Spin Theory Based on the Extended Least Action Principle and Information Metrics: Quantization, Entanglement, and Bell Test With Time Delay

Jianhao M. Yang*
Qualcomm, San Diego, CA 92321, USA
(Dated: April 23, 2024)

A theory of electron spin is developed here based on the extended least action principle and assumptions of intrinsic angular momentum of an electron with random orientations. By incorporating appropriate relative entropy for the random orientations of intrinsic angular momentum in the extended least action principle, the theory recovers the quantum formulation of electron spin. The two-level quantization of spin measurement is a natural mathematical consequence instead of a postulate. The formulation of measurement probability when a second Stern-Gerlach apparatus is rotated relative to the first Stern-Gerlach apparatus, and the Schrödinger-Pauli equation, are also derived successfully. Furthermore, we provide an intuitive physical model and formulation to explain the entanglement phenomenon between two electron spins. In this model, spin entanglement is the consequence of correlation between the random orientations of the intrinsic angular momenta of the two electrons. Since the orientation is an intrinsic local property of electron, the correlation of orientations can be preserved even when the two electrons are remotely separated. Such a correlation can be manifested without causal effect. Owing to this orientation correlation, the Bell-CHSH inequality is shown to be violated in a Bell test. The standard quantum theory of electron spin can be considered as an ideal approximation of the present theory when certain conditions are taken to the limits. To test the difference between the present theory and the standard quantum theory, we propose two potential experiments. First, the observation of spin quantization depends on the its interactions with the measuring magnetic field; Second, in a typical Bell test that confirms the violation of Bell-CHSH inequality, the theory suggests that by adding a sufficiently large time delay before Bob's measurement, the Bell-CHSH inequality can become non-violated.

I. INTRODUCTION

In quantum mechanics, the spin of an electron is one of the most important physical observables to demonstrate the essences of quantum theory. These essential features include, for instance, 1.) the quantization of measurement outcome - measuring the spin of an electron always obtains two possible values, spin-up or spin-down, along the direction of measuring magnetic field; 2.) the need of multi-components wave functions in the Schrödinger-Pauli equation to describe the dynamics of an electron with spin; 3.) quantum entanglement - the singlet state of an electron pair is often used as example to illustrate the conceptual challenges such as in the Bohm version EPR thought experiment [1, 2], or to verify the violation of Bell inequality [3]. However, these important quantum features are either not derived from first principles or still have conceptual difficulty. Quantization of spin measurement outcome is normally introduced as postulate in standard quantum mechanics; Whether the wave function is epistemological or associated with physical reality is still under debate; The violation of Bell inequality confirms that spin correlation can be non-local in the sense that the correlation is inseparable even if the two electrons are space-like separated. Without an intuitive physical model, such inseparability is not possible to be comprehended in classical terms and still puzzling the physics community [4]. A more intuitive physical model and a clearer theory of electron spin to better explain these challenges are still much desirable in modern quantum physics.

Historically, a number of physical models for electron spin were proposed [5, 6]. Here we will briefly review two prominent ones. The first model is to consider the electron as a rotating rigid body with uniform distributed charge, and has been adopted in stochastic mechanism [7–9]. This model can calculate the relation between angular momentum and magnetic moment of an electron. By applying the theory of stochastic mechanics, the electron spin is treated as the averaged of hidden angular momentum variables with random orientations. Formulations equivalent to quantum mechanics can be recovered. However, one of the problems for this model is that when the radius of the rigid body is sufficiently small, the edge is moving faster than the speed of light [5]. In addition, the velocity fields in stochastic mechanics retain the non-local characteristics, a drawback that we attempt to avoid. Another model is the so-called "Zitterbewegung" model where the origin of spin is proposed as a result of helical motion of a "light-like" particle [10–13]. The particle is moving automatically in a circle with the speed of light. There is also a simpler model that considers an electron as a point particle with intrinsic magnetic moment and intrinsic angular momentum. Such a model offers no explanation on how the intrinsic angular moment or intrinsic magnetic moment originate.

Recent interests in searching for foundational principles of quantum theory from the information perspective [14–43] can offer new perspectives for the investigations of spin theory. One of the motivations is to reformu-

* jianhao.yang@alumni.utoronto.ca

late quantum mechanics with information theoretic principles such that some of the conceptual challenges can be resolved. Inspired by this program, an extended least action principle has been proposed [44]. The principle extends the classical least action principle by incorporating information metrics about vacuum fluctuations. Based on this principle, the non-relativistic quantum mechanics [44] and quantum scalar field theory [45] have been recovered, demonstrating the general applicability of the theoretical framework.

The goal of this paper is to apply the extended least action principle to develop the theory for electron spin with the goal of resolving some of the challenges mentioned earlier. This is achieved by introducing an additional assumption of intrinsic angular momentum with random orientations, and choosing proper relative entropy for the random orientations. We will show the results are indeed quite fruitful. Discreteness of measurement outcome is a natural mathematical result when the order of relative entropy approaches a limit. Formulation equivalent to quantum mechanics is obtained when the direction of magnetic field of a second Stern-Gerlach apparatus is rotated with an angle from the direction of magnetic field of the first Stern-Gerlach apparatus. Recursively applying the extended least action principle, we can derive the Schrödinger-Pauli equation.

With respect to spin entanglement between two electron spins, we will show that the root cause of entanglement is the correlation between the random variables, the orientations of the intrinsic angular momenta, of the two electrons. The correlation can be established during preparation. Crucially, since the orientation of the angular momentum is an intrinsic local property of the electron, the correlation can be maintained even though the two electrons are space-like separated and without interaction. There is no "spooky action in the remote" in the EPR experiment. Mathematically, we give an equivalent formulation of the four Bell states using probability density function instead of wave function. Based on the probability density function, Bell-CHSH inequality is proved to be violated in our theory due to the same root cause of entanglement. Since the correlation of the orientations of the angular momenta is preserved even when the two electrons are separated, the joint probability for the measurement outcome cannot be factorized into a product of two individual terms - a requirement for the Bell inequality to hold.

The standard quantum theory of electron spin can be considered as an idealized approximation of the present theory when certain conditions are taken to limits. When these conditions are not taken to the limits, the theory predicts results that are different from standard quantum mechanics. Two potential experiments are proposed to test the difference. In the first experiment, the Stern-Gerlach apparatus is configured with a sufficiently weak gradient of the inhomogeneous magnetic field along the z-axis; We expect the electron detector screen will exhibit continuous distribution of displacements along the z-axis around two areas, instead of only two discrete lines. The second experiment is to modify a typical Bell test experiment that has already confirmed the violation of Bell-CSHS inequality. Here, we purposely let Alice and Bob perform their measurements at different times on each entangled pair of electrons. While Alice performs her measurement at time t_a, Bob purposely delays his measurement at time $t_b = t_a + \Delta t$. The present theory predicts that when Δt is chosen properly, the Bell-CSHS inequality will not be violated.

The rest of the paper is organized as follows. In Section II, we review the extended least action principle and its underlying assumptions. Detailed calculation showing how the basic quantum theory is derived from the principle is provided in Appendix A. Section III is devoted to developing the spin model, where the extended least action principle is incorporated with Tsallis relative entropy, resulting in a probability density function for the random orientation. The probability density function becomes two Dirac delta functions when the order of relative entropy approaches a limit, corresponding to the quantization of measurement outcomes. The theory for rotation of the spin is developed in this section as well. In Section IV, the Schrödinger-Pauli equation is derived by combining the effects of random translational fluctuations and random rotational fluctuations when applying the extended least action principle. Section V develops the theory for spin entanglement based on the correlations between the random orientations of intrinsic angular momenta, which allows us to prove that Bell inequality is violated with this spin theory. To test the difference between our theory and the standard quantum mechanics, we propose two potential experiments in Section VI. Lastly, the conceptual implications and the limitations of the present theory are discussed in Section VII.

II. THE EXTENDED LEAST ACTION PRINCIPLE

The theoretical framework in this paper is built on the extended least action principle proposed in Ref. [44]. We will briefly review the principle before introducing additional assumptions in order to develop the spin theory. In [44], the least action principle in classical mechanics is extended to derive quantum formulation by factoring in the following two assumptions.

> *Assumption 1 – A quantum system experiences vacuum fluctuations constantly. The fluctuations are local and completely random.*

> *Assumption 2 – There is a lower limit to the amount of action that a physical system needs to exhibit in order to be observable. This basic discrete unit of action effort is given by $\hbar/2$ where $\hbar$ is the Planck constant.*

The first assumption is generally accepted in mainstream quantum mechanics, which is responsible for the intrinsic randomness of the dynamics of a quantum object. Even though we do not know the physical details of the vacuum fluctuation, the crucial assumption here is the locality of the vacuum fluctuation. This implies that for a composite system, the fluctuation of each subsystem is independent of each other.

The implications of the second assumption need more elaborations. Historically the Planck constant was first introduced to show that energy of radiation from a black body is discrete. One can consider the discrete energy unit as the smallest unit to be distinguished, or detected, in the black body radiation phenomenon. In general, it is understood that Planck constant is associated with the discreteness of certain observables in quantum mechanics. Here, we instead interpret the Planck constant from an information measure point of view. Assumption 2 states that there is a lower limit to the amount of action that the physical system needs to exhibit in order to be observable or distinguishable in potential observation, and such a unit of action is determined by the Planck constant.

Making use of this understanding of the Planck constant conversely provides us a new way to calculate the additional action due to vacuum fluctuations. That is, even though we do not know the physical details of vacuum fluctuations, the vacuum fluctuations manifest themselves via a discrete action unit determined by the Planck constant as an observable information unit. If we are able to define an information metric that quantifies the amount of observable information manifested by vacuum fluctuations, we can then multiply the metric with the Planck constant to obtain the action associated with vacuum fluctuations. Then, the challenge to calculate the additional action due to vacuum fluctuation is converted to define a proper new information metric I_f, which measures the additional distinguishable, hence observable, information exhibited due to vacuum fluctuations. Even though we do not know the physical details of vacuum fluctuations (except that as Assumption 1 states, these vacuum fluctuations are completely random and local), the problem becomes less challenged since there are information-theoretic tools available.

The first step is to assign a transition probability distribution due to vacuum fluctuation for an infinitesimal time step at each position along the classical trajectory. The distinguishability of vacuum fluctuation then can be defined as the information distance between the transition probability distribution and a uniform probability distribution. Uniform probability distribution is chosen here as reference to reflect the complete randomness of vacuum fluctuations. In information theory, the common information metric to measure the information distance between two probability distributions is relative entropy. Relative entropy is more fundamental to Shannon entropy since the latter is just a special case of relative entropy when the reference probability distribution is a uniform distribution. But there is a more important reason to use relative entropy. As shown in later sections, when we consider the dynamics of the system for an accumulated time period, we assume the initial position is unknown but is given by a probability distribution. This probability distribution can be defined along the position of classical trajectory without vacuum fluctuations, or with vacuum fluctuations. The information distance between the two probability distributions gives the additional distinguishability due to vacuum fluctuations. It is again measured by a relative entropy. Thus, relative entropy is a powerful tool allowing us to extract meaningful information about the dynamic effects of vacuum fluctuations. Concrete form of I_f will be defined later as a functional of relative entropy that measures the information distances of different probability distributions caused by vacuum fluctuations. Thus, the total action from classical path and vacuum fluctuation is

$$S_t = S_c + \frac{\hbar}{2} I_f, \tag{1}$$

where S_c is the classical action. Non-relativistic quantum theory can be derived through a variation approach to minimize such a functional quantity, $\delta S_t = 0$. When $\hbar \to 0$, $S_t = S_c$. Minimizing S_t is then equivalent to minimizing S_c, resulting in Newton's laws in classical mechanics. However, in quantum mechanics, $\hbar \neq 0$, the contribution from I_f must be included when minimizing the total action. We can see I_f is where the quantum behavior of a system comes from. These ideas can be condensed as

> ***Extended Least Action Principle*** *– The law of physical dynamics for a quantum system tends to exhibit as little as possible the action functional defined in (1).*

Non-relativistic quantum formulation can be derived based on the extended least action principle if we only consider the translational component of vacuum fluctuations [44]. A brief description of the derivation is given in Appendix A for self reference. It has also been shown that the quantum scalar field theory can be formulated based on the principle [45], which further demonstrates its general applicability. We will next apply the extended least action principle to derive the quantum theory of electron spin by considering not only translational, but rotational components of vacuum fluctuations. It turns out that the results are more fruitful than the results already obtained in [44, 45].

III. QUANTUM THEORY OF SPIN

A. Spin Model

In order to explain the existence of electron spin, additional assumptions on electron properties are needed. Historically, several models have been proposed to explain the existence of spin. The most popular model is to consider the electron as a rotating rigid body with uniform distributed charge. Denote the electron magnetic moment as $\vec{\mu}$ and the angular momentum as $\vec{L}$. The relation between $\vec{\mu}$ and $\vec{L}$ can be derived if we consider the electron as a localized electrical distribution with current density $\vec{j}(\vec{r})$, and the magnetic moment of electron is the magnetic dipole moment, which is calculated as

$$\vec{\mu} = \frac{1}{2}\int \vec{r} \times \vec{j}(\vec{r}) d^3\vec{r}. \tag{2}$$

Assuming the electrical density and mass density is identical, the above expression can be rewritten as

$$\vec{\mu} = -\frac{e}{2m}\int \vec{r} \times \vec{p}(\vec{r}) d^3\vec{r} = -\frac{e}{2m}\vec{L}, \tag{3}$$

where $\vec{p}(\vec{r})$ is the momentum density. However, one of the problems for this model is that when the radius of the rigid body is sufficiently small, the edge is moving faster than the speed of light [5]. There is also a simpler model that considers an electron as a point particle with intrinsic magnetic moment and intrinsic angular momentum and they still satisfy the relation (3). Such a model offers no explanation on how the intrinsic angular moment or intrinsic magnetic moment originate.

Built upon the success of applying the extended least action principle in deriving quantum theories [44, 45], we wish to explain the origin of the intrinsic magnetic moment and angular momentum by applying the same principle but adding additional refinement on the vacuum fluctuations assumption in Section II. The electron exhibits random displacement due to vacuum fluctuation and such displacement is described as a vector. The displacement vector is intrinsically local in the sense that it is independent of the reference of origin of the coordinate system. If we further assume the displacement vector contains not only translational components, but also rotational components, then the electron will exhibit intrinsic angular momentum due to such random rotation. In Appendix B, using the extended least action principle, we show that if the rotation of the displacement vector is a circular movement, the averaged magnitude of intrinsic angular momentum is $L_s = \hbar/2$, while the orientations of the intrinsic intrinsic angular momentum is completely random. That is, the intrinsic angular momentum is isotropic and its orientation is randomly fluctuating. Although our model is very different from the rigid body model, we assume that the relation (3) still holds except with a factor g_s to be determined further

$$\vec{\mu} = -g_s\frac{e}{2m}\vec{L}_s. \tag{4}$$

Since the orientation of the magnetic moment is always opposite to the orientation of intrinsic angular momentum, the orientation of the intrinsic magnetic moment is randomly fluctuating as well. Although the result in Appendix B is impressive enough, the assumption of circular motion of displacement vectors due to vacuum fluctuation is rather strong. Instead, we will just assume the existence of intrinsic angular momentum with magnitude of $\hbar/2$ and with random orientation for an electron. To summarize, in addition to Assumption 1 and 2 as described in Section II, we need

> *Assumption 3 – An electron has intrinsic angular momentum with magnitude of $\hbar/2$ and completely random orientation in free space.*

As to be shown next, with this additional assumptions and applying the extended least action principle, we can recover the predictions of quantum theory of the electron spin, including the discreteness of spin measurement results, the Pauli-Schrodinger equation, and entanglement between two correlated electron spins[1].

Suppose we choose a reference frame such that the electron is at rest, that is, the average translational momentum is zero. The electron still exhibits an intrinsic angular momentum with the random orientation according to Assumption 3. We want to derive the probability distribution of the intrinsic angular momentum orientations. If there is no external magnetic field applied, the probability distribution is simply an uniform distribution according to Assumption 3. Now, suppose an external magnetic field along the direction of the z-axis, B_z, is applied, the probability distribution of the magnetic moment orientations is no longer uniform. We show next that this probability distribution can be derived from the extended least action principle.

Due to the interaction of electron magnetic moment and the external magnetic field, the electron is experiencing the Larmor precession around the z-axis, as shown in Figure 1. Denote the probability density of the intrinsic angular momentum orientations as $p(\theta, \varphi)$ where θ is the angle between the direction of $\vec{L}_s$ and the z-axis, and φ is the angle between the projection of $\vec{L}_s$ in the X-Y plane and the x-axis. There is no reason to assume the probability density p depends on φ since the external field is along the z-axis. We can simply drop the parameter φ from $p(\theta, \varphi)$. The Larmor angular frequency is $\omega = eB_z/m$, which is independent of angle θ. The corresponding potential energy for Larmor precession is

$$\begin{aligned} U &= -\vec{\mu}\cdot\vec{B} \\ &= \frac{e}{2m}g_s B_z L_s\cos(\theta) \\ &= \frac{1}{2}g_s\omega L_s\cos(\theta). \end{aligned} \tag{5}$$

where $L_s = \hbar/2$ is the magnitude of the intrinsic angular momentum. The above expression allows us to calculate the expectation value of classical action for an infinitesi-

[1] It is desirable to develop a better model than that described in Appendix B to derive the $\hbar/2$ magnitude of intrinsic angular momentum, but we leave it as a future research topic.

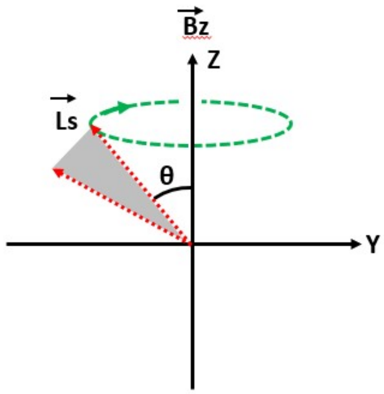

FIG. 1. Precession of the intrinsic angular momentum $\vec{L}_s$ around the magnetic field $\vec{B}_z$. The shading area indicates that the orientation of $\vec{L}_s$ is randomly fluctuating.

mal time period Δt as

$$\begin{aligned} A_c &= -\int_0^{\pi}\int_0^{\Delta t} p(\theta) U d\theta dt \\ &= -\frac{1}{2} g_s \int_0^{\pi}\int_0^{\Delta t} p(\theta)\omega L_s \cos(\theta) d\theta dt. \end{aligned} \tag{6}$$

Note that $d\varphi = \omega dt$ where φ is the angle of circular rotation due to Larmor precession. In the infinitesimal period of time Δt, the electron intrinsic angular momentum rotates $\Delta\varphi = \omega\Delta t$. We can rewrite

$$A_c = -\frac{1}{2} g_s L_s \int_0^{\pi}\int_0^{\Delta\varphi} p(\theta)\cos(\theta) d\theta d\varphi. \tag{7}$$

To compute I_f, we recall that the framework based on the extended least action principle [44] allows us to choose general relative entropy definitions such as Kullback-Leibler divergence, Rényi divergence, or Tsallis divergence [50–52]. Here, we will choose Tsallis divergence for the reason to be clear later. For the infinitesimal period of time Δt, we define

$$I_f = \frac{1}{\alpha-1}\{\int_0^{\pi}\int_0^{\Delta\varphi} \frac{p^{\alpha}(\theta)}{\sigma^{\alpha-1}} d\theta d\varphi - 1\}, \tag{8}$$

where $\alpha \in (0,1)\cup(1,\infty)$ is the order of Tsallis divergence, and σ is an uniform probability density to reflect the total ignorance of knowledge due to complete randomness of orientations. Then, the total action as defined in (1) is

$$\begin{aligned} A_t = &-\frac{1}{2} g_s L_s \int\int p(\theta)\cos(\theta) d\theta d\varphi \\ &+ \frac{\hbar}{2(\alpha-1)}\{\int \frac{p^{\alpha}(\theta)}{\sigma^{\alpha-1}} d\theta d\varphi - 1\}. \end{aligned} \tag{9}$$

Taking the variation of A_t over the functional variable $p(\theta)$, and demanding $\delta A_t = 0$, we obtain

$$-\frac{1}{2} g_s L_s \cos(\theta) + \frac{\alpha\hbar}{2(\alpha-1)}[\frac{p(\theta)}{\sigma}]^{\alpha-1} = 0. \tag{10}$$

This gives

$$\begin{aligned} p(\theta) &= \sigma[\frac{(\alpha-1) g_s L_s}{\alpha\hbar}\cos(\theta)]^{\frac{1}{\alpha-1}} \\ &= \frac{1}{Z_\alpha}[\cos(\theta)]^{\frac{1}{\alpha-1}}, \end{aligned} \tag{11}$$

where Z_α is a normalization factor that is dependent on α. Now for $p(\theta)$ to be a probability density number, it must be real and non-negative. This imposes restrictions on the value of α. Since $\cos(\theta)$ can be a negative number, $1/(\alpha-1)$ needs to be an even integer. Let $1/(\alpha-1) = 2m$ where $m \in \mathbb{N}$. This gives

$$\alpha = 1 + \frac{1}{2m}. \tag{12}$$

Thus, the probability density is rewritten as

$$p_m(\theta) = \frac{1}{Z_m}[\cos(\theta)]^{2m}, m \in \mathbb{N}. \tag{13}$$

Thus, from the extended least action principle and Assumption 3, we obtain a family of probability density functions $p_m(\theta)$. Mathematically, each of the probability density functions is a legitimate solution. We will further impose physical constraints in the next subsection and single out the solution that matches measurement results.

B. Discreteness of Spin Measurement

The probability density functions $p_m(\theta)$ in (13) are valid only relative to the context of magnetic field B_z since B_z defines the direction of the space to measure the spin. Without the external magnetic field, i.e., $B_z = 0$, then the spin orientation is completely random. This corresponds to the case $m = 0$ (or, $\alpha \to \infty$) such that $p_m(\theta)$ is an uniform distribution. Recognizing the fact that spin measurement result is context dependent is itself an important result[2].

When an inhomogeneous field B_z is applied, m must be a number larger than zero. This means m increases when the spatial gradient of B_z is non-zero. We can assume further that m monotonically increases as the electron travels along the inhomogeneous field B_z. This is intuitively explained as follows. The key is not to consider the electron as an idealized point particle. Instead it has non-zero size and is distributed with a spatial volume. We do not assume it is a rigid rotating body. Instead, it can be divided into many small mass elements δm. Each

[2] More accurately, we should label the probability density functions as $p_m[(\theta)|B_z]$ to show its contextual dependency. For simplicity of notation, we will not adopt such labeling in this paper. However, conceptual implication on such contextual dependency will be given in the Section VII.

mass element is also a charge element δe, and experiences Larmor precession. Note that in Larmor precession, the angular frequency is $\omega = \delta e B_z/\delta m$. In a homogeneous magnetic field, ω is a constant for each charge element and independent of the angle θ. In this case, the probability density function $p_m(\theta)$ should remain unchanged for the whole electron. But in an inhomogeneous field B_z, because the electron has spatial size, each charge element will experience a different magnitude of magnetic field. Consequently, different parts of the electron will have different angular frequencies. The upper parts of the electron along the direction of B_z tend to precess faster than the lower parts of the electron. Effectively, the orientation of overall intrinsic angular momentum is "pulled" closer to align with the z-axis. This corresponds to choosing a larger value of m.

Although the above physical interpretation is intuitive, it is nevertheless an assumption and need to be called out explicitly as

> *Assumption 4 – Parameter m in (13) increases monotonically as the electron travels along an inhomogeneous magnetic field.*

Now we ask, what happens if the gradient of B_z is sufficiently large, or the electron has been traveling in the inhomogeneous magnetic field long enough, such that m can be approximated as infinity? Since $\theta \in [0, \pi]$ and $\cos(\theta) \in [-1, 1]$, we have

$$\lim_{m\to\infty}[\cos(\theta)]^{2m} = \begin{cases} 0, & \text{for } \theta \in (0, \pi) \\ 1, & \text{for } \theta = 0, \pi \end{cases} \tag{14}$$

Thus, when ∇B_z is sufficiently large, the probability density function is

$$\lim_{m\to\infty} p_m(\theta) \propto \begin{cases} 0, & \text{for } \theta \in (0, \pi) \\ 1, & \text{for } \theta = 0, \pi \end{cases} \tag{15}$$

This shows that when ∇B_z is sufficiently large, or the electron has been traveling in the inhomogeneous magnetic field for a distance long enough, measurement on the electron spin can only obtain two discrete outcomes: spin up and spin down, along the z-axis, as shown in Figure 2. Impressively, the quantum phenomenon of quantization in spin measurement outcome is a natural mathematical result here instead of being a postulate as usually taught in orthodox quantum mechanics textbooks.

However, (15) itself cannot be a valid probability density function. Its integral is zero since it is only non-zero when $\theta = 0, \pi$. The Dirac delta function is the proper mathematical tool to describe such a scenario. We can rewrite

$$\bar{p}(\theta) := \lim_{m\to\infty} p_m(\theta) = \frac{1}{2}\{\delta(\theta) + \delta(\theta - \pi)\}. \tag{16}$$

The factor of 1/2 is due to two facts, 1.) normalization requirement, and 2.) the measurement outcome of spin-up or spin-down is completely random. The second point here needs more elaboration. Before the external inhomogeneous field B_z is applied, the intrinsic angular momentum of the electron is oriented completely randomly. At the moment when B_z is applied, with half of chance that the initial direction of the angular momentum forms an angle with the z-axis θ such that $\theta \in [0, \pi/2]$. As the electron travels along B_z, the angular momentum vector evolves closer and closer to the z-axis and eventually becomes spin-up. With the other half of chance the initial direction of the angular momentum forms an angle with the z-axis θ such that $\theta \in [\pi/2, \pi]$. As the electron travels along B_z, the angular momentum evolves to be spin-down. For simpler notation, define

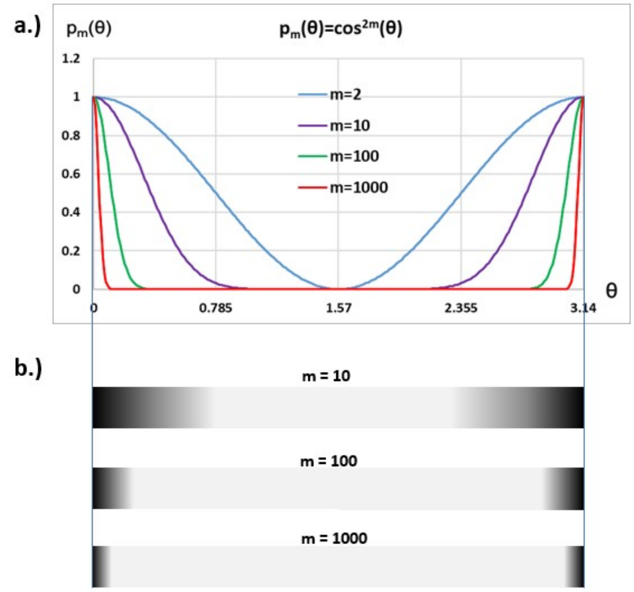

FIG. 2. a.) As parameter m increases, the probability distribution converges to only two directions at $\theta = 0$ and $\theta = \pi$. b.) The envisioned Stern-Garlach measurements with an inhomogeneous magnetic field. As the electrons travel along B_z, the electrons reach the detectors that eventually show two converging discrete lines.

$$\Theta^+ := \forall\theta \in [0, \pi/2], \text{ and } \Theta^- := \forall\theta \in [\pi/2, \pi]. \tag{17}$$

Then, $\theta \in \Theta^+$ means the orientation is pointing to the upper half of the orientation sphere. Similarly, $\theta \in \Theta^-$ means the orientation is pointing to the lower half of the orientation sphere. The factor 1/2 is thus due to initial complete randomness of orientation of intrinsic angular momentum such that the probability of $\theta \in \Theta^+$ is the same as the probability of $\theta \in \Theta^-$.

The above analysis implies that if the initial probability density before measurement is $\sigma(\theta)$, the probability of measurement outcome as spin-up can be calculated as

$$p(\uparrow) := \varrho(\Theta^+) = \int_0^{\pi/2} \sigma(\theta) d\theta. \tag{18}$$

Similarly, the probability of measurement outcome as spin-down is

$$p(\downarrow) := \varrho(\Theta^-) = \int_{\pi/2}^{\pi} \sigma(\theta) d\theta, \tag{19}$$

and $\varrho(\Theta^+) + \varrho(\Theta^-) = 1$. The generalized form of (16) is

$$\bar{p}(\theta) = \varrho(\Theta^+)\delta(\theta) + \varrho(\Theta^-)\delta(\theta - \pi). \tag{20}$$

We will derive (20) more rigorously in later sections. For an initially uniformly distributed orientation, one gets $p(\uparrow) = p(\downarrow) = 1/2$.

The probability of 1/2 for either spin-up and spin-down is due to the complete randomness of the initial orientations of the intrinsic angular momentum. It is important to note that the measurement results are relative to the direction defined by the magnetic field B_z. Given the same initial random orientations, if one measures the spin along a different direction, say z'-axis, one will still get either spin-up and spin-down with probability of 1/2. However, if the initial orientation of the intrinsic angular momentum is not completely random, the probability of obtaining spin-up and spin-down will be different from 1/2. This is the subject for study in the next section.

C. Subsequent Measurement with a Rotated Stern-Gerlach Apparatus

Suppose after the electron passes through a Stern-Gerlach apparatus and it is observed the measurement result is spin-up. Now let the electron pass a subsequent Stern-Gerlach apparatus with the exact same direction of inhomogeneous magnetic field, one will obtain the result of spin-up for the electron. This means the electron stays in the spin-up state after passing through the first Stern-Gerlach apparatus[3]. However, in our model, after the electron passes through the first Stern-Gerlach apparatus and the external magnetic field is removed, due to the nature of random fluctuations, the orientation of electron intrinsic angular momentum can start to deviate from the z-axis with an angle θ between the angular momentum and the z-axis. In other words, due to fluctuations, the orientation of the angular momentum is relaxing away from pointing along the z-axis. This process is important in order to explain the behavior if the second Stern-Gerlach apparatus is set up such that the direction of the magnetic field is tilted with an angle β from the original z-axis. Denote this new direction as z'-axis, and the magnetic field as $B_{z'}$.

Let the probability density conditioned on the initial measurement outcome of spin-up as $p(\theta'|\theta)$, where θ' is the angle between the orientation of electron intrinsic angular momentum and the z'-axis. From Figure 3, it can be seen that $\theta = \theta' + \beta$. We will apply the extended least action principle again to find out $p(\theta'|\theta)$. With the magnetic field $B_{z'}$, the classical action can be calculated similarly to that in (6) - (7) as

$$A'_c = -\frac{1}{2} g_s L_s \int_0^{\pi} \int_0^{\Delta\varphi} p(\theta'|\theta) \cos(\theta') d\theta' d\varphi. \tag{21}$$

To compute I_f, we choose Tsallis divergence again. For the infinitesimal period of time Δt, we define

$$I_f = \frac{1}{\alpha - 1}\{\int_0^{\pi} \int_0^{\Delta\varphi} \frac{p^{\alpha}(\theta'|\theta)}{\sigma^{\alpha-1}(\theta)} d\theta' d\varphi - 1\}, \tag{22}$$

where $\sigma(\theta)$ is no longer an uniform probability density since the initial condition is that the electron has passed through the first Stern-Gerlach apparatus. Applying the same variation principle as that to derive (11) and (13), we obtain

$$p_m(\theta'|\theta) = \frac{1}{Z_m} \sigma(\theta) [\cos(\theta')]^{2m}, m \in \mathbb{N}. \tag{23}$$

When $\nabla B_{z'}$ is sufficiently large, or the electron travels sufficiently long distance in the magnetic field, we get

$$\bar{p}(\theta'|\theta) := \lim_{m\to\infty} p_m(\theta'|\theta) = \sigma(\theta)\{\delta(\theta') + \delta(\theta' - \pi)\}, \tag{24}$$

where the normalization factor is omitted. Noting that $\theta = \theta' + \beta$, we can compute the probability of obtaining measurement of spin-up along the z'-axis as

$$p(\uparrow|\beta) \propto \int_0^{\pi/2} \bar{p}(\theta'|\theta) d\theta' = \sigma(\beta). \tag{25}$$

Similarly, the probability of obtaining spin-down is

$$p(\downarrow|\beta) \propto \int_{\pi/2}^{\pi} \bar{p}(\theta'|\theta) d\theta' = \sigma(\beta + \pi). \tag{26}$$

To impose the normalization requirement, we must have $p(\uparrow|\beta) + p(\downarrow|\beta) = 1$. Denote $Z := \sigma(\beta) + \sigma(\beta + \pi)$, the normalization can be achieved by setting

$$p(\uparrow|\beta) = \sigma(\beta)/Z := \varrho(\beta) \tag{27}$$

$$p(\downarrow|\beta) = \sigma(\beta + \pi)/Z := \varrho(\beta + \pi) \tag{28}$$

Then, (24) can be rewritten as

$$\bar{p}(\theta'|\beta) = \varrho(\beta)\delta(\theta') + \varrho(\beta + \pi)\delta(\theta' - \pi). \tag{29}$$

Note that (29) is the same as (20) if we set $\varrho(\beta) = \varrho(\Theta^+)$ and $\varrho(\beta + \pi) = \varrho(\Theta^-)$. This effectively proves (20).

The average of angular momentum along the z'-axis will be $(\varrho(\beta) - \varrho(\beta + \pi))L_s$. On the other hand, from Figure 3, one can deduce that this average angular momentum must be $L_s \cos(\beta)$, thus[4]

$$\varrho(\beta) - \varrho(\beta + \pi) = \cos(\beta). \tag{30}$$

[3] In standard quantum mechanics, this is explained as that the electron stays in the eigen-state of spin-up.

[4] If the initial condition is spin-down after the electron passing the first Stern-Gerlach apparatus, the average angular momentum becomes $-L_s \cos(\beta)$, resulting $\varrho(\beta) - \varrho(\beta + \pi) = -\cos(\beta)$.

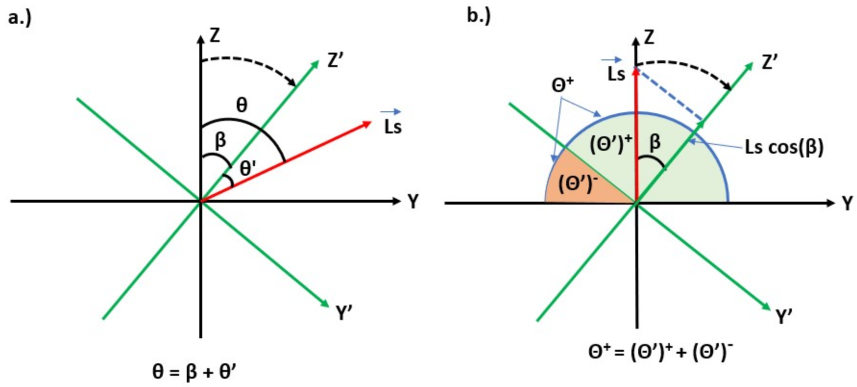

FIG. 3. a.) When the direction of inhomogeneous magnetic field in the second Stern-Gerlach apparatus is rotated with an angle β with the respect to the z-axis, the orientation of intrinsic angular momentum is given by $\theta' = \theta + \beta$. b.) The range of upward angle with respect to the z-axis, Θ^+, is split into two parts, $(\Theta')^+$, and $(\Theta')^-$, with respect to the z'-axis. Thus, the second Stern-Gerlach apparatus gives both possible measurement outcomes of spin-up and spin-down.

Together with the normalization condition $\varrho(\beta) + \varrho(\beta + \pi) = 1$, we obtain

$$\begin{cases} \varrho(\beta) = (1+\cos(\beta))/2 = \cos^2(\beta/2) \\ \varrho(\beta+\pi) = (1-\cos(\beta))/2 = \sin^2(\beta/2) \end{cases}, \tag{31}$$

Thus, we reproduce the same result as that from the standard quantum mechanics,

$$\begin{cases} p(\uparrow|\beta) = \cos^2(\beta/2) \\ p(\downarrow|\beta) = \sin^2(\beta/2) \end{cases}. \tag{32}$$

Eq.(29) now becomes

$$\bar{p}(\theta'|\beta) = \cos^2(\beta/2)\delta(\theta') + \sin^2(\beta/2)\delta(\theta' - \pi). \tag{33}$$

Suppose $\beta < \pi/2$, we wonder why there is possibility that the measurement result of the second Stern-Gerlach apparatus is spin-down. That is, why $p(\downarrow|\beta)$ is non-zero? The question arises because if the initial condition is spin-up after the electron passes through the first Stern-Gerlach apparatus, and if the orientation of the intrinsic angular momentum is kept unchanged, the initial angle $\theta' = \beta \in (0, \pi/2)$. Then, according to the discussions leading to (18), we should only obtain spin-up as the only results when $B_{z'}$ is applied. However, as pointed out at the beginning of this section, after the electron passing through the first Stern-Gerlach apparatus, the direction of the intrinsic angular momentum starts to deviate from the z-axis due to vacuum fluctuations. Suppose after a sufficient time of relaxation τ_r, the orientation of electron intrinsic angular momentum can be at any angle θ from the z-axis with $\theta \in [0, \pi/2]$. When the electron enters the second Stern-Gerlach with the magnetic field $B_{z'}$ is tilted with an angle $\beta \in (0, \pi/2)$, the initial orientation is now measured with respect to the z'-axis as angle θ'. Since $\theta' = \theta - \beta$, the orientation of the intrinsic angular momentum can point to either upward with $\theta' \in (\Theta')^+$, or downward with $\theta' \in (\Theta')^-$, as shown in Figure 3b. Thus, there is possibility of spin-down when $\nabla B_{z'}$ is sufficiently large. We see the relaxation of intrinsic angular momentum orientation is necessary in order to explain the result in (32), which has been confirmed experimentally.

There is still one puzzle to clarify. The second Stern-Gerlach apparatus can be rotated such that $\theta - \theta'$ is β or $-\beta$, the probability of measurement outcome for spin-up or spin-down (33) are the same for both cases. This suggests that (33) does not completely specify the behavior of electron spin in a single measurement. To remedy that, we need to specify the probability of measurement outcome along another axis, say, the y-axis. Suppose the $y - z$ plane is rotated along the x-axis such that z'-axis is deviated from z-axis by an angle β, then the angle between y'-axis and z-axis is $\pi/2 + \beta$. The probability density of measurement outcome along the y'-axis can be obtained by replacing parameter β in (33) with $\pi/2 + \beta$,

$$\begin{aligned} \bar{p}(\theta'_y|\frac{\pi}{2}+\beta) &= \frac{1}{2}[\cos(\frac{\beta}{2}) - \sin(\frac{\beta}{2})]^2\delta(\theta'_y) \\ &+ \frac{1}{2}[\cos(\frac{\beta}{2}) + \sin(\frac{\beta}{2})]^2\delta(\theta'_y - \pi). \end{aligned} \tag{34}$$

On the other hand, if the z'-axis is deviated from z-axis by an angle $-\beta$, then the angle between y'-axis and z-axis is $\pi/2 - \beta$. The probability density of measurement outcome along the y'-axis is

$$\begin{aligned} \bar{p}(\theta'_y|\frac{\pi}{2}-\beta) &= \frac{1}{2}[\cos(\frac{\beta}{2}) + \sin(\frac{\beta}{2})]^2\delta(\theta'_y) \\ &+ \frac{1}{2}[\cos(\frac{\beta}{2}) - \sin(\frac{\beta}{2})]^2\delta(\theta'_y - \pi). \end{aligned} \tag{35}$$

It is clear that (34) and (35) are different. Thus, (33) and (34) together give a complete description of spin behavior when then second Stern-Gerlach is rotated by an angle β, while (33) and (35) together give a complete description of spin behavior when then second Stern-Gerlach is rotated by an angle $-\beta$.

We can generalize the result by assuming the magnetic field of the first Stern-Gerlach apparatus is along a direction tilted an angle β_1 from the z-axis, and the second Stern-Gerlach apparatus is along a direction tilted an angle β_2 from the z-axis. Supposed measurement outcome of the first Stern-Gerlach apparatus is spin-up, what is the probability of measurement outcome with spin-up from the second Stern-Gerlach apparatus? The calculation steps from (21) to (32) can be repeated but with $\theta = \theta' + \beta_2 - \beta_1$. Thus, we can just replace β with $\beta_2 - \beta_1$ in steps from (21) to (32), and the final result will be

$$\begin{cases} p(\uparrow|\beta_2, \beta_1) = \cos^2((\beta_2-\beta_1)/2) \\ p(\downarrow|\beta_2, \beta_1) = \sin^2((\beta_2-\beta_1)/2) \end{cases}. \tag{36}$$

In Appendix C, we verify that (36) is exactly the same result predicted by standard quantum mechanics.

IV. DERIVATION OF THE SCHRÖDINGER-PAULI EQUATION

The framework to derive the law of dynamics for an electron in a external magnetic field B_z, for a cumulative period from t_a to t_b, is similar to that described in Appendix A, except that we need to add a new term in the Lagrangian due to the the interaction of spin and the external magnetic field. Based on the results shown in (16) and (29), the probability density can be written as

$$p(\theta) = \sum_i \sigma_i \delta(\theta - \theta_i). \tag{37}$$

For an electron, we only need to consider the two-level case where $i \in \{0, 1\}$ and $\theta_0 = 0, \theta_1 = \pi$. The expectation value of the potential energy due to the the interaction of spin and B_z is

$$\begin{aligned} \bar{U} &= -\int p(\theta)\vec{\mu}\cdot\vec{B}d\theta \\ &= \frac{e\hbar}{2m}\int p(\theta)B_z\cos(\theta)d\theta, \end{aligned} \tag{38}$$

where in the second step, we substitute the magnetic moment μ with (4) and choose the g-factor $g_s = 2$.

More generically, we need to specify the probability density with the space-time coordinates. Denote

$$\rho(\mathbf{x}, t, \theta) = \sum_i \sigma_i(\mathbf{x}, t, \theta_i)\delta(\theta - \theta_i). \tag{39}$$

Then the expectation value of the potential energy is

$$\begin{aligned} \bar{U} &= \frac{e\hbar}{2m}\int \rho(\mathbf{x}, t, \theta)B_z\cos(\theta)d\theta d^3\mathbf{x} \\ &= \frac{e\hbar}{2m}\sum_i\int \sigma_i(\mathbf{x}, t, \theta_i)\delta(\theta-\theta_i)B_z\cos(\theta)d\theta d^3\mathbf{x}. \end{aligned} \tag{40}$$

In the case of electrons, there are only two values $\theta_0 = 0, \theta_1 = \pi$. Therefore,

$$\begin{aligned} \bar{U} &= \frac{e\hbar}{2m}\{\int \sigma_0(\mathbf{x}, t, 0)\delta(\theta)\mu B_z\cos(\theta)d\theta d^3\mathbf{x} \\ &+ \int \sigma_1(\mathbf{x}, t, \pi)\delta(\theta-\pi)\mu B_z\cos(\theta)d\theta d^3\mathbf{x}\} \\ &= \frac{e\hbar}{2m}\{\int \sigma_0(\mathbf{x}, t, 0)\mu B_z d^3\mathbf{x} - \int \sigma_1(\mathbf{x}, t, \pi)\mu B_z d^3\mathbf{x}\}. \end{aligned} \tag{41}$$

Here we run into a problem. In the Stern-Gerlach experiment, it is confirmed that electrons with spin-up follow a different trajectory from that of electrons with spin-down. An electron with either spin-up or spin-down will follow separated laws of dynamics. We cannot simply add the two terms in (41) together. Instead, a more proper notation should treat the probability density as a two-component vector. Thus,

$$\bar{U} = \frac{e\hbar}{2m}\int B_z \begin{pmatrix} \sigma_0(d^3\mathbf{x}, t, 0) \\ -\sigma_1(d^3\mathbf{x}, t, \pi) \end{pmatrix} dx := \begin{pmatrix} \bar{U}_+ \\ \bar{U}_- \end{pmatrix}. \tag{42}$$

For further simplification of notation, denote $\rho_+(x, t) = \sigma_0(x, t, 0)$ and $\rho_-(x, t) = \sigma_0(x, t, 0)$. They should satisfy the normalization condition

$$\int [\rho_+(d^3\mathbf{x}, t) + \rho_-(d^3\mathbf{x}, t)]dx = 1. \tag{43}$$

First we will derive the law of dynamics for the electron with spin up. Without considering the contribution from the spin magnetic interaction, the classical action of an electron in an external electromagnetic field described by the magnetic vector potential[5] $\vec{A}$ and electric scalar potential ϕ is given by (see Appendix C of [44])

$$A_c = \int \rho_+\{\frac{\partial S_+}{\partial t} + \frac{1}{2m}(\nabla S_+ + e\vec{A})^2 - e\phi\}d^3\mathbf{x}dt \tag{44}$$

Now we need to add the additional term of potential energy $\bar{U}_+$ due to spin magnetic interaction into (44)

$$A_c^+ = \int \rho_+\{\frac{\partial S_+}{\partial t} + \frac{1}{2m}(\nabla S_+ + e\vec{A})^2 + \frac{e\hbar}{2m}B_z - e\phi\}d^3\mathbf{x}dt \tag{45}$$

The definition of information metrics for the vacuum translational fluctuation is the same as (A7), and similar to the calculation shown in [44], when $\Delta t \to 0$ it becomes

$$I_f^+ = \int d^3\mathbf{x}dt\frac{\hbar}{4m}\frac{1}{\rho_+}\nabla\rho_+\cdot\nabla\rho_+. \tag{46}$$

---

[5] Note that $\vec{B} = \nabla\times\vec{A}$. Since we only consider the case that $\vec{B}$ is along the z-axis, $\nabla\times\vec{A}$ only has the z component.

Inserting (45) - (46) into (1), we have

$$A_t^+ = \int \rho_+\{\frac{\partial S_+}{\partial t} + \frac{1}{2m}(\nabla S_+ + e\vec{A})^2 + \frac{e\hbar}{2m}B_z - e\phi\}d^3\mathbf{x}dt + \frac{\hbar^2}{8m}\int \frac{1}{\rho_+}\nabla\rho_+\cdot\nabla\rho_+ d^3\mathbf{x}dt. \tag{47}$$

Performing the variation procedure on A_t with respect to S_+ gives

$$\frac{\partial\rho_+}{\partial t} + \frac{1}{m}\nabla\cdot(\rho_+(\nabla S_+ + e\vec{A})) = 0, \tag{48}$$

which is the continuity equation for ρ_+. Performing the variation procedure on A_t with respect to ρ_+ leads to the spin-up version of extended Hamilton-Jacobi equation

$$\frac{\partial S_+}{\partial t} + \frac{1}{2m}(\nabla S_+ + e\vec{A})^2 + \frac{e\hbar}{2m}B_z - e\phi - \frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{\rho_+}}{\sqrt{\rho_+}} = 0. \tag{49}$$

Defined a complex function $\Psi_+ = \sqrt{\rho_+}e^{iS_+/\hbar}$, the continuity equation (48) and the extended Hamilton-Jacobi equation (49) can be combined into a single differential equation,

$$i\hbar\frac{\partial\Psi_+}{\partial t} = [\frac{1}{2m}(i\hbar\nabla + e\vec{A})^2 + \frac{e\hbar}{2m}B_z - e\phi]\Psi_+, \tag{50}$$

which is the Schrödinger equation for a spin-up electron.

The derivation of dynamics equations for a spin-down electron is exactly the same, except there is a sign difference for the potential energy term $\bar{U}_-$. The resulting Schrödinger equation is

$$i\hbar\frac{\partial\Psi_-}{\partial t} = [\frac{1}{2m}(i\hbar\nabla + e\vec{A})^2 - \frac{e\hbar}{2m}B_z - e\phi]\Psi_-, \tag{51}$$

where $\Psi_- = \sqrt{\rho_-}e^{iS_-/\hbar}$.

To combine equations (50) and (51), we introduce a two-component vector wave function

$$\Psi = \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \tag{52}$$

and a two dimensional matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{53}$$

then (50) and (51) can be combined into a compact form

$$i\hbar\frac{\partial\Psi}{\partial t} = [\frac{1}{2m}(i\hbar\nabla + e\vec{A})^2 + \frac{e\hbar}{2m}\sigma_z B_z - e\phi]\Psi, \tag{54}$$

which is the Schrödinger-Pauli equation for an electron in an external magnetic field along the z-axis B_z. The normalization condition (43) is rewritten as

$$\int \Psi^+\Psi dx = \int(\Psi_+^*\Psi_+ + \Psi_-^*\Psi_-)d^3\mathbf{x} = 1. \tag{55}$$

Generalizing the Schrödinger-Pauli equation for an electron in an external magnetic field $\vec{B}$ along arbitrary orientation is possible. We will leave it for future research and instead focus next on the more interesting quantum phenomenon of entanglement.

V. ENTANGLEMENT

A. Spin Correlation

In order to investigate the entanglement phenomenon between two electron spins, we need to take a closer look at the fluctuations of spin orientations from the time perspective. As shown earlier, when measuring the spin of an electron using an inhomogeneous external magnetic field along any axis, one will obtain two discrete results, spin-up or spin-down along the axis. Assign this axis as z-axis. Whether the result is spin-up or spin-down is random, depending on the initial orientation relative to the z-axis, described by the angle θ. To calculate the probability of the initial orientation for $\theta\in\Theta^+$ or $\theta\in\Theta^-$, we need to analyze how the orientation of intrinsic angular momentum fluctuates before an external magnetic field B_z is applied. Recall that we denote $\theta\in\Theta^+$ as the orientation of angular momentum trending upward with angle $\theta\in[0,\pi/2]$, and $\theta\in\Theta^-$ as the orientation trending downward. It is intuitive to assume that it takes a non-zero amount of time for the orientation to change from $\theta\in\Theta^+$ to $\theta\in\Theta^-$. That is, it takes a non-zero amount of time for the orientation of intrinsic angular momentum to change from trending upward to trending downward, and vice versa. Suppose at t_0, $\theta\in\Theta^+$. At time t_1^+, θ changes to trending downward, $\theta\in\Theta^-$. So the angular momentum stays trending upward for a period of time $\Delta t_1^+ = t_1^+ - t_0$. Then the orientation stays trending downward for a duration Δt_1^- till at time t_1^-, it moves back to point upward. After the orientation stays trending upward for a period of time Δt_2^+, it changes to point downward for a period of time Δt_2^-. The process of switching the orientation directions continues till $t=\Delta T$, as shown in Figure 4. Statistically, $\{\Delta t_i^+, i\in\mathbb{N}\}$ forms a random variable, and similarly for $\{\Delta t_-^+, i\in\mathbb{N}\}$. Define

$$\Delta T^+ = \sum_i \Delta t_i^+, \text{ and } \Delta T^- = \sum_i \Delta t_i^-, \tag{56}$$

we have $\Delta T^+ + \Delta T^- = \Delta T$. Suppose a spin measurement experiment is conducted at t_m. If t_m occurs at the moment when $\theta\in\Theta^+$, that is, t_m falls into one of the time periods Δt_i^+ as depicted in Figure 4, the measurement outcome will be spin-up. On the other hand, if t_m falls into one of the time periods Δt_i^-, the measurement outcome will be spin-down. Since Δt_i^+ and Δt_i^- are random variables, the measurement outcome are random as well, and the probabilities of orientation trending upward and downward are

$$\varrho(\Theta^+) = \lim_{\Delta T\to\infty}\frac{\Delta T^+}{\Delta T}, \text{ and } \varrho(\Theta^-) = \lim_{\Delta T\to\infty}\frac{\Delta T^-}{\Delta T}, \tag{57}$$

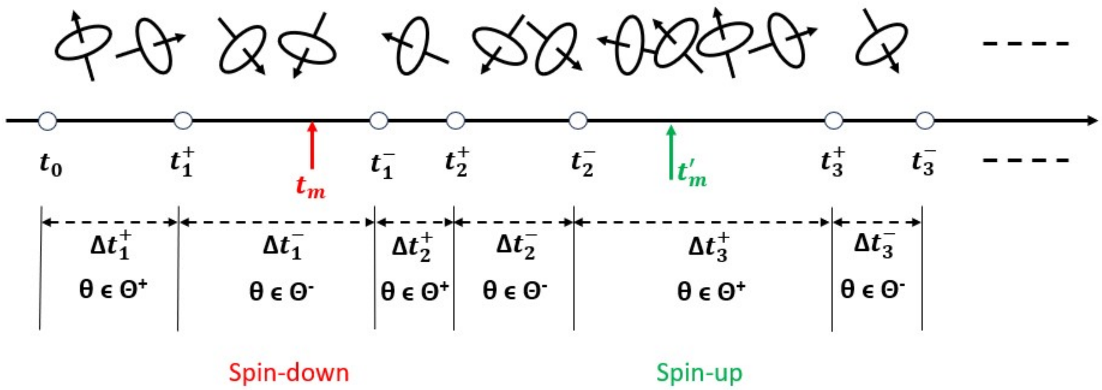

FIG. 4. Randomness of orientation of intrinsic angular momentum from time perspective. If the measurement occurs at a time t_m that when the initial orientation is $\theta \in \Theta^-$, the measurement outcome is spin-down. On the other hand, if the measurement occurs at a time t'_m that when the initial orientation $\theta \in \Theta^+$, the measurement outcome is spin-up

respectively[6]. For an intrinsic angular momentum with completely random orientations, we must have $\Delta T^+ = \Delta T^-$, and therefore, $\varrho(\Theta^+) = \varrho(\Theta^-) = 1/2$.

The random variable $\{\Delta t_i^+ \geq 0, i \in \mathbb{N}\}$ can follow a probability distribution. Here we will not investigate the actual probability distribution. What is relevant to our investigation is the mean value, denoted as

$$\tau^+ = \langle \Delta t_i^+ \rangle, \text{ and similarly, } \tau^- = \langle \Delta t_i^- \rangle. \tag{58}$$

Again, for an intrinsic angular momentum with completely random orientation, it is intuitive to assume $\tau^+ = \tau^-$.

In summary, the spin model for an electron presented here has the following characteristics:

- Applying an external inhomogeneous magnetic field results in two discrete measurement outcomes, spin-up or spin-down;
- The statistical probability of obtaining spin-up or spin-down is determined by the initial probability $\varrho(\Theta^+)$ and $\varrho(\Theta^-)$;
- For a particular measurement instant occurred at t_m, whether the outcome is spin-up or spin-down depends on $\theta \in \Theta^+$ or $\theta \in \Theta^-$ at t_m, respectively.

Now we consider the case of two electrons A and B. Suppose in the preparation stage of experiment setup they together go through interactions with a common source of external field before t_0. As a result, they share some kinds of correlation even though there is no external field applied to them after t_0. Consider a particular type of correlation between these two electrons that their intrinsic angular momenta always point to the opposite orientations. This implies that $\theta_A + \theta_B = \pi$ at any time and the reference axis can be along any direction, till a measurement event occurs. The correlation $\theta_A + \theta_B = \pi$ is maintained even though the orientations of the intrinsic angular momenta are randomly fluctuating. Consequently, we have $\varrho(\Theta_A^+) = \varrho(\Theta_B^-)$, $\varrho(\Theta_A^-) = \varrho(\Theta_B^+)$, and $\tau_A^+ = \tau_B^-$, $\tau_A^- = \tau_B^+$.

Next we want to see what happens when a measurement event occurs. There are two cases here, measurements on A and B occur at the same time, or at different times. First, suppose the two electrons are measured at the same time, which is the usual experiment setup for a typical Bell test. For electron A, the probability of measurement outcome can be described as

$$\tilde{p}(\theta_A) = \varrho(\Theta_A^+)\delta(\theta_A) + \varrho(\Theta_A^-)\delta(\theta_A - \pi). \tag{59}$$

For electron B, because of the correlation $\theta_B + \theta_A = \pi$, the measurement results will be exactly anti-correlated to the results of A. For example, if the measurement result of A is spin-up, it implies the initial condition $\theta_A \in \Theta_A^+$, meaning the orientation of angular momentum is trending upward with any angle $\theta_A \in [0, \pi/2)$ relative to the z-axis. Since $\theta_B + \theta_A = \pi$, θ_B must be also trending downward relative to the same direction. Thus, the measurement outcome must be spin-down. Similarly, if the measurement result of A is spin-down, then the measurement result of B must be spin-up. But note that the measurement outcome of A is random, depending on whether the initial condition θ_A is trending upward or downward relative to the z-axis at t_m. The correlation of measurement results can be described by a joint prob-

---

[6] Eq. (57) can be considered as the frequency interpretation of $\varrho(\cdot)$, while Eqs. (18) and (19) are the classical interpretation. Both interpretations are considered equivalent.

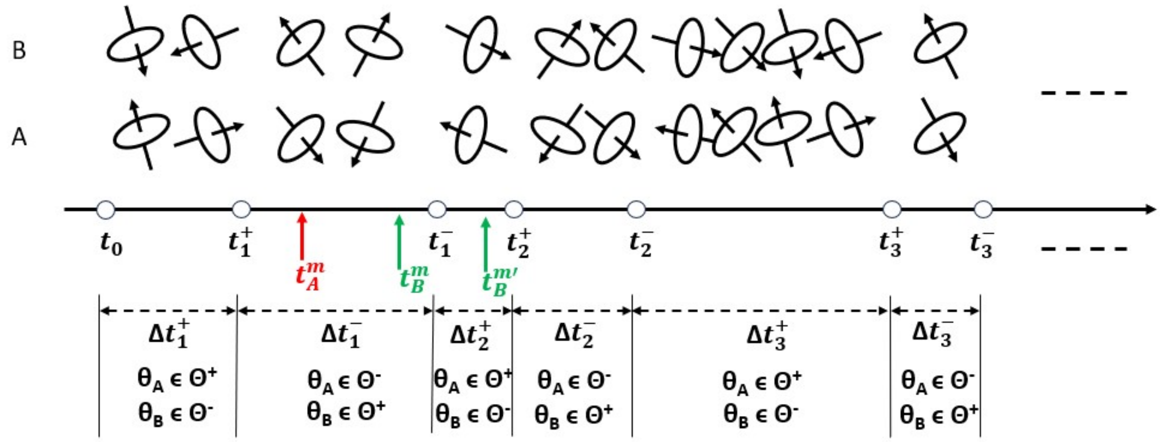

FIG. 5. Spin entanglement of two electrons from time perspective. Here, the orientation of intrinsic angular momentum for each electron is completely random, but orientations for both A and B are always opposite to each other. If Alice performs the measurement at a time t_m^A that when the initial orientation is $\theta_A \in \Theta^-$, the measurement outcome is spin-down. If Bob performs his measurement at $t_m^B < t_1^-$, he will obtain spin-up, as predicted by the singlet spin state. On the other hand, If Bob performs his measurement at $t_m^{B'} > t_1^-$ the measurement outcome can be spin-down or spin-up.

ability density

$$\tilde{p}(\theta_A, \theta_B) = \varrho(\Theta_A^+)\delta(\theta_A)\delta(\theta_B-\pi)+\varrho(\Theta_A^-)\delta(\theta_A-\pi)\delta(\theta_B). \quad (60)$$

One can drop the subscript A in the expression $\varrho(\Theta_A^+)$ since $\varrho(\Theta_A^+) = \varrho(\Theta_B^-)$.

Before considering the more complicated case that measurements of A and B occur at different times, we need to examine how the correlation $\theta_B + \theta_A = \pi$ is maintained when the two electrons are separated. Recall that in our spin model, the electron intrinsic angular momentum is not due to rotation as a rigid body, but due to the rotational components of vacuum fluctuations of the electrons, as shown in Appendix B. When the two electrons move away from each other, their motions can be translational such that they have no impact on the orientations of intrinsic angular momenta. Or their motions can have the same rotational effects as well such that the impacts on the orientations of the intrinsic angular momenta are the same. Either case, the correlation $\theta_B + \theta_A = \pi$ is preserved when the two electrons are physically moved apart. In the modern Bell test experiment [53], the two electrons are already separated remotely before being prepared to be entangled. The electron entanglement is achieved by entanglement swapping with photons [53]. Thus, there is no need to worry about the impacts on the correlation $\theta_B + \theta_A = \pi$ due to the movement of electrons. Nevertheless, even if $\theta_B + \theta_A = \pi$ is preserved before measurement, if the measurements on A and B are performed at different times, there is a possibility that the entanglement effect can be diminished. Let's consider this more subtle situation next.

Suppose the electron pair, with the correlation $\theta_B + \theta_A = \pi$ established, are separated far away. Electron A is with observer Alice, while electron B is with observer Bob. The correlation $\theta_B + \theta_A = \pi$ is maintained even though the electron pair are space-like separated. At time t_A^m, Alice performs measurement on A. The probability density is given by (59). Alice will get the result randomly. Suppose that she obtains spin-down. At this point, from Bob's point of view, if he measures electron B along any direction, he will obtain spin-up or spin-down randomly. Now Alice sends Bob her measurement results. The information on the measurement direction is critical. With the information from Alice, Bob infers that θ_A is trending downward relative the measurement direction, and therefore θ_B is trending upward relative the measurement direction due to the correlation $\theta_B + \theta_A = \pi$. Consequently, he predicts he will obtain spin-up if he perform the measurement on B along the same orientation as Alice. However, due to fluctuation, θ_B can move to trending downward after some time. Bob's measurement must be performed at a time t_B^m before θ_B becomes tending downward. This constraint $t_B^m < t_1^-$ can be approximately expressed as

$$t_B^m - t_A^m < \tau_B^+, \quad (61)$$

where τ_B^+ is the average time θ_B stays in Θ_B^+ before it switches to Θ_B^-. The scenario is depicted in Figure 5.

Eq. (61) implies that Bob's prediction on his measurement outcome of spin-up for B is valid only for a period of time. Even though the entanglement between the electron pair is preserved after they are remotely separated, once Alice performs her measurement on A, she must send the result to Bob and Bob must confirm his prediction within the time constraint (61). This is different from the prediction of standard quantum mechanics where there is no such time constraint. If τ_B^+ is sufficiently large, our model will practically give the same prediction as standard quantum mechanics.

One may argue that if $t_B^m > t_A^m + \tau_B^+$, Bob will obtain measurement results of spin-down or spin-up, what is the difference between an entangled pair and a non-entangled electron pair? The answer is that for non-entangled electron pairs, there is no correlation between the measurement results of Alice and Bob. Bob cannot make any prediction on his measurement outcome of B even after he receives the information from Alice.

The analysis here shows that Alice measurements on A will obtain outcome of either spin-up or spin-down randomly, but once the measurement outcome is known, she can predict Bob's measurement outcome on the remote electron B along the same measurement direction, even though her measurement outcome does not cause physical impact on B. There is no mysterious "spooky action in the remote". The root of the entanglement phenomenon is that the random variables θ_A and θ_B are correlated. Furthermore, these random variables are intrinsically local to the electron, it is possible to maintain the correlation even though the two electrons are remotely separated.

B. Bell States

If the correlation for the entangled electron pair is described by $\theta_A + \theta_B = \pi$ instead, the measurement outcome on A will be exactly opposite to the measurement outcome on B. The joint probability density is given by

$$\tilde{p}(\theta_A, \theta_B) = \varrho(\Theta_A^+)\delta(\theta_A)\delta(\theta_B - \pi) + \varrho(\Theta_A^-)\delta(\theta_A - \pi)\delta(\theta_B). \tag{62}$$

In the case that the initial orientations of angular momentum of electron A are uniformly distributed for all $\theta \in [0, \pi]$, one has $\varrho(\Theta_A^+) = \varrho(\Theta_A^-) = 1/2$, (60) and (62) become

$$\tilde{p}(\theta_A^z, \theta_B^z) = \frac{1}{2}\delta(\theta_A^z)\delta(\theta_B^z) + \frac{1}{2}\delta(\theta_A^z - \pi)\delta(\theta_B^z - \pi) \tag{63}$$

$$\tilde{p}(\theta_A^z, \theta_B^z) = \frac{1}{2}\delta(\theta_A^z)\delta(\theta_B^z - \pi) + \frac{1}{2}\delta(\theta_A^z - \pi)\delta(\theta_B^z), \tag{64}$$

where the superscript z is added to explicitly show the context dependency on the z-axis. However, (64) itself does not uniquely specify to the effect of correlation when the angular momentum orientations of the electron pair are always opposite to each other, because (64) can correspond to two possible correlations, as shown in Figure 6. To uniquely specify the effect of correlation when the angular momentum orientations of the electron pair are always opposite to each other, one needs to show the probability density along another axis perpendicular z-axis, say, the y-axis. Consider the case that the correlations between the electron pair are such that not only $\theta_A^z + \theta_B^z = \pi$ along the z-axis, but also $\theta_A^y + \theta_B^y = \pi$ along

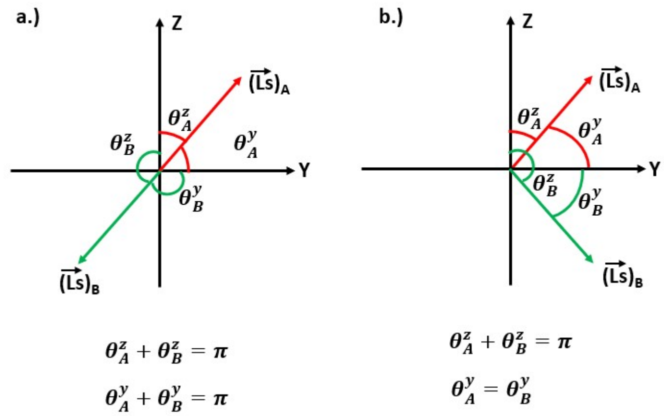

FIG. 6. Correlation of intrinsic angular momenta between two electrons. a.) Correlation corresponds to the singlet state $|\Psi^-\rangle$; b.) Correlation corresponds to Bell state $|\Psi^+\rangle$. Note that for both a.) and b.), we have $\theta_A^z + \theta_B^z = \pi$. It shows $\theta_A^z + \theta_B^z = \pi$ alone is unable to distinguish the correlations between a.) and b.).

the y-axis. This implies that we must also have

$$\tilde{p}(\theta_A^y, \theta_B^y) = \frac{1}{2}\delta(\theta_A^y)\delta(\theta_B^y - \pi) + \frac{1}{2}\delta(\theta_A^y - \pi)\delta(\theta_B^y). \tag{65}$$

When a pair of electrons are described by both (64) and (65), the correlation between the two electrons is equivalent to that is described by the spin singlet state in standard quantum mechanics.

$$|\Psi^-\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B). \tag{66}$$

We will further verify the equivalence in the next section when calculating the Bell-CHSH inequality.

On the other hand, if he correlations between the electron pair are such that $\theta_A^z + \theta_B^z = \pi$ along the z-axis, and $\theta_A^y = \theta_B^y$ along the y-axis, we must also have

$$\tilde{p}(\theta_A^y, \theta_B^y) = \frac{1}{2}\delta(\theta_A^y)\delta(\theta_B^y) + \frac{1}{2}\delta(\theta_A^y - \pi)\delta(\theta_B^y - \pi). \tag{67}$$

Thus, measurement along the y-axis results in the same spin directions for both electrons, while measurement along the z-axis yields opposite spin direction. A pair of electrons described by both (63) and (67) shares the same correlation specified by the spin triplet states in standard quantum mechanics

$$|\Psi^+\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\downarrow\rangle_B + |\downarrow\rangle_A|\uparrow\rangle_B). \tag{68}$$

Similarly, the combination of (63) and (65) is equivalent to the Bell state

$$|\Phi^-\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\uparrow\rangle_B - |\downarrow\rangle_A|\downarrow\rangle_B), \tag{69}$$

and the combination (63) and (67) is equivalent to

$$|\Phi^+\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle_A|\uparrow\rangle_B + |\downarrow\rangle_A|\downarrow\rangle_B). \tag{70}$$

We therefore recover the four Bell states.

C. Bell Inequality

Bell inequality is developed to prove that local hidden variable theory cannot produce the predictions of standard quantum mechanics. Bell inequality is violated when tested with an entangled pair of photons or electrons, demonstrating there is non-local correlation in standard quantum mechanics. Even though modern Bell experiments [53] confirm that such non-local correlation does not imply non-local causality, such Bell non-locality is still daunting to the quantum physics community because there is no intuitive physical picture to explain the phenomenon. Here we will investigate what insights the present theory can bring to this challenge.

Consider the experiment performed by Alice and Bob discussed in the previous section using an entangled pair of electrons. The CHSH version of Bell inequality reads

$$|E(a,b) - E(a,b') + E(a',b) + E(a',b')| \le 2. \tag{71}$$

Here a, a' are detector settings for Alice, and b, b' are detector settings for Bob. They usually refer to the directions of the magnetic field in the spin measurement and can be considered as unit vectors for the corresponding directions. The term $E(a,b)$ is the expectation value of measurement outcomes. That is, the statistical average of $S_A(a)S_B(b)$ where $S_A(a)$ and $S_B(b)$ are the spin measurement outcomes for setting a and b, respectively. For the setting a, $S_A(a) = +1$ for spin-up and $S_A(b) = -1$ for spin-down. Similar meanings can be inferred for the other $E(\cdot)$ terms in (71).

The expectation value $E(\cdot)$ depends on the state of the entangled pair. Suppose the entangled electron pair is in the singlet described by the joint probability density (63) and (65). It is important to note that both (63) and (65) are needed to give a complete description of the singlet state. In Appendix D, we prove that given the joint probability density (63) and (65),

$$E(a,b) = -(a\cdot b) = -\cos(\gamma), \tag{72}$$

where γ is the angle between unit vectors a and b. This result is exactly the same as the result from standard quantum mechanics. If Alice chooses $a = 0$ and $a' = \pi/2$, and Bob chooses $b = \pi/4$ and $b' = 3\pi/4$, we will have $|E(a,b) - E(a,b') + E(a',b) + E(a',b')| = 2\sqrt{2} > 2$. Thus, the CHSH inequality is violated.

The reason for the violation of the CHSH inequality is due to the fact that the joint probability density (63) and (65) cannot be factorized. In other words, they cannot be written as a product of two factors, one only contains variable for A and the other only contains variable for B. Such correlation can be maintained even though the two electrons are remotely separated. The root cause of such correlation is that the random variables are correlated, $\theta_A^z + \theta_B^z = \pi$ and $\theta_A^y + \theta_B^y = \pi$. These random variables are intrinsically local to the electrons. The term non-local correlation, or Bell non-locality, is misleading. Instead, as Hall pointed out [54], a better term is Bell non-separability.

VI. POSSIBLE EXPERIMENTS TO CONFIRM THE SPIN THEORY

The spin theory presented here so far recovers many of the results from standard quantum mechanics. In fact, the standard quantum theory of electron spin can be considered as an ideal approximation of the present theory when certain conditions are taken to the limits. When these conditions are not taken to the limits, the present theory will predict some results different from standard quantum mechanics. In this section, we will give two possible experiments to confirm the difference.

A. Dependency of Quantization on Interaction with Magnetic Field

One of the crucial assumption in the present theory is the derivation of (15), which assumes that the parameter m, which is related to the order of Tsallis relative entropy as shown in (12), monotonically increases when the electron travels along an inhomogeneous magnetic field B_z. Standard quantum mechanics postulates that quantization of electron spin is an intrinsic property. Practically, it corresponds to the case that whenever there is an external field B_z is applied, quantization occurs instantaneously, such that only two discrete results can be observed. However, the theory presented here allows m taking a finite number when the magnetic field B_z is applied. In that case, the probability density (13) should be used instead of the probability density (15) with two quantized outcomes.

Suppose the weak magnetic field in the Stern-Gerlach apparatus is given by

$$\vec{B} = B_0\hat{z} - \eta z\hat{z}, \tag{73}$$

where $B_0\hat{z}$ is a constant field, and the coefficient η determines the strength of the magnetic field. Since the orientation of the intrinsic magnetic moment of the electron follows a probability density distribution (13), the electron passing the Stern-Gerlach apparatus will experience a force that also depends on the orientation. Between angles θ and $\theta + d\theta$, the force is

$$\vec{F}(\theta) = -\nabla(p_m(\theta)\vec{\mu}\cdot\vec{B}) = \frac{e\hbar\eta}{2m_eZ_m}\cos^{2m+1}(\theta)\hat{z}, \tag{74}$$

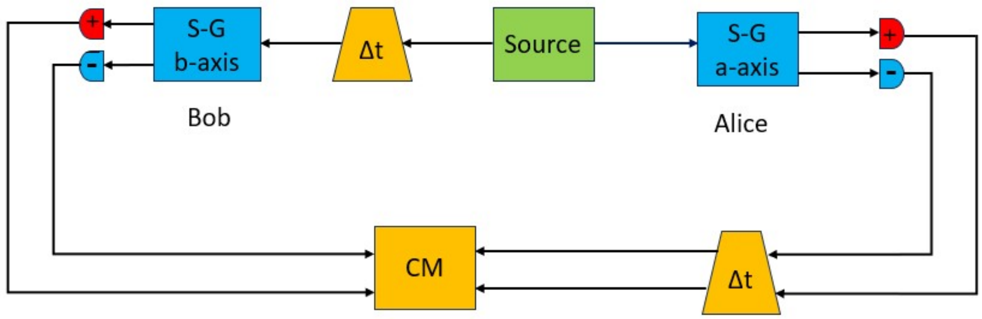

FIG. 7. Bell test with time delay. Starting with a Bell test experiment setup that has already successfully confirmed the violation of Bell-CHSH inequality. Then, we add a time delay in the measurement. Bob delays his measurement by a time period Δt, while the signals from Alice's Stern-Gerlach detectors are delayed by the same Δt before it is sent to the coincident monitor. When Δt is chosen to be sufficiently large, we expect the Bell-CHSH inequality to become non-violated.

where m_e is the mass of electrons. After traveling in the Stern-Gerlach apparatus for a period of time ΔT, the electron exhibits a displacement along the z-axis given by

$$\Delta z = \frac{e\hbar\eta}{4m_e^2 Z_m}(\Delta T)^2 \cos^{2m+1}(\theta). \tag{75}$$

The above calculation ignores the effect of the translational component of vacuum fluctuations and calculates the displacement along the z-axis using classical mechanics. In theory, this approximation is reasonable when ΔT is sufficiently large so that Δz is much larger than the effect of the translational component of vacuum fluctuations. Eq. (75) shows that in a magnetic field with extremely small η, the electrons passing the Stern-Gerlach apparatus will reach the detector screen with a continuous distribution, as shown in Fig. 2b.

However, as ΔT increases, parameter m monotonically increases as well, the measurement outcomes may rapidly converge to two discrete lines, and yield the same prediction as standard quantum mechanics. The possibility to set up an experiment to test the deviation from standard quantum mechanics depends on the balance of two factors, 1.) choosing the small η, and 2.) choosing a proper time duration ΔT that determines how long the electron is interacting with $\vec{B}$. Such an experiment can be challenged to implement. The experiment proposed next can be more realistic.

B. Bell Test with Time Delay

In Section V A, we show that for an entangled electron pair described by the joint probability density (63) and (65), the entangled relation can be maintained without time constraint. However, when Alice performs her measurement on the spin of electron A, obtains spin-up, and makes her prediction on Bob's measurement result of spin-up along the same direction, the prediction is only valid for a limited time. That is, there is a time constraint specified in (61) that Bob needs to perform his measurement on the spin of electron B in order to confirm Alice's prediction. If $t_m^B > t_m^A + \Delta t_+^B$, Bob will obtain the measurement result of spin-down instead. Furthermore, a singlet spin state is described by two probability density functions (64) and (65). After Alice performs her measurement on electron A and obtains result of spin-up, if Bob delays his measurement by a time such that the correlation described by either (64) or (65) is no longer satisfied, the singlet state is degraded and no longer valid.

We can explore the effect of entanglement degradation for a singlet state due to Bob's delayed measurement in the Bell test experiment, as illustrated in Figure 7. In a Bell test experiment, the electron source generates N copies of entangled electron pairs in a singlet state. For each pair, electrons A and B are sent to two Stern-Gerlach apparatuses that are remotely separated[7]. Alice measures the spin of electron A along an axis labeled by unit vector a. At the same time, Bob measures the spin of electron B along an axis labeled by unit vector b. The detectors for each Stern-Gerlach apparatus detect the spin-up and spin-down results, and generate corresponding signals that are sent to the coincidence monitor. Counting the four types $(++, +-, -+, --)$ of coincidences for the N copies of the singlet spin pairs, one can calculate whether the Bell-CHSH inequality is violated. As shown in Section V C, for the singlet spin pair described by (64) or (65), if Alice chooses $a = 0$ and

[7] As mentioned earlier, in modern Bell test experiment [53], the two electrons are already separated remotely before being prepared to be entangled. The electron entanglement is achieved by entanglement swapping with photons. There is no need to move the electrons and therefore, no need to worry about the impacts on the correlation due to the movement of electrons.

$a' = \pi/2$, and Bob chooses $b = \pi/4$ and $b' = 3\pi/4$, maximum violation of the Bell-CHSH inequality is achieved.

Using exactly the same experiment setup, we add an extra step here. Instead that Bob performs the measurement on electron B at the same time as Alice's measurement, he delays his measurement by a time period of Δt. When Δt is sufficiently large, the entanglement correlation of a singlet state is degraded, and we expect that the Bell-CHSH inequality will not be violated. Since we do not know how long to delay in order to observe the non-violation of Bell-CHSH inequality, we can start with a very small delay Δt_1 such that the Bell-CHSH inequality is still violated, and then gradually increase the delay. We repeat the experiments M times, but choose the delay times such that $\Delta t_1 < \Delta t_2 < \ldots < \Delta t_M$. Each experiment consumes N copies of singlet electron pairs. According to the present theory, we expect the violation of Bell-CHSH inequality will be gradually reduced, and at some time the results will satisfy the Bell-CHSH inequality. The reason for this is that with a sufficient large delay Δt, the orientation of intrinsic angular momentum of electron B can fluctuate to switch from trending upward to trending downward, or vice versus, as depicted in Figure 4. When a orientation switching occurs, either $\theta_B \in \Theta_B^+ \to \theta_B \in \Theta_B^-$, or $\theta_B \in \Theta_B^- \to \theta_B \in \Theta_B^+$, the initial correlation between θ_A and θ_B is changed. As Δt increases, the orientation of electron B appears as a random variable again for the N copies of electron pairs, so that the initial correlation between θ_A and θ_B is no longer hold. Consequently, the Bell-CHSH inequality becomes satisfied. Appendix D gives the mathematical calculation of such a scenario.

The reason the entanglement relation between A and B is degraded is due to the orientation fluctuations of the intrinsic angular momentum. This is different from the decoherence theory where the entanglement relation between A and B can be destroyed by environmental disturbance. Here, however, the degradation of entanglement during measurement is intrinsic to the electron pair, nothing to do with the environment. Without the measurement event, the entanglement relation between A and B can be maintained without a time limit. Certainly, if there is also environmental disturbance, the entanglement relation will be destroyed without measurement. So the present theory is not incompatible with the decoherence theory. However, if confirmed by experiment, the intrinsic degradation of entanglement during measurement can have a non-negligible implication on its applications in quantum computing and quantum information.

The challenges to perform the above experiment come from choosing the proper time delay Δt. To meet the locality condition of Bell test, Δt needs to be smaller than the time for a light signal traveling from Alice to Bob. This is to eliminate the potential loophole of hidden signals sent from Alice to Bob. But Δt can not be too small either so that the orientation of intrinsic angular momentum of B switches from $\theta_B \in \Theta_B^+ \to \theta_B \in \Theta_B^-$, or $\theta_B \in \Theta_B^- \to \theta_B \in \Theta_B^+$. To resolve this contention, one can separate electrons A and B with a very long distance so that the time window to comply with locality condition is large enough. A more subtle argument is that since the goal here is to show that the Bell-CHSH inequality is not violated, the requirement of locality is not applicable. Thus, Δt does not need to be smaller than the time for a light signal traveling from Alice to Bob.

VII. DISCUSSION AND CONCLUSIONS

A. Conceptual Implications

One advantage of the present theory is it provides an intuitive physical picture to understand the phenomenon of entanglement between two electron spins. As shown in Section V, entanglement is simply the manifestation of the correlation between the orientations of intrinsic angular momenta of two electrons. Even though the orientations are random variables, they can have correlation due to previous interactions between them, or due to previous interactions with common external fields. More importantly, the orientation of intrinsic angular momentum is an intrinsically local property. Once the correlation of orientations between two electrons is established, it can be maintained even if the two electrons are remotely separated, until there is measurement performed on them. Performing measurement on one electron will not cause any change on the other electron. However, the measurement outcome reveals the orientation of the measured electron's intrinsic angular momentum, which enables us to infer the spin measurement outcome of the other electron because of the correlation on the orientations. The seemingly non-local correlation has nothing to do with any non-local causal relation.

To further clarify the point that there is no non-causal effect in the Bell experiment with a pair of electrons in the singlet state, suppose after Alice performs her experiment and obtains the result of spin-up for electron A, she does not send the measurement outcome to Bob. Since the measurement action of Alice is local and has no influence on electron B, from Bob's point of view, he still perceives both A and B in the original singlet state described by (64) and (65). Alice's and Bob's descriptions on B are different but both are valid. This is consistent with the relational quantum mechanics (RQM) interpretation advocated by Rovelli [15, 55]. Since Bob does not know the measurement outcome from Alice, he predicts that if he measures electron B along any direction, he will obtain spin-up and spin-down with equal probability. However, if Alice sends the measurement results to Bob, Bob can infer the orientation of intrinsic angular momentum of electron B. Bob's knowledge on electron B is updated with the new information from Alice. Thus, he can predict if he measures electron B along the same direction as Alice's measurement, he will obtain the result of spin-down. This analysis also illustrates that the wave function in quantum mechanics is just a mathematical tool and not associated with certain ontic properties.

We see the spin model and theory presented here gives an intuitive physical picture that may help to consolidate several quantum physics interpretations. In addition, our theory supports the idea that measurement outcome is context dependent. Measurement outcome for the spin of an electron depends on the measurement context, that is, the direction of magnetic field used in the Stern-Gerlach apparatus. The two level discreteness of measurement outcome is the result of a sufficiently strong magnetic field. For an electron with initially completely random orientation of intrinsic angular momentum, we can measure the spin along any direction but always get spin-up and spin-down with equal probability. There is no predefined spin property without specifying an experimental context.

B. Limitations

There are several limitations we need to point out for future investigations. The first challenge is to explain why the magnitude of intrinsic angular momentum of an electron is $\hbar/2$. Although Appendix B gives a reasonable derivation, which is also based on the extended least action principle, there is a strong assumption on random circular motion due to vacuum fluctuation. A more intuitive model with weaker assumptions is desirable.

Secondly, assumption 4 suggests that when the interaction between the electron and an inhomogeneous magnetic field in a Stern-Gerlach apparatus takes place, the order of relative entropy monotonically increases, so that the direction of intrinsic angular momentum tends to align with the direction of the magnetic field. We give an intuitive explanation by taking into consideration that the electron is not an idealized point particle. A more rigorous theory to confirm this assumption is needed. The spin theory based on stochastic mechanics [9] gives similar results. However, stochastic mechanics relies on velocity fields which are non-local in nature, which makes stochastic mechanics a less favorable theory.

Lastly, the Schrödinger-Pauli equation derived in Section IV is only valid when the external magnetic field is along the z-axis. We need to extend the derivation for an external magnetic field with arbitrary direction.

C. Conclusions

From the extended least action principle, and with additional assumptions on the intrinsic angular momentum of an electron, particularly its random orientations, we are able to develop a theory that recovers the quantum formulation for electron spin. The key component of the theoretical framework here is the introduction of relative entropy for both translational and rotational random fluctuations. The two-level quantization of spin measurement outcomes is no longer a postulate. Instead, it is a mathematical consequence based on Assumption 4. We also obtain the same formulation as standard quantum mechanics when the direction of magnetic field of a second Stern-Gerlach apparatus is rotated with an angle from the direction of magnetic field of the first Stern-Gerlach apparatus. Recursively applying the extended least action principle, we have derived the Schrödinger-Pauli equation.

An important result is that we provide an intuitive physical model and formulation to explain the entanglement phenomenon between two electron spins. The root cause of entanglement is the correlation between the orientations of the intrinsic angular momenta of the two electrons. Essentially, the correlation is established between two random variables due to previous interactions. Since the orientation of the angular momentum is an intrinsic local property of the electron, the correlation can be maintained even though the two electrons are space-like separated and have no further interaction. Mathematically, we give an equivalent formulation of the four Bell states using probability density function instead of wave function. Using the probability density functions for a pair of entangled electrons, we prove that the Bell-CHSH inequality is violated. The violation of Bell-CHSH inequality is due to the same root cause of entanglement, that is, correlation of the orientations of the angular momenta which is preserved even when the two electrons are separated.

The standard quantum theory of electron spin can be considered as an ideal approximation of the present theory when certain conditions are taken to the limits. When these conditions are not taken to the limits, the present theory predicts results that are different from standard quantum mechanics. For instance, if the Stern-Gerlach apparatus is set up with a sufficiently weak gradient of the magnetic field along the z-axis, we expect the electron detector screen will exhibit continuous distribution of displacements along the z-axis, instead of only two discrete lines. Another more interesting experiment proposed is to add a delay before Bob's measurement in the typical Bell test experiment, which could result in the non-violation of the Bell-CHSH inequality. The second experiment shows the effect of time dependency in measurements involving spin entanglement. If confirmed by experiment, such an effect can impose a limitation of the application of spin entanglement in quantum computing and quantum information.

The interplay of the extended least action principle and the information metrics for vacuum fluctuations prove to be very fruitful. We have shown the theoretical framework based on the extended least action principle has successfully recovered the non-relativistic quantum mechanics [44], the relativistic quantum scalar field theory [45], and in the paper, the quantum theory for electron spin. We believe the results in this paper are more interesting and conceptually more important because we can explain the entanglement phenomenon with an intuitive physical picture, and predict new results different from standard quantum mechanics that can be potentially tested in the proposed experiments.

[1] A. Einstein, B. Podolsky, N. and Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Phys. Rev.* **47**, 777-780 (1935)

[2] D. Bohm, Y. Aharonov, "Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky", *Phys. Rev.* **108**(4), 1070 (1957)

[3] J. Bell, "On the Einstein Podolsky Rosen paradox", *Physics Physique Fizika* **1**, 195 (1964)

[4] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, "Bell Nonlocality", Rev. Mod. Phys. 86, 419 (2014) arXiv:1303.2849

[5] Tomonaga, S: "The Story of Spin." *Univ. of Chicago Press* Chicago (1997)

[6] C. T. Sebens, "Particles, fields, and the measurement of electron spin", *Synthese* **198**:11943-11975 (2021)

[7] T. G. Dankel, "Higher spin states in the stochastic mechanics of the Bopp–Haag spin model", *J. Math. Phys.* **18** 253 (1997)

[8] W. G. Fairs, "Spin Correlation in Stochastic Mechanics", *Found. of Phys.* **12** 825 (1982)

[9] Beyer, M., Paul, W. "Stern–Gerlach, EPRB and Bell Inequalities: An Analysis Using the Quantum Hamilton Equations of Stochastic Mechanics". *Found. of Phys.* **54** 20 (2024)

[10] J. Weyssenhof, "On Two Relativistic Models of Dirac's Electron." *Acta Phys. Pol.***9** 47 (1947)

[11] Barut, A.O., Sanghi. N.: "Classical Model of the Dirac Electron." *Phys. Rev. Lett.***52** 2009-2012 (1984)

[12] A. Niehaus, "A Probalilistic Model of Spin and Spin Measurements", *Found. of Phys.* **46** 3 (2016)

[13] A. Niehaus, "Trying an Alternative Ansatz to Quantum Physics", *Found. of Phys.* **52** 41 (2022)

[14] A. Zeilinger, "A foundational principle for quantum mechanics," *Found. Phys.* **29** no. 4, (1999) 631–643.

[15] C. Rovelli, "Relational quantum mechanics," *Int. J. Theor. Phys.* **35** 1637–1678 (1996), `arXiv:quant-ph/9609002 [quant-ph]`.

[16] C. Brukner and A. Zeilinger, "Information and fundamental elements of the structure of quantum theory," *in "Time, Quantum, Information", edited by L.. Castell and O. Ischebeck (Springer, 2003)* , `quant-ph/0212084`. `http://arxiv.org/abs/quant-ph/0212084`.

[17] C. Brukner and A. Zeilinger, "Operationally invariant information in quantum measurements," *Phys. Rev. Lett.* **83** (1999) 3354–3357, `quant-ph/0005084`. `http://arxiv.org/abs/quant-ph/0005084`.

[18] C. Brukner and A. Zeilinger, "Young's experiment and the finiteness of information," *Phil. Trans. R. Soc. Lond. A* **360**, 1061 (2002) `quant-ph/0201026`. `http://arxiv.org/abs/quant-ph/0201026`.

[19] C. A. Fuchs, Quantum Mechanics as Quantum Information (and only a little more). arXiv:quant-ph/0205039, (2002)

[20] Č. Brukner and A. Zeilinger, "Information invariance and quantum probabilities," *Found. Phys.* **39** no. 7, 677–689 (2009).

[21] C. Brukner, M. Zukowski, and A. Zeilinger, "The essence of entanglement," `quant-ph/0106119`. `http://arxiv.org/abs/quant-ph/0106119`.

[22] R. W. Spekkens, "Evidence for the epistemic view of quantum states: A toy theory," *Phys. Rev. A* **75** no. 3, 032110 (2007).

[23] R. W. Spekkens, "Quasi-quantization: classical statistical theories with an epistemic restriction," `1409.5041`. `http://arxiv.org/abs/1409.5041`.

[24] T. Paterek, B. Dakic, and C. Brukner, "Theories of systems with limited information content," *New J. Phys.* **12**, 053037 (2010) `0804.1423`. `http://arxiv.org/abs/0804.1423`.

[25] T. Görnitz and O. Ischebeck, *An Introduction to Carl Friedrich von Weizsäcker's Program for a Reconstruction of Quantum Theory.* Time, Quantum and Information. Springer (2003)

[26] H. Lyre, "Quantum theory of ur-objects as a theory of information," *International Journal of Theoretical Physics* **34** no. 8, 1541–1552 (1995)

[27] L. Hardy, "Quantum theory from five reasonable axioms," `arXiv:quant-ph/0101012 [quant-ph]`.

[28] L. Masanes and M. P. Müller, "A derivation of quantum theory from physical requirements," *New J. Phys.* **13** no. 6, 063001 (2011)

[29] M. P. Müller and L. Masanes, "Information-theoretic postulates for quantum theory," `arXiv:1203.4516 [quant-ph]`.

[30] L. Masanes, M. P. Müller, R. Augusiak, and D. Perez-Garcia, "Existence of an information unit as a postulate of quantum theory," *PNAS vol 110 no 41 page 16373 (2013)* (08, 2012) , `1208.0493`. `http://arxiv.org/abs/1208.0493`.

[31] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Phys. Rev. A* **84** no. 1, 012311 (2011)

[32] M. P. Müller and L. Masanes, "Three-dimensionality of space and the quantum bit: how to derive both from information-theoretic postulates," *New J. Phys. 15,* **053040** (2013) , `arXiv:1206.0630 [quant-ph]`.

[33] L. Hardy, "Reconstructing quantum theory," `1303.1538`. `http://arxiv.org/abs/1303.1538`.

[34] S. Kochen, "A reconstruction of quantum mechanics," *arXiv preprint arXiv:1306.3951* (2013)

[35] P. Goyal, "From Information Geometry to Quantum Theory," *New J. Phys.* **12**, 023012 (2010) `0805.2770`. `http://arxiv.org/abs/0805.2770`.

[36] M. Reginatto and M.J.W. Hall, "Information geometry, dynamics and discrete quantum mechanics," *AIP Conf. Proc.* **1553**, 246 (2013); arXiv:1207.6718.

[37] P. A. Höhn, "Toolbox for reconstructing quantum theory from rules on information acquisition," *Quantum* **1**, 38 (2017) `arXiv:1412.8323 [quant-ph]`.

[38] P. A. Höhn, "Quantum theory from questions," *Phys. Rev. A* **95** 012102, (2017) `arXiv:1517.01130 [quant-ph]`.

[39] W. Stuckey, T. McDevitt, and M. Silberstein, "No preferred reference frame at the foundation of quantum me-

chanics," *Entropy* **24**, 12 (2022)

[40] M. Mehrafarin, "Quantum mechanics from two physical postulates," *Int. J. Theor. Phys.*, **44**, 429 (2005); arXiv:quant-ph/0402153.

[41] A. Caticha, "Entropic Dynamics, Time, and Quantum Theory," *J. Phys. A: Math. Theor.* **44**, 225303 (2011); arXiv.org: 1005.2357.

[42] B. R. Frieden, "Fisher Information as the Basis for the Schrödinger Wave Equation," *American J. Phys.* **57**, 1004 (1989)

[43] M. Reginatto, "Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information," *Phys. Rev. A* **58**, 1775 (1998)

[44] J. M. Yang, "Quantum Mechanics Based on an Extended Least Action Principle and Information Metrics of Vacuum Fluctuations", accepted on Feb. 6th, 2024 and will appear in *Found. Phys.*, arXiv:2302.14619, (2024)

[45] J. M. Yang, "Quantum Scalar Field Theory Based on an Extended Least Action Principle", *Int. J. Theor. Phys.* **63** 15 (2024)

[46] B. R. Frieden, "Physics from Fisher Information," Cambridge University Press, Cambridge (1999)

[47] E. T. Jaynes, "Prior information." *IEEE Transactions on Systems Science and Cybernetics* **4**(3), 227–241 (1968)

[48] D. Bohm, "A suggested interpretation of the quantum theory in terms of hidden variables, I and II," *Phys. Rev.* **85**, 166 and 180 (1952).

[49] Stanford Encyclopedia of Philosophy: Bohmian Mechanics, (2021)

[50] C. Tsallis, "Possible generalization of Boltzmann–Gibbs statistics," *J. Stat. Phys.* **52**, 479–487 (1998)

[51] T. van Erven, P. Harremoës, "Rényi divergence and Kullback-Leibler divergence," *IEEE Transactions on Information Theory* **70**, 7 (2014)

[52] F. Nielsen, and R. Nock, "On Rényi and Tsallis entropies and divergences for exponential families," *J. Phys. A: Math. and Theo.* **45**, 3 (2012)

[53] B. Hensen, *et al*, "Experimental loophole-free violation of a Bell inequality using entangled electron spins separated by 1.3 km", *Nature* **526**, 682-686 (2015)

[54] M. J. W. Hall, "The significance of measurement independence for Bell inequalities and locality", In: At the Frontier of Spacetime, ed. T. Asselmeyer-Maluga (Springer, Switzerland, 2016), ch. 11 (pp. 189-204)

[55] Smerlak M., and Rovelli, C.: Relational EPR, Found. Phys., 37, 427-445 (2007)

Appendix A: Deriving Basic Quantum Formulation From the Extended Least Action Principle

Basic quantum formulation can be derived by recursively applying the extended least action principle in two steps. First, we consider the dynamics of a system with an infinitesimal time internal Δt due to vacuum fluctuation only. Define the probability for the system to transition from a 3-dimensional space position $\mathbf{x}$ to another position $\mathbf{x}+\mathbf{w}$, where $\mathbf{w}=\Delta\mathbf{x}$ is the displacement in 3-dimensional space due to fluctuations, as $\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})d^3\mathbf{w}$. The expectation value of classical action is $S_c=\int\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})Ld^3\mathbf{w}dt$. Since we only consider the vacuum fluctuations, the Lagrangian L only contains the kinetic energy, $L=\frac{1}{2}m\mathbf{v}\cdot\mathbf{v}$. For an infinitesimal time internal Δt, one can approximate the velocity $\mathbf{v}=\mathbf{w}/\Delta t$. This gives

$$S_c=\frac{m}{2\Delta t}\int_{-\infty}^{+\infty}\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})\mathbf{w}\cdot\mathbf{w}d^3\mathbf{w}. \tag{A1}$$

The information metrics I_f is defined as the Kullback–Leibler divergence, to measure the information distance between $\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})$ and a uniform prior probability distribution μ that reflects the vacuum fluctuations are completely random with maximal ignorance [47**?**],

$$\begin{aligned} I_f &=: D_{KL}(\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})||\mu) \\ &= \int\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})ln[\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})/\mu]d^3\mathbf{w}. \end{aligned}$$

Insert both S_c and I_f into (1) and perform the variation procedure, one obtain

$$\wp(\mathbf{x}+\mathbf{w}|\mathbf{x})=\frac{1}{Z}e^{-\frac{m}{\hbar\Delta t}\mathbf{w}\cdot\mathbf{w}}, \tag{A2}$$

where Z is a normalization factor. Equation (A2) shows that the transition probability density is a Gaussian distribution. The variance for the vector component w_i is $\langle w_i^2\rangle=\hbar\Delta t/2m$, where $i\in\{1,2,3\}$ denotes the spatial index. Recalling that $w_i/\Delta t=v_i$ is the approximation of velocity due to the vacuum fluctuations, one can deduce

$$\langle\Delta x_i\Delta p_i\rangle=\frac{\hbar}{2}. \tag{A3}$$

Applying the Cauchy–Schwarz inequality gives

$$\langle\Delta x_i\rangle\langle\Delta p_i\rangle\geq\hbar/2. \tag{A4}$$

In the second step, we will derive the dynamics for a cumulative period from $t_A\to t_B$. In classical mechanics, the equation of motion is described by the Hamilton-Jacobi equation,

$$\frac{\partial S}{\partial t}+\frac{1}{2m}\nabla S\cdot\nabla S+V=0. \tag{A5}$$

Suppose the initial condition is unknown, and define $\rho(x,t)$ as the probability density for finding a particle in a given volume of the configuration space. The probability density must satisfy the normalization condition $\int\rho(\mathbf{x},t)d^3\mathbf{x}=1$, and the continuity equation

$$\frac{\partial\rho(\mathbf{x},t)}{\partial t}+\frac{1}{m}\nabla\cdot(\rho(\mathbf{x},t)\nabla S)=0.$$

The pair (S,ρ) completely determines the motion of the classical ensemble. As pointed out by Hall and Reginatto [**?** **?**], the Hamilton-Jacobi equation, and the continuity equation, can be derived from classical action

$$A_c=\int\rho\{\frac{\partial S}{\partial t}+\frac{1}{2m}\nabla S\cdot\nabla S+V\}d^3\mathbf{x}dt \tag{A6}$$

through fixed point variation with respect to ρ and S, respectively. Note that A_c and S are different physical variables, where A_c can be considered as the ensemble average of classical action and S is a generation function that satisfied $\mathbf{p}=\nabla S$ [44].

To define the information metrics for the vacuum fluctuations, I_f, we slice the time duration $t_A\to t_B$ into N short time steps $t_0=t_A,\ldots,t_j,\ldots,t_{N-1}=t_B$, and each step is an infinitesimal period Δt. In an infinitesimal time period at time t_j, the particle not only moves according to the Hamilton-Jacobi equation but also experiences random fluctuations. Such additional revelation of distinguishability due to the vacuum fluctuations on top of the classical trajectory is measured by the following definition,

$$I_f =: \sum_{j=0}^{N-1} \langle D_{KL}(\rho(\mathbf{x}, t_j)||\rho(\mathbf{x}+\mathbf{w}, t_j))\rangle_w \tag{A7}$$

$$= \sum_{j=0}^{N-1} \int d^3\mathbf{w} d^3\mathbf{x} \wp(\mathbf{x}+\mathbf{w}|\mathbf{x})\rho(\mathbf{x}, t_j) ln \frac{\rho(\mathbf{x}, t_j)}{\rho(\mathbf{x}+\mathbf{w}, t_j)}. \tag{A8}$$

When $\Delta t \to 0$, I_f turns out to be [44]

$$I_f = \int d^3\mathbf{x} dt \frac{\hbar}{4m}\frac{1}{\rho}\nabla\rho\cdot\nabla\rho. \tag{A9}$$

Eq. (A9) contains the term related to Fisher information for the probability density [46] but bears much more physical significance than Fisher information. Inserting (A6) and (A9) into (1), and performing the variation procedure on I with respect to S gives the continuity equation, while variation with respect to ρ leads to the quantum Hamilton-Jacobi equation,

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\nabla S\cdot\nabla S + V - \frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}} = 0, \tag{A10}$$

Defined a complex function $\Psi = \sqrt{\rho}e^{iS/\hbar}$, the continuity equation and the extended Hamilton-Jacobi equation (A10) can be combined into a single differential equation,

$$i\hbar\frac{\partial\Psi}{\partial t} = [-\frac{\hbar^2}{2m}\nabla^2 + V]\Psi, \tag{A11}$$

which is the Schrödinger Equation.

The last term in (A10) is the Bohm quantum potential [48]. The Bohm potential is considered responsible for the non-locality phenomenon in quantum mechanics [49]. Historically, its origin is mysterious. Here we show that it originates from the information metrics related to relative entropy, I_f.

As pointed out in [44], the choice of relative entropy for I_f can be Renyi divergence or Tsallis divergence, which results in a family of Schrödinger equations that depends on the order of relative entropy α. When $\alpha = 1$, the regular Schrödinger equation is recovered. The flexibility to choose general relative entropy definitions is a very useful mathematical tool for further theoretical investigation of more advanced quantum theory, as we have shown in the present work.

Appendix B: Intrinsic Angular Momentum Due to Vacuum Fluctuations

Assumption 1 on vacuum fluctuations in Section II does not specify the details of the fluctuating motion. The electron exhibits random displacement due to vacuum fluctuation and such displacement is described as a vector. The displacement vector is intrinsically local in the sense that it is independent of the reference of origin of the global coordinate system. Given a local point and a infinitesimal time period, the displacement vector can be just a translational motion, i.e., the direction of velocity is in parallel with the displacement vector. However, there is no reason that the direction of velocity must be in parallel with the displacement vector. When the velocity vector has a component perpendicular to the displacement vector, the electron will exhibit rotational movement that can be characterized by an angular momentum. Here we will further assume that rotational movement is circular. Thus, the vacuum fluctuation not only causes the translational displacement, but also comprises components of motion as circular rotations. Essentially, the vacuum fluctuation produces virtual circular rotations with random radius. Such a model has been proposed in [12, 13], but no derivation was provided on the magnitude of intrinsic angular momentum. Here, with the help of the extended least action principle, we will show that such random circular rotations give rise to the intrinsic angular momentum with magnitude of $\hbar/2$.

Denote the radius vector of circular motion in an infinitesimal time period Δt as $\vec{u}$ and the velocity vector as $\dot{\vec{u}}$. Here we will only consider the circular motion of the electron such that $\dot{\vec{u}}$ is perpendicular to $\vec{u}$. The angular momentum is the cross product of $\vec{u}$ and $m\dot{\vec{u}}$

$$\vec{L}_s = \vec{u}\times m\dot{\vec{u}}. \tag{B1}$$

The orientation of $\vec{L}_s$ is completely random. We choose the coordinate plane perpendicular to $\vec{L}_s$ for further analysis of the circular motion. Here, the radius u is a random variable. Let ω be the angular frequency. Then, the magnitude of velocity is $\dot{u} = \omega u$, and the magnitude of angular momentum is $L_s = m\omega u^2$. Since the radius u is a random variable, we denote $p(u)$ as the probability density for radius being u. It must satisfy the normalization condition

$$\int_0^\infty p(u)du = 1. \tag{B2}$$

For this circular motion, the Lagrangian L only contains the kinetic energy, $L = \frac{1}{2}m(\omega u)^2$. For an infinitesimal period of time Δt, the expectation value of classical action is

$$A_c = \frac{m}{2}\int_0^\infty \int_0^{\Delta t} p(u)(\omega u)^2 du dt. \tag{B3}$$

Note that $d\varphi = \omega dt$ where φ is the angle of circular rotation. In the infinitesimal period of time Δt, the electron rotates $\Delta\varphi = \omega\Delta t$. We can rewrite (B3) as

$$A_c = \frac{m}{2}\int_0^\infty \int_0^{\Delta\varphi} p(u)\omega u^2 du d\varphi. \tag{B4}$$

Define the information metrics exhibited during this time period as a relative entropy

$$I_f := \int_0^\infty \int_0^{\Delta\varphi} p(u) ln\frac{p(u)}{\mu} du d\varphi, \tag{B5}$$

where μ is a uniform probability density to reflect the total ignorance of knowledge due to complete randomness of fluctuations. Then, the total action, per (1), is

$$A_t = A_c + \frac{\hbar}{2}I_f = \frac{m}{2}\int p(u)\omega u^2 du d\varphi + \frac{\hbar}{2}\int p(u) ln\frac{p(u)}{\mu} du d\varphi. \tag{B6}$$

Taking the variation of A_t over the functional variable $p(u)$, and demanding $\delta A_t = 0$, we obtain

$$\frac{m}{2}\omega u^2 + \frac{\hbar}{2}(ln\frac{p(u)}{\mu} + 1) = 0. \tag{B7}$$

This gives

$$p(u) = \frac{1}{Z}e^{-\frac{m\omega}{\hbar}u^2}, \tag{B8}$$

where Z is the normalization factor. We can then compute the variance of u as

$$\langle u^2 \rangle = \int p(u)u^2 du = \frac{\hbar}{2m\omega}. \tag{B9}$$

This gives the expectation value of local angular momentum

$$\langle L_s \rangle = \langle m\omega u^2 \rangle = \frac{\hbar}{2}. \tag{B10}$$

Thus, by assuming that the vacuum fluctuations cause random circular motions for the electron, we show that the electron possesses an intrinsic angular momentum with magnitude of $\hbar/2$. Such intrinsic angular momentum is local in the sense that it is independent of the global orbital movement. Its orientation is completely random. These properties have been summarized as Assumption 3 in Section III. But based on the derivation showing in this appendix, we can replace Assumption 3 with

> *Assumption 3a – The vacuum fluctuations cause an electron in free space to exhibit random circular movements in addition to random translational movements.*

Since the assumption of circular movements appears to be a very strong assumption, it is more prudent to just adopt Assumption 3. We speculate that classical field theory could be a better framework to describe these random movements more accurately, which is beyond the scope of this work but has been studied in [6].

Appendix C: Proof of (36) in Standard Quantum Mechanics

In standard quantum mechanics, if an electron is measured with result of spin-up along a direction that is tilt an angle β_1 from the z-axis, its state can be described as

$$|\psi^+_{\beta_1}\rangle = \cos(\beta_1/2)|\uparrow\rangle + \sin(\beta_1/2)|\downarrow\rangle. \tag{C1}$$

This state can be obtained by rotating the state of spin-up along the z-axis by an angle β_1. The rotation operator is $\hat{R}(\beta_1) = e^{-i\beta_1\sigma_z/2}$ where σ_z is the z-component of the Pauli operator. Suppose the electron is in this initial state, and passes a Stern-Gerlach apparatus that is configured with a magnetic field with direction tilt an angle β_2 from the z-axis. The measurement operator for spin-up along the β_2 direction is $\hat{O}_+ = |\psi^+_{\beta_2}\rangle\langle\psi^+_{\beta_2}|$ where

$$|\psi^+_{\beta_2}\rangle = \cos(\beta_2/2)|\uparrow\rangle + \sin(\beta_2/2)|\downarrow\rangle. \tag{C2}$$

The probability is then given by

$$p(\uparrow|\beta_2,\beta_1) = |\langle\psi^+_{\beta_1}|\psi^+_{\beta_2}\rangle|^2 = \cos^2((\beta_2-\beta_1)/2). \tag{C3}$$

The measurement operator for spin-down along the β_2 direction is $\hat{O}_- = |\psi^-_{\beta_2}\rangle\langle\psi^-_{\beta_2}|$ where

$$|\psi^-_{\beta_2}\rangle = \sin(\beta_2/2)|\uparrow\rangle - \cos(\beta_2/2)|\downarrow\rangle. \tag{C4}$$

The probability is then given by

$$p(\downarrow|\beta_2,\beta_1) = |\langle\psi^+_{\beta_1}|\psi^-_{\beta_2}\rangle|^2 = \sin^2((\beta_2-\beta_1)/2). \tag{C5}$$

Eqs. (C3) and (C5) are the same as (36).

Appendix D: Proof of (72)

For a spin singlet state, the correlation can be described as $\theta^z_A + \theta^z_B = \pi$ and $\theta^y_A + \theta^y_B = \pi$. The joint probability density functions are given by (63) and (65). The physical implication is that if measurement of A along the z-axis gives result of spin-up, then measurement of B along the z-axis will obtain result of spin-down. Similarly, if measurement of A along the y-axis gives result of spin-up, then measurement of B along the y-axis will obtain result of spin-down. Both measurements are mutual exclusive and are needed to provide a complete description of the entangled pair. The expectation value of Alice measuring A along direction of unit vector a and obtaining spin-up, and Bob measuring B along direction of unit vector b and obtaining spin-up, is denoted as $E(+,+|a,b,\tilde{p})$ where $\tilde{p}$ is the initial joint probability density for a singlet state. Since $\tilde{p}$ comprises two mutual exclusive components given by (63) and (65), we have

$$E(+,+|a,b,\tilde{p}) = E(+,+|a,b,\tilde{p}_z) + E(+,+|a,b,\tilde{p}_y) \tag{D1}$$

First we consider initial condition described by the joint probability density functions are given by (63). This initial condition can be further decomposed such that half of the chance that the joint probability density $\tilde{p}_{z,1}(\theta^z_A,\theta^z_B) = \delta(\theta^z_A)\delta(\theta^z_B-\pi)$ and the other half chance that $\tilde{p}_{z,2}(\theta^z_A,\theta^z_B) = \delta(\theta^z_A-\pi)\delta(\theta^z_B)$. Now for the initial condition that $\tilde{p}_{z,1}(\theta^z_A,\theta^z_B) = \delta(\theta^z_A)\delta(\theta^z_B-\pi)$, suppose Alice measures A along a direction determined by unit vector a which forms an angle with z-axis by α, and Bob measures B along a direction determined by unit vector b which forms an angle with z-axis by β. In typical Bell experiments, the measurements of Alice and Bob are performed almost at the same time. Therefore, we will not consider the time dependency of measurement outcomes as discussed in Section V A.

The probability density of measurement outcome for Alice is given by (29) as

$$\bar{p}(\theta^z_a|\alpha) = \cos^2(\alpha/2)\delta(\theta^z_a) + \sin^2(\alpha/2)\delta(\theta^z_a-\pi), \tag{D2}$$

while the probability density of measurement outcome for Bob is given by

$$\bar{p}(\theta^z_b|\beta) = \sin^2(\beta/2)\delta(\theta^z_b) + \cos^2(\beta/2)\delta(\theta^z_b-\pi), \tag{D3}$$

Then, the statistical average of $S_A(a)S_B(b)$ for $(+,+)$, $(+,-)$, $(-,+)$, $(-,-)$ are

$$\begin{aligned} E(+,+|a,b,\tilde{p}_z) &= \cos^2(\alpha/2)\sin^2(\beta/2) \\ E(+,-|a,b,\tilde{p}_z) &= -\cos^2(\alpha/2)\cos^2(\beta/2) \\ E(-,+|a,b,\tilde{p}_z) &= -\sin^2(\alpha/2)\sin^2(\beta/2) \\ E(-,-|a,b,\tilde{p}_z) &= \sin^2(\alpha/2)\cos^2(\beta/2). \end{aligned} \tag{D4}$$

The overall expectation value for this initial condition is

$$E(a,b|\tilde{p}_z) = E(+,+|a,b,\tilde{p}_z) + E(+,-|a,b,\tilde{p}_z) + E(-,+|a,b,\tilde{p}_z) + E(-,-|a,b,\tilde{p}_z) = -\cos(\alpha)\cos(\beta). \tag{D5}$$

The initial condition can also be described by joint probability density $\tilde{p}_{z,2}(\theta_A^z,\theta_B^z) = \delta(\theta_A^z-\pi)\delta(\theta_B^z)$. Performing the same calculation with this initial condition, one will find the resulting $E(a,b|\tilde{p}_z)$ is the same. Thus, the statistical average of $E(a,b|\tilde{p}_z)$ for both initial conditions is still given by (D5).

Next we consider the initial condition described by the joint probability density functions are given by (65). This initial condition can be further decomposed such that half of the chance that the joint probability density $\tilde{p}_{y,1}(\theta_A^y,\theta_B^y) = \delta(\theta_A^y)\delta(\theta_B^y-\pi)$ and the other half chance that $\tilde{p}_{y,2}(\theta_A^y,\theta_B^y) = \delta(\theta_A^y-\pi)\delta(\theta_B^y)$. Now for the initial condition that $\tilde{p}_{y,1}(\theta_A^y,\theta_B^y) = \delta(\theta_A^y)\delta(\theta_B^y-\pi)$, Alice's measurement direction of a forms an angle with the y-axis that is given by $\alpha' = \pi/2-\alpha$. Similarly, Bob's measurement direction of b forms an angle with the y-axis that is given by $\beta' = \pi/2-\beta$. The calculation for $E(a,b|y)$ is exactly the same as calculation for $E(a,b|z)$ but with the replacement of $\alpha \to \alpha'$ and $\beta \to \beta'$. Thus,

$$E(a,b|\tilde{p}_y) = -\cos(\alpha')\cos(\beta') = -\sin(\alpha)\sin(\beta). \tag{D6}$$

Similarly, the initial condition can also be described by joint probability density $\tilde{p}_{y,2}(\theta_A^y,\theta_B^y) = \delta(\theta_A^y-\pi)\delta(\theta_B^y)$, but the calculation result for $E(a,b|\tilde{p}_y)$ is the same. Thus, the statistical average of $E(a,b|\tilde{p}_y)$ for both initial conditions is still given by (D6).

According to (D1), the overall expectation value $E(a,b)$ is obtained by adding (D5) and (D6),

$$E(a,b|\tilde{p}) = E(a,b|\tilde{p}_z) + E(a,b|\tilde{p}_y) = -\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) = -\cos(\alpha-\beta). \tag{D7}$$

Let $\gamma = \alpha-\beta$, which is the angle formed by unit vectors a and b. Thus, $E(a,b|\tilde{p}) = -\cos(\gamma) = -(a\cdot b)$ as desired.

Now consider the situation that Bob delays his spin measurement on electron B by a time period Δt. When Δt is sufficiently large, even though Alice performs measurement on A and obtains spin-up, Bob's measurement after Δt can obtain spin-down or spin-up because the initial orientation for B can be either $\theta_B^z \in \Theta_z^+$ or $\theta_B^z \in \Theta_z^-$. Thus, Bob's measurement outcome can be

$$\bar{p}(\theta_b^z|\beta) = \sin^2(\beta/2)\delta(\theta_b^z) + \cos^2(\beta/2)\delta(\theta_b^z-\pi), \tag{D8}$$

or, with equal probability,

$$\bar{p}(\theta_b^z|\beta) = \cos^2(\beta/2)\delta(\theta_b^z) + \sin^2(\beta/2)\delta(\theta_b^z-\pi). \tag{D9}$$

Therefore, the expectation value $E(a,b|\tilde{p}_z)$ can be $E(a,b|\tilde{p}_z) = -\cos(\alpha)\cos(\beta)$, or $E(a,b|\tilde{p}_z) = \cos(\alpha)\cos(\beta)$ with equal probability. This resulting the average $E(a,b|\tilde{p}_z) = 0$. Suppose the correlation between θ_A^y and θ_B^y still holds, so that (D6) is still valid. The overall expectation value becomes

$$E(a,b|\tilde{p}) = E(a,b|\tilde{p}_z) + E(a,b|\tilde{p}_y) = -\sin(\alpha)\sin(\beta). \tag{D10}$$

For the typical chosen experiment configurations $a=0$ and $a'=\pi/2$, and Bob chooses $b=\pi/4$ and $b'=3\pi/4$, one can calculate that

$$|E(a,b) - E(a,b') + E(a',b) + E(a',b')| = \sqrt{2} < 2. \tag{D11}$$

Thus, the Bell-CHSH inequality is not violated.

If the correlation between θ_A^y and θ_B^y does not hold either after the delay Δt, and $\theta_B^y \in \Theta_y^+$ or $\theta_B^y \in \Theta_y^-$ randomly, then one can calculate that $E(a,b|\tilde{p}_y) = 0$. This results in $E(a,b|\tilde{p}) = 0$ regardless of the configuration of a and b. Clearly, in such a scenario, the Bell-CHSH inequality is not violated.