LIMITS OF CLASSICAL WORLD WITH FINITE INFORMATION

SUBMITTED

BY

RAY FELLIX GANARDI

DIVISION OF PHYSICS AND APPLIED PHYSICS

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

A final year project report

presented to

Nanyang Technological University

in partial fulfillment of the

requirements for the

Bachelor of Science (Hons) in Physics

Nanyang Technological University

June 2015

ABSTRACT

Computer simulations are getting more and more common in physics. Here we examine the underlying assumption that Nature can be simulated with classical bits. We first postulate that every physical object can be encoded into a finite number of classical bits. We allow the bits to have an unknown but fixed probability distribution. The second postulate is that measurements can be computed as deterministic functions on these bits. It is shown that we can model exponentially many measurements with n bits. We also derive the minimum precision that one needs in order to disprove this model in an experiment. Finally, imposing quantum mechanical restrictions on measurement devices we show that disproving the classical models with only about 100 bits is already practically impossible.

ACKNOWLEDGMENTS

I would like to thank Assistant Professor Tomasz Paterek, for his guidance and ideas. It has been a fun year working on this project. I thank Matthew Ho Shu Hui and Ian Yap Chang Jie for numerous discussions. I would also like to thank every members in the group, especially Tran Cong Minh for being extra helpful during hours of need.

To Felicia Ganardi, for being super extra awesome. To Johanes Eldwin Pradipta, for ever and ever. To NTU Jazz and Blues and NUS Jazz Band members, for the random jam sessions and jokes. It made my fun year even more fun. To myself, for waking up in time to go to meetings. To Doraemon, for teaching me the meaning of life. To everyone that I bugged to proof read my thesis. Thank you for spending the time, I really appreciate the effort.

I would also like to thank the CN Yang Programme for the continual support during my years in NTU.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

It is known that quantum mechanics gives only statistical predictions. The early forefathers of quantum mechanics thought that the statistical nature of quantum mechanics stems from some ignorance of some *hidden variables* (HV), and ultimately the results can be explained classically [1]. One of the more famous and successful HV model is the de Broglie-Bohm theory, also known as the pilot wave theory [2, 3]. However, in 1964, Bell showed that no local hidden variable theory can explain all quantum mechanical predictions [4] if we have at least two particles.

More recently, there are attempts to study the correspondence between HV models and quantum mechanics, by creating HV models of quantum systems [5]. Hardy showed that we need *infinitely* many HV states to model infinitely many measurements on one qubit [6, 7] (see also the paper by Montina [8]). In this work, we will show how much resources we need to model one qubit with a finite number of measurements. In the limit of infinitely many measurement settings, our results agree with Hardy's [6].

We are investigating a model in which the universe is modeled with a finite number of classical bits. If it is, it should be possible to model quantum mechanical behavior with classical bits. The model that we are investigating is based on the following assumptions (see Figure 1):

1. There is a finite amount of classical information in every volume. We assume that this classical information is encoded in a finite number of classical bits r. We allow the bits to have an unknown, but fixed probability distribution $p_r \in [0, 1]$, $\sum_r p_r = 1$. Therefore all physical objects carry a finite number of classical bits.

2. Measurement outcomes are Boolean functions of bits contained in the input, i.e. $m = f(r, s)$, where s denotes the setting of our measuring device. We assume the functions are deterministic, i.e. given the same input bits and setting, the measurement should return the same output. For simplicity we will assume $m \in \{0, 1\}$.

We assume that the functions are deterministic, since if they are allowed to be non-deterministic, we can easily reproduce quantum predictions. As an example, suppose we want to explain an expectation value $\overline{m}$. If we allow the functions to be non-deterministic, then in principle there is a function $f(r)$ that returns 1 with probability $\overline{m}$ and returns

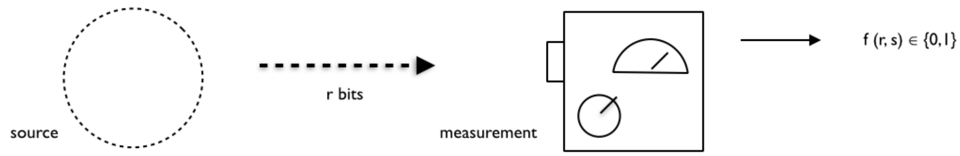

FIG. 1. An illustration of the model. Any volume in space with finite energy can be characterized by some bits r. A measurement will be some deterministic function $f(r,s)$ on the bits carrying the information contained in the physical system, where s is the setting of our measurement device.

0 with probability $1-\overline{m}$, independent of the input bit r. Thus we can recover quantum mechanical predictions quite trivially.

For simplicity, we will reduce the question of the validity of the model to the following problem.

Problem 1. *Given a set of expectation values $\overline{m}_i$, find a probability distribution of the bits $p_r \in [0,1]$ and the functions on those bits $f_i(r) \in \{0,1\}$ such that $\overline{f}_i = \sum_r p_r f_i(r) = \overline{m}_i$.*

Intuitively, this means the following. Suppose we have some setup with a source that can be characterized by n bits and a measuring device. By performing the experiment with different measurement settings i, we will measure a list of expectation values $\{\overline{m}_i\}$. Now, suppose this setup can be described by a model introduced above. Then there exists $p_r \in [0,1]$ and $f_i(r) \in \{0,1\}$ such that $\overline{f}_i(r) = \sum_r p_r f_i(r) = \overline{m}_i$. This is precisely the statement in Problem 1.

In the following section, we will explore the simplest case when our system can be described by one bit. In addition to the solution to this problem, we will also investigate another related question, what will happen if we allow uncertainties in our measurements (see Problem 2).

In view of Bell's inequality [4], we would expect local hidden variable models to be unable to model quantum mechanics if we have at least two particles. We will show that if the HV model has a finite number of states, we cannot even fully model a one qubit system. Our results show that to model one particle, the HV model has to carry infinite amount of information, e.g. using continuous variables [2, 3].

We will show that the number of explainable measurements grows exponentially with the number of bits in the model (see Theorem 2). We will also deduce that the maximum tolerable error to disprove an n bits model decreases exponentially with n (see Theorem 4). Finally, we will show that by imposing quantum mechanical restrictions, disproving classical models with about 100 bits is practically impossible (see Equation (15)).

II. PRELIMINARY DISCUSSIONS

The reason that we are interested in this class of models is because our intuition tells us that quantum systems will not admit any classical model. This is based on the well known Bell's inequality [4]. In this section we will formalize this intuition and show that indeed there is a quantum system (in this case one qubit) that does not admit a classical model described in the previous section. We will also explore related questions such as what would happen if we allow our measurements to have errors.

First we note that if we only have one measurement m on one qubit, we can always explain the expectation value with one bit by simply tweaking p_0, p_1 so that $p_1 = \overline{m}$, and assigning the function $f(r) = r$ so that $\overline{f}(r) = \sum_r f(r)\, p_r = \overline{m}$. Thus the scenario becomes interesting when we have at least two measurements.

Theorem 1. *There is no one bit model that can explain one qubit with two arbitrary measurements.*

Proof. Suppose we have a qubit in the state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Define the family of projectors

$$P_\theta = 1\,|+,\theta\rangle\,\langle +,\theta| + 0\,|-,\theta\rangle\,\langle -,\theta|\,,$$

where

$$\begin{aligned}|+,\theta\rangle &= \cos\,\theta\,|0\rangle + \sin\,\theta\,|1\rangle \\ |-,\theta\rangle &= -\sin\,\theta\,|0\rangle + \cos\,\theta\,|1\rangle\,.\end{aligned}$$

We can easily check that P_θ returns 0 if the system is in the state $|-,\theta\rangle$, and returns 1 if the system is in the state $|+,\theta\rangle$. The expectation value of P_θ is then

$$\langle P_\theta\rangle_\psi = |\alpha|^2\,\cos^2\,\theta + |\beta|^2\,\sin^2\,\theta + \cos\,\theta\,\sin\,\theta\;(\alpha^*\beta + \alpha\beta^*).$$

It is not easy to see that $\langle P_\theta\rangle_\psi \in [0,1]$. For simplicity, let us prepare our qubit in the state $|\psi\rangle = |0\rangle$. Then, $\langle P_\theta\rangle_\psi = \cos^2\,\theta$. Suppose we perform measurements $m_1 = P_{\frac{\pi}{4}}$ and $m_2 = P_{\frac{\pi}{6}}$. It is easy to check that $\langle m_1\rangle_\psi = \frac{1}{2}$, and $\langle m_2\rangle_\psi = \frac{3}{4}$.

r	$f_1(r)$	$f_2(r)$	$f_3(r)$	$f_4(r)$
0	0	0	1	1
1	0	1	0	1

TABLE I. All Boolean functions on one bit are shown.

Now suppose the qubit can be modeled by one bit with some probability distribution. Naturally, we expect that no one bit model will be able to explain the two expectation values that we measured earlier. Let us formalize this intuition.

Let the probability of the bit being in state 0 (1) to be p_0 (p_1). The four Boolean functions on one bit are listed in Table I.

Suppose we have two measurements m_1, m_2. By assigning m_1, m_2 to different functions in the Table I and calculating their expectation values, we get every possible two measurement statistics that a one bit model would explain. Plotting the expectation values for every assignment, $m_1 = f_i,\ m_2 = f_j,$ with $i, j \in \{1, 2, 3, 4\}$, we get Figure 2. Note that the point $\overline{m}_1 = \frac{1}{2}, \overline{m}_2 = \frac{3}{4}$ is *not* in the set of points explainable by the one bit classical model, whereas as shown earlier we can easily prescribe a qubit model that can explain these numbers. Since we have considered all the possible two measurements statistics a one bit model can have, we conclude that there is no one bit model that can model one qubit with two arbitrary measurements. □

We can justify this result by the following argument. Suppose we make a graph with different components of the probability vector $\vec{p}$ as the axes (see Figure 3). The requirement that probabilities have to be non-negative restricts the possible probability distributions to the positive quadrant where $p_i \geq 0$. The requirement that they have to sum up to unity puts another restriction that is equivalent to stating that the possible distributions have to lie on a plane, defined by $\vec{p} \cdot \binom{1}{1} = 1$. If we want $\vec{p}$ to explain both $\overline{m}_1$ and $\overline{m}_2$, this is equivalent to saying that $\vec{p}$ has to lie in planes defined by $\vec{p} \cdot \vec{f}_1 = \sum_r p_r f_1(r) = \overline{m}_1$ and $\vec{p} \cdot \vec{f}_2 = \sum_r p_r f_2(r) = \overline{m}_2$, where $\vec{f}_1$ ($\vec{f}_2$) is the vector representing the function f_1 (f_2). As we can see in Figure 3, there is no point that lies on the blue, red, and green line. Therefore no one bit model can explain the expectation values $\overline{m}_1 = \frac{1}{2},\ \overline{m}_2 = \frac{3}{4}$.

It is interesting to ask what is the most non-classical measurement that we can perform. We will define this as the point that is the farthest from the set of explainable points. Since a

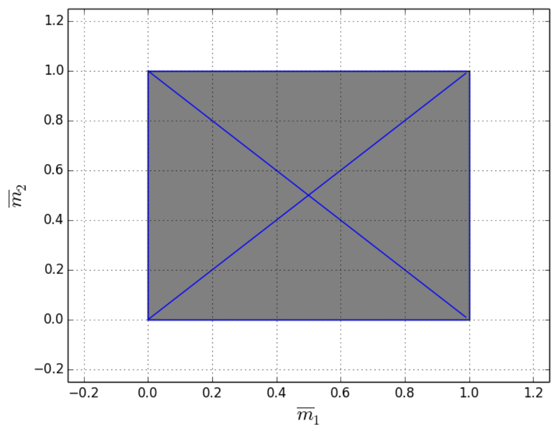

FIG. 2. All possible expectation values of two measurements explainable by a one bit model are shown in blue. The shaded area represents the possible expectation values of two measurements explainable by a one qubit model. See Problem 1.

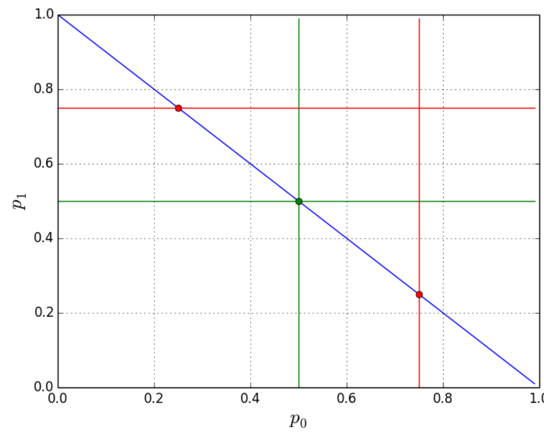

FIG. 3. The visual proof of Theorem 1 is drawn. The blue line represents the restriction arising from normalization, while the green (red) lines represent the restriction arising from the expectation value $\overline{m}_1 = \frac{1}{2}$ ($\overline{m}_2 = \frac{3}{4}$). The probability distribution that explains $\overline{m}_1$ ($\overline{m}_2$) are marked with green (red) dots.

qubit system can explain any measurement that lies inside the box $\overline{m}_1, \overline{m}_2 \in [0,1]$, the most non-classical point then is the center of the largest circle that can fit inside the structure (see Figure 4). Elementary calculations show that one of the largest circles is centered at $(\frac{1}{2}, \frac{1}{2(1+\sqrt{2})})$, and the radius of the circle is $\frac{1}{2(1+\sqrt{2})}$.

Now that we have shown that no one bit model can model one qubit with two arbitrary

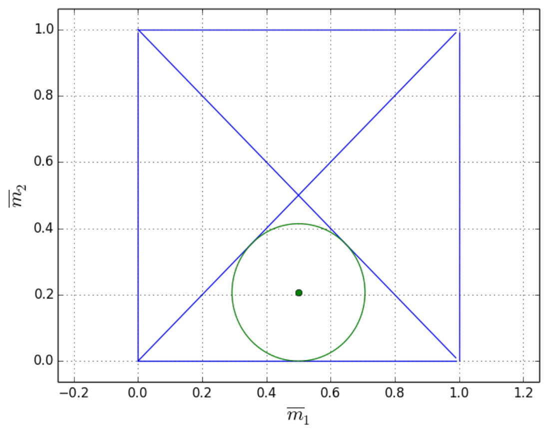

FIG. 4. The largest circle that fits inside the structure is shown. The center of the circle $\overline{m}_1 = \frac{1}{2}, \overline{m}_2 = \frac{1}{2(1+\sqrt{2})}$ is shown as a dot. All possible expectation values of two measurements explainable by one bit are shown in blue.

measurements, one might ask if introducing error will change this result. By introducing errors, we have to rephrase the problem into the following.

Problem 2. *If we allow our measurements to have an uncertainty Δm, what is the maximum Δm that we can tolerate such that we can disprove any n bit model?*

To make sense of the statement above, first we need to define circumstances under which we call the fuzzy measurement explainable by the classical model. If a single point explainable by the classical model is within one standard deviation Δm away from the expectation value $\overline{m}_i$, we say that expectation value $\overline{m}_i$ can be explained. For example, suppose we measured $\overline{m} = 0.5$ and the uncertainty is $\Delta m = 0.1$. We will say that $\overline{m}$ is explainable if there is point in the interval $[0.4, 0.6]$ that can be explained by the classical model. Thus expectation values within the error bars might be explainable by the classical model although the actual mean of the measurement cannot be explained.

Since we have drawn all the possible statistics of two measurements on one bit (see Figure 2), and recalling that a qubit system can explain any measurement that lies inside the box $\overline{m}_1, \overline{m}_2 \in [0, 1]$, solving Problem 2 is equivalent to drawing the largest box inside the structure (see Figure 5). Elementary calculations will show that the maximum error that we can tolerate is $\Delta m_{max} = \frac{1}{6}$, centered at $\overline{m}_1 = \frac{1}{2}$ and $\overline{m}_2 = \frac{1}{6}$.

It is interesting to contrast the most non-classical measurement ($\overline{m}_1 = \frac{1}{2}, \overline{m}_2 = \frac{1}{2(1+\sqrt{2})}$)

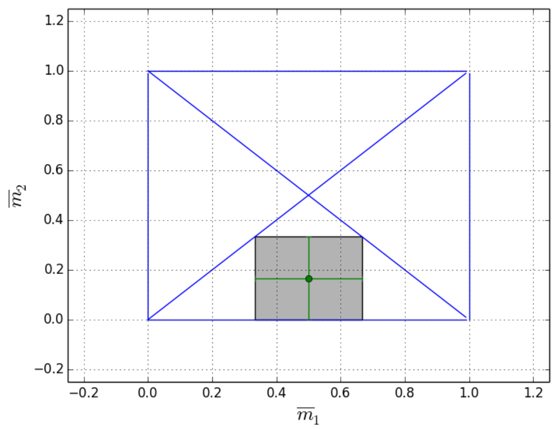

FIG. 5. The largest box that fits inside the structure is shown. All possible expectation values of two measurements explainable by one bit are shown in blue. The measurement that will disprove any one bit model with the largest allowance in uncertainty ($\overline{m}_1 = \frac{1}{2}, \overline{m}_2 = \frac{1}{6}$) is indicated by the green dot. See Problem 2.

to the measurement where we can have the largest allowable error ($\overline{m}_1 = \frac{1}{2}, \overline{m}_2 = \frac{1}{6}$). Both of these questions are equivalent to drawing some maximal structure within the boundaries of the model. However for the box we have a clear interpretation that the sides of the box represent the uncertainty of the measurement. By that, we mean the actual expectation value of the system is in the interval $[\overline{m} - \Delta m, \overline{m} + \Delta m]$ with uniform probability (see Figure 6, in green). One might suggest a similar interpretation for the circle by using a Gaussian model of error (i.e. the probability density of the actual expectation value of the system is distributed according to a Gaussian around $\overline{m}$ with standard deviation Δm, see Figure 6, in blue).

Now that we have solved Problem 1 and 2 for one bit, we can ask the effect of having more bits in our model.

III. GENERALIZATIONS

Since we know that we cannot model two arbitrary measurements with one bit, we might ask the question, what is the maximum number of arbitrary measurements that one can explain using n bits? We know that we can model one arbitrary measurement with a single bit, thus it is easy to see that we can easily model n arbitrary measurements with n bits.

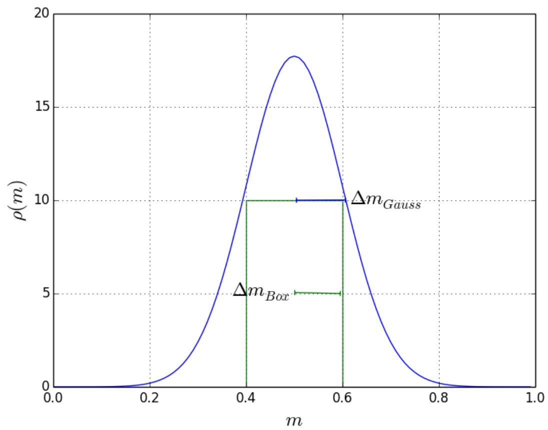

FIG. 6. An illustration of a Gaussian model of error (blue) and box model of error (green). Δm is illustrated for both model. $\rho(m)dm$ is the likelihood that the outcome is between m and $m + dm$.

r	$f_1(r)$	$f_2(r)$	$f_3(r)$	$f_4(r)$	$f_5(r)$	$f_6(r)$	$f_7(r)$	$f_8(r)$	$f_9(r)$	$f_{10}(r)$	$f_{11}(r)$	$f_{12}(r)$	$f_{13}(r)$	$f_{14}(r)$	$f_{15}(r)$	$f_{16}(r)$
00	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
01	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
10	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
11	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

TABLE II. All Boolean functions on two bits are shown.

We can explain $\overline{m}_i$ by bit i. However, parameter counting suggest that one can perform $2^n - 1$ arbitrary measurements. Thus in the following sections we will try to model $2^n - 1$ arbitrary measurements with n bits.

A. The naive way

Following the same strategy as the previous section, we can list all Boolean functions on two bits (see Table II). We know that we can always model any two measurements with two bits. Thus we need to consider three measurements. Seeing that there are $2^{2^2} = 16$ functions on two bits and $\binom{2^{2^2}}{3} = 560$ possibilities of choosing three measurements, we see that an exhaustive search is doomed to fail. Thus we need to develop some tools to tackle this problem.

Let us write down the equations that we have to solve.

$$
\begin{aligned}
p_1 + p_2 + \ldots \quad + p_4 &= 1 \\
p_1 f_1(1) + p_2 f_1(2) + \ldots + p_4 f_1(4) &= \overline{m}_1 \\
p_1 f_2(1) + p_2 f_2(2) + \ldots + p_4 f_2(4) &= \overline{m}_2 \\
p_1 f_3(1) + p_2 f_3(2) + \ldots + p_4 f_3(4) &= \overline{m}_3,
\end{aligned} \tag{1}
$$

with $f_i(r) \in \{0, 1\}$, and $p_i \in [0, 1]$. Note $f_i(r)$ in equation (1) can be different from $f_i(r)$ in Table II. Let us define

$$
M = \begin{pmatrix} 1 \\ \overline{m}_1 \\ \overline{m}_2 \\ \overline{m}_3 \end{pmatrix}, \; P = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_4 \end{pmatrix}, \text{ and } F = \begin{pmatrix} 1 & 1 & 1 & 1 \\ f_1(1) & f_1(2) & f_1(3) & f_1(4) \\ f_2(1) & f_2(2) & f_2(3) & f_2(4) \\ f_3(1) & f_3(2) & f_3(3) & f_3(4) \end{pmatrix}, \tag{2}
$$

so we can write equations (1) as $FP = M$. It is easy to see that if we assign appropriate functions $f_i(r)$ to the expectation values, we can calculate the probabilities by calculating $P = F^{-1}M$.

The astute reader will notice the flaws in this approach. There is no guarantee that F^{-1} exists, or even if it does exist, there is no guarantee that $F^{-1}M$ is a valid probability distribution (the vector $F^{-1}M$ might have a negative component). We can refute the first argument by noting that the functions in Table II form a vector space over $\mathbb{Z}_2$. By assigning linearly independent functions to $f_i(r)$, we can ensure that $det(F) \neq 0 \implies F^{-1}$ exists.

Since we want linearly independent functions as $f_i(r)$, the natural first guess for F in equation (2) would be

$$
F_{guess1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{3}
$$

This would imply that

$$
\begin{aligned}
P &= F^{-1}M \\
&= \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \overline{m}_1 \\ \overline{m}_2 \\ \overline{m}_3 \end{pmatrix}.
\end{aligned}
$$

However this will not explain any three expectation values since setting $\overline{m}_1 = \overline{m}_2 = \overline{m}_3 = 0$ will imply that $P = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$, which is not a valid probability distribution. Thus we can only do the assignment in equation (3) if and only if

$$-1 + \overline{m}_1 + \overline{m}_2 + \overline{m}_3 \geq 0 \tag{4a}$$

$$1 - \overline{m}_1 + \overline{m}_2 - \overline{m}_3 \geq 0 \tag{4b}$$

$$1 + \overline{m}_1 - \overline{m}_2 - \overline{m}_3 \geq 0 \tag{4c}$$

$$1 - \overline{m}_1 - \overline{m}_2 + \overline{m}_3 \geq 0. \tag{4d}$$

Since our first guess has proved to be limited, we might ask is there any assignment of functions $f_i(r)$ that will satisfy our requirements? Let us try another assignment

$$F_{guess2} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

This would imply that

$$\begin{aligned} P &= F^{-1}M \\ &= \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \overline{m}_1 \\ \overline{m}_2 \\ \overline{m}_3 \end{pmatrix}. \end{aligned}$$

Again we arrive at the same problem. This time, setting $\overline{m}_1 = \overline{m}_2 = \overline{m}_3 = 1$ will make $P = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. The inequalities that we have to satisfy for the assignment in equation (5) are

$$1 - \overline{m}_1 - \overline{m}_2 - \overline{m}_3 \geq 0 \tag{6a}$$

$$\overline{m}_1 \geq 0 \tag{6b}$$

$$\overline{m}_2 \geq 0 \tag{6c}$$

$$\overline{m}_3 \geq 0. \tag{6d}$$

Although three out of the four inequalities are trivial, the first one is not. However, note that the inequalities implied by F_{guess1} and F_{guess2} are complementary, i.e. if we violate one we have to satisfy the other. What is also interesting is the inequalities (4) and (6) are blind to index-swapping, i.e. $\overline{m}_1 \leftrightarrow \overline{m}_2$. This is what we expect since Nature should not care about our labeling. Thus we can always model three arbitrary measurements with two bits.

Let us try to generalize these assignments for an arbitrary number of bits. If we denote the matrix F_{guess1} as H_{2^2}, we can enumerate the elements of the sequence $\{H_{2^k}\}$ by

$$H_{2^k} = \begin{pmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & 1 \oplus H_{2^{k-1}} \end{pmatrix},$$

where $(1 \oplus H_{2^k})_{ij} = 1 \oplus (H_{2^k})_{ij}$, and $\oplus$ is addition mod 2. This seems to tell us that we can indeed model $2^n - 1$ measurements.

Suppose we have three bits, and seven measurements. The inequalities implied by H_{2^3} are

$$-3 + \overline{m}_1 + \overline{m}_2 + \overline{m}_3 + \overline{m}_4 + \overline{m}_5 + \overline{m}_6 + \overline{m}_7 \geq 0 \tag{7a}$$

$$1 - \overline{m}_1 + \overline{m}_2 - \overline{m}_3 + \overline{m}_4 - \overline{m}_5 + \overline{m}_6 - \overline{m}_7 \geq 0 \tag{7b}$$

$$1 + \overline{m}_1 - \overline{m}_2 - \overline{m}_3 + \overline{m}_4 + \overline{m}_5 - \overline{m}_6 - \overline{m}_7 \geq 0 \tag{7c}$$

$$1 - \overline{m}_1 - \overline{m}_2 + \overline{m}_3 + \overline{m}_4 - \overline{m}_5 - \overline{m}_6 + \overline{m}_7 \geq 0 \tag{7d}$$

$$1 + \overline{m}_1 + \overline{m}_2 + \overline{m}_3 - \overline{m}_4 - \overline{m}_5 - \overline{m}_6 - \overline{m}_7 \geq 0 \tag{7e}$$

$$1 - \overline{m}_1 + \overline{m}_2 - \overline{m}_3 - \overline{m}_4 + \overline{m}_5 - \overline{m}_6 + \overline{m}_7 \geq 0 \tag{7f}$$

$$1 + \overline{m}_1 - \overline{m}_2 - \overline{m}_3 - \overline{m}_4 - \overline{m}_5 + \overline{m}_6 + \overline{m}_7 \geq 0 \tag{7g}$$

$$1 - \overline{m}_1 - \overline{m}_2 + \overline{m}_3 - \overline{m}_4 + \overline{m}_5 + \overline{m}_6 - \overline{m}_7 \geq 0 \tag{7h}$$

Note that the inequalities are not invariant under index-swapping. You can check that $\{\overline{m}_1 = 0.8, \overline{m}_2 = 0.7, \overline{m}_3 = 0.9, \overline{m}_4 = 0.1, \overline{m}_5 = 0.3, \overline{m}_6 = 0.4, \overline{m}_7 = 0.2\}$ satisfies the inequalities (7), but $1 - \overline{m}_1 - \overline{m}_2 - \overline{m}_3 + \overline{m}_4 + \overline{m}_5 + \overline{m}_6 - \overline{m}_7 = -0.8 < 0$. This is contrary to what we expect.

Let us see what happens with the generalization of F_{guess2}. The generalization would be

$$F_{2^n} = \begin{pmatrix} 1 & 1 \\ 0 & \mathbb{1}_{2^n-1} \end{pmatrix}$$

Again, supposing that we have three bits and seven measurements, we get the inequalities

$$1 - \overline{m}_1 - \overline{m}_2 - \overline{m}_3 - \overline{m}_4 - \overline{m}_5 - \overline{m}_6 - \overline{m}_7 \geq 0 \tag{8a}$$

$$\overline{m}_1 \geq 0 \tag{8b}$$

$$\overline{m}_2 \geq 0 \tag{8c}$$

$$\overline{m}_3 \geq 0 \tag{8d}$$

$$\overline{m}_4 \geq 0 \tag{8e}$$

$$\overline{m}_5 \geq 0 \tag{8f}$$

$$\overline{m}_6 \geq 0 \tag{8g}$$

$$\overline{m}_7 \geq 0 \tag{8h}$$

Here we note that although seven out of the eight inequalities are trivially satisfied, inequality (8a) is non-trivial. Also, there is no complementary relation between inequality (7a) and inequality (8a), although satisfying inequality (8a) will imply inequalities (7b) - (7h).

Thus the generalizations of assignments in equations (3) and (5) does not seem to be useful. This motivates us to do a numerical search of 2^3x2^3 matrices that will explain seven arbitrary measurements. However to do an exhaustive search we need to check $\binom{2^{2^3}}{7} \approx 10^{13}$ matrices. This is well beyond the computational power of home computers.

Note that the requirement of invariability significantly reduces our search space of F. To take advantage of the reduction, we need to formalize the notion. We say that F is blind to index-swapping if and only if for all permutation matrices $\Pi \in P_{n-1}$, we require $\Pi F = F\Pi'$ for some $\Pi' \in P_{n-1}$, where P_{n-1} denotes the set of all permutation matrices of n elements with the first element fixed. In simple terms, it means that permuting $\{\overline{m_i}\}$ corresponds to some permutation of columns in F, and thus, a permutation of the inequalities. We only consider the permutations P_{n-1} since we do not want to swap out the first row of F or M.

We have searched the $\binom{16}{3}$ 4x4 matrices and found 12 matrices that commute with the permutation group. However even with this reduction in search space, it will still take weeks to look for 8x8 matrices that commutes with permutation group. Thus it appears that this approach is not scaling well. Nevertheless we have proved that we can indeed model three arbitrary measurements with two bits.

B. The smart way

As mentioned before, parameter counting suggests that we can model $2^n - 1$ arbitrary measurements. However the naive way is not effective due to the sheer size of the search space. In this section we present an explicit construction that proves that n bits can model $2^n - 1$ arbitrary measurements.

Theorem 2. *A model with n bits can explain $2^n - 1$ arbitrary expectation values.*

Proof. 1. Order $\overline{m}_i$ so that $\overline{m}_1 \leq \overline{m}_2 \leq \ldots \leq \overline{m}_{2^n-1}$. We can always do this since Nature does not care about our labeling and thus exchanging labels are allowed.

2. Let

$$\begin{aligned} p_1 &= \overline{m}_1 \\ p_2 &= \overline{m}_2 - \overline{m}_1 \\ p_3 &= \overline{m}_3 - \overline{m}_2 \\ &\vdots \\ p_{2^n-1} &= \overline{m}_{2^n-1} - \overline{m}_{2^n-2} \\ p_{2^n} &= 1 - \overline{m}_{2^n-1} \end{aligned} \tag{9}$$

and

$$f_i(r) = \begin{cases} 1, \text{if } r \leq i \\ 0, \text{otherwise} \end{cases} \tag{10}$$

3. Then

$$\begin{aligned} \overline{f}_i &= \sum_r f_i(r)\; p_r \\ &= \overline{m}_1 + \overline{m}_2 - \overline{m}_1 + \ldots + \overline{m}_i - \overline{m}_{i-1} \\ &= \overline{m}_i \end{aligned}$$

□

Remark 1. Seeing that we get the result we expected from parameter counting, it is easy to discount this result as trivial. However as seen in the previous section, it is nontrivial that there is an assignment of functions and probabilities that will explain $2^n - 1$ measurements. We also note that this solution is blind to renaming of variables, i.e. $\overline{m}_i \leftrightarrow \overline{m}_j$.

It is easy to check that the assignments in equations (9) and (10) satisfies our requirements that $p_i \geq 0$ and $f_i(r) \in \{0, 1\}$.

C. The error bars

Now that we proved we can model any $2^n - 1$ measurements, we can direct our attention to Problem 2 again. Since any $2^n - 1$ measurements can be modeled by n bits, we have to consider cases where we have 2^n measurements on n bits. However, we cannot draw a picture like Figure 2. Thus it is hard to derive a tight bound for tolerable errors. However even with the results that we have we can derive a reasonable upper bound for the error.

First, we need to define what do we mean by independent measurements. We say that m_i and m_j are independent if $\overline{m}_i$ and $\overline{m}_j$ are distinct. We say $\overline{m}_i$ and $\overline{m}_j$ are distinct if $[\overline{m}_i - \Delta m, \overline{m}_i + \Delta m]$ and $[\overline{m}_j - \Delta m, \overline{m}_j + \Delta m]$ are completely disjoint. In other words, we cannot ascribe one function to explain *both* $\overline{m}_i$ and $\overline{m}_j$. Equiped with these definitions, we can prove the following statement.

Lemma 3. *To distinguish between l measurements, we need $\Delta m = O((2l)^{-1})$.*

Proof.

1. If $\overline{m}_i$ and $\overline{m}_j$ are distinguishable, then for $i \neq j$, $[\overline{m}_i - \Delta m, \overline{m}_i + \Delta m]$ and $[\overline{m}_j - \Delta m, \overline{m}_j + \Delta m]$ are completely disjoint.

2. It follows that Δm is maximized when $\{\overline{m}_i\}$ are equally spaced (see Figure 7)

3. If $\overline{m}_i$ are equally spaced, then $\Delta m = (2l)^{-1}$. Since this is the upper bound, we can say that $\Delta m = O((2l)^{-1})$.

□

This naturally implies the following Theorem.

Theorem 4. *To disprove an n bit model, we need $\Delta m = O(2^{-(n+1)})$.*

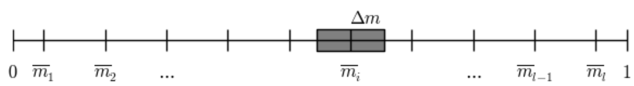

FIG. 7. It is shown that Δm will be maximized when $\{\overline{m}_i\}$ are equally spaced.

Proof. 1. By Theorem 2, an n bit model can explain any $2^n - 1$ independent measurements. Thus to disprove an n bit model we need to perform at least 2^n independent measurements.

2. By Lemma 3, we need $\Delta m = O(2^{-(n+1)})$ to distinguish 2^n measurements.

□

We see that Theorem 4 agrees with the bound that we derived earlier, i.e. $\Delta m_{max}(1 \text{ bit}) = \frac{1}{6} \leq \frac{1}{4}$.

D. The implications

It is easy to imagine a physical scenario where Theorem 4 along with the uncertainty principle implies that we will *not* be able to disprove a model with $n \geq n_{max}$ bits. A rough estimation of n_{max} follows. First we need to calculate Δm for a generic setup. We will follow the derivation by Kofler et. al [9], but using polarizers instead of magnets.

Suppose we have a source that emits photons . It follows that the amount of information that the photon carries cannot be greater than the number of bits that characterizes the source. Furthermore, assume that the classical bits are encoded in the polarization of the photons. Suppose we want to set the polarizer at some angle θ with some uncertainty $\Delta\theta$. We know that for wave functions that are 2π-periodic [10],

$$[\theta, L] = i\hbar. \tag{11}$$

Since we are using the polarizer as a measuring device, we will assume that its wave function is sharply peaked at θ, and thus there is no concern of periodicity. Now, assume that the photon starts interacting with the polarizer at $t = 0$ and stops interacting at $t = \tau$. The Hamiltonian of a freely rotating polarizer is given by $H = \frac{L^2}{2I}$, where L is the angular momentum of the polarizer, $I \approx MR^2$ is the moment inertia of the polarizer, M is the mass of the polarizer, and R is the radius of the polarizer. In the Heisenberg picture, an observable will evolve as $\frac{d\theta}{dt} = -i\frac{[\theta,H]}{\hbar} = \frac{L}{I}$. Hence,

$$\theta(\tau) = \theta(0) + \frac{L\tau}{I}. \tag{12}$$

Recall the generalized uncertainty principle,

$$\Delta A \Delta B \geq \tfrac{1}{2}|\langle[A, B]\rangle|, \tag{13}$$

where A and B are arbitrary operators. Combining equations (11), (12), and (13), we get

$$\Delta\theta(0)\Delta\theta(\tau) \geq \frac{\hbar\tau}{MR^2}.$$

It follows that either $\Delta\theta(0)$ or $\Delta\theta(\tau)$ obeys

$$\Delta\theta \geq \sqrt{\frac{\hbar\tau}{MR^2}}.$$

If we assume that the interaction time $\tau \geq R/c$ (we require that an observer at the rim of the polarizer knows that the photon is passing through), we get

$$\Delta\theta \geq \sqrt{\frac{\hbar}{cMR}} \tag{14}$$

Equation (14) along with Theorem 4 implies that

$$n_{max} = \tfrac{1}{2}log_2\frac{cMR}{\hbar} - 1 \tag{15}$$

Substituting nonsensical numbers ($M =$ total mass of the universe, $R =$ diameter of the universe) into equation (15), we get $n_{max} \approx 60$. For comparison, the Bekenstein bound (the limit of information that a volume can carry, see [11]) for a hydrogen atom is on the order of 10^6 bits. If we take a proton, the Bekenstein bound is on the order of 44 bits [12]. Thus in a practical setting, there is little hope of measuring a violation of this model.

Remark 2. It is important to note that this model investigates what happens after a measurement. We merely investigate if there is any fundamental restriction on these models. We found that these models cannot model quantum mechanics, however with the presence of uncertainty, there are effectively no such restriction. We are *not* claiming that such a model exists.

IV. CONCLUSION AND FUTURE WORKS

We asked if it is possible to model the universe with classical bits, i. e. assuming that every physical object carries a finite number of classical bits and measurement outcomes can be calculated as functions of those bits. We have answered this question in the negative by showing that quantum predictions for a single spin-1/2 system are not compatible with such classical models. However, we also showed that limitations on measurement precision imposed by the very quantum mechanics make it practically impossible to disprove the classical model operating even on small number of bits. Since typically the number of bits in a physical object scales with its size we conclude that the bigger the system gets the harder it is to observe the difference between the quantum and the classical models. This supports similar observations made by Kofler et. al. [9].

We have left open the question of deriving the exact bound for maximum error allowed to disprove the classical model with n bits since even with our bounds we can deduce sensible predictions. We are also still working on concrete example of quantum expectation values that allow one to disprove the n-bit classical model. Present results demonstrate that there must exist such an example but its closed form would definitely complete the picture.

Finally, it would be very interesting to find a physical situation where the models we discussed would naturally emerge. We were thinking that such a possibility can occur in relation to the Bekenstein bound [11]. This bound gives the maximal number of bits that can be encoded in a volume. We would like to take a closer look at this bound and check if it indeed discusses classical bits. If yes, then any object emerging from a volume (say a photon from hydrogen atom) cannot carry more bits that the bound and we could verify it experimentally by performing sufficiently many precise measurements on the photons. We have done preliminary calculations for various systems but all of them turn out to be outside the reach of experiments.

---

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[2] D. Bohm, Phys. Rev. **85**, 166 (1952).

[3] D. Bohm, Phys. Rev. **85**, 180 (1952).

[4] J. S. Bell, Physics **1**, 195 (1964).

[5] B. Dakić, M. Šuvakov, T. Paterek, and i. c. v. Brukner, Phys. Rev. Lett. **101**, 190402 (2008).

[6] L. Hardy, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **35**, 267 (2004).

[7] E. F. Galvão and L. Hardy, Phys. Rev. Lett. **90**, 087902 (2003).

[8] A. Montina, Phys. Rev. A **83**, 032107 (2011).

[9] J. Kofler and C. Brukner, ArXiv e-prints (2010), arXiv:1009.2654 [quant-ph].

[10] A. Peres, *Quantum Theory: Concepts and Methods*, Fundamental Theories of Physics (Springer, 1995).

[11] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).

[12] F. J. Tipler, (2001), arXiv:astro-ph/0111520 [astro-ph].