_n(A_{(k_1)}) \geq 1 - \frac{1}{10q'_{S,n}}.$$

The proof then proceeds by contradiction, essentially in the same style as the proof of Theorem 3.2. Specifically, if $A_{(k_0)}$ does not contain a convex body of Gaussian measure at least $1/2$ for some sufficiently large $k_0 \in \mathbb{N}$, then one can construct a symmetric 1-subgaussian vector $X \in \mathbb{R}^n$ with

$$\operatorname{supp} X \subset \mathbb{R}^n \setminus A_{(k_0)}, \tag{3.1}$$

where k_0 and k_1 are related by an integer multiple. Then, there are standard Gaussians provided by (iv) and a random vector R, also as in (iv), such that

$$G_1 + \ldots + G_{q'_{S,n}} + R \in \operatorname{supp}(X)$$

almost surely. With positive probability, it follows that $G_j \in A_{(k_1)}$ and $\|R\| \leq 1$ so that there exists $x \in \operatorname{supp}(X)$ that can be written as

$$x = g_1 + \ldots + g_{q'_{S,n}} + r,$$

where $g_1, \ldots, g_{q'_{S,n}} \in A_{(k_1)}$ and $\|r\| \leq 1$. The Steinhaus lemma (c.f. [Son26, Lemma 1.2]) can be used to show that because $\|r\| \leq 1$, we have that $r \in A_{(k_2)}$ for some integer $k_2 \in \mathbb{N}$ such that $q_{S,n'} k_1 + k_2 \leq k_0$, a contradiction of Equation (3.1)

References

[MS24] Dan Mikulincer and Yair Shenfeld, *The Brownian transport map*, Probability Theory and Related Fields **190** (2024), no. 1-2, 379–444.

[MSS15] Adam Marcus, Daniel Spielman, and Nikhil Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem*, Annals of Mathematics (2015), 327–350.

[Son26] Antoine Song, *Sum of Gaussian vectors and large sets*, 2026.

[Ver18] Roman Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge University Press, September 2018.

[vH17] Ramon van Handel, *Structured Random Matrices*, p. 107–156, Springer New York, 2017.

A Additional Proofs

In this section, we clean up some of the technical details from the exposition above. First, we prove closure Property (3), concerning vectors of bounded support. First, we have the following result from [MS24].

Proposition A.1 (Theorem 1.3 in [MS24]). *Let $p := \gamma_d * \nu$ be the convolution of the standard Gaussian measure γ_d with a probability measure ν on $\mathbb{R}^d$ that is supported on a ball of radius R. Then, the Brownian transport map between γ and p is an almost-sure contraction with Lipschitz constant $\left(\frac{e^{2R^2}-1}{2}\right)^{1/2}\frac{1}{R}$.*

With this in hand, the proof of Property (3) is relatively direct.

Proof of Property (3). By Proposition A.1, we immediately observe that if $X \in \mathbb{R}^n$ is a random vector such that $\|X\| \le 1$ almost surely, then $X - G \overset{d}{=} F(G')$, where F is a transport map with Lipschitz constant at most $\left(\frac{e^2-1}{2}\right)^{1/2}$ and $G, G' \sim \mathcal{N}(0, I_n)$. Here, we use the fact that $-G \overset{d}{=} G$. By Property (4), we have that

$$X - G = \frac{1}{2}\left(\frac{e^2-1}{2}\right)^{1/2}(G_1 + G_2),$$

where $G_1, G_2 \sim \mathcal{N}(0, I_n)$ as well. By Property (2), the lefthand side belong to $\mathcal{G}_4^n$, and Property (3) follows. □

Next, we have the following technical lemma that complements the proof of Lemma 2.3.

Lemma A.2. *Suppose that $d > 1$, and take*

$$v_1 := \sqrt{\log d}\left(e_1 - \frac{1}{d}\mathbb{1}\right), \dots, v_d := \sqrt{\log d}\left(e_d - \frac{1}{d}\mathbb{1}\right),$$

the same collection of vectors from the proof of Lemma 2.3 and partition the $(d-1)$-dimensional subspace $V \subset \mathbb{R}^d$ spanned by $\{v_i\}_{i\in[d]}$ into the regions

$$R_j := \{z \in V : \langle z, v_j\rangle \ge \langle v, z_k\rangle \text{ for all } k \ne j\}.$$

If G_1 is a standard Gaussian supported on V and Ω_j, $j \in [d]$, is a partition of the probability space Ω over which G_1 is defined into the events $G_1 \in R_j$, then there exists universal $M > 0$ such that

$$\mathbb{E}[G_1 \mid \mathbb{1}_{\Omega_j} = 1] = C_d v_j$$

for $C_d \in (1/M, M)$.

Proof. To construct such G_1, we may take $G = (G^{(1)}, \dots, G^{(d)}) \sim \mathcal{N}(0, I_d)$ and then take

$$G_1 = G - \frac{1}{d}\sum_{i=1}^{d} G^{(i)}.$$

Then, we compute

$$\langle G_1, v_j \rangle = \sqrt{\log d}(G^{(j)} - \frac{1}{d}\sum_{i=1}^{d} G^{(i)}).$$

Hence, $G_1 \in R_j$ if and only if $G^{(j)} \geq G^{(k)}$ for all $k \neq j$. In other words, $\mathbb{1}_{\Omega_j} = 1$ if and only if $G^{(j)} = \max_{j\in[d]} G^{(j)}$. By the symmetry of the Gaussian and the fact that each region R_j is a convex cone centered around the ray defined by v_j, it is clear that $\mathbb{E}[G_1 \mid \mathbb{1}_{\Omega} = 1] = C_d v_j$ for some constant $C_d > 0$ that may, a priori, be dimension dependent. Now, observe that by linearity of conditional expectation, an easy computation shows that

$$\mathbb{E}[\langle G_1, v_j \rangle \mid \mathbb{1}_{\Omega_j} = 1] = C_d \|v_j\|^2 = C_d \log d \left(1 - \frac{1}{d}\right).$$

On the other hand,

$$\begin{aligned}\mathbb{E}[\langle G_1, v_j \rangle \mid \mathbb{1}_{\Omega_j} = 1] &= \sqrt{\log d}\mathbb{E}[G^{(j)} \mid G^{(j)} \geq G^{(k)} \text{ for all } k \neq j] \\ &= \sqrt{\log d}\mathbb{E}\left[\max_{j\in[d]} G^{(j)}\right],\end{aligned}$$

using the above characterization of Ω_j and the fact that the mean $\frac{1}{d}\sum_{i=1}^{d} G^{(i)}$ is independent of the increments $G^{(j)} - G^{(k)}$ for all $j, k \in [d]$, as can be seen by a direct covariance calculation. Hence, it suffices to show that for $G \sim \mathcal{N}(0, I_d)$, we have that

$$C_d := \frac{d\mathbb{E}\left[\max_{j\in[d]} G^{(j)}\right]}{\sqrt{\log d}(d-1)} \in (1/M, M)$$

for some universal constant $M > 0$ independent of d. Now, we can conclude due to the fact that if $G \sim \mathcal{N}(0, I_d)$, then

$$\mathbb{E}\left[\max_{j\in[d]} G^{(j)}\right] \in \left[\sqrt{\frac{\log d}{\pi \log 2}}, \sqrt{2\log d}\right].$$

For a proof of this fact, see [vH17, p. 5]. In particular, it seems that it suffices to take $M = 2\sqrt{2\pi \log 2}$, which still leaves a bit of room to spare. □

Next, we provide a bit more detail about Theorem 2.4, including the key result from [MSS15] that was cited above. In particular, [Son26] makes use of the following generalization of Weaver's KS_2 conjecture, which was proven in [MSS15] and implies the Kadison-Singer conjecture.

Theorem A.3 (Corollary 1.5 in [MSS15]). *Let $r \in \mathbb{N}$ and $u_1, \ldots, u_m \in \mathbb{C}^d$ be vectors such that*

$$\sum_{i=1}^{m} u_i u_i^* = I_d, \quad \|u_i\|^2 \leq \delta, \quad i \in [m].$$

Then, there exists a partition $\{S_1, \ldots, S_r\}$ of $[m]$ such that

$$\left\| \sum_{i \in S_j} u_i u_i^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^{1/2}, \quad j \in [r].$$

Proof of Theorem 2.4. Take $e_1, \ldots, e_k$ to be the standard basis of $\mathbb{R}^k$, and choose $u_1, \ldots, u_m \in \mathbb{R}^k$ such that $u_i = e_j$ if $i \equiv j \mod k$. Then, we observe that

$$\frac{1}{2k} I_k \preceq \frac{1}{m} \sum_{i=1}^{m} u_i u_i^\top \preceq \frac{2}{k} I_k,$$

as the eigenvectors of this frame operator are simply the u_i, with corresponding eigenvalues that are either $\frac{1}{m}\lfloor \frac{m}{k} \rfloor$ or $\frac{1}{m}\lceil \frac{m}{k} \rceil$. With this in mind, take $w_i := (u_i, v_i) \in \mathbb{R}^{k+n}$, observing that $1 \leq \|w_i\|^2 \leq 2$ by construction. Moreover, by the Cauchy-Schwarz inequality,

$$\frac{1}{m} \sum_{i=1}^{m} w_i w_i^\top \preceq \frac{6}{k} I_{k+n},$$

so as long as $k \geq 6$, we can apply Theorem A.3. Taking $r = \lfloor \frac{m}{k} \rfloor$ and $\delta = 2$, Theorem A.3 tells us that there exists a partition $S_1, \ldots, S_r$ of $[m]$ such that

$$\left\| \sum_{i \in S_j} w_i w_i^\top \right\| \leq M$$

for some universal constant $M > 0$ independent of k and n ([Son26] takes $M = 50$). To gain control of the size of each set in the partition, we use the fact that each of w_i is a unit vector and take traces:

$$\sum_{i \in S_j} u_i u_i^\top \preceq M I_k \implies |S_j| = \sum_{i \in S_j} \|u_i\|^2 \leq Mk.$$

Similarly, we have that

$$\sum_{j=1}^{r} |S_j| = m \geq rk.$$

As a result, if we relabel the sets S_i so that $|S_j|$ is monotone nondecreasing in j, then for each $t = 1, \ldots, \lfloor \frac{r}{100} \rfloor + 1$, it must be the case that $|S_t| \geq \frac{k}{3}$. To conclude, we partition $\{1, \ldots, r\} \setminus \{1, \ldots, \lfloor \frac{r}{100} \rfloor + 1\}$ into sets of size at most 100, indexed as $O_1, \ldots, O_{\lfloor \frac{r}{100} \rfloor + 1}$, so that

$$\left| \bigcup_{j \in O_t} S_j \right| \leq 100Mk.$$

For $t = 1, \dots, \lfloor \frac{r}{100} \rfloor + 1$, we may then define $T_t := S_t \cup \bigcup_{j \in O_t} S_j$, which fully partitions $[m]$ and satisfies $|T_t| \leq 101Mk$ and $|T_t| \geq \frac{k}{3}$. Moreover, because $|O_t| \leq 100$, we have that

$$\left\| \frac{1}{|T_t|} \sum_{i \in T_t} v_i v_i^\top \right\| \leq \left\| \frac{1}{|T_t|} \sum_{i \in S_t} v_i v_i^\top \right\| + \left\| \frac{1}{|T_t|} \sum_{j \in O_t} \sum_{i \in S_j} v_i v_i^\top \right\| \leq \frac{3M}{k} + \frac{300M}{k} = \frac{303M}{k},$$

concluding the proof. □

Finally, we prove a technical result that was used in the proof of Theorem 1.3 to construct the linear maps F_j that allow us to describe the support of a random vector as simple vectors to which Lemma 2.3 applies.

Lemma A.4. *Suppose that $\{v_i\}_{i \in T}$ is a collection of vectors in $\mathbb{R}^n$, indexed by a set T with $|T| = d \geq 2$, such that for some $\tilde{C} > 0$,*

$$\left\| \sum_{i \in T} v_i v_i^\top \right\| \leq \tilde{C}^2 (\log d + \log \tilde{C}).$$

Then, there exists a constant $C > 0$, depending only on $\tilde{C}$, such that we can find a linear map $F : \mathbb{R}^d \to \mathbb{R}^n$ with $\|F\| \leq C$ and $v_i = F(\sqrt{\log d} e_i)$ for all $i \in T$, where we have ordered the elements of T arbitrarily as $T = \{i_1, \dots, i_d\}$ and $(e_1, \dots, e_d)$ is the standard basis of $\mathbb{R}^d$.

Proof. We construct this map directly, simply defining it on the standard basis of $\mathbb{R}^d$ and extending by linearity. In particular, let

$$F(e_j) := \frac{1}{\sqrt{\log d}} v_{i_j}, \quad j \in [d].$$

Then, by linearity, we see that

$$F(\sqrt{\log d} e_j) = v_{i_j}, \quad j \in [d],$$

as desired. To compute the operator norm of F, it suffices to compute the largest singular of F, which is the square root of the largest eigenvalue of $FF^\top$. Noting that $F(e_j)$ is simply the jth column of the matrix $F \in \mathbb{R}^{n \times d}$, we see that

$$FF^\top = \frac{1}{\log d} \sum_{j=1}^{d} v_{i_j} v_{i_j}^\top = \frac{1}{\log d} \sum_{i \in T} v_i v_i^\top,$$

so

$$\left\| FF^\top \right\| \leq \frac{1}{\log d} \left\| \sum_{i \in T} v_i v_i^\top \right\| \leq \frac{\tilde{C}^2 (\log d + \log \tilde{C})}{\log d} = \tilde{C}^2 \left(1 + \frac{\log \tilde{C}}{\log d} \right) \leq \tilde{C}^2 \left(1 + \frac{\log \tilde{C}}{\log 2} \right) =: C^2,$$

completing the proof. □