DISCRETE QUANTUM THEORIES AND COMPUTING

Tai, Yu-Tsung

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in the Department of Mathematics

and the Department of Computer Science,

Indiana University

February 2019

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Amr A. Sabry, PhD

Dylan Paul Thurston, PhD

Gerardo Ortiz, PhD

Andrew J. Hanson, PhD

Shouhong Wang, PhD

November 29, 2018

To my parents, Cheng-Tien Tai (戴振泏) and Feng-Ming Chang (張鳳鳴).

ACKNOWLEDGMENTS

I want to thanks all my professors co-authored with me, suggested me as a committee members, and taught me so that our work or studies became the bases or inspired any part of my dissertation as listed in Table 1. Especially, Prof. Andrew J. Hanson, Prof. Gerardo Ortiz, and Prof. Amr A. Sabry (عمرو صبري) meet me weekly for the pass few years, read and tried to understand what I typed which sometimes even myself cannot understand, taught me how to organize them in "English" so that general audience might be interested and have some chance to understand, and encouraged me to participate seminars, workshops, and conferences. Prof. Lawrence Moss invited me to present our results in the interdisciplinary logic seminar and theory seminar in Computer Science and wrote an assessment letter for me. Although not directly related to this dissertation, the Weil conjecture on Grassmannians studied with Prof. Dylan Paul Thurston and the quantum interpretation discussed with Prof. Shouhong Wang (汪守宏) also broadened my understanding on discrete and conventional quantum theories. I would like express my appreciation to all of these.

To have the current the content of the dissertation, I would also like to thank John Gardiner for some inspiring discussion about quantum probability measures over finite fields [8]. His results were improved and motivated the transition from DQT to QIVPM in Sec. 4.7. Also, the unpublished computation for the discrete Deutsch algorithm [4, 9, 10] is organized in the term paper for SLST-T501 Academic Writing course. I want to thank Prof. Tom Lewis for proof-reading them, and they become the bases of Sec. 4.4.1. Besides of Prof. Tom Lewis, my other English teachers, especially Traci Nagle, Elizabeth "Betsy" Merceron, and Kexin "Casey" Chen (谌可心), taught me a lot on how to communicate in English as well. Others who help me indirectly on the content of the dissertation

Table 1: I want to thank the professors co-authored with me, suggested me as a committee member, and taught me, and the collaboration or supervised works or studies become the basis of or inspire this dissertation.

Committee or course	Gerardo Ortiz	Amr A. Sabry (عمرو صبري)	Andrew J. Hanson	Dylan Paul Thurston	Shouhong Wang (汪守宏)	Lawrence S. Moss	Relation with the dissertation
	Co-author	Co-author	Co-author				Most of Chapters 2, 3, and 4 is based on [1] and [2]; most of Chapter 5 is based on [3].
Research committee	Co-chair	Co-chair	Member	Co-chair	Member		Some questions in the dissertation proposal are answered in Chapter 5, and some of the remaining ones are explained in Chapter 6.
		Supervisor					The application for "Rethinking Foundations of Physics" 2017 workshop inspires Chapter 1.
Advisory committee	Member	Member	Member	Member			The survey of real computation in the Computer Science qualifying exam (chaired by Prof. Sabry) inspires Chapter 1.
PHYS-P700	Instructor						The textbook [4] is cited in Chapters 1, 2, 4, and 5.
MATH-M781						Instructor	The presented paper [5] is cited in Sec. 2.2.1.
Tier 3 committee	Co-chair	Co-chair		Member		Co-chair	Two of the assigned papers in the Tier 3 exam, [6] and [7], are cited in Sec. 3.2.

including Kin Wai Chan, Prof. Huah Chu (朱樺), Weihua Liu (刘伟华), Caren Liu (劉鈺平), Meng-Wei Chen (陳孟瑋), Hao-Chun Lee (李浩君), Chao-Hong Chen (陳昭宏)…, etc. This could be listed on and on. I still thank you even if your name is not listed here, and sorry for my bad memory.

Many people in both the departments of Mathematics and Computer Science have done everything to make my double-major smooth. **TODO. Acknowledgments cleaned up until here.** I need to thank all of them, especially directors of graduate study in Mathematics, especially Prof. Christopher Judge and Prof. Matthias Weber; directors of Ph.D. study in Computer Science, especially Prof. Yuqing Melanie Wu (吴愈青) and Prof. Funda Ergun.

Thank all my friends for your support, including:

- Meng-Wei Chen (陳孟瑋) as my roommate, AT&T family plan member, and emergency contact for all kinds of emergency and non-emergency situations.
- My other AT&T family plan members, especially Kexin "Casey" Chen (谌可心), Yao-yu Chih (池耀宇), and Xinyi "Diana" Gong (龚欣怡), for you always pay our telephone bill on time so that……**TODO. Maybe I don't have to thank them as AT&T family plan members, but I should thank some of them for support and sharing information…**
- My other roommates,
- Hao-Chun Lee (李浩君) and Hsien-Ching Kao (高憲慶) for discussing what I did from the perspective of a general PhD. in Physics and working later in more computational industry.

Finally, I thank Jin-Ru Yang (楊謹如) for everything that could not be written here.

PREFACE

- The latest PDF version of this thesis is in GitHub: `https://git.io/fxbuG`.
- Its source code is typeset using LyX [11] in `https://git.io/fxbuE` and the exported LaTeX file is in `https://git.io/fhZBW`.
- Any comments can be left in the Issue part of the GitHub repository: `https://git.io/fxbug`.

Following explains how to compile the PDF from the LaTeX source:

1. Clone the git repository of this thesis `https://git.io/fhZP5` into your local directory and the thesis document class `https://github.com/yuttai/iuphd` in a location where the system can find.
2. In the local directory of this thesis, run `xelatex dissertation.tex`.
3. Run `biber dissertation`.
4. Run `xelatex dissertation.tex` several times until all the references are resolved.

The previous steps only tested in Microsoft Windows with MiKTeX 2.9. Please report me if there is any difficulty to compile it. Besides, to compile a PDF file from the LyX source, just finish step 1, open `dissertation.lyx` in LyX, and click the toolbar button  .

Tai, Yu-Tsung

DISCRETE QUANTUM THEORIES AND COMPUTING

Most quantum computing models are based on the continuum of real numbers, while classical digital computers faithfully realize only discrete computational models. Analog computers appear to be an option, but in reality are far weaker than would be needed for computational models requiring real numbers. One approach to resolving this conflict is to find consistent mathematical ways to limit measurement precision to computable contexts that do not require incomputable real numbers. Our goal is to build more philosophically consistent models by investigating discrete quantum computing using finite number systems, and, alternatively, by incorporating finite precision measurement using intervals into quantum theory.

We begin by replacing the continuum of complex numbers by discrete finite fields in quantum theory. The simplest theory, defined over unrestricted finite fields, is so weak that it cannot express Deutsch's algorithm, but, paradoxically, is also so powerful that it can be used to solve the UNIQUE-SAT problem, which is as hard as a general NP-complete problem.

Our second framework employs only finite fields over prime numbers of the form $4\ell + 3$, which possess no solutions to $x^2 + 1 = 0$, and thus permit an elegant complex representation of the finite field by adjoining $\mathrm{i} = \sqrt{-1}$. Because the states of a discrete n-qubit system are in principle enumerable, we can count the number of states, and determine the proportions of entangled and unentangled states. Depending on how we model the measurement process, this improved framework can be used to implement deterministic Deutsch's algorithm and the probabilistic Grover search algorithm in a

local region, but we still haven't found a consistent way to treat quantum probability measures in general.

Finally, we shift our attention to consider quantum interval-valued probability measures (IVPMs), which potentially embody both finite precision measurement and a sensible correspondence to standard quantum probability. This interval-valued framework not only provides a natural generalization of both classical IVPMs and conventional quantum probability measures but also allows us to establish bounds on the validity of the Kochen-Specker and Gleason theorems in realistic experimental environments.

Amr A. Sabry, PhD

Dylan Paul Thurston, PhD

Gerardo Ortiz, PhD

Andrew J. Hanson, PhD

Shouhong Wang, PhD

CONTENTS

Chapter 1

INTRODUCTION

Since no human being can distinguish quantities differing only by an arbitrarily small extent [12, 13], we want to incorporate this limitation into our quantum mechanical model to better capture our ability to predict quantum phenomenon and the power of realistic quantum computers. Since we are agnostic about whether the reality is ultimately discrete or continuous, we consider two types of quantum models. By assuming the quantum states are ultimately discrete and distinguishable, our first type of quantum theories replaces the complex numbers by finite fields. We then focus on the limitation of distinguishability itself, incorporate the finite precision measurement into our model, and consider quantum interval-valued probability measures.

After first suggested by Richard Feynman 40 years ago [14], IBM, Rigetti, Google, Microsoft, and many other companies and countries start to build realistic quantum computers recently [15, 16] and seek more efficient ways to simulate chemical molecules and to break down RSA [17, 18]. No matter how to, classical or quantum, attack these two tasks, they require two different types of quantities. On one hand, a question and its answer of RSA are both discrete integers. On the other hand, molecules are described by quantities including the domain of their position wave functions and the probability amplitudes of these wave functions, which look like continuous at first sight. However, either these quantities might later be discovered to be discrete, or their exact values may never be pinned down. Hence, modeling quantum computing based on the exact values of continuous quantities cannot really describe the computational power of a realistic quantum computer. To better

understand this issue without struggling with the quantum theory, it is instructive to review how classical quantities are characterized and computed.

Discrete quantities can be easily characterized by integers [19]. Even if the proportion of integers are fractions, the operations among fractions are essentially the same as those among integers over the common denominator. In general, discrete computation can be reliably repeated with exact equivalence, and any digital computer can faithfully simulate a Turing machine limited only by the available memory and time [12].

In contrast, the non-apparently discrete quantities confused many cultures since the very beginning. For example, the ancient Chinese philosopher Chuang-tzu (庄子) didn't believe these quantities have a minimum unit and expressed "If a one-foot-long stick is cut into halves every day, the cutting will never come to an end"[1]. As another example, Greek mathematicians in the school of Pythagoras believed every mathematical model must be ultimately discrete, and they in legend murdered Hippasus because he proved there is no minimum unit between a side of a square and its diagonal. Comparing the idea that the length of a square's diagonal is not a physical quantity or there is no square physically, working with a mathematical model without minimum unit seems more appealing. Since the mathematical models are continuous, their corresponding physical quantities are naturally assumed to be continuous. Later, these continuous quantities become the basic component of analog computers. After logarithm was discovered, a slide rule, also known as a slipstick, is used to compute multiplication, division, and more complex operations as a mechanical analog computer [22]. Charge amplifiers are used to build electrical analog computers and compute the integration in calculus [23, 24].

Although it is convenient to assume physical quantities is as continuous as their mathematical models, each physical quantity discussed in the previous paragraph could not have a simple one-to-one correspondence with a real number. Chuang-tzu's stick is made by molecules and cannot be cut

[1]"一尺之棰，日取其半，萬世不竭。" in Chinese [20, 21].

in halves endlessly. The input charge of an integrator must be an integer multiple of the elementary charge. A perfect square cannot be drawn physically. Even we don't draw it, but merely ask whether it exists in the space. This question is still beyond our ability to answer because whether space itself is ultimately discrete or continuous is still an unsolved question. Even if a perfect square exists physically, we may not able to decide whether the "square" we discussed is really perfect or almost perfect differing only by an arbitrarily small extent. Similar situation applying to the length of a slide rule, it is beyond our ability to decide whether its length is ultimately discrete or continuous. Even if its length is continuous, identifying which real number represents its length requires infinite precision, but the precision in reality is limited generally three or four significant figures [25]. As the results, even if analog computers really store real numbers, they cannot be precisely read, written, and used to branch the computation. The opposite of the last statement is sometimes assumed by real computational models. For example, Blum, Cucker, Shub, and Smale modeled their BCSS machine to allow the operations of deciding whether a number is greater or equal to zero over real numbers and whether a number is exactly zero over complex numbers [26–28]. Hence, when the quantities used to branch the computation closes to zero, the branch chosen by a BCSS machine cannot reliably predict the branch chosen by a realistic analog computer. Since two branches of computations may not have any relations, the follow-up computational paths may be significantly different, and this difference is unlikely to be compensated by error analysis techniques easily like the butterfly effect. Therefore, we might not be able to utilize a realistic analog computer to acquire the computational power predicted by this kind of theoretical models.

The situations for classical quantities and the computational power above them are summarized in Table 1.1, and two possibilities for classical quantities inspires two types of quantum models. Like its classical counterparts, quantum circuit model, the most widely used model of quantum computing, manipulates quantities which are not discrete apparently: the probability amplitudes of quantum states, which are assumed in the field of complex numbers in the "conventional quantum

Table 1.1: Comparison between the classical quantities and their mathematical models with the inspired quantum models.

Physical quantity	Discrete	Discrete with extreme small units	Continuous, or no evidence to support they are discrete
Computational device	Digital computer	Analog computer	Analog computer
Mathematical representation	Integer	Real number	Real number
Computational model	Turing machine	BCSS machine	BCSS machine
How does the theorical model predict the behavior of the physical device?	Reliably	Not reliably, but after we have better technology to manipulate the minimum units, we can more reliably predict the physical phenomenon and their computational power by a better discrete model.	Not reliably, because precision can never be high enough, and the difference between their computational power might not be able compensated by error analysis techniques.
Inspired quantum model		Quantum theories and computing over finite fields	Quantum interval-valued probability measures (QIVPM)

theory" (CQT)[2]. If the probability amplitudes are assumed to be ultimately discrete, we choose to replace the field of complex numbers by discrete finite fields for two reasons. On one hand, CQT is built upon the operations among the probability amplitudes, and the probability amplitudes have the required operations since the complex numbers form a field. On the other hand, based on finite fields, we will define the fraction-like cardinal probability which has extremely small units when the size of the field is extremely large. Alternatively, even if the probability amplitudes might never be found quantized, considering the precision of quantities could still result in a better model. However, comparing to the precision of probability amplitudes, it is easier to consider the precision of their inducing probabilities because the idea of "imprecise probability" has well-studied classically [29–34]. This idea will be extended to quantum domain as quantum interval-valued probability measure (QIVPM).

The geometrical structure of states in CQT is first reviewed in Sec. 2.1, where a pure state is represented as a vector in a Hilbert space, and an irreducible state is actually a point in a projective Hilbert space [35, 36] also known as a complex projective space [37, 38]. In Sec. 2.2, we will construct a quantum probability space based on classical probability spaces, apply Gleason's theorem [39–43] to recover the Born rule [4, 36, 44], and define the expectation value of an observable. The expectation value will then be used in Sec. 2.3 to define purity which provides an easy entanglement test.

In Chapter 3, we examine previously-introduced quantum theories defined over unrestricted finite fields [2, 6, 45]. It is called the modal quantum theory because it can only predict whether a measurement result is possible or impossible, but not its probability. Although a version of quantum theory defined over the two-valued field can express simple algorithms such as quantum teleportation [4, 6, 36, 41, 46], it is so weak that it cannot express Deutsch's algorithm [2, 4, 9]. This quantum theory is, however, also so powerful that it can be used to efficiently solve the UNIQUE-SAT prob-

[2]Alternative terminology in the literature includes "actual," "standard," and "ordinary" quantum theory.

lem [47], which is as hard as a general NP-complete problem [48].

In Sec. 4.1 to 4.4, we improve on this by showing that for finite fields of order p^2, $\mathbb{F}_{p^2}$, with the prime p of the form $4\ell + 3$ (ℓ a non-negative integer), the complex numbers then have extremely compelling and natural discrete analogs that preserve a great many of the standard requirements of quantum computing. In this model, the state of a discrete D-dimensional system is a vector in the D-dimensional vector space over $\mathbb{F}_{p^2}$ and can be reduced to a point in the discrete complex projective space. Among these irreducible states, the product states and entangled states could not only be identified as in CQT but also be counted for different p. Despite the similarity between this model and CQT, it still only predicts whether a measurement result is possible or impossible because it is not possible to define an inner product in the usual sense due to the modular arithmetic [49]. Under suitable conditions, we can have deterministic quantum algorithms such as the algorithms of Deutsch, Simon [4, 36, 50], and Bernstein-Vazirani [4, 51], but this still leads to excessive computational power for partially solving the UNIQUE-SAT problem efficiently.

Since many important quantum algorithms are probabilistic, we want to find a notion of probability for quantum theories over finite fields. The cardinal probability defined in Sec. 4.5 restricts states in some local regions of the vector spaces, where a notion of inner product and probability could be recovered. In Sec. 4.6, this discrete quantum theory is applied to two representative algorithms, the deterministic Deutsch-Jozsa algorithm [36, 52] and the probabilistic Grover algorithm [4, 36, 53]. Despite the great success of cardinal probabilities, it is difficult to define their arithmetic operations and expectation values. Since they could be defined easily with real-valued probabilities in CQT, we attempt to consider real-valued quantum probability measures over finite fields (QPMFF) in Sec. 4.7. In contrast to CQT, our Gleason-like conditions over finite fields cannot deduce a discrete Born rule in general. This impossible result suggests that the discrete and finite state spaces might not be compatible with infinitely precise real-valued probability. Thus, no matter whether the probability amplitudes are ultimately discrete or continuous, we should study imprecise quantum

probabilities formulated by quantum interval-valued probability measures.

Sec. 5.1 starts from reviewing how to formulate classical "imprecise probabilities" [29–34] as interval-valued probability measures (IVPMs). Given a classical IVPM, its core is the set of the real-valued probability measures consistent with it [30, 34, 54, 55], and the expectation value with respect to it is just classical Choquet integral [30, 34, 56]. These classical properties can naturally be extended for QIVPMs when we extend classical IVPMs to QIVPMs. In Sec. 5.2, we utilize the expectation values while proving the finite precision variants of the Kochen-Specker theorem [36, 40, 41, 57, 58], and establish bounds on its validity in realistic experimental environments. In Sec. 5.3, we are able to show that, while a QIVPM incorporating the effects of finite precision might not be consistent with Gleason's unique state ρ on all projectors defined on a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, it is possible to construct a class of QIVPMs representing bounded resources that is parameterized by the size of the intervals, and the original Gleason theorem could be recovered for this class of QIVPMs asymptotically.

The organization of this thesis is summarized in Table 1.2, and finally, Chapter 6 contains further extended questions, especially, briefly explaining why it might be hard to recover Gleason's theorem for general QIVPMs.

Table 1.2: Comparison among our models, where the upper half of the table lists their quantum mechanical definitions and properties, and the lower half of the table lists their computational power.

	Conventional	Modal	Discrete (I)	Discrete (II)	QPMFF	QIVPM
	Chapter 2	Sec. 3.2	Sec. 4.1 to 4.3	Sec. 4.5	Sec. 4.7	Chapter 5
States space	$\mathbb{C}^D$	$\mathbb{F}_q^D$	$\mathbb{F}_{p^2}^D$	Local region in $\mathbb{F}_{p^2}^D$	$\mathbb{F}_{p^2}^D$	$\mathbb{C}^D$
Likelihood of events is predicted by	Real-valued probability	Possible or impossible	Possible or impossible	Cardinal probability	Real-valued probability, but no sensible Born rule	Interval-valued probability
Expectation value	Defined	Undefined	Formally defined	Undefined		Defined
		Sec. 3.3	Sec. 4.4	Sec. 4.6		
Deutsch's algorithm	Yes	Maybe no	Yes	Yes		
Efficiently solve UNIQUE-SAT	Unlikely	Yes	Partially	Unlikely		
Grover search algorithm	Yes			Yes		

Chapter 2

CONVENTIONAL QUANTUM THEORY

The part of conventional quantum theory (CQT) used by quantum circuit model is described by the following:

(i) D orthonormal basis vectors for a Hilbert space of dimension D,

(ii) D complex probability amplitude coefficients describing the contribution of each basis vector,

(iii) a set of probability-conserving unitary matrix operators that suffice to describe all required state transformations of a quantum circuit,

(iv) and a measurement framework.

In Sec. 2.1, we focus on the geometric issues raised by the properties (i) and (ii) given above for CQT. In Sec. 2.2, we introduce the important issues of (iv) and the foundations of quantum probability space. The expectation values defined in Sec. 2.2 will be used to understand the geometry of entangled states in Sec. 2.3. The property (iii) directly related to quantum circuits will be introduced later in Chapters 3 and 4 when we dive into quantum algorithms.

2.1 GEOMETRICAL STRUCTURE OF STATES

There are many things that are assumed in CQT, such as the absence of zero norm states for non-zero vectors, and the decomposition of complex amplitudes into a pair of ordinary real numbers. One

also typically assumes the existence of a D-dimensional Hilbert space with an orthonormal basis, allowing us to write *pure* states in general as Hilbert space vectors with a Hermitian inner product:

$$|\Psi\rangle = \sum_{i=0}^{D-1} \alpha_i \, |i\rangle \ . \tag{2.1}$$

Here $\alpha_i \in \mathbb{C}$ are complex probability amplitudes, $\vec{\boldsymbol{\alpha}} \in \mathbb{C}^D$, and the $\{|i\rangle\}$ is an orthonormal basis of states obeying $\langle i|k\rangle = \delta_{ik}$.

The meaning of this is that any state $|\Phi\rangle = \sum_{i=0}^{D-1} \beta_i \, |i\rangle$ can be projected onto another state $|\Psi\rangle$ by writing

$$\langle \Phi|\Psi\rangle = \sum_{i=0}^{D-1} \beta_i^* \alpha_i \, , \tag{2.2}$$

thus quantifying the proximity of the two states. (Here * denotes complex conjugation.) This is one of many properties we take for granted in continuum quantum mechanics that challenge us in defining a discrete quantum geometry. To facilitate the transition to DQT carried out in later sections, we concern ourselves first with the properties of the simplest possible abstract state object in CQT, the single qubit state.

2.1.1 TWO-DIMENSIONAL HILBERT SPACE

A state in a two-dimensional Hilbert space, known as a qubit, already provides access to a wealth of geometric information and context. When we write the single qubit state as $|\psi_1\rangle = \alpha_0 \, |0\rangle + \alpha_1 \, |1\rangle$, a convenience for computing probability and relative state properties is the normalization condition

$$\|\psi_1\|^2 = |\alpha_0|^2 + |\alpha_1|^2 = \alpha_0^* \alpha_0 + \alpha_1^* \alpha_1 = 1 \, , \tag{2.3}$$

which identifies α_0 and $\alpha_1 \in \mathbb{C}$ as probability amplitudes and implies the conservation of probability in the closed world spanned by $\{|0\rangle \, , |1\rangle\}$. Note that we distinguish for future use the *norm* $\|\cdot\|$ of a vector from the *modulus* $|\cdot|$ of a complex number. Continuing, we see that if we want only the

irreducible state descriptions, we must supplement the process of computing Eq. (2.3) by finding a way to remove the distinction between states that differ only by an overall phase transformation $\mathrm{e}^{\mathrm{i}\theta}$, that is, $\alpha_0 |0\rangle + \alpha_1 |1\rangle$ and $\mathrm{e}^{\mathrm{i}\theta}\alpha_0 |0\rangle + \mathrm{e}^{\mathrm{i}\theta}\alpha_1 |1\rangle$ are representing the same physical state. This can be accomplished by the Hopf fibration [35, 37, 38, 59–61], which can be written down as follows: let $\alpha_0 = x_0 + \mathrm{i}y_0$ and $\alpha_1 = x_1 + \mathrm{i}y_1$. Then Eq. (2.3) becomes the condition that the four real variables describing a qubit denote a point on the three-sphere $\mathbf{S}^3$ (a 3-manifold) embedded in $\mathbb{R}^4$:

$$ {x_0}^2 + {y_0}^2 + {x_1}^2 + {y_1}^2 = 1 \,. \tag{2.4} $$

We can reduce 3 degrees of freedom in Eq. (2.4) to 2 degrees of freedom by effectively removing $\mathrm{e}^{\mathrm{i}\theta}$ ("fibering out by the circle $\mathbf{S}^1$"). The standard form of these maps ("the Hopf fibration") is

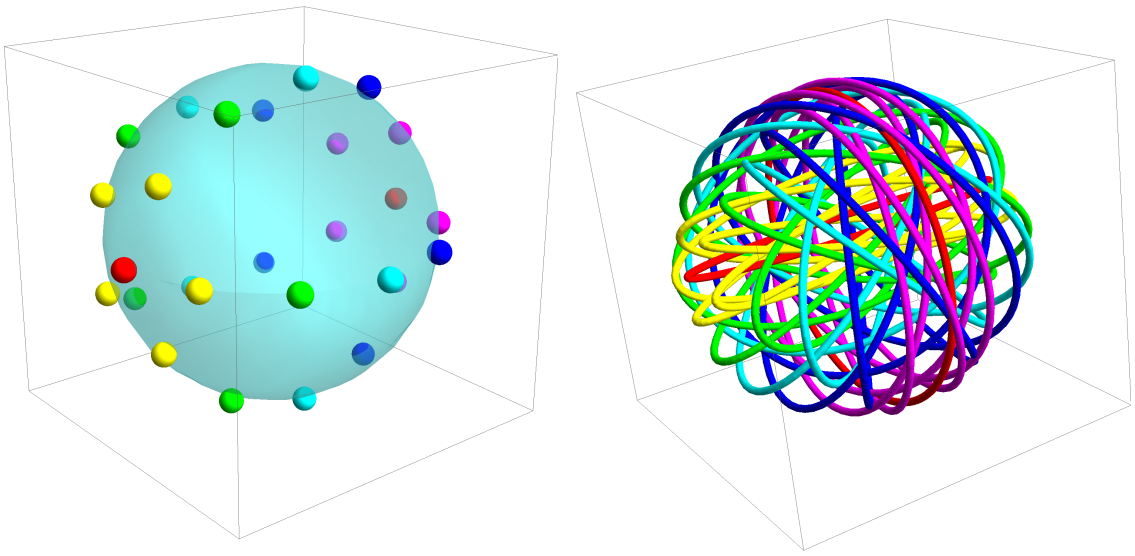
$$ \begin{aligned} X &= 2\,\mathrm{Re}\;\alpha_0\alpha_1^* = 2x_0x_1 + 2y_0y_1 \,, \\ Y &= 2\,\mathrm{Im}\;\alpha_0\alpha_1^* = 2x_1y_0 - 2x_0y_1 \,, \\ Z &= |\alpha_0|^2 - |\alpha_1|^2 = {x_0}^2 + {y_0}^2 - {x_1}^2 - {y_1}^2 \,. \end{aligned} \tag{2.5} $$

By denoting the three-dimensional vector (X, Y, Z) as $\hat{a}$, Eq. (2.4) implies these transformed coordinates obeying

$$ \|\hat{a}\|^2 = X^2 + Y^2 + Z^2 = \left(|\alpha_0|^2 + |\alpha_1|^2\right)^2 = 1 \tag{2.6} $$

and therefore have only two remaining degrees of freedom describing all possible distinct one-qubit quantum states. In Fig. 2.1, we illustrate schematically the family of circles *each one of which is collapsed to a point* (ϕ, ψ) on the surface $X^2 + Y^2 + Z^2 = 1$ by the Hopf map.

The resulting manifold is the two-sphere $\mathbf{S}^2$ (a 2-manifold) embedded in $\mathbb{R}^3$. If we choose one



(a) (b)

Figure 2.1: (a) The two-sphere $\mathbf{S}^2$ represented by Eq. (2.6), which is the irreducible space of one-qubit states, along with a representative set of points on the sphere. Each single point on the sphere in (a) corresponding to a circle in (b), and a whole family of circles (the paths of $\mathrm{e}^{\mathrm{i}\theta}$) on the three-sphere $\mathbf{S}^3$ represents the Hopf fibration, Eq. (2.5). Although $\mathbf{S}^3$ cannot be directly embedded in $\mathbb{R}^3$, three-sphere $\mathbf{S}^3$ can be regarded as attaching two three-dimensional balls on two sides of two-sphere $\mathbf{S}^2$. In this way, each circle in $\mathbf{S}^3$ can be represented as a circle in the three-dimensional ball as shown in (b). Moreover, points in (a) are color-coded corresponding to circles in (b), e.g., one pole contains the red elliptical circle that would become an infinite-radius circle by a slightly different way to represent $\mathbf{S}^3$ in $\mathbb{R}^3$, and the opposite pole corresponds to the large perfectly round red circle at the equator.

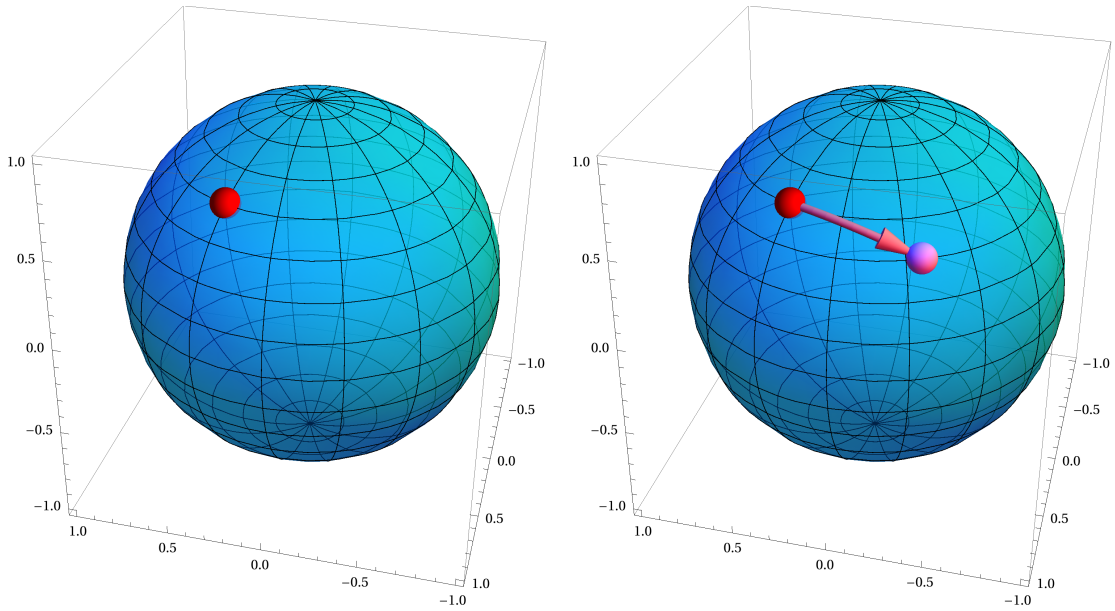
of many possible coordinate systems describing $\mathbf{S}^3$ via Eq. (2.4) such as

$$(x_0, y_0, x_1, y_1) = (\cos(\theta+\phi)\cos\psi,\ \sin(\theta+\phi)\cos\psi,\ \cos(\theta-\phi)\sin\psi,\ \sin(\theta-\phi)\sin\psi)\ , \tag{2.7}$$

where $0 \le \psi \le \frac{\pi}{2}$, with $0 \le \theta+\phi < 2\pi$ and $0 \le \theta-\phi < 2\pi$, we see that

$$(X, Y, Z) = (\cos(2\phi)\sin(2\psi)\,,\ \sin(2\phi)\sin(2\psi)\,,\ \cos(2\psi))\ . \tag{2.8}$$

Thus, the one-qubit state is independent of θ, and we can choose $\theta = \phi$ without loss of generality, reducing the formula of the unique one-qubit states to $|\psi_1\rangle = \mathrm{e}^{2\mathrm{i}\phi}\cos\psi\,|0\rangle + \sin\psi\,|1\rangle$, and an



(a) (b)

Figure 2.2: (a) The conventional Bloch sphere with a unique state represented by the point at the red sphere. (b) The geodesic shortest-distance arc connecting two one-qubit quantum states.

irreducible state can be represented as a point on a sphere called the Bloch sphere, as shown in Fig. 2.2 (a).

Thus, the geometry of a single qubit reduces to transformations among points on $\mathbf{S}^2$, which can be parametrized in an infinite one-parameter family of transformations, one of which is the geodesic or minimal-length transformation. Explicitly, given two one-qubit states denoted by points $\hat{a}$ and $\hat{b}$ on $\mathbf{S}^2$, the shortest rotation carrying $\hat{a}$ to $\hat{b}$ is the SLERP (spherical linear interpolation) [62, 63]

$$S\left(\hat{a},\hat{b},t\right) = \hat{a}\,\frac{\sin((1-t)\omega)}{\sin\omega} + \hat{b}\,\frac{\sin(t\omega)}{\sin\omega}\,, \tag{2.9}$$

where $\hat{a}\cdot\hat{b} = \cos\omega$. Figure 2.2 (b) illustrates the path traced by a SLERP between two irreducible one-qubit states on the Bloch sphere. Because states in CQC are defined by infinite precision real numbers, it is not possible, even in principle, to make an exact state transition as implied by Fig. 2.2 (b). In practice, one must be content with approximate, typically exponentially expensive, transitions from state to state.

2.1.2 D-DIMENSIONAL HILBERT SPACE

The irreducible states in a D-dimensional Hilbert space are encoded in a similar family of geometric structures known technically as the complex projective space $\mathbb{C}\mathbf{P}^{D-1}$. We obtain these structures starting with the D initially unnormalized complex coefficients of the D-dimensional basis $|\Psi\rangle = \sum_{i=0}^{D-1} \alpha_i |i\rangle$. We then follow the analog of the two-dimensional procedure: Conservation of probability requires that the norm of the vector $\vec{\boldsymbol{\alpha}}$ be normalized to unity:

$$\langle\Psi|\Psi\rangle = \|\vec{\boldsymbol{\alpha}}\|^2 = \sum_{i=0}^{D-1} |\alpha_i|^2 = 1\,. \tag{2.10}$$

Thus, the initial equation for the geometry of a quantum state describes a *topological sphere* $\mathbf{S}^{2D-1}$ embedded in $\mathbb{R}^{2D}$. To see this, remember that we can write the real and imaginary parts of α_i as $\alpha_i = x_i + \mathrm{i}y_i$, so

$$\sum_{i=0}^{D-1} |\alpha_i|^2 = \sum_{i=0}^{D-1} {x_i}^2 + {y_i}^2 = 1 \tag{2.11}$$

describes the locus of a $2D$-dimensional real unit vector in $\mathbb{R}^{2D}$, which is by definition $\mathbf{S}^{2D-1}$, the $(2D-1)$-sphere.

This $\mathbf{S}^{2D-1}$ is ambiguous up to the usual overall phase, inducing an $\mathbf{S}^1$ symmetry action, and identifying $\mathbf{S}^{2D-1}$ as an $\mathbf{S}^1$ bundle, whose base space is the $(D-1)$-complex-dimensional projective space $\mathbb{C}\mathbf{P}^{D-1}$. There are thus $2D-2$ irreducible real degrees of freedom ($D-1$ complex degrees of freedom) for a quantum state with a D-dimensional basis, $\{|i\rangle \mid i = 0, \ldots, D-1\}$.

In summary, the full space of a D-dimensional quantum state, including its overall phase defining its relationship to other quantum states, is the topological space $\mathbf{S}^{2D-1}$. For an isolated system, the overall phase is not measurable, and eliminating the phase dependence corresponds to identifying $\mathbf{S}^{2D-1}$ as a circle bundle over the base space $\mathbb{C}\mathbf{P}^{D-1}$, and therefore $\mathbb{C}\mathbf{P}^{D-1}$ defines the $2D-2$ intrinsic, irreducible, degrees of freedom of the isolated D-dimensional state's dynamics. In mathematical notation, this would be written $\mathbf{S}^1 \hookrightarrow \mathbf{S}^{2D-1} \to \mathbb{C}\mathbf{P}^{D-1}$. For $D = 2$, the single qubit, we

have $2-1=1$, and the base space of the circle bundle is $\mathbb{C}\mathbf{P}^1 = \mathbf{S}^2$, the usual Bloch sphere. Note that only for $D=2$ is this actually a sphere-like geometry due to an accident of low-dimensional topology.

2.1.3 EXPLICIT GENERALIZATION OF THE HOPF FIBRATION CONSTRUCTION

For a two-dimensional system, we could easily solve the problem of reducing the full unit-norm space to its irreducible components $\hat{a} = (X, Y, Z)$ characterizing the Bloch sphere. We have just argued that essentially the same process is possible for D-dimensional system: in the abstract argument, we simply identify the family of coefficients $\{\alpha_i\}$ as being the same if they differ only by an overall phase $\mathrm{e}^{\mathrm{i}\theta}$. However in practice, this is not a construction that is easy to realize in a practical computation. We now outline an explicit algorithm for accomplishing the reduction to the irreducible D-dimensional state space $\mathbb{C}\mathbf{P}^{D-1}$; this construction will turn out to be useful for the validation of our discrete results to follow below.

Given a normalized pure state $|\Psi\rangle = \sum_{i=0}^{D-1} \alpha_i |i\rangle$, a natural quantity characterizing a D-dimensional system is its *density matrix*, $\rho = |\Psi\rangle\langle\Psi|$, or

$$\rho = \begin{pmatrix} |\alpha_0|^2 & \alpha_0\alpha_1^* & \cdots & \alpha_0\alpha_{D-1}^* \\ \alpha_1\alpha_0^* & |\alpha_1|^2 & \cdots & \alpha_1\alpha_{D-1}^* \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{D-1}\alpha_0^* & \cdots & \alpha_{D-1}\alpha_{D-2}^* & |\alpha_{D-1}|^2 \end{pmatrix}. \tag{2.12}$$

We can now use the complex generalization of the classical Veronese coordinate system for projective geometry to remove the overall phase ambiguity $\mathrm{e}^{\mathrm{i}\theta}$ from the D-dimensional states. If we take a particular weighting of the elements of the density matrix ρ, we can construct a *unit vector* of real

dimension D^2 with the form:

$$\hat{a} = \left(|\alpha_i|^2, \ldots, \sqrt{2}\,\mathrm{Re}\;\alpha_i\alpha_j^*, \ldots, \sqrt{2}\,\mathrm{Im}\;\alpha_i\alpha_j^*, \ldots \right) , \tag{2.13}$$

where

$$\hat{a}\cdot\hat{a} = \sum_{i=0}^{D-1}\left(|\alpha_i|^2\right)^2 + \sum_{i=0}^{D-1}\sum_{\substack{j=0\\ j\neq i}}^{D-1}\left(\mathrm{Re}\;\alpha_i\alpha_j^*\right)^2 + \left(\mathrm{Im}\;\alpha_i\alpha_j^*\right)^2 = \left(\sum_{i=0}^{D-1}|\alpha_i|^2\right)\left(\sum_{j=0}^{D-1}|\alpha_j|^2\right) = 1\,.$$

This construction gives an explicit embedding of the $(D-1)$-dimensional complex, or $(2D-2)$-dimensional real, object in a real space of dimension D^2. However, this is somewhat subtle because the vector is of unit length, so technically the embedding space is a sphere of dimension D^2-1 embedded in $\mathbb{R}^{D^2}$. For example, the two-dimensional irreducible states could be represented in a four-dimensional embedding, but the magnitude of every coordinate would be one. Furthermore, the object embedded in the resulting $\mathbf{S}^3$ is indeed $\mathbf{S}^2$ because we can fix one complex coordinate to be unity, and let one vary, giving a total of two irreducible dimensions. In fact, one must choose *two* coordinate patches, one covering one pole of $\mathbf{S}^2$ with coordinates $\alpha_0 = 1 + \mathrm{i}0$ and $\alpha_1 = x_1 + \mathrm{i}y_1$, and the other patch covering the other pole of $\mathbf{S}^2$ with coordinates $\alpha_0 = x_0 + \mathrm{i}y_0$ and $\alpha_1 = 1 + \mathrm{i}0$.

We finally see that the irreducible D-dimensional state space $\mathbb{C}\mathbf{P}^{D-1}$ is described by D projectively equivalent coordinates, one of which can always be scaled out to leave $(D-1)$ actual (complex) degrees of freedom. We must choose, in turn, D different local sets of complex variables defined by taking the value $\alpha_k = 1$, with $k = 0, \ldots, D-1$, and allowing the remaining $D-1$ complex (or $2D-2$ real) variables to run free. No single set of coordinates will work, since the submanifold including $\alpha_k = 0$ is undefined and another coordinate system must be chosen to cover that coordinate patch. This is a standard feature of the topology of non-trivial manifolds such as $\mathbb{C}\mathbf{P}^{D-1}$ (see any textbook on geometry [64]).

2.2 QUANTUM PROBABILITY

A *probability space* is a mathematical abstraction specifying the necessary conditions for reasoning coherently about collections of uncertain events [10, 34, 65, 66]. Although they can be used to describe an individual quantum experiment, to describe a family of quantum experiments, we would like to glue their probability spaces together to define a quantum probability space. The glued quantum probability space is well-behaved since not only we can define the expectation values, but the quantum probability measure defined on the whole space can be simply induced by the Born rule according to Gleason's theorem [39–43].

2.2.1 CLASSICAL AND QUANTUM PROBABILITY SPACES

Given a finite sample space Ω representing all possible outcomes of a process, and its power set 2^{Ω} as the classical event space, a classical probability measure μ maps every event $E \subseteq \Omega$ to a number $\mu(E) \in [0, 1]$ specifying how likely one of the outcomes in E will happen. To maintain the coherence, μ is subject to the following constraints: $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$, and $\mu\left(\overline{E}\right) = 1 - \mu(E)$, where $\overline{E}$ is the complement of E. Moreover, we require $\mu(E_0 \cup E_1) = \mu(E_0) + \mu(E_1)$ for each pair of disjoint events $E_0 \subseteq \Omega$ and $E_1 \subseteq \Omega$.

Since the previous abstraction doesn't specify the process to generate the outcomes, this process could well be a quantum experiment. Let us prepare a beam of one kind of spin 1 particles whose state can be characterized by a vector in three-dimensional Hilbert space with basis vectors $|0\rangle$, $|1\rangle$, and $|2\rangle$. In principle, this beam can be split by a Stern-Gerlach type experiment according to the eigenvalues of the observable $\mathbf{O}_0$ with spectral decomposition

$$\mathbf{O}_0 = 0\ |0\rangle\langle 0| + 1\ |1\rangle\langle 1| + 2\ |2\rangle\langle 2|\ , \tag{2.14}$$

and the states corresponding to the split beams are $|0\rangle$, $|1\rangle$, and $|2\rangle$ [41, 67]. Instead of sending a

beam of particles, if only one particle is sent to the beam splitter, the state of the particle after the experiment is one of $|0\rangle$, $|1\rangle$, and $|2\rangle$, and there is a probability to get each post-experimental state. In the language of our abstraction, the set of outcomes is $\Omega_0 = \{|0\rangle, |1\rangle, |2\rangle\}$. All possible events are $\emptyset$, $\{|0\rangle\}$, $\{|1\rangle\}$, $\{|2\rangle\}$, $\{|0\rangle, |1\rangle\}$, $\{|0\rangle, |2\rangle\}$, $\{|1\rangle, |2\rangle\}$, and Ω_0. And this experiment defines a classical probability measure $\mu_0 \colon 2^{\Omega_0} \to [0, 1]$.

One special feature of quantum experiments is that the probabilities in different experiments are correlated. Consider another experiment sending exactly the same particle as the previous one but using a different beam splitter corresponding to the observable $\mathbf{O}_1$ with spectral decomposition $\mathbf{O}_1 = 0\ |+\rangle\langle+| + 1\ |-\rangle\langle-| + 2\ |2\rangle\langle 2|$, where $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$. Although the sample space $\Omega_1 = \{|+\rangle, |-\rangle, |2\rangle\}$ and the probability measure $\mu_1 \colon 2^{\Omega_1} \to [0, 1]$ defined by this experiment is different from the previous one, these two experiments may produce the same post-experimental state $|2\rangle$, and the probability of the common event $\{|2\rangle\}$ is believed to be the same, i.e.,

$$\mu_0(\{|2\rangle\}) = \mu_1(\{|2\rangle\}) . \tag{2.15}$$

In general, as long as sending the same particle, the probability of the same event in different experiments should always be the same. This fact is equivalent to the fact that commuting observables could be measured simultaneously which is essential to define contextuality and will be explained in Sec. 5.2.

Since the probability induced by the different beam splitters are correlated, it is more natural to define one quantum event space $\mathcal{E}$ containing all possible classical event spaces using different beam splitters. However, simply taking the union of all event spaces is a bad idea because two events might appear different but representing the same situation. For example, if we take the complement on both sides of Eq. (2.15), we will have

$$\mu_0(\{|0\rangle, |1\rangle\}) = 1 - \mu_0(\{|2\rangle\}) = 1 - \mu_1(\{|2\rangle\}) = \mu_1(\{|+\rangle, |-\rangle\}) , \tag{2.16}$$

that is, the probabilities of events $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ are always the same, and these events should be identified as the same quantum event to simplify our discussion. This identification can be achieved by mapping a classical event E to the projector generated by E,

$$\varphi(E) = \sum_{|j\rangle \in E} |j\rangle\langle j| \tag{2.17}$$

with the convention $\varphi(\emptyset) = \mathbb{0}$, because $\varphi(\{|0\rangle, |1\rangle\})$ is equal to $\varphi(\{|+\rangle, |-\rangle\})$ as operators,

$$\varphi(\{|0\rangle, |1\rangle\}) = |0\rangle\langle 0| + |1\rangle\langle 1| = |+\rangle\langle +| + |-\rangle\langle -| = \varphi(\{|+\rangle, |-\rangle\}) \ . \tag{2.18}$$

In general, if two classical events E and E' are mapped to the same projector, i.e., $\varphi(E) = \varphi(E')$, then the probability of E is the same as the probability of E'. Therefore, for any classical event E, its corresponding quantum event is defined to be the projector $\varphi(E)$, and the set of all projectors on a given Hilbert space is called a quantum event space $\mathcal{E}$.

This function φ not only respects the probability of events but also naturally sends the set structure to the corresponding projector structure:

$$\varphi(\Omega) = \mathbb{1}, \qquad \varphi\left(\overline{E}\right) = \mathbb{1} - \varphi(E) \ . \tag{2.19}$$

Given two *commuting* projectors P_0 and P_1, there exists a pair of events E_0 and E_1 in the same sample space Ω such that $P_0 = \varphi(E_0)$ and $P_1 = \varphi(E_1)$. Conversely, given a pair of events E_0 and E_1 in the same sample space Ω, their corresponding quantum events $\varphi(E_0)$ and $\varphi(E_1)$ are commuting and satisfying the following properties:

$$\varphi(E_0 \cap E_1) = \varphi(E_0)\,\varphi(E_1) \ , \quad \varphi(E_0 \cup E_1) = \varphi(E_0) + \varphi(E_1) - \varphi(E_0)\,\varphi(E_1) \ . \tag{2.20}$$

Moreover, E_0 and E_1 are disjoint if and only if $\varphi(E_0)$ and $\varphi(E_1)$ are orthogonal, where two pro-

jectors P_0 and P_1 are called *orthogonal* if $P_0 P_1 = \mathbb{0}$.

Then, a quantum probability space can be defined as a quantum event space $\mathcal{E}$ together with a quantum probability measure $\mu \colon \mathcal{E} \to [0, 1]$ subject to the corresponding constraints [5, 39, 40, 68, 69]:

$$\mu(\mathbb{0}) = 0\,, \qquad \mu(\mathbb{1}) = 1\,, \qquad \mu(\mathbb{1} - P) = 1 - \mu(P)\,, \tag{2.21}$$

and for each pair of orthogonal projectors P_0 and P_1:

$$\mu(P_0 + P_1) = \mu(P_0) + \mu(P_1)\,. \tag{2.22}$$

Because φ respects the probability of events, if we restrict the domain of φ on a classical event space 2^Ω, the function $\varphi^* \mu \colon 2^\Omega \to [0, 1]$ defined by precomposition

$$(\varphi^* \mu)(E) = \mu(\varphi(E)) \tag{2.23}$$

is a classical probability measure and called the pullback of μ by $\varphi \colon 2^\Omega \to \mathcal{E}$.

Given a Hilbert space $\mathcal{H}$ of dimension D and a probability assignment for every projector P, we can define the expectation value of an observable $\mathbf{O}$ having spectral decomposition $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i P_i$, with eigenvalues $\lambda_i \in \mathbb{R}$, as [10, 36]:

$$\langle \mathbf{O} \rangle_\mu = \sum_{i=0}^{D-1} \lambda_i \mu(P_i)\,, \tag{2.24}$$

where the subscript μ might be omitted if it is clear according to the context. This definition is also consistent with the classical expectation values because we can pullback an observable to a classical random variable, and the expectation values are invariant.

Definition 2.1 (Pullback of Observables)**.** Consider an observable $\mathbf{O}$ diagonalizable by an orthonor-

mal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ so that $\mathbf{O}$ has spectral decomposition $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i |i\rangle\langle i|$. If we restrict φ on the classical event space 2^Ω and consider $\varphi\colon 2^\Omega \to \mathcal{E}$, then the pullback of $\mathbf{O}$ by φ is a random variable $\varphi^*\mathbf{O}\colon \Omega \to \mathbb{R}$ defined by $\varphi^*\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i \mathbf{1}_{\{|i\rangle\}}$, where $\mathbf{1}_E$ is the indicator function defined by

$$\mathbf{1}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E\,; \\ 0 & \text{if } \omega \notin E\,. \end{cases} \tag{2.25}$$

Lemma 2.1. *Consider an observable* $\mathbf{O}$ *diagonalizable by an orthonormal basis* $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ *with* $\varphi\colon 2^\Omega \to \mathcal{E}$ *defined by Eq. (2.17). Given a quantum probability measure* $\mu\colon \mathcal{E} \to [0,1]$*, the expectation value of* $\mathbf{O}$ *relative to* μ *is exactly the expectation value of the pullback of* $\mathbf{O}$ *relative to the pullback of* μ*, i.e.,*

$$\langle \mathbf{O} \rangle_\mu = \int (\varphi^*\mathbf{O})\, \mathrm{d}\,(\varphi^*\mu)\ . \tag{2.26}$$

Proof. By Eqs. (2.24), (2.17), and (2.23), we have

$$\langle \mathbf{O} \rangle_\mu = \sum_{i=0}^{D-1} \lambda_i \mu(|i\rangle\langle i|) = \sum_{i=0}^{D-1} \lambda_i \mu(\varphi(\{|i\rangle\})) = \sum_{i=0}^{D-1} \lambda_i (\varphi^*\mu)(\{|i\rangle\}) = \int (\varphi^*\mathbf{O})\, \mathrm{d}\,(\varphi^*\mu)\ .$$

□

2.2.2 GLEASON'S THEOREM AND THE BORN RULE

After introducing quantum probability measures, we might follow the convention to introduce the Born rule which is the only way corresponding a state to a quantum probability measure in CQT. However, since the variants of quantum probability measures in Sec. 4.7 and Chapter 5 could not be constructed by a "Born rule" easily, searching for a Born rule step-by-step here might provide a better idea of what we should do in Sec. 4.7 and Chapter 5. While the situations will become more delicate in the later sections, we will fortunately find the unique Born rule for CQT here.

Although the quantum probability measure constructed by gluing together classical ones looks complex, it could be induced by an operator according to Gleason's theorem [39–43].

Theorem 2.1 (Gleason's theorem)**.** *In a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, given a quantum probability measure $\mu \colon \mathcal{E} \to [0,1]$, there exists a unique mixed state $\rho = \sum_{j=1}^{N} q_j \left|\Phi_j\right\rangle \left\langle\Phi_j\right|$ such that $\mu(P) = \mathrm{Tr}(\rho P)$ for any D-dimensional projector P, where $\left|\Phi_j\right\rangle \in \mathcal{H}$ are normalized, $q_j > 0$, and $\sum_{j=1}^{N} q_j = 1$.*

If we follow the discussion of Stern-Gerlach type experiments in the previous section to interpret Gleason's theorem, we can find Gleason's theorem doesn't specify whether its unique mixed state ρ characterizes the state of particle sending to the quantum experiment or not. Consider an extreme example: a quantum theory could ignore the input state and predict the experimental outcomes are always equally probable. Even if these predictions form a quantum probability measure, this kind of prediction is so different from the prediction of CQT that they can be easily distinguished experimentally.

Let a pure unnormalized state $|\Phi\rangle \in \mathcal{H}$ characterize the particle sending to a Stern-Gerlach type experiment, and μ_Φ^{B} be the quantum probability measure of the resulting events sensibly corresponding to $|\Phi\rangle$. For a correspondence $|\Phi\rangle \mapsto \mu_\Phi^{\mathrm{B}}$ to be sensible, we hope that if the state $|\Phi\rangle$ is one of the outcomes of a quantum event P, $P|\Phi\rangle = |\Phi\rangle$, then the event P always happens,

$$\mu_\Phi^{\mathrm{B}}(P) = 1\,, \tag{2.27}$$

and vice versa. Moreover, since the physical phenomena exist and should be the same no matter how we describe them, the probability of an event should be invariant despite how we choose the basis. Because changing to another basis is the same as applying a unitary map U, we should have

$$\mu_{U|\Phi\rangle}^{\mathrm{B}}\left(UPU^\dagger\right) = \mu_\Phi^{\mathrm{B}}(P)\ . \tag{2.28}$$

Table 2.1: Two fragments of valid probability measures μ_1 and μ_2.

$\|\Psi\rangle$	$\|0\rangle$	$\|1\rangle$	$\|2\rangle$	$\frac{\|0\rangle+\|1\rangle}{\sqrt{2}}$	$\frac{\|0\rangle+\mathrm{i}\|1\rangle}{\sqrt{2}}$	$\frac{\|0\rangle+\|2\rangle}{\sqrt{2}}$	$\frac{\|0\rangle+\mathrm{i}\|2\rangle}{\sqrt{2}}$	$\frac{\|1\rangle+\|2\rangle}{\sqrt{2}}$	$\frac{\|1\rangle+\mathrm{i}\|2\rangle}{\sqrt{2}}$	…
$\mu_1\left(\|\Psi\rangle\langle\Psi\|\right)$	$\frac{1}{2}$	$\frac{1}{2}$	0	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	…
$\mu_2\left(\|\Psi\rangle\langle\Psi\|\right)$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	…

It is easy to check that the correspondence satisfying these conditions is unique,

$$\mu_{\Phi}^{\mathrm{B}}(P)=\frac{\langle\Phi|P|\Phi\rangle}{\langle\Phi|\Phi\rangle} \tag{2.29}$$

and called the Born rule [4, 36, 44]. If $|\Phi\rangle$ is normalized, Eq. (2.29) could be simplified as $\mu_{\Phi}^{\mathrm{B}}(P)=\langle\Phi|P|\Phi\rangle$. Since a mixed state $\rho=\sum_{j=1}^{N}q_j\left|\Phi_j\right\rangle\left\langle\Phi_j\right|$ is a weighted average of projectors $\left|\Phi_j\right\rangle\left\langle\Phi_j\right|$ with weights $q_j>0$ and $\sum_{j=1}^{N}q_j=1$, the generalized Born rule of ρ, $\mu_{\rho}^{\mathrm{B}}\left(P\right)$, is also a weighted average of $\mu_{\Phi_j}^{\mathrm{B}}\left(P\right)$,

$$\mu_{\rho}^{\mathrm{B}}\left(P\right)=\sum_{j=1}^{N}q_j\mu_{\Phi_j}^{\mathrm{B}}\left(P\right)=\mathrm{Tr}\left(\rho P\right)\ , \tag{2.30}$$

which is also a quantum probability measure and consistent with Gleason's theorem.

As an example, consider a three-dimensional Hilbert space with orthonormal basis $\{|0\rangle\,,|1\rangle\,,|2\rangle\}$ and the observable $\mathbf{O}_0$ defined in Eq. (2.14). Two fragments of valid probability measures μ_1 and μ_2 that can be associated with this space are defined in Table 2.1. By the Born rule, the first probability measure corresponds to the quantum system being in the pure state $|+\rangle=\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and the second corresponds to the quantum system being in the state $\frac{|0\rangle\langle 0|+|2\rangle\langle 2|}{2}$. The expectation values of the observable $\mathbf{O}$, $\langle\mathbf{O}\rangle_{\mu_{1,2}}$, are 1.5 in the first case and 2 in the second. The quantum expectation value can also be used to decide whether a state is entangled or not for multipartite systems as we describe in the following section.

2.3 THE GEOMETRY OF ENTANGLEMENT

Entanglement may be regarded as one of the main characteristics distinguishing quantum from classical mechanics. Entanglement involves quantum correlations such that the measurement outcomes in one subsystem are related to the measurement outcomes in another one. To discuss entanglement, we consider a D-dimensional quantum system composed of n-qubit subsystems, i.e., $D = 2^n$. A pure state of the total system $|\Psi\rangle$ is said to be entangled if it cannot be written as a product of states of each subsystem [10, 36, 41]. That is, a state $|\Psi\rangle$ is entangled if $|\Psi\rangle \neq |\psi_1\rangle \otimes \cdots \otimes \left|\psi_j\right\rangle \otimes \cdots \otimes |\psi_n\rangle$, where $\left|\psi_j\right\rangle$ refers to an arbitrary state of the j-th qubit, and $\otimes$ represents the tensor product. This is equivalent to saying that if one calculates the reduced density operator ρ_j of the j-th subsystem by tracing out all the other subsystems,

$$\rho_j = \mathrm{Tr}_{\{1,\ldots,j-1,j+1,\ldots,n\}}\left(\rho\right) , \tag{2.31}$$

with $j = 1, \ldots, n$ and $\rho = |\Psi\rangle\langle\Psi|$, the normalized state $|\Psi\rangle$ is entangled if and only if at least one subsystem state is *mixed*, i.e., $\mathrm{Tr}_j\left(\rho_j^2\right) < 1$ [10, 36].

The reduced density operator could be expressed explicitly by the expectation value of the Pauli operators. Therefore, we can decide whether a system is entangled or not by examining these expectation values. Let σ_η^j be the Pauli operators acting on the j-th spin [10],

$$\sigma_\eta^j = \overbrace{\sigma_0 \otimes \cdots \otimes \sigma_0 \otimes \underbrace{\sigma_\eta}_{j^{\text{th}}\text{ factors}} \otimes \sigma_0 \otimes \cdots \otimes \sigma_0}^{n\text{ factors}} , \tag{2.32}$$

and $\left\langle \sigma_\eta^j \right\rangle$ be the corresponding expectation value, $\left\langle \sigma_\eta^j \right\rangle = \left\langle \Psi \middle| \sigma_\eta^j \middle| \Psi \right\rangle$, where $\eta = x$, y, and z, and

$$\sigma_0 = |0\rangle\langle 0| + |1\rangle\langle 1| , \qquad \sigma_x = |1\rangle\langle 0| + |0\rangle\langle 1| , \tag{2.33a}$$

$$\sigma_y = \mathrm{i}\,|1\rangle\langle 0| - \mathrm{i}\,|0\rangle\langle 1| , \qquad \sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1| . \tag{2.33b}$$

For example, given a normalized two-qubit system $|\Psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$, some of its expectation values are

$$\begin{aligned}
\langle\sigma_0^1\rangle &= |\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1\,, \\
\langle\sigma_x^1\rangle &= \alpha_{00}\alpha_{10}^* + \alpha_{01}\alpha_{11}^* + \alpha_{10}\alpha_{00}^* + \alpha_{11}\alpha_{01}^*\,, \\
\langle\sigma_y^1\rangle &= -\alpha_{00}\alpha_{10}^*\mathrm{i} - \alpha_{01}\alpha_{11}^*\mathrm{i} + \alpha_{10}\alpha_{00}^*\mathrm{i} + \alpha_{11}\alpha_{01}^*\mathrm{i}\,, \\
\langle\sigma_z^1\rangle &= |\alpha_{00}|^2 + |\alpha_{01}|^2 - |\alpha_{10}|^2 - |\alpha_{11}|^2\,.
\end{aligned} \tag{2.34}$$

Then, the reduced density operator ρ_1 can be expressed by these expectation values as following:

$$\begin{aligned}
\rho_1 =& \mathrm{Tr}_{\{2\}}\left(|\Psi\rangle\langle\Psi|\right) \\
=& \left(|\alpha_{00}|^2 + |\alpha_{01}|^2\right)|0\rangle\langle 0| + \left(\alpha_{00}\alpha_{10}^* + \alpha_{01}\alpha_{11}^*\right)|0\rangle\langle 1| \\
&+ \left(\alpha_{10}\alpha_{00}^* + \alpha_{11}\alpha_{01}^*\right)|1\rangle\langle 0| + \left(|\alpha_{10}|^2 + |\alpha_{11}|^2\right)|1\rangle\langle 1| \\
=& \frac{\langle\sigma_0^1\rangle\,\sigma_0 + \langle\sigma_x^1\rangle\,\sigma_x + \langle\sigma_y^1\rangle\,\sigma_y + \langle\sigma_z^1\rangle\,\sigma_z}{2}\,.
\end{aligned} \tag{2.35}$$

In general, the reduced density operator ρ_j of the j-th subsystem can always be expressed as

$$\rho_j = \frac{1}{2}\sum_{\eta=0,x,y,z}\left\langle\sigma_\eta^j\right\rangle\sigma_\eta\,, \tag{2.36}$$

and its coefficients can be summarized as the vector

$$\mathbf{X}_j = \left(\left\langle\sigma_x^j\right\rangle, \left\langle\sigma_y^j\right\rangle, \left\langle\sigma_z^j\right\rangle\right) \in \mathbb{R}^3 \tag{2.37}$$

that allows a geometric representation of each reduced state in $\mathbb{R}^3$, satisfying $0 \leq \left\|\mathbf{X}_j\right\| \leq 1$. Since $\mathrm{Tr}_j\left(\rho_j^2\right) = \frac{1}{2}\left(1 + \left\|\mathbf{X}_j\right\|^2\right)$, the state $|\Psi\rangle$ is entangled if $\left\|\mathbf{X}_j\right\| < 1$ for at least one j, represented by a point *inside* the corresponding local Bloch sphere embedded in $\mathbb{R}^3$. Therefore, one may consider $|\Psi\rangle$ to be maximally entangled if $\left\|\mathbf{X}_j\right\| = 0$ for all j. On the other hand, the state $|\Psi\rangle$ is unentangled

(i.e., a product state) if $\left\|\mathbf{X}_j\right\| = 1$ for all j, corresponding to points lying on the surface of the Bloch spheres.

A natural geometric measure of multipartite entanglement is obtained by defining the *purity of a state relative to a set of observables* [70, 71]. If the set is chosen to be the set of *all local observables*, i.e., corresponding to each of the subsystems that compose the actual system, one recovers the standard notion of entanglement for multipartite systems. For example, if the system consists of n qubits, we obtain a measure of conventional entanglement by calculating the purity relative to the semi-simple Lie algebra $\mathfrak{h}$ spanned by $\{\sigma_x^1, \sigma_y^1, \sigma_z^1, ..., \sigma_x^n, \sigma_y^n, \sigma_z^n\}$,

$$P_{\mathfrak{h}} = \frac{1}{n}\sum_{j=1}^{n}\sum_{\eta=x,y,z}\left\langle\sigma_\eta^j\right\rangle^2 = \frac{1}{n}\sum_{j=1}^{n}\left\|\mathbf{X}_j\right\|^2 . \tag{2.38}$$

Since the norm of the geometric representation state $\left\|\mathbf{X}_j\right\|$ defined in Eq. (2.37) is between 0 and 1, we have $0 \leq P_{\mathfrak{h}} \leq 1$, where $\frac{1}{n}$ in Eq. (2.38) is just a normalization factor. All the product states of the form $|\Psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$, have maximum purity (i.e., $P_{\mathfrak{h}} = 1$). Other states such as the Greenberger-Horne-Zeilinger state $|\Psi\rangle = |\mathsf{GHZ}_n\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \otimes \cdots \otimes |0\rangle + |1\rangle \otimes \cdots \otimes |1\rangle\right)$ are (maximally) entangled relative to the set of local observables (i.e., $P_{\mathfrak{h}} = 0$).

Different entanglement measures are obtained when an algebra $\mathfrak{h}$ different from the local observables is chosen. An obvious example is given by the set of all observables. In this case, the purity takes its maximum value independently of the pure quantum state [70, 71], expressing the fact that any state is a generalized coherent state of the Lie algebra of all observables.

Chapter 3

QUANTUM THEORIES AND COMPUTING OVER UNRESTRICTED FINITE FIELDS

3.1 FUNDAMENTALS OF FINITE FIELDS

3.1.1 BACKGROUND

A field $\mathbb{F}$ is an algebraic structure consisting of a set of elements equipped with the operations of addition, subtraction, multiplication, and division [72–74]. Fields may contain an infinite or a finite number of elements. The rational $\mathbb{Q}$, real $\mathbb{R}$, and complex numbers $\mathbb{C}$ are examples of infinite fields, while the set $\mathbb{F}_3 = \{0, 1, 2\}$, under multiplication and addition modulo 3, is an example of a finite field.

There are two distinguished elements in a field, the addition identity 0, and the multiplication identity 1. Given the field $\mathbb{F}$, the closed operations of addition, "$+$," and multiplication, "$*$," satisfy the following set of axioms:

1. $\mathbb{F}$ is an Abelian group under the addition operation $+$ (additive group);

2. The multiplication operation $*$ is associative and commutative. The field has a multiplicative identity and the property that every nonzero element has a multiplicative inverse;

3. Distributive laws: For all $a, b, c \in \mathbb{F}$

$$a * (b + c) = a * b + a * c\,, \qquad (b + c) * a = b * a + c * a\,. \tag{3.1}$$

From now on, unless specified, we will omit the symbol $*$ whenever we multiply two elements of a field.

Finite fields of q elements, $\mathbb{F}_q = \{0, \ldots, q-1\}$, will play a special role in this work. A simple explicit example is $\mathbb{F}_3$ with the following addition and multiplication tables:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$*$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

3.1.2 CYCLIC PROPERTIES OF FINITE FIELDS

The characteristic of a field is the least positive integer m such that $m = 1 + 1 + 1 + \cdots + 1 = 0$, and if no such m exists we say that the field has characteristic zero (which is the case for $\mathbb{R}$ for example). It turns out that if the characteristic is non-zero, it must be a prime p. For every prime p and positive integer r there is a finite field $\mathbb{F}_{p^r}$ of size $q = p^r$ and characteristic p, which is unique up to field isomorphism [59, 72]. The exponent r is known as the *degree* of the field over its prime subfield[1] [75]. If the characteristic p is an arbitrary prime number, we call the field *unrestricted*.

For every $a \in \mathbb{F}_q$, $a \neq 0$, then $a^{q-1} = 1$, implying the Frobenius endomorphism (also a consequence of Fermat's little theorem) $a^q = a$, which in turn permits us to write the multiplicative inverse of any non-zero element in the field as $a^{-1} = a^{q-2}$, since $a^{q-2}a = a^{q-1} = 1$. Every subfield of the field $\mathbb{F}_q$, of size $q = p^r$, has $p^{r'}$ elements with some r' dividing r, and for a given r' it is unique.

[1] Fields $\mathbb{F}_q$ where q is a power of a prime p, i.e., $q = p^r$, are known as Galois fields.

3.2 MODAL QUANTUM THEORY

Recently, Schumacher and Westmoreland [6, 45] and Chang et al. [7, 76] defined versions of quantum theory over *unrestricted* finite fields, which they call modal quantum theories (MQT) or Galois field quantum theories. Such theories retain several key quantum characteristics including notions of superposition, interference, entanglement, and mixed states, along with time evolution using invertible linear operators, complementarity of incompatible observables, exclusion of local hidden variable theories, impossibility of cloning quantum states, and the presence of natural counterparts of quantum information protocols such as superdense coding and teleportation. These modal theories are obtained by collapsing the Hilbert space structure over the field of complex numbers to that of a vector space over an *unrestricted* finite field. In the resulting structure, all non-zero vectors represent valid quantum states, and the evolution of a closed quantum system is described by *arbitrary* invertible linear maps.

Specifically, consider a one-qubit system with basis vectors $|0\rangle$ and $|1\rangle$. In conventional quantum theory, there exists an infinite number of states for a qubit of the form $\alpha_0 |0\rangle + \alpha_1 |1\rangle$, with α_0 and α_1 elements of the underlying field of complex numbers subject to the normalization condition $|\alpha_0|^2 + |\alpha_1|^2 = 1$. Moving to a finite field immediately limits the set of possible states as the coefficients α_0 and α_1 are now drawn from a finite set. In particular, in the field $\mathbb{F}_2 = \{0, 1\}$ of Booleans, there are exactly four possible vectors: the zero vector, the vector $|0\rangle$, the vector $|1\rangle$, and the vector $|0\rangle + |1\rangle = |+\rangle$. Since the zero vector is considered non-physical, a one-qubit system can be in one of only three states. The dynamics of these one-qubit states is realized by any invertible linear map, i.e., by any linear map that is guaranteed never to produce the zero vector from a valid state. There are exactly 6 such maps, and their matrix representations with respect to the standard

basis are:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \tag{3.2a}$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S^\dagger = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.2b}$$

For example,

$$S\left|0\right\rangle = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \left|+\right\rangle, \quad S\left|+\right\rangle = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left|0\right\rangle. \tag{3.3}$$

This set of maps is clearly quite impoverished compared to the full set of one-qubit unitary maps in conventional quantum theory. In particular, it does not include the Hadamard transformation. However, this set also includes non-unitary maps such as S and $S^\dagger$ that are not allowed in conventional quantum computation.

Measurement in the standard basis is straightforward: measuring $\left|0\right\rangle$ or $\left|1\right\rangle$ deterministically produces the same state while measuring $\left|+\right\rangle$ nondeterministically produces $\left|0\right\rangle$ or $\left|1\right\rangle$ with no assigned probability distribution. When measuring an arbitrary state $\left|\phi\right\rangle \in \{\left|0\right\rangle, \left|1\right\rangle, \left|+\right\rangle\}$ in other bases $\{\left|\psi_0\right\rangle, \left|\psi_1\right\rangle\}$, we first represent $\left|\phi\right\rangle$ as the linear combination of the basis vectors $\beta_0 \left|\psi_0\right\rangle + \beta_1 \left|\psi_1\right\rangle$, where β_0 and β_1 are elements in the field $\mathbb{F}_2$. If β_i is zero, measuring $\left|\phi\right\rangle$ is impossible to produce $\left|\psi_i\right\rangle$; otherwise, measuring $\left|\phi\right\rangle$ is possible to produce $\left|\psi_i\right\rangle$. Since only possibility and impossibility is predicted by the theory, modal quantum theories are named after these "modal" concepts.

Notice that the measurement process is complicated by the fact that the possibility to produce a basis vector $\left|\psi_i\right\rangle$ depending on the measurement basis. For example, measuring $\left|+\right\rangle$ is possible to produce $\left|0\right\rangle$ in the standard basis $\{\left|0\right\rangle, \left|1\right\rangle\}$ but is impossible to produce $\left|0\right\rangle$ in another basis $\{\left|+\right\rangle, \left|0\right\rangle\}$. In contrast, when measuring a state $\left|\phi\right\rangle$ in CQT, the probability to produce a basis vector $\left|\psi_i\right\rangle$ is completely determined by $\left|\psi_i\right\rangle$ and $\left|\phi\right\rangle$ no matter $\left|\psi_i\right\rangle$ is in which measurement basis.

This phenomena of the measurement basis dependence in CQT only exists when discussing quantum contextuality. Despite this kind of "supercontextuality" of MQT, its computational model, modal quantum computing (MQC), having "supernatural" computational power is also far from conventional quantum computing as we will describe next.

3.3 MODAL QUANTUM COMPUTING

To understand the computational implications of the modal quantum theory defined over the field $\mathbb{F}_2$ of Booleans, we developed a quantum computing model and established its correspondence to a classical model of logical programming with a feature that has quantum-like behavior [77]. In a conventional logic program, answers produced by different execution paths are collected in a sequence with *no* interference. However, in this modal quantum computing model over $\mathbb{F}_2$, these answers may interfere destructively with one another.

Our computations with this "toy" modal quantum theory showed that it possesses "supernatural" computational power. For example, one can solve a black box version of the UNIQUE-SAT problem [78] in a way that outperforms conventional quantum computing. The classical UNIQUE-SAT problem (also known as USAT or UNAMBIGUOUS-SAT) is the problem of deciding whether a given Boolean formula has a satisfying assignment, assuming that it has at most one such assignment [47]. This problem is, in a precise sense [48], just as hard as the general satisfiability problem and hence all problems in the NP complexity class. Our black-box version of the UNIQUE-SAT problem replaces the Boolean formula with an arbitrary black box. Solutions to this generalized problem can be used to solve an unstructured database search of size N using $O(\log N)$ black box evaluations by binary search on the database. This algorithm then outperforms the known asymptotic bound $O\left(\sqrt{N}\right)$ for unstructured database search in conventional quantum computing.

We can prove the unreasonable power of the arbitrary-function UNIQUE-SAT starting with a classical function $f\colon \mathsf{Bool}^n \to \mathsf{Bool}$ that takes n bits and returns at most one `true` result. To build

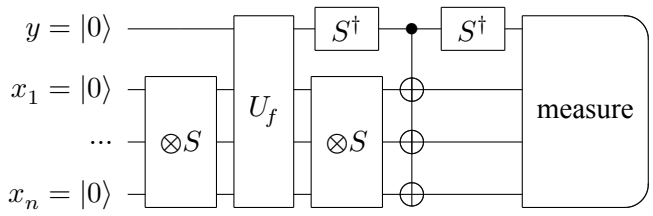


Figure 3.1: Circuit for the black box UNIQUE-SAT in modal quantum theory over the field $\mathbb{F}_2$. For further notation see text.

a quantum algorithm, f is first represented as the Deutsch quantum black box U_f with [9, 10]

$$U_f \left|y\right\rangle \left|\overline{x}\right\rangle = \left|y \oplus f\left(\overline{x}\right)\right\rangle \left|\overline{x}\right\rangle = \begin{cases} \left|y\right\rangle \left|\overline{x}\right\rangle & \text{if } f\left(\overline{x}\right) = \texttt{false}\,; \\ \left|\mathsf{not}\left(y\right)\right\rangle \left|\overline{x}\right\rangle & \text{if } f\left(\overline{x}\right) = \texttt{true}\,, \end{cases} \tag{3.4}$$

where $\overline{x}$ denotes a sequence $x_1, x_2, \ldots, x_n$ of n bits, $\oplus$ is exclusive disjunction, and 0 and 1 are identified as `false` and `true`, respectively. Then, we can give an algorithm (see Fig. 3.1) taking as input such a classical function that decides, deterministically and in a constant number of black box evaluations, whether f is satisfiable or not:

Case 1. *f is unsatisfiable; the measurement deterministically produces $\left|0\right\rangle \left|\overline{0}\right\rangle$.*

Proof. The state is initialized to $\left|0\right\rangle \left|\overline{0}\right\rangle$, with $\left|\overline{0}\right\rangle = \left|0\right\rangle \left|0\right\rangle \cdots \left|0\right\rangle$, i.e., the tensor product of n $\left|0\right\rangle$ states. As Eq. (3.3), applying the map S to each qubit in the second component of the state produces $\left|0\right\rangle \left|\overline{+}\right\rangle$ where $\left|\overline{+}\right\rangle$ denotes the sequence $\left|+\right\rangle \ldots \left|+\right\rangle$ of length n. Applying U_f to the entire state has no effect since U_f is the identity when f is unsatisfiable. Applying S to each qubit in the second component of the state produces $\left|0\right\rangle \left|\overline{0}\right\rangle$. Applying $S^\dagger$ to the first component leaves the state unchanged. As the first component of the state is 0, applying the map σ_0 (which is the identity) leaves the state unchanged. Applying $S^\dagger$ to the first component leaves the state unchanged. Measuring the state will deterministically produce $\left|0\right\rangle \left|\overline{0}\right\rangle$. □

Case 2. *f is satisfiable; the measurement produces some state other than $\left|0\right\rangle \left|\overline{0}\right\rangle$.*

Proof. Assume the function f is satisfiable at some input $a_1, a_2, ..., a_n$ denoted $\overline{a}$, and where $|\overline{a}\rangle = |a_1\rangle ... |a_n\rangle$. In the second step, the state becomes $|0\rangle |\overline{+}\rangle$ as above. We can write this state as $|0\rangle |\overline{a}\rangle + \Sigma_{\overline{x}\neq\overline{a}} |0\rangle |\overline{x}\rangle$. Applying U_f produces $|1\rangle |\overline{a}\rangle + \Sigma_{\overline{x}\neq\overline{a}} |0\rangle |\overline{x}\rangle$. We can rewrite this state as $|+\rangle |\overline{a}\rangle + \Sigma_{\overline{x}} |0\rangle |\overline{x}\rangle = |+\rangle |\overline{a}\rangle + |0\rangle |\overline{+}\rangle$, where the summation is now over all vectors (notice that $|0\rangle |\overline{a}\rangle + |0\rangle |\overline{a}\rangle$ is the zero vector). Applying S to each qubit in the second component produces $|+\rangle \left|\overline{S(a)}\right\rangle + |0\rangle \left|\overline{0}\right\rangle$. Applying $S^\dagger$ to the first component produces: $|1\rangle \left|\overline{S(a)}\right\rangle + |0\rangle \left|\overline{0}\right\rangle$. Applying control-not gate, which applying σ_0 or σ_x on the second component depending on the first component of the state, and produces

$$|1\rangle \left(\sigma_x \left|\overline{S(a)}\right\rangle\right) + |0\rangle \left(\sigma_0 \left|\overline{0}\right\rangle\right) = |1\rangle \left|\overline{\mathsf{not}\,(S(a))}\right\rangle + |0\rangle \left|\overline{0}\right\rangle \ . \tag{3.5}$$

Applying $S^\dagger$ to the first component produces $|+\rangle \left|\overline{\mathsf{not}\,(S(a))}\right\rangle + |0\rangle \left|\overline{0}\right\rangle$. For the measurement of $|+\rangle \left|\overline{\mathsf{not}\,(S(a))}\right\rangle + |0\rangle \left|\overline{0}\right\rangle$ to be guaranteed to never be $|0\rangle \left|\overline{0}\right\rangle$, we need to verify that $|+\rangle \left|\overline{\mathsf{not}\,(S(a))}\right\rangle$ has one occurrence $|0\rangle \left|\overline{0}\right\rangle$. This can be easily proved as follows. Since each a_i is either 0 or 1, then each $S(a_i)$ is either $+$ or 1, and hence each $\mathsf{not}\,(S(a_i))$ is either $+$ or 0. The result follows since any state with a combination of $+$ and 0, when expressed in the standard basis, would consist of a superposition containing the state $|0 ...\rangle$. □

Chapter 4

QUANTUM THEORIES AND COMPUTING OVER COMPLEXIFIED FINITE FIELDS

4.1 DISCRETE QUANTUM THEORY (I)

4.1.1 COMPLEXIFIED FINITE FIELDS

Our next objective is to develop more realistic discrete quantum theory variants that exclude "supernatural" algorithms such as the one presented above. Our first such plausible framework [79] is based on complexifiable finite fields. To incorporate complex numbers for quantum amplitudes, we exploit the fact that the polynomial $x^2 + 1$ is *irreducible* ($x^2 + 1 = 0$ has no solution) over a prime field $\mathbb{F}_p$ with p odd if and only if p is of the form $4\ell + 3$, with ℓ a non-negative integer [59, 72, 74]. For example, when $p = 3$, x could be 0 or ± 1. Since $0^2 + 1 \neq 0$ and $(\pm 1)^2 + 1 \neq 0$, none of the element in $\mathbb{F}_3$ solves $x^2 + 1 = 0$, and $x^2 + 1$ is irreducible over $\mathbb{F}_3$. In contrast, $2^2 + 1 = 0$ over $\mathbb{F}_5$ so that $x^2 + 1$ is reducible.

Since $x^2 + 1 = 0$ has no solution in any field $\mathbb{F}_p$ with $p = 4\ell + 3$, we can extend $\mathbb{F}_p$ to a field $\mathbb{F}_{p^2}$ whose elements can be viewed as discrete complex numbers with the real and imaginary parts in $\mathbb{F}_p$. Therefore, every element in $\mathbb{F}_{p^2}$ can be expressed as $a + b\mathrm{i}$ with $a, b \in \mathbb{F}_p$, and a $\mathbb{F}_{p^2}$ is called a complexified finite field. Since the multiplicative group of any finite field is cyclic [59], there is a generator $g \in \mathbb{F}_{p^2}$ such that every non-zero element $a + b\mathrm{i}$ can also be represented as the power of a generator, i.e., $a + b\mathrm{i} = g^j$ for some j. For example, $1 - \mathrm{i}$ is a generator in $\mathbb{F}_{3^2}$ so another element

Table 4.1: Generators in $\mathbb{F}_{3^2}$

j	0	1	2	3	4	5	6	7
$(1+\mathrm{i})^j$	1	$1+\mathrm{i}$	$-\mathrm{i}$	$1-\mathrm{i}$	-1	$-1-\mathrm{i}$	i	$-1+\mathrm{i}$
$(1-\mathrm{i})^j$	1	$1-\mathrm{i}$	i	$1+\mathrm{i}$	-1	$-1+\mathrm{i}$	$-\mathrm{i}$	$-1-\mathrm{i}$
$(-1+\mathrm{i})^j$	1	$-1+\mathrm{i}$	i	$-1-\mathrm{i}$	-1	$1-\mathrm{i}$	$-\mathrm{i}$	$1+\mathrm{i}$
$(-1-\mathrm{i})^j$	1	$-1-\mathrm{i}$	$-\mathrm{i}$	$-1+\mathrm{i}$	-1	$1+\mathrm{i}$	i	$1-\mathrm{i}$

$1+\mathrm{i} \in \mathbb{F}_{3^2}$ can be expressed as $1+\mathrm{i} = 1-3\mathrm{i}-3+\mathrm{i} = (1-\mathrm{i})^3$. All possible choices of generators in $\mathbb{F}_{3^2}$ are listed in Table 4.1.

In Table 4.1, one can notice that $(a+b\mathrm{i})^3 = a-b\mathrm{i}$. In general, the p-th power $(a+b\mathrm{i})^p = a-b\mathrm{i}$ is called the *Frobenius automorphism* acts like complex conjugation $(a+b\mathrm{i})^* = a-b\mathrm{i}$ [49, 59, 72]. Then, we define the *field norm* $\mathsf{N}(\cdot) : \mathbb{F}_{p^2} \to \mathbb{F}_p$ as an element $a+\mathrm{i}b$ multiplying its complex conjugation $(a+b\mathrm{i})^*$ [80],

$$\mathsf{N}(a+\mathrm{i}b) = (a+b\mathrm{i})(a+b\mathrm{i})^* = (a+b\mathrm{i})^{p+1} = a^2+b^2\,, \tag{4.1}$$

where the square root in the usual definition of the norm is avoided because, unlike the continuous case, the square root does not always exist, and the field norm of an element $\mathsf{N}(\cdot)$ should be the direct counterpart of the norm-squared $|\cdot|^2$ in the conventional quantum theory. For example, the field norm of every generator g in Table 4.1 is the same number $\mathsf{N}(g) = g^{3+1} = -1$. In fact, these four generators are the only elements in $\mathbb{F}_{3^2}$ whose field norm is $-1 \in \mathbb{F}_3$. Generally, given any $c \in \mathbb{F}_p$, let $\mathsf{N}^{-1}(\{c\})$ denote the set of elements whose field norm is c, i.e., $\mathsf{N}^{-1}(\{c\}) = \left\{\alpha \in \mathbb{F}_{p^2} \,\middle|\, \mathsf{N}(\alpha) = c\right\}$. The set $\mathsf{N}^{-1}(\{c\})$ is the discrete analog of phase-equivalence under the modulus-preserving transformation $z \to \mathrm{e}^{\mathrm{i}\phi}z$, and the number of its elements is characterized by the following proposition.

Proposition 4.1. *Given any $c \in \mathbb{F}_p$, the number of elements in $\mathsf{N}^{-1}(\{c\})$ is $p+1$, i.e., there are always $p+1$ elements in $\mathbb{F}_{p^2}$ whose field norm is c.*

Proof. To prove Proposition 4.1, we start by proving $\mathsf{N}^{-1}(\{c\})$ is non-empty. Consider a special case of the field norm $\mathsf{N}(.)$, namely the real quadratic map $Q(e) = e^2$ taking an arbitrary element $e \in \mathbb{F}_p$ to its square in the field. Since $(\pm 1)^2 = 1$, the image of $Q(e)$ has only $\frac{p+1}{2}$ elements in $\mathbb{F}_p$, including the zero element. We let A be the image of the map $Q(e)$ in $\mathbb{F}_p$, and note that the set A_c resulting from displacing an element $x = b^2$ of A to $c - x = c - b^2$ with $c \in \mathbb{F}_p$ also has $\frac{p+1}{2}$ elements because the result is simply a cyclic shift of element labels. We now observe that for any non-zero $c \in \mathbb{F}_p$, the sum of the elements in two sets A and A_c is $\frac{p+1}{2} + \frac{p+1}{2} = p+1$, which is greater than the size p of $\mathbb{F}_p$, and so there must be at least one common element such that $a^2 = c - b^2$. Thus, every element $c \in \mathbb{F}_p$ is the field norm of some element $\alpha = a + b\mathrm{i} \in \mathbb{F}_{p^2}$ such that $\mathsf{N}(\alpha) = a^2 + b^2 = c$, and $\mathsf{N}^{-1}(\{c\})$ is non-empty.

We then want to show for all non-zero $c \in \mathbb{F}_p$, the size of $\mathsf{N}^{-1}(\{c\})$ is always the same. Given a particular non-zero $c_0 \in \mathbb{F}_p$ and $\alpha_0 \in \mathbb{F}_{p^2}$ with $\mathsf{N}(\alpha_0) = c_0$, consider the map $f(\alpha) = \alpha_0 \alpha$. When $\mathsf{N}(\alpha) = 1$, we have [72]

$$\mathsf{N}(f(\alpha)) = \mathsf{N}(\alpha_0 \alpha) = \mathsf{N}(\alpha_0)\,\mathsf{N}(\alpha) = c_0 \tag{4.2}$$

so that $f(\alpha) \in \mathsf{N}^{-1}(\{c_0\})$. Since $\mathsf{N}(a + b\mathrm{i}) = 0$ only for $a = b = 0$, α_0 is non-zero, and f is a bijection between $\mathsf{N}^{-1}(\{1\})$ and $\mathsf{N}^{-1}(\{c_0\})$. This means the number of elements in $\mathsf{N}^{-1}(\{1\})$ and $\mathsf{N}^{-1}(\{c_0\})$ is the same. Because c_0 can be any non-zero element, the number of elements in the equivalence classes $\mathsf{N}^{-1}(\{c\})$ is always the same.

We can now compute the size of the equivalence class of complex unit-modulus phases corresponding to the Hopf fibration circle. Since $\mathbb{F}_{p^2}$ has $p^2 - 1$ non-zero values, and the map $\mathsf{N}(\alpha)$ distributes these equally across the domain of $p-1$ non-zero elements $c \in \mathbb{F}_p$, there are $\frac{p^2-1}{p-1} = p+1$

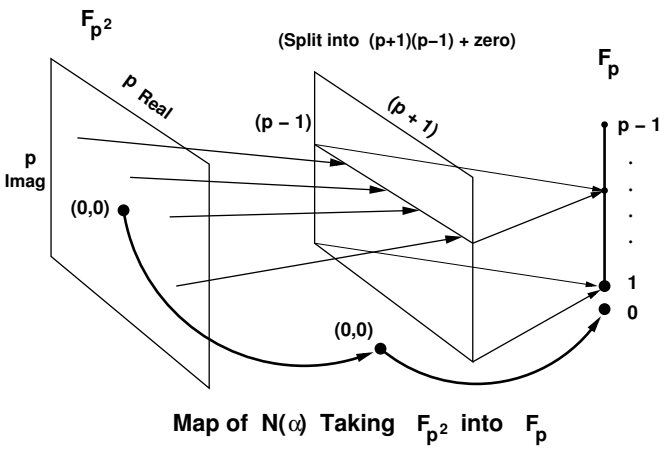


Figure 4.1: Sketch of the map from $\mathbb{F}_{p^2}$ to $\mathbb{F}_p$ using $\mathsf{N}(\alpha)$, showing the decomposition of $\mathbb{F}_{p^2}$ into the zero element $(0,0)$ and the $p^2-1=(p+1)(p-1)$ non-zero elements that map onto the $p-1$ non-zero elements of $\mathbb{F}_p$ with multiplicity $p+1$.

(non-zero) domain elements in $\mathbb{F}_{p^2}$ for each (non-zero) image element in $\mathbb{F}_p$. We illustrate this graphically in Figure 4.1. Thus, the Hopf circle always has the size $p+1$, corresponding essentially to a discrete projective line, and that is the size of each equivalence class of the map $\mathsf{N}(\alpha)$ for non-vanishing α, including the map to the unit norm value $c=1\in\mathbb{F}_p$. □

4.1.2 VECTOR SPACES

In this section, we want to build a theory of discrete vector spaces that approximates as closely as possible the features of the conventional quantum theory. Such a structure would ideally consist of the following: (i) a vector space over the field of complex numbers, and (ii) an inner product $\langle\Phi|\Psi\rangle$ associating to each pair of vectors a complex number, and satisfying the following properties:

(A) $\langle\Phi|\Psi\rangle$ is the complex conjugate of $\langle\Psi|\Phi\rangle$;

(B) $\langle\Phi|\Psi\rangle$ is conjugate linear in its first argument and linear in its second argument;

(C) $\langle\Psi|\Psi\rangle$ is always non-negative and is equal to 0 only if $|\Psi\rangle$ is the zero vector.

It turns out that a vector space defined over a finite field cannot have an inner product satisfying the properties above. However, we will introduce a Hermitian "dot product" satisfying some of those

properties.

We are interested in the vector space $\mathcal{H}$ of dimension D defined over the complexified field $\mathbb{F}_{p^2}$. Let $|\Psi\rangle = \begin{pmatrix}\alpha_0 & \alpha_1 & \cdots & \alpha_{D-1}\end{pmatrix}^T$ and $|\Phi\rangle = \begin{pmatrix}\beta_0 & \beta_1 & \cdots & \beta_{D-1}\end{pmatrix}^T$ represent vectors in $\mathcal{H}$, with numbers α_i and β_i drawn from $\mathbb{F}_{p^2}$, and where $(\cdot)^T$ is the transpose.

Definition 4.1 (Hermitian dot product)**.** Given vectors $|\Phi\rangle$ and $|\Psi\rangle \in \mathcal{H}$, it can be shown [49, 81] the Hermitian dot product is always reducible to the form

$$\langle\Phi|\Psi\rangle = \sum_{i=0}^{D-1} \beta_i^p \alpha_i \,. \tag{4.3}$$

If we use $\langle\Phi|$ to represent the dual vector of $|\Phi\rangle$, i.e., the conjugate transport of $|\Phi\rangle$, then the matrix multiplication of $\langle\Phi|$ and $|\Phi\rangle$ results a 1×1 matrix whose only entry is their Hermitian dot product $\langle\Phi|\Psi\rangle$.

Two vectors $|\Phi\rangle$ and $|\Psi\rangle \in \mathcal{H}$ are said to be orthogonal if $\langle\Phi|\Psi\rangle = 0$. This product satisfies conditions (A) and (B) for inner products but violates condition (C) since in every finite field there always exists a non-zero vector $|\Psi\rangle$ such that $\langle\Psi|\Psi\rangle = 0$. The reason is that addition in finite fields eventually "wraps around" (because of their cyclic or modular structure), allowing the sum of non-zero elements to be zero. The fraction of non-zero vectors satisfying $\langle\Psi|\Psi\rangle = 0$ decreases with the order p.

For any vector $|\Psi\rangle = \begin{pmatrix}\alpha_0 & \alpha_1 & \cdots & \alpha_{D-1}\end{pmatrix}^T$, the Hermitian dot product $\langle\Psi|\Psi\rangle$ is equal to $\sum_{i=0}^{D-1} \mathsf{N}(\alpha_i)$, which is the sum of the field norms for the complex coefficients. For convenience, we now extend the field norm to include vector arguments by defining

$$\mathsf{N}(|\Psi\rangle) = \langle\Psi|\Psi\rangle = \sum_{i=0}^{D-1} \mathsf{N}(\alpha_i) \,. \tag{4.4}$$

Although the field norm of a vector can vanish for non-vanishing vectors, if a vector $|\Psi\rangle$ has a non-vanishing field norm c, then $|\Psi\rangle$ can be normalized by utilizing its field norm. Recalled

in Sec. 4.1.1, we defined $\mathsf{N}^{-1}(\{c\})$ to be the set of elements whose field norm is c. Given any $\alpha \in \mathsf{N}^{-1}(\{c\})$, the field norm of $\frac{|\Psi\rangle}{\alpha}$ is

$$\mathsf{N}\left(\frac{|\Psi\rangle}{\alpha}\right) = \frac{\mathsf{N}(|\Psi\rangle)}{\mathsf{N}(\alpha)} = \frac{c}{c} = 1\,, \tag{4.5}$$

i.e., $\frac{|\Psi\rangle}{\alpha}$ is normalized. However, since the size of $\mathsf{N}^{-1}(\{c\})$ is $p+1$, we cannot identify a "unique" way to normalize any given vector.

The similar problem has already happened in conventional quantum theory. For example, assume we want to normalize $|\Psi\rangle = |0\rangle + |1\rangle$. Its inner product with itself is $\langle\Psi|\Psi\rangle = 2$. Since $\left(\pm\sqrt{2}\right)^2 = 2$, both $\frac{|\Psi\rangle}{\sqrt{2}}$ and $\frac{|\Psi\rangle}{-\sqrt{2}}$ are normalized and representing the same state as $|\Psi\rangle$. In this case, we systematically choose the state divided by the positive square root as "the" normalized vector of $|\Psi\rangle$ in conventional quantum theory, and the positive square root function is called the *principal* branch of $\sqrt{w}$ [82]. In the discrete case, we can also systematically choose the *principal inverse field norm* by utilizing a generator $g \in \mathbb{F}_{p^2}$ discussed in Sec. 4.1.1. Because g is a generator, any non-zero element $c \in \mathbb{F}_p \setminus \{0\}$ can be expressed as $g^{(p+1)k}$ where k is an integer and $0 \leq k < p-1$, so we can define the principal inverse field norm $\mathsf{N}^{-1}\left(g^{(p+1)k}\right)$ as g^k. For example, the inverse field norm over $\mathbb{F}_{3^2}$ with respect to the generator $1 - \mathrm{i}$ is shown in Table 4.2. Given the non-normalized state $|\Psi\rangle = |0\rangle + |1\rangle$, since its field norm is $\mathsf{N}(|\Psi\rangle) = \mathsf{N}(1) + \mathsf{N}(1) = -1$, it can be normalized as

$$\frac{|\Psi\rangle}{\mathsf{N}^{-1}(-1)} = \frac{|0\rangle + |1\rangle}{1 - \mathrm{i}} = (1+\mathrm{i})|0\rangle + (1+\mathrm{i})|1\rangle\,. \tag{4.6}$$

Table 4.2: Inverse field norm over $\mathbb{F}_{3^2}$ with respect to the generator $1-\mathrm{i}$

$c = g^{(p+1)k}$	$(p+1)\,k$	k	$\mathsf{N}^{-1}\left(g^{(p+1)k}\right) = g^k$
-1	4	1	$1-\mathrm{i}$
1	0	0	1

4.2 DISCRETE QUANTUM THEORY (I): IRREDUCIBLE DISCRETE D-DIMENSIONAL STATES

In the one-qubit state with coefficients in $\mathbb{F}_{p^2}$, the discrete analog of the Bloch sphere is constructed by exact analogy to the continuous case: we first require that the coefficients of the single qubit basis obey

$$\mathsf{N}\left(|\psi_1\rangle\right) = \mathsf{N}\left(\alpha_0\right) + \mathsf{N}\left(\alpha_1\right) = 1 \tag{4.7}$$

in the discrete field. We show that there are $p\left(p^2-1\right)$ such values later in the general theorem, Proposition 4.2. Given this requirement, which is similar in form to the conservation of probability, but not as useful due to the lack of orderable probability values, we can immediately conclude that the discrete analog of the Hopf fibration is again

$$\begin{aligned} X &= 2\,\mathrm{Re}\;\alpha_0\alpha_1^* = 2x_0x_1 + 2y_0y_1\,, \\ Y &= 2\,\mathrm{Im}\;\alpha_0\alpha_1^* = 2x_1y_0 - 2x_0y_1\,, \\ Z &= \mathsf{N}\left(\alpha_0\right) - \mathsf{N}\left(\alpha_1\right) = {x_0}^2 + {y_0}^2 - {x_1}^2 - {y_1}^2\,. \end{aligned} \tag{4.8}$$

but now with all computations in $(\mathrm{mod}\ p)$. At this point, one can simply write down all possible discrete values for the complex numbers (α_0, α_1) satisfying Eq. (4.7) and enumerate those that project to the same value of (X, Y, Z). This equivalence class is the discrete analog of the circle in the complex plane that was eliminated in the continuous case. In Proposition 4.1, we show that $p+1$ discrete

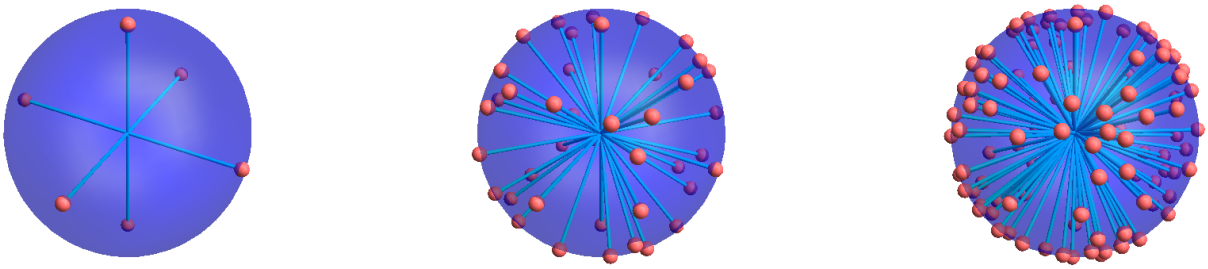


Figure 4.2: Schematically normalized plots of the elements of the discrete Bloch sphere, the irreducible single-qubit (two-dimensional) state vectors with unit norm over the field $\mathbb{F}_{p^2}$. We show the results for $p = 3$, 7, and 11. For example, in $\mathbb{F}_{3^2}$, there are 24 vectors of unit norm, but only the 6 inequivalent classes appear in the plot. The $p + 1 = 4$ equivalent vectors in each class differ only by a complex discrete phase.

values of (α_0, α_1) with unit norm map to the same point under the Hopf map Eq. (4.8); we may think of these as discrete circles or projective lines of equivalent, physically indistinguishable, complex phase. The surviving $p\,(p-1)$ values of (α_0, α_1) correspond to irreducible physical states of the discrete single qubit system. Thus, for example, choosing the underlying field to be $\mathbb{F}_{3^2}$, there are exactly 6 single-qubit state vectors to populate the Bloch sphere; the four equivalent phase-multiples mapping to each of the six points on the $\mathbb{F}_{3^2}$ Bloch sphere are collapsed and regarded as physically indistinguishable. In Figure 4.2, we plot the irreducible states on the Bloch sphere for $p = 3$, 7, and 11. Note that the Cartesian lengths of the real vectors corresponding to the points on the Bloch sphere vary considerably due to the nature of discrete fields; we have artificially normalized them to a "continuous world" unit radius sphere for conceptual clarity.

4.2.1 COUNTING STATES ON THE DISCRETE BLOCH SPHERE

We have the unique opportunity in the finite-field approach to quantum computing to precisely identify and enumerate the physical states. In the conventional theory, as we have seen in Sec. 2.1.3, we employ a generalized Hopf fibration on the normalized states to project out a circle of phase-

equivalent states, yielding the generalized Bloch sphere.

In the introduction to this subsection, we sketched the counting of the irreducible single-qubit discrete states. To count the number of inequivalent discrete states for the general n-qubit case with coefficients in $\mathbb{F}_{p^2}$, we first must find the set of unit-norm states, and then determine the equivalence classes of unit-norm states under discrete phase transformations; we can then enumerate the list of states on the discrete generalized Bloch sphere. By executing computer searches of these spaces, we discovered a hypothesis for a closed-form solution for the counting of the states and find a rigorous proof of the enumeration.

This process of describing the discrete D-dimensional irreducible states can again be understood geometrically by following the discrete analog of the Hopf fibration. First, we construct the discrete version of the quadratic unit-length form that automatically annihilates the distinction among states differing only by a discrete phase,

$$\hat{a} = \left(\mathsf{N}\left(\alpha_i\right), ..., \sqrt{2}\,\mathrm{Re}\;\alpha_i \alpha_j^*, ..., \sqrt{2}\,\mathrm{Im}\;\alpha_i \alpha_j^*, ... \right) , \tag{4.9}$$

where

$$\hat{a} \cdot \hat{a} = \left(\sum_{i=0}^{D-1} \mathsf{N}\left(\alpha_i\right) \right)^2 = 1 . \tag{4.10}$$

From Proposition 4.1, we know that $p+1$ elements of this discrete $\mathbf{S}^{2D-1}$ structure map to the *same point* in $\hat{a}$. Each set of $p+1$ redundant points is, geometrically speaking, the *discrete Hopf fibration circle* living above each *irreducible* point of the D-dimensional state description. These $p+1$ points are interpretable as the p finite points plus the single point at infinity of the projective discrete line (see, e.g., [83]).

The next part of this argument is the determination of the unit-norm states, effectively the space of allowed discrete partitions of unity; we cannot exactly call these "probability-conserving" sectors of the state coefficients since we do not have a well-defined notion of probability, but we do have a

well-defined notion of the partition of unity. Compared to the total number p^{2D} of possible complex integer state vectors that could be chosen, the number of unit-norm states is given by the following proposition. This unit-norm state structure is the discrete analog of $\mathbf{S}^{2D-1}$.

Proposition 4.2. *The number of unit-norm states described by a D-dimensional vector* $(\alpha_0, \ldots, \alpha_{D-1})$ *with coefficients* $\alpha_i \in \mathbb{F}_{p^2}$ *is* $p^{D-1}\left(p^D - (-1)^D\right)$.

Proof. Proposition 11.27 in Grove [49] provides the count of the zero-norm states $\zeta(D,p) = p^{D-1}\left(p^D + (-1)^D(p-1)\right)$. Since there are p^2 elements $\alpha \in \mathbb{F}_{p^2}$, we must have $\left(p^2\right)^D = p^{2D}$ possible values of a D-dimensional vector $(\alpha_0, \ldots, \alpha_{D-1})$. There are $p^2 - 1$ non-zero values of $\alpha \in \mathbb{F}_{p^2}$, and we showed in Proposition 4.1 that $\mathsf{N}(\alpha)$ maps exactly $p+1$ values in that set to each of the $p-1$ non-zero values in $\mathbb{F}_p$. Therefore, the *unit-norm case* has a count of domain elements that is $\frac{1}{p-1}$ of the total number of non-zero-norm cases,

$$\frac{p^{2D} - \zeta(D,p)}{p-1} = \frac{p^{2D} - p^{2D-1} - (-1)^D p^{D-1}(p-1)}{p-1} = p^{D-1}\left(p^D - (-1)^D\right). \tag{4.11}$$

□

Finally, we repeat the last step of the D-dimensional continuous Hopf fibration process for discrete D-dimensional states, eliminating the discrete set of $p+1$ equivalent points that map to the same point $\hat{a}$ on the generalized Bloch sphere. Dividing the tally $p^{D-1}\left(p^D - (-1)^D\right)$ of unit norm states by the $p+1$ elements of each phase-equivalent discrete circle, we find

$$\frac{p^{D-1}\left(p^D - (-1)^D\right)}{p+1} \tag{4.12}$$

as the total count of unique irreducible states in a discrete D-dimensional configuration. The resulting object is precisely the discrete version of $\mathbb{C}\mathbf{P}^{D-1}$, which we might call a *discrete complex projective space* or $\mathbf{D}\mathbb{C}\mathbf{P}^{D-1}$.

4.3 DISCRETE QUANTUM THEORY (I): GEOMETRY OF ENTANGLED STATES

To discuss entanglement, we consider a D-dimensional quantum system composed of n-qubit subsystems, i.e., $D = 2^n$ as usual. Without regard to uniqueness, an n-qubit state with discrete complex coefficients in $\mathbb{F}_{p^2}$ will have the total possible space of coefficients with dimension $p^{2\times 2^n}$ (including the null state). Imposing the condition of a length-one norm in $\mathbb{F}_p$, this number is reduced to $p^{2^n-1}\,(p^{2^n}-1)$. The ratio of all the states to the unit-norm states is asymptotically p:

$$\frac{p^{2^n+1}}{p^{2^n}-1} \to p\,, \tag{4.13}$$

so there are roughly p sets of coefficients, for any number of qubits n, that are discarded for each retained unit-length state vector. A factor of $p+1$ more states are discarded in forming the discrete Bloch sphere of irreducible states. Selected plots of the full space compared to both the unit-norm space and the irreducible space for a selection of complexified finite fields are shown in Figure 4.3 for 1, 2, 3, and 4 qubits.

4.3.1 UNENTANGLED VS ENTANGLED DISCRETE STATES

For a given p and the corresponding complexified field $\mathbb{F}_{p^2}$, the n-qubit discrete quantum states with coefficients in $\mathbb{F}_{p^2}$ can be classified by their degree of entanglement to a level of precision that is unavailable in the continuous theory. We look first at the unentangled n-qubit states, which are direct product states of the form

$$|\Psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_j\rangle \otimes \cdots \otimes |\psi_n\rangle\ . \tag{4.14}$$

Without regard to normalization, there are $\left(p^4\right)^n$ possible unentangled states out of the total of $p^{2\times 2^n}$ states noted above. When we normalize the individual product states to unit norm, the norm of the entire n-qubit state becomes the product of those unit norms and is automatically normalized to one.

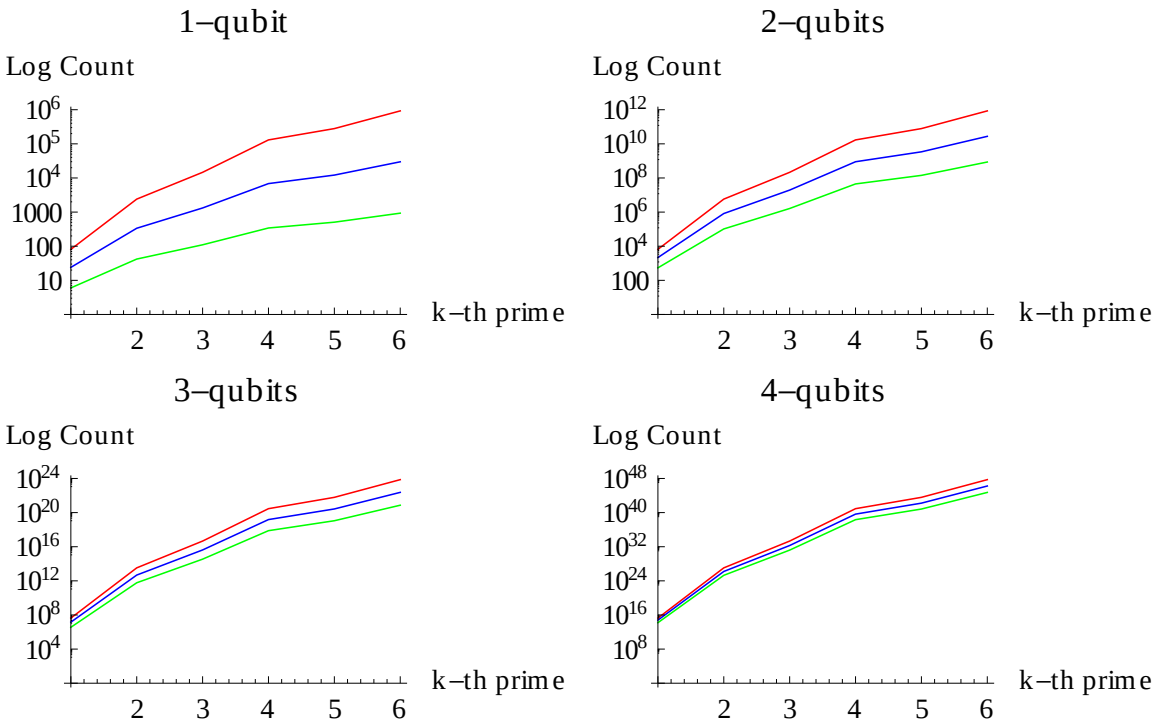


Figure 4.3: Logarithmic plot of the number of discrete unnormalized states (top, in red), vs the number of normalized discrete states (middle, in blue), vs the irreducible states (bottom, in green) for the first 6 $\mathbb{F}_{p^2}$-compatible primes, (3, 7, 11, 19, 23, 31), for the number of qubits 1, 2, 3, and 4.

We have already seen that each single-qubit normalized state in the tensor product Eq. (4.14) has precisely $p\,(p-1)$ irreducible components due to $D = 2$ case in Eq. (4.12).

4.3.2 COMPLETELY UNENTANGLED STATES AND THE DISCRETE BLOCH SPHERE

In effect, the irreducible states for unentangled n-qubit configurations reduce to a single Bloch sphere for each one-qubit component $|\psi_j\rangle$, and thus the whole set of states is defined by an n-tuple of discrete Bloch sphere coordinates. Since each Bloch sphere in $\mathbb{F}_{p^2}$ has $p\,(p-1)$ distinct irreducible components, we have

$$\textbf{Count of Unentangled States} = p^n\,(p-1)^n\ . \tag{4.15}$$

According to Eq. (4.12), we know that the total number of irreducible states (points in the generalized $\mathbf{D}\mathbb{C}\mathbf{P}^{2^n-1}$ Bloch sphere) for an n-qubit state is $\frac{p^{2^n-1}\left(p^{2^n}-1\right)}{p+1}$, and so the number of states containing some measure of entanglement is

$$\textbf{Count of Entangled States} = \frac{p^{2^n-1}\left(p^{2^n}-1\right)}{p+1} - p^n\left(p-1\right)^n \ . \tag{4.16}$$

Therefore, a very small fraction of the unit norm states are unentangled.

4.3.3 MAXIMAL ENTANGLEMENT

Equation (2.38) for $P_{\mathfrak{h}}$ includes a normalization factor $\frac{1}{n}$. In the discrete case, this normalization factor is undefined when $p \mid n$. Equation (2.38) also includes a summation of n terms. In the discrete case, certainly when $p \mid n$ but also in other cases, this summation may vanish in the field even if the individual summands are non-zero. These anomalies are irrelevant for the classification of unentangled states as this computation is performed by directly checking the possibility of direct decomposition into product states, disregarding equation (2.38).

For maximally entangled states, the purity calculation in conventional quantum mechanics using equation (2.38) produces 0. Given the above observations, in a discrete field, equation (2.38) may be undefined or may report a purity of 0 even for partially entangled states. For example, the normalized 5-qubit state $|\Psi\rangle = (1-\mathrm{i})\left(|00\rangle + |11\rangle\right) \otimes |000\rangle$ has $P_{\mathfrak{h}} = 0$ for $p = 3$, and is not maximally entangled because only the first two qubits are entangled. In the discrete case, we therefore check for maximally entangled states using the following equations [8],

$$\forall j, \forall \eta \in \{x, y, z\}\,, \left\langle \sigma_\eta^j \right\rangle^2 = 0\,, \tag{4.17}$$

which avoids the normalization factor and simply checks that each summand is 0, where $\left\langle \sigma_\eta^j \right\rangle$ is defined as $\left\langle \Psi \middle| \sigma_\eta^j \middle| \Psi \right\rangle$.

We now implement these procedures to enumerate the maximally entangled states for the specific cases for $n = 2, 3$ and compare these to the counts for product states. We have verified explicitly in Eq. (4.15) that the numbers of unit-norm product states for $n = 2$, $p = \{3, 7, 11, 19, ...\}$ are

$$(p+1)p^2(p-1)^2 = \{144, 14\,112, 145\,200, 2339\,280, ...\}\,, \tag{4.18}$$

and for general n, $(p+1)p^n(p-1)^n$. The irreducible state counts are reduced by $(p+1)$, giving

$$p^2(p-1)^2 = \{36, 1764, 12\,100, 116\,964, ...\}\,, \tag{4.19}$$

and in general for n-qubits, there are $p^n\,(p-1)^n$ instances of pure product states.

Performing the computation using equation (4.17), we find the numbers of maximally entangled states for two qubits to be

$$p\,(p^2-1)\,(p+1) = \{96, 2688, 15\,840, 136\,800, ...\}\ . \tag{4.20}$$

The irreducible state counts for maximal entanglement are reduced by $(p+1)$, giving, for $n = 2$,

$$p\,(p^2-1) = \{24, 336, 1320, 6840, ...\}\ . \tag{4.21}$$

For three qubits, there are $p^3\,(p^4-1)\,(p+1)$ (total) and $p^3\,(p^4-1)$ (irreducible) instances of pure maximally entangled states, while the general formula for 4-qubit states remains unclear.

Therefore, the ratio of maximally entangled to product states is

$$\frac{\textbf{Max entangled}}{\textbf{Product}} = \frac{p+1}{p\,(p-1)} \text{ and } \frac{(p^2+1)\,(p+1)}{(p-1)^2} \tag{4.22}$$

for $n = 2$ and 3, respectively.

4.4 DISCRETE QUANTUM COMPUTING (I)

Given a complexified finite field $\mathbb{F}_{p^2}$ and its Hermitian dot product (Eq. (4.3)) much of the structure of conventional quantum computing can be recovered. For example, the smallest field $\mathbb{F}_{3^2}$ is already rich enough to express the standard Deutsch-Jozsa algorithm [10, 36, 52], which requires only normalized versions of vectors or matrices with the scalars 0, 1, and -1. Similarly, other deterministic quantum algorithms (algorithms for which we may determine the outcome with certainty), such as Simon's [4, 36, 50] and Bernstein-Vazirani [4, 51], perform as desired. In the following subsection, we will present the discrete Deutsch algorithm as an example. However, this quantum computing model is still different from the conventional one. On one hand, algorithms such as Grover's search [4, 36, 53] will not work in the usual way because we lack (the notion of) ordered angles and probability in general. On the other hand, this computational model still leads to excessive computational power for the unstructured database search problem for certain database sizes.

4.4.1 DISCRETE DEUTSCH ALGORITHM

Although having no realistic application, the Deutsch algorithm is the first quantum algorithm which outperforms any possible classical algorithm for the Deutsch problem [4, 9, 10]. The Deutsch problem is to decide whether a function $f : \mathsf{Bool} \to \mathsf{Bool}$ is constant or balanced. As listed in Table 4.3, we have only 4 different f: 2 of them are constant while another 2 are balanced. Similar to our UNIQUE-SAT algorithm in Sec. 3.3, we start by representing f as a Deutsch black box U_f in the middle of the quantum circuit, Figure 4.4. To explain why this circuit solves the Deutsch problem, we then compute the state in each step explicitly and express them in the Dirac bracket notation and its matrix representation in the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

First, a 2-qubit pure state is initialized to $|\Phi_1\rangle = |1\rangle|0\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \otimes \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$.

Second, on both initialized qubits, we apply the Hadamard matrix $H = \frac{1}{\alpha}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$ over $\mathbb{F}_{p^2}$, where $\alpha = \mathsf{N}^{-1}(2)$ is the principal inverse field norm which is used to replace the square root $\sqrt{2}$ in

Table 4.3: All Possible $f : \mathsf{Bool} \to \mathsf{Bool}$.

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
`false`	`false`	`false`	`true`	`true`
`true`	`false`	`true`	`false`	`true`
constant or balanced?	constant	balanced	balanced	constant

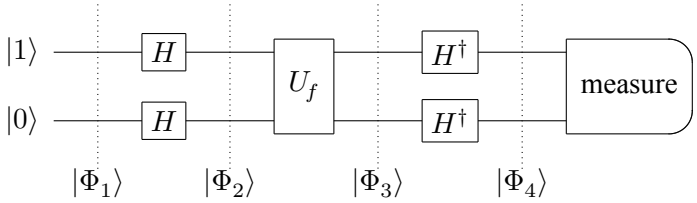


Figure 4.4: Quantum Circuit for the Deutsch Algorithm.

the conventional Hadamard matrix $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$ as discussed in Sec. 4.1.2. The second step produces

$$|\Phi_2\rangle = (H \otimes H)\,|\Phi_1\rangle = \left[\frac{1}{\alpha}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix}\right] \otimes \left[\frac{1}{\alpha}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}\right] = |-\rangle\,|+\rangle\ ,$$

where

$$|+\rangle = \frac{1}{\alpha}\begin{pmatrix}1\\ 1\end{pmatrix} = \frac{|0\rangle + |1\rangle}{\alpha}\,, \qquad |-\rangle = \frac{1}{\alpha}\begin{pmatrix}1\\ -1\end{pmatrix} = \frac{|0\rangle - |1\rangle}{\alpha}\,. \tag{4.23}$$

Third, the Deutsch black box U_f is applied to the state $|\Phi_2\rangle$. According to Eq. (3.4), these Deutsch black box will be used to apply exclusive disjunction on the first qubit, where we respectively identify `false` and `true` as 0 and 1 as usual. Since our first qubit is $|-\rangle$, the value $f(x)$ could

be moved outside as a phase no matter $f(x)$ is `false` or `true`:

$$\begin{aligned} U_f |-\rangle |x\rangle &= \tfrac{1}{\alpha} \left[|0 \oplus f(x)\rangle |x\rangle - |1 \oplus f(x)\rangle |x\rangle \right] \\ &= \begin{cases} \tfrac{1}{\alpha} \left[|0\rangle |x\rangle - |1\rangle |x\rangle \right] & \text{if } f(x) = \texttt{false} = 0\,; \\ \tfrac{1}{\alpha} \left[|1\rangle |x\rangle - |0\rangle |x\rangle \right] & \text{if } f(x) = \texttt{true} = 1 \end{cases} \\ &= (-1)^{f(x)} |-\rangle |x\rangle \ . \end{aligned} \tag{4.24}$$

Then, $|\Phi_3\rangle$ can be evaluated as follow:

$$\begin{aligned} |\Phi_3\rangle = U_f |-\rangle |+\rangle &= \tfrac{1}{\alpha} \left[U_f |-\rangle |0\rangle + U_f |-\rangle |1\rangle \right] \\ &= \tfrac{1}{\alpha} \left[(-1)^{f(0)} |-\rangle |0\rangle + (-1)^{f(1)} |-\rangle |1\rangle \right] \\ &= \begin{cases} (-1)^{f(0)} |-\rangle |+\rangle & \text{if } f(0) = f(1)\ ; \\ (-1)^{f(0)} |-\rangle |-\rangle & \text{if } f(0) \neq f(1)\ . \end{cases} \end{aligned} \tag{4.25}$$

Finally, $|\Phi_4\rangle$ can then be obtained by applying the Hermitian conjugate of Hadamard matrix $H^\dagger = \frac{1}{\alpha^*} \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right)$ on both qubits. If $f(0) = f(1)$, i.e., f is constant, we have

$$|\Phi_4\rangle = (-1)^{f(0)} \left[\frac{1}{\alpha^*} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\alpha} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right] \otimes \left[\frac{1}{\alpha^*} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\alpha} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right] = (-1)^{f(0)} |1\rangle |0\rangle \ ;$$

if $f(0) \neq f(1)$, i.e., f is balanced, we have

$$|\Phi_4\rangle = (-1)^{f(0)} \left[\frac{1}{\alpha^*} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\alpha} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right] \otimes \left[\frac{1}{\alpha^*} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\alpha} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right] = (-1)^{f(0)} |1\rangle |1\rangle \ .$$

Hence, we can decide whether f is constant or balanced by measuring $|\Phi_4\rangle$ in the computational basis.

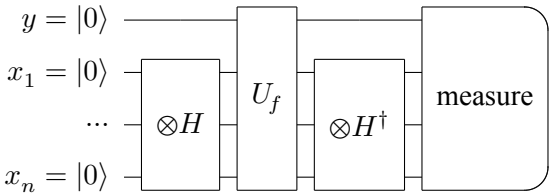


Figure 4.5: Circuit for the black box UNIQUE-SAT in discrete quantum computing.

4.4.2 PARTIAL UNIQUE-SAT ALGORITHM

It is possible, in some situations, to exploit the cyclic behavior of the field to creatively cancel probability amplitudes and solve problems with what again appears to be "supernatural" efficiency. We illustrate this behavior with the algorithm in Fig. 4.5, which is a variant of the one in Fig. 3.1. Unlike the modal quantum algorithm, the new algorithm does not always succeed deterministically using a constant number of black box evaluations. We can, however, show that supernatural behavior occurs if the characteristic p of the field divides $2^N - 1$. For a database of fixed size N, matching the conditions becomes less likely as the size of the field increases. Nevertheless, for a *given* field, it is always possible to expand any database with dummy records to satisfy the divisibility property. Physically, we are taking advantage of additional interference processes that happen because of the possibility of "wrapping around" due to modular arithmetic. We do not know, in general, whether this version of discrete quantum computing actually enables the rapid solution of NP-complete problems.

4.5 DISCRETE QUANTUM THEORY (II)

4.5.1 INNER PRODUCT SPACE

We next discuss an approach using finite complexifiable fields that conditionally resolves the inner product condition (C) discussed in Sec. 4.1.2, which is violated by the theory just presented. A possible path is suggested by the work of Reisler and Smith [84]. The general idea is that while the cyclic properties of arithmetic in finite fields make it impossible to *globally* obtain the desired

properties of the conventional Hilbert space inner product, it *is* possible to recover them *locally*, thereby restoring, with some restrictions, all the usual properties of the inner product needed for conventional quantum mechanics and conventional quantum computing. As the size of the discrete field becomes large, the size of the locally valid computational framework grows as well, leading to the *effective emergence of conventional quantum theory*. We next briefly outline such a context for local orderable subspaces of a finite field and introduce an improvement on the original method [84] suggested by recent number theory resources [85].

Let us first note that the range of the quadratic map, $\left\{x^2 \bmod p \;\middle|\; x \in \mathbb{F}_p\right\}$, is always one-half of the non-zero elements of $\mathbb{F}_p$, and is the set of elements with square roots in the field. This is the set of *quadratic residues*, and the complementary set (the other half of the non-zero field elements) is the set of *quadratic non-residues*. For example in $\mathbb{F}_7$, the elements $\{1, 2, 4\}$ are considered positive as they have the square roots $\{1, 3, 2\}$ respectively; the remaining elements $\{3, 5, 6\}$ do not have square roots in the field. What is interesting is that if we have a uninterrupted sequence of numbers that are all quadratic residues, then we can define a *transitive order*, with $a > c$ if $a > b$ and $b > c$, provided $a - b$, $b - c$, and $a - c$ are all quadratic residues.

As a concrete example, consider a finite field in which the sequential elements 0, 1, 2, 3, …, and $k - 1$ are all quadratic residues (including 0). Then any sequence of odd length k and centered around an arbitrary $x \in \mathbb{F}_p$, i.e., $S_x(k) = x - \frac{k-1}{2}$, …, $x - 2$, $x - 1$, x, $x + 1$, $x + 2$, …, $x + \frac{k-1}{2}$, is *transitively ordered*. Indeed, we have $(x + 1) - x = 1$ which is a quadratic residue and hence $x + 1 > x$. Similarly, $x - (x - 1) = 1$ and hence $x > x - 1$. Also, $(x + 1) - (x - 1) = 2$ which is a quadratic residue and hence $x + 1 > x - 1$. Clearly, this process may be continued to show that the sequence $S_x(k)$ is transitively ordered. We can construct examples using the sequence A000229 in the encyclopedia of integer sequences [85][1]. The n-th element of that sequence (which must be

[1]For computational purposes, this sequence is preferable to the one proposed by Reisler and Smith [84] because it produces smaller primes. Their work showed that a sufficient condition on finite fields to produce sequences of quadratic residues is to further constrain the underlying prime numbers to be of the form $8\prod_{i=1}^{m} q_i - 1$, where q_i is the i-th odd

Table 4.4: Number k of transitively ordered elements for a given field $\mathbb{F}_p$.

p	3	7	23	71	311	479	1559	5711	10559	18191	…
k	**2**	**3**	**5**	**7**	**11**	**13**	**17**	**19**	**23**	**29**	…
$\pi(k)$	1	2	3	4	5	6	7	8	9	10	…

prime) is the least number such that the n-th prime is the *least* quadratic non-residue for the given element. The first few elements of this sequence are listed in the top row of Table 4.4. The next row lists the number k of transitively ordered consecutive elements in that field, and $\pi(k)$ in the bottom row is the prime counting function (the number of primes up to k).

As an example, consider the field $\mathbb{F}_{23}$. Looking at the squares of the numbers $\mathbb{F}_{23} = \{0, \ldots, 22\}$ modulo 23, we find the 2-centered uninterrupted sequence $S_2(5) = \{0, 1, 2, 3, 4\}$, followed by 5, which is both the smallest quadratic non-residue and the size of the uninterrupted sequence of quadratic residues (including 0) of interest. In particular, it is possible to construct a total order for the elements $S_0(5) = \{-2, -1, 0, 1, 2\}$ in the fields $\mathbb{F}_{23}$, $\mathbb{F}_{71}$, $\mathbb{F}_{311}$, etc., but not in the smaller fields $\mathbb{F}_3$ and $\mathbb{F}_7$.

Given a D-dimensional vector space over $\mathbb{F}_{p^2}$ where p is one of the primes above, it is possible to define a *region* over which an inner product and norm can be identified. Let the length of the sequence of quadratic residues be k. The region of interest includes all vectors $|\Psi\rangle = \sum_{i=0}^{D-1} \alpha_i |i\rangle = \begin{pmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_{D-1} \end{pmatrix}^T$, for which $D < p - \frac{k-1}{2}$ and each α_i satisfies

$$D\,\mathsf{N}(\alpha_i) = D\,(a_i^2 + b_i^2) \leq \frac{k-1}{2}\,, \tag{4.26}$$

with a_i and b_i drawn from the set $S_0(k)$. Consider, for example, $\mathbb{F}_{311^2}$ ($p = 311$, $k = 11$). We find that we can trade off the dimension D of the vector space against the range of probability amplitudes

prime. While all such primes are of the form $4\ell + 3$, the set is severely restricted to astronomical numbers because the first few such primes are 7, 23, 839, 9239, 2042039, …

Table 4.5: Allowed probability amplitudes for different vector space dimensions D and $k = 11$.

	allowed probability amplitudes $F^D(k)$
$D = 1$	$F^1(11) = \{0, \pm 1, \pm 2, \pm \mathrm{i}, \pm 2\mathrm{i}, (\pm 1 \pm \mathrm{i}), (\pm 1 \pm 2\mathrm{i}), (\pm 2 \pm \mathrm{i})\}$
$D = 2$	$F^2(11) = \{0, \pm 1, \pm \mathrm{i}, (\pm 1 \pm \mathrm{i})\}$
$D = 3$	$F^3(11) = \{0, \pm 1, \pm \mathrm{i}\}$
$D = 4$	$F^4(11) = \{0, \pm 1, \pm \mathrm{i}\}$
$D = 5$	$F^5(11) = \{0, \pm 1, \pm \mathrm{i}\}$
$D \geq 6$	$F^D(11) = \{0\}$

available for each α_i in Table 4.5.

We can now verify, by using Table 4.5, that for any vector $|\Psi\rangle$ in the selected region the value of $\langle\Psi|\Psi\rangle$ is ≥ 0 and vanishes precisely when $|\Psi\rangle$ is the zero vector. Thus, in the selected region, condition (C) is established. Although the set of vectors defined over that region is not closed under addition, and hence the set is not a vector subspace, we can still have a theory by restricting our computations. In other words, *as long as our computation remains within the selected region*, we may pretend to have an inner product space. The salient properties of conventional quantum mechanics emerge, but the price to be paid is that the state space is no longer a vector space. This is basically a rigorous formulation of Schwinger's intuition (See, Chapter 1, Section 1.16. in [86]).

Readers with backgrounds in computer science or numerical analysis will notice, significantly, that this model for discrete quantum computing is reminiscent of practical computing with a classic microprocessor having only integer arithmetic and a limited word length. We cannot perform a division having a fractional result at all since there are no fractional representations; we do have the basic constants zero and one, as well as positive and negative numbers, but multiplications or additions producing results outside the integer range wrap-around modulo the word length and typically yield nonsense. This implies that, for the local discrete model, we must accept an operational world

view that *has no awareness of the value of* p, and depends on having set up in advance an environment with a field size, analogous to the word size of a microprocessor, that happily processes *any* calculation we are prepared to perform. This is the key step, though it may seem strange because we are accustomed to arithmetic with real numbers: we list the calculations that must be performed in our theory, discover an *adequate size of the processor word* —implying a possibly ridiculously large value of p chosen as described above —and from that point on, we calculate necessarily valid values within that processor, never referring in any way to p itself in the sequel.

4.5.2 CARDINAL PROBABILITY

The final issue that must be addressed in the discrete theory put forward in Section 4.5.1 concerns measurement. To recap, within the theory, states are D-dimensional vectors with complex discrete-valued amplitudes drawn from a totally-ordered range, $F^D(k)$, in the underlying finite field. These states possess, by construction, having field norms in the non-negative integers, all in the ordered range of Eq. (4.26), and hence potentially produce probabilities that can be ordered. Our point is that, although the mathematical framework of conventional quantum mechanics relies on infinite precision probabilities, it is impossible in practice to measure exact equality of real numbers —we can only achieve an approximation within measurement accuracy. Significantly, when we use finite fields, this measurement accuracy will be encoded in the size of the finite field used for measurements.

Given a D-dimensional Hilbert space in conventional quantum theory, although we can measure the probability for every eigenprojector of an observable as discussed in Sec. 2.2, our previous quantum circuits in Figures 3.1, 4.4, and 4.5 always measure in the computational basis $\{|0\rangle, ..., |i\rangle, ..., |D-1\rangle\}$. Indeed, it is sufficient to only consider measuring a quantum circuit in the computational basis because measuring in another basis is the same as applying a quantum gate and measuring in the computational basis. In this situation, the Born rule for pure states, Eq. (2.29),

Table 4.6: Field norms and probabilities for one-qubit states $|\Psi\rangle$ in $F^2(11)$.

$\lvert\Psi\rangle$	$\mathsf{N}(\langle 0\lvert\Psi\rangle)$	$\mathsf{N}(\langle 1\lvert\Psi\rangle)$	$\mathsf{N}(\lvert\Psi\rangle)$	$\mu_\Psi^{\mathrm{C}}(0)$	$\mu_\Psi^{\mathrm{C}}(1)$
$1\lvert 0\rangle$	1	0	1	$1 /\!\!/ 1$	$0 /\!\!/ 1$
$1\lvert 0\rangle + 1\lvert 1\rangle$	1	1	2	$1 /\!\!/ 2$	$1 /\!\!/ 2$
$1\lvert 0\rangle + (1+\mathrm{i})\lvert 1\rangle$	1	2	3	$1 /\!\!/ 3$	$2 /\!\!/ 3$
$(1-\mathrm{i})\lvert 0\rangle + (1+\mathrm{i})\lvert 1\rangle$	2	2	4	$2 /\!\!/ 4$	$2 /\!\!/ 4$

can be simplified as

$$\mu_\Psi^{\mathrm{B}}(|i\rangle\langle i|) \equiv \mu_\Psi^{\mathrm{B}}(i) = \frac{\langle\Psi|i\rangle\langle i|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{|\langle i|\Psi\rangle|^2}{\langle\Psi|\Psi\rangle} = \frac{|\alpha_i|^2}{\langle\Psi|\Psi\rangle}, \tag{4.27}$$

where $|\Psi\rangle = \begin{pmatrix}\alpha_0 & \alpha_1 & \cdots & \alpha_{D-1}\end{pmatrix}^T$ is an unnormalized state. Hereafter, we will simply call $\mu_\Psi^{\mathrm{B}}(i)$ the probability of measuring $|i\rangle$.

Although division is not an allowed operation for the elements in an ordered region, following the standard procedure to define a conventional fraction as a pair of integers [59, 72, 87, 88], we could define a *cardinal probability* as a pair of order-region elements as well:

$$\mu_\Psi^{\mathrm{C}}(i) = \mathsf{N}(\langle i|\Psi\rangle) /\!\!/ \mathsf{N}(|\Psi\rangle) = \mathsf{N}(\alpha_i) /\!\!/ \mathsf{N}(|\Psi\rangle), \tag{4.28}$$

where every probability amplitude $\alpha_i \in F^D(k)$ so that both $\mathsf{N}(\alpha_i)$ and $\mathsf{N}(|\Psi\rangle)$ are within the order range $S_0(k)$, and $\mu_\Psi^{\mathrm{C}}(i)$ is called the cardinal probability of measuring $|i\rangle$. For example, let $p = 311$, $k = 11$, and $D = 2$. The permitted range is $S_0(11) = \{-5, \ldots, -1, 0, 1, \ldots, 5\}$, given the dimension $D = 2$, the allowed probability amplitude coefficients are $F^2(11) = \{0, \pm 1, \pm\mathrm{i}, (\pm 1 \pm \mathrm{i})\}$ (see Table 4.5). Now the cardinal probabilities of several representative one-qubit states are listed in Table 4.6.

When measuring cardinal probabilities, *inequalities* can be preserved with appropriate resources

(in the form of a sufficiently large choice of the field), while *equalities* cannot be guaranteed in the theory, and in fact, can be represented as *inequalities of any order*. That is, given two cardinal probabilities $\mathsf{N}(\alpha_i) /\!\!/ \mathsf{N}(|\Psi\rangle)$ and $\mathsf{N}(\alpha_j) /\!\!/ \mathsf{N}(|\Psi\rangle)$ with the same "denominator", $\mathsf{N}(\alpha_i) > \mathsf{N}(\alpha_j)$ physically means it is more likely to measure $|i\rangle$ than $|j\rangle$ if we have enough resource; $\mathsf{N}(\alpha_i) = \mathsf{N}(\alpha_j)$ means the experimental results might not always favor $|i\rangle$ or $|j\rangle$ no matter how many resources we use. The same principle can also apply to two cardinal probabilities with the different "denominator". The details of the comparison can be easily formulated by following the standard procedure [59, 72, 87, 88] and will leave as an exercise for the readers.

4.6 DISCRETE QUANTUM COMPUTING (II)

We now examine two particularly important types of examples within the discrete theory of the previous section: the first is the deterministic Deutsch-Jozsa algorithm [10, 36, 52], which determines the balanced or unbalanced nature of an unknown function with a single measurement step ($O(1)$), and the second is the (normally) probabilistic Grover algorithm [4, 36, 53], determining the result of an unstructured search in $O\left(\sqrt{N}\right)$ time. In the following, we use k to denote the upper bound of the ordered range of integers needed to perform a given calculation; this in turn is assumed to be implemented using a choice of a finite prime number p that supports calculation in the range of k.

4.6.1 DISCRETE DEUTSCH-JOZSA ALGORITHM: DETERMINISTIC

To examine the Deutsch-Jozsa algorithm in the discrete theory of the previous section, we assume we are given a classical function $f : \mathsf{Bool}^n \to \mathsf{Bool}$ and are told that f is either constant or balanced [10, 36, 52]. The algorithm is expressed in a space of dimension $D = 2^{n+1}$: it begins with the $n+1$ qubit state $|1\rangle\left|\overline{0}\right\rangle$ where the overline denotes a sequence of length n. A straightforward calculation [10]

shows that the final state is[2]

$$\sum_{\overline{z}\in\{0,1\}^n}\sum_{\overline{x}\in\{0,1\}^n}(-1)^{f(\overline{x})+\overline{x}\cdot\overline{z}}\left(|0\rangle\,|\overline{z}\rangle-|1\rangle\,|\overline{z}\rangle\right)\,, \tag{4.29}$$

and that its field norm is 2^{n+1}. To make sure that the algorithm works properly, we note that all the probability amplitudes involved in the calculation are in the range -2^n, ..., 2^n and therefore, by Eq. (4.26), we get the following constraint on the size of the ordered region in the finite field:

$$2^{n+1}\left(2^n\right)^2\leq\frac{k-1}{2}\quad\Leftrightarrow\quad k\geq 2^{3n+2}+1\,. \tag{4.30}$$

Now we need to choose a prime number p that supports calculation in the range of k. Assume that k is the least prime satisfying $k\geq 2^{3n+2}+1$, and let p be the $\pi\left(k\right)$-th element of the sequence A000229 [85]. We argue that no prime less than this value of p can support calculation in the ordered range of k and that this p is sufficient to support such calculation. Since k is the least quadratic non-residue of p, every number less than k is a quadratic residue, and thus 0, 1, 2, 3, ..., 2^{3n+2} are all quadratic residues. Hence, the numbers -2^n, ..., 2^n are all inside the ordered range $S_0\left(k\right)$. On the other hand, if we choose any prime smaller than p, there is a quadratic non-residue smaller than k, and we also know that the least quadratic non-residue is a prime [74]. Thus, there is a quadratic non-residue in 0, 1, 2, 3, ..., 2^{3n+2}, and therefore, for this smaller p, there would be a number in -2^n, ..., 2^n that is not in the ordered range $S_0\left(k\right)$.

When f is constant, the cardinal probability of measuring $|0\rangle\left|\overline{0}\right\rangle$ or $|1\rangle\left|\overline{0}\right\rangle$ is $\left(2^n\right)^2+\left(2^n\right)^2=2^{2n+1}\mathbin{/\!\!/}2^{2n+1}$; i.e., the cardinal probability of measuring any other state is $0\mathbin{/\!\!/}2^{2n+1}$. When f is balanced, the cardinal probability of measuring $|0\rangle\left|\overline{0}\right\rangle$ or $|1\rangle\left|\overline{0}\right\rangle$ is $0\mathbin{/\!\!/}2^{2n+1}$. Therefore, if we find that the post-measurement state is either $|0\rangle\left|\overline{0}\right\rangle$ or $|1\rangle\left|\overline{0}\right\rangle$, we know f is constant; otherwise, f is

[2]Note that the algorithm in reference [10] makes use of the Hadamard matrix. We have eliminated the factor $\frac{1}{\sqrt{2}}$ to ensure that all quantities are expressed in terms of integers. Also, notice that the positioning of the initial qubit state $|1\rangle$ is reversed from [10].

Table 4.7: Extension of transitively ordered elements.

p	…	422231	…	196265095009	…		…		…
k	…	**37**	…	**131**	…	**257**	…	**32771**	…
$\pi(k)$	…	12	…	32	…	55	…	3513	…

balanced.

For a single qubit Deutsch problem, the probability amplitudes are between -2 and 2, and the dimension $D = 2^{1+1} = 4$, so we want to have

$$k \geq 2^{3\cdot 1+2} + 1 = 2^5 + 1 = 33\,. \tag{4.31}$$

The least prime satisfying the above condition is $k = 37$, and thus $\pi(37) = 12$ and $p = 422231$, as shown in the extended elements Table 4.7.

For the 2-qubit Deutsch-Jozsa, the computation is already quite challenging. Now the probability amplitudes are between -4 and 4, and the dimension $D = 2^{2+1} = 8$, so we need

$$k \geq 2^{3\cdot 2+2} + 1 = 2^8 + 1 = 257\,. \tag{4.32}$$

Because 257 is a prime, we can pick $k = 257$ and $\pi(257) = 55$. The actual value of p is already outside the range of the published sequence A000229.

These examples illustrate that the value of p plays an essential role: its size grows with the numerical range of the intermediate and final results of the algorithms being implemented. Therefore, we naturally recover a deterministic measure of the intrinsic resources required for a given level of complexity; this measure is normally completely hidden in computations with real numbers, and explicitly exposing it is one of the significant achievements of our discrete field analysis of quantum computation. This solves the conundrum that the conventional Deutsch-Jozsa algorithm

mysteriously continues to work for larger and larger input functions without any apparent increase in resources. Our analysis of this problem reveals that as the size of the input increases, it is necessary to increase the size of p and hence the size of the underlying available numeric coefficients. This observation does not fully explain the power of quantum computing over classical computing, but at least it explains that some of the power of quantum computing depends on increasingly larger precision in the underlying field of numbers.

4.6.2 DISCRETE GROVER SEARCH: NONDETERMINISTIC

As an example of how to apply our cardinal probability framework to a nondeterministic algorithm, we consider discrete Grover's algorithm searching an unstructured database of size $N = 2^n$ [4, 36, 53]. Let $f : \mathsf{Bool}^n \to \mathsf{Bool}$ be the function we want to search. To simplify our discussion, we will only consider

$$f(\overline{x}) = \begin{cases} 1\,, & \text{if } \overline{x} = \overline{0}\,; \\ \\ 0\,, & \text{if } \overline{x} \neq \overline{0}\,, \end{cases} \tag{4.33}$$

where we identify `false` and `true` as 0 and 1 as usual, and $\overline{0} = (0, \ldots, 0)$.

To solve Grover's problem, we represent the search states as n-qubit states as usual. However, instead of the Deutsch black box U_f, f is represented as the $N \times N$ "phase rotation" matrix

$$R = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{4.34}$$

where the "marked" element is in the first position. Beside R, we also need the $N \times N$ "diffusion"

matrix

$$\Delta = \begin{pmatrix} 1-\frac{N}{2} & 1 & 1 & \cdots & 1 \\ 1 & 1-\frac{N}{2} & 1 & \cdots & 1 \\ 1 & 1 & 1-\frac{N}{2} & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1-\frac{N}{2} \end{pmatrix}, \tag{4.35}$$

where we have eliminated, in matrix Δ, the scaling factor $\frac{2}{N}$ to enforce the requirement that all matrix coefficients in our framework are integer-valued. By applying the transformation ΔR repeatedly

$$j = \text{round}\left(\frac{\pi}{4 \arccos\sqrt{1-\frac{1}{N}}} - \frac{1}{2}\right) \approx \text{round}\left(\frac{\pi}{4}\sqrt{N}\right) \tag{4.36}$$

times, we can find the target element $\overline{0}$. In our context, we also need to choose a prime number that is large enough to ensure that all the numbers that occur during the calculation and after measurement are within the transitively-ordered subrange.

Let's walk through the state in each iteration to make sure the algorithm works. Because the probability amplitudes of $|\overline{x}\rangle$ are all the same for $\overline{x} \neq \overline{0}$, we can let a_l be the probability amplitude of $|\overline{0}\rangle$, with b_l the probability amplitude of each of the other possibilities, which are all the same, after the operators ΔR is applied l times. Beginning at $l = 0$ with the information-less state, the normalization scaled to integer values as usual, the state after l-th iteration can be written as

$$\begin{pmatrix} a_l \\ b_l \\ \vdots \\ b_l \end{pmatrix} = (\Delta R)^l \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \Delta R \begin{pmatrix} a_{l-1} \\ b_{l-1} \\ \vdots \\ b_{l-1} \end{pmatrix}. \tag{4.37}$$

We can solve the above iteration by the following recurrence relation for the successive coefficients:

$$a_0 = 1\,, \qquad a_{l+1} = \left(\frac{N}{2} - 1\right) a_l + (N-1)\, b_l\,, \tag{4.38a}$$

$$b_0 = 1\,, \qquad b_{l+1} = (-1)\, a_l + \left(\frac{N}{2} - 1\right) b_l\,. \tag{4.38b}$$

We also know $|a_j| > |b_j|$, so we can estimate an upper bound for the maximum cardinal probability as $\max \mathsf{N}\left(a_j\right) \leq 2\left(\frac{N}{2}\right)^{2j+1}$. By applying Eq. (4.26) with $D = N = 2^n$, we can estimate k using $k \geq 8\left(\frac{N}{2}\right)^{2j+2} + 1$. We can then pick a prime k, and choose the $\pi(k)$-th prime in the sequence represented by Table 4.4 guaranteeing that every number we need for the computation is within the transitively ordered range $F^D(k)$.

For the 2-qubit Grover search, we have $N = D = 4$ and $j = 1$, with the maximum cardinal probability

$$\max \mathsf{N}\left(a_j\right) \leq 2\left(\frac{4}{2}\right)^{2+1} = 16\,, \tag{4.39}$$

so we need

$$k \geq 8\left(\frac{4}{2}\right)^{2\cdot 1+2} + 1 = 8 \cdot 2^4 + 1 = 129\,. \tag{4.40}$$

The least prime k satisfying the above condition is $k = 131$, and so $\pi(131) = 32$ and $p = 196265095009$.

When $p = 196265095009$, we assume that $f(\overline{x}) = 1$ if and only if $|\overline{x}\rangle = |0\rangle|0\rangle$, and so the final state is $\begin{pmatrix} 4 & 0 & \cdots & 0 \end{pmatrix}^T$ with the field norm of 16. Then, the cardinal probability of obtaining $|0\rangle|0\rangle$ as the post-measurement state is $16 \mathbin{/\!\!/} 16$, and it is $0 \mathbin{/\!\!/} 16$ for the rest of the states.

For the 3-qubit Grover search, we have $N = D = 8$ and $j = 2$, with an upper bound $\max \mathsf{N}\left(a_j\right) \leq 2\left(\frac{8}{2}\right)^{4+1} = 2048$ on the cardinal probability. Thus

$$k \geq 8\left(\frac{8}{2}\right)^{6} + 1 = 32769\,. \tag{4.41}$$

The nearest prime greater than this number is 32771, so we can pick $k = 32771$ and $\pi\left(32771\right) = 3513$, and so if we use the 3513-th prime, we can implement Grover's algorithm for a database of size 8.

Continuing with the 3-qubit Grover example, we show how the cardinal probabilities evolve to single out the target state. First, assume that $f\left(\overline{x}\right) = 1$ if and only if $|\overline{x}\rangle = |0\rangle|0\rangle|0\rangle$. The initial information-less 8-dimensional state vector evolves under the application of ΔR as follows:

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 10 \\ 2 \\ \vdots \\ 2 \end{pmatrix} \rightarrow \begin{pmatrix} 44 \\ -4 \\ \vdots \\ -4 \end{pmatrix}. \tag{4.42}$$

These states have differing field norm, so we cannot compare their cardinal probability directly. Since states differ by only scalar multiplication representing the same state, if we multiply the first and second states by 16 and 4, respectively, their field norms become the same value of 2048. The now-consistently-normalized states become

$$\begin{pmatrix} 16 \\ 16 \\ \vdots \\ 16 \end{pmatrix} \rightarrow \begin{pmatrix} 40 \\ 8 \\ \vdots \\ 8 \end{pmatrix} \rightarrow \begin{pmatrix} 44 \\ -4 \\ \vdots \\ -4 \end{pmatrix}. \tag{4.43}$$

Therefore, the cardinal probabilities of measuring $|0\rangle|0\rangle|0\rangle$ in each state are

$$256 \mathbin{/\!\!/} 2048 \qquad 1600 \mathbin{/\!\!/} 2048 \qquad 1936 \mathbin{/\!\!/} 2048, \tag{4.44}$$

while the cardinal probabilities of measuring the other states become

$$256 \mathbin{/\!\!/} 2048 \qquad 64 \mathbin{/\!\!/} 2048 \qquad 16 \mathbin{/\!\!/} 2048. \tag{4.45}$$

We may thus conclude that the cardinal probability of measuring the satisfying assignment of f increases as we apply the diffusion Δ and phase rotation R matrices repeatedly.

Clearly, the required size of k increases systematically with the problem size, and the corresponding size of the required prime number p defining the discrete field increases in the fashion illustrated in Tables 4.4 and 4.7.

4.7 QUANTUM PROBABILITY MEASURES OVER FINITE FIELDS

Although cardinal probabilities can be ordered as discussed in 4.5.2, it is still unclear how to define their arithmetic operations, which is required to further define the mixed states and expectation values. Notice that the cardinal probability over finite fields defined in Eq. (4.28) is based on the conventional Born rule, Eq. (4.27) which could be derived axiomatically according to Gleason's theorem as in Sec. 2.2.2. To overcome these issues, we want to follow the steps in Sec. 2.2 to define events and probability measures, and see whether we could have a Gleason-like theorem which induces a correspondence between states and probability measures qualified as a discrete Born rule.

Definition 4.2 (Quantum Probability Measures over Finite Fields)**.** Given a vector space $\mathcal{H}$ of dimension D over the complexified field $\mathbb{F}_{p^2}$, the set of events $\mathcal{E}_{p^2}$ is recursively defined as follows:

- $\mathbb{0}$ and $\mathbb{1} \in \mathcal{E}_{p^2}$.
- If $|\Psi\rangle$ is a unit-norm state in $\mathcal{H}$, i.e., $|\Psi\rangle \in \mathcal{H}$ with $\langle\Psi|\Psi\rangle = \mathsf{N}(|\Psi\rangle) = 1$, then the projector of the form $|\Psi\rangle\langle\Psi| \in \mathcal{E}_{p^2}$.
- For each pair of *orthogonal* events $P_0 \in \mathcal{E}_{p^2}$ and $P_1 \in \mathcal{E}_{p^2}$, i.e., $P_0 P_1 = \mathbb{0}$, their sum $P_0 + P_1$ is an event, i.e., $P_0 + P_1 \in \mathcal{E}_{p^2}$.

Then a quantum probability measure over a finite field (QPMFF) $\mu \colon \mathcal{E}_{p^2} \to [0, 1]$ assigns a probability to each event (projection operator P) subject to $\mu(\mathbb{0}) = 0$, $\mu(\mathbb{1}) = 1$, and satisfying for each pair of *orthogonal* events P_0 and P_1, $\mu(P_0 + P_1) = \mu(P_0) + \mu(P_1)$.

After defining a QPMFF μ, we would like to follow the idea of Gleason's theorem to see whether we might trace back from μ to a pure or mixed state and summarize this correspondence as a discrete Born rule. When $p = D = 3$, there is indeed some correspondence between QPMFFs and states. However, when $D = 3$ and $p = 7$, we have numerically verified that the only QPMFF $\mu\colon \mathcal{E}_{7^2} \to [0, 1]$ is

$$\mu(P) = \begin{cases} 0 & \text{if } P = \mathbb{0}\,; \\ \frac{1}{3} & \text{if } P \text{ is a one-dimensional projector;} \\ \frac{2}{3} & \text{if } P \text{ is a two-dimensional projector;} \\ 1 & \text{if } P = \mathbb{1}\,. \end{cases} \tag{4.46}$$

As we computed in Eq. (4.12), the number of unique irreducible pure states in $\mathbb{F}_{7^2}^3$ is

$$\frac{7^{3-1}\left(7^3 - (-1)^3\right)}{7+1} = 2107 \tag{4.47}$$

which is much more than one, the number of QPMFF. Since we don't have enough QPMFF to correspond even pure states, there is no discrete Born rule in this case. In general, we conjecture that QPMFF is always unique for $D \geq 3$ except $D = p = 3$.

Although we couldn't prove there is no discrete Born rule by investigating QPMFFs alone, we could prove there is no "sensible" Born rule which also satisfies an analog of Eqs. (2.27) and (2.28).

Theorem 4.1. *For $D \geq 3$ except $p = D = 3$, there is no "sensible" Born rule μ^{F} parametrized by unit-norm states $|\Phi\rangle$ in $\mathcal{H}$ such that μ_Φ^{F} is a QPMFF;*

$$\mu_\Phi^{\mathrm{F}}(P) = 1 \tag{4.48}$$

if and only if $P|\Phi\rangle = |\Phi\rangle$*; and*

$$\mu^{\mathrm{F}}_{U|\Phi\rangle}\left(UPU^{\dagger}\right) = \mu^{\mathrm{F}}_{\Phi}\left(P\right) , \tag{4.49}$$

where U *is any unitary map.*

Proof. [3]Assume such a map μ^{F} satisfying all properties exists, we will use the listed properties to build a contradiction. Consider the computational basis $\{|0\rangle, |1\rangle, |2\rangle, ...\}$, and the projectors formed by these vectors, $P_0 = |0\rangle\langle 0|$, $P_1 = |1\rangle\langle 1|$, and $P_2 = |2\rangle\langle 2|$. Since $p \geq 7$, $1+1+1$ is not 0 in $\mathbb{F}_{p^2}$, and has the principal inverse field norm $\gamma = \mathsf{N}^{-1}(3)$ as defined in Sec. 4.1.2. The unit-norm state $|\oplus\rangle = \frac{1}{\gamma}\left[|0\rangle + |1\rangle + |2\rangle\right]$ can then be used as a parameter of μ^{F} and induces a QPMFF $\mu^{\mathrm{F}}_{\oplus}$. To compute the probability values of $\mu^{\mathrm{F}}_{\oplus}$, we want to utilize Eq. (4.49) by letting U_i be the unitary map that permutes the basis vectors $|0\rangle$ and $|i\rangle$ and acts as the identity for the rest for $i = 1$ and 2. Note that $|\oplus\rangle$ is invariant under U_i, so we have $\mu^{\mathrm{F}}_{\oplus}(P_0) = \mu^{\mathrm{F}}_{U_i|\oplus\rangle}\left(U_i P_0 U_i^{\dagger}\right) = \mu^{\mathrm{F}}_{\oplus}(P_i)$ by Eq. (4.49). Let $P' = P_0 + P_1 + P_2$. Since $P'|\oplus\rangle = |\oplus\rangle$, Eq. (4.48) implies

$$1 = \mu^{\mathrm{F}}_{\oplus}(P') = \mu^{\mathrm{F}}_{\oplus}(P_0) + \mu^{\mathrm{F}}_{\oplus}(P_1) + \mu^{\mathrm{F}}_{\oplus}(P_2) = 3\mu^{\mathrm{F}}_{\oplus}(P_0) , \tag{4.50}$$

and thus for $i = 0$, 1, and 2, we always have $\mu^{\mathrm{F}}_{\oplus}(P_i) = \frac{1}{3}$.

Following the similar computation to the previous paragraph, we now compute the probabilities induced by $|\Phi'\rangle = U'|\oplus\rangle = \frac{2\alpha}{\gamma}|0\rangle + \frac{1}{\gamma}|2\rangle$, where $\alpha = \mathsf{N}^{-1}\left(\frac{1}{2}\right) \in \mathbb{F}_{p^2}$ and

$$U' = \begin{pmatrix} \alpha & \alpha & 0 \\ -\alpha & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{4.51}$$

Since $|2\rangle$ is invariant under U', Eq. (4.49) implies $\mu^{\mathrm{F}}_{\oplus}(P_2) = \mu^{\mathrm{F}}_{U'|\oplus\rangle}\left(U'P_2U'^{\dagger}\right) = \mu^{\mathrm{F}}_{\Phi'}(P_2)$. Also,

[3]This proof assumes $p \geq 7$, and there is a simpler proof for $D \geq 4$ applying to $p = 3$ case [8].

Eq. (4.48) implies $\mu^{\mathrm{F}}_{\Phi'}(P_0)+\mu^{\mathrm{F}}_{\Phi'}(P_2) = \mu^{\mathrm{F}}_{\Phi'}(P_0+P_2) = 1$. By moving $\mu^{\mathrm{F}}_{\Phi'}(P_2)$ to the right-hand side of the equation, we have

$$\mu^{\mathrm{F}}_{\Phi'}(P_0) = 1-\mu^{\mathrm{F}}_{\Phi'}(P_2) = 1-\mu^{\mathrm{F}}_{\oplus}(P_2) = 1-\tfrac{1}{3} = \tfrac{2}{3}\,. \tag{4.52}$$

Iterating the similar process with $|\Phi''\rangle = U''|\Phi'\rangle = \frac{2\alpha}{\gamma}(|0\rangle+|1\rangle)+\frac{\beta}{\gamma}|2\rangle$, where $\beta = \mathsf{N}^{-1}(-1) \in \mathbb{F}_{p^2}$ and

$$U'' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \beta^* & 2\alpha \\ 0 & 2\alpha^* & \beta \end{pmatrix}. \tag{4.53}$$

Because $|\Phi''\rangle$ is invariant under U_1, invoking Eq. (4.49) in $\mu^{\mathrm{F}}_{\Phi''}(P_1) = \mu^{\mathrm{F}}_{U_1|\Phi''\rangle}\left(U_1P_1U_1^\dagger\right) = \mu^{\mathrm{F}}_{\Phi''}(P_0)$, we know $\mu^{\mathrm{F}}_{\Phi''}(P_1)$ and $\mu^{\mathrm{F}}_{\Phi''}(P_0)$ are same. They are both equal to $\frac{2}{3}$ because $\mu^{\mathrm{F}}_{\Phi''}(P_0) = \mu^{\mathrm{F}}_{U''|\Phi'\rangle}(U''P_0U'') = \mu^{\mathrm{F}}_{\Phi'}(P_0) = \frac{2}{3}$ by Eqs. (4.49) and (4.52). Since $P'|\Phi''\rangle = |\Phi''\rangle$, Eq. (4.48) implies

$$1 = \mu^{\mathrm{F}}_{\Phi''}(P') = \mu^{\mathrm{F}}_{\Phi''}(P_0)+\mu^{\mathrm{F}}_{\Phi''}(P_1)+\mu^{\mathrm{F}}_{\Phi''}(P_2) = \tfrac{2}{3}+\tfrac{2}{3}+\mu^{\mathrm{F}}_{\Phi''}(P_2) > \tfrac{4}{3}\,. \tag{4.54}$$

This is inconsistent with the requirement that the probabilities for orthogonal outcomes add up to 1, and build a contradiction. □

The reason why there is no "sensible" Born rule might be that the state spaces are now discrete and finite, but we still consider mapping probability assignments to infinitely precise values in the unit interval $[0,1]$. To overcome this issue, we want to consider a discrete Born rule mapping to a finite number of intervals called interval-valued probability [3, 32]. To adopting the idea of interval-valued probability step-by-step, before attempting to study quantum interval-valued probability over finite fields, we will first review the classical interval-valued probability and extend it with the conventional quantum theory in the next chapter.

Chapter 5

TOWARD A QUANTUM MEASUREMENT THEORY WITH ERROR: QUANTUM INTERVAL-VALUED PROBABILITY

5.1 INTERVALS OF UNCERTAINTY

5.1.1 DEFINITIONS OF CLASSICAL AND QUANTUM IVPMS

We will start by reviewing classical IVPMs and then propose our quantum generalization. In the classical setting, there are several proposals for "imprecise probabilities" [29–34, 89, 90]. Although these proposals differ in some details, they all share the fact that the probability $\mu(E)$ of an event E is generalized from a single *real number* to an *interval* $[\ell, r]$, where ℓ intuitively corresponds to the strength of evidence for the event E and $1-r$ corresponds to the strength of the evidence against the same event. Under some additional assumptions, this interval could be interpreted as the Gaussian width of a probability distribution.

We next introduce probability axioms for IVPMs. First, for each interval $[\ell, r]$ we have the natural constraint $0 \leq \ell \leq r \leq 1$ that guarantees that every element of the interval can be interpreted as a conventional probability. We also include $\mathsf{F} = [0, 0]$ and $\mathsf{T} = [1, 1]$ as limiting intervals that refer, respectively, to the probability interval for impossible events and for events that are certain. We can write the latter as $\mu(\emptyset) = \mathsf{F}$ and $\mu(\Omega) = \mathsf{T}$, where $\emptyset$ is the empty set and Ω is the event covering the entire sample space. For each interval $[\ell, r]$, we also need the dual interval $[1 - r, 1 - \ell]$ so that if one interval refers to the probability of an event E, the dual refers to the probability of the event's

complement $\overline{E}$. For example, if we discover as a result of an experiment that $\mu(E) = [0.2, 0.3]$ for some event E, we may conclude that $\mu\left(\overline{E}\right) = [0.7, 0.8]$ for the complementary event $\overline{E}$. In addition to these simple conditions, there are some subtle conditions on how intervals are combined, which we discuss next.

Let E_0 and E_1 be two disjoint events with probabilities $\mu(E_0) = [\ell_0, r_0]$ and $\mu(E_1) = [\ell_1, r_1]$. A first attempt at calculating the probability of the combined event that *either* E_0 or E_1 occurs might be $\mu(E_0 \cup E_1) = [\ell_0 + \ell_1, r_0 + r_1]$. In some cases, this is indeed a sensible definition. For example, if $\mu(E_0) = [0.1, 0.2]$ and $\mu(E_1) = [0.3, 0.4]$ we get $\mu(E_0 \cup E_1) = [0.4, 0.6]$. But consider an event E such that $\mu(E) = [0.2, 0.3]$ and hence $\mu\left(\overline{E}\right) = [0.7, 0.8]$. The two events E and $\overline{E}$ are disjoint; the naïve addition of intervals would give $\mu\left(E \cup \overline{E}\right) = [0.9, 1.1]$, which is not a valid probability interval. Moreover, the event $E \cup \overline{E}$ is the entire space; its probability interval should be $\mathbf{T}$ which is sharper than $[0.9, 1.1]$. The problem is that the two intervals are correlated: there is more information in the combined event than in each event separately so the combined event should be mapped to a sharper interval. In our example, even though the "true" probability of E can be anywhere in the range $[0.2, 0.3]$ and the "true" probability of $\overline{E}$ can be anywhere in the range $[0.7, 0.8]$, the values are not independent. Any value of $\mu(E) \leq 0.25$ will force $\mu\left(\overline{E}\right) \geq 0.75$. To account for such subtleties, the axioms of interval-valued probability do not use a strict equality for the combination of disjoint events. The correct constraint enforcing coherence of the probability assignment for $E_0 \cup E_1$ when E_0 and E_1 are disjoint is taken to be:

$$\mu(E_0 \cup E_1) \subseteq [\ell_0 + \ell_1, r_0 + r_1]\,. \tag{5.1}$$

Note that for any event E with $\mu(E) = [\ell, r]$, we always have

$$\mu(\Omega) = \mathbf{T} \subseteq [\ell, r] + [1 - r, 1 - \ell] = \mu(E) + \mu\left(\overline{E}\right)\,. \tag{5.2}$$

When combining non-disjoint events, there is a further subtlety whose resolution will give us the final general condition for IVPMs. For events E_0 and E_1, not necessarily disjoint, we have:

$$\mu(E_0 \cup E_1) + \mu(E_0 \cap E_1) \subseteq \mu(E_0) + \mu(E_1)\,, \tag{5.3}$$

which is a generalization of the classical inclusion-exclusion principle that uses $\subseteq$ instead of $=$ for the same reason as before. The new condition, known as *convexity* [30, 31, 34, 54, 55, 91], reduces to the previously motivated Eq. (5.1) when the events are disjoint, i.e., when $\mu(E_0 \cap E_1) = 0$.

Previous discussions can be summarized as the following definition [32].

Definition 5.1 (IVPM)**.** Assume a collection of intervals $\mathcal{I}$ including $\mathbf{F}$ and $\mathbf{T}$ with addition and scalar multiplication defined as follows:

$$[\ell_0, r_0] + [\ell_1, r_1] = [\ell_0 + \ell_1, r_0 + r_1] \text{ and} \tag{5.4a}$$

$$x[\ell, r] = \begin{cases} [x\ell, xr] & \text{for } x \geq 0\,; \\ [xr, x\ell] & \text{for } x \leq 0\,. \end{cases} \tag{5.4b}$$

Given a finite sample space Ω, and its power set 2^Ω as the classical event space, a classical IVPM $\bar{\mu} \colon 2^\Omega \to \mathcal{I}$ is a function subject to the following constraints:

$$\bar{\mu}(\emptyset) = \mathbf{F}\,, \tag{5.5a}$$

$$\bar{\mu}(\Omega) = \mathbf{T}\,, \tag{5.5b}$$

$$\bar{\mu}\left(\overline{E}\right) = \mathbf{T} - \bar{\mu}\left(E\right)\,, \tag{5.5c}$$

and satisfying the convexity condition, Eq. (5.3), for each pair of events E_0 and E_1.

Note that the minus sign appearing in Eq. (5.5c) is accommodated by the $x \leq 0$ case in Eq. (5.4b).

We now have the necessary ingredients to define the quantum extension, QIVPMs, as a gener-

alization of both classical IVPMs and conventional quantum probability measures in Sec. 2.2. We will show that QIVPMs reduce to classical IVPMs when the space of quantum events $\mathcal{E}$ is restricted to mutually commuting events $\mathcal{E}_{\mathrm{C}}$, i.e., to compatible events that can be measured simultaneously. In Sec. 5.3 we will discuss the connection between QIVPMs and conventional quantum probability measures in detail.

Definition 5.2 (QIVPM)**.** We take a QIVPM $\bar{\mu}$ to be an assignment of an interval to each event (projection operator P) subject to the following constraints:

$$\bar{\mu}(\mathbb{0}) = \mathbf{F}\,, \tag{5.6a}$$

$$\bar{\mu}(\mathbb{1}) = \mathbf{T}\,, \tag{5.6b}$$

$$\bar{\mu}\left(\mathbb{1} - P\right) = \mathbf{T} - \bar{\mu}\left(P\right)\,, \tag{5.6c}$$

and satisfying for each pair of *commuting* projectors P_0 and P_1 with $P_0 P_1 = P_1 P_0$,

$$\bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) \subseteq \bar{\mu}\left(P_0\right) + \bar{\mu}\left(P_1\right)\,. \tag{5.7}$$

The first three constraints, Eqs. (5.6), are the direct counterpart of the corresponding ones for classical IVPMs, Eqs. (5.5). With the understanding that the union of classical sets $E_0 \cup E_1$ is replaced by $P_0 + P_1 - P_0 P_1$ in the case of quantum projection operators [66], the last condition, Eq. (5.7), is a direct counterpart of the convexity condition of Eq. (5.3). Thus, our definition of QIVPMs merges aspects of both classical IVPMs and quantum probability measures.

5.1.2 STATES CONSISTENT WITH CLASSICAL AND QUANTUM IVPMS

Our definition of QIVPMs is consistent with classical IVPMs in the sense that a restriction of QIVPMs to mutually commuting subspace of events, $\mathcal{E}_{\mathrm{C}}$, recovers the definition of classical IVPMs. To see this, we first define the idea of subspace of events. $\mathcal{E}'$ is called a *subspace* of the set of events $\mathcal{E}$

if $\mathcal{E}'$ contains the projectors $\mathbb{0}$ and $\mathbb{1}$ and is closed under complements, sums, and products. In particular, for any projector $P \in \mathcal{E}'$, we have $\mathbb{1}-P \in \mathcal{E}'$ and for each pair of commuting projectors $P_0 \in \mathcal{E}'$ and $P_1 \in \mathcal{E}'$, we have P_0P_1 and $P_0 + P_1 - P_0P_1 \in \mathcal{E}'$. Given a mutually commuting subspace $\mathcal{E}_{\mathrm{C}}$ whose elements are diagonalizable by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, ..., |D-1\rangle\}$, the function $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined in Eq. (2.17) maps any set E to the sum of the projectors formed by elements in E, and preserve all operations used to defined classical and quantum IVPMs as in Eqs. (2.19) and (2.20). According to the same reason as the classical counterpart, Eq. (2.23), given a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$, the function $\varphi^*\bar{\mu}\colon 2^{\Omega} \to \mathcal{I}$ defined by precomposition $(\varphi^*\bar{\mu})(E) = \bar{\mu}(\varphi(E))$ is the pullback of $\bar{\mu}$ by φ and is a classical IVPM naturally.

Since we can pull back a QIVPM to a classical one, known properties of classical IVPMs directly hold for QIVPMs when one restricts to mutually commuting events $\mathcal{E}_{\mathrm{C}}$, and one of them is *core*. Recall in Sec. 5.1.1, given a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathcal{I}$ and an event $E \subseteq \Omega$ such that $\bar{\mu}(E) = [0.2, 0.3]$ and $\bar{\mu}(\overline{E}) = [0.7, 0.8]$, we discussed the "true" probabilities, $\mu(E)$ and $\mu(\overline{E})$, can be anywhere in the range $[0.2, 0.3]$ and $[0.7, 0.8]$, respectively, as long as they satisfy $\mu(E)+\mu(\overline{E}) = 1$. These "true" probabilities can be represented as a probability measure $\mu\colon 2^{\Omega} \to [0, 1]$ satisfying

$$\mu(E) \in [0.2, 0.3] = \bar{\mu}(E) , \qquad \mu(\overline{E}) \in [0.7, 0.8] = \bar{\mu}(\overline{E}) . \tag{5.8}$$

Together with the trivial fact

$$\mu(\emptyset) = 0 \in \mathbf{F} = \bar{\mu}(\emptyset) , \qquad \mu(\Omega) = 1 \in \mathbf{T} = \bar{\mu}(\Omega) , \tag{5.9}$$

we conclude the "true" probability of each event belongs to the interval probability of the corresponding event. This conclusion gives the following definition of core.

Definition 5.3 (Classical Consistency and Core)**.** We say an IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathcal{I}$ is *consistent* with a probability measure $\mu\colon 2^{\Omega} \to [0, 1]$ on an event $E \subseteq \Omega$ if the interval $\bar{\mu}(E)$ contains the precise

probability calculated by μ, i.e., $\mu(E) \in \bar{\mu}(E)$. The *core* of an IVPM $\bar{\mu}$, core $(\bar{\mu})$, is the collection of all probability measures μ that are *consistent* with $\bar{\mu}$ on every event, that is,

$$\text{core}(\bar{\mu}) = \left\{\mu \colon 2^{\Omega} \to [0,1] \;\middle|\; \forall E \subseteq \Omega,\ \mu(E) \in \bar{\mu}(E)\right\} . \tag{5.10}$$

One fundamental question of these imprecise interval-valued probabilities is whether they *always* have underlying precise probabilities. A positive answer is given by the following theorem.

Theorem 5.1 (Shapley [30, 34, 54, 55])**.** *Every classical IVPM has a non-empty core.*

Although it is *impossible* in the classical world to have an empty core, it is possible in the quantum world for the imprecise probabilities associated with some events to be inconsistent with *any* quantum state, i.e., a QIVPM might have an empty core as we will show in Sec. 5.3. In that case, one cannot guarantee non-empty cores for finite-precision attempts at proving Gleason's theorem by extending the Born measure $\mu_{\rho}^{\mathrm{B}}(P)$ to QIVPMs $\bar{\mu}(P)$. However, if we restrict ourselves to the set $\mathcal{E}_{\mathrm{C}}$ of mutually commuting events, the situation reverts to the classical case in which probabilities always determine at least one state.

We now give the necessary technical definition and lemma to prove this non-empty core property.

Definition 5.4 (Quantum Consistency and Core)**.** We say a QIVPM $\bar{\mu}$ is *consistent* with a state ρ on a projector P if the interval $\bar{\mu}(P)$ contains the exact probability calculated by the Born rule [4, 36, 44], i.e.,

$$\mu_{\rho}^{\mathrm{B}}(P) = \mathrm{Tr}(\rho P) \in \bar{\mu}(P) . \tag{5.11}$$

The *core* $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}')$ of $\bar{\mu}$ relative to a subspace of events $\mathcal{E}'$ is the collection of all states ρ that are *consistent* with $\bar{\mu}$ on every projector in $\mathcal{E}'$, that is,

$$\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}') = \left\{\rho \;\middle|\; \forall P \in \mathcal{E}',\ \mu_{\rho}^{\mathrm{B}}(P) \in \bar{\mu}(P)\right\} . \tag{5.12}$$

In contrast with the classical Thm. 5.1, there is *no guarantee* that there exists a state ρ that satisfies Eq. (5.12) and therefore is in the core of a QIVPM. However, for the special case of commuting events, since the classical core corresponds to the quantum one by the following lemma, there will be a quantum theorem naturally corresponding to classical Thm. 5.1.

Lemma 5.1. *Consider a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$ and a commuting subspace of events $\mathcal{E}_{\mathrm{C}}$ diagonalizable by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ with the function $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined by Eq. (2.17). For any classical probability measure μ consistent with the pullback of $\bar{\mu}$, i.e., $\mu \in \mathrm{core}\,(\varphi^{*}\bar{\mu})$, there is a density matrix ρ consistent with $\bar{\mu}$ relative to $\mathcal{E}_{\mathrm{C}}$, $\rho \in \overline{\mathcal{H}}\,(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$, such that the pullback of μ_{ρ}^{B} is μ, i.e.,*

$$\left(\varphi^{*}\mu_{\rho}^{\mathrm{B}}\right)(E) = \mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) = \mu\left(E\right) \tag{5.13}$$

for all $E \subseteq \Omega$.

Proof. Consider $\rho = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right)|j\rangle\langle j|$ which is a density matrix because ρ is a positive operator and has trace equal to one [10]:

$$\mathrm{Tr}\left(\rho\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right)\mathrm{Tr}\left(|j\rangle\langle j|\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right) = 1\,. \tag{5.14}$$

Then, Eq. (5.13) can be inductive proved on E. By Eq. (2.17) and the generalized Born rule, Eq. (2.30), we have the base case

$$\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(\{|i\rangle\}\right)\right) = \mu_{\rho}^{\mathrm{B}}\left(|i\rangle\langle i|\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right)\mu_{|j\rangle}^{\mathrm{B}}\left(|i\rangle\langle i|\right) = \mu\left(\{|i\rangle\}\right) \tag{5.15}$$

for all i. By Eq. (2.20) and the induction hypothesis $\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) = \mu\left(E\right)$, the inductive case is also

valid:

$$\mu_\rho^{\mathrm{B}}\left(\varphi\left(E\cup\{|i\rangle\}\right)\right)=\mu_\rho^{\mathrm{B}}\left(\varphi\left(E\right)\right)+\mu_\rho^{\mathrm{B}}\left(\varphi\left(\{|i\rangle\}\right)\right)=\mu\left(E\right)+\mu\left(\{|i\rangle\}\right)=\mu\left(E\cup\{|i\rangle\}\right) \quad (5.16)$$

for all $i\notin E$. After we inductively proved $\varphi^*\mu_\rho^{\mathrm{B}}=\mu$, we can finally prove $\rho\in\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_{\mathrm{C}}\right)$. Since μ is consistent with the pullback of $\bar{\mu}$, $\varphi^*\bar{\mu}$, we have

$$\mu_\rho^{\mathrm{B}}\left(\varphi\left(E\right)\right)=\mu\left(E\right)\in\left(\varphi^*\bar{\mu}\right)\left(E\right)=\bar{\mu}\left(\varphi\left(E\right)\right)\ . \quad (5.17)$$

Together with the fact that the image of φ contains $\mathcal{E}_{\mathrm{C}}$, we have $\mu_\rho^{\mathrm{B}}\left(P\right)\in\bar{\mu}\left(P\right)$ for all $P\in\mathcal{E}_{\mathrm{C}}$, i.e., $\rho\in\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_{\mathrm{C}}\right)$. □

Then, the quantum version of Thm. 5.1 is a natural consequence of the previous lemma.

Theorem 5.2 (Non-empty Core for Compatible Measurements)**.** *For every QIVPM $\bar{\mu}\colon\mathcal{E}\to\mathcal{I}$, if a subspace of events $\mathcal{E}_{\mathrm{C}}\subseteq\mathcal{E}$ commutes, then $\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_{\mathrm{C}}\right)\neq\emptyset$.*

Proof. Since $\mathcal{E}_{\mathrm{C}}$ is a set of mutually commuting projections, they can be diagonalized by a common orthonormal basis $\Omega=\{|0\rangle,|1\rangle,\ldots,|D-1\rangle\}$. We can then follow Eq. (2.17) to define $\varphi\colon 2^\Omega\to\mathcal{E}$ such that the pullback of $\bar{\mu}$ by φ, $\varphi^*\bar{\mu}$, is a classical IVPM. By Thm. 5.1, there is a classical probability measure $\mu\colon 2^\Omega\to[0,1]$ consistent with $\varphi^*\bar{\mu}$, and thus there must be a density matrix $\rho\in\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_{\mathrm{C}}\right)$ satisfying $\varphi^*\mu_\rho^{\mathrm{B}}=\mu$ according to Lemma 5.1. □

5.1.3 CLASSICAL CHOQUET INTEGRALS AND EXPECTATION VALUES OF OBSERVABLES

We conclude this section with a generalization of expectation values of observables in the context of QIVPMs. In conventional quantum mechanics, the expectation value of an observable as defined in Eq. (2.24) is a unique real number. The generalization to QIVPMs implies that this expectation

value should be bounded by an interval. We will start from the classical notion of the *Choquet integral* which is used to calculate the expectation value of random variables as a weighted average [30, 34, 56]. Then, we will generalize this notation from classical IVPMs to QIVPMs and prove the generalized definition consistent with our intuition. For example, if $\bar{\mu}$ is a conventional (Born) probability measure induced by a state ρ, then the interval expectation value collapses to a point, thus reducing the interval expectation value to the conventional definition of Eq. (2.24). We will also show the expectation value of an observable relative to a QIVPM $\bar{\mu}$ lies between two possible outcomes, which themselves lie between the minimum and maximum bounds of the probability intervals associated with each state ρ that is consistent with $\bar{\mu}$ on every projector in the spectral decomposition of the observable.

Like how we pulled back classical ideas to the quantum world in Secs. 5.1.1 and 5.1.2, we start from defining the classical expectation value relative to IVPMs. Consider a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathcal{I}$ and an event $E \subseteq \Omega$ such that

$$\bar{\mu}(E) = [0.2, 0.3]\ , \qquad\qquad \bar{\mu}\left(\overline{E}\right) = [0.7, 0.8]\ , \tag{5.18}$$

and a random variable $X\colon \Omega \to \mathbb{R}$ defined by

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in E\,; \\ 2 & \text{if } \omega \notin E\,. \end{cases} \tag{5.19}$$

As we have discussed, the "true" probability is a probability measure $\mu\colon 2^{\Omega} \to [0, 1]$ consistent with $\bar{\mu}$ on every event. On one extreme case, $\mu(E) = 0.2$, the expectation value of X relative to μ is

$$1 \cdot \mu(E) + 2 \cdot \mu\left(\overline{E}\right) = 1 \cdot 0.2 + 2 \cdot 0.8 = 1.8\,. \tag{5.20}$$

On the other extreme case, $\mu(E) = 0.3$, the expectation value of X relative to μ is

$$1 \cdot \mu(E) + 2 \cdot \mu\left(\overline{E}\right) = 1 \cdot 0.3 + 2 \cdot 0.7 = 1.7\,. \tag{5.21}$$

In other word, if μ is in the core of $\bar{\mu}$, the expectation value of X belongs to $[1.7, 1.8]$ which should be the expectation value of X with respect to $\bar{\mu}$.

Although it is intuitive to define the expectation value as the minimum and maximum bounds of the probability intervals associated with each probability measure in the core of $\bar{\mu}$, the core of a general IVPM is hard to be described and computed. Fortunately, the above description is equivalent to the well-known Choquet integral [30, 34, 56, 92], which can be computed step-by-step like a weighted average.

Definition 5.5 (Classical Expectation Values)**.** Consider a classical sample space Ω with a random variable $X\colon \Omega \to \mathbb{R}$. Since we only consider finite sample space, X can always be decomposed into the sum of step functions $\mathbf{1}_E$ defined in Def. 2.1. For our convenience to define the Choquet integral later, we order these step functions by their coefficients from the smallest to the largest, and express X as follow:

$$X = x^-_{N^-}\mathbf{1}_{E^-_{N^-}} + \cdots + x^-_1\mathbf{1}_{E^-_1} + x^+_1\mathbf{1}_{E^+_1} + \cdots + x^+_{N^+}\mathbf{1}_{E^+_{N^+}} = \sum_{s\in\{+,-\}}\sum_{i=1}^{N^s} x^s_i\mathbf{1}_{E^s_i}\,, \tag{5.22}$$

where $\{E^-_i\}_{i=1}^{N^-}$ and $\{E^+_i\}_{i=1}^{N^+}$ are all disjoint and $x^-_{N^-} < \cdots < x^-_1 < 0 \leq x^+_1 < \cdots < x^+_{N^+}$. Then, the expectation values of X relative to real-valued and interval-valued probability measures can be defined as weighted averages as follow.

- Given a classical probability measure $\mu\colon 2^\Omega \to [0,1]$, the expectation value of X relative to

μ is an average of x_i^s with the weight being the measure of their step functions $\mu\left(E_i^s\right)$, i.e.,

$$\int X \mathrm{d}\mu = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \mu\left(E_i^s\right) . \tag{5.23}$$

- Consider a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathcal{I}$. The expectation value of X relative to $\bar{\mu}$ still looks like a weighted average of x_i^s, where s is either $+$ or $-$, but this time the "weights" are not just the measure of their step functions $\mu\left(E_i^s\right)$. Instead, by computing the accumulative sets $E_{i\uparrow}^s = \bigcup_{j\geq i} E_j^s$ and their interval-valued probabilities $\bar{\mu}\left(E_{i\uparrow}^s\right)$, the "weights" are the differences of the both-end of the interval probabilities, $\Delta\mu^t\left(E_{i\uparrow}^s\right) = \mu^t\left(E_{i\uparrow}^s\right) - \mu^t\left(E_{(i+1)\uparrow}^s\right)$, where $\mu^t\left(E_{i\uparrow}^s\right)$ is either the left-end $\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right)$ or right-end $\mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)$ of $\bar{\mu}\left(E_{i\uparrow}^s\right)$, that is, $\bar{\mu}\left(E_{i\uparrow}^s\right) = \left[\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right), \mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)\right]$. Then, the Choquet integral of X relative to $\bar{\mu}$ is averaging x_i^s with the given "weights,"

$$\int X \mathrm{d}\bar{\mu} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\Delta\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right), \Delta\mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)\right] . \tag{5.24}$$

Suppose $\bar{\mu}$ is just a real-valued probability measure μ, i.e., $\bar{\mu}\left(E\right) = \left[\mu\left(E\right), \mu\left(E\right)\right]$ for all E. Since the "weights" in Eq. (5.24) are the difference on the probability of the accumulative sets, they have the same magnitude as ones in Eq. (5.23), and the right-hand side of Eq. (5.24) can be easily

Table 5.1: This table lists the immediate values to compute the Choquet integral $\int X \mathrm{d}\bar{\mu}$ according to Def. 5.5.

i	x_i^+	E_i^+	$E_{i\uparrow}^+$	$\mu^{\mathrm{L}}\left(E_{i\uparrow}^+\right)$	$\Delta\mu^{\mathrm{L}}\left(E_{i\uparrow}^+\right)$	$\mu^{\mathrm{R}}\left(E_{i\uparrow}^+\right)$	$\Delta\mu^{\mathrm{R}}\left(E_{i\uparrow}^+\right)$
1	1	E	Ω	1	0.3	1	0.2
2	2	$\overline{E}$	$\overline{E}$	0.7	0.7	0.8	0.8
3			$\emptyset$	0		0	

simplified to the right-hand side of Eq. (5.23) as follow:

$$
\begin{aligned}
&\int X \mathrm{d}\bar{\mu} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\Delta\mu\left(E_{i\uparrow}^s\right), \Delta\mu\left(E_{i\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(E_{i\uparrow}^s\right) - \mu\left(E_{(i+1)\uparrow}^s\right), \mu\left(E_{i\uparrow}^s\right) - \mu\left(E_{(i+1)\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(\bigcup_{j\geq i} E_j^s\right) - \mu\left(\bigcup_{j\geq i+1} E_j^s\right), \mu\left(\bigcup_{j\geq i} E_j^s\right) - \mu\left(\bigcup_{j\geq i+1} E_j^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\sum_{j\geq i} \mu\left(E_j^s\right) - \sum_{j\geq i+1} \mu\left(E_j^s\right), \sum_{j\geq i} \mu\left(E_j^s\right) - \sum_{j\geq i+1} \mu\left(E_j^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(E_i^s\right), \mu\left(E_i^s\right)\right] = \left[\int X \mathrm{d}\mu, \int X \mathrm{d}\mu\right].
\end{aligned}
\tag{5.25}
$$

The Choquet integral defined in Eq. (5.24) is consistent with not only Eq. (5.23) but also our previous intuition. For example, if $\bar{\mu}$ and X defined in Eqs. (5.18) and (5.19) have an intuitive expectation interval $[1.7, 1.8]$, their Choquet integral $\int X \mathrm{d}\bar{\mu}$ should give the same interval $[1.7, 1.8]$. To verify our idea, we first represent X as the sum $1 \cdot \mathbf{1}_E + 2 \cdot \mathbf{1}_{\overline{E}}$. Since the coefficients are all positive, we have $N^- = 0$, $N^+ = 2$, $x_1^+ = 1$, $E_1^+ = E$, $x_2^+ = 2$, and $E_2^+ = \overline{E}$ as listed in Table 5.1. Step-by-step we compute the accumulative sets $E_{i\uparrow}^+$, their measures $\mu^t\left(E_{i\uparrow}^+\right)$, and their differences

$\Delta\mu^t\left(E^s_{i\uparrow}\right)$. These values can then be plugged into Eq. (5.24) giving the Choquet integral

$$\begin{aligned}\int X\mathrm{d}\bar{\mu} &= x_1^+\left[\Delta\mu^{\mathrm{L}}\left(E^+_{1\uparrow}\right),\Delta\mu^{\mathrm{R}}\left(E^+_{1\uparrow}\right)\right]+x_2^+\left[\Delta\mu^{\mathrm{L}}\left(E^+_{2\uparrow}\right),\Delta\mu^{\mathrm{R}}\left(E^+_{2\uparrow}\right)\right]\\ &= 1\cdot[0.3,0.2]+2\cdot[0.7,0.8]=[1.7,1.8]\ .\end{aligned} \tag{5.26}$$

In general, the Choquet integral is always equal to the minimum and maximum of expectation values relative to probability measures in the core according to the following theorem [30, 34, 93].

Theorem 5.3. *For every classical IVPM $\bar{\mu}\colon 2^{\Omega}\to\mathcal{I}$ and any random variable $X\colon\Omega\to\mathbb{R}$, we have*

$$\int X\mathrm{d}\bar{\mu}=\left[\min_{\mu\in\mathrm{core}(\bar{\mu})}\int X\mathrm{d}\mu,\max_{\mu\in\mathrm{core}(\bar{\mu})}\int X\mathrm{d}\mu\right]. \tag{5.27}$$

After we defined the expectation values of random variables relative to classical IVPMs, we can merge this definition with the expectation values to quantum probability measures, Eq. (2.24), and define the expectation values of observables relative to QIVPMs as follow.

Definition 5.6 (Expectation Values relative to QIVPMs)**.** Since we want to define the expectation value of observables relative to QIVPMs parallel to classical definition 5.5, we order the eigenvalues of an observable $\mathbf{O}$ in the spectral decomposition from the smallest to the largest as well

$$\mathbf{O}=\lambda^-_{N^-}P^-_{N^-}+\cdots+\lambda^-_1P^-_1+\lambda^+_1P^+_1+\cdots+\lambda^+_{N^+}P^+_{N^+}=\sum_{s\in\{+,-\}}\sum_{i=1}^{N^s}\lambda^s_iP^s_i\,, \tag{5.28}$$

where $\lambda^-_{N^-}<\cdots<\lambda^-_1<0\leq\lambda^+_1<\cdots<\lambda^+_{N^+}$ and each P^s_i is the projector onto the eigenspace of distinct eigenvalue λ^s_i. Consider a QIVPM $\bar{\mu}\colon\mathcal{E}\to\mathcal{I}$ with its left-end and right-end denoted by $\mu^{\mathrm{L}}\colon\mathcal{E}\to[0,1]$ and $\mu^{\mathrm{R}}\colon\mathcal{E}\to[0,1]$, respectively, i.e., $\bar{\mu}(P)=[\mu^{\mathrm{L}}(P),\mu^{\mathrm{R}}(P)]$. Let s be either $+$ or $-$ in the superscript in the following discussion. We denote $P^s_{i\uparrow}$ as the accumulative projectors on the sub-index i, i.e., $P^s_{i\uparrow}=\sum_{j\geq i}P^s_j$, and $\Delta\mu^t\left(P^s_{i\uparrow}\right)$ as the difference of the measure of these accumulative projectors, i.e., $\Delta\mu^t\left(P^s_{i\uparrow}\right)=\mu^t\left(P^s_{i\uparrow}\right)-\mu^t\left(P^s_{(i+1)\uparrow}\right)$, where $t\in\{\mathrm{L},\mathrm{R}\}$. Then, the

expectation value of $\mathbf{O}$ relative to $\bar{\mu}$ is

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \sum_{s \in \{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\mu^{\mathrm{L}} \left(P_{i\uparrow}^s \right), \Delta\mu^{\mathrm{R}} \left(P_{i\uparrow}^s \right) \right] , \tag{5.29}$$

which looks like a weighted average of the difference of the measure of these accumulative projectors.

Recall Lemma 2.1 proved real expectation values are invariant when pulling back observables and probability measures. The following lemma shows interval expectation values are invariant under pullback as well.

Lemma 5.2. *Consider an observable* $\mathbf{O}$ *diagonalizable by an orthonormal basis* $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ *with* $\varphi \colon 2^{\Omega} \to \mathcal{E}$ *defined by Eq. (2.17). Given a QIVPM* $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$*, the expectation value of* $\mathbf{O}$ *relative to* $\bar{\mu}$ *is exactly the expectation value of the pullback of* $\mathbf{O}$ *relative to the pullback of* $\bar{\mu}$*, i.e.,*

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \int (\varphi^* \mathbf{O}) \, \mathrm{d} \, (\varphi^* \bar{\mu}) \ . \tag{5.30}$$

Proof. Consider expressing the observable $\mathbf{O}$ in the spectral decomposition specialized to compute the expectation value relative to a QIVPM, $\sum_{s \in \{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s P_i^s$, where $\lambda_{N^-}^- < \cdots < \lambda_1^- < 0 \leq \lambda_1^+ < \cdots < \lambda_{N^+}^+$ and each P_i^s is the projector onto the eigenspace of distinct eigenvalue λ_i^s. Since $\mathbf{O}$ can be diagonalized by an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$, each P_i^s can be expressed as the sum of some projectors formed by the elements in Ω, i.e., there is a subset $E_i^s \subseteq \Omega$ such that $P_i^s = \sum_{|j\rangle \in E_i^s} |j\rangle\langle j| = \varphi(E_i^s)$. To compute the expectation values, we want to not only map E_i^s to P_i^s but also map the other ingredients of the expectation values, including the accumulative sets or

projectors, their measures, and their differences:

$$
\begin{aligned}
P^s_{i\uparrow} &= \sum_{j\geq i} P^s_j = \sum_{j\geq i} \sum_{|k\rangle\in E^s_j} |k\rangle\langle k| = \sum_{|k\rangle\in\bigcup_{j\geq i} E^s_j} |k\rangle\langle k| = \varphi\left(\bigcup_{j\geq i} E^s_j\right) = \varphi\left(E^s_{i\uparrow}\right) , \\
\mu^t\left(P^s_{i\uparrow}\right) &= \mu^t\left(\varphi\left(E^s_{i\uparrow}\right)\right) = \left(\varphi^*\mu^t\right)\left(E^s_{i\uparrow}\right) , \qquad (5.31a) \\
\Delta\mu^t\left(P^s_{i\uparrow}\right) &= \mu^t\left(P^s_{i\uparrow}\right) - \mu^t\left(P^s_{(i+1)\uparrow}\right) = \left(\varphi^*\mu^t\right)\left(E^s_{i\uparrow}\right) - \left(\varphi^*\mu^t\right)\left(E^s_{(i+1)\uparrow}\right) \\
&= \Delta\left(\varphi^*\mu^t\right)\left(E^s_{i\uparrow}\right) , \qquad (5.31b)
\end{aligned}
$$

where $t \in \{\mathrm{L},\mathrm{R}\}$ and $\left[\mu^{\mathrm{L}}(P),\mu^{\mathrm{R}}(P)\right] = \bar{\mu}(P)$ as in Defs. 5.5 and 5.6. Notice that right-most side of Eq. (5.31a) and (5.31b) are actually the pullback of μ^{L} and μ^{R} which can be combined into the pullback of $\bar{\mu}$ as follows:

$$
\begin{aligned}
\left[\left(\varphi^*\mu^{\mathrm{L}}\right)(E),\left(\varphi^*\mu^{\mathrm{R}}\right)(E)\right] &= \left[\mu^{\mathrm{L}}\left(\varphi(E)\right),\mu^{\mathrm{R}}\left(\varphi(E)\right)\right] = \bar{\mu}\left(\varphi(E)\right) \\
&= \left(\varphi^*\bar{\mu}\right)(E) = \left[\left(\varphi^*\bar{\mu}\right)^{\mathrm{L}}(E),\left(\varphi^*\bar{\mu}\right)^{\mathrm{R}}(E)\right]
\end{aligned} \qquad (5.32)
$$

for all $E \subseteq \Omega$.

To compute the expectation value, we also need to know the pullback of $\mathbf{O}$. Since $\mathbf{O}$ can be expressed as the sum of one-dimensional projectors

$$
\mathbf{O} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda^s_i P^s_i = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda^s_i \sum_{|j\rangle\in E^s_i} |j\rangle\langle j| , \qquad (5.33)
$$

the pullback of $\mathbf{O}$ should be

$$
\varphi^*\mathbf{O} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda^s_i \sum_{|j\rangle\in E^s_i} \mathbf{1}_{\{|j\rangle\}} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda^s_i \mathbf{1}_{E^s_i} \qquad (5.34)
$$

according to Def. 2.1. Then, the expectation value can be computed by spelling the definitions and

applying Eqs. (5.31b), (5.32), and (5.34)

$$\begin{aligned}
\langle \mathbf{O} \rangle_{\bar{\mu}} &= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\mu^{\mathrm{L}} \left(P_{i\uparrow}^s \right), \Delta\mu^{\mathrm{R}} \left(P_{i\uparrow}^s \right) \right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta \left(\varphi^* \mu^{\mathrm{L}} \right) \left(E_{i\uparrow}^s \right), \Delta \left(\varphi^* \mu^{\mathrm{R}} \right) \left(E_{i\uparrow}^s \right) \right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta \left(\varphi^* \bar{\mu} \right)^{\mathrm{L}} \left(E_{i\uparrow}^s \right), \Delta \left(\varphi^* \bar{\mu} \right)^{\mathrm{R}} \left(E_{i\uparrow}^s \right) \right] \\
&= \int \left(\varphi^* \mathbf{O} \right) \mathrm{d} \left(\varphi^* \bar{\mu} \right) .
\end{aligned} \tag{5.35}$$

□

Since Def. 5.6 is an extension of both quantum real expectation value and classical interval expectation value, not only the property of quantum real expectation value could be extended to the previous lemma, but the properties of classical interval expectation value, like Eq. (5.25) and Thm. 5.3, could be extended to quantum interval expectation value as well. In particular, Eq. (5.25) can be extended to the following theorem.

Theorem 5.4. *Given a quantum probability measure $\mu \colon \mathcal{E} \to [0,1]$ satisfying Eqs. (2.21) and (2.22), if we define a QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$ by $\bar{\mu}(P) = [\mu(P), \mu(P)]$ for all projectors P, then the expectation value of any observable $\mathbf{O}$ relative to $\bar{\mu}$, $\langle \mathbf{O} \rangle_{\bar{\mu}}$, is just $\left[\langle \mathbf{O} \rangle_{\mu}, \langle \mathbf{O} \rangle_{\mu} \right]$.*

Although we can faithfully extend Eq. (5.25) to Thm. 5.4, this theorem is weaker than its classical counterpart, Thm. 5.3 because Lemma 5.1 hasn't established enough correspondence between classical and quantum cores.

Theorem 5.5. *Consider an observable $\mathbf{O}$ diagonalizable by an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ and a QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$.*

- *Express $\mathbf{O}$ in the spectral decomposition specialized to compute the expectation value relative to a QIVPM, $\sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s P_i^s$, where $\lambda_{N^-}^- < \cdots < \lambda_1^- < 0 \leq \lambda_1^+ < \cdots < \lambda_{N^+}^+$ and*

each P_i^s is the projector onto the eigenspace of distinct eigenvalue λ_i^s. Given any commuting subspace of events $\mathcal{E}_{\mathrm{C}}$ containing all projectors P_i^s, we have

$$\langle \mathbf{O} \rangle_{\bar{\mu}} \subseteq \left[\min_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{5.36}$$

- *Let $\mathcal{E}_\Omega$ be the set of projectors generated from Ω, $\{\varphi(E) \mid E \subseteq \Omega\}$, we have*

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \left[\min_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{5.37}$$

Proof. By Eq. (5.30) and Thm. 5.3, we have

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \int (\varphi^* \mathbf{O}) \, \mathrm{d}(\varphi^* \bar{\mu}) = \left[\min_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu, \max_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu \right], \tag{5.38}$$

where $\varphi \colon 2^\Omega \to \mathcal{E}$ is defined by Eq. (2.17). Hence, to prove Eq. (5.36), it is sufficient to prove

$$\left[\min_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu, \max_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu \right] \subseteq \left[\min_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{5.39}$$

Since $\mathcal{E}_\Omega$ contains all projectors P_i^s no matter how Ω is picked, $\mathcal{E}_\Omega$ is one of the possible choices of $\mathcal{E}_{\mathrm{C}}$. Thus, if Eq. (5.39) is true, to prove Eq. (5.37), it is sufficient to verify

$$\left[\min_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu, \max_{\mu \in \mathrm{core}(\varphi^* \bar{\mu})} \int (\varphi^* \mathbf{O}) \, \mathrm{d}\mu \right] \supseteq \left[\min_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{5.40}$$

PROOF OF EQ. (5.39) According to Lemma 5.1, for all $\mu \in \mathrm{core}(\varphi^* \bar{\mu})$, there is a density matrix $\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$ such that $\mu = \varphi^* \mu_\rho^{\mathrm{B}}$. Together with the properties of extremum and Eq. (2.26), we

have

$$\min_{\mu\in\text{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \geq \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}\ , \quad (5.41)$$

$$\max_{\mu\in\text{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \leq \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}\ , \quad (5.42)$$

which implies Eq. (5.39).

PROOF OF EQ. (5.40) Given a density matrix $\rho\in\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_{\Omega}\right)$, it must satisfy $\mu_\rho^{\mathrm{B}}\left(P\right)\in\bar{\mu}\left(P\right)$ for all $P\in\mathcal{E}_{\Omega}$. Since $\varphi\left(E\right)\in\mathcal{E}_{\Omega}$ for all $E\subseteq\Omega$, we have

$$\left(\varphi^*\mu_\rho^{\mathrm{B}}\right)(E) = \mu_\rho^{\mathrm{B}}\left(\varphi\left(E\right)\right)\in\bar{\mu}\left(\varphi\left(E\right)\right) = \left(\varphi^*\bar{\mu}\right)(E)\ , \quad (5.43)$$

i.e., $\varphi^*\mu_\rho^{\mathrm{B}}\in\text{core}\left(\varphi^*\bar{\mu}\right)$.

Together with the properties of extremum and Eq. (2.26), the previous paragraph implies

$$\min_{\mu\in\text{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \leq \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\Omega})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\Omega})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}\ , \quad (5.44)$$

$$\max_{\mu\in\text{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \geq \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\Omega})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\Omega})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}\ , \quad (5.45)$$

which then implies Eq. (5.40). □

5.2 THE KOCHEN-SPECKER THEOREM AND CONTEXTUALITY

Our generalization of quantum probability measures to QIVPMs allows us to strengthen the scope of one of the fundamental theorems of quantum physics: the Kochen-Specker theorem [36, 40, 41, 57, 58, 94, 95]. Our finite-precision extension of that theorem will suggest a resolution to the debate initiated by Meyer and Mermin on the relevance of the Kochen-Specker to experimental, and hence finite-precision, quantum measurements [96–108]. Specifically, the original Kochen-Specker theo-

rem is formulated using a model quantum mechanical system that *always has definite values* [95], i.e., its observables have infinitely precise values at all times. Our interval-valued probability framework will allow us to state and prove, a stronger version of the theorem that holds even if the observables have values that are only definite up to some precision specified by a parameter δ. Our approach provides a quantitative realization of Mermin's intuition [98]:

> …although the outcomes deduced from such imperfect measurements will occasionally differ dramatically from those allowed in the ideal case, if the misalignment is very slight, the statistical distribution of outcomes will differ only slightly from the ideal case.

5.2.1 FINITE-PRECISION EXTENSION OF THE KOCHEN-SPECKER THEOREM

The first step in our formalization is to introduce a family of QIVPMs parameterized by an uncertainty δ, which we call δ-deterministic QIVPMs.

Definition 5.7 (δ-Determinism)**.** A QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$ is δ-deterministic if, for every event $P \in \mathcal{E}$, we have that either $\bar{\mu}(P) \subseteq [0, \delta]$ or $\bar{\mu}(P) \subseteq [1-\delta, 1]$.

This definition puts no restrictions on the set of intervals itself, only on which intervals are assigned to events. When $\delta = 0$, every event must be assigned a probability either in $\mathbf{F}$ or in $\mathbf{T}$, i.e., whether every event happens is completely determined with certainty. As δ gets larger, the QIVPM allows for more indeterminate behavior.

The expectation value of an observable $\mathbf{O}$ in a Hilbert space H of dimension D relative to a 0-deterministic QIVPM is fully determinate and is equal to one of the eigenvalues λ_i of that observable. To see this, note that given an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$, a 0-deterministic QIVPM must map exactly one of the projectors $|i\rangle\langle i|$ to $\mathbf{T}$ and all others to $\mathbf{F}$. This is because, by Eq. (5.6b), we have $\bar{\mu}\left(\sum_{j=0}^{D-1} |j\rangle\langle j|\right) = \mathbf{T}$ and by inductively applying Eq. (5.7), we must have one of the $\bar{\mu}(|i\rangle\langle i|) = \mathbf{T}$ and all others mapped to $\mathbf{F}$. This unique projector mapping to $\mathbf{T}$ will be

denoted by $|\mathbf{T}\rangle\langle\mathbf{T}|$. Given any state ρ that is consistent with $\bar{\mu}$ on all the projectors in Ω, we have by Eq. (5.11) that μ_ρ^{B} must also map $|\mathbf{T}\rangle\langle\mathbf{T}|$ to 1 and all other projectors formed by elements in Ω to 0. If an observable has a spectral decomposition along Ω then, by Eq. (2.24), its expectation value relative to μ_ρ^{B} is the eigenvalue $\lambda_{\mathbf{T}}$ whose projector is $|\mathbf{T}\rangle\langle\mathbf{T}|$. It therefore follows, by Eq. (5.37), that the expectation value relative to the 0-deterministic $\bar{\mu}$ is fully determinate and lies in the interval $[\lambda_{\mathbf{T}}, \lambda_{\mathbf{T}}]$.

Given a Hilbert space H of dimension D, the expectation value of the product of a sequence of commuting observables $\{\mathbf{O}_j\}$ relative to the 0-deterministic QIVPM $\bar{\mu}$ is the product of the expectation value of individual $\mathbf{O}_j$. Since $\{\mathbf{O}_j\}$ is a sequence of commuting observables, they can be diagonalized by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ with spectral decompositions

$$\mathbf{O}_j = \sum_{i=0}^{D-1} \lambda_{j,i} |i\rangle\langle i| \ . \tag{5.46}$$

Because of the orthogonality of Ω, the product of Eqs. (5.46) can also be simplified as a spectral decomposition $\prod_j \mathbf{O}_j = \sum_{k=0}^{d-1} \prod_j \lambda_{j,i} |i\rangle\langle i|$. According to our discussion in the previous paragraph, there is a unique projector such that $\bar{\mu}(|\mathbf{T}\rangle\langle\mathbf{T}|) = \mathbf{T}$, and the expectation values relative to $\bar{\mu}$ are their eigenvalues of $|\mathbf{T}\rangle\langle\mathbf{T}|$, i.e., $\langle \mathbf{O}_j \rangle_{\bar{\mu}} = [\lambda_{j,\mathbf{T}}, \lambda_{j,\mathbf{T}}]$ and

$$\left\langle \prod_j \mathbf{O}_j \right\rangle_{\bar{\mu}} = \left[\prod_j \lambda_{j,\mathbf{T}}, \prod_j \lambda_{j,\mathbf{T}} \right] = \prod_j \langle \mathbf{O}_j \rangle_{\bar{\mu}} \tag{5.47}$$

with the understanding that the product of the singleton sets, $\prod_j [\lambda_{j,\mathbf{T}}, \lambda_{j,\mathbf{T}}]$, is defined to be the product of their elements, $\left[\prod_j \lambda_{j,\mathbf{T}}, \prod_j \lambda_{j,\mathbf{T}}\right]$.

We can now proceed with the main technical result of this section. We first observe that the original Kochen-Specker theorem is a statement regarding the non-existence of a 0-deterministic QIVPM and generalize to a corresponding statement about δ-deterministic QIVPMs.

Theorem 5.6 (0-Deterministic Variant of the Kochen-Specker Theorem)**.** *Given a Hilbert space $\mathcal{H}$*

of dimension $D \geq 3$, there is no 0-deterministic measure $\bar{\mu}$ mapping every event to either **F** *or* **T**.

To explain why this result is equivalent to the original Kochen-Specker theorem and to prove it at the same time, we proceed by assuming a 0-deterministic QIVPM $\bar{\mu}$ and derive the same contradiction as the original Kochen-Specker theorem. Instead of adapting the more complicated proof for $D = 3$, the counterexample presented below uses the simpler proof for a Hilbert space of dimension $D = 4$ and is constructed as follows.

Proof of Thm. 5.6. We consider a two spin-$\frac{1}{2}$ Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ of dimension $D = 4$. We use the same nine observables $\mathbf{O}_{ij}$ with i and j ranging over $\{0, 1, 2\}$ from the Mermin-Peres "magic square" used to prove the Kochen-Specker theorem [41, 66, 94]:

$\mathbf{O}_{ij}$	$j = 0$	$j = 1$	$j = 2$
$i = 0$	$\mathbb{1} \otimes \sigma_z$	$\sigma_z \otimes \mathbb{1}$	$\sigma_z \otimes \sigma_z$
$i = 1$	$\sigma_x \otimes \mathbb{1}$	$\mathbb{1} \otimes \sigma_x$	$\sigma_x \otimes \sigma_x$
$i = 2$	$\sigma_x \otimes \sigma_z$	$\sigma_z \otimes \sigma_x$	$\sigma_y \otimes \sigma_y$

The observables are constructed using the Pauli matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ whose eigenvalues are all either 1 or -1 [4, 10, 36, 40, 66]. They are arranged such that in each row and column, *except the column $j = 2$*, every observable is the product of the other two. In the $j = 2$ column, we have instead that $(\sigma_z \otimes \sigma_z)(\sigma_x \otimes \sigma_x) = -\sigma_y \otimes \sigma_y$. Now assume a 0-deterministic QIVPM $\bar{\mu}$; the expectation values of the observables in each row relative to this 0-deterministic QIVPM are fully determinate and must lie in either the interval $[1, 1]$ or the interval $[-1, -1]$ depending on which eigenvalue is the one whose associated projector is certain. Since the product of any two observables in a row is

equal to the third, the product of any two their expectation values in a row is also equal to the third by Eq. (5.47), and there must be an even number of occurrences of the interval $[-1,-1]$ in each row and hence in the entire table. However, looking at the expectation values of the observables in each column, by the same reason, there must be an even number of occurrences of the interval $[-1,-1]$ in the first two columns and an odd number in the $j=2$ column and hence in the entire table. The contradiction implies the non-existence of the assumed 0-deterministic QIVPM. □

Our framework allows us to generalize the above theorem to state that for small enough δ, it is impossible to have δ-deterministic QIVPMs, which is a stronger statement of contextuality that includes the effects of finite-precision. Every QIVPM must map some events to truly uncertain intervals, not just "almost definite intervals." The proof requires two simple lemmas that we present first.

The first lemma shows a simpler way to prove the convexity condition. Recall that the convexity condition for a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$ states that for each pair of *commuting* projectors P_0 and P_1 with $P_0 P_1 = P_1 P_0$, the following equation holds:

$$\bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) \subseteq \bar{\mu}\left(P_0\right) + \bar{\mu}\left(P_1\right) . \tag{5.48}$$

Lemma 5.3. *To verify the convexity condition of a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$, it is sufficient to check that:*

$$\bar{\mu}\left(P' + P''\right) = \bar{\mu}\left(P'\right) + \bar{\mu}\left(P''\right) \tag{5.49}$$

for all orthogonal projectors P' and P''.

Proof. The proof follows the outline of the proof of the classical inclusion-exclusion principle. From the commuting projectors P_0 and P_1, we construct the following three orthogonal projectors: $P_0 P_1$, $P_0\left(\mathbb{1} - P_1\right)$, and $\left(\mathbb{1} - P_0\right) P_1$. Then we proceed as follows:

$$
\begin{aligned}
&\bar{\mu}\left(P_0 + P_1 - P_0P_1\right) + \bar{\mu}\left(P_0P_1\right) \\
&\quad = \bar{\mu}\left(P_0P_1 + P_0(\mathbb{1} - P_1) + P_1 - P_0P_1\right) + \bar{\mu}\left(P_0P_1\right) && \text{(because } P_0 = P_0P_1 + P_0(\mathbb{1} - P_1)\text{)} \\
&\quad = \bar{\mu}\left(P_0(\mathbb{1} - P_1) + P_1\right) + \bar{\mu}\left(P_0P_1\right) \\
&\quad = \bar{\mu}\left(P_0(\mathbb{1} - P_1) + P_0P_1 + (\mathbb{1} - P_0)P_1\right) + \bar{\mu}\left(P_0P_1\right) && \text{(because } P_1 = P_0P_1 + (\mathbb{1} - P_0)P_1\text{)} \\
&\quad = \bar{\mu}(P_0(\mathbb{1} - P_1)) + \bar{\mu}(P_0P_1) + \bar{\mu}((\mathbb{1} - P_0)P_1) + \bar{\mu}\left(P_0P_1\right) && \text{(using Eq. (5.49) twice)} \\
&\quad = \bar{\mu}(P_0(\mathbb{1} - P_1) + P_0P_1) + \bar{\mu}((\mathbb{1} - P_0)P_1 + P_0P_1) && \text{(using Eq. (5.49) twice)} \\
&\quad = \bar{\mu}(P_0) + \bar{\mu}(P_1)
\end{aligned}
$$

□

The next lemma relates δ-deterministic QIVPMs with $\delta < \frac{1}{3}$ to 0-deterministic QIVPMs.

Lemma 5.4. *From any δ-deterministic QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$ with $\delta < \frac{1}{3}$, we can construct a 0-deterministic QIVPM $\bar{\mu}^{\mathrm{D}} \colon \mathcal{E} \to \{\mathsf{F}, \mathsf{T}\}$ defined as follows:*

$$
\bar{\mu}^{\mathrm{D}}\left(P\right) = \begin{cases} \mathsf{F} & \text{if } \bar{\mu}\left(P\right) \subseteq [0, \delta] \ ; \\ \\ \mathsf{T} & \text{if } \bar{\mu}\left(P\right) \subseteq [1 - \delta, 1] \ . \end{cases} \tag{5.50}
$$

Proof. The most important part of the proof is to verify the convexity condition for $\bar{\mu}^{\mathrm{D}}$. By Lemma 5.3, it is sufficient to verify the following equation for orthogonal projectors P' and P'',

$$
\bar{\mu}^{\mathrm{D}}\left(P' + P''\right) = \bar{\mu}^{\mathrm{D}}\left(P'\right) + \bar{\mu}^{\mathrm{D}}\left(P''\right) \ , \tag{5.51}
$$

for two cases, which we now examine in detail.

When one of $\bar{\mu}^{\mathrm{D}}\left(P'\right)$ and $\bar{\mu}^{\mathrm{D}}\left(P''\right)$ is T, say $\bar{\mu}^{\mathrm{D}}\left(P'\right) = \mathsf{F}$ and $\bar{\mu}^{\mathrm{D}}\left(P''\right) = \mathsf{T}$, we have $\bar{\mu}\left(P'\right) \subseteq [0, \delta]$ and $\bar{\mu}\left(P''\right) \subseteq [1 - \delta, 1]$ which implies $\bar{\mu}\left(P' + P''\right) \subseteq [1 - \delta, 1 + \delta]$. Since $\bar{\mu}\left(P' + P''\right)$ is a subset of $[0, 1]$, $\bar{\mu}\left(P' + P''\right)$ must be a subset of $[1 - \delta, 1]$, which implies $\bar{\mu}^{\mathrm{D}}\left(P' + P''\right)$ is also T,

Table 5.2: Possible probability measures on a Hilbert space of dimension $D = 3$, where $\bar{\mu}_2'$ and $\bar{\mu}_3$ are QIVPMs while $\bar{\mu}_0$, $\bar{\mu}_1$, and $\bar{\mu}_2$ are not. Events are listed in the column labeled by P.

P	$\bar{\mu}_0(P)$	$\bar{\mu}_1(P)$	$\bar{\mu}_2(P)$	$\bar{\mu}_2'(P)$	$\bar{\mu}_3(P)$
$\mathbb{0}$	**F**	**F**	**F**	**F**	**F**
All one-dimensional projectors	$[0,0]$	$[0,\frac{1}{4}]$	$[0,\frac{1}{3}]$	$[\frac{1}{3},\frac{1}{3}]$	$[0,\frac{1}{2}]$
All two-dimensional projectors	$[1,1]$	$[\frac{3}{4},1]$	$[\frac{2}{3},1]$	$[\frac{2}{3},\frac{2}{3}]$	$[\frac{1}{2},1]$
$\mathbb{1}$	**T**	**T**	**T**	**T**	**T**

thus satisfying Eq. (5.51).

When both $\bar{\mu}^{\mathrm{D}}(P')$ and $\bar{\mu}^{\mathrm{D}}(P'')$ are **F**, we have both $\bar{\mu}(P')$ and $\bar{\mu}(P'') \subseteq [0,\delta]$ which implies $\bar{\mu}(P'+P'') \subseteq [0,2\delta]$. Since we assume $\delta < \frac{1}{3}$, the intervals $[0,2\delta]$ and $[1-\delta,1]$ are disjoint, which implies $\bar{\mu}(P'+P'')$ and $[1-\delta,1]$ are disjoint. Together with the fact that $\bar{\mu}(P'+P'')$ is a subset of either $[0,\delta]$ or $[1-\delta,1]$, $\bar{\mu}(P'+P'')$ must be a subset of $[0,\delta]$, which implies $\bar{\mu}^{\mathrm{D}}(P'+P'') =$ **F**, and hence also Eq. (5.51) is again satisfied. □

Theorem 5.7 (Finite-precision Extension of the Kochen-Specker Theorem)**.** *Given a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, there is no δ-deterministic QIVPM for $\delta < \frac{1}{3}$.*

Proof by Contradiction. Suppose there is a δ-deterministic QIVPM $\bar{\mu} : \mathcal{E} \to \mathcal{I}$. By Lemma 5.4, we can construct a 0-deterministic QIVPM; however, by Thm. 5.6, such 0-deterministic QIVPMs do not exist. □

The bound $\delta < \frac{1}{3}$ is tight as it is possible to construct a $\frac{1}{3}$-deterministic QIVPM $\bar{\mu} : \mathcal{E} \to \mathcal{I}$. For example, $\bar{\mu}_2'$ defined in Table 5.2 is a valid $\frac{1}{3}$-deterministic QIVPM. When $\delta \geq \frac{1}{3}$, i.e., when the uncertainty in measurements becomes so large, it becomes possible to map every observable to some (quite inaccurate) probability interval, thus invalidating the Kochen-Specker theorem. We can summarize and illustrate the above arguments using Fig. 5.1.

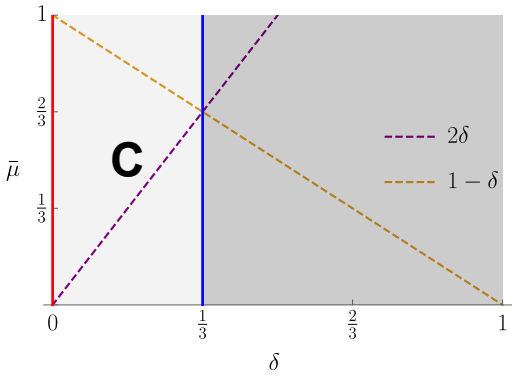


Figure 5.1: The region to the left of the vertical line at $\delta = \frac{1}{3}$ is where we assume small measurement degradation; in that region, our extension of the KS theorem demonstrates contextuality (C). In the region to the right, the degradation of the data is large, and our extension of the KS theorem no longer refutes other explanations for the experimental data.

As is the case for conventional, infinitely-precise, quantum probability measures, the theorem is only applicable to dimensions $D \geq 3$. Indeed, when the Hilbert space has dimension 2, it is straightforward to construct a 0-deterministic QIVPM as follows. Consider a non-contextual hidden variable model for $D = 2$ (e.g., as proposed by Bell or Kochen-Specker [57, 58]). Such a two-dimensional model always assigns definite values to all observables and hence assigns a *determinate* probability (0 or 1) to each event. This probability measure directly induces a 0-deterministic QIVPM by changing 0 to **F** and 1 to **T**. It follows that every 0-deterministic QIVPM is δ-deterministic.

5.2.2 EXPERIMENTAL DATA AND δ-DETERMINISM

We have thus quantified one important aspect of uncertainty in quantum mechanics—the effect of the imprecise nature of devices—which is a novel addition to the theory of measurement. Indeed, as Heisenberg emphasized in his famous microscope example [109], the conventional theory of measurement states that it is impossible to precisely measure any property of a system without disturbing it somewhat. Thus, there are fundamental limits to what one can measure and these limits have traditionally been attributed to complementarity. Our imprecision represents an *additional* source of indeterminacy beyond the inherent probabilistic nature of quantum mechanics.

In an experimental setup, δ is calculated as follows. To determine the probability of any event, we typically repeat an experiment m times and count the number of times we witness the event. This assumes that for each run of the experiment we can determine, using our apparatus, whether the event occurred or not. Assume an event has an ideal mathematical probability of 0, and we repeat the experiment 100 times. In a perfect world, we should be able to refute the event 100 times and calculate that the probability is 0. We might also observe the event 2 times and refute it 98 times and therefore calculate the probability to be 0.02. Note that this situation assumes perfect measurement conditions and remains within the context of conventional (real-valued) probability theory. The question we focus on is what happens if we are only able to refute it 97 times and are *uncertain* 3 times? This is quite common in actual experiments. Mathematically we can model this idea by stating that the probability of the event is in the range $[0, 0.03]$ which says that the probability of the event could be 0, 0.01, 0.02, or 0.03 as each the three uncertain records could either be evidence for the event or against it. We just cannot nail it down given the current experimental results and therefore represent the evidence as a ($\delta =$)0.03-deterministic probability measure. The interesting observation is that the axioms of probability theory (like additivity and convexity) impose enough constraints on the structure of interval-valued quantum probability measures to make them robust in the face of small non-vanishing δ's.

To see this idea in the context of a quantum experiment, consider a three-dimensional Hilbert space with one-dimensional projectors P_ρ, two-dimensional projectors $P_\rho + P_\sigma$, and an experiment that is repeated 12 times. By the Kochen-Specker theorem, it is impossible to build a probability measure that maps every projection to either $0 = \frac{0}{12}$ or $1 = \frac{12}{12}$. That is, the assignment $\bar{\mu}_0$ defined in Table 5.2 is not a QIVPM.

Now consider what happens if $\frac{1}{4}$ of the data for *every* one-dimensional projector is uncertain. A potential account of this degradation is to assign to each event P the entire range of possibilities $\bar{\mu}_1(P)$ as defined in Table 5.2. This measure is not a valid QIVPM because it does not satisfy

the convexity condition: for any two orthogonal one-dimensional events P_0 and P_1, the convexity condition requires $\bar{\mu}_1(P_0 + P_1) \subseteq \bar{\mu}_1(P_0) + \bar{\mu}_1(P_1)$, but $\bar{\mu}_1(P_0 + P_1) = [\frac{3}{4}, 1]$ which is not a subset of $[0, \frac{1}{2}] = \bar{\mu}_1(P_0) + \bar{\mu}_1(P_1)$. Interestingly, it is impossible to find any probability measure that would be consistent with these observations, as the interval $[\frac{3}{4}, 1]$ is completely disjoint from the interval $[0, \frac{1}{2}]$ and no amount of shifting of assumptions regarding the precise outcome of the uncertain observations could change that disjointness. However, as shown next, a sharp transition occurs when $\delta = \frac{1}{3}$.

When the proportion of uncertain data reaches $\frac{1}{3}$, the probability measure that assigns to each event the entire range of possibilities is $\bar{\mu}_2$ defined in Table 5.2. This is also not a valid probability measure by the same argument as above. However, in this case, $\bar{\mu}_2(P_0 + P_1) = [\frac{2}{3}, 1]$ and $[0, \frac{2}{3}] = \bar{\mu}_2(P_0) + \bar{\mu}_2(P_1)$ have a *common point*. Hence, by assuming that the uncertain data for one-dimensional projectors always support the associated event, while those for two-dimensional projectors always refute the event, we can find the probability measure $\bar{\mu}_2'$ that can be verified as a valid QIVPM and is consistent with the experimental data.

A similar situation happens when more than $\frac{1}{3}$ of data is uncertain. In particular, if half of the data is uncertain, the probability measure $\bar{\mu}_3$ that assigns to each event the entire range of possibilities is already a QIVPM.

5.3 THE BORN RULE AND GLEASON'S THEOREM

A conventional quantum probability measure can be easily constructed from a state ρ according to the Born rule [4, 36, 44]. According to Gleason's theorem [39–41], this state ρ is also the unique state consistent with any possible probability measure.

5.3.1 FINITE-PRECISION EXTENSION OF GLEASON'S THEOREM

In order to re-examine these results in our framework, we first reformulate Gleason's theorem in QIVPMs using infinitely precise uncountable intervals $\mathcal{I}_\infty = \{[x,x] \mid x \in [0,1]\}$:

Theorem 5.8 ($\mathcal{I}_\infty$ Variant of the Gleason Theorem)**.** *In a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, given a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}_\infty$, the state ρ consistent with $\bar{\mu}$ on every projector is unique, i.e., there exists a unique state ρ such that $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}) = \{\rho\}$.*

Now let us consider relaxing $\mathcal{I}$ to a countable set of finite-width intervals. As the intervals in the image of a QIVPM become less and less sharp, we expect more and more states to be consistent with it. In the limit of minimal sharpness, all states ρ are consistent with the QIVPM

$$\bar{\mu}(P) = \begin{cases} \mathbf{F} & \text{if } P = \mathbb{0}\,; \\ \mathbf{T} & \text{if } P = \mathbb{1}\,; \\ \mathbf{U} = [0,1] & \text{otherwise} \end{cases} \tag{5.52}$$

mapping nearly all projections to the *unknown* interval $\mathbf{U}$. There is however a subtlety: as we will show in Thm. 5.9 later, it is possible for an arbitrary assignment of intervals to projectors to be globally inconsistent, but before proving Thm. 5.9, we need other two lemmas to simplify the proof of the convexity condition again.

Lemma 5.5. *Given a Hilbert space $\mathcal{H}$ of dimension 3, to verify $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$ is a QIVPM, it is sufficient to check Eqs. (5.6) and*

$$\bar{\mu}(P' + P'') \subseteq \bar{\mu}(P') + \bar{\mu}(P'') \tag{5.53}$$

for each pair of orthogonal projectors P' and P''.

Proof. The most important part of the proof is to verify the convexity condition for $\bar{\mu}$. Given a pair of commuting projectors P_0 and P_1 on a three-dimensional Hilbert space, they can be diagonalized

by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, |2\rangle\}$. Consider the function $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined in Eq. (2.17), there are two set of basis vectors E_0 and $E_1 \subseteq \Omega$, such that $\varphi(E_0) = P_0$ and $\varphi(E_1) = P_1$. Since E_0 and E_1 are both subsets of a three-element set, their relation has only three possibilities. The first possibility is that one of them is a subset of the other one, $E_0 \subseteq E_1$ or $E_1 \subseteq E_0$. The second possibility is that they are disjoint, $E_0 \cap E_1 = \emptyset$. If neither of the previous possibilities is true, i.e., they have some intersections, but no subset relation, then $E_0 \cap E_1$, $E_0 \backslash E_1$, and $E_1 \backslash E_0$ are all non-empties. Together with the fact that Ω has only three elements, they are all singleton sets. These three possibilities are going to be discussed as follow.

- When one of them is a subset of the other one, say $E_0 \subseteq E_1$, we have $P_0 P_1 = \varphi(E_0 \cap E_1) = P_0$ and $P_0 + P_1 - P_0 P_1 = P_1$. Thus,

$$\bar{\mu}(P_0 + P_1 - P_0 P_1) + \bar{\mu}(P_0 P_1) = \bar{\mu}(P_1) + \bar{\mu}(P_0) \ . \tag{5.54}$$

- When $E_0 \cap E_1 = \emptyset$, we have $P_0 P_1 = \varphi(E_0 \cap E_1) = \mathbb{0}$ and

$$\bar{\mu}(P_0 + P_1 - P_0 P_1) + \bar{\mu}(P_0 P_1) = \bar{\mu}(P_0 + P_1) \subseteq \bar{\mu}(P_0) + \bar{\mu}(P_1) \tag{5.55}$$

  by Eq. (5.53).

- When $E_0 \cap E_1$, $E_0 \backslash E_1$, and $E_1 \backslash E_0$ are all singleton sets, say $E_0 \backslash E_1 = \{|0\rangle\}$, $E_1 \backslash E_0 = \{|1\rangle\}$, and $E_0 \cap E_1 = \{|2\rangle\}$, proving an equivalent condition for the convexity condition, Eq. (5.7), is easier than proving Eq. (5.7) directly. Since one minus an interval maps this interval to its mirror image, and reflection preserves the subset relations, the convexity condition holds if and only if

$$\mathbf{T} - \bar{\mu}(P_0 P_1) + \mathbf{T} - \bar{\mu}(P_0 + P_1 - P_0 P_1) \subseteq \mathbf{T} - \bar{\mu}(P_0) + \mathbf{T} - \bar{\mu}(P_1) \tag{5.56}$$

which is equivalent to

$$\bar{\mu}\left(\mathbb{1}-P_0P_1\right)+\bar{\mu}\left(\mathbb{1}-\left(P_0+P_1-P_0P_1\right)\right)\subseteq\bar{\mu}\left(\mathbb{1}-P_0\right)+\bar{\mu}\left(\mathbb{1}-P_1\right) \tag{5.57}$$

because of Eq. (5.6c). The last equation holds because we can apply Eq. (5.53) on the following chain of equations:

$$\begin{aligned}&\bar{\mu}\left(\mathbb{1}-P_0P_1\right)+\bar{\mu}\left(\mathbb{1}-\left(P_0+P_1-P_0P_1\right)\right)=\bar{\mu}\left(|0\rangle\langle 0|+|1\rangle\langle 1|\right)+\bar{\mu}\left(\mathbb{0}\right)\\ \subseteq\ &\bar{\mu}\left(|0\rangle\langle 0|\right)+\bar{\mu}\left(|1\rangle\langle 1|\right)=\bar{\mu}\left(\mathbb{1}-P_0\right)+\bar{\mu}\left(\mathbb{1}-P_1\right)\ .\end{aligned} \tag{5.58}$$

Since the convexity condition holds for all three possibilities, $\bar{\mu}$ is a QIVPM. □

Lemma 5.6. *Given a Hilbert space $\mathcal{H}$ of dimension 3, to verify $\bar{\mu}\colon \mathcal{E}\to\mathcal{I}$ is a QIVPM, it is sufficient to check Eqs. (5.6) and*

$$\bar{\mu}\left(|\psi'\rangle\langle\psi'|+|\psi''\rangle\langle\psi''|\right)\subseteq\bar{\mu}\left(|\psi'\rangle\langle\psi'|\right)+\bar{\mu}\left(|\psi''\rangle\langle\psi''|\right) \tag{5.59}$$

for each pair of orthogonal states $|\psi'\rangle$ and $|\psi''\rangle$.

Proof. Since any projectors can be expressed as the sum of orthogonal one-dimensional projectors, Eq. (5.59) implies Eq. (5.53) by induction, and this lemma holds because of Lemma 5.5. □

After we proved the lemmas, we can state and prove the theorem that some assignment of intervals to projectors can be globally inconsistent..

Theorem 5.9 (Empty Cores Exist for General QIVPMs)**.** *There exists a Hilbert space $\mathcal{H}$ and a QIVPM $\bar{\mu}\colon \mathcal{E}\to\mathcal{I}$ such that $\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}\right)=\emptyset$.*

Proof. To prove this theorem, we need to construct a QIVPM on some Hilbert space and verify that there are no states that are consistent (see Def. 5.4) with it on all possible events. Assume a Hilbert

Table 5.3: QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}_0$ on a Hilbert space of dimension $D = 3$. Events are listed in the column labeled by P.

P	$\bar{\mu}(P)$
$\mathbb{0}$, $\|0\rangle\langle 0\|$, $\|+\rangle\langle +\|$, $\|+'\rangle\langle +'\|$	**F**
$\mathbb{1}$, $\mathbb{1} - \|0\rangle\langle 0\|$, $\mathbb{1} - \|+\rangle\langle +\|$, $\mathbb{1} - \|+'\rangle\langle +'\|$	**T**
All other projectors	**U**

space of dimension $D = 3$ with orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$, let $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, $|+'\rangle = \frac{|0\rangle+|2\rangle}{\sqrt{2}}$, and assign

$$\mathcal{I}_0 = \{\mathbf{T}, \mathbf{F}, \mathbf{U}\} \ . \tag{5.60}$$

Consider the map $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}_0$ defined in Table 5.3. We want to prove $\bar{\mu}$ is a QIVPM. Since it is easy to verify $\bar{\mu}$ satisfies Eqs. (5.6), it is sufficient by Lemma 5.6 to verify

$$\bar{\mu}\left(|\psi'\rangle\langle\psi'| + |\psi''\rangle\langle\psi''|\right) \subseteq \bar{\mu}\left(|\psi'\rangle\langle\psi'|\right) + \bar{\mu}\left(|\psi''\rangle\langle\psi''|\right) \tag{5.61}$$

for each pair of orthogonal states $|\psi'\rangle$ and $|\psi''\rangle$. Since $|0\rangle$, $|+\rangle$, and $|+'\rangle$ are not orthogonal to each other, at least one of $\bar{\mu}\left(|\psi'\rangle\langle\psi'|\right)$ and $\bar{\mu}\left(|\psi''\rangle\langle\psi''|\right)$ is unknown **U**, which implies $\mathbf{U} \subseteq \bar{\mu}\left(|\psi'\rangle\langle\psi'|\right) + \bar{\mu}\left(|\psi''\rangle\langle\psi''|\right)$. Together with the fact that every interval in $\mathcal{I}_0$ is a subset of **U**, Eq. (5.61) holds, and $\bar{\mu}$ is a QIVPM.

Next, we will prove by contradiction that $\overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}\right)$ is the empty set. Suppose there is a state $\rho = \sum_{j=1}^{N} q_j \left|\phi_j\right\rangle\left\langle\phi_j\right| \in \overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}\right)$, where $\sum_{j=1}^{N} q_j = 1$ and $q_j > 0$. Since we assumed the core $\overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}\right)$ is non-empty, so $\mu_\rho^{\mathrm{B}}(P) \in \bar{\mu}(P)$, and Table 5.3 tells us that $\bar{\mu}(|0\rangle\langle 0|) = \mathbf{F} = [0, 0]$, we must conclude that $\mu_\rho^{\mathrm{B}}(|0\rangle\langle 0|) = 0 \in [0, 0]$, and similarly for $|+\rangle\langle +|$ and $|+'\rangle\langle +'|$. If this is true, then

$\langle 0|\phi_j\rangle = \langle +|\phi_j\rangle = \langle +'|\phi_j\rangle = 0$ for all j, and thus

$$\langle 1|\phi_j\rangle = \sqrt{2}\,\langle +|\phi_j\rangle - \langle 0|\phi_j\rangle = 0\,, \qquad \langle 2|\phi_j\rangle = \sqrt{2}\,\langle +'|\phi_j\rangle - \langle 0|\phi_j\rangle = 0\,. \tag{5.62}$$

The above equations imply $|\phi_j\rangle = |0\rangle\,\langle 0|\phi_j\rangle + |1\rangle\,\langle 1|\phi_j\rangle + |2\rangle\,\langle 2|\phi_j\rangle = 0$, violating the assumption that $|\phi_j\rangle$ is a normalized state, and thus the theorem is proved. □

The fact that a collection of poor measurements on a quantum system cannot reveal the underlying state is not surprising. Under certain conditions, we can however guarantee that the uncertainty in measurements is consistent with *some* non-empty collection of quantum states. Furthermore, we can relate the uncertainty in measurements to the volume of quantum states such that, in the limit of infinitely precise measurements, the volume of states collapses to a single state.

To that end, we introduce the concept of *interval maps*, which we can use to construct a consistent family of QIVPMs. An interval map $f\colon [0,1] \to \mathcal{I}$ maps every real-valued probability $x \in [0,1]$ to a set of intervals $f(x) = [\ell, r]$ containing x, where $[0,1]$ denotes the set of real-valued probabilities (this should not be confused with the interval-valued probability $\mathbf{U}$). We also need a notion of *norm* to quantify the uncertainty in measurements and the distance between (pure or mixed) states. The norm of a collection of intervals $\mathcal{I}$, $\|\mathcal{I}\|$, is defined as the maximum length of intervals in it. The norm of a pure state $\rho = |\psi\rangle\langle\psi|$ is defined as usual by $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$. For any given Hermitian operator A, we choose the operator norm $\|A\| = \max_{\||\psi\rangle\|=1} \|A\,|\psi\rangle\|$, which is also known as the 2-norm or the spectral norm [41, 110–112]. In fact, for any such matrix, including the density matrix ρ, this norm is the maximum absolute value of its eigenvalues. Then, a finite-precision extension of Gleason's theorem can be stated as follows.

Theorem 5.10 (Finite-Precision Extension of the Gleason Theorem)**.** *Let $f\colon [0,1] \to \mathcal{I}$ be an interval map and let the composition $f \circ \mu_\rho^{\mathrm{B}}$ be a QIVPM, where μ_ρ^{B} is the probability measure induced by the Born rule for a given state ρ. If a state ρ' is consistent with $f \circ \mu_\rho^{\mathrm{B}}$ on all events, i.e.,*

$\rho' \in \overline{\mathcal{H}}\left(f \circ \mu_\rho^{\mathrm{B}}, \mathcal{E}\right)$*, then the norm of their difference is bounded by* $\|\mathcal{I}\|$*, i.e.,* $\|\rho - \rho'\| \leq \|\mathcal{I}\|$.

Proof. Given a state ρ' consistent with $f \circ \mu_\rho^{\mathrm{B}}$, we have $\mu_{\rho'}^{\mathrm{B}}\left(|\psi\rangle\langle\psi|\right) \in f\left(\mu_\rho^{\mathrm{B}}\left(|\psi\rangle\langle\psi|\right)\right)$ for any one-dimensional projector $P = |\psi\rangle\langle\psi|$. Since the maximum length of the intervals in $\mathcal{I}$ is $\|\mathcal{I}\|$, it is also the upper bound of the difference:

$$\left|\mu_{\rho'}^{\mathrm{B}}\left(|\psi\rangle\langle\psi|\right) - \mu_\rho^{\mathrm{B}}\left(|\psi\rangle\langle\psi|\right)\right| = |\langle\psi|\rho - \rho'|\psi\rangle| \leq \|\mathcal{I}\| \ . \tag{5.63}$$

Since $\rho - \rho'$ is Hermitian, $\max_{\|\psi\|=1} |\langle\psi|\rho - \rho'|\psi\rangle|$ is the maximum absolute value of the eigenvalues of $\rho - \rho'$ [10], and equal to $\|\rho - \rho'\|$ [111, 112]. Hence, $\|\rho - \rho'\| \leq \|\mathcal{I}\|$. □

5.3.2 ULTRAMODULAR FUNCTIONS

Theorem 5.10 generalizes Gleason's theorem in the sense that it accounts for a larger class of probability measures that includes the conventional one as a limit. The theorem is however "special" in the sense that it only applies to the particular class of QIVPMs constructed by composing an interval map with a conventional quantum probability measure. QIVPMs constructed in this manner have some peculiar properties that we examine next.

An interval map is called *ultramodular* if it satisfies the following properties.

Definition 5.8 (Ultramodular Functions)**.** Given a collection of intervals $\mathcal{I}$ including $\mathbf{F}$ and $\mathbf{T}$, an interval map $\mathcal{M} : [0,1] \to \mathcal{I}$ is called ultramodular if

$$\mathcal{M}(0) = \mathbf{F}\,, \qquad \mathcal{M}(1) = \mathbf{T}\,, \qquad \mathcal{M}\left(1 - x\right) = \mathbf{T} - \mathcal{M}\left(x\right)\,, \tag{5.64}$$

and for any three numbers x_0, x_1, and $x_2 \in [0,1]$ such that $y = x_0 + x_1 + x_2 \in [0,1]$, we have

$$\mathcal{M}\left(y\right) + \mathcal{M}\left(x_2\right) \subseteq \mathcal{M}\left(x_0 + x_2\right) + \mathcal{M}\left(x_1 + x_2\right)\,. \tag{5.65}$$

The first three constraints, Eqs. (5.64), are the direct counterpart of the corresponding QIVPM constraints, Eqs. (5.6); the last condition, Eq. (5.65), is the direct counterpart of the convexity conditions, Eqs. (5.3) and (5.7) [54–56, 91]. Therefore, these conditions guarantee that for any conventional quantum probability measure μ, the composition $\mathcal{M} \circ \mu$ defines a valid QIVPM. Conversely, if for every quantum probability measure μ, it is the case that $f \circ \mu$ is a QIVPM, then the interval map f is an ultramodular function. Formally, we have the following result:

Theorem 5.11 (Equivalence of Ultramodular Functions and IVPMs)**.** *The following three statements are equivalent:*

1. *A function $\mathcal{M} : [0,1] \to \mathcal{I}$ is ultramodular.*
2. *The composite function $\mathcal{M} \circ \mu : 2^{\Omega} \to \mathcal{I}$ is a classical IVPM for all classical probability measures $\mu : 2^{\Omega} \to [0,1]$.*
3. *The composite function $\mathcal{M} \circ \mu : \mathcal{E} \to \mathcal{I}$ is a QIVPM for all quantum probability measures $\mu : \mathcal{E} \to [0,1]$.*

Proof. Statement 1 implies 2 and 3 as we have outlined above. Conversely, for the quantum case, we want to show that if $\mathcal{M}$ is not ultramodular, then for some quantum probability measure μ, the composite $\mathcal{M} \circ \mu$ might not be a QIVPM. Suppose there are three particular numbers x_0, x_1, and $x_2 \in [0,1]$ such that $y = x_0 + x_1 + x_2 \in [0,1]$, but they don't satisfy Eq. (5.65). Consider the state:

$$\rho = x_0 \left|0\right\rangle\!\left\langle 0\right| + x_1 \left|1\right\rangle\!\left\langle 1\right| + x_2 \left|2\right\rangle\!\left\langle 2\right| + (1-y) \left|3\right\rangle\!\left\langle 3\right| . \tag{5.66}$$

The induced map $\mathcal{M} \circ \mu_{\rho}^{\mathrm{B}}$ constructed using the Born rule and blurred by $\mathcal{M}$ fails to satisfy Eq. (5.7) when $P_0 = \left|0\right\rangle\!\left\langle 0\right| + \left|2\right\rangle\!\left\langle 2\right|$ and $P_1 = \left|1\right\rangle\!\left\langle 1\right| + \left|2\right\rangle\!\left\langle 2\right|$. In other words, this induced map fails to be a QIVPM.

For the classical case, if $\mathcal{M}$ is not ultramodular, we also want to find a classical probability measure $\mu : 2^{\Omega} \to [0,1]$ such that $\mathcal{M} \circ \mu$ is not a classical IVPM. Consider an orthonormal basis

$\Omega = \{|0\rangle, |1\rangle, ..., |D-1\rangle\}$ and $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined by Eq. (2.17). Notice that the pullback of our previous quantum probability measure μ_{ρ}^{B}, $\varphi^{*}\mu_{\rho}^{\mathrm{B}}$, is a classical probability measure. If we pick μ as $\varphi^{*}\mu_{\rho}^{\mathrm{B}}$, then the induced map $\mathcal{M} \circ \mu$ fails to be a classical IVPM for the same reason as in the quantum case. □

In other words, the essential properties of QIVPMs constructed using interval maps can be gleaned from the properties of ultramodular functions. The following is the most interesting property in our setting.

Theorem 5.12 (Range of Ultramodular Functions)**.** *For any ultramodular function $\mathcal{M}\colon [0,1] \to \mathcal{I}$, either $\mathcal{I} = \mathcal{I}_0$ as defined in Eq. (5.60) or $\mathcal{I}$ contains uncountably many intervals.*

Proof. Since $\mathcal{M}$ maps to intervals, we can decompose it into two functions: its left-end and right-end, where $[\mathcal{M}^L(x), \mathcal{M}^R(x)] = \mathcal{M}(x)$. By Eq. (5.65), the left-end function $\mathcal{M}^L\colon [0,1] \to [0,1]$ is Wright-convex [110, 113, 114], i.e.,

$$\mathcal{M}^L(y) + \mathcal{M}^L(x_2) \geq \mathcal{M}^L(x_0 + x_2) + \mathcal{M}^L(x_1 + x_2) \tag{5.67}$$

for three numbers x_0, x_1, and $x_2 \in [0,1]$ with $y = x_0 + x_1 + x_2 \in [0,1]$. Together with the fact that $\mathcal{M}^L$ maps to a bounded interval $[0,1]$, the left-end function $\mathcal{M}^L$ must be continuous on the unit open interval $(0,1)$ [91]. Therefore, either $\mathcal{M}$ maps every number in $(0,1)$ to the same interval, or the number of intervals to which $\mathcal{M}$ maps must be uncountable. □

To summarize, a conventional quantum probability measure has an uncountable range $[0,1]$. A QIVPM constructed by blurring such a conventional quantum probability measure must also have an uncountable range of intervals. Of course, any particular QIVPM, or any particular experiment, will use a fixed collection of intervals appropriate for the resources and precision of the particular experiment.

Chapter 6

FURTHER DISCUSSION

6.1 GLEASON'S THEOREM FOR GENERAL QIVPMS

As we discussed in the previous section, Thm. 5.10 only applies to the QIVPMs constructed by composing an interval map with a conventional quantum probability measure, and the states consistent with the composite QIVPM collapses to a single state as the maximum length of intervals in $\mathcal{I}$, $\|\mathcal{I}\|$, shrinks to 0. In contrast, the globally inconsistent QIVPM defined in Table 5.3 has the least sharp range $\mathcal{I}_0$ with $\|\mathcal{I}_0\| = 1$. This suggests a possibility that shrinking the length $\|\mathcal{I}\|$ might help to regularize general QIVPMs, and it is natural to ask whether there is a short enough length ε such that QIVPMs mapping to intervals not longer than ε always have non-empty cores.

Question 6.1. Given a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, is there an $\varepsilon > 0$ such that for all QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathcal{I}$ satisfying $\|\mathcal{I}\| \leq \varepsilon$, $\bar{\mu}$ must have a non-empty core, i.e., $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}) \neq \emptyset$?

To better understand this question, consider the $D = 3$ situation, where any two-dimensional projectors can be expressed as the complement of a one-dimensional projector, and by Eq. (5.6c) so does their interval-valued probabilities, i.e.,

$$\bar{\mu}(\mathbb{1} - |\psi\rangle\langle\psi|) = \mathsf{T} - \bar{\mu}(|\psi\rangle\langle\psi|) \ . \tag{6.1}$$

Hence, a QIVPM is completely determined by its values on the one-dimensional projectors which are one-to-one corresponding to the irreducible states, and these irreducible states are encoded in the

complex projective space $\mathbb{C}\mathbf{P}^2$ as we discussed in Sec. 2.1.3. Therefore, to study a QIVPM $\bar{\mu}$, we can just study a pair of functions $f^L\colon \mathbb{C}\mathbf{P}^2 \to [0,1]$ and $f^R\colon \mathbb{C}\mathbf{P}^2 \to [0,1]$ defined by $\left[f^L\left(|\psi\rangle\right), f^R\left(|\psi\rangle\right)\right] = \bar{\mu}\left(|\psi\rangle\langle\psi|\right)$ for any irreducible state $|\psi\rangle \in \mathbb{C}\mathbf{P}^2$. According to Lemma 5.6, $\bar{\mu}\colon \mathcal{E} \to \mathcal{I}$ is a QIVPM if and only if $\bar{\mu}$ satisfies Eqs. (5.6) and

$$\bar{\mu}\left(\mathbb{1} - |\psi_0\rangle\langle\psi_0|\right) \subseteq \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) + \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \tag{6.2}$$

for all orthonormal basis $\{|\psi_i\rangle\}_{i=0}^2$ because $|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| = \mathbb{1} - |\psi_0\rangle\langle\psi_0|$. By applying Eq. (6.1) on the left-hand side, Eq. (6.2) is equivalent to the following interval-inclusion

$$\begin{aligned} &\mathsf{T} - \left[f^L\left(|\psi_0\rangle\right), f^R\left(|\psi_0\rangle\right)\right] \subseteq \left[f^L\left(|\psi_1\rangle\right), f^R\left(|\psi_1\rangle\right)\right] + \left[f^L\left(|\psi_2\rangle\right), f^R\left(|\psi_2\rangle\right)\right] \qquad (6.3)\\ \Leftrightarrow\quad &\left[1 - f^R\left(|\psi_0\rangle\right), 1 - f^L\left(|\psi_0\rangle\right)\right] \subseteq \left[f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right), f^R\left(|\psi_1\rangle\right) + f^R\left(|\psi_2\rangle\right)\right] . \end{aligned}$$

This interval-inclusion can be rephrased as a long inequality

$$f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) \leq 1 - f^R\left(|\psi_0\rangle\right) \leq 1 - f^L\left(|\psi_0\rangle\right) \leq f^R\left(|\psi_1\rangle\right) + f^R\left(|\psi_2\rangle\right) . \tag{6.4}$$

Since $\|\mathcal{I}\| \leq \varepsilon$, the length of every interval $\left[f^L\left(|\psi\rangle\right), f^R\left(|\psi\rangle\right)\right]$ is bounded by ε as well, which implies the largest term in the previous inequality $f^R\left(|\psi_1\rangle\right) + f^R\left(|\psi_2\rangle\right)$ is bounded by $f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) + 2\varepsilon$. In other word, the left-end function f^L satisfies the following inequalities

$$\begin{aligned} &f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) \leq 1 - f^L\left(|\psi_0\rangle\right) \leq f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) + 2\varepsilon \qquad (6.5a)\\ \Leftrightarrow\quad &1 - 2\varepsilon \leq \sum_{i=0}^{2} f^L\left(|\psi_i\rangle\right) \leq 1 . \qquad (6.5b) \end{aligned}$$

In this language, Gleason's theorem basically states that when $\varepsilon = 0$, given any function $f^L\colon \mathbb{C}\mathbf{P}^2 \to$

$[0, 1]$ satisfying Eq. (6.5b), there exists a unique mixed state ρ such that

$$f^L\left(|\psi\rangle\right) = \langle\psi|\rho|\psi\rangle \tag{6.6}$$

for any state $|\psi\rangle \in \mathbb{C}\mathbf{P}^2$. Our Question 6.1 then ask how f^L would look like with a positive ε.

With different settings, whether there is an approximate version of Gleason's theorem was asked by Sam Sanders in constructivenews on 2013 [115], and there is no clear answer for his question as far as I understand. To have an idea of how hard this question could be, recall in Sec. 2.2 we state that a quantum probability space is glued by a family of classical probability spaces. **TODO. Check the differential geometry analog with Andy, since I don't really understand the differential geometry…** This is like the situation that a manifold is glued by many local coordinates. When each small piece has *exactly* the same and positive curvature, the Killing-Hopf theorem asserts this manifold is a sphere, but little can we say even if the curvature has a small deviation from constant. A similar situation might happen when approximating Gleason's theorem, but this time the whole space is glued by "local" classical probability space defined by each orthonormal basis $\{|\psi_i\rangle\}_{i=0}^2$. When the sum of f^L, $\sum_{i=0}^2 f^L\left(|\psi_i\rangle\right)$, is *exactly* the same and equal to 1, Gleason's theorem asserts that f^L can be expressed as Eq. (6.6). However, when each local classical probability space becomes imprecise, a general f^L might be as wild as we can imagine, and we might need to know a little bit more, like its QIVPM is a composite function, to deduce its global property.

6.2 AND BEYOND…

After we build the quantum interval-valued probability model, we might want to know how powerful a quantum computer could be based on this model. Since the conventional quantum circuit model manipulates the probability amplitudes instead of the measured probabilities, either a quantum computing model above QIVPMs needs to simultaneously manipulate all states in the core of a QIVPM, or we need to find a way to manipulate a QIVPM directly. However, both strategies are not

straightforward. On one hand, as we proved in Thm. 5.9, a QIVPM might have an empty core which cannot be evolved over time. On the other hand, if we want to manipulate and compute QIVPMs directly for multi-qubit algorithms, we need to glue the QIVPM for each qubit or subsystem together to get the QIVPM of the whole system, and this is not straightforward either.

A successful interval-valued theory might be further extended to the quantum theory over finite fields based on the following definition.

Definition 6.1 (Quantum Interval-valued Probability Measures over Finite Fields)**.** Consider a vector space $\mathcal{H}$ of dimension D over the complexified field $\mathbb{F}_{p^2}$, its set of events $\mathcal{E}_{p^2}$ as defined in Def. 4.2, and a collection of intervals $\mathcal{I}$. A quantum interval-valued probability measure over finite field $\bar{\mu}\colon \mathcal{E}_{p^2} \to \mathcal{I}$ assigns an interval to each event P subject to $\bar{\mu}(\mathbb{0}) = \mathbf{F}$, $\bar{\mu}(\mathbb{1}) = \mathbf{T}$, $\bar{\mu}\left(\mathbb{1} - P\right) = \mathbf{T} - \bar{\mu}\left(P\right)$, and satisfying for each pair of *commuting* events P_0 and P_1 with $P_0 P_1 = P_1 P_0$,

$$\bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) \subseteq \bar{\mu}\left(P_0\right) + \bar{\mu}\left(P_1\right) . \tag{6.7}$$

Understanding the properties of these probability measures and whether we could define a "sensible" Born upon them combine two approaches for dealing with the continuous quantities used in the conventional quantum theory and will be the next natural extension for our discrete quantum theories and computing.

BIBLIOGRAPHY

[1] Andrew J. Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. "Geometry of discrete quantum computing". In: *J. Phys. A: Math. Theor.* 46.18 (2013), p. 185301. Erratum "Corrigendum: Geometry of discrete quantum computing". In: *J. Phys. A: Math. Theor.* 49.3 (Dec. 2016), p. 039501.

[2] Andrew J. Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. "Discrete quantum theories". In: *J. Phys. A: Math. Theor.* 47.11 (2014), p. 115305.

[3] Yu-Tsung Tai, Andrew J. Hanson, Gerardo Ortiz, and Amr Sabry. "Quantum interval-valued probability: Contextuality and the Born rule". In: *Phys. Rev. A* 97.5 (5 May 2018), p. 052121.

[4] N. David Mermin. *Quantum Computer Science*. Cambridge University Press, 2007.

[5] Samson Abramsky. "Big toy models: Representing physical systems as Chu spaces". In: *Synthese* 186.3 (2012), pp. 697–718.

[6] Benjamin Schumacher and Michael D. Westmoreland. "Modal Quantum Theory". In: *Foundations of Physics* 42.7 (2012), pp. 918–925.

[7] Lay Nam Chang, Zachary Lewis, Djordje Minic, and Tatsu Takeuchi. "Galois Field Quantum Mechanics". In: *Modern Physics Letters B* 27.10 (2013), p. 1350064. eprint: `http://www.worldscientific.com/doi/pdf/10.1142/S0217984913500644`.

[8] John Gardiner. "Notes on Quantum Mechanics over a Finite Field". In: *Research Experience for Undergraduates. Research Reports*. Ed. by Chris Connell. Indiana University, Bloomington, 2014, pp. 5–18.

[9] David Deutsch. “Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 400.1818 (July 1985), pp. 97–117.

[10] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. New York, NY, USA: Cambridge University Press, 2000.

[11] The LyX Team. *LyX 2.3.0 - The Document Processor [Computer software and manual]*. Internet: http://www.lyx.org. Retrieved June 27, 2018, from http://www.lyx.org. 2018.

[12] A. M. Turing. “On Computable Numbers, with an Application to the Entscheidungsproblem”. In: *Proceedings of the London Mathematical Society* s2-42.1 (Jan. 1937), pp. 230–265. Erratum “On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction”. In: *Proceedings of the London Mathematical Society* s2-43.6 (Jan. 1938), pp. 544–546.

[13] Nicolas Gisin. “Time Really Passes, Science Can’t Deny That”. In: *Time in Physics*. Ed. by Renato Renner and Sandra Stupar. Cham: Springer International Publishing, 2017, pp. 1–15.

[14] Richard P. Feynman. “Simulating physics with computers”. In: *Int. J. Theor. Phys.* 21.6–7 (June 1982), pp. 467–488.

[15] Jonathan Shieber. *The reality of quantum computing could be just three years away*. Sept. 7, 2018. URL: `https://techcrunch.com/2018/09/07/the-reality-of-quantum-computing-could-be-just-three-years-away/`.

[16] Jeremy Kahn. “Quantum Computing. Hard-to-explain topics, explained simply”. In: *QuickTake* (10 Jan. 2019). Ed. by Leah Harrison Singer, p. 60.

[17] Dan Boneh. “Twenty years of attacks on the RSA cryptosystem”. In: *Notices of the AMS* 46.2 (Feb. 1999), pp. 203–213.

[18] Wikipedia. *RSA Factoring Challenge — Wikipedia, The Free Encyclopedia*. [Online; accessed 12-November-2016]. 2016.

[19] Thomas H. Parker and Scott J. Baldridge. *Elementary Mathematics for Teachers*. Sefton-Ash Publishing, 2004.

[20] Zhuangzi (庄子). *Zhuangzi (庄子)*. English. Trans. Chinese by Rongpei (榕培) Wang (汪), Xuqing (旭卿) Qin (秦), and Yongchang (雍长) Sun (孙). Changsha: Hunan People's Publishing House, 1999.

[21] Jinyan (金岩) Liu (刘). "Chinese Physicists' Construction of the Straton Model in Social Context". In: (June 2018), pp. 85–122.

[22] Wikipedia contributors. *Slide rule — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Slide_rule&oldid=861863336`. [Online; accessed 17-October-2018]. 2018.

[23] Wikipedia contributors. *Integrator — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Integrator&oldid=855449964`. [Online; accessed 17-October-2018]. 2018.

[24] Wikipedia contributors. *Charge amplifier — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Charge_amplifier&oldid=841714174`. [Online; accessed 27-October-2018]. 2018.

[25] Wikipedia contributors. *Analog computer — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Analog_computer&oldid=865018733`. [Online; accessed 23-October-2018]. 2018.

[26] Lenore Blum, Mike Shub, and Steve Smale. "On a theory of computation and complexity over the real numbers: NP- completeness, recursive functions and universal machines". In: *Bull. Amer. Math. Soc.* 21 (1989), pp. 1–46.

[27] Martin Ziegler. “Real Computability and Hypercomputation”. Habilitationsschrift. University of Paderborn, 2007.

[28] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. SpringerLink : Bücher. Springer New York, Dec. 6, 2012.

[29] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Apr. 11, 1976. 314 pp.

[30] Itzhak Gilboa and David Schmeidler. “Additive representations of non-additive measures and the Choquet integral.” In: *Annals of Operations Research* 52.1–4 (1994), pp. 43–65.

[31] Massimo Marinacci. “Limit Laws for Non-additive Probabilities and Their Frequentist Interpretation”. In: *Journal of Economic Theory* 84.2 (Feb. 1999), pp. 145–195.

[32] Kenneth David Jamison and Weldon A. Lodwick. *Interval-Valued Probability Measures*. Tech. rep. 213. Center for Computational Mathematics, University of Colorado Denver, 2004.

[33] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. English. 2nd ed. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Mar. 6, 2009. 354 pp.

[34] Michel Grabisch. *Set functions, games and capacities in decision making*. Theory and Decision Library C 46. Springer International Publishing, 2016.

[35] Rémy Mosseri and Rossen Dandoloff. “Geometry of entangled states, Bloch spheres and Hopf fibrations”. In: *Journal of Physics A: Mathematical and General* 34.47 (2001), p. 10243.

[36] Gregg Jaeger. *Quantum Information*. Springer New York, Apr. 3, 2007.

[37] Allen Hatcher. *Algebraic Topology*. Cambridge University Pr., 2001. 556 pp.

[38] I. Bengtsson and K. Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2007.

[39] Andrew Gleason. “Measures on the Closed Subspaces of a Hilbert Space”. In: *Indiana Univ. Math. J.* 6 (4 1957), pp. 885–893.

[40] Michael Redhead. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford University Press, 1987.

[41] Asher Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, Sept. 30, 1995. 464 pp.

[42] Fred Richman and Douglas Bridges. “A Constructive Proof of Gleason’s Theorem”. In: *Journal of Functional Analysis* 162.2 (1999), pp. 287–312.

[43] Jan Hamhalter. *Quantum Measure Theory*. Vol. 134. The Fundamental Theories of Physics. Springer Science & Business Media, Oct. 31, 2003. 420 pp.

[44] Max Born. “On the Quantum Mechanics of Collisions”. English. In: *Quantum Theory and Measurement*. Trans. by John Archibald Wheeler and Wojciech Hubert Zurek. Princeton University Press, 1983, pp. 52–55.

[45] Benjamin Schumacher and Michael D. Westmoreland. *Non-contextuality and free will in modal quantum theory*. Oct. 26, 2010. arXiv: `1010.5452v1 [quant-ph]`.

[46] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”. In: *Phys. Rev. Lett.* 70 (13 Mar. 1993), pp. 1895–1899.

[47] Christos H. Papadimitriou. *Computational Complexity*. Addison-Wesley Publishing Company, Dec. 11, 1993.

[48] L. G. Valiant and V. V. Vazirani. “NP is as easy as detecting unique solutions”. In: *Theoretical Computer Science* 47 (1986), pp. 85–93.

[49] Larry C. Grove. *Classical Groups and Geometric Algebra*. Fields Institute Communications. American Mathematical Society, 2002.

[50] D. R. Simon. “On the Power of Quantum Computation”. In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. SFCS ’94. Washington, DC, USA: IEEE Computer Society, 1994, pp. 116–123.

[51] Ethan Bernstein and Umesh Vazirani. “Quantum Complexity Theory”. In: *Proceedings of the Twenty-fifth Annual ACM Symposium on Theory of Computing*. STOC ’93. San Diego, California, USA: ACM, 1993, pp. 11–20. Updated Version “Quantum Complexity Theory”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1411–1473. eprint: `http://dx.doi.org/10.1137/S0097539796300921`.

[52] David Deutsch and Richard Jozsa. “Rapid Solution of Problems by Quantum Computation”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 439.1907 (1992), pp. 553–558. eprint: `http://rspa.royalsocietypublishing.org/content/439/1907/553.full.pdf`.

[53] Lov K. Grover. “A Fast Quantum Mechanical Algorithm for Database Search”. In: *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*. STOC ’96. Philadelphia, Pennsylvania, USA: ACM, 1996, pp. 212–219. Updated Version *A fast quantum mechanical algorithm for database search*. Nov. 19, 1996. arXiv: `quant-ph/9605043v3 [quant-ph]`.

[54] Lloyd S. Shapley. “Cores of convex games”. In: *International Journal of Game Theory* 1.1 (1971), pp. 11–26.

[55] Man-Chung(民忠) Ng(吳), Chi-Ping(寄屏) Mo(莫), and Yeong-Nan(永南) Yeh(葉). “On the Cores of Scalar Measure Games”. In: *Taiwanese Journal of Mathematics* 1.2 (1997), pp. 171–180.

[56] Gustave Choquet. “Theory of capacities”. In: *Annales de l’institut Fourier* 5 (1954), pp. 131–295.

[57] John S. Bell. “On the Problem of Hidden Variables in Quantum Mechanics”. In: *Rev. Mod. Phys.* 38.3 (3 July 1966), pp. 447–452.

[58] S. Kochen and E. Specker. “The Problem of Hidden Variables in Quantum Mechanics”. In: *Indiana Univ. Math. J.* 17 (1 1968), pp. 59–87.

[59] M. Artin. *Algebra*. Prentice Hall, 1991.

[60] Andrew J. Hanson. *Visualizing Quaternions*. Elsevier LTD, Oxford, Jan. 11, 2006. 600 pp.

[61] Wikipedia contributors. *Hopf fibration — Wikipedia, The Free Encyclopedia*. [Online; accessed 3-March-2018]. 2017.

[62] Ken Shoemake. “Animating Rotation with Quaternion Curves”. In: *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’85. New York, NY, USA: ACM, 1985, pp. 245–254.

[63] Wikipedia contributors. *Slerp — Wikipedia, The Free Encyclopedia*. [Online; accessed 3-March-2018]. 2018.

[64] Marcel Berger and Bernard Gostiaux. *Differential Geometry: Manifolds, Curves, and Surfaces*. Graduate Texts in Mathematics. Springer New York, 1988.

[65] Andrei Nikolaevich Kolmogorov. *Foundations of the Theory of Probability*. English. Trans. from the German by Nathan Morrison. New York: Chelsea Publishing Company, 1950.

[66] Robert B. Griffiths. *Consistent quantum theory*. Cambridge University Press, 2003.

[67] Arthur R. Swift and Ron Wright. “Generalized Stern-Gerlach experiments and the observability of arbitrary spin operators”. In: *Journal of Mathematical Physics* 21.1 (1980), pp. 77–82.

[68] George W. Mackey. “Quantum Mechanics and Hilbert Space”. In: *The American Mathematical Monthly* 64.8 (1957), pp. 45–57.

[69] Hans Maassen. “Quantum probability and quantum information theory”. In: *Quantum information, computation and cryptography*. Springer, 2010, pp. 65–108.

[70] Howard Barnum, Emanuel Knill, Gerardo Ortiz, and Lorenza Viola. “Generalizations of entanglement based on coherent states and convex sets”. In: *Physical Review A* 68.3 (Sept. 2003), p. 032308.

[71] Howard Barnum, Emanuel Knill, Gerardo Ortiz, Rolando Somma, and Lorenza Viola. “A Subsystem-Independent Generalization of Entanglement”. In: *Physical Review Letters* 92.10 (10 Mar. 2004), p. 107902.

[72] David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley & Sons, Inc., 2004.

[73] G. L. Mullen and C. Mummert. *Finite Fields and Applications*. American Mathematical Society, Rhode Island, 2007.

[74] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, Oxford, 2006.

[75] I. Stewart. *Galois theory*. Chapman and Hall/CRC, Boca Raton, 2004.

[76] Lay Nam Chang, Zachary Lewis, Djordje Minic, and Tatsu Takeuchi. “Quantum $\mathbb{F}_{\text{un}}$: the $q = 1$ limit of Galois field quantum mechanics, projective geometry and the field with one element”. In: *Journal of Physics A: Mathematical and Theoretical* 47.40 (2014), p. 405304.

[77] Roshan P. James, Gerardo Ortiz, and Amr Sabry. *Quantum Computing over Finite Fields*. Jan. 19, 2011. arXiv: `1101.3764v1 [quant-ph]`.

[78] Jeremiah Willcock and Amr Sabry. *Solving UNIQUE-SAT in a Modal Quantum Theory*. Feb. 17, 2011. arXiv: `1102.3587v1 [quant-ph]`.

[79] Andrew J. Hanson, Gerardo Ortiz, Amr Sabry, and Jeremiah Willcock. *The Power of Discrete Quantum Theories*. Apr. 8, 2011. arXiv: `1104.1630v1 [quant-ph]`.

[80] Wikipedia contributors. *Field norm — Wikipedia, The Free Encyclopedia*. `https://en.wikipedia.org/w/index.php?title=Field_norm&oldid=804256096`. [Online; accessed 21-April-2018]. 2017.

[81] Zhexian (哲先) Wan (万). *Geometry of Classical Groups over Finite Field (*有限域上典型群的几何学*)*. 2nd ed. Beijing : Science Press, 2006.

[82] Theodore W. Gamelin. *Complex Analysis*. Springer New York, July 17, 2003. 500 pp.

[83] V. I. Arnold. *Dynamics, Statistics and Projective Geometry of Galois Fields*. Cambridge University Press, 2010.

[84] Donald L Reisler and Nicholas M Smith. *Geometry Over a Finite Field*. Tech. rep. AD0714115. Defense Technical Information Center, Jan. 1969.

[85] N. J. A. Sloane and Simon Plouffe. *The Encyclopedia of Integer Sequences*. Academic Press, San Diego, 2005.

[86] Julian Schwinger. *Quantum Mechanics*. Ed. by Berthold-Georg Englert. Physics and astronomy online library. Springer-Verlag GmbH, Feb. 27, 2003.

[87] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. Institute for Advanced Study: `https://homotopytypetheory.org/book`, 2013.

[88] Stephen Wolfram. *An Elementary Introduction To The Wolfram Language*. Wolfram Media Inc, Jan. 14, 2016. 328 pp.

[89] A. P. Dempster. “Upper and Lower Probabilities Induced by a Multivalued Mapping”. In: *Ann. Math. Statist.* 38.2 (Apr. 1967), pp. 325–339.

[90] Kurt Weichselberger. “The theory of interval-probability as a unifying concept for uncertainty”. In: *Int. J. Approximate Reasoning* 24.2 (May 2000), pp. 149–170.

[91] Massimo Marinacci and Luigi Montrucchio. “Ultramodular Functions”. In: *Mathematics of Operations Research* 30.2 (2005), pp. 311–332.

[92] Giuseppe Vitali. “Sulla definizione di integrale delle funzioni di una variabile”. In: *Annali di Matematica Pura ed Applicata* 2.1 (Dec. 1925), pp. 111–121.

[93] Joachim Rosenmüller. “On core and value”. In: *Operations Research-Verfahren. Methods of operations research* 9 (1971), pp. 84–104.

[94] N. David Mermin. “Simple unified form for the major no-hidden-variables theorems”. In: *Phys. Rev. Lett.* 65.27 (27 Dec. 1990), pp. 3373–3376.

[95] Carsten Held. “The Kochen-Specker Theorem”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2016. Metaphysics Research Lab, Stanford University, 2016.

[96] David Meyer. “Finite Precision Measurement Nullifies the Kochen-Specker Theorem”. In: *Phys. Rev. Lett.* 83 (19 Nov. 1999), pp. 3751–3754.

[97] Hans Havlicek, Guenther Krenn, Johann Summhammer, and Karl Svozil. *On coloring the rational quantum sphere*. Nov. 9, 1999. arXiv: `quant-ph/9911040v1 [quant-ph]`. Published as Hans Havlicek, Günther Krenn, Johann Summhammer, and Karl Svozil. “Colouring the rational quantum sphere and the Kochen-Specker theorem”. In: *J. Phys. A: Math. Gen.* 34.14 (2001), p. 3071.

[98] N. David Mermin. *A Kochen-Specker Theorem for Imprecisely Specified Measurement*. Dec. 16, 1999. arXiv: `quant-ph/9912081v1 [quant-ph]`.

[99] Adrian Kent. “Noncontextual hidden variables and physical measurements”. In: *Phys. Rev. Lett.* 83.19 (1999), p. 3755.

[100] Christoph Simon, Časlav Brukner, and Anton Zeilinger. “Hidden-Variable Theorems for Real Experiments”. In: *Phys. Rev. Lett.* 86 (20 May 2001), pp. 4427–4430.

[101] Adán Cabello. “Finite-precision measurement does not nullify the Kochen-Specker theorem”. In: *Phys. Rev. A* 65 (5 Apr. 2002), p. 052101.

[102] J.-Å. Larsson. “A Kochen-Specker inequality”. In: *Europhys. Lett.* 58.6 (2002), p. 799.

[103] D. M. Appleby. “Existential contextuality and the models of Meyer, Kent, and Clifton”. In: *Phys. Rev. A* 65 (2 Jan. 2002), p. 022105.

[104] Jonathan Barrett and Adrian Kent. “Non-contextuality, finite precision measurement and the Kochen–Specker theorem”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 35.2 (2004), pp. 151–176.

[105] D. M. Appleby. “The Bell−Kochen−Specker theorem”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 36.1 (Mar. 2005), pp. 1–28.

[106] Robert W. Spekkens. “Contextuality for preparations, transformations, and unsharp measurements”. In: *Phys. Rev. A* 71.5 (2005), p. 052108.

[107] Otfried Gühne, Matthias Kleinmann, Adán Cabello, Jan-Åke Larsson, Gerhard Kirchmair, Florian Zähringer, Rene Gerritsma, and Christian F Roos. “Compatibility and noncontextuality for sequential measurements”. In: *Phys. Rev. A* 81.2 (2 Feb. 2010), p. 022121.

[108] Michael D. Mazurek, Matthew F. Pusey, Ravi Kunjwal, Kevin J. Resch, and Robert W. Spekkens. “An experimental test of noncontextuality without unphysical idealizations”. In: *Nat. Commun.* 7 (June 2016), ncomms11780.

[109] W. Heisenberg. *The actual content of quantum theoretical kinematics and mechanics*. English. Tech. rep. NASA-TM-77379. Washington, D.C.: National Aeronautics and Space Administration, Dec. 1, 1983. Trans. from “Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik”. In: *Zeitschrift für Physik* 43.3–4 (Mar. 1927), pp. 172–198.

[110] A. Wayne Roberts and Dale E. Varberg. *Convex Functions*. Pure & Applied Mathematics. Academic Press Inc, 1973.

[111] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. 3rd ed. Johns Hopkins Studies in Mathematical Sciences. Johns Hopkins University Press, 1996.

[112] Simon Foucart. "Lecture 6: Matrix Norms and Spectral Radii". lecture notes for the course NSTP187 at Drexel University, Philadelphia, PA, Fall 2012. 2012.

[113] E. M. Wright. "An Inequality for Convex Functions". In: *Amer. Math. Monthly* 61.9 (1954), pp. 620–622.

[114] Josip E. Pečarić, Frank Proschan, and Yung Liang Tong. *Convex Functions, Partial Orderings and Statistical Applications*. Vol. 187. Mathematics in Science and Engineering. Academic Press Inc, May 11, 1992. 467 pp.

[115] Sam Sanders. *Gleason's Theorem*. constructivenews - Google Groups. Aug. 28, 2013. URL: `https://groups.google.com/forum/#!topic/constructivenews/R7oZqVuVrDM` (visited on 01/29/2017).

Yu-Tsung Tai (戴湞琮)
LinkedIn: https://www.linkedin.com/in/yu-tsung-tai-9aa30551
GitHub: https://github.com/yuttai

EDUCATION

Indiana University Bloomington (IUB) (GPA: 3.802/4.0) **2010 – Present**

- Ph.D. double-majoring in Mathematics and Computer Science (expected) February 2019
- Master of Science in Computer Science May 2016
- Master of Arts in Mathematics December 2012

National Taiwan University (NTU) (GPA: 3.68/4.0) **2002 – 2006**

- Bachelor of Science in Mathematics (Rank: 4/48) June 2006

PUBLICATIONS

[1] Y.-T. Tai, A. J. Hanson, G. Ortiz, and A. Sabry, "Quantum interval-valued probability: Contextuality and the Born rule," *Phys. Rev. A,* vol. 97, no. 5, p. 052121, May 2018.

[2] A. J. Hanson, G. Ortiz, A. Sabry, and Y.-T. Tai, "Discrete Quantum Theories," *J. Phys. A: Math. Theor.,* vol. 47, p. 115305, 2014.

[3] A. J. Hanson, G. Ortiz, A. Sabry, and Y.-T. Tai, "Geometry of Discrete Quantum Computing," *J. Phys. A: Math. Theor.,* vol. 46, p. 185301, 2013. Erratum "Corrigendum: Geometry of Discrete Quantum Computing," *J. Phys. A: Math. Theor.,* vol. 49, p. 039501, 12 2016.

CONFERENCES AND SEMINARS

Quantum Interval-Valued Probability: Contextuality and the Born Rule

- Talk in Interdisciplinary Logic Seminar, IUB August 2017
- Poster Session in Contextuality: Conceptual Issues, Operational Signatures, and Applications, Perimeter Institute for Theoretical Physics July 2017

Introduction to Discrete Quantum Theories and Computing

- Talk in Theory Seminar, Department of Computer Science, IUB March 2017

Real Computation

- Talk in Theory Reading Group, Department of Computer Science, IUB Feb 2016

TEACHING EXPERIENCE

Taught with Full Responsibility

- MATH-T101 Mathematics for Elementary Teachers I, IUB — Fall 2017
- MATH-M216 Calculus II (Online), Indiana University East — Summer 2012

Designed and Edited Online Courses, Data Science Program, IUB

- Introduction to C++ (Designer) — Summer 2016 – Present
- Basic Linear Algebra and Calculus with Python (Designer) — Summer 2017 – Spring 2018
- Machine Learning with Python (Editor) — Fall 2016 – Spring 2018

Taught Recitation Sessions, IUB

- MATH-M211 Calculus I — Fall 2016
- MATH-M212 Calculus II — Summer 2014, Fall 2014, Fall 2015
- CSCI-B501 Theory of Computing — Spring 2015

Assisted and Graded, IUB

- CSCI-B609 Topics in Algorithms and Computing Theory (AlphaGo) — Spring 2018
- INFO-I231 Introduction to the Mathematics of Cybersecurity — Spring 2017
- CSCI-B503 Algorithms Design and Analysis — Spring 2016
- MATH-M119 Brief Survey of Calculus I — Fall 2013, Spring 2014
- MATH-M303 Linear Algebra for Undergraduates — Spring 2013
- MATH-M118 Finite Mathematics — Fall 2010, Fall 2012
- MATH-M301 Linear Algebra and Applications — Spring 2012
- MATH-M365 Introduction to Probability and Statistics — Fall 2011
- MATH-M120 Brief Survey of Calculus II — Spring 2011
- MATH-S312 Honors Course in Calculus IV — Spring 2011

Taught Mini-Courses in NTU Math Camps

- There is No Formula for General Quintic Equations in Terms of Radicals — 2005
- Game Theory — 2004

RESEARCH APPOINTMENTS

- Research Assistant, Department of Computer Science, IUB — July 2018 – Present
- Research Assistant, Kelley School of Business, IUB — May 2016
- Research Assistant, Department of Computer Science, IUB — Summer 2015
- Research Associate, Department of Computer Science, IUB — Summer 2013
- Research Assistant, Department of Economics, NTU — January 2008 – July 2009

TECHNICAL SKILLS

- Programming Languages:
  Python (with NumPy, matplotlib, SymPy, pandas, and TensorFlow), Mathematica, Visual Basic for Application, HTML, C/C++, LaTeX, MATLAB, Isabelle, Agda, Scheme, SQL
- Platforms:
  Microsoft Windows (7, 10, 8, XP, 98, 95, 3.1), Cygwin, Android, Red Hat Linux, MS-DOS 6.22
- Office Software:
  Microsoft Outlook, Microsoft PowerPoint, Microsoft Excel, Microsoft Word,
  Adobe Acrobat, Adobe Dreamweaver CC, Adobe Captivate 9,
  LyX, Trello, Google Docs, Google Sheets, Slack, emacs, Overleaf, Foxit Reader
- Version Control and Project Management:
  Git/GitHub, Apache Subversion, Agile, Trello, Slack
- Integrated Development Environments:
  Eclipse, PyCharm, Visual Studio 2013
- Fluency of Languages:
  Chinese (Native), English (Fluent), Japanese (Beginning), French Reading (Beginning)

AWARDS AND HONORS

- Studying Abroad Scholarship, Ministry of Education, Taiwan, R.O.C. 2010 – 2012
- Presidential Award, NTU Spring 2005, Fall 2005, Spring 2006
- Distinction Award, 1st Taiwan Mathematical Contest of Modeling for Undergraduate Students September 2003

CLUB ACTIVITIES

- Account Administrator of ptt2.cc 2006 – 2009
- NTU Go Club 2002 – 2006