Sparse sets

Alexander Shen, Laurent Bienvenu, Andrei Romashchenko

Journées Automates Cellulaires 2008 (Uzès), pp. 18-28

SPARSE SETS

LAURENT BIENVENU [1], ANDREI ROMASHCHENKO [2], AND ALEXANDER SHEN [1]

[1] LIF, Aix-Marseille Université, CNRS, 39 rue Joliot Curie, 13013 Marseille, France
E-mail address: Alexander.Shen@lif.univ-mrs.fr
URL: http://www.lif.univ-mrs.fr/

[2] LIP, ENS Lyon, CNRS, 46 allée d'Italie, 69007 Lyon, France
E-mail address: Andrei.Romashchenko@ens-lyon.fr
URL: http://www.ens-lyon.fr/

Abstract. For a given $p > 0$ we consider sequences that are random with respect to p-Bernoulli distribution and sequences that can be obtained from them by replacing ones by zeros. We call these sequences *sparse* and study their properties. They can be used in the robustness analysis for tilings or computations and in percolation theory.

This talk is a report on an (unfinished) work and is based on the discussions with many people in Lyon and Marseille (B. Durand, P. Gacs, D. Regnault, G. Richard a.o.) and Moscow (at the Kolmogorov seminar: A. Minasyan, M. Raskin, A. Rumyantsev, N. Vereshchagin, D. Hirschfeldt a.o.).

1. Motivation

There are several results which say, intuitively speaking, that "if errors are sparse enough, they do not destroy the intended behavior of the system". For example, percolation theory says that if every edge of a planar conducting grid is cut with probability ε, cuts of different edges are independent and ε is small enough, then the grid remains mostly connected (there is an infinite connected component). Similar statements can be done for cellular automata computations with independent random errors, for tiling with errors etc.

It would be nice to translate these results into the language of algorithmic information theory and define the notion of "individual sparse set" (for a given ε). This notion could be used to split the above mentioned results into two parts: first, we note that with probability 1 with respect to the Bernoulli distribution the resulting set and any its subset is sparse; second, we prove that a sparse set of errors does not destroy the desired behavior (connectivity of the grid, computation of the error-correcting automaton etc.)

2000 ACM Subject Classification: 68Q30, 82B43.

Key words and phrases: algorithmic randomness, percolation theory, error correction.

Thanks to all members and guests of the ESCAPE team at LIF who contributed to the discussions reported in this paper.

Note that the notion of sparse set should take into account not only the overall density of errors but also their distribution. For example, if we cut all the conductors along some line in a grid, we may destroy the connectivity though the fraction of destroyed conductors is negligible (density in a $N \times N$ square tends to 0 as $N \to \infty$).

In this paper we suggest such a definition of p-sparse set and outline some its properties as well as possible applications.

2. Definition of sparse sets

Let $p > 0$ be some rational number. Consider a Bernoulli distribution B_p on the space Ω of all infinite sequences of zeros and ones where bits are independent and each bit equals 1 with probability p. This is a computable distribution on Ω, so the notion of Martin-Löf random sequence with respect to this distribution is well defined (see, e.g., [2, 6]).

Identifying each sequence ω with a subset of $\mathbb{N}$ that consists of all i such that $\omega_i = 1$, we get a notion of a random subset of $\mathbb{N}$ according to this distribution. Let us call *p-sparse* all these sets and all their subsets. The following proposition justifies this definition:

Theorem 2.1. *Let $0 < p_1 < p_2 < 1$ be two rational numbers. Then every p_1-sparse set is p_2-sparse.*

Proof. We need to prove that any p_1-random sequence α can be converted to a p_2-random sequence by replacing some zeros by ones. Informally, we want to replace each 0 by 1 with probability $q = (p_2 - p_1)/(1 - p_1)$; if this replacement is directed by a q-random (with respect to α-oracle) sequence β then the result will be p_2-random. To prove this, we may use van Lambalgen theorem; it says that the pair (α, β) is random with respect to the product distribution, and the replacement result is therefore random with respect to the image distribution, i.e., p_2-Bernoulli distribution. (We do not go into the details since we prove more general result about coupling below.) ∎

Remark. We have defined the notion of a p-sparse set for subsets of $\mathbb{N}$, but the Bernoulli distribution is invariant under permutations and Bernoulli-random sequences remain Bernoulli-random after computable permutations of the domain. Therefore this notion can be defined for subsets of $\mathbb{Z}$, $\mathbb{Z}^k$, sets of strings etc.

By definition, every p-Bernoulli random sequence α is p-sparse. The next theorem shows that there could be another reason of α to be p-sparse:

Theorem 2.2. *For every computable $p \in (0, 1)$ and every p-Bernoulli random sequence α one could replace infinitely many zeros in α by ones but still have p-Bernoulli random sequence (for the same p).*

Proof. (Discovered independently by Peter Gacs.) Consider a diverging series with converging squares, say, $q_i = 1/i$. Then consider a sequence β that is random with respect to the distribution of independent bits where ith bit is 1 with probability q_i. Moreover, let us assume that β is random even relative to the oracle α. Then we can check that $\alpha_i \vee \beta_i$ will contain infinitely many additional ones compared to α (since $\sum q_i$ diverges and β is independent of α). On the other hand, $\alpha_i \vee \beta_i$ is still random with respect to p-Bernoulli distribution since its natural distribution $(p + (1 - p)q_i)$ is equivalent to p-Bernoulli distribution due to the effective version of Kakutani's theorem (because the sum of *squares* of the differences between probabilities of the two Bernoulli measures *converges*). ∎

3. Criteria

One would like to have a more straightforward definition for p-sparse sets that does not involves the existential quantifier ("there exists a random sequence such that..."). Such a criterion indeed can be obtained if we restrict ourselves to monotone cylinders in the Martin-Löf definition of randomness.

Let X be a finite subset of $\mathbb{N}$. Consider the set Γ_X of all $\omega \in \Omega$ that have ones at all positions in X. The p-Bernoulli measure of Γ_X equals p^k where k is the cardinality of X.

Let us call a set $N \subset \Omega$ *effectively p-monotone null* set if there exists an algorithm that given rational $\varepsilon > 0$ enumerates a sequence $X_1, X_2, \ldots$ of finite subsets of $\mathbb{N}$ such that the union of corresponding monotone cylinders Γ_{X_i} has measure at most ε (with respect to p-Bernoulli distribution) and covers N.

Theorem 3.1.

A. *For every rational $p > 0$ there exists the maximal effectively p-monotone null set that contains all effectively p-monotone null sets.*

B. *A sequence ω is p-sparse if and only if ω does not belong to the maximal effectively p-monotone null set.*

Proof. **A**. This can be proved in the same way as Martin-Löf theorem about the existence of the maximal effectively null set: one can enumerate all algorithms, modify them to guarantee the bound for the measure and then take the union of all corresponding sets.

B. (See [3].) Let us note that every effectively p-monotone null set N is an effectively null set by definition; moreover, all the sequences obtained from the elements of N by replacing zeros with ones still form an effectively null set. Therefore, the elements of N are not p-sparse.

On the other hand, we have to prove that the set U of all sequences that cannot be made p-random by replacing zeros by ones is an effectively p-monotone null set. By definition, a sequence ω belongs to U if the set E_ω of all sequences that can be obtained from ω by such a replacement is a subset of the maximal effectively null set M (with respect to p-Bernoulli measure). We need (for a given $\varepsilon > 0$) to cover U by the sequence of monotone cylinders of total measure at most ε. Consider the enumerable union of (non-monotone) cylinders that covers M and has total measure at most ε. This union covers the closed (and therefore compact) set E_ω, therefore a finite union of these cylinders also covers E_ω. Those cylinders deal with finitely many positions in E_ω, therefore there exists a monotone cylinder that contains ω and is covered by that union. The set of all monotone cylinders covered by some finite union of (non-monotone) cylinders in the sequence is enumerable, and we get the required enumerable family of monotone cylinders of total measure at most ε. ∎

Remark. This criteria can be used to prove Theorem 2.1: it remains to show that for a monotone set X its p-Bernoulli measure $B_p(X)$ is a monotone function of p. (This is a standard result in classical probability theory that is proved essentially in the same way, by coupling p_1- and p_2-Bernoulli random variables, see below Theorem 4.4.)

It would be interesting to find another equivalent definition of p-sparse sets. One may look for a martingale criterion of sparseness. Recall [4] that a sequence is random if and only if every lower semicomputable martingale is bounded on its prefixes. Trying to characterize sparse sequences, one may try to find the corresponding class of martingales.

Formally speaking, a *p-martingale* (in the algorithmic information theory) is defined as a function $x \mapsto m(x)$ defined on binary strings with non-negative real values such that

$$m(x) = pm(x1) + (1-p)m(x0).$$

If $m(x)$ is interpreted as the capital of the player after bits of x appear during the game (from left to right), this equation says that the game is fair, i.e., the expected capital after the next round equals the capital before it. We say that martingale m *makes bets only on ones* if the outcome 1 is always more profitable than 0, i.e., if

$$m(x0) \leq m(x1)$$

for every binary string x. A stronger condition: a martingale m is *monotone* if $m(x) \leq m(y)$ when $x \prec y$, i.e., when y can be obtained from x by replacing some zeros by ones. It turns out that monotone martingales can be used for defining sparse sets (sequences) while martingales that bet only on 1's are not suitable (there are too many of them).

Recall that a martingale m is *lower semicomputable* if there exists a computable function $(x, n) \mapsto M(x, n)$ with rational values (here x is a string and n is a natural number) such that for every x the sequence $M(x, 0), M(x, 1), \ldots$ is non-decreasing and converges to $m(x)$.

Theorem 3.2.

A. *If a sequence α is not p-sparse, there exists a monotone lower semicomputable martingale m that tends to infinity on the prefixes of α.*

B. *If there exists a monotone lower semicomputable martingale m that is unbounded on the prefixes of α, then α is not p-sparse.*

C. *There exists a p-sparse sequence α and a lower semicomputable martingale m that makes bets only on ones and still tends to infinity on the prefixes of α.*

Proof. **A**. For every enumerable union T of monotone cylinders we may consider a lower semicomputable martingale m_T by letting $m_T(x)$ be the conditional probability to get a sequence in T starting from x (and adding independent p-Bernoulli bits). Since T is monotone, this martingale is also monotone; it starts from the measure of T and reaches 1 at any sequence in T. Adding up these martingales (say, take T_n of measure at most 4^{-n} and multiply the corresponding martingale by 2^n), we get a lower semicomputable martingale that tends to infinity on the prefixes of all elements of given monotone effectively null set.

B. Let m be any lower semicomputable monotone martingale. For a given c we may consider the enumerable set of all x such that $m(x) > c$, and all (non-monotone) cylinders rooted at these x. The union of these cylinder has measure at most $1/c$ due to martingale inequality. Since the martingale is monotone, we can replace non-monotone cylinders by the monotone ones not changing the union. Doing this for large c, we get the required covering by an enumerable union of monotone cylinders that has small measure.

C. To construct a required counterexample, let us consider a p-Bernoulli random sequence and split its elements into pairs. Then let us modify this sequence replacing pairs 01 and 10 by 00 (and leaving 00 and 11 unchanged). This gives us a p-sparse sequence since we replace ones by zeros in a random sequence. However, the player can make a safe bet on the second bit of each pair if she sees that the first bit is 1, and this happens infinitely many times since we have started from a random sequence. ∎

4. Coupling

In this section we review a well known technique that can be then adapted to prove that some operation produce sparse sets.

Let A and B be two finite sets and let $R \subset A \times B$ be a binary relation; we write aRb if $\langle a, b\rangle \in R$. Consider two random variables α and β that range over A and B. defined on unrelated probability spaces. We say that $\alpha R \beta$ (abusing slightly the notation) if there exist random variables α' and β' defined on some common probability space M such that

- α' has the same distribution as α;
- β' has the same distribution as β;
- $\alpha'(m) R \beta'(m)$ for every $m \in M$.

This definition refers not to α and β themselves, but only to the corresponding probability distributions on A and B, so it defines a relation between probability distributions on A and B.

In other terms, we are looking for a matrix of non-negative reals (a distribution on $A \times B$) that has given sums for all rows and columns. This task can be reformulated in terms of network flows. Indeed, consider a bipartite graph that has A on the left and B on the right. The source s is connected with all elements of A by edges whose capacities are given probabilities of the elements in A; all elements of B are connected to the sink t by edges whose capacities are probabilities of the elements of B. The edges between A and B have unlimited capacities and correspond to the elements of R. Then $\alpha R \beta$ means that this network has a flow of size 1. The Ford–Fulkerson theorem provides a criterion for the existence of such a flow; this criterion describes the obstacles for $\alpha R \beta$. A pair of sets $S \subset A$ and $T \subset B$ is an *obstacle* if

- all R-neighbors of all elements in S belong to T
- $\Pr[\alpha \in S] > \Pr[\beta \in T]$

It is evident that if such an obstacle exists, then $\alpha R \beta$ is not possible, and Ford–Fulkerson duality says that this condition is necessary and sufficient:

Theorem 4.1. *$\alpha R \beta$ if and only if there are no obstacles of described type.*

Now we extend this result to infinite sequences. Consider two random variables α and β that range over A^∞ and B^∞ (here X^∞ stands for the set of all sequences $x_0, x_1, \ldots$ where $x_i \in X$). We say that $\alpha R \beta$ if there exist α' and β' that share the same probability space M, have the same distribution as α and β, and $\alpha'_i(m) R \beta_i(m)$ for every $m \in M$ and for every $i \in \mathbb{N}$.

Note that this definition refers only to the distributions on A^∞ and B^∞ and that the relation on $A^\infty \times B^\infty$ is defined coordinate-wise (separately for each i). It turns out that the statement about possible obstacles remains true for this definition.

Theorem 4.2. *The relation $\alpha R \beta$ is false if and only if there exist Borel sets $S \subset A^\infty$ and $T \subset B^\infty$ such that:*

- *for every $a_0 a_1 \ldots \in S$ every $b_0 b_1 \ldots \in B^\infty$ such that $\forall i\, (a_i R b_i)$ belongs to T;*
- $\Pr[\alpha \in S] > \Pr[\beta \in T]$

Proof. It is evident that the existence of S and T with these properties is indeed an obstacle for $\alpha R \beta$. In the other direction we cannot use just the Ford–Fulkerson argument since the spaces are infinite. However, the relation is defined coordinate-wise, and we may for every N find a joint distribution on $A^N \times B^N$ that has the same projections on A^N and B^N as α

and β, and has the required property with respect to R. (Indeed, an obstacle to this finite task as described in Theorem 4.1 at the same time is an obstacle in our current sense.) It remains to find limit point as $N \to \infty$ using the standard compactness argument. ∎

Note that the existential quantifier for one of the sets S and T can be eliminated: the obstacle can be defined as the set S such that the set of all its neighbors have β-probability less than the α-probability of T. This evident remark shows that defined relation is transitive:

Theorem 4.3. *Let A, B, C be finite sets; let $R_1 \subset A \times B$ and $R_2 \subset B \times C$ be two binary relations and $R = R_1 \circ R_2$ be its composition. Then $\alpha R_1 \beta \wedge \beta R_2 \gamma \Rightarrow \alpha R \gamma$ for any random variables α, β, γ that range over $A^\infty, B^\infty, C^\infty$ respectively.*

We will mostly deal with the special case where $A = B = \{0,1\}$ and the relation R is linear order $\leq$. Let us denote the corresponding relation between two random variables α and β that range over $\Omega = \{0,1\}^\infty$ by $\alpha \preceq \beta$. (The same notation will be used for corresponding distributions.) In this case the description of obstacle can be simplified further:

Theorem 4.4. *$\alpha \preceq \beta$ if and only if $\Pr[\alpha \in Z] \leq \Pr[\beta \in Z]$ for every monotone Borel set $Z \subset \Omega$.*

(A set $Z \subset \Omega$ is monotone if $\alpha \in Z$ and $\forall i\, (\alpha_i \leq \beta_i)$ implies $\beta \in Z$.)

Proof. One could argue that the set $T(S)$ of neighbors (as defined above) for every set $S \subset \Omega$ is monotone and contains S, so S can be replaced by $T(S)$. (Note that $T(T(S)) = T(S)$.)

However, there is a technical problem since we need T to be a Borel set. This can be avoided if we use this argument for finite case $\{0,1\}^N$ and take a limit point after that. ∎

Remark. Similar statements can be made for any finite set and partial preordering on it (instead of the standard ordering of $\{0,1\}$).

5. Coupling and algorithmic randomness

We want to apply coupling technique to establish results about random and sparse sets. The following two statements will be used as main technical tools for this.

Let $\mathcal{M}$ be an oracle machine that accepts or rejects its input (natural number) n asking questions of type "$m \in L$?" for different natural numbers m and oracle L. Such a machine defines a mapping that maps the oracle set L into the set accepted by $\mathcal{M}$ when oracle L is used.

More precisely, this is a partial map, since it may happen that for some oracles L the machine $\mathcal{M}^L$ does not terminate for some inputs. We are not interested in the partial functions and consider $\mathcal{M}(L)$ as undefined in these cases. Recall also that we identify sets $X \subset \mathbb{N}$ and their characteristic sequences, so an oracle machine defines a (partial) mapping $\Omega \to \Omega$.

Let P be a computable probability distribution on Ω and let $\mathcal{M}$ be an oracle machine. For a random variable α that ranges over Ω we consider the image distribution $\mathcal{M}(P)$ on Ω, i.e., the distribution for the variable $\mathcal{M}(\alpha)$. This distribution is well defined if $\mathcal{M}(\omega)$ is defined with probability 1. We assume that this is the case; then $Q = \mathcal{M}(P)$ is a computable distribution on Ω. The following natural connection between P-randomness and Q-randomness [5] holds:

Theorem 5.1.

A. *If ω is Martin-Löf P-random sequence, then $\mathcal{M}(\omega)$ is Martin-Löf Q-random sequence.*

B. *Any Martin-Löf Q-random sequence τ equals $\mathcal{M}(\omega)$ for some Martin-Löf P-random sequence ω.*

Proof. **A**. If ω is Martin-Löf P-random and $\mathcal{M}(\omega)$ is defined, this is a direct corollary of the definitions: if Z is a Q-effectively null set containing $\mathcal{M}(\omega)$ then its preimage $\mathcal{M}^{-1}(Z)$ is a P-effectively null set containing ω (it is an effectively null set since the preimage of an effectively open set of Q-measure less than ε is an effectively open set of the same P-measure).

However, there is one more thing to prove: we need to show that $\mathcal{M}(\omega)$ is defined on every P-random ω. Indeed, for every n the set of all oracles ω such that $\mathcal{M}^\omega$ is defined on n is an effectively open set of full measure. It is easy to see that its complement is an effectively null set and therefore cannot contain a Martin-Löf random sequence.

B. Let N be the maximal P-effectively null subset of Ω. We need to prove that every Q-random sequence has a $\mathcal{M}$-preimage outside N. In other terms, we have to prove that the set N' of sequences that do not have preimages outside N is a Q-effectively null set.

Assume that some rational $\varepsilon > 0$ is given. We have to find an effectively open set of Q-measure less than ε that covers N'. First, let us consider an effectively open set N_ε that covers N and has P-measure less than ε. We want to show that the set of sequences that have no preimages outside N_ε is an effectively open set. (Note that Q-measure of this set is less than ε since Q is the image of P.)

Indeed, consider a sequence ω that has no preimages outside N_ε. As we have seen, $\mathcal{M}$ is defined everywhere outside N (and therefore outside N_ε). So for every $\tau \notin N_\varepsilon$ the infinite sequence $\mathcal{M}(\tau)$ deviates from ω at some place. The set of all τ for which this happens at ith place is open, and these sets together with the open set N_ε form a covering of Ω. Compactness allows to replace this covering by its finite part, and this gives some neighborhood of ω; all elements of this neighborhood have no preimages outside N_ε. Moreover, we can search for all finite coverings of this type, so the set of all sequences that have no preimage outside N_ε is effectively open (and has Q-measure less than ε as we have discussed).

This finishes the argument (which is an effective version of a classical argument proving that if a continuous function is defined everywhere on a compact set, then the image of this set is also compact). ∎

This statement has a simple intuitive meaning: if we have a source of random bits composed of a P-random bit source and an oracle machine $\mathcal{M}$ using those bits as an oracle, what sequences should we expect at the output? There are two possible answers. First, we can say that the internal P-random source can produce any P-random sequence, so we can get images of those sequences at the output. On the other hand, we can ignore the internal structure and say that we altogether have a random bits generator with distribution Q, so Q-random sequences are expected at its output. Theorem 5.1 says that these two classes coincide.

Remark. This question is more difficult (and is not solved yet, as far as we know) if the machine $\mathcal{M}$ is undefined with positive probability. Then we get only a semimeasure as output distribution; one would like to prove that the image of the set of P-random sequences

is still determined by this semimeasure, but it is not clear how to prove this and whether it is possible.

The second tool connects coupling with algorithmic randomness.

Theorem 5.2. *Let P and Q be two computable distributions such that $P \preceq Q$ and, moreover, there exists a computable distribution on $\Omega \times \Omega$ that has projections P and Q and $\alpha_i \le \beta_i$ has probability one according to this distribution for every i.*

Then every P-random sequence can be transformed into a Q-random sequence by replacing some zeros by ones. Similarly, every Q-random sequence can be transformed into a P-random one be replacing some ones by zeros.

Proof. It would be nice to get rid of the additional condition: the computability of the distribution on pairs. But with this condition (that is easy to check in all applications below) the statement of the theorem is a direct consequence of Theorem 5.1. Using projection as $\mathcal{M}$, we see that every P-random sequence is a first term of a random pair, and the second term of this pair is a Q-random sequence we looked for. (The same argument can be used for choosing a P-random sequence for a given Q-random one.) ∎

Note that in this way we get one more proof of Theorem 2.1. More interesting applications will be given in the next section.

6. Operations that preserve sparseness

In this section we want to prove that some transformations preserve sparseness (though may change the value of parameter p). Let us start with a simple example.

Theorem 6.1. *Let $a_0, a_1, a_2, \dots$ be a p-sparse sequence. Then the sequence*

$$a_0, a_0, a_1, a_1, a_2, a_2, \dots$$

(each bit is doubled) is $\sqrt{p}$-sparse.

Proof. Since bit doubling is a monotone transformation, we may assume that $a_0, a_1, \dots$ is a random sequence with respect to the Bernoulli distribution B_p. Theorem 5.1 says that the sequence $a_0, a_0, a_1, a_1, a_2, a_2, \dots$ is then random with respect to the image distribution $D(B_p)$ where D doubles each bit.

Theorem 5.2 shows that it remains to prove that $D(B_p) \preceq B_{\sqrt{p}}$. Since both distributions are products of (independent) distributions on bit pairs, it is enough to consider these distributions on pairs. They have the same probability of 11 combination (p in both cases) while 01 and 10 have zero probability in $D(B_p)$ and positive probability $\sqrt{p}(1-\sqrt{p})$ for $B_{\sqrt{p}}$. So we can start with $B_{\sqrt{p}}$ distribution and then replace 1 by 0 if the other bit is 0. ∎

More complicated tools are needed in the other example.

Theorem 6.2. *Let $a_0, a_1, a_2, \dots$ be a p-sparse sequence. Then the sequence*

$$a_0 \vee a_1, a_1 \vee a_2, a_2 \vee a_3, \dots$$

is $2\sqrt{p}$-sparse.

Note that this sequence is an upper bound for $a_1, a_1, a_3, a_3, a_5, a_5, \dots$, so this result implies that the latter sequence is also sparse (though for a larger value of p compared with Theorem 6.1).

Proof. The sequence in question can be represented as bitwise OR of two sequences:

$$\begin{array}{cccccccccc} a_0 & a_2 & a_2 & a_4 & a_4 & a_6 & a_6 & a_8 & \dots \\ a_1 & a_1 & a_3 & a_3 & a_5 & a_5 & a_7 & a_7 & \dots \\ a_0{\vee}a_1 & a_1{\vee}a_2 & a_2{\vee}a_3 & a_3{\vee}a_4 & a_4{\vee}a_5 & a_5{\vee}a_6 & a_6{\vee}a_7 & a_7{\vee}a_8 & \dots \end{array}$$

If $a_0, a_1, a_2, \dots$ is a random sequence with Bernoulli distribution, then the first two sequences are independent and their distributions (as we have seen above) are $\preceq$-below $B_{\sqrt{p}}$. It follows easily that the disjuction of these two sequences is $\preceq$-below the disjuction of two Bernoulli distributions $B_{\sqrt{p}}$ and therefore is below $B_{2\sqrt{p}}$-distribution. It remains to apply Theorem 5.2 (the computability condition is easy to check). ∎

Similar argument can be applied to the sequence

$$a_0 \vee a_1 \vee \dots \vee a_{k-1}, a_1 \vee a_2 \vee \dots \vee a_k, \dots$$

for any k and shows that it is $k\sqrt[k]{p}$-sparse if the initial sequence is p-sparse. In terms of sets we get the following result (note that d-neighborhood of a point in $\mathbb{N}$ or $\mathbb{Z}$ consists of $2d+1$ points):

Theorem 6.3. *For every positive integer $d > 0$ and every p-sparse set A the d-neighborhood of A is q-sparse for $q = (2d+1)\sqrt[2d+1]{p}$.*

For simplicity we may ignore the exact value of the probability and say something like "the neighborhood of a sparse set is sparse". The exact meaning of this claim is that for every $q > 0$ there exists $p > 0$ such that the neighborhood of every p-sparse set is q-sparse. This is true for subsets of $\mathbb{Z}^2$ or $\mathbb{Z}^k$ (the proof works for L_∞-neighborhoods as before; then the statement can be extended to any type of neighborhood).

One would like to strengthen this theorem and prove that the union of two p-sparse sets is q-sparse if p is much less than q. Unfortunately, this cannot be done:

Theorem 6.4. *For every $p, q \in (0,1)$ there exist two p-sparse sets whose union is not q-sparse.*

Proof. (Discovered independently by Denis Hirschfeldt.) Note that for every $q \in (0,1)$ (even very close to 1) and every N (even very large) every q-sparse sequence α splitted into N-bit blocks must contain a block of N zeros.

On the other hand, for every p for large enough N there are two p-sparse sequences α and β whose union $\alpha \vee \beta$ does not have this property. If $p = 1/2$, it is trivial: take $N = 1$ and two complementary 1/2-random sequences.

In the general case, take N large enough so that $(1-p)^N < 1/2$. Then two events (for a fixed N-bit block) "all N α-bits are zeros" and "all N β-bits are zeros" can be made disjoint keeping the p-Bernoulli distributions for α and β in the block unchanged. Making all blocks independent, we get a computable distribution on pair of sequences that has p-Bernoulli projections (marginal distributions) and is concentrated on pairs of sequences that have the required property. ∎

Another source of sparse sequences is provided by the following theorem:

Theorem 6.5. *Let P be a computable distribution on Ω and let $\alpha = a_0a_1a_2\dots$ be a random sequence with respect to this distribution. If all the conditional probabilites of ones along this path (i.e., $\Pr[x_i = 1 | x_0 \dots x_{i-1} = a_0 \dots a_{i-1}]$) do not exceed some rational p, then α is p-sparse.*

Proof. Consider the following randomized machine. It uses uniform Bernoulli distribution to generate a sequence of independent random variables $\xi_0, \xi_1, \xi_2, \ldots$ distributed uniformly in $[0, 1]$ (by splitting random bits into countably many groups). Then these random variables are used to produce bits that have distribution P: we compare ξ_0 with the probability of 1 at the first place to get a_0, then we compare ξ_1 with the (conditional) probability of 1 after a_0 to get a_1, etc.

Theorem 5.1 guarantees that in this way we can obtain exactly P-random sequences. Replacing all thresholds by p, we in parallel get p-Bernoulli random sequence that guarantees that the sequence $a_0a_1\ldots$ is p-sparse.

Note a subtle point in this argument: we do not claim that the $P \preceq B_p$ since the inequality for conditional probabilities is guaranteed only along the path $a_0a_1a_2\ldots$; however, the sequence generated in parallel with $a_0a_1a_2\ldots$ is p-random since the image of every random sequence $\xi_0\xi_1\ldots$ is p-random. ∎

7. Using sparse sets in error analysis

Now several results about robustness may be reformulated in terms of sparse sets. Let us give an example from percolation theory.

Consider a grid of vertical and horizontal lines (as in the cell paper) that splits the plane into unit squares. Each node (line crossing) is a contact, and each edge (of a unit square) is a conductor.

Percolation theory says that if each conductor is independently cut with sufficiently small probability, the network remains essentially connected (has an infinite connected component) with probability 1. Now this can be reformulated in terms of sparse sets:

Theorem 7.1. *If the set of deleted edges is p-sparse for small enough p, the remaining network has an infinite connected component.*

Proof. It is enough to show this for the case of p-random network failures (since the property of having an infinite connected component is monotone).

Therefore, we need to show that the set of measure zero (of all networks that have no infinite connected components or have more than one connected component) is in fact an effectively null set. This can be done by analyzing one of the classical proofs of this result (we omit the details since this is another story). ∎

Now we can apply our results about sparse sets to derive the similar result for node failures (instead of edge failures): failure of a node is equivalent to cutting all four edges that are adjacent to this node.

Theorem 7.2. *If the set of deleted nodes is p-sparse for small enough p, the remaining network has an infinite connected component.*

Proof. It is easy to see that the set of edges adjacent to deleted nodes can be covered by a neighborhood of fixed size of the set of deleted nodes and therefore is also sparse. (Technically, we have twice more edges than nodes, so some technical changes are needed.) ∎

The notion of sparse set can be used in the analysis of tilings that are robust to errors: the results of [1] can be expressed in terms of sparse error sets.

8. Questions

There are some open (as far as we know) questions related to the topic of this paper.

1. Is it possible to replace in the criterion of sparseness the measure of the union of the monotone cylinders by the sum of their measures?

2. One would expect there exists some criterion of p-sparseness in terms of complexity or randomness deficiency. How to define "sparseness deficiency"? One possibility is to take minimum of randomness deficiency (with respect to p-Bernoulli distribution) over all strings obtained from a given one by $0 \to 1$ replacement. Another natural option is to consider lower semicomputable deficiency functions $d_n(x)$ defined on n-bit strings (uniformly in n) that are monotone (i.e., replacement $0 \to 1$ may only increase the deficiency) such that $2^{d_n(x)}$ has average at most 1 (over all n-bit strings).

Are these two definitions connected? close to each other? Can any of them be used to characterize p-sparse sets?

3. Can one get rid of the computability condition in the statement of Theorem 5.2?

4. What can be said about two computable measures P and Q if every P-random sequence can be made Q-random by replacing some zeros with ones? We cannot expect $P \preceq Q$ since this may be not the case even for equivalent measures P and Q (that have the same set of random sequences), but can we claim something weaker in this direction (e.g., P is equivalent to $P' \preceq Q$ or something like this)?

5. What can be said about sets that are p-sparse for every $p > 0$? Can we eliminate the universal quantifier ("for every p") in the definition?

References

[1] Durand B., Romashchenko A., Shen A., *Fixed point and aperiodic tilings*, arXiv:0802.2432v2 (19 Feb. 2008)

[2] Li M., Vitányi P., *An Introduction to Kolmogorov Complexity and Its Applications*, Second Edition, Springer, 1997. (638 pp.)

[3] A. Minasyan, manuscript, 2007 (the course project at the Moscow State University).

[4] Schnorr C.P., *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*, Lecture notes in mathematics, v. 218. IV+212 S. Springer, 1971.

[5] Shen A., One more definition of random sequence with respect to computable measure. *First World Congress of the Bernoulli Society on Math. Statictics and Probability theory*, Tashkent, USSR, 1986

[6] Uspensky V.A., Semenov A.L., Shen A., Can a single sequence of zeros and ones be random? *Uspekhi matem. nauk*, 1990, 45, 1, p. 105–162. (Russian; for English translation see *Russian Math. Surveys*, 45:1 (1990), 121–189.; MR 91f:03043.)