Theory Comput Syst
DOI 10.1007/s00224-009-9203-9

Limit Complexities Revisited

Laurent Bienvenu · Andrej Muchnik · Alexander Shen · Nikolay Vereshchagin

Abstract The main goal of this article is to put some known results in a common perspective and to simplify their proofs.

We start with a simple proof of a result from Vereshchagin (Theor. Comput. Sci. 271(1–2):59–67, 2002) saying that $\limsup_n C(x|n)$ (here $C(x|n)$ is conditional (plain) Kolmogorov complexity of x when n is known) equals $C^{\mathbf{0}'}(x)$, the plain Kolmogorov complexity with $\mathbf{0}'$-oracle.

Then we use the same argument to prove similar results for prefix complexity, a priori probability on binary tree and measure of effectively open sets, and also to improve results of Muchnik (Theory Probab. Appl. 32:513–514, 1987) about limit frequencies. As a by-product, we get a criterion of $\mathbf{0}'$ Martin-Löf randomness (called

The research of L. Bienvenu and A. Shen was supported in part by ANR Sycomore and NAFIT ANR-08-EMER-008-01 grants.
The research of N. Vereshchagin was supported in part by RFBR 05-01-02803-CNRS-a, 06-01-00122-a.

Andrej Muchnik (24.02.1958–18.03.2007) worked in the Institute of New Technologies in Education (Moscow). For many years he participated in Kolmogorov seminar at the Moscow State (Lomonosov) University. N. Vereshchagin and A. Shen (also participants of that seminar) had the privilege to know Andrej for more than two decades and are deeply indebted to him both as a great thinker and noble personality. The text of this paper was written after Andrej's untimely death but it (like many other papers written by the participants of the seminar) develops his ideas.

L. Bienvenu · A. Shen (✉)
Laboratoire d'Informatique Fondamentale, CNRS & Université Aix-Marseille, Marseille, France
e-mail: Alexander.Shen@lif.univ-mrs.fr

N. Vereshchagin
Moscow State Lomonosov University, Moscow, Russia

A. Shen · N. Vereshchagin
CNRS Poncelet Laboratory, Moscow, Russia

A. Shen
IITP RAS, Moscow, Russia

Published online: 17 March 2009

also 2-randomness) proved in Miller (J. Symb. Log. 69(2):555-584, 2004): a sequence ω is 2-random if and only if there exists c such that any prefix x of ω is a prefix of some string y such that $C(y) \geq |y| - c$. (In the 1960ies this property was suggested in Kolmogorov, IEEE Trans. Inf. Theory IT-14(5):662–664, 1968, as one of possible randomness definitions; its equivalence to 2-randomness was shown in Miller, J. Symb. Log. 69(2):555-584, 2004.) Miller (J. Symb. Log. 69(2):555-584, 2004) and Nies et al. (J. Symb. Log. 70(2):515–535, 2005) proved another 2-randomness criterion: ω is 2-random if and only if $C(x) \geq |x| - c$ for some c and infinitely many prefixes x of ω.

We show that the low-basis theorem can be used to get alternative proofs of our results on Kolmogorov complexity and to improve the result about effectively open sets; this stronger version implies the 2-randomness criterion mentioned in the previous sentence.

Keywords Kolmogorov complexity · Relativization · Limit frequency

1 Plain Complexity

We denote by $\{0, 1\}^*$ the set of binary strings and by $\{0, 1\}^\infty$ the set of infinite binary sequences. For $x \in \{0, 1\}^*$, we denote by $C(x)$ the plain complexity of x (the length of the shortest description of x when an optimal description method is fixed, see Li and Vitanyi [3]; no requirements about prefixes). By $C(x|n)$ we mean conditional complexity of x when n is given, see for example Li and Vitanyi [3]. Superscript $\mathbf{0}'$ in $C^{\mathbf{0}'}$ means that we consider the relativized version of complexity to the oracle $\mathbf{0}'$, the universal computably enumerable set.

The following result was proved in Vereshchagin [10]. We provide a simple proof for it.

Theorem 1 *For all* $x \in \{0, 1\}^*$:

$$\limsup_{n\to\infty} C(x|n) = C^{\mathbf{0}'}(x) + O(1).$$

(In this theorem and below "$f(x) = g(x) + O(1)$" means that there is a constant c such that $|f(x) - g(x)| \leq c$ for all x.)

Proof We start in the easy direction. Let $\mathbf{0}_n$ be the (finite) set consisting of the elements of the universal enumerable set $\mathbf{0}'$ that have been enumerated after n steps of computation (note that $\mathbf{0}_n$ can be computed from n). If $C^{\mathbf{0}'}(x) \leq k$, then there exists a description (program) of size at most k that generates x using $\mathbf{0}'$ as an oracle. Only finite part of the oracle can be used, so $\mathbf{0}'$ can be replaced by $\mathbf{0}_n$ for all sufficiently large n, and oracle $\mathbf{0}_n$ can be reconstructed if n is given as a condition. Therefore, $C(x|n) \leq k + O(1)$ for all sufficiently large n, and

$$\limsup_{n\to\infty} C(x|n) \leq C^{\mathbf{0}'}(x) + O(1).$$

For the reverse inequality, fix k and assume that $\limsup C(x|n) < k$. This means that for all sufficiently large n the string x belongs to the set

$$U_n = \{u \mid C(u|n) < k\}.$$

The family U_n is an enumerable family of sets (given n and k, we can generate U_n); each of these sets has at most 2^k elements. We need to construct a $\mathbf{0}'$-computable process that given k generates at most 2^k elements including all elements that belong to U_n for all sufficiently large n. (Then strings of length k may be assigned as $\mathbf{0}'$-computable codes of all generated elements.)

To describe this process, consider the following operation: for some u and N add u to all U_n such that $n \geq N$. (In other terms, we add a horizontal ray starting from (N, u) to the set $\mathcal{U} = \{(n, u) \mid u \in U_n\}$.) This operation is *acceptable* if all U_n still have at most 2^k elements after it (i.e., if before this operation all U_n such that $n \geq N$ either contain u or have less than 2^k elements).

For any given triple u, N, k, we can find out using $\mathbf{0}'$-oracle whether this operation is acceptable or not. Indeed, the operation is not acceptable if and only if some U_n for $n \geq N$ contains at least 2^k elements that are distinct from u. Formally, the operation is not acceptable if

$$(\exists n \geq N) \quad |U_n \setminus \{u\}| \geq 2^k,$$

and this is an enumerable condition as the U_n are themselves enumerable. Now for all pairs (N, u) (in some computable order) we perform the (N, u)-operation if it is acceptable. (The elements added to some U_i remain there and are taken into account when next operations are attempted.) This process is $\mathbf{0}'$-computable since after any finite number of operations the set $\mathcal{U}$ is enumerable (without any oracle) and its enumeration algorithm can be $\mathbf{0}'$-effectively found (uniformly in k).

Therefore the set of all elements u that participate in acceptable operations during this process is uniformly $\mathbf{0}'$-enumerable. This set contains at most 2^k elements (otherwise U_n would become too big for large n). Finally, this set contains all u such that u belongs to the (original) U_n for all sufficiently large n. Indeed, the operation is always acceptable if the element we want to add is already present! □

The proof has the following structure. We have an enumerable family of sets U_n that all have at most 2^k elements. This implies that the set

$$U_\infty = \liminf_{n \to \infty} U_n$$

has at most 2^k elements where, as usual, the lim inf of a sequence of sets is the set of elements that belong to almost all sets of the sequence. If U_∞ were $\mathbf{0}'$-enumerable, we would be done. However, this may be not the case: the criterion

$$u \in U_\infty \quad \Longleftrightarrow \quad \exists N\, (\forall n \geq N) \quad [u \in U_n]$$

has $\exists\forall$ prefix before an enumerable (not necessarily decidable) relation, that is, one quantifier more than we want (to guarantee that U_∞ is $\mathbf{0}'$-enumerable). However, in our proof we managed to cover U_∞ by a set that is $\mathbf{0}'$-enumerable and still has at most 2^k elements.

2 Prefix Complexity and a Priori Probability

We now prove a similar result for prefix complexity (or, in other terms, for a priori probability). Let us recall the definition. The function $a(x)$ on binary strings (or integers) with non-negative real values is called a *semimeasure* if $\sum_x a(x) \le 1$. The function a is *lower semicomputable* if there exists a computable total function $(x,n) \mapsto a(x,n)$ with rational values such that for every x the sequence $a(x,0), a(x,1), \dots$ is a nondecreasing sequence that has limit $a(x)$.

There exists a maximal (up to a constant factor) lower semicomputable semimeasure m (see, e.g., Li and Vitanyi [3]). The value $m(x)$ is sometimes called the *a priori probability* of x. In the same way we can define *conditional* a priory probability $m(x|n)$ and $\mathbf{0}'$-*relativized* a priori probability $m^{\mathbf{0}'}(x)$ (which is a maximal semimeasure among the $\mathbf{0}'$-lower semicomputable ones).

Theorem 2 *For all* $x \in \{0,1\}^*$:

$$\liminf_{n\to\infty} m(x|n) = m^{\mathbf{0}'}(x)$$

up to a $\Theta(1)$ *multiplicative factor* (*in other terms*, *two inequalities with* $O(1)$ *factors hold*).

Proof If $m^{\mathbf{0}'}(x)$ is greater than some ε, then for sufficiently large n the value $m^{\mathbf{0}_n}(x)$ is also greater than ε. (Indeed, this inequality is established at some finite stage when only a finite part of $\mathbf{0}'$ is used.) We may assume without loss of generality that the function $x \mapsto m^A(x)$ is a semimeasure for any A (recalling the construction of the maximal semimeasure). Then, similarly to the previous theorem, we have

$$\liminf_{n\to\infty} m(x|n) \ge \liminf_{n\to\infty} m^{\mathbf{0}_n}(x) \ge m^{\mathbf{0}'}(x)$$

up to constant multiplicative factors. Indeed, for the first inequality, notice that we can define a conditional lower semicomputable semimeasure μ by $\mu(x|n) = m^{\mathbf{0}_n}(x)$. By maximality of m, we have $\mu(x|n) \le m(x|n)$ for all x, n, up to a multiplicative factor. For the second inequality, recall that $m^{\mathbf{0}'}(x)$ is the nondecreasing limit of an $\mathbf{0}'$-computable sequence $m^{\mathbf{0}'}(x,0), m^{\mathbf{0}'}(x,1), \dots$. Let s be such that $m^{\mathbf{0}'}(x,s) \ge \frac{1}{2} m^{\mathbf{0}'}(x)$. Since the computation of $m^{\mathbf{0}'}(x,s)$ only uses finitely many bits of $\mathbf{0}'$, we have for all large enough n: $m^{\mathbf{0}_n}(x,s) = m^{\mathbf{0}'}(x,s) \ge \frac{1}{2} m^{\mathbf{0}'}(x)$ and thus $m^{\mathbf{0}_n}(x) \ge \frac{1}{2} m^{\mathbf{0}'}(x)$.

The other direction of the proof is also similar to the second part of the proof of Theorem 1. Instead of enumerable finite sets U_n we now have a sequence of (uniformly) lower semicomputable functions $x \mapsto m_n(x) = m(x|n)$. Each of the m_n is a semimeasure. We need to construct an $\mathbf{0}'$-lower semicomputable semimeasure m' such that

$$m'(x) \ge \liminf_{n\to\infty} m_n(x).$$

Again, the $\liminf$ itself cannot be used as m': we do have $\sum_x \liminf_n m_n(x) \le 1$ as $\sum_x m_n(x) \le 1$ for all n, but unfortunately the equivalence

$$r < \liminf_{n\to\infty} m_n(x) \quad \Longleftrightarrow \quad (\exists r' > r) \quad (\exists N)\ (\forall n \ge N) \quad [r' < m_n(x)]$$

has too many quantifier alternations (one more than needed; note that the quantity $m_n(x)$ is lower semicomputable making the [...] condition enumerable). The similar trick helps. For a triple (r, N, u) consider an *increase operation* that increases all values $m_n(u)$ such that $n \ge N$ up to a given rational number r (not changing them if they were greater than or equal to r). This operation is *acceptable* if all m_n remain semimeasures after the increase.

The question whether the increase operation is acceptable is $\mathbf{0}'$-decidable. And if it is acceptable, by performing it we get a new (uniformly) lower semicomputable sequence of semimeasures. We can then try to perform an increase operation for some other triple. Doing that for all triples (in some computable ordering), we can then define $m'(u)$ as the upper bound of r for all successful (r, N, u) increase operations (for all N). This gives a $\mathbf{0}'$-lower semicomputable function; it is a semimeasure since we verify the semimeasure inequality for every successful increase attempt; finally, $m'(u) \ge \liminf m_n(u)$ since if $m_n(u) \ge r$ for all $n \ge N$, then the (r, N, u)-increase does not change anything and is guaranteed to be acceptable at any step. □

The expression $-\log m(x)$, where m is the maximal lower semicomputable semimeasure, equals the so-called *prefix* complexity $K(x)$ (up to an additive $O(1)$ term; see for example Li and Vitanyi [3]). The same is true for relativized and conditional versions, and we get the following reformulation of the last theorem:

Theorem 3

$$\limsup_{n\to\infty} K(x|n) = K^{\mathbf{0}'}(x) + O(1).$$

Another corollary improves a result of Muchnik [5]. For any (partial) function f from $\mathbb{N}$ to $\mathbb{N}$ let us define the *limit frequency* $q_f(x)$ of an integer x as

$$q_f(x) = \liminf_{n\to\infty} \frac{\#\{i < n \mid f(i) = x\}}{n}.$$

In other words, we look at the fraction of terms equal to x among the first n values $f(0), \dots, f(n-1)$ of f (undefined values are also listed) and take the $\liminf$ of these fractions. It is easy to see that for a total computable f the function q_f is a lower $\mathbf{0}'$-semicomputable semimeasure. Moreover, it is shown in Muchnik [5] that any $\mathbf{0}'$-semicomputable semimeasure μ can be represented as $\mu = q_f$ for some computable function f. In particular this implies that there exists a total computable function f such that $q_f = m^{\mathbf{0}'}$.

We would like to extend Muchnik's result to partial computable functions f. The problem is that if f is only partial computable, the function q_f is no longer guaranteed to be lower semicomputable. Using the second part of the proof of Theorem 2, we can nonetheless prove:

Theorem 4 *For any partial computable function f, the function q_f is upper bounded by a lower $\mathbf{0}'$-semicomputable semimeasure.*

Indeed, given a partial computable function f, we can define for all n a semimeasure μ_n as

$$\mu_n(x) = \frac{\#\{i < n \mid f(i) = x\}}{n};$$

μ_n is lower semicomputable uniformly in n. Then $q_f = \liminf \mu_n$; on the other hand we know from the proof of Theorem 2 that the lim inf of a sequence of (uniformly) lower semicomputable semimeasures is bounded by a $\mathbf{0}'$-lower semicomputable semimeasure. The result follows.

The same type of argument also is applicable to the so-called *a priori complexity* defined as negative logarithm of a maximal lower semicomputable semimeasure on the binary tree (see Zvonkin and Levin [11]). This complexity is sometimes denoted as $KA(x)$ and we get the following statement:

Theorem 5

$$\limsup_{n\to\infty} KA(x|n) = KA^{\mathbf{0}'}(x) + O(1).$$

(To prove this we define an increase operation in such a way that, for a given lower semicomputable semimeasure on the binary tree a, it increases not only $a(x)$ but also $a(y)$ for y that are prefixes of x, if necessary. The increase is acceptable if $a(\Lambda)$ still does not exceed 1.)

It would be interesting to find out whether similar results are true for monotone complexity or not (the authors do not know this).

3 Open Sets of Small Measure

In Sect. 1 we covered the lim inf of a sequence of finite uniformly enumerable sets U_i by a $\mathbf{0}'$-enumerable set V that is essentially no bigger than the U_i. It was done in a uniform way, i.e., V can be effectively constructed given the enumerations of the U_i and an upper bound for their cardinalities. We now look at the continuous version of this problem where the U_i are open sets of small measure.

We consider open sets in the Cantor space $\{0, 1\}^\infty$ (the set of all infinite sequence of zeros and ones). An *interval* $[x]$ (for a binary string x) is formed by all sequences that have prefix x. Open sets are unions of intervals. An *effectively open* subset of $\{0, 1\}^\infty$ is an enumerable union of intervals, i.e., the union of intervals $[x]$ where strings x are taken from some enumerable set.

We consider standard (uniform Bernoulli) measure on $\{0, 1\}^\infty$: the interval $[x]$ has measure 2^{-l} where l is the length of x.

A classical theorem of measure theory says: *if $U_0, U_1, U_2, \ldots$ are open sets of measure at most ε, then $\liminf_n U_n$ has measure at most ε*, and this implies that *for* *every $\varepsilon' > \varepsilon$ there exists an open set of measure at most ε' that covers* $\liminf_n U_n$. Indeed,

$$\liminf_{n\to\infty} U_n = \bigcup_N \bigcap_{n\ge N} U_n,$$

and the measure of the union of an increasing sequence

$$V_N = \bigcap_{n\ge N} U_n,$$

equals the limit of measures of V_N, and all these measures do not exceed ε since $V_N \subset U_N$. It remains to note that for any measurable subset X of $\{0,1\}^\infty$ its measure $\mu(X)$ is the infimum of the measures of open sets that cover X.

We now can try to "effectivize" this statement in the same way as we did before. In Sect. 1 we started with an (evident) statement: *if U_n are finite sets of at most 2^k elements, then* $\liminf_n U_n$ *has at most 2^k elements* and proved its effective (in the halting problem) version: *for a uniformly enumerable family of finite sets U_n that have at most 2^k elements, the set* $\liminf_n U_n$ *is contained in a uniformly $\mathbf{0}'$-enumerable set that has at most 2^k elements.*

In Sect. 2 we did a similar thing with semimeasures. Again, the non-effective version is trivial: it says that if $\sum_x m_n(x) \le 1$ for every n, then $\sum_x \liminf_n m_n(x) \le 1$. We have proved the effective version that provides a $\mathbf{0}'$-semicomputable semimeasure that is an upper bound for $\liminf m_n$.

For the statement about lim inf of open sets the effective version could look like this. *Let $\varepsilon > 0$ be a rational number and let $U_0, U_1, \ldots$ be an enumerable family of effectively open sets of measure at most ε each. Then for every rational $\varepsilon' > \varepsilon$ there exists a $\mathbf{0}'$-effectively open set V of measure at most ε' that contains* $\liminf_{n\to\infty} U_n = \bigcup_N \bigcap_{n\ge N} U_n$.

We cannot prove this general statement and do not know whether it is true (see Sect. 8 for some partial negative results). However, some weaker statements can be proven if we put extra requirements on the sets U_n or weaken the conclusion. Let us start with the simple case where the U_n form a computable family of clopen (closed and open) sets. Such a set is a finite union of intervals, and we assume that the list of these intervals can be computed given n.

Theorem 6 *Let U_n be a uniformly computable family of clopen sets. Suppose also that for all n the set U_n has measure at most ε for some rational ε. Then for every rational $\varepsilon' > \varepsilon$ there exists a $\mathbf{0}'$-effectively open set V of measure at most ε' such that*

$$U_\infty = \liminf_{n\to+\infty} U_n \subseteq V.$$

Proof By definition

$$U_\infty = \bigcup_N \bigcap_{n\ge N} U_n,$$

therefore U_∞ is a union of the pairwise disjoint sets

$$F_0 = \bigcap_i U_i, \qquad F_1 = \bigcap_{i \geq 1} U_i \setminus U_0, \qquad F_2 = \bigcap_{i \geq 2} U_i \setminus U_1, \ldots$$

(in other terms, a given $x \in U_\infty$ belongs to F_k if and only if the last U_i to which x does not belong is U_{k-1}). Each of F_k is an effectively closed set (recall that each U_i is a finite union of intervals hence is closed). Since the sets F_k are disjoint and

$$\liminf_{n \to +\infty} U_n = \bigcup_k F_k,$$

we conclude that

$$\mu\Big(\liminf_{n \to +\infty} U_n\Big) = \sum_k \mu(F_k).$$

For each k the value $\mu(F_k)$ is the limit over r of the (non-increasing) quantity $\mu(\bigcap_{i=k}^r U_i \setminus U_{k-1})$ which is computable uniformly in (i, r). Thus, with oracle $\mathbf{0}'$, one can compute $\mu(F_k)$ for every k (with arbitrary precision) and find a stage r_k such that

$$\mu\left(\bigcap_{i=k}^{r_k} U_i \setminus U_{k-1}\right) < \mu(F_k) + (\varepsilon' - \varepsilon)/2^{k+1}.$$

Set $F'_k = \bigcap_{i=k}^{r_k} U_i \setminus U_{k-1}$. Notice that F'_k contains F_k, and is itself an clopen set, and the list of corresponding intervals can be $\mathbf{0}'$-effectively computed given k. Thus,

$$V = \bigcup_k F'_k$$

is a $\mathbf{0}'$-effectively open set, contains $\bigcup_k F_k = U_\infty$ and has measure at most

$$\underbrace{\sum_k \mu(F_k)}_{= \mu(U_\infty) \leq \varepsilon} + \underbrace{\sum_k (\varepsilon' - \varepsilon)/2^{k+1}}_{= \varepsilon' - \varepsilon},$$

i.e. at most ε'. Hence, V satisfies all the requirements. □

As we have said, instead of putting additional requirements on the sequence U_i (requiring it to be a computable sequence of clopen sets) we can weaken the conclusion. The techniques presented in the previous sections allow us to prove the following:

Theorem 7 *Let $\varepsilon > 0$ be a rational number and let U_n be a uniformly enumerable family of effectively open sets of measure at most ε each. Then there exists a $\mathbf{0}'$-effectively open set of measure at most ε that contains*

$$\bigcup_N \mathrm{Int}\Big(\bigcap_{n \geq N} U_n\Big)$$

(*and the construction is uniform given ε and an index for the sequence U_n*).

Here $\mathrm{Int}(X)$ denotes the interior part of X, i.e., the largest open subset of X. In this case we do not need to consider $\varepsilon' > \varepsilon$ (not a surprise, since the union of open sets is open).

Proof Following the same scheme as in Sects. 1 and 2, for every string x and integer N we consider (x, N)-operation that adds $[x]$ to all U_n such that $n \geq N$. This operation is *acceptable* if the measures of all the U_n remain at most ε for each n. This can be checked using $\mathbf{0}'$ as an oracle (if the operation is not acceptable, this fact becomes known after a finite number of steps).

We attempt to perform this operation (if acceptable) for all pairs in some computable order. The union of all added intervals for all accepted pairs is $\mathbf{0}'$-effectively open. If some sequence belongs to the union of the interior parts, then it is covered by some interval $[u]$ that is a subset of U_n for all sufficiently large n. Then some (u, N)-operation is acceptable since it actually does not change anything and therefore $[u]$ is a part of an $\mathbf{0}'$-open set that we have constructed. □

In Sect. 6 we will return to this topic and state in Theorem 10 one more result about the lim inf of small sets.

4 Kolmogorov and 2-Randomness

Theorem 7 has an historically remarkable corollary. When Kolmogorov tried to define randomness in 1960ies, he started with the following approach. A string x of length n is "random" if its complexity $C(x)$ (or conditional complexity $C(x|n)$; in fact, these requirements are almost equivalent) is close to n: its *randomness deficiency* $d(x)$ is defined as

$$d(x) = |x| - C(x)$$

(here $|x|$ stands for the length of x). This sounds reasonable, but if we then define an infinite random sequence as a sequence whose prefixes have deficiencies bounded by a constant, such a sequence does not exist at all: Martin-Löf showed that every infinite sequence has prefixes of arbitrarily large deficiency, and suggested a different definition of randomness using effectively null sets. Later more refined versions of randomness deficiency (using monotone or prefix complexity) appeared that make the criterion of randomness in terms of deficiencies possible. But before that, in 1968, Kolmogorov wrote: "The most natural definition of infinite Bernoulli sequence is the following: x is considered m-Bernoulli type if m is such that all x^i are *initial segments* of the finite m-Bernoulli sequences. Martin-Löf gives another, possibly narrower definition", see Kolmogorov [2, p. 663].

Here Kolmogorov speaks about "m-Bernoulli" finite sequence x (this means that $C(x|n, k)$ is greater than $\log \binom{n}{k} - m$ where n is the length of x and k is the number of ones in x). We restrict ourselves to the case of uniform Bernoulli measure where $p = q = 1/2$. In this case Kolmogorov's idea can be described as follows: an infinite sequence is random if each its prefix also appears as a prefix of some random string (= string with small randomness deficiency). More formal, let us define

$$\bar{d}(x) = \inf\{d(y) \mid x \text{ is a prefix of } y\}$$

and require that $\bar{d}(x)$ is bounded for all prefixes of an infinite sequence ω. It is shown by Miller [4] that this definition is equivalent to Martin-Löf randomness relativized to $\mathbf{0}'$ (called also 2-*randomness*):

Theorem 8 *A sequence ω is Martin-Löf $\mathbf{0}'$-random if and only if the quantities $\bar{d}(x)$ for all prefixes x of ω are bounded from above by a common constant.*

It turns out that the forward direction of the equivalence stated in Theorem 8 follows easily from Theorem 7.

Proof Assume that $\bar{d}$-deficiencies for prefixes of ω are not bounded. According to Martin-Löf's definition, we have to construct for a given c an $\mathbf{0}'$-effectively open set that covers ω and has measure at most 2^{-c}.

Fix some c. For each n consider the set D_n^c of all sequences u of length n such that $C(u) < n - c$ (i.e., sequences u of length n such that $d(u) > c$). It has at most 2^{n-c} elements. By definition of $\bar{d}$, the requirement $\bar{d}(x) > c$ means that every string extension y of x belongs to D_m^c where m is the length of y. This implies that $[x]$ (= set of sequences with prefix x) is contained in every U_m for $m \geq |x|$, where

$$U_m^c = \bigcup_{u \in D_m^c} [u]$$

(in other words U_m^c is the set of all sequences that have prefixes in D_m^c). Therefore, in this case the interval $[x]$ is a subset of $\bigcap_{m \geq |x|} U_m^c$ and (being open) is a subset of its interior. To sum up, we have proven that if an infinite sequence ω has a prefix x such that $\bar{d}(x) > c$, then

$$\omega \in \operatorname{Int}\left(\bigcap_{m \geq |x|} U_m^c\right).$$

Now note that each U_m^c is effectively open uniformly in (m, c) as D_m^c is enumerable uniformly in (m, c). Moreover, there are at most 2^{m-c} strings in D_m^c, hence the measure of U_m^c is at most 2^{-c}. Applying Theorem 7, we conclude that there exists a $\mathbf{0}'$-effectively open (uniformly in c) set V_c that has measure at most 2^{-c} such that

$$\bigcup_N \operatorname{Int}\left(\bigcap_{m \geq N} U_m^c\right) \subseteq V_c.$$

Note that this tells us in particular that the sequence $(V_c)_{c \in \mathbb{N}}$ forms an $\mathbf{0}'$-Martin-Löf test. Thus, if a sequence ω is $\mathbf{0}'$-Martin-Löf random, it only belongs to finitely many V_c. Let then b be such that $\omega \notin V_b$. By the above argument, this means that ω has no prefix x such that $\bar{d}(x) > b$, or equivalently that for every prefix x of ω, $\bar{d}(x) \leq b$. This proves the forward direction of the equivalence.

For the sake of completeness, we give the proof of the reverse implication in terms of Martin-Löf tests (Miller [4] provided a proof solely in terms of Kolmogorov complexity). Assume that a sequence ω is covered (for each c) by a $\mathbf{0}'$-computable sequence of intervals $I_0, I_1, \ldots$ of total measure at most 2^{-c}. (We omit c in our notation, but the construction below depends on c.)

Using the approximations $\mathbf{0}_n$ of $\mathbf{0}'$ (obtained by performing at most n steps of computation for each n) we get another (now computable) family of intervals $I_{0,n}, I_{1,n}, \ldots$ such that $I_{i,n} = I_i$ for every i and sufficiently large n. We may assume without loss of generality that $I_{i,n}$ either has size at least 2^{-n} (i.e., is determined by a string of length at most n) or equals $\perp$ (a special value that denotes the empty set) since only the limit behavior is prescribed. Moreover, we may also assume that $I_{i,n} = \perp$ for $n < i$ and that the total measure of all $I_{0,n}, I_{1,n}, \ldots$ does not exceed 2^{-c} for every n (the latter is achieved by deleting the excessive intervals in this sequence starting from the beginning; the stabilization guarantees that all limit intervals will be eventually let through).

Since $I_{i,n}$ is defined by intervals of size at least 2^{-n}, we get at most 2^{n-c} strings of length n covered by intervals $I_{i,n}$ for any given n and all i. This set of strings is decidable (recall that only i not exceeding n are used), therefore each string in this set can be defined, assuming c is known, by a string of length $n - c$, the binary representation of its ordinal number in this set. Note that this string also determines n if c is known.

Returning to the sequence ω, we note that it is covered by some I_i and therefore is covered by $I_{i,n}$ for this i and all sufficiently large n (after the value of $I_{i,n}$ is stabilized), say, for all $n \geq N$. Let u be the prefix of ω of length N. All extensions of u of any length n are covered by $I_{i,n}$ and thus have complexity less than $n - c + O(1)$, conditional to c, hence their complexity is at most $n - c + 2\log c + O(1)$. This means that $\bar{d}(u) \geq c - 2\log c - O(1)$.

Such a string u can be found for every c, therefore ω has prefixes of arbitrarily large $\bar{d}$-deficiency. □

In fact a stronger statement than Theorem 8 is proved in Miller [4] and Nies et al. [6]; our tools are still too weak to get this statement. However, the low basis theorem helps.

5 The Low Basis Theorem

The low basis theorem is a classical result in recursion theory (see, for example, Odifreddi [7]). It was used in Nies et al. [6] to prove 2-randomness criterion; analyzing this proof, we get theorems about limit complexities as byproducts. For the sake of completeness, we state the low-basis theorem and its simple proof.

Theorem 9 *Let $U \subset \{0,1\}^\infty$ be an effectively open set that does not coincide with $\{0,1\}^\infty$. Then there exists a sequence $\omega \notin U$ which is low, i.e., $\omega' =_{\mathrm{T}} \mathbf{0}'$.*

Here ω' is the jump of ω; the equation $\omega' =_{\mathrm{T}} \mathbf{0}'$ means that the universal ω-enumerable set is $\mathbf{0}'$-decidable.

Theorem 9 says that any effectively closed nonempty set contains a low element. For example, if $P, Q \subset \mathbb{N}$ are enumerable inseparable sets, then the set of all separating sequences is an effectively closed set that does not contain computable sequences. We conclude, therefore, that there exists a non-computable low separating sequence.

Proof Assume that an oracle machine M and an input x are fixed. The computation of M with oracle ω on x may terminate or not depending on oracle ω. Let us consider the set $T(M,x)$ of all ω such that $M^{\omega}(x)$ terminates (for fixed machine M and input x). This set is an effectively open set: if termination happens, it happens due to finitely many oracle values. This set together with U may cover the entire space $\{0,1\}^{\infty}$; this means that $M^{\omega}(x)$ *terminates for all* $\omega \notin U$. If it is not the case, we can add $T(M,x)$ to U and get a bigger effectively open set $\widehat{U}$ that still has nonempty complement such that $M^{\omega}(x)$ *does not terminate for all* $\omega \notin \widehat{U}$. Either way, this operation guarantees that the termination of the computation $M^{\omega}(x)$ does not depend on the choice of ω in the remaining nonempty effectively closed set (meaning that for all ω_1, ω_2 in the remaining effectively closed set, $M^{\omega_1}(x)$ terminates if and only if $M^{\omega_2}(x)$ terminates).

This increase operation of the effectively open set can be performed for all pairs (M,x) sequentially. At each stage the effectively open set U stays the same or is increased but in any case its complement remains nonempty. Hence, by compactness of $\{0,1\}^{\infty}$, the open set U_{∞} obtained in the limit will have nonempty complement. Note that the set U_{∞} does not have to be *effectively* open: though at any stage the current U is an effectively open set, the construction is not effective (we need to find out which of the two cases happens).

We claim that any sequence $\omega \notin U_{\infty}$ is low. Indeed, by construction of U_{∞}, for every M and x the termination of the computation of $M^{\omega}(x)$ is independent on the choice of ω in the complement of U_{∞} and is determined at some point of the construction. And the construction is $\mathbf{0}'$-effective: if during the increase operation $U \cup T(M,x)$ covers the entire space $\{0,1\}^{\infty}$, this happens on some finite stage (compactness), so $\mathbf{0}'$ is enough to find out whether this happens or not. Therefore, for every M and x we can $\mathbf{0}'$-effectively find out whether $M^{\omega}(x)$ terminates or not. This precisely means that $\omega' =_{\mathrm{T}} \mathbf{0}'$, i.e., that ω is low. □

6 Using the Low Basis Theorem

Let us show how Theorem 1 can be proved using the low basis theorem. As we have seen, we have an enumerable family of sets U_n; each of U_n has at most 2^k elements (say, strings). We need to construct effectively a $\mathbf{0}'$-enumerable set that has at most 2^k elements and contains $U_{\infty} = \liminf_n U_n$.

In the special case where the sets U_n happen to be (uniformly) decidable, U_{∞} is $\mathbf{0}'$-enumerable and we do not need any other set. The low basis theorem allows us to reduce the general case to this special one.

First, we may assume without loss of generality that for all n the set U_n contains only strings of length at most n. To see why we can do this, consider for all n the set $\widehat{U}_n$ of strings in U_n that have length at most n. The sequence of $\widehat{U}_n$ is uniformly enumerable and $\liminf U_n = \liminf \widehat{U}_n$ (any string $x \in \liminf_n U_n$ belongs to almost all U_n and will be allowed to enter $\widehat{U}_n$ for $n \geq |x|$).

Having imposed this restriction on the U_n, let us consider the family of all "upper bounds" for U_n: by an upper bound we mean a sequence V_n of finite sets such that for all n, we have (1) $U_n \subseteq V_n$; (2) $\#V_n \leq 2^k$ and (3) V_n contains strings of length at most n. The sequence $V_0, V_1, \ldots$ can be encoded as an infinite binary sequence: each V_i contains only strings of length at most n (there are $2^{n+1} - 1$ of them) and can be encoded as a binary string of length $2^{n+1} - 1$. Then the sequence $V_0, V_1, \ldots$ can be encoded by concatenation of the individual encodings of the V_i.

For a binary sequence the property "to be an encoding of an upper bound for U_n" is effectively closed (the restriction $\#V_n < 2^k$ is decidable and the restriction $U_n \subset V_n$ is co-enumerable). Therefore the low basis theorem can be applied. We get an upper bound V that is low. Then $V_\infty = \liminf V_n$ is (uniformly in k) V'-enumerable (as we have said: with V-oracle the family V_n is uniformly decidable), but since V is low, the V'-oracle can be replaced by the $\mathbf{0}'$-oracle, and we get the desired result.

This proof though being simple looks rather mysterious: we get something almost out of nothing! (As far as we know, this idea appeared in a slightly different context in Nies et al. [6].)

The same trick can be used to prove Theorem 2: here "upper bounds" are distributions M_n with rational values and finite support that are greater than $m(x|n)$ but still are semimeasures. (Technical correction: first we have to assume that $m(x|n) = 0$ if x is large, and then we have to weaken the restriction $\sum M_n(x) \leq 1$ replacing 1 by, say, 2; this is needed since the values $m(x|n)$ may be irrational.)

Theorem 5 can be also proved in this way (upper bounds should be semimeasures on tree with rational values and finite support).

Returning to the topic of Sect. 3, we can use the low basis theorem to improve Theorem 6:

Theorem 10 *Let $\varepsilon > 0$ be a rational number and let U_n be a sequence sets that are effectively open* (*uniformly in n*). *Assume that U_n has measure at most ε for every n. Assume also that U_i has "effectively bounded granularity", i.e., all strings x that define the intervals in U_n have length at most $c(n)$ where c is a total computable function. Then for every $\varepsilon' > \varepsilon$ there exists a $\mathbf{0}'$-effectively open set V of measure at most ε' that contains* $\liminf_{n\to\infty} U_n$ *and this construction is uniform.*

Proof We use the low basis theorem to reduce the general case to the case where the U_n form a computable family of finitely generated open sets.

Indeed, define an "upper bound" as a sequence W of sets W_n where W_n is a set of strings of length at most $c(n)$ such that U_n is covered by the intervals generated by elements of W_n. Again W can be encoded as an infinite sequence of zeros and ones, and the property "to be an upper bound" is effectively closed. Applying the low basis theorem, we choose a low W and add it is an oracle; evidently, W_n is a W-computable family of finitely generated open sets. By Theorem 6 (relativized to oracle W) for every $\varepsilon' > \varepsilon$ there exists a W'-effectively open set V covering $\liminf_n W_n$, hence covering $\liminf_n U_n$. And since W' is Turing-equivalent to $\mathbf{0}'$, we are done. □

7 Corollary on 2-Randomness

Theorem 10 can be used to prove 2-randomness criterion from [4, 6]. In fact, this gives exactly the proof from [6]; the only thing we did is structuring the proof in two parts (formulating Theorem 10 explicitly and putting it in the context of other results on limits of complexities). For the sake of completeness, let us reproduce this proof.

Theorem 11 [4, 6] *A sequence ω is $\mathbf{0}'$ Martin-Löf random if and only if*

$$C(\omega_0\omega_1\cdots\omega_{n-1}) \geq n - c$$

for some c and for infinitely many n.

Proof Let us first understand the relation between this theorem and Theorem 8. If

$$C(\omega_0\omega_1\cdots\omega_{n-1}) \geq n - c$$

for infinitely many n and given c, then $\bar{d}(x) \leq c$ for every prefix x of ω (indeed, one can find the required continuation of x among prefixes of ω). As we know, this guarantees that ω is $\mathbf{0}'$ Martin-Löf random.

It remains to prove that if for all c we have

$$C(\omega_0\omega_1\cdots\omega_{n-1}) < n - c$$

for all sufficiently large n, then ω is not $\mathbf{0}'$-random. Using the same notation as in the proof of Theorem 8, we can say that ω has a prefix in D_n and therefore belongs to U_n for all sufficiently large n. We can apply then Theorem 10 since U_n is defined using strings of length n (so $c(n) = n$) and cover U_∞ (and therefore ω) by a $\mathbf{0}'$-effectively open set of small measure. Since this can be uniformly done for all c, the sequence ω is not $\mathbf{0}'$-random. □

8 The General Case: $\mathbf{\Sigma_3^0}$ Sets and lim inf of Open Sets

As we have said before, it would be nice to prove the following statement: if U_n are (uniformly) effectively open sets in Cantor space, all U_n have measure less than ε, and $\varepsilon' > \varepsilon$, then there exists a $\mathbf{0}'$-effectively open set of measure less than ε' that covers $\liminf U_n$.

We do not know whether this is true or not. However, some partial negative result could be obtained. In this section, we prove that there exists a lim inf of uniformly effectively open sets U_n that has small measure but cannot be covered by a $\mathbf{0}'$-effectively open set of small measure. (The difference is that only this limit set has small measure while the sets U_n itself can have any measure.)

The first step towards this result is to prove that every Σ_3^0 subset of $\{0,1\}^\infty$ can be written as the lim inf of a sequence of uniformly effectively open sets. The term Σ_3^0 refers to the standard effective Borel hierarchy: effectively open sets are the Σ_1^0 sets, effectively closed sets (i.e., the complements of effectively open sets) are Π_1^0

sets etc.: by induction, a Σ^0_n set is an effective countable union of Π^0_{n-1} sets and a Π^0_n set is an effective (countable) intersection of Σ^0_{n-1} sets. It is easy to see from the definition that a lim inf of a sequence of uniformly effectively open sets is a Σ^0_3 set. Simpson [9] noted that the converse is also true:

Theorem 12 (Simpson) *For every Σ^0_3 set $S \subset \{0,1\}^\infty$ there exists a sequence of uniformly effectively open sets V_n such that $S = \liminf_{n\to+\infty} V_n$.*

Proof It is sufficient to construct a procedure which, given $\omega \in \{0,1\}^\infty$ as an oracle, enumerates a set $X^\omega \subset \mathbb{N}$ such that $\omega \in S \Leftrightarrow X^\omega$ is co-finite. Then we let $V_n = \{\omega \mid n \in X^\omega\}$.

Such a procedure is in fact provided by a proof of Σ^0_3-completeness of the cofiniteness property for enumerable sets (see, e.g., Rogers' textbook [8], Corollary XIV in Sect. 14.8). We give a sketch of this proof here.

As S is Σ^0_3, there exists a collection U^n_k of uniformly effectively open sets such that

$$S = \bigcup_n \bigcap_k U^n_k.$$

For a given ω, let

$$Z^\omega = \{(n,k) \mid \omega \in U^n_k\}.$$

Then ω is in S if and only if

$$\text{there exists an } n \quad \text{such that } (n,k) \in Z^\omega \quad \text{for all } k. \tag{$*$}$$

We have to reduce this characterization to a characterization via cofiniteness of some ω-enumerable set X^ω, i.e., we have to transform (effectively and uniformly) the enumeration of Z^ω into the enumeration of an X^ω in such a way that

$$\omega \in S \quad \Longleftrightarrow \quad \exists n\ \forall k \quad \left[(n,k) \in Z^\omega\right] \quad \Longleftrightarrow \quad X^\omega \text{ is cofinite.}$$

To do so, we use a so-called "movable markers" construction. We first consider a countable series of counters: at each stage, the nth counter contains the maximal k such that all k pairs $(n,0),\ldots,(n,k-1)$ have already appeared in Z^ω. The property $(*)$ now means that some counter increases indefinitely.

We also use "markers" that will locate the missing elements in X^ω. Markers are numbered $0,1,2\ldots$. Initially the ith marker is located under the number i; then markers can be moved to the right, but the ith marker is always on the left of $(i+1)$th one (so all markers mark different numbers). When the nth counter increases, the number that was marked by the nth marker is added to X^ω, and all the markers $n, n+1, n+2, \ldots$ are moved to the right (the ith marker moves to the previous place of the $(i+1)$th one). The markers $0,1,\ldots,n-1$ do not move. Note the invariant relation: the currently enumerated part of X^ω is the set of all non-marked numbers.

If the nth counter increases indefinitely, then the nth marker is moved infinitely many times, so X^ω is cofinite (its complement consists of the final positions of the markers $0,1,\ldots,n-1$). Conversely, if every counter increases only finitely many times, then each marker eventually reaches a final position, and these positions form an infinite complement of X^ω. □

Theorem 12 implies the following result, due to Kjos-Hanssen [1]:

Theorem 13 (Kjos-Hanssen) *There exists a sequence U_n of uniformly effectively open sets such that*:

- $\mu(\liminf_n U_n) < 1/2$
- $\liminf_n U_n$ *cannot be covered by any* $\mathbf{0}'$*-effectively open set of measure less than* 1.

Proof This can be obtained by relativizing standard results about Martin-Löf random sequences. The set of non-random sequences can be covered by an effectively open set of measure less that $1/2$ (by definition). Therefore, the set of $\mathbf{0}''$-nonrandom sequences can be covered by an $\mathbf{0}''$-effectively open set, i.e., by Σ^0_3-set (as the standard results about the arithmetic hierarchy say, see, e.g., Rogers [8]).

On the other hand, for every effectively open set that has measure less than 1 there exists a $\mathbf{0}'$-computable sequence outside this set (e.g., the leftmost path of the binary tree representing its complement). And this can be relativized: for every $\mathbf{0}'$-effectively open set of measure less than 1 there is an $\mathbf{0}''$-computable sequence outside it.

Now, combining these two remarks and Theorem 12, we get the desired result. The Σ^0_3 set mentioned above can be represented as lim inf of effectively open sets. If this set could be covered by an $\mathbf{0}'$-effectively open set of measure less than 1, we would be able to find a $\mathbf{0}''$-computable sequence outside it; this sequence is not $\mathbf{0}''$-random but is outside the set that has to cover all nonrandom (relative to $\mathbf{0}''$) sequences. □

Another related question: even if the most general statement mentioned at the beginning of the section is not true, may be it is enough to require that U_n are clopen sets (thus removing the hypothesis of "effectively bounded granularity" from Theorem 10)?

9 Fatou's Lemma

It would be nice to find some general result that could unite several cases that we have treated separately. Indeed, these results may be considered as constructive version of classical Fatou's lemma.

This lemma guarantees that if $\int f_n(x)\,d\mu(x) \le \varepsilon$ for μ-measurable functions $f_0, f_1, f_2, \ldots$, then

$$\int \liminf_{n\to+\infty} f_n(x)\,d\mu(x) \le \varepsilon.$$

The constructive version may require that f_i are lower semicomputable functions (probably with some additional conditions), and the statement could say that for every $\varepsilon' > \varepsilon$ there exists a lower $\mathbf{0}'$-semicomputable function φ such that $\liminf f_n(x) \le \varphi(x)$ for every x and $\int \varphi(x) d\mu(x) \le \varepsilon'$.

Special case of this lemma appears when f_i are indicator functions of some effectively open sets.

However, as we have seen in Sect. 8, some additional requirements may be needed and we don't know how to formulate them in a natural and general form. In Theorem 10 instead of computably bounded granularity we may require that for each U_i we can provide a finite list of "simple" sets that have measure at most ε and guarantee that U_i is contained in one of them; a similar thing can be done for functions (a list of upper bounds having small integrals). The source of difficulties here is the low basis theorem: it uses the compactness of Cantor space, so when choosing upper bound for the U_n or f_n we need to have in advance a finite list of possibilities.

Additional complications appear when the measure of the space where f_n are defined is infinite (and this is needed for the results of Sects. 1 and 2). Then we should artificially cut f_n in such a way that lim inf is not changed; this is possible in the special cases we need, but it is not clear how one can combine all these considerations into one (preferably not very boring) theorem.

Another open question: classical Fatou lemma usually is formulated in a stronger form:

$$\int \liminf_{n\to+\infty} f_n(x)\,d\mu(x) \le \liminf_{n\to+\infty} \int f_n(x)\,d\mu(x)$$

there the right hand side has also lim inf. This motivates the question: what happens if we weaken the condition and require only that $\int f_n(x)\,d\mu(x) \le \varepsilon$ for infinitely many ε? (For the classical version this is not important, since we can delete all the terms of the sequence that are not bounded by ε; this could only increase lim inf in the conclusion.)

Acknowledgements The authors are thankful to Steve Simpson and Bjorn Kjos-Hanssen for allowing them to include Theorem 12 and Theorem 13, to Peter Cholak for useful discussions, to the members of LIF and Poncelet laboratory, to the participants of Kolmogorov seminar and to two anonymous referees for their numerous comments and suggestions.

References

1. Kjos-Hanssen, B.: Private communication, May 2008
2. Kolmogorov, A.N.: Logical basis for information theory and probability theory. IEEE Trans. Inf. Theory **IT-14**(5), 662–664 (1968). (Russian version was published in 1969.)
3. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications, 2nd edn. Springer, Berlin (1997)
4. Miller, J.: Every 2-random real is Kolmogorov random. J. Symb. Log. **69**(2), 555–584 (2004)
5. Muchnik, A.A.: Lower limits of frequencies in computable sequences and relativized a priori probability. Theory Probab. Appl. **32**, 513–514 (1987)
6. Nies, A., Stephan, F., Terwijn, S.: Randomness, relativization and Turing degrees. J. Symb. Log. **70**(2), 515–535 (2005)
7. Odifreddi, P.: Classical Recursion Theory. North-Holland, Amsterdam (1989)
8. Rogers, H. Jr.: Theory of Recursive Functions and Effective Computability. McGraw Hill, New York (1967)
9. Simpson, S.: Private communication, May 2008
10. Vereshchagin, N.K.: Kolmogorov complexity conditional to large integers. Theor. Comput. Sci. **271**(1–2), 59–67 (2002)
11. Zvonkin, A.K., Levin, L.: The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. Russ. Math. Surv. **25**(6), 83–124 (1970)