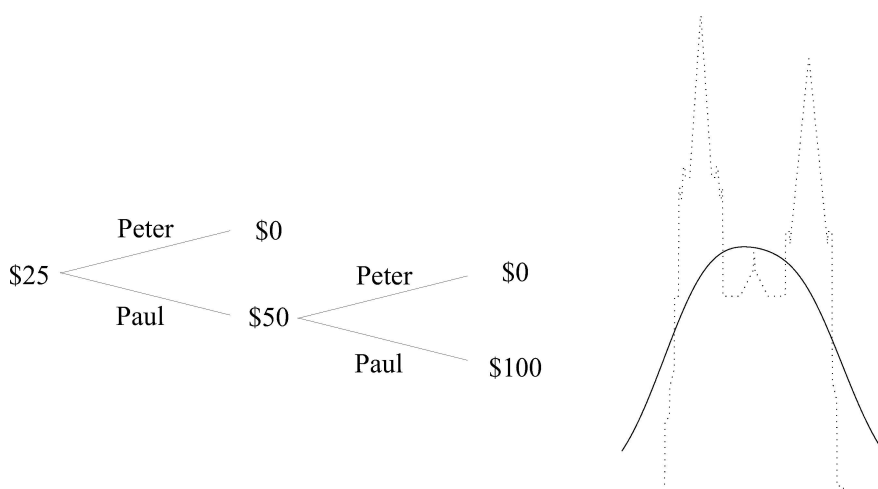


The universal measure of nonstochastic objects

Vladimir Vovk



The Game-Theoretic Probability and Finance Project

Working Paper #70

First posted August 27, 2026. Last revised September 30, 2026.

Project web site:

<http://www.probabilityandfinance.com>

Abstract

The usual settings for discussing the notion of stochasticity are very different from Kolmogorov's original setting as recorded by Shen in November 1981. This paper simplifies Kolmogorov's setting and shows that the universal machine generates nonstochastic objects with surprisingly high probability.

A version of this paper (which is updated less often and is more stable and polished) is published as an arXiv technical report: <http://arxiv.org/abs/2608.27098>.

Contents

1	Introduction	1
2	Kolmogorov's questions	3
3	Main result	5
4	Nonstochastic objects of high and low complexity	7
5	Probability measures as models	8
6	Conclusion	10
	References	10
A	Proofs	11
A.1	The tail of the universal measure	12
A.2	Proof of Theorem 2	13
A.3	Proof of Theorem 4	15
A.4	Proof of Theorem 7	18
A.5	Proof of Theorem 8	18

1 Introduction

The notion of stochastic objects was first introduced (in the currently recorded history) in Kolmogorov’s talk at a seminar at Moscow State University on 26 November 1981. In the same talk he also posed the problem of estimating the abundance of stochastic and nonstochastic objects. (Alexander Shen’s notes from that seminar have been published as [9, note 5.12].) Various solutions to this problem have been given by Shen in 1983 [10], Vladimir V’yugin in 1985 [14], Andrei Muchnik in 1998 [8, Theorems 10.10 and 10.11], and Nikolai Vereshchagin and Shen [12, Proposition 22]. See [11, Sect. 14.2] and [12, Sect. 2] for excellent reviews.

A surprising feature of the existing literature on stochastic objects is that it always assumes some externally given nested structure on the objects whose stochasticity we are interested in.¹ In many cases the objects are the binary strings, with their lengths playing an important role. In other cases (e.g., in [10]), the objects are the nonnegative integers with the nested structure consisting of the sets $\{0, \dots, 2^n - 1\}$, $n = 0, 1, \dots$. Kolmogorov’s original problem [9, note 5.12] is completely different: no externally given nested structure is used in measuring sets of nonstochastic objects. Kolmogorov’s setting will be discussed in Sect. 2. This paper simplifies it further. While Kolmogorov’s setting still involves a nested structure, albeit intrinsic, the main result of this paper (Theorem 4) dispenses with it altogether.

The standard problem of measuring the set of all (α, β) -nonstochastic objects (where α and β are parameters, to be defined in Sects. 2 and 3) has an important special case: do such objects even exist in a given level of the nested structure? (Such as the question of existence of (α, β) -nonstochastic strings of length n .) Important results in this direction are due to Shen [10, Theorem 2] and Gács et al. [6, Theorem IV.2]. Such problems do not arise in the setting of this paper.

Different authors and publications use different ways of measuring sets of nonstochastic objects. Shen [10, Theorem 3] bounds the uniform probability measure of nonstochastic objects. Using the uniform probability measure for measuring the amount of nonstochastic objects appears somewhat unnatural, and [10, Theorem 4] covers non-uniform probability measures; Shen’s Theorem 4, however, only gives a one-sided bound. V’yugin [14, Corollary and Theorem 3] bounds the universal semimeasure of nonstochastic objects. The existing results that are closest to the main result of this paper are Theorem 10.11 in [8] and Theorem 251 in [11, Sect. 14.2]; both these results involve a different version of the universal semimeasure. In the last three of these results, those in [14, 8, 11], the bounds on the measure of the (α, β) -nonstochastic objects are governed by the term $2^{-\alpha}$. Kolmogorov’s setting described in Sect. 2 is very different, but we still have a similar prominent role for $2^{-\alpha}$. The nature of the main result of this paper is even more different, and $2^{-\alpha}$ disappears.

One interpretation of the existing results on the universal semimeasure of

¹In this paper I will follow Shen [10] and Gács et al. [6] and make a terminological distinction between stochastic objects, which are defined using finite sets as models, and quasi-stochastic objects, which are defined using probability measures.

nonstochastic objects is that “the probability of obtaining a non-stochastic object in a random process is negligible” [12, Sect. 4.6]. According to this paper’s results, whether this is true in the structure-free setting depends on the details of the definition of stochasticity (and happens to be true for Vereshchagin and Shen’s official definition [12, Definition 1]). If the definition uses finite sets as models, the situation in the structure-free setting is the opposite, and non-stochastic objects are generated extremely often by the universal semimeasure (Theorem 4 below). But if the definition uses probability measures as models, their counterparts become rare (Corollary 9). In the terminology of [10, 6] and this paper, nonstochastic objects are frequent while non-quasi-stochastic objects are rare.

The notions of stochasticity and quasi-stochasticity are regarded as close in the standard structured settings (see, e.g., [10, Theorem 1] and [12, Proposition 2]). Therefore, nowadays the same term, “stochasticity”, is often used for both stochasticity and quasi-stochasticity. In the structure-free setting a chasm opens up between stochasticity and quasi-stochasticity.

Kolmogorov’s original questions asked in 1981 concern the number of non-stochastic objects of complexity not exceeding a given level k . His specific questions are answered in Sect. 2. The main result of this paper (Theorem 4) is stated in Sect. 3; it does not involve a threshold k and evaluates the likelihood of the universal machine producing nonstochastic objects. The result is very simple and implies that nonstochastic objects are produced very often. In view of Kolmogorov’s questions, it is also natural to ask for the likelihood of producing nonstochastic objects above and below a given complexity level k . The “above” part is answered in the first part of Sect. 4; the answer is almost as simple as the main result and says that complex nonstochastic objects are produced very often. The “below” part is discussed in the second part of Sect. 4; the result here is more complicated and less satisfactory. All nontrivial proofs are collected in Appendix A.

In the case of quasi-stochastic objects some interesting structure-free results exist, and they are in stark contrast with the results of this paper for stochastic objects. The ones that I am aware of originate from [4, Proposition 1]; for proofs, see [7, Proposition 1] and [3, Lemma 10]. One of these results will be discussed in Sect. 5 and used to derive a corollary demonstrating the paucity of non-quasi-stochastic objects. The reason for the contrast between the abundance of nonstochastic objects and the paucity of non-quasi-stochastic objects is that the complexity of stochastic objects is simple (i.e., of low complexity), while this is not true for quasi-stochastic objects (and the key Lemma 14, due to Gács et al. [6], fails for quasi-stochastic objects).

Let $\mathbb{N} := \{1, 2, \dots\}$ be the natural numbers, $\mathbb{N}_m := \{m, m + 1, \dots\}$ for each integer m (we will be mostly interested in $m = 0$), $[m] := \{1, \dots, m\}$ for each $m \in \mathbb{N}$, $\log$ be binary logarithm, C be plain Kolmogorov complexity [11, Sect. 1.1], K be prefix complexity [11, Sect. 4.3], and $\mathbf{m}$ be a fixed universal semimeasure on $\mathbb{N}$ [11, Sect. 4.2]. For a subset $A \subseteq \mathbb{N}$, we set $\mathbf{m}(A) := \sum_{a \in A} \mathbf{m}(a)$ (therefore, $\mathbf{m}(\{a\}) = \mathbf{m}(a)$ for all $a \in \mathbb{N}$). This makes $\mathbf{m}$ a measure, so let us also refer to it as the *universal measure*. The words

“increasing” and “decreasing” are used in the wide sense allowing intervals of constancy.

In this paper we take $\mathcal{X}$ to be our basic space of finite objects whose (α, β) -stochasticity we are interested in. In all formal statements $\mathcal{X}$ can be replaced by $\mathbb{N}$ (but in informal discussion using $\mathbb{N}$ in place of $\mathcal{X}$ can cause confusion, since natural numbers play many other roles in this paper). The basic space $\mathcal{X}$ can be any set in computable bijection with $\mathbb{N}$, such as the binary strings $\{0, 1\}^*$. We will often refer to elements of the basic space $\mathcal{X}$ as objects and will usually denote them by ω . The complexities $C(\omega)$ and $K(\omega)$ and the universal measure $\mathbf{m}(A)$ for $A \subseteq \mathcal{X}$ are defined, naturally, as the corresponding functions (C , K , or $\mathbf{m}$) applied to the images of ω or A under the computable bijection. It is also convenient to define, especially in the appendix, $[m] \subseteq \mathcal{X}$, where $m \in \mathbb{N}$, to be the image of $[m] \subseteq \mathbb{N}$ under the computable bijection; similarly, $\min A$ for $A \subseteq \mathcal{X}$ is the image of $\min A'$, where $A' \subseteq \mathbb{N}$ is the image of A . For concreteness, the reader can interpret $\mathcal{X}$ as $\mathbb{N}$ or $\{0, 1\}^*$ throughout the paper (and in the appendix the interpretation of $\mathcal{X}$ as $\mathbb{N}$ is particularly convenient).

For a brief summary of the main results about finitary randomness, including connections with universal p-values and e-values, see [13, Appendix 1].

2 Kolmogorov’s questions

This section is a historical interlude discussing the motivation for this paper. In the next paragraph I reproduce Kolmogorov’s questions posed in his talk on 26 November 1981 and briefly recorded by Shen [9, note 5.12].

A finite object ω is (α, β) -stochastic if there exists a finite set Ω such that

- $\omega \in \Omega$;
- $K(\Omega) \leq \alpha$;
- $K(\omega) \geq \log |\Omega| - \beta$.

Let $N(k, \alpha, \beta)$ be the number of (α, β) -stochastic objects of complexity at most k . Typical cases are $\alpha, \beta = \epsilon k$ and $\alpha, \beta = \sqrt{k}$. What can we say about the asymptotics of $N(k, \alpha, \beta)/2^k$ as $k \rightarrow \infty$? Is it true that $N(k, \alpha, \beta)/2^k \rightarrow 0$, if k increases and $\alpha/k, \beta/k \rightarrow 0$?

It is very likely that by K Kolmogorov meant plain Kolmogorov complexity, nowadays usually denoted by C . It is also likely that there is a typo in the definition of $N(k, \alpha, \beta)$ and in fact it means the number of (α, β) -nonstochastic objects of complexity at most k . Therefore, in this section we set

$$N(k, \alpha, \beta) := |D_k \setminus \mathcal{S}_{\alpha, \beta}|,$$

where

$$\begin{aligned} D_k &:= \{\omega \in \mathcal{X} \mid C(\omega) \leq k\}, \\ \mathcal{S}_{\alpha, \beta} &:= \{\omega \in \mathcal{X} \mid \exists \Omega \subseteq \mathcal{X} : \omega \in \Omega, C(\Omega) \leq \alpha, C(\omega) \geq \log |\Omega| - \beta\}; \end{aligned} \tag{1}$$

in this section we refer to elements of $\mathcal{S}_{\alpha,\beta}$ as (α, β) -stochastic (unlike in the rest of the paper, where stochasticity will be defined in terms of prefix complexity).

Remark 1. Reading “stochastic” literally, i.e., replacing $N(k, \alpha, \beta)$ by

$$|D_k \cap \mathcal{S}_{\alpha,\beta}| = |D_k| - N(k, \alpha, \beta),$$

makes both of Kolmogorov’s questions degenerate. Let c be a constant such that $C(\omega) \leq n + c$ for all n and all $\omega \in [2^n]$. For $n := k - c$, the set $\Omega := [2^n]$ is a subset of D_k with $C(\Omega) \leq \log k + O(1)$, and more than half of its elements satisfy $C(\omega) \geq n - 1 = \log |\Omega| - 1$. Hence $|D_k \cap \mathcal{S}_{\alpha,\beta}| \geq 2^{k-c-1}$ as soon as $\beta \geq 1$ and $\alpha \geq \log k + O(1)$, and the answer to the last question would be negative already in the typical case $\alpha = \beta = \sqrt{k}$. Moreover, by (2) below, in both typical cases $|D_k \cap \mathcal{S}_{\alpha,\beta}|/2^k$ differs by $o(1)$ from $|D_k|/2^k \in [2^{-c}, 2)$, which involves neither α nor β ; so the first question would not be about stochasticity at all. With “nonstochastic”, both questions become substantive: see Theorem 2 and Corollary 3.

Two other facts point the same way. Shen’s 1983 paper, presented to *Doklady* by Kolmogorov himself, bounds the number of nonstochastic objects [10, Theorem 3], and V’yugin, another participant in the seminar, recalls Kolmogorov asking in the talk whether nonstochastic strings exist and, if so, how many of them there are among the strings of a given length [15, Sect. 2]. (Another discrepancy between Shen’s record and V’yugin’s recollection is that V’yugin mentions lengths of strings while Shen’s record is structure-free, but contemporaneous records are usually more reliable.) Slips in the record are not surprising: its definition of stochasticity contains another one, $K(\Omega)$ in place of $\log |\Omega|$, flagged by the authors [9, note 5.12] and corrected above.

Kolmogorov used the condition $C(\omega) \geq \log |\Omega| - \beta$ rather than the more standard $C(\omega | \Omega) \geq \log |\Omega| - \beta$ or $K(\omega | \Omega) \geq \log |\Omega| - \beta$, which we will use in the following sections. The suggestion to measure the randomness deficiency of ω as an element of Ω by $\log |\Omega| - C(\omega | \Omega)$ rather than $\log |\Omega| - C(\omega)$ is due to Per Martin-Löf (personal communication), and it leads to an elegant theory: see, e.g., Propositions 11 and 13 in the summary [13, Appendix 1].

The following result answers Kolmogorov’s specific questions; its core part is given as (4) after its statement.

Theorem 2. *There is a positive constant c such that, for all integers $k \geq 2$ and $\alpha \geq 0$ and all $\beta \in [0, \infty]$,*

$$N(k, \alpha, \beta) \leq 2^{k-\alpha+c \log k} \tag{2}$$

and, provided $\alpha + K(\alpha) + c \leq k$,

$$N(k, \alpha, \beta) \geq 2^{k-\alpha-K(\alpha)-c}. \tag{3}$$

In particular,

$$N(k, \alpha, \beta) = 2^{k-\alpha+O(\log k)} \tag{4}$$

uniformly in $\beta \in [0, \infty]$ whenever $\alpha + K(\alpha) + c \leq k$. The form of (4) is familiar: the fraction of nonstochastic objects is about $2^{-\alpha}$.

The theorem will be proved in Appendix A.2, but now we will use it to answer Kolmogorov's questions. The following corollary uses $N(k, \alpha, \beta) = N(k, \lfloor \alpha \rfloor, \beta)$ for α that are not necessarily integer.

Corollary 3. *For Kolmogorov's two typical cases $\alpha, \beta = \epsilon k$ (with $\epsilon \in (0, 1)$ fixed) and $\alpha, \beta = \sqrt{k}$ we have, for large k ,*

$$\frac{N(k, \epsilon k, \epsilon k)}{2^k} = 2^{-\epsilon k + O(\log k)}, \quad \frac{N(k, \sqrt{k}, \sqrt{k})}{2^k} = 2^{-\sqrt{k} + O(\log k)},$$

so in both cases the ratio tends to 0 faster than any power of k .

Proof. In both cases $\alpha + K(\alpha) + c \leq k$ for large k , so (4) applies. $\square$

Kolmogorov's last question appears to contain another slip. A negative answer is provided already by a fixed (arbitrarily large) α . This can be seen both from (3) and directly, without using the theorem. Indeed, for a fixed α the family $\{\Omega \mid C(\Omega) \leq \alpha\}$ is finite (of size at most $2^{\alpha+1}$); let F_α be its union, a finite set. Every (α, β) -stochastic object is in F_α , whatever β may be, so $N(k, \alpha, \beta) \geq |D_k| - |F_\alpha| \geq 2^{k-O(1)} - O(1)$ and, therefore,

$$\liminf_{k \rightarrow \infty} \frac{N(k, \alpha, \beta)}{2^k} \geq c$$

for a positive constant c . Probably Kolmogorov meant the question “when do we have $N(k, \alpha, \beta)/2^k \rightarrow 0$?” According to (2), “ $\log k = o(\alpha)$ ” is a sufficient condition. This condition, however, is far from being necessary (e.g., by (2), it suffices to require $\alpha - c \log k \rightarrow \infty$).

To conclude this section, let us discuss in greater detail what appears to be the main difference of Kolmogorov's approach to stochasticity from the standard one relying on a given structure on the objects. Kolmogorov's definitions and questions are structure-free in that they only involve the complexities of the objects (the complexities can be, e.g., plain or prefix, conditional or unconditional) and other objects derived from the given objects (such as finite sets composed of the given objects). Complexities can be replaced by randomness deficiencies (which do not have to be defined via complexities); the point is that structure-free statements should involve only functions that are invariant, to within an additive constant, w.r.t. computable permutations of the objects. An example of a non-invariant function is provided by the length of a binary string. In this paper we will be interested only in structure-free statements. Examples are Theorem 2 and Corollary 3, and we will see many other examples in the following sections (such as Theorems 4, 7, and 8).

3 Main result

Starting from this section, we say that an object $\omega \in \mathcal{X}$ is (α, β) -stochastic, where α is a nonnegative integer and $\beta \in [0, \infty]$, if there exists a finite set

$\Omega \subseteq \mathcal{X}$ for which

- $\omega \in \Omega$;
- $K(\Omega) \leq \alpha$;
- $K(\omega \mid \Omega) \geq \log |\Omega| - \beta$.

Otherwise ω is (α, β) -nonstochastic. In our notation we follow Kolmogorov (Kolmogorov's setting is described in the previous section).

Let $\mathcal{N}_{\alpha, \beta} \subseteq \mathcal{X}$ be the set of all (α, β) -nonstochastic objects. We are interested in the asymptotics of $\mathbf{m}(\mathcal{N}_{\alpha, \beta})$ as $\alpha, \beta \rightarrow \infty$. Let

$$\tau(t) := \mathbf{m}(\{n \in \mathbb{N} \mid n > t\})$$

be the tail of $\mathbf{m}$; set $\tau(\infty) := 0$.

Theorem 4. *There exists $c > 1$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \leq 2^{2^\alpha}$,*

$$\tau(\alpha)/c \leq \mathbf{m}(\mathcal{N}_{\alpha, \beta}) \leq c\tau(\alpha). \quad (5)$$

The condition $\beta \leq 2^{2^\alpha}$ in the statement of the theorem can be replaced by $\beta \leq g(\alpha)$ for a fixed computable function g ("fixed" in the sense of c depending on g). Moreover, this condition is only needed for the lower bound. See Sect. A.3 for an even stronger version of Theorem 4 and its proof.

The theorem says that $\mathbf{m}(\mathcal{N}_{\alpha, \beta}) = \Theta(\tau(\alpha))$ under mild conditions on β . Informally, nonstochastic objects are generated with probability that decays extremely slowly. The function $n \mapsto \tau(n)$ monotonically decreases to 0 but does so more slowly than any computable function monotonically decreasing to 0, such as $1/\log \log n$ (and we can take any number of logarithms). Let us state this well-known observation formally.

Lemma 5. *Let $f : \mathbb{N} \rightarrow (0, \infty)$ be a computable decreasing function such that $f(n) \rightarrow 0$ as $n \rightarrow \infty$. Then $\tau(n)/f(n) \rightarrow \infty$ as $n \rightarrow \infty$.*

Proof. Redefine $f := \min(f, 1)$. Set $h(t) := \sqrt{f(t)}$ for $t \in \mathbb{N}$ and $h(0) := 1$. Since f is decreasing and $f(t) \rightarrow 0$, $P(n) := h(n-1) - h(n)$, $n \in \mathbb{N}$, is a computable probability measure on $\mathbb{N}$ with tails $P(\{n \mid n > t\}) = h(t)$ for all $t \geq 0$. The universality of $\mathbf{m}$ gives $\mathbf{m} \geq cP$ for some $c > 0$, and so, as $t \rightarrow \infty$,

$$\frac{\tau(t)}{f(t)} \geq \frac{ch(t)}{f(t)} = \frac{c}{\sqrt{f(t)}} \rightarrow \infty. \quad \square$$

Now let us briefly discuss relations between τ and other well-known functions. Let

$$\sigma(t) := \mathbf{m}(\{n \mid K(n) > t\}); \quad (6)$$

we will need this function in the following section. In Lemma 12 in Sect. A.1 we will see that $\sigma = \Theta(\tau)$. (This is a well-known relation, but I have not seen it stated explicitly in print.)

Another interesting connection is with the version of the busy beaver function as defined by Chaitin [2, Definition 5.3] (see also [11, Sect. 1.2.2]). However, we only prove a trivial one-sided relation (an optimal two-sided relation, due to Andreev, is stated after the proof). Let us set

$$\mu(t) := \min\{K(n) \mid n > t\},$$

so that μ is, essentially, the inverse to the busy beaver function.

Lemma 6. *For some constant $c > 0$,*

$$\tau \geq c2^{-\mu}.$$

Proof. The inequality follows from

$$\tau(t) = \mathbf{m}(\{n \mid n > t\}) \geq \mathbf{m}(n^*) \geq c2^{-K(n^*)} = c2^{-\mu(t)},$$

where $n^* > t$ is any number that achieves $K(n^*) = \mu(t)$. $\square$

Lemma 6 is the trivial part of Andreev's [1, Theorem 1] two-sided estimate

$$\mu - O(\log \mu) \leq -\log \tau \leq \mu + O(1).$$

(Andreev's result uses inverse functions to μ and $-\log \tau$, and in [1, Theorem 2] he shows its optimality crediting Gács [5] with part of his Theorem 2).

4 Nonstochastic objects of high and low complexity

Theorem 4 measures the whole of $\mathcal{N}_{\alpha,\beta}$. In view of Kolmogorov's questions discussed in Sect. 2, it is natural to ask how that mass is spread over the complexities of the objects carrying it. The first theorem of this section answers the question for the mass of a range (k, ∞) : requiring $K(\omega) > k$ in addition to nonstochasticity simply replaces α by $\max(\alpha, k)$, with no condition on k whatsoever. Theorem 7 below therefore contains Theorem 4, which is its special case for $k := 0$.

Theorem 7. *There exists $c > 1$ such that, for all α , all $\beta \leq 2^{2^\alpha}$, and all $k \in \mathbb{N}_0$,*

$$\tau(\max(\alpha, k))/c \leq \mathbf{m}(\mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) > k\}) \leq c\tau(\max(\alpha, k)). \quad (7)$$

Only the lower bound in (7) uses a condition on β .

We can see from (7) that α and k do not interact, and the larger of the two wins outright. Suppose $k \geq \alpha$ (the case $k < \alpha$ is almost vacuous, since every nonstochastic object satisfies $K(\omega) \geq \alpha - O(1)$). Comparing (7) with $\sigma(k) = \Theta(\tau(k))$ (see (12)) shows that a significant part of the mass that $\mathbf{m}$ assigns to the objects of complexity exceeding k is assigned to nonstochastic ones. Theorem 4

says that nonstochastic objects are generated with a probability that decays extremely slowly; Theorem 7 adds that, above the threshold α , nonstochastic objects are not merely plentiful: above any level $k \geq \alpha$ they carry a constant fraction of the universal measure.

Theorem 7 describes the mass of the nonstochastic objects above a complexity level k . The following theorem describes the mass below it. Perhaps this result is less interesting, and its statement is more complicated, but it answers a natural question. Set

$$\mathcal{L}_{\alpha,\beta,k} := \mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) \leq k\}.$$

The picture here is, in important respects, the opposite of Theorem 7.

Theorem 8. *There exists $c > 0$ such that, for all integers $\alpha \geq 0$ and $k \geq 1$ and all $\beta \in [0, \infty]$,*

$$\mathcal{L}_{\alpha,\beta,k} = \emptyset \quad \text{if } k \leq \alpha - c; \quad (8)$$

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \leq 2^{-\alpha+K(\alpha)+\log k+c} \quad \text{if } \beta \geq K(\alpha) + c; \quad (9)$$

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \geq 2^{-\alpha-K(\alpha)-c} \quad \text{if } k \geq \alpha + K(\alpha) + c. \quad (10)$$

In particular, whenever $\beta \geq K(\alpha) + c$ and $k \geq \alpha + K(\alpha) + c$,

$$2^{-\alpha-K(\alpha)-c} \leq \mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \leq 2^{-\alpha+K(\alpha)+\log k+c}$$

and, therefore,

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) = 2^{-\alpha+O(\log k)}.$$

Since $\tau(\alpha) \geq c2^{-K(\alpha+1)}$ for some constant $c > 0$, $\mathbf{m}(\mathcal{N}_{\alpha,\beta}) = \Theta(\tau(\alpha))$ is only polynomially small in α , whereas, by (9) (assuming $\beta \geq K(\alpha) + c$), the part of the mass $\mathbf{m}(\mathcal{N}_{\alpha,\beta})$ below complexity level k is exponentially small as long as $\log k = o(\alpha)$; the dominant term $2^{-\alpha}$ is familiar from Sect. 2. Therefore, almost all of the mass of the nonstochastic objects lies above every such level; this agrees with Theorem 7, according to which the mass above k is $\Theta(\tau(\max(\alpha, k))) = \Theta(\tau(\alpha))$ whenever $k \leq g(\alpha)$ for a computable g .

See Sects. A.4–A.5 for the proofs of Theorems 7 and 8.

5 Probability measures as models

A popular alternative to Kolmogorov’s definition of stochastic objects uses probability measures in place of finite sets as models (see, e.g., [10] or [6, Sect. V]). For binary strings of length n , the two classes of models lead to the same notion of stochasticity to within $O(\log n)$ ([10, Theorem 1], [6, Proposition I.2]). In our structure-free setting there is no n , and changing the class of models has a dramatic effect on the universal measure of the nonstochastic objects. In this section we only consider the simplest probability measures: a probability measure Q on $\mathcal{X}$ is *elementary* if it has a finite support and takes only rational values, as in [10]. (Our discussion is also applicable to general computable

probability measures Q as models, as in [6, Sect. V].) Elementary probability measures are finite objects, and so $K(Q)$ and $K(\omega \mid Q)$ are defined.

An object $\omega \in \mathcal{X}$ is (α, β) -quasi-stochastic if there exists an elementary probability measure Q such that

- $K(Q) \leq \alpha$;
- $Q(\{\omega\}) > 0$ and $K(\omega \mid Q) \geq -\log Q(\omega) - \beta$.

Let $\mathcal{N}_{\alpha, \beta}^q \subseteq \mathcal{X}$ be the set of all objects that are not (α, β) -quasi-stochastic (or (α, β) -non-quasi-stochastic). Regarding a finite set Ω as the uniform probability measure on Ω , we can see that $\mathcal{N}_{\alpha+c, \beta+c}^q \subseteq \mathcal{N}_{\alpha, \beta}^q$ for some constant c . The following immediate corollary of a result by Epstein and Levin shows that, in terms of the universal measure, $\mathcal{N}_{\alpha, \beta}^q$ is much smaller than $\mathcal{N}_{\alpha, \beta}$.

Corollary 9. *There exists $c > 0$ such that, for all $\alpha \in \mathbb{N}_0$ and $\beta \geq 2^\alpha$,*

$$\mathbf{m}(\mathcal{N}_{\alpha, \beta}^q) \leq 2^{-\alpha+c \log(\alpha+2)}.$$

By Theorem 4, $\mathbf{m}(\mathcal{N}_{\alpha, \beta}) = \Theta(\tau(\alpha))$ is only polynomially small in α whenever $\beta \leq 2^{2^\alpha}$. Therefore, for $2^\alpha \leq \beta \leq 2^{2^\alpha}$, the contrast between the two classes of models is stark: with probability measures as models, non-quasi-stochastic objects are generated by the universal machine with an exponentially small probability, and the term $2^{-\alpha}$, familiar from Sect. 2 and from the results discussed in the introduction, reappears.

Corollary 9 is deduced from the following result by Epstein and Levin [4] (see [3, Lemma 10] for a published proof). Set

$$\Lambda(\omega) := \min_Q (K(Q) + 3 \log \max(-\log Q(\omega) - K(\omega \mid Q), 1)),$$

where Q ranges over the elementary probability measures, and let

$$I(\omega : \mathcal{H}) := K(\omega) - K(\omega \mid \mathcal{H})$$

be the amount of information that ω carries about the halting problem $\mathcal{H}$ (regarded as an infinite binary sequence; the nature of $\mathcal{H}$ is completely irrelevant for us). Epstein and Levin's result is that, for all $\omega \in \mathcal{X}$,

$$\Lambda(\omega) \leq I(\omega : \mathcal{H}) + 6K(I(\omega : \mathcal{H})) + O(1). \quad (11)$$

Proof of Corollary 9. Let $\omega \in \mathcal{N}_{\alpha, \beta}^q$. We have $\Lambda(\omega) > \alpha$, since otherwise there would be an elementary probability measure Q such that

$$K(Q) \leq \alpha \quad \text{and} \quad -\log Q(\omega) - K(\omega \mid Q) \leq 2^{\alpha/3} \leq \beta,$$

contradicting $\omega \in \mathcal{N}_{\alpha, \beta}^q$. We also have $K(I(\omega : \mathcal{H})) = O(\log(\alpha + 2))$ unless $I(\omega : \mathcal{H}) > \alpha$, and so (11) gives $I(\omega : \mathcal{H}) > \alpha - c \log(\alpha + 2)$ for some constant $c > 0$. On the other hand, by the Kraft inequality,

$$\sum_{\omega \in \mathcal{X}} \mathbf{m}(\omega) 2^{I(\omega : \mathcal{H})} = \Theta \left(\sum_{\omega \in \mathcal{X}} 2^{-K(\omega \mid \mathcal{H})} \right) = O(1)$$

(cf. [6, Sect. III.E.3]). By Markov's inequality,

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}^q) \leq \mathbf{m}(\{\omega \mid I(\omega : \mathcal{H}) > \alpha - c \log(\alpha + 2)\}) = O\left(2^{-\alpha + c \log(\alpha + 2)}\right). \quad \square$$

6 Conclusion

Let me summarize this paper's structure-free picture. The universal measure of (α, β) -nonstochastic objects hardly depends on β (unless β is huge, in a busy-beaver sense, as compared with α) and tends to 0 extremely slowly as $\alpha \rightarrow \infty$, more slowly than any decreasing computable function that tends to 0. Above any complexity level $k \geq \alpha$ a significant part of $\mathbf{m}$ is allocated to (α, β) -nonstochastic objects:

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) > k\}) = \Theta(\mathbf{m}(\{\omega \mid K(\omega) > k\}))$$

(provided β is computably bounded in α). The situation below a complexity level k is more complex.

The most glaring open question for nonstochastic objects is whether the addend $\log k$ in (9) is really necessary (preliminary results suggest that it is). A more detailed study of non-quasi-stochastic objects is also left for future work.

Acknowledgments

Many thanks to Sasha Shen for several discussions of Kolmogorov's definition of stochastic objects. I acknowledge the use of Claude Opus 5, Fable 5.1, and Opus 5.5 in exploring proof ideas, which I reviewed carefully, and in checking the paper. I take full responsibility for this paper's claims and statements, including mathematical statements and their proofs.

References

- [1] Mikhail Andreev. Busy beavers and Kolmogorov complexity. In Arnold Beckmann, Laurent Bienvenu, and Nataša Jonoska, editors, *Pursuit of the Universal, Proceedings of the Twelfth Conference on Computability in Europe, CiE 2016*, volume 9709 of *Lecture Notes in Computer Science*, pages 195–204, 2016. Also published as arXiv report (more accessible, and so all references in the paper are to it) [arXiv:1703.05170 \[cs.CC\]](#) (March 2017).
- [2] Gregory J. Chaitin. Information-theoretic limitations of formal systems. *Journal of the Association for Computing Machinery*, 21:403–424, 1974.
- [3] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67:7568–7578, 2021.
- [4] Samuel Epstein and Leonid A. Levin. On sets of high complexity strings. Technical Report [arXiv:1107.1458 \[cs.CC\]](#), [arXiv.org](#) e-Print archive, July

2011. Latest version (v7): July 2012; withdrawn in February 2014 and reposted as [arXiv:1403.4539 \[cs.CC\]](#) by Levin.
- [5] Péter Gács. On the relation between descriptive complexity and algorithmic probability. *Theoretical Computer Science*, 22:71–93, 1983.
 - [6] Péter Gács, John T. Tromp, and Paul Vitányi. Algorithmic statistics. *IEEE Transactions on Information Theory*, 47:2443–2463, 2001.
 - [7] Leonid A. Levin. Occam bound on lowest complexity of elements. *Annals of Pure and Applied Logic*, 167:897–900, 2016.
 - [8] Andrei A. Muchnik, Alexei L. Semenov, and Vladimir A. Uspensky. Mathematical metaphysics of randomness. *Theoretical Computer Science*, 207:263–317, 1998.
 - [9] Alexey Semenov, Alexander Shen, and Nikolai Vereshchagin. Kolmogorov’s last discovery? (Kolmogorov and algorithmic statistics). *Theory of Probability and its Applications*, 68:582–606, 2024. Also published as [arXiv:2303.13185 \[math.LO\]](#), October 2023.
 - [10] Alexander Shen. The concept of Kolmogorov (α, β) -stochasticity and its properties. *Soviet Mathematics Doklady*, 28:295–299, 1983.
 - [11] Alexander Shen, Vladimir A. Uspensky, and Nikolai Vereshchagin. *Kolmogorov Complexity and Algorithmic Randomness*. American Mathematical Society, Providence, RI, 2017.
 - [12] Nikolai Vereshchagin and Alexander Shen. Algorithmic statistics: forty years later. In Adam Day, Michael Fellows, Noam Greenberg, Bakhadyr Khoussainov, Alexander Melnikov, and Frances Rosamond, editors, *Computability and Complexity: Essays Dedicated to Rodney G. Downey on the Occasion of His 60th Birthday*, volume 10010 of *Lecture Notes in Computer Science*, pages 669–737. Springer, Cham, 2017.
 - [13] Vladimir Vovk. Testing exchangeability in the batch mode with e-values and Markov alternatives. *Machine Learning*, 114:article 99, 2025.
 - [14] Vladimir V. V’yugin. On nonstochastic objects. *Problems of Information Transmission*, 21:77–83, 1985.
 - [15] Vladimir V. V’yugin. Kolmogorov complexity in the USSR (1975–1982): Isolation and its end. *Information and Computation*, 307:article 105374, 2025. Also published as [arXiv:1907.05056 \[cs.GL\]](#), July 2019.

A Proofs

In this appendix we will prove all the theorems stated in this paper. In the proofs, $c_1, c_2, \dots$ will stand for various positive constants.

The techniques used in the proofs of Theorems 2, 4, 7, and 8 are either well known or trivial. For Theorems 2 and 8,

- the upper bounds rely on the standard technique of “bounded complexity lists”; see, e.g., [6, Sect. III.E] and [12, Sect. 4];
- the lower bounds involve the technique of the “first uncovered object”, i.e., the smallest (α, ∞) -nonstochastic object; this is essentially Shen’s trick in [10, Theorem 2(b)].

For Theorems 4 and 7,

- the upper bounds are trivial;
- the lower bounds rely on the observation that the complexity of stochastic objects is simple (while the universal measure may well generate objects of complex complexity); see [6, Lemmas III.8 and III.10].

In principle, as Theorem 4 is a special case of Theorem 7, we only need to prove the latter. However, Theorem 4 appears to be more important (and I regard it as the main result of this paper), so its shorter proof will be given separately (and then used in the proof of Theorem 7).

A.1 The tail of the universal measure

This preliminary section explores the function τ , which plays such a prominent role in Theorems 4 and 7. Some of the proofs employ the great stability of τ ; it hardly changes over huge regions of its domain. We start with the trivial observation that the composition of a computable Markov kernel with the universal measure is dominated by the universal measure.

Lemma 10. *Let $(\kappa_n)_{n \in \mathbb{N}}$ be probability measures on $\mathbb{N}$ such that the mapping $(n, i) \in \mathbb{N}^2 \mapsto \kappa_n(\{i\})$ is computable. There is $c_\kappa > 0$ such that, for all $A \subseteq \mathbb{N}$,*

$$\mathbf{m}(A) \geq c_\kappa \sum_{n \in \mathbb{N}} \mathbf{m}(n) \kappa_n(A).$$

Proof. The function $i \mapsto \sum_n \mathbf{m}(n) \kappa_n(\{i\})$ is a lower semicomputable semimeasure, and the universality of $\mathbf{m}$ gives $\mathbf{m}(i) \geq c_\kappa \sum_n \mathbf{m}(n) \kappa_n(i)$ for all i . It remains to sum over $i \in A$. $\square$

It is clear that Lemma 10 is also applicable to families $(\kappa_n)_{n \in \mathbb{N}}$ of probability measures on $\mathcal{X}$.

Our first application of Lemma 10 is to demonstrate the great stability of τ .

Lemma 11. *Let $g : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ be a computable increasing function. There is $c_g > 0$ such that $\tau(g(t)) \geq c_g \tau(t)$ for all $t \in \mathbb{N}_0$.*

Proof. Applying Lemma 10 to κ_n set to $\delta_{g(n)+1}$, the point mass at $g(n) + 1$, we have

$$\begin{aligned}\tau(g(t)) &= \mathbf{m}(\{i \mid i > g(t)\}) \geq c_\kappa \sum_{n \in \mathbb{N}} \mathbf{m}(n) 1_{\{g(n) \geq g(t)\}} \\ &\geq c_\kappa \sum_{n \in \mathbb{N}} \mathbf{m}(n) 1_{\{n \geq t\}} = c_\kappa \sum_{n \geq t} \mathbf{m}(n) \geq c_\kappa \tau(t). \quad \square\end{aligned}$$

To conclude this section, let us check the relation between τ and σ mentioned earlier (right after the definition (6)).

Lemma 12. *For some constant $c_1 > 1$,*

$$\tau/c_1 \leq \sigma \leq c_1 \tau. \quad (12)$$

Proof. The lower bound in (12) follows from Lemma 10 applied to $A := \{n \mid K(n) > t\}$ and the uniform probability measures κ_k on $[2^k]$. At most 2^t numbers n have $K(n) \leq t$, so for $k \geq t + 1$ we have $\kappa_k(A) \geq 1/2$; therefore,

$$\sigma(t) = \mathbf{m}(\{n \mid K(n) > t\}) = \mathbf{m}(A) \geq c_2 \sum_{k=t+1}^{\infty} \frac{\mathbf{m}(k)}{2} = \frac{c_2}{2} \tau(t).$$

The upper bound follows from

$$\begin{aligned}\sigma(t) &= \mathbf{m}(\{n \mid K(n) > t\}) \leq \mathbf{m}(\{n \mid 2 \log n > t\}) \\ &= \mathbf{m}(\{n \mid n > 2^{t/2}\}) \leq \mathbf{m}(\{n \mid n > t\}) = \tau(t),\end{aligned}$$

where the inequalities in this chain hold for all $t \geq c_3$ (smaller t are covered by $\sigma \leq 1$ and $\tau(t) > 0$). $\square$

A.2 Proof of Theorem 2

We prove the two parts of Theorem 2, (2) and (3), separately, starting with (2). Since $\mathcal{S}_{\alpha, \beta}$ is increasing in (both α and) β , $N(k, \alpha, \beta)$ is decreasing in β . Therefore, it suffices to prove (2) for $\beta = 0$ and (3) for $\beta = \infty$.

Proof of (2). This proof introduces the technique of “bounded complexity lists” in a simple setting. Fix a jointly computable way of enumerating all elements of D_k (defined by (1)) given k (running the optimal plain decoder on all inputs of length at most k in parallel and listing the outputs in the order in which they first appear). Cutting the enumeration into consecutive blocks of length 2^{k-a} and taking the block containing ω means, intuitively, spending about a bits on the block index and leaving $k - a$ bits of noise inside the block.

The idea behind the technique is that the completed blocks consist, roughly, of stochastic objects. Namely, there is a constant c_4 such that the following holds for each integer $k \geq 0$ (the *level*) and each $a \in \{0, \dots, k\}$ (the *parameter*). Let

$\omega \in D_k$ lie in a complete block of size 2^{k-a} (i.e., a block that the enumeration of D_k lists in full). Then there is a finite $\Omega \ni \omega$ with

$$\log |\Omega| = k - a, \quad C(\Omega) \leq a + K(k) + c_4. \quad (13)$$

Indeed, let Ω be the block containing ω (so that the first equality in (13) follows from $|\Omega| = 2^{k-a}$). As the block is complete, Ω is computable from k , a , and the ordinal number $i < 2^{a+1}$ of the block. Its description may consist of a self-delimiting description of k followed by exactly $a+1$ further bits for i (those further bits also determining a).

Choose $k \in \mathbb{N}_2$ and put $a := \alpha - \lfloor c_5 \log k \rfloor$ for a large c_5 as compared with c_4 ; make c (in Theorem 2) large as compared with c_5 . If $a < 0$ then $\alpha < \lfloor c_5 \log k \rfloor$, and the asserted bound is at least $2^{k+\log k} \geq 2^{k+1} > |D_k|$, so there is nothing to prove; assume $a \geq 0$.

Let $\omega \in D_k$ and $k' := C(\omega) \leq k$. If $k' < a$, apply (13) with the level k' and the parameter k' : the block size is 1, so $\Omega = \{\omega\}$,

$$C(\Omega) \leq k' + K(k') + c_4 \leq \alpha,$$

and $\log |\Omega| = 0 \leq C(\omega)$; therefore, ω is $(\alpha, 0)$ -stochastic. If $k' \geq a$ and the block of ω at level k' with parameter a is complete, (13) gives

$$C(\Omega) \leq a + K(k') + c_4 \leq \alpha$$

and $\log |\Omega| = k' - a \leq k' = C(\omega)$, so again ω is $(\alpha, 0)$ -stochastic.

Hence every $(\alpha, 0)$ -nonstochastic $\omega \in D_k$ satisfies $k' = C(\omega) \geq a$ and lies in the last, incomplete block at its own level k' , and there are fewer than $2^{k'-a}$ such ω . Summing over $k' \in \{a, \dots, k\}$,

$$N(k, \alpha, 0) \leq \sum_{k'=a}^k 2^{k'-a} < 2^{k-a+1} = 2^{k-\alpha+\lfloor c_5 \log k \rfloor+1} \leq 2^{k-\alpha+c \log k}. \quad \square$$

Another proof technique that we need is the “first uncovered object”. The next lemma presents this technique both for plain complexity C (which we need now) and for prefix complexity K (which we will need later).

Lemma 13. *For integers $\alpha \geq 0$, let*

$$F_\alpha^C := \cup\{\Omega \mid C(\Omega) \leq \alpha\}, \quad \text{and} \quad F_\alpha^K := \cup\{\Omega \mid K(\Omega) \leq \alpha\},$$

the unions being over the finite $\Omega \subseteq \mathcal{X}$. Set

$$\omega_\alpha^{*C} := \min(\mathcal{X} \setminus F_\alpha^C) \quad \text{and} \quad \omega_\alpha^{*K} := \min(\mathcal{X} \setminus F_\alpha^K).$$

There is a constant c_6 such that, for all α ,

$$K(\omega_\alpha^{*C}) \leq \alpha + K(\alpha) + c_6 \quad \text{and} \quad K(\omega_\alpha^{*K}) \leq \alpha + K(\alpha) + c_6.$$

Proof. The proofs for ω_α^{*C} and ω_α^{*K} are identical; for concreteness, let us talk about plain complexity C . Since F_α^C is a finite union of finite sets, ω_α^{*C} is well defined. Let $M_\alpha < 2^{\alpha+1}$ be the number of finite sets Ω with $C(\Omega) \leq \alpha$. The following is a self-delimiting description of ω_α^{*C} : a shortest self-delimiting description of α followed by the binary representation of M_α padded with leading 0s to exactly $\alpha + 1$ bits. This gives $K(\omega_\alpha^{*C}) \leq K(\alpha) + \alpha + 1 + c_7$. $\square$

Since ω_α^{*C} lies in no finite set Ω with $C(\Omega) \leq \alpha$, it is (α, β) -nonstochastic for every β , including $\beta = \infty$: $\omega_\alpha^{*C} \notin \mathcal{S}_{\alpha, \infty}$.

Proof of (3). Fix a computable bijection $\langle \cdot, \cdot \rangle : \mathcal{X} \times \{0, 1\}^* \rightarrow \mathcal{X}$ and let $\pi(\langle \nu, q \rangle) := \nu$, applied element-wise to finite subsets of $\mathcal{X}$. Let c_8 be such that $C(\pi(\Omega)) \leq C(\Omega) + c_8$ for every finite Ω . Let $\omega^* := \omega_{\alpha+c_8}^{*C}$ be the first uncovered object of Lemma 13 for plain complexity at level $\alpha + c_8$. Thus ω^* lies in no finite Ω with $C(\Omega) \leq \alpha + c_8$, and

$$K(\omega^*) \leq \alpha + c_8 + K(\alpha + c_8) + c_6 \leq \alpha + K(\alpha) + c_9.$$

Let $\Delta := k - \alpha - K(\alpha) - c_{10}$, where c_{10} is a large integer (see below), and, in accordance with the statement of the theorem, assume $\Delta \geq 0$. For $q \in \{0, 1\}^\Delta$ put $\omega_q := \langle \omega^*, q \rangle$; these 2^Δ objects are pairwise distinct. Since a shortest self-delimiting description of ω^* followed by q is a plain description of ω_q ,

$$\begin{aligned} C(\omega_q) &\leq K(\omega^*) + \Delta + c_{11} \leq \alpha + K(\alpha) + \Delta + c_9 + c_{11} \\ &= k - c_{10} + c_9 + c_{11} \leq k, \end{aligned}$$

provided $c_{10} \geq c_9 + c_{11}$; thus $\omega_q \in D_k$. Suppose some finite $\Omega \ni \omega_q$ satisfies $C(\Omega) \leq \alpha$. Then $\pi(\Omega) \ni \omega^*$ and $C(\pi(\Omega)) \leq \alpha + c_8$, which contradicts the choice of ω^* . Hence $\omega_q \notin \mathcal{S}_{\alpha, \infty}$ for all q , and $N(k, \alpha, \infty) \geq 2^\Delta$. $\square$

The “in particular” bit after the theorem follows from $K(\alpha) \leq 2 \log(\alpha + 2) + c \leq 2 \log(k + 2) + c$.

A.3 Proof of Theorem 4

We will prove Theorem 4 in the form

$$\tau(\max(\alpha, \beta))/c_{12} \leq \mathbf{m}(\mathcal{N}_{\alpha, \beta}) \leq c_{12}\tau(\alpha). \quad (14)$$

This implies the stronger version of Theorem 4 mentioned right after its statement. Indeed, let g be a computable function; we assume, without loss of generality, that g is increasing and $g(n) \geq n$ for all n in its domain. Lemma 11 gives

$$\tau(\max(\alpha, \beta)) \geq \tau(g(\alpha)) \geq c_g \tau(\alpha).$$

It remains to combine this with (14).

Theorem 4 in the form (14) continues to hold when prefix complexity K is replaced by plain complexity C . Let us call the statement about C the *C-version*

of Theorem 4; let us also apply the expression “ C -version” in the analogous sense to other statements in the rest of this section. The C -version of Theorem 4 will be treated at the end of the section (with one exception, where the K -version is messy while the C -version is almost trivial).

Let us start with the easy upper bound in (14). For some constant c_{13} ,

$$C(\{\omega\}) \leq C(\omega) + c_{13} \quad \text{and} \quad K(\{\omega\}) \leq K(\omega) + c_{13}. \quad (15)$$

Every $\omega \in \mathcal{N}_{\alpha,\beta}$ satisfies $K(\omega) > \alpha - c_{13}$ (otherwise $\Omega := \{\omega\}$ would have witnessed its $(\alpha, 0)$ -stochasticity). Therefore, every $\omega \in \mathcal{N}_{\alpha,0}$ satisfies $2 \log \omega \geq \alpha - c_{14}$, which implies

$$\omega \geq 2^{\alpha/2 - c_{14}/2} \geq \alpha + 1$$

for all $\alpha \geq c_{15}$. This implies

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) \leq \mathbf{m}(\mathcal{N}_{\alpha,0}) \leq \mathbf{m}(\{\alpha + 1, \alpha + 2, \dots\}) = \tau(\alpha)$$

from $\alpha = c_{15}$ on. (This is sufficient since the remaining finitely many α s can be absorbed into c_{12} .)

It remains to prove the lower bound. As mentioned earlier, the key observation (cf. [6, Lemmas III.8 and III.10]) is that the complexity of stochastic objects is simple.

Lemma 14. *Every (α, β) -stochastic $\omega \in \mathcal{X}$ satisfies*

$$K(K(\omega)) \leq \alpha + 2 \log(\alpha + \beta + 2) + c_{16}.$$

Proof. The proof uses the fact that $K(\omega) \approx \log |\Omega|$:

$$\begin{aligned} K(\omega) &\leq K(\omega \mid \Omega) + K(\Omega) + c_{17} \leq \log |\Omega| + \alpha + c_{18} \\ K(\omega) &\geq K(\omega \mid \Omega) - c_{19} \geq \log |\Omega| - \beta - c_{19}, \end{aligned} \quad (16)$$

where Ω witnesses the (α, β) -stochasticity of ω . Since $\log |\Omega|$ is computable given Ω ,

$$K(K(\omega)) \leq K(\Omega) + 2 \log(\alpha + \beta + 2) + c_{20} \leq \alpha + 2 \log(\alpha + \beta + 2) + c_{16}. \quad \square$$

Lemma 14 gives us the following sandwich, which is the key to the proof: for all $\alpha \in \mathbb{N}_0$ and $\beta \in [0, \infty)$,

$$\{\omega \mid K(K(\omega)) > t_{\alpha,\beta}\} \subseteq \mathcal{N}_{\alpha,\beta} \subseteq \{\omega \mid K(\omega) > \alpha - c_{13}\}, \quad (17)$$

where $t_{\alpha,\beta} := \lceil \alpha + 2 \log(\alpha + \beta + 2) \rceil + c_{16}$ and c_{13} is from (15). (The second inclusion was already used in the proof of the upper bound.)

The following two lemmas give us a handle on the first inclusion in (17). The first of them says that an object that is random in $[2^n]$ for a complex n has complex complexity.

Lemma 15. *There is a constant c_{21} such that the following holds for all integers $t \geq 0$ and all $n \in \mathbb{N}$ with $K(n) > 2t + c_{21}$: every $\omega \in [2^n]$ with $K(\omega) \geq n$ satisfies $K(K(\omega)) > t$.*

Proof sketch. The full proof will be given at the end of this section, but in this sketch we will only check the C -version of Lemma 15, which will be very easy. To make the C -version of Lemma 15 useful, we replace the condition $K(\omega) \geq n$ by $C(\omega) \geq n - 1$ (to ensure that there are many such ω).

Let $\omega \in [2^n]$ satisfy $C(\omega) \geq n - 1$. The n -bit binary representation (with leading 0s) of $\omega - 1$ describes ω , and so $C(\omega) \leq n + c_{22}$. The difference $n - C(\omega)$ takes one of finitely many values, and so $C(C(\omega)) \geq C(n) - c_{23} > 2t \geq t$ (for a suitable choice of the constants). $\square$

Lemma 16. *There is a constant $c_{24} > 0$ such that, for all integers $t \geq 0$,*

$$\mathbf{m}(\{\omega \mid K(K(\omega)) > t\}) \geq c_{24}\tau(t).$$

Proof. Let $B := \{\omega \mid K(K(\omega)) > t\}$. Apply Lemma 10 to the uniform probability measures κ_n on $[2^n]$. Suppose $K(n) > 2t + c_{21}$; at least half of the objects $\omega \in [2^n]$ satisfy $K(\omega) \geq n$ (this is true for any n since at most 2^{n-1} objects satisfy $K(\omega) \leq n - 1$), and those objects belong to B by Lemma 15. Therefore, $\kappa_n(B) \geq 1/2$ whenever $K(n) > 2t + c_{21}$, and

$$\mathbf{m}(B) \geq \frac{c_{25}}{2} \sum_{n: K(n) > 2t + c_{21}} \mathbf{m}(n) = \frac{c_{25}}{2} \sigma(2t + c_{21}) \geq c_{26}\tau(t)$$

by Lemmas 12 and 11. $\square$

To complete the proof of the lower bound, from (17) and Lemma 16 we derive

$$\mathbf{m}(\mathcal{N}_{\alpha,\beta}) \geq \mathbf{m}(\{\omega \mid K(K(\omega)) > t_{\alpha,\beta}\}) \geq c_{24}\tau(t_{\alpha,\beta}).$$

It remains to apply Lemma 11.

There is one gap in the proof of the K -version of the theorem: we only established the C -version of Lemma 15. Let us prove its K -version. Let $\omega \in [2^n]$ satisfy $K(\omega) \geq n$. A self-delimiting description of n followed by the n -bit binary representation (with leading 0s) of $\omega - 1$ describes ω , and so $K(\omega) \leq K(n) + n + c_{27}$. Therefore, $d := K(\omega) - n \in \{0, \dots, K(n) + c_{27}\}$. Since $n = K(\omega) - d$ is computable from $K(\omega)$ and d , we have

$$\begin{aligned} K(n) &\leq K(K(\omega)) + K(d) + c_{28} \leq K(K(\omega)) + 2\log(d + 2) + c_{29} \\ &\leq K(K(\omega)) + 2\log(K(n) + 2) + c_{30}. \end{aligned}$$

Let $v := K(n)$, so that

$$K(K(\omega)) \geq v - 2\log(v + 2) - c_{30}.$$

The function $v \mapsto v - 2\log(v + 2)$ is increasing over $[1, \infty)$. Therefore, for $v \geq 2t + c_{21} + 1$,

$$K(K(\omega)) \geq 2t + c_{21} + 1 - 2\log(2t + c_{21} + 3) - c_{30} > t,$$

where the last inequality holds for all $t \geq 0$ provided c_{21} is large enough.

The C -version of Theorem 4 in the form (14) (where ω is (α, β) -stochastic if $C(\Omega) \leq \alpha$ and $C(\omega \mid \Omega) \geq \log |\Omega| - \beta$ for some finite $\Omega \ni \omega$) is proved in the same way. The proof of the upper bound works verbatim. In the proof of Lemma 14 we replace the α in (16) by 2α , which gives the C -version of Lemma 14 identical to the K -version except that K is replaced by C and perhaps with a different c_{16} . This gives us the C -version of the sandwich (17). Lemma 15 was proved in its C -version with $C(\omega) \geq n - 1$ in place of $K(\omega) \geq n$. Therefore, the proof of Lemma 16 carries over with C in place of K except that $K(\omega) \geq n$ is replaced by $C(\omega) \geq n - 1$; in its last step we use $\{n \mid C(n) > 2t + c_{21}\} \supseteq \{n \mid K(n) > 3t + c_{31}\}$ together with Lemmas 12 and 11.

A.4 Proof of Theorem 7

Set

$$A := \mathcal{N}_{\alpha, \beta} \cap \{\omega \mid K(\omega) > k\}.$$

Let us check that, analogously to (14),

$$\tau(\max(\alpha, \beta, k))/c_{32} \leq \mathbf{m}(A) \leq c_{32}\tau(\max(\alpha, k)).$$

Under the condition $\beta \leq g(\alpha)$, Lemma 11 then gives

$$\mathbf{m}(A) \geq \tau(\max(\alpha, \beta, k))/c_{32} \geq \tau(\max(g(\alpha), k))/c_{32} \geq c_{33}\tau(\max(\alpha, k))$$

for some $c_{33} > 0$ (again assuming, without loss of generality, $g(n) \geq n$ for all n).

We start with the upper bound. Since τ is decreasing, $\tau(\max(\alpha, k)) = \min(\tau(\alpha), \tau(k))$, so it suffices to bound $\mathbf{m}(A)$ above by a multiple of each of $\tau(\alpha)$ and $\tau(k)$. From $A \subseteq \mathcal{N}_{\alpha, \beta}$ and the upper bound in (5) we get $\mathbf{m}(A) \leq c\tau(\alpha)$; this part of Theorem 4 needs no condition on β , as $\mathcal{N}_{\alpha, \beta} \subseteq \mathcal{N}_{\alpha, 0}$. And since $A \subseteq \{\omega \mid K(\omega) > k\}$, (12) gives $\mathbf{m}(A) \leq \sigma(k) \leq c_{34}\tau(k)$.

It remains to establish the lower bound. If $K(\omega) \leq k$, then $K(K(\omega)) < k + c_{35}$ for a constant c_{35} . In combination with (17) this gives

$$\begin{aligned} A &= \mathcal{N}_{\alpha, \beta} \cap \{\omega \mid K(\omega) > k\} \\ &\supseteq \{\omega \mid K(K(\omega)) > t_{\alpha, \beta}\} \cap \{\omega \mid K(K(\omega)) \geq k + c_{35}\} \\ &\supseteq \{\omega \mid K(K(\omega)) > t'\}, \end{aligned}$$

where $t' := \max(t_{\alpha, \beta}, k + c_{35})$. By Lemma 16, $\mathbf{m}(A) \geq c_{24}\tau(t')$. Now Lemma 11 gives $\mathbf{m}(A) \geq \tau(\max(\alpha, \beta, k))/c_{32}$ for a large enough c_{32} .

A.5 Proof of Theorem 8

Proof of (8). This part was established at the beginning of Sect. A.3 (see (15)): every $\omega \in \mathcal{N}_{\alpha, \beta}$ satisfies $K(\omega) \geq \alpha - c_{13}$. $\square$

Proof of (9). We continue using the technique of “bounded complexity lists”, this time for prefix complexity. Set $E_j := \{\omega \mid K(\omega) \leq j\}$. We will need Chaitin’s inequality

$$|E_j| \leq 2^{j-K(j)+c_{36}} \quad (18)$$

(see, e.g., [11, Theorem 64]).

Set $\mathcal{L}^j := \mathcal{N}_{\alpha,\beta} \cap \{\omega \mid K(\omega) = j\}$. We will show that, for every integer $j \geq 0$,

$$\mathbf{m}(\mathcal{L}^j) \leq 2^{-\alpha+K(\alpha)+c_{37}}; \quad (19)$$

summing (19) over $j \in \{0, \dots, k\}$ will then give

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \leq (k+1)2^{-\alpha+K(\alpha)+c_{37}} \leq k2^{-\alpha+K(\alpha)+c}.$$

Fix j and set $s := j - \alpha + K(\alpha) + c_{38}$, where $c_{38} \geq c_{13}$ is a large integer constant (see the proof of (8) for c_{13}). If $s < 0$, then $\mathcal{L}^j = \emptyset$ by the proof of (8); therefore, we can assume $s \geq 0$. Enumerate E_j given j as in the proof of (2) (run the optimal prefix decoder on all programs of length at most j in parallel and list the outputs in the order of their first appearance), and cut the list into consecutive blocks of 2^s elements. The number of complete blocks is at most $|E_j|/2^s \leq 2^b$, where, by (18),

$$b := j - K(j) + c_{36} - s = \alpha - K(\alpha) - K(j) + c_{36} - c_{38};$$

we assume that c_{36} and, therefore, b are integers.

Suppose a complete block Ω exists (so that $b \geq 0$). It can be described by a shortest self-delimiting description of j (which gives both j and its length $K(j)$), followed by a shortest self-delimiting description of α (which gives α and $K(\alpha)$), followed by exactly $b+1$ bits giving the ordinal number of Ω among the complete blocks. Therefore,

$$K(\Omega) \leq K(j) + K(\alpha) + b + c_{39} = \alpha + c_{36} - c_{38} + c_{39} \leq \alpha,$$

provided $c_{38} \geq c_{36} + c_{39}$. If $\omega \in \Omega$ and $K(\omega) = j$, then

$$\begin{aligned} K(\omega \mid \Omega) &\geq K(\omega) - K(\Omega) - c_{40} \geq K(\omega) - \alpha - c_{40} \\ &= \log |\Omega| - K(\alpha) - c_{38} - c_{40} \geq \log |\Omega| - \beta, \end{aligned}$$

provided $\beta \geq K(\alpha) + c$ with $c \geq c_{38} + c_{40}$; therefore, ω is (α, β) -stochastic. Consequently, $\mathcal{L}^j$ is contained in the part of the list following the last complete block, and this part has fewer than 2^s elements; therefore,

$$\mathbf{m}(\mathcal{L}^j) < c_{41}2^s2^{-j} = c_{41}2^{-\alpha+K(\alpha)+c_{38}}.$$

(If there are no complete blocks, the whole list has fewer than 2^s elements, and the same bound holds.)

To conclude the proof, let us check that there are no vicious circles in the choice of the constants c_i used in it. None of these constants are chosen depending on the values of other constants (they are *basic constants*) except for c_{38} (chosen depending on the basic constants only), c_{37} (chosen depending on the basic constants and c_{38}), and c (used in the statement of (9) and depending on all other constants in this proof). $\square$

Proof of (10). By Lemma 13, $\omega^* := \omega_\alpha^{*K} \in \mathcal{N}_{\alpha,\infty} \subseteq \mathcal{N}_{\alpha,\beta}$ and $K(\omega^*) \leq \alpha + K(\alpha) + c_6$. If $k \geq \alpha + K(\alpha) + c_6$, then $\omega^* \in \mathcal{L}_{\alpha,\beta,k}$ and

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \geq \mathbf{m}(\omega^*) \geq c_{42} 2^{-\alpha-K(\alpha)-c_6}. \quad \square$$

This completes the proof of the theorem. Notice that in this proof the lower bound in (10) is witnessed by a single object, the first object not covered by the simple models; the pairing used in the proof of (3) to fill a whole complexity level is not needed here.

The condition on β in (9) can be dropped at the cost of an additional term: for a constant c and all α , β , and $k \geq 1$,

$$\mathbf{m}(\mathcal{L}_{\alpha,\beta,k}) \leq k 2^{-\alpha+K(\alpha)+c} + c 2^{-\beta}.$$

Indeed, without the condition on β , a complete block Ω in the proof of (9) may also contain elements ω of complexity j with $K(\omega \mid \Omega) < \log |\Omega| - \beta$. There are fewer than $2^{s-\beta}$ of them in each complete block (there are fewer than $2^{s-\beta}$ self-delimiting descriptions of length less than $s - \beta$), and so fewer than $2^{b+s-\beta} = 2^{j-K(j)-\beta+c_{36}}$ in total. Their mass is at most $c_{43} 2^{-K(j)-\beta}$, and summing over j (using $\sum_j 2^{-K(j)} \leq 1$) adds at most $c_{43} 2^{-\beta}$ to the bound.