Approximate Gaussian isoperimetry for k sets

Gideon Schechtman[*]

August 16, 2011

Abstract

Given $2 \le k \le n$, the minimal $(n-1)$-dimensional Gaussian measure of the union of the boundaries of k disjoint sets of equal Gaussian measure in $\mathbb{R}^n$ whose union is $\mathbb{R}^n$ is of order $\sqrt{\log k}$. A similar results holds also for partitions of the sphere S^{n-1} into k sets of equal Haar measure.

1 Introduction

Consider the canonical Gaussian measure on $\mathbb{R}^n$, γ_n. Given $k \in \mathbb{N}$ and k disjoint measurable subsets of $\mathbb{R}^n$ each of γ_n measure $1/k$ we can compute the $(n-1)$-dimensional Gaussian measure of the union of the boundaries of these k sets. Below (see Definition 1) we shall make clear what exactly we mean by the $(n-1)$-dimensional Gaussian measure but in particular our normalization will be such that the $(n-1)$-dimensional Gaussian measure of a hyperplane at distance t from the origin will be $e^{-t^2/2}$ (and not $\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$ which is also a natural choice). The question we are interested in is what is the minimal value that this quantity can take when ranging over all such partitions of $\mathbb{R}^n$. As is well known, the Gaussian isoperimetric inequality ([Bo, ST] implies that, for $k = 2$, the answer is 1 and is attained when the two sets are half spaces. The answer is also known for $k = 3$ and $n \ge 2$ and is given by 3 $2\pi/3$-sectors in $\mathbb{R}^2$ (product with R^{n-2}) ([CCH]). The value in question is then $3/2$. If the k sets are nice enough (for example if, with respect to the $(n-1)$-dimensional Gaussian measure, almost every point in

the union of the boundaries of the k sets belongs to the boundary of only two of the sets) then the quantity in question is bounded from below by $c\sqrt{logk}$ for some absolute $c > 0$. This was pointed out to us by Elchanan Mossel. Indeed, by the Gaussian isoperimetric inequality, the boundary of each of the sets has measure at least $e^{-t^2/2}$ where t is such that $\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-s^2/2}ds = 1/k$. If k is large enough t satisfies

$$\frac{e^{-t^2/2}}{\sqrt{2\pi}2t} < \frac{1}{k} < \frac{e^{-t^2/2}}{\sqrt{2\pi}t}$$

which implies $\sqrt{\log k} \le t \le \sqrt{2\log k}$ and so the boundary of each of the k sets has $(n-1)$-dimensional Gaussian measure at least $e^{-t^2/2} \ge \sqrt{2\pi}t/k \ge \sqrt{2\pi\log k}/k$. Under the assumption that the sets are nice we then get a lower bound of order $\sqrt{2\pi\log k}$ to the quantity we are after[1].

Of course the minimality of the boundary of each of the k sets cannot occur simultaneously for even 3 of the k sets (as the minimal configuration is a set bounded by an affine hyperplane) so it may come as a surprise that one can actually achieve a partition with that order of the size of the boundary. To show this is the main purpose of this note. It is natural to conjecture that, for $k - 1 \le n$ the minimal configuration is that given by the Voronoi cells of the k vertices of a simplex centered at the origin of $\mathbb{R}^n$. So it would be nice to compute or at least estimate well what one gets in this situation. This seems an unpleasant computation to do. However, in Corollary 1 below we compute such an estimate for a similar configuration - for even k with $k/2 \le n$, we look at the k cells obtained as the Voronoi cells of $\pm e_i$, $i = 1, \ldots, k/2$ and show that the order of the $(n-1)$-dimensional Gaussian measure of the boundary is of order $\sqrt{\log k}$ and we deduce the main result of this note:

Main Result *Given even k with $k \le 2n$, the minimal $(n-1)$-dimensional Gaussian measure of the union of the boundaries of k disjoint sets of equal Gaussian measure in $\mathbb{R}^n$ whose union is $\mathbb{R}^n$ is of order $\sqrt{\log k}$.*

In Corollary 2 we deduce analogue estimates for the Haar measure on the sphere S^{n-1}.

This note benefitted from discussions with Elchanan Mossel and Robi Krauthgamer. I first began to think of the subject after Elchanan and I

---

[1]One may think that the right quantity should be $\sqrt{2\pi\log k}/2$ since (almost) every boundary point is counted twice but our Definition 1 is such that almost every boundary point is counted with multiplicity of the number of sets in the partition it is on the boundary of. In any case, absolute constants do not play a significant role here.

spent some time trying (alas in vain) to use symmetrization techniques to gain information on the (say, Gaussian) "k-bubble" conjecture and some variant of it (see Conjecture 1.4 in [IM]). Robi asked me specifically the question that is solved here, with some possible applications to designing some algorithm in mind (but apparently the solution turned out to be no good for that purpose). I thank Elchanan and Robi also for several remarks on a draft of this note. I had also a third motivation to deal with this question. It is related to the computation of the dependence on ε in (the probabilistic version of) Dvoretzky's theorem. It is too long to explain here, especially since it does not seem to lead to any specific result.

2   Approximate isoperimetry for k sets

We begin with the formal definition of the $(n-1)$-dimensional Gaussian measure of the boundary of a partition of $\mathbb{R}^n$ into k sets.

Definition 1 *Let $A_1, A_2, \ldots, A_k$ be a partition of $\mathbb{R}^n$ into k measurable sets. Put $A = \{A_1, A_2, \ldots, A_k\}$ and denote*

$$\partial_\varepsilon A = \cup_{i=1}^k ((\cup_{j \neq i} A_j)_\varepsilon \setminus \cup_{j \neq i} A_j)$$

(where B_ε denotes the ε-neighborhood of the set B). We shall call $\partial_\varepsilon A$ the ε-boundary of A. The $(n-1)$-dimensional Gaussian measure of the boundary of A will be defined and denoted by

$$\gamma_{n-1}(\partial A) = \liminf_{\varepsilon \to 0} \frac{\gamma_n(\partial_\varepsilon A) - \gamma_n(A)}{\sqrt{2/\pi}\varepsilon}.$$

The reason we are using the above definition of $\partial_\varepsilon A$ rather than what might look more natural, $\cup_{i=1}^k ((A_i)_\varepsilon \setminus A_i)$, is that in the former the sets in the union are disjoint, making the computation of the measure of the union easier. Note that we do not define the boundary of the partition, only the measure of the boundary. However, in simple cases when the boundary and its $(n-1)$-dimensional Gaussian measure are well understood, this definition coincides with the classical one (modulo normalization by absolute constant). In particular notice that if the partition is into two sets which are separated by a hyperplane at distance t from the origin the definition says that the $(n-1)$-dimensional Gaussian measure of the boundary is $e^{-t^2/2}$ and in particular

when $t = 0$ the measure is 1 which coincides with what we understand as the classical γ_{n-1} measure of a hyperplane through 0. This is why the factor $\sqrt{2/\pi}$ is present in the definition above.

The main technical tool here is Proposition 1 below for its proof we need the following simple inequality.

Lemma 1 *For all $\varepsilon > 0$ if C is large enough (depending on ε) then for all $k \in \mathbb{N}$*

$$\frac{1}{\sqrt{2\pi}} \int_{\sqrt{2\log \frac{k}{C}}-1}^{\sqrt{2\log Ck}} \left(\frac{1}{\sqrt{2\pi}} \int_{-s}^{s} e^{-t^2/2}dt\right)^{k-1} e^{-s^2/2}ds \geq \frac{1}{(2+\varepsilon)k}$$

Proof: Let $g_1, g_2, \ldots, g_k$ be independent identically distributed $N(0,1)$ variables. Then by symmetry,

$$\frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-s}^{s} e^{-t^2/2}dt\right)^{k-1} e^{-s^2/2}ds = P(g_1 \geq |g_2|, \ldots, |g_k|) = \frac{1}{2k}. \quad (1)$$

Also,

$$\frac{1}{\sqrt{2\pi}} \int_{0}^{\sqrt{2\log \frac{k}{C}}-1} \left(\frac{1}{\sqrt{2\pi}} \int_{-s}^{s} e^{-t^2/2}dt\right)^{k-1} e^{-s^2/2}ds$$

$$= \frac{1}{2k}\left(\frac{2}{\sqrt{2\pi}} \int_{0}^{s} e^{-t^2/2}\right)^{k}\Big]_{s=0}^{\sqrt{2\log \frac{k}{C}}-1} \quad (2)$$

$$\leq \frac{1}{2k}(1 - \frac{2}{\sqrt{2\pi}}e^{-\log \frac{k}{C}})^k \leq \frac{1}{2k}e^{-\frac{2C}{\sqrt{2\pi}}},$$

and, for C large enough,

$$\frac{1}{\sqrt{2\pi}} \int_{\sqrt{2\log Ck}}^{\infty} \left(\frac{1}{\sqrt{2\pi}} \int_{-s}^{s} e^{-t^2/2}dt\right)^{k-1} e^{-s^2/2}ds$$

$$= \frac{1}{2k}\left(\frac{2}{\sqrt{2\pi}} \int_{0}^{s} e^{-t^2/2}\right)^{k}\Big]_{s=\sqrt{2\log Ck}}^{\infty} \quad (3)$$

$$\leq \frac{1}{2k}\left(1 - \left(1 - \frac{2}{\sqrt{2\pi}} \int_{\sqrt{2\log Ck}}^{\infty} e^{-s^2/2}\right)^{k}\right) \leq \frac{1}{2k}(1 - e^{-1/C}).$$

The Lemma now follows from (1),(2) and (3). ∎

The next proposition is the main technical tool of this note. The statement involves the (appropriately normalized) $(k-1)$-dimensional Gaussian

4

measure of a certain subset of $\mathbb{R}^k$. The set is the common boundary of the Voronoi cells corresponding to e_1 and e_2 in the partition of $\mathbb{R}^k$ obtained by the Voronoi cells of $\pm e_i$, $i = 1, \ldots, k/2$. This set is a subset of a hyperplane (through the origin of $\mathbb{R}^k$) and for such a set the measure in question coincides with the canonical Gaussian measure associated with this subspace (appropriately normalized).

Proposition 1 *For each $\varepsilon > 0$ there is a C such that for all $k \geq 2$, the $(k-1)$-dimensional Gaussian measure of the set $\{(t_1, t_2, \ldots, t_k);\ t_1 = t_2 \geq |t_3|, \ldots, |t_k|\}$ is bounded between $\frac{\sqrt{\pi \log \frac{k}{C}} - 1}{(1+\varepsilon)2k(k-1)}$ and $\frac{(1+\varepsilon)\sqrt{\pi \log Ck}}{2k(k-1)}$.*

Proof: The measure in question is

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\frac{1}{\sqrt{2\pi}} \int_{-s}^s e^{-t^2/2} dt \right)^{k-2} e^{-s^2} ds.$$

Integration by parts (with parts $\left(\frac{2}{\sqrt{2\pi}} \int_0^s e^{-t^2/2} dt \right)^{k-2} e^{-s^2/2}$ and $e^{-s^2/2}$) gives that this it is equal to

$$\frac{1}{2(k-1)} \int_0^\infty \left(\frac{2}{\sqrt{2\pi}} \int_0^s e^{-t^2/2} dt \right)^{k-1} s e^{-s^2/2} ds. \tag{4}$$

Now,

$$\int_{\sqrt{2\log Ck}}^\infty \left(\frac{2}{\sqrt{2\pi}} \int_0^s e^{-t^2/2} dt \right)^{k-1} s e^{-s^2/2} ds \tag{5}$$

$$= - \int_s^\infty \left(\frac{2}{\sqrt{2\pi}} \int_0^u e^{-t^2/2} dt \right)^{k-1} e^{-u^2/2} du \Big]_{s=\sqrt{2\log Ck}}^\infty$$

$$+ \int_{\sqrt{2\log Ck}}^\infty \int_s^\infty \left(\frac{2}{\sqrt{2\pi}} \int_0^u e^{-t^2/2} dt \right)^{k-1} e^{-u^2/2} du\, ds$$

$$\leq \frac{\sqrt{2\pi}}{2k}(1 - e^{-1/C})\sqrt{2\log Ck} + \int_{\sqrt{2\log Ck}}^\infty \frac{\sqrt{2\pi}}{2k}(1 - e^{-ke^{-s^2/2}}) ds, \tag{6}$$

where the estimate for the first term in (6) follows from (3) and of the second term follows from a similar computation to (3). Now (6) is at most

$$\frac{\sqrt{2\pi}}{2Ck}\sqrt{2\log Ck} + \int_{\sqrt{2\log Ck}}^\infty \frac{\sqrt{2\pi}}{2} e^{-s^2/2} ds \leq \frac{\sqrt{2\pi}(\sqrt{2\log Ck} + 1)}{2Ck} \tag{7}$$

and we conclude that

$$\int_{\sqrt{2\log Ck}}^{\infty} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} se^{-s^2/2}ds \le \frac{\sqrt{2\pi}(\sqrt{2\log Ck}+1)}{2Ck}. \qquad (8)$$

On the other hand

$$\int_0^{\sqrt{2\log Ck}} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} se^{-s^2/2}ds \qquad (9)$$

$$\le \sqrt{2\log Ck} \int_0^{\infty} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} e^{-s^2/2}ds = \frac{\sqrt{2\pi}\sqrt{2\log Ck}}{2k}.$$

Now, (4),(8) and (9) gives the required upper bound. The lower bound (which also follows from the Gaussian isoperimetric inequality) is easier. By Lemma 1

$$\frac{1}{2(k-1)}\int_0^{\infty} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} se^{-s^2/2}ds \qquad (10)$$

$$\ge \frac{1}{2(k-1)}\int_{\sqrt{2\log\frac{k}{C}}-1}^{\sqrt{2\log Ck}} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} se^{-s^2/2}ds \qquad (11)$$

$$\ge \frac{\sqrt{2\log\frac{k}{C}}-1}{2(k-1)}\int_{\sqrt{2\log\frac{k}{C}}-1}^{\sqrt{2\log Ck}} \Big(\frac{2}{\sqrt{2\pi}}\int_0^s e^{-t^2/2}dt\Big)^{k-1} e^{-s^2/2}ds \qquad (12)$$

$$\ge \frac{\sqrt{\pi\log\frac{k}{C}}-1}{(1+\varepsilon)2k(k-1)}. \qquad (13)$$

∎

The next Corollary is the main result here, in the most general setting we need the definition of the $(n-1)$-dimensional Gaussian measure of the boundary of a partition of $\mathbb{R}^n$ into k sets given in Definition 1 above. Note that this is the first time we use the full details of the definition; Until now we dealt only with subsets of hyperplanes for which simpler definitions could suffice.

Corollary 1 *For some universal constants $0 < c < C < \infty$ and all $k = 2,3,\ldots$,*
(1) If $A = \{A_1, A_2, \ldots, A_k\}$ is a partition of $\mathbb{R}^n$ into k measurable sets each of γ_n measure $1/k$. Then $\gamma_{n-1}(\partial A) \ge c\sqrt{\log k}$.
(2) If $k \le n$, there is a partition $A = \{A_1, A_2, \ldots, A_{2k}\}$ of $\mathbb{R}^n$ into $2k$ measurable sets each of γ_n measure $1/2k$ such that $\gamma_{n-1}(\partial A) \le C\sqrt{\log k}$.

6

(1) follows very similarly to the argument in the introduction, except that there is no need for the boundary to be nice anymore: By the Gaussian isoperimetric inequality, for each $\varepsilon > 0$ and each $i = 1, \ldots, k$,

$$\gamma_n((\cup_{j \neq i} A_j)_\varepsilon \setminus \cup_{j \neq i} A_j) \geq \frac{1}{\sqrt{2\pi}} \int_t^{t+\varepsilon} e^{-s^2/2} ds,$$

where t is such that $\frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-s^2/2} ds = 1/k$. If ε is small enough, the argument in the introduction gives that the integral in question is of order $\varepsilon \frac{\sqrt{\log k}}{k}$. Since the k sets $(\cup_{j \neq i} A_j)_\varepsilon \setminus \cup_{j \neq i} A_j$ are disjoint, we deduce (1). (2) follows directly from Proposition 1 since the boundary of the partition into the Voronoi cells corresponding to $\{\pm e_i\}_{i=1}^k$ is contained in the union of $k(k-1)$ hyperplans through zero and thus $\gamma_{n-1}(\partial A)$ coincide with the classical $\gamma_{n-1}(\partial A)$ which is what is estimated in Proposition 1.

A similar result to Corollary 1 holds on the n-dimensional sphere, S^{n-1} with its normalized Haar measure σ_n. One defines the ε-boundary of a partition A of the sphere in a similar way to the first part of Definition 1 (using, say, the geodesic distance to define the ε-neighborhood of a set). Then one defines the $(n-1)$-*dimensional Haar measure of the boundary of A* by

$$\sigma_{n-1}(\partial A) = \liminf_{\varepsilon \to 0} \frac{\sigma_n(\partial_\varepsilon A) - \sigma_n(A)}{\sqrt{2n/\pi}\varepsilon}.$$

The choice of the normalization constant $\sqrt{2n/\pi}$ was made so that if the partition is into two sets separated by a hyperplane then the measure of the boundary (which "is" S^{n-2}) will be 1. The proof of the upper bound ((2) of Corollary 2) can be obtained from that of Corollary 1 by a standard reduction, using the fact that if $(g_1, \ldots, g_n)$ is a standard Gaussian vector then the distribution of $(\sum g_i^2)^{-1}(g_1, \ldots, g_n)$ is σ_n. Note that we deal only with subsets of centered hyperplanes here so there is no problem with the reduction. The lower bound (1) can also be achieved using this reduction although one needs to be a bit more careful. Alternatively, a similar argument to that of the Gaussian case (given in the introduction), replacing the Gaussian isoperimetric inequality with the spherical isoperimetric inequality can be used.

Corollary 2 *For some universal constants $0 < c < C < \infty$ and all $k = 2, 3, \ldots,$*

(1) If $A = \{A_1, A_2, \ldots, A_k\}$ is a partition of S^{n-1} into k measurable sets each of σ_n measure $1/k$. Then $\sigma_{n-1}(\partial A) \geq c\sqrt{\log k}$.
(2) If $k \leq n$, there is a partition $A = \{A_1, A_2, \ldots, A_{2k}\}$ of S^{n-1} into $2k$ measurable sets each of σ_n measure $1/2k$ such that $\sigma_{n-1}(\partial A) \leq C\sqrt{\log k}$.

Remark 1 *It may be interesting to investigate what happens when $k >> n$. In particular, if $k = 2^n$ then the partition of R^n into its $k = 2^n$ quadrants satisfy that the γ_{n-1} measure of its boundary (consisting of the coordinates hyperplanes) is $n = \log k$. Is that the best (order) that can be achieved?*

References

[Bo] Borell, Christer, Convex measures on locally convex spaces. Ark. Mat. 12 (1974), 239–252.

[CCH] Corneli, J.; Corwin, I.; Hurder, S.; Sesum, V.; Xu, Y.; Adams, E.; Davis, D.; Lee, M.; Visocchi, R.; Hoffman, N, Double bubbles in Gauss space and spheres. Houston J. Math. 34 (2008), no. 1, 181–204.

[IM] Isaksson, M; Mossel, E., Maximally stable Gaussian partitions with discrete applications. To appear.

[ST] Sudakov, V. N.; Cirel'son, B. S., Extremal properties of half-spaces for spherically invariant measures. (Russian) Problems in the theory of probability distributions, II. Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 41 (1974), 14–24, 165.

Gideon Schechtman
Department of Mathematics
Weizmann Institute of Science
Rehovot, Israel
E-mail: gideon.schechtman@weizmann.ac.il