My favorite problems

Michel Talagrand[1]

Abstract

I discuss some problems I studied but could not solve.

1 Introduction

I am not a theory builder or a visionary. Much of my mathematical activity was trying to solve specific problems which I learned from others. I studied so many problems that I can almost make statistics. When studying an open problem, most of the time, the effort is largely wasted, as the problem has a negative solution, which is provided by a counter-example to the original question. Sometimes, however, one needs to make some kind of new observation to solve the problem, and if one is lucky this observation develops into a powerful technique or machinery. My experience is that, quite amazingly, the probability that such a miracle occurs is not smaller when the problems look special and specific. It did happen to me several times that such problems were the starting points of very fruitful research directions. For this reason I am not shy to propose several such problems which look more like puzzles, in the sense that the odds seem that once the solution is found (possibly after much struggle) one will not be much wiser. This will be the case for the problems of Section 5 and especially Section 4. On the other hand, the problems of Section 2 seem more likely to touch central issues, although it could very well happen that even there a simple counter example dismisses the whole story. Fellow mathematicians, such is our life, there is no guaranteed return on investment. As for Section 3, there could also be some potential.

As this paper proposes problems on which I failed to make progress, I assume that the reader is motivated, and some easy statements will not be proved.

[1]Paris, France, michel.talagrand@gmail.com

2 Creating convexity in a few steps

Consider a subset A of $\mathbb{R}^N$. Then an easy result of Caratheodory asserts that every point in the convex hull of A is a convex combination of $N+1$ points of A. In flowery words, we can recover the convex hull of A in $N+1$ operations and we cannot do it with fewer operations. But if A is large, can we create a large convex set from A in a bounded number of operations? The canonical notion of "large" is for the natural Gaussian measure γ_N on $\mathbb{R}^N$. To obtain cleaner statements, let us assume that A is balanced, that is $tx \in A$ for $x \in A$ and $|t| \leq 1$.

Problem 2.1 *[T2] Does there exist an integer n such that whenever A is a balanced*[2] *set in $\mathbb{R}^N$ with $\gamma_N(A) \geq 3/4$ then the sum $A_{(n)} = A + \cdots + A$ (n times) contains a convex set C with $\gamma_N(C) \geq 1/2$?*

You notice that I did not dare calling this a conjecture, because it would be such an extraordinary fact if true. In this statement, the condition $\gamma_N(A) \geq 3/4$ can be replaced by $\gamma_N(A) > \alpha$ for any $\alpha > 1/2$ (changing if necessary the value of n). One simply cannot believe that such a result could be true, until one tries to disprove it. One then realizes how little is understood about the condition $\gamma_N(A) \geq 3/4$. Basically the only examples of non trivial sets A of large measure which I know are based on (extensions of) the following idea: Under the law γ_N the components of a point of $\mathbb{R}^N$ are i.i.d. of law γ_1 so the set A of points $(x_i)_{i \leq N}$ such that the empirical measure $N^{-1} \sum_{i \leq N} \delta_{x_i}$ is close to γ_1 in a suitable sense will be of probability close to 1.[3] This idea is used both in [T2] and [J]. In [T2] it is shown that $n = 2$ does not work. More precisely, given a number $L > 0$ one can find N and a balanced set A with $\gamma_N(A) \geq 3/4$ such that $L(A + A)$ does not contain a large convex set. In [J] it is a more restrictive problem which is considered: Given n the author constructs N and a balanced set $A \subset \mathbb{R}^N$ such that the set $A(n)$ of convex combinations of n points of A does not contain a large convex set. To provide a negative answer to Problem 2.1 one should require that $nA(n)$ does not contain a large convex set. This is a much stronger requirement because blowing up a very small (in the sense of γ_N) convex set by a factor n (or even 2!) may produce a very large one. In fact, A. Song [S] recently completed the proof that sets A constructed by the previous method cannot provide a negative answer to Problem 2.1, see Theorem 2.5 below.

It seems to me that Problem 2.1 is fundamental. It is all the more remarkable that 30 years after it was printed, there has been no sign of interest whatsoever from the rather large group of people focusing on the study of convex sets. However very

[2] A. Song [S] has recently shown that one obtains the same problem is one removes the word "balanced" but out of inertia we will consider only balanced sets.

[3] Such sets are not balanced, but then one replaces them by their balanced hull.

recently Antoine Song [S] has made some remarkable progress on this problem, and we will mention a few of his results.

Denoting by $(\cdot,\cdot)$ the dot product on $\mathbb{R}^N$ we say that a $\mathbb{R}^N$-valued r.v. X is *subgaussian* if for any $x \in \mathbb{R}^N$ with $\|x\| \leq 1$ and any $t > 0$ we have $\mathsf{P}(|(x,X)| \geq t) \leq 2\exp(-t^2)$.

Problem 2.2 *[S] Does there exist an integer n such that whenever X is an $\mathbb{R}^N$-valued centered subgaussian r.v., then we can find standard Gaussian $\mathbb{R}^N$-valued r.v.s $G_1,\dots,G_n$ such that $X = \sum_{i\leq n} G_i$.*

Of course here $G_1,\dots,G_n$ are not independent.

Theorem 2.3 *[S] The answer to Problem 2.1 is positive if and only if the answer to Problem 2.2 is positive.*

A considerable obstacle to a positive solution of Problem 2.1 is that one wonders by which mechanism one could produce the convex set C. Such an ominous obstacle is not present in Problem 2.2, but the difficulty there is that we understand very little about subgaussian r.v.s.

A. Song could solve Problem 2.2 when $N = 1$. The proof is surprisingly difficult.

Theorem 2.4 *[S]. There exists a number $c > 0$ such that whenever X is a real-valued centered subgaussian r.v., one may find standard real-valued Gaussian r.v.s G_1, G_2, G_3 such that $cX = G_1 + G_2 + G_3$.*

Combining this deep result with some simple observations of [T2], one obtains the following.

Theorem 2.5 *There exists a number L such that if A is a balanced permutation invariant subset of $\mathbb{R}^N$ with $\gamma_N(A) > 2/3$ then $L(A+A+A+A+A)$ contains a convex set C with $\gamma_N(C) > 1/2$.*

One reason behind this partial success is that the set C is absolutely explicit and independent of A. The significance of this result is that it shows that the most obvious approach of constructing large sets A as the sets of points $(x_i)_{i\leq N}$ such that the empirical measure $N^{-1}\sum_{i\leq N}\delta_{x_i}$ is close in some sense to the measure γ_1 cannot provide a negative answer to Problem 2.1 because such sets are invariant by permutation of the coordinates.

When faced with an absolutely impregnable problem as Problem 2.1, a good strategy is to invent a simpler problem of the same nature. A central difficulty in Problem 2.1 is that we cannot visualize the space $\mathbb{R}^N$. So, let us replace $\mathbb{R}$ by a set as simple as possible: $\{0,1\}$. Instead of balanced sets in $\mathbb{R}^N$, identifying $\{0,1\}^N$ with the set of subsets of $\{1,\dots,N\}$, we say that the set A is hereditary if

$X \in A, Y \subset X \Rightarrow Y \in A$. And we will now define $A_{(n)}$ as the collections of sets of the type $X_1 \cup \cdots \cup X_n$ for $X_1, \ldots, X_n \in A$: the operation of taking union replaces the sum in $\mathbb{R}^N$. To measure the size of A we use the measure $\mu_p = ((1-p)\delta_0 + p\delta_1)^{\otimes N}$ where $0 < p < 1$. (The value of N remains implicit in the notation μ_p. The case of interest is the case p small.) To complete the formulation of the problem, we keep in mind the following very remarkable property of Gaussian measures, see [T4]. p. 73.

Theorem 2.6 *There exists a number L with the following property. For any N and any balanced convex set $A \subset \mathbb{R}^N$ with $\gamma_N(A) \geq 3/4$ then the complement of $LA = \{Lx; x \in A\}$ is covered by a countable union of half-spaces H_ℓ with $\sum_{\ell \geq 1} \gamma_N(H_\ell) \leq 1/2$.*

For a subset I of $\{1, \ldots, N\}$ define $H_I = \{X; I \subset X\}$. When working in $\{0,1\}^N$ rather than in $\mathbb{R}^N$ it does not take much imagination to thinks that the sets H_I are appropriate substitutes for half-spaces. Also, $\mu_p(H_I) = p^{\mathrm{card} I}$. We are then led to the following definition (which is given in [T2] under a different name).

Definition 2.7 *[T2] A subset A of $\{0,1\}^N$ is p-small if $A \subset \bigcup_{\ell \geq 1} H_{I_\ell}$ with*

$$\sum_{\ell \geq 1} \mu_p(H_{I_\ell}) = \sum_{\ell \geq 1} p^{|I_\ell|} \leq 1/2.$$

For a subset A of $\{0,1\}^N$ let us now denote by $A^{(n)}$ the complement of $A_{(n)}$, that is the collection of subsets of $\{1, \ldots, N\}$ which cannot be covered by n sets in A. We can now state a combinatorical version of Problem 2.1.

Problem 2.8 *Does there exist a number n such that $A^{(n)}$ is p-small whenever $A \subset \{0,1\}^N$ satisfies $\mu_p(A) \geq 3/4$?*

Of course, it is the same problem whether or not one requires A to be hereditary. This problem is discussed at length in [T3]. It did attract some attention from experts in combinatorics, but, sadly, it remains as mysterious as ever. The one progress which has been made is that J. Park and H. Pham provided in [P-P] a magnificent proof of an exact analogue to Theorem 2.6.

Theorem 2.9 ([P-P]) *There exists a number L with the following property. Consider a family T of sequences $t = (t_i)_{i \leq N}$ with $t_i \geq 0$. Define the function f_T on $\{0,1\}^N$ by $f_T(X) = \sup_{t \in T} \sum_{i \in X} t_i$, and let $\mathsf{E} f_T = \int f_T(X) \mathrm{d}\mu_p(X)$. Then the set $\{f \geq L\mathsf{E} f_T\}$ is p-small.*

A positive solution to Problem 2.8 would imply Theorem 2.9, because the set $A = \{f \leq 4\mathsf{E}f_T\}$ satisfies $\mu_p(A) \geq 3/4$ and $\{f > 4n\mathsf{E}f_T\} \subset A^{(n)}$. However there seems to be no way to go the other direction, as the set A is of a very special type. One could say that Theorem 2.9 assumes that "one already has convexity" but gives no information about "creating convexity".

My feeling is that the main reason that Problem 2.8 remains so mysterious is that we do not know how to approach the operation " taking the union of two sets in A". It seems clear that the resulting set is in a sense "much larger than A" and the challenge is to find a way to express this. In Problem 2.8 we take the union of n sets of A, but I feel that the key would be to understand what happens for $n = 2$. This motivates the following.

Problem 2.10 *[[T2]] Does there exist a number $\alpha > 0$ with the following property: Given any subset A of $\{0,1\}^N$ with $\mu_p(A) \geq 3/4$, then the set $A^{(2)}$ is (αp)-small?*

We leave the reader convince herself that a positive solution of this problem would imply a positive solution of Problem 2.8 (maybe at the cost of replacing 3/4 by a larger number in Problem 2.8.)

There is a whole line of problems in the same direction. Consider a family $\mathcal{X}$ of subsets of $\{1,\dots,N\}$. What is the smallest value of p such that whenever $A \subset \{0,1\}^N$ satisfies $\mu_p(A) \geq 3/4$ then there is an element of $A_{(2)}$ which contains an element of $\mathcal{X}$, or equivalently, there is an element of $\mathcal{X}$ contained in the union of two elements of A? For example, assume that $\{1,\dots,N\}$ is made of q blocks of length k, and that $\mathcal{X}$ consists of the sets which meet each block exactly once. We leave it as a teaser to the reader to prove that any p large enough that $(1-p)^k \leq 1/2$ works. (A stronger result will be proved below, but the present claim is simpler.) Suppose now that $N = M^2$ is a square and think of the N points as a $M \times M$ grid. What happens when $\mathcal{X}$ is the family of sets which meet each line and each column in exactly one point? (Equivalently $\mathcal{X}$ consists of sets of the type $\{(i,\sigma(i)); 1 \leq i \leq M\}$ where σ is a permutation of $\{1,\dots,M\}$.)

These questions are connected to the idea of "weakly p-small sets" developed in [T4]. To save some space and energy let us define:

Definition 2.11 *A probability measure on $\{0,1\}^N$ is α-spread if $\nu(H_I) \leq \alpha^{\mathrm{card}I}$ for each I.*

If $\mathcal{X}$ carries a p-spread probability measure ν and $\mathcal{X} \subset \bigcup_{\ell\geq 1} H_\ell$ then $1 \leq \sum_{\ell\geq 1} \nu(H_\ell) \leq \sum_{\ell\geq 1} \mu_p(H_\ell)$. Thus X is not p-small. A p-small set does not carry a p-spread probability measure, and we may ask the following weaker version of Problem 2.10:

Problem 2.12 *Does there exist a number $\alpha > 0$ with the following property: Given any subset A of $\{0,1\}^N$ with $\mu_p(A) \geq 3/4$, then the set $A^{(2)}$ does not carry an αp-spread probability measure?*

As equivalent formulation is as follows.

Problem 2.13 *Does there exist a number $\alpha > 0$ with the following property: Given any subset A of $\{0,1\}^N$ with $\mu_p(A) \geq 3/4$ and any family $\mathcal{X} \subset \{0,1\}^N$ which carries an αp-spread probability measure we can find $X \in \mathcal{X}$ which can be covered by two elements of A?*

At first sight, the definition of α-spread probability measures looks like an appealing concrete starting point for our investigations, but in fact such probability measures are very hard to understand in general. But one can study examples. For the two classes $\mathcal{X}$ considered in the previous paragraph, the uniform probability on $\mathcal{X}$ is $1/k$-spread in the first case and in the second case it is C/M-spread for some number C. It follows from stronger results stated in the next section that Problem 2.13 has a positive solution in both cases.

3 Projections

This section is a continuation of the previous one, but it develops a new direction.

As there is no obvious way to handle the set $A^{(2)}$ one may remember the elementary fact that if $\mu_{1/2}(A) \geq 1/2$ then $\{1,\dots,N\} \in A^{(2)}$. This suggests a more general approach to Problem 2.10. For $X \in \{0,1\}^N$ let us denote by P_X the projection from $\{0,1\}^N$ to $\{0,1\}^X$ where now X is thought of as a subset of $\{1,\dots,N\}$. Let us denote by θ_X the uniform measure on $\{0,1\}^X$.

Problem 3.1 *Does there exist a number $\alpha > 0$ such that whenever $A \subset \{0,1\}^N$ is hereditary and satisfies $\mu_p(A) \geq 3/4$ then the set $\{X \in \{0,1\}^N; \theta_X(P_X(A)) < 1/2\}$ is αp-small?*

A positive solution would imply a positive solution to Problem 2.10. To provide a position answer to Problem 3.1 one has to show that if a class $\mathcal{X} \subset \{0,1\}^N$ is not αp-small, then it contains a set X such that $\theta_X(P_X(A)) \geq 1/2$. The best known result in that direction seems to be that (if α is small enough) one can find $X \in \mathcal{X}$ and $Y \in A$ such that $\mathrm{card}(Y \cap X) \geq \mathrm{card}\, Y/2$.

Despite the fact that I spent a whole year studying the problems of [T2] I thought that a little more day-dreaming might be profitable. I first must make a disclaimer. The length of time I have though about the forthcoming problems is several orders of magnitude shorter than for the previous ones, so I might very well have missed an obvious reason for which the answer to these problems is a resounding "NO".

Generally speaking, I know very little about the size of the projections of a set. An obvious line of approach is to replace in Problem 3.1 the requirement "αp-small" by "does not carry an αp-spread probability measure", that is to consider the following problem.

Problem 3.2 *Does there exist a number $\alpha > 0$ such that whenever $A \subset \{0,1\}^N$ is hereditary and satisfies $\mu_p(A) \geq 3/4$ then for each family $\mathcal{X} \subset \{0,1\}^N$ which carries an αp-spread probability measure we can find $X \in \mathcal{X}$ such that $\theta_X(P_X(A)) > 1/2$?*

As we already mentioned the difficulty here is that the structure of αp-spread measures is very mysterious, but at least we may start investigating what happens for the two concrete choices we met in the previous section. In these results, ν is the uniform measure on $\mathcal{X}$.

Proposition 3.3 *Consider two integers k, q, set $N = kq$ and think of $\{1, \ldots, N\}$ as made of q blocks of size k. Consider the class $\mathcal{X}$ of sets that meet every block in exactly one point. Then as soon as $(1-p)^k \leq 1/2$, for each set hereditary set A we have*

$$\int \theta_X(P_X(A)) \mathrm{d}\nu(X) \geq \mu_p(A) \; . \tag{3.1}$$

In particular we can find $X \in \mathcal{X}$ such that $\theta_X(P_X(A)) \geq \mu_p(A)$.

The way to prove a statement such as (3.1) is to prove that the measure μ_p can be "pushed down" to the measure η given by $\eta(A) = \int \theta_X(P_X(A)) \mathrm{d}\nu(X)$, that is to prove that there is a disintegration $\eta = \int \eta_X \mathrm{d}\mu_p(X)$ where the probability measure η_X on $\{0,1\}^N$ is supported by the subsets of X. It is then clear that the proof reduces to the case $q = 1$ where is is really easy. The proof of the following is far less obvious and is left as a challenge to the reader.

Proposition 3.4 *There exists a number L with the following property. Assume that $N = M^2$ and think of $\{1, \ldots, N\}$ as $\{1, \ldots, M\}^2$. Consider the class $\mathcal{X}$ of subsets of $\{1, \ldots, M\}^2$ which meet each line and each column is exactly one point. Then as soon as $p \geq L/M$, if $A \subset \{0,1\}^N$ is hereditary then (3.1) holds.*

Of course, based on these examples, one may become more greedy and ask the following.

Problem 3.5 *Does there exist a number $\alpha > 0$ such that whenever ν is an αp-spread probability measure (3.1) holds for each hereditary set A?*

Even when $p = 1/2$, I know very little about the size of the projections of a large set $A \subset \{0,1\}^N$.

Problem 3.6 *For $0 < \varepsilon < 1/2$, does there exist a number $\alpha(\varepsilon) > 0$ with the following property. Whenever $A \subset \{0,1\}^N$ satisfies $\mu_{1/2}(A) \geq 3/4$, then the set $\{X \in \{0,1\}^N; \theta_X(P_X(A)) < 1 - \varepsilon\}$ is $\alpha(\varepsilon)$-small?*

Note that here A is not assumed to be hereditary. Howevr it is enough to consider the case where A is hereditary.

Lemma 3.7 *We do not weaken Problem 3.6 by assuming A hereditary.*

This is proved using the time-honored "push-down method". Given $1 \leq i \leq N$ and $A \subset \{0,1\}^N$ we define A^i as follows. Whenever $i \in X \in A$ and $X \setminus \{i\} \notin A$ we remove X from A and we add $X \setminus \{i\}$ to A. Thus cardA^i = card$A$. On the other hand, it is not difficult to show that $\nu_X(P_X(A^i)) = \nu_X(P_X(A))$.[4]Iterating the push-down method for each i proves the lemma.

In order to understand what is going on in Problem 3.6 a first goal should be to find examples proving upper bounds for $\alpha(\varepsilon)$. Noga Alon showed me the following, which proves only the very weak bound $\alpha(\varepsilon) \leq L/(\log\log(1/\varepsilon))$ for small ε. Consider an integer k not too small and $N = k2^k$. Think of $\{1,\ldots,N\}$ as the union of 2^k blocks of length k. Consider the set A consisting of subsets of $\{1,\ldots,N\}$ which miss at least one these blocks, so that for large k we have $\mu_{1/2}(A) = 1-(1-2^{-k})^{2^k} \geq 1/2$. Consider then the class $\mathcal{X}$ of subsets of $\{1,\ldots,N\}$ which meet each block in exactly one point. Then for each $X \in \mathcal{X}$ we have $\nu_X(P_X(A)) = 1 - 2^{-2^k}$.The set X is not $1/k$-small because the uniform probability on X is $1/k$-spread.

Certainly one should ask what happens when in Problem 3.6 one replaces $\{0,1\}$ by a more general probability space, say $[0,1]$ with Lebesgue's measure. For $X \subset \{1,\ldots,N\}$ we denote P_X the projection from $[0,1]^N$ to $[0,1]^X$. We denote by λ_N and λ_X the natural measures on $[0,1]^N$ and $[0,1]^X$.

Problem 3.8 *Does there exist a number $\alpha > 0$ such that whenever $A \subset [0,1]^N$ satisfies $\lambda_N(A) \geq 3/4$ then the set of X for which $\lambda_X(P_X(A)) < 9/10$ is α-small?*

To understand better this problem, it helps to consider the case where $A = [0,1-1/(5N)]^N$. This example shows that if in Problem 3.8 we rather consider the set of X for which $\lambda_X(P_X(A)) < 1-\theta$ we cannot expect better that this set being α-small for α of order θ.

These problems are quite far from Problem 2.1, and it seems a worthy project to investigate what happens if we come back to $\mathbb{R}^N$, using now projections on a subspace of lower dimension. For a subspace W of $\mathbb{R}^N$ we denote by P_W the orthogonal projection of $\mathbb{R}^N$ on W and by γ_W the canonical Gaussian measure on W.

Problem 3.9 *Is it true the for some constant L, if $A \subset \mathbb{R}^N$ satisfies $\gamma_N(A) \geq 3/4$, then for $M \leq N - L\sqrt{N}$ the set of spaces W of $\mathbb{R}^N$ of dimension M for which $\gamma_W(P_W(A)) \leq 9/10$ is extremely small?*

[4]In fact, for $i \in X$ the operation "pushdown at i" commutes with the projection P_X.

That the condition $M \leq N - L\sqrt{N}$ is necessary is shown by the case where A is a ball centered at the origin. Part of the problem is to invent the appropriate concept of "extremely small" which is relevant here, and which should replace the concept of p-small sets previously used. A first step would be to show that the proportion of such spaces is very small, but possibly a much stronger result holds.

And now the best part of this story. It seems to me that nothing is known about Problem 3.9 even when one assumes A to be convex balanced. Isn't that embarrassing?

4 Matchings

The basic object of this section is a sequence $(X_i)_{i\leq N}$ of independent r.v.s uniformly distributed in the unit square $[0,1]^2$. There are typically regions of the unit square which have a deficit or an excess of points and our goal is to quantify this in different ways. This material is extensively discussed in the monograph [T4] to which we refer for all references. Our goal is simply to provide an overview of the main remaining problems. These problems touch some of the central issues of the theory of stochastic processes, so it is not possible to have an in depth discussion of the underlying issues. Rather, we simply try to hook a reader in considering these fascinating questions.

The basic idea to measure the irregularities of the set $\{X_i, i \leq N\}$ is to consider another independent sequence $(Y_i)_{i\leq N}$ and to try to match the points X_i with the points Y_j in a way that two points that get matched are close to each other. Such a matching is given by a permutation π of the set $\{1,\dots,N\}$.[5] The goal is to prove the existence of a matching in a manner that the distances $d(X_i, Y_{\pi(i)})$ are "small". Two different objectives are to minimize the sum of these distances, or their maximum.

Theorem 4.1 (The Ajtai-Komlós-Tusnády matching theorem) *With high probability we have* $\inf_\pi \sum_{i\leq N} d(X_i, Y_{\pi(i)}) \leq L\sqrt{N\log N}$.

Here and below L is a universal constant (i.e. a number independent of N) and "with high probability" means "with probability going to 1 as $N \to \infty$" (much more precise statements are available). So, the average distance between X_i and $Y_{\pi(i)}$ is $\leq L\sqrt{\log N}/\sqrt{N}$. The factor $1/\sqrt{N}$ is the natural scaling factor, and the $\sqrt{\log N}$ is reflects the irregularities in the distribution.

Theorem 4.2 (The Leighton-Shor matching theorem) *With high probability we have* $\inf_\pi \max_{i\leq N} d(X_i, Y_{\pi(i)}) \leq L(\log N)^{3/4}/\sqrt{N}$.

[5] We will lighten the expositions by calling such a permutation "a matching"

These unexpected powers of $\log N$ are optimal. A fairly obvious question is whether we can simultaneously control the sum of the distances $d(X_i, Y_{\pi(i)})$ and their maximum.

Problem 4.3 *Is it true that with high probability we can find a matching π such that $\sum_{i\le N} d(X_i, Y_{\pi(i)}) \le L\sqrt{N\log N}$ and $\max_{i\le N} d(X_i, Y_{\pi(i)}) \le L(\log N)^{3/4}/\sqrt{N}$?*

Later we will ask a considerably more ambitious question (whose solution is possibly far more difficult).

Let us now denote by X_i^1 and X_i^2 the coordinates of X_i (etc.). Peter Shor has discovered the following striking improvement of Theorem 4.1.

Theorem 4.4 (Shor's matching theorem) *With high probability we can find a matching π such that $\sum_{i\le N} |X_i^1 - Y_{\pi(i)}^1| \le L\sqrt{N\log N}$ and $\max_{i\le N} |X_i^2 - Y_{\pi(i)}^2| \le L\sqrt{\log N}/\sqrt{N}$.*

This improves Theorem 4.1 which only asserts that $\sum_{i\le N} |X_i^2 - Y_{\pi(i)}^2| \le L\sqrt{N\log N}$. We have replaced the control of the sum of these quantities by the control of each of them. Concerning the quantities $|X^1 - Y_{\pi(i)}^1|$ it is possible to do far better than controlling just their sum.

Conjecture 4.5 *With high probability there exists a matching π such that*

$$\sum_{i\le N} \exp\Big(\frac{\sqrt{N}}{L\sqrt{\log N}}|X_i^1 - Y_{\pi(i)}^1|\Big)^2 \le 2N \tag{4.1}$$

and $\max_{i\le N} |X_i^2 - Y_{\pi(i)}^2| \le L\sqrt{\log N}/\sqrt{N}$.

One can use the inequality $\exp x \ge 1 + x$ to compare this with Theorem 4.4. The best I can do in the direction of Conjecture 4.5 is to prove a result where in (4.1) the exponent 2 is replaced by $\alpha < 1/2$ (and where L is replaced by a number depending on α.)

Conjecture 4.5 is a special case of the Ultimate Matching conjecture stated below, but there is a very specific reason why we mention it separately: there is a clear road to attack it, and the road block which prevents its solution looks more technical than conceptual.

There is a deep link between matching problems and discrepancy problems. We explain first what these are. Consider a class $\mathcal{F}$ of functions on a probability space (Ω, P), all of mean 0, and an i.i.d. sequence $(X_i)_{i\le N}$ in Ω distributed according to P. A discrepancy bound is a bound on $\sup_{f\in\mathcal{F}} |\sum_{i\le N} f(X_i)|$. Such bounds quantify "how well the empirical mean approaches the true mean" uniformly on the class $\mathcal{F}$,

and their study is a central topic of probability. The standard method to obtain discrepancy bounds is to use tail inequalities (in particular Bernstein's inequality) and chaining. The possibility of using chaining requires a certain control of the size of $\mathcal{F}$ seen as a subset of $L^2(\mathsf{P})$. It is a very beautiful fact that both Theorems 4.1 and 4.2 are obtained through a discrepancy result (and duality arguments) and that in both cases the required smallness of the corresponding class $\mathcal{F}$ is deduced from the same abstract result on ellipsoids in Hilbert space (the "Ellipsoid theorem" of [T4]).

The appeal of Conjecture 4.5 is that it boils down to a discrepancy problem, as is explained in [T4] (showing this requires non trivial work). We state this problem now. Let us consider an integer $p \geq 1$ and the grid $G = \{1, \ldots, 2^p\}^2$. The components of a point u in G are denoted by k, ℓ. We denote by P the uniform probability on G. Consider the function θ on $\mathbb{R}^+$ given by $\theta(x) = x \log(x+3)$. Consider the class $\mathcal{F}$ of functions f on G for which $\sum_{u \in G} f(u) = 0$ and which satisfy the condition

$$\sum \theta(|f(k+1,\ell) - f(k,\ell)|) + \sum |f(k,\ell+1) - f(k,\ell)| \leq 2^{2p} \, . \tag{4.2}$$

Here, the first sum of over $1 \leq k \leq 2^p - 1$ and $1 \leq \ell \leq 2^p$, whereas the second sum is over $1 \leq k \leq 2^p$ and $1 \leq \ell \leq 2^p - 1$.

It is claimed in [T4] that to prove Conjecture 4.5 it suffices to prove[6] (an appropriate version of) the following, where the random variables U_i are independent uniform on G.

Conjecture 4.6 *For $N \geq 2^{2p}$ with probability $\geq 1 - 2^{-p}$ we have*

$$\sup_{f \in \mathcal{F}} \Big| \sum_{i \leq N} f(U_i) \Big| \leq L \sqrt{Np} 2^p \, .$$

The difficulty in approaching this result is that I do not even understand what is the true size of $\mathcal{F}$ seen as a subset of $L^2(\mathsf{P})$. We can formulate the question as follows, where $(g_u)_{u \in G}$ are independent standard Gaussian r.v.s.

Conjecture 4.7 *It holds that*

$$\mathsf{E} \sup_{f \in \mathcal{F}} \Big| \sum_{u \in G} g_u f(u) \Big| \leq L \sqrt{p} 2^{2p} \, . \tag{4.3}$$

This is a question about the supremum of a concrete Gaussian processes. The theory of Gaussian processes has reached a very satisfactory state, in the sense that very

[6] It might require some work to figure out all the details, but the claim is plausible!

precise results are known in complete generality, see [T4]. Unfortunately, as the previous question shows there is no magic wand to understand the combinatorics in concrete examples. Related to Conjecture 4.7 is the question of evaluating the left-hand side of (4.3) when the condition (4.2) is replaced by

$$\sum \theta_1(|f(k+1,\ell)-f(k,\ell)|)+\sum \theta_2(|f(k,\ell+1)-f(k,\ell)|)\le 2^{2p}$$

for two functions θ_1,θ_2. To show how matters are subtle, when $\theta_1(x)=\theta_2(x)=x$, in the right-hand side of (4.3) it is not possible to do better then $Lp^{3/4}2^{2p}$. This can be shown using the same construction which proves that the power $3/4$ is necessary in Theorem 4.2.

We are finally ready to state the following lovely (and beloved) main conjecture of this section.

Conjecture 4.8 (The Ultimate Matching Conjecture) *Consider $\alpha_1,\ \alpha_2>0$ with $1/\alpha_1+1/\alpha_2=1/2$. Then with high probability there exists a matching π such that for $j=1,2$ we have*

$$\sum_{i\le N}\exp\Big(\frac{\sqrt{N}}{L\sqrt{\log N}}|X_i^j-Y^j_{\pi(i)}|\Big)^{\alpha_j}\le 2N.$$

Conjecture 4.5 is the special case $\alpha_1=2,\alpha_2=\infty$. The case $\alpha_1=\alpha_2=4$ would provide a very neat generalization of Theorems 4.1 and 4.2.

When working on an open problem, there is always the chance that one is lead to discover a powerful new method, but still the odds seems that Conjecture 4.8 is just a very tough puzzle to crack.

5 Regularization from L^1 by convolution

My bet is again that this problem is not important. But it is very pretty.

It is well known that "convolution spreads regularity". For example, if one want to approximate a continuous function on $\mathbb{R}$ by a $\mathcal{C}^\infty$ function, one takes convolution with a $\mathcal{C}^\infty$ function with (small) compact support. However, some regularization is possible even when one takes convolution with a singular measure, and even when convolution is applied to L^1 functions. Here we consider only convolution on the group $G=\{-1,1\}^{\mathbb{N}}$ provided with its Haar measure λ. The group operation is denoted as a product.

Given a positive, finite measure μ on G we consider the operator T_μ on $L^1=L^1(G,\mathrm{d}\lambda)$ given by

$$T_\mu(f)(x)=\int f(xy)\mathrm{d}\mu(y)\ .$$

Let us recall that an Orlicz function φ is a convex function from $\mathbb{R}^+$ to itself with $\varphi(0) = 0$ and that the corresponding Orlicz norm is defined by $\|f\|_\varphi = \inf\{C > 0; \int \varphi(f/C)\mathrm{d}\lambda \leq 1\}$. As the following shows, it is quite a requirement on μ that T_μ "improves the integrability".

Proposition 5.1 *For some Orlicz function φ we have*

$$\|T_\mu(f)\|_\varphi \leq C\|f\|_1 \tag{5.1}$$

for all $f \in L^1$ if and only if μ is absolutely continuous with respect to λ and the Radon-Nikodym derivative $\theta = \mathrm{d}\mu/\mathrm{d}\lambda$ satisfies $\|\theta\|_\varphi \leq C$.

The moral is that (5.1) is a far too stringent requirement, so we shall consider weaker conditions. We define the function

$$\psi_\mu(u) = \sup\left\{u\lambda(\{T_\mu(f) \geq u\})\ ;\ f \geq 0, \|f\|_1 = 1\right\}.$$

Since

$$\|T_\mu(f)\|_1 = \mu(G)\|f\|_1,$$

from Markov inequality we see that

$$\psi_\mu(u) \leq \mu(G)\ .$$

Here is another simple fact, which is almost obvious.

Proposition 5.2 *If μ is absolutely continuous with respect to λ then*

$$\lim_{u\to\infty} \psi_\mu(u) = 0\ .$$

The interesting phenomenon is that as we shall see it can happen that μ is singular with respect to λ but that $\lim_{u\to\infty} \psi_\mu(0) = 0$. As the following simple fact shows, this does not happen when μ has a finite support.

Proposition 5.3 *If μ has a finite support then* $\limsup_{u\to\infty} \psi_\mu(u) \geq \mu(G)$.

To see this, may assume that μ is a probability. The support of μ is finite, so it generates a finite subgroup H of G. Consider then a subgroup H' of G, so that HH' is a subgroup of G, which is invariant under translations by elements of the support of μ. Thus if f is the indicator of HH' it is invariant under translations by elements of the support of μ, and thus $T_\mu(f) = f$. Since the measure of HH' can be as small as we wish, the result should be obvious. □

Here is another simple fact showing that when μ is singular (i.e. is supported by a set of λ-measure 0) the function ψ_μ cannot decrease too fast.

Proposition 5.4 *If μ is singular then*

$$\int_1^\infty \frac{\psi_\mu(u)}{u}\mathrm{d}u = \infty \ . \tag{5.2}$$

Proof. For a function $g \geq 0$ and a subset V of G we have, for any number $A \geq 0$,

$$\int_V g\mathrm{d}\lambda = \int_0^\infty \lambda(\{g \geq u\} \cap V)\mathrm{d}u \leq A\lambda(V) + \int_A^\infty \lambda(\{g \geq u\})\mathrm{d}u \ .$$

Consequently, if $f \geq 0$ and $\|f\|_1 = 1$, using the previous inequality for $g = T(f)$ we obtain

$$\int_V T_\mu(f)\mathrm{d}\lambda \leq A\lambda(V) + \int_A^\infty \frac{\psi_\mu(u)}{u}\mathrm{d}u \ .$$

Consider an open set U and $f \in L^1$, $f \geq 0$, $\int f\mathrm{d}\lambda = 1$, f supported by U. Consider a compact set K. Set $V = UK$. Then for $y \in K$ we have

$$\int_V f(xy)\mathrm{d}\lambda(x) \geq \int_{Uy} f(xy)\mathrm{d}\lambda(x) = \int_U f(x)\mathrm{d}\lambda(x) = 1 \ .$$

Consequently

$$\int_V T_\mu(f)(x)\mathrm{d}\lambda(x) = \int \mathrm{d}\mu(y) \int_V f(xy)\mathrm{d}\lambda(x) \geq \mu(K)$$

and finally

$$\mu(K) \leq A\lambda(UK) + \int_A^\infty \frac{\psi_\mu(u)}{u}\mathrm{d}u \ .$$

Since we assume that μ is singular, we can find a compact set K with $\mu(K) > 0$ and $\lambda(K) = 0$. We can find U with $\lambda(UK)$ as small as we wish. We then see that for each A we have $\mu(K) \leq \int_A^\infty \psi_\mu(u)/u\mathrm{d}u$. □

We are now ready to state our main problem. Given $0 \leq a \leq 1$ we consider the "biased coin" probability μ_a on G. It is the product measure on G that on each factor gives weight $(1 + a)/2$ to the point 1 (and weight $(1 - a)/2$ to the point -1), that is

$$\mu_a = \Big(\frac{1+a}{2}\delta_1 + \frac{1-a}{2}\delta_{-1}\Big)^{\otimes\mathbb{N}} \ .$$

Here is a simple fact.

Proposition 5.5 *Given $0 < a \leq 1$ there exists $C(a) > 0$ such that for $u \geq 2$ we have $\psi_{\mu_a} \geq C(a)/\sqrt{\log u}$.*

Proof. We simply look at the density of μ_a when we replace G by $\{-1,1\}^n$. On a sequence with k terms equal to 1 this density g is $(1+a)^k(1-a)^{n-k}$, so that

$$\lambda(\{g \geq (1+a)^k(1-a)^{n-k}\}) \geq 2^{-n}\binom{n}{k} .$$

We then choose k as close as possible to $(1+a)n/2$. Letting $c^2 = (1+a)^{1+a}(1-a)^{1-a}$ then $(1+a)^k(1-a)^{n-k}$ is about c^n while, using Stirling's formula, $2^{-n}\binom{n}{k}$ is about $1/(c^n\sqrt{n})$. □

Conjecture 5.6 *Given $a > 0$ there exists $C(a) > 0$ such that for $u \geq 2$ we have*

$$\psi_{\mu_a}(u) \leq \frac{C(a)}{\sqrt{\log u}} .$$

The following is proved in [T1]. In view of (3) this is about as fast a decrease as can be expected.

Theorem 5.7 *For $u \geq 3$ the probability measure $\mu = \int_0^1 \mu_{\exp(-t)}\mathrm{d}t$ satisfies*

$$\varphi_\mu(u) \leq \frac{C \log\log u}{\log u} .$$

While this paper was under review, Conjecture 5.6 was proved within an extraneous $\log\log$ factor by Y. Chen.

Theorem 5.8 ([C]) *Given $a > 0$ there exists $C(a) > 0$ such that for $u \geq 3$ we have*

$$\psi_{\mu_a}(u) \leq \frac{C(a)\log\log u}{\sqrt{\log u}} .$$

One may formulate in Gaussian space a conjecture similar to Conjecture 5.6. Within an extraneous $\sqrt{\log\log u}$ it was proved by R. Eldan and J. Lee in [E-L], and the extraneous factor was removed by J. Lehec [L].

References

[AHPT] Ascoli, Ruben; He, Xiaoyu; Park, Jinyoung, Talagrand, Michel.

[C] Chen, Yuansi. Talagrand's convolution conjecture up to loglog via perturbed reverse heat, arXiv:2511.19374

[E-L] Eldan, Ronen; Lee, James R. Talagrand's convolution conjecture on Gaussian space, IEEE Computer Society, Los Alamitos, CA, 2015, 1395–1408.

[J] Johnston, Samuel, On creating convexity in high dimensions, arXiv:2502.10382.

[L] Lehec, Joseph, Regularization in L_1 for the Ornstein-Uhlenbeck semigroup. Annales de la Faculté des Sciences de Toulouse, Mathématiques **25**, 2016, no 1, 191–204.

[P-P] Park, Jinyoung; Pham, Huy Tuan. On a conjecture of Talagrand on selector processes and a consequence on positive empirical processes. Ann. of Math. (2) **199**, 2022, no. 3, 1293–1321.

[S] Song, Antoine: Sum of Gaussian vectors and large sets, manuscript, 2026.

[T1] Talagrand, Michel. A conjecture on convolution operators, and operators from L^1 to a Banach lattice. Israel J. of Math. 68, 1989, 82–88.

[T2] Talagrand, Michel. Are all sets of positive measure essentially convex? Oper. Theory Adv. Appl., 77 Birkhäuser Verlag, Basel, 1995, 295–310.

[T3] Talagrand, Michel. Are many small sets explicitly small? Association for Computing Machinery (ACM), New York, 2010, 13–35.

[T4] Talagrand, Michel. Upper and lower bounds for stochastic processes–decomposition theorems. Ergeb. Math. Grenzgeb. (3), 60 Springer, Cham, 2021, xviii+726 pp.