A refinement of the argument of Bell's inequality versus quantum mechanics by algorithmic randomness

Kohtaro Tadaki
Department of Computer Science,
Chubu University

Abstract. The notion of probability plays a crucial role in quantum mechanics. It appears in quantum mechanics as the *Born rule*. In modern mathematics which describes quantum mechanics, however, probability theory means nothing other than measure theory, and therefore any operational characterization of the notion of probability is still missing in quantum mechanics. In our former works [K. Tadaki, arXiv:1804.10174], based on the toolkit of *algorithmic randomness*, we presented a refinement of the Born rule, called the *principle of typicality*, for specifying the property of results of measurements *in an operational way*. In this paper, we make an application of our framework to the argument of Bell's inequality versus quantum mechanics for refining it, in order to demonstrate how properly our framework works in practical problems in quantum mechanics.

1 Introduction

The notion of probability plays a crucial role in quantum mechanics. It appears in quantum mechanics as the so-called *Born rule*, i.e., *the probability interpretation of the wave function* [8, 32, 16]. In modern mathematics which describes quantum mechanics, however, probability theory means nothing other than *measure theory*, and therefore any *operational characterization of the notion of probability* is still missing in quantum mechanics. In this sense, the current form of quantum mechanics is considered to be *imperfect* as a physical theory which must stand on operational means.

1.1 The previous works

In our former works [20, 23, 24, 25, 27, 28, 29], based on the toolkit of *algorithmic randomness*, we presented a refinement of the Born rule as an alternative rule to it, for aiming at making quantum mechanics *operationally perfect*. i.e., for specifying the property of the results of quantum measurements *in an operational way*. 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 [19, 14, 4, 15, 18, 5, 17, 9]. We used the notion of *Martin-Löf randomness with respect to Bernoulli measure* [15] to present the refinement of the Born rule. We then presented an operational refinement of the Born rule for mixed states, as an alternative rule to it, based on algorithmic randomness. In particular, we gave a precise definition for the notion of *mixed state*. We then showed that all of the refined rules of the Born rule for both pure states and mixed states can be derived from a single postulate, called the *principle of typicality*, in a unified manner. We did this from the point of view of the *many-worlds interpretation of quantum mechanics*. Finally, we made an application of our framework to the BB84 quantum key distribution protocol in order to demonstrate how properly our framework works in practical problems in quantum mechanics, based on the principle of typicality. See the work [29] for the detail of the framework based on the principle of typicality.

1.2 Contribution of the paper

In this paper, we make an application of our framework based on the principle of typicality to the argument of Bell's inequality versus quantum mechanics to refine it, in order to demonstrate how properly our framework works in *practical* problems in quantum mechanics. Sepecifically, in this paper, we *refine* and *reformulate* the argument of Bell's inequality versus quantum mechanics, which is described in Section 2.6 "EPR and the Bell inequality" of Nielsen and Chuang [16]. Thus, on the one hand, we refine and reformulate the *assumptions of local realism* to lead to *Bell's inequality*

$$\langle RS\rangle + \langle QS\rangle + \langle RT\rangle - \langle QT\rangle \leq 2,$$

i.e., the inequality (23) below, in terms of the theory of *operational characterization of the notion of probability by algorithmic randomness* developed by Tadaki [21, 22, 26]. On the other hand, we refine and reformulate the corresponding argument of quantum mechanics to violate Bell's inequality, resulting in the equality

$$\langle RS\rangle + \langle QS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2},$$

i.e., the equality (18) below, based on the principle of typicality.

1.3 Organization of the paper

The paper is organized as follows. We begin in Section 2 with some mathematical preliminaries, in particular, about measure theory and Martin-Löf randomness with respect to an arbitrary probability measure. We then review the notion of Martin-Löf randomness with respect to an arbitrary Bernoulli measure, called the *Martin-Löf P-randomness* in this paper, in Section 3. In Section 4 we summarize the theorems and notions on Martin-Löf P-randomness from Tadaki [21, 22, 26], which are need to establish the contributions of this paper presented in Sections 7 and 8. In Section 5 we review the central postulates of the *conventional* quantum mechanics according to Nielsen and Chuang [16]. In Section 6 we review the framework of the *principle of typicality*, which was introduced by

Tadaki [24, 29]. In Sections 7 and 8, we describe the contributions of this paper. On the one hand, in Section 7, we refine and reformulate the argument of quantum mechanics to violate Bell's inequality, based on the principle of typicality. On the other hand, in Section 8, we refine and reformulate the *assumptions of local realism* to lead to Bell's inequality, in terms of the theory of operational characterization of the notion of probability by algorithmic randomness developed by Tadaki [21, 22, 26]. We conclude this paper with summary in Section 9.

2 Mathematical 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*.

An *alphabet* is a non-empty finite set. Let Ω be an arbitrary alphabet throughout the rest of this subsection. A *finite string over* Ω is a finite sequence of elements from the alphabet Ω. We use Ω^* to denote the set of all finite strings over Ω, which contains the *empty string* denoted by λ. 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.

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 use Ω^∞ to denote 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 for any $n \in \mathbb{N}^+$ we denote 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 $[\lambda]^\prec = \Omega^\infty$.

2.2 Martin-Löf randomness with respect to an arbitrary probability measure

We briefly review measure theory according to Nies [17, Section 1.9]. See also Billingsley [2] for measure theory in general.

Let Ω be an arbitrary alphabet. A real-valued function μ defined on the class of all subsets of Ω^∞ is called an *outer measure on* Ω^∞ if the following conditions hold:

(i) $\mu(\emptyset) = 0$;

(ii) $\mu(\mathcal{C}) \le \mu(\mathcal{D})$ for every subsets $\mathcal{C}$ and $\mathcal{D}$ of Ω^∞ with $\mathcal{C} \subset \mathcal{D}$;

(iii) $\mu\left(\bigcup_i \mathcal{C}_i\right) \le \sum_i \mu(\mathcal{C}_i)$ for every sequence $\{\mathcal{C}_i\}_{i\in\mathbb{N}}$ of subsets of Ω^∞.

A *probability measure representation over* Ω is a function $r\colon \Omega^* \to [0, 1]$ such that

(i) $r(\lambda) = 1$ and

(ii) for every $\sigma \in \Omega^*$ it holds that

$$r(\sigma) = \sum_{a \in \Omega} r(\sigma a). \tag{1}$$

A probability measure representation r over Ω *induces* an outer measure μ_r on Ω^∞ in the following manner: A subset $\mathcal{R}$ of Ω^∞ is called *open* if $\mathcal{R} = [S]^\prec$ for some $S \subset \Omega^*$. Let r be an arbitrary probability measure representation over Ω. For each open subset $\mathcal{A}$ of Ω^∞, we define $\mu_r(\mathcal{A})$ by

$$\mu_r(\mathcal{A}) := \sum_{\sigma \in E} r(\sigma),$$

where E is a prefix-free subset of Ω^* with $[E]^\prec = \mathcal{A}$. Due to the equality (1) the sum is independent of the choice of the prefix-free set E, and therefore the value $\mu_r(\mathcal{A})$ is well-defined. Then, for any subset $\mathcal{C}$ of Ω^∞, we define $\mu_r(\mathcal{C})$ by

$$\mu_r(\mathcal{C}) := \inf\{\mu_r(\mathcal{A}) \mid \mathcal{C} \subset \mathcal{A} \;\&\; \mathcal{A} \text{ is an open subset of } \Omega^\infty\}.$$

We can then show that μ_r is an *outer measure* on Ω^∞ such that $\mu_r(\Omega^\infty) = 1$.

A class $\mathcal{F}$ of subsets of Ω^∞ is called a *σ-field on* Ω^∞ if $\mathcal{F}$ includes Ω^∞, is closed under complements, and is closed under the formation of countable unions. The *Borel class* $\mathcal{B}_\Omega$ is the σ-field *generated by* all open sets on Ω^∞. Namely, the Borel class $\mathcal{B}_\Omega$ is defined as the intersection of all the σ-fields on Ω^∞ containing all open sets on Ω^∞. A real-valued function μ defined on the Borel class $\mathcal{B}_\Omega$ is called a *probability measure on* Ω^∞ if the following conditions hold:

(i) $\mu(\emptyset) = 0$ and $\mu(\Omega^\infty) = 1$;

(ii) $\mu\left(\bigcup_i \mathcal{D}_i\right) = \sum_i \mu(\mathcal{D}_i)$ for every sequence $\{\mathcal{D}_i\}_{i \in \mathbb{N}}$ of sets in $\mathcal{B}_\Omega$ such that $\mathcal{D}_i \cap \mathcal{D}_i = \emptyset$ for all $i \neq j$.

Then, for every probability measure representation r over Ω, we can show that the restriction of the outer measure μ_r on Ω^∞ to the Borel class $\mathcal{B}_\Omega$ is a probability measure on Ω^∞. We denote the restriction of μ_r to $\mathcal{B}_\Omega$ by μ_r just the same.

Then it is easy to see that

$$\mu_r\left([\sigma]^\prec\right) = r(\sigma) \tag{2}$$

for every probability measure representation r over Ω and every $\sigma \in \Omega^*$. The probability measure μ_r is called a *probability measure induced by the probability measure representation r*.

Now, we introduce the notion of *Martin-Löf randomness* [15] in a general setting, as follows.

Definition 2.1 (Martin-Löf randomness with respect to a probability measure)**.** Let Ω be an alphabet, and let μ be a probability measure on Ω^∞. A subset $\mathcal{C}$ of $\mathbb{N}^+ \times \Omega^*$ is called a *Martin-Löf test with respect to μ* if $\mathcal{C}$ is a recursively enumerable set, and

$$\mu\left([\mathcal{C}_n]^\prec\right) < 2^{-n} \tag{3}$$

for every $n \in \mathbb{N}^+$, where $\mathcal{C}_n$ denotes the set $\{\,\sigma \mid (n, \sigma) \in \mathcal{C}\,\}$.

For any $\alpha \in \Omega^\infty$, we say that α is *Martin-Löf random with respect to* μ if

$$\alpha \notin \bigcap_{n=1}^{\infty} [\mathcal{C}_n]^{\prec}$$

for every Martin-Löf test $\mathcal{C}$ with respect to μ. □

3 Martin-Löf P-randomness

The principle of typicality, Postulate 5 below, is stated by means of the notion of Martin-Löf randomness with respect to an arbitrary probability measure introduced in the preceding section. However, in many situations of the applications of the principle of typicality, such as in a contribution of this paper described in Section 7, a more restricted notion is used where the probability measure is chosen to be a Bernoulli measure. Specifically, the notion of *Martin-Löf randomness with respect to a Bernoulli measure* is used in many situations of the applications. Thus, in order to introduce this notion, we first review the notions of *finite probability space* and *Bernoulli measure*.

Definition 3.1 (Finite probability space)**.** Let Ω be an alphabet. A *finite probability space on* Ω is a function $P\colon \Omega \to [0,1]$ such that

(i) $P(a) \geq 0$ for every $a \in \Omega$, and

(ii) $\sum_{a\in\Omega} P(a) = 1$.

The set of all finite probability spaces on Ω is denoted by $\mathbb{P}(\Omega)$.

Let $P \in \mathbb{P}(\Omega)$. The set Ω is called the *sample space* of P, and elements of Ω are called *sample points* or *elementary events* of P. For each $A \subset \Omega$, we define $P(A)$ by

$$P(A) := \sum_{a\in A} P(a).$$

A subset of Ω is called an *event* on P, and $P(A)$ is called the *probability* of A for every event A on P. □

Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. For each $\sigma \in \Omega^*$, we use $P(\sigma)$ to denote $P(\sigma_1)P(\sigma_2)\dots P(\sigma_n)$ where $\sigma = \sigma_1\sigma_2\dots\sigma_n$ with $\sigma_i \in \Omega$. Therefore $P(\lambda) = 1$, in particular. For each subset S of Ω^*, we use $P(S)$ to denote

$$\sum_{\sigma\in S} P(\sigma).$$

Therefore $P(\emptyset) = 0$, in particular.

Consider a function $r\colon \Omega^* \to [0,1]$ such that $r(\sigma) = P(\sigma)$ for every $\sigma \in \Omega^*$. It is then easy to see that the function r is a probability measure representation over Ω. The probability measure μ_r induced by r is called a *Bernoulli measure on* Ω^∞, denoted λ_P. The Bernoulli measure λ_P on Ω^∞ satisfies that

$$\lambda_P\left([\sigma]^{\prec}\right) = P(\sigma) \tag{4}$$

for every $\sigma \in \Omega^*$, which follows from (2).

The notion of *Martin-Löf randomness with respect to a Bernoulli measure* is defined as follows. We call it the *Martin-Löf P-randomness*, since it depends on a finite probability space P. This notion was, in essence, introduced by Martin-Löf [15], as well as the notion of Martin-Löf randomness with respect to Lebesgue measure.

Definition 3.2 (Martin-Löf P-randomness, Martin-Löf [15])**.** Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. For any $\alpha \in \Omega^\infty$, we say that α is *Martin-Löf P-random* if α is Martin-Löf random with respect to λ_P. □

4 The properties of Martin-Löf P-randomness

In order to obtain the results in this paper, we need several theorems and notions on Martin-Löf P-randomness from Tadaki [21, 22, 26]. Originally, these theorems and notions played a key part in developing the theory of *operational characterization of the notion of probability* based on Martin-Löf P-randomness in Tadaki [21, 22, 26]. We enumerate these theorems and notions in this section.

4.1 The law of large numbers

The following theorem shows that the law of large numbers holds for an arbitrary Martin-Löf P-randomness infinite sequence. For the proof of Theorem 4.2, see Tadaki [26, Theorem 11].

Theorem 4.1 (The law of large numbers, Tadaki [21])**.** *Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. For every $\alpha \in \Omega^\infty$, if α is Martin-Löf P-random then for every $a \in \Omega$ it holds that*

$$\lim_{n\to\infty} \frac{N_a(\alpha\restriction_n)}{n} = P(a),$$

where $N_a(\sigma)$ denotes the number of the occurrences of a in σ for every $a \in \Omega$ and every $\sigma \in \Omega^*$. □

4.2 Conditional probability

The notion of *conditional probability* in a finite probability space can be represented by the notion of Martin-Löf P-randomness in a natural manner as follows.

First, we recall the notion of conditional probability in a finite probability space. Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. Let $B \subset \Omega$ be an event on 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 $P(A|B)$, 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 an infinite sequence in B^∞ obtained from α by eliminating all elements of $\Omega \setminus B$ occurring in α. If α is Martin-Löf P-random for the finite probability space P

and $P(B) > 0$, then α contains infinitely many elements from B due to Theorem 4.1 above. Therefore, $\mathrm{Filtered}_B(\alpha)$ is properly defined in this case. Note that the notion of $\mathrm{Filtered}_B(\alpha)$ in our theory is introduced by Tadaki [21], suggested by the notion of *partition* in the theory of *collectives* introduced by von Mises [31] (see Tadaki [26] for the detail).

We can then show Theorem 4.2 below, which states that Martin-Löf P-random sequences are *closed under conditioning*. For the proof of Theorem 4.2, see Tadaki [26, Theorem 18].

Theorem 4.2 (Closure property under conditioning, Tadaki [21])**.** *Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. Let $B \subset \Omega$ be an event on the finite probability space P with $P(B) > 0$. For every $\alpha \in \Omega^\infty$, if α is Martin-Löf P-random then* $\mathrm{Filtered}_B(\alpha)$ *is Martin-Löf P_B-random for the finite probability space P_B.* □

4.3 Independence of Martin-Löf *P*-random infinite sequences

Tadaki [22] proposed the notion of *independence* of Martin-Löf P-random infinite sequences. This notion is introduced in the following manner: Let $\Omega_1, \dots, \Omega_K$ be alphabets. For any $\alpha_1 \in \Omega_1^\infty, \dots, \alpha_K \in \Omega_K^\infty$, we use

$$\alpha_1 \times \cdots \times \alpha_K$$

to denote an infinite sequence α over $\Omega_1 \times \cdots \times \Omega_K$ such that

$$\alpha(n) = (\alpha_1(n), \dots, \alpha_K(n))$$

for every $n \in \mathbb{N}^+$. On the other hand, for any $P_1 \in \mathbb{P}(\Omega_1), \dots, P_K \in \mathbb{P}(\Omega_K)$, we use

$$P_1 \times \cdots \times P_K$$

to denote a finite probability space $Q \in \mathbb{P}(\Omega_1 \times \cdots \times \Omega_K)$ such that

$$Q(a_1, \dots, a_K) = P_1(a_1) \cdots P_K(a_K)$$

for every $a_1 \in \Omega_1, \dots, a_K \in \Omega_K$.

Definition 4.3 (Independence of Martin-Löf P-random infinite sequences, Tadaki [22])**.** Let $\Omega_1, \dots, \Omega_K$ be alphabets, and let $P_1 \in \mathbb{P}(\Omega_1), \dots, P_K \in \mathbb{P}(\Omega_K)$. For each $k = 1, \dots, K$, let α_k be a Martin-Löf P_k-random infinite sequence over Ω_k. We say that $\alpha_1, \dots, \alpha_K$ are *independent* if $\alpha_1 \times \cdots \times \alpha_K$ is Martin-Löf $P_1 \times \cdots \times P_K$-random. □

Note that the notion of the independence of Martin-Löf P-random infinite sequences is introduced by Tadaki [22], suggested by the notion of *independence* of *collectives* in the theory of collectives introduced by von Mises [31] (see Tadaki [26] for the detail).

5 Postulates of quantum mechanics

In this section, we review the central postulates of (the conventional) quantum mechanics. For simplicity, in this paper we consider the postulates of quantum mechanics for a *finite-dimensional* quantum system, i.e., a quantum system whose state space is a *finite-dimensional* Hilbert space. Nielsen and Chuang [16] treat thoroughly the postulates of (the conventional) quantum mechanics in the finite-dimensional case, as a textbook of the field of quantum computation and quantum information in which such a case is typical. In this paper we refer to the postulates of the *conventional* quantum mechanics in the form presented in Nielsen and Chuang [16, Chapter 2].

The first postulate of quantum mechanics is about *state space* and *state vector*.

Postulate 1 (State space and state vector)**.** Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *state vector*, which is a unit vector in the system's state space. □

The second postulate of quantum mechanics is about the *composition* of systems.

Postulate 2 (Composition of systems)**.** The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state $|\Psi_i\rangle$, then the joint state of the total system is $|\Psi_1\rangle \otimes |\Psi_2\rangle \otimes \cdots \otimes |\Psi_n\rangle$. □

The third postulate of quantum mechanics is about the *time-evolution* of *closed* quantum systems.

Postulate 3 (Unitary time-evolution)**.** The evolution of a *closed* quantum system is described by a *unitary transformation*. That is, the state $|\Psi_1\rangle$ of the system at time t_1 is related to the state $|\Psi_2\rangle$ of the system at time t_2 by a unitary operator U, which depends only on the times t_1 and t_2, in such a way that $|\Psi_2\rangle = U|\Psi_1\rangle$. □

The forth postulate of quantum mechanics is about *measurements* on quantum systems. This is the so-called *Born rule*, i.e, *the probability interpretation of the wave function*.

Postulate 4 (The Born rule)**.** A quantum measurement is described by a collection $\{M_m\}_{m\in\Omega}$ of *measurement operators* which satisfy the *completeness equation*,

$$\sum_{m\in\Omega} M_m^\dagger M_m = I.$$

These are operators acting on the state space of the system being measured. The set of possible outcomes of the measurement equals the finite set Ω. If the state of the quantum system is $|\Psi\rangle$ immediately before the measurement then the probability that result m occurs is given by

$$\langle\Psi|M_m^\dagger M_m|\Psi\rangle,$$

and the state of the system after the measurement is

$$\frac{M_m|\Psi\rangle}{\sqrt{\langle\Psi|M_m^\dagger M_m|\Psi\rangle}}.$$

□

Thus, the Born rule, Postulate 4, uses the *notion of probability*. However, the operational characterization of the notion of probability is not given in the Born rule, and therefore the relation of its statement to a specific infinite sequence of outcomes of quantum measurements which are being generated by an infinitely repeated measurements is *unclear*. Tadaki [20, 23, 24, 25, 27, 28, 29] fixed this point.

In this paper as well as our former works [20, 23, 24, 25, 27, 28, 29], *we keep Postulates 1, 2, and 3 in their original forms without any modifications*. The principle of typicality, Postulate 5 below, is proposed as a *refinement* of Postulate 4 to replace it, based on the notion of *Martin-Löf randomness with respect to a probability measure*.

6 The principle of typicality

In what follows, we review the framework of the *principle of typicality* introduced by Tadaki [24]. It is a refinement of the *many-worlds interpretation of quantum mechanics* (*MWI*, for short) introduced by Everett [11] in 1957. More specifically, Tadaki [24] refined the argument of MWI by adding the principle of typicality to it. For the detail of the difference between the framework of the principle of typicality and that of the original MWI, see Tadaki [29].

To begin with, we review the framework of MWI introduced by Everett [11]. Actually, we review the *reformulation of the original framework of MWI in a form of mathematical rigor*, which was developed by Tadaki [29] from a modern point of view. The point of the rigorous treatment of the framework of MWI by Tadaki [29] is the use of the notion of *probability measure representation* and its *induction of probability measure*, as presented in Subsection 2.2.

We stress that MWI is more than just an interpretation of quantum mechanics. It aims to derive Postulate 4, the Born rule, from the remaining postulates, i.e., Postulates 1, 2, and 3. In this sense, Everett [11] proposed MWI as a “metatheory” of quantum mechanics. The point is that in MWI the measurement process is fully treated as the interaction between a system being measured and an apparatus measuring it, based only on Postulates 1, 2, and 3. Then MWI tries to derive Postulate 4 in such a setting.

Now, according to Tadaki [29], let us investigate the setting of MWI *in terms of our terminology in a form of mathematical rigor*. Let $\mathcal{S}$ be an arbitrary quantum system with state space $\mathcal{H}$ of finite dimension. Consider a measurement over $\mathcal{S}$ described by arbitrary measurement operators $\{M_m\}_{m\in\Omega}$ satisfying the completeness equation,

$$\sum_{m\in\Omega} M_m^\dagger M_m = I. \tag{5}$$

Here, Ω is the set of all possible outcomes of the measurement.[1] Let $\mathcal{A}$ be an apparatus performing the measurement described by $\{M_m\}_{m\in\Omega}$, which is a quantum system with state space $\overline{\mathcal{H}}$.[2] According to Postulates 1, 2, and 3, the measurement process of the

[1]The set Ω is finite, and therefore it is an alphabet.

[2]The dimension of the state space $\overline{\mathcal{H}}$ of the apparatus $\mathcal{A}$ is not necessarily finite. Even in the case where the state space $\overline{\mathcal{H}}$ is of infinite dimension, the mathematical subtleness which arises from the infinite dimensionality does not matter, to the extent of our treatment of $\overline{\mathcal{H}}$ and operators on it in this paper.

measurement described by the measurement operators $\{M_m\}_{m\in\Omega}$ is described by a unitary operator U such that

$$U|\Psi\rangle \otimes |\Phi_{\mathrm{init}}\rangle = \sum_{m\in\Omega}(M_m|\Psi\rangle) \otimes |\Phi[m]\rangle \tag{6}$$

for every $|\Psi\rangle \in \mathcal{H}$ (von Neumann [32], Tadaki [29]). Actually, U describes the interaction between the system $\mathcal{S}$ and the apparatus $\mathcal{A}$. The vector $|\Phi_{\mathrm{init}}\rangle \in \overline{\mathcal{H}}$ is the initial state of the apparatus $\mathcal{A}$, and $|\Phi[m]\rangle \in \overline{\mathcal{H}}$ is a final state of the apparatus $\mathcal{A}$ for each $m \in \Omega$, with $\langle\Phi[m]|\Phi[m']\rangle = \delta_{m,m'}$. For every $m \in \Omega$, the state $|\Phi[m]\rangle$ indicates that *the apparatus $\mathcal{A}$ records the value m as the measurement outcome*. By the unitary interaction (6) as a measurement process, a correlation (i.e., entanglement) is generated between the system and the apparatus.

In the framework of MWI, we consider countably infinite copies of the system $\mathcal{S}$, and consider a countably infinite repetition of the measurements described by the identical measurement operators $\{M_m\}_{m\in\Omega}$ performed over each of such copies in a sequential order, where each of the measurements is described by the unitary time-evolution (6). As repetitions of the measurement progressed, correlations between the systems and the apparatuses are being generated in sequence in the superposition of the total system consisting of the systems and the apparatuses. The detail is described as follows.

For each $n \in \mathbb{N}^+$, let $\mathcal{S}_n$ be the nth copy of the system $\mathcal{S}$ and $\mathcal{A}_n$ the nth copy of the apparatus $\mathcal{A}$. Each $\mathcal{S}_n$ is prepared in a state $|\Psi_n\rangle$ while all $\mathcal{A}_n$ are prepared in an identical state $|\Phi_{\mathrm{init}}\rangle$. The measurement described by the measurement operators $\{M_m\}_{m\in\Omega}$ is performed over each $\mathcal{S}_n$ one by one in the increasing order of n, by interacting each $\mathcal{S}_n$ with $\mathcal{A}_n$ according to the unitary time-evolution (6). For each $n \in \mathbb{N}^+$, let $\mathcal{H}_n$ be the state space of the total system consisting of the first n copies $\mathcal{S}_1, \mathcal{A}_1, \mathcal{S}_2, \mathcal{A}_2, \dots, \mathcal{S}_n, \mathcal{A}_n$ of the system $\mathcal{S}$ and the apparatus $\mathcal{A}$. These successive interactions between the copies of the system $\mathcal{S}$ and the apparatus $\mathcal{A}$ as measurement processes proceed in the following manner:

The starting state of the total system, which consists of $\mathcal{S}_1$ and $\mathcal{A}_1$, is $|\Psi_1\rangle \otimes |\Phi_{\mathrm{init}}\rangle \in \mathcal{H}_1$. Immediately after the measurement described by $\{M_m\}_{m\in\Omega}$ over $\mathcal{S}_1$, the total system results in the state

$$\sum_{m_1\in\Omega} (M_{m_1}|\Psi_1\rangle) \otimes |\Phi[m_1]\rangle \in \mathcal{H}_1$$

by the interaction (6) as a measurement process. In general, immediately before performing the measurement described by $\{M_m\}_{m\in\Omega}$ over $\mathcal{S}_n$, the state of the total system, which consists of $\mathcal{S}_1, \mathcal{A}_1, \mathcal{S}_2, \mathcal{A}_2, \dots, \mathcal{S}_n, \mathcal{A}_n$, is

$$\sum_{m_1,\dots,m_{n-1}\in\Omega} (M_{m_1}|\Psi_1\rangle) \otimes \cdots \otimes (M_{m_{n-1}}|\Psi_{n-1}\rangle) \otimes |\Psi_n\rangle \otimes |\Phi[m_1]\rangle \otimes \cdots \otimes |\Phi[m_{n-1}]\rangle \otimes |\Phi_{\mathrm{init}}\rangle$$

in $\mathcal{H}_n$, where $|\Psi_n\rangle$ is the initial state of $\mathcal{S}_n$ and $|\Phi_{\mathrm{init}}\rangle$ is the initial state of $\mathcal{A}_n$. Immediately after the measurement described by $\{M_m\}_{m\in\Omega}$ over $\mathcal{S}_n$, the total system results in the state

$$\begin{aligned}
&\sum_{m_1,\dots,m_n\in\Omega} (M_{m_1}|\Psi_1\rangle)\otimes\cdots\otimes(M_{m_n}|\Psi_n\rangle)\otimes|\Phi[m_1]\rangle\otimes\cdots\otimes|\Phi[m_n]\rangle \\
=&\sum_{m_1,\dots,m_n\in\Omega} (M_{m_1}|\Psi_1\rangle)\otimes\cdots\otimes(M_{m_n}|\Psi_n\rangle)\otimes|\Phi[m_1\dots m_n]\rangle \qquad (7)
\end{aligned}$$

in $\mathcal{H}_n$, by the interaction (6) as a measurement process between the system $\mathcal{S}_n$ prepared in the state $|\Psi_n\rangle$ and the apparatus $\mathcal{A}_n$ prepared in the state $|\Phi_{\mathrm{init}}\rangle$. The vector $|\Phi[m_1\dots m_n]\rangle$ denotes the vector $|\Phi[m_1]\rangle\otimes\cdots\otimes|\Phi[m_n]\rangle$ which represents the state of $\mathcal{A}_1,\dots,\mathcal{A}_n$. This state indicates that *the apparatuses $\mathcal{A}_1,\dots,\mathcal{A}_n$ record the values $m_1\dots m_n$ as the measurement outcomes over $\mathcal{S}_1,\dots,\mathcal{S}_n$, respectively.*

In the superposition (7), on letting $n\to\infty$, the length of the records $m_1\dots m_n$ of the values as the measurement outcomes in the apparatuses $\mathcal{A}_1,\dots,\mathcal{A}_n$ diverges to infinity. The consideration of this limiting case results in the definition of a *world*. Namely, a *world* is defined as an infinite sequence of records of the values as the measurement outcomes in the apparatuses. Thus, in the case described so far, a world is an infinite sequence over Ω, and the finite records $m_1\dots m_n$ in each state $|\Phi[m_1\dots m_n]\rangle$ in the superposition (7) of the total system is a *prefix* of a world.

Then, for aiming at deriving Postulate 4, MWI assigns "weight" to each of worlds. Namely, it introduces a *probability measure* on the set of all worlds in the following manner: First, MWI introduces a probability measure representation on the set of prefixes of worlds, i.e., the set Ω^* in this case. This probability measure representation is given by a function $r\colon\Omega^*\to[0,1]$ with

$$r(m_1\dots m_n)=\prod_{k=1}^{n}\langle\Psi_k|M_{m_k}^{\dagger}M_{m_k}|\Psi_k\rangle, \qquad (8)$$

which is the square of the norm of each state $(M_{m_1}|\Psi_1\rangle)\otimes\cdots\otimes(M_{m_n}|\Psi_n\rangle)\otimes|\Phi[m_1\dots m_n]\rangle$ in the superposition (7). Using the completeness equation (5), it is easy to check that r is certainly a probability measure representation over Ω. We call the probability measure representation r *the measure representation for the prefixes of worlds*. Then MWI tries to derive Postulate 4 by adopting the probability measure *induced by* the measure representation r for the prefixes of worlds as *the probability measure on the set of all worlds*.

We summarize the above consideration and clarify the definitions of the notion of *world* and the notion of *the measure representation for the prefixes of worlds* as in the following.

Definition 6.1 (The measure representation for the prefixes of worlds, Tadaki [29]). Consider an arbitrary finite-dimensional quantum system $\mathcal{S}$ and a measurement over $\mathcal{S}$ described by arbitrary *measurement operators* $\{M_m\}_{m\in\Omega}$, where the measurement process is described by (6) as an interaction of the system $\mathcal{S}$ with an apparatus $\mathcal{A}$. We suppose the following situation:

(i) There are countably infinite copies $\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3\dots$ of the system $\mathcal{S}$ and countably infinite copies $\mathcal{A}_1,\mathcal{A}_2,\mathcal{A}_3,\dots$ of the apparatus $\mathcal{A}$.

(ii) For each $n \in \mathbb{N}^+$, the system $\mathcal{S}_n$ is prepared in a state $|\Psi_n\rangle$,[3] while the apparatus $\mathcal{A}_n$ is prepared in a state $|\Phi_{\text{init}}\rangle$, and then the measurement described by $\{M_m\}_{m\in\Omega}$ is performed over $\mathcal{S}_n$ by interacting it with the apparatus $\mathcal{A}_n$ according to the unitary time-evolution (6).

(iii) Starting the measurement described by $\{M_m\}_{m\in\Omega}$ over $\mathcal{S}_1$, the measurement described by $\{M_m\}_{m\in\Omega}$ over each $\mathcal{S}_n$ is performed in the increasing order of n.

We then note that, for each $n \in \mathbb{N}^+$, immediately after the measurement described by $\{M_m\}_{m\in\Omega}$ over $\mathcal{S}_n$, the state of the total system consisting of $\mathcal{S}_1, \mathcal{A}_1, \mathcal{S}_2, \mathcal{A}_2, \dots, \mathcal{S}_n, \mathcal{A}_n$ is

$$|\Theta_n\rangle := \sum_{m_1,\dots,m_n\in\Omega} |\Theta(m_1,\dots,m_n)\rangle,$$

where

$$|\Theta(m_1,\dots,m_n)\rangle := (M_{m_1}|\Psi_1\rangle) \otimes \cdots \otimes (M_{m_n}|\Psi_n\rangle) \otimes |\Phi[m_1]\rangle \otimes \cdots \otimes |\Phi[m_n]\rangle.$$

The vectors $M_{m_1}|\Psi_1\rangle, \dots, M_{m_n}|\Psi_n\rangle$ are states of $\mathcal{S}_1, \dots, \mathcal{S}_n$, respectively, and the vectors $|\Phi[m_1]\rangle, \dots, |\Phi[m_n]\rangle$ are states of $\mathcal{A}_1, \dots, \mathcal{A}_n$, respectively. The state vector $|\Theta_n\rangle$ of the total system is normalized while each of the vectors $\{|\Theta(m_1,\dots,m_n)\rangle\}_{m_1,\dots,m_n\in\Omega}$ is not necessarily normalized. Then, *the measure representation for the prefixes of worlds* is defined as a function $p\colon \Omega^* \to [0,1]$ such that

$$p(m_1 \dots m_n) = \langle\Theta(m_1,\dots,m_n)|\Theta(m_1,\dots,m_n)\rangle \tag{9}$$

for every $n \in \mathbb{N}^+$ and every $m_1, \dots, m_n \in \Omega$. Moreover, an infinite sequence over Ω, i.e., an infinite sequence of possible outcomes of the measurement described by $\{M_m\}_{m\in\Omega}$, is called a *world*. $\square$

As mentioned above, the original MWI by Everett [11] aimed to derive the Born rule, i.e., Postulate 4, from the remaining postulates. However, it would seem impossible to do this for several reasons. See Tadaki [29] for these reasons. Instead, it is appropriate to introduce an additional postulate in the framework of MWI, in order to overcome the defect of the original MWI and to make quantum mechanics *operationally perfect*. Thus, we put forward the following postulate.

Postulate 5 (The principle of typicality, Tadaki [24])**.** Our world is *typical*. Namely, our world is Martin-Löf random with respect to the probability measure on the set of all worlds, induced by the measure representation for the prefixes of worlds, in the superposition of the total system which consists of systems being measured and apparatuses measuring them. $\square$

For the comprehensive arguments of the validity of the principle of typicality, see Tadaki [29]. For example, based on the results of Tadaki [21, 22, 26], we can see that Postulate 5 is certainly a refinement of Postulate 4, the Born rule, from the point of view of our intuitive understanding of the notion of probability.

[3] In Definition 6.1, all $|\Psi_n\rangle$ are not required to be an identical state. In the application of the principle of typicality in this paper, which is presented in Section 7, all $|\Psi_n\rangle$ are chosen to be an identical state.

7 Refinement of the argument of quantum mechanics to violate Bell's inequality

In this section, based on the *principle of typicality*, we *refine* and *reformulate* that argument of quantum mechanics to violate Bell's inequality which is given in Section 2.6 of Nielsen and Chuang [16]. Thus, according to Nielsen and Chuang [16, Section 2.6], we consider Protocol 7.1 below due to Bell [1], Clauser, et al. [6], and Nielsen and Chuang [16].

First, we fix some notation. Let $|0\rangle$ and $|1\rangle$ be an orthonormal basis of the state space of a single qubit system. Based on them we define a state $|+\rangle$ of a system of a single qubit by

$$|+\rangle := \frac{|0\rangle + |1\rangle}{\sqrt{2}},$$

and define the *Bell state* $|\beta_{11}\rangle$ of a system of two qubits by

$$|\beta_{11}\rangle := \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle}{\sqrt{2}}.$$

Pauli matrices X, Y, Z are defined by

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We deal with four observables R, Q, S, T of a system of a single qubit defined by the following way:[4]

$$R := X, \quad Q := Z, \quad S := -\frac{1}{\sqrt{2}}X - \frac{1}{\sqrt{2}}Z, \quad T := -\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Z.$$

Protocol 7.1. The protocol involves three parties, Charlie, Alice, and Bob. They together repeat the following procedure *forever*.

Step 1: Charlie prepares a quantum system of two qubits in the state $|\beta_{11}\rangle$.

Step 2: Charlie passes the first qubit to Alice, and the second qubit to Bob.

Then Alice and Bob do the following respectively. On the one hand, Alice does the following:

Step A3: Alice tosses a fair coin C to get outcome $c \in \{0, 1\}$.

Step A4: Alice performs the measurement of either R or Q over the first qubit to obtain outcome $m \in \{-1, 1\}$, depending on $c = 0$ or 1.

On the other hand, Bob does the following:

Step B3: Bob tosses a fair coin D to get outcome $d \in \{0, 1\}$.

Step B4: Bob performs the measurement of either S or T over the second qubit to obtain outcome $n \in \{-1, 1\}$, depending on $d = 0$ or 1.

[4] We intentionally put R before Q at variance with Alphabetical order, but we use the same notation exactly as in Section 2.6 of Nielsen and Chuang [16].

While repeating the above procedure forever, Alice and Bob together calculate the conditional averages $\langle RS\rangle$, $\langle QS\rangle$, $\langle RT\rangle$, and $\langle QT\rangle$, which are defined by the equations (17) or the equations (21) below and whose meaning is explained below. $\square$

In what follows, we analyze Protocol 7.1 in our framework of quantum mechanics based on the principle of typicality, Postulate 5, together with Postulates 1, 2, and 3. To complete this, *we have to implement everything in Steps A1 and A2 and Steps B1 and B2 of Protocol 7.1 by unitary time-evolution*.

7.1 Unitary implementation of all steps by Alice and Bob

We denote the system of the first qubit which Charlie sends to Alice in Step 2 by $\mathcal{Q}_{\mathrm{A}}$ with state space $\mathcal{H}_{\mathrm{A}}$, and denote the system of the second qubit which Charlie sends to Bob in Step 2 by $\mathcal{Q}_{\mathrm{B}}$ with state space $\mathcal{H}_{\mathrm{B}}$. We analyze Steps A1 and A2 and Steps B1 and B2 of Protocol 7.1 in our framework of quantum mechanics based on the principle of typicality. In particular, we realize each of the coin tossings in Steps A3 and B3 by a measurement of the observable $|1\rangle\langle 1|$ over a system of a single qubit in the state $|+\rangle$. We will then describe all the measurement processes during Steps A1 and A2 and Steps B1 and B2 as a single unitary interaction between systems and apparatuses.

On the one hand, each of Steps A3 and A4 by Alice is implemented by a unitary time-evolution in the following manner:

Unitary implementation of Step A3 by Allice. To realize the coin tossing by Alice in Step A3 of Protocol 7.1 we make use of a measurement over a single qubit system. Namely, to implement the Step A3 we introduce a single qubit system $\mathcal{Q}_{\mathrm{A3}}$ with state space $\mathcal{H}_{\mathrm{A3}}$, and perform a measurement over the system $\mathcal{Q}_{\mathrm{A3}}$ described by a unitary time-evolution:

$$U_{\mathrm{A3}}|c\rangle \otimes |\Phi_{\mathrm{A3}}^{\mathrm{init}}\rangle = |c\rangle \otimes |\Phi_{\mathrm{A3}}[c]\rangle$$

for every $c \in \{0,1\}$, where $|c\rangle \in \mathcal{H}_{\mathrm{A3}}$. The vector $|\Phi_{\mathrm{A3}}^{\mathrm{init}}\rangle$ is the initial state of an apparatus $\mathcal{A}_{\mathrm{A3}}$ measuring $\mathcal{Q}_{\mathrm{A3}}$, and $|\Phi_{\mathrm{A3}}[c]\rangle$ is a final state of the apparatus $\mathcal{A}_{\mathrm{A3}}$ for each $c \in \{0,1\}$.[5] Prior to the measurement, the system $\mathcal{Q}_{\mathrm{A3}}$ is prepared in the state $|+\rangle \in \mathcal{H}_{\mathrm{A3}}$.

Unitary implementation of Step A4 by Alice. Let $\{E^A_{0,m}\}$ and $\{E^A_{1,m}\}$ be the collections of projectors corresponding to the observables R and Q, respectively. Namely, we have

$$R = E^A_{0,+1} - E^A_{0,-1}, \quad Q = E^A_{1,+1} - E^A_{1,-1}.$$

The switching of these two types of measurements in Step A4, depending on the outcome c in Step A3, is realized by a unitary time-evolution:

$$U_{\mathrm{A4}}|\Theta\rangle \otimes |\Phi_{\mathrm{A3}}[c]\rangle = (V_c^{\mathrm{A}}|\Theta\rangle) \otimes |\Phi_{\mathrm{A3}}[c]\rangle \tag{10}$$

for every $c \in \{0,1\}$ and every state $|\Theta\rangle$ of the composite system consisting of the system $\mathcal{Q}_{\mathrm{A}}$ and the apparatus $\mathcal{A}_{\mathrm{A4}}$ explained below. The operator V_c^{A} appearing in (10) describes

[5] We assume, of course, the orthogonality of the final states $|\Phi_{\mathrm{A3}}[c]\rangle$, i.e., the property that $\langle\Phi_{\mathrm{A3}}[c]|\Phi_{\mathrm{A3}}[c']\rangle = \delta_{c,c'}$. Furthermore, we assume the orthogonality of the finial states for each of all apparatuses which appear in the rest of this section.

a unitary time-evolution of the composite system consisting of the system $\mathcal{Q}_{\rm A}$ and an apparatus $\mathcal{A}_{\rm A4}$ measuring $\mathcal{Q}_{\rm A}$, and is defined by the equation:

$$V_c^A|\psi\rangle \otimes |\Phi_{\rm A4}^{\rm init}\rangle = \sum_{m=\pm 1} (E_{c,m}^A|\psi\rangle) \otimes |\Phi_{\rm A4}[m]\rangle$$

for every $c \in \{0,1\}$. The vector $|\Phi_{\rm A4}^{\rm init}\rangle$ is the initial state of the apparatus $\mathcal{A}_{\rm A4}$, and $|\Phi_{\rm A4}[m]\rangle$ is a final state of the apparatus $\mathcal{A}_{\rm A4}$ for each $m \in \{0,1\}$. Thus, the operator $V_c^{\rm A}$ describes the alternate measurement process of the qubit $\mathcal{Q}_{\rm A}$ sent from Charlie, depending on the outcome c, on Alice's side. Note that the unitarity of $U_{\rm A4}$ is confirmed by the following theorem.[6]

Theorem 7.2. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be complex Hilbert spaces of finite dimension. Let $\{|\Psi_1\rangle,\dots,|\Psi_N\rangle\}$ and $\{|\Phi_1\rangle,\dots,|\Phi_N\rangle\}$ be arbitrary two orthonormal bases of $\mathcal{H}_2$, and let $U_1,\dots,U_N$ be arbitrary N unitary operators on $\mathcal{H}_1$. Then*

$$U := U_1 \otimes |\Psi_1\rangle\langle\Phi_1| + \cdots + U_N \otimes |\Psi_N\rangle\langle\Phi_N|$$

is a unitary operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$, and

$$U(|\Theta\rangle \otimes |\Phi_k\rangle) = (U_k|\Theta\rangle) \otimes |\Psi_k\rangle$$

for every $|\Theta\rangle \in \mathcal{H}_1$ and every $k = 1,\dots,N$.

Proof. See Tadaki [29, Theorem 14.1] for the proof. □

The sequential applications of $U_{\rm A3}$ and $U_{\rm A4}$ result in:

$$U_A|\Psi\rangle \otimes |\Phi_{\rm A3}^{\rm init}\rangle \otimes |\Phi_{\rm A4}^{\rm init}\rangle = \sum_{c=0,1}\sum_{m=\pm 1} ((E_c \otimes E_{c,m}^A)|\Psi\rangle) \otimes |\Phi_{\rm A3}[c]\rangle \otimes |\Phi_{\rm A4}[m]\rangle \tag{11}$$

for every $|\Psi\rangle \in \mathcal{H}_{\rm A3} \otimes \mathcal{H}_{\rm A}$, where $E_c := |c\rangle\langle c|$.

On the other hand, each of Steps B3 and B4 by Bob is implemented by a unitary time-evolution in a similar manner as follows:

Unitary implementation of Step B3 by Bob. To realize the coin tossing by Bob in Step B3 of Protocol 7.1 we make use of a measurement over a single qubit system. Namely, to implement the Step B3 we introduce a single qubit system $\mathcal{Q}_{\rm B3}$ with state space $\mathcal{H}_{\rm B3}$, and perform a measurement over the system $\mathcal{Q}_{\rm B3}$ described by a unitary time-evolution:

$$U_{\rm B3}|d\rangle \otimes |\Phi_{\rm B3}^{\rm init}\rangle = |d\rangle \otimes |\Phi_{\rm B3}[d]\rangle$$

for every $d \in \{0,1\}$, where $|d\rangle \in \mathcal{H}_{\rm B3}$. The vector $|\Phi_{\rm B3}^{\rm init}\rangle$ is the initial state of an apparatus $\mathcal{A}_{\rm B3}$ measuring $\mathcal{Q}_{\rm B3}$, and $|\Phi_{\rm B3}[d]\rangle$ is a final state of the apparatus $\mathcal{A}_{\rm B3}$ for each $d \in \{0,1\}$. Prior to the measurement, the system $\mathcal{Q}_{\rm B3}$ is prepared in the state $|+\rangle \in \mathcal{H}_{\rm B3}$.

[6] As we mentioned in Footnote 2, the state space of an apparatus commonly has infinite dimension. Thus, to be precise, the unitarity of $U_{\rm A4}$ is confirmed by a theorem which is obtained by an *immediate* generalization of Theorem 7.2 over Hilbert spaces of infinite dimension.

Unitary implementation of Step B4 by Bob. Let $\{E^B_{0,m}\}$ and $\{E^B_{1,m}\}$ be the collections of projectors corresponding to the observables S and T, respectively. Namely, we have

$$S = E^B_{0,+1} - E^B_{0,-1}, \quad T = E^B_{1,+1} - E^B_{1,-1}.$$

The switching of these two types of measurements in Step B4, depending on the outcome d in Step B3, is realized by a unitary time-evolution:

$$U_{\mathrm{B4}}|\Theta\rangle \otimes |\Phi_{\mathrm{B3}}[d]\rangle = (V^{\mathrm{B}}_d|\Theta\rangle) \otimes |\Phi_{\mathrm{B3}}[d]\rangle \tag{12}$$

for every $d \in \{0,1\}$ and every state $|\Theta\rangle$ of the composite system consisting of the system $\mathcal{Q}_{\mathrm{B}}$ and the apparatus $\mathcal{A}_{\mathrm{B4}}$ explained below. The operator V^{B}_d appearing in (12) describes a unitary time-evolution of the composite system consisting of the system $\mathcal{Q}_{\mathrm{B}}$ and an apparatus $\mathcal{A}_{\mathrm{B4}}$ measuring $\mathcal{Q}_{\mathrm{B}}$, and is defined by the equation:

$$V^B_d|\psi\rangle \otimes |\Phi^{\mathrm{init}}_{\mathrm{B4}}\rangle = \sum_{n=\pm 1} (E^B_{d,n}|\psi\rangle) \otimes |\Phi_{\mathrm{B4}}[n]\rangle$$

for every $d \in \{0,1\}$. The vector $|\Phi^{\mathrm{init}}_{\mathrm{B4}}\rangle$ is the initial state of the apparatus $\mathcal{A}_{\mathrm{B4}}$, and $|\Phi_{\mathrm{B4}}[n]\rangle$ is a final state of the apparatus $\mathcal{A}_{\mathrm{B4}}$ for each $n \in \{0,1\}$. Thus, the operator V^{B}_c describes the alternate measurement process of the qubit $\mathcal{Q}_{\mathrm{B}}$ sent from Charlie, depending on the outcome d, on Bob's side.

The sequential applications of U_{B3} and U_{B4} result in:

$$U_B|\Psi\rangle \otimes |\Phi^{\mathrm{init}}_{\mathrm{B3}}\rangle \otimes |\Phi^{\mathrm{init}}_{\mathrm{B4}}\rangle = \sum_{d=0,1}\sum_{n=\pm 1} ((E_d \otimes E^B_{d,n})|\Psi\rangle) \otimes |\Phi_{\mathrm{B3}}[d]\rangle \otimes |\Phi_{\mathrm{B4}}[n]\rangle \tag{13}$$

for every $|\Psi\rangle \in \mathcal{H}_{\mathrm{B3}} \otimes \mathcal{H}_{\mathrm{B}}$, where E_d is the same as before.

Now, let us consider a single unitary time-evolution U_{AB} which describes all the measurement processes over the composite system consisting of $\mathcal{Q}_{\mathrm{A3}}$, $\mathcal{Q}_{\mathrm{A}}$, $\mathcal{Q}_{\mathrm{B3}}$, and $\mathcal{Q}_{\mathrm{B}}$. According to Postulates 2 and 3, we have that

$$U_{AB}(|\Theta_A\rangle \otimes |\Theta_B\rangle) = (U_A|\Theta_A\rangle) \otimes (U_B|\Theta_B\rangle)$$

for every state $|\Theta_A\rangle$ of the composite system consisting of the systems $\mathcal{Q}_{\mathrm{A3}}$ and $\mathcal{Q}_{\mathrm{A}}$ and the apparatus $\mathcal{A}_{\mathrm{A3}}$ and $\mathcal{A}_{\mathrm{A4}}$ and every state $|\Theta_B\rangle$ of the composite system consisting of the systems $\mathcal{Q}_{\mathrm{B3}}$ and $\mathcal{Q}_{\mathrm{B}}$ and the apparatus $\mathcal{A}_{\mathrm{B3}}$ and $\mathcal{A}_{\mathrm{B4}}$. Therefore, for each $|\Psi_A\rangle \in \mathcal{H}_{\mathrm{A3}} \otimes \mathcal{H}_{\mathrm{A}}$ and $|\Psi_B\rangle \in \mathcal{H}_{\mathrm{B3}} \otimes \mathcal{H}_{\mathrm{B}}$, using (11) and (13) we see that

$$\begin{aligned} &U_{AB}\left(|\Psi_A\rangle \otimes |\Psi_B\rangle \otimes |\Phi^{\mathrm{init}}\rangle\right) \\ &= \left(U_A|\Psi_A\rangle \otimes |\Phi^{\mathrm{init}}_{A3}\rangle \otimes |\Phi^{\mathrm{init}}_{A4}\rangle\right) \otimes \left(U_B|\Psi_B\rangle \otimes |\Phi^{\mathrm{init}}_{B3}\rangle \otimes |\Phi^{\mathrm{init}}_{B4}\rangle\right) \\ &= \sum_{(c,d,m,n)\in\Omega} \left(\left(E^A_c \otimes E^B_d \otimes E^A_{c,m} \otimes E^B_{d,n}\right)\left(|\Psi_A\rangle \otimes |\Psi_B\rangle\right)\right) \otimes |\Phi[c,d,m,n]\rangle, \end{aligned}$$

where Ω denotes the alphabet $\{0,1\}^2 \times \{+1,-1\}^2$, $|\Phi^{\mathrm{init}}\rangle$ denotes $|\Phi^{\mathrm{init}}_{A3}\rangle \otimes |\Phi^{\mathrm{init}}_{B3}\rangle \otimes |\Phi^{\mathrm{init}}_{A4}\rangle \otimes |\Phi^{\mathrm{init}}_{B4}\rangle$, and $|\Phi[c,d,m,n]\rangle$ denotes $|\Phi_{A3}[c]\rangle \otimes |\Phi_{B3}[d]\rangle \otimes |\Phi_{A4}[m]\rangle \otimes |\Phi_{B4}[n]\rangle$ for each $(c,d,m,n) \in \Omega$. It follows from the linearity of U_{AB} that

$$U_{AB}\left(|\Psi\rangle \otimes |\Phi^{\mathrm{init}}\rangle\right) = \sum_{(c,d,m,n)\in\Omega} \left(\left(E^A_c \otimes E^B_d \otimes E^A_{c,m} \otimes E^B_{d,n}\right)|\Psi\rangle\right) \otimes |\Phi[c,d,m,n]\rangle.$$

for every $|\Psi\rangle \in \mathcal{H}_{\mathrm{A3}} \otimes \mathcal{H}_{\mathrm{B3}} \otimes \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$. This U_{AB} describes the unitary time-evolution of Alice and Bob in the repeated once of the infinite repetition of that procedure in Protocol 7.1, which consists of the four steps: Steps A3 and A4 on Alice's side and Steps B3 and B4 on Bob's side. Totally, prior to the application of U_{AB}, the total system consisting of $\mathcal{Q}_{\mathrm{A3}}$, $\mathcal{Q}_{\mathrm{B3}}$, $\mathcal{Q}_{\mathrm{A}}$ and $\mathcal{Q}_{\mathrm{B}}$ is prepared in the state

$$|\Psi^{\mathrm{init}}\rangle := |+\rangle \otimes |+\rangle \otimes |\beta_{11}\rangle.$$

7.2 Application of the principle of typicality

The operator U_{AB} applying to the initial state $|\Psi^{\mathrm{init}}\rangle$ describe the *repeated once* of the infinite repetition of the measurements in Protocol 7.1, where the execution of Steps A3 and A4 and Steps B3 and B4 is infinitely repeated. Actually, we can check that a collection

$$\left\{E_c^A \otimes E_d^B \otimes E_{c,m}^A \otimes E_{d,n}^B\right\}_{(c,d,m,n)\in\Omega} \tag{14}$$

forms *measurement operators*. Thus, the total application U_{AB} of U_{A3}, U_{A4}, U_{B3}, and U_{B4} can be regarded as a *single measurement* which is described by the measurement operators (14) and whose all possible outcomes form the set Ω.

Hence, we can apply Definition 6.1 to this scenario of the setting of measurements. Therefore, according to Definition 6.1, we can see that a *world* is an infinite sequence over Ω and the probability measure induced by the *measure representation for the prefixes of worlds* is a Bernoulli measure λ_P on Ω^∞, where P is a finite probability space on Ω such that $P(c,d,m,n)$ is the square of the norm of the state

$$\left(\left(E_c^A \otimes E_d^B \otimes E_{c,m}^A \otimes E_{d,n}^B\right)|\Psi^{\mathrm{init}}\rangle\right) \otimes |\Phi[c,d,m,n]\rangle$$

for every $(c,d,m,n) \in \Omega$. Here Ω is the set of all possible records of the apparatus in the *repeated once* of the measurements. Let us calculate the explicit form of $P(c,d,m,n)$. It is easy to see that

$$P(c,d,m,n) = \frac{1}{16}\left[1 + \frac{(-1)^{cd}mn}{\sqrt{2}}\right] \tag{15}$$

for every $(c,d,m,n) \in \Omega$.

Now, let us apply Postulate 5, the *principle of typicality*, to the setting of measurements developed above. Let α be *our world* in the infinite repetition of the measurements in the above setting. This α is an infinite sequence over Ω consisting of records in the apparatuses which is being generated by the infinite repetition of the measurements described by the measurement operators (14) in the above setting. Since the Bernoulli measure λ_P on Ω^∞ is the probability measure induced by the measure representation for the prefixes of worlds in the above setting, it follows from Postulate 5 that *α is Martin-Löf P-random*.

7.3 Equality among the conditional averages by quantum mechanics

For each $c,d \in \{0,1\}$, we use $H(c,d)$ to denote the set $\{(c,d,m,n) \mid m = \pm 1 \ \&\ n = \pm 1\}$. Let $c,d \in \{0,1\}$. The set $H(c,d)$ consists of all records of the apparatuses $\mathcal{A}_{\mathrm{A3}}, \mathcal{A}_{\mathrm{B3}}, \mathcal{A}_{\mathrm{A4}}$, and $\mathcal{A}_{\mathrm{B4}}$, in a repeated once of the procedure in Protocol 7.1, where *Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3.* It follows from (15) that $P(H(c,d)) = 1/4$, as expected from the point of view of the conventional quantum mechanics, and moreover

$$P_{H(c,d)}(c,d,m,n) = \frac{P(c,d,m,n)}{P(H(c,d))} = \frac{1}{4}\left[1 + \frac{(-1)^{cd}mn}{\sqrt{2}}\right] \tag{16}$$

for every $m, n \in \{+1, -1\}$. Let $\alpha_{c,d} := \mathrm{Filtered}_{H(c,d)}(\alpha)$. Then, since α is Martin-Löf P-random, using Theorem 4.2 we have that $\alpha_{c,d}$ is Martin-Löf $P_{H(c,d)}$-random for the finite probability space $P_{H(c,d)}$ on $H(c,d)$. Recall that $\mathrm{Filtered}_{H(c,d)}(\alpha)$ is defined as an infinite sequence over the alphabet $H(c,d)$ obtained from α by eliminating all elements of $\Omega \setminus H(c,d)$ occurring in α. In other words, $\alpha_{c,d}$ is the subsequence of α consisting only of results that Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3. For each $k \in \mathbb{N}^+$, we denote the kth element $\alpha_{c,d}(k)$ of the subsequence $\alpha_{c,d}$ as

$$\alpha_{c,d}(k) = (c, d, m_{c,d}(k), n_{c,d}(k)).$$

In this manner, we introduce infinite sequences $m_{c,d}$ and $n_{c,d}$ over $\{+1, -1\}$. The sequence $m_{c,d}$ is *the infinite sequence of outcomes of the measurements by Alice in Step A4 over the infinite repetition of the procedure in Protocol 7.1 in our world, conditioned that Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3.* Similarly, the sequence $n_{c,d}$ is *the infinite sequence of outcomes of the measurements by Bob in Step B4 over the infinite repetition of the procedure in Protocol 7.1 in our world, conditioned that Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3.*

Based on the sequences $m_{c,d}$ and $n_{c,d}$, the *conditional averages* $\langle RS\rangle$, $\langle QS\rangle$, $\langle RT\rangle$, and $\langle QT\rangle$ are defined as follows:

$$\begin{aligned}
\langle RS\rangle &:= \lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L} m_{0,0}(k)n_{0,0}(k), \qquad & \langle QS\rangle &:= \lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L} m_{1,0}(k)n_{1,0}(k), \\
\langle RT\rangle &:= \lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L} m_{0,1}(k)n_{0,1}(k), \qquad & \langle QT\rangle &:= \lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L} m_{1,1}(k)n_{1,1}(k).
\end{aligned} \tag{17}$$

The value $\langle RS\rangle$ can be interpreted as *the average value of the product of outcome of the measurement of the observable R by Alice in Step A4 and outcome of the measurement of the observable S by Bob in Step B4, conditioned that Alice gets the outcome* 0 *in Step A3 and Bob gets the outcome* 0 *in Step B3.* Note here that, in Protocol 7.1, whenever Alice gets the outcome 0 she performs the measurement of R over the first qubit in Step A4, and whenever Bob gets the outcome 0 he performs the measurement of S over the second qubit in Step B4. An analogous interpretation can be made for each of $\langle QS\rangle$, $\langle RT\rangle$, and $\langle QT\rangle$.

The conditional averages are calculated as follows: Let $c, d \in \{0, 1\}$. Since the subsequence $\alpha_{c,d}$ is Martin-Löf $P_{H(c,d)}$-random, it follows from Theorem 4.1 and (16) that for every $m, n = \pm 1$ it holds that

$$\lim_{L\to\infty}\frac{N_{(c,d,m,n)}(\alpha_{c,d}\restriction_L)}{L} = \frac{1}{4}\left[1 + \frac{(-1)^{cd}mn}{\sqrt{2}}\right].$$

Recall here that $N_{(c,d,m,n)}(\alpha_{c,d}\restriction_L)$ denotes the number of the occurrences of (c, d, m, n) in the prefix of $\alpha_{c,d}$ of length L. Therefore, we have that

$$
\begin{aligned}
&\lim_{L\to\infty}\frac{1}{L}\sum_{k=1}^{L} m_{c,d}(k)n_{c,d}(k)\\
=&\lim_{L\to\infty}\frac{1}{L}\left[N_{(c,d,1,1)}(\alpha_{c,d}\restriction_L)+N_{(c,d,-1,-1)}(\alpha_{c,d}\restriction_L)-N_{(c,d,1,-1)}(\alpha_{c,d}\restriction_L)-N_{(c,d,-1,1)}(\alpha_{c,d}\restriction_L)\right]\\
=&(-1)^{cd}\frac{1}{\sqrt{2}}.
\end{aligned}
$$

Thus, the conditional averages $\langle RS\rangle$, $\langle QS\rangle$, $\langle RT\rangle$, and $\langle QT\rangle$ are calculated as follows:

$$
\langle RS\rangle=\frac{1}{\sqrt{2}},\quad \langle QS\rangle=\frac{1}{\sqrt{2}},\quad \langle RT\rangle=\frac{1}{\sqrt{2}},\quad \langle QT\rangle=-\frac{1}{\sqrt{2}}.
$$

Eventually, we have the following equality for the conditional averages as a result of the analysis of Protocol 7.1 in our framework of quantum mechanics based on the principle of typicality:

$$
\langle RS\rangle+\langle QS\rangle+\langle RT\rangle-\langle QT\rangle=2\sqrt{2}. \tag{18}
$$

This equality has exactly the same form as expected from the aspect of the conventional quantum mechanics, i.e., as the equation (2.230) in Section 2.6 of Nielsen and Chuang [16].

8 Refinement of the argument of local realism to lead to Bell's inequality

Nielsen and Chuang [16, Section 2.6] describes an analysis for Protocol 7.1 to lead to Bell's inequality, based on the *assumptions of local realism*. In this section, we *refine* and *reformulate* their analysis and the assumptions of local realism, in the framework of the *operational characterization of the notion of probability by algorithmic randomness* developed via our former works [21, 22, 26]. For that purpose, first we review the framework of the operational characterization of the notion of probability in the following subsection.

8.1 Operational characterization of the notion of probability

The notion of probability plays an important role in almost all areas of science and technology. 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 our former works [21, 22, 26, 30], based on the toolkit of algorithmic randomness we presented an operational characterization of the notion of probability for *discrete probability spaces*, including *finite probability spaces*. We used the notions of *Martin-Löf P-randomness* [26] and *its extension over Baire space* [30] to present the operational characterization.

According to Tadaki [26], in order to clarify our motivation and standpoint, and the meaning of the operational characterization, let us consider a familiar example of a probabilistic phenomenon. Specifically, we consider the repeated throwing of a fair die.

In this probabilistic phenomenon, as throwing progressed, a specific infinite sequence such as

$$3, 5, 6, 3, 4, 2, 2, 3, 6, 1, 5, 3, 5, 4, 1, \dots\dots\dots$$

is being generated, where each number is the outcome of the corresponding throwing of the die. Then the following naive question may arise naturally.

> **Question**: What property should this infinite sequence satisfy as a probabilistic phenomenon?

Via a series of works [21, 22, 26, 30], we tried to answer this question. We characterized the notion of probability as an infinite sequence of outcomes in a probabilistic phenomenon of a *specific mathematical property*. We called such an infinite sequence of outcomes the *operational characterization of the notion of probability*. As the specific mathematical property, we adopted the notion of *Martin-Löf P-randomness*.

In our former works [21, 22, 26, 30], we put forward this proposal as a thesis, i.e., as Thesis 1 below. We checked the validity of the thesis based on our intuitive understanding of the notion of probability. Furthermore, we characterized *equivalently* the basic notions in probability theory in terms of the operational characterization. Namely, we equivalently characterized the notion of the *independence* of random variables/events in terms of the operational characterization, and we represented the notion of *conditional probability* in terms of the operational characterization in a natural way. The existence of these equivalent characterizations confirms further the validity of the thesis. See Tadaki [26, 30] for the detail of the operational characterization of the notion of probability based on Martin-Löf P-randomness.

In this manner, as revealed by Tadaki [21, 22, 26, 30], the notion of Martin-Löf P-randomness is thought to reflect all the properties of the notion of probability based on our intuitive understanding of the notion of probability. Thus, *we propose that a Martin-Löf P-random sequence of elementary events gives an operational characterization of the notion of probability*, as follows.

Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. Accroding to Tadaki [26], consider an infinite sequence $\alpha \in \Omega^\infty$ of outcomes which is being generated by infinitely repeated 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 (Tadaki [21, 22, 26])**.** *Let Ω be an alphabet, and let $P \in \mathbb{P}(\Omega)$. An infinite sequence of outcomes in Ω which is being generated by infinitely repeated trials* described by *the finite probability space P on Ω is a Martin-Löf P-random sequence over Ω.* $\square$

Tadaki [26] demonstrated the validity of Thesis 1 from a variety of aspects.

8.2 Refinement of the assumptions of local realism

In what follows, we refine and reformulate the analysis for Protocol 7.1 to derive Bell's inequality, given in Section 2.6 of Nielsen and Chuang [16], in the framework of the operational characterization of the notion of probability reviewed in the preceding subsection.

Basically, we follow the flow of the argument of Nielsen and Chuang [16, Section 2.6], while refining it appropriately in terms of the operational characterization of the notion of probability. Hence, we proceed according to Nielsen and Chuang [16, Section 2.6], as follows:

We first forget all the knowledge of quantum mechanics. To obtain Bell's inequality, we analyze Protocol 7.1 based on "our common sense notions of how the world works." Thus, we perform "the common sense analysis" for Protocol 7.1. In doing so, we are implicitly assuming the following two assumptions:

The assumption of realism: The assumption that the observables R, Q, S, T have definite values r, q, s, t, respectively, which exist independent of observation.

The assumption of locality: The assumption that Alice performing her measurement does not influence the result of Bob's measurement, and vice versa.

These two assumptions together are known as the *assumptions of local realism*. See Nielsen and Chuang [16, Section 2.6] for the considerations for the assumptions of local realism.

Now, according to Nielsen and Chuang [16, Section 2.6], let us make "the common sense analysis" for Protocol 7.1, based the assumptions of local realism. In Protocol 7.1, Alice performs the measurement of either the observable R or Q over the first qubit in Step A4, while Bob performs the measurement of either the observable S or T over the second qubit in Spep B4. Based on the assumptions of local realism, we assume that each of the observables R, Q, S, and T has a specific value before the measurement, which is merely revealed by the measurement. In particular, *in the terminology of the conventional probability theory*, we assume that

$$p(r, q, s, t)$$

is the "probability" that, before the measurements are performed, the system is in a state where $R = r, Q = q, S = s$, and $T = t$. This "probability" may depend on how Charlie performs his preparation, and on experimental noise. In the framework of the operational characterization of the notion of probability, *based on Thesis 1*, the assumption above is refined and reformulated in the following form:

Assumption 1. *Let ω be an infinite sequence of the values (r, q, s, t) of the observables R, Q, S, T which is being generated by the infinite repetition of the procedure in Protocol 7.1. Then there exists a finite probability space H on $\{1, -1\}^4$ such that ω is a* Martin-Löf H-random *infinite sequence over $\{1, -1\}^4$.* □

Assumption 1 is an *operational refinement* of one of the consequences of the assumptions of local realism.

In Protocol 7.1, Alice tosses a fair coin C to get outcome $c \in \{0, 1\}$ in Step A3, while Bob tosses a fair coin D to get outcome $d \in \{0, 1\}$ in Step B3. In the framework of the operational characterization of the notion of probability, these probabilistic phenomena are refined and reformulated in the following form:

Assumption 2. *Let γ be an infinite binary sequence which is being generated by infinitely repeated tossing of the fair coin C by Alice in Protocol 7.1. Then the infinite sequence γ*

is a Martin-Löf U-random *sequence over* $\{0,1\}$*, where* U *is a finite probability space on* $\{0,1\}$ *such that* $U(0) = U(1) = 1/2$*. Namely,* γ *is* Martin-Löf random*.*

Similarly, let δ *be an infinite binary sequence which is being generated by infinitely repeated tossing of the fair coin* D *by Bob in Protocol 7.1. Then the infinite sequence* δ *is also a* Martin-Löf U-random *sequence over* $\{0,1\}$*.* □

Assumption 2 is just an implementation of Thesis 1 in an infinite repetition of tossing of a fair coin.

In order to advance "the common sense analysis" for Protocol 7.1 further in a rigorous manner, however, we need to make an additional assumption for the relation among the infinite sequences ω, γ, and δ. Namely, the infinite sequences ω, γ, and δ need to be *independent*. Thus, based on the notion of independence given in Definition 4.3, we assume the following:

Assumption 3. *The infinite sequences* ω*,* γ*, and* δ *are* independent*. Equivalently, the infinite sequence* $\omega \times \gamma \times \delta$ *over* $\{+1,-1\}^4 \times \{0,1\} \times \{0,1\}$ *is* Martin-Löf $H \times U \times U$-random*.* □

Assumption 3 is an *operational refinement* of one of the consequences of the assumption of locality.

Remark 8.1. In the context of the *relativized computation*, we can consider the notion of *Martin-Löf P-randomness relative to infinite sequences*. Theorem 38 of Tadaki [26] states that the notion of independence of Martin-Löf P-random infinite sequences can be equivalently represented by the notion of Martin-Löf P-randomness relative to infinite sequences. Thus, using Theorem 38 of Tadaki [26], Assumption 3 can be *equivalently* rephrased as the condition that

(i) the infinite sequence ω is Martin-Löf H-random,

(ii) the infinite sequence γ is Martin-Löf random *relative to* ω, and

(iii) the infinite sequence δ is Martin-Löf random *relative to* γ *and* ω. □

8.3 Refined derivation of Bell's inequality among the conditional averages

Based on Assumptions 1, 2, and 3, let us derive Bell's inequality in the framework of the operational characterization of the notion of probability. Let $\alpha := \omega \times \gamma \times \delta$. Note that

$$(H \times U \times U)((r,q,s,t),c,d) = \frac{1}{4}H(r,q,s,t) \tag{19}$$

for every $r,q,s,t \in \{+1,-1\}$ and every $c,d \in \{0,1\}$.

For each $c,d \in \{0,1\}$, we use $G(c,d)$ to denote the set $\{(x,c,d) \mid x \in \{+1,-1\}^4\}$. Let $c,d \in \{0,1\}$. The set $G(c,d)$ consists of all possible results in a repeated once of the procedure in Protocol 7.1, where *Alice gets the outcome* c *in Step A3 and Bob gets the outcome* d *in Step B3. In the terminology of the conventional probability theory,*

$(H \times U \times U)(G(c,d))$ is the "probability" that *Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3.* Actually, it follows from (19) that

$$(H \times U \times U)(G(c,d)) = \frac{1}{4},$$

as expected from the conventional probability theory. Thus, it follows from (19) again that

$$(H \times U \times U)_{G(c,d)}((r,q,s,t),c,d) = \frac{(H \times U \times U)((r,q,s,t),c,d)}{(H \times U \times U)(G(c,d))} = H(r,q,s,t) \quad (20)$$

for every $r,q,s,t \in \{+1,-1\}$. Let $\alpha_{c,d} := \mathrm{Filtered}_{G(c,d)}(\alpha)$. Then, since α is Martin-Löf $H \times U \times U$-random by Assumption 3, using Theorem 4.2 we have that $\alpha_{c,d}$ is Martin-Löf $(H \times U \times U)_{G(c,d)}$-random for the finite probability space $(H \times U \times U)_{G(c,d)}$ on $G(c,d)$. Recall that $\mathrm{Filtered}_{G(c,d)}(\alpha)$ is defined as an infinite sequence over the alphabet $G(c,d)$ obtained from α by eliminating all elements of $(\{+1,-1\}^4 \times \{0,1\} \times \{0,1\}) \setminus G(c,d)$ occurring in α. In other words, $\alpha_{c,d}$ is the subsequence of α consisting of results that Alice gets the outcome c in Step A3 and Bob gets the outcome d in Step B3. For each $k \in \mathbb{N}^+$, we denote the kth element $\alpha_{c,d}(k)$ of the subsequence $\alpha_{c,d}$ as

$$\alpha_{c,d}(k) = ((r_{c,d}(k), q_{c,d}(k), s_{c,d}(k), t_{c,d}(k)), c, d).$$

In this manner, we introduce infinite sequences $r_{c,d}$, $q_{c,d}$, $s_{c,d}$, and $t_{c,d}$ over $\{+1,-1\}$. The sequence $r_{c,d}$ is *the infinite sequence of values of the observable R over the infinite repetition of the procedure in Protocol 7.1 in our world, conditioned that Alice gets the outcome c at Step A3 and Bob gets the outcome d at Step B3.* The sequences $q_{c,d}$, $s_{c,d}$, and $t_{c,d}$ have an analogous meaning with respect to the observables Q, S, and T, respectively.

Based on the sequences $r_{c,d}$, $q_{c,d}$, $s_{c,d}$, and $t_{c,d}$, the *conditional averages* $\langle RS \rangle$, $\langle QS \rangle$, $\langle RT \rangle$, and $\langle QT \rangle$ are defined as follows:

$$\begin{aligned}
\langle RS \rangle &:= \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} r_{0,0}(k) s_{0,0}(k), \qquad & \langle QS \rangle &:= \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} q_{1,0}(k) s_{1,0}(k), \\
\langle RT \rangle &:= \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} r_{0,1}(k) t_{0,1}(k), \qquad & \langle QT \rangle &:= \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} q_{1,1}(k) t_{1,1}(k).
\end{aligned} \quad (21)$$

The value $\langle RS \rangle$ can be interpreted as *the average value of the product of outcome of the measurement of the observable R by Alice in Step A4 and outcome of the measurement of the observable S by Bob in Step B4, conditioned that Alice gets the outcome* 0 *in Step A3 and Bob gets the outcome* 0 *in Step B3.* Note here that, in Protocol 7.1, whenever Alice gets the outcome 0 she performs the measurement of R over the first qubit in Step A4, and whenever Bob gets the outcome 0 he performs the measurement of S over the second qubit in Step B4. An analogous interpretation can be made for each of $\langle QS \rangle$, $\langle RT \rangle$, and $\langle QT \rangle$.

The conditional averages are calculated as follows: Let $c,d \in \{0,1\}$. Since the subsequence $\alpha_{c,d}$ is Martin-Löf $(H \times U \times U)_{G(c,d)}$-random, it follows from Theorem 4.1 and (20) that for every $r,q,s,t = \pm 1$ it holds that

$$\lim_{L\to\infty} \frac{N_{((r,q,s,t),c,d)}(\alpha_{c,d}\restriction_L)}{L} = H(r,q,s,t).$$

Recall here that $N_{((r,q,s,t),c,d)}(\alpha_{c,d}\restriction_L)$ denotes the number of the occurrences of

$$((r,q,s,t),c,d)$$

in the prefix of $\alpha_{c,d}$ of length L. Thus, the conditional averages $\langle RS\rangle$, $\langle QS\rangle$, $\langle RT\rangle$, and $\langle QT\rangle$ are calculated as follows:

$$\begin{aligned}
\langle RS\rangle &= \lim_{L\to\infty}\frac{1}{L}\sum_{r,q,s,t\pm1} N_{((r,q,s,t),c,d)}(\alpha_{0,0}\restriction_L)rs = \sum_{r,q,s,t=\pm1} H(r,q,s,t)rs,\\
\langle QS\rangle &= \lim_{L\to\infty}\frac{1}{L}\sum_{r,q,s,t\pm1} N_{((r,q,s,t),c,d)}(\alpha_{1,0}\restriction_L)qs = \sum_{r,q,s,t\pm1} H(r,q,s,t)qs,\\
\langle RT\rangle &= \lim_{L\to\infty}\frac{1}{L}\sum_{r,q,s,t\pm1} N_{((r,q,s,t),c,d)}(\alpha_{0,1}\restriction_L)rt = \sum_{r,q,s,t\pm1} H(r,q,s,t)rt,\\
\langle QT\rangle &= \lim_{L\to\infty}\frac{1}{L}\sum_{r,q,s,t\pm1} N_{((r,q,s,t),c,d)}(\alpha_{1,1}\restriction_L)qt = \sum_{r,q,s,t\pm1} H(r,q,s,t)qt.
\end{aligned} \tag{22}$$

Note that $rs+qs+rt-qt=(r+q)s+(r-q)t$. Since $q,r=\pm1$, we have that either $(r+q)s=0$ or $(r-q)t=0$. It follows that $rs+qs+rt-qt=\pm2$. Thus, using (22) we see that

$$\begin{aligned}
&\langle RS\rangle+\langle QS\rangle+\langle RT\rangle-\langle QT\rangle\\
&=\sum_{r,q,s,t=\pm1} H(r,q,s,t)rs+\sum_{r,q,s,t\pm1} H(r,q,s,t)qs+\sum_{r,q,s,t\pm1} H(r,q,s,t)rt\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\sum_{r,q,s,t\pm1} H(r,q,s,t)qt\\
&=\sum_{r,q,s,t\pm1} H(r,q,s,t)(rs+qs+rt-qt)\\
&\le\sum_{r,q,s,t\pm1} H(r,q,s,t)\ \times 2\\
&=2.
\end{aligned}$$

Hence, we finally obtain the *Bell's inequality* (also known as the *CHSH inequality*),

$$\langle RS\rangle+\langle QS\rangle+\langle RT\rangle-\langle QT\rangle\le 2. \tag{23}$$

This inequality has the same form as expected from the conventional probability theory, by means of "the common sense analysis" based on the assumptions of local realism, which is performed in Nielsen and Chuang [16, Section 2.6] in a vague manner.

9 Conclusion

In this paper, we have refined and reformulated the argument of Bell's inequality versus quantum mechanics by algorithmic randomness. On the one hand, we have refined and reformulated local realism to lead to Bell's inequality, based on our operational characterization of the notion of probability by algorithmic randomness. On the other hand, we have refined and reformulated the corresponding argument of quantum mechanics to violate Bell's inequality in our framework of quantum mechanics based on the principle of typicality. Hence, in terms of algorithmic randomness, we have refined the derivation of the following fact: Local realism cannot recover the prediction of quantum mechanics.

The principle of typicality is a *unfined principle* which refines the Born rule, Postulate 4, and its related postulates about quantum measurements *in one lump* [29]. In this paper, we have successfully made an application of the principle of typicality to the argument of local realism versus quantum mechanics, *demonstrating how properly our framework based on the principle of typicality works in practical problems in quantum mechanics.* Thus, it seems further confirmed that *the principle of typicality together with Postulates 1, 2, and 3 forms quantum mechanics.*

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 18K03405. A version of the oral presentation of this work was made in LA Symposium 2019 held at the Research Institute for Mathematical Sciences, Kyoto University, on February 5–7, 2020.

References

[1] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, pp. 195–200, 1964.

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

[3] D. Bohm, *Quantum Theory.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1951.

[4] G. J. Chaitin, "On the length of programs for computing finite binary sequences," *J. Assoc. Comput. Mach.*, vol. 13, pp. 547–569, 1966.

[5] G. J. Chaitin, "A theory of program size formally identical to information theory," *J. Assoc. Comput. Mach.*, vol. 22, pp. 329–340, 1975.

[6] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, no. 15, pp. 880–884, 1969; Erratum, *Phys. Rev. Lett.*, vol. 24, no. 10, p. 549, 1970.

[7] B. S. DeWitt and N. Graham (eds.), *The Many-Worlds Interpretation of Quantum Mechanics.* Princeton University Press, Princeton, 1973.

[8] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. Oxford University Press, London, 1958.

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

[10] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, pp. 777–780, 1935.

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

[12] J. B. Hartle, "Quantum mechanics of individual systems," *Am. Journ. Phys.*, vol. 36, no. 8, pp. 704–712, 1968.

[13] A. N. Kolmogorov, *Foundations of the theory of probability*, Chelsea Publishing Company, New York, 1950.

[14] A. N. Kolmogorov, "Three approaches to the quantitative definition of information," *Problems Inform. Transmission*, vol. 1, no. 1, pp. 1–7, 1965.

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

[16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

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

[18] C.-P. Schnorr, "Process complexity and effective random tests," *J. Comput. System Sci.*, vol. 7, pp. 376–388, 1973.

[19] R. J. Solomonoff, "A formal theory of inductive inference. Part I and Part II," *Inform. and Control*, vol. 7, pp. 1–22, 1964; vol. 7, pp. 224–254, 1964.

[20] K. Tadaki, Reformulating quantum mechanics by algorithmic randomness. Presentation at Ninth International Conference on Computability, Complexity and Randomness (CCR 2014), June 9-13, 2014, Institute for Mathematical Sciences, National University of Singapore, Singapore.

[21] 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. Available at: `http://www2.odn.ne.jp/tadaki/5_4_1.pdf`

[22] K. Tadaki, An operational characterization of the notion of probability by algorithmic randomness and its application to cryptography. Proceedings of the 32nd Symposium on Cryptography and Information Security (SCIS2015), 2D4-3, January 20-23, 2015, Kokura, Japan. Available at: `http://www2.odn.ne.jp/tadaki/2D4-3.pdf`

[23] K. Tadaki, "A refinement of quantum mechanics by algorithmic randomness: extended abstract," *RIMS Kokyuroku* 1952, pp. 112–116, June 2015.

[24] K. Tadaki, The principle of typicality. Presentation at Eleventh International Conference on Computability, Complexity and Randomness (CCR 2016), January 4-8, 2016, University of Hawaii at Manoa, Honolulu, USA.

[25] K. Tadaki, A refinement of quantum mechanics by algorithmic randomness. Poster Presentation at 19th Conference on Quantum Information Processing (QIP 2016), January 10-15, 2016, The Banff Centre, Banff, Canada.

[26] K. Tadaki, "An operational characterization of the notion of probability by algorithmic randomness and its application," arXiv:1611.06201v1 [math.PR], November 2016.

[27] K. Tadaki, A refinement of quantum mechanics by algorithmic randomness. Proceedings of the 35th Quantum Information Technology Symposium (QIT35), November 24-25, 2016, High Energy Accelerator Research Organization, Tsukuba, Japan.

[28] K. Tadaki, A refinement of quantum mechanics by algorithmic randomness and its application to quantum cryptography. Proceedings of the 2017 Symposium on Cryptography and Information Security (SCIS2017), 4A1-2, January 24-27, 2017, Naha, Japan. Available at: `http://www2.odn.ne.jp/tadaki/4A1-2.pdf`

[29] K. Tadaki, "A refinement of quantum mechanics by algorithmic randomness," arXiv:1804.10174v1 [quant-ph], April 2018.

[30] K. Tadaki, "An operational characterization of the notion of probability by algorithmic randomness II: Discrete probability spaces," arXiv:1909.02854v1 [math.PR], August 2019.

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

[32] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, 1955.

Department of Computer Science
Chubu University
1200 Matsumoto-cho, Kasugai-shi
Aichi 487-8501
JAPAN
E-mail address: tadaki@isc.chubu.ac.jp
http://www2.odn.ne.jp/tadaki/