Randomness and Pseudo-Randomness in Discrete Mathematics

Noga Alon *

The discovery, demonstrated in the early work of Paley, Zygmund, Erdős, Kac, Turán, Shannon, Szele and others, that *deterministic* statements can be proved by *probabilistic* reasoning, led already in the first half of the century to several striking results in Analysis, Number Theory, Combinatorics and Information Theory. It soon became clear that the method, which is now called *the probabilistic method*, is a very powerful tool for proving results in Discrete Mathematics. The early results combined combinatorial arguments with fairly elementary probabilistic techniques, whereas the development of the method in recent years required the application of more sophisticated tools from probability. The books [10], [54] are two recent texts dealing with the subject.

Most probabilistic proofs are existence, non-constructive arguments. The rapid development of theoretical Computer Science, and its tight connection to Combinatorics, stimulated the study of the algorithmic aspects of these proofs. In a typical probabilistic proof, one establishes the existence of a combinatorial structure satisfying certain properties by considering an appropriate probability space of structures, and by showing that a randomly chosen point of this space is, with positive probability, a structure satisfying the required properties. Can we find such a structure *efficiently*, that is, by a (deterministic or randomized) polynomial time algorithm ? In several cases the probabilistic proof provides such a randomized efficient algorithm, and in other cases the task of finding such an algorithm requires additional ideas. Once an efficient randomized algorithm is found, it is sometimes possible to *derandomize* it and convert it into an efficient deterministic one. To this end, certain explicit *pseudo-random* structures are needed, and their construction often requires tools from a wide variety of mathematical areas including Group Theory, Number Theory and Algebraic Geometry.

The application of probabilistic techniques for proving deterministic theorems, and the application of deterministic theorems for derandomizing probabilistic existence proofs, form an interesting combination of mathematical ideas from various areas, whose intensive study in recent years led to the development of fascinating techniques. In this paper I survey some of these developments and mention several related open problems.

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: noga@math.tau.ac.il. Research supported in part by a USA Israeli BSF grant and by the Fund for Basic Research administered by the Israel Academy of Sciences.

1 Probabilistic methods

The applications of probabilistic techniques in Discrete Mathematics, initiated by Paul Erdős who contributed to the development of the method more than anyone else, can be classified into three groups. The first one deals with the study of certain classes of random combinatorial objects, like random graphs or random matrices. The results here are essentially results in Probability Theory, although most of them are motivated by problems in Combinatorics. The second group consists of applications of probabilistic arguments in order to prove the existence of combinatorial structures which satisfy a list of prescribed properties. Existence proofs of this type often supply extremal examples to various questions in Discrete Mathematics. The third group, which contains some of the most striking examples, focuses on the application of probabilistic reasoning in the proofs of deterministic statements whose formulation does not give any indication that randomness may be helpful in their study.

The above classification is, of course, somewhat arbitrary, and there are results that can fit more than one of the above groups. Most of the combinatorial results obtained by applying probabilistic arguments belong, however, naturally to one of these groups.

There has been recent interesting progress in all three groups. This chapter contains a brief description of several typical results in each of them.

1.1 Random structures

Although there have been several papers by various researchers in the late 50's that deal with the statistical aspects of graphs, the systematic study of Random Graphs was initiated by Erdős and Rényi whose first two papers on the subject are [21], [22]. Formally, $G(n,p)$ denotes the probability space whose points are graphs on a fixed set of n labelled vertices, where each pair of vertices forms an edge, randomly and independently, with probability p. The term "the random graph $G(n,p)$" means, in this context, a random point chosen in this probability space. Each graph property A (that is, a family of graphs closed under graph isomorphism) is an event in this probability space, and one may study its probability $Pr[A]$, that is, the probability that the random graph $G(n,p)$ lies in this family. In particular, we say that A holds *almost surely* if the probability that $G(n,p)$ satisfies A tends to 1 as n tends to infinity. There are numerous papers dealing with random graphs, and the book of Bollobás [13], is an excellent extensive account of the known results in the subject proved before its publication in 1985.

One of the most important discoveries of Erdös and Rényi was the discovery of *threshold functions*. A function $r(n)$ is called a threshold function for a graph property A, if when $p(n)/r(n)$ tends to 0, then $G(n,p(n))$ does not satisfy A almost surely, whereas when $p(n)/r(n)$ tends to infinity, then $G(n,p(n))$ satisfies A almost surely. Thus, for example, it is shown in [21] that the function $r(n) = \ln n/n$ is a threshold function for the property "G is connected. " (In fact, a much more

precise estimate follows from the results in [21]: if $p(n) = \frac{\ln n}{n} + \frac{c}{n}$, then, as n tends to infinity, the probability that $G(n, p(n))$ is connected tends to $e^{-e^{-c}}$).

A graph property is *monotone* if it is closed under the addition of edges. Note that many interesting graph properties, like hamiltonicity, non-planarity, connectivity or containing at least 10 vertex disjoint triangles are monotone.

Bollobás and Thomason [15] proved that *any* monotone graph property has a threshold function. Their proof applies to any monotone family of subsets of a finite set, and relies on the Kruskal-Katona Theorem that describes the possible number of subsets of each cardinality in a monotone family. By viewing a monotone graph property as a family of subsets of the set of all potential edges, this yields the result for random graphs. Their theorem shows that for any monotone property A, if the probability that a random graph $G(n, p)$ satisfies A exceeds ϵ, then for $q \geq C(\epsilon)p$, the probability that $G(n, q)$ satisfies A is at least $1 - \epsilon$. This result applies even without the assumption that the property A is closed under graph isomorphism. In fact, if one is not interested in the precise behaviour of $C(\epsilon)$ this can be deduced simply by observing that if $(1 - \epsilon)^k < \epsilon$ then the probability that at least one of k graphs G_i chosen independently according to the distribution $G(n, p)$ satisfies A is more than $1 - \epsilon$, and hence so is the probability that their union satisfies A.

Friedgut and Kalai showed that the symmetry of graph properties can be applied to obtain a sharper result, as follows.

Theorem 1.1 ([24]) *For any monotone graph property A, if $G(n, p)$ satisfies A with probability at least ϵ, then $G(n, q)$ satisfies A with probability at least $1 - \epsilon$, for $q = p + O(\log(1/2\epsilon)/\log n)$.*

The proof follows by combining two results. The first is a simple but fundamental lemma of Margulis [41] and Russo [51], which is useful in Percolation Theory. This lemma can be used to express the derivative with respect to p of the probability that $G(n, p)$ satisfies A as a sum of contributions associated with the single potential edges. The second result is a theorem of [17] that asserts that at least one such contribution is always large. The symmetry implies that all contributions are the same and the result follows.

Another interesting early discovery in the study of Random Graphs was that of the fact that many interesting graph invariants are highly concentrated. A striking result of this type was first proved by Matula [40] and strengthened by various researchers; for fixed values of p almost all graphs $G(n, p)$ have the same *clique number*. The clique number of a graph is the maximum number of vertices in a clique of it, that is , in a subgraph in which any two vertices are adjacent. It turns out that for every fixed positive value of $p < 1$ and every n, there is a real number $r_0 = r_0(n, p)$ which is roughly $2 \log n / \log(1/p)$, such that the clique number of $G(n, p)$ is either $\lfloor r_0 \rfloor$ or $\lceil r_0 \rceil$ almost surely. Moreover, $r_0(n, p)$ can be chosen to be an integer for most values of n and p. The proof of this result is not difficult, and is based on the second moment method. One estimates the expectation and the variance of the number of cliques of a given size contained in $G(n, p)$ and applies the inequalities of

Markov and Chebyshev.

An *independent set* of vertices in a graph G is a set of vertices no two of which are adjacent. The *chromatic number* $\chi(G)$ of G is the minimum number of independent sets needed to cover all its vertices. This is a more complicated quantity than the clique number, and its behaviour for the random graph $G(n,p)$ is much less understood than the corresponding behaviour of the clique number.

Answering a problem suggested by Erdős and Rényi, Bollobás [14] showed that the chromatic number of $G(n,0.5)$ is almost surely $(1+o(1))n/2\log_2 n$. His proof applies a Martingale Inequality to show that almost surely, every set of at least, say, $n/\log^2 n$ vertices of $G(n,0.5)$ contains an independent subset of size nearly as large as the maximum independent set in the whole graph, implying that a greedy approach of omitting maximum independent sets from the graph one by one yields a nearly optimal coloring.

How concentrated is the chromatic number of $G(n,p)$? Shamir and Spencer [53] proved that there is always a choice of an interval $I = I(n,p)$ of length roughly $\sqrt{n}$, such that the chromatic number of $G(n,p)$ lies, almost surely, in I. More surprisingly, if $p(n) < n^{-5/6-\epsilon}$, then there is always such an interval containing only four distinct values. This was improved by Łuczak [38], who showed that for such values of $p(n)$ the chromatic number is actually, almost surely, one of two consecutive values. In a very recent joint work of the author and Krivelevich it is shown that this is the case whenever $p(n) \leq n^{-1/2-\epsilon}$. This implies the following.

Proposition 1.2 *For every $\alpha < 1/2$ and every integer valued function $r(n) < n^{\alpha}$, there exists a function $p(n)$ such that the chromatic number of $G(n,p(n))$ is precisely $r(n)$ almost surely.*

Therefore, for such values of $p(n)$, almost all graphs $G(n,p(n))$ have the same chromatic number ! The proofs of all these results start by applying a Martingale Inequality to show that if $\delta > 0$ is an arbitrarily small real, and t is the smallest integer for which the chromatic number of $G(n,p(n))$ is at least t with probability that exceeds δ, then one can omit, with probability at least $1-\delta$, a set of at most $C(\delta)\sqrt{n}$ vertices from $G(n,p(n))$ to get a t-colorable subgraph. This can be combined with several additional combinatorial and probabilistic tools to deduce the above results.

1.2 Probabilistic constructions

The *Ramsey number* $R(k,t)$ is the minimum number n such that every graph on n vertices contains either a clique of size k or an independent set of size t. By a special case of the celebrated theorem of Ramsey (cf., e.g., [28]), $R(k,t)$ is finite for every positive integers k and t, and satisfies $R(k,t) \leq \binom{k+t-2}{k-1}$. In particular, $R(k,k) < 4^k$. The problem of determining or estimating the numbers $R(k,t)$ received a considerable amount of attention, and seems to be very difficult in general.

In one of the first applications of the probabilistic method in Combinatorics, Erdős [18] proved that if $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then $R(k,k) > n$, that is, there exists a graph on n vertices containing neither

a clique of size k nor an independent set of size k. The proof is extremely simple; Every fixed set of k vertices in the random graph $G(n, 0.5)$ is a clique or an independent set with probability $2^{1-\binom{k}{2}}$. Thus $\binom{n}{k}2^{1-\binom{k}{2}}$ (< 1) is an upper bound for the probability that the random graph $G(n, 0.5)$ contains a clique or an independent set of size k. Despite the simplicity of this proof, there is no constructive version of it in the sense that there is no known deterministic algorithm that constructs a graph on $n > (1+\epsilon)^k$ vertices with neither a clique nor an independent set of size k, in time which is polynomial in n, where $\epsilon > 0$ is any positive absolute constant.

Ajtai, Komlós and Szemerédi [1] showed that $R(3,t) \leq O(t^2/\log t)$. Their proof is probabilistic (and can thus fit the next subsection). In a recent paper, Kim [33] proves that this is tight, up to a constant factor. This provides the correct asymptotic behaviour of $R(3,t)$:

Theorem 1.3 ([1], [33]) *There are two positive constants c_1, c_2 such that*

$$c_1 \frac{t^2}{\log t} \leq R(3,t) \leq c_2 \frac{t^2}{\log t},$$

for every t.

The proof of Kim is based on a clever "semi-random" construction and proceeds in stages. Starting from the empty graph on n vertices, in each stage choose every potential edge which does not form a triangle with two of the edges picked so far, randomly and independently, with probability $1/(\sqrt{n}\log^2 n)$. If triangles are formed, omit a maximal collection of pairwise edge disjoint triangles, thus completing the stage. The process, which clearly generates a triangle-free graph, terminates after some n^δ stages. It is shown in [33], by combining subtle combinatorial and probabilistic arguments, that with positive probability this process produces a graph whose independence number does not exceed $t = c\sqrt{n\log n}$ for an appropriate choice of an absolute positive constant c. Therefore, $R(3,t) > n = \Omega(t^2/\log t)$, as needed. As is the case with the Ramsey numbers $R(k,k)$, there is no known deterministic efficient algorithm that constructs a triangle-free graph on n vertices which contains no independent sets of size $n^{1/2+o(1)}$.

The above mentioned semi-random approach for constructing the required combinatorial structure in stages, where in each stage some correction may be applied, is influenced by a method developed by Rödl in [51], following some similar ideas that appeared already in [1]. This technique, which is sometimes called the "Rödl Nibble", was initiated by Rödl in order to solve a packing and covering problem of Erdős and Hanani [19]. His result forms another interesting example of a probabilistic construction. It asserts that for every fixed $k \geq l \geq 2$, there is a collection of at most $\binom{n}{l}/\binom{k}{l} + o(n^l)$ subsets of cardinality k of an n-element set, so that each l-element subset is contained in at least one k-tuple. Note that this means that most l-subsets are covered precisely once, that is, are contained in exactly one of the k-tuples in the collection. The proof is by repeatedly picking a small random subset of the k-tuples that do not intersect any of the ones picked already by more than $l-1$ points. By a careful analysis it can be shown that this produces, with positive probability,

a collection of at most $\binom{n}{l}/\binom{k}{l} + o(n^l)$ k-tuples that cover all but at most $o(n^l)$ l-subsets. Covering the uncovered l-subsets by additional k-sets, one obtains a collection with the desired properties.

The main part of the proof here, as well as in [33], is to maintain certain regularity properties of the combinatorial structure which is being constructed in stages, during the whole process.

This technique has been developed by several researchers, who applied it to prove various interesting results about packing, covering and coloring problems for hypergraphs. Some of these results are mentioned in the next subsection.

Probabilistic constructions have been used extensively in Combinatorial Geometry and Combinatorial Number Theory. A recent geometric example, answering a question of Füredi and Stanley [27], appears in [11], where it is shown that for every k and d there are collections of at least $d^{\Omega(\log(k+2)/\log\log(k+2))}$ nonzero vectors in R^d, in which any $k+1$ members contain an orthogonal pair.

1.3 Proving deterministic theorems

A *hypergraph* H is a pair (V, E), where V is a finite set whose members are called *vertices* and E is a finite collection of subsets of V, called *edges*. If every edge contains precisely k vertices, the hypergraph is *k-uniform*. Thus, 2-uniform hypergraphs are graphs. A *matching* in H is a subset of its edges no two of which share a vertex. A *proper coloring* of the edges of H is an assignment of colors to the edges of H so that each color class forms a matching. The *chromatic index* of H is the smallest number of colors used in a proper edge coloring of it.

Several researchers noticed that the Nibble technique developed in [51] can be applied for tackling various packing, covering and coloring problems for hypergraphs. See [25], [47] and [32] for some interesting examples. The results in all these papers are deterministic theorems about hypergraphs, and therefore belong to this subsection. The strongest result of this type, due to Kahn, deals with proper edge colorings of hypergraphs.

Theorem 1.4 ([32]) *For every $\epsilon > 0$ and every k there is a finite $D_0 = D_0(k, \epsilon)$ with the following property. Let H be a k-uniform hypergraph with maximum degree D, where $D > D_0$. If no two vertices of H share more than ϵD common edges, then for any assignment of a list of at least $D(1+\epsilon)$ colors for each edge of H, there is a proper edge coloring of H assigning to each edge a color from its list.*

In particular, this implies that the chromatic index of H does not exceed $(1+\epsilon)D$, as proved already in [47].

The proofs in the above mentioned papers and in several related ones are based on the Nibble technique, and usually combine it with several martingale inequalities or other large deviation inequalities like the one of Talagrand in [58].

A *proper k-coloring* of a graph is an assignment of a color from a set of k colors to each of its vertices so that adjacent vertices get distinct colors. Such a coloring is *acyclic* if there is no two-colored cycle. The *acyclic chromatic number* of a graph is the minimum number of colors in an acyclic coloring of it. The Four Color Theorem, which is the best known result in Graph Theory, asserts that the chromatic number of every planar graph is at most 4. Answering a problem of Grünbaum and improving results of various authors, Borodin [16] showed that every planar graph has an acyclic 5-coloring. He conjectured that for any surface but the plane, the maximum possible chromatic number of a graph embeddable on the surface, is equal to the maximum possible acyclic chromatic number of a graph embeddable on it. The Map Color Theorem (see [50]) determines precisely the maximum possible chromatic number of any graph embeddable on a surface of genus g and shows this maximum is

$$\lfloor \frac{7+\sqrt{1+48g}}{2} \rfloor = \Theta(g^{1/2}).$$

The following result shows that the maximum possible acyclic chromatic number of a graph on such a surface is asymptotically different.

Theorem 1.5 ([9]) *The acyclic chromatic number of any graph embeddable on a surface of genus g is at most $O(g^{4/7})$. This is nearly tight in the sense that for every $g > 0$ there is a graph embeddable on a surface of genus g whose acyclic chromatic number is at least $\Omega(g^{4/7}/(\log g)^{1/7})$.*

Therefore, the above mentioned conjecture of Borodin is false for all surfaces with large genus.

The proof of the $O(g^{4/7})$ upper bound is probabilistic, and combines some combinatorial arguments with the Lovász Local Lemma. This Lemma, first proved in [20], is a tool for proving that under suitable conditions, with positive probability, none of a large finite collection of nearly independent, low probability events in a probability space holds. This positive probability is often exponentially small, and yet the Local Lemma can be used to show it is positive. The proof of the $\Omega(g^{4/7}/(\log g)^{1/7})$ lower bound is also probabilistic, and is based on an appropriate random construction.

Among the deterministic theorems proved by probabilistic arguments, there are examples of probability theorems. An interesting example of this type is a derivation of a large deviation inequality of Janson ([30], [29], see also [10]). Another example is the 123-theorem proved in [12]; For every two independent identically distributed real random variables X and Y

$$Pr[|X-Y| \leq 2] < 3Pr[|X-Y| \leq 1].$$

2 Pseudo-randomness

The rapid development of theoretical Computer Science and its tight connection to Discrete Mathematics motivated the study of the algorithmic aspects of probabilistic techniques. Can a combinatorial structure whose existence is proved by probabilistic means be constructed *explicitly* (that

is, by an efficient deterministic algorithm)? Can the algorithmic problems corresponding to existence probabilistic proofs be solved by efficient procedures? The area of *randomized algorithms* has been developed tremendously during the last decade, when it has been realized that for numerous computational problems, the simplest and fastest algorithms are often randomized ones. Can such algorithms be derandomized, that is, can they be converted into efficient deterministic ones? The investigation of these questions in recent years led to fascinating techniques which are often related to other branches of Mathematics. In this section I briefly describe some of the highlights.

2.1 Expanders

An (n, d, c)-*expander* is a d-regular graph on n vertices, such that every set X of at most $n/2$ of its vertices has at least $c|X|$ neighbors outside the set. Infinite families of such graphs with fixed positive values of d and c and growing number of vertices have numerous applications in Combinatorics and Theoretical Computer Science. The simplest way of proving the existence of such families is by a probabilistic construction first described by Pinsker [46]; For every $d \geq 3$ there is some $c = c(d) > 0$ such that a random bipartite graph obtained by choosing d random permutations between the two parts is a $(2n, d, c)$-expander almost surely.

The problem of constructing such families of graphs explicitly is more complicated. Most known constructions rely on the tight relationship between the expansion properties of a graph and the ratio between its largest and second largest eigenvalues. The *adjacency matrix* of a graph $G = (V, E)$ is the matrix $A = (a_{u,v} : \ u, v \in V)$ in which $a_{u,v}$ is the number of edges between u and v. This is a symmetric matrix, and thus it has real eigenvalues and an orthonormal basis of eigenvectors. If the graph is d-regular, then the largest eigenvalue is d, and the second largest eigenvalue, which is denoted by $\lambda(G)$, is strictly smaller than d iff the graph is connected. It is not too difficult to see that any d regular graph with n vertices and second eigenvalue λ is an (n, d, c)-expander for $c = (d - \lambda)/(2d)$. This (in a slightly stronger form) has been proved, independently, by Tanner in [57] and by the author and Milman in [8]. The proof is simple and applies the variational definition of the second eigenvalue to an appropriate test function.

The converse is more complicated, but is also true, and has been proved in [6].

Theorem 2.1 *For any (n, d, c)-expander G, $\lambda(G) \leq d - \frac{c^2}{4+2c^2}$.*

Therefore, a d regular graph is highly expanding iff its second eigenvalue is far from the first. Combining this fact with some known results about Kazhdan's Property T of group representations, it is possible to give some explicit families of expanders. These are not, however, the best known constructions.

It is known (see [6]) that for any infinite family of d-regular graphs, the limsup of the second largest eigenvalue is at least $2\sqrt{d-1}$. Lubotzky, Phillips and Sarnak [37], and independently, Margulis [42], constructed, for every $d = p + 1$ where p is a prime congruent to 1 modulo 4, explicit infinite families

of d-regular graphs in which the second largest eigenvalue is at most $2\sqrt{d-1}$. Thus, at least in terms of the second eigenvalue, these expanders are best possible. Moreover, in these graphs all the eigenvalues, besides the first, are bounded in absolute value by $2\sqrt{d-1}$. This fact implies certain strong pseudo-random properties, which are useful for some of the applications.

The graphs of [37] and [42] are Cayley graphs of factor groups of the group of all 2 by 2 matrices over a finite field. Their spectral properties are proved by applying results of Eichler and Igusa on the Ramanujan conjectures concerning the number of ways an integer can be represented as a sum of four squares of some special form. Eichler's proof is based on Weil's famous theorem known as the Riemann Hypothesis for curves. More details can be found in [36].

Expanders have numerous applications. They form the basic building blocks of various interconnection and sorting networks, including the sorting network of Ajtai, Komlós and Szemerédi [2] that sorts n elements in $O(\log n)$ parallel steps. They are useful for parallel sorting, merging and selection, and for various variants of the sorting problem, like the "nuts and bolts sorting problem" considered in several papers including [4], [34]. Expanders have recently been used in the construction of Spielman [55] of linear time encodable and decodable error-correcting codes which correct a linear number of errors. Considered as (finite) metric spaces, such graphs cannot be embedded in the Banach spaces ℓ_p with low distortion, as shown by Matoušek [39]. They are also useful in amplification of probabilities, as the random walks on them converge quickly to a uniform distribution. The connection between the expansion properties of graphs and the rate of convergence of random walks on them forms the basis for several algorithms for approximating difficult combinatorial quantities using rapidly mixing Markov chains, developed by Jerrum and Sinclair, see, e.g., [56].

2.2 Derandomization

The tremendous recent development of randomized algorithms, described, among other places, in the comprehensive recent book of Motwani and Raghavan [43], motivated the study of the possibility to convert such algorithms into deterministic ones. Although this is not known in many cases, there are several general techniques that often supply the desired derandomization.

One of the general techniques is the *method of conditional probabilities*. An early instance of this method is implicit in a paper of Erdős and Selfridge [23], but the explicit description of the method is due to Spencer (see, e.g., [54] or [10]), and further developments are due to Raghavan [49]. The basic approach is the following; given a random variable X defined on a finite probability space, the objective is to find deterministically and efficiently a point s of the sample space in which the value of X does not exceed its expectation $E(X)$. To do so, assume the points of the sample space are represented by binary vectors, and try to determine the bits of an appropriate point s one by one, where each bit is chosen in a way that ensures that the conditional expectation of X given the bits chosen so far does not exceed $E(X)$. This process, which can be viewed as a variant of binary search,

is possible only when the required conditional expectations can be computed efficiently. In some cases precise computation is difficult, and one may use estimates that satisfy certain requirements. These estimates, introduced in [49] and called *pessimistic estimators*, are often useful in applications of this method. Several illustrations of the method appear, among other places, in [49], [10], [54].

Another general technique relies on the fact that many randomized algorithms run successfully even when the random choices they utilize are not fully independent. For the analysis some limited amount of independence, like *k-wise independence* for some fixed k, often suffices. In these cases, it is possible to replace the appropriate exponentially large sample spaces required to simulate all random choices of the algorithms by ones of polynomial size. The algorithms can then be converted into deterministic ones, by searching the relatively small sample spaces deterministically.

A simple construction of small sample spaces supporting k-wise independent random variables, appears in [31]. For the case of binary, uniform random variables this is treated under the name *orthogonal arrays* in the Coding Theory literature, see, e.g., [44]. These constructions, as well as some others, are based on some simple properties of polynomials over a finite field or on certain explicit error correcting codes.

Several researchers realized that constructions of this type are useful for derandomizing *parallel* algorithms, since one may simply check all points of the sample space in parallel. The following simple result supplies a lower bound for the size of any sample space supporting n k-wise independent nonconstant random variables.

Proposition 2.2 *Let $\mathcal{S}$ be a sample space supporting n nontrivial k-wise independent random variables. Then, if k is even, $\mathcal{S}$ has at least $\sum_{i=0}^{k/2}\binom{n}{i}$ points, and if k is odd $\mathcal{S}$ has at least $\sum_{i=0}^{(k-1)/2}\binom{n}{i}+\binom{n-1}{(k-1)/2}$ points.*

Note that this implies that for fixed k and large n, the size of $\mathcal{S}$ is $\Omega(n^{\lfloor k/2\rfloor})$. For the binary uniform case this proposition is essentially the Rao bound [48], whereas for the general case it is shown in [3], where it is also observed that this is nearly tight in several cases including the binary uniform one. It follows that polynomial size sample spaces suffice only for handling k-wise independence for fixed k. There are, however, several ways to achieve a higher amount of independence. The most promising way, initiated by Naor and Naor in [45] and improved in [5], constructs sample spaces that support random variables any k of which are *nearly* independent. The constructions here are based on certain error-correcting codes together with some simple properties of the Fourier transform of a distribution on an Abelian group.

The above techniques have been applied in numerous papers dealing with derandomization of parallel as well as sequential algorithms and I make no attempt to include a comprehensive list of references here.

There are several additional derandomization techniques, including ones that rely on crypto-

graphic assumptions to generate pseudo-random sequences and including more specific methods, that are not described here.

2.3 Explicit constructions

There have been many attempts to convert some known probabilistic proofs of existence of combinatorial structures into explicit constructions. To consider these problems systematically, the notion of an *explicit construction* should first be defined precisely. There are several definitions of this notion and the most natural one is probably the existence of an algorithm for constructing the desired structure in time which is polynomial in its size.

Since the early work of Shannon it has been known that randomly chosen codes have powerful error correcting properties. A major part of the work in the theory of error correcting codes is focused on attempts to try and construct explicit codes that are (nearly) as good as random ones. The basic, simply stated problem of determining or estimating the maximum number of vectors of length n over an alphabet of size q so that the Hamming distance between any two vectors is at least d, is still wide open. Let $A_q(n,d)$ denote this maximum. There is, of course, a large number of known upper and lower bounds for $A_q(n,d)$ (cf., e.g., [44]), but even the correct asymptotic behaviour of its logarithm in the binary case is not known. The problem of finding explicit large collections of vectors providing lower bounds for $A_q(n,d)$ is also very difficult, and there are several explicit constructions that rely on some simple properties of polynomials over finite fields as well as on certain deep estimates of character sums. The most exciting explicit constructions are the Algebraic-Geometric codes introduced by Goppa in 1981. Tsfasman, Vladut and Zink proved in [59] that for alphabets that are even powers of primes and exceed 49, these codes yield explicit collections of vectors providing lower bounds for $A_q(n,d)$ which are exponentially better than the best bounds obtained by a random construction (or, equivalently, by the Gilbert-Varshamov bound). Therefore, in coding theory there are interesting cases where explicit constructions beat the best known random ones.

Another example of explicit constructions which are better than the best known random ones is the construction of dense graphs without short cycles- see, e.g., [37]. A more recent example, due to Kollár, Rónyai and Szabó [35], is a construction of dense bipartite graphs that do not contain some fixed complete bipartite subgraph. The properties of these graphs are proved by applying some basic tools from Algebraic Geometry.

The best known problem of finding an explicit construction of a combinatorial structure is probably that of constructing explicit Ramsey graphs. As described in subsection 1.2, it is very simple to prove, by a probabilistic argument, the existence of graphs with at least $2^{k/2}$ vertices which contain neither a clique nor an independent set of size k. Yet, the largest known explicit graphs with this property contain only $2^{\Omega(\log^2 k/\log\log k)}$ vertices. These graphs have been constructed by Frankl and

Wilson [26], using certain results on intersections of finite sets, which are proved by applying some linear algebra techniques.

Another Ramsey-type question mentioned in subsection 1.2 deals with the existence of large triangle-free graphs with no large independent sets. Kim [33] proved by an appropriate random construction that there are triangle-free graphs on n vertices whose largest independent sets are of size $O(\sqrt{n}\sqrt{\log n})$. There is no known explicit construction of such a graph. The best known explicit construction, described in [7], gives explicit triangle-free graphs on n vertices whose largest independent set is of size $O(n^{2/3})$. The properties of these graphs, which are Cayley graphs of Abelian groups, are deduced from their spectral properties, which are proved by applying some estimates on character sums.

Combinatorial examples like the last two, in which random constructions give much better results than explicit ones, seem to be much more frequent than examples in which the constructive approach wins. This could be viewed as a victory of the probabilistic method and a sign for its power in the study of problems in Discrete Mathematics, or as a sign for our lack of imagination and ability to find more constructive solutions. In any case, I am convinced that the study and application of probabilistic arguments, and the related study of pseudo-random structures, will keep playing a crucial role in the development of Combinatorics and Theoretical Computer Science in the future.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, A note on Ramsey numbers, *J. Combinatorial Theory Ser. A* 29 (1980), 354-360.

[2] M. Ajtai, J. Komlós and E. Szemerédi, Sorting in $c \log n$ parallel steps, *Combinatorica* 3 (1983), 1-19.

[3] N. Alon, L. Babai and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem, *J. Alg.* 7 (1986), 567–583.

[4] N. Alon, M. Blum, A. Fiat, S. K. Kannan, M. Naor and R. Ostrovsky, Matching nuts and bolts, *Proc. of the Fifth Annual ACM-SIAM SODA* (1994), ACM Press, 690-696.

[5] N. Alon, O. Goldreich, J. Håstad and R. Peralta, Simple constructions of almost k–wise independent random variables, *Random Structures and Algorithms* 3 (1992), 289–303.

[6] N. Alon, Eigenvalues and expanders, *Combinatorica* 6 (1986), 83-96.

[7] N. Alon, Explicit Ramsey graphs and orthonormal labelings, *The Electronic J. Combinatorics* 1 (1994), R12, 8pp.

[8] N. Alon and V. D. Milman, Eigenvalues, expanders and superconcentrators, *Proc.* 25^{th} *Annual Symp. on Foundations of Computer Science*, Singer Island, Florida, IEEE (1984), 320-322. (Also: λ_1, isoperimetric inequalities for graphs and superconcentrators, *J. Combinatorial Theory, Ser. B* 38 (1985), 73-88.)

[9] N. Alon, B. Mohar and D. P. Sanders, On acyclic colorings of graphs on surfaces, *Israel J. Math.* 94 (1996), 273-283.

[10] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley, New York, 1992.

[11] N. Alon and M. Szegedy, Large sets of nearly orthogonal vectors, to appear.

[12] N. Alon and R. Yuster, The 123 Theorem and its extensions, *J. Combinatorial Theory Ser. A* 72 (1995), 322- 331.

[13] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[14] B. Bollobás, The chromatic number of random graphs, *Combinatorica* 8 (1988), 49-55.

[15] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* 7 (1987), 35-38.

[16] O.V. Borodin, On acyclic colorings of planar graphs, *Discrete Math.* 25 (1979), 211-236.

[17] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson and N. Linial, The influence of variables in product spaces, *Israel J. Math.* 77 (1992), 55-64.

[18] P. Erdős, Some remarks on the theory of graphs, *Bulletin of the Amer. Math. Soc.* 53 (1947), 292-294.

[19] P. Erdős and H. Hanani, On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen*, 10 (1963), 10-13.

[20] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in *Infinite and Finite Sets*, A. Hajnal et. al. eds, North Holland (1975), 609-628.

[21] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959), 290-297.

[22] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5 (1960), 17-61.

[23] P. Erdős and J. L. Selfridge, On a combinatorial game, *J. Combinatorial Theory , Ser. A* 14 (1973), 298-301.

[24] E. Friedgut and G. Kalai, Every monotone graph property has a sharp threshold, Proc. AMS, to appear.

[25] P. Frankl and V. Rödl, Near perfect coverings in graphs and hypergraphs, *Europ. J. Combinatorics* 6 (1985), 317-326.

[26] P. Frankl and R. M. Wilson, Intersection theorems with geometric consequences, Combinatorica 1 (1981), 357-368.

[27] Z. Füredi and R. Stanley, Sets of vectors with many nearly orthogonal pairs (Research Problem), *Graphs and Combinatorics* 8 (1992), 391-394.

[28] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory*, Second Edition, Wiley, New York, 1990.

[29] S. Janson, Poisson approximation for large deviations, *Random Structures and Algorithms* 1 (1990), 221-230.

[30] S. Janson, T. Łuczak and A. Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph in a random graph, in *Random Graphs 87* (M. Karonski et. al. eds.), Wiley (1990), 73-87.

[31] A. Joffe, On a set of almost deterministic k-independent random variables, *Annals of Probability* 2 (1974), 161–162.

[32] J. Kahn, Asymptotically good list-colorings, *J. Combinatorial Theory, Ser. A* 73 (1996), 1–59.

[33] J. H. Kim, The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures and Algorithms* 7 (1995), 173-207.

[34] J. Komlós, Y. Ma and E. Szemerédi, Matching nuts and bolts in $O(n \log n)$ time, *Proc. of the* 7^{th} *Annual ACM-SIAM SODA* (1996), ACM Press, 232–241.

[35] J. Kollár, L. Rónyai and T. Szabó, Norm-graphs and bipartite Turán numbers, *Combinatorica*, to appear.

[36] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures,* Birkhäuser Verlag, 1994.

[37] A. Lubotzky, R. Phillips and P. Sarnak, Explicit expanders and the Ramanujan conjectures, *Proc. of the* 18^{th} *ACM Symp. on the Theory of Computing*, (1986), 240-246; (Also: Ramanujan graphs, *Combinatorica* 8 (1988), 261-277).

[38] T. Łuczak, A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica* 11 (1991), 295-297.

[39] J. Matoušek, On embedding expanders into ℓ_p spaces, to appear.

[40] D. W. Matula, On the complete subgraph of a random graph, *Combinatory Mathematics and its Applications*, Chapel Hill, North Carolina (1970), 356-369.

[41] G. A. Margulis, Probabilistic characteristics of graphs with large connectivity, *Prob. Peredachi Inform.* 10 (1974), 101-108.

[42] G. A. Margulis, Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators, *Problemy Peredachi Informatsii*, 24 (1988), 51-60 (in Russian). (English translation in *Problems of Information Transmission*, 24 (1988), 39-46).

[43] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, New York, 1995.

[44] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland, Amsterdam, 1977.

[45] J. Naor and M. Naor, Small–bias probability spaces: efficient constructions and applications, *SIAM J. Comput.* 22 (1993), 838–856.

[46] M. Pinsker, On the complexity of a concentrator, 7^{th} *Internat. Teletraffic Conf.*, (1973), Stockholm, 318/1-318/4.

[47] N. Pippenger and J. H. Spencer, Asymptotic behaviour of the chromatic index for hypergraphs, *J. Combinatorial Theory, Ser. A* 51 (1989), 24-42.

[48] C. R. Rao, Factorial experiments derivable from combinatorial arrangements of arrays, *J. Royal Stat. Soc.* 9 (1947), 128–139.

[49] P. Raghavan, Probabilistic construction of deterministic algorithms: approximating packing integer programs, *J. Comput. Syst. Sci.*, 37 (1988), 130–143.

[50] G. Ringel and J. W. T. Youngs, Solution of the Heawood map coloring problem, *Proc. Nat. Acad. Sci. U.S.A.* 60 (1968), 438-445.

[51] V. Rödl, On a packing and covering problem, *European Journal of Combinatorics* 5 (1985), 69-78.

[52] L. Russo, On the critical percolation probabilities, *Z. Wahrsch. werw. Gebiete* 43 (1978), 39-48.

[53] E. Shamir and J. H. Spencer, Sharp concentration of the chromatic number on random graphs $G_{n,p}$, Combinatorica 7 (1987), 124-129.

[54] J. H. Spencer, *Ten lectures on the Probabilistic Method*, Second Edition, SIAM, Philadelphia, 1994.

[55] D. Spielman, Linear-Time Encodable and Decodable Error-Correcting Codes, *Proc. of the* 27^{th} *ACM Symp. on the Theory of Computing* 1995, ACM Press, 388-397.

[56] A. Sinclair and M. R. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, *Information and Computation* 82 (1989), 93-133.

[57] R. M. Tanner, Explicit construction of concentrators from generalized N-gons, *SIAM J. Alg. Disc. Meth.* 5 (1984), 287-293.

[58] M. Talagrand, A new isoperimetric inequality for product measure and the tails of sums of independent random variables, *Geometric and Functional Analysis* 1 (1991), 211-223.

[59] M. A. Tsfasman, S. G. Vladut and T. Zink, Modular curves, Shimura curves and Goppa codes, better than the Varshamov-Gilbert bound, *Math. Nachr.* 104 (1982), 13-28.