SCIS 2015 The 32nd Symposium on
Cryptography and Information Security
Kokura, Japan, Jan. 20 - 23, 2015
The Institute of Electronics,
Information and Communication Engineers

An Operational Characterization of the Notion of Probability by Algorithmic Randomness and Its Application to Cryptography

Kohtaro Tadaki *

Abstract: The notion of probability plays an important role in almost all areas of science, including cryptography. In modern mathematics, however, probability theory means nothing other than measure theory, and the operational characterization of the notion of probability is not established yet. In this paper, based on the toolkit of algorithmic randomness we present an operational characterization of the notion of probability. Algorithmic randomness, also known as algorithmic information theory, is a field of mathematics which enables us to consider the randomness of an individual infinite sequence. We use the notion of Martin-Löf randomness with respect to Bernoulli measure to present the operational characterization. As the first step of the research of this line, in this paper we only consider the case of finite probability space, i.e., the case where the sample space of the underlying probability space is finite, for simplicity. This case is enough to study modern cryptography since all probability spaces which modern cryptography considers are finite. In the paper we make an application of our formalism to cryptography by presenting new equivalent characterizations of the notion of perfect secrecy in terms of our formalism.

Keywords: probability, algorithmic randomness, operational characterization, Martin-Löf randomness, Bernoulli measure, perfect secrecy

1 Introduction

The notion of probability plays an important role in almost all areas of science, including cryptography. In modern mathematics, however, probability theory means nothing other than measure theory, and an operational characterization of the notion of probability is not established yet.

In the past century, however, there was a comprehensive attempt to provide such a characterization. Namely, von Mises developed a mathematical theory of repetitive events which is aimed at reformulating the theory of probability and statistics based on an operational characterization of the notion of probability [17, 18]. In a series of comprehensive works which began in 1919, von Mises developed this theory and, in particular, introduced the notion of *collective* as a mathematical idealization of a long sequence of outcomes of experiments or observations repeated under a set of invariable conditions, such as the repeated tossing of a coin or of a pair of dice.

The collective plays a role as an operational characterization of the notion of probability, and is an infinite sequence of sample points in the sample space of a probability space. As the randomness property of the collective, von Mises assumes that all "reasonable" infinite subsequences of a collective satisfy the law of large numbers with the identical limit value, where the subsequences are selected using "acceptable selection rules." Wald [19, 20] later showed that for any countable collection of selection rules, there are sequences that are collectives in the sense of von Mises, but at the time it was unclear exactly what types of selection rules should be acceptable. There seemed to von Mises to be no canonical choice.

Later, with the development of computability theory and the introduction of generally accepted precise mathematical definitions of the notions of algorithm and computable function, Church [5] made the first explicit connection between computability theory and randomness by suggesting that a selection rule be considered acceptable if and only if it is computable. In 1939, however, Ville [16] revealed the defect of the notion of collective. Namely, he showed that for any countable collection of selection rules, there is a sequence that is random in the sense of von Mises but has properties that make it clearly nonrandom. (For the development of the theory of collectives from the point of view of the definition of randomness, see Downey and Hirschfeldt [6].)

In 1966, Martin-Löf [10] introduced the definition of random sequences, which is called *Martin-Löf randomness* nowadays, and plays a central role in the recent development of algorithmic randomness. At the same time, he introduced the notion of *Martin-Löf randomness with respect to Bernoulli measure* [10]. He then pointed out that this notion overcomes the defect of the collective in the sense of von Mises, and this can be regarded precisely as the collective which von Mises wanted to define. However, he did not develop probability theory based on Martin-Löf random sequence

* Research and Development Initiative, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan. E-mail: tadaki@kc.chuo-u.ac.jp WWW: http://www2.odn.ne.jp/tadaki/

with respect to Bernoulli measure.

Algorithmic randomness is a field of mathematics which studies the definitions of random sequences and their property [11, 6]. However, the research on algorithmic randomness would seem only interested in the notions of randomness and their property, and not seem to have tried to develop probability theory based Martin-Löf randomness with respect to Bernoulli measure in an operational manner so far.

In our former work [14] we started such an attempt. Namely, in the work [14] we presented an operational characterization of the notion of probability based on Martin-Löf randomness with respect to Bernoulli measure. As the first step of the research of this line, in the work [14] we only considered the case of finite probability space, i.e., the case where the sample space of the underlying probability space is finite, for simplicity.

In this paper we further develop and refine our theory of operational characterization of the notion of probability started by the work [14] *while reviewing the results of the work*. In this paper we consider the case of finite probability space, as with the work [14]. This case is enough to study modern cryptography since all probability spaces which modern cryptography considers are finite. In particular, we make an application of our theory to cryptography by presenting new equivalent characterizations of the notion of perfect secrecy in terms of our formalism.

Due to the 8-page limit, we omit all proofs of the new results. A full paper which describes all the proofs and other related results is in preparation.

2 Preliminaries

2.1 Basic Notation and Definitions

We start with some notation about numbers and strings which will be used in this paper. $\#S$ is the cardinality of S for any set S. $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ is the set of natural numbers, and $\mathbb{N}^+$ is the set of positive integers. $\mathbb{Q}$ is the set of rationals, and $\mathbb{R}$ is the set of reals.

An *alphabet* is a nonempty finite set. We suppose that any alphabet which we consider in this paper has at least two elements. Let Ω be an alphabet. A *finite string over* Ω is a finite sequence of elements from the alphabet Ω. We denote by Ω^* the set of all finite strings over Ω, which contains the *empty string* denoted by λ. We denote by Ω^+ the set $\Omega - \{\lambda\}$. For any $\sigma \in \Omega^*$, $|\sigma|$ is the *length* of σ. Therefore $|\lambda| = 0$. A subset S of Ω^* is called *prefix-free* if no string in S is a prefix of another string in S. We write "r.e." instead of "recursively enumerable."

An *infinite sequence over* Ω is an infinite sequence of elements from the alphabet Ω, where the sequence is infinite to the right but finite to the left. We denote by Ω^∞ is the set of all infinite sequences over Ω.

Let $\alpha \in \Omega^\infty$. For any $n \in \mathbb{N}$, we denote by $\alpha\restriction_n \in \Omega^*$ the first n elements in the infinite sequence α and by $\alpha(n)$ the nth element in α. Thus, for example, $\alpha\restriction_4 = \alpha(1)\alpha(2)\alpha(3)\alpha(4)$, and $\alpha\restriction_0 = \lambda$. For any $S \subset \Omega^*$, the set $\{\alpha \in \Omega^\infty \mid \exists n \in \mathbb{N}\ \alpha\restriction_n \in S\}$ is denoted by $[S]^\prec$. Note that (i) $[S]^\prec \subset [T]^\prec$ for every $S \subset T \subset \Omega^*$, and (ii) for every set $S \subset \Omega^*$ there exists a prefix-free set $P \subset \Omega^*$ such that $[S]^\prec = [P]^\prec$. For any $\sigma \in \Omega^*$, we denote by $[\sigma]^\prec$ the set $[\{\sigma\}]^\prec$, i.e., the set of all infinite sequences over Ω extending σ. Therefore $[\lambda]^\prec = \Omega^\infty$.

We briefly review measure theory. For the detail, see Billingsley [3]. A subset R of Ω^∞ is *open* if $R = [S]^\prec$ for some $S \subset \Omega^*$. In this paper we consider *the σ-field $\mathcal{F}$ generated by all open sets on Ω^∞*, which is defined as the intersection of all the σ-fields containing all open sets on Ω^∞. A *probability measure representation over* Ω is a function $r\colon \Omega^* \to [0, 1]$ such that (i) $r(\lambda) = 1$ and (ii) $r(\sigma) = \sum_{a \in \Omega} r(\sigma a)$ for every $\sigma \in \Omega^*$. A probability measure representation r induces the measure μ_r on the σ-field $\mathcal{F}$. In this paper, we use the following properties of the measure μ_r.

Proposition 1 (Properties of measure on Ω^∞)**.**

(i) $\mu_r\left([P]^\prec\right) = \sum_{\sigma \in P} r(\sigma)$ *for every prefix-free set* $P \subset \Omega^*$. *Therefore* $\mu_r(\emptyset) = \mu_r\left([\emptyset]^\prec\right) = 0$ *and* $\mu_r(\Omega^\infty) = \mu_r\left([\{\lambda\}]^\prec\right) = 1$.

(ii) $\mu_r(\mathcal{C}) \leq \mu_r(\mathcal{D})$ *for every* $\mathcal{C}, \mathcal{D}$ *in the σ-field $\mathcal{F}$ with* $\mathcal{C} \subset \mathcal{D}$.

(iii) $\mu_r\left(\bigcup_i \mathcal{C}_i\right) = \sum_i \mu_r(\mathcal{C}_i)$ *for every sequence* $\{\mathcal{C}_i\}_{i \in \mathbb{N}}$ *in the σ-field $\mathcal{F}$.* □

A function $f\colon \mathbb{N} \to \Omega^*$ or $f\colon \mathbb{N} \to \mathbb{Q}$ is called *computable* if there exists a deterministic Turing machine which on every input $n \in \mathbb{N}$ halts and outputs $f(n)$. A computable function is also called a *total recursive function*. A real a is called *computable* if there exists a computable function $g\colon \mathbb{N} \to \mathbb{Q}$ such that $|a - g(k)| < 2^{-k}$ for all $k \in \mathbb{N}$. We say that $\alpha \in \Omega^\infty$ is *computable* if the mapping $\mathbb{N} \ni n \mapsto \alpha\restriction_n$ is a computable function, which is equivalent to that the real $0.\alpha$ in base-$\#\Omega$ notation is computable.

2.2 Algorithmic Randomness

In the following we concisely review some definitions and results of algorithmic randomness [4, 11, 6].

We use $\mathcal{L}$ to denote Lebesgue measure on $\{0, 1\}^\infty$. Namely, $\mathcal{L} = \mu_r$ where the probability measure representation r is defined by the condition that $r(\sigma) = 2^{-|\sigma|}$ for every $\sigma \in \{0, 1\}^*$. The idea in algorithmic randomness is to think of an infinite binary sequence as random if it is in no *effective null set*. An effective null set is a subset $\mathcal{S}$ of $\{0, 1\}^\infty$ such that $\mathcal{L}(\mathcal{S}) = 0$ and $\mathcal{S}$ has some type of effective property. To specify an algorithmic randomness notion, one has to specify a type of effective null set, which is usually done by introducing a test concept. Failing the test is the same as being in the null set. In this manner, various randomness notions, such as 2-randomness, weak 2-randomness, Demuth randomness, Martin-Löf randomness, Schnorr randomness, Kurtz randomness, have been introduced so far, and a hierarchy of algorithmic randomness notions has been developed (see [11, 6] for the detail).

Among all randomness notions, *Martin-Löf randomness* is a central one. This is because in many respects, Martin-Löf randomness is well-behaved, in that the many properties of Martin-Löf random infinite sequences do match our intuition of what random infinite sequence should look like. Moreover, the concept of Martin-Löf randomness is robust in the sense that it admits various equivalent definitions that are all natural and intuitively meaningful (see e.g., [11, 6] for the detail). Martin-Löf randomness is defined as follows based on the notion of *Martin-Löf test*.

Definition 2 (Martin-Löf randomness, Martin-Löf [10])**.** *A subset $\mathcal{C}$ of $\mathbb{N}^+ \times \{0,1\}^*$ is called a* Martin-Löf test *if $\mathcal{C}$ is an r.e. set and for every $n \in \mathbb{N}^+$, $\mathcal{L}\left([\mathcal{C}_n]^{\prec}\right) \leq 2^{-n}$, where $\mathcal{C}_n = \left\{\, \sigma \;\middle|\; (n,\sigma) \in \mathcal{C} \,\right\}$.*

For any $\alpha \in \{0,1\}^\infty$, we say that α is Martin-Löf random *if for every Martin-Löf test $\mathcal{C}$ there exists $n \in \mathbb{N}^+$ such that $\alpha \notin [\mathcal{C}_n]^{\prec}$.* □

Let $\mathcal{C}$ be a Martin-Löf test. Then, for each $k \in \mathbb{N}^+$, using (ii) of Proposition 1 we see that $\mathcal{L}\left(\bigcap_{n=1}^{\infty} [\mathcal{C}_n]^{\prec}\right) \leq \mathcal{L}\left([\mathcal{C}_k]^{\prec}\right) \leq 2^{-k}$. On letting $k \to \infty$, we have

$$\mathcal{L}\left(\bigcap_{n=1}^{\infty} [\mathcal{C}_n]^{\prec}\right) = 0.$$

Thus, the set $\bigcap_{n=1}^{\infty} [\mathcal{C}_n]^{\prec}$ forms an effective null set in the notion of Martin-Löf randomness. Definition 2 says that an infinite binary sequence α is Martin-Löf random if α is not in the effective null set $\bigcap_{n=1}^{\infty} [\mathcal{C}_n]^{\prec}$ for any Martin-Löf test $\mathcal{C}$.

3 Martin-Löf Randomness with respect to Bernoulli Measure

In order to provide an operational characterization of the notion of probability we use a generalization of Martin-Löf randomness over Bernoulli measure.

Let Ω be an alphabet through out the rest of this paper. It plays a role of the set of all possible outcomes of experiments or observations. The *probability simplex on* Ω, denoted by $\mathbb{P}(\Omega)$, is the set of all functions $P\colon \Omega \to \mathbb{R}$ such that $P(a) \geq 0$ for every $a \in \Omega$ and $\sum_{a\in\Omega} P(a) = 1$. Bernoulli measure is given as follows.

Let $P \in \mathbb{P}(\Omega)$. Consider a function $r\colon \Omega^* \to [0,1]$ such that $r(a_1 \dots a_n) = \prod_{i=1}^{n} P(a_i)$ for every $n \in \mathbb{N}$ and $a_1, \dots, a_n \in \Omega$. The function r is a probability measure representation. The measure μ_r induced by r is *Bernoulli measure on* Ω^∞, denoted λ_P. Then Bernoulli measure λ_P on Ω^∞ has the following property: For every $\sigma \in \Omega^*$,

$$\lambda_P\left([\sigma]^{\prec}\right) = \prod_{a\in\Omega} P(a)^{N_a(\sigma)}, \tag{1}$$

where $N_a(\sigma)$ is the number of the occurrences of the element a in the finite string σ.[1]

Martin-Löf randomness with respect to Bernoulli measure is defined as follows. This notion was, in essence, introduced by Martin-Löf [10], as well as the notion of Martin-Löf randomness, which we describe in Definition 2.

Definition 3 (Martin-Löf randomness with respect to Bernoulli measure, Martin-Löf [10])**.** *Let $P \in \mathbb{P}(\Omega)$. A subset $\mathcal{C}$ of $\mathbb{N}^+ \times \Omega^*$ is called a* Martin-Löf P-test *if $\mathcal{C}$ is an r.e. set such that, for every $n \in \mathbb{N}^+$, $\lambda_P\left([\mathcal{C}_n]^{\prec}\right) \leq 2^{-n}$, where $\mathcal{C}_n = \left\{\, \sigma \;\middle|\; (n,\sigma) \in \mathcal{C} \,\right\}$.*

For any $\alpha \in \Omega^\infty$, we say that α is Martin-Löf P-random *if for every Martin-Löf P-test $\mathcal{C}$ there exists $n \in \mathbb{N}^+$ such that $\alpha \notin [\mathcal{C}_n]^{\prec}$.* □

Note that in Definition 3 we do not require that $P(a) > 0$ for all $a \in \Omega$. Therefore, $P(a_0)$ may be 0 for some $a_0 \in \Omega$. In the case where $\Omega = \{0,1\}$ and $P \in \mathbb{P}(\Omega)$ satisfies that $P(0) = P(1) = 1/2$, the Martin-Löf P-randomness results in the Martin-Löf randomness.

4 Ensemble

In this section, according to our former work [14] we give an operational characterization of the notion of probability for a finite probability space. We will identify the substance of the notion of probability for a finite probability space. For that purpose, we first review the notion of finite probability space, based on the notion of probability simplex. Let $P \in \mathbb{P}(\Omega)$. For each $A \subset \Omega$, we define $P(A)$ by $P(A) := \sum_{a\in A} P(a)$. Then, P can be regarded as a *finite probability space* $(\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is the set of all subset of Ω. The set Ω is the *sample space*, and elements in Ω are called *sample points* or *elementary events*. A subset of Ω is called an *event*, and $P(A)$ is called the *probability* of A for every event A. In what follows, we regard each element in $\mathbb{P}(\Omega)$ as a finite probability space in this manner.

We propose to regard a Martin-Löf P-random sequence of sample points as an operational characterization of the notion of probability for a finite probability space. Thus, since the Martin-Löf P-randomness plays a central role in our formalism, in particular we call it *ensemble* for a finite probability space, as in Definition 4. The name "ensemble" comes from physics.

Definition 4 (Ensemble, Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$. A Martin-Löf P-random sequence in Ω^∞ is called an* ensemble *for the finite probability space P.* □

Let $P \in \mathbb{P}(\Omega)$. Consider an infinite sequence $\alpha \in \Omega^\infty$ of outcomes which is obtained by an infinite reputation of trials described by the finite probability space P. The operational characterization of the notion of probability for the finite probability space P is thought to be completed if the property which the infinite sequence α has to satisfy is determined. We thus propose the following thesis.

Thesis 1. *Let $P \in \mathbb{P}(\Omega)$. An ensemble for P is an operational characterization of the finite probability space P.* □

[1] 0^0 is defined as 1 in the equation (1).

Let us consider the validity of Thesis 1. In what follows we check that the notion of ensemble satisfies the necessary conditions which the notion of probability is considered to have to satisfy from our intuitive understanding of the notion of probability. Let $P_0 \in \mathbb{P}(\Omega)$, and consider an infinite sequence $\alpha_0 \in \Omega^\infty$ of outcomes which is obtained by an infinite reputation of trials described by the finite probability space P_0.

The first necessary condition which the notion of probability is considered to have to satisfy is that *the law of large numbers holds for* α_0. Theorem 5 below confirms that this certainly holds. Note that we have to check whether the law of large numbers holds for *any* Martin-Löf P-random sequence since P is not computable reals, in general. However, we can certainly prove it using the Chernoff bound as follows.

Theorem 5 (The law of large numbers)**.** *Let $P \in \mathbb{P}(\Omega)$. For every $\alpha \in \Omega^\infty$, if α is an ensemble for P then, for every $a \in \Omega$, $\lim_{n\to\infty} N_a(\alpha\restriction_n)/n = P(a)$.* □

In order to prove Theorem 5, we need the following theorem, Chernoff bound, which is a modification of the form given in Section 1.2.2 of Goldreich [8].

Theorem 6 (Chernoff bound)**.** *Let P in $\mathbb{P}(\{0,1\})$. Then for each ε with $0 < \varepsilon \leq P(0)P(1)$ and each $n \in \mathbb{N}^+$, we have $\lambda_P\left([S_n]^\prec\right) < 2e^{-\frac{\varepsilon^2}{2P(0)P(1)}n}$, where S_n is the set of all $\sigma \in \{0,1\}^n$ such that $|N_1(\sigma)/n - P(1)| > \varepsilon$.* □

Proof of Theorem 5. Let $a \in \Omega$. We define $Q \in \mathbb{P}(\{0,1\})$ such that $Q(1) = P(a)$ and $Q(0) = 1 - P(a)$. Let β be the infinite binary sequence obtained from α by replacing all a by 1 and all symbols other than a by 0 in α. It follows from Theorem 13 below that β is Martin-Löf Q-random and $N_1(\beta\restriction_n) = N_a(\alpha\restriction_n)$ for every n.

Assume contrarily that $\lim_{n\to\infty} N_a(\alpha\restriction_n)/n \neq P(a)$. Then $\lim_{n\to\infty} N_1(\beta\restriction_n)/n \neq P(a)$ and therefore there exists $\varepsilon > 0$ such that $|N_1(\beta\restriction_n)/n - P(a)| > 2\varepsilon$ for infinitely many n. It follows from Theorem 6 that

$$\Pr\left[\left|\frac{N_1(\beta\restriction_n)}{n} - P(a)\right| > \varepsilon\right] < 2e^{-\frac{\varepsilon^2}{2P(a)(1-P(a))}n}.$$

Since p_i is not necessarily computable, we choose $r_L, r_R \in \mathbb{Q}$ with $p_i - 2\varepsilon < r_L < p_i - \varepsilon$ and $p_i + \varepsilon < r_R < p_i + 2\varepsilon$. For each $n \in \mathbb{N}^+$, let S_n be the set $\{x \in \{0,1\}^n \mid r_L < N_1(x)/n < r_R\}$ and let $T_n = \bigcup_{m=n}^\infty S_m$. Then $\beta \in [T_n]^\prec$ and

$$\lambda_Q\left([T_n]^\prec\right) \leq \sum_{m=n}^\infty 2e^{-cm} = 2e^{-cn}/(1-e^{-c}),$$

where $c \in \mathbb{Q}$ with $0 < c < \varepsilon^2/2P(a)(1-P(a))$. Then it is easy to show that there exists a total recursive function $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ such that $2e^{-cf(n)}/(1-e^{-c}) \leq 2^{-n}$. Thus, β is Martin-Löf Q-random since the set $\{(n,\sigma) \mid n \in \mathbb{N}^+ \,\&\, \sigma \in T_{f(n)}\}$ is Martin-Löf Q-test and $\beta \in [T_n]^\prec$ for every n. Hence we have a contradiction, and the result follows. □

The following is immediate from Theorem 5.

Corollary 7. *Let $P, Q \in \mathbb{P}(\Omega)$. If there exists $\alpha \in \Omega^\infty$ which is both an ensemble for P and an ensemble for Q, then $P = Q$.* □

The second necessary condition which the notion of probability is considered to have to satisfy is that *an elementary event with probability zero never occurs in the infinite sequence* α_0. Note that the notion of probability is more than the law of large numbers. To see this, consider the finite probability space $P \in \mathbb{P}(\{a,b\})$ such that $P(a) = 0$ and $P(b) = 1$, and consider the infinite sequence $\alpha = b,a,b,b,b,b,b,b,b,b,b,b,\ldots\ldots$. Since $\lim_{n\to\infty} N_a(\alpha\restriction_n)/n = 0 = P(a)$, the law of large numbers certainly holds for α. However, the elementary event a with probability zero has occurred in α once. This contradicts our intuition that an elementary event with probability zero never occurs. The example shows that the law of large numbers is insufficient to characterizes the notion of probability. Thus, *the notion of probability is more than the law of large numbers.*

Theorem 8 below states that an elementary event with probability zero never occurs in an ensemble, and thus shows that the notion of ensemble coincides with our intuition about the notion of probability in this respect. The result was, in essence, pointed out by Martin-Löf [10].

Theorem 8. *Let $P \in \mathbb{P}(\Omega)$, and let $a \in \Omega$. Suppose that α is an ensemble for the finite probability space P and $P(a) = 0$. Then α does not contain a.*

Proof. Assume contrarily that α contains a. Then there exists a prefix $\sigma \in \Omega^+$ of α which contains a. For each n, we define T_n as $\{\sigma\}$. Then, since $P(a) = 0$, we have $\lambda_P\left([T_n]^\prec\right) = 0$ for all $n \in \mathbb{N}^+$, and T is r.e., obviously. Thus, $\{T_n\}$ is Martin-Löf P-test. On the other hand, $\alpha \in [T_n]^\prec$ for all n, and therefore α is not Martin-Löf P-random. Hence, we have a contradiction, and the proof is completed. □

The following corollary is immediate from Theorem 8. It states that an elementary event with probability one always happens in an ensemble, and thus the notion of ensemble coincides with our intuition about the notion of probability in this respect.

Corollary 9. *Let $P \in \mathbb{P}(\Omega)$, and let $a \in \Omega$. Suppose that α is an ensemble for the finite probability space P and $P(a) = 1$. Then α consists only of a.* □

In what follows we consider the third necessary condition which the notion of probability is considered to have to satisfy. Assume that an observer A performs an infinite reputation of trials described by a finite probability space $P \in \mathbb{P}(\Omega)$, and thus is generating an infinite sequence $\alpha \in \Omega^\infty$ of outcomes of observations: $\alpha = a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, \ldots\ldots$. According to our thesis, Thesis 1, α is an ensemble for P. Consider another observer B who wants to adopt the following

subsequence β of α as the outcomes of the observations: $\beta = a_2, a_3, a_5, a_7, a_{11}, a_{13}, a_{17}, \ldots\ldots$, where the observer B only takes into account the nth elements in the original sequence α such that n is a prime number. According to Thesis 1, β has to be an ensemble for P, as well. However, is this true?

Consider this problem in a general setting. Assume as before that an observer A performs an infinite reputation of trials described by a finite probability space $P \in \mathbb{P}(\Omega)$, and thus is generating an infinite sequence $\alpha \in \Omega^\infty$ of outcomes of observations: $\alpha = a_1, a_2, a_3, a_4, a_5, a_6, a_7, \ldots\ldots$. According to Thesis 1, α is an ensemble for P. Now, let $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ be an injection. Consider another observer B who wants to adopt the following sequence β of α as the outcomes of the observations: $\beta = a_{f(1)}, a_{f(2)}, a_{f(3)}, a_{f(4)}, a_{f(5)}, \ldots\ldots$. According to our thesis, β has to be an ensemble for P, as well. However, is this true?

We can confirm this by restricting the ability of B, that is, by assuming that every observer can select elements from the original sequence α *only in an effective manner*. This means that the function $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ has to be a computable function. Theorem 10 below shows this result.

Theorem 10 (Closure property under a computable shuffling, Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$, and let α be an ensemble for P. Then, for every injective function $f\colon \mathbb{N}^+ \to \mathbb{N}^+$, if f is computable, then the infinite sequence $\alpha_f := \alpha(f(1))\alpha(f(2))\alpha(f(3))\alpha(f(4))\ldots\ldots$ is an ensemble for P.*

Proof. We show the contraposition. Suppose that α_f is not Martin-Löf P-random. Then there exists a Martin-Löf P-test $S \subset \mathbb{N}^+ \times \Omega^*$ such that $\alpha_f \in [S_n]^\prec$ for every n. For each $\sigma \in \Omega^+$, let $F(\sigma)$ be the set of all $\tau \in \Omega^+$ such that $|\tau| = \max f(\{1, 2, \ldots, |\sigma|\})$ and $\sigma = \tau(f(1))\tau(f(2))\ldots\tau(f(|\sigma|))$. We then define T to be $\{(n, F(\sigma)) \mid n \in \mathbb{N}^+ \ \&\ \sigma \in S_n\}$. Since f is an injection and $\sum_{a\in\Omega} P(a) = 1$, it is easy to see that $\lambda_P\left([F(\sigma)]^\prec\right) = \lambda_P\left([\sigma]^\prec\right)$. Therefore $\lambda_P\left([T_n]^\prec\right) = \lambda_P\left([S_n]^\prec\right) \leq 2^{-n}$. Thus, since T is r.e., we see that T is Martin-Löf P-test. On the other hand, $\alpha \in [T_n]^\prec$ for every n, and therefore α is not Martin-Löf P-random. This completes the proof. □

In other words, Theorem 10 states that ensembles for P are closed under a *computable shuffling*.

As the forth necessary condition which the notion of probability is considered to have to satisfy, we can consider the condition that the infinite sequence $\alpha_0 \in \Omega^\infty$ of outcomes which is obtained by an infinite reputation of trials described by the finite probability space $P_0 \in \mathbb{P}(\Omega)$ is closed under *the selection by a computable selection function*, as considered in the theory of collectives [17, 18, 19, 20, 5]. Theorem 11 below confirms that this condition certainly holds for every ensemble. Thus, ensembles for P are closed under the selection by a computable selection function. For the notion of the selection by a computable selection function and its meaning, see e.g., Downey and Hirschfeldt [6].

Theorem 11 (Closure property under the selection by a computable selection function, Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$, and let α be an ensemble for P. Let g be a computable selection function, that is, let $g\colon \Omega^* \to \{\mathrm{YES}, \mathrm{NO}\}$ be a computable function. Suppose that $g(\alpha\restriction_k)$ is defined for every $k \in \mathbb{N}$ and $\{k \in \mathbb{N} \mid g(\alpha\restriction_k) = \mathrm{YES}\}$ is an infinite set. Then the infinite sequence $\alpha_f := \alpha(f(1))\alpha(f(2))\alpha(f(3))\alpha(f(4))\ldots\ldots$ is an ensemble for P, where the computable function $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ is defined by*

$$f(n) = \min\{m \in \mathbb{N}^+ \mid \#\{k \leq m \mid g(\alpha\restriction_k) = \mathrm{YES}\} = n\}+1.$$

□

5 Conditional Probability and the Independence between Events

In this section, we operationally characterize the notions of conditional probability and the independence between events in a finite probability space in terms of ensembles.

Let $P \in \mathbb{P}(\Omega)$, and let $A \subset \Omega$ be an event in the finite probability space P. For each ensemble α for P, $\mathrm{C}_A(\alpha)$ is defined as the infinite binary sequence such that, for every i, its ith element $\mathrm{C}_A(\alpha)(i)$ is 1 if $\alpha(i) \in A$ and 0 otherwise. The pair (P, A) induces a finite probability space $\mathcal{C}(P, A) \in \mathbb{P}(\{0, 1\})$ such that $\mathcal{C}(P, A)(1) = P(A)$ and $\mathcal{C}(P, A)(0) = 1 - P(A)$. Note that the notions of $\mathrm{C}_A(\alpha)$ and $\mathcal{C}(P, A)$ in our theory together correspond to the notion of *mixing* in the theory of collectives by von Mises [18]. We can then show the following theorem.

Theorem 12 (Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$, and let $A \subset \Omega$. Suppose that α is an ensemble for the finite probability space P. Then $\mathrm{C}_A(\alpha)$ is an ensemble for the finite probability space $\mathcal{C}(P, A)$.* □

In order to prove Theorem 12, it is convenient to prove the following theorem first, from which Theorem 12 follows. For the proof of Theorem 13, see our former work [14].

Theorem 13 (Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$. Let α be an ensemble for P, and let a and b be distinct elements in Ω. Suppose that β is the infinite sequence in $(\Omega - \{b\})^\infty$ obtained by replacing all occurrences of b by a in α. Then β is an ensemble for Q, where $Q \in \mathbb{P}(\Omega - \{b\})$ such that $Q(d) = P(a) + P(b)$ if $d = a$ and $Q(d) = P(d)$ otherwise.* □

We show that the notion of conditional probability in a finite probability space can be represented by an ensemble in a natural manner. For that purpose we recall the notion of conditional probability in a finite probability space.

Let $P \in \mathbb{P}(\Omega)$, and let $B \subset \Omega$ be an event in the finite probability space P. Suppose that $P(B) > 0$. Then, for each event $A \subset \Omega$, the *conditional probability of A given B*, denoted by $P(A|B)$, is defined as $P(A \cap B)/P(B)$. This notion defines a finite probability space

$P_B \in \mathbb{P}(B)$ such that $P_B(a) = P(\{a\}|B)$ for every $a \in B$.

When an infinite sequence $\alpha \in \Omega^\infty$ contains infinitely many elements from B, $\mathrm{Filtered}_B(\alpha)$ is defined as the infinite sequence in B^∞ obtained from α by eliminating all elements in $\Omega - B$ occurring in α. If α is an ensemble for the finite probability space P and $P(B) > 0$, then α contains infinitely many elements from B due to Theorem 5. Therefore, $\mathrm{Filtered}_B(\alpha)$ is defined in this case. Note that the notion of $\mathrm{Filtered}_B(\alpha)$ in our theory corresponds to the notion of *partition* in the theory of collectives by von Mises [18].

We can then show Theorem 14 below, which states that ensembles are closed under conditioning. For the proof of the theorem, see Tadaki [14].

Theorem 14 (Closure property under conditioning, Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$, and let $B \subset \Omega$ be an event in the finite probability space P with $P(B) > 0$. For every ensemble α for P, $\mathrm{Filtered}_B(\alpha)$ is an ensemble for the finite probability space P_B.* □

As an application of Theorem 14, we can consider the Von Neumann extractor as follows.

Example 15 (Von Neumann extractor)**.** *Consider a Bernoulli sequence in the sense of normal probability theory. Recall that the Von Neumann extractor takes successive pairs of consecutive bits from the Bernoulli sequence. If the two bits matches, no output is generated. If the bits differs, the value of the first bit is output. The Von Neumann extractor can be shown to produce a uniform binary output. For the detail, see [21].*

In our framework, the Von Neumann extractor operates as follows: Let $P \in \mathbb{P}(\{0,1\})$ and let α be an ensemble for P. Then α can be regarded as an ensemble for $Q \in \mathbb{P}(\{00,01,10,11\})$ where $Q(ab) = P(a)P(b)$ for every $a, b \in \{0,1\}$. Consider the event $B = \{01, 10\}$. It follows from Theorem 14 that $\mathrm{Filtered}_B(\alpha)$ is an ensemble for $P_B \in \mathbb{P}(\{01,10\})$ with $P_B(01) = P_B(10) = 1/2$. Namely, α is, in essence, Martin-Löf random. Hence, a random individual infinite sequence is certainly extracted by the Von Neumann extractor in our framework. □

Let $P \in \mathbb{P}(\Omega)$. For any events $A, B \subset \Omega$ in the finite probability space P, we say that A and B are *independent* if $P(A \cap B) = P(A)P(B)$. In the case of $P(B) > 0$, A and B are independent if and only if $P(A|B) = P(A)$.

Theorem 16 below gives operational characterizations of the notion of the independence between two events in terms of ensembles. Let $\alpha, \beta \in \Omega^\infty$. We say that α and β are *equivalent* if there exists $P \in \mathbb{P}(\Omega)$ such that α and β are both an ensemble for P. For the proof of Theorem 16, see Tadaki [14].

Theorem 16 (Tadaki [14])**.** *Let $P \in \mathbb{P}(\Omega)$, and let $A, B \subset \Omega$ be events in the finite probability space P. Suppose that $P(B) > 0$. Then the following conditions are equivalent to one another.*

(i) The events A and B are independent.

(ii) For every ensemble α for the finite probability space P, $\mathrm{C}_A(\alpha)$ is equivalent to $\mathrm{C}_{A\cap B}(\mathrm{Filtered}_B(\alpha))$.

(iii) There exists an ensemble α for the finite probability space P such that $\mathrm{C}_A(\alpha)$ is equivalent to $\mathrm{C}_{A\cap B}(\mathrm{Filtered}_B(\alpha))$. □

6 Independence of Random Variables

In this section, we operationally characterize the notion of the independence of random variables in a finite probability space in terms of ensembles.

A *random variable* on Ω is a function $X\colon \Omega \to \Omega'$ where Ω' is an alphabet. Let $X_1\colon \Omega \to \Omega_1, \dots, X_n\colon \Omega \to \Omega_n$ be random variables on Ω. For any predicate $F(v_1, \dots, v_n)$ with variables $v_1, \dots, v_n$, we use $F(X_1, \dots, X_n)$ to denote the event $\{a \in \Omega \mid F(X_1(a), \dots, X_n(a))\}$. We say that the random variables $X_1, \dots, X_n$ are *independent* if for every $x_1 \in \Omega_1, \dots, x_n \in \Omega_n$ it holds that $P(X_1 = x_1 \ \& \ \dots \ \& \ X_n = x_n) = P(X_1 = x_n) \cdots P(X_n = x_n)$.

Let $\alpha \in \Omega^\infty$, and $X\colon \Omega \to \Omega'$ be a random variable on Ω. We define $X(\alpha)$ as an infinite sequence β over Ω' such that $\beta(i) = X(\alpha(i))$ for every $i \in \mathbb{N}^+$. Using Theorem 13 we can show the following theorem.

Theorem 17 (Closure property under the mapping by a random variable)**.** *Let $X\colon \Omega \to \Omega'$ be a random variable on Ω, and let $P \in \mathbb{P}(\Omega)$. If α is an ensemble for P then $X(\alpha)$ is an ensemble for $P' \in \mathbb{P}(\Omega')$ where $P'(x) = P(X = x)$ for every $x \in \Omega'$.* □

We introduce the notion of the *independence* of ensembles as follows. Let $\Omega_1, \dots, \Omega_n$ be alphabets. For any $\alpha_1 \in \Omega_1^\infty, \dots, \alpha_n \in \Omega_n^\infty$, we use $\alpha_1 \times \cdots \times \alpha_n$ to denote an infinite sequence over $\Omega_1 \times \cdots \times \Omega_n$ such that $\alpha(i) = (\alpha_1(i), \dots, \alpha_n(i))$ for every $i \in \mathbb{N}^+$.

Definition 18 (Independence of ensembles)**.** *Let $\Omega_1, \dots, \Omega_n$ be alphabets, and let $P_1 \in \mathbb{P}(\Omega_1), \dots, P_n \in \mathbb{P}(\Omega_n)$. Let $\alpha_1, \dots, \alpha_n$ be ensembles for $P_1, \dots, P_n$, respectively. We say that $\alpha_1, \dots, \alpha_n$ are* independent *if $\alpha_1 \times \cdots \times \alpha_n$ is an ensemble for $P \in \mathbb{P}(\Omega_1 \times \cdots \times \Omega_n)$ where $P(a_1, \dots, a_n) = P_1(a_1) \cdots P_n(a_n)$ for every $a_1 \in \Omega_1, \dots, a_n \in \Omega_n$.* □

Note that the notion of the independence of ensembles in our theory corresponds to the notion of *independence* of collectives in the theory of collectives by von Mises [18]. The following theorem gives equivalent characterizations of the notion of the independence of random variables in terms of that of ensembles.

Theorem 19. *Let $X_1\colon \Omega \to \Omega_1, \dots, X_n\colon \Omega \to \Omega_n$ be random variables on Ω, and let $P_1 \in \mathbb{P}(\Omega_1), \dots, P_n \in \mathbb{P}(\Omega_n)$. Then the following conditions are equivalent to one another.*

(i) The random variables $X_1, \dots, X_n$ are independent.

(ii) For every ensemble α for the finite probability space P, the ensembles $X_1(\alpha), \ldots, X_n(\alpha)$ are independent.

(iii) There exists an ensemble α for P such that the ensembles $X_1(\alpha), \ldots, X_n(\alpha)$ are independent. □

In the rest of this section we consider the notion of *Martin-Löf P-randomness relative to an oracle*. The *relativized computation* is a generalization of normal computation. Let $\beta_1, \ldots, \beta_m$ be arbitrary infinite sequences over an alphabet. In the relativized computation, a (deterministic) Turing machine is allowed to refer to $(\beta_1, \ldots, \beta_m)$ as an oracle during the computation. Namely, in the relativized computation, a Turing machine can query (k, i) at any time and then obtains the response $\beta_k(i)$ during the computation. Such a Turing machine is called an *oracle Turing machine*. The relativized computation is more powerful than normal computation, in general.

We can define the notion of *Martin-Löf P-test relative to $\beta_1, \ldots, \beta_m$* where the Turing machine which computes the Martin-Löf P-test is an oracle Turing machine which can refer to the sequence $\beta_1, \ldots, \beta_m$ during the computation. Using the notion of Martin-Löf P-tests relative to $\beta_1, \ldots, \beta_m$, we can define the notion of *Martin-Löf P-randomness relative to $\beta_1, \ldots, \beta_m$* in the same manner as the second part of Definition 3. Obviously, the following holds.

Proposition 20. *Let $\beta_1, \ldots, \beta_m$ be infinite sequences over an alphabet, and let $P \in \mathbb{P}(\Omega)$. For every $\alpha \in \Omega^\infty$, if α is Martin-Löf P-random relative to $\beta_1, \ldots, \beta_m$, then α is Martin-Löf P-random.* □

The converse does not necessarily hold. In the case where α is Martin-Löf P-random, the converse means that the Martin-Löf P-randomness of α is independent of $\beta_1, \ldots, \beta_m$ in a certain sense.

For any $P \in \mathbb{P}(\Omega)$, we say that P is *computable* if $P(a)$ is a computable real for every $a \in \Omega$. The following theorem gives an equivalent characterization of the notion of the Independence of ensembles in terms of Martin-Löf P-randomness relative to an oracle. Its proof is obtained by modifying the proof of van Lambalgen's Theorem [15].

Theorem 21. *Let $P_1 \in \mathbb{P}(\Omega_1), \ldots, P_n \in \mathbb{P}(\Omega_n)$. Let $\alpha_1, \ldots, \alpha_n$ be ensembles for $P_1, \ldots, P_n$, respectively. Suppose that $P_1, \ldots, P_n$ are computable. Then the ensembles $\alpha_1, \ldots, \alpha_n$ are independent if and only if for every $k = 1, \ldots, n$ it holds that α_k is Martin-Löf P_k-random relative to $\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_n$.* □

Combining Theorem 19 with Theorem 21 we obtain the following theorem.

Theorem 22. *Let $X_1 \colon \Omega \to \Omega_1, \ldots, X_n \colon \Omega \to \Omega_n$ be random variables on Ω, and let $P \in \mathbb{P}(\Omega)$. For each $k = 1, \ldots, n$, let $P_k \in \mathbb{P}(\Omega_k)$ be a finite probability space such that $P_k(x) = P(X_k = x)$ for every $x \in \Omega_k$. Suppose that P is computable. Then the following conditions are equivalent to one another.*

(i) The random variables $X_1, \ldots, X_n$ are independent.

(ii) For every ensemble α for P and every $k = 1, \ldots, n$ it holds that $X_k(\alpha)$ is Martin-Löf P_k-random relative to $X_1(\alpha), \ldots, X_{k-1}(\alpha), X_{k+1}(\alpha), \ldots, X_n(\alpha)$.

(iii) There exists an ensemble α for P such that for every $k = 1, \ldots, n$ it holds that $X_k(\alpha)$ is Martin-Löf P_k-random relative to $X_1(\alpha), \ldots, X_{k-1}(\alpha)$, $X_{k+1}(\alpha), \ldots, X_n(\alpha)$. □

7 Application to Cryptography

In this section, we make an application of our formalism to cryptography by presenting new equivalent characterizations of the notion of perfect secrecy in terms of our formalism.

The notion of *perfect secrecy* plays a basic role in cryptography. First, we review the definition of encryption schemes to which the notion of perfect secrecy is applied.

Definition 23 (Encryption scheme)**.** *Let $\mathcal{M}$, $\mathcal{K}$, and $\mathcal{C}$ be alphabets. An* encryption scheme *over a message space $\mathcal{M}$, a key space $\mathcal{K}$, and a ciphertext space $\mathcal{C}$, is a tuple $\Pi = (P_{\mathrm{K}}, \mathsf{Enc}, \mathsf{Dec})$ such that (i) $P_{\mathrm{K}} \in \mathbb{P}(\mathcal{K})$, (ii) $\mathsf{Enc} \colon \mathcal{M} \times \mathcal{K} \to \mathcal{C}$, (iii) $\mathsf{Dec} \colon \mathcal{C} \times \mathcal{K} \to \mathcal{M}$, and (iv) $\mathsf{Dec}(\mathsf{Enc}(m, k), k) = m$ for every $m \in \mathcal{M}$ and $k \in \mathcal{K}$.* □

Let $\Pi = (P_{\mathrm{K}}, \mathsf{Enc}, \mathsf{Dec})$ be as in Definition 23, and let $Q \in \mathbb{P}(\mathcal{M})$, which serves as a probability distribution over message space $\mathcal{M}$ for the encryption scheme Π. We consider a finite probability space $P_{\Pi,Q} \in \mathbb{P}(\mathcal{M} \times \mathcal{K})$ defined by the condition that $P_{\Pi,Q}(m, k) = Q(m)P_{\mathrm{K}}(k)$ for every $m \in \mathcal{M}$ and $k \in \mathcal{K}$. We then define random variables $M_{\Pi,Q}$ and $C_{\Pi,Q}$ on $\mathcal{M} \times \mathcal{K}$ by $M_{\Pi,Q}(m, k) = m$ and $C_{\Pi,Q}(m, k) = \mathsf{Enc}(m, k)$, respectively. The notion of perfect secrecy is then defined as follows.

Definition 24 (Perfect secrecy, Shannon [13])**.** *Let $\mathcal{M}$, $\mathcal{K}$, and $\mathcal{C}$ be alphabets. Let $\Pi = (P_{\mathrm{K}}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme over a message space $\mathcal{M}$, a key space $\mathcal{K}$, and a ciphertext space $\mathcal{C}$. The encryption scheme Π is* perfectly secret *if for every $Q \in \mathbb{P}(\mathcal{M})$ it holds that the random variables $M_{\Pi,Q}$ and $C_{\Pi,Q}$ are independent.* □

Using Theorems 19 and 22 we can show the following theorem, which characterizes the notion of perfect secrecy in terms of the notions of the independence of ensembles and Martin-Löf P-randomness relative to an oracle.

Theorem 25 (New equivalent characterizations of perfect secrecy)**.** *Let $\mathcal{M}$, $\mathcal{K}$, and $\mathcal{C}$ be alphabets. Let $\Pi = (P_{\mathrm{K}}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme over a message space $\mathcal{M}$, a key space $\mathcal{K}$, and a ciphertext space $\mathcal{C}$. For each $Q \in \mathbb{P}(\mathcal{M})$, let $R_Q \in \mathbb{P}(\mathcal{C})$ be a finite probability space such that $R_Q(c) = P_{\Pi,Q}(C_{\Pi,Q} = c)$ for every $c \in \mathcal{C}$. Suppose that P_{K} is computable. Then the following conditions are equivalent to one another.*

(i) The encryption scheme Π *is perfectly secret.*

(ii) For every $Q \in \mathbb{P}(\mathcal{M})$ *and every ensemble* α *for* $P_{\Pi,Q}$*, the ensembles* $M_{\Pi,Q}(\alpha)$ *and* $C_{\Pi,Q}(\alpha)$ *are independent.*

(iii) For every $Q \in \mathbb{P}(\mathcal{M})$ *there exists an ensemble* α *for* $P_{\Pi,Q}$ *such that the ensembles* $M_{\Pi,Q}(\alpha)$ *and* $C_{\Pi,Q}(\alpha)$ *are independent.*

(iv) For every computable $Q \in \mathbb{P}(\mathcal{M})$ *and every ensemble* α *for* $P_{\Pi,Q}$ *it holds that* $M_{\Pi,Q}(\alpha)$ *is Martin-Löf* Q*-random relative to* $C_{\Pi,Q}(\alpha)$*, and* $C_{\Pi,Q}(\alpha)$ *is Martin-Löf* R_Q*-random relative to* $M_{\Pi,Q}(\alpha)$

(v) For every computable $Q \in \mathbb{P}(\mathcal{M})$ *there exists an ensemble* α *for* $P_{\Pi,Q}$ *such that* $M_{\Pi,Q}(\alpha)$ *is Martin-Löf* Q*-random relative to* $C_{\Pi,Q}(\alpha)$*, and* $C_{\Pi,Q}(\alpha)$ *is Martin-Löf* R_Q*-random relative to* $M_{\Pi,Q}(\alpha)$*.* □

Note that the finite probability space P_{K}, which serves as a probability distribution over key space $\mathcal{K}$, is normally computable in modern cryptography.

Acknowledgements

This work was partially supported by "Research and Development of the Public Key Systems for Secure Communication between Organizations", the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan and by JSPS KAKENHI Grant Number 24540142. This work was partially done while the author was visiting the Institute for Mathematical Sciences, National University of Singapore in 2014.

References

[1] L. Bienvenu, W. Merkle, and A. Nies, Solovay functions and K-triviality, Proceedings of the 28th Symposium on Theoretical Aspects of Computer Science (STACS 2011), pp.452–463, 2011.

[2] V. Brattka, J. Miller, and A. Nies, "Randomness and differentiability," preprint, 2012.

[3] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, Inc., New York, 1995.

[4] G. J. Chaitin, *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.

[5] A. Church, "On the concept of a random sequence," *Bulletin of the American Mathematical Society*, vol. 46, pp. 130–135, 1940.

[6] R. G. Downey and D. R. Hirschfeldt, *Algorithmic Randomness and Complexity*. Springer-Verlag, New York, 2010.

[7] H. Everett, III, ""Relative State" formulation of quantum mechanics," *Rev. Mod. Phys.*, vol. 29, no. 3, pp. 454–462, 1957.

[8] O. Goldreich, *Foundations of Cryptography: Volume 1 – Basic Tools*. Cambridge University Press, New York, 2001.

[9] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*. Chapman & Hall/CRC Press, 2007.

[10] P. Martin-Löf, "The definition of random sequences," *Information and Control*, vol. 9, pp. 602–619, 1966.

[11] A. Nies, *Computability and Randomness*. Oxford University Press, Inc., New York, 2009.

[12] M. B. Pour-El and J. I. Richards, *Computability in Analysis and Physics*. Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1989.

[13] C. E. Shannon, "Communication theory of secrecy systems," *Bell Systems Technical Journal*, vol. 28, pp. 656–715, 1949.

[14] K. Tadaki, An operational characterization of the notion of probability by algorithmic randomness. Proceedings of the 37th Symposium on Information Theory and its Applications (SITA2014), 5.4.1, pp. 389–394, December 9-12, 2014, Unazuki, Toyama, Japan.

[15] M. van Lambalgen, *Random Sequences*. Ph.D. dissertation, University of Amsterdam, 1987.

[16] J. Ville, "Étude Critique de la Notion de Collectif," Monographies des Probabilités. Calcul des Probabilités et ses Applications. Gauthier-Villars, Paris, 1939.

[17] R. von Mises, *Probability, Statistics and Truth*. Dover Publications, Inc., New York, 1957.

[18] R. von Mises, *Mathematical Theory of Probability and Statistics*. Academic Press Inc., New York, 1964.

[19] A. Wald, "Sur la notion de collectif dans la calcul des probabilités," *Comptes Rendus des Seances de l'Académie des Sciences*, vol. 202, pp. 180–183, 1936.

[20] A. Wald, "Die Wiederspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung," *Ergebnisse eines Mathematischen Kolloquiums*, vol. 8, pp. 38–72, 1937.

[21] Wikipedia contributors, "Randomness extractor," Wikipedia, The Free Encyclopedia, `http://en.wikipedia.org/wiki/Randomness_extractor` (accessed December 17, 2014).