Kolmogorov Complexity with Error

Lance Fortnow
University of Chicago
fortnow@cs.uchicago.edu

Troy Lee
CWI and University of Amsterdam
tlee@cwi.nl

Nikolai Vereshchaigin
Moscow State University
kolya@ver.mccme.ru

September 5, 2004

Abstract

We introduce the study of Kolmogorov complexity with error. For a metric d, we define $C_a(x)$ to be the length of a shortest program p which prints a string y such that $d(x,y) \leq a$. We also study a conditional version of this measure $C_{a,b}(x|y)$ where the task is, given a string y' such that $d(y,y') \leq b$, print a string x' such that $d(x,x') \leq a$. This definition admits both a uniform, the *same* program should work given any y' such that $d(y,y') \leq b$, and nonuniform measures, where we take the length of a program for the worst case y'. We study the relation of these measures in the case where d is Hamming distance, and show an example where the uniform measure is exponentially larger than the nonuniform one. We also show an example where symmetry of information does not hold for complexity with error.

1 Introduction

Kolmogorov complexity measures the information content of a string typically by looking at the size of the smallest program generating that string. Suppose we received that string over a noisy or corrupted channel. Such a channel could change random bits of string, possibly increasing its Kolmogorov complexity without adding any real information.

In this paper we explore a variation of Kolmogorov complexity designed to help us measure information over a noisy channel. We define Kolmogorov complexity with error by defining the complexity of a string x with error a as the smallest program generating a string x' that differs from x in at most a bits. We give tight bounds (up to logarithmic factors) on the maximum complexity of such strings and also look at time-bounded variations.

We also look at conditional Kolmogorov complexity with errors. Traditional conditional Kolmogorov complexity looks at the smallest program that converts a string y to a string x. In our context both x and y could be corrupted. We want the smallest program that converts a string close to y to a string close to x. We consider two variations of this definition, a uniform version where we have a single program that that converts any y' close to y to a string x' close to x and a nonuniform version where the program can depend on y'. We show examples giving a large separation between the uniform and nonuniform definitions.

Finally we consider symmetry of information for Kolmogorov complexity with error. Traditionally the complexity a a pair (x,y) is roughly equal to the sum of the complexity of x and the complexity of y given x. We show that for any values of d and a the complexity of (x,y) with error

d is at most the sum of the complexity of x with error a and the complexity of converting a string y with $d - a$ error given x with a bits of error. We show the other direction fails in a strong sense, we do not get equality for any a.

2 Preliminaries

We use $|x|$ to denote the length of a string x, and $\|A\|$ to denote the cardinality of a set A. All logarithms are base 2.

We use $d_H(x, y)$ to denote the Hamming distance between two binary strings x, y, that is the number of bits on which they differ. For $x \in \{0,1\}^n$ we let $B_n(x, R)$ denote the set of n-bit strings within Hamming distance R from x, and $V(n, R) = \sum_{i=1}^{R} \binom{n}{i}$ denote the volume of a Hamming ball of radius R over n-bit strings. We will use the following approximation of $V(n, R)$ (see [1]) on several occasions.

Lemma 1. *Suppose that $0 < \lambda < 1/2$ and λn is an integer. Then*

$$\frac{2^{nH(\lambda)}}{\sqrt{8n\lambda(1-\lambda)}} \leq V(n, \lambda n) \leq 2^{nH(\lambda)}.$$

3 Defining Kolmogorov Complexity with Error

We consider several possible ways of defining Kolmogorov complexity with error. In this section we present these alternatives in order to evaluate their relative merits in the coming sections. First, we review the standard definition of Kolmogorov complexity. More details can be found in [5].

For a Turing machine T, the Kolmogorov complexity $C_T(x|y)$ of x given y is the length of a shortest program p such that $T(p, y) = x$. The theory of Kolmogorov complexity begins from the following invariance theorem: there is a universal machine U such that for any other Turing machine T, there exists a constant c_T such that $C_U(x|y) \leq C_T(x|y) + c_T$, for all x, y. We now fix such a U and drop the subscript.

Definition 1. *Let $d : \{0,1\}^n \to R$ be a metric, and $a \in R$. The complexity of x with error a, denoted $C_a(x)$ is $C_a(x) = \min_{x'}\{C(x') : d(x', x) \leq a\}$.*

We will also consider a time bounded version of this definition, $C_a^t(x) = \min_{x'}\{C^t(x') : d(x, x') \leq a\}$, where $C^t(x)$ is the length of a shortest program which prints x in less than $t(|x|)$ time steps.

A relative version of Kolmogorov complexity with error is defined by Impagliazzo, Shaltiel and Wigderson [3]. That is, they use the definition $C_\delta(x) = \min\{C(y) : d_H(x, y) \leq \delta|x|\}$. We prefer using absolute distance here as it behaves better with respect to concatenations of strings—using relative distance has the disadvantage of of severe nonmonotonicity over prefixes. Take, for example, $x \in \{0,1\}^n$ satisfying $C(x) \geq n$. Let $y = 0^{2n}$. Then $C_{1/3}(x) \geq n - \log V(n, n/3)$ while $C_{1/3}(xy) \leq \log n + O(1)$. Using absolute error we have that $C_a(xy) \geq C_a(x) - O(\log n)$, that is it only suffers from logarithmic dips as with standard definition.

Defining conditional complexity with error is somewhat more subtle. We introduce both uniform and nonuniform versions of conditional complexity with error.

Definition 2. *Uniform conditional complexity, denoted $C_{a,b}^u(x|y)$, is the length of a shortest program p such that,* for any *y' satisfying $d(y, y') \leq b$ it holds that $U(p, y')$ outputs a string whose Hamming distance from x is less than a.*

Definition 3. *Nonuniform conditional complexity, which we denote $C_{a,b}(x|y)$ is defined as $C_{a,b} = \min_{x'} \max_{y'} \{C(x|y) : d(x', x) \leq a$ and $d(y', y) \leq b\}$.*

In section 5 we study the difference between these two measures.

4 Strings of Maximal Complexity

One of the most famous applications of Kolmogorov complexity is the incompressibility method (see [5], Chapter 6). To prove there exists an object with a certain property, we consider an object with maximal Kolmogorov complexity and show that it could be compressed if it did not possess this property.

This method relies on a simple fact about strings of maximal complexity: for every length n, there is a string x of complexity at least n. This follows from simple counting. It is also easy to see that, up to an additive constant, every string has complexity at most its length. What is the behavior of maximal complexity strings in the error case?

Again by a counting argument, we see that for every n there is an x of length n with $C_a(x) \geq \log 2^n/V(n,a) = n - \log V(n,a)$. Upper bounding the complexity of strings in the error case requires a bit more work, and has a close connection with the construction of covering codes. A covering code $\mathcal{C}$ of radius a is a set of strings such that for every $x \in \{0,1\}^n$ there is an element $y \in \mathcal{C}$ such that $d_H(x,y) \leq a$. Thus an upper bound on the maximum complexity strings will be given by the existence of covering codes of small size. The following Lemma is well known in the covering code literature, (see [1] or [4]).

Lemma 2. *For any n and integer $R \leq n$, there exists a set $\mathcal{C} \subseteq \{0,1\}^n$ with the following properties:*

1. *$\|\mathcal{C}\| \leq n2^n/V(n,R)$*
2. *for every $x \in \{0,1\}^n$, there exists $c \in \mathcal{C}$ with $d_H(x,c) \leq R$*
3. *The set $\mathcal{C}$ can be enumerated in time poly(2^n)*

Proof. For the first two items we argue by the probabilistic method. The third item will be obtained by derandomizing this argument with the method of conditional probabilities.

Fix a point $x \in \{0,1\}^n$. We uniformly at random choose k distinct elements $x_1, \ldots, x_k$ of $\{0,1\}^n$. The probability P_x that x is not contained in $\cup_{i=1}^k B(x_i, R)$ is precisely

$$\begin{aligned} P_x &= \frac{\binom{2^n - V(n,R)}{k}}{\binom{2^n}{k}} && (1) \\ &\leq (1 - V(n,R)/2^n)^k \leq e^{-kV(n,R)/2^n} && (2) \end{aligned}$$

For the inequality we have used the fact that $e^{-z} \geq 1 - z$ for any $0 \leq z \leq 1$. Taking k to be $n2^n/V(n,R)$ makes this probability strictly less than 2^{-n}. Thus the probability of the union of the events P_x over $x \in \{0,1\}^n$ is, by the union bound, less than 1 and there exists a set of $n2^n/V(n,R)$ centers which cover $\{0,1\}^n$. This gives items 1 and 2.

For item 3 we now derandomize this argument. Let $t = n2^n/V(n,R)$ be the desired size of our covering, and let $x_1, x_2, \ldots, x_{2^n}$ be a list of $x \in \{0,1\}^n$ in lexicographical order. Roughly speaking, we will consider each x_i in turn and decide if the covering is better with or without it, given the partial covering selected from $x_1, \ldots, x_{i-1}$. Initially, we know that the probability that t randomly chosen points do not form a covering is less than one. We add points to our covering in such a way that this probability does not increase.

Say that we have considered the points $x_1, \dots, x_{i-1}$ and selected some subset of them X_{i-1} to be part of the covering, where $\|X_{i-1}\| = k$. At stage i we decide whether or not to include x_i in our partial covering. To do this we consider the two cases:

- We take $X_i := X_{i-1} \cup \{x_i\}$ and the remaining $t - k - 1$ elements of the covering are chosen uniformly from $x_{i+1}, \dots, x_{2^n}$.
- We take $X_i := X_{i-1}$, and the remaining $t - k$ elements of the covering are chosen uniformly from $x_{i+1}, \dots, x_{2^n}$.

Following equation 1 we can calculate the probability that our result is not a covering in these two cases. If the first is not larger than the second, we take x_i as part of our covering. As these are disjoint events, and we know that given the partial covering X_{i-1} the probability of not getting a covering when choosing $t - k$ elements at random is less than one, by an averaging argument, the probability of one of these events must also be less than one. □

Theorem 1. *For every n, a and $x \in \{0,1\}^n$, $C_a(x) \leq n - \log V(n,a) + O(\log n)$.*

Proof. By Lemma 2 we know that a covering code with radius a of cardinality less than $n2^n/V(n,a)$ exists. Let $\mathcal{C}$ be the lexicographically first such covering. Such a covering can be described by saying "look for the lexographically first covering over $\{0,1\}^n$ of radius a", and thus has a description of size $O(\log n)$. For any $x \in \{0,1\}^n$ there is an element $c \in \mathcal{C}$ such that $d_H(x,c) \leq a$. Once we know the covering, this element c can be described by index in the covering, of size $n + \log n - \log V(n,a)$. Thus the total description is of size $n - \log V(n,a) + O(\log n)$. □

One nice property of covering codes is that they behave very well under concatenation. Let $\mathcal{C}_1$ be a covering code of $\{0,1\}^{n_1}$ of radius R_1 and $\mathcal{C}_2$ be a covering code of $\{0,1\}^{n_2}$ of radius R_2. Now let $\mathcal{C} = \{cc' : c \in \mathcal{C}_1, c' \in \mathcal{C}_2\}$ be the set of all ordered concatenations of codewords from $\mathcal{C}_\infty$ with codewords from $\mathcal{C}_\in$. Then $\mathcal{C}$ is a covering code over $\{0,1\}^{n_1+n_2}$ of radius $R_1 + R_2$.

We can use this idea in combination with item 3 of Lemma 2 to efficiently construct near-optimal covering codes. This construction has already been used for a complexity-theoretic application in [2].

Theorem 2. *Let $0 < \lambda < 1/2$ and set $d = \lambda n$. There is a polynomial time bound $p(n)$ such that $C^p_d(x) \leq n - \log V(n,d) + O(n \log\log n / \log n)$ for every $x \in \{0,1\}^n$.*

Proof. We construct a covering code over $\{0,1\}^n$ with radius d such that the ith element of the covering can be generated in time polynomial in n. For some constant c, we let $\ell = c \log n$ and divide n into $\lceil n/\ell \rceil$ blocks of length ℓ. Let $r = \lambda \ell$. Now by item 3 of Lemma 2 we can in time polynomial in n construct a covering code over $\{0,1\}^\ell$ of radius r and of cardinality $\ell 2^\ell / V(\ell, r)$. Call this covering $\mathcal{C}_\ell$. Our covering code $\mathcal{C}$ over $\{0,1\}^n$ will be the set of codewords $\{c_1 c_2 \cdots c_{\lceil n/\ell \rceil} : c_i \in \mathcal{C}_\ell\}$. The size of this code will be:

$$\|\mathcal{C}\| \leq (2^{\ell - \log V(\ell,r) + \log \ell})^{n/\ell} = 2^{n - n/\ell \log V(\ell,r) + (n/\ell)\log \ell}.$$

As $r = \lambda \ell$ we can use the estimates of Lemma 1 to obtain:

$$n/\ell \log V(\ell, r) \geq nH(\lambda) - (n/2\ell)\log 8\ell\lambda(1-\lambda) \geq \log V(n,d) - O(n \log\log n / \log n).$$

Thus $\|\mathcal{C}\| \leq 2^{n - \log V(n,d) + O(n \log\log n/\log n)}$. □

5 Uniform vs. Nonuniform Conditional Complexity

In this section we show an example where the uniform version of conditional complexity can be exponentially larger than the nonuniform one. Our example will be for $C_{0,d}(x|x)$. Notice that this example is the error correction problem. Given some x' such that $d_H(x, x') \leq d$, we want to recover x exactly. The intuition behind the proof is the following: say we have some computable family S of Hamming balls of radius d, and let x be the center of one of these balls. Given any x' such that $d(x, x') \leq d$, there may be other centers of the family S which are also less than distance d from x'. Say there are t of them. Then x has a nonuniform description of size about $\log t$ by giving the index of x in the t balls which are of distance less than d from x'.

In the uniform case, on the other hand, our program can no longer be tailored for a particular x', it must work for any x' such that $d(x, x') \leq d$. That is, intuitively, the program must be able to distinguish the ball of x from any other ball intersecting the ball of x. To create a large difference between the nonuniform and uniform conditional complexity measures, therefore, we wish to construct a large family of Hamming balls, every two of which intersect, yet that no single point is contained in the intersection of too many balls. The next lemma shows the existence of such a family.

Lemma 3. *Given large enough length m of strings and $d < m/2$ satisfying the inequality*

$$m\big(2H(d/m) - 1 - H(1 - 2d/m)\big) \geq 4 \log m + 2 \tag{3}$$

there is a family of at least $N = 2^{m(1-H(d/m))}$ Hamming balls of radius d such that every two balls intersect but no string belongs to more than $2m^2$ balls.

Proof. The proof is by probabilistic arguments. Take m^2N independent random balls $B_1, \dots, B_{m^2N}$ of radius d. We will prove that with high probability at least N of them are different and satisfy the statement.

First we estimate the probability that of the m^2N centers at least N are distinct. For a fixed set U of size N the probability all m^2N centers fall in U is $(N2^{-m})^{m^2N}$. The number of different sets of size N is less than 2^{mN}, thus by a union bound we obtain the probability these m^2N centers lie in any set of size N is at most $(N2^{-m})^{m^2N}2^{mN}$. Taking the logarithm of this number and dividing it by Nm we obtain

$$m \log N - m^2 + 1 = -m^2H(d/m) + 1.$$

The inequality (3) implies that $H(d/m) > 1/2$ hence for large enough m the estimated probability is close to 0.

Estimate now the probability that there are two disjoint balls. Fix two indexes $i < j \leq mN$. If balls B_i, B_j are disjoint then the center x_j of B_j is at distance at least $2d$ from the center x_i of B_i. The latter means that x_j is at distance at most $m - 2d$ from the string $\bar{x}_i$, that is obtained from x_i by flipping all bits. The probability of this is equal to the ratio of cardinality of the ball of radius $m - 2d$ and 2^m:

$$\frac{V(m, m-2d)}{2^m} \leq \frac{2^{mH(1-2d/m)}}{2^m} = 2^{m(H(1-2d/m)-1)}.$$

Multiplying this probability by the number $N(N-1)/2$ of different pairs j, j we get less than

$$2^{m(H(1-2d/m)-1)+4\log m+2m(1-H(d/m))-1} \leq 1/2,$$

(the inequality is just a reformulation of (3)).

It remains to estimate the probability that there is a string that belongs to more than $2m^2$ balls B_i's. Fix x. For every i the probability that x lands in B_i is equal to $p = |B_i|/2^m$. Using Lemma 1 we estimate p as:

$$2^{mH(d/m)-m-(\log m)/2-O(1)} \leq p \leq 2^{mH(d/m)-m}.$$

So the average number of i with $x \in B_i$ is at most $pm^2N \leq 2^{mH(d/m)-m}(m^2N) = m^2$. By Chernoff inequality the probability that the number of i such that x lands in B_i exceeds twice the average is at most

$$e^{-pm^2N/2} \leq \exp(-2^{mH(d/m)-m-(\log m)/2-O(1)}m^2N/2) = \exp(-2^{(3\log m)/2-O(1)}) \ll 2^{-m}.$$

Thus even after multiplying it by 2^m the number of different x's we get a number close to 0.

Thus with positive probability every balls B_i, B_j with $i < j$ intersect, every string belongs to at most $2m^2$ B_i's and there are at least $2^{m(1-H(d/m))}$ different B_i's. □

THEOREM 3. *Given $d \leq n/2$ let m be the largest number $m \in \{2d, \dots, n\}$ satisfying* (3). *Then there is x of length n such that $C_{0,d}(x|x) \leq O(\log n)$ while $C^u_{0,d}(x|x) \geq m(1 - H(d/m))$.*

Ignoring additive terms of order $O(\log n)$ the number m is equal to $\min\{n, d/\alpha\}$ where $\alpha = 0.353...$ is the solution of the equation $2H(\alpha) = 1+H(1-2\alpha)$ and $C^u_{0,d}(x|x) \geq m(1-\max\{H(\alpha), H(d/n)\})$, where $H(\alpha) = 0.93....$

Proof. Given n, d find m and find the first family satisfying the lemma. The list of balls has complexity at most $C(d, n) = O(\log n)$. Append 0^{n-m} to all centers to get another family of balls, this time of strings of length n. Obviously the new family also satisfies the lemma. For the center x of every ball in the family we have $C_{0,d}(x|x) = O(\log n)$, as given any x' at distance at most d from x we can specify x by specifying its index among centers of the balls in the family containing x' in $\log(2m^2)$ bits and specifying the family itself in $O(\log n)$ bits.

It remains to show that there is a center x with $C^u_{0,d}(x|x) \geq m(1 - H(d/m))$. Assume the contrary and choose for every center x a program p_x of length less than such that $U(p, x') = x$ for every x' at distance at most d from x. As the number of different centers is strictly greater than the number of strings of length less than $m(1 - H(d/m))$, by the Pigeon Hole Principle there are different centers x_1, x_2 with $p_{x_1} = p_{x_2}$. However the balls with those centers intersect and there is x' at distance at most d both from x_1, x_2. Hence $x_1 = U(p, x') = x_2$, a contradiction. □

6 Symmetry of Information

The principle of symmetry of information, independently proven by Kolmogorov and Levin [6], is one of the most beautiful and useful theorems in Kolmogorov complexity. It states $C(x, y) = C(x)+C(y|x)+O(\log n)$ for any $x, y \in \{0, 1\}^n$. The direction $C(x, y) \leq C(x)+C(y|x)+O(\log n)$ is easy to see—given a program for x, and a program for y given x, and a way to tell these programs apart, we can print the pair x, y. The other direction of the inequality requires a clever proof.

Looking at symmetry of information in the error case, the easy direction is again easy: The inequality $C_d(x, y) \leq C_a(x) + C_{d-a,a}(y|x) + O(\log n)$ holds for any a — let p be a program of length $C_a(x)$ which prints a string x^* within Hamming distance a of x. Let q be a shortest program which, given x^*, prints a string y^* within Hamming distance $d - a$ of y. By definition, $C_{d-a,a}(y|x) = \min_{y'} \max_{x'} C(y'|x') \geq \min_{y'} C(y'|x^*) = |q|$. Now given p and q and a way to tell them apart, we can print the pair (x, y) within d errors.

For the converse direction we would like to have the statement

$$\text{For every } d \text{ there exists } a \text{ such that } C_d(x,y) \geq C_a(x) + C_{d-a,a}(y|x) - O(\log n). \qquad (*)$$

We do not expect this statement to hold for every a, as the shortest program for x, y will have a particular pattern of errors which might have to be respected in the programs for x and y given x. We now show, however, that even the formulation $(*)$ is too much to ask.

Theorem 4. *For every n there exists x, y such that $C_d(x,y) = n - \log V(n,d)$ yet $C_a(y) \geq n - \log V(n,a)$ and $C_{d-a,a}(x|y) \geq \log V(n,d+a) - \log V(n,d)V(n,d-a)$.*

Proof. Coverings will again play an important role in the proof. Let $\mathcal{C}$ be the lex first minimal size covering of radius d. Choose y of length n with $C(y) \geq n$, and let x be the lex least element of the covering within distance d of y. It must be the case that $C(x) \geq n - \log V(n,d) - 2\log n$, as otherwise we could obtain a shorter description of y. Notice that $C(x,y) \leq n - \log V(n,d)$, as the string xx is within distance d of xy, and can be described by giving a shortest program for x and a constant many more bits saying "repeat".

It now remains to lower bound $C_a(y) + C_{d-a,a}(x|y)$. As y has maximal complexity, for any $0 \leq a \leq d$ we have $C_a(y) \geq n - \log V(n,a)$, thus it remains to lower bound $C_{d-a,a}(x|y)$. We do this by using symmetry of information in the nonerror case.

Let d_1 be the Hamming distance between x and y, and let y' be obtained from y by changing a random set of a bits on which x and y agree. Thus $C(y'|y,x) = \log \binom{n-d_1}{a}$. Notice this means $d_H(x,y') = d_1 + a$, and so $C(y',x) \leq \log V(n, d_1 + a)$. We show the converse also holds. We show the following:

Claim 1. $C(y'|x) \geq \log V(n, d_1 + a)$

Proof. We use symmetry of information to turn the task of lower bounding $C(y'|x)$ into the task of upper bounding $C(y|y',x)$. This works as follows: by symmetry of information,

$$C(y'y|x) = C(y|x) + C(y'|y,x) = C(y'|x) + C(y|y',x).$$

We know that $C(y|x) \geq \log V(n,d_1)$ and $C(y'|y,x) \geq \log \binom{n-d_1}{a}$, thus we obtain $C(y,y'|x) = \log V(n,d_1) + \log \binom{n-d_1}{a}$. Now using the second part of the equality we have $C(y'|x) \geq \log V(n,d_1) + \log \binom{n-d_1}{a} - C(y|y',x)$. It thus remains to upper bound $C(y|y',x)$. The string y differs from y' on a bits out of the $d_1 + a$ bits on which y' and x differ. Thus $C(y|y',x) \leq \log \binom{d_1+a}{a}$. Hence, $C(y'|x) \geq \log V(n,d_1) + \log \binom{n-d_1}{a} - \log \binom{d_1+a}{a} = \log \binom{n}{d_1+a}$. □

It follows from the claim that $C(y'|x') \geq \log V(n, d_1 + a) - \log V(n, d-a)$, for any x' with $d_H(x,x') \leq d - a$.

We are now ready to bound $C_{d-a,a}(x|y) \geq C(x'|y') = C(x') + C(y'|x') - C(y') \geq \log V(n, d_1 + a) - \log V(n, d-a) - \log V(n,d)$. □

We can achieve the formulation $(*)$ within an error term of $\log V(2n,d)$.

Theorem 5. *For any $n, d < n$ and $x, y \in \{0,1\}^n$,*

$$C_d(x,y) \geq C_0(x) + C_{d,0}(y|x) - \log V(2n,d) - O(\log n)$$

Proof. First note that $C_d(x,y) \geq C(x,y) - \log V(2n,d) - O(\log n)$, as given the program p which prints x', y' such that $d_H(xy, x'y') \leq d$ we can then describe x, y with $\log V(2n,d)$ more bits. Now applying ordinary symmetry of information we obtain $C_d(x,y) \geq C(x) + C(y \mid x) - \log V(2n,d) - O(\log n)$. As $C(y|x) \geq C_{d,0}(y|x)$ we obtain $C_d(x,y) \geq C_0(x) + C_{d,0}(y|x) - \log V(2n,d)$. □

Acknowledgment

We thank Harry Buhrman for several useful discussions.

References

[1] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes*. North-Holland, Amsterdam, 1997.

[2] E. Dantsin, A. Goerdt, E. Hirsch, and U. Schöning. Deterministic algorithms for k-sat based on covering codes and local search. In *Proceedings of the 27th International Colloquium On Automata, Languages and Programming*, Lecture Notes in Computer Science, pages 236–247. Springer-Verlag, 2000.

[3] R. Impagliazzo, R. Shaltiel, and A. Wigderson. Extractors and pseudo-random generators with optimal seed length. In *Proceedings of the 32nd ACM Symposium on the Theory of Computing*, pages 1–10. ACM, 2000.

[4] M. Krivelevich, B. Sudakov, and V. Vu. Covering codes with improved density. In *IEEE Transactions on Information Theory*, volume 49, pages 1812–1815, 2003.

[5] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, second edition, 1997.

[6] A. Zvonkin and L. Levin. The complexity of finite objects and the algorithmic concepts of information and randomness. *Russian Mathematical Surveys*, 25:83–124, 1970.